F36 Wavefront analysis with a Shack-Hartmann sensor

of the Advanced Physics Lab for Physicists of the University of Heidelberg

Responsible MPIA staff contact:
Markus Feldt (feldt@mpia.de)

Latest update
September 2023

The instructions to this lab course are a rework of the former directions, largely done by Dr. Stefan Hippler, the former coordinator of this experiment, as well as Dr. Wolfgang Brandner and Prof. Dr. Thomas Henning. Further credit goes to Michael Newberry (1998) for the explanation on how to determine the system gain of a CMOS camera. To provide information on Zernike polynomials, an excerpt of the work of James C. Wyatt (2003) has been used. The latest contribution to this script is by Noa Bergmann.

Contents

1 Task overview

While the directions to this experiment seem long, the experiment in and of itself is rather short and can be done in a days work. It is advised to carefully read the pages, as they provide important insights into the astronomical context for this lab course as well as the working principle of adaptive optics.

In this experiment, we want to investigate the way adaptive optics works. To do so your **first task** will be to determine the system gain of a CMOS detector from the statistical properties of photons. This procedure is commonly referred to as the determination/measurement of the so-called photon transfer curve. The result is the system gain in units of electrons per ADU (analog-digital-unit). To perform the first task, an adjustable flat-field lamp is used.

For the **second task** of the experiment, a Shack-Hartmann sensor will be set up on an optical bench and put into operation. Subsequently, the optical quality of an eyeglass lens will be determined. From the determined Shack-Hartmann gradients, the wavefront for the correction of a typical eye defect will be calculated. Zernike polynomials are to be fitted to the measured Shack-Hartmann gradients, however this is required for the long report only.

1.1 Things to know and be familiar with

After reading these directions and doing your own research you should be prepared to answer the following questions. These are only a rough guideline for what you should take away from the following pages, you are very welcome to dive deeper into the subject.

1. What is the idea behind adaptive optics, why do we need it?
2. What is seeing? What causes it and what does it look like?
3. How does the principle of adaptive optics work?
4. What is a CMOS detector? What are its advantages?
5. What is the system gain of a CMOS detector? Which quantities and assumptions are used to calculate it?
6. What is the flat field effect? How does it come into play when determining the system gain?
7. What is a Shack-Hartmann sensor? How does it work?
8. What do wavefront gradients tell us about an incoming wavefront? How do we determine them (mathematically)?
9. What are polynomials? What limits to their use exist?

2 Introduction

Everyone has looked up into a cloudless sky at night at some point, to gaze at twinkling stellar constellations or even used a telescope to observe astronomical bodies.
However, if you want to make observations of the night-sky, this twinkling of the stars actually poses a problem. It is turbulence in the earth's atmosphere, causing fluctuating diffraction of the light of a star, and as a result significantly limiting the resolving powers of telescopes no matter their size. In astronomy, this blurring of single light sources through the earth's atmosphere is called *seeing*.
We often see this as the appearance of multiple 'copies' of the same star close together (called speckles) and slight variations in the brightness (called scintillation). The latter is what we know as the 'twinkling' of the stars. These phenomena also are wavelength-dependent, which is why the twinkling of stars is also often accompanied by variations in their color. This hindered astronomical measurements for a long time, which is why adaptive optics was developed. First envisioned in 1953 by Horace W. Babcock, it only came into usage in the 1990s when technological advances made its use feasible. Nowadays, Shack-Hartmann sensors are widely used to measure the aberrations coming from the refraction in the atmosphere. Where the size of telescopes increases the 'depth' with which one can look into the cosmos, adaptive optics increase the spatial/angular resolutions. The most commonly used sensors in astronomy are the so-called Curvature sensor and the Pyramid wavefront sensor, as well as the 'Shack-Hartmann-Sensor' which will be investigated in this experiment.

3 Fundamental principles

To understand the context of this lab course and the meaning behind adaptive optics, it is important to look at the astronomical phenomena leading to its development. This chapter will give an overwiev over 'seeing', the working principle of adapted optics, the components and characteristics of a CMOS detector as well as the math behind wavefront reconstruction from measured gradients.
To start, you will get some context as to why adaptive optics are needed in the first place.

3.1 Seeing

Seeing is the phenomenon resulting from the refractive index of the earths atmosphere not being constant, but dependent on the density of air which changes with different temperatures. Since cold and warm air is constantly and chaotically mixed in atmospheric turbulence, patches of air ('turbulence cells') with different temperature lead to slight refraction. These turbulence cells diffract the light of a single source into rapidly moving speckles, which smear the perceived image of the source.
The cells are randomly distributed with diameters of 10 to 20cm, which act like weak lenses. Due to their light-collecting effect, they cause intensity fluctuations over areas of

a few centimetres near the ground.
When observing with a telescope several meters in diameter, the turbulence is less noticable through intensity fluctuations of the stars ('twinkling'). Instead one often sees the disturbances as the granular structure ('speckles') on the image of an individual star.
Since many of these turbulence cells are captured simultaneously by the telescope aperture, the image fragments into many randomly distributed single images, the speckles. The number of speckles corresponds approximately to the number of turbulence cells over the telescope aperture, and the size of the star image made by those speckles is about one arcsecond under good atmospheric conditions.

The shape and position of the speckle patterns change depending on how fast the turbulence passes over the telescope aperture. In general, the exposure time must be less than 1/10 second to see a »sharp« speckle image. If several exposures with short exposure times are strung together like in a film, the temporal change of the turbulences can be seen as a teeming movement of the individual speckles and as the star image moving sideways left and right as well as up and down.
At exposure times of a few seconds, both effects contribute to the blurring of the speckle image, and a uniformly illuminated image »disc« is created, whose width FWHM (Full Width Half Maximum) is called »seeing« in astronomical terminology. Depending on the observation site and meteorological conditions, the seeing is 0.5 to 2.5 arcseconds.

The definition of the resolving power of a telescope can be used to establish a relationship between the seeing and the size of the turbulence cells. For the diffraction limited resolving power α (in radians) of a telescope with diameter D at the wavelength λ the following holds:

$$\alpha = \lambda/D \tag{1}$$

For example, in the visible, at a wavelength of 500μm, a 3.5m telescope, has a theoretical resolving power of 1.4310^{-7} rad or about 0.03 arcseconds. This means that two point-like stars that are 0.03 arcseconds apart can just about be resolved as single stars.
Thus the known seeing limited angular resolution power of 1 arcsecond corresponds to the diffraction limited angular resolution of a telescope with a diameter of 10cm at 500nm. This is just about the relevant size of the turbulence cells in the atmosphere.

In the context of describing turbulence as a statistical phenomenon, instead of turbulence cells, one speaks of the coherence length r_0, which is also called the *Fried parameter* after an American physicist. This parameter denotes the diameter of the wavefront over which the wavefront disturbance, i.e. the deviation from the plane wave, is negligible (standard deviation $<$ 1rad). Therefore, the vivid picture of turbulence cells with the diameter r_0, within which the light can propagate undisturbed, is absolutely correct. The greater the coherence length becomes, i.e. the more r_0 approaches the telescope diameter D, the closer the resolving power comes to the diffraction limit α. Note that r_0 is wavelength dependant.

The wavefront can be imagined as a mountain and valley landscape created by the turbulence, whose typical height difference is in the range of $3 - 6\mu$m, regardless of the wavelength. Relevant for the image quality is the ratio of the deviation of the wavefront from the plane wave to the observation wavelength; for example, a numerical value of $3-6\mu$m at a wavelength of 10μm is a relatively insignificant value of 0.3 to 0.6 wavelengths units. On the other side, in the visible at a wavelength of 0.5μm, the deviation is 6 - 12 wavelengths units, a much stronger effect. A formula for the coherence length r_0 describes this effect quantitatively. The Fried parameter depends on the wavelength as $r_0 \sim \lambda^{6/5}$. This means that the range over which wavefront distortions are negligible increases with wavelength. A r_0 of typically 10cm at 0.5μm is then equivalent to an r_0 of 360cm at 10μm!

Accordingly, the speckle patterns at 0.5μm and at 10μm look completely different under identical atmospheric conditions (see figure 1). On a 3.5m telescope, there is a cloud of $D/r_0 \approx 1000$ in the visible at 0.5μm, which are in a wild swarming movement and whose envelope has a FWHM of 1 arcsecond. Each speckle has the size of the diffraction disk of the telescope, i.e. 0.036 arcseconds. In the infrared at 10μm there is only one speckle with a FWHM of 0.6 arcseconds, the size of the diffraction disk at 10μm, because r_0 is larger than the telescope diameter D. The only effect of the turbulence is that the image moves relatively slowly back and forth, and up and down.

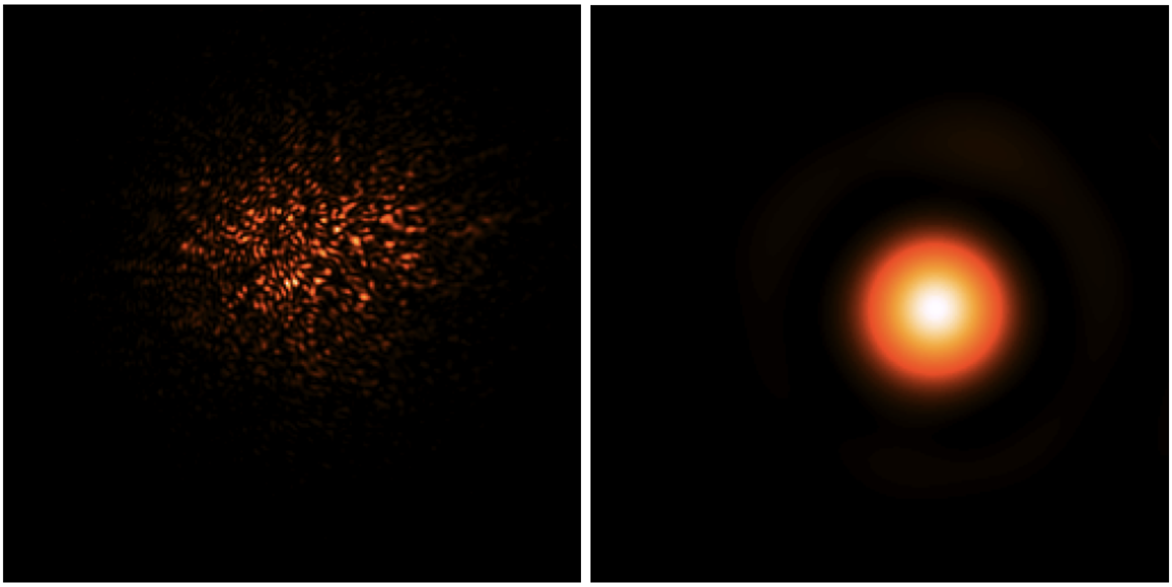

Figure 1: Star images simulated at $\lambda = 0.5\mu$m (left) and $\lambda = 10\mu$m (right) for identical atmospheric conditions on a 3.5m telescope. While at $\lambda = 0.5\mu$m there is a speckle pattern with about 1000 single speckles and an envelope of one arcsecond, at $\lambda = 10\mu$m one can already see the diffraction disk with only slightly deformed diffraction rings.

3.2 Adaptive optics

Adaptive optics typically includes a tip-tilt mirror and a deformable mirror, which continuously compensate for the atmospheric wavefront deformations caused by turbulence cells. The (compensated) deformations are endlessly measured with a wavefront sensor. A real-time control system (real-time computer) converts the signals from the wavefront sensor into control signals for the correcting element. A part of the light is decoupled and sent to the observation, which is usually an infrared camera.
Figure 2 shows a scheme of an adaptive optics system often used in astronomy and its components. There are a number of methods to measure wavefront disturbances. The most common in astronomy are the curvature wavefront sensing method, the Shack-Hartmann wavefront analysis and the analysis with pyramid wavefront sensors. The principle of wavefront measurements with the Shack-Hartmann sensor is explained in section 3.4

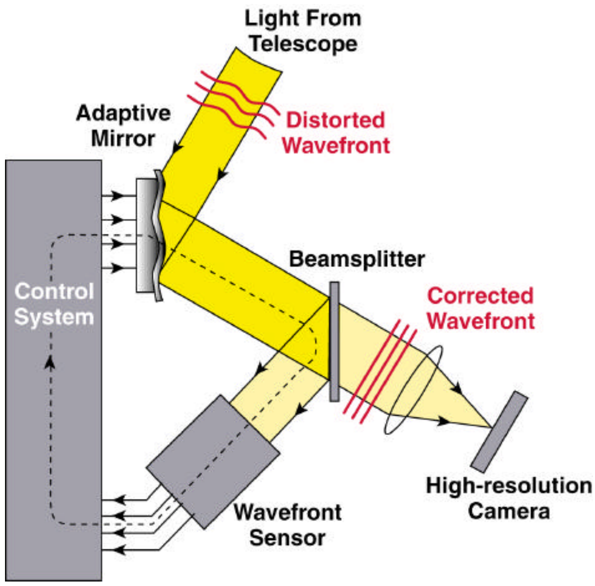

Figure 2: Schematic representation of an adaptive optics system in astronomy (Hippler and Kasper, 2004)

3.3 The CMOS camera

To measure wavefronts, a 2D detector for the incoming light is necessary. In this experiment a CMOS camera of the type ZWO ASI183 MM, which comes with an IMX183CLK CMOS detector is used. The main part of the first half of the experiment is to determine the system gain of the CMOS camera. To do so, one must first understand the different components and characteristics of a CMOS camera.

3.3.1 Basic working principle

A CMOS camera is usually a camera which is equipped with a CMOS detector. It has the ability to capture two dimensional images and give them out digitally. The camera or the sensor within the camera is an image detector with active pixels, i.e. each pixel is a photon detector (typically a pinned photodiode) with active amplifier. The generic category of such sensors is often referred to as Active Pixel Sensor (APS). The sensors contain large amounts of pixels, which are in their simplest realisation composed of a photodiode and three n-channel-MOSFETs (figure 3).
Before each measurement, the photodiode in a pixel is set to a defined starting voltage $V_{DD} - U_{th}$, where U_{th} is the threshold voltage of the backset-transistor. During the measurement, when light hits the photodiode, the resulting photocurrent discharges it, decreasing the voltage on the photodiode. After the measurement, the remaining voltage is read out over an amplification transistor, with the signal being fed to an analog to digital converter via the selection transistor. Since the discharge of the photodiode is proportional to the radiant flux density and integration time, the detected voltage indicates 'how much' light hit the photodiode, thus giving an image of the detected light.

The Sony chip used in our experiment is manufactured in the Complementary metal-oxide-semiconductor technology, hence the name CMOS. As opposed to a charge-coupled device (CCD) image sensor, a CMOS sensor requires only a single power source, and is able to read off the charges much faster than a CCD. This is due to its main advantage over the CCD, which cannot read out pixels individually, as the CMOS can. It emerged as an alternative to CCD image sensors, and largely replaced them by the mid 2000s, after initial disadvantages such as high noise and low sensitivity were overcome.

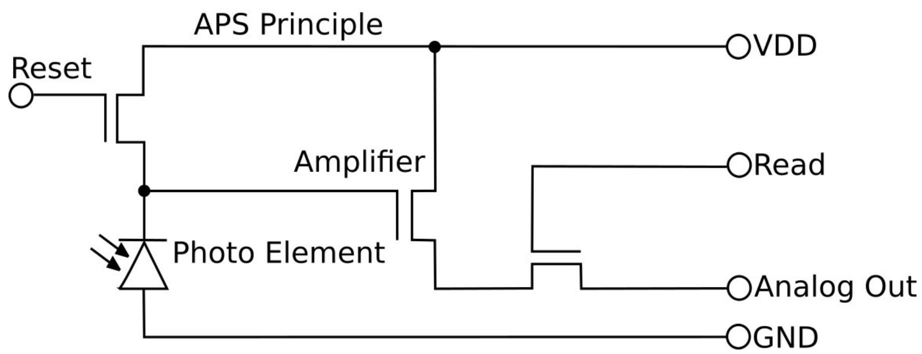

Figure 3: Schematic structure of a basic single active pixel in a CMOS sensor (Stohr)

Please note that there are many similarities but also differences between CCD and CMOS detectors. However, the method for determining the system gain of those detectors is the same for both.

Technical specifications of the camera are given in the appendix.

3.3.2 Features of the CMOS camera

Introduction and explanation of terms A CMOS camera is distinguished from other optical detectors, such as eye, film or photo-multiplier, by two characteristics, in particular:

Excellent sensitivity The quantum efficiency (QE) is around 50% - for some wavelengths even up to 90% - and is thus one order of magnitude larger than in the human eye or in film, where the QE is only a few percent.

Linearity between measurement signal and quantity of light collected Unlike photographic plates, which only cover a comparatively small range of linear response, the CMOS detector behaves linearly over a range starting from zero noise levels to saturation levels. This makes a CMOS camera particularly suitable for photometric applications.

For the development of the CCD, Willard Boyle and George E. Smith received the 2009 Nobel Prize in Physics.

The CMOS detector is further characterized by the following values, **of which some shall be determined during the lab course**.

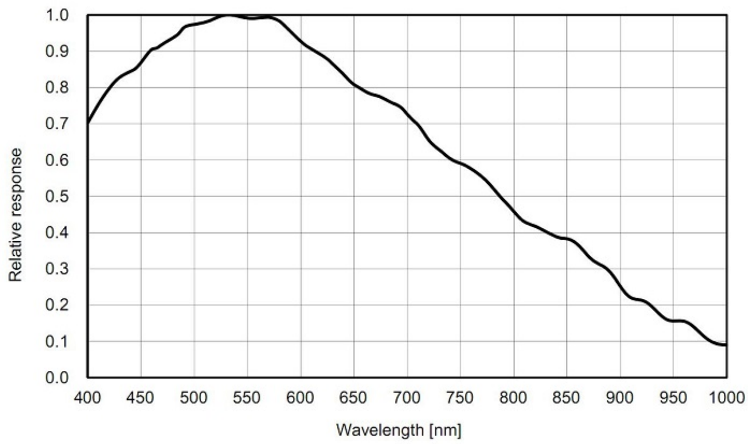

Figure 4: Relative QE of the IMX183 detector. The maximum value is ≈ 0.84. Source: ZWO ASI183 manual. (Image Credit: Dr. Stefan Hippler) Kamera Manual Quelle

Full well capacity is defined as the amount of charge a pixel can hold without being saturated. A larger pixel results in a higher full well capacity. In the Sony IMX183 CMOS detector each pixel has a size of 2.4μm x 2.4μm, resulting in a full well capacity of 15-16.000 electrons.

Quantum efficiency, QE determines the sensitivity of the detector to incident electromagnetic radiation. The quantum efficiency is defined as the number of processes that an absorbed photon creates on average. The QE is determined by the energy of the photon and thus from the wavelength of light or the electromagnetic radiation. The QE indicates the probability with which the inner photoelectric effect releases a photo-electron, which can be measured.

The quantum yield of the Sony IMX183 CMOS detector can be determined from figure 4.

Bias level is the average pixel value that the camera electronics reads out at shortest exposure time and without flux on the detector; this is ideally a dark image with an exposure time of 0 s. Since it is an electronic offset, this value is usually subtracted from the measured data.

Readout noise consists of fluctuations of the signal, which are caused by the electronics during readout, i.e. analog to digital conversion.The readout noise can be expressed as standard deviation in analog to digital units (ADU) per pixel or electrons per pixel. For the conversion from ADU to electrons the camera amplification factor is needed (see below).

Dark current consists of electrons which are generated by the temperature of the detector (i.e. thermal noise, Fermi distribution, Boltzmann distribution), i.e. electrons that move from the valence band to the conduction band and thus add to the measured signal.

Flat-field is an image where the pixels are all equally illuminated with either monochromatic light or white light. The flat-field image shows the differences in sensitivity of the individual pixels. Especially in astronomy, with its very low-contrast objects, the flat-field image often shows additional features. These can be caused by dust on optical elements in front of the detector or on the detector itself.

Camera system gain describes the relationship between the photo-electrons measured and the signal in ADU, that you get on the computer. The system gain is not identical to the programmable gain amplifier (PGA) of the CMOS camera. This programmable gain is used to match the maximum amplitude of the CMOS signal to the full voltage range of the analog to digital converter (ADC).

Signal-to-noise ratio SNR characterizes the quality of an image. In general it should be kept as high as possible, because it reduces the measurement noise of the Shack-Hartmann sensor.

The last two values are particularly important. The camera's system gain represents the connection between a pure number resulting from the ADC and the physically interesting number of photo-electrons. In addition, if you know the quantum efficiency of the detector and the wavelength at which it was observed, the number of photons absorbed by a pixel is known. The detector has a higher full well capacity than the maximal bit resolution of the ADU, so it is sensible to convert it with the system gain. More on that in section 3.3.3

3.3.3 The system gain

To understand what the system gain is, we must first introduce some of the cameras components.

Analog to digital converter The analog to digital converter, ADC for short, is the component which transforms an analog signal to a digital one. The analog signal in this case is the remaining voltage measured on the pixels of the sensor, as the CMOS discharges electrons when being exposed to light. The conversion from analog to digital usually implies transforming from a continuous measurement to a discrete one, since we cannot measure continuous events digitally.
This is were the resolution of the converter comes into play. It indicates how many gradations there are, or how close to each other two analog signals can be before the converter is able to differentiate between the two. In other words: The higher the resolution of the converter, the finer the stepsize in which the events are measured.

For example: If one wants to digitally measure the intensity of light in arbitrary units from 0 to 1, they would have convert the incoming analog light signal to a digital one. If we have a 1-bit (2^1) ADC, we would be able to distinguish between two different intensities: From 0 to 0.5 and from 0.5 to 1. A 2-bit (2^2) ADC could already sort the intensities into 4 different gradations.
So in terms of intensities, where 0 is pitch black and 1 is super white, two to the power of the bit-number of the ADC tells you how many shades of grey can be distinguished in the digital data of the measurement. Thus the 12bit ADC used in the experiment can resolve $2^{12} \approx 4096$ shades in the image.

Pixel binning With a given ADC at hand, there still is a way to increase the resolution of our digital image in the software. This is called pixel binning. Binning means that an arbitrary amount of neighbour pixels are interpreted as one. This proportionally reduces the total amount of pixels in the image, but also increases the number of shades by the same factor.

Example: If you set the pixel binning to 4, this means that squares of 4 pixels are regarded as one. This in turn means, that if you had 1600 pixels beforehand, your new effective number of pixels is 400. However, if you had a 4-bit ADC, the amount of shades would also increase by the factor 4, from 16 to 64. This also increases the maximal value of our measurements, due to the signals of the binned pixels being added. However since in a CMOS each pixel is read out individually, the binning is just a calculation, meaning we can easily rescale the intensities, without any physical meaning behind it.
Note: Due to the individual readout of the pixels in a CMOS, binning really is just a trick to get around technical problems, it has no physical impact on the measurement. In a CCD however, where all pixels are readout simultaneously the binning has physical impact.

CMOS camera system gain The gain of a CMOS camera is the conversion between the number of electrons ("e-") recorded by the CMOS and the number of digital units (ADU, "counts") contained in the CMOS image. It is useful to know this conversion for evaluating the performance of the CMOS camera, and it is expressed in units of electrons per count.

Since quantities in the CMOS image can only be measured in units of counts, knowing the gain permits the calculation of quantities such as readout noise and full well capacity in the fundamental units of electrons.
The gain value is required by some types of image deconvolution such as Maximum Entropy since, in order to properly do the statistical part of the calculation, the processing needs to convert the image into units of electrons. Calibrating the gain is also useful for detecting electronic problems in a CMOS camera, including gain change at high or low signal level, and the existence of unexpected noise sources.

The gain value is set by the electronics that read out the CMOS chip. For example, a gain of $1.8e^{-}$/count means that the camera produces 1 count for every 1.8 recorded electrons. Of course, we cannot split electrons into fractional parts, as in the case for a gain of $1.8e^{-}$/count. What this number means is that 4/5 of the time 1 count is produced from 2 electrons, and 1/5 of the time 1 count is produced from 1 electron. This number is an average conversion ratio, based on changing large numbers of electrons into large numbers of counts.

Note: This use of the term "gain" is in the opposite sense to the way a circuit designer would use the term. In electronic design, gain is considered to be an *increase* in the number of output units compared with the number of input units.
The conversion used in this experiment might sound confusing at first, but due to the different noise sources (see table 1) it is important to not simply amplify the electron signal, and in turn amplifying the noise, and rather put several electrons into one count. Another reason for this particular conversion is the full well capacity of the detector (3.3.2). The maximum amout of electrons in the pixel is usually higher than the maximum amount of counts in the image. Since we want to be able to display the full capacity of the pixel it is thus useful to scale it that way.

For example, if the readout noise of the CMOS detector is around 3 electrons per pixel, it is sensible to avoid too fine noise sampling by picking a system gain of 1-3 electrons per ADU. Given a 12-bit ADC this would result in a analog measuring range of 4000 to 12000 electrons per pixel. Since we further want to set the system gain such that the analog range of a pixel (full well capacity) is covered digitally to the fullest extent possible, this gain might need to be adjusted. Given a detector with a full well capacity of 16000 electrons as used in this lab course, a system gain of 1-4 electrons per ADU might be more suitable.
It is important to note that every measurement you make in a CMOS image uses units of counts. Since one camera may use a different gain than another camera, count units

do not provide a straightforward comparison to be made. For example, suppose two cameras each record 24 electrons in a certain pixel. If the gain of the first camera is 2.0 and the gain of the second camera is 8.0, the same pixel would measure 12 counts in the image from the first camera and 3 counts in the image from the second camera. Without knowing the gain, comparing 12 counts against 3 counts is pretty meaningless.

3.3.4 Theory to determining the system gain

To determine the system gain the photon transfer curve (figure 5) is used. The photon transfer curve represents the measured detector signal vs. the variance of the signal for different illumination levels. Flat-field effects are usually corrected to erase possibly large errors in the gain calculation. The system gain can be identified as the slope of the linear part of the curve.

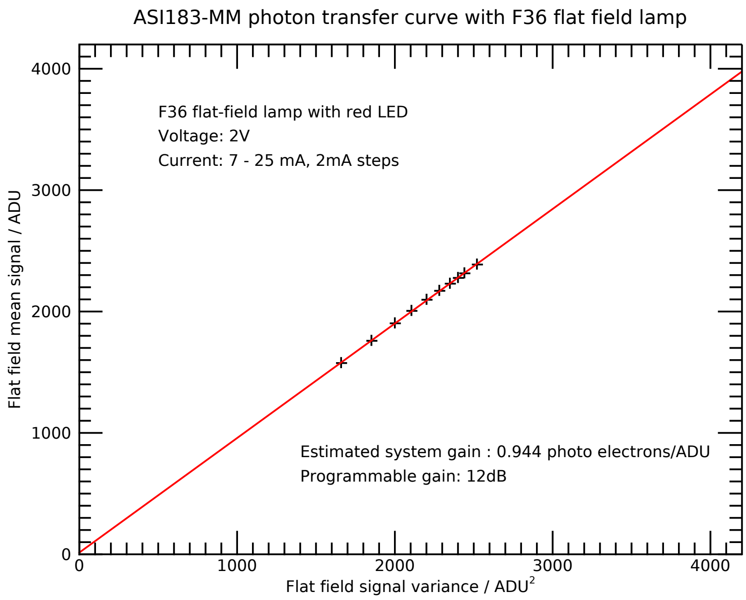

Figure 5: Example of a photon transfer curve (Image Credit: Dr. Stefan Hippler)

The signal recorded by a CMOS and its conversion from units of electrons to counts (system gain) can be mathematically described in a straightforward way. Understanding the mathematics validates the gain calculation technique described in the next section, and it shows why neglecting the flat field effect fails to give the correct answer.

Signal and noise This derivation uses the concepts of ”signal” and ”noise”. CMOS performance is usually described in terms of signal to noise ratio, or ”S/N”, but we shall deal with them separately here. The signal is defined as the quantity of information you measure in the image - in other words, the signal is the number of electrons recorded by the CMOS detector or the number of counts present in the CMOS image. The noise is the uncertainty in the signal. Since the photons recorded by the CMOS arrive in random packets (courtesy of nature), observing the same source many times records a different number of electrons every time. This variation is a random error, or ”noise” that is added to the true signal. You measure the gain of the CMOS by comparing the signal level to the amount of variation in the signal. This works because the relationship between counts and electrons is different for the signal and the variance. There are two ways to make this measurement:

1. Measure the signal and variation within the same region of pixels at many intensity levels.
2. Measure the signal and variation in a single pixel at many intensity levels.

Both of these methods have the same mathematical foundation. To keep these directions a bit shorter, only the preferred method 1 which requires far less image processing effort will be explained in detail.

To derive the relationship between signal and variance in a CMOS image, let us define the following quantities:

S_C	The signal measured in count units in the CMOS image.
S_E	The signal recorded in electron units by the CMOS chip. This quantity is unknown.
N_C	The total noise measured in count units in the CMOS image.
N_E	The total noise in terms of recorded electrons. This quantity is unknown.
g	The gain, in units of electrons per count. This will be calculated.
R_E	The readout noise of the CCD chip, measured in electrons. This quantity is unknown.
σ_E	The photon noise in the signal N_E.
σ_o	An additional noise source in the image. This is described below.

Table 1: Important quantities

We need an equation to relate the number of electrons, which is unkown, to quantities we measure in the CMOS image in units of counts. The signals and noises are simply related through the gain factor as

$$S_E = gS_C \quad \text{and} \quad N_E = gN_C \tag{2}$$

these can be inverted to give

$$S_C = \frac{1}{g} S_E \quad \text{and} \quad N_C = \frac{1}{g} N_E \tag{3}$$

The noise is contributed by various sources. We consider these to be readout noise, R_E, photon noise attributable to the nature of light, σ_E, and some additional noise, σ_0, which is important to characterize the flat field effect. Remembering that the different noise sources are independent of each other, they add in quadrature. This means that they add as the square of their noise values. If we could measure the total noise in units of electrons, the various noise sources would combine in the following way:

$$N_E^2 = R_E^2 + \sigma_E^2 + \sigma_{o,E}^2 \tag{4}$$

The random arrival of photons controls the photon noise, σ_E. Photon noise obeys the laws of Poissonian statistics, which makes the square of the noise equal to the signal, or $\sigma_E^2 = S_E$. Therefore, we can make the following substitution.

$$N_E^2 = R_E^2 + S_E + \sigma_{o,E}^2 \tag{5}$$

Knowing how the gain relates units of electrons and counts, we can modify this equation to read as follows:

$$g^2 N_C^2 = g^2 R_C^2 + g S_C + g^2 \sigma_{o,C}^2 \tag{6}$$

which then gives

$$N_C^2 = R_C^2 + \frac{1}{g} S_C + \sigma_{o,C}^2 \tag{7}$$

We can rearrange this to get the final equation:

$$N_C^2 = \frac{1}{g} S_C + (R_C^2 + \sigma_{o,C}^2) \tag{8}$$

Flat field effect $\sigma_{o,C}$ is the noise caused by the built in pixel to pixel variation in the image, also known as the flat field effect. The flat field effect produces a pattern of apparently "random" scatter in a CMOS image. Even an exposure with infinite signal to noise ratio ("S/N") shows the flat field pattern. Despite its appearance, the pattern is not actually random because it repeats from one image to another. Changing the color of the light source changes the details of the pattern, but the pattern remains the same for all images exposed to light of the same spectral makeup. The importance of this effect is that, although the flat field variation is not a true noise, unless it is removed from the image it contributes to the noise you actually measure.

The characterization of the noise contributed by the flat field effect is rather simple: Since the flat field pattern is a fixed percentage of the signal, the standard deviation, or "noise" you measure from it is always proportional to the signal.

For example, a pixel might be 1% less sensitive than its left neighbor, but 3% less sensitive than its right neighbor. Therefore, exposing this pixel at the 100 count level produces the following 3 signals: 101, 100, 103. However, exposing at the 10,000 count level gives these results: 10,100, 10,000, 10,300. The standard deviation for these 3 pixels is $\sigma_{0,C} = 2,3333$ counts for the low signal case but is $\sigma_{0,C} = 233.333$ counts for the high signal case. Thus the standard deviation is 100 times larger when the signal is also 100 times larger.

This proportionality can be expressed in a mathematical way:

$$\sigma_{o,C} = kS_C \tag{9}$$

In the present example, we have $k = 0.02333$. Substituting this expression for the flat field variation into our master equation, we get the following result:

$$N_C^2 = R_C^2 + \frac{1}{g} S_C + k^2 S_C^2 \tag{10}$$

This reveals a nice quadratic function of signal.

When plotted with the Signal on the x axis, this equation describes a parabola that opens upward. Since the Signal - Variance plot is actually plotted with Signal on the y axis, we need to invert this equation to solve for S_C:

$$S_C = \frac{-1 + \sqrt{1 + 4g^2k^2(N_C^2 - R_C^2)}}{2gk^2} \tag{11}$$

This final equation describes the classic Signal - Variance plot. In this form, the equation describes a family of horizontal parabolas that open toward the right. The strength of the flat field variation, k, determines the curvature. When k = 0, the curvature goes away and it gives the straight line relationship we desire. The curvature to the right of the line means that the stronger the flat field pattern, the more the variance is inflated at a given signal level. This result shows that it is impossible to accurately determine the gain from a Signal - Variance plot unless we know one of two things: Either 1) we know the value of k, or 2) we setup our measurements to avoid flat field effects. Option 2 is the correct strategy.
Essentially, adjusting the measurement in such a way that flat field effects can be neglected results in a straight line relationship for signal and variance. If the measuring method does not take flat field effects into account, which is the weakness of formula 10, where we assumed $\sigma_{o,C}$ to be constant, the result is always an underestimated gain.

To illustrate the effect of flat field variations, mathematical models were constructed using the equation above with parameters typical of commonly available CCD cameras. These include readout noise $R_E = 15e^-$ and gain $g = 2.0e^-/count$. Three models were constructed with flat field parameters $k = 0$, $k = 0.005$, and $k = 0.01$. Flat field variations of this order are not uncommon. These models are shown in the figure 6 below.

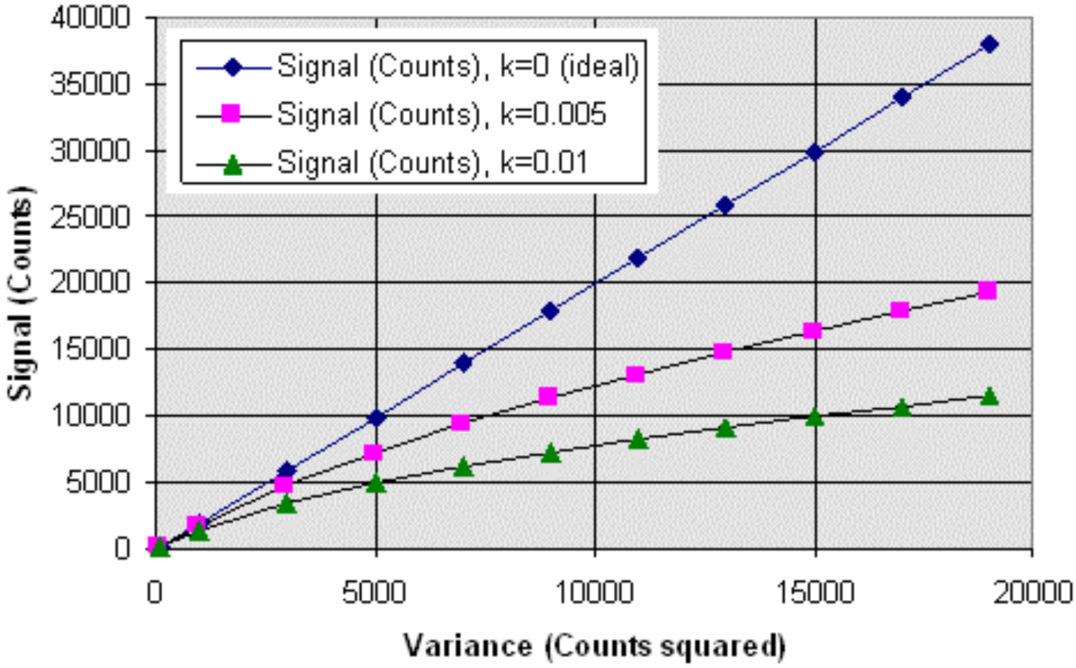

Figure 6: Signal-variance plot for $g = 2.0\frac{e^-}{count}$, $R = 15e^-$ (Image Credit: Dr. Stefan Hippler)

Increasing values of k correspond to progressively larger flat field irregularities in the CMOS chip. The amplitude of flat field effects, k, tends to increase with shorter wavelength, particularly with thinned CMOS's. The flat field pattern is present in every image exposed to light. Clearly, it can be seen from the models that if one simply obtains images at different signal levels and measures the variance in them, then fitting a line through any part of the curve yields a slope lower than its true value. Thus the simple method of section 4 always underestimates the gain.

The best strategy for doing the Signal - Variance method is to find a way to produce a straight line by properly compensating for flat field effects. This is important by the "virtue of straightness": Deviation from a straight line is completely unambiguous and easy to detect. It avoids the issue of how much curvature is attributable to what cause. The electronic design of a CMOS camera is quite complex, and problems can occur, such as gain change at different signal levels or unexplained extra noise at high or low signal levels. Using a "robust" method for calculating gain, any significant deviation from a line is a diagnostic of possible problems in the camera electronics. The method used in this lab course is described in the following section.

Method for determining the system gain This section presents a robust method to correct the flat field effects in the Signal-Variance relationship to yield the desired straight-line relationship. This permits an accurate gain value to be calculated. Instead of removing flat field effects, one could also either attept to use a low signal level where flat field effects are believed not to be important or to measure and compensate for the flat

field parameter k. However it is better to find a way of measuring the signal in such a way that flat field effects do not need to be taken in to account in later calculations.
In both of the two mentioned ways to make the measurement, one must consider some procedural issues:

- Both methods measure sets of 2 or more images at each signal level. An image set is defined as 2 or more successive images taken under the same illumination conditions. To obtain various signal levels, it is better to change the intensity received by the CMOS than to change the exposure time. This may be achieved either by varying the light source intensity or by altering the amount of light passing into the camera. The illumination received by the CMOS should not vary too much within a set of images, but it does not have to be truly constant.

- Cool the CCD camera to reduce the dark current to as low as possible. This prevents you from having to subtract dark frames from the images (doing so adds noise, which adversely affects the noise measurements at low signal level). In addition, if the bias varies from one frame to another, be sure to subtract a bias value from every image. For very short or constant integration times, the dark current does not have to be measured.

- The CMOS should be illuminated the same way for all images within a set. Irregularities in illumination within a set are automatically removed by the image processing methods used in the calibration. It does not matter if the illumination pattern changes when you change the intensity level for a different image set.

- Within an image set, variation in the light intensity is corrected by normalizing the images so that they have the same average signal within the same pixel region. The process of normalizing multiplies the image by an appropriate constant value so that its mean value within the pixel region matches that of other images in the same set. Multiplying by a constant value does not affect the signal to noise ratio or the flat field structure of the image.

- Do not estimate the CMOS camera's readout noise by calculating the noise value at zero signal. This is the square root of the variance where the gain line intercepts the y axis. Especially do not use this value if bias is not subtracted from every frame. To calculate the readout noise, use the "Two Bias" method and apply the gain value determined from this test. In the Two Bias Method, 2 bias frames are taken in succession and then subtracted from each other. Measure the standard deviation inside a region of, say 100x100 pixels and divide by 1.4142. This gives the readout noise in units of counts. Multiply this by the gain factor to get the Readout Noise in units of electrons. If bias frames are not available, cool the camera and obtain two dark frames of minimum exposure, then apply the Two Bias Method to them.

Now on to the steps to measure the signal without flat field effects at multiple signal levels. From the two methods we chose the strategy where the flat field effects are removed by subtracting one image from another at each signal level. For each intensity level, do the following:

1. Obtain 2 images in succession at the same light level. Call these images A and B.
2. Subtract the bias level from both images. Keep the exposure short so that the dark current is negligibly small. If the dark current is large, you should also remove it from both frames.
3. Measure the mean signal level S in a region of pixels on images A and B. Call these mean signals S_A and S_B. It is best if the bounds of the region change as little as possible from one image to the next. The region might be as small as 50x50 to 100x100 pixels but should not contain obvious defects such as cosmic ray hits, dead pixels, etc.
4. Calculate the ratio of the mean signal levels as $r = S_A/S_B$.
5. Multiply image B by the number r. This corrects image B to the same signal level as image A without affecting its noise structure or flat field variation.
6. Subtract image B from image A. The flat field effects present in both images should be cancelled to within the random errors.
7. Measure the standard deviation over the same pixel region you used in step 3. Square this number to get the Variance. In addition, divide the resulting variance by 2.0 to correct for the fact that the variance is doubled when you subtract one similar image from another.
8. Use the Signal from step 3 and the Variance from step 7 to add a data point to your Signal - Variance plot.
9. Change the light intensity and repeat steps 1 through 8.

After repeating these sets for an appropriate amount of times, you will have obtained the corrected data to determine the system gain with the help of formula 10.

3.4 The shack hartmann sensor

In a Shack-Hartmann sensor the aperture of a telescope is divided into sub-apertures by a microlens mask, also called lenslet array. In each sub-aperture the relative position of a star – and thus the local wavefront tilt – is measured. This allows the deformation of the light wave to be determined.

A microlens mask contains several small lenses, which are used to reconstruct the wavefront in the respective sub-aperture. These lenses are usually placed equidistantly in a two-dimensional lateral arrangement in the optical beam path, which creates a matrix of focal points. The displacement of each focal point is a measure of the tilt of the wavefront within the associated lens, and from the measurements of the individual displacements, the entire wavefront can be reconstructed using an approximation algorithm.
Each microlens creates a focal point at a distance of the focal length of the lens. A local tilt of the wavefront within the sub-aperture causes a shift of the focal point from position x_1 to x_2 by the amount Δx. In a first approximation this displacement is exactly proportional to the tilt of the wavefront across the microlens.

Figure 7 shows this displacement. A Shack-Hartmann wavefront sensor detects this displacement (gradient) and thus concludes the amount of tilt of the wavefront across the microlens. After appropriate detection of a two-dimensional matrix of focal point shifts, the complete wavefront can be reconstructed. Possible errors in the reconstruction of the wavefront tilt across a microlens arise from the uncertainties in the determination of the focal point shift. The determination of the shift is discussed in more detail below.

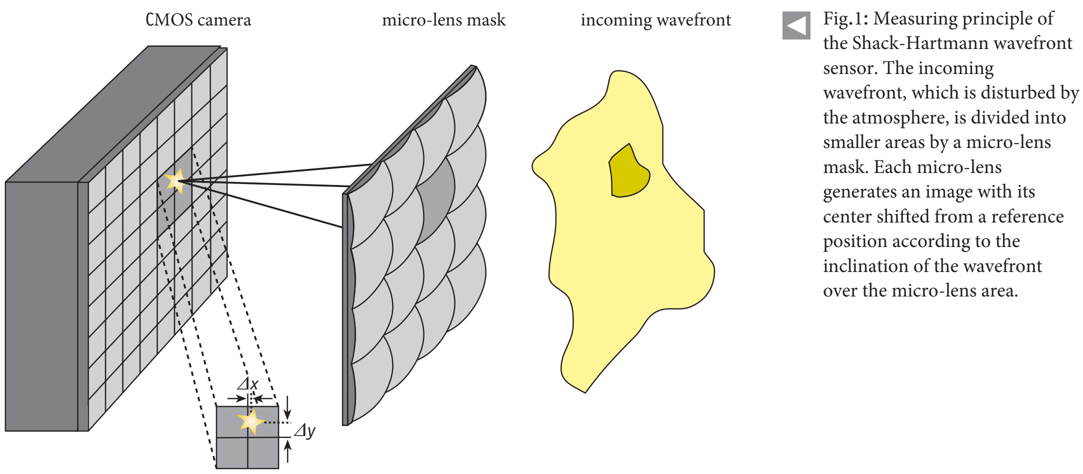

Figure 7: A non-planar wavefront falling on a microlens mask causes a shift of the light spots on the CMOS detector. (Credit: Hippler and Kasper (2004))

Applying simplifications of paralaxial optics, the following one-dimensional relationship can be defined:

$$\Delta x = f\Theta \tag{12}$$

Here Θ is the tilt angle in radians of the wavefront over a microlens, and Δx is the shift of the focal point on the detector (gradient).
In the experiment, a so-called relay optic is located between the lens array with the focal length and the CMOS detector. The relay optic images the focal plane of the microlenses onto the CMOS detector and additionally causes an optical magnification / reduction M. The microlens array with diameter d used in the lab course has 28 microlenses arranged in two rings (keystone arrangement). The horizontal distance between two adjacent microlenses is therefore equal to the diameter of a sub-aperture d_{sub}. With the pixel size of the CMOS detector, the magnification M of the relay optics can be determined.

For the optical path length difference OPD applies:

$$OPD = \Theta \cdot d_{sub} \tag{13}$$

This is the gradient of the wavefront over the sub-aperture. In a one-dimensional case, one can derive a relation between the measured shift of the focal point Δx and the optical path length difference OPD in units of the used wavelength λ:

$$\Delta x = f \cdot M \cdot \lambda \cdot \frac{OPD}{d_{sub}} \tag{14}$$

This allows the gradients of the wavefront to be determined and plotted from the measured shifts of the focal points.

For each sub-aperture i, the mean path length differences can be determined by

$$\frac{d_{sub}\Delta x_i}{fM\lambda} = OPD_{x_i} \quad \text{and} \quad \frac{d_{sub}\Delta y_i}{fM\lambda} = OPD_{y_i} \tag{15}$$

With this information, it is possible to reconstruct the wavefront across the entire microlens mask.

To detect the images of a Shack-Hartmann lens array, a CCD or CMOS camera is usually used. For the detection of the image of each individual sub-aperture, a square section of at least 2 x 2 pixels is required. To avoid optical crosstalk between the sub-apertures, a safety distance of at least one pixel width is inserted between the sub-apertures. According to this principle, some Shack-Hartmann systems work with 3 x 3 pixels per sub-aperture. A common alternative is 4 x 4 to 8 x 8 pixels per sub- aperture (for example at Starfire Optical Range or the adaptive optics system ALFA at the Calar Alto observatory). To increase the readout speed and reduce the readout noise, the 8 x 8 pixels can be "binned" to so-called super-pixels, which creates/simulates a smaller number of (larger) pixels.

3.5 Zernike polynomials

In general, a wavefront can be represented as a sum of higher order polynomials with respective coefficients. A representation frequently used in optics was developed by Frits Zernike in 1934 to describe the quality of concave mirrors. They come in handy, since they are made up of terms that are of the same form as the types of aberrations often observed in optical tests, such as $Z_{3,5}$ = astigmatism, Z_4 = defocus, or $Z_{7,8}$ = coma (see table in figure 8).
The Zernike representation of a wavefront W(x,y) is the sum of the products of these polynomials $Z_n(x, y)$ with the respective coefficients Cn

$$W(x, y) = \sum_{0}^{\infty} C_n Z_n(x, y) \tag{16}$$

Although widely used, this is not to say that Zernike polynomials are the best polynomials for fitting test data. Sometimes Zernike polynomials give a poor representation of the wavefront data. For example, Zernikes have little value when air turbulence is present. Likewise, fabrication errors in the single point diamond turning process cannot be represented using a reasonable number of terms in the Zernike polynomial. In the testing of conical optical elements, additional terms must be added to Zernike polynomials to accurately represent alignment errors. Thus, the blind use of Zernike polynomials to represent test results can lead to disastrous results.

Zernike polynomials are one of an infinite number of complete sets of polynomials in two variables, ρ and Θ, that are orthogonal in a continuous fashion over the interior of a unit circle. It is important to note that the Zernikes are orthogonal over a discrete set of data points within a unit circle.
Zernike polynomials have three properties that distinguish them from other sets of orthogonal polynomials. First, they have simple rotational symmetry properties that lead to a polynomial product of the form

$$r[\rho] \cdot g[\Theta] \tag{17}$$

where $g[\Theta]$ is a continious function that repeats self every 2π radians and satisfies the requirement that rotating the coordinate system by an angle α does not change the form of the polynomial. That is

$$g[\Theta + \alpha] = g[\Theta] \cdot g[\alpha] \tag{18}$$

The set of trigonometric functions

$$g[\Theta] = e^{\pm im\Theta} \tag{19}$$

where m is any positive integer or zero, meets these requirements. The second property of Zernike polynomials is that the radial function must be a polynomial in ρ of degree 2n and contain no power of ρ less than m. The third property is that $r[\rho]$ must be even if m is even, and odd if m is odd.

The radial polynomials can be derived as a special case of the Jacobi polynomials, and tabulated as $r[n, m, \rho]$. Their orthogonality and normalization properties are given by

$$\int_0^1 r[n, m, \rho] \cdot r[n', m, \rho]\rho d\rho = \frac{1}{2(n+1)}\delta[n - n'] \tag{20}$$

and

$$r[n, m, 1] = 1 \tag{21}$$

As stated above, $r[n, m, \rho]$ is a polynomial of order 2n and it can be written as

$$r[n, m, \rho] := \sum_{s=0}^{n-m} (-1)^s \frac{(2n - m - s)!}{s!(n - s)!(n - m - s)!} \rho^{2(n-s)-m} \tag{22}$$

In practice, the radial polynomials are combined with sines and cosines rather than with a complex exponential. It is convenient to write

$$rcos[n, m, \rho] := r[n, m, \rho] \cdot cos[m\Theta] \tag{23}$$

and

$$rsin[n, m, \rho] := r[n, m, \rho] \cdot sin[m\Theta] \tag{24}$$

The final Zernike polynomial series for the wavefront $OPD\Delta w$ can be written as

$$\Delta w[\rho, \Theta] := \bar{\Delta w} + \sum_{n=1}^{n_{max}} \left(a[n] \cdot r[n, 0, \rho] + \sum_{m=1}^{n} (b[n, m] \cdot rcos[n, m, \rho] + c[n, m] \cdot rsin[n, m, \rho]) \right) \tag{25}$$

The first few Zernike polynomials in cartesian coordinates are shown in figure 8.
Although widely used, this is not to say that Zernike polynomials are the best polynomials for fitting test data. Sometimes Zernike polynomials give a poor representation of the wavefront data.
To understand why that is the case, we have to look at how the polynomials are determined from the measurement in this lab course.
As a first step, the polynomials are scaled to the size of the sub-aperture over which the wavefront is measured. This scaling allows for the comparison of the measured wavefront gradients to the gradients of the Zernike polynomials. By doing so numerically, the coefficients which make up the wavefront over the circular aperture can be reconstructed.
An important keyword here is 'circular', since one of the main properties of Zernike polynomials is that they are orthogonal over the unit circle. However, the turbulence cells which function as sub-apertures in astronomical measurements are not circular, but randomly shaped. Thus, Zernike polynomials have little value when air turbulence is present, as they do not correctly represent the shape.
Likewise, fabrication errors in the single point diamond turning process cannot be represented using a reasonable number of terms in the Zernike polynomial. In the testing of conical optical elements, additional terms must be added to Zernike polynomials to accurately represent alignment errors. Thus, the blind use of Zernike polynomials to represent test results can lead to disastrous results.

$Z_j(x, y)$	n, m	$Z_n^m(x, y)$	Name
$Z_0(x, y)$	0, 0	1	piston
$Z_1(x, y)$	1, -1	y	tilt
$Z_2(x, y)$	1, 1	x	tip
$Z_3(x, y)$	2, -2	2xy	astigmatism
$Z_4(x, y)$	2, 0	$2x^2 + 2y^2 - 1$	defocus
$Z_5(x, y)$	2, 2	$x^2 - y^2$	astigmatism
$Z_6(x, y)$	3, -3	$3x^2y - y^3$	trefoil
$Z_7(x, y)$	3, -1	$3x^2y + 3y^3 - 2y$	coma
$Z_8(x, y)$	3, 1	$3x^3 + 3xy^2 - 2x$	coma
$Z_9(x, y)$	3, 3	$x^3 - 3xy^2$	trefoil
$Z_{10}(x, y)$	4, -4	$4x^3y - 4xy^3$	secondary astigmatism
$Z_{11}(x, y)$	4, -2	$8x^3y + 8xy^3 - 6xy$	secondary astigmatism
$Z_{12}(x, y)$	4, 0	$6x^4 + 12x^2y^2 + 6y^4 - 6x^2 - 6y^2 + 1$	spherical aberration

Figure 8: Some Zernike polynomials in Cartesian coordinates (Credit: (Schwiegerling, 2016))

4 Experimental prerequisites

4.1 Python

To evaluate the data measured with the CMOS camera and the Shack-Hartmann sensor, you will be given prewritten python code by your tutor, which has to be modified to analyse the date you measure in the lab. There are several packages incorporated in order to get the wanted results, which will be glossed over briefly. If you want to write your own code feel free to do so, although that is not necessary.

4.1.1 Numpy & Matplotlib

Since the python code is prewritten and the most basic knowledge is assumed these directions will not further go into the commands available with the libraries numpy and matplotlib. In case any of the commands and functions in the code are unclear to you, you can look them up here and here. For the zernike polynomials, mpl_toolkits is used to extend matplotlib to 3Daxes.

4.1.2 Glob

The glob module, which is short for global is a function used to search for files that match a specific file pattern or name. In your given python scripts it allows you to simply put the names of the files which need to be evaluated in your code, so that they can then be analyzed according to the task.

4.1.3 Astropy

Astropy is a module with a broad range of functions, specifially designed to help with common tasks in astronomy. It contains an ecosystem of interoperable packages, such as astropy.io.fits which provides access to FITS files. FITS stands for Flexible Image Transport System and is a flexible open data format developed by NASA, thus commmonly used in astronomy. Since the CMOS in the experiment produces FITS data it is needed to evaluate the images.
Astropy.stats contains statistical functions and algorithms used in astronomy. It does not replace scipy.stats, but rather adds functionalities according to the astronomers needs. In the experiment, the package 'sigma_clipped_stats' is used to find the centroids in the image which will b analyzed in task two. Further information on that can be found in the link above.

4.1.4 Photutils

Photutils is an affiliated Astropy package for detection and photometry of astronomical sources. The provided script to find the centroids of the microlens mask makes use of the modules datasets, DAOstarfinder, Circular Aperture and find_peaks. You can find more on that here.

Datasets gives easy access to load or make a few example datasets, Circular Aperture is used to create an aperture object from coordinates.

Modules like DAOstarfinder, find_peaks are part of photutils.detection. DAOstarfinder searches images for local density maxima that have a peak amplitude greater than threshold, while find_peaks is a function to identify local peaks in an image that are above a specified threshold value. Peaks are the local maxima above a specified threshold that are separated by a specified minimum number of pixels, defined by a box size or a local footprint.

4.2 Navigating a terminal and asicap

The software used to readout the CMOS camera is called **asicap**. Detailed instructions on how to operate it can be found in section 5.1.2. To navigate to the software and use all python scripts you will use the terminal. In case you have never used one or are just not familiar with the commands, the basic ones necessary for this lab course will be briefly discussed.
To navigate through your folders in the terminal, you use the command '*cd*', which means change directory, followed by the pathname of the folder you want to move into. In the experiment, your general folder where you'll be working from can be opened by typing:

cd home/fprakt/F36/supervisor/your-folder/F36examples

If you want to see the folders and files inside the directory which you are currently in, simply type '*ls*', which stands for list, in the terminal. This will give you a list containing all possible paths. You can then go 'down' into a folder:

```
cd lower_folder_name
```

If you want to move back 'up' one 'step', you can type two dots:

```
cd ..
```

To move back to home, just put a tilde:

```
cd ~
```

Alternatively, you can always navigate to the folder you need to work in through the graphical interface, right click on it and chose 'Open terminal here'.

To run a python script in the terminal, you will need to use the following command:

```
python your_script.py
```

Before you run your scripts you have to make sure that all data called in the script is located in the folder where your script lies, which you are currently in. You can always check by using the *'ls'* command.

4.3 Experimental setup

Many measurements made with the CMOS camera should be done in the dark because the camera reacts very sensitively to all light levels. As light sources a diode laser and a special LED flat-field lamp are available. The supervisor will tell you how these are controlled and used.

The setup for the first part of the experiment is shown in figure 10. It only consists of the flat field lamp and the CMOS camera on an optical bench.

We use a CMOS camera of the type ZWO ASI183 MM. This camera comes with an IMX183CLK CMOS detector from Sony with 5640 x 3694 active pixels. Each camera pixel has a size of 2.4 μm x 2.4 μm.

The Shack-Hartmann wavefront sensor to be set-up in the second part of the experiment makes uses of the microlens mask as its central optical element (see 3.4)

In the second part of the experiment, the Shack-Hartmann wavefront sensor to be set-up in the experiment has a central optical element; a microlens mask (so called lenslet array).

This is set up on an optical bench, alongside a collimator to create coherent light and a relay lens to focus the focal points of the microlens masks on the CMOS camera, as well as an astigmatism lens to create a deformation of the wavefront. The setup for the measurement without the lens can be seen in figure 11.

The relay optic is an Apo-Rodagon-N f/2.8 objective with 50mm focal length. The focal length of the lens array is $f = 45$mm.
The microlens array used in the lab course has a diameter of 5mm and has 28 microlenses arranged in two rings (keystone arrangement). The horizontal distance between two adjacent microlenses is therefore approximately equal to the diameter of a sub-aperture (approximately 755 microns). One pixel of the CMOS detector has a size of 2.4μm x 2.4μm.

Figure 9: Your place of work

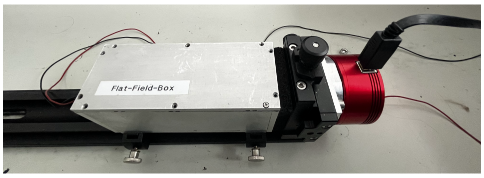

Figure 10: Setup for the measurement of the system gain

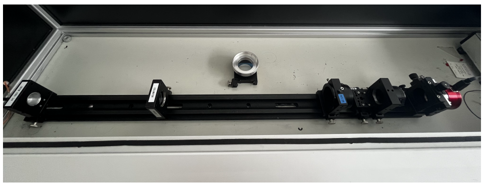

Figure 11: Setup for the measurement of the Shack-Hartmann gradients

5 The experiment

5.1 Part 1 - Determining system gain

5.1.1 Getting started

Start by taking care of the experimental setup. Open the lid of the box and make sure that all components are placed correctly. Put the flat field lamp right in front of the CMOS-camera as shown in figure 10, so that no additional light is being captured by the camera. Before turning on the power supply make sure that the pins on the flat field lamp are connected properly and in right order. The longer pin should be connected to the positive (red) cable and the shorter pin should be connected to the negative (black) cable. **Connecting them the other way around will break the LED.**

Now operate the flat-field lamp at a voltage of about 2,2 - 2,4 V and max. 25 mA. Make sure that you **do not go over 5V**, as this can damage the LED. When you are satisfied with the setup, close the lid and get to the computer.

The camera can be operated with different programs. The program of the camera manufacturer is called **asicap**. Another method to readout the camera is based on the open source software **INDI** (https://- indilib.org). In combination with a python interface, python3 scripts can be used to read the camera.
In this experiment however we will use **asicap**, as it has the advantage of a graphical interface which gives you easy control over the CMOS camera.

As a first step, you should create a folder for all your data to go in, and also for you to modify the given python scripts. To do so, open the *'Home'* folder on the desktop, followed by the *'F36'* folder and find your supervisors name. Open it and create a new folder called:
'number_your-name_partners_name'. Then take the *'F36examples'* folder and copy it into the folder with your names. You can delete the data in the 'system gain' folder, as it is old data, and use the folder to store your own measurements. Afterwards, either right-click and chose 'open terminal here', or change the directory from the terminal as explained in section 4.2.

5.1.2 Operating the CMOS camera

To operate the software open a second terminal, type 'asicap' into the command line and hit enter. The interface of the software should open after a few moments. Under 'Capture' you can change the path so that future data is saved right to the folder you just created.

First take care of the settings in the program. Set 'Format' to RAW16 and select the exposure time such that detector in the histogram is not saturated (< 65000 ADU). You should find a value somewhere between $50 - 90$ms.

Set the programmable gain to 12dB (note that this corresponds to 120 in the asicap GUI). The amount of images you want to take can be chosen under 'Limit'. Measurements are taken by clicking on the camera symbol.

Shortly investigate the effect of pixel binning on images taken with the same setup. For the following steps the binning should be set to 4. Why are we using it in this experiment? Discuss with your tutor.

5.1.3 Recording the photon transfer curve

The measurement of the photon transfer curve will go according to the steps described in the theoretical part in section 3.3.4. Take two pictures at the almost saturated light intensity (max. 25mA) which you set earlier. Then set a lower light intensity (1 - 2 mA less) and take two pictures again. The exposure time should remain the same throughout all of those measurements. Repeat this as often as necessary and with appropriate light levels such that in the photon transfer curve one can later see the saturation effect as well as the linearity of the Poisson noise and the effect of the system noise. Consider beforehand whether a dark current and offset/bias correction should be taken into account. Section 3.3.2 should help you with that decision. Discuss with your tutor and measure the correction image.

Next, open up the script **'gainSH_v1.py'** in your folder and get familiar with what it does. Section 3.3.4 explains how to calculate the photon transfer curve, how is that implemented in the code? What part of the code do you have to modify so it reads your data? Exchange the according code and run the script in the terminal.
If all went well, you will get the photon transfer curve, i.e. the mean values against the variance.
How can you now determine the system gain? Calculate it in electrons per ADU and display it in the plot. Explain the non-linear parts of the curve, if available.

5.2 Part 2a - Shack-Hartmann gradients

For the measurement of the Shack-Hartmann gradients a new setup is needed. Remove the flat field lamp from the optical bench and place the collimator in front of the laser. Make sure to put it at the right distance, you can find the focal length on the side of collimator. Next put the microlens mask between the collimator and the camera. Between the microlens mask and the camera comes the relay lens. Your setup should look like figure 11. What do the individual components do? Turn the small key in the laser in order to switch it on.

Warning: do not look directly into the laser light with your eyes.

Get to the software and have a look at the new picture. You should see 28 points, with 2 circles containing 11 and 18 dots. Reduce the exposure time to the lowest possible setting, such that the points are barely visible. This will make sure that the points are not oversaturated, and will not be counted double when the code looks for the centroids. Now you will have to adjust the position of the microlens mask and the relay lens so the dots look as sharp as possible and the circles have a good size. When you are done with that your image should somewhat look like figure 12.

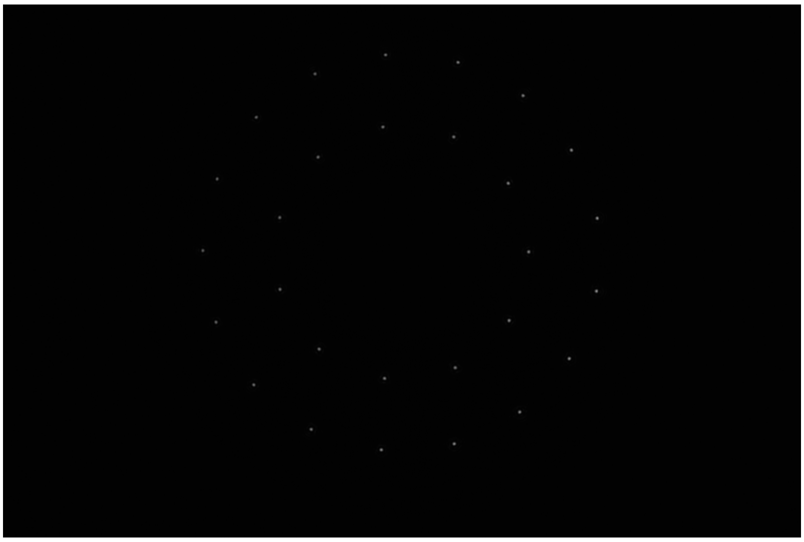

Figure 12: Laserlight through a microlens mask without any aberration. Note that the points really are barely visible (Image credit: Noa Bergmann)

5.2.1 Finding the centroids

As soon as you are satisfied with all the adjustements, take an image using the software. This will be your reference photo to determine the Shack-Hartmann gradients later.
Open the script **'find_centroids_and_save_results.py'** in **'F36_task2_scripts'** in your folder and get familiar with it. Modify it so it reads and saves your data.
Next, change the directory to the right folder in the terminal and run the script. Did the code find all 28 points and circle them? If not, one or more points might have been counted multiple times. To solve this problem, have a look at the parameters defined under 'adapt to set-up'. What do they do? Vary them and see if you can get better results.

After you managed to find all 28 points, move on to the next measurement and place the astigmatism lens between the collimator and the microlens mask. Does its exact position on the optical bench make a difference? Repeat the imaging process for the aberration.

5.2.2 Displaying the gradients

As soon as you found all 28 centroids for the reference and eyeglass setup you can get to determining the gradients. To do so, open the script **'display_gradients.py'** and try to understand what it does. Yet again, modify it to read your data and run it in the terminal. Now have a look at the displayed gradients. What do they tell you?

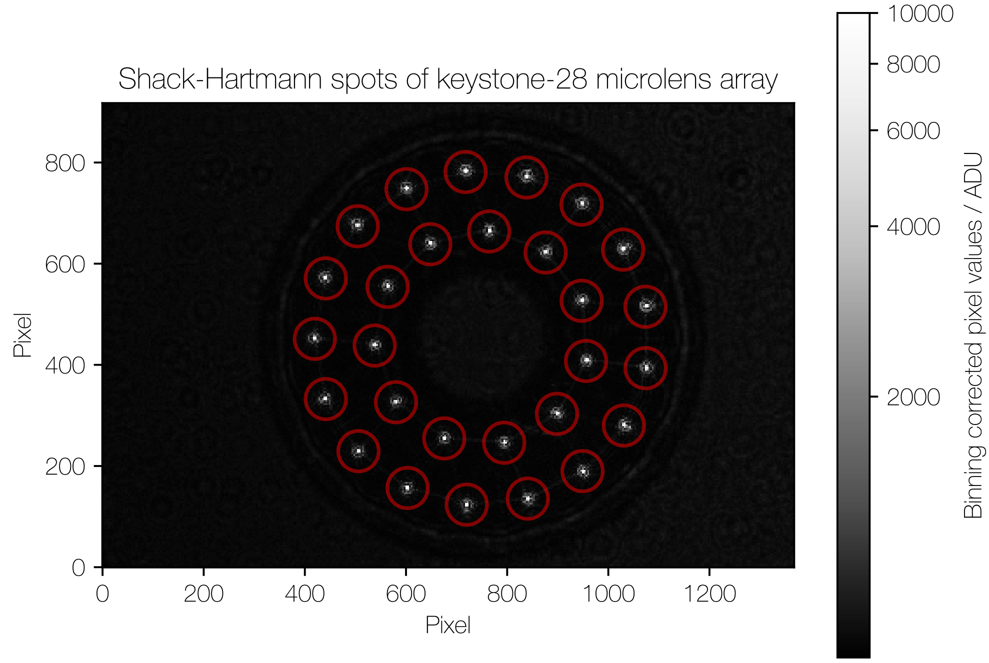

Figure 13: Reference image without aberrations, all 28 centroids have been found

Incase your plot looks like figure 15 try to rotate the astigmatism lens. The algorithm sorts the coordinates of the centroids by y-coordinate, so depending on the direction of the gradients there might be mixups.

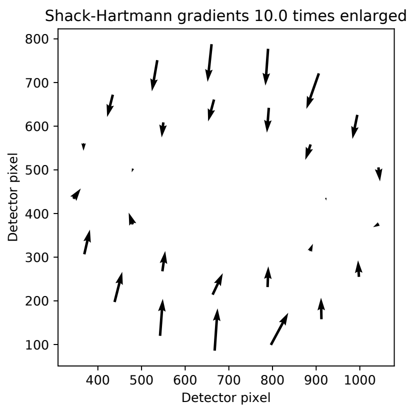

Figure 14: Successful plot of gradients

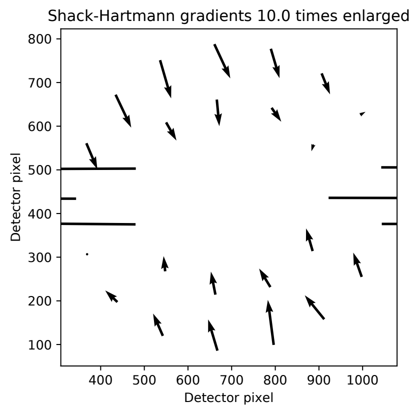

Figure 15: Unsuccessful plot of gradients

If you do not plan on doing the long report, you can now clean up your workspace, turn off the laser and call your tutor to briefly discuss your findings. **Do not turn off the computer.**

5.3 Part 2b - Zernike Polynomials

This part of the experiment is only required for the long report - however you are very welcome to do it voluntarily if you'd like.
Go to the folder **'F36_task2_zernike_fit_script/cart_zernike/tests'** and open the file **'f36_tests.py'**. Once again have a look at the code and modify it according to your data. Do not forget to change directory in the terminal and run the script. Have a look at your results. What do you expect aberration from an astigmatism lens to look like? In the output from the script, you will see two colored plots next to each other. The yellow/green/blue background shows the reconstructed wavefront from Zernike polynomials. The plot on the right, shows the reconstructed Zernike coefficients in units of microns. Do these coefficients to the different polynomials match your expectations?

6 Report instruction

Since all of the data analysis for the experiment happened in the lab, your reports will mostly consist of summarizing the theories and discussing your results and possible errors. If you do not chose to do the long or short report then you are done with this experiment.

6.1 Short report

Your short report should include a short introduction or an abstract as well as a brief theoretical background to understand your findings. This includes but is not restricted to a chapter to describe the readout scheme of the used CMOS imager. You should further summarize and discuss your results from the determination of the system gain and the Shack-Hartmann gradients. Include plots and graphics where you see fit.
Research about astigmatism and shortly explain why this aberration results in the found gradients. Add a chapter about possible errors in the determination of the gradients and draw a conclusion on the lab course.

6.2 Long report

The long report has to contain all the elements of the short report. Add a concise section on Zernike Polynomials to your theory chapter. In addition to the description of the readout scheme of the CMOS, add a detailed error discussion regarding the determination of the system gain in the first task. Further, a detailed wavefront reconstruction of the wavefront has to be included, e.g. the fit of Zernicke polynomials to the gradients which you did in the lab. Decide whether the found coefficients are sensible or not. Discuss your results as well as possible errors in the determination of the Shack-Hartmann gradients.

References

Michael Newberry. Pixel response effects on ccd camera gain calibration. 1998. URL http://www.phy.cuhk.edu.hk/djwang/teachlab/projects/CCD/CCD%20Camera%20Gain%20Measurement.pdf.

James C. Wyatt. Zernike polynomials. 2003. URL http://www.phy.cuhk.edu.hk/djwang/teachlab/projects/CCD/CCD%20Camera%20Gain%20Measurement.pdf.

Stefan Hippler and Markus Kasper. Adaptive optik - sieg über die luftunruhe. *Sterne & Weltraum*, (10):32–42, 2004.

Alexander Stohr. Schematic working principle of an active pixel sensor. *Wikipedia*. URL https://commons.wikimedia.org/w/index.php?curid=3643538.

Jim Schwiegerling. Description of zernike polynomials. 2016. URL https://wp.optics.arizona.edu/visualopticslab/wp-content/uploads/sites/52/2016/08/Zernike-Notes-15Jan2016.pdf.

Appendix

4. Camera technical specifications

Sensor	1″ CMOS IMX183CLK-J/CQJ-J
Diagonal	15.9mm
Resolution	20.18Mega Pixels 5496*3672
Pixel Size	2.4µm
Image area	13.2mm*8.8mm
Max FPS at full resolution	19FPS
Shutter	Rolling shutter
Exposure Range	32µs-2000s
Read Noise	1.6e @30db gain
QE peak	84%
Full well	15ke
ADC	12 bit or 10 bit
DDR3 buffer	256MB(Cooled)
Interface	USB3.0/USB2.0
Adapters	2″ / 1.25″ / M42X0.75 Uncooled 2″ / M42X0.75 Cooled
Protect window	AR window
Dimensions	Uncooled 62mm/Cooled 78mm
Weight	Uncooled 140g/Cooled 410g
Back Focus Distance	6.5mm
Cooling:	Regulated Two Stage TEC
Delta T	40°C -45°C below ambient
Cooling Power consumption	12V at 3A Max
Supported OS	Windows, Linux & Mac OSX
Working Temperature	-5°C—45°C
Storage Temperature	-20°C—60°
Working Relative Humidity	20%—80%
Storage Relative Humidity	20%—95%