

Dissertation
submitted to the
Combined Faculties of the Natural Sciences and Mathematics
of the Ruperto-Carola-University of Heidelberg, Germany
for the degree of
Doctor of Natural Sciences

Put forward by
Christian Schwab
born in: Hanau
Oral examination: 16th July 2010

A Calibration System for a Rayleigh
Laser Guide Star Constellation
&
A Planet around a Giant Star

Referees:

Prof. Dr. Andreas Quirrenbach
Prof. Dr. Joachim Krautter

Für meinen Vater Stephan,
dessen Wunsch das war.

Abstract

My thesis consists of two parts. In Part I, I describe my contributions to the laser guide star adaptive optics facility at the Large Binocular Telescope, the *Advanced Rayleigh guided Ground layer adaptive Optics System*, ARGOS. I investigate the effect of scattered light from propagating the laser light across the aperture, and find that contamination is not a concern for the science instruments. I present a study of reflective laser launch telescopes based on commercial optics, and compare this to the refractive design chosen as baseline for ARGOS. The discussed options present an interesting alternative with only small additional light loss at substantially reduced cost and procurement risk. I develop a calibration scheme for the full adaptive optics system, based on artificial light sources that illuminate the deformable mirror, imitating the laser guide star beacons. This enables the interaction between deformable mirror and wavefront sensor to be calibrated at any time, greatly enhancing the possibilities and time available for engineering on the installed system. The light source has to be placed in the prime focus, as ARGOS uses the adaptive secondary mirrors of LBT. The optical design of the calibration light source is not trivial as the guide star constellation comprises of three beacons forming an equilateral triangle on a circle with a radius of 2 arcminutes. The images of the beacons that have to be reproduced suffer from strong aberrations caused by the large off-axis distance. To match the wavefront of the beacons' images to the desired precision of 50 nanometres rms, I designed a custom objective, incorporating a computer generated hologram to shape the wavefronts of three optical fibres forming the light sources. The elliptical front surface of the objective is used in reflection to generate a central, diffraction limited spot serving as an alignment aid and truth sensor for the measurements. A thorough tolerance analysis including the assembly and the alignment at the telescope ensures that the design specifications can be met during operation.

The second part of the thesis concerns the search for planets around a sample of nearby giant stars with the Doppler technique carried out at Lick Observatory's CAT telescope. I analyse new data from the previous three years, which together with the existing data from our survey form a database covering 11 years. The radial velocity measurements reveal the presence of a planetary companion around one of our target stars, which was previously unknown. This is the lightest planet found around a giant star to date with a minimum mass of 1.92 Jupiter masses in a 785 day orbit.

Zusammenfassung

Die vorliegende Arbeit setzt sich aus zwei Teilen zusammen. Im ersten Teil erläutere ich meine Beiträge zu ARGOS, dem *Advanced Rayleigh guided Ground layer adaptive Optics System*, einer Laserleitstern-gestützten adaptiven Optik (AO) am Large Binocular Telescope (LBT). Ich untersuche die Auswirkungen des Projizierens eines Laserstrahls quer über die Teleskopöffnung und folgere, daß das gestreute Licht kein Problem für die wissenschaftlichen Instrumente darstellt. Desweiteren stelle ich Entwürfe für reflektive Laserprojektionsteleskope basierend auf kommerziell erhältlichen Amateurteleskopen vor. Die erläuterten Optionen stellen eine interessante Alternative da, die im Ausgleich für einen kleinen zusätzlichen Lichtverlust einen deutlichen günstigeren Preis und vermindertes Beschaffungsrisiko aufweist. Weiterhin entwickle ich ein Kalibrationsschema für das komplette AO System, welches auf künstlichen Lichtquellen basiert, die den deformierbaren Spiegel (DM) beleuchten und die Laserleitsterne imitieren. Dies ermöglicht, die gegenseitige Abhängigkeit zwischen deformierbarem Spiegel und Wellenfrontsensor jederzeit zu rekalisieren, und vervielfacht so die Möglichkeiten und die verfügbare Zeit, um das System zu optimieren. Die Lichtquelle muß im Primärfokus untergebracht werden, da ARGOS die adaptiven Sekundärspiegel des LBT benutzt. Das optische Design der Kalibrationslichtquelle ist nicht trivial, da die Konstellation der Laserleitsterne aus drei Leitsternen gebildet wird, die ein gleichseitiges Dreieck auf einem Kreis mit zwei Bogenminuten Radius formen. Die Bilder der Laserleitsterne, die reproduziert werden müssen, sind durch die wegen des großen Abstandes von der optische Achse bedeutenden Aberrationen stark deformiert. Um ihre Wellenfronten mit der geforderten Genauigkeit von 50 nm rms nachzubilden, habe ich ein spezielles Objektiv konstruiert, das ein computergeneriertes Hologramm enthält, um die von drei optischen Fasern ausgehenden Wellenfronten der gewünschten Form anzupassen. Die Frontfläche des Objektivs ist elliptisch und wird in Reflexion benutzt, um einen zentralen, beugungsbegrenzten Lichtpunkt zu erzeugen, der zum Justieren des Objektivs sowie als Referenzquelle zum Überprüfen der erreichten Leistungsfähigkeit des AO Systems dient. Eine ausführliche Toleranzanalyse, die den Zusammenbau und die Justage am Teleskop einschließt, stellt sicher, daß die Spezifikationen auch im Betrieb erreicht werden.

Der zweite Teil der Arbeit befaßt sich mit der Suche nach extrasolaren Planeten um eine Auswahl von nahen Riesensternen, die mit der Dopplermethode am Lick Observatorium durchgeführt wird. Ich werte neue Daten der letzten drei Jahre aus, die zusammen mit den schon existierenden Daten unserer Durchmusterung einen Zeitraum von elf Jahren umspannen. Die Radialgeschwindigkeitsmessungen enthüllen die Existenz eines vorher unentdeckten planetarischen Begleiters um einen der von uns untersuchten Sterne. Dies ist der bisher leichteste Planet eines Riesensternes, mit einer minimalen Masse von 1.92 Jupitermassen in einem 785 Tage dauernden Orbit.

Contents

Part I: The ARGOS Calibration Unit.....	1
1 Introduction.....	3
1.1 Theoretical framework.....	7
1.1.1 Atmospheric turbulence.....	7
1.1.2 Isoplanatic angle.....	10
1.1.3 The atmospheric time constant.....	11
1.1.4 The cone effect.....	11
1.1.5 Performance simulations for GLAO.....	13
1.1.6 Spot elongation.....	14
1.2 Overview over existing LGS facilities.....	16
1.2.1 The William Herschel Telescope (WHT).....	16
1.2.2 The Multi Mirror Telescope (MMT).....	18
1.2.3 The Southern Astrophysical Research Telescope (SOAR).....	21
1.2.4 Sodium Laser Guide Star facilities.....	22
1.2.5 Summary.....	26
1.3 Description of LBT and Lucifer.....	30
1.4 Description of ARGOS project.....	31
1.5 Overview of Part I of the thesis.....	35
2 Technical contributions.....	36
2.1 Scattered light impact from open air propagation.....	36
2.2 Laser launch telescope design.....	39
2.2.1 Introduction.....	39
2.2.2 Overview of options.....	39
2.2.3 Throughput calculation.....	39
2.2.4 Options based on commercial telescopes.....	41
2.2.5 Optical designs.....	43
2.2.6 Comparison with the refractor design.....	49
2.2.7 Discussion and conclusions.....	52
3 The ARGOS calibration scheme.....	55
3.1 Introduction.....	55
3.2 Calibration strategy.....	56
3.3 Top level requirements.....	57
3.4 Investigated optical solutions.....	59
3.4.1 First Light AO calibration scheme and hardware.....	59
3.4.2 On-axis double pass with periscopes.....	64
3.4.3 Off-axis double pass.....	67
3.4.4 Off-axis single pass.....	68
3.4.5 Discussion and comparison of all options.....	77
3.5 Calibration Unit Design.....	78
3.5.1 PDR Design.....	80
3.5.2 FDR design.....	89
3.5.3 On-axis Source.....	100
3.6 Alignment on the telescope.....	104
3.7 Conclusions and outlook.....	109

Part II: A planetary companion around HIP31592	111
4 Historical introduction	113
5 Detection methods	117
5.1 Orbital elements	117
5.2 Astrometry.....	119
5.3 Microlensing.....	119
5.4 Transits	122
5.5 Direct imaging.....	122
5.6 Doppler searches	125
5.6.1 Instrumentational challenges	127
6 Data analysis	130
6.1 Data acquisition and extraction	130
6.2 Doppler shift measurement	130
6.3 Doppler code	132
7 A new planet around HIP31592.....	133
7.1 Known properties of the host star	134
7.2 Observations.....	135
7.3 Orbital solution.....	137
7.4 Astrometric signature	142
7.5 Comparison to known systems.....	143
7.6 Conclusions	145

Part I: The ARGOS Calibration Unit

1 Introduction

The night sky has always held fascination for humankind. For thousands of years we have observed the sky more or less systematically. From early times on, instruments were invented to make the observations more precise. Specially designed buildings and structures consisting of stones in specific positions, for example, were set up to measure the positions of the stars on the sky, initially to regulate the yearly cycles and later to map the sky. The single most important invention in observational astronomy was the telescope by Hans Lipperhey in 1608. Since then telescopes have developed from being mere magnifying glasses to enormous instruments with which we explore the Universe to depths unimagined by the earliest Astronomers.

Most observations taken on Earth have been impeded by the effects of the atmosphere on the light from the object before it reaches the observer. The twinkling of the stars as seen by the naked eye, an effect we are all familiar with, distorts images of stars to blurred, washed-out specs of light on big telescopes. Skilled visual observers, objective scientists, have been fooled into believing in artificial phenomena, due to atmospheric turbulence combined with our brains' age old training to recognise structures. The most prominent example of this may be discovery of the “canali” on Mars by Schiaparelli in 1877 (see Figure 1) and the numerous arguments between those who verified and contradicted the claims thereafter (Figure 2), an argument that went on for nearly 20 years.

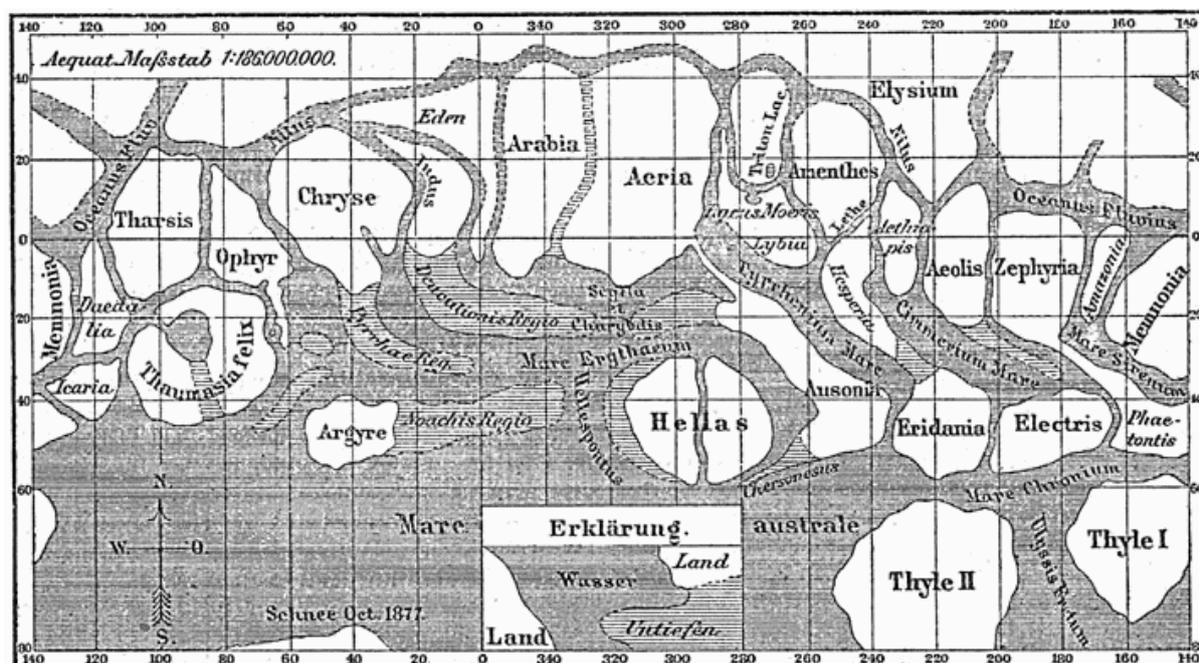


Figure 1: A map of Mars' surface with the “canali” drawn by Giovanni Schiaparelli at the Milan Observatory in 1877 (from Meyers Konversationslexikon 1888).

Every optical telescope above a very moderate size of approximately 15 cm is limited in its resolving power because of atmospheric turbulence in the last 10 km of distance the starlight travels. After a long period of scientific discoveries based primarily on improvements in the light-gathering power of telescopes, we are now in an era where the resolving power of the telescope is greatly improving and playing an increasing role. The emergence of two technologies has allowed us to overcome the adverse effects of the atmosphere: by operating telescopes in space we can avoid the atmosphere completely, while adaptive optics (AO)

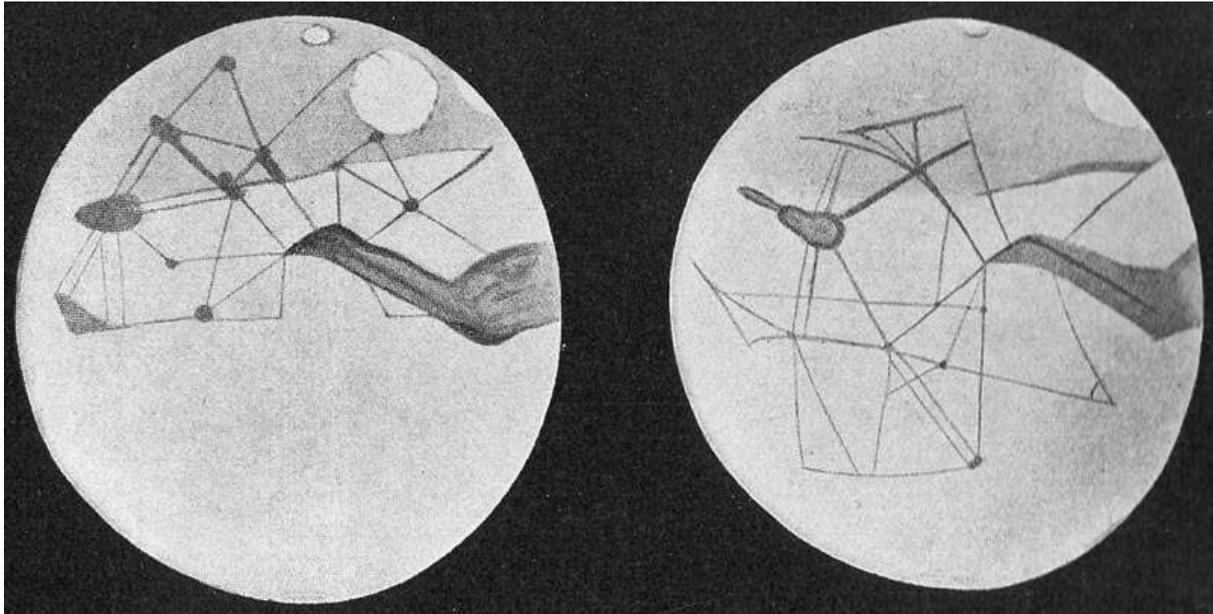


Figure 2: The “canali” as Percival Lowell perceived them through the 24” Clark Refractor at Lowell Observatory in Flagstaff, Arizona (Perelman 1914).

systems on ground-based telescopes cancel some of the distortions introduced by the atmosphere and sharpen the images dramatically. The Hubble Space Telescope (HST) has been a very successful facility - its unprecedented resolution opened a whole new window for scientific exploration. Satellite observatories are very costly to operate, however, and exceedingly hard to maintain. Ground-based telescopes of the 10-metre class equipped with AO are now able to deliver sharper images than HST, at least in the near infrared. Naturally, ground-based observatories are much more versatile and accessible for keeping instruments up to date. While satellite observations remain important for some scientific cases, many programs can be conducted more efficiently from the ground.

AO technology has grown and become so widespread that in today’s scientific landscape it has become an indispensable tool for exploring the full potential of any major observational facility. This can be seen on the steadily growing number of AO related publications, for example. Large observatories are devoting a substantial part of the observing time to AO guided science programs. As telescopes become larger, the potential gain of AO corrected images grows too. This is due to the relation between aperture diameter D and peak intensity I of the image of a point source. For seeing-limited observations, the image of a point source has a constant angular diameter, and the intensity grows with the collecting area of the telescope, $I \propto D^2$, while for diffraction-limited observations, the diameter of the image is inversely proportional to the diameter of the telescope, and hence $I \propto D^4$.

Adaptive optics was proposed as early as the 1950s, by Horace Babcock, an astronomer at Mt. Wilson and Palomar Observatories (Babcock 1953). The available technology only became powerful enough for a first attempt to realize such a system in the late 60s. The early systems were developed for military and aerospace applications, and progress was driven by these communities, who had considerable resources available to them. Early milestones in the advancement of AO systems were the development of the first functioning deformable mirrors, and subsequent experiments with natural guide stars, by John Hardy and co-workers at ITEK company in 1972, and the atmospheric compensation achieved with the

Compensated Imaging System on Mt. Maleakala, Hawaii, using stars and satellites as guide stars during the years 1982 -85 (Hardy 1998).

The first AO systems used the light of a bright star (termed natural guide star, NGS) to measure deformations in the wavefront caused by atmospheric seeing. Advancements in this field were coupled to progress in key technologies, most importantly the availability of deformable mirrors with many actuators and accurate response, low noise, high fidelity wavefront sensors and fast computers. Each of these areas has seen astounding growth over the last few decades. These technologies are now so well-developed that the main limitation is the availability of a sufficiently bright guide star in close proximity to the object of interest. The fraction of the sky containing bright enough stars that the uncertainty of the wavefront measurement induced by the photon noise is small enough to permit seeing correction to the diffraction limit is only a few percent.

To overcome this obstacle, artificial guide stars were conceived. Powerful lasers are employed to generate a beacon in the upper atmosphere. The light propagated downward is used to measure the atmospheric wavefront distortion. The two possible methods for creating laser guide stars, Rayleigh and Sodium lasers, were developed in the early 1980s. The use of Rayleigh scattering based systems had been first proposed by Feinlieb (AOA) and Hutchin (ITEK). A similar idea was proposed by Hunter of Western Res. Corp. To use lasers tuned to the sodium line has been proposed by W. Happer in 1982 (Happer et al 1994). Both methods were investigated experimentally in parallel, the Rayleigh guide star by Fugate et al. (Air Force Weapons Laboratory/ Starfire Optical Range, SOR, in New Mexico) (Fugate et al. 1991), and the sodium approach in White Sands, close to the SOR site, by Humphreys and Primmerman (MIT/Lincoln Labs) (Primmerman et al. 1991). The Rayleigh system went on sky in 1983, followed by the Sodium laser in 1984.

All the references concerning these systems are from the early 1990s, as up until then the research was classified and could not be made public. The technology became declassified in 1991 and the wealth of information collected with the military systems became available for wider use. The first AO systems for astronomical use came online soon after the release of this information.

Sodium laser guide stars

The Sodium laser technology was adopted for the first astronomical laser guide star facilities, and has become the choice for major observatories around the world. The functioning principle is based on the excitation of Sodium that has been deployed in very high layers of the atmosphere, by the ablation of meteors. Typical altitudes are around 92 km. The laser must be tuned to the Sodium D2a-line at 589 nm wavelength. The large absorption cross section makes it possible to generate a relatively bright spot by exciting a column of Sodium atoms with a laser focussed on this altitude. The laser beacon is high enough to be well above any atmospheric turbulence, allowing the full vertical extend of the atmosphere to be sampled. A drawback of sodium layer beacons is the fact that the necessary lasers are complicated systems which are costly and hard to maintain. Different technologies to generate laser light at 589 nm have been employed more or less successfully. The first systems were based on dye lasers, which need a lot of maintenance. The handling of the dye, typically Rhodamin, is particularly messy. Solid state lasers would be more favourable, but the common laser materials have no transitions at the desired wavelength. However, by a quirk of nature, the sum frequency of the 1st and 2nd laser transition of neodymium doped

yttrium aluminium garnet (Nd:YAG), the most common laser medium today, matches the sodium D line. Nonlinear sum frequency conversion is used to achieve 589 nm laser emission from Nd:YAG lasers. Unfortunately, these systems are highly customized, and very expensive. Commercial systems with sufficient output power became available only recently, for the price of several million USD per unit. A third avenue to explore is fibre lasers. Their inherent advantage is the excellent beam quality and that they can deliver the output at a convenient location. However, the achieved output powers do not yet match the promised values.

Rayleigh laser guide stars

Rayleigh Laser Guide Star (RLGS) Systems have been investigated continuously for more than 25 years (Fugate et al. 1991) as an alternative to sodium laser guide star systems. These systems are based on the Rayleigh scattering of photons on air molecules. Some photons of a light beam propagating upwards in the atmosphere are scattered backwards at each layer. The percentage of scattered light is linearly dependent on the density of the air. Due to the exponential drop in density of the atmosphere, quite powerful lasers are needed to achieve substantial scattered flux at high beacon altitudes, even though low power lasers, like laser pointers, may produce visible scatter at sea level altitudes. An altitude of about 25 km is the limit at which backscatter can be detected and used to measure the atmospheric turbulence. Rayleigh scattering is further dependent on the wavelength of the light. The Rayleigh scattering cross-section is proportional to $1/\lambda^4$ (for this and calculations of the returned photon flux, see van der Hulst 1981). This is the well-known reason for the blue daylight sky, or the perceived colour of cigarette smoke, for example. As the Rayleigh scattering takes place along the whole path through the atmosphere, these systems need to be pulsed and range gated to achieve an approximately point-like source. Effectively, an exposure is taken while the pulse travels only a short distance at the desired altitude. Obviously very fast shutters are needed as well as good synchronisation with the firing of the laser pulse.

A great advantage of the Rayleigh guide star principle is that one does not need a particular wavelength. Hence, laser systems developed for other applications can be used with only slight modifications, if any at all. At the moment, the most widespread laser systems are solid state lasers with Nd:YAG as the lasing medium. These lasers are robust, have a good power to price ratio, and deliver excellent beam quality. The natural wavelength of the first laser transition in Nd:YAG is 1064 nm. Units with twice or even three times the frequency are common, too. The light is converted by nonlinear frequency conversion, typically in nonlinear crystals.

Ground layer adaptive optics

We are currently witnessing a renaissance of Rayleigh Laser Guide star systems, possibly related to the progress of commercial lasers but also because the advantages of a concept called Ground Layer Adaptive Optics have been realized (GLAO, Rigaut 2002).

Ground Layer Adaptive Optics is a concept that aims to correct only for the turbulence occurring at low altitudes, in the so-called ground layer, up to about 1.5km above the telescope. The efficiency of this method was recognized with the realization of the relative strength of the turbulence in this atmospheric layer. Up to 80 % of the phase distortions have been measured to originate in the ground layer. As the ground layer is close to the telescope, the effect of angular anisoplanatism is less pronounced than in higher layers. In the idealized case that the turbulent layer is directly on top of the aperture, the AO system would correct for the whole field of view of the telescope at once. Of course, the image quality cannot be

diffraction limited, or even approach the level a natural guide star system achieves in the direction of the guide star, as some turbulence in the higher layers remains uncorrected or even unsampled by the AO system. A GLAO system therefore has been called a ‘seeing-improver’, as it reduces the seeing over a wide field of view without producing diffraction limited cores in the stellar images. More than 90% of all ground based astronomy is surely done in seeing-limited mode. For all these programs GLAO can improve resolution, but more importantly, raise efficiency by a substantial amount. Spectrographs with small slits especially benefit from the rise in encircled energy. This gain is provided over a large field of view, ideal for multi-object units.

1.1 Theoretical framework

1.1.1 Atmospheric turbulence

In this section I give a short overview of the theoretical background for the generation of turbulence in the atmosphere, and the wavefront deformation this generates, based on Quirrenbach (2006). Many detailed treatises can be found in the literature, see for example (Roddier 81, Fried 94, Hardy 98). To assess the turbulent behaviour, one first needs to characterize the medium under investigation. Based on the theory of fluid flows, we use the Reynolds number (Stokes 1851, Reynolds & Osborne 1883), defined as

$$R = \frac{VL}{\nu}$$

where V is the velocity, L a characteristic scale length and ν the kinematic viscosity. The Reynolds number is a measure of the ratio between the inertial forces and the viscous forces. If the Reynolds number is high, the motion on the scale L is undamped, as viscosity plays no dominant role. When the Reynolds number becomes of the order of 1, the influence of viscosity becomes high enough that energy can be dissipated. L is set by the outer geometry of the flow, and the Reynolds number can be interpreted as the ratio between this length and the structure size at which the energy is dissipated. For atmospheric flows, $R \geq 10^6$. They can always be considered turbulent. The transfer of energy between the large scale and the dissipative scale can be explained as follows: the turbulence in the atmosphere is introduced by large eddies of a scale L_0 . These eddies cannot dissipate, as indicated by the Reynolds number, but break up in a hierarchical cascade to smaller and smaller eddies, until they reach a scale l_0 , and the turbulent energy is dissipated. The range between l_0 and L_0 is called the inertial range; L_0 is called the outer scale of the turbulence.

Kolmogorov developed a theory that allows us to assess the turbulent behaviour on different scales over the inertial range, that is, to calculate the power spectrum. His result shows that the spectrum has a universal form, and the calculations can be easily retraced.

A good way to describe the turbulence as a function of position is a structure function. A structure function of the form

$$D_\xi(r_1, r_2) \equiv \langle |\xi(r_1) - \xi(r_2)|^2 \rangle$$

describes the expectation value of the difference of the variable ξ evaluated at two positions r_1 and r_2 . We now want to describe the velocity field of the air. If we consider the turbulence to be homogenous and isotropic, the structure function for the velocity D_v can only depend on the distance between two points, and so the structure function assumes the form:

$$D_v(r_1, r_2) \equiv \langle |v(r_1) - v(r_2)|^2 \rangle = \alpha \cdot f(|r_1 - r_2|/\beta)$$

For the units to comply, the argument of f needs to be dimensionless, so the dimension of β must be a length, and that of α velocity squared.

In addition to the scale parameters, the turbulence depends only on the kinematic viscosity ν and the energy generation rate, ε . ε has units of m^2s^{-3} and ν of m^2s^{-1} . As α and β depend only on ν and ε , a simple dimensional analysis yields

$$\alpha = \sqrt[2]{\nu\varepsilon} \text{ and } \beta = \nu^{3/4} \varepsilon^{-1/4}.$$

Together with the assumptions that dissipation can be neglected in the inertial range, and so D_v must be independent of ν , we find for f :

$$f = k \cdot (|r_1 - r_2|/\beta)^{2/3}$$

with k as a scale factor. With this form, ν cancels out in the final form of the structure function:

$$D_v(r_1, r_2) = \alpha \cdot k (|r_1 - r_2|/\beta)^{2/3} = C_v^2 \cdot |r_1 - r_2|^{2/3}$$

with $C_v^2 = \alpha \cdot k/\beta^{2/3} = k \cdot \varepsilon^{2/3}$.

From very fundamental assumptions, Kolmogorov could derive a statistical description of the turbulence over the full intrinsic scale!

The relevant effect of the turbulence is the deformation of the incoming wavefront. For this, we need the structure equation for the index of refraction. The turbulent velocity field causes mixing of air of different temperatures; the air cells forming the turbulence are in pressure equilibrium, hence their density must be different. Pressure variations due to wind can be neglected in our case. The refractive index variations originate in the density perturbations. With the dependences of the index of refraction on the density, and applying the Wiener-Khinchin theorem, we find for the structure function:

$$C_N^2 \cdot r^{2/3} = D_N(r) = 2 \int d\kappa (1 - e^{2\pi i \kappa r}) \Phi(\kappa)$$

where $r = r_1 - r_2$, $N = n - 1$, κ is the wavenumber of the perturbations, and Φ denotes the power spectral density. From there, a rather complicated calculation (Tatarski 1961) leads to the power spectrum of the turbulence:

$$\Phi(\kappa) = 0.0365 \cdot C_N^2 \cdot \kappa^{-5/3} \quad |^1$$

Now, we calculate the perturbation of a wavefront. For a layer of thickness δh with spatial variations of the refractive index n , the local phase at position x imprinted onto a transmitted wavefront with wavenumber k is:

¹ These calculations have all been done in one dimension; the case in three dimensions is sometimes found in the literature and yields an exponent of -11/3.

$$\phi(x) = k \int_h^{h+\delta h} n(x, z) dz$$

With the structure function of the refractive index, $D_N(r)$, one can derive the structure function for the wavefront perturbation, or the phase shift, introduced by a thin layer at height h :

$$D_\phi^h(r) = 2.914 k^2 \delta h C_N^2 r^{-5/3}$$

This is valid for one layer of turbulence. We integrate over the whole height of the atmosphere and take a possible zenith angle, γ , into account. This yields:

$$D_\phi(r) = 2.914 k^2 r^{-5/3} \cdot \sec \gamma \int C_N^2(h) \delta h$$

With the definition of the so-called Fried parameter, r_0 ,

$$r_0 = \left[0.423 k^2 \sec \gamma \int C_N^2(h) \delta h \right]^{-\frac{3}{5}}$$

the expression for the phase shift structure function becomes

$$D_\phi(r) = 6.88 \left(\frac{r}{r_0} \right)^{\frac{5}{3}}$$

The relative phase shift between two points on a wavefront due to atmospheric turbulence grows rapidly with distance to the power of 5/3, and is dependent on the wavelength, the zenith angle, and the turbulence strength, all combined in the Fried parameter r_0 .

The Fried parameter is commonly used to describe the quality of an astronomical observing site. The value of r_0 for a given spectral band is a measure the integrated spatial properties of the atmosphere. The definition is chosen such that the mean-square phase variation over an aperture of the size r_0 is ~ 1 rad². In practical terms, the resolution of telescopes with apertures larger than r_0 is equivalent to the resolution of a diffraction limited aperture with diameter r_0 . The relation

$$r_0 \propto \lambda^{6/5}$$

is the reason why AO systems are much more common at longer wavelength. The ratio of the telescope aperture D to the Fried parameter, D/r_0 , is much better at infrared wavelength than in the visible part of the spectrum. From the discussion above it is also clear that r_0 is the approximate spacing of the actuators of the AO system's deformable mirror, projected onto the entrance pupil of the telescope. The number of actuators, and with it the complexity and cost of the system, scales with $(D/r_0)^2$. To give a value, for the best observing sites, r_0^{vis} is on the order of 20 cm in the green part of the spectrum. Hence, for an 8m telescope, D/r_0^{vis} is about 40, and the number of actuators needed would be on the order of 1600, very challenging and costly with today's technology. In the K -band, in contrast, r_0^{IR} is about 1m, according to the relation between Fried parameter and wavelength. So $D/r_0^{IR} \cong 8$, and the number of actuators needed is only about 60.

1.1.2 Isoplanatic angle

A common parameter to describe the turbulent behaviour of the atmosphere and its effect on image deterioration is the isoplanatic angle, θ_0 .

From our derivation of the structure function it follows that the turbulence is statistically independent of the viewing direction (at least for negligible variations in zenith angle). Hence the long exposure atmospheric PSF of a telescope is uniform over the field of view. The wavefront distortion at a given moment, however, is dependent on the viewing direction. As illustrated in Figure 3, the beams from two objects, overlapping at the entrance pupil of the telescope, do not overlap completely anymore, depending on angle θ between them and height h .

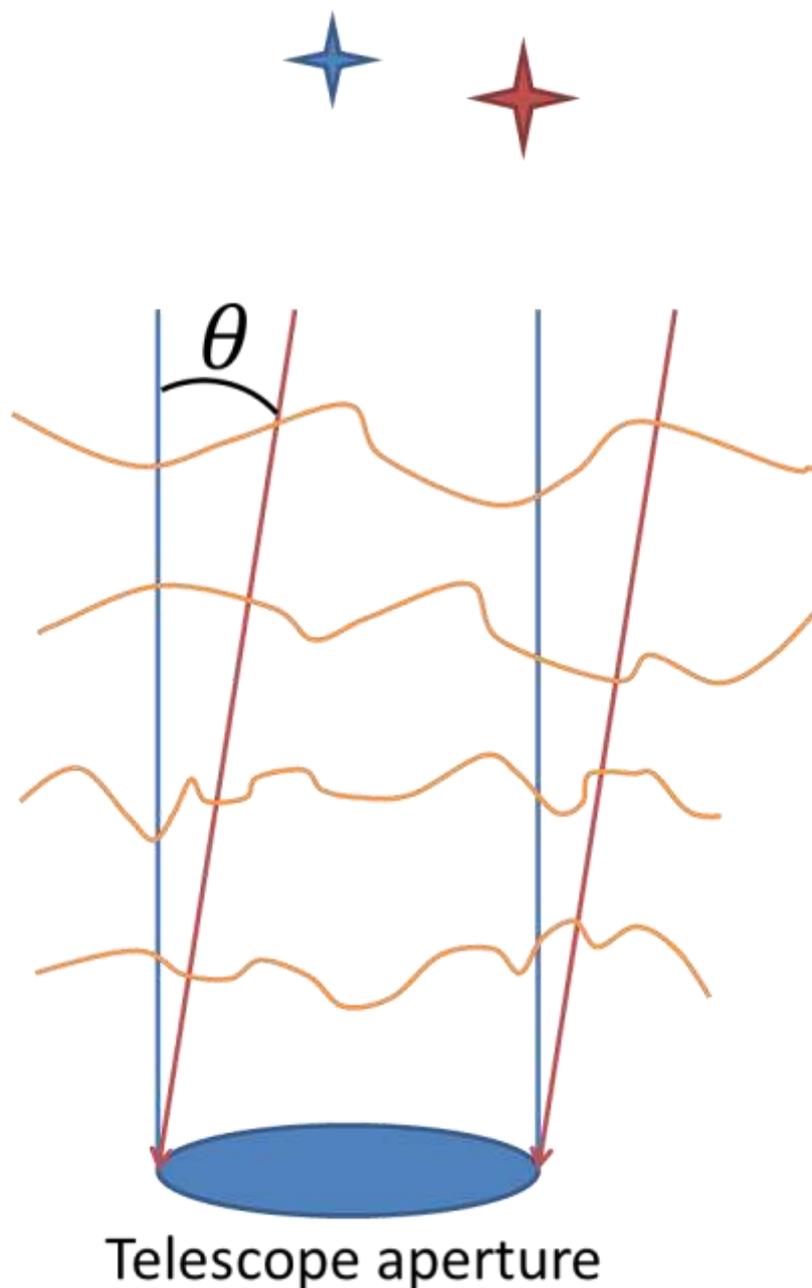


Figure 3: Illustration of the isoplanatic angle. The light from two stars separated by an angle θ , one of which is the AO guide star, samples different parts of the atmosphere.

Turbulent layers at higher altitudes have different contributions to both beams, and an ideal correction for one beam (the beam coming from the guide star) is not perfect anymore for an object at a different field position. This effect is called angular anisoplanatism. The isoplanatic angle θ_0 is defined such that the phase variation of objects closer together than this angle is correlated, and the correlation drops for field points further apart. More precisely the definition is such that within θ_0 the variance of the wavefront is 1 rad. Using r_0 the angle can be expressed as

$$\theta_0 = r_0 \cos \gamma \frac{r_0}{H}$$

Where γ is zenith angle, and H is a characteristic mean effective turbulence height, which is derived by weighting the turbulence strength C_N with $h^{5/3}$:

$$H = \left\{ \frac{\int C_N^2(h) h^{5/3} dh}{\int C_N^2(h) dh} \right\}^{3/5}$$

For typical conditions, this leads to a value for H of some kilometres.

In practical terms, θ_0 is the angle of the field benefitting from AO correction around the guide star's position.

1.1.3 The atmospheric time constant

The above thoughts cover the spatial structure of the turbulent layers at one given moment. The rate at which an AO system has to update the shape of the DM, and consequently re-evaluate the wavefront is one of the technical challenging limits. The time behaviour of the turbulence is therefore of great importance. Most often, the time in which the local shape of a turbulent layer changes is much longer than the time it takes for the wind to shift the layer across the telescope aperture. Turbulence can then be modelled as a frozen sheet, a phase screen, which is transported with the wind speed in the respective layer. This idea is called the *Taylor hypothesis of frozen turbulence*. For widely varying speeds in different altitudes and different turbulence strength in these layers, the precise temporal form of turbulence can be complex. From the approximation that the integrated turbulence can be described by r_0 , the definition of a time constant t_0 is commonly used to approximate the temporal behaviour:

$$t_0 = r_0/v$$

with v designating is the wind speed in the dominant layer. This definition leads to the same wavelength dependence for t_0 and r_0 . Wind speeds can on average be around 20 m/s. With typical values for r_0 of approximately 20 cm at good observing sites, we find ~ 10 ms for t_0 . $1/t_0$ is an estimate of the minimum frequency at which all control loops have to run. This has far reaching implications for the necessary guide star brightness, hardware specifications, and computation power.

1.1.4 The cone effect

The cone effect, sometimes called focal anisoplanatism, is attributed to the fact that guide stars of finite height sample a cone of the atmosphere, not a cylinder, like a star at infinity would. Dependent on the aperture size and the guide star altitude, the wavefront sensor signal

does not contain meaningful information to correct for the turbulence in layers above a certain height. A schematic of the effect is shown in Figure 4: The shaded area of the blue object's beam is unsampled by the red guide star. While the lowest layer of turbulence is well sampled by the guide star cone, the layers close to the guide star are poorly sampled, and turbulence even higher is not sampled at all.

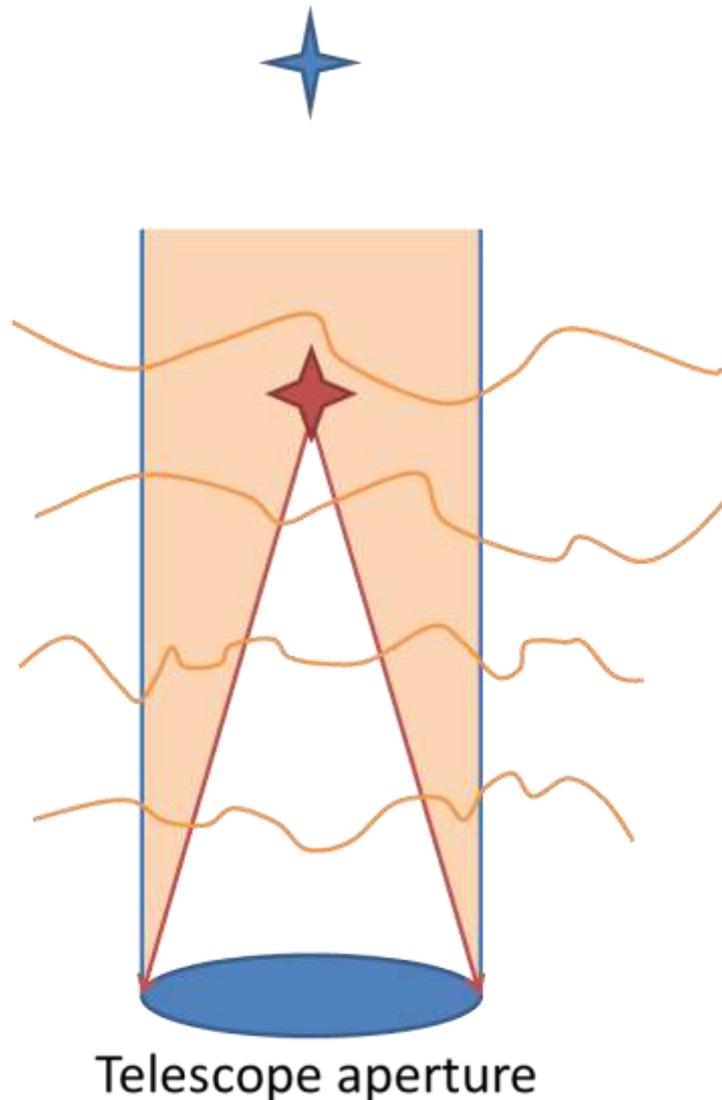


Figure 4: The light from an object at infinity (blue) that falls onto the telescope aperture forms a cylinder in the atmosphere. For a laser guide star at a finite distance (red), the sampled volume is a cone. The shaded volume is not sampled by the laser guide star.

This is one of the strongest arguments for Sodium laser beacons: with a mean altitude of ~92 km, they are well above all atmospheric turbulence and allow for a much better sampling of the high altitude layers than Rayleigh beacons with altitudes around 12 km can provide. For ground layer AO, on the other hand, the effect is less important, but for aperture sizes on the order of ten meter still not negligible. A straight forward way to improve the sampling of the atmosphere is to use multiple guide stars in an asterism. For high altitudes, it is complicated and very computationally expensive to disentangle the contributions of the different layers in different directions. This is referred to as laser tomography in the literature. For GLAO,

however, the computations become much simpler. The contribution of low layers is the common part of all the wavefronts from the guide stars, and can be derived by averaging over the measured signals.

1.1.5 Performance simulations for GLAO

During the Phase A of the Argos project, we carried out simulations assessing the performance of different asterism configurations and guide stars' altitudes for 4 different bins of atmospheric quality, based on the LBT's optical parameters and measured turbulence profiles. The simulations addressed the question of which constellation achieves the best compromise of on-axis performance and PSF homogeneity across the full field of Lucifer, an IR imager and multi object spectrograph with 4 x 4 arcminutes field of view, the main instrument to benefit from ARGOS. The simulations take into account all the effects discussed above, r_0 , θ_0 , and focal anisoplanatism.

I quote here simulations done by Sebastian Rabien, MPE. The geometry is shown in Figure 5. The results for the investigated setup closest to the finally chosen configuration are shown in Figure 6. The asterism radius and altitude are the same, but the simulation is performed for 4 guide stars while it was finally decided that ARGOS will use three. The difference in performance is negligible.

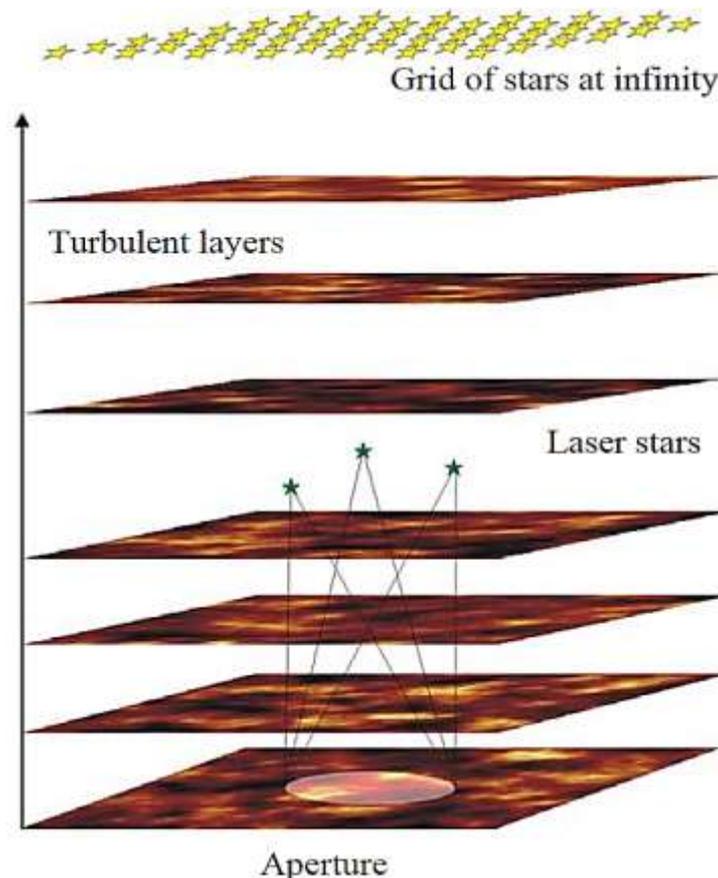


Figure 5: The optical geometry used in the simulations. A grid of background stars is propagated through a set of turbulent layers in the atmosphere. The same is done for laser guide stars at a given altitude and constellation geometry. Plot taken from ARGOS Phase A study.

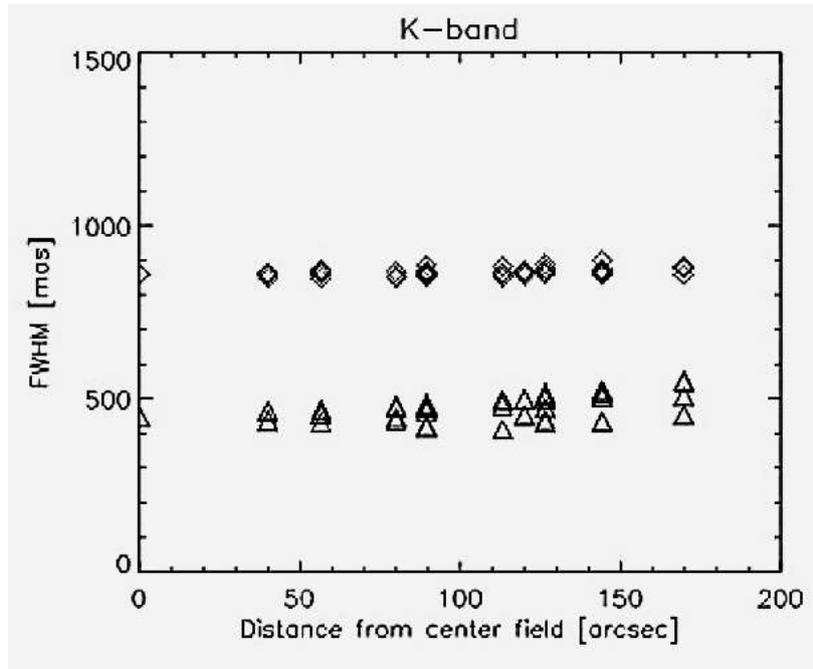


Figure 6: Field dependence of the FWHM over a 4x4 arcmin field. As an example, the simulation with the profile ‘bad’ is shown for the K-band. 4 laser guide stars, 12km height and 2’ radius were assumed. The top row of points shows the simulated uncorrected FWHM, and the bottom row the result of GLAO correction.

1.1.6 Spot elongation

The vertical extension of the laser beacon in the atmosphere is the cause for what is commonly called spot elongation. From the schematic in Figure 7 it is clear that a cylindrical beacon has an angular extension if not seen from directly on the propagation axis.

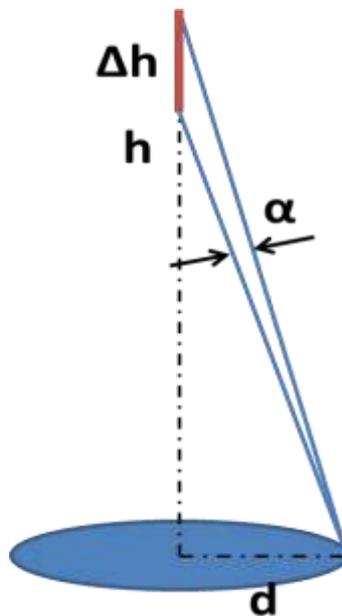


Figure 7: The spot elongation effect. A pencil shaped laser beacon at height h and vertical extension Δh appears under an angle α for a subaperture located at a distance d from the propagation axis.

For the image formed by the telescope aperture as a whole, this leads to the superposition of defocussed images, and so a blurred spot. As wavefront sensors sample the pupil plane, the effect is different. Each subaperture corresponds to a single aperture with a distance l from the propagation axis. The image of the laser beacon has an angular extension in the radial direction that grows for subapertures further away from the propagation axis, defined by the laser launch telescope. Obviously, a launch telescope in the centre of the telescope's aperture, for example behind the secondary mirror, minimises the overall effect. Figure 8 shows the magnitude of the effect for the subapertures of ARGOS, assuming a vertical beacon length of 120 m at an altitude of 12 km.

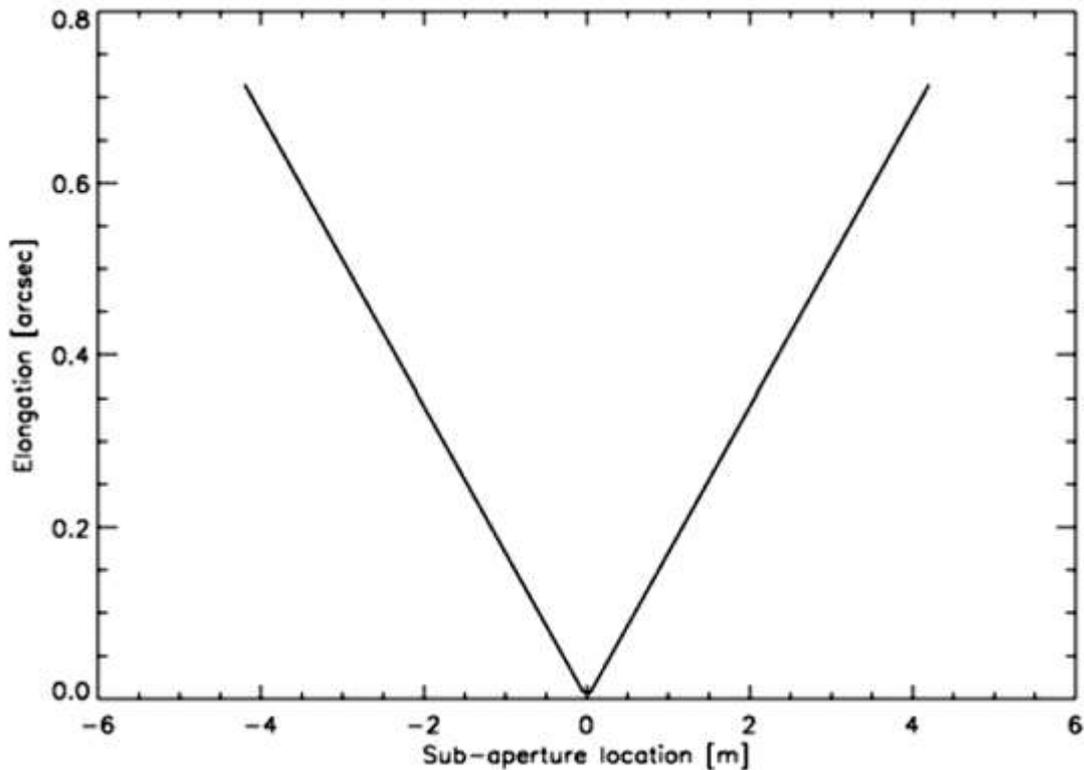


Figure 8: Spot elongation as a function of sub-aperture location relative to the primary centre. The plot is drawn for a 120m gating range at 12km gating height from a simple geometric consideration.

For a Rayleigh laser beacon, the spot elongation is determined by the optical setup (namely the distances of the subapertures to the laser launch axis) and the gating range. As the laser pulse travels up, a shutter in front of the wavefront sensor is only opened when the pulse is traveling a short distance at the desired altitude. The open time of the shutter sets the vertical extent of the pulse that is recorded by the wavefront sensor. Longer gating leads to more photons, but also to a pronounced spot elongation in the outer subapertures. For a given subaperture size, guide star altitude, and laser power an optimum has to be found.

Spot elongation causes the following problems:

- The light can spill into the region on the wavefront detector corresponding to the adjacent subaperture
- The centroiding accuracy, the sensitivity with which the spot's position can be measured, drops with the extension.

- The radial shift of the spot's centre with respect to a point source introduces a false focus term in the wavefront measurement. The range gate timing needs to be held constant on the 1% level to ensure this is static so it can be calibrated.

We carried out calculations during the preliminary design phase to assess the optimum range gate for ARGOS. A beacon with a Gaussian profile in the propagation direction was ray-traced with ZEMAX to generate the expected spot patterns on the Shack-Hartman wavefront sensor. The result is shown in Figure 9. We selected a gating range of 1% of the altitude, 120m, as the baseline.

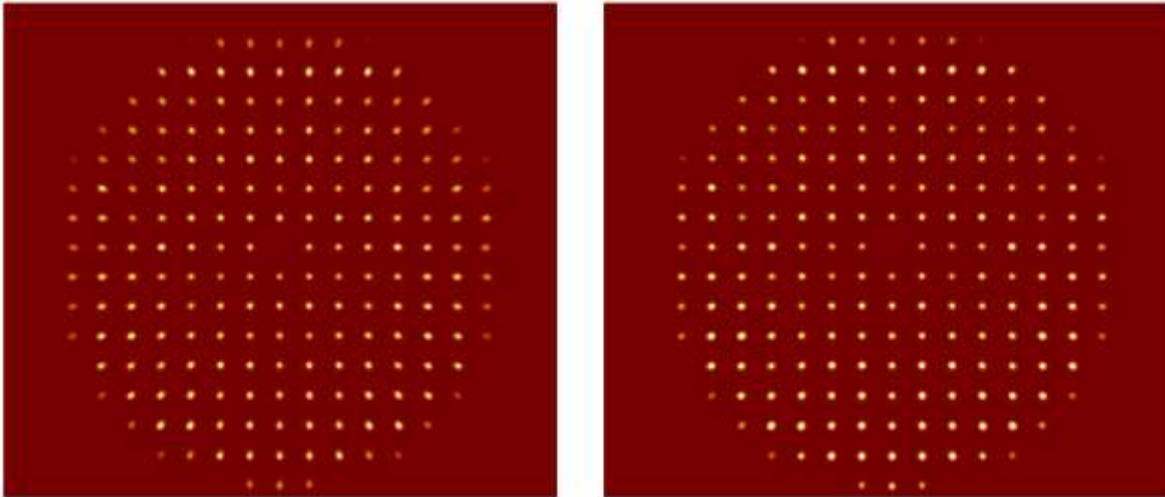


Figure 9: Spot patterns as expected for the Shack-Hartmann sensor for a given gating range. To the left a gate range of 200m travelling time is shown, to right the same for a 100m range. For the 200m case the elongation in the outer sub-apertures becomes visible, but is still below 1.5 arcsec. For the 100m range gate the maximum ellipticity is 86%. Taken from the ARGOS PDR.

1.2 Overview over existing LGS facilities

During the Phase A of the ARGOS project, I had the task to visit major laser guide star facilities in order to learn from their experiences. In this section I present the main characteristics of the other facilities developing Rayleigh laser guide stars. For comparison, I describe some features of prominent Sodium facilities. This was included in the Phase A to transfer the knowledge to the ARGOS team. A summary of the main LGS systems in operation is provided in Table 2. The lessons learned are summarized in the end of this chapter.

1.2.1 The William Herschel Telescope (WHT)

The WHT is a 4.2 m telescope, from the Isaac Newton Group (ING), located on La Palma. It has been equipped with a GLAO system (known as GLAS) developed by the ING in collaboration with the University of Durham (UK), the University of Leiden (NL) and the ASTRON institute (NL) (see Figure 10). The project started in 2004, with first light and commissioning taking place at the end of 2007². Scientific exploitation of the system is underway (Figure 11).

² ING Press release 4/2007.

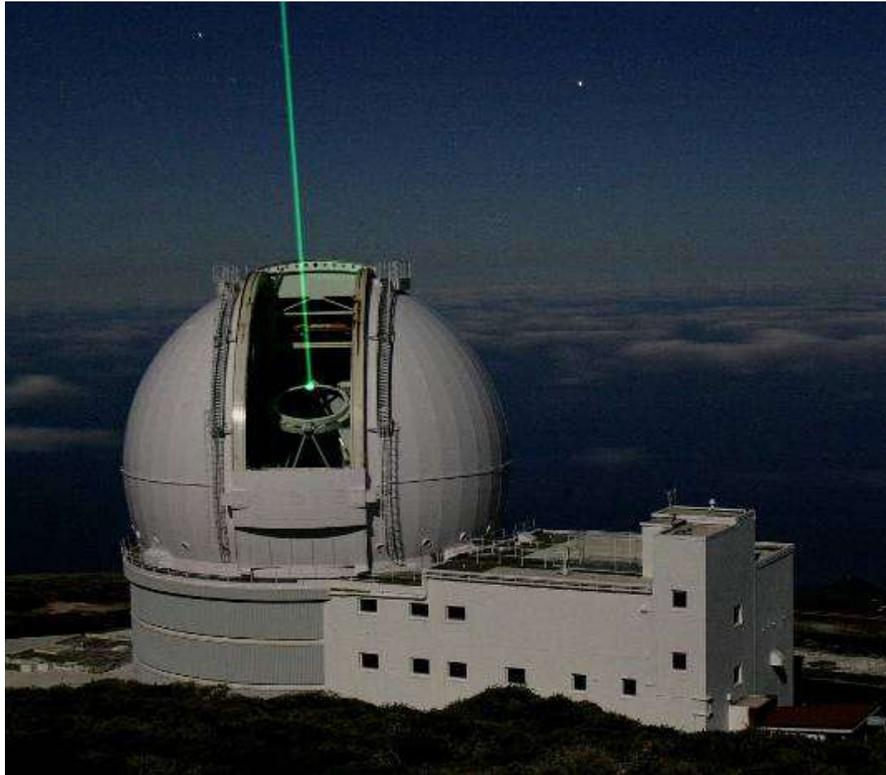


Figure 10: The Wilhelm Herschel Telescope on La Palma with its green laser beacon launched from behind the secondary mirror

The scientific drivers for GLAS were mainly to increase the sky coverage of the AO instrument NAOMI to nearly 100%, coupled with point-and-shoot capabilities. The instruments foreseen to be used with GLAS are firstly OASIS, a visual spectrograph employing an integral field unit with 0.2" lenslets and 20" FOV, and then INGRID, a 40" FOV NIR camera with 0.04" pixel scale. It is understood that GLAS will provide an improved image over a moderately wide field of view, but will not be able to deliver high Strehl ratios, as is the case with all GLAO systems. The expected gain in FWHM is a factor of two or better.

GLAS uses a single Rayleigh beacon, which creates an artificial star at an altitude of 15 km. The light source is a commercial frequency doubled Yb:YAG disc laser³, emitting an output power of 20 W at a wavelength of 515 nm. Given the telescope's transmission and the quantum efficiency of available CCDs, the use of a UV laser was decided against since for an equivalent price, it would not have been able to deliver a decent signal (a sufficient number of photons) on the wavefront sensor.

To keep the laser head fixed with respect to the gravity vector, it is mounted on the top ring of the telescope, in a cradle rotating around an axis parallel to the elevation axis. Some problems occurred in the procurement of the custom built laser unit, causing a delay of several months; and the goal of 30 W output power could not be achieved. The laser launch telescope (LLT) is a folded Galilean refractor design with single lenses and an aperture of

3 ELS GmbH, Germany.

350 mm. It is fully mounted behind the secondary mirror of the telescope. There are no fast steering optics in the uplink path of the beam. On the downlink side, a Shack-Hartmann array based on a CCD39⁴ and self-made camera electronics is used as a wave front sensor (WFS). The range gating is realized with Pockels cells (PC) with crossed polarizations, to be independent of the incoming light's polarization.

Design and integration of GLAS were straightforward, in the sense that readily available commercial parts were used wherever possible (with only minor exclusions). Also, care was taken to minimize risk by identifying critical components and performance necessities, and taking the appropriate steps to reduce the risk. This policy proved to be successful with the exception of the laser head itself.

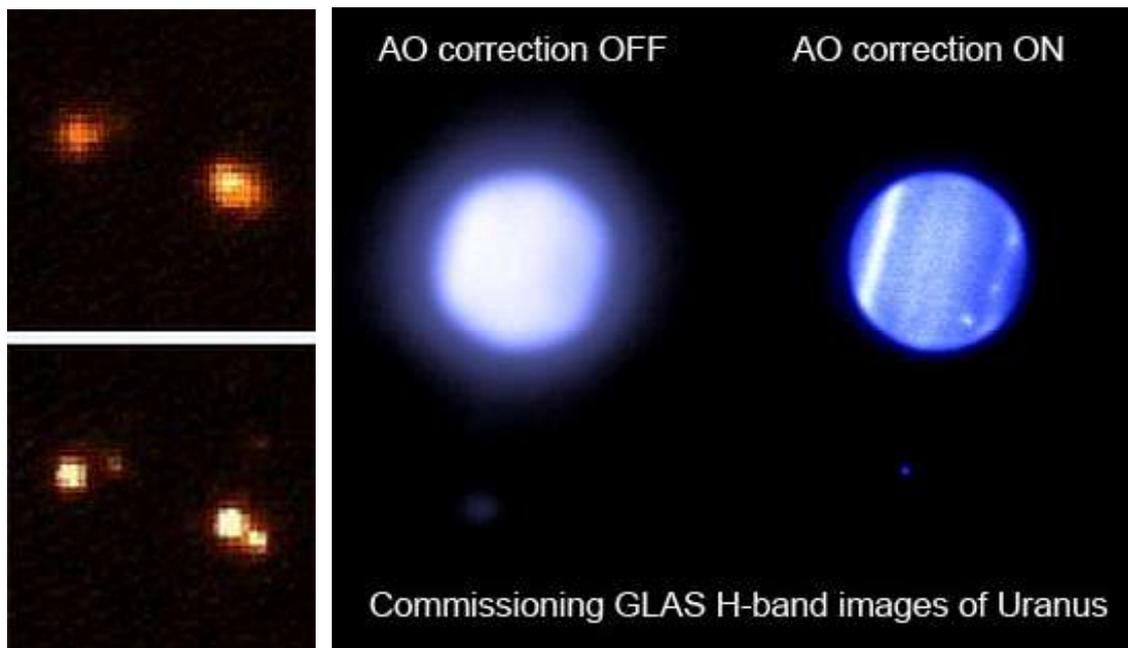


Figure 11: Left: comparison of stars in M15; top: seeing limited; bottom: with GLAS. Right: Uranus with and without AO

1.2.2 The Multi Mirror Telescope (MMT)

The MMT, a 6.5 m telescope located on Mt. Hopkins, AZ is the only facility that uses a multiple beacon RLGS and an adaptive secondary. In this respect, it is similar to the proposed LBT RLGS system. The participation of some MMT laser adaptive optics specialists in the LBT RLGS project guarantees a good transfer of experience. As the two sites are separated by only a small distance, the prediction of atmospheric conditions at LBT has also been simplified.

For the MMT LGS, the outputs of two frequency doubled commercial Nd:YAG Laser heads⁵ are combined by a polarizing beam splitter to form one single beam with 24 W power. These

4 E2V ltd, UK

5 JDSU Q-Series

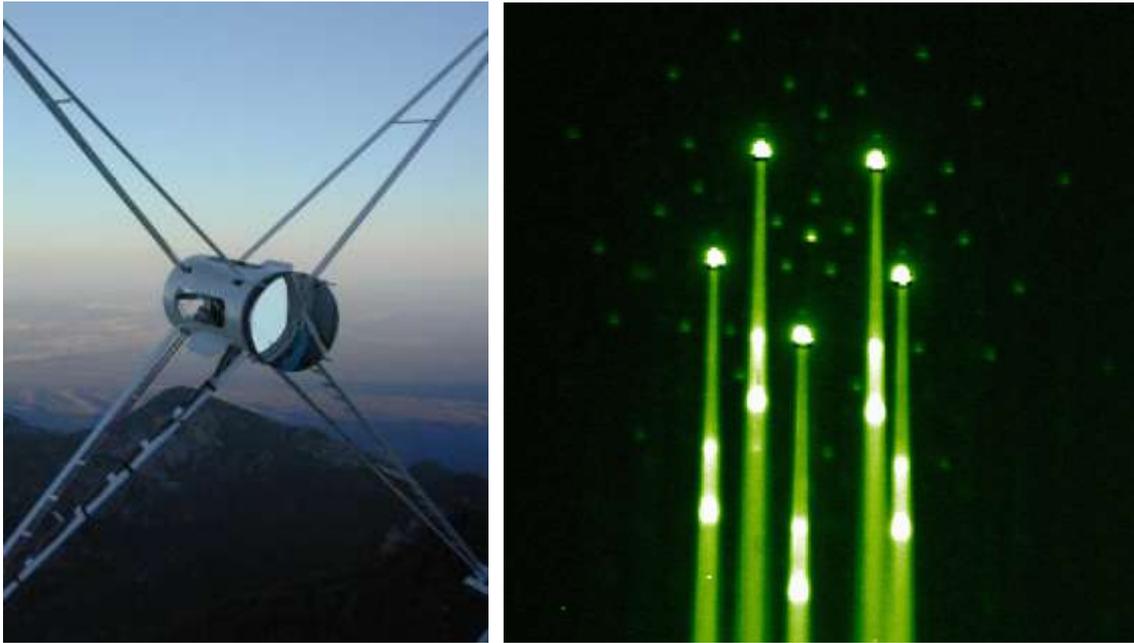


Figure 12: To the left a photograph of the MMT's adaptive secondary in dawn. To the right the asterism of the Rayleigh guide stars on sky can be seen.

lasers use the widespread technology of diode-pumped Nd:YAG rods as a lasing medium and are built-in quantities. The laser heads are mounted in an isolated aluminium box at the side of the telescope, above the elevation axis. The beam is relayed in free air to a so-called pupil box at the upper ring of the telescope, where a holographic phase plate is used to produce a fivefold asterism out of the single incoming beam. The beams are then put through to the folded refractive LLT, mounted behind the secondary mirror.

On the downlink side of the system, MMT employs unique techniques. A mechanical resonator is used to refocus the laser beacon while the pulse is travelling upwards through the atmosphere. This allows for the integration of the scattered light over a longer altitude interval, collecting more photons, and reaching higher altitudes with the same laser power. The resonator itself is an aluminum cylinder with an eigenfrequency of approximately 5 kHz, driven by electronics based on an audio amplifier and a voice coil. In its present form, the resonator is a stepped design for ease of manufacturing; nevertheless, the Q factor is very high so the resonance peak has a width of less than 1 Hz, with a stroke of up to 150 μm . Thanks to a well-thought out optical design, a single WFS (a gated CCD manufactured by the Lawrence Livermore National Laboratory, LLNL) is used to measure all five beacons. This unique feature saves the need for external optical gating mechanisms such as Pockels cells. Note that MMT had to struggle with performance problems of the CCD chips; two chips could be obtained, one of which was unexpectedly noisy, and the other had considerable different gain between the two output amplifiers.

A lesson learned is that small beam diameters on optical surfaces, especially those not inside the instrument, should be avoided. Moths crossing the beam get burned and the adjacent optics destroyed (Figure 14). Furthermore, safety concerns imply using larger beam diameters in free air to lower the power densities to an uncritical value.

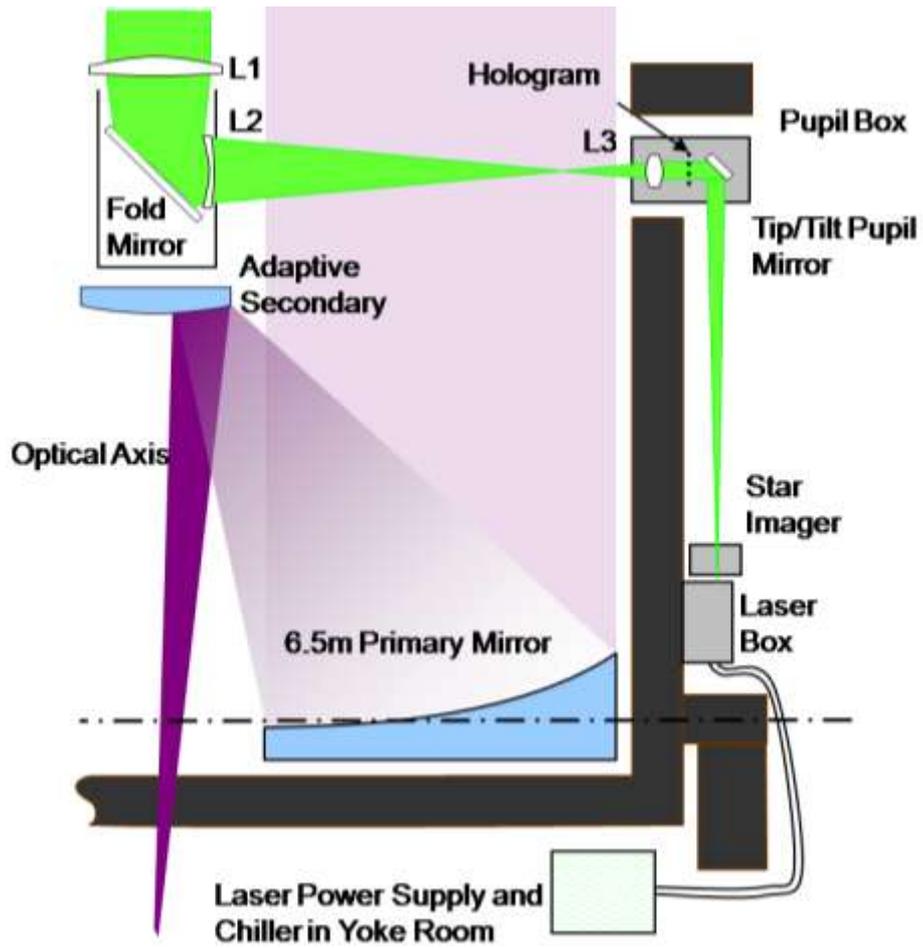


Figure 13: Schematic of the MMT laser guide star facility.



Figure 14: A moth that was killed by the 30W at 523nm laser beam of the MMT, and burned into the surface of an exit window. Image courtesy of MMT.

1.2.3 The Southern Astrophysical Research Telescope (SOAR)

SOAR with its 4.1 m aperture is a recently completed telescope located on Cerro Pachon in Chile (Figure 15). It is a joint project between the University of North Carolina, Michigan State University, NOAO and Brazil. The LGS program for SOAR, the SOAR adaptive module (SAM), differs from the other projects discussed thus far since it also fully includes the development of the AO part of the facility, not only the laser guide star facility. SOAR aims for high spatial resolution imaging and spectroscopy in the visible, with both requirements driven by competition with other facilities on-site. Based on extensive seeing studies on Cerro Pachon, and with the boundary condition of limited funds, GLAO was identified as the best option. The clear goal is to build a robust, low-cost AO instrument.

The actual status of the project is somewhat less advanced than foreseen. Integration of the optical and mechanical parts started in 2008. At the time of the Phase A study for ARGOS, no on-sky experiences could thus be shared. However this project has useful well-documented and very detailed engineering aspects. Nearly all the critical components were thoroughly studied, and in many cases simple but considerable improvements of commercially available solutions were implemented. Furthermore, breadboard tests were carried out, and the proper functioning of the subsystems verified. Andrei Tokovinin, leading the SAM project, shared very interesting insights about ways to characterize the instrument, to optimize the performance and to train the operators. For this purpose, a light source with a turbulence simulator (TurSim) will be integrated into SAM (Thomas 2004).

SAM uses a UV laser with 8 W output power at 355 nm⁶. This model has the same characteristics as the laser used by MMT and is from the same company, but is frequency tripled. The global efficiency of the system from the light source to the WFS differs only minimally from a solution with a green laser. The main advantage of the UV beacon is that it may be possible to abandon the use of aircraft spotters employed by the observatory every time the laser is operated. Andrei Tokovinin argues that at the speed an aircraft moves, the irradiance by the laser is much lower than the damage threshold of tissue, even for direct exposure, as the exposure time will be very short. In addition, at the given UV wavelength, the cornea of the eye is opaque, so the energy is absorbed before reaching the retina. Damage of the cornea or lens occurs at the same energy as in normal tissue, thus the laser should not be able to dazzle a pilot and should not present a hazard to aircrafts. A further argument is the fact that the (plastic) windows of airplanes are opaque at UV wavelengths. It remains to be seen whether these arguments will be accepted by the relevant authorities, however. Nevertheless precautions for laser safety for SAM comply with the measures adopted by the other observatories. The use of this short wavelength imposes consequences for the employed optics; transmission and chromaticity characteristics were chosen accordingly. The laser launch telescope works with an aperture of 350 mm. The beam path will be enclosed up to the Cassegrainian launch telescope that will be mounted behind the secondary. Both the NGS and the LGS WFS will be housed together with re-imaging optics and a bimorph DM in a sealed aluminum enclosure mounted on one of the Nasmyth foci of the telescope. The enclosure will rotate to compensate field rotation. The Shack-Hartman wavefront sensing will use a CCD39 with Leach electronics. The gating will be carried out by a single Pockels cell. Fiber-coupled avalanche photodiodes (APD) will be used for the tip-tilt measurement.

6 JDSU Q-series

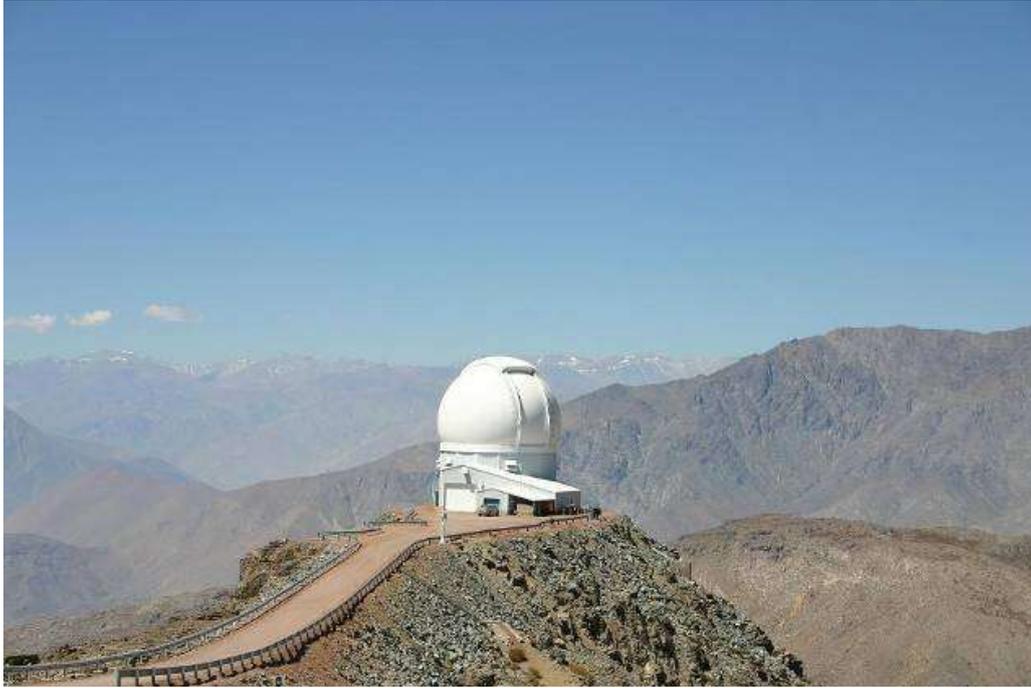


Figure 15: The SOAR dome at Cerro Pachon

An instrument foreseen to be used with SAM is an imaging camera with 3' FOV. There is a second port available for visitor instruments, where for e.g., the Brazilian collaborators could install their spectrograph.

1.2.4 Sodium Laser Guide Star facilities

For comparison, I describe some of the main Sodium Laser Guide Star (NaLGS) facilities below.

Lick Observatory

The LGS facility at the 120" Shane telescope, Lick Observatory (California), was the first astronomical laser guided AO system in operation. The technology and experience in operating the systems developed there was then used to build the Laser facilities for Keck I on Hawaii. It took eight years from closing the loop at Lick in 1996 to successful operation at the Keck telescope, however. The Lick Observatory laser system is a pulsed dye laser developed by the Lawrence Livermore National Laboratories (LLNL). Four flashlamp pumped solid-state lasers operating at 532 nm are located beneath the main floor of the telescope dome and pump a dye laser that converts the light from green to yellow, i.e. tunes it to 589nm (sodium's resonant wavelength). The light travels through fibre-optic lines to a launch telescope mounted on the side of the telescope, before it gets launched into the atmosphere.

W. M. Keck Observatories

In 2004, Keck started using a LGS, and began producing substantial scientific output with it. As a consequence, nowadays more than 20% of the nights on Keck are scheduled for LGS operations. The instruments available for these observations are two imaging spectrographs, also compatible with NGS AO. While the AO correction is slightly worse than in NGS mode (with a best Strehl of 0.5 compared to 0.7), the LGS drastically improved sky coverage. (van

Dam et al. 2007) The Laser itself, custom build by LLNL, is a seeded two-stage dye laser amplifier, pumped by six frequency-doubled Nd:YAG lasers. A dye master oscillator in a laser room is locked to a sodium gas cell and delivers a few milliwatts of light into an optical fibre; the fibre output is then amplified by a laser mounted on the side of the telescope. The whole chain draws 50 kW of electric power to deliver 14 W laser light on sky (Wizinowich et al. 2006).

PARSEC at the VLT

Most major 8 to 10m size telescopes facilities are now becoming equipped with LGS facilities and most of them, with the exception of the three RLGS systems discussed above, are using the atmospheric sodium layer to create artificial guide stars. ESO's VLT at Paranal does not escape this trend and has been offering a laser guide star facility to its community for scientific use since October 2006. After a slow start, the first scientific results are now coming out. The PARSEC laser is installed at Yepun (UT4) and is used for both an integral field unit near-infrared instrument (SINFONI) equipped with a curvature sensing adaptive optics system, and a multi-purpose near-infrared camera (CONICA) fed by the Nasmyth Adaptive Optics System (NAOS) (Lenzen et al. 2003, Rousset et al. 2003), an adaptive optics system based on Shack-Hartmann wavefront sensing. This camera allows imaging, long-slit spectroscopy, polarimetry, coronagraphy and simultaneous-differential imaging to be performed. The laser is a continuous wave (cw) single frequency dye laser, divided into a master laser that defines the frequency and a resonant cw amplifier delivering the high power. Optical pumping of the dye molecules takes place in the green spectral region with a total of five 532 nm solid state lasers, each delivering a 10W cw beam. Unlike the Lick system, the sodium tuned laser light is first created in the laser lab and then propagated through fibres to the launch telescope placed behind the secondary mirror of the VLT UT4. Under average atmospheric conditions, for stars of V magnitude between 10 and 12, the Strehl ratio measured is the same when using the NGS and LGS systems, as expected for an equivalent laser guide star magnitude of $V \sim 12$. Note however that the performance of the AO systems is limited by telescope vibrations. Moreover, in LGS mode, both AO systems suffer from the poor performance of the STRAP unit used for tip-tilt sensing. On top of that, in the case of NAOS, neither the pixel size nor the aperture of the WFS is well suited for the LGS. The system is presently under-correcting, compared to what it could achieve with an appropriate lenslet array (under construction).

Gemini South

The first telescope to use multiple NaLGS beacons to overcome some of the limitations of existing LGS AO systems, i.e. limited sky coverage, restricted field of view and negative impact of the cone effect, will be Gemini South. With the multi-conjugate AO (MCAO), the goal is to achieve AO correction over more than $80''$ FOV with a stable PSF. Expected performance is around 25% Strehl in the K-band. Gemini's MCAO design uses five laser guide stars and three natural guide stars and thus eight wavefront sensors. The idea is to use a continuous wave laser of 50 W that will be split into five equivalent beams of 10 W. As of the end of 2007, all loops had been closed in the laboratory, and integration at the telescope had started (Figure 16). Commissioning on the telescope was planned for the second half of 2008, but has been delayed due to problems with the laser system on the manufacturer's side. The MCAO module will feed two instruments: The Gemini South Adaptive Optics Imager (GSAOI), a $4K \times 4K$ near-infrared camera with a field of view of $80'' \times 80''$; and FLAMINGOS-2, a near-infrared multi-object spectrograph, with a resolution of up to $R=3000$ over a 2 arcminutes field of view (quite similar to the LUCIFER capabilities).



Figure 16: Gemini LGS parts (the top of the beam relay housing), in the integration hall at the telescope site in December 2007

Major effort and cost went into development of a reliable laser system with an output power of more than 50 W. A solid-state Nd:YAG laser with sum-frequency mixing was expected to be ready in April 2008 (information Nov. 2007). Integration of the other components started in 2007 at the telescope. The latest information is that the laser will be delivered by mid-2010.

GALACSI at the VLT

ESO is developing an AO instrument called GALACSI for the VLT UT4 (on which PARSEC is mounted). It makes use of the so called four laser guide star facility (4LGSF) comprising four Sodium LGS on a variable size asterism, and a deformable secondary mirror, a convex shell with 1.2m diameter and 1170 actuators, based on the same technology as LBT's deformable mirrors (Arsenault et al. 2008).

The development of GALACSI is part of the AO facility initiative aiming to convert UT4 into a specialized AO telescope. (Arsenault et al. 2006). The science instrument to benefit from GALACSI is the Multi Unit Spectroscopic Explorer (MUSE), a intergarl field spectrograph in the optical waveband (465 – 930 nm). MUSE comprises 24 spectroscopic channels, each equipped with an image slicing integral field unit (IFU). The full field of view delivered by UT4 + GALACSI is divided between the 24 channles. GALACSI and MUSE together (called the MUSE facility) are designed to operate in two different modes, a wide field mode (WFM) and a narrow field mode (NFM). Their respective basic parameters are summarized in Table 1. The optical layout of the GALACSI bench is shown in Figure 17.

Table 1: Specifications for GALACSI + MUSE

	Wide Field Mode (WFM)	Narrow Field Mode (NFM)
MUSE field of view	1' by 1'	7.5" by 7.5"
Spatial pixel size	0.2" by 0.2"	0.025" by 0.025"
AO performance specification	double the ensquared energy in one spatial pixel at 750nm	5% Strehl at 650nm (goal 10%)
AO correction type	Ground Layer AO	Laser Topographic AO
Tip Tilt star brightness	>17.5 R-mag	>15 J-H mag
Tip Tilt star location	52" to 105" off axis --> >70% Sky coverage	on axis, within 7.5" diameter
LGS separation	4 LGS 64" off axis	4 LGS 10" off axis

In both modes, the Laser guide stars are located just outside the science field. In the wide field mode, these are picked up by an annular mirror to minimize the impact on the science field of view, while in narrow field mode a dichroic is inserted to redirect the alser light to the wavefront sensor, based on a Shack-Hartmann lenslet array. The tip-tilt signal is measured with an optical wavelength sensor on a star located outside the science field to avoid obstruction. In narrow field mode, an IR low order sensor (IRLOS) is used to sense remaining aberrations. It operates from 980 to 1800 nm and is based on a 2 by 2 subaperture Shack-Hartmann array.

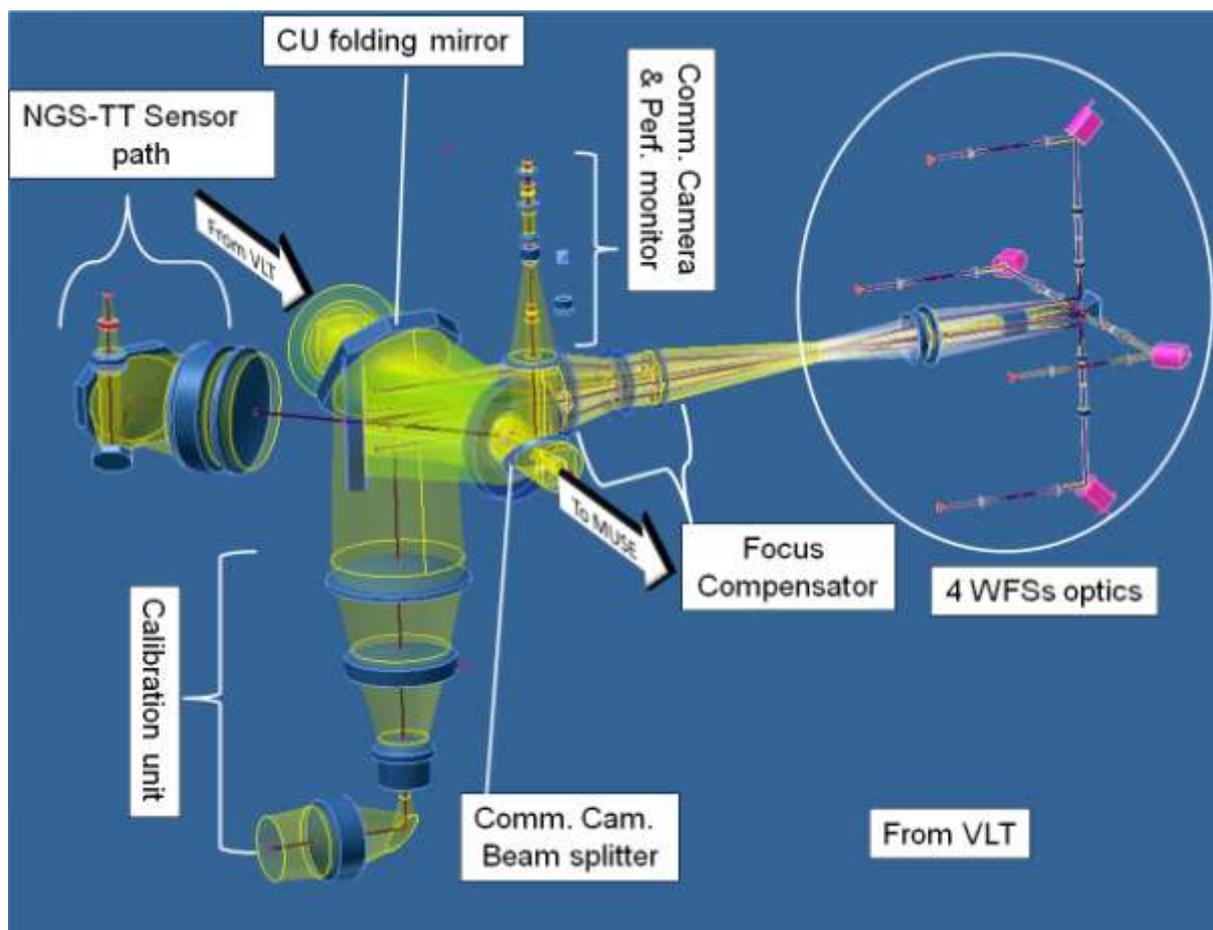


Figure 17: Schematic of all optical components of the GALACSI main assembly. ESO

GALACSI also contains a calibration unit. Optical fibres at different wavelengths are positioned precisely in a metal plate (Figure 18). This so-called source plate is imaged by a dedicated objective onto the Nasmyth focal plane. The calibration unit is fixed within the GALACSI structure on the Nasmyth platform and deployed with a flat folding mirror. The point sources will provide the means to calibrate the wavefront sensors as well as the field selector positioning. A diffraction limited point source for performance assessment in the narrow field mode is included as well. The unit can be focussed to simulate changes in the sodium layer altitude. The calibration system is located after the deformable secondary mirror and therefore cannot be used to measure the interaction matrix or run the system, including the DM, in closed loop. A way to introduce wavefront distortion to simulate atmospheric turbulence is not foreseen.

After passing the FDR in June 2009, procurement of all components is underway. System integration is planned to start in mid-2010 and commissioning at the end of 2012.

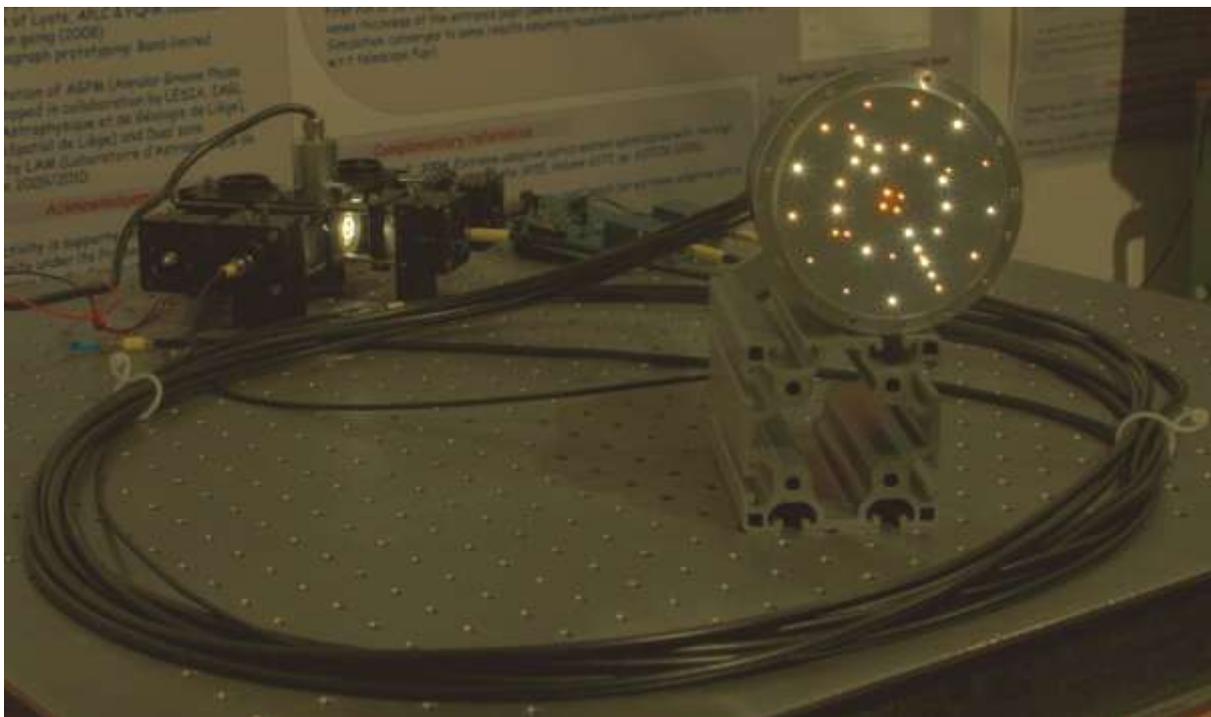


Figure 18: Source module for the GALACSI Calibration Unit. Fibers of different diameters, fed with light of different wavelengths to simulate NGSs and LGSs are precisely arranged in a metal plate. *Picture courtesy ESO.*

1.2.5 Summary

As reflected in the discussion above, most laser guide star efforts have focused on two methods of creating artificial stars. The first method uses visible or ultraviolet light to reflect off air molecules (Rayleigh scattering) in the lower atmosphere, creating an artificial star at an altitude of about 10km. The other method uses yellow laser light to excite sodium atoms at 90km. Ideally, the artificial star should be as high as possible to have the laser star and the observed star passing through the same part of the atmosphere. However recent studies have shown that most of the atmospheric turbulence affecting observations is concentrated in the

first 15 km of the atmosphere. Only rarely is the turbulence dominated by high atmosphere turbulence, e.g. the jet stream. This explains the enthusiasm for ground layer AO and its capacity to significantly improve the image resolution over large field of views.

Much effort has been put into LGS AO systems but thus far only a few of them are fully functional i.e. delivering refereed science papers. On large telescopes, it is now important to have an adaptive optics system to be competitive. Most 8-10m class telescopes opted for a sodium laser solution with on-axis correction. Presently only smaller telescopes have opted for a GLAO solution. This thus provides a unique niche for the LBT, which will be able to deliver higher resolution over a wide field and has a unique NIR multi-object spectrograph capability.

The observatories discussed above that employ a RLGS want to use the LGS AO as a way to observe with spatial resolution approximately twice as good as in seeing limited mode, everywhere on the sky, with the additional benefit of reaching fainter objects due to the improvement in encircled energy (EE). All three groups plan to employ the LGS on a substantial number of nights each year. The three investigated RLGS projects have different political environments, as well as technical capabilities. This clearly leads to different approaches for the same problems. Still, there is very good agreement on the capabilities of GLAO, and a large overlap on how to implement it with the available technology.

Each of the systems described above are different, as each team responds to the existing conditions at the telescope and wants to find their own niche. Each system is hence optimized for different kinds of science or instruments to be fed by AO. While the prospects have improved over the last few years, obtaining the necessary laser power at the sodium wavelength is still a major obstacle for the sodium facilities. For all the described projects, the Sodium lasers come with high costs, require a considerable amount of maintenance and are custom built systems, and hence impose a substantial amount of risk on a project. The advantage is obviously the better sampling of the atmosphere, and the resulting high Strehl ratios.

Rayleigh AO on the other hand, cannot correct for some hundred nanometres phase distortion RMS that is not sampled by the low altitude beacon. The performance is necessarily different - RLGS systems are seen more as facilities to improve the seeing. The Rayleigh laser technology, however, is much easier to use. The necessary power can be bought more or less off-the-shelf, with only slight modifications from commercial units necessary (e.g. to the cooling and pulse rate). Frequency doubled or tripled solid state lasers are the workhorses for the laser lithography industry, for example, and are very robust. These very systems were identified to be the most attractive commercial alternatives. Spare parts are also likely be available for a long time. As outlined earlier, GLAO can foster interesting science, and hence such a system makes sense if coupled to the right instrument and if one wants to pursue a low risk strategy for implementing AO.

From the visited facilities, some lessons learned were universal:

- The Laser projection system is important, and requires a substantial amount of attention to find the right engineering solutions. The spot size achieved on sky is a very relevant parameter for the performance of the whole system. Concerns are the beam quality, vibrations and throughput.
- Insects are attracted by the laser beam at night, and can seriously disrupt operations, and even destroy parts of the system. Large beam diameters when propagating in free

air, enclosed beam relays wherever possible and easy access to exposed surfaces (to remove insects) are important means to prevent possible loss of a large amount of time on sky.

- Scattered light is important, as the flux levels of the propagated light are so high. This requires regular cleaning of optical surfaces and attention to baffling or notch filters in front of detectors.
- Laser safety is an important issue for installation and operations. Astronomers are not familiar with the safety regulations, protocols, and what could be called ‘good laser lab practice’. Therefore, rigorous safety standards are necessary.
- The matter of range gating is complicated. Pockels cells are a reliable, if somewhat complicated and alignment-effort intense solution. They show ringing, which leads to spurious secondary peaks in the spot profile. The high voltages and fast switching next to the science detectors make good shielding a must. The necessary field of view in the astronomical application is much larger than in the typical application of Pockels cells as Q-switches in laser cavities. This becomes more serious with larger aperture diameters. As a last point, Pockels cells reject one linear polarization state. Further research is needed to make sure the backscattered light is linearly polarized. Other techniques, the foremost of which is on-chip gating, promise much easier integration but have other drawbacks, like higher noise levels, inappropriate pixel sizes and lower quantum efficiency, or are hard to obtain. The LLNL CCD system is seemingly not a good option, based on the results from MMT. As a good, reliable range gating is so important, a considerable amount of investigation is needed on this topic.
- The choice of Rayleigh laser wavelength is not immediately clear. UV lasers gain from the λ^{-4} dependence of the Rayleigh scattering, but have lower output due to the efficiency difference in the frequency conversion process between doubling and tripling. Also, the telescope optics are not optimized for these short wavelengths, so the light experiences more losses. The WFS and LLT optics are more complicated as they need to be made from UV transparent materials. Green lasers on the other hand are visible from considerable distances; this may be disapproved of by third parties. The visibility makes them safer for the staff handling the systems, though, a very important point as observatory staff might not expect laser radiation, and cannot see the UV laser, even at very dangerous flux levels.
- One of the main cost drivers is the staff required to look for aircraft (the laser spotters). All possible effort should go into preventing one from having to hire and organize the spotters. UV lasers might have an interesting advantage, if the argument can be made that airplane windows are opaque and the lasers therefore present therefore no danger to the pilots. An automated system that triggers an interlock if airplanes come close to the beam is another good alternative
- Custom laser systems should be avoided as much as possible, to mitigate the risk of power degradation, unavailable spare parts, or service personnel with little experience with the system in question.
- It is noteworthy that RLGS will allow for upgrade, either by combining it with a low power Sodium laser, or by using tomography, possibly with the aid of a dynamic refocusing system. Interfaces for those options should be planned in as an option.

Table 2: Key parameters of existing or planned LGS facilities and their instrumentation.

Telescope	Type of LGS AO	Attached Instrumentation	Strehl/FWHM
Functional systems			
Lick (120'')	Na LGS – SCAO ^a	IRCAL: 0.9 - 2.5 μm imaging - 19.4'' FOV - H & K spectroscopy (R=500)	4 \times better than seeing
Keck (10m)	Na LGS - SCAO ^a	<ul style="list-style-type: none"> NIRSPEC: 46'' FOV <ul style="list-style-type: none"> 0.95–5.4 μm high resolution spectrograph 0.95-2.5 μm imaging NIRC2: 1-5 μm imager, 10-40'' FOV OSIRIS: NIR integral field spectrograph R~ 3800 	50% K strehl peak – on axis
VLT (8m)	Na LGS - SCAO ^a	<ul style="list-style-type: none"> NACO: 1.0-3.5 μm spectrograph and imager, FOV 13'' to 54'' SINFONI: 0.95-2.5μm integral field unit 0.8-8'' FOV MUSE: 0.46 -0.93 μm IFU spectrograph, 7.5'' or 1' FOV 	Median K strehl 25% 2x EE gain / 8% Strehl in R band
Gemini North (8m)	Na LGS - SCAO ^a	<ul style="list-style-type: none"> NIRI: <ul style="list-style-type: none"> 1.0 – 5.0 μm imaging, 22'' and 51'' FOV 1.05–2.41 μm moderate resolution spectrograph, 22'' FOV NIFS: integral field spectrograph <ul style="list-style-type: none"> R~5000, 0.95–2.40 μm, 3.0'' FOV Spatially resolved spectrograph on 0.1'' scale Coronagraphic spectroscopy 	20% K strehl
Systems installed & undergoing commissioning			
Subaru (8m)	Na LGS - SCAO ^a	IRCS: 0.9-5.6 μm imager & Echelle spectrograph	Diffraction limit [2-5] μm
MMT (6m)	532nm Rayleigh LGS SCAO ^a adaptive secondary 5 LGS, 1 WFS	<ul style="list-style-type: none"> ARIES: 1.1-2.5 μm imager & spectrograph, 20 & 40'' FOV MIRAC 3/BLINC: 5-25 μm imager, diffraction limited up to 6 microns BLINC: nulling interferometer which uses MIRAC as its imager 	30% K strehl on axis
WHT (4m)	Rayleigh LGS – GLAO ^b	<ul style="list-style-type: none"> OASIS: 0.43-1.0 μm imager & integral field spectrograph, 38'' FOV INGRID: 0.8-2.5 μm imager, 4.3' FOV 	< 0.2'' FWHM
Systems under construction			
Gemini South (8m)	Sodium LGS - MCAO ^c 5 LGS, 3 NGS, 8 WFS	<ul style="list-style-type: none"> GSAOI: NIR high resolution imager, 80'' FOV FLAMINGOS-2: <ul style="list-style-type: none"> 0.95-2.4 μm imaging Multi-object spectrograph, R=1200-3000, 2' FOV (MCAO) 	Expected 25% K strehl

SOAR (4m)	UV Rayleigh - GLAO ^b	<ul style="list-style-type: none"> • SOI: 0.31-1.05 μm imager, 5.25' FOV • OSIRIS: <ul style="list-style-type: none"> - 1.0-2.2 μm imager, 3.3' FOV - Spectrograph, 1"x175" maximum FOV 	< 0.2" FWHM
LBT (2x 8.4m)	Rayleigh LGS GLAO ^b adaptive secondary	LUCIFER: 1-2.5 μm spectrograph and imager, with multi-object spectroscopy capacity, 4' FOV	< 0.3" FWHM (median)

^a SCAO = Single conjugate adaptive optics i.e. on-axis correction. These systems typically use one LGS and one NGS (for tip-tilt correction) and single WFS. The quality of the correction at a given point in the field of view depends on the anisoplanatic angle.

^b GLAO = Ground layer adaptive optics

^c MCAO = Multi-conjugate adaptive optics

1.3 Description of LBT and Lucifer

The Large Binocular Telescope is located on Mt. Graham in Arizona. Two 8.4 m primary mirrors are mounted on the same azimuthal mount, as shown in Figure 19. These two "eyes" can be used with individual instruments or coupled to form an interferometer with a maximum baseline of 22.8 m. Adaptive Optics capabilities have been a priority in the design of the telescope from the start, and consequently it will be equipped with adaptive secondary mirrors (ASM), sporting 672 actuators driven by voice coils.

It is currently planned to explore the telescope's capabilities in four phases: firstly operated as two telescopes in seeing limited mode, then with the addition of NGS AO. In the next step, the interferometric modes will be exploited, and finally LGS AO will be added.

The first generation instrument suite consists of two prime focus cameras (Large Binocular Cameras, LBC), an interferometer bench called LBTI (LBT Interferometer), the imaging Fizeau interferometer LINC-NIRVANA (for LBT Interferometric Camera and the Near-IR / Visible Adaptive Interferometer for Astronomy), PEPSI, (the Potsdam Echelle Polarimetric and Spectroscopic Instrument), a spectrograph with polametry option and a maximum resolution of $R=320,000$; further two MODS units, a visible light (330-1100nm) Multi Object Double Spectrograph, and two near infrared multi-object spectrograph and imager, LUCIFER I&II. This is the instrument the laser guide star facility is developed for.

Two basically identical LUCIFER units will be mounted on Gregorian ports of the two eyes of LBT. LUCIFER I was commissioned in 2009, and delivers a resolution of up to $R=37100$ over up to 4 x 4 arcmin field of view (see Table 3). The large FOV and MOS unit coupled to the sensitivity provided by the 8.4 m primary make the instrument highly competitive. With these unique capabilities, the LUCIFER units will most likely be the workhorses of the LBT in the upcoming years.

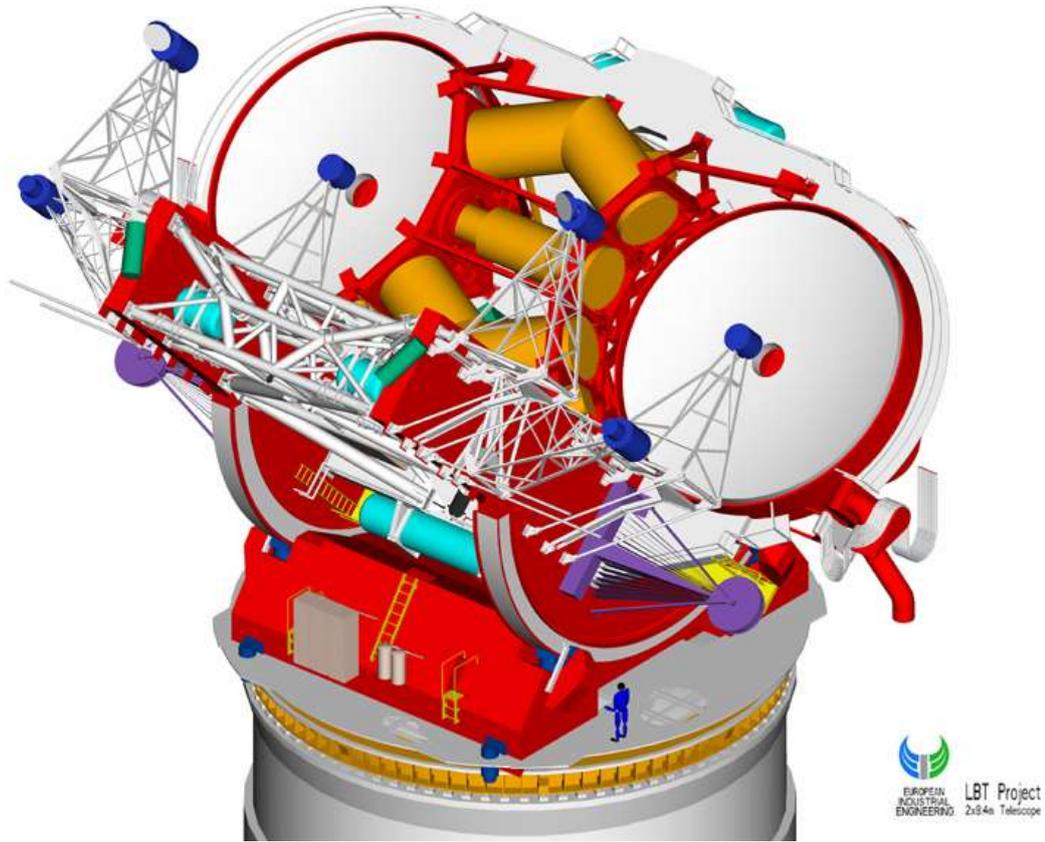


Figure 19: A model view of LBT

Table 3: Observing modes of Lucifer

	Seeing Limited		Diffraction Limited
N_{cam}	N3.75	N1.8	N30
FOV	4 × 4 arcmin	4 × 4 arcmin	0.5 × 0.5 arcmin
f_{Coll}	1500 mm	1500 mm	1500 mm
f_{Cam}	375 mm	180 mm	3000 mm
f_{eff}	30940 mm	14850 mm	247540 mm
Scale	0.12 arcsec/px	0.25 arcsec/px	0.015 arcsec/px
Slit length	< 4 arcmin	< 4 arcmin	≤ 0.5 arcmin
R_{lim}	10000	5000	37100

1.4 Description of ARGOS project

On the background of the current competition, the LBT consortium decided to equip the telescopes with an LGS facility that is tailored to improve on the unique capabilities of Lucifer, namely by providing substantial gain in ensquared energy (EE) over the full field of view, rather than aiming for high Strehl ratios in a narrow field.

The design drivers we identified in the Phase A study for ARGOS were:

- The wide field capabilities of LUCIFER MOS and imaging, leading to unique observations when combined with AO.
- The need for a reliable and low maintenance system, minimizing the technical risk and changes to existing telescope systems.
- The goal to realize an AO system as quickly as possible.
- The inclusion of possible upgrade paths towards diffraction limited operation.
- The aim to keep the AO system working even when the seeing is considerably worse than the average.

The available technical solutions led to the conception of a ground layer AO (GLAO) system based on multiple Rayleigh beacons per eye, delivering moderate but uniform Strehl ratios over a 4 arcmin FOV, and effectively at least doubling the efficiency of Lucifer for spectroscopy. This exciting feature is thus far unavailable on 8-m class telescopes. Furthermore, the gain in resolution from a GLAO system will enable a range of additional science cases to be addressed, e.g. in kinematics of high redshift galaxies or in crowded fields.

Another attractive feature of a GLAO system is the wide range of usable atmospheric conditions. Nearly independently of the seeing, GLAO should provide a stable reduction in PSF size. This was demonstrated in simulations of the system's performance for the Phase A study. The results are summarized in Figure 20, which shows a prediction of the seeing statistics of the LBT site with and without the AO system (top right panel). Here the seeing statistics without AO are estimated based on the statistics from the MMT. These are then used as the basis for the simulations to determine how the situation changes with AO. The top-left panel shows the predicted K-band Strehl ratio as a function of seeing FWHM. The plot on the bottom left translates this to a number for the encircled energy (EE) in H-band, and in the bottom right plot, one can see the difference in PSF shape with and without AO correction. Rather than aiming for diffraction limited PSF cores, GLAO improves the average seeing, effectively transforming a good site into an excellent site.

The layout of the baseline system is shown in Figure 21. ARGOS uses three Rayleigh beacons per eye, launched from behind the secondary mirror to minimize spot elongation in the outer subapertures. Industry proven commercial frequency doubled solid state (Nd:YAG) lasers are employed. Range gating is performed by Pockels cells, which are followed by one wavefront sensor (WFS) camera for all three beacons. A custom-made pnCCD from the Max-Planck-Institut Halbleiterlabor in Garching was chosen for the sensor as it delivers excellent sensitivity and noise performance. As deformable mirror ARGOS uses the adaptive secondary mirror (ASM). The conjugate height of the ASM is 96 m above the primary mirror, suitable for GLAO operation.

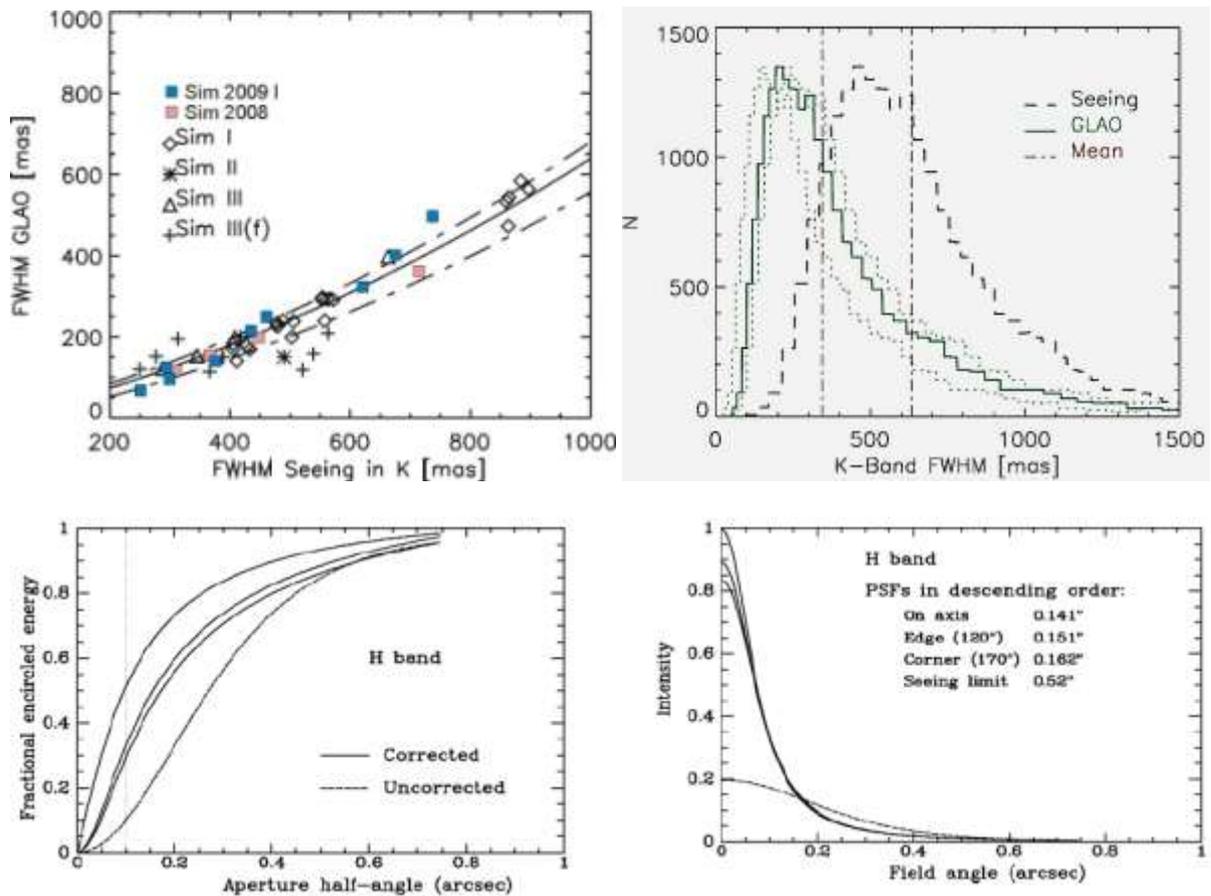


Figure 20: Simulations of the ground layer adaptive optics performance in various seeing conditions result in a range of performance estimates. On the top left a summary of all simulation results is shown. For a given seeing, GLAO with laser guide stars results in a decreased PSF size over the full field of view. The mean of these simulations and the seeing statistics from the MMT are shown in the plot on the right. With GLAO the mean FWHM of the distribution is reduced from 0.63 to 0.34 arcsec in K-band. The dotted lines denote the range of results that have been obtained in the simulations, showing that even better performance is possible. In the lower left panel, a comparison of the encircled energy as function of aperture radius is shown. The gain for apertures with radii of up to 0.4 arcsec is substantial. In the lower right panel the resulting PSFs from the simulations are compared to a seeing limited intensity distribution.

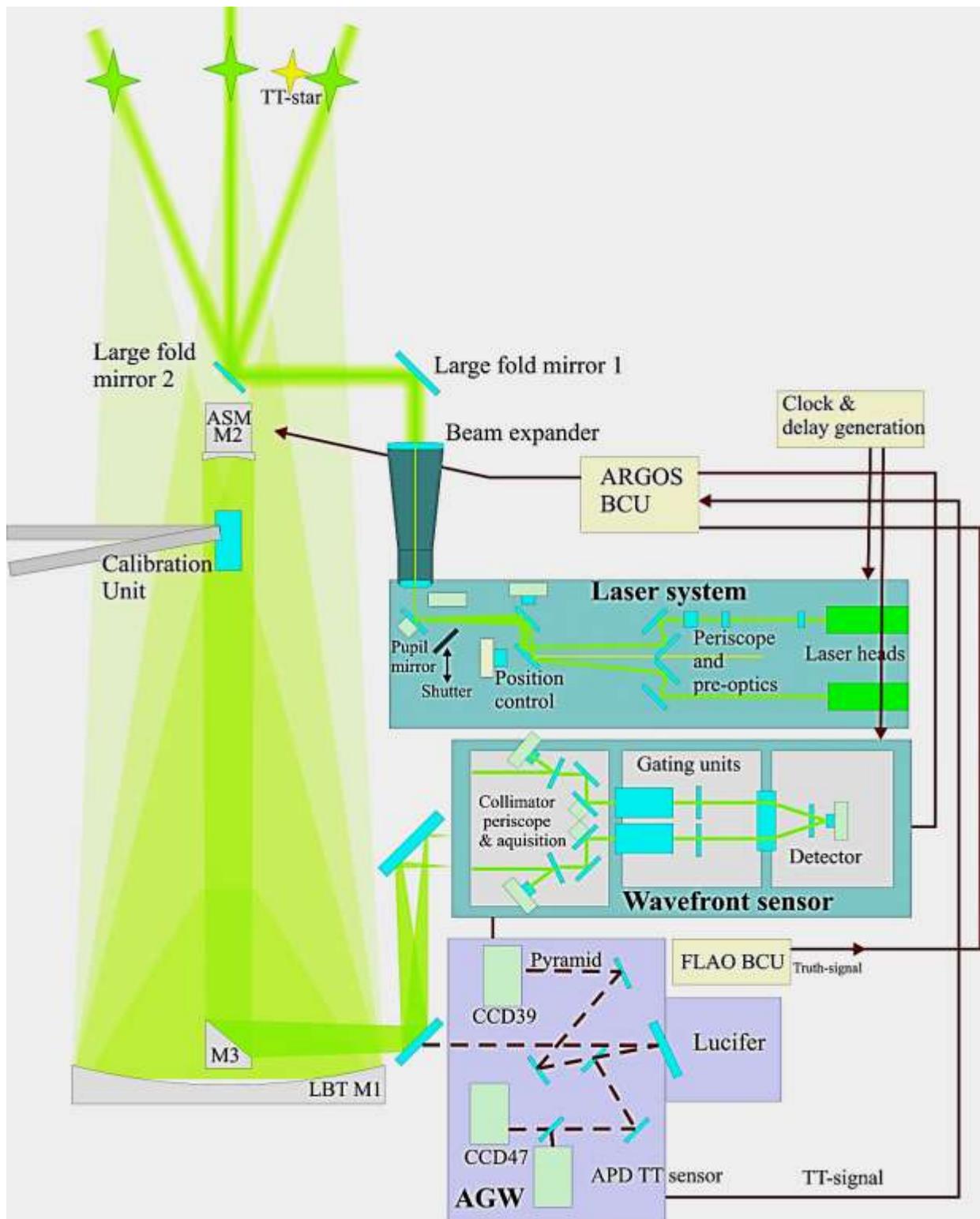


Figure 21: Schematic of the Argos system. A laser system contains all units to generate the laser beams. Followed by a launch beam expander and large fold flats to direct the laser beams to sky. On the detection side a dichroic beamsplitter separates the green light out and directs it to the Argos LGS wavefront sensor. The WFS images are transferred to the slope BCU (basic computation unit) and finally the BCU's on LBTs ASM calculate the required mirror shape.

1.5 Overview of Part I of the thesis

In Part I of this thesis I present the technical contributions I have made to the ARGOS project. In addition to an investigation into the launch of the laser beacons, I have been responsible for the development of a calibration scheme and its optics. In Section 2.1 I explore the effect of scattered light on the laser propagation. Possible solutions for the laser launch telescope are presented and compared in Section 2.2. The calibration concept I developed, with a detailed discussion of the optical and optomechanical design is described in Section 3.

2 Technical contributions

2.1 Scattered light impact from open air propagation

The lasers used to generate the beacons some 10 km above the telescope are necessarily very bright. The Rayleigh scattering we rely on to get a measurable return flux does not only happen in the high layers of the atmosphere; on the contrary, the air near the ground scatters even more than higher layers, due to the higher density and more solid particles (dust) per unit volume. The central launch makes it necessary to transfer all the laser light across the aperture to the launch optics behind the secondary mirror, causing a considerable amount of scattered light falling on the primary mirror.

To investigate the impact of the laser light scattered in the dome from the free air propagation to behind the M2 unit, a thorough calculation of the flux to be expected has been carried out.

The following is a detailed calculation to retrieve the Rayleigh volume scattering coefficient β , dependent of the viewing angle θ and the polarization direction α . Figure 22 shows the resulting scattering coefficient.

The calculations were carried out following the works of Penndorf (1957), Chandrasekhar (1950) and Hayano et al. (2003).

In accordance with the references, we define as follows:

β :	<i>Rayleigh volume scattering coefficient [cm⁻¹]</i>
β_θ :	<i>Angular Rayleigh volume scattering coefficient [cm⁻¹]</i>
n :	<i>Refractive index [1]</i>
σ :	<i>Rayleigh scattering cross section [cm²]</i>
N :	<i>Particle number density [cm⁻³]</i>
ρ_n :	<i>Depolarization factor, value for air is 0.035</i>

The indices $_s$ and $_0$ indicate standard conditions ($T_s=15^\circ C$, $p_s=1035 mb$, dry air) and zero conditions ($T_0=0^\circ C$, $N_0=2.68731 \times 10^{19} cm^{-3}$), respectively.

Penndorf gives the Rayleigh scattering cross section σ as

$$\sigma = (8\pi^3(n_s^2-1)^2/3\lambda^4N_s^2) \times ((6+3\rho_n)/(6-7\rho_n))$$

where the term in the second brackets, describing the influence of the anisotropic air molecules, is called the depolarization term.

From this the scattering coefficient is derived straightforwardly from the relation

$$\beta=\sigma N.$$

For a given scattering angle the angular Rayleigh cross section is

$$\sigma_{\theta} = (2\pi^2(n_s^2-1)^2/\lambda^4 N_s^2) \times (2+\rho_n)/(6-7\rho_n) \times P(\theta, \alpha)$$

which is derived from σ by dividing it by 4π , reducing the denominator $(6+3\rho_n)$ by 3, and multiplying σ by the normalized Rayleigh phase function $P(\theta)$.

The angular scattering coefficient follows directly:

$$\beta_{\theta} = (2\pi^2(n_s^2-1)^2/\lambda^4 N_s^2) \times (2+\rho_n)/(6-7\rho_n) \times P(\theta)$$

Substituting $\gamma = \rho_n/(2-\rho_n)$, $P(\theta)$ can be expressed as

$$P(\theta) = \{3/(4(1+2\gamma))\} \times \{(2\gamma+(1-\gamma)\cos^2\theta) + (1+\gamma)\}$$

where the term $(2\gamma+(1-\gamma)\cos^2\theta)$ corresponds to the parallel polarized fraction of the incident light and $(1+\gamma)$ to the perpendicular polarization.

For linear polarized light, the two polarization components can be expressed as

$$E_{\parallel} = 1/2 E(1+\cos 2\alpha) \quad \text{and}$$

$$E_{\perp} = 1/2 E(1-\cos 2\alpha).$$

By multiplying the corresponding terms in the equation for $P(\theta)$, we derive the phase function as a function of the polarization angle; the factor 1/2 is not to be considered, as already in the original equation, unpolarized light can be expressed by $E_{\parallel} = E_{\perp} = 1/2 E$.

$$\begin{aligned} P(\theta, \alpha) &= \{3/(4(1+2\gamma))\} \times \{(2\gamma + (1-\gamma)\cos^2\theta)(1+\cos 2\alpha) + (1+\gamma)(1-\cos 2\alpha)\} \\ &= \{3/(4(1+2\gamma))\} \times \{(2\gamma + (1-\gamma)\cos^2\theta) + (1+\gamma) + (2\gamma \cos 2\alpha) + \\ &\quad (1-\gamma)\cos^2\theta \cos 2\alpha - (1+\gamma)\cos 2\alpha\} \\ &= \{3/(4(1+2\gamma))\} \times \{(2\gamma + (1-\gamma)\cos^2\theta) + (1+\gamma) - \cos 2\alpha (1-\gamma)(\cos^2\theta - 1)\} \end{aligned}$$

This differs from the formula given by Hayano by a sign, most probably a typographic error in the reference. The shape of this function is shown in Figure 22

The scattering perpendicular to the beam propagation, $\theta=90^\circ$, is a sinusoidal function of the polarization angle α . It does not go down to zero, but reaches a lower limit due to the depolarization term. The fraction of light scattered varies with the polarization angle by a factor of 28.5 between $3.93 \times 10^{-19} \text{ cm}^{-1}$ and $1.38 \times 10^{-20} \text{ cm}^{-1}$. Given the propagation distance of 3.3 m across the primary and an assumed maximum of 60 W at a wavelength of 532 nm, this amounts to 723 photons per second and arcsecond squared of scattered light, for the polarization parallel to the scattering direction, from propagating the laser to optics behind the secondary. On the LGS wavefront sensor, running at a kHz frame rate and distributing the light over the re-imaged pupil the additional background will be negligible.

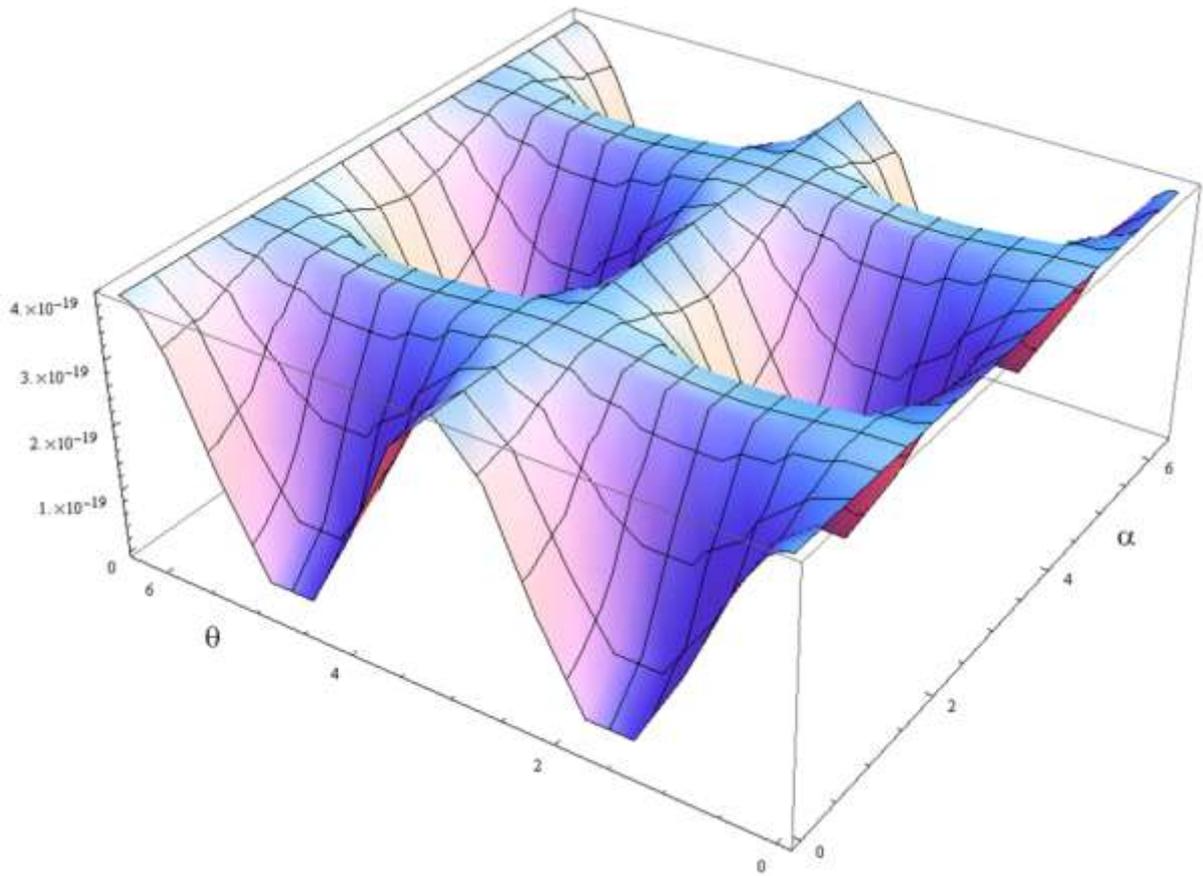


Figure 22: Scattering coefficient as a function of polarization angle α and viewing angle θ . The fraction of light scattered varies with the polarization angle by a factor of 28.5 between $3.93 \times 10^{-19} \text{ cm}^{-1}$ and $1.38 \times 10^{-20} \text{ cm}^{-1}$.

2.2 Laser launch telescope design

2.2.1 Introduction

During the Phase A study, I developed an alternative concept for a laser launch telescope (LLT). The LLT is a key part of a laser guide star facility. It projects the laser into the atmosphere, at the desired position in the field, and acts as a beam expander for the small Gaussian beam the laser delivers. The main concerns are the transmission of the system, the achieved spot size in the atmosphere, and the stability of optical alignment.

Simulations show that a beam diameter at the exit aperture of the order of 3 times the Fried parameter r_0 is optimal for the projected spot size, balancing the seeing effects against the diffraction. Depending on the seeing, this leads to an aperture size of up to 450 mm for typical conditions at the LBT site (compare also Figure 42). At the time of the Phase A study, we planned to adapt the beam size to the atmospheric conditions, and estimated a factor of three in necessary expansion ratios.

To minimize the effects of spot elongation due to the finite vertical length of the measured beam in the atmosphere, projection from the centre of the aperture that is to be corrected is optimal. The natural solution is to project the laser beam from behind the secondary mirror, as done at the VLT, the WHT or the MMT, for example. As space and load is limited at that location, the LLT has to be either folded, or mounted somewhere else, and the beam then transferred to a fold mirror behind the secondary mirror. The other mentioned facilities all opted for a folded design behind the secondary. At the WHT and the MMT folded refractors are used, while the VLT uses a Cassegrain type reflector. Difficulties with the VLT design have been reported, probably due to its very fast primary mirror.

2.2.2 Overview of options

The baseline design for the Phase A study was a big refractor with a slow focal ratio, to

- make alignment easier and more stable, and
- provide excellent performance over a large field of view.

As the laser is monochromatic, chromatic aberrations are to first order not a concern in the design. The front lens for a refractor of these dimensions is not trivial, though.

To be able to compare the expected performance and the estimated design effort to that of a reflector, I made an alternative study based on commercially available optics and telescope assemblies of suitable size.

Telescopes in this size range are available off the shelf for advanced amateurs and small observatories. Designs include Newton, Cassegrain, Ritchey-Chretien and Dall-Kirkham, and modified versions of these. Even refractors are offered, but are very expensive and heavy. Prices for the reflectors range from approximately 15 kUSD to 60 kUSD.

2.2.3 Throughput calculation

All these options have in common that they have a large central obstruction if they are designed to deliver a large field of view. To assess the light loss due to central obstruction of

a Gaussian beam, I calculated the integrated transmission of an annular aperture with a transmission of 100%, when illuminated with a Gaussian intensity distribution, as a function of relative beam size and relative obstruction diameter.

The expression for the relative transmission of an annular aperture for a Gaussian beam is:

$$T(d, D) = \frac{1}{\pi} \int_0^{2\pi} \int_{d \times D \times \xi}^{D \times \xi} e^{-r^2/2\sigma^2} r \, dr \, d\varphi$$

where $\xi = 2\sigma$, D denotes the aperture diameter in units of the $1/e^2$ -diameter of the Gaussian beam, and d the relative obstruction diameter. The shape of the function $T(d, D)$ is shown in Figure 23.

For example, the light loss due to truncation at a circular aperture of 1.4 times the $1/e^2$ -diameter is

$$L(D = 1.4) = 1 - \frac{1}{\pi} \int_0^{2\pi} \int_0^{1.4 \times \xi} e^{-r^2/2\sigma^2} r \, dr \, d\varphi = 1.8\%$$

One can see that for every value of central obstruction there is an optimal truncation radius D to balance the losses on the inner and outer limit of the aperture. For obstruction sizes around 15%, D is roughly between 1.8 and 2.2 sigma (see Figure 24).

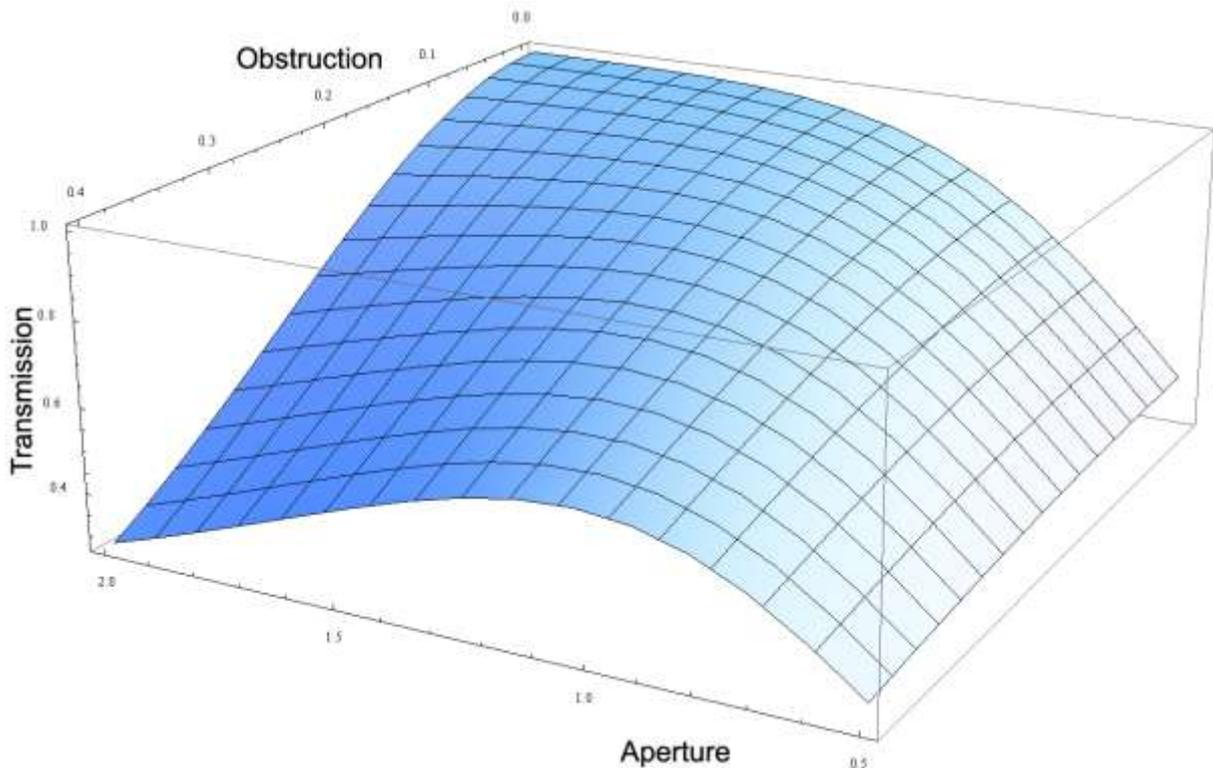


Figure 23: Transmission as function of truncation radius and central obstruction.

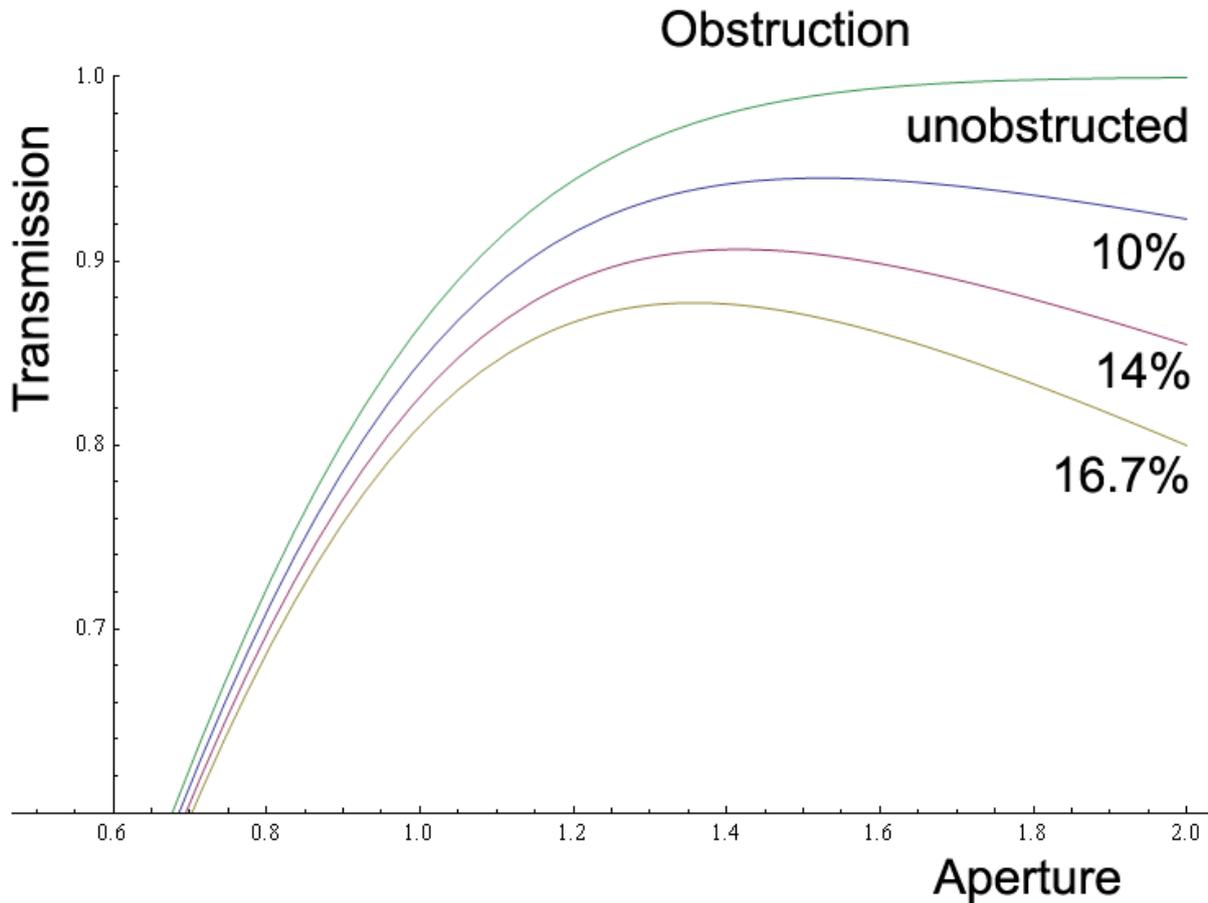


Figure 24: Transmission as function of truncation diameter for obstructions of 16.7, 14 and 10%, and for the unobstructed case.

For the aperture size, primary focal ratio and field of view we are looking for, the two mirror systems become exceedingly troublesome for obstructions of less than 10%, due to the high magnification introduced by the secondary mirror. A reasonable compromise between the unavoidable losses and complexity of the optical design seems to be a system with better than 85% transmission at the aperture.

2.2.4 Options based on commercial telescopes

There are no commercial telescopes available with such a small secondary mirror. I tried to adapt a commercial design to our needs by changing only the secondary. Of the commercially available options, a modified Dall-Kirkham seemed to be suited best for this, as the secondary is spherical. This makes the system less prone to misalignment, as there is no preferred tilt orientation of the secondary, but just a lateral shift to adjust. Another positive side effect is that the custom secondary becomes an easily manufacturable piece of optics.

One possible vendor who manufactures suitable telescopes is PlaneWave. Joe Haberman of PlaneWave was willing to cooperate with us, and do a custom version of their standard 20" telescope. He shared the optical system data for further investigation.

Mechanically, the optical tube assembly (OTA) is based on a double truss construction of carbon fibre reinforced plastic (CFRP) tubes (Figure 25). This design is very stiff and shows little thermally induced focus drift. Mounting is provided by one or two dovetails, which are thermally compensated to minimize mechanical stress introduced by coefficient of thermal expansion (CTE) mismatch. The two dovetail option, made for mounting the OTA to a fork mount, would be an ideal interface for installing the telescope on top of LBT's ASM. As the 20" primary makes up nearly all the mass of the OTA, the centre of gravity can be kept very close to the ASM support structure, leading to reduced torque in the ASM swing arm for different elevation angles. Indeed, as a custom version made for us would necessarily have a much smaller secondary mirror, the top ring of the telescope would be lighter, providing even better stiffness and alignment stability. In the standard version, the OTA weighs about 62kg.

In the following, I present options for modifying the 20" PlaneWave telescope to our needs. I investigate a design with a modified spherical secondary, a design with a conicoid secondary, and a design with a spherical secondary mirror, but an elliptical primary with different conic constant. All these designs include a custom subaperture corrector in front of the focal plane. Lastly, I included the design of the refractor that was chosen at the preliminary design review (PDR), for comparison.

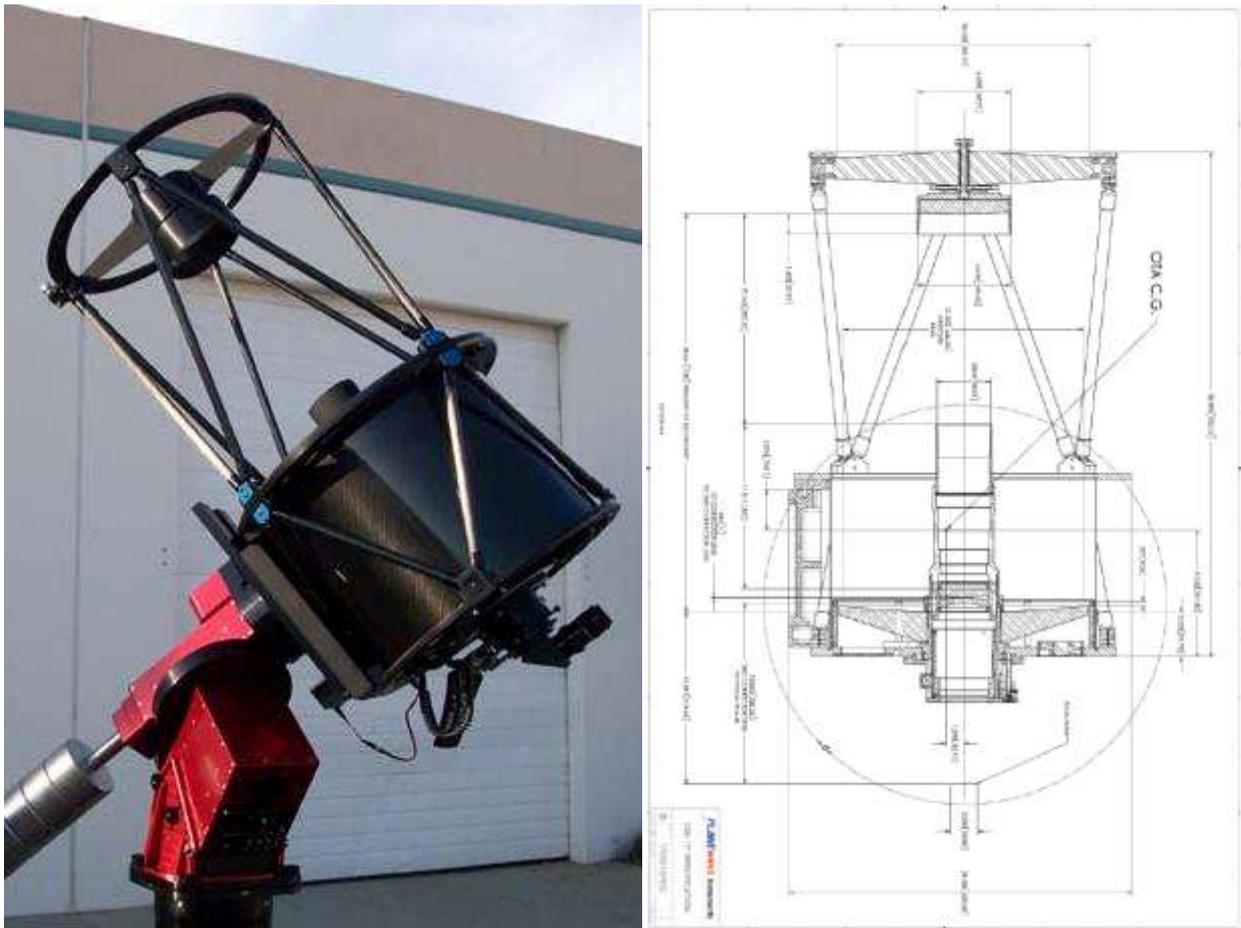


Figure 25: The 20" PlaneWave telescope in standard configuration. Picture courtesy of PlaneWave.

2.2.5 Optical designs

The optical design of the standard telescope from PlaneWave, called 20" CDK, consists of an elliptical primary with 508 mm aperture and a focal ratio of $f/3$, a spherical secondary mirror with 190 mm diameter, and a refractive two element corrector in front of the focal plane, resulting in a focal ratio of $f/6.8$. The primary is made out of Pyrex, and the substrate is oversized by 0.5". The large obstruction of 39% makes the standard version unsuitable for our purpose.

If we restrict the field of view to a few arcminutes, a much smaller secondary mirror becomes viable. First I investigated the performance that is possible if one keeps the primary's curvature and changes the secondary mirror and the corrector, leaving the surfaces spherical. Of course, the distances need to be changed accordingly.

Modified Dall Kirkham with 17% obstruction

With the constraints above I found a system achieving excellent performance with a central obstruction of 85 mm (16.7%, see Figure 26). The two-element corrector is located near the primary and can be mounted in a baffle tube (Figure 27). The corrector can conveniently seal the beam path up to that location, to keep the system free of dust and insects. This is particularly important where the beam size is small and the power densities consequently very high, especially in the focus. The distance of the lenses to the focus is such that the beam footprint on the lens surfaces is large enough for the power density to have dropped to an uncritical level. The optical design is diffraction limited over the envisaged FOV (see Figure 28). The wavefront error in forwards propagation is less than $\lambda/10$ PV for a field point 2 arcminutes off-axis (Figure 29), and still better than $\lambda/4$ for 4 arcminutes distance. The same field is unvignetted (Figure 31). The focal ratio is $f/22$.

With this design, the nice feature that a Dall-Kirkham is fairly insensitive to misalignment of the secondary due to its spherical form is preserved. All the custom optics are easy to manufacture to very high precision. The total cost for the additional optics (one 85 mm mirror and two 50 mm lenses), should not exceed a few thousand USD, and the necessary mechanical modifications (mainly the secondary spider and longer struts due to the changed distance between primary and secondary mirror) would also cost a fraction of the price of the optical tube assembly. Joe Haberman stated that the company would be willing to manufacture a custom model based on a design like this, and as a first estimate informed me the additional cost would probably range between 5000 and 10 000 USD.

The plots show that the performance of the modified Dall-Kirkham is sufficient for our needs. The drawback is the still substantial obstruction of 16.7%.

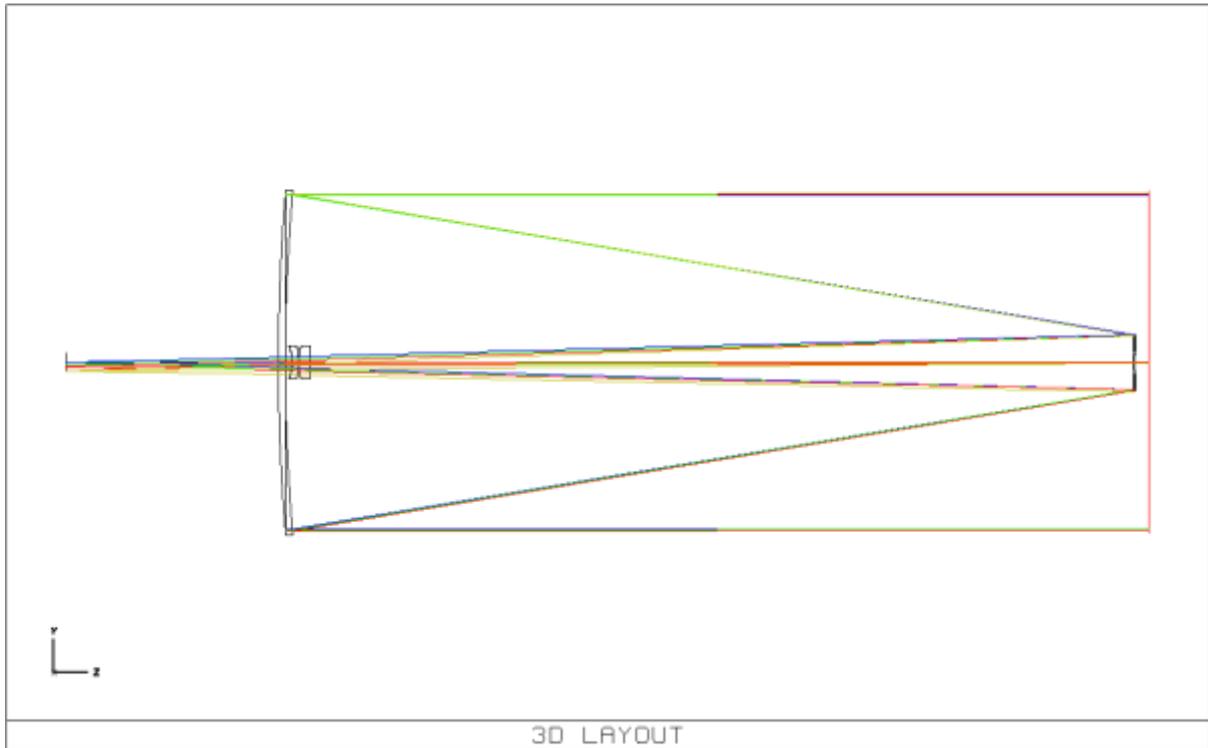


Figure 26: Optical layout drawing of the modified Dall Kirkham with 16.7% obstruction. On the right hand side, the Entrance aperture and the secondary mirror, on the left the primary mirror and the field corrector directly in front of it.

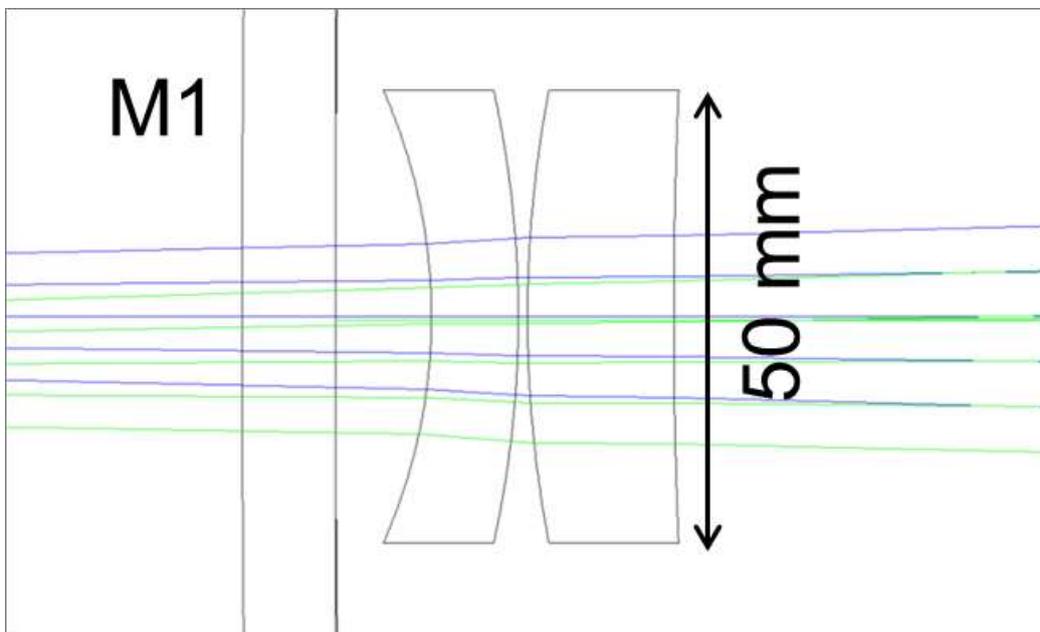


Figure 27: Detail of the two-element field corrector in front of the primary mirror. The surfaces are all spherical, the lenses made out of BK7.

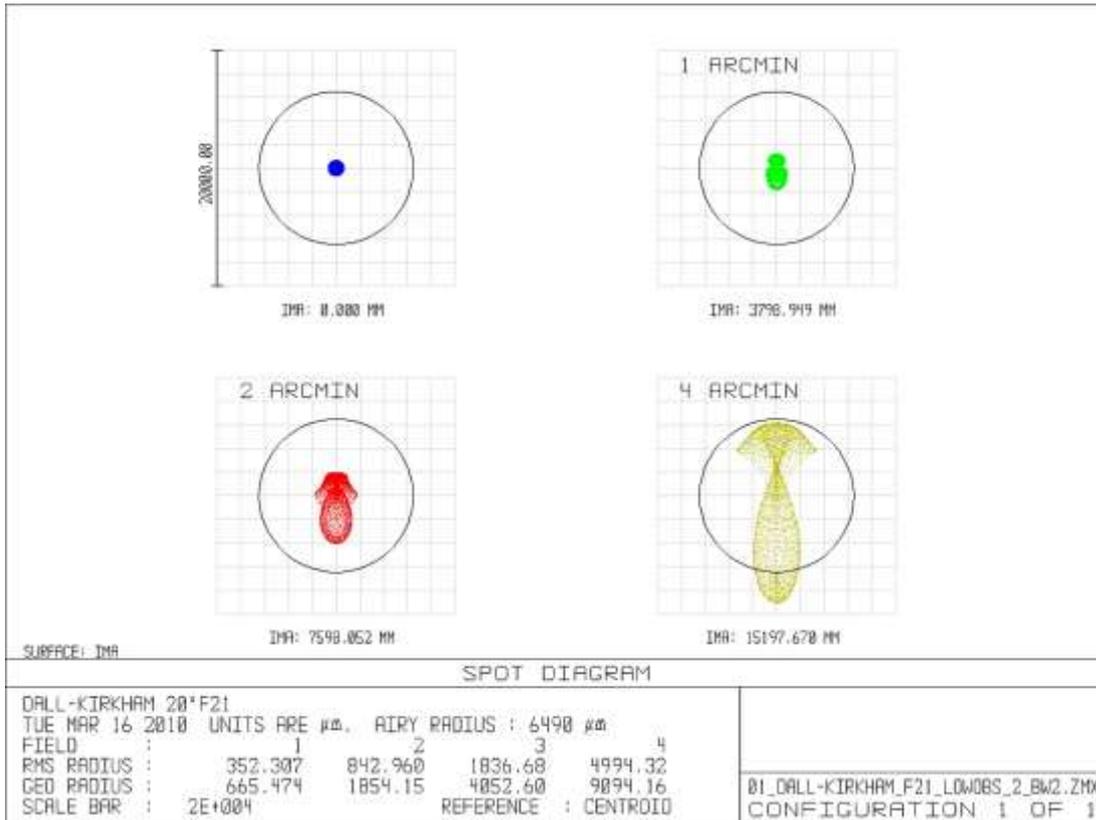


Figure 28: Spot diagrams over a field of view with 4 arcminutes radius. The design is diffraction limited.

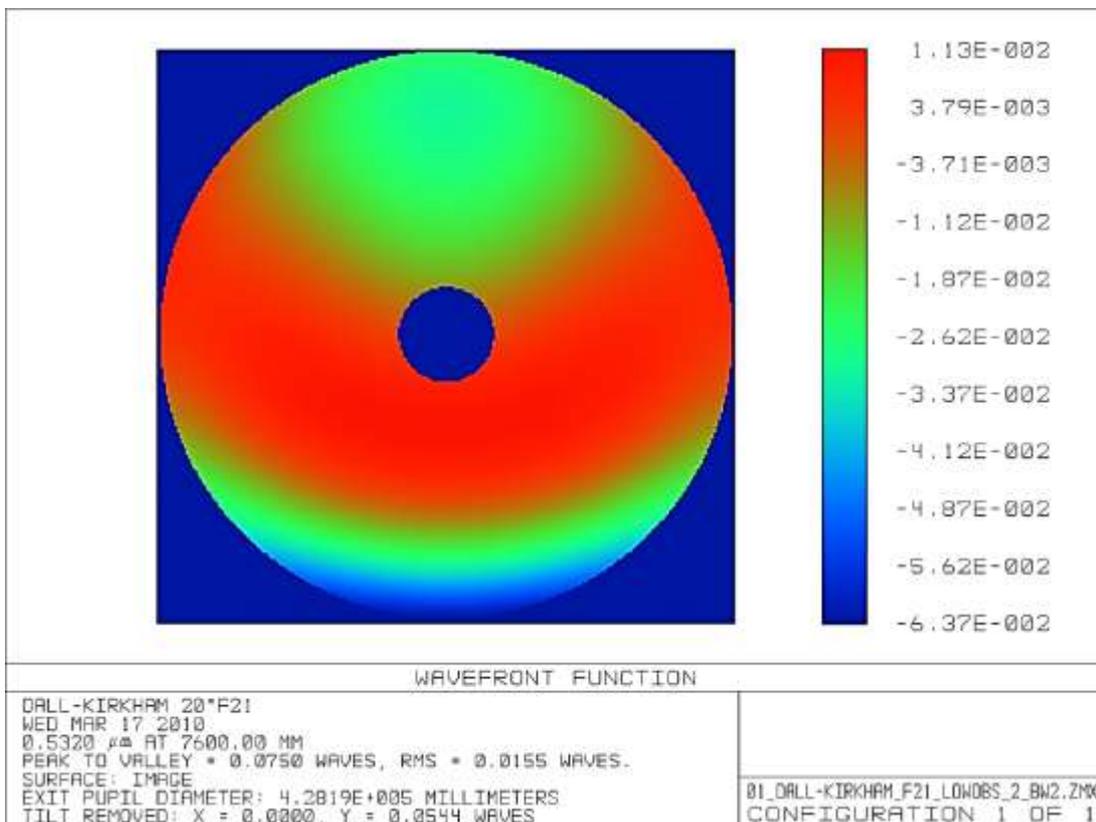


Figure 29: Wavefront of the image in 12km altitude, for the spot 2 arcminutes off axis. The wavefront error is $\lambda/13$ peak to valley.

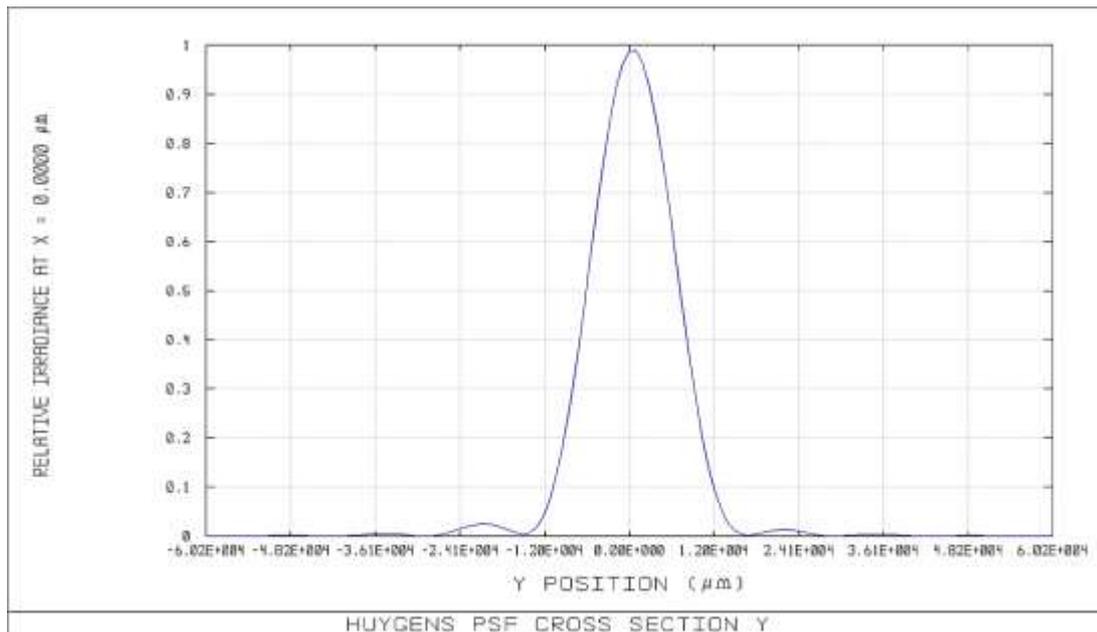


Figure 30: The corresponding PSF to Figure 29. Crosscut in radial direction.

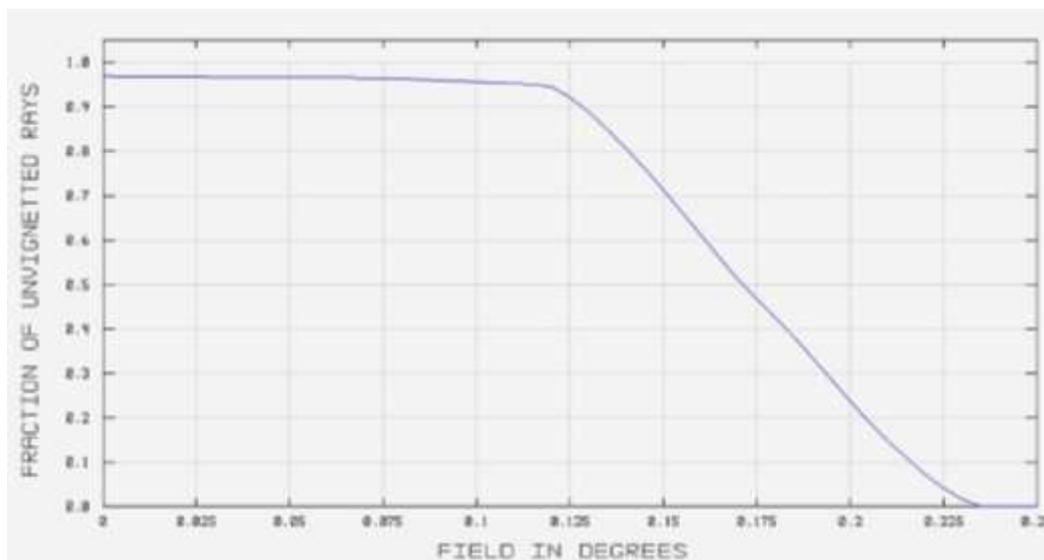


Figure 31: Vignetting due to the small secondary mirror. Light loss sets on at 7.5 arcminutes radius.

Modifications with elliptic secondary mirror and 14% obstruction

To achieve similar good performance with smaller obstructions, the design must be modified further. Giving up on the spherical shape of the secondary mirror allowed me to sketch a design with 14% obstruction. The primary mirror is still the same as in the standard CDK telescope, and I again use a two element corrector. The secondary is a prolate ellipse with a conic constant of about 0.5. The optical performance of the system is very good (Figure 32 and Figure 33); however, the convex ellipse is much harder to manufacture to high precision than a spherical mirror, partly because it is complicated to measure. This leads to considerably higher costs. Furthermore, the system's alignment is not invariant to tilt about the secondary mirror's centre of curvature anymore.

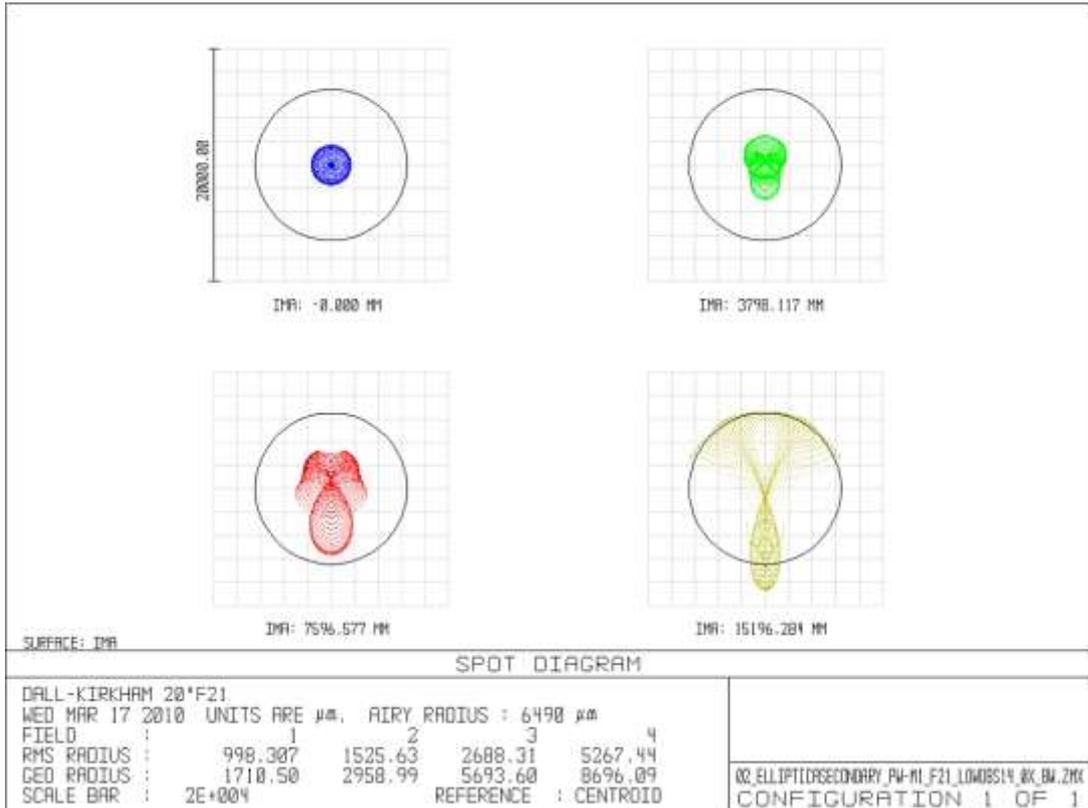


Figure 32: Spot diagrams for the modified system with an elliptical secondary and 14% obstruction. Again, the design is diffraction limited.

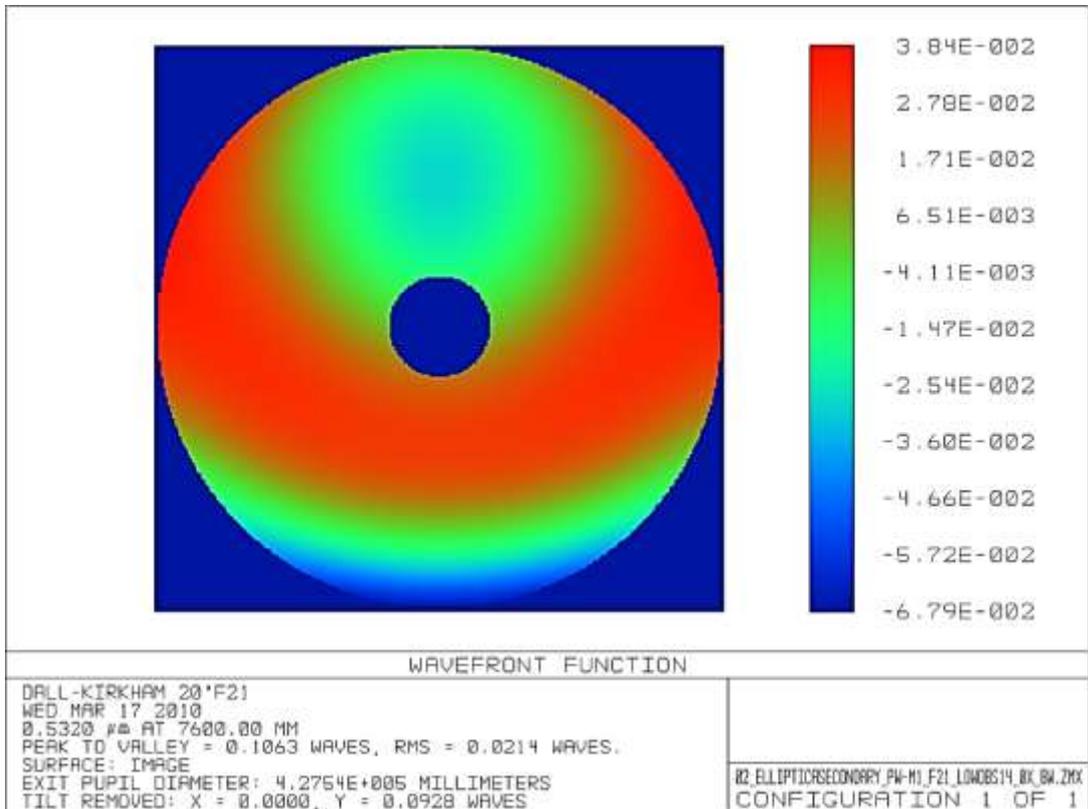


Figure 33: The resulting wavefront for a spot 2 arcminutes off axis. The wavefront error became slightly larger compared to the design before, and is $\lambda/10$ peak to valley.

Dall-Kirkham with modified primary mirror

To be able to reduce the central obstruction even further, one must deviate from the original primary mirror's shape. Of course, this leads to many options for the design of the system. Here I include a design very close to the CDK design. I re-optimised only the primary's conic constant for a system with 10% obstruction and a focal ratio of $f/36$. Naturally, it is still an ellipsoid, now with a conic constant of 0.9. The performance is comparable to the first design I showed (Figure 34). As the primary's diameter and focal ratio stay the same, PlaneWave's optical tube assembly could still be employed. The custom made primary adds to the manufacturing costs, but not necessarily very much more than for an elliptical secondary mirror, as the large concave ellipse should be easier to test. Due to the spherical secondary, this design is again robust against misalignment. The dependence on a custom made elliptical primary mirror imposes a risk, which has to be balanced against the advantages of a small obstruction.

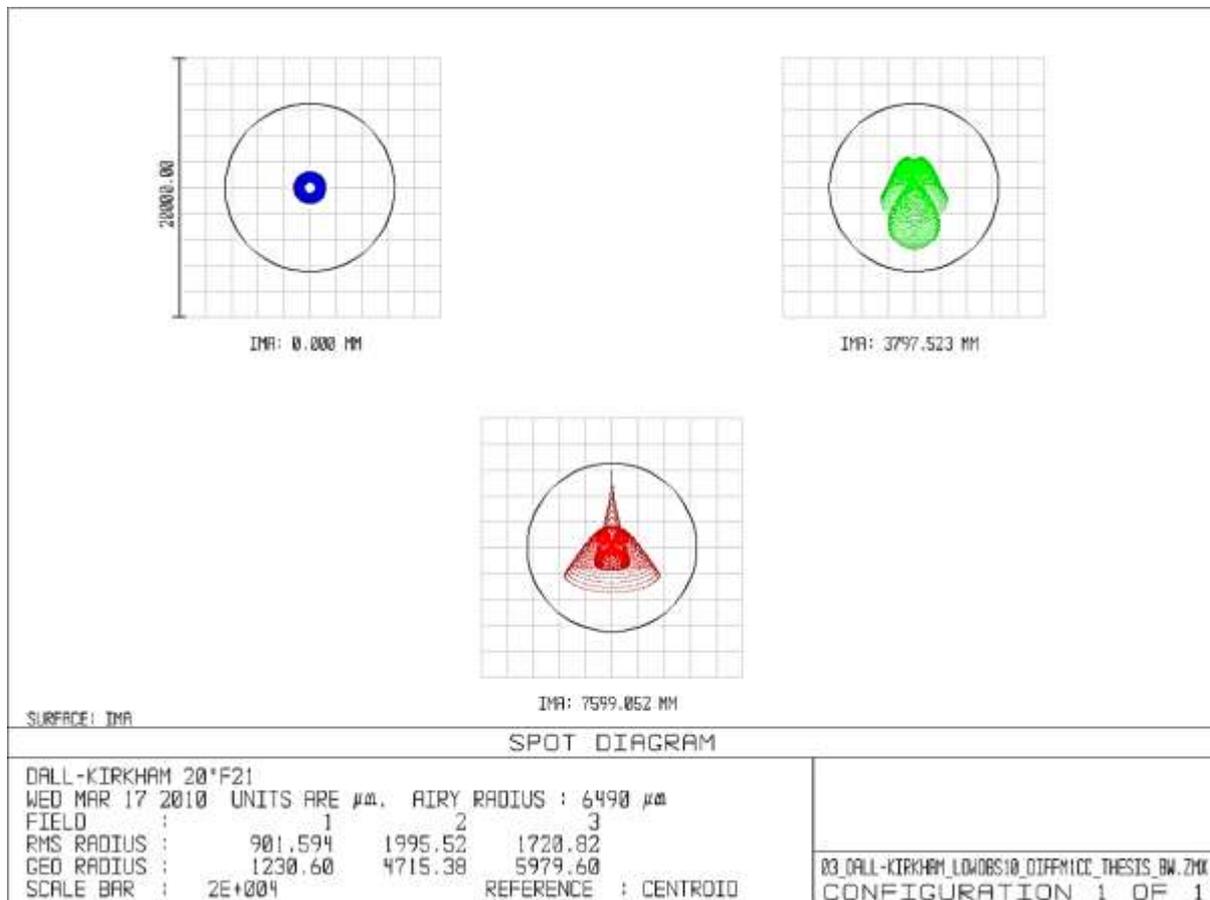


Figure 34: Spot diagram for the modified Dall-Kirkham with 10% obstruction. The image quality degrades quickly for larger field radii than 2.5 arcminutes; hence the spot at 4 arcminutes field radius is not shown. For the relevant position at 2 arcminutes the design is diffraction limited.

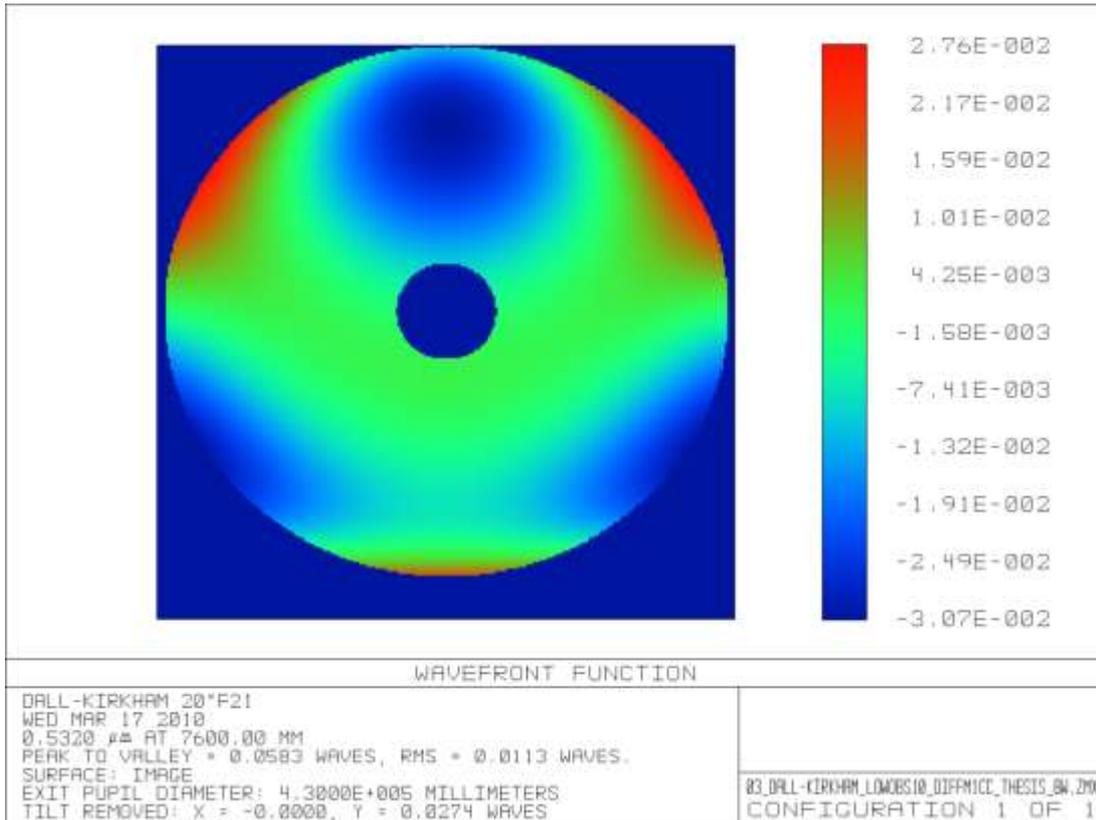


Figure 35: The wavefront for the spot at 2 arcminutes radius. The wavefront error is the lowest so far, $\lambda/17$ peak to valley.

2.2.6 Comparison with the refractor design

For comparison, I show the optical design of the refractive LLT (see Figure 36 -Figure 41). The design is based on a single, custom made Fused Silica front lens of 450mm diameter (and 400 mm clear aperture). One surface is aspheric, while the other was kept plane for testing and alignment purposes (see Figure 37). The focal ratio of this large refractor is f/21. The LLT is planned to be mounted vertically in the central truss structure of LBT, between the primaries. Optical performance for the refractor is excellent. The mounting of the large front lens is problematic, however. The tight tolerances require a custom made lens cell, counteracting the changing gravity vector in different elevations, and athermalized. Testing the aspheric surface with such a long radius of curvature is not trivial, and neither is the polishing, which results in a price for the two lenses of 160.000 USD.

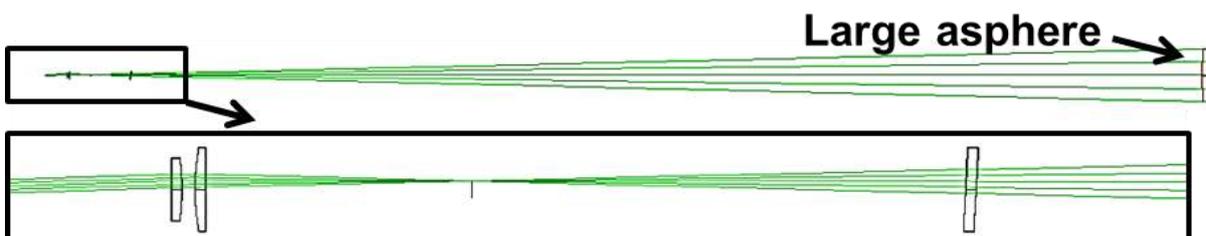


Figure 36: Layout of the refractive beam expander telescope. The plano-convex asphere with 450mm diameter is on the right and the beam insertion optics with the intermediate focus on the left.

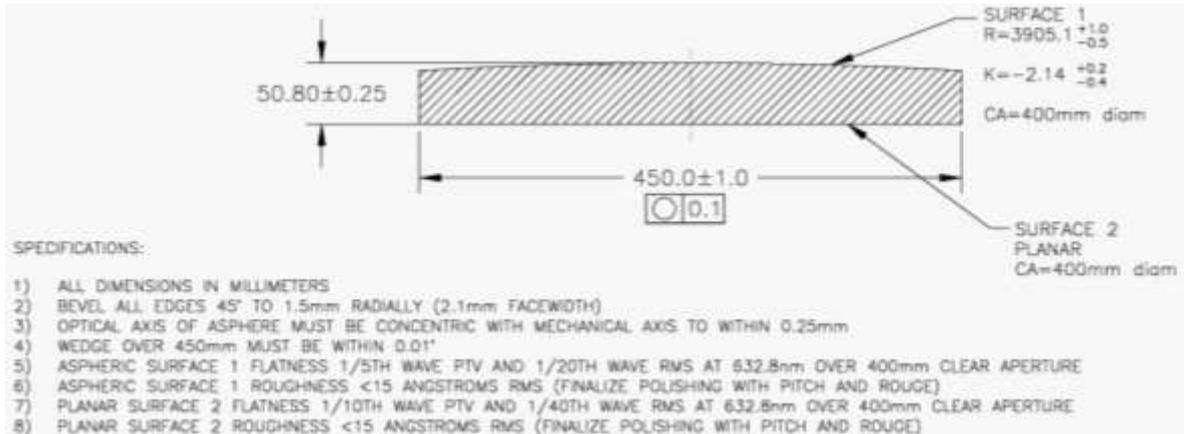


Figure 37: Manufacturing drawing for the aspheric front lens of the refractive laser launch telescope.

Chromatic defocus is huge as the design is not colour corrected. While not a problem for the launch of the monochromatic laser, a bore-sighted camera becomes harder to employ. A possible upgrade to a sodium laser can be launched, as the focus for another single wavelength can be adapted. As the launch telescope expands the beam before it gets transferred to the volume behind the secondary, two large folding flats are necessary. These have approximately the dimensions of LBT's tertiary mirrors, and have similar specifications for the surface figure. The tertiary mirrors use Hextek borosilicate glass substrates, with a thin facesheet fusion-bonded to a honeycomb-like core (Figure 38). Using the tertiary mirrors as examples, the folding flats will cost about 100.000 USD in total.

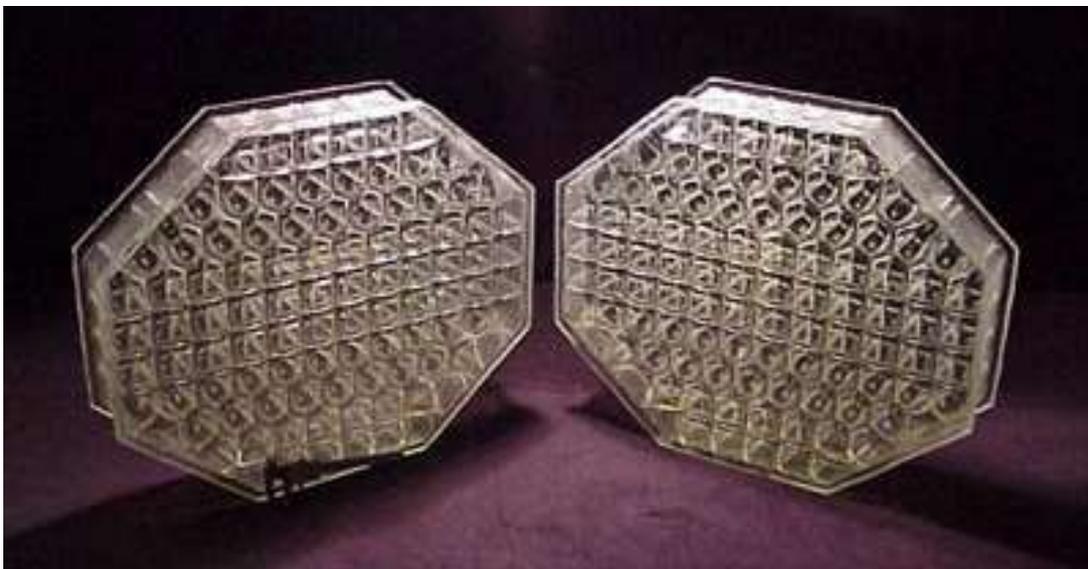


Figure 38: Light-weighted Hextek borosilicate substrates as used for M3.

The main advantages of the refractor design are its unsurpassed throughput and relaxed alignment tolerances. As it turned out later, with the optical components separated over the whole telescope structure, it is hard to capitalize on that fact. It is complicated to adjust the position of the large optics in front. In addition, a new and customized interferometer was deemed necessary to perform and check the alignment of the front lens with respect to the field lens optics. This alone would add additional costs of roughly 100 000 USD.

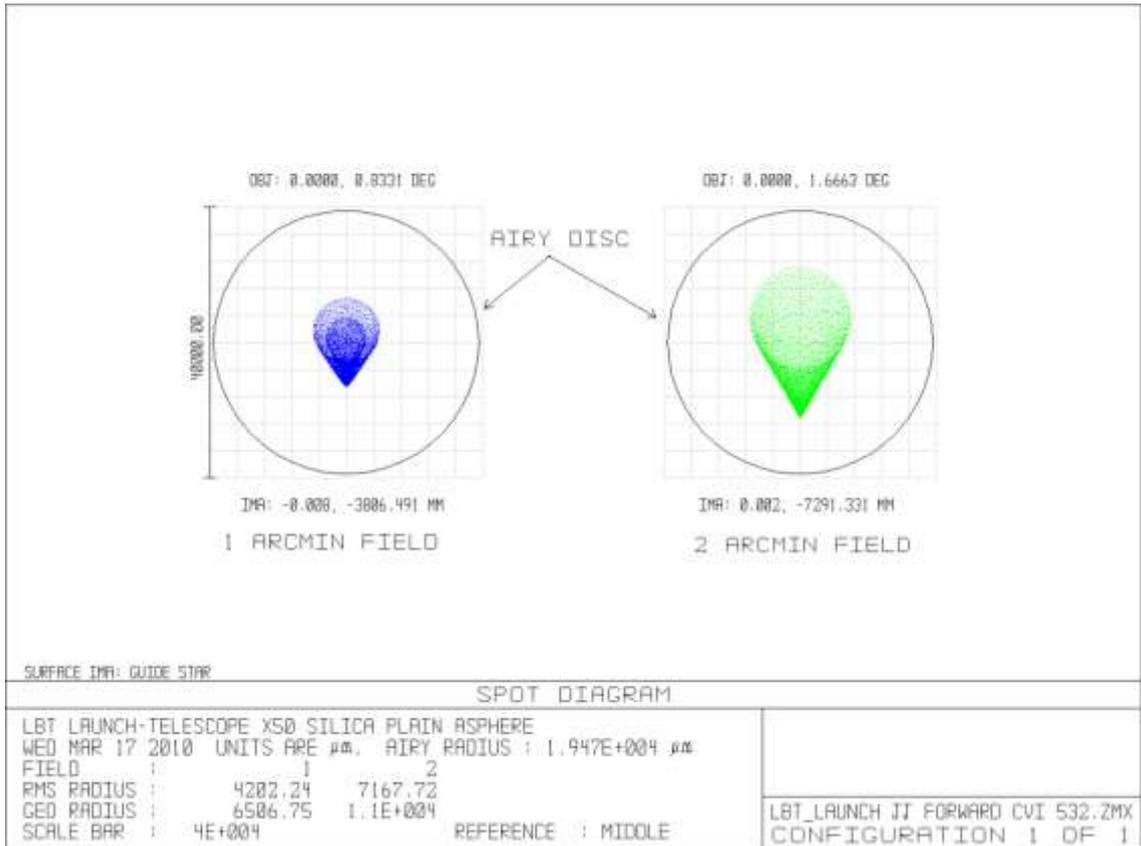


Figure 39: Spot diagram of the refractive beam expander as taken from the ARGOS documentation. It is diffraction limited.

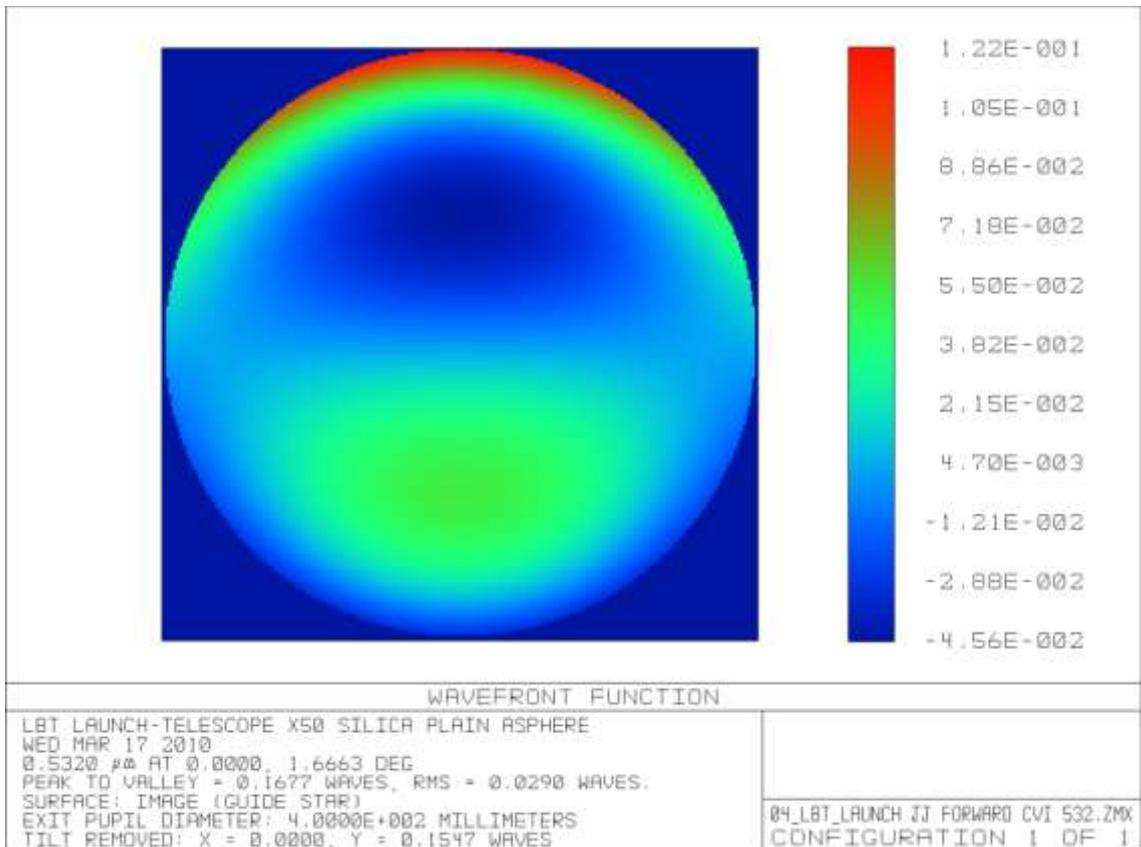


Figure 40: Wavefront at 2 arcminutes field radius; the wavefront error is $\lambda/6$.

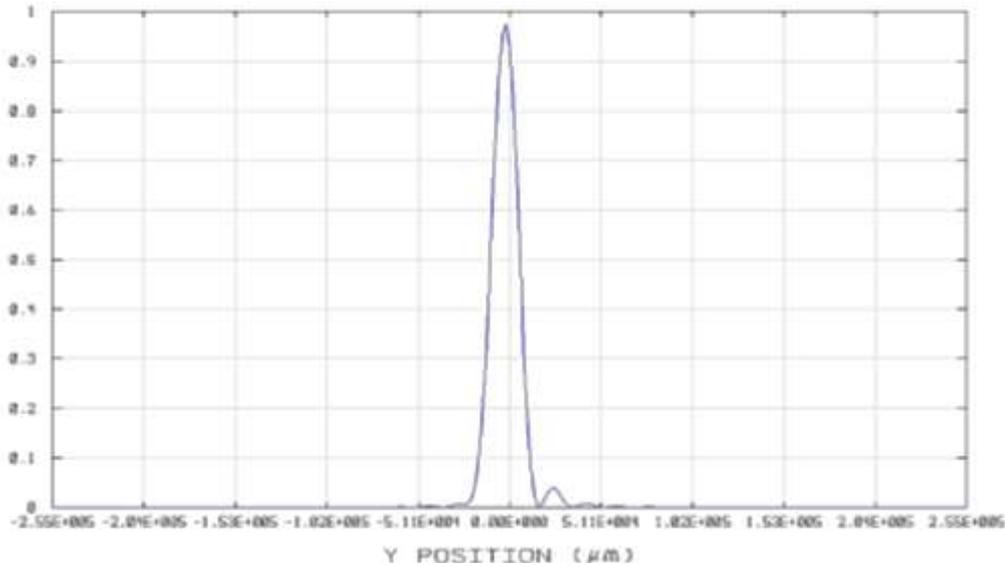


Figure 41: The corresponding PSF of the refractor.

2.2.7 Discussion and conclusions

This investigation was carried out during the Phase A study, after which it was decided to adopt a refractor as design. The reflector design was not updated subsequently, although some critical design parameters have since changed. The most important of these is the beam size. Further investigation of the achievable FWHM of the laser beacon on sky, as a function of the beam size at the LLT exit and the seeing, and simulations of the results of the wavefront fitting error on the WFS led to the conclusion that changing the beam diameter to adapt for different atmospheric conditions is not necessary. Figure 42 shows the results from these simulations. A fixed beam size of 250mm $1/e^2$ was adopted.

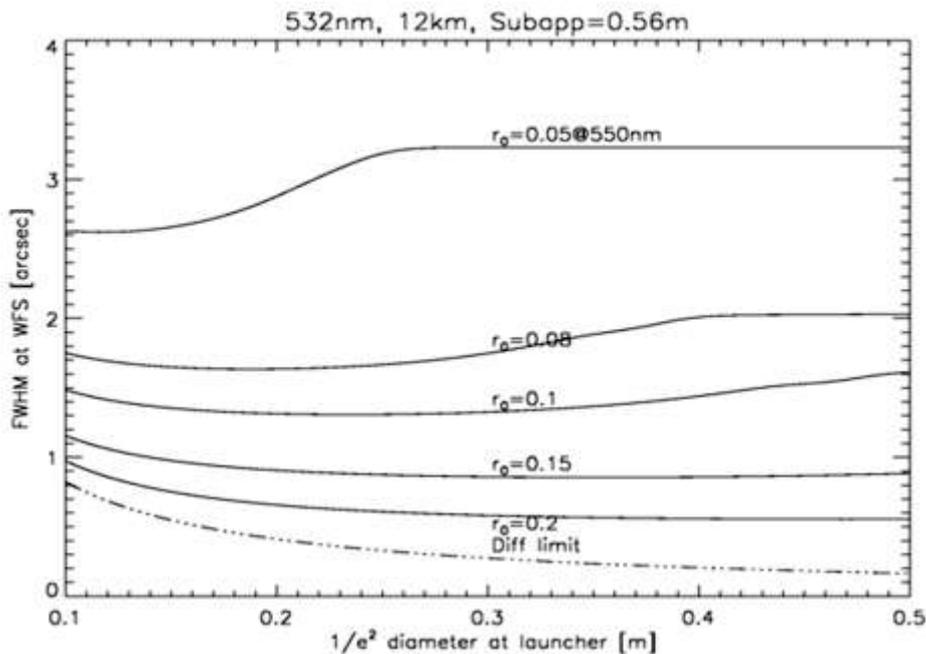


Figure 42: Spot diameter as seen on the wavefront sensor for various r_0 values. In bad seeing conditions a smaller launch beam gives smaller spots on the wavefront sensor. In good seeing conditions the differences are rather negligible. *From the Phase A study.*

To achieve negligible light loss the aperture of the refractor was set to 400 mm, but with an aperture of about 330 mm the losses due to clipping already amount to less than 2%. With these constraints, a reflective system becomes even easier to design, and there are more commercial options. To give an example similar to the options detailed earlier: a 12.5 inch (320mm) model of the CDK system is available with a closed CFRP tube for less than 10,000 USD. The ratio between the desired $1/e^2$ beam diameter of 250 mm and the aperture of a 12.5 inch reflector is 1.3, close to the optimum value for obstructions of 14%.

As the aperture gets smaller, it becomes easier to correct for the aberrations, so the relative central obstruction can be made smaller with the same quality of the wavefront. The whole optical tube assembly becomes very light, and reasonably compact. At this aperture size, and considering the fully enclosed optical tube assembly, a plane parallel plate as a window, sealing the tube and holding the secondary mirror, becomes an attractive option. This would prevent a small beam diameter from being accessible to insects, for example.

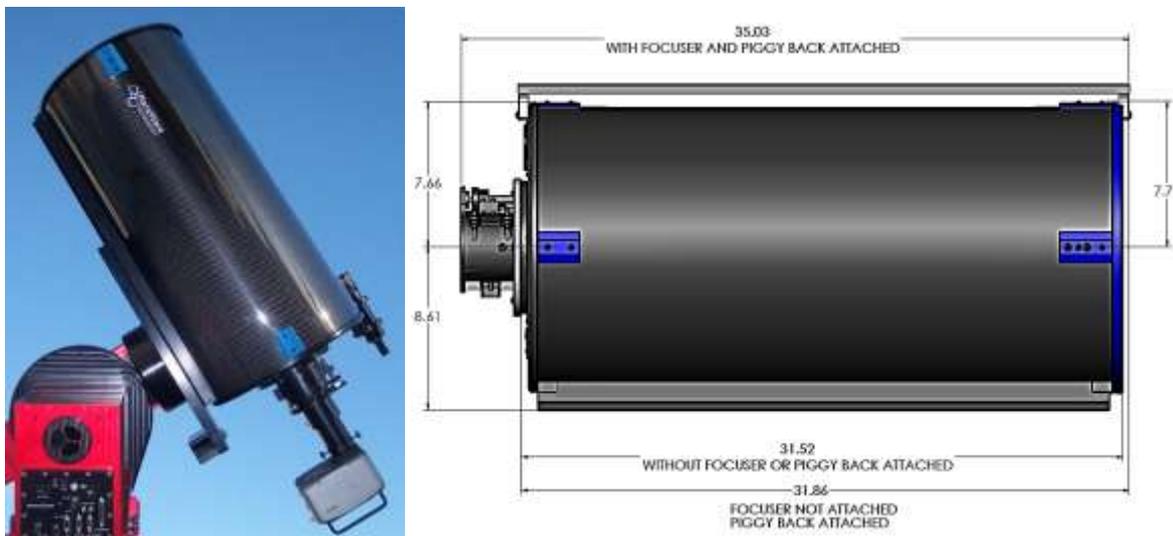


Figure 43: A picture and drawing of the 12.5 inch CDK model from PlaneWave. The OTA has a length of 800mm and weight about 20kg. The tube is made from CFRP. *Courtesy of PlaneWave.*

Based on this investigation, a reflective laser launch telescope seems like a good alternative to the refractor design. The central obstruction imposes an additional light loss that can be kept to about 8% with modifications from a commercial design. This is balanced by a considerable reduction in the cost of the optical components. Depending on the available coatings, especially for the large folding flats of the refractor design, some of the light loss could also be regained due to the fact that the smaller folding mirrors of the reflective LLT can be obtained with dielectric laser line coatings of highest reflectivity. Furthermore, a laser launch telescope based on a commercially available telescope carries much lower risk as its key parameters (for example, the mechanical behaviour of the optical tube assembly and transmission) can easily be tested. In contrast, the large optical elements of the refractor design are all custom items with a long lead time. This imposes a considerable risk for delays, should anything go wrong during the procurement process. Damage during transport and especially a possible fatal failure of a dielectric coating were identified as critical concerns. The price for the custom optics prohibits the procurement of spares to mitigate the risks. On

the other hand, the custom parts of a modified reflective LLT are all comparatively easy to manufacture and inexpensive, allowing for spare items to be bought and reducing the likelihood of delays. The reflector design is more sensitive to misalignment of the optical surfaces; however, it is much more compact, and the optical components are not distributed across the telescope structure.

Another difference between the two designs is that in case of the refractor, the expanded beam has to be propagated in free air across the primary mirror. If the expansion takes place behind the ASM, the beam can be enclosed in a thin tube. While we don't expect problems from the increased background (see Section 2.1), an enclosed beam would allow for an arbitrary orientation of the polarization. The expansion behind the ASM has another subtle advantage: a beam expander demagnifies angles of incoming to outgoing rays with respect to the optical axis by the expansion factor, reducing the effect of vibrations of the ASM swing arm, as the expansion takes place after the vibrations have affected the beam.

The setup has the drawback that more optomechanical parts, including active components like tip-tilt mirrors and cameras, would have to be located behind the secondary, where they are hardly accessible, and space is very limited.

A reflective design has several attractive advantages over a large refractor. To finally conclude how this compares to the additional light loss, a full trade-off study, including the implications for engineering, installation, and operations, would be necessary.

3 The ARGOS calibration scheme

3.1 Introduction

The current and future generations of adaptive optics systems on very large telescopes are very complex systems, using a great number of optical, mechanical and electronic interfaces to control the wavefront deformation by atmospheric turbulence to a very high degree (Cuby et al. 2008, Boyer et al. 2008, Arsenault et al. 2006). To achieve the desired performance, in some cases up to the diffraction limit, it is essential to be able to calibrate the full system at the telescope. Given the number of actuators, subapertures and cycles per second, together with the desired precision, it is not surprising that small changes in flexure, temperature, or simply initial alignment can render pre-calculated predictions, especially of the influence function between the deformable mirror (DM) shape and the wavefront sensor (WFS) readout (called the interaction matrix, IM), useless, or at least compromise performance substantially. The sheer expense of night-time on a 8-10 m or bigger telescope (Baade1996), on the other hand, restricts the possibility of working on the AO system and tweaking it under realistic conditions, namely on sky.

The obvious remedy, common to all the systems in operation, is to use a calibration light source that illuminates the DM and the optical path to the WFS, and allows for calibrations to be performed during the day. This strategy has been fairly successful, although some deviations between the calibration source IM and optimal IM at night occur. The typical realisations for these calibration light sources are optical fibres in an intermediate focus, before the optical relay that re-images the pupil onto a flat DM with a diameter in the order of 10 inches. All this is usually contained in a laboratory environment, on an optical bench. While the principle of such a system stays the same for the new generation of AO systems, the implementation difficulties become substantially more severe. With a large, curved DM early in the optical train, the calibration source is no longer located on an optical bench, but contained within the telescope structure, in a location that is not gravitationally invariant and experiences huge temperature variations. On top of that, multiple laser guide star beacons, substantially off-axis in terms of the field of view of the telescope, and hence with strong aberrations, need to be feigned simultaneously.

In our specific case, we have the following setup: The LBT consists of two Gregorian telescopes, with an 8.4 m f/1.14 primary mirror each. The concave adaptive secondary mirror (ASM) has 910 mm diameter, and 672 voice coil actuators. Three Rayleigh LGSs, 120 degree rotated, are generated 2 arcmin off-axis in 12 km distance from the telescope. For the LGSs, 220 waves of coma are introduced by the parabolic primary, as well as 70 waves of spherical aberration because of the non-infinite distance (Figure 44). These characteristics must be reproduced by artificial calibration light sources in order to allow calibration during the day.

In this section I first describe the strategy we adopted for calibrating the ARGOS system at the telescope. I present possible optical designs that I investigated. The design history of the selected solution is described in detail, followed by the mechanical design, and the assembly plan. A procurement and test plan is outlined, and I conclude with performance expectations and an upgrade path.

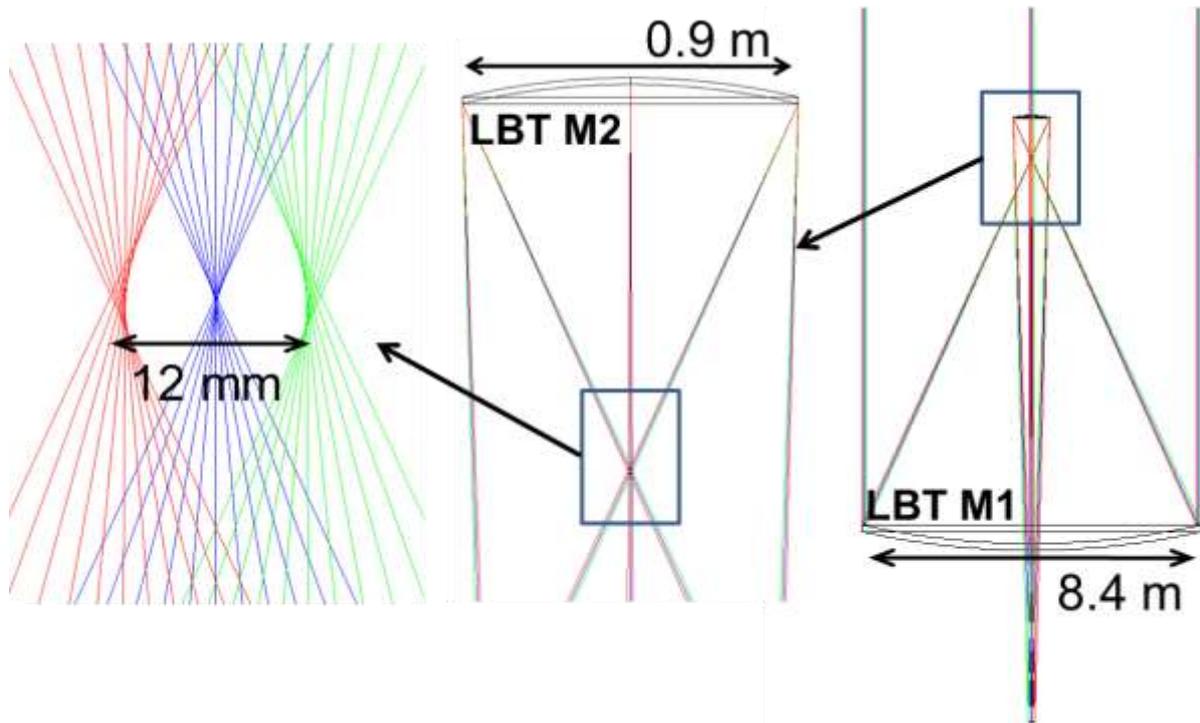


Figure 44: The ray bundles from the three laser beacons in the prime focus.

3.2 Calibration strategy

The most important calibrations of the AO system are the interaction between the DM and WFS and determining the non-common path aberrations between sensor and science instrument.

One way to determine an interaction matrix is to model the optical system and the DM's response as precisely as possible and then calculate the influence function between the WFS and DM. Accurate knowledge of the system, especially the non-common path aberrations, the mapping of the pupil to the WFS and the electromechanical behaviour of the DM, is crucial.

The approach we follow here is to measure the influence function and the non-common path aberrations at the telescope. The strategy is as follows: First, the shape of the DM is optimized to deliver the best possible image of a point source at infinity. Then the spot positions of the Shack-Hartmann sensor are measured and stored. This defines the null position. The set of point coordinates, or more accurately, the difference from an idealised system, is generally called the slope offset. In the next step, the DM is perturbed in a known way and the corresponding differences in the spot positions are measured. The resulting matrix that relates the Shack-Hartmann positions to the DM shape is called the influence or interaction matrix. Ideally, the necessary commands (e.g., the voltages applied to the actuators) to set the DM to the inverse position are already known (this is the case for a perfectly linear system). If so, it is straight-forward to compute a command matrix relating WFS signal to the commands needed to drive the DM in order to flatten the wavefront; this matrix is called the reconstruction matrix, and is basically the inverse of the interaction matrix.

Any change in static aberrations makes it necessary to refine the slope offset to achieve an optimal PSF. Likewise, any change in DM control behaviour (for example, a difference in the relation between actuator voltage and deformation) requires an update of the interaction matrix. If the PSF on the science detector can be evaluated, these refinements can be performed iteratively.

The stability of the calibration, and hence the intervals at which calibrations have to be performed, is a very important factor for the operation of an AO system. Other observatories report very different experiences on this issue; the AO facilities at the Subaru telescope and at the MMT, for example, run with control matrices that have been calculated at some point during installation of the system, and do not need to be updated. ALFA, the decommissioned NaLGS system at the 3.5 m telescope on Calar Alto in Spain, was at the other end of the spectrum, with recalibrations needed on a nightly basis. Part of the difference between these systems can likely be attributed to the difference in actuator technology. Piezostacks, as used on ALFA for example, show much higher dependence on temperature than voice coils. LBT's ASMs have position controlled actuators; the deflection at each actuator is measured with a capacitive sensor between the deformable shell and a reference body. Hence we expect changes in the control matrix to be minimal. However, to avoid unnecessary downtime, the ARGOS consortium adopted the conservative view that a calibration system that can be operated during the day is highly desirable. It has to be easily deployable so it can be used whenever the need arises. Additional advantages of such a system will be the possibility to increase engineering time on the system running in closed loop, and the chance to train the operators without the risk of losing observation time.

3.3 Top level requirements

A real intermediate focus in front of the ASM, as provided by the Gregorian design, is a basic prerequisite for employing a calibration light source in front of such a big mirror. Naturally, we choose to place the light source close to that focus where the beam size is small. To investigate possible optical solutions for this, we identified top level requirements for the system. These were driven by the wavefront correction accuracy we hope to achieve on the nights with the best seeing, as well as external constraints implied by the foreseen operation scheme and location of the calibration unit at the telescope. The main design drivers are the following:

- Wave front error (WFE) – We assume the AO system can in the best case provide seeing correction with a residual WFE of 400 nm, of which only 50 nm are contributed by the wavefront fitting. To reliably test the system's behaviour at this error level with a calibration source, we aimed for an RMS WFE of 50 nm.
- Wavelength range – As the laser beacons are monochromatic, the WFS is not designed to be achromatic. To avoid chromatic effects, especially in the Pockels cells, the calibration light source has to operate at the same wavelength as the laser beacons, 532 nm.
- Reliable alignment – The calibration system can be used to its full potential only if it can be deployed quickly and without complicated preparation. Precise and automated alignment is a prerequisite to provide reliable calibration whenever needed, for example to identify the cause of a system break-down during observations.

- Truth sensor feature – To adjust the WFS slope offsets for non-common path aberrations and for tweaking the system with the help of the calibration source, the calibration system ideally provides the means to measure the performance of the correction by analysing the science PSF, either directly on the science camera or with the aid of the NGS wavefront sensor located in the science field.
- Outer dimensions – To avoid obstruction of the beam, any device placed in the prime focus volume has to stay below a diameter of 83 mm. In that plane, this corresponds to the projected central obstruction of the beam caused by the secondary. We want to deploy the unit by means of a dedicated swing arm made from carbon fibre reinforced plastic. Due to the long moment arm we defined the maximum weight to be 2 kg.
- Temperature range – The calibration system should be available under all operation conditions in which AO can be run. This sets the temperature range from –15 to 25 degrees Centigrade.
- Pupil illumination homogeneity – To ensure even spot brightness across the Shack Hartmann sensor, the overall variation in light flux in the pupil plane should be less than 10%. Furthermore, the illumination homogeneity within the footprint of a given actuator should be better than 3%, to avoid errors due to the overestimation of local curvature.

The following Table 4, taken from the FDR documentation, summarizes the most important requirements identified for the system:

Table 4: Summary of the top level requirements for the calibration unit from the FDR documentation.

Goal	
Minimize night time calibration and setup time	
Minimize day time setup time	
Tasks	
Testing of the system in closed loop	
Measurement of Interaction Matrix	
Measurement of non common path aberrations – slope offsets	
Global Requirements	
Must have a diffraction limited on-axis source conjugated to infinity.	
Should be remotely deployable	
Should provide a reproduction of the LGS image on the WFS detector close to the real LGS image from sky.	
Should leave option to conjugation to 100km on-axis later on.	
Wavefront error	Requirements
On-axis source (from 0.5 to 2.4 um)	WFE 50nm RMS
Off-axis source [Goal] for 532nm	WFE 50nm RMS
Off-axis source [Spec] for 532nm	WFE 500nm RMS
Illumination homogeneity in pupil plane	
Overall homogeneity	10% PV
Inter-subaperture contrast (flux variation between two subapertures)	10%
Flux variation within the area of one actuator	3%

Mechanical envelope	
CalUnit diameter	< 82.5 mm
CalUnit height	~180 mm
CalUnit weight	≤ 2 kg
Operation conditions	
Temperature	
Operation	-15 to +25 °C
Storage	-30 to +50 °C
Temperature changes within 1h (max. calibration time assumed)	< 1K
Temperature change over night	< 5K
Humidity	
Storage	5 to 95%
Operation	5 to 95%
Pressure	
Storage	500 to 760 Torr
Operation	500 to 600 Torr

3.4 Investigated optical solutions

Several possible ideas to illuminate the DM and project the light onto the LGS WFS were investigated and compared to arrive at a baseline concept at PDR. A concept and some of the hardware for the FLAO system were already in place by this time. The FLAO calibration source thus served as a starting point for our considerations. In this section, the different options are outlined, with a description of the FLAO calibration system in Section 3.4.1 as reference. The calibration system design presented at the PDR is detailed in Section 3.5.1. The tolerancing analysis (Section 3.5.1.1) and a subsequent investigation into the procurement of the optical parts showed that modifications would be necessary. The final design adopted at FDR is described in Section 3.5.2.

The options were investigated using ZEMAX. The general optical layout of the LBT together with the LGS wavefront sensor is depicted in Figure 45. To illustrate the different optical paths of the discussed solutions, I show raytrace diagrams; the tertiary mirror M3 is generally omitted as it only folds the beam and has no relevant optical function for our problem.

3.4.1 First Light AO calibration scheme and hardware

The first light AO (FLAO) system is a natural guide star based AO system using the ASM as the corrective element. The wavefront sensing is done with the so-called acquisition, guiding and wavefront sensing units (AGW), developed mainly at the AIP (Astrophysikalisches Institut Potsdam). The AGW units are mounted on the Gregorian ports of LBT, in front of the instruments. Lucifer's front window is tilted and has a dichroic coating, reflecting the light blueward of 900 nm towards the high order sensor of the AGW unit (Figure 46). The sensor is mounted on stages, and can pick up a guide star over a patrol field of 3 x 2 arcminutes, roughly determined by the field of view of LUCIFER. The wavefront sensor is based on the pyramid concept. A four-sided pyramid provides four pupil images, which are imaged by one common CCD detector (CCD39), providing a sampling of up to 30 x 30 subapertures. This allows up to 496 mirror modes of the 672 actuator ASM to be reconstructed.

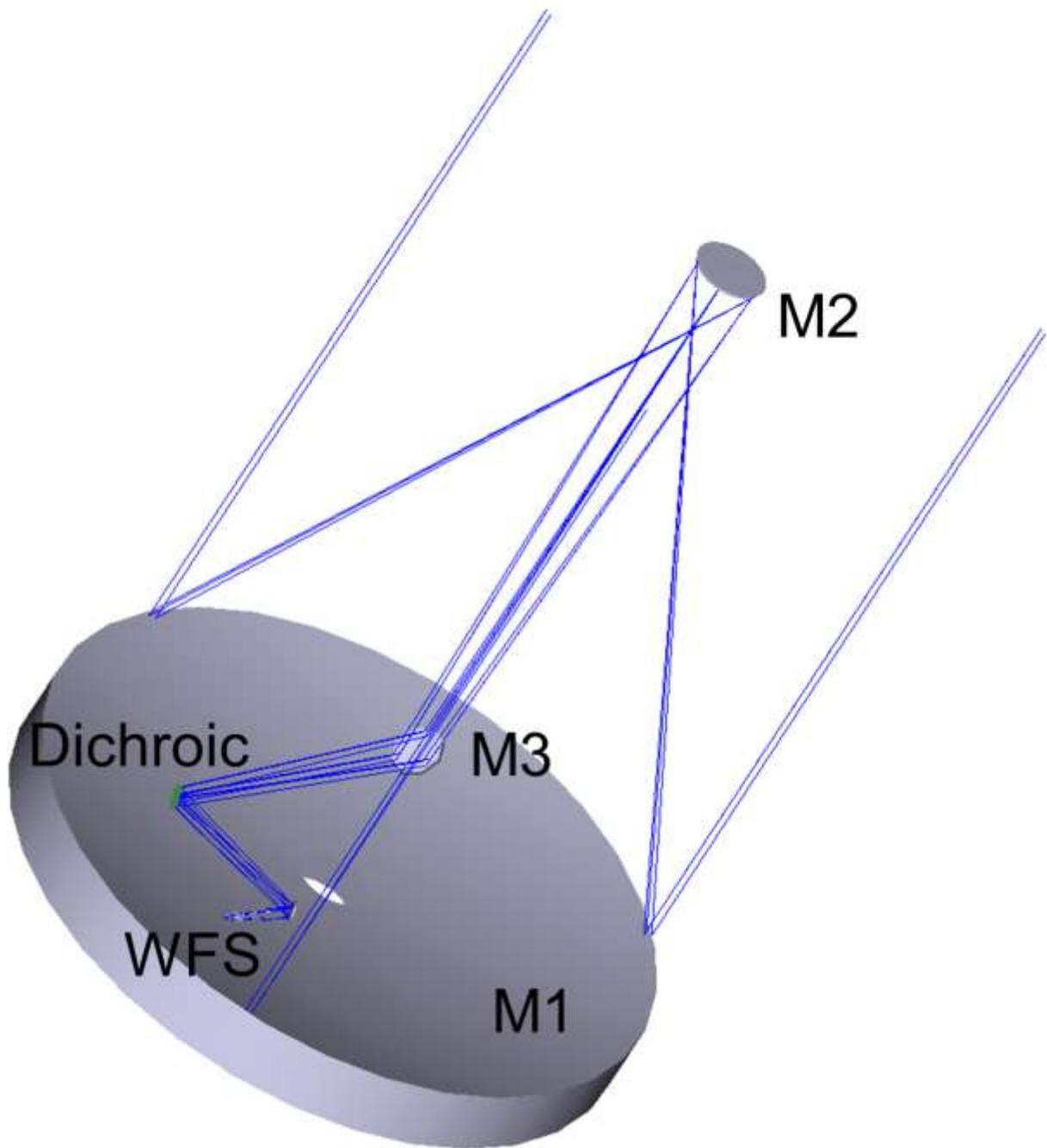


Figure 45: The beam path of the laser guide stars from the entrance aperture to the wavefront sensors. The light gets focussed by M1 into an intermediary focus below M2, the adaptive secondary mirror. M2 reimages the focus into the Gregorian focus. The beam is send towards the Nasmyth platform by M3. A large dichroic mirror in front of the science instrument separates the laser light and sends it via another folding flat towards the wavefront sensor optics.

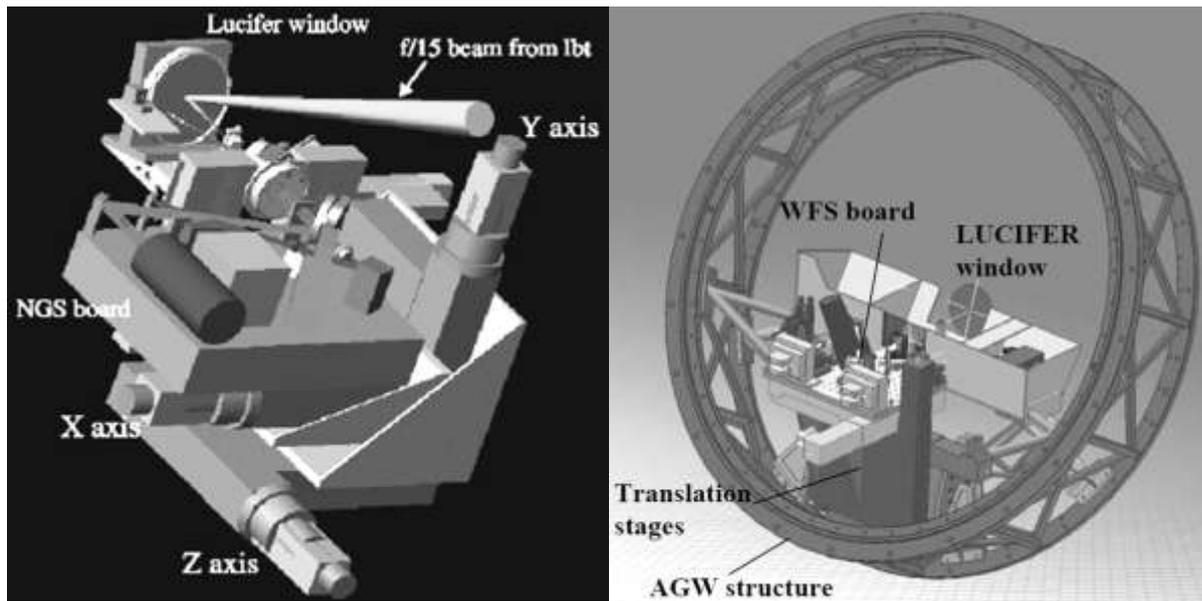


Figure 46: Left: Schematic of the opto-mechanical layout of the FLAO wavefront sensor pickup geometry. The light bluer than 900nm from the guide star is reflected off Lucifer's dichroic coated front window, and directed onto the pyramid sensor on the NGS wavefront sensor board (from Esposito et al. 2003a). Right: The location of the wavefront sensor inside the AGW unit (from Esposito et al. 2003b).

The FLAO system has its own calibration hardware, to be installed at the telescope during commissioning of the first adaptive secondary mirror, which started in March 2010. The concept is to illuminate the secondary in double pass, with a light source at the Gregory focal plane, and a retro-reflector at the prime focus. The light source is comprised of an incandescent lamp with a spatial filter and an objective matching the focal ratio in the secondary focus. The retro-reflector sends the light back to the secondary mirror, and then to the Gregory focus and the FLAO WFS. Figure 47 illustrates the light path.

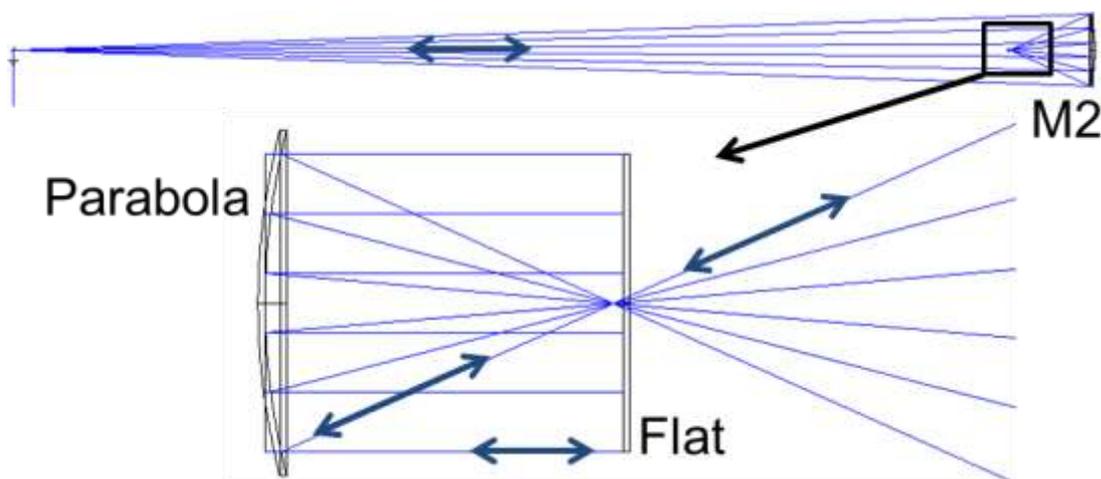


Figure 47: Top: The lightpath of the FLAO calibration light source. The ASM is on the right. Bottom: A close up view of the retro-reflector below the ASM. On the left is a parabolic mirror, in the middle a plane mirror, with a small hole in the centre to permit the beam to enter. Together they form a retro-reflector.

The light source in the Gregory focus is located on the optical axis. The secondary mirror's ellipsoidal shape images it perfectly into the prime focus, and the following retro-reflector also has a stigmatic image. Hence, if M2 has its nominal shape, the aberrations measured at the FLAO must be zero. As the DM is seen twice, any remaining aberrations must be divided by a factor of two when analysing the DM shape. The aberration free prime focus image in the nominal case is a necessary condition for this scheme to work, as otherwise it becomes very hard to ensure that the same parts of the beam see the same subaperture on the ASM both ways. The necessary precision of the registration is given by the spatial scale of the actuators on the ASM's shell (about one inch). The registration has to be considerably better to provide reliable results for the calibration of the mirror shape.

The employed retro-reflector deserves special attention (see Figure 47). It uses a parabola, confocal with the ASM ellipsoid's short focus, collimating the light coming from the prime focus. A plane mirror follows the parabola, reflecting the beam into itself, and the parabola reimaging it onto the prime focus. This slightly more complicated design was chosen to make sure the pupil lies on the secondary mirror, as it is for the whole telescope. A single, very fast, but nevertheless much easier to manufacture spherical mirror could also be used as retro-reflector in that position, but leads to a different pupil location. A spherical mirror would also be insensitive to tilt.

The fact that the parabola in the FLAO retro-reflector assembly is used twice leads to doubled surface figure specifications, as well as tight alignment tolerances between the parabola and the flat mirror. Table 5 shows the relevant specifications of the parabola as used for the call for tender.

Table 5: The specifications of the FLAO retro-reflector parabolas.

<i>Specification item</i>		
REQ01	Material	Zerodur (standard grade)
REQ02	Shape	On-axis parabolic concave mirror
REQ03	External optical diameter	30.0mm
REQ04	Diameter of unused circular patch centered on parabola vertex	3% of external optical diameter
REQ05	Focal length	30.0mm±1%
REQ06	External mechanical diameter	32.0mm±0.2mm
REQ07	Centering of parabola vertex with respect to external mechanical diameter	0.2mm on radius
REQ08	Optical surface quality.WFE.	WFE< λ /10PtV
REQ09	Low order aberrations (first 8 Zernike-Noll coeffs). Total WFE.	WFE< λ /4PtV (goal WFE< λ /10PtV)
REQ10	Scratch/dig quality	<40/20

During procurement it became apparent that the parabola is a complicated piece of optics. The envisaged tight wavefront error specifications could not be met by the manufacturer. The goal for the surface error on the parabola was split into low order and high order aberrations, defining low order as the first eight Zernike coefficients considering the Zernike-Noll convention. It turned out that despite the considerable cost of the item, the desired specification of $\lambda/10$ wavefront error in single pass was not met for the high order aberration regime, and indeed was missed by quite a substantial margin. The best substrate delivered by the vendor was measured with a surface error of $\lambda/6$ (Figure 48 and Figure 49). As the parabola is used in double pass, this translates to a final WFE of 0.6λ . The high order aberrations are particularly troublesome, as they affect only several subapertures, and their effect on the interaction matrix measurement is hard to model. With a total of 672 actuators across the DM, high order aberrations on the reference surface cannot be easily ignored during calibration. The footprint of a single actuator corresponds to less than a millimetre on the retro-reflector, which consequently should be smooth on small spatial scales. For large aspheric departures, as in the case of the used f/1 parabola, this is a challenge for the manufacturer. One can clearly see that the large surface error comes from residual tooling marks, mostly concentric. A possible explanation would be residuals from single point turning of the surface shape. Nevertheless, optics can be smoothed out by different techniques, for example localised polishing with magneto-rheological finishing machines or ion beam figuring, or simply with a flexible tool.

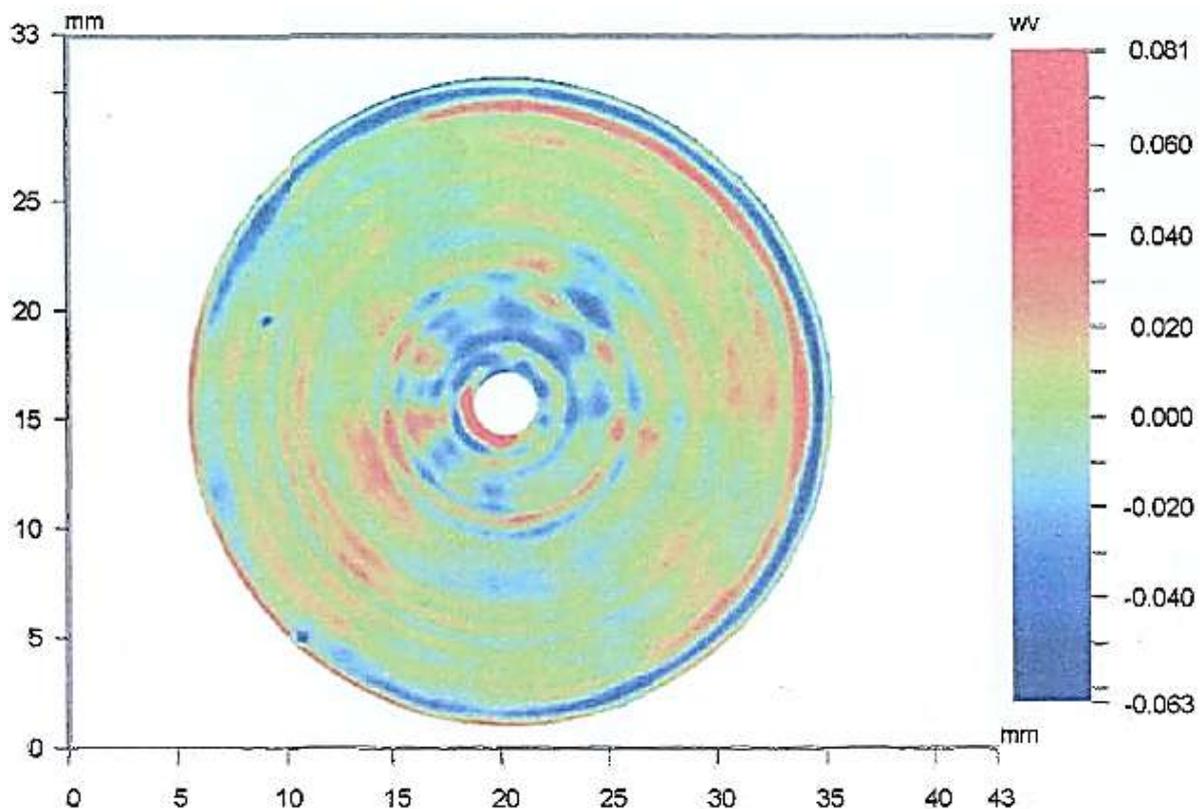


Figure 48: Surface error (SFE) of parabola number 2 for the FLAO retro-reflector. As the mirror is used in double pass, this translates to 4 times the value shown here as wavefront error. SFE is 0.144 waves PV and 0.015 waves RMS.

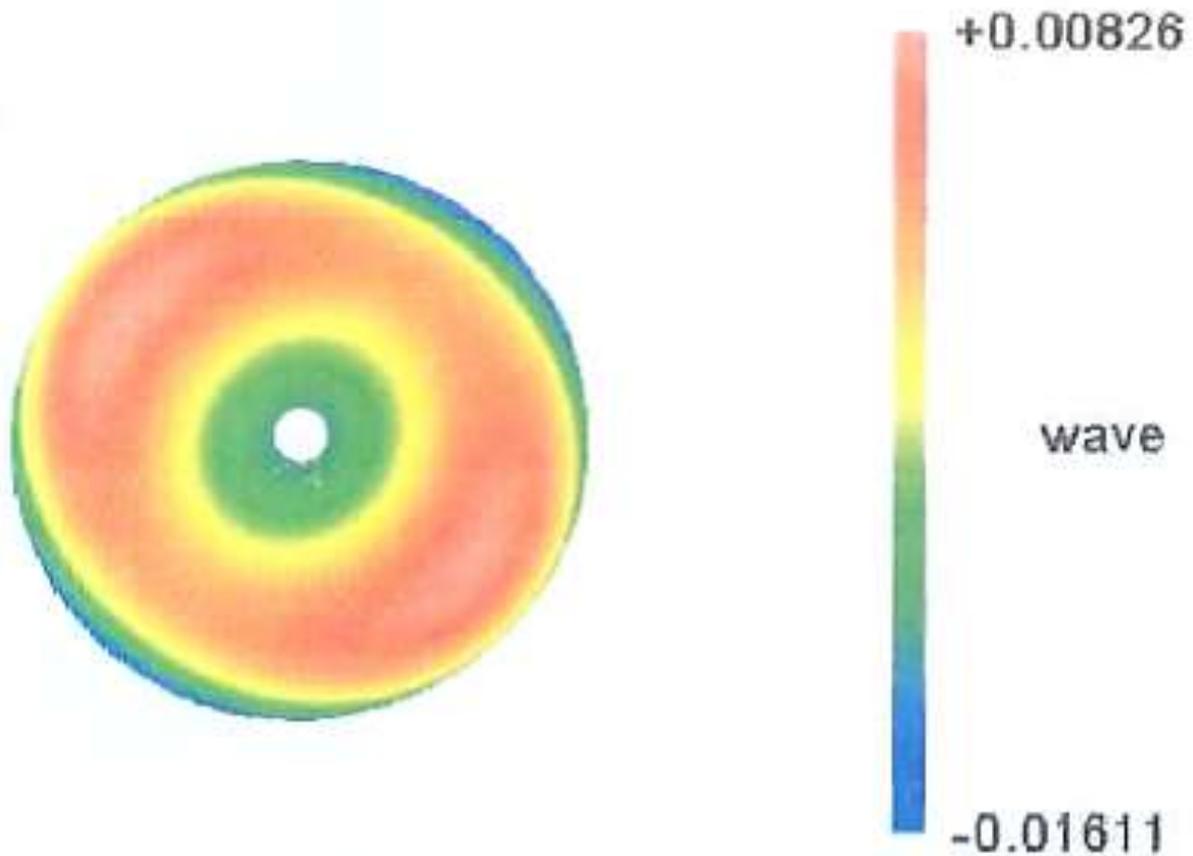


Figure 49: Surface error of the same parabola, but only taking the first 8 Zernike terms into account. SFE is 0.024 waves PV, and 0.005 waves RMS, an excellent value. As can be seen by comparison to Figure 48, the large surface error comes from residual tool marks on small spatial scales. Seemingly the surface was not smoothed out, for example with an elastic full aperture tool.

3.4.2 On-axis double pass with periscopes

I now consider how the design of the FLAO calibration source and the experiences from its development can be applied to the ARGOS calibration system. One solution to illuminate the LGS WFS with light coming from the ASM is to adopt the same strategy and use only one on-axis source to illuminate the DM in double pass using a retro-reflector. Pickup periscopes in front of the off-axis WFSs then direct the light onto the Shack-Hartmann lenslet array.

This option was considered non-optimal because of the additional problems arising from the off-axis position of the WFSs. In addition to the concerns about the correct mapping for the two reflections off the ASM, more optics are needed in conjunction with the periscopes. First, the retro-reflector works only at infinity conjugate, as the ASM's ellipse has its focus in that plane. The 12 km conjugated focus is far enough behind the infinity focus to produce aberrations that lead to problems with a light source that is imaged onto that plane. If the light source is placed in the focus at infinity, an optical relay system is needed in conjunction with the periscopes, to move the focus to the 12 km conjugate plane. Additionally, there need to be optics that introduce the aberrations of the real laser beacons.

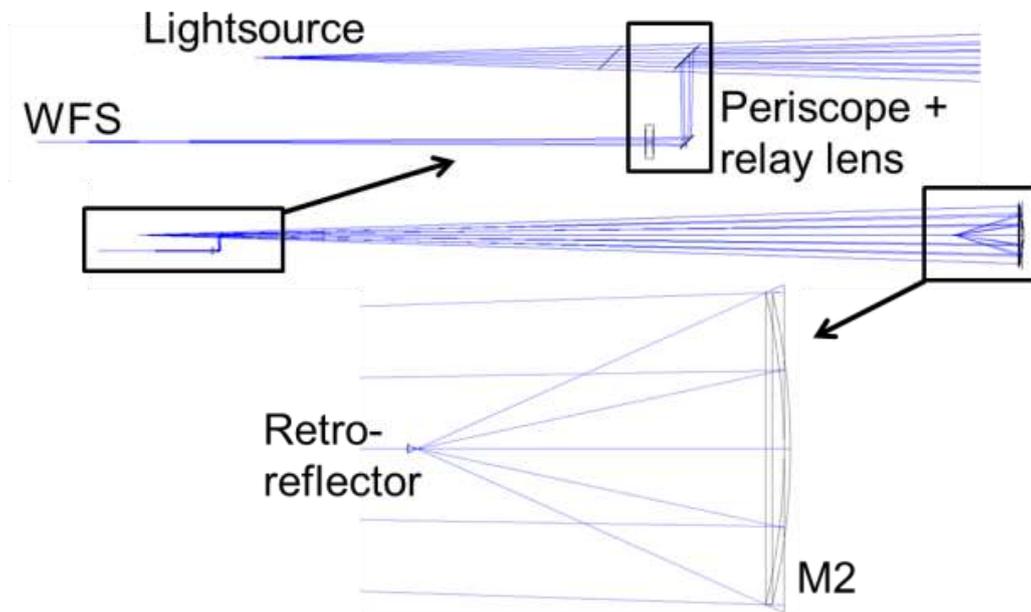


Figure 50: The layout for a double pass solution with an on-axis lightsource and a periscope to pick off the beam. Top: The full optical train. Middle: periscope. Bottom: the retroreflector and M2. Note that M3 was omitted in these plots.

Using the existing hardware from the FLAO calibration system (the light source and retro-reflector) would not be straight forward. The mounting procedure for the retro-reflector takes much longer than the setup time specified in the requirements of the ARGOS calibration system. The spectrum of the FLAO light source is obviously very different from the monochromatic laser beacons, which is undesirable, for example because of chromatic effects in the WFS pupil imaging optics.

When operating the LGS system, a large pickoff dichroic is mounted in front of the acquisition and guiding unit. When the calibration unit was designed, it was not yet clear if this dichroic mirror will have some optical power to balance the astigmatism it introduces to the science light path, or if this is offset with the secondary mirror. In both cases, the substrate thickness of the dichroic mirror leads to a lateral shift of the light source position as seen by the secondary mirror (Figure 51). This shift cannot be compensated by moving M2, as the lateral movement of the ASM hexapod is used to align M2 with respect to the retro-reflector, and furthermore the necessary shift exceeds the range of the hexapod. In principle, an axial shift of M3 can compensate for the displacement of the optical axis, but again the range is exceeded. The most practical solution would be a new, separate light source. It could be fed by a broadband light source (an incandescent lamp), and a monochromatic source simultaneously.

If this unit would be mounted together with the WFSs, seeing the dichroic in reflection, an optical element would have to include the same aberrations as the dichroic – astigmatism, in the case of a flat substrate, which could possibly be introduced with a plane-parallel plate. For the retro-reflector to work correctly, a point-like image of the light source is necessary in the prime focus. Consequently, the broadband part of the light could be used to feed the FLAO WFSs and the science instrument, independent of the mounting position.

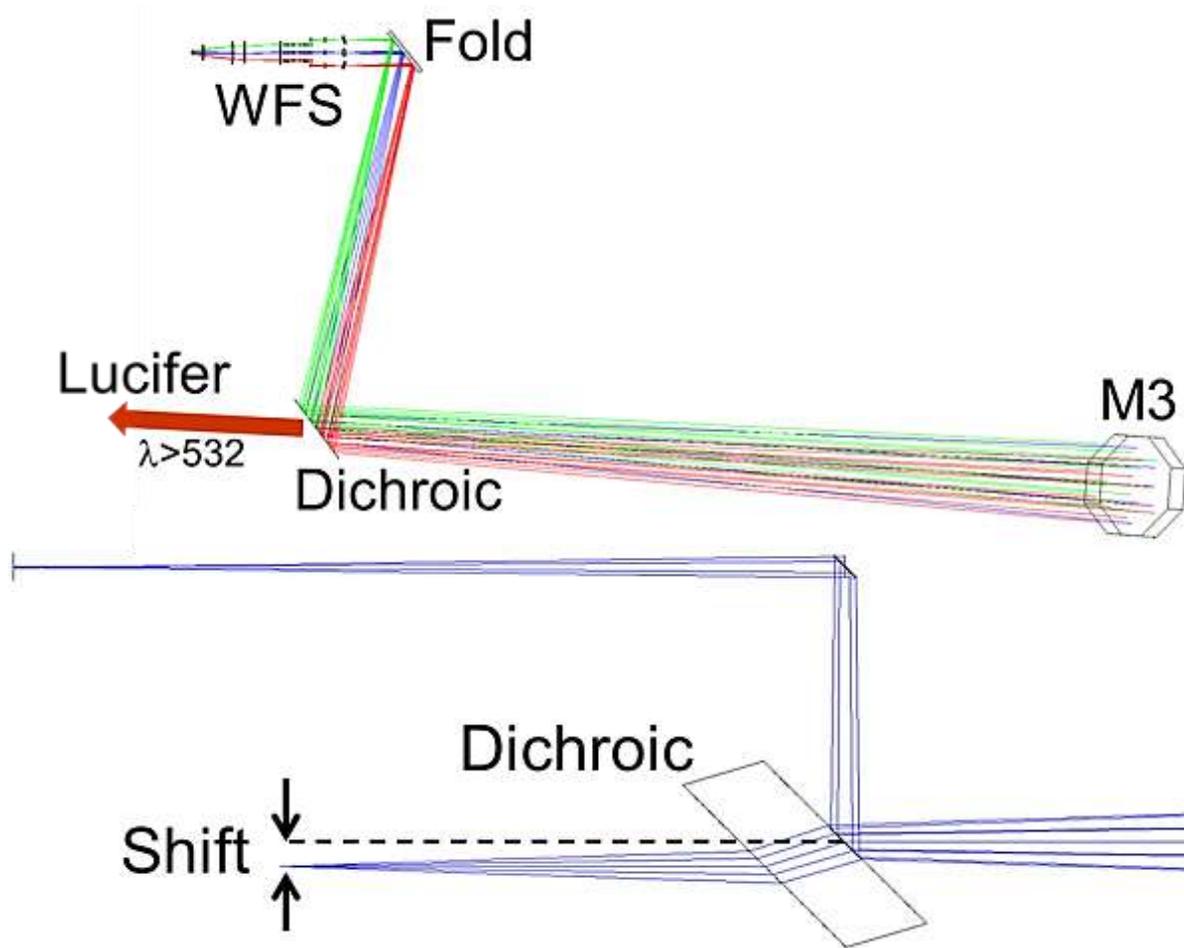


Figure 51: Lateral shift of the beam from a calibration light source caused by the tilted dichroic that separates the science from the laser light (see Figure 45 for where this fits into the larger telescope structure). The lower ray-trace shows how the thickness of the dichroic causes the shift.

The periscope assembly could either be one periscope that rotates to feed the WFS successively, or three periscopes behind each other, with beam-splitters to pick off the light. This latter option would be required to run the system in full.

In summary, the use of periscopes seems very complicated, with many optical elements in the light path, and automated opto-mechanical parts in the restricted space of the WFS assembly. The off-axis aberrations have to be introduced after reflection on the secondary mirror, which leads to problems with the correct mapping. Adopting the light source and retro-reflector design of the FLAO calibration unit has only slight advantages, as additional units would have to be built. The retro-reflector has to be mounted on a mechanical structure that can be inserted quickly.

3.4.3 Off-axis double pass

Another investigated option was to illuminate M2 in double pass, but with dedicated light sources that are directly off axis, to avoid the periscopes. This approach immediately leads to problems with the aberrations at that field position and conjugate distance. As there would be no well-defined focus without specific correction of the secondary's aberrations, collinear retro-reflection (where the rays are reflected into themselves) becomes very complicated. A free-form surface with a shape such that the surface is everywhere perpendicular to the incoming rays would serve this purpose, but at this precision it is beyond the capabilities of available manufacturing technologies.

A corner cube reflector reflects all incoming rays antiparallel to themselves. The lateral displacement of the beam inherent to this type of reflector is problematic. Also, corner cube prisms have a maximum opening angle of 70 degrees. A single corner cube for all three beams would lead to a displacement comparable to the actuator separation, which is unacceptable. Three very small reflectors, one per beam, would minimize this effect. But due to the high magnification between the conjugated planes, and the quickly changing aberrations of the fast M2 as a function of field position, even very small shifts are problematic. The lateral shift needs to be at least of the order of the spot size in the prime focus plane if one wants to avoid seeing the edges of the corner cube. In our case, this gives a lower limit to the shift of 1 mm, corresponding to 21 arcseconds at the plate scale in prime focus (Figure 52). Lastly, the corner cubes would have to be of very high accuracy in terms of their wavefront error and perpendicularity of the reflecting surfaces.

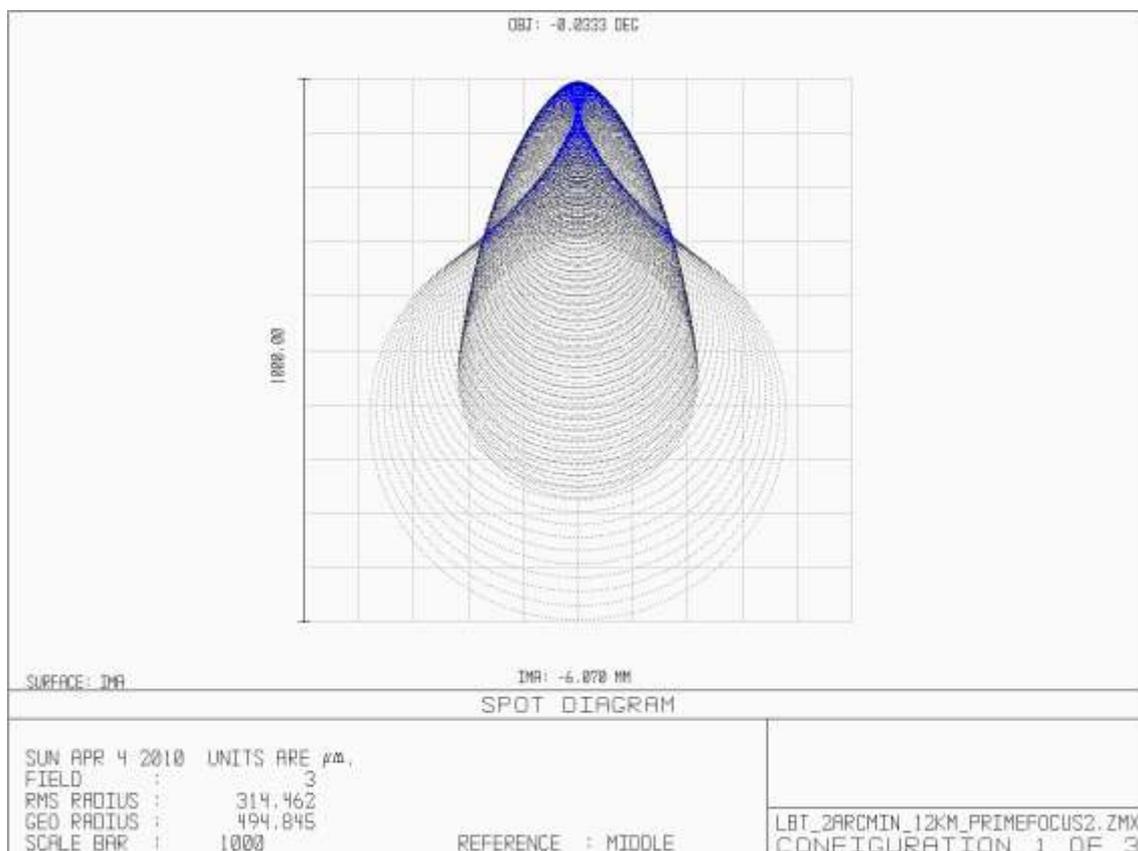


Figure 52: Spot of a point source at the location of the laser beacons as seen in the primefocus. The box is 1 x 1 mm².

3.4.4 Off-axis single pass

A solution to illuminate the ASM in a way more closely resembling the light path used at night is to place a light source in front of it, and use the ASM in single path, that is, with only one reflection. The challenge here is to find optics that generate the same wavefront as the images of the laser beacons at the prime focus, coming from the primary mirror's fast parabola. A second challenge is to find a way to mount these optical components and the light source so that they can be accurately aligned and repeatedly inserted. Different reflective and refractive objectives were investigated to achieve this. Below I consider the alternatives and the difficulties that arise. No satisfactory solution using a reflective surface was found. With transmissive objectives, however, a number of possible ways to generate the desired wavefront could be derived. For the lens elements I considered tilted spherical elements, rotationally symmetric aspheric surfaces, and diffractive optical elements (DOEs) in the various optical designs. Where possible, I tried to incorporate a way to have simultaneously an on-axis light source that could serve as truth sensor, as stated in the top level requirements. One idea to do this was to adapt the idea of the FLAO calibration unit and use a reflector on-axis. The different explored options, reflective and transmissive, are presented below.

3.4.4.1 Reflective objectives

The first idea that comes to mind is to use a small concave (parabolic) mirror imaging an off-axis light source, for example an optical fibre, into the prime focus. In principle, this is the configuration of the primary in front of the ASM. We need to arrive with the same amount of coma at the same field position and f-ratio.

It turns out that one cannot compensate for the aberrations with a simple parabola in this way under the constraint that the diameter must be smaller than the ASM's obstruction. This constraint leads immediately to a maximal distance of the parabola to the prime focus of about 65mm, and the resulting angle of the incoming beam places the light source at an inconvenient position. A light source mounted on the secondary would be >80 mm off axis, obstructing the beam (Figure 53). A light source closer to the parabola could be mounted such that no additional obstruction results, but can also not correct for the aberrations. A third way would be to place the light source somewhere on the telescope structure, some 4.2 m from the mirror, and fold the light in with a beamsplitter in front of the parabola.

A quick ZEMAX analysis shows that not even a free conic constant and an additional lens in front of the light source provide enough freedom to correct the wavefront for any of these configurations. It is interesting to note that a parabolic or nearly parabolic conicoid delivers the best results for the configurations without an additional lens, even with the relatively short object distances examined. A convex mirror instead of the concave one in front of the prime focus has the same problems, and delivers similar results.

However, the presented scheme lends itself to a very elegant solution for the on-axis source that can serve as truth sensor. An ellipse with one of its foci in the prime focus and the other one at an optical fibre in the centre of the ASM will image the fibre into the prime focus and on to the secondary focus (Figure 54). We bear that in mind for the following options. To summarize, no satisfying solution for a single pass configuration, off axis, with a mirror was found.

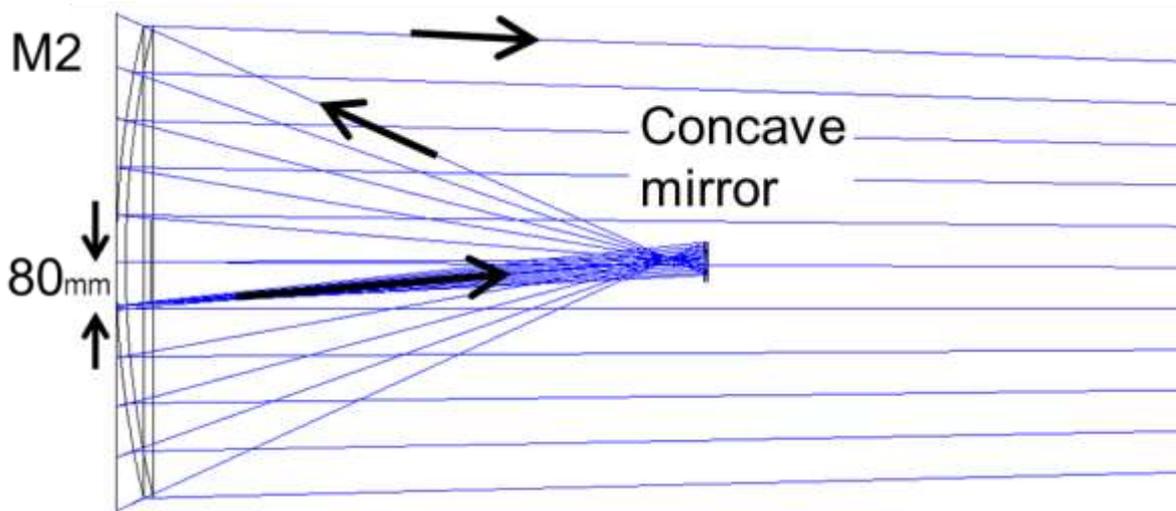


Figure 53: A small, nearly parabolic mirror (conic constant -0.97) in front of the ASM (on the left). This illustrates the configuration with the largest allowed diameter for the mirror before it obstructs the beam. The resulting location of the light source is about 80mm off-axis, far outside the central obstruction of the ASM, and therefore cannot be mounted.

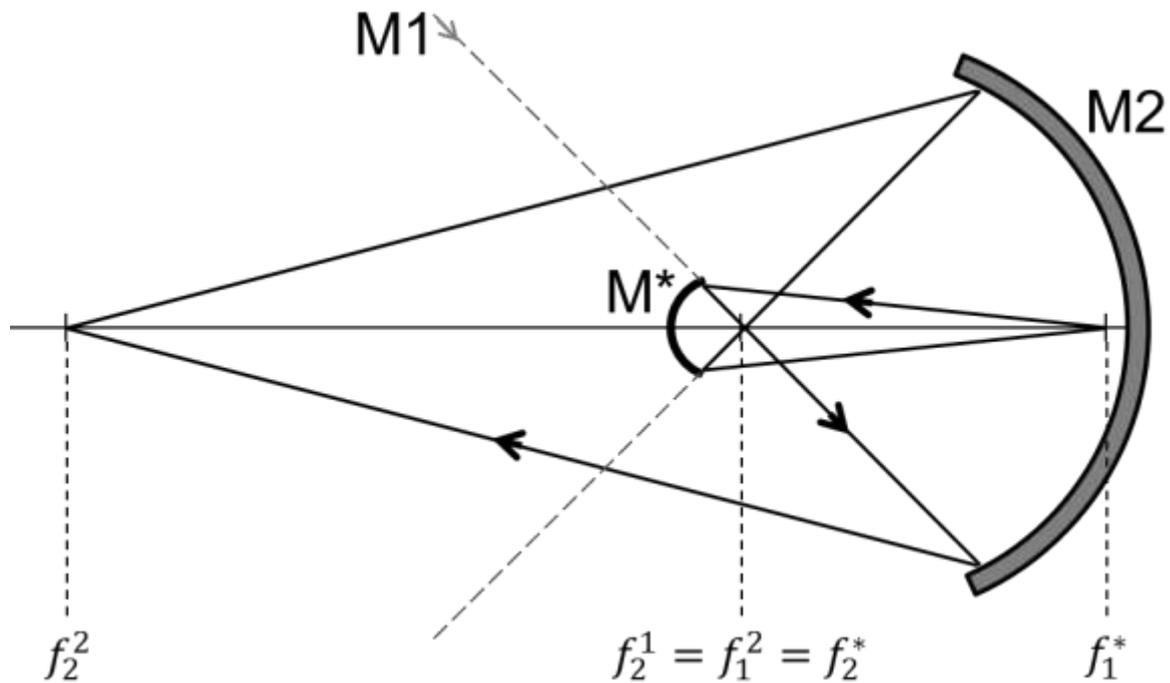


Figure 54: The foci of two confocal ellipses. A small prolate elliptical mirror M^* can image a source from its first focus f_1^* onto its second focus f_2^* , which is identical with the first focus f_1^2 of the secondary mirror $M2$ (and consequently the prime focus f_2^1 of $M1$). From there on the secondary mirror, itself a prolate ellipse, re-images the source into the secondary focus, f_2^2 .

3.4.4.2 Aspheric design

Aspherical surfaces have become more and more common in optical design, due to greatly improved manufacturing and testing capabilities. As long as the rays do not intersect, a free-form surface can, in principle, correct any aberration, as it can locally add the phase difference between desired and real wavefront. However, these surfaces are considerably (orders of magnitude) more expensive than spherical optics, and cannot be manufactured with the same precision as tightly specified spheroids.

Rotationally symmetric aspheres are easier to fabricate than freeform surfaces and are often found in commercial high performance optics. These may be either conicoids or surfaces with higher polynomial (often up to 10th order) deviations and are described by the following equation:

$$Z(s) = \frac{R^{-1}s^2}{1 + \sqrt{1 - (1 + k)R^{-2}s^2}} + As^4 + Bs^6 + Cs^8 + Ds^{10}$$

Z is the surface sag at a distance s from the centre; A , B , C , D are the aspheric coefficients, and k is the conic constant. R is the surface radius (R^{-1} the curvature).

Such lenses provide a reasonable compromise between flexibility, precision and cost.

I investigated if a single, rotationally symmetric objective with aspheric lenses can correct the aberrations at one particular field radius, corresponding to the LGS position of 2 arcminutes, and project three fibres into the prime focus. With four lenses and two aspheric surfaces, a satisfying solution can be found.

I also looked at designs with the additional constraint of using an elliptical retro-reflector as the front surface, to potentially use it for the on-axis light source (see Figure 54). With one additional asphere (a 10th order even asphere), the wavefront quality is already reasonable ($\lambda/2$ PV), and a design with a second additional asphere (see Figure 55) achieves an excellent wavefront correction of better than $\lambda/10$. The aspheric departures are 0.1 and 0.3 mm, respectively, and both aspheres have a rather relaxed radius. This amount of departure on about 2" diameter is hard to manufacture, but achievable.

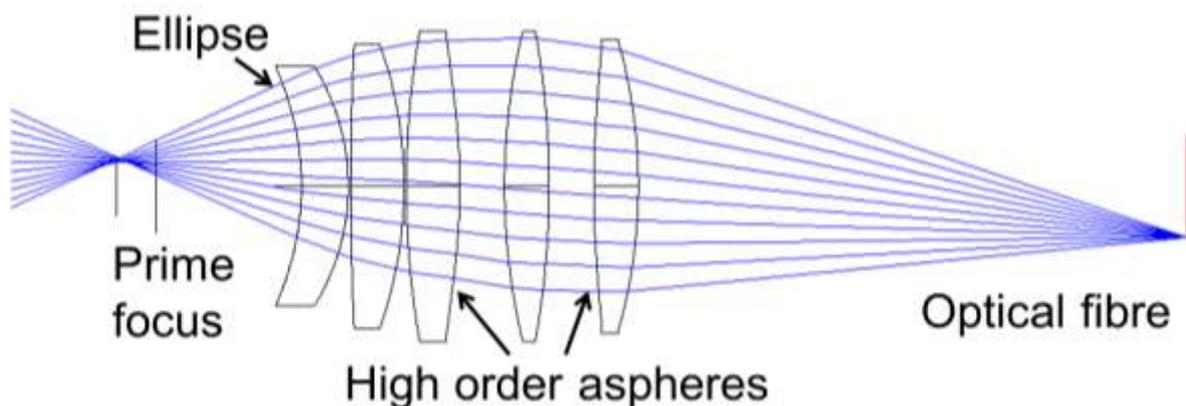


Figure 55: Projection objective with an elliptic front surface (on the left) and two additional high order aspheres (on the 3rd and 5th lens). The power is evenly distributed between the elements, the correction is very good.

3.4.4.3 Tilted elements

For one single field position (that is, not axisymmetric), allowing for tilt and shift of the lens with respect to the secondary mirror's optical axis, as well as to the fibre axis, provides additional possibilities to correct the aberrations. I investigated the possibilities with a single lens element and two lenses, with spherical surfaces. Using a single negative lens would be the most attractive choice, as it is mechanically not as complex and can provide the necessary matching of the fibre's numerical aperture to the $f/1$ focal ratio of the ASM. To my knowledge, the use of a negative lens in front of a fibre had been discussed for an on-axis light source before, but dismissed, as the necessary object diameter (equivalent to the virtual image formed by the lens) for a diffraction limited spot is considerably smaller than even single mode fibres. In our case, however, this is of no importance as we need a spot corresponding to 1 arcsecond on sky. The setup with one tilted lens does not comply with our goal to run all three WFSs simultaneously. To do so, each of the three fibres has to be mounted behind its own tilted lens system, and everything co-aligned.

An important question is if the system can be assembled and aligned at the telescope. The alignment tolerances with respect to the ASM are very challenging, but this is necessarily true for all solutions. However, the internal mechanical alignment of an assembly consisting of three tilted optical sub-assemblies is indeed more complicated than for a solution with optics common to the three light sources. As the field of view of the wavefront sensors is very small, the individual light sources need to be aligned to about 50 micron with respect to each other (corresponding to about 1 arcsecond) and the optical axes must be collinear to a few arcminutes.

I investigated the possible performance of the described setup with ZEMAX, optimizing tilt, shift and distances as well as radii of one or two lenses to produce the desired wavefront. No satisfying solution could be found.

3.4.4.4 CGH corrected wavefront

I considered computer generated holograms (CGHs) to shape the wavefront of an optical fibre and match it to the desired form. CGHs, also known as diffractive optical elements (DOEs), are, as the name implies, optical elements based not on refraction or reflection, but on interference. They act as phase masks to alter an incoming wavefront. In principle, the most simple CGHs in use are optical gratings, where periodic $\pi/2$ phase steps with a pre-calculated spatial frequency are put onto a surface to achieve constructive interference in a specific direction, and hence seeing reflection in this direction. CGHs are generalisations of this idea, where the phase pattern is not as regular, but can take on a much more complicated shape. It is again calculated to achieve the desired wavefront, and then transferred onto a substrate.

Types of holograms

Two different types of holograms are used: amplitude and phase holograms. They work like their analogues from optical gratings. A transmissive amplitude hologram is opaque where the desired phase shift is an odd multiple of $\lambda/2$, and transparent where the phase shift corresponds to an even multiple. Effectively the destructive parts of the wavefront are removed, leaving a large amplitude of positive interference. For a binary structure, the

efficiency is low, though. Obviously, half the light is lost directly on the opaque parts of the structure. Then, binary holograms divide the light evenly between positive and negative orders. The net result is a maximum efficiency of ~40%. An advantage of amplitude gratings if there is enough light to tolerate the losses is the fact that they can be produced more precisely, as no three-dimensional structure is necessary.

The other type, phase gratings, adds a phase shift of π where the amplitude hologram is opaque. This immediately doubles the efficiency. But the phase step is typically achieved by etching a transparent substrate material until the step height times the refractive index produces the desired phase shift. This additional step lowers the achievable precision.

A variety derived from the binary phase hologram is the grayscale hologram, where the phase shift can assume any value between 0 and π . This effectively blazes the structure, and leads to higher efficiencies. In practice, the smooth surface is approximated by a finite number of steps, between 2 and 256.

Manufacturing technologies

CGH manufacturing is most often based on lithographic techniques. While in the infrared the structures are large enough to be fabricated directly (mechanically) into the substrate, in the visible wavelength range they are photochemically written into a lacquer layer, and then either hardened or etched into the underlying substrate. The lacquers are called photoresist; they absorb photons of specific energy, which transforms the molecular structure of the material. Much like with a photographic emulsion, in a developing step the exposed material is removed (this is called positive resist; the opposite exists, too). In this way, a structure corresponding to the exposure pattern is generated on the substrate. It can be used like this, or transferred further onto the substrate. If a layer of chrome has been deposited on the substrate before the photoresist, an amplitude hologram can be made by wet etching of the exposed chrome. If a phase hologram is desired, the pattern can be transferred into the substrate material with appropriate etching technologies, which makes the structure more durable. One technique is reactive ion beam etching, also called dry etching. Blazed structures are made by approximating them with step functions. They may either be generated in multiple process steps, or directly with varying exposure intensities on photoresist that shows a relatively linear response to exposure levels. Photoresist is sensitive to UV radiation, typically below 450 nm. UV lasers around 350 nm are most commonly used as light sources. They offer good beam quality, resulting in small focus spot diameter and precise structures.

Holographic structures for the applications summarized above have structure sizes of a few microns. The challenge in manufacturing lies in addressing the writing laser beam to a precision considerably better than this, typically some 10 – 100 nm. It becomes immediately clear that holograms with complicated structures also present a formidable challenge for the data handling. A single addressable point with a laser writer has a size of below 2 micron. A CGH on a 20 x 20 mm substrate already has 100 Megapixels.

For the highest accuracy, so called E-beam writers that use a focussed electron beam instead of the laser to expose a special photoresist material are used. Their addressable features have sizes on the order of some 10 nm, leading to data fields that cannot be described on a pixel grid any more. Reduction techniques describing the pattern without loss of information have been developed for this.

There are two basic methods to describe the position of an addressable bit on the substrate, and point the writing beam there: Cartesian and polar coordinates. Both are used in the

industry. For rotationally symmetric holograms, often used for testing rotationally symmetric aspheres, polar coordinates obviously make sense. The writing process is similar to the writing of a compact disc – the structure gets written in one long spiral on a rotating substrate, with a writer head that is moving radially. The largest computer generated holograms (used for testing telescope mirrors) were produced like this. In our case, the holograms are not rotationally symmetric, and they are even located off centre on the substrate. This is a complication for the radial writing technique, and also for the verification of the structure.

Applications

Common applications include beam-shaping, typically from a collimated laser source. A hologram placed in the beam modifies the intensity distribution in the far field to generate everything from simple intensity fields like lines or grids to complicated shapes like company logos. Spot patterns or modified shapes have applications in laser material processing, like laser drilling, for example.

The other common application is in optical testing, or more precisely, in measuring aspheric surfaces. Here DOEs act as null lenses to form an optical system together with the surface under test that generates a spherical wavefront, which can be measured with an interferometer. In principle, a CGH can generate any difference between a given and a desired wavefront if it is placed at a pupil or at least at a position where rays do not intersect. A CGH can locally generate any deviation from an incoming ray by adding an arbitrary phase shift, but if rays are intersecting on the CGH plane, they get confused, as two rays with different directions would originate at the same point on the CGH surface. This property of adding an arbitrary phase shift is of course very attractive for us, as we need a very specific wavefront in monochromatic light that is hard to achieve with conventional optics.

The strong chromatic aberration of diffractive lenses is used to colour-correct infrared objectives. The chromatic aberration of a diffractive structure has the opposite sign to the one of a positive glass lens. By combining the two, the overall chromatic aberration can be greatly reduced without increasing the focal length! This leads to larger lens radii and reduced spherical aberration. An issue is the efficiency of the DOEs in this application. If fabricated as a binary structure, as a Fresnel zone lens, multiple diffractive orders are generated and the light distributed between those. This leads to an enhanced background and low throughput. If the structures are blazed, nearly all of the energy can be regained. But the blazing needs much more complicated manufacturing processes, for example grayscale lithography. In the infrared region, the rotationally symmetric blazed structures can be single point diamond turned directly into the lens substrate with the necessary precision and surface roughness. Recently, this method to correct colour has also been transferred to high-end consumer optics in the visible range. Canon offers a 400 mm f/4 photographic lens that makes use of a DOE to reduce chromatic aberrations, which leads to a much shorter and lighter design than conventional techniques based on the use of glass with abnormal chromatic dispersion.

CGHs are also used for the correction of monochromatic aberrations of lenses in laser processing applications. A CGH on the plane side of a plano-convex bulk lens can correct for spherical aberration, for example. CGHs are sometimes also used to correct for off axis aberrations of bulk lenses, an application that comes closest to our needs. The required accuracy, size, and the amount of aberrations are substantially different, though.

3.4.4.5 Only transmissive CGH

The most direct way to use a hologram would be to mount it in front of a single mode fibre, which delivers a well-defined spherical wavefront, and let the CGH add the local phase difference between this spherical wavefront and the one the real laser beacons will create in the same plane. If one wants real images in the focal plane, the fast f-ratio is problematic. The fringe structure to achieve a power in the hologram that generates an f/1 beam is very fine, and a challenge for manufacturing. If a virtual image is sufficient, a CGH in front of a fibre can adjust the aberrations basically perfectly. It could also convert the f/1 beam to match the numerical aperture of the used fibre. The layout of this geometry can be seen in Figure 56. In this case the implementation of a truth sensor is problematic. A monochromatic on-axis reference source can be generated with the hologram as well, but this was deemed undesirable for use with the FLAO wavefront sensor. Ideally, it should have roughly the same spectral coverage as a real star for the FLAO wavefront sensor, and should extend into the infrared, so it can be seen on the science cameras.

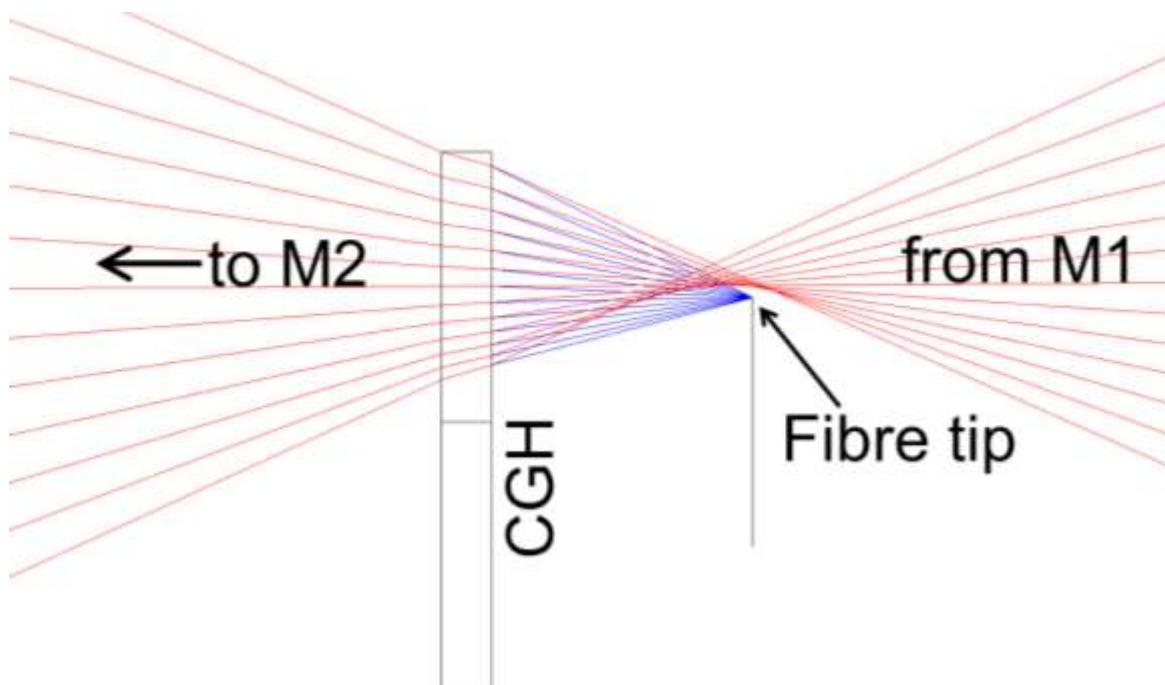


Figure 56: A possible CGH correcting the spherical wavefront shape of a fibre (blue) to match the wavefront of the laser beacons (red). On the CGH surface, the rays coming from the fibre are tilted so they overlap from there on with the laser beacons rays.

3.4.4.6 CGH corrected objective

One possibility is to have a conventional refractive objective imaging a fibre into the prime focus, thereby transforming the numerical aperture of the fibre to the focal ratio of the primary mirror. Instead of correcting the aberrations with the aspheric surfaces, spherical lenses can be used and the aberrations corrected with a dedicated CGH within the objective. In this way, the necessary tasks are split: the optical power is provided by the lenses, while the hologram takes care of the wavefront shape. This is more complicated than using a CGH alone but provides real instead of virtual images of the calibration beacons in prime focus and again opens up the possibility of employing an elliptical front surface for the truth sensor.

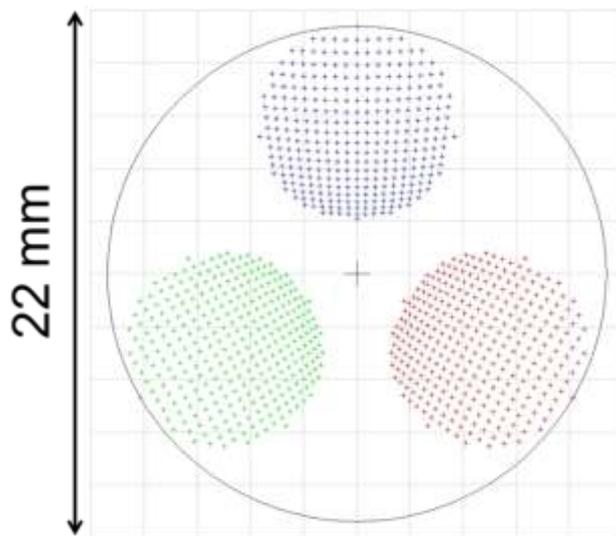


Figure 57: Footprint diagram for the layout shown in Figure 56. The footprints for the three beacons are well separated.

3.4.4.7 CGH on ellipse

A very interesting and elegant alternative to generate the off-axis spots is to use a holographic beamsplitter on a curved surface. The hologram works just like the CGH in the aforementioned setups, in the sense that it transforms a smooth wavefront into the desired shape. It can be used as a beamsplitter with the zeroth order still used for another purpose. If the CGH is used in reflection and put on the surface of an ellipse it can be used to generate the off-axis spots while allowing the on-axis light to pass through unaffected. Such a setup is illustrated in Figure 58.

The on-axis source should be broadband, while the CGH works only for one wavelength. The idea here is to make use of the different spectral power densities of the broadband light source (an incandescent lamp) and the single line (or narrowband) light source, typically a laser. Both sources can be delivered with the same fibre, with the help of a fibre-optical beam combiner. While the CGH generates undesired, washed out spots from the broadband light, these spots are much dimmer than the single, correct spot from the narrowband source. If the contrast ratio is high enough, the light coming from the broadband source is negligible. Key here is to put enough carrier tilt (the angular separation between the zeroth and first order) into the CGH. This will smear out the light distribution from a continuum source over a large region because the tilt angle is linearly dependent on the wavelength.

For the diffraction angle of the first order resulting from the geometry of the ASM and a confocal ellipse with 2" diameter, a wavelength shift of only 2 nm moves the spot by its own diameter. Hence only the light from the broadband source in a spectral band of about +/- 2 nm around the central wavelength of 532nm contributes to the background.

The holographic structures for the three off-axis spots should ideally not overlap on the surface, in order to avoid the mixed orders originating from the additional structuration of crossing fringe systems (see Figure 57).

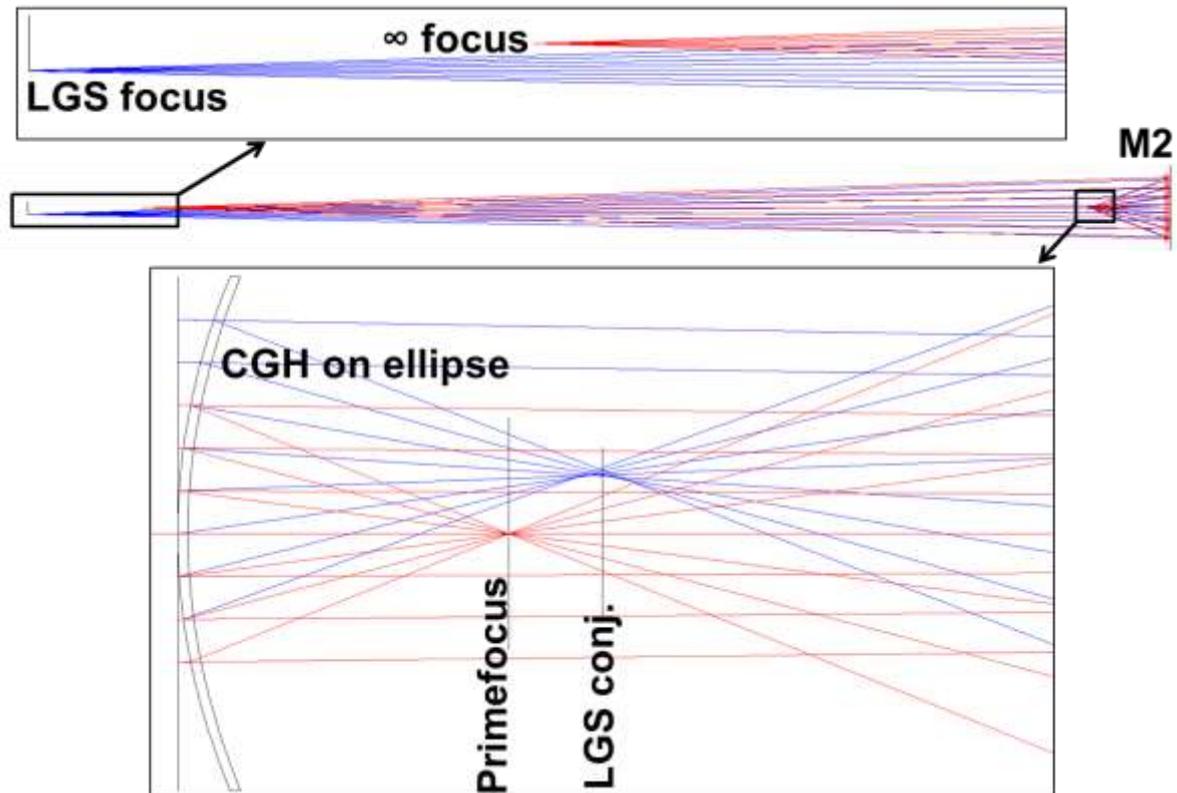


Figure 58: Optical layout of a CGH directly written onto an elliptical substrate. A lightsource is mounted in the vertex of the secondary mirror. The normal reflection (zeroth order) is used for the truth sensor as explained before (red). The CGH acts as beamsplitter and generates the beam resembling the laser beacon (blue).

The solution with the CGH on the ellipse's surface has the following advantages:

- provides a fixed alignment between off-axis spots and truth sensor
- thermal stability; as the substrate can have a very low CTE, and there is only one element, the mechanical assembly cannot become misaligned by thermal cycling
- robust against vibration, for the same reason
- very lightweight
- mechanically easy to mount
- only one fibre needs to be routed
- the remaining wavefront error is nearly static and can be precisely measured

As we want to use the light reflected on the ellipse's surface (the zeroth order) for the truth sensor, an amplitude hologram seems more appropriate, as this should prevent scattered light from the etched edges of a phase hologram. Furthermore, amplitude holograms can always be made more precise, as the etching introduces room for errors.

Care must be taken to ensure that the hologram's fringes are considerably smaller than the projected size of an ASM subaperture everywhere, as otherwise the illumination of the subapertures varies with the ratio of black to white lines on the corresponding area of the ellipse.

One drawback of CGHs on curved surfaces is that they have much larger manufacturing errors than their counterparts on plane substrates. This comes partly from the writing technique (on curved substrates laser writers are used) and partly from the mechanical difficulty of positioning the substrate with respect to the writing beam. Positioning accuracies of down to 150 nm are reported (Radtke & Zeitner 2007). For our purpose, the wavefront error resulting from this accuracy would be acceptable.

Unfortunately, the machines currently available can only write the holographic structure on substrates with a maximum local surface tilt of about 10 degrees, and therefore cannot write a CGH on the steeply curved ellipse we require, which has a maximum surface vector tilted by around 20 degrees. This is nearly independent of the surface radius. The area covered by one single CGH (for one of the three beacons) has a surface tilt between opposite sides of only 12 degrees, which might be possible to write on, but then the substrate would have to be rotated 120 degrees around a tilted axis (the ellipse's vertex) to very high accuracy, for the next CGH. While this problem surely has technical solutions, it would probably be uneconomical to pursue this for the calibration unit.

3.4.5 Discussion and comparison of all options

Several of the outlined solutions have the potential to meet all the requirements. The selection of one was based on cost, performance, risk, and available resources at MPIA and LSW (mainly on the background of assembly and testing).

The double pass options were discarded, as we want to trace the path of the real beacons as closely as possible and, in case of an off-axis scheme, we doubt that the WFE requirements can be reliably met.

For the single pass options, both the aspherical design and the spherical design incorporating a CGH are attractive as they use common optics to all three calibrations beacons, and incorporate a clever way to provide an on-axis source as truth sensor. This also has implications for the alignment at the telescope, which is complicated as the tolerances are very tight. As we aim for a WFE of $\lambda/10$ RMS, a CGH is favorable over the aspheres because two identical CGHs are cheaper than twice the two necessary aspheres with this precision. Basically, the standard way to test aspherical surfaces with the desired precision is to use a custom CGH as null lens, immediately raising the cost to the same as directly using the CGH. On top of that, additional holographic structures on the CGH can facilitate very precise assembly and alignment for very little extra cost, as will be described later. This presents another important advantage of this design. Lastly, I found a vendor for CGHs who was willing to share his expertise and work closely with me on developing a design, and deliver the CGHs. Given these advantages, we decided at the end of Phase A to pursue a transmissive single pass design comprising a conventional refractive objective that incorporates a CGH.

3.5 Calibration Unit Design

The investigation outlined above led to the preliminary design of the calibration unit presented at the PDR of the ARGOS project in March 2009. This included a detailed tolerance analysis that specifies the accuracy required for each of the optical elements. Follow-up on acquiring the required components and assessment of the capabilities to assemble the system in-house showed that modifications were necessary. The design was then adjusted appropriately for the FDR in March 2010. In this section I describe the optical designs for the PDR (Section 3.5.1) and adjusted design for the FDR (Section 3.5.2) in detail. The tolerance analysis of the PDR design is shown in Section 3.5.1.1. In addition to the optical components, the mechanical design (Section 3.5.2.3), the assembly and testing plan (Section 3.5.2.6) and the alignment of the unit on the telescope (Section 3.6) are discussed.

The calibration light source for the LGS WFS was designed to meet the top level requirements. In particular, this means to:

- project three beacons for the three WFS simultaneously,
- give a light distribution and spot position only marginally different from the situation on-sky,
- trace the light path of the real Rayleigh beacons through the telescope as closely as possible - that is, single path, with the same footprints, and comparable wavelength,
- be easily deployable to facilitate routine calibrations and testing of the system,
- allow the image quality to be checked on the science detector.

The main component of the calibration unit, referred to as the CalUnit hereafter, is a custom objective projecting three off-axis multimode fibers into the prime focus, matching the primary's focal ratio. A computer generated hologram adjusts for the aberrations and introduces the severe coma the real LGS spots will have in that field position. Naturally, the CGH works only for one well-defined wavelength, which in our case is chosen to be the same as the laser wavelength of the real beacons. As the beacons are monochromatic, this presents no restriction. The assembly consists of four lenses plus the hologram substrate and a fibre holder, mounted in a custom lens barrel. Three multimode fibres are used as the off-axis light sources, feeding the light of a small, frequency-stable laser onto the CGH.

As we need some kind of truth sensor to check image quality in the science field, we want to have access to an on-axis source to be used with the first light AO wavefront sensor (FLAO WFS). The first light AO light source and retro-reflector combination does not meet our need for easy and quick employment; therefore the calibration unit also has its own on-axis mode to be used with the FLAO WFS. Consequently, the on-axis light source is operated in single pass configuration, and simultaneously with the off-axis sources, allowing real-time monitoring of the mirror shape by the truth sensor (the FLAO WFS) during LGS wavefront sensor calibration. For this purpose, an optical fibre light source is mounted on the adaptive secondary mirror (ASM). Additionally, the well-defined light source fixed to the ASM acts as reference for alignment of the ASM with respect to the calibration optics.

The on-axis reference light source uses a fibre permanently mounted in the central bore of the ASM. The light is reflected off the dichroic front surface of the objective, an ellipse with one of its foci on the infinity conjugated prime focus and the other one on the fibre surface in the

ASM and illuminates M2 (see Figure 54). The fact that only the central ~5% of the fibre's light cone are reflected back towards the secondary ensures that there is no drop in illumination towards the edge of M2. As we use a broadband light source, speckles are not problematic. The fibre is imaged onto the prime focus, with a demagnification of ~25, relaxing centering tolerances of the light source by the same amount. This enables a diffraction limited spot size with a multi-mode fibre. To achieve this, the image of the light source in the prime focus must not exceed 1.6 μm in diameter because of the fast focal ratio of the primary. A fibre of up to 40 μm diameter can be used.

To facilitate automated and quick deployment of the CalUnit components in the prime focus volume, a dedicated swing arm made from Carbon fibre reinforced plastic (CFRP) has been constructed at MPIA (Figure 59). I have been involved in the definition of the mechanical properties, the interfaces and the design of the front part of the swing arm that obstructs the beam after reflection on the ASM (Figure 60). A static structure with an optimized footprint ensures not too many actuators are obstructed. Simulations by Lorenzo Busoni (Arcetri) show the residual obstruction has only negligible influence on the wavefront measurement.

To ensure the critical positioning of the CalUnit within an envelope of 1 x 1 x 1 mm^3 is achieved over the operational temperature range, the structure of the CFRP tubes has been thermally adapted to the coefficient of thermal expansion (CTE) of the ASM's swingarm. Finite element analyses ensured us that vibrations and static flexure are within an acceptable range. The structure is very stiff; the sag (in the z direction) of the CalUnit when the telescope is pointing at zenith is less than 0.5mm.

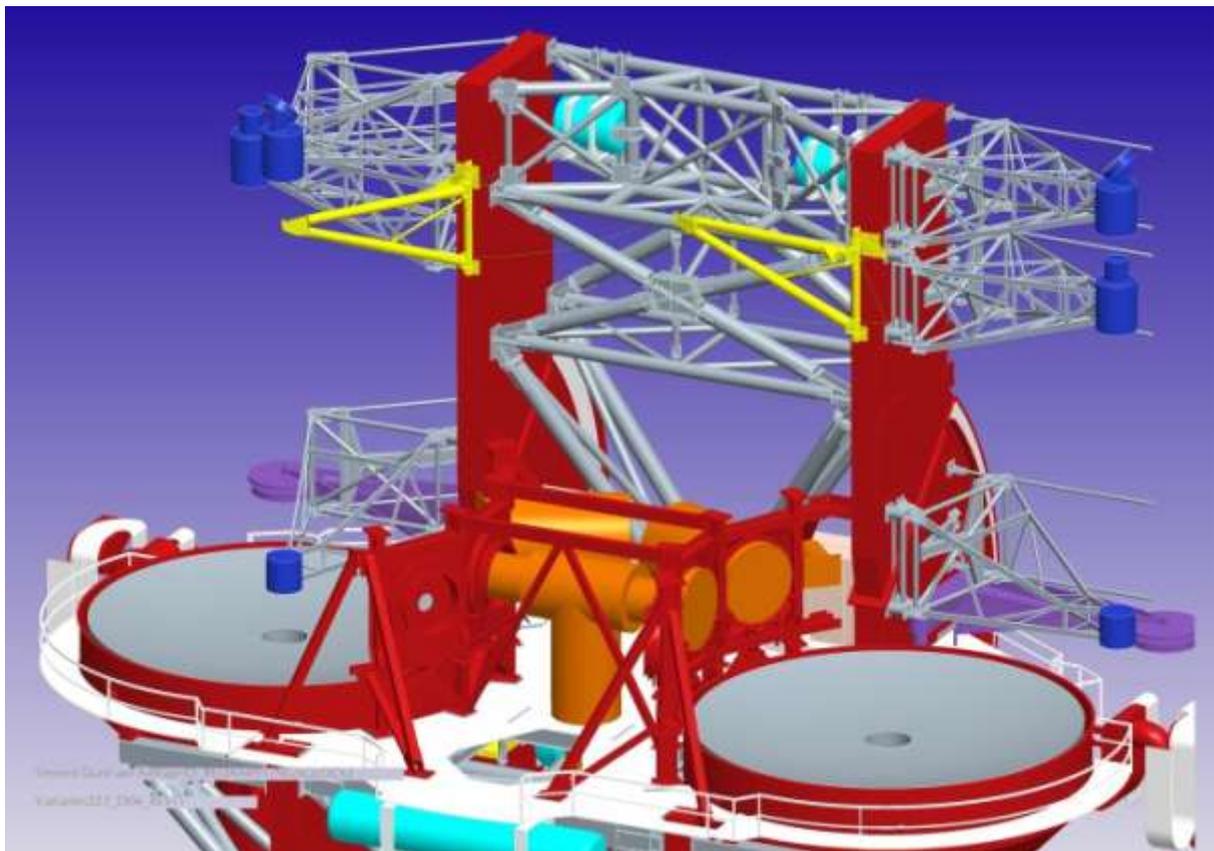


Figure 59: The location of the CalUnit swing arms (yellow) at LBT.

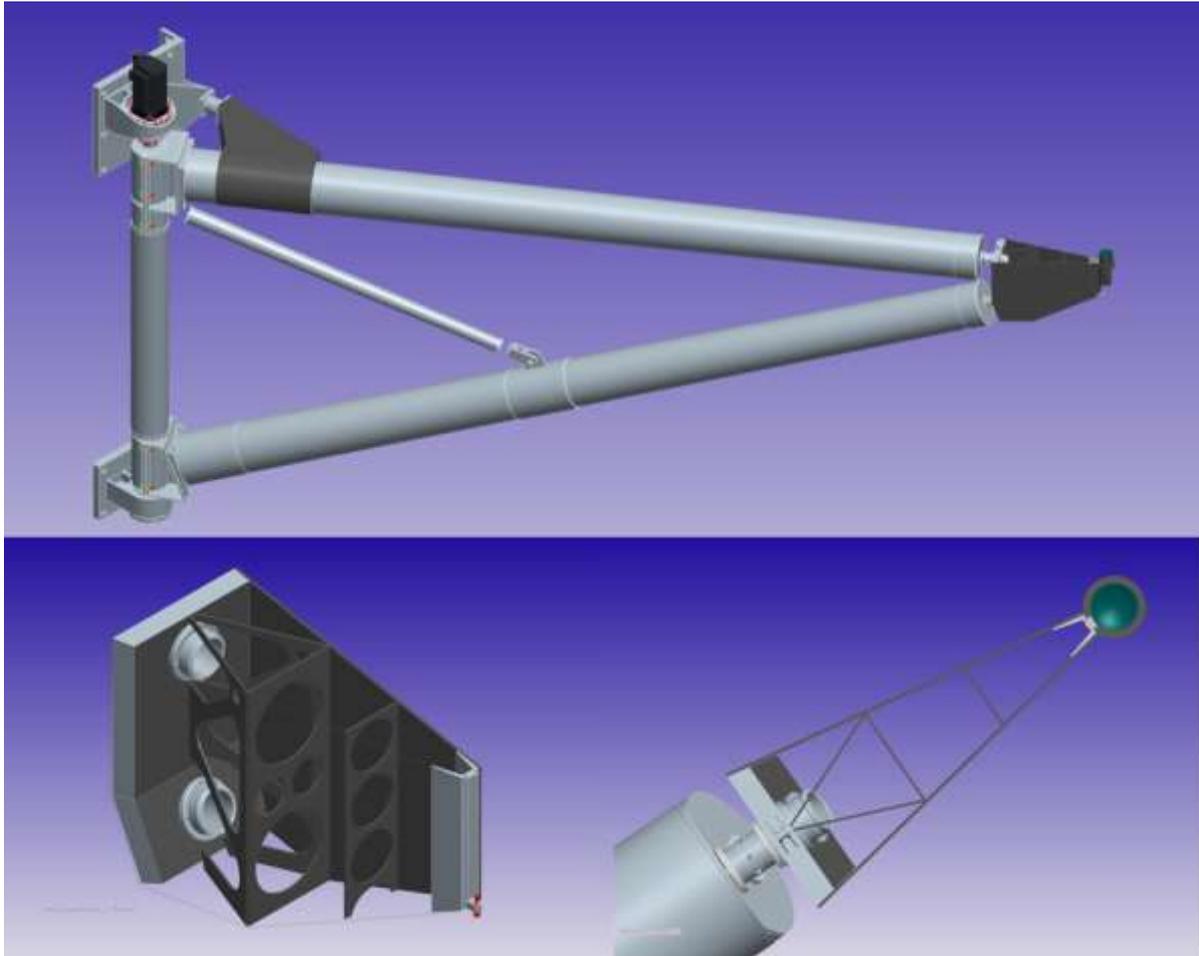


Figure 60: Drawings of the swing arm. Top: The whole arm, ~4.50m long. Bottom: the front end interface, designed to minimize the obstruction of the beam.

3.5.1 PDR Design

The layout of the CalUnit optics as described above is shown in Figure 61. The corresponding prescription of each of the components in the ZEMAX design is given in Table 5. The CGH was calculated by DIOPTIC, a possible provider of the CGH. Optimization for minimal optical path difference (a flat wavefront) was performed to demonstrate the optical quality with a CGH. For the final design of the whole system, the CGH was optimized to include the residual wavefront aberrations from the sky, mainly coma as shown in Figure 63. Figure 62 depicts the wavefront of the CGH calculated with ZEMAX. The spot diagram is still dominated by the uncorrected coma. The CGH is able to control the aberrations of the CalUnit and M2 down to less than $\lambda/10$ (RMS = 0.0738 waves), as shown in Figure 62. Introducing the necessary additional aberrations in a further design step was not problematic.

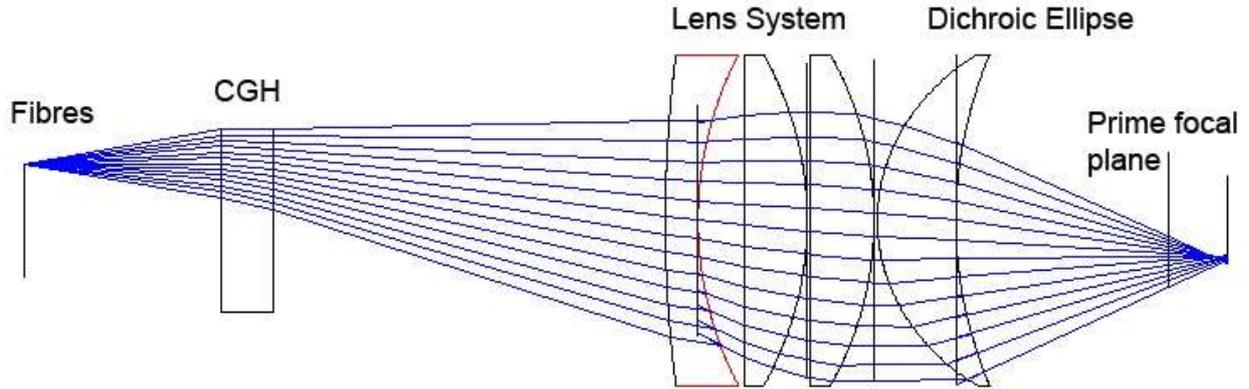


Figure 61: Layout of the CalUnit optics from the off-axis fiber source (left) to the 12km conjugated prime focus (right).

Table 6: Optical components and prescription taken from the ZEMAX lens data editor.

Surf	Radius	Thickn.	Glass	Diam.	Conic	Comment
OBJ	Infinity	-15135.2		159.8		BFL 12km conj.
STO		1974.3	MIRROR	911.0	-0.7328	M2
2	∞	1054.6		959.6		Dist. to M2
3	∞	9.1		13.4		Dist. to 12km focus
4	∞	32.5		20.6		Dist. to Primefocus
5	-63	12	BK7	50.8	-0.8860	On-axis LS Mirror
6	-26.7	0		50.8	-0.2620	1st Lens custom
7	∞	0.5		49.2		Distance 1st-2nd
8	51.5	9.7	BK7	49.6		2nd Lens TH LA1050-A
9	∞	0		49.6		
10	∞	0.5		48.3		Distance 2nd-3rd
11	51.5	9.7	BK7	50.8		3rd Lens TH LA1050-A
12	∞	0		50.8		
13	∞	7		42.2		Distance 3rd-4th
14	-55.3	5	BK7	50.8		4th Lens TH LF1829-A
15	-200	0		50.8		
16	∞	60		35.9		Distance to CGH
17	∞	8	F_SILICA	28.0		Front CGH Substrate
18	∞	29.9		27.8		CGH - Binary1
IMA	∞	-		17.3		

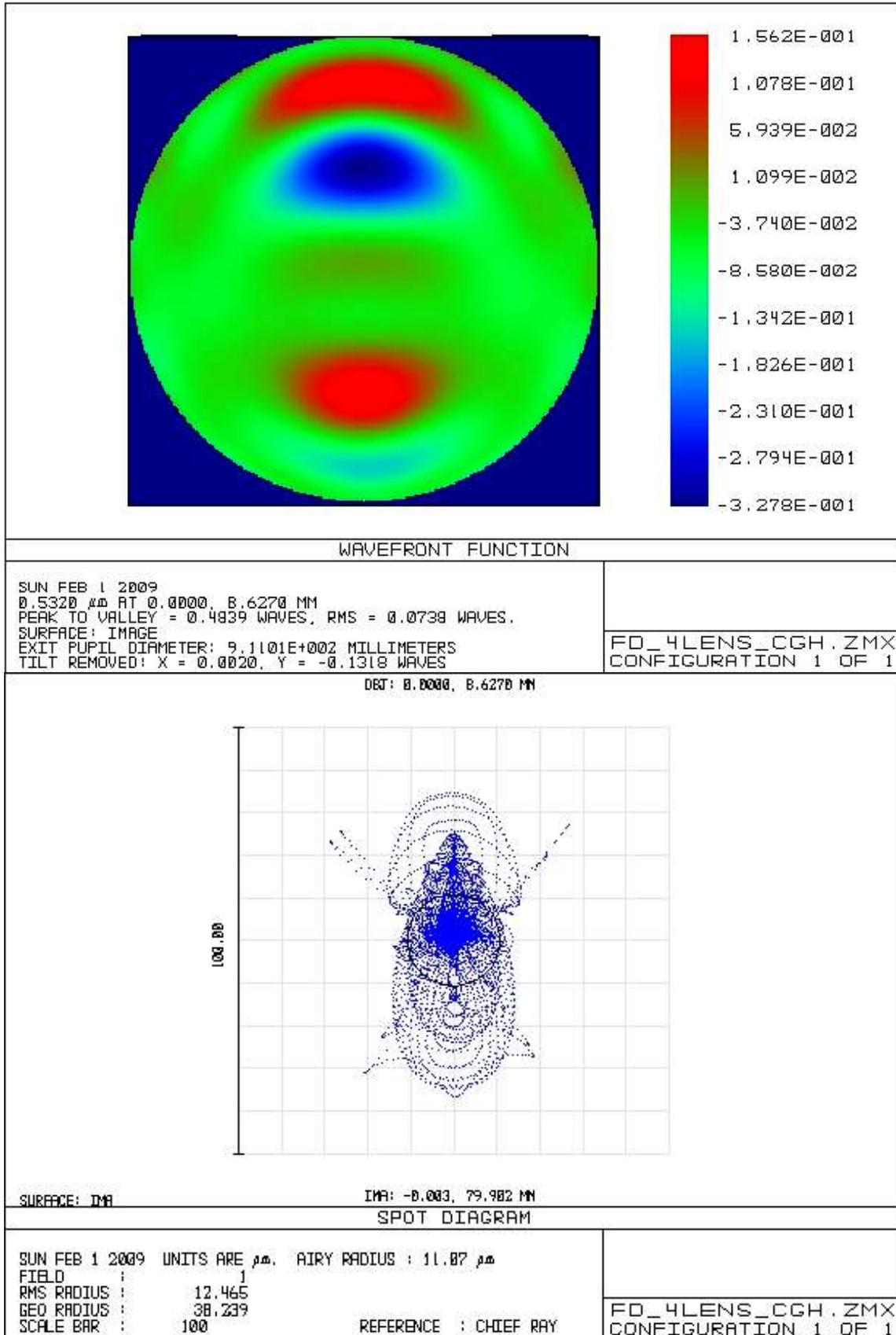


Figure 62: Top: The wavefront of the CalUnit with the CGH, optimized for a flat wavefront. It demonstrates the control of the aberrations to better $\lambda/10$ (RMS WFE is 0.073 waves). Bottom: The spot diagram corresponding to the wavefront shown above.

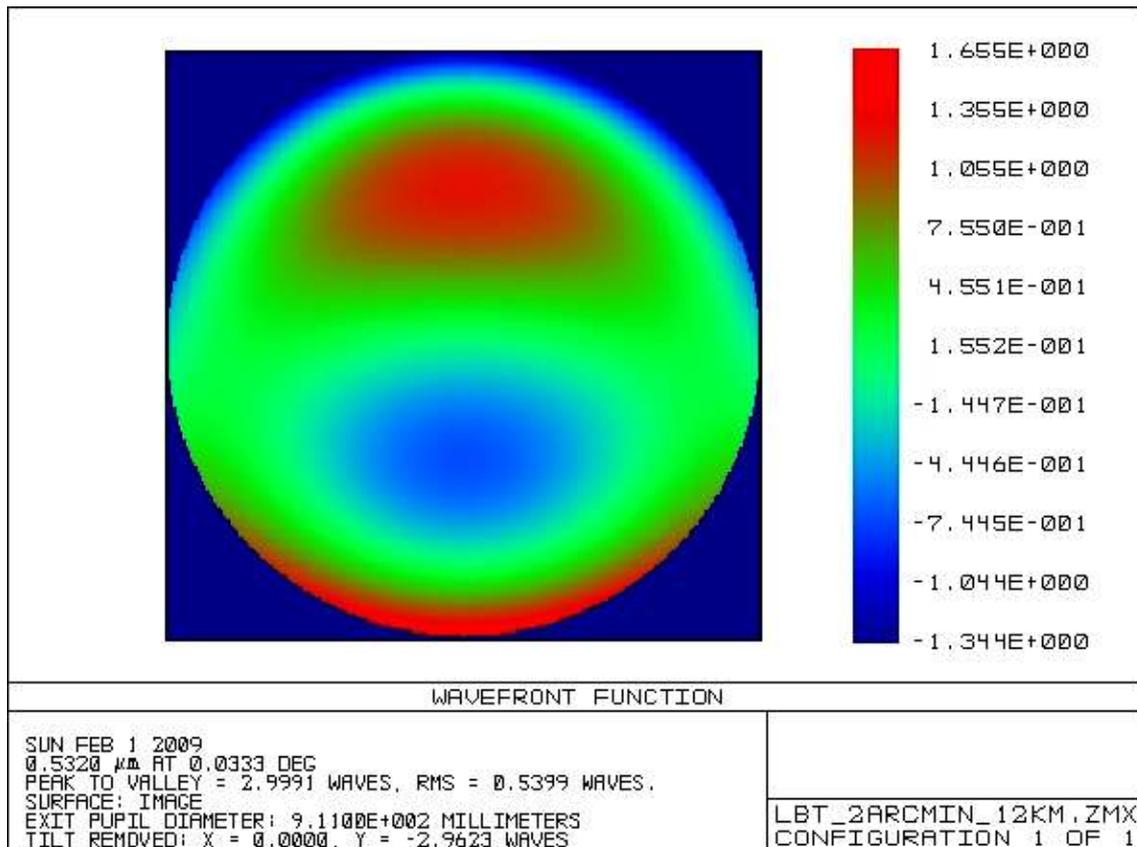


Figure 63: The wavefront of a point source at 12km altitude, 2 arcminutes off axis.

3.5.1.1 Optical Tolerancing

With the advice of Andrew Rakich (LBTO), a thorough tolerance analysis for the CalUnit was carried out. The parameters of the custom lens and airspaces were foreseen as compensators. To assess the expected performance, Monte Carlo analyses of the possible configurations with given tolerances were performed with ZEMAX. The Monte Carlo runs consisted of 100 configurations each, with a damped least square optimization of the respective compensators, using the Merit Function (calculating the wavefront error in the three off-axis beams) as criterion. The distance between M2 and the lens was kept fixed during the whole procedure and not used as a compensator. This ensures that the distance necessary for on-axis operation can be maintained, to facilitate the simultaneous operation of the on- and off-axis modes. For the same reason, the elliptical front surface of the CalUnit, which makes the dichroic mirror, was not used as a compensator and toleranced very tightly. With these constraints we were able to disentangle the tolerancing for the on- and the off-axis parts.

Taking a conservative approach, after each subsequent Monte Carlo run a representative configuration one standard deviation worse than the mean of the distribution was chosen as input to the next run. We consider this a typical system with these tolerances. Care was taken to ensure that the footprint on the CGH was still homogeneously illuminated. The tolerancing consisted of the following steps:

- 1) Tolerancing of radii, thicknesses, surface figures and wedges of the spherical stock lenses with the values given in the Thorlabs catalog, using the radius and conic constant of the convex asphere and the airspaces between elements as compensators.
- 2) Radius and conic constant of convex asphere with airspaces as compensators.
- 3) Tolerancing of airspaces without any compensation.
- 4) Tolerancing of decenter and tilt of elements without compensation.

In Table 7, Table 8 and Table 9, the assumed tolerances are listed. For the stock lenses, numbers were taken from the catalog. The mechanical tolerances for the spacers are routinely achieved in our workshop.

Note in particular the tolerances on the convex asphere: 0.5% for the radius and 10E-4 for the conic constant. The tolerance on the conic constant corresponds to a deviation from the designed shape of more than 700 nm, which should be directly measurable. The surface figure tolerance was made half as tight as for the other surfaces, because of the steepness of the curve. We have identified possible manufacturers.

The combined effect of these tolerances is shown in the comparison of a wavefront map of the unaltered system and the wavefront maps for a typical system with the tolerances in Table 7 (Figure 64). The WFE increases from about 1/12th to a maximum of 1/3rd of a wave.

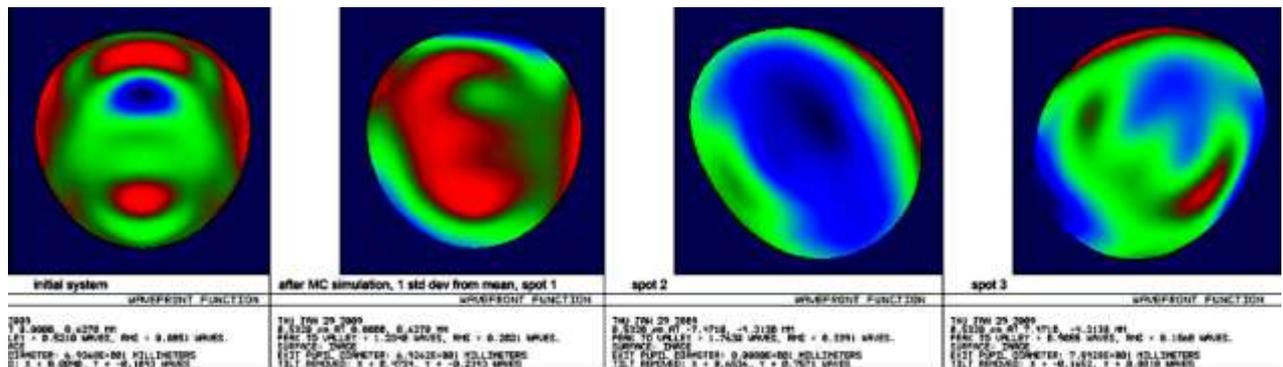


Figure 64: WFE comparison for a perfect system (leftmost) and the three LGS wavefronts with tolerances applied to the lens elements. The WFE increases from about 1/12th of a wave to a maximum of 1/3rd of a wave.

Table 7: Zemax tolerancing of lens parameters

#	Type	Int1	Nominal	Min/Max	Comment
Compensators					
2	COMP0				Distance fibre - CGH
3	COMP15				Airgap 1st-2nd lens
4	COMP20				Airgap 2nd-3rd lens
5	COMP25				Airgap 3rd-4th lens
6	COMP27				Radius convex asphere
7	COMP27				Conic convex asphere

CGH Substrate				
10	TEZI 6	0.00	$\pm 2.5E-005$	Surface figure Zernike 4-11 front
11	TFRN 6	0.00	± 1.000	Flatness (± 1 fringe)
12	TPAR 5	0.00	± 0.025	Wedge (x, ± 1.5 arcmin)
13	TPAR 5	0.00	± 0.025	Wedge (y, ± 1.5 arcmin)
14	TTHI 4	8.00	± 0.100	Thickness CGH substrate
16	TTHI 9	60.00	± 0.050	Thickness airgap CGH -1st lens
Lens 2 - Thorlabs TH-LF1829				
19	TEZI 11	0.00	$\pm 2.5E-005$	Surface figure Zernike 4-11 front
20	TEZI 13	0.00	$\pm 2.5E-005$	Surface figure Zernike 4-11 back
21	TRAD 11	200.00	± 2.000	Radius front ($\pm 1\%$)
22	TRAD 13	55.30	± 0.550	Radius back ($\pm 1\%$)
23	TPAR 12	0.00	± 0.025	Wedge (x, ± 1.5 arcmin)
24	TPAR 12	0.00	± 0.025	Wedge (y, ± 1.5 arcmin)
25	TTHI 11	5.00	± 0.100	Thickness of lens
Lens 3 & 4 - Thorlabs TH-LA1050				
28	TEZI 16	0.00	$\pm 2.5E-005$	Surface figure Zernike 4-11 front
29	TEZI 18	0.00	$\pm 2.5E-005$	Surface figure Zernike 4-11 back
30	TRAD 18	-51.50	± 0.500	Radius front ($\pm 1\%$)
31	TFRN 16	0.00	± 1.000	Flatness back (1 fringe)
32	TPAR 17	0.00	± 0.025	Wedge (x, ± 1.5 arcmin)
33	TPAR 17	0.00	± 0.025	Wedge (y, ± 1.5 arcmin)
34	TTHI 16	9.70	± 0.100	Thickness of lens
Lens 5 - custom asphere				
46	TEZI 27	0.00	$\pm 5.0 E-005$	Surface figure Zernike 4-11 back
47	TEZI 29	0.00	$\pm 2.0 E-005$	Surf. fig. Z4-Z11 dichroic mirror
48	TRAD 29	63.00	± 0.120	Radius dichroic mirror ($\pm 0.2\%$)
49	TCON 29	-0.886015	$\pm 1.0E-003$	Conic dichroic mirror
50	TPAR 28	0.00	± 0.025	Wedge (x, ± 1.5 arcmin)
51	TPAR 28	0.00	± 0.025	Wedge (y, ± 1.5 arcmin)
52	TTHI 27	12.00	± 0.100	Thickness of lens

Table 8: Tolerancing step 2, convex asphere

#	Type	Int1	Nominal	Min/Max	Comment
1	COMP0				Distance fibre - CGH
2	COMP15, 20, 25				Airgaps lens 1 - 4
6	TRAD 27		26.648134	± 0.130000	Radius convex asphere ($\pm 0.5\%$)
7	TCON 27		-0.265703	$\pm 1.0E-004$	Conic convex asphere

Table 9: Tolerancing step 3&4: airgap thicknesses, airgap wedges and centring.

Airgap thicknesses	± 0.01	all airgaps
Airgap wedges	± 0.005	0.3 arcmin ($\pm 5 \mu\text{m}$ runout)
Centring	± 0.01	

The configuration shown in Figure 64 was subsequently used for the tolerancing steps 2) to 4). The assumed tolerances are given in Table 8 and Table 9. I want to point out that compensation was applied like described before.

To show the results of the tolerancing steps 2 to 4 we again present a wavefront map of a configuration one standard deviation worse than the mean to show the wavefront degradation (Figure 65). This configuration includes all the listed tolerances of Table 7 to Table 9. The WFE is still less than 0.8 waves RMS and therefore meeting the specification of 1λ RMS, leaving room for positioning tolerances of the whole unit w.r.t. the ASM.

About half of the WFE is contributed by the deviation of the conic constant from the optimum of the convex asphere, which undercompensates spherical aberration. This is rotationally symmetric, and can easily be compensated with the CGH itself. During tolerancing, the CGH was not used to compensate aberrations caused by the tolerances. This would remove rotation invariant aberrations to a high degree, effectively cutting the residual aberrations in half. A quick inspection using ZEMAX confirmed this. A list of the power in the first 15 Zernike modes for the three wavefronts shown is given in Table 10, and an example for the compensation of spherical aberration with the CGH is shown in Figure 66. The beam that produced the first map in Figure 65 was used as input, and the RMS wavefront error minimized by optimizing the CGH prescription in ZEMAX. The remaining residual RMS could be minimized to 0.275 waves on the focal plane of the WFS. While this has to be checked with the manufacturer, it shows the correction achievable with the CGH. Note that the CGH was altered only very little! This level of correction is achievable in practice, in conformity to my tolerancing procedure. Measurements of lens radii with the necessary precision can be obtained with the common tools of an optics shop, for example by the lens manufacturer. Distances and diameters can be measured with a tactile coordinate measuring machine. The information from the measurements can directly be used to update the CGH design.

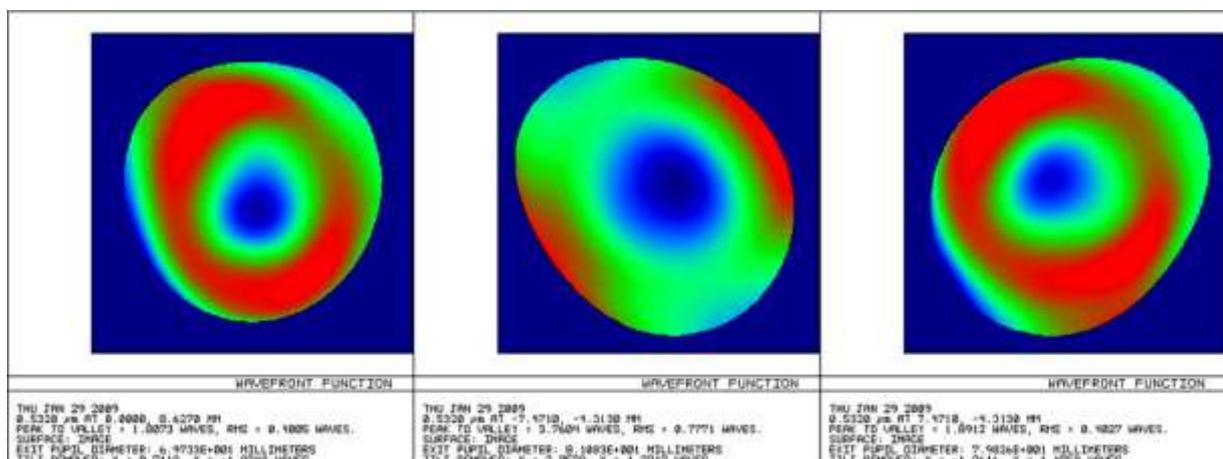


Figure 65: Wavefront maps of the three laser beams after applying all tolerances.

Table 10: Zernike coefficients for the three wavefront maps in Figure 65.

RMS (to chief) 0.954	RMS (to chief) 1.246	RMS (to chief) 0.900
RMS (to centroid) 0.399	RMS (to centroid) 0.779	RMS (to centroid) 0.403
Variance (waves ²) 0.159	Variance (waves ²) 0.607	Variance (waves ²) 0.162
Z 1 0.794	Z 1 1.871	Z 1 0.768
Z 2 0.327	Z 2 1.112	Z 2 -0.591
Z 3 -0.992	Z 3 0.600	Z 3 0.674
Z 4 -0.186	Z 4 0.574	Z 4 -0.177
Z 5 -0.199	Z 5 0.474	Z 5 -0.037
Z 6 0.150	Z 6 0.282	Z 6 -0.236
Z 7 -0.038	Z 7 0.036	Z 7 0.037
Z 8 0.099	Z 8 0.098	Z 8 0.092
Z 9 0.151	Z 9 0.066	Z 9 0.125
Z 10 0.029	Z 10 0.023	Z 10 0.014
Z 11 -0.431	Z 11 -0.339	Z 11 -0.430
Z 12 0.130	Z 12 -0.008	Z 12 -0.109
Z 13 -0.025	Z 13 -0.001	Z 13 0.083
Z 14 -0.061	Z 14 0.001	Z 14 0.009
Z 15 0.010	Z 15 0.001	Z 15 0.062

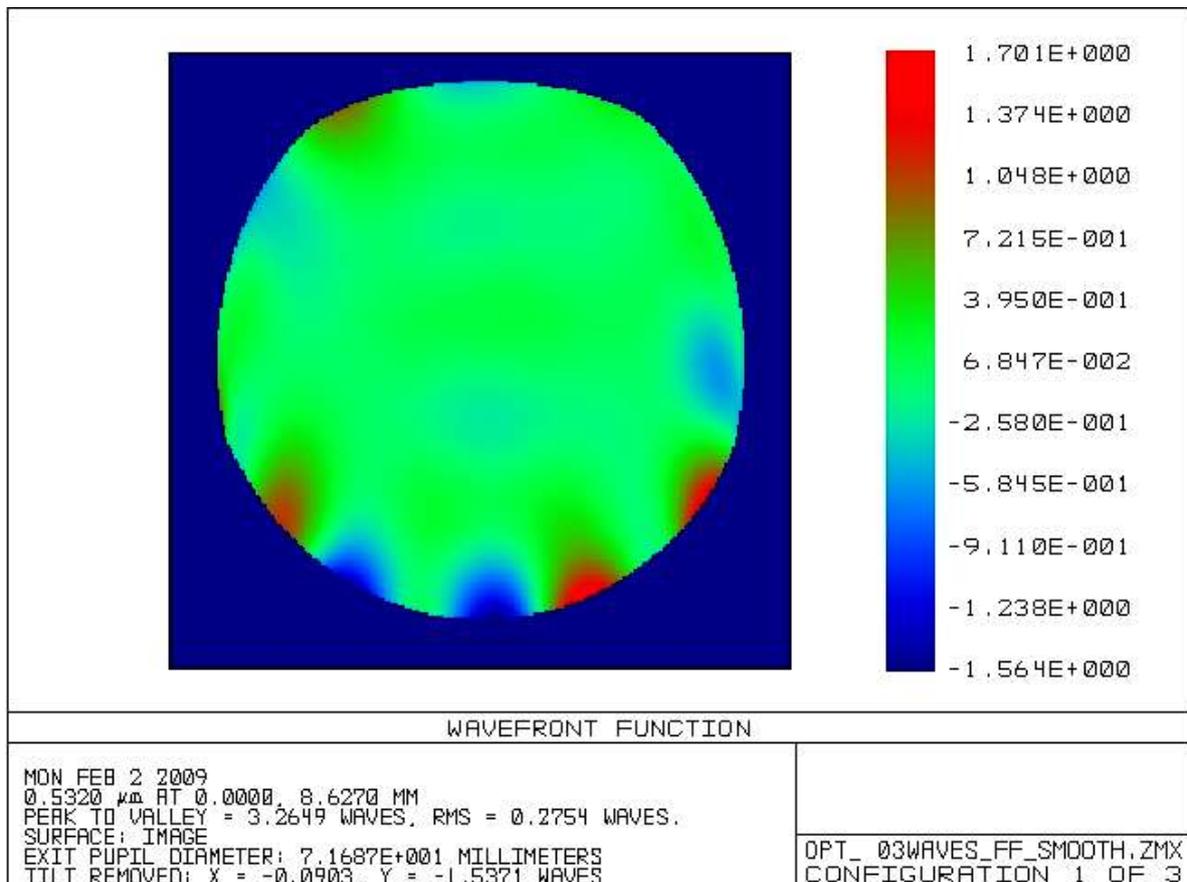


Figure 66: Wavefront if the CGH is reoptimised to correct for rotationally symmetric aberrations assuming real values for the lens parameters after tolerancing.

3.5.1.2 Assembly and Testing plan of CalUnit

For the PDR design it was planned to assemble the CalUnit in the labs at LSW or MPIA. Some of the necessary tools are available in-house, including an interferometer to aid the alignment. However, fast spherical interferometer lenses suitable for measuring the steep curvature of the optical elements, and air-bearing spindles with tactile probes to measure centration would need to be bought. Moreover, no one at either of the institutes has significant experience with such precise alignment of a lens assembly. This presented a risk that had to be mitigated. Despite these concerns, I developed the following assembly strategy for the CalUnit. To relax the tolerances necessary for the optical and mechanical elements, specific components will be obtained and measured, with the parameters of custom made components subsequently adjusted to compensate for deviations from the designed values. This includes especially the mechanical parts that are being manufactured in the MPIA workshops. These components will be tested on the 3D tactile coordinate measurement machine in the MPIA with a precision of about 5 μ m. Included parts are the lens barrel, fiber holder plate and lens distance rings.

As the tolerancing shows (Section 3.5.1.1), this strategy would allow us to use stock lenses except for the element with the ellipsoidal dichroic mirror. The CalUnit should be assembled in the following order, based on how I used the compensators during tolerancing:

1. Obtain and measure the stock lenses.
2. Adjust the conic constant and radius of the custom lens' convex surface.
3. Measure the custom lens.
4. Adjust the airspace values accordingly.
5. Machine and measure the spacers.
6. Adjust the CGH prescription to optimize the wavefront aberrations.
7. Obtain the CGH (together with a DOE to measure the whole unit).
8. Assemble the lens.
9. Verify performance in the lab.

Measurements would need to be carried out in-house (LSW/MPIA) and by the lens manufacturer. The measurement of the spherical lenses with the necessary precision is a standard task with common lens measuring systems. The aspheric surfaces are more problematic to measure, however. One has to separate between the elliptical concave front surface, that is used as a mirror and has to give diffraction limited performance, and the back surface, which is seemingly much harder to measure but has less tight tolerances if we accept a higher systematic wavefront error on the off-axis sources, which then can be calibrated on sky. For the front surface, a either diffractive optical element could be used to verify the specifications, or it could be tested in the conic conjugates, while for the back surface a profilometric measurement (e.g. tactile) could be sufficient.

The CGH is inherently easier to measure than the asphere as the measurement is only 2D, one of the great advantages of diffractive optical elements. A test report is included by the manufacturer as standard. The transmitted wavefront of the substrate can be checked with the interferometer in house.

The fully assembled CalUnit should be tested interferometrically by the manufacturer with a dedicated diffractive optical element acting as a null corrector.

3.5.2 FDR design

The concerns about the assembly together with the desire to achieve the goal of a WFE of $\lambda/10$ RMS led to a revisit of the design, considering custom spherical elements throughout the objective. As the convex asphere proved to be expensive and difficult to manufacture, I looked into ways of exchanging it against a spherical surface, and balance the system with the additional freedom due to the custom radii of the lenses. Lastly, I wanted to adapt fused silica as the material for all the lenses, to make the unit less sensitive to thermal fluctuations. The modified design was presented at the FDR in March 2010 and passed the review.

3.5.2.1 Optical design

The layout of the CalUnit optics for the FDR design is shown in Figure 67. Table 11 contains all the components and their prescriptions taken from the ZEMAX design. The ideal wavefront at the CGH surface was determined by tracing the rays from an ideal spot at the WFS plane backwards through the system, and optimizing the CGH which was represented as a so-called *Binary 1* surface in Zemax.

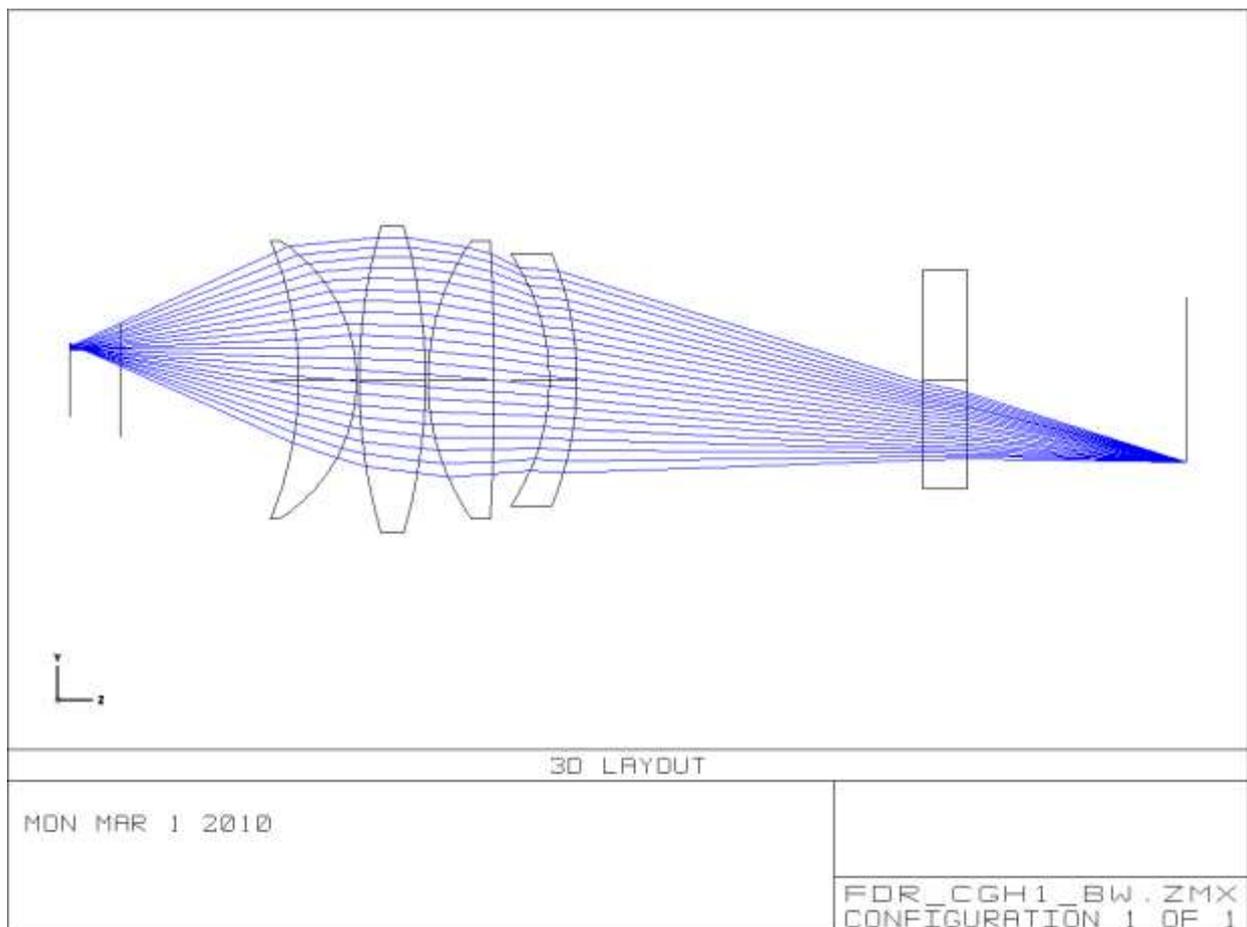


Figure 67: Layout of the CalUnit optics (FDR design) from the off-axis fiber source (right) to the 12km conjugated prime focus (left).

Table 11: Optical components and prescription taken from the ZEMAX lens data editor for the FDR design

Surf	Radius	Thickn.	Glass	Diam.	Conic	Comment
OBJ	Infinity	-15135.2		159.8		BFL 12km conj.
STO	1974.3	0	MIRROR	911.0	-0.7328	M2
2	∞	1054.598		959.6		Dist. to M2
3	∞	9.102		13.4		Dist. to 12km focus
4	∞	32.460		20.6		Dist. to Primefocus
5	-63	10.5	FS	50.8	-0.8850	On-axis LS Mirror
6	-30	0		50.8		L1
7	∞	0.5		51.1		Distance 1st-2nd
8	100.555	12	FS	56.0		L2
9	-100.555	0		56.0		
10	∞	0.5		50.7		Distance 2nd-3rd
11	44.784	12	FS	50.8		L3
12	-464.5	0		50.8		
13	∞	10		45.3		Distance 3rd-4th
14	-41.31	5	FS	46		L4
15	-60.352	0		46		
16	∞	60		38.6		Distance to CGH
17	∞	8	F_SILICA	29.2		Front CGH Substrate
18	∞	39.7		28.8		CGH - Binary1
IMA	∞	-		30		

The CGH structure used for performance analysis of the design was calculated in close cooperation with DIOPTIC. Figure 68 depicts the wavefront of the FDR design including the coma. The CGH is able to control the CalUnit and ASM aberrations down to less than $\lambda/10$ RMS.

The main difference to the PDR design is the exclusive use of custom lens elements made from fused silica. This allowed us to remove the complicated aspheric convex surface on lens one in favor of different spherical surfaces on the other elements. Achievable tolerances on radii, surface figure and especially the critical tolerance on the wedges are considerably tighter (a factor of five for the wedge), facilitating a simpler mechanical design without radial adjustment of the lens elements and at the same time providing better wavefront accuracy. The use of fused silica throughout the design makes the objective less susceptible to temperature changes. Changing the first lens material from Clearceram-Z to fused silica, which has a lower index of refraction, made a slightly larger second lens necessary, but the diameter is still well within the central obstruction. The effect of the fused silica's higher CTE on the on-axis light source wavefront is negligible. The bulk scatter inherent to Clearceram-Z at visible wavelength is removed. The radii were fitted to a testplate inventory (of Tucson Optical Research Corp., TORC), to keep the additional cost for the custom elements low.

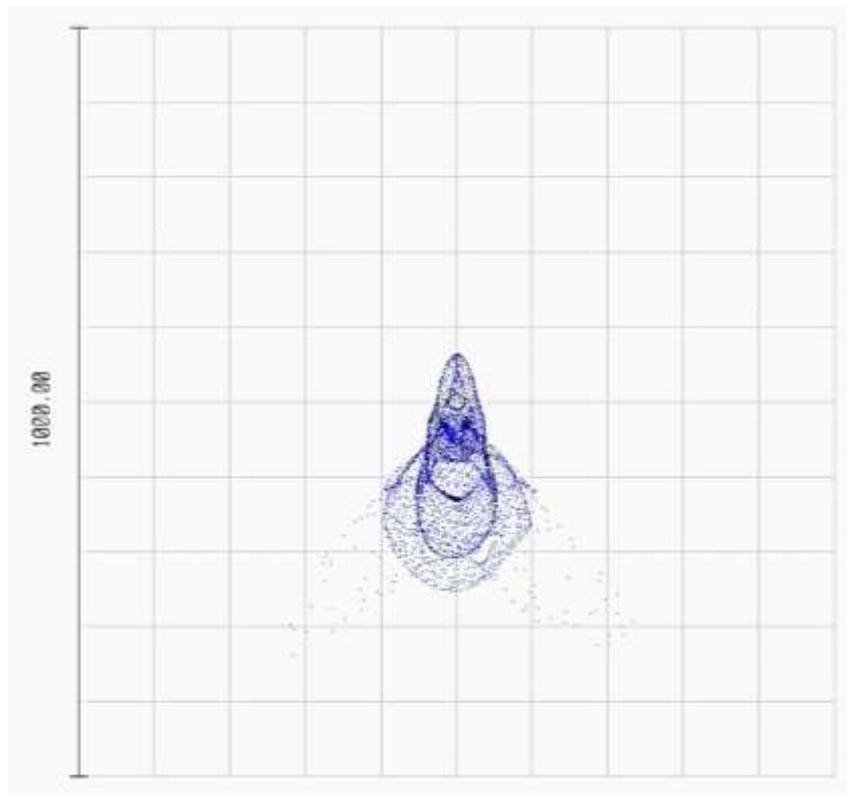
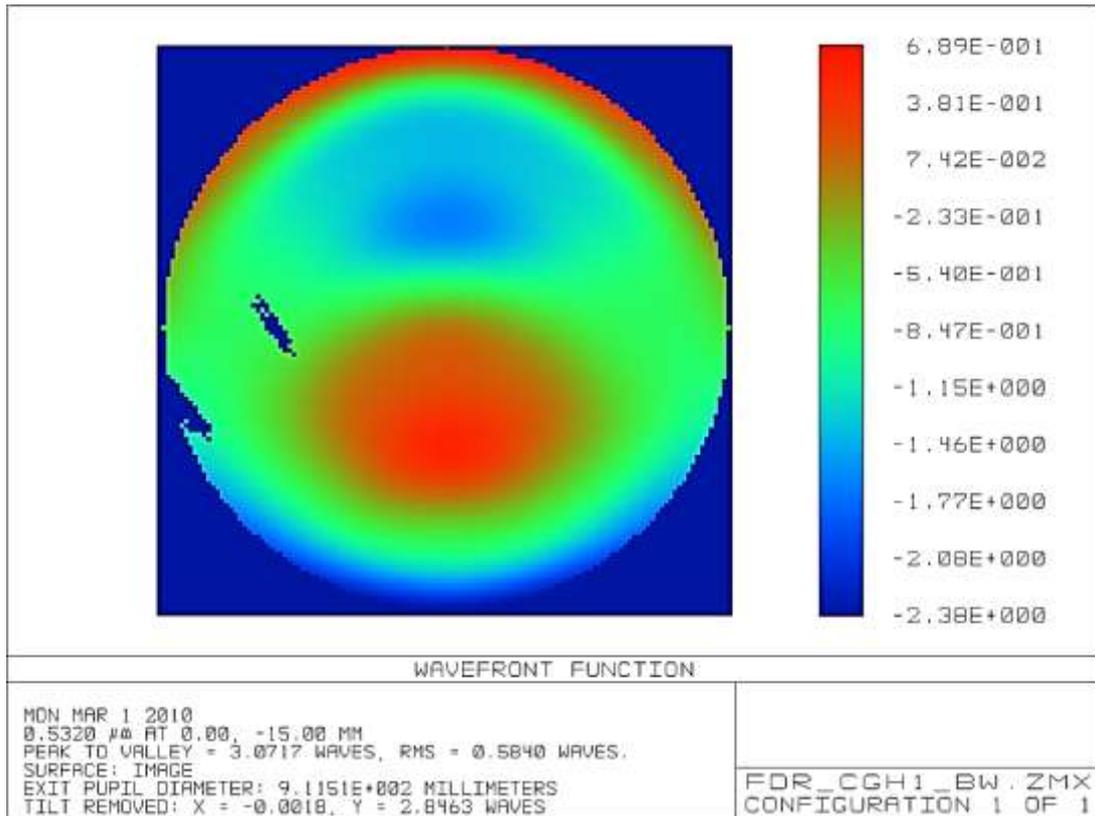


Figure 68: Top: The wavefront of the CalUnit with the CGH. The missing parts are artifacts from ZEMAX. WFE difference to the real beacon is 0.09 waves RMS. **Bottom:** The spot diagram of the CalUnit corresponding to the wavefront show above. The RMS spot size matches the beacon's spots well; the diagram shows some stray rays. These artifacts will be removed for manufacturing.

3.5.2.2 Tolerancing

A full tolerance analysis of the lens specifications as well as the assembly specifications was carried out for the PDR. The new design is very similar, and as the achievable tolerances of the components are known, I did not repeat the full investigation including the sensitivity analysis, but did forward modeling to assess the likely performance of the CalUnit with the new component specifications.

Specifications for the optical elements

Here I summarize the specifications for the optical elements:

Common specifications for the spherical lenses:

Material: Fused Silica

Surface Finish: 40 / 20 scratch-dig

Radius: 2 fringes of power in the test plate fit

Surface figure: ¼ fringe

Centre thickness: +/- 0.1 mm

Wedge: 20 arcsec

Specifications for the asphere:

Material: Fused Silica

Surface finish: 40/20 scratch-dig

Radius of convex sphere: 2 fringes of power in the test plate fit

Surface figure of convex sphere: $\lambda/10$ in transmission

Radius of concave asphere: +/- 0.05%

Surface figure of concave asphere: $\lambda/8$

Centre thickness: +/- 0.1 mm

Wedge: 20 arcsec

Centration of ellipse' vertex: 20 micron

Specifications for the CGH substrate:

Material: Fused Silica

Surface finish: 40/20 scratch-dig

Wedge: 5 arcsec

Surface irregularity: $\lambda/10$

Flatness: 1/4 fringe

Thickness: +/- 0.05 mm

Assembly tolerances

The tolerances for the assembling of the lens body and the CGH substrate in the same Invar lens barrel must conform to the following values (Table 12).

Table 12: Assembly tolerances

Airgap thicknesses	± 0.01 mm	all airgaps
Airgap wedges	± 0.005 mm	0.3 arcmin (± 5 μ m over full aperture)
Centring	± 0.01 mm	

Performance estimation

The new tolerances are tighter than the tolerances of the stock lenses in the PDR design. The wedge and radius tolerances in particular made it necessary to adjust the (x,y,z) position of each lens in the PDR design to match the as-built specifications. Both these parameters will be made considerably tighter (0.3 arcmin wedge vs. 1.5 arcmin, 2 fringes radius vs. 1%) with the custom lenses, rendering the adjustment unnecessary, as shown by the tolerance analysis.

To assess the resultant WFE of a system with these specifications, we did a Monte Carlo simulation of 100 configurations, perturbing the respective parameters in a random way. Based on the resultant distribution of the figure of merit for the 100 disturbed systems, we picked one with a figure of merit one sigma worse than the mean as representative. This is a slightly pessimistic assumption, giving us confidence that we will achieve this performance. The wavefront of this configuration matches the desired shape to a precision of $\lambda/5$ RMS. A comparison of the spots in the secondary focal plane for the design with and without tolerances applied is shown in Figure 70, the corresponding wavefront with the tolerances applied is shown in Figure 69. This facilitates straight-forward design and manufacturing of the mechanical components as well as the CGH, as no compensation is needed.

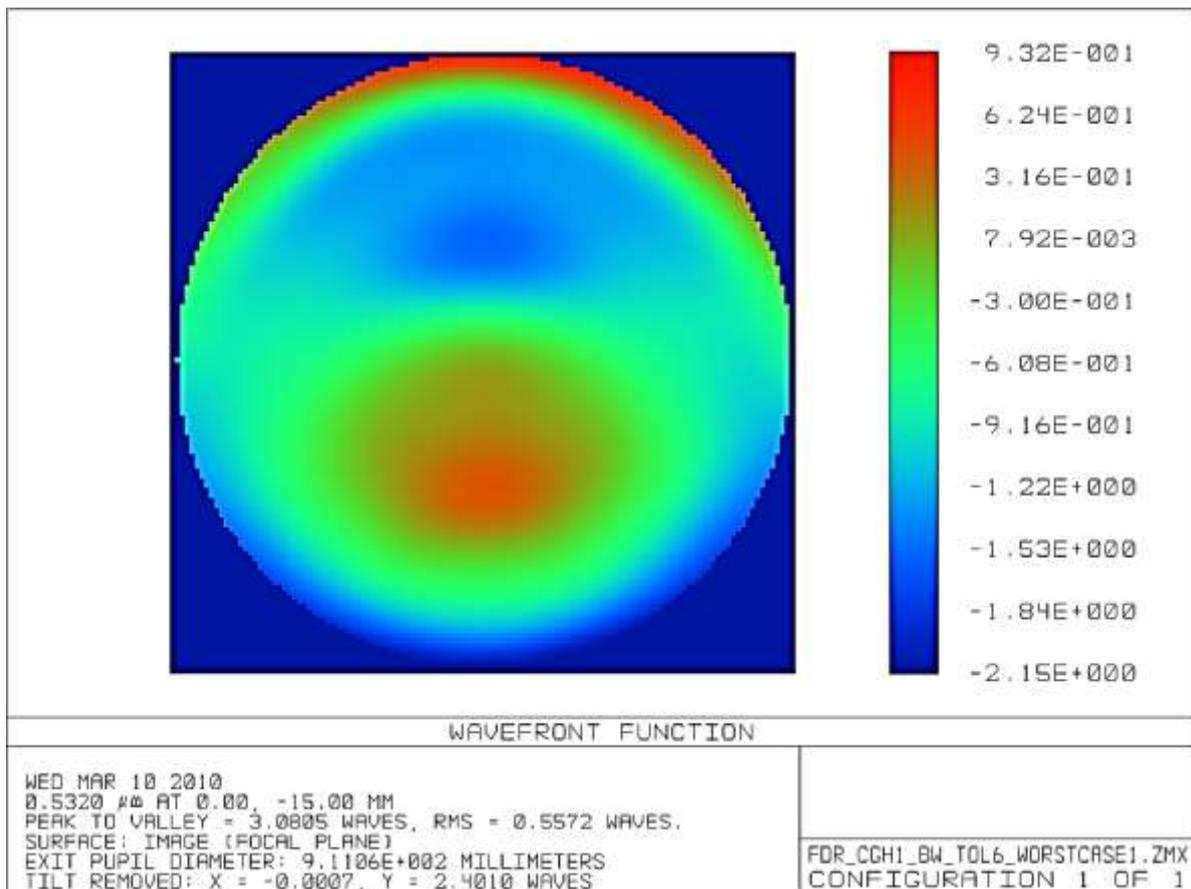


Figure 69: Wavefront of the CalUnit with tolerances applied.

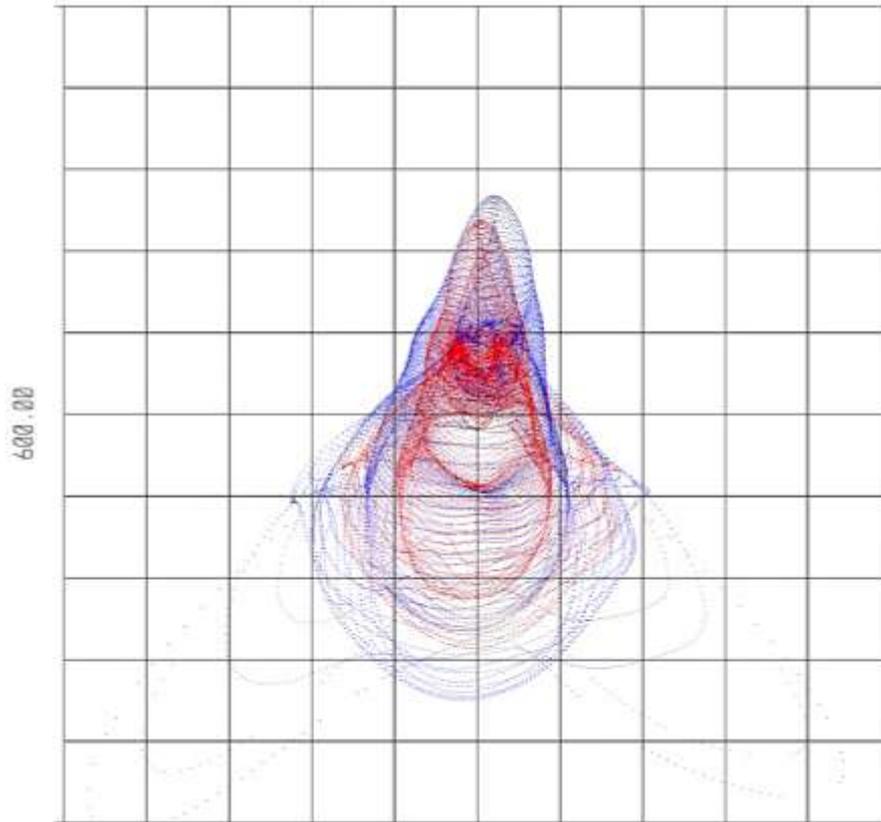


Figure 70: Overlay of the spot diagrams of the CalUnit with (blue) and without (red) tolerances applied.

3.5.2.3 Mechanical design

The mounting concept for the lenses was adjusted after PDR as well. As neither MPIA nor LSW have the necessary optical measurement capabilities to verify the alignment of the lenses, the assembly of the lens body will be done by a lens manufacturer.

The custom lenses are made to such tight specifications that no adjustment is necessary. The lens elements are set into a precision-made lens barrel with slip fits. Joe Appels of Tucson Optical Research Company (TORC), who has extensive experience with this technique, agreed to manufacture the lens system. TORC will grind the diameter of the lenses with high precision to ensure a tight fit with less than 5 micron gap. As the lenses L2 to L4 have decreasing diameters, their fits are directly and consecutively accessible during assembly. L1 will be set into a mounting ring, which then has a precision fit to the lens barrel. The centricity of the mounted ring will be verified after manufacturing. The lenses are held in place by lens rings.

All the mechanical components of the lens barrel will be made of Invar 36. Due to the small CTE of the Invar alloy as well as the fused silica of the lens elements, no athermalisation is necessary. The CTEs at room temperature of “Freecut Invar 36” and fused silica are 1.64×10^{-6} and 0.55×10^{-6} , respectively. This leads to a maximal radial shrinkage of a lens element over the full operating temperature range of less than 1.2 micron. The differential shrinkage between lens and barrel is less than 2.5 micron.

The integration of the CGH and the fibres into the lens barrel will be done by the vendor delivering the CGH. With an appropriate adjustment CGH, the substrate can be precision ground to a centration of better than 5 micron (personal communication S. Aigner, DIOPTIC). This makes any adjustment of the CGH unnecessary. MPIA will test if the fibre plate delivers tight enough position tolerances of the fibres to the mounted CGH to mount the plate directly in a fit, too. The positions of the fibres with respect to the CGH will be verified by an adjustment grating on the CGH substrate. If the specs cannot be met in this way, the fibre plate will be mounted in an x-y adjuster with solid-body joints, providing accurate alignment capability with good dimensional stability.

A mechanical concept was prepared for the FDR. Details of the lens barrel design will have to be verified by the lens manufacturer and the company providing the CGH. The outer dimensions of the lens barrel are approximately D: 67 mm, l: 180mm, the weight including lenses is less than 1900 g. This is compliant with the specifications and the capabilities of the swing arm. A summary of the requirements is given in Table 13. A concept drawing is shown in Figure 71.

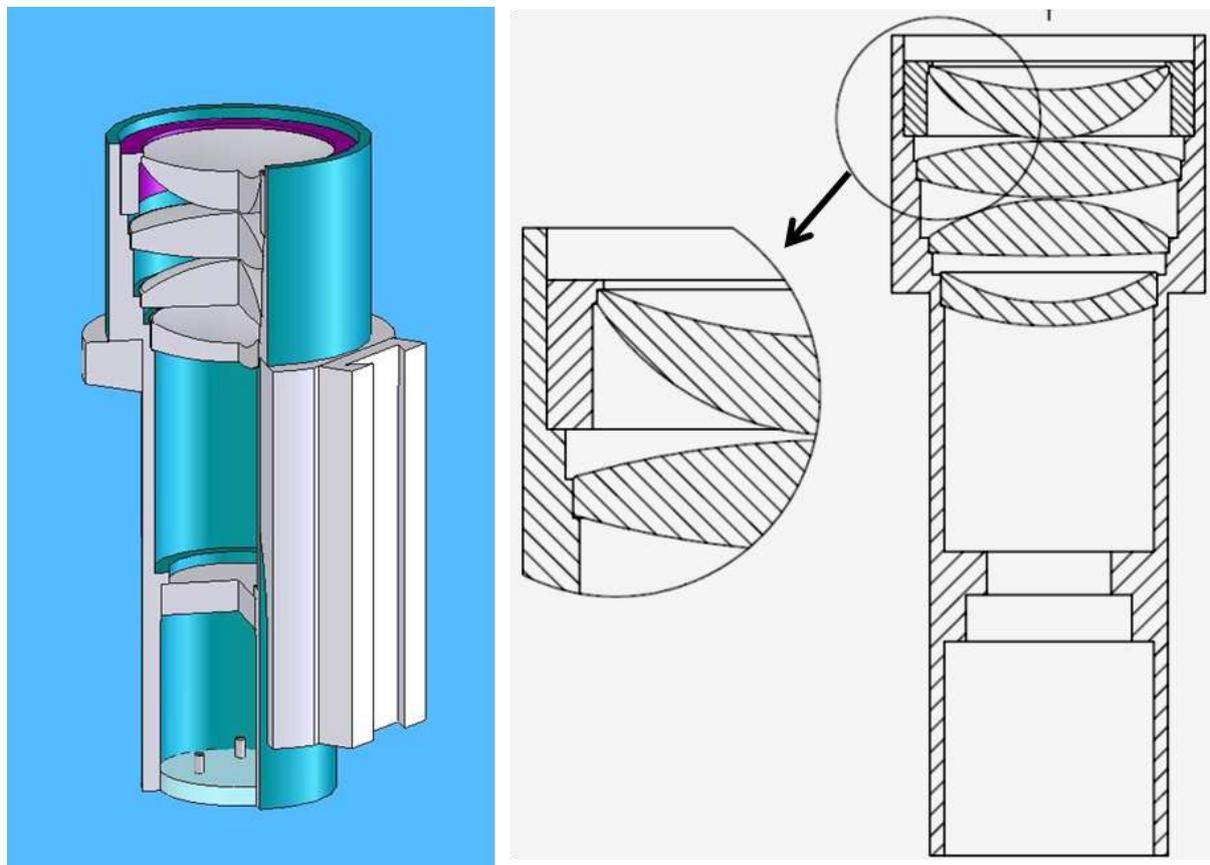


Figure 71: Conceptual drawing of the Invar lens barrel with the lenses inserted.

Table 13: Summary of the dimensions, loads and positioning of the CalUnit.

Dimensions		Source
CalUnit diameter	70 mm	Optical design
CalUnit height	180 mm	Optical design
Swing Arm load		
CalUnit	< 2 kg	Mechanical draft

Working Position (interface CalUnit – Swing Arm)		
Z (Reference: Primary mirror main plane)	9417.54 mm	Optical design
X (Reference: Primary mirror centre)	0 mm	Optical design
Y (Reference: Primary mirror centre)	0 mm	Optical design
Positioning repeatability x,y,z	<1 mm	Hexapod alignment
Positioning repeatability rot x, rot y	0.017°	Hexapod alignment
Stability after positioning (w/o telescope movement)	<10 μm	Optical design

3.5.2.3.1 Interface to Swingarm

We plan to mount the CalUnit with a dovetail to the mounting surface provided by the swingarm, to be able to easily detach and remount it without losing the alignment (see Figure 60). The interface between CalUnit and swingarm has to provide means to align the CalUnit w.r.t. M2 within the operating range of the hexapod. We defined this to be 1 mm³ in lateral movement and 1.5 arcmin tilt. Tilt adjustment to an axis parallel to the swingarm’s horizontal beam is provided by push-pull screws in the interface (Figure 72). The perpendicular direction can be adjusted by differentially changing the nominal length of the beams (using screws at the end of the beams). This also provides one adjustment for the lateral position, by changing both beams’ length in the same direction. Shift perpendicular to the swingarm is provided by adjusting the end stop of the swingarm. Focus is adjusted with a fine adjuster screw setting the z position of the CalUnit on a dovetail.

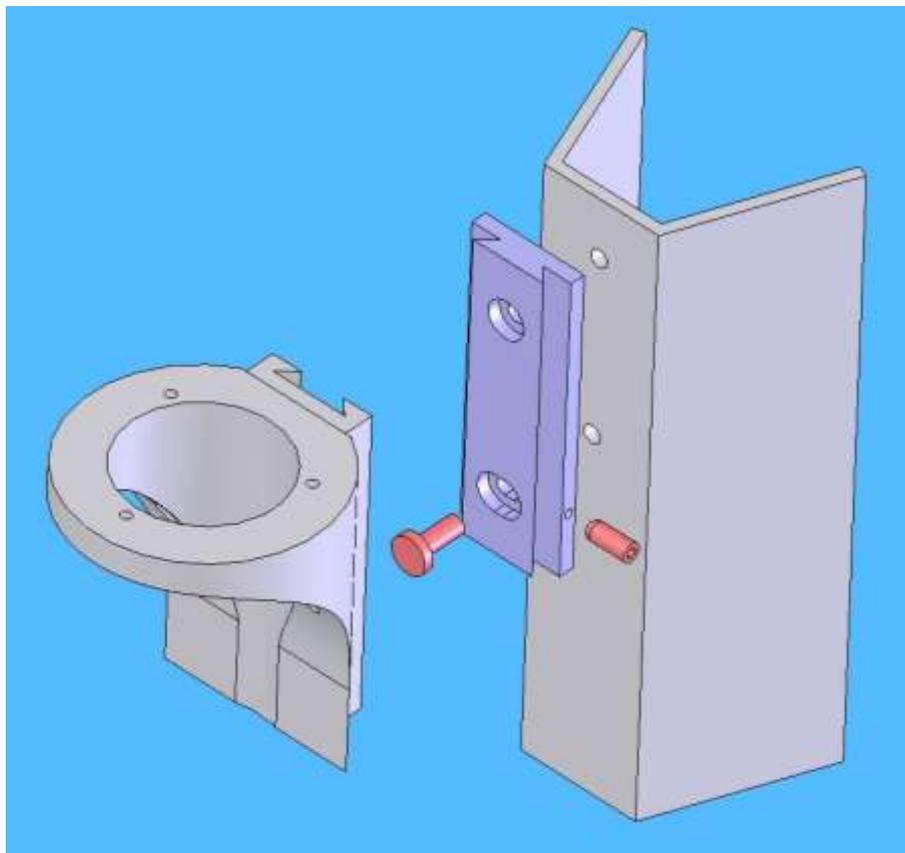


Figure 72: Concept for the CalUnit to swing arm interface. The CalUnit is mounted on a dovetail joint (blue), which also provides tilt adjustment in one axis (red bolts).

3.5.2.4 Light sources

In our case, the wavelength accuracy and stability of the light source illuminating the CGH, determined by the amount of WFE allowed, is rather relaxed, while for the highest precision measurements with diffractive optical elements, narrowband, frequency-stabilized light sources are typical. Stabilized gas lasers are normally used for that purpose. Unfortunately, at the design wavelength of 532 nm no gas lasers are available. Helium-Neon lasers have a transition at 543 nm, but this is already off the peak of the Pockels' Cell adjustment. At 532 nm, frequency doubled solid state lasers are readily available. A ZEMAX analysis showed that a wavelength stability of ± 0.3 nm is sufficient, before chromatic effects lead to unacceptable wavefront aberrations. This requirement is met by common units.

One concern with using a laser source are speckles. As we need to use a multimode fibre to resemble an extended spot on sky, and sample the wavefront in the pupil, small scale speckles can lead to varying actuator illumination. Several strategies to mitigate this effect have been discussed. The speckles could be temporally averaged by agitating the fibre, or modulating the wavefront with high order modes before feeding it into the fibre; commercial units based on membrane deformable mirrors can be readily obtained and reduce the speckle contrast to a few percent⁷. One has to keep in mind, though, that the exposure times set by the Pockels cell hold time are very short, and we average over three gating periods in one exposure. The frequency of the speckle modulation has to be at least 10 kHz (the laser pulse rate) for every snapshot within the exposure time of 1 millisecond to have a different pattern. Another possibility is to modulate the laser wavelength, as possible with diode lasers. Care must be taken to stay within the 0.6 nm waveband allowed. The same considerations as described above for the modulation frequency apply. A third idea is simply to employ a spectrally broader lightsource, using the full 0.6 nm waveband.

As the wavefront sensor is very sensitive, we do not need the high fluxes provided by laser sources. Additionally, the above mentioned modulation techniques are all costly, as is the laser source. For these reasons we chose as baseline design to use an LED for each fibre with narrowband filters to reduce the FWHM to the desired 0.6 nm. Using an LED has the advantage of a relatively short coherence length, reducing speckles. MPE has set up a fibre coupled LED, used for calibrations of the Pockels cell, that delivers 7 μ W in a 3 nm bandpass out of a 400 μ m fibre. Our calculations show this to be sufficient for our purposes. Other LED sources with better coupling are commercially available.

To reduce the spectral width of the light to 0.6 nm or less, two filters with broader transmission are tilted against each other, achieving the narrow pass band. This concept can be verified with a spectral scan of such an assembly, providing measurements of FWHM and absolute peak transmission. The filters under investigation are Thorlabs FL05532-1 types with 1 nm FWHM and similar filters from Semrock with a FWHM of 1.6 nm. An example of the resulting transmission from two filters with a relative tilt in series is given in Figure 73. I simply calculated the combined transmission of two filters with a relative wavelength shift between them by multiplying the transmission at each wavelength. This setup will be tested at MPIA. The filters that are used in the end should have an edge steepness at the half maximum point of better than $dT = 0.5$ per nm to ensure good throughput and a sharp transmission spectrum. Peak transmission of each of the filters should be no less than 80%.

⁷ Dyoptyka brochure 2009, www.dyoptyka.com

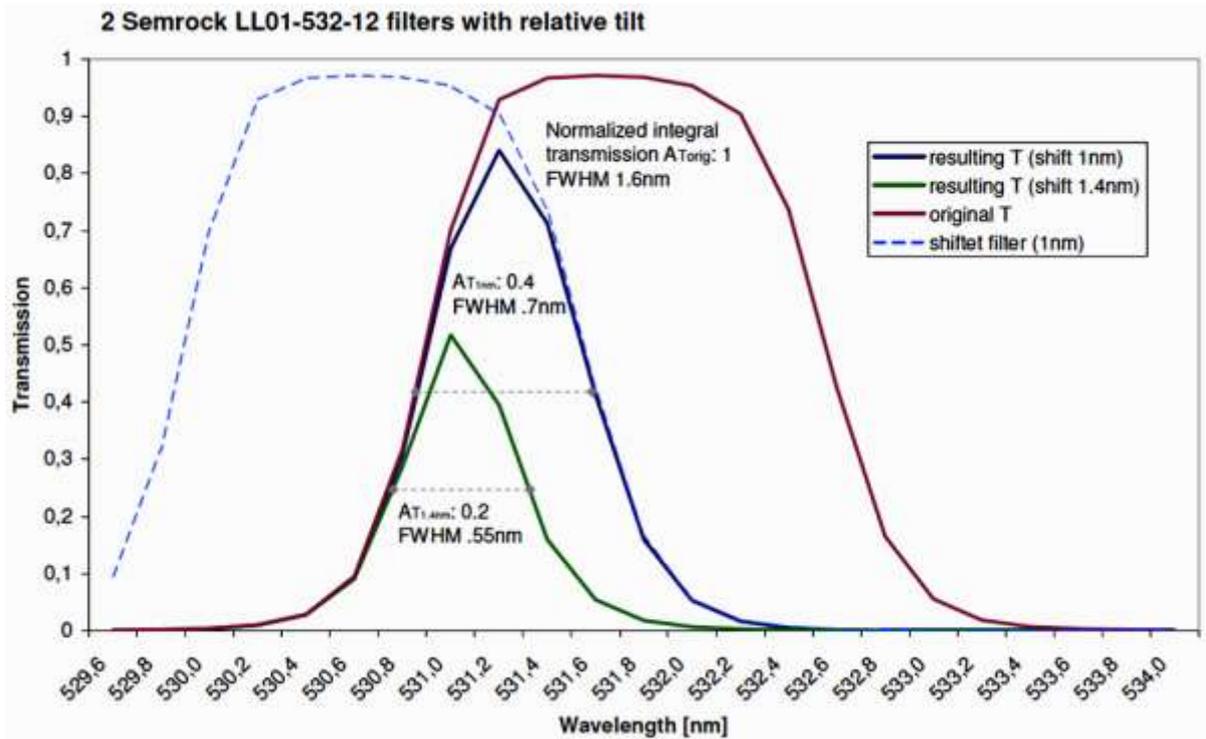


Figure 73: Combined exemplary transmission spectrum of two narrowband filters with a tilt between them.

3.5.2.4.1 Fibres and connectors

For the off-axis lightsources, step index multimode fibres made of silica and housed in protective buffers are planned. The numerical aperture (NA) needs to be higher than 0.2 to make sure the CGH is evenly illuminated, but should stay small to minimize losses. Connectors will be chosen to facilitate easy attachment to the fibre holder plate. Ferrule size should be no less than 2.5 mm diameter and the material ceramic. Pre-terminated fibres are preferred. Standard polish is sufficient.

3.5.2.4.2 Throughput calculation and necessary flux

The throughput of the calibration system from the fibre output to the WFS detector is estimated as follows:

CGH efficiency:	0.13 (Amplitude CGH)
CGH illumination geometry	0.60 (Fibre NA: 0.22, Lens system NA: 0.15)
CGH Substrate:	0.96 (one side uncoated)
7 FS Lens Surfaces (AR@532):	0.98
Ellipse:	0.98
Telescope and WFS:	0.50

Total	0.036

The necessary flux at the WFS is 1800 photons per subaperture and integration. With this, the minimum required power from the fibre can be calculated:

WFS has 15 subapertures diameter →	177 subapertures
1800 ph/subap/exposure →	318.600 ph/exp
1ms exposure time →	318.600.000 ph/sec
Duty cycle 0.2%, repetition rate of laser 10kHz	1.593×10^{11} photons
$h * c/\lambda = 3.736 * 10^{-19}$ →	59.5 nW
T(CalUnit): 3.6% →	1.67 μW

With the current coupling of the LED to the fibre we can get approximately 50 nW out of a 100 μ m fibre through a filter of 0.6 nm FWHM. For calibrating, we can raise the duty cycle of the Pockels cells by a factor of 100 (20 μ s gate hold time instead of 200 ns), to be a factor of three within margin.

Another possibility is to use an optimized LED. A vendor offering a green LED coupled to the desired 105 μ m fibre (NA 0.22) with a typical output power of 43 μ W was found. The spectral width of these LEDs is approximately 35 nm. Considering a typical spectral distribution for an LED, I assume a fraction of about 2.5% of the light is within a band of 0.6 nm FWHM around 532 nm. Together with the transmission of the filter assembly shown in Figure 73, this results in 270 nW out of the fibre, and the duty cycle would have to be raised only by a factor of ten.

3.5.2.5 Software requirements

Software will be needed to

- operate the CalUnit, that is to insert it with the swing arm, to power up, to select the mode of operation (on-/off-axis).
- align the ASM with respect to the CalUnit by using the on-axis spot positions on the science camera and the wavefront sensor output. A model will predict the relative position from the sensor signals and a new position will be set. This step is repeated until satisfactory alignment is achieved. The software should calculate and drive this automatically as an alignment routine.
- switch the light sources on and off. On- and off-axis sources should be able to be operated independently.
- read out the aberrations measured by the FLAO WFS and feed them into the alignment routine

3.5.2.6 Testing, Assembly and Alignment

The manufacturing plan for the mechanics, especially the lens barrel with the slip fits, was adapted from the PDR. All the mechanical parts will be manufactured in house at MPIA, and the tolerances verified with the 3D tactile coordinate measurement machine. This includes the lens barrel and the fibre holder plate. The slip-fit mount in the lens barrel will be tested by verifying the achievable diameter and concentricity tolerances of cylindrical fits with ~50 mm diameter in Invar. The available tactile coordinate measuring machine has sufficient accuracy (< 2 micron). The fibre holder plate accuracy will be tested in a similar way, by measuring the coordinates of three fibre ferrules set in an Invar plate.

The assembly of the lenses will be done by the lens manufacturer. The system is to be delivered without further adjustments necessary. The specifications of the individual lenses must be verified by the manufacturer with a suitable test report. For the ellipse, an interferogram is needed. The wedge and decentre direction has to be marked on the substrate.

Integration of the CGH and the fibres will be done by the CGH vendor. The fully assembled CalUnit will also be tested by the CGH manufacturer, preferably with a dedicated diffractive optical element acting as a null corrector in an interferometer setup. As a diffractive null imposes high costs, other means to verify the correct alignment of the assembly and the residual wavefront error will be investigated.

Test at Arcetri in Solar Tower

It may be very useful to test the calibration unit with the full wavefront sensor in the solar tower at Arcetri observatory, where LBT's second DM is currently being assembled. This will enable us to understand and debug the system before going to the telescope, reducing the cost and commissioning time dramatically. The CalUnit could also be very helpful for testing the DM and FLAO system at Arcetri before it is commissioned on the telescope. It needs to be seen if this is feasible without implications for the delivery schedule of the second DM.

3.5.3 On-axis Source

An on-axis calibration source will be placed in the central bore of the ASM. An aluminum tenon was designed to act as a cable feed and attachment point for the sealing cap, instead of the existing piece of printed circuit board (PCB). A first one has already been manufactured. It was machined to tight tolerances, providing a precise interface for alignment and measurement purposes. The cap is held not with one central screw (as in the previous design), but three M3 screws. The light source, an optical fibre coupled to an incandescent lamp, will be mounted from the inside and visible through a hole in the cap. The whole assembly is sealed, and interchanges with the current design without any change for the sealing lip at the contact point to the shell; this has been verified with Roberto Biasi from Microgate. The fibre could be routed through the ASM, or fed from the side in front of the mirror, with negligible resulting obscuration. Two fibres (one spare) have been routed through the central hole of one of the ASMs but feeding the fibre from the side may be easier for the other ASM, which is already mounted at the telescope.

3.5.3.1 Optical design on axis source

The optical concept for the on-axis truth sensor was adapted as described in Section 3.4.4.1. The layout follows the schematic in Figure 54, and is also shown in Figure 78.

As the confocal ellipses deliver a stigmatic image, the wavefront of the on-axis light source is perfect by design. Accounting for manufacturing tolerances of the dichroic ellipse and alignment errors, we still achieve a wavefront accuracy of better than $\lambda/4$ PV. The wavefront and corresponding spot diagram are shown in Figure 74.

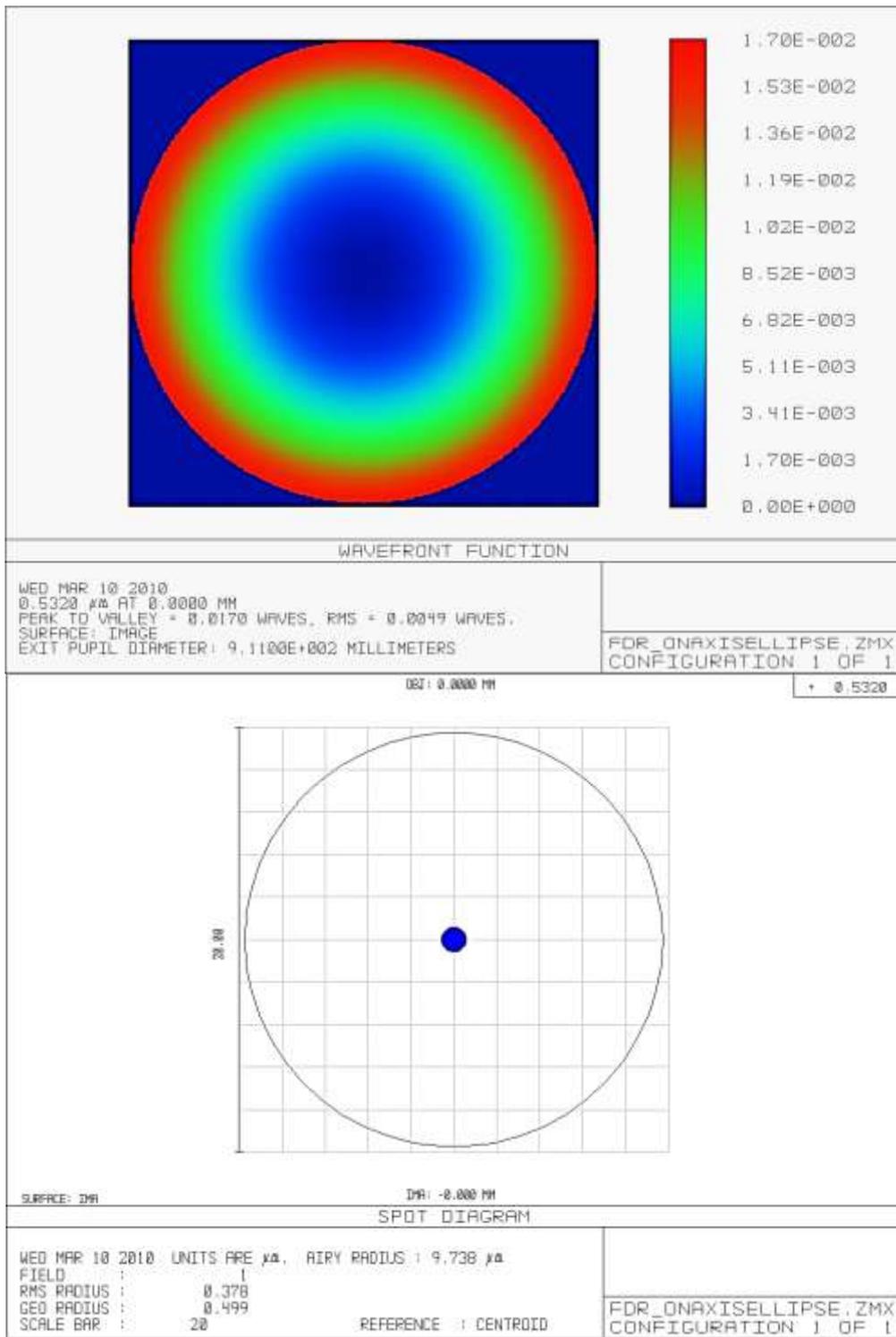


Figure 74: Wavefront and spot of the on-axis source.

3.5.3.2 Mechanical design of the on-axis source

The tenon is a monolithic aluminum piece. The first unit has been made at MPIA and is being tested at ADS, the company responsible for the mechanical assembly of the ASM. It is inserted in the centring hole in the membrane support cover and attached with three screws (the same as used now for the PCB). Using set screws will define the rotation of the tenon.

The achieved tolerances are much higher than the specifications for the optical centering of the shell itself. Six cable feed-throughs are provided at the same position as in the current design. A 3D representation is shown in Figure 75. The dimensions of the cap (Figure 76) have to be verified by Microgate. The manufacturing tolerances are (~ 0.03 mm). The centering tolerance on the fibre plug is rather loose (we can allow for > 0.5 mm decenter); this is important as the optical centre of the shell is not known with high precision.

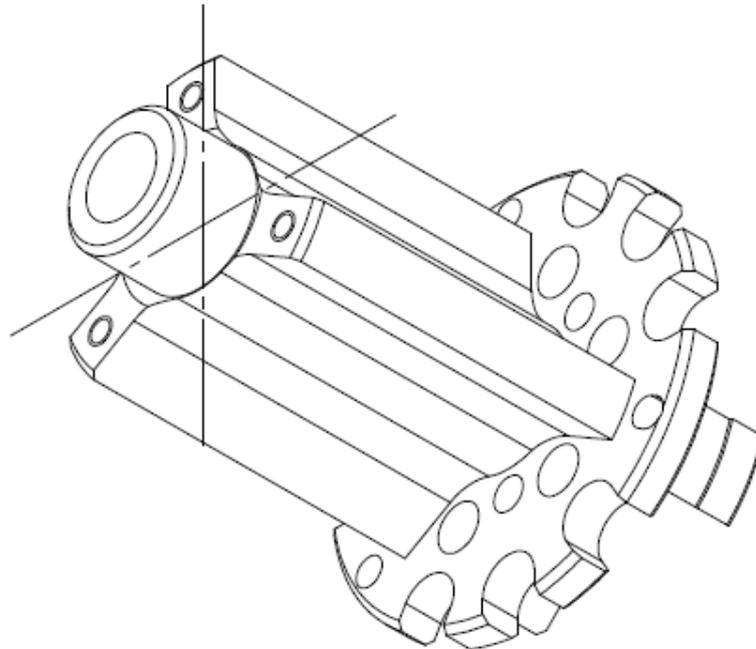


Figure 75: 3D image of the tenon that will be inserted into the central bore of the ASM.

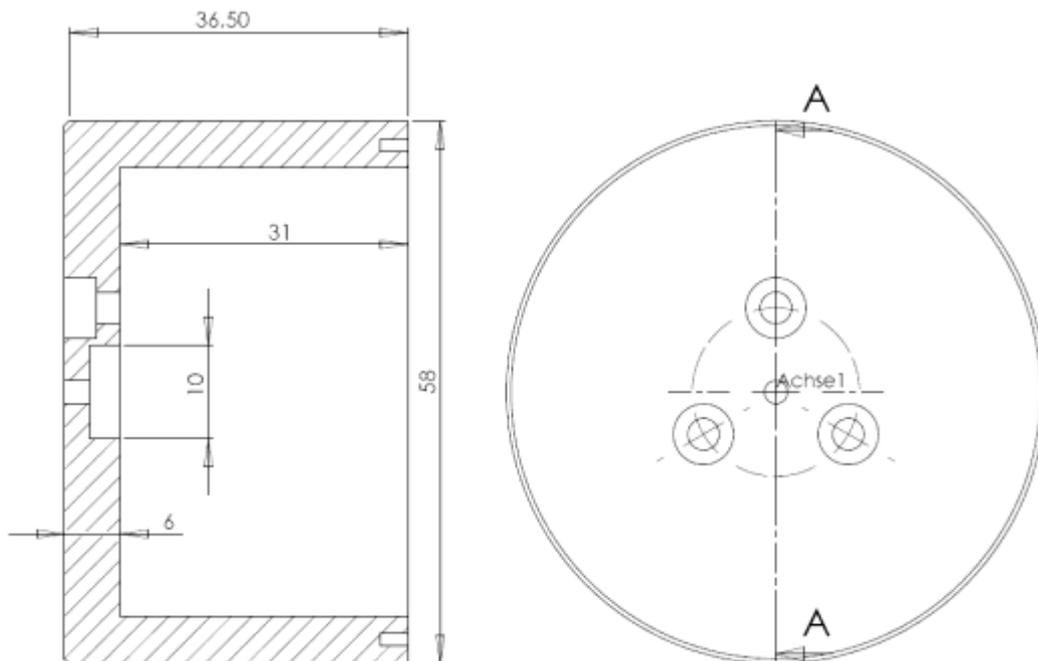


Figure 76: The dust cover cap to be mounted on top of the tenon; outer dimensions are to be decided.

3.5.3.3 Electromechanical design of on-axis source

The optical fibre to be used as the on-axis light source is a Corning SMF 28e IR single-mode fibre, an industry standard. Below its cut-off wavelength it behaves like a multimode fibre with $<10\ \mu\text{m}$ core diameter. On the ASM's side, the fibre will be terminated with a single ceramic ferrule with 2.5 mm diameter. To use only the ferrule and not a full plug saves space. The other side of the fibre has a FC/PC connector. The same type of fibre can be used to deliver the light from the light source to the hub. Fibres with FC/PC connectors are available in various lengths. Highly flexible steel tubing for protection is commonly used. From an optical point of view, lengths of some 10 meters are acceptable, which allows us to place the rack for the light source in a convenient position. I anticipate a power consumption of a few watts for an incandescent lamp.

Fibre Specifications

Type: Corning SMF28e or equivalent
Diameter with buffer: 0.9 mm
Coating diameter: 250 micron
Cladding diameter: 125 micron
Attenuation: $< 0.4\ \text{dB/km}$
NA: 0.13

3.5.3.4 Assembly and testing

The assembly of the cable feed through the tenon will be done by Microgate and the integration of the tenon by ADS. The routing of the fibres from the electronic rack to the ASM needs to be coordinated with LBT. The termination of the fibres depends on the routing scheme. Fibres routed through the ASM have to be connectorised during integration of the tenon. This can be done by a subcontractor of Microgate. The work consists of glueing and polishing a ceramic ferrule to the fibre end. Care has to be taken for the stress relief, as only 0.9 mm outer buffer can be used. Fibres routed in front of the ASM are best obtained with connectors already attached. During the integration of the tenon, the used fibre should be inserted in the plug and the axial distance to the front of the tenon adjusted with a depth gauge to $\pm 0.1\ \text{mm}$. The cap to seal the assembly can be mounted easily. The same procedures as for the existing cap apply.

The tenon's dimensions will be verified and protocolled at the MPIA workshop. ADS will test the fit to the mounting flange at the ASM. Sufficient time to rework the diameter of the flange at the tenon's side needs to be planned for, in case the diameter of the actual flange differs from the manufacturing drawings. The transmission of the connectorised fibres will be measured at Microgate.

3.6 Alignment on the telescope

To reliably achieve the desired performance, not only theoretically, but at the telescope, a robust alignment procedure is necessary. The alignment of the calibration unit in the prime focus has six degrees of freedom. Relevant here are five that have to be realigned every time the swing arm is deployed (the lateral displacement, Δx , Δy , Δz , and tip-tilt $rot\ x$, $rot\ y$). Rotation around the optical axis, $rot\ z$, has to be aligned only during installation to make sure the calibration spots hit the apertures of the wavefront sensor and is neglected hereafter.

First I assess the necessary precision. The wavefront error budget for the alignment is set to 35 nm rms. If the alignment parameters are investigated independently the problem becomes rotationally symmetric. The allowable ranges for the resulting parameters that lead to a wavefront degradation of 35 nm rms are:

- Δr (radial displacement) : 20 micron
- $\Delta\theta$ (tilt) : 3 arcminutes
- Δz (axial displacement) : 5 micron (without focus compensation)

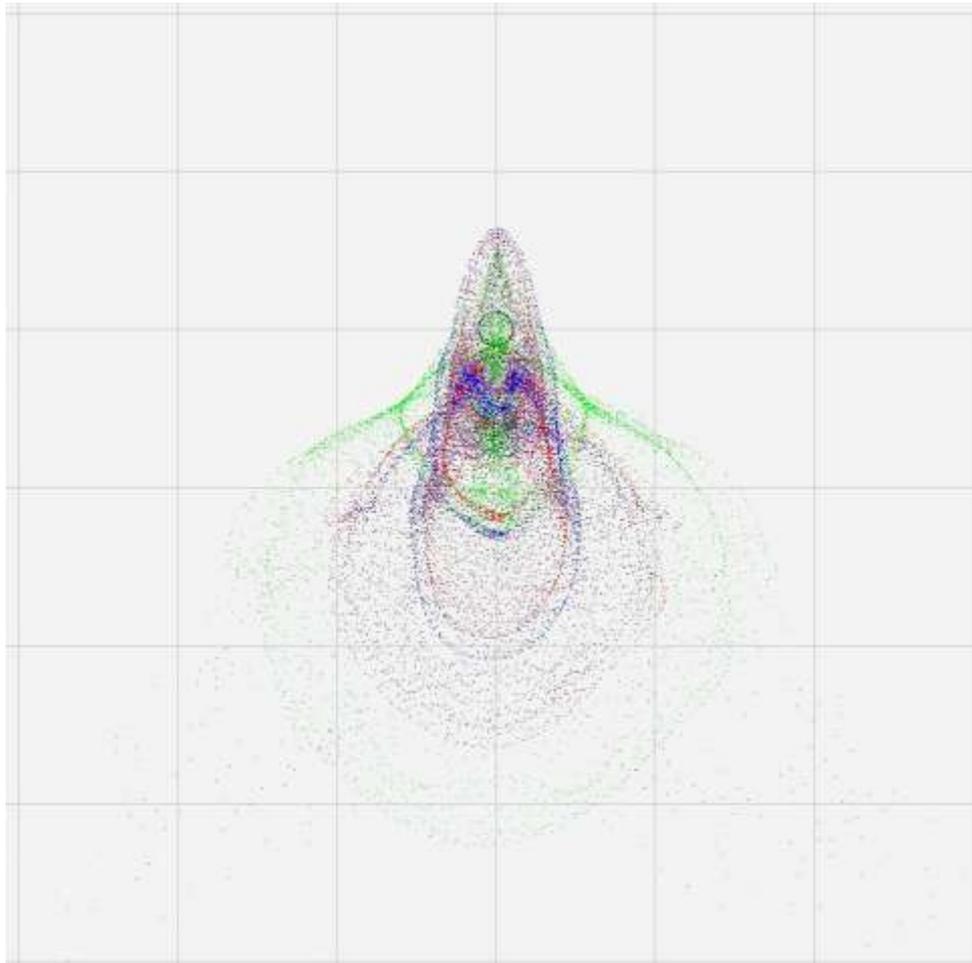


Figure 77: Spots of the LGS calibration sources resulting from 10 micron displacement in z of the CalUnit. Blue the regular spot, red the spot after displacing it, but compensating focus on the sensor, green without compensation.

These tolerances are very tight, especially considering the location of the calibration unit. Key factors for achieving the alignment with this precision are

- the means to change the position with an accuracy better than the values above,
- the capability of measuring the position unambiguously.

The first problem is solved by adjusting the ASM's position and keeping the CalUnit fixed. The ASM hexapod is extremely accurate and its position can be set considerably better than our tolerances. To address the second point, I looked at different strategies for measuring the position of the CalUnit with respect to the secondary mirror. The following methods seemed promising:

- Using position sensitive diodes (PSDs) and a reference surface mounted on the cap in the central bore. This is a good way to measure the tilt by employing an autocollimator setup. The displacement, on the other hand, is much harder to measure this way.
- Using additional alignment features on the CGH, for example fiducial marks projected onto a sensor or recorded with a camera. This is affected by the objective's aberrations and the placement of the necessary sensors must be practical. Its advantage is that it provides a reference directly to the optical system of the CalUnit. The method is very flexible.
- Using the laser tracker system available at the telescope. For this, the retroreflectors (corner cube prisms) used as reference points need to be mounted on the CalUnit and the secondary mirror. The accuracy of the tracker is impressive; however, it is not clear whether this option could be integrated into CalUnit operations so that it can be quickly and easily deployed (even at night, for example) without buying a dedicated tracker unit that would be permanently mounted for that purpose. The laser tracker seems to be an ideal way to carry out the initial alignment during installation.
- Using the on-axis source as the reference. If the ASM's shape can be reliably reset (or flattened) with its capacitive sensors, the aberrations of the on-axis source's image can inform the system about the CalUnit's position error. For the flattening, the lower order aberrations are most important. The aberrations can be measured with the first light AO wavefront sensor or on the science camera with phase diversity techniques. The wavefront of the on axis spot is more sensitive to misalignment, and better defined. Consequently, the aberrations on the on-axis source can be measured to higher accuracy than the residual wavefront error on the LGS calibrations spots. The relevant terms are coma and spherical aberration. The great advantages of this method are that it needs no additional hardware and the necessary measurements are already part of the truth-sensing scheme.

The most sensible solution seems to be a combination of the aforementioned methods, building on their respective strengths. Using PSDs was discarded to avoid the additional design and integration efforts and the risk that one of the substantial numbers of electronic subsystems fails.

For installation and commissioning, provisions will be taken to guide the initial alignment with the laser tracker. Corner cube prisms are necessary on the CalUnit. It is relatively straight forward to install these or provide a mounting interface during this design phase. The case for the ASM, to which it must be referenced, is more complicated.

- Installation of any hardware near the sensitive thin shell needs to be discussed and orchestrated with a number of people.
- Few usable mounting surfaces for the corner cube prisms exist. The most practical may be the top of the sealing cap in the centre of the ASM. The manufacture of this part is carried out by the ARGOS team.
- There is no mechanical reference that has the necessary tight tolerances with respect to the shell. The centring error, which defines the radial displacement, is given as “of the order of 0.5 mm”. The tilt of the shell with respect to the mounting flange in the central bore that acts as a mechanical interface for the tenon (part of the on-axis light source) as well as the tolerance in z , should be somewhat better.

Inferring the position error of the CalUnit from the on-axis spot aberrations and then aligning the ASM in a closed loop is a fast and accurate method. This could be well integrated in the operational environment. All the necessary subsystems need to be running for the calibration procedure.

Tilt and decentre of the elliptical front surface of the CalUnit with respect to the secondary mirror, which is also elliptical, leads to coma in the on-axis light source’s image. The first light AO wavefront sensor is able to measure this better than $\lambda/10$ peak-to-valley. An axial displacement (Δz), on the other hand, can be seen as spherical aberration and can be detected at the same level. There remains degeneracy between tilt and radial displacement: the coma coming from tilt (around the surface vertex) can be compensated by displacing the surface.

Stated differently, the coma is very insensitive to rotation of the CalUnit around the vertex of the ASM (it is not invariant to rotation around the prime focus, though, which is an important difference to a spherical retroreflector). Rotating the CalUnit around the ASM’s vertex results in an offset of the intermediate image from the optical axis of the CalUnit. This fact is utilised to break the degeneracy. A dedicated alignment structure on the CGH generates an additional spot at the correct location of the prime focus with respect to the CalUnit (~32 mm in front of the surface on the optical axis).

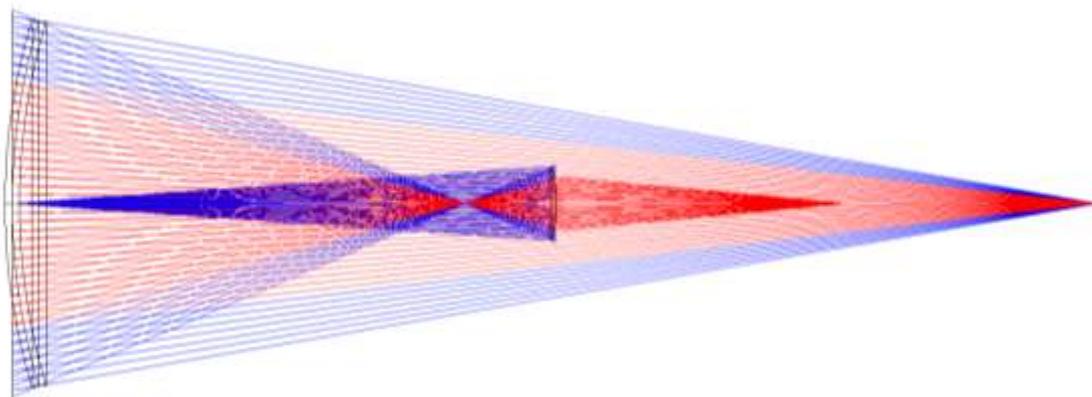


Figure 78: Schematic ray diagram of the on axis source (blue) and the transmissive alignment source (red). On the far left the ASM, in the middle the confocal elliptical reflector, on the far right the secondary focal plane. The reflected on axis spot as well as the alignment spot overlap in the prime focus and consequently in the secondary focus.

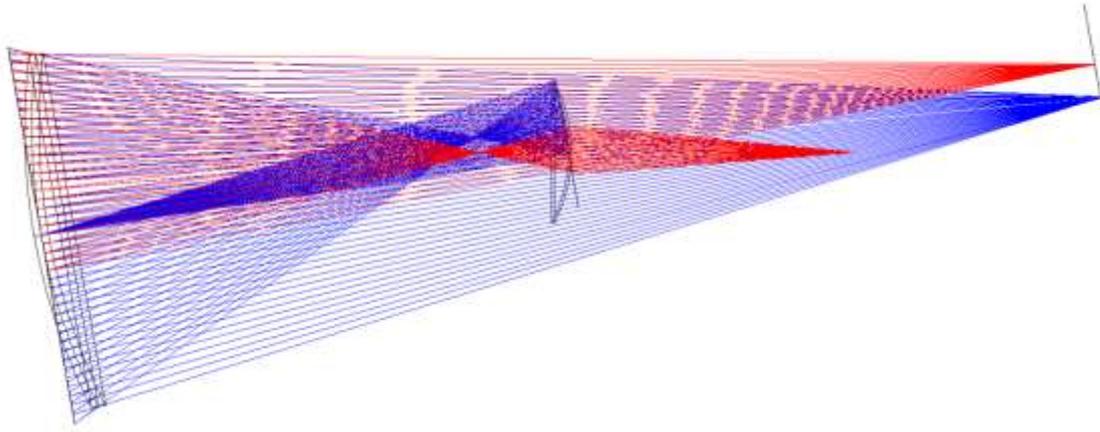


Figure 79: The same rays as in Figure 78, but if the ASM is shifted and tilted with respect to the CalUnit. Coma is compensated in the blue beam, but the spots don't overlap anymore.

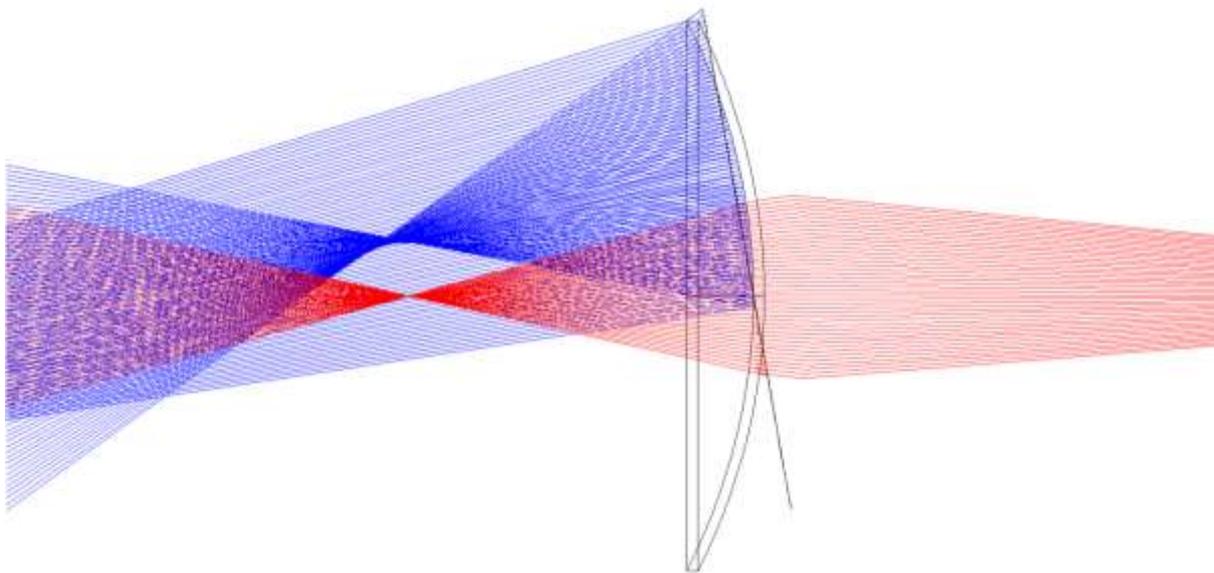


Figure 80: Zoom in on the prime focus in Figure 79. The focus spot of the on-axis source shows aberrations (coma), which is then corrected by the ASM for this field position. The alignment spot (red) is fixed to the CalUnit and still in the nominal position on the optical axis.

Due to space constraints on the CGH, the spot will show diffraction artefacts and will not illuminate the whole secondary mirror. This makes it unsuitable for use as the truth sensor but it can act as a reference mark for the on-axis light source spot. Both spots are imaged by the ASM onto the science camera, where they must overlap if the alignment is correct. The distance between the spots on the science detector puts an additional constraint on the position of the CalUnit, which leads to a unique solution. The centre of gravity of each of the spots can be determined to better than 1 pixel on the science detector. This is precise enough to derive the position information with the desired accuracy. The displacement of the spots is a very sensitive instrument to measure misalignments. A tilt of 30 arcseconds, which is fully compensated by shift and focus to yield a perfectly unaberrated spot of the on-axis source,

leads to a distance between the spots of 60 micron. Figure 81 shows the location of the spots in the secondary focal plane, in comparison to the Airy disc diameter. The displacement is independent on the fact if the tilt has been compensated by shift or not.

OBJ: 0.0000, 0.0000 DEG

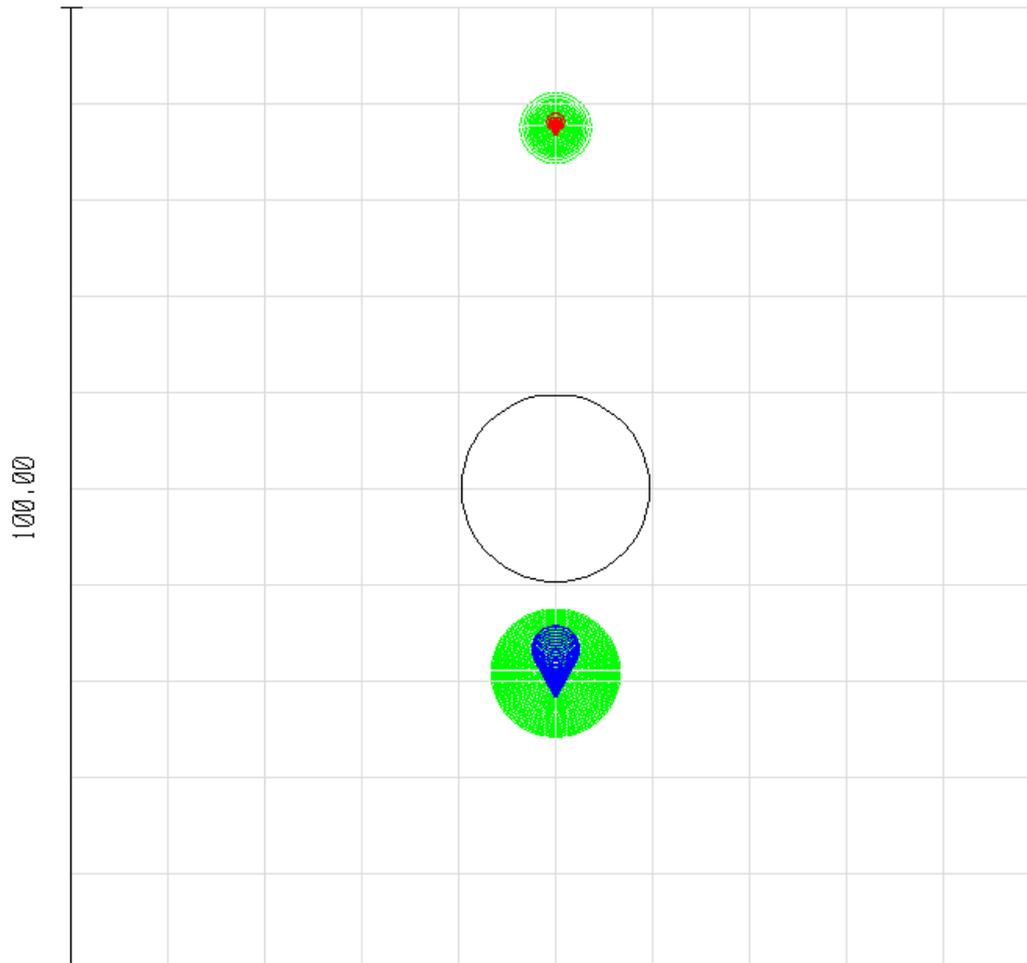


Figure 81: The spots of the real on-axis source (blue) and the alignment feature (red) in the secondary focal plane, for an uncompensated tilt of the CalUnit of 30 arcseconds. In green the spot locations if the tilt has been compensated by shifting M2. 100 micron defocus has been added to make the spots visible in this plot.

3.7 Conclusions and outlook

I designed a wavefront sensor calibration unit for the Rayleigh laser guide star AO system on LBT (ARGOS). The unit was required to feign the 3 laser beacons, producing the same aberrations that would be seen on-sky, in order to allow the complete system to be calibrated at any time. The beams are each strongly aberrated because of their positions (2 arcmin off-axis at a height of 12 km) so the main challenge was to imitate these characteristics with very high wavefront precision. A truth sensor to check the performance of the science camera had to be provided. In addition, the calibration unit was required to be quickly and easily deployable during the day as well as during observations, and work over a wide temperature range.

I thoroughly investigated modern optical concepts and manufacturing techniques, and identified the optimal solution. The final design makes novel use of binary optics to generate a highly aberrated wavefront with very high accuracy, in combination with high precision lens elements providing the necessary optical power to match the fast $f/1$ focal ratio of the telescope in prime focus. The solution uses the telescope optics in single pass to match the real lightpath and footprints of the laser beacons as closely as possible. A dichroic aspherical surface is cleverly combined into the lens system and doubles as an alignment aid and truth sensor. This is helpful as it allows self-consistent calibration and minimizes additional optoelectric elements (like position sensors). It requires a dedicated light source to be located in the centre of the ASM, on the optical axis. I designed the mechanical components to serve as a stable, high precision interface to the ASM hub.

To research the necessary tolerances for the optical and mechanical elements as well as the alignment, I did a thorough tolerance analysis of the whole system. The practice of modelling the full system to be calibrated, including the tolerances on the calibration source and deriving from this the necessary tolerances on single elements, is all too often neglected in astronomical instrumentation. This work demonstrates how the tools available for parameter sensitivity analysis, combined with the knowledge of what is achievable with state of the art manufacturing and measuring techniques, can be used to give confidence in the design of an instrument and predict its performance. The optical design has been finalized and is ready for production; vendors for the subcomponents have been identified, and a testing and assembly plan formulated. The conceived calibration system is unique in its capabilities, technical realisation and performance assessment.

This concept is an innovative approach to calibrating WFSs. It is very flexible and can easily be adapted to future instruments. The next generation of extremely large telescopes (ELTs) will have large, curved deformable mirrors as an integral part of the telescope, for which my calibration concept is a highly attractive possibility. As an example, I investigated the transfer of this idea to the European ELT (E-ELT) (Schwab et al. 2010). The calibration unit also has many applications on the LBT beyond the calibration of the laser guide star wavefront sensors alone, for example the illumination and calibration of the interferometric instruments with the on-axis source. The latter is already being investigated.

The calibration concept, with its flexibility to exactly produce nearly arbitrary wavefronts, and the combination of optical elements with CGH structures to facilitate alignment, has wider applications. It is potentially useful for aligning and measuring optical systems that would normally employ a CGH as null lens in an interferometric test setup.

The calibration source I designed for the LBT will be an integral part of the ARGOS facility, allowing reliable calibration and extensive testing of the whole system, and so enabling it to reach the expected performance. Furthermore, the capability of producing a well-aligned, diffraction limited on-axis source will most likely become an indispensable tool for setting up and commissioning new instruments. The ARGOS facility will be one of the first AO systems with multiple laser guide stars, and due to the use of Rayleigh technology, has the highest aberrations on the beacons. The experiences with the system, including the vital aspect of calibration, will serve as a demonstrator for next generation AO systems.

Part II: A planetary companion around HIP31592

4 Historical introduction

The question of whether planets around other stars exist has concerned humans for at least 2 millennia. The first record of such thoughts is from Ancient Greece, when Epicure (342-270 BC) speculated about other worlds, and living beings on them, in his letter to Herodot. In fact, the word planet originates from the Greek for “wanderer”. In the modern astronomical context, these ideas became meaningful when Nikolaus Kopernikus announced the view that the earth is a planet orbiting the sun and that the stars themselves are suns, similar to our own (*De revolutionibus orbium coelestium*, 1543). Giordano Bruno, a Domenican friar, was the first to formulate the idea that many more planets might be harboured by the uncountable stars we see, and that life might evolve on them. He was deemed a heretic and burnt at the stake in 1600. The existence of other planets invariably raises the question of whether life in other forms may exist on them. This may be the reason for the great fascination that the search for extrasolar planets, as they are now called, holds for a broad public audience. Finding planets is a major step towards the possibility of probing the Universe for signatures of life.

The first claims that an extrasolar planet had been observed were made in the late 19th century (e.g., Jacob 1855, See 1890). All of these were found to be false alarms. Indeed, the precision necessary to detect a planetary body around a star only became available technologically more than 100 years later. Nevertheless, discussion on the detection of extrasolar planets, numbers of planets and possible formation mechanisms began in the 1950s and became the foundation of today’s theoretical framework for exoplanet science (e.g. Struve 1952).

During this time, the question of what characteristics a planet must possess for life to evolve, and the likelihood of this happening was revisited, and investigated against the background of new knowledge on the makeup of our galaxy and the galaxies in the Universe beyond. Three names particularly stand out in this discussion: Enrico Fermi, Carl Sagan, and Frank Drake. In 1950 Fermi asked some of his colleagues over lunch "Where is everybody?", referring to intelligent extraterrestrial life. The reformulated form of this question became known as the Fermi Paradox: if a planet like our Earth is common, or mediocre, the chances that life will evolve on a planet around one of the billions of stars in a galaxy are high. What is the reason then that we do not observe this? This way to put it draws a much too simple picture of the situation, considering the vast distances and timescales involved, and extreme complexities involved in developing life, let alone intelligent life. But Sagan and Drake built on this idea and suggested the active and scientific search for extraterrestrial intelligence, SETI for short. The first SETI program was conducted by Frank Drake in 1960. Drake formulated his famous equation in 1961, estimating the likelihood of what he believed to be the essential ingredients for the formation of intelligent life, and as such assessing the probability that there is a planet within reach that harbours intelligent life and transmits signals into the sky. He concluded that the probability for another life-bearing planet in our galaxy is on the order of unity, and looked for a way to search for signals originating from a distant civilisation. Several observational programs have been conducted since then, making use of a number of radio telescopes, amongst them the big radio dish at Arecibo, Puerto Rico, to filter the measurements of the telescope for a narrowband signal of high spectral power density. From May 1999 on, after NASA retreated from the project, the computational analysis was carried out on private PCs volunteered by interested individuals under the name “SETI@home” and the program became very popular amongst interested people outside scientific astronomy.

Drake's equation has been challenged as meaningless, as some of the terms are factually unknown. A relatively new development on this question is the so called "Rare Earth hypothesis", that states that conditions for the development of life on a planet are exceedingly rare, and Earth is probably the only planet in our galaxy that bears life (Ward, Brownlee 2000). For example, the authors argue that there is a galactic habitable zone (Lineweaver, Fenner & Gibson 2004) in which planets can develop life, and exclude the galactic bulge with the highest density of stars from this zone, substantially lowering the numbers of stars that could possibly harbour life-bearing planets.

The concept of a habitable zone first emerged in the context of planetary systems. This zone, sometimes called the Goldilocks zone⁸, normally means the distance range from a host star at which water on a planet's surface could be in its liquid phase due to the surface temperature. While this is dependent on the existence and makeup of an atmosphere, it can still be determined with reasonable accuracy based simply on the temperature of the star. These thoughts can be traced back at least 150 years (Whewell 1853), to a time when Pluto had not yet been found and the discussion about life on Mars that started with Schiaparelli's observations had not yet begun (see Part I), but have gained immensely in importance since the discovery of exoplanets. The concrete question then arose: could life have formed on these planets?

Ever since the detection of the first exoplanets, the holy grail for planet hunters has been finding an earth-like, that is, rocky planet in the habitable zone of a star. The public resonance of such a discovery would be enormous. Scientifically, it seems at least equally as important to extend our sample of known planets in all directions - to lower mass planets, different types of stars, and different orbital radii (and hence periods).

The first exoplanet was discovered in 1988 (Campbell & Walker), but the detection had to be withdrawn due to the limited evidence. Campbell and Walker used the radial velocity technique to infer the presence of a planet around γ Cephei. The detection was later verified by Hatzes (2003). This was followed by the detection of a planet around a pulsar using very precise pulsar timing (Wolszczan & Frail 1992). The field really took off, however, when the first planet around a sun-like star on the main sequence was announced by Mayor & Queloz (1995), using high resolution spectra to measure the Doppler motion of the host star. Other groups were at very similar stages in their search programs, and could confirm the detection within weeks (Marcy & Butler 1995). Since then, the number of known exoplanets steadily rose to more than 450 today (April 2010), detected with various methods, including Doppler searches, microlensing, transits and direct imaging (Figure 82). I give a short overview of these techniques in the next section. In the plot, two other methods are mentioned: pulsar timing, which looks for planets around pulsars, and hence is somewhat different in its scope and implications, and timing of double stars, for which only 5 discoveries are claimed. It has not yet been established whether these are unambiguous planet detections.

Currently the field of planet searches is changing from looking for the most likely targets of opportunity to broader searches, and we finally have a sample big enough to infer statistics about the properties of planetary systems. This gives important feedback for the development of theories of planet formation.

⁸ Named after a tale in which a girl, Goldilocks, has to choose between three pots of porridge and chooses the middle one, which is neither too hot nor too cold, but 'just right'.

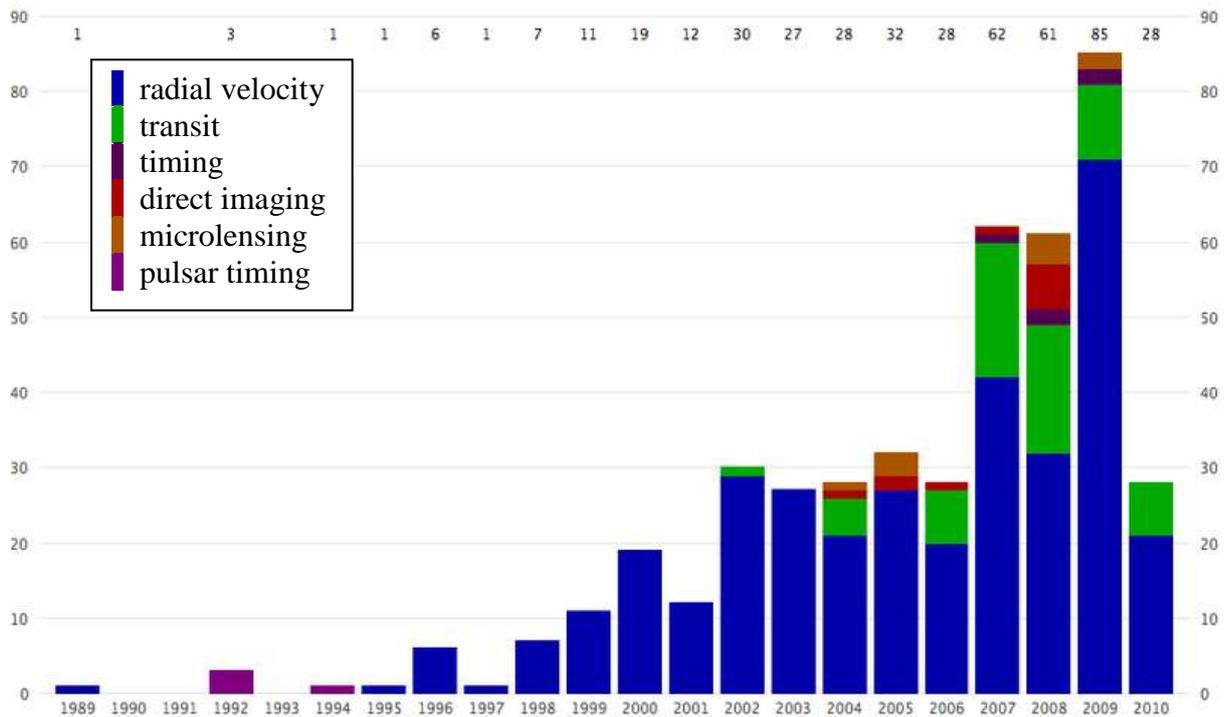


Figure 82: Bar chart of exoplanet discoveries by year, through 2010-03-19, indicating the discovery method in different colours (see legend). Base details from the following sources, and references therein: Extrasolar Planets Encyclopaedia (2010-03-19), Planetary Society: Catalog of Exoplanets (2010-01-07), Summary Table of Kepler Discoveries, NASA (2010-03-15). Note that some of the discoveries are controversial. However, this plot shows the general trend for the numbers of detected exoplanets, as well as the ratio between the detection methods. *Source of plot: Wikimedia Commons.*

At the moment, the vast majority of planets have been detected with the Doppler method. The most successful, largest and oldest surveys are the efforts of the California planet search team at Lick and Keck observatory, and the program of the Geneva team, now conducted with the HARPS spectrograph at La Silla. These programs target mainly G-type stars similar to our sun. Many planets have been found in short orbits, from the order of a few days, up to a few years. Most of them are very massive, with a typical mass of a few Jupiter masses. Together with the close orbits this gave them the name “hot Jupiters” (see Figure 83).

Some newer programs target other types of stars, like M dwarfs or later types. M dwarfs have the advantage that smaller planets can be found with the same instrumental sensitivity, as the host star is lighter in mass. The drawback is the faintness of these stars in the visual spectral band. K-giants are also studied, for example by a program conducted at the Hobby Eberly Telescope (HET), a program at the Karl Schwarzschild Observatorium (Tautenburg), as well as our own survey.

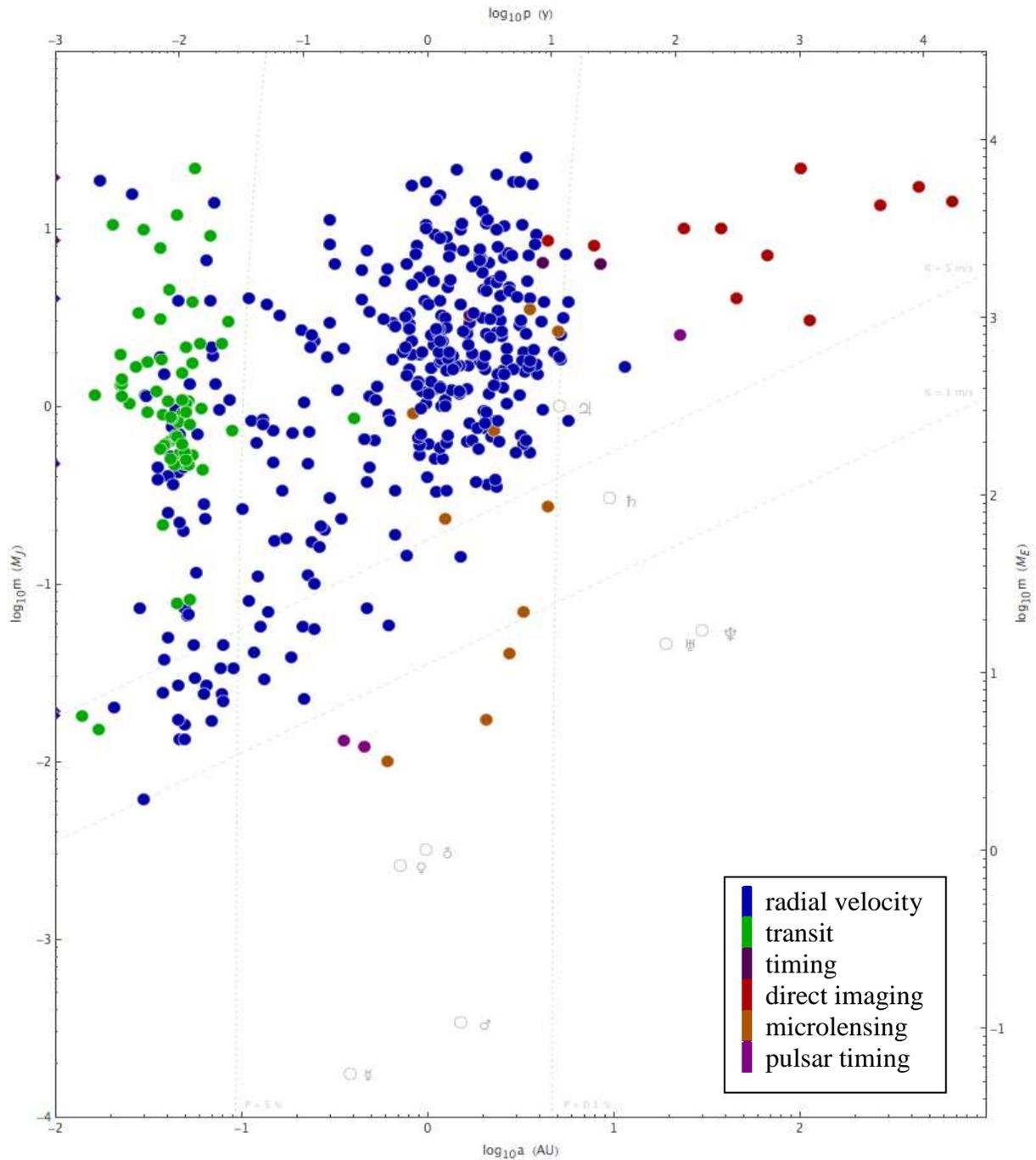


Figure 83: Planetary mass against distance from the star or period of the orbit for the exoplanets discovered up until 19th of March 2010.

Base details from the following sources, and references therein: Extrasolar Planets Encyclopaedia (2010-03-19), Planetary Society: Catalog of Exoplanets (2010-01-07), Summary Table of Kepler Discoveries, NASA (2010-03-15). Note that some of the discoveries are controversial. Source of plot: Wikimedia Commons.

Our planet hunting program looks at a sample of 376 giant stars with spectral types between early G and late K using the CAT (Coudé Auxiliary Telescope, 24 inch f/36) at Lick observatory. The program is conducted with the Hamilton spectrograph (Vogt 1987). The target stars are all brighter than 6th magnitude to account for the relatively small aperture of the telescope used. They are observed, depending on their visibility, during monthly observation runs. The sample was built up in three steps: the oldest members of the list have

been observed for more than 10 years now, since 1999, with additions of stars in 2000 and 2004. This puts our search program amongst the oldest in the field, and gives us outstanding opportunities with the excellent phase coverage of even very long orbital periods. We can probe for planets with orbits of up to two decades, a region largely unexplored as yet.

5 Detection methods

The most successful of the detection methods used do not directly measure the planet, but the induced reflex motion of the host star due to the gravitational pull of the second body; the corresponding methods are astrometry and the Doppler technique. Two more methods, transit detection and microlensing, are also indirect, using photometry to measure changes in the host or a foreground star's apparent brightness due to the effects of the smaller body. Only direct imaging and secondary transits measure a signal coming directly from the planet itself. I give a short overview over the different methods and then explain the technique we use in greater detail.

5.1 Orbital elements

An orbit is fully defined by six independent parameters. There are slightly different ways to specify these six parameters. A common set consists of the semi-major axis, eccentricity, inclination angle, longitude of periastron, longitude of the ascending node, and time of periastron (mean anomaly). The inclination angle is the angle between the tangent plane and the orbital plane of the planet. The longitude of the ascending node is the angle between the direction north and the nodal line, in the direction of the ascending node. The longitude of periastron is the angle between the nodal line and the periastron, and the time of periastron (or the mean anomaly) is the point in time where the planet is at its closest approach to the main body.

The various detection methods measure different subsets of these orbital parameters. The Doppler method, for example, can determine four of these - semi-major axis, eccentricity, longitude of periastron, and time of periastron. The inclination and the longitude of the ascending node cannot be recovered from radial velocity measurements alone.

One physically very important parameter is the mass of the planet. However, the mass cannot be measured directly, but only the so called mass function of primary and secondary components' masses (see Section 5.6). Only if one can deduce the mass of the star, for example by determining its spectral type and using stellar evolution modelling, one can determine the planet's mass once the orbit of the host star has been measured. The first time the mass of an exoplanet could be accurately determined was in 2000, for the transiting exoplanet around HD 209458 (Henry et al. 2000).

Apart from the mass, the other physically important parameter is the semi-major axis of the orbit. If the star's mass has been derived, the semi-major axis follows directly from Kepler's third law. Figure 84 shows the masses and semi-major axis of the planets found by 2004, together with the sensitivity limits of different techniques and missions. The distribution of planets in this fundamental plane comes largely from selection effects, and not from the physical properties of planetary systems. It is interesting to note that the predicted sensitivity for radial velocity searches by 2010 has been achieved and even surpassed (Pepe & Lovis 2008). By comparison to Figure 83, which shows the same plane with all the detections up to 2010, we can see the expansion of the region in which we can detect planets.

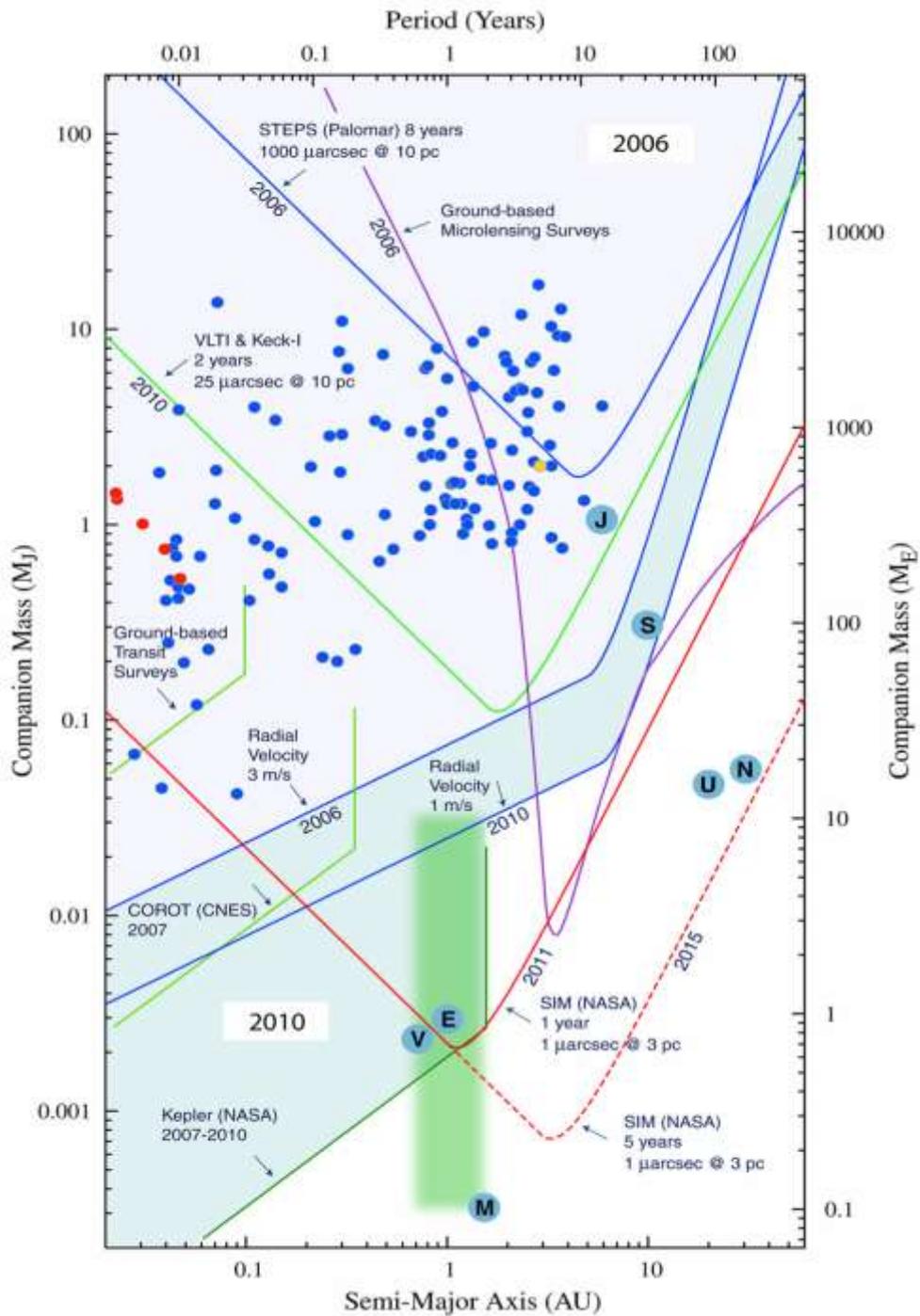


Figure 84: Limiting sensitivity of planet finding techniques, including radial velocity surveys, astrometric surveys, microlensing surveys, and ground and space-based transit techniques. The lines show 5- σ limits. The shaded areas show the expected progress towards the detection of Earth-like planets by 2006 and 2010. Planets in our Solar System are indicated by the blue circles (labelled with the first letter of each planet). The filled circles indicate the planets found by RV surveys (blue), transit surveys (red), and microlensing surveys (yellow). The extrasolar planets shown in this plot represent the reported findings up until 31 August 2004. *Courtesy NASA/JPL-Caltech*

5.2 Astrometry

For an astrometric detection, the motion of the host star due to the gravitational pull by the planet is directly measured as a change in position with respect to a background grid of the ‘fixed stars’, in the classical sense. If planet and star form a two-body system, both must revolve around their common centre of mass, and each describes an ellipse projected onto the background. Obviously the period of these movements is the same. The semi-major axes of the ellipses are related inversely to the ratio of the masses of the two bodies. The actual shape of the ellipse is a function of the projection to our line of sight.

The great complication for astrometry is that the position of the star needs to be determined with extreme precision and this over a substantial period of time. For a system like the Sun and Jupiter, with the same distances and masses, 100 parsec away from us, the diameter of the star’s orbit has an angular radius of only 0.05 mas. The signal descends linearly with the distance to the star. Consequently, it is very hard to find small planets and planets around far away stars. The precision of astrometric facilities has only very recently become good enough to reliably detect such feeble signals. So far, no exoplanets have been detected by using astrometry. The first astrometric follow-up observation looked at the planet around Gliese 876 using HST (Benedict et al. 2002). Space based observatories specially developed for highest precision astrometry, like the Hipparcos satellite; deliver a resolution of less than 1 milliarcsecond (ESA, Perryman 1997). Ground based, dedicated interferometers are set up to reach even higher resolution. Prima at the VLT is designed to reach 10 microarcseconds (AQ1998), for example.

5.3 Microlensing

With the microlensing technique, researchers look for the amplification of the light of a background star due to the gravitational lensing effect of a compact body in the foreground. Paczynski (1986) was the first to note that this technique could be used to look for compact objects making up the dark matter in the halo of the Milky Way. The technique can be exploited to find exoplanets if the foreground object is a star-planet system. In order to detect a planetary system, one has to monitor a large number of potential background stars, as the chances of a single event occurring are very slim. This is due to the fact that the geometry - the host star and the planet passing the line of sight between the observer and the background star - is very critical for a signal to be produced. Because of the low probability of observing microlensing events, most surveys monitor dense star fields, for example towards the Magellanic clouds or Galactic bulge. For a typical event, with the source star located in the bulge of the Milky Way (distance 8 kpc) and a lensing star of 1 solar mass at half that distance, the Einstein ring radius is 1 milliarcsecond (see Figure 85 for the description of an Einstein ring). The radius of the ring in radians is calculated as

$$\theta_E = \sqrt{\frac{4GM}{c^2} \frac{D_{LS}}{D_L D_S}}$$

where G is the gravitational constant, M is the mass of the lens, c is the speed of light, D_{LS} is the angular diameter distance between the lens and the source, D_L is the angular diameter distance to the lens, and D_S is the angular diameter distance to the source. This provides a natural scale for the problem (for a thorough mathematical treatise see Gould 2000).

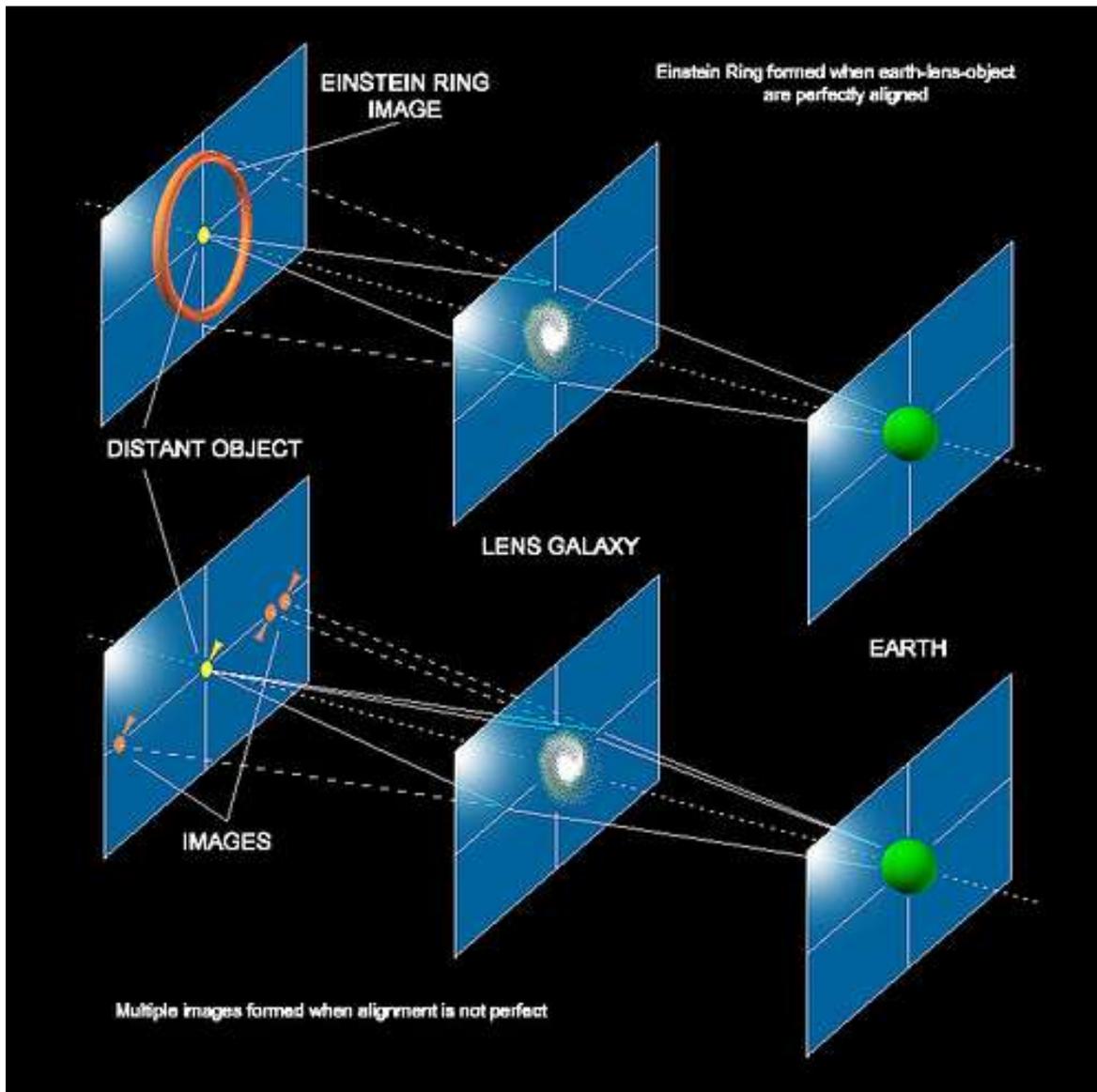


Figure 85: A massive object between the observer and a light source can bend the light from the distant light source in such a way that it reaches the observer from a given angular distance to the line of sight. If the geometry is rotationally symmetric, that is, the mass distribution is rotationally symmetric and the observer, massive object and light source form a line, the observer sees a ring like image of the source around the foreground object (top panel). This phenomenon is called Einstein ring. Image credit: NASA/JPL.

The most important microlensing surveys are MACHO, OGLE, EROS and DUO, together with the follow-up networks PLANET, MOA and μ FUN. So far 10 exoplanets have been detected in this way.

The advantages of the microlensing technique are that it is sensitive up to great distances between earth and the planetary system, and that it is sensitive to low planetary masses (but in a relatively narrow range of semi-major axes). A substantial field of view is required in order to observe many stars simultaneously. Consequently, small sized telescopes equipped with standard scientific grade CCDs are employed. The microlensing method complements

other techniques by sampling another part of the parameter space and increasing the statistics on exoplanets.

The main disadvantage, on the other hand, is that the measurement cannot be repeated. While many measurements can be taken of one single event, the event will not recur. The great distances involved make it difficult to investigate and characterise the host star.⁹ The background source and host star are not separable at the time of the lensing event, but the lens star may be observable after they have moved far enough apart. The first host star which could be spectroscopically characterised was OGLE-2003-BLG-235L/MOA-2003-BLG-53L. Two years after the lensing event (Bond et al. 2004), HST observations revealed the lensing star to be a K dwarf of about 0.63 solar masses at a distance of 5.8 kpc (Bennett et al. 2006). An example of a microlensing light curve from Beaulieu et al. 2006 is shown in Figure 86. This was the lowest mass exoplanet that had been detected at that time (5.5 Earth masses).

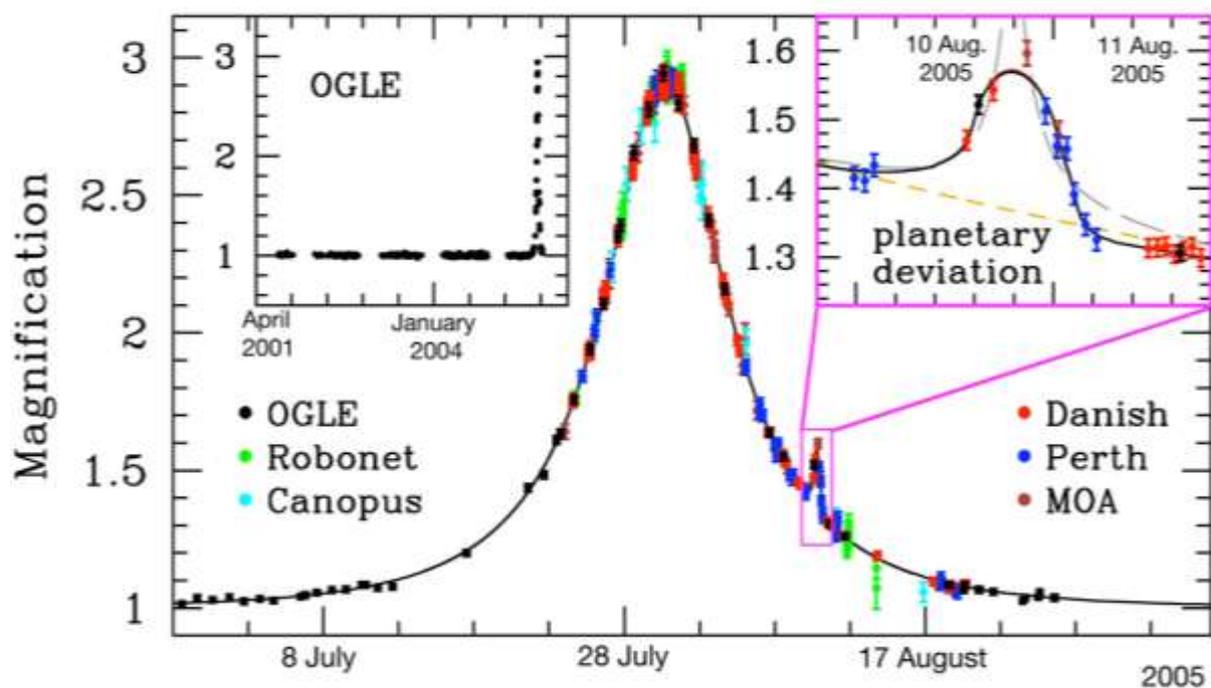


Figure 86: A microlensing light curve (OGLE-2005-BLG-390) from Beaulieu et al. (2006) showing a planet detection (enlarged in inset). The data was obtained using a number of different telescopes, colour-coded as shown in the legend: Danish 1.54m at ESO La Silla (Chile), Perth 0.6m (Bickley, Western Australia), Canopus 1.0m (Hobart, Tasmania, Australia), Faulkes North 2.0m (part of RoboNet, Haleakela, Hawaii, USA), OGLE 1.3m (Las Campanas, Chile), and MOA 1.8m at Mt John Observatory (New Zealand). The best-fitting lensing models for a planet (solid curve), a single lensing star with a binary background object (long-dashed) and a single lens star and single background object (short-dashed) are also shown.

⁹ D. Fischer is quoted as having said „So you look for planets you can’t see around stars you can’t see?” about a decade ago, in reference to the microlensing technique.

5.4 Transits

Like microlensing, the transit technique uses photometry to probe for planets, but in this case, the light from the host star is directly measured. As with microlensing, the observational strategy for a transit search is to monitor as many stars as possible at the same time, as the chances of a transit to occur for any particular star at a given point in time are very small, and the planet must be in the line of sight. If the system is seen edge on, the planet can pass in front of the star, and eclipse part of the stellar surface. This induces a tiny drop in apparent magnitude that may be measurable. The depth of the dip in the lightcurve depends on the ratio of the diameters of the star and the planet. Planets with large diameters are easier to detect. To detect planets transiting in front of giant stars is particularly challenging; while the chance of seeing a transit grows because the star's diameter increases dramatically at this stage of stellar evolution, the decrease in magnitude during the eclipse gets smaller. When multiple transits around a single star are observed, the period can be determined. If the diameter of the star can be derived, for example by stellar evolution modelling, the planet's diameter can be directly measured. Together with its mass, derived from the measured orbital parameters (period and mass ratio), its mean density can be calculated. Most planets found to date are gas giants and have low densities, a result from the selection effect due to the sensitivities of the used detection methods. Only two transiting exoplanets that could be detected have densities close to the density of the inner four bodies of our own solar system. (COROT-7b Queloz et al. 2009; GJ 1214b, Charbonneau et al. 2009).

The transiting method also allowed the spectrum of a planet's atmosphere to be measured for the first time (Richardson et al. 2007). This is possible as the planet becomes invisible during the secondary eclipse (when the star eclipses the planet), and therefore the star's spectrum can be subtracted from the combined spectrum to reveal the planet's contribution. A recent result that received a lot of attention is the measurement of the spectrum of HD189733b (Swain et al. 2010). The spectral emission features measured around 3.25 microns cannot be explained with a simple atmosphere model assuming local thermal equilibrium (see Figure 87). The optimistic interpretation would be to assume this indicates more complex non-equilibrium chemistry processes in the atmosphere. However, these measurements are extremely sensitive and complex, and have not been confirmed independently. It remains to be seen if additional investigations can put further constraints on the model.

5.5 Direct imaging

The attractiveness of direct imaging lies in the fact that it records light from the planet itself. In a sense this makes it the least ambiguous form of planet detection – the planet shows up as a tiny dot of light next to its host star. However, the technical complications are enormous. The challenge lies on one hand in that the small angular distances have to be resolved. To use the example of the Sun-Jupiter system again: at a distance of 100 pc, the angular radius of Jupiter's orbit (5AU) is only 50 mas – approximately the Airy disc diameter of the biggest ground based telescopes- hardly resolvable, even if they were not hampered by the atmosphere. The second, more important problem is the extreme contrast between planet and star. The diffraction effects on the telescopes apertures, as well as unavoidable stray light, produce glare that is many times brighter than the planet even several Airy disc diameters away from the centre of the star's PSF. One technical solution is to concentrate the light of

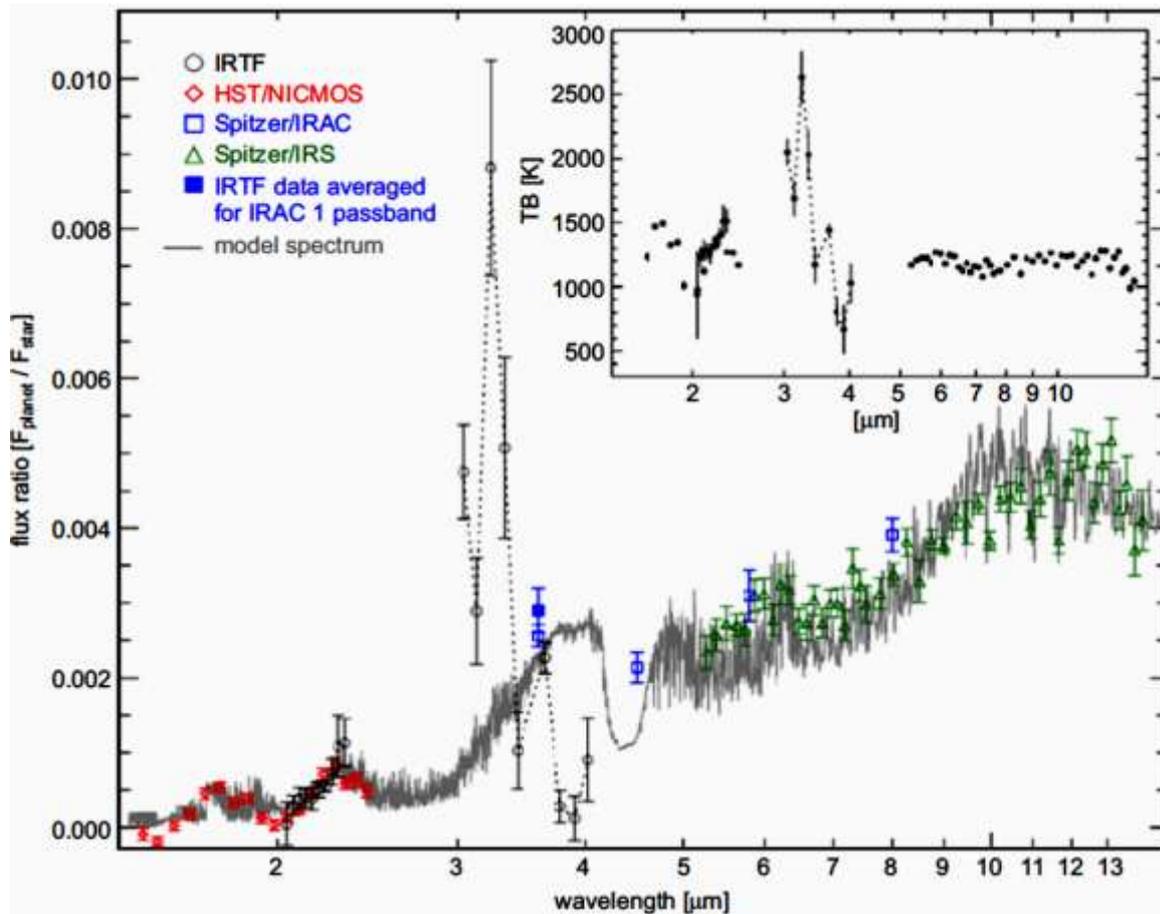


Figure 87: This plot from Swain et al 2010 shows a dayside emission spectrum of the planet around HD 189733b. It is compared to a radiative transfer model (grey) assuming LTE conditions, which cannot explain the strong emission at 3.25 micron. The inset plot displays the brightness temperature at each wavelength and shows the large temperature change needed to produce the 3.25 μm emission if LTE conditions hold.

the host star as well as possible and suppress the core of the PSF to reveal a dim companion below the scatter of the host star. This is the focus of efforts in the fields of the so called extreme adaptive optics (ExAO), and coronagraphy. In the planned extreme adaptive optics systems, one or more deformable mirrors, of which at least one has a very high density of actuators, aims to correct the atmospherically distorted wavefront to an unprecedented accuracy (a few nanometres RMS), and thereby removing the seeing limited halo around the diffraction limited core that becomes already visible with much larger correction errors. This diffraction limited core, that concentrates most of the starlight (>98%), is then masked out with a clever arrangement of masks and baffles, sometimes polarization optics, and recently also diffraction based (so called binary) optics that comprise the coronagraph. One recent result of such a system is given in Figure 88. The AO team at the Hale telescope on Mount Palomar maps the existing deformable mirror onto a 1.5m subsection of the 5m primary, achieving a projected actuator spacing of $\sim 9\text{cm}$. A system upgrade with a new DM with 3388 actuators, achieving the same actuator footprint on the full aperture of the telescope, is planned to come online in 2012.

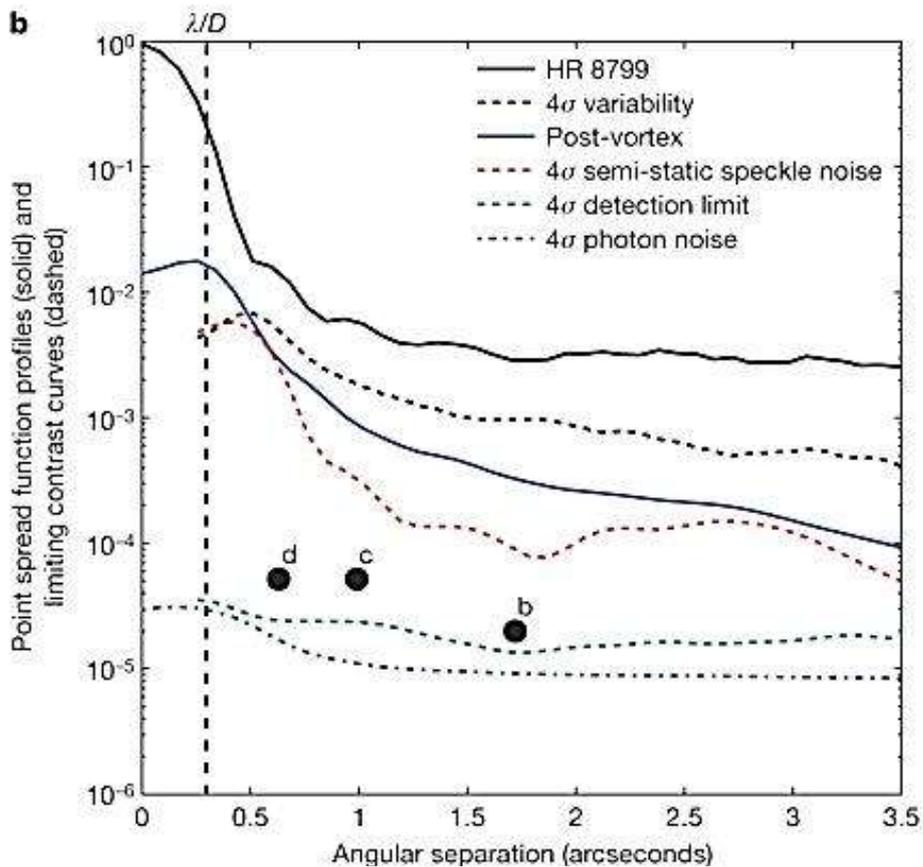
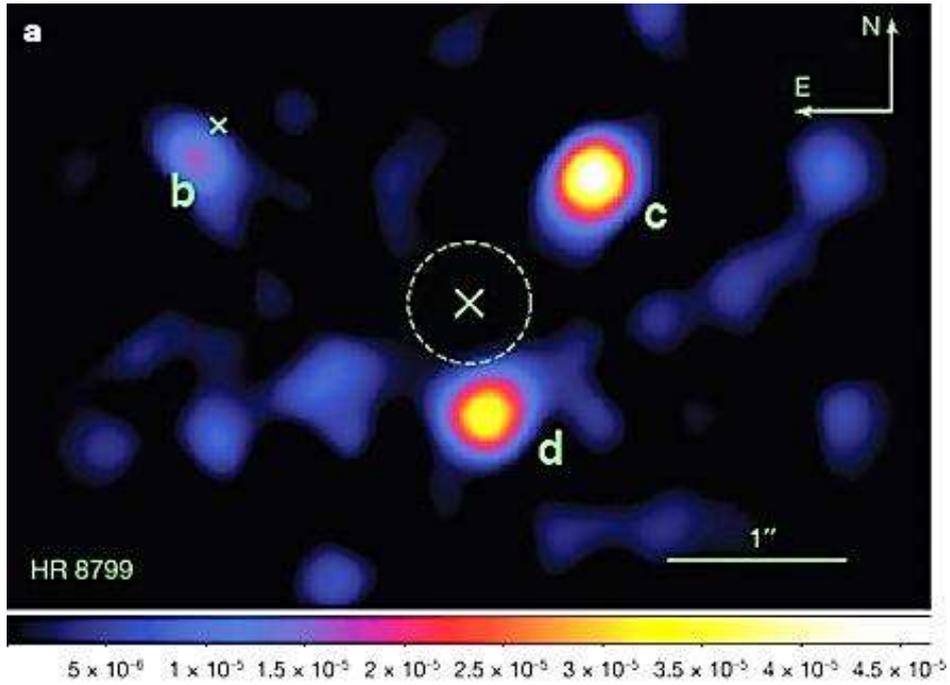


Figure 88: This image shows the three planetary companions to HR8799. The star is located at the spot marked with X. The planets' orbits are 24, 38, 68 AU, respectively. It was generated using the so called well-corrected aperture mode of the AO system, and an optical vortices coronagraph, operating in K_s -band. The achieved contrast level is shown in the lower plot. Taken from E. Serabyn et al. 2010. Image credit: NASA/JPL-Caltech/Palomar Observatory.

5.6 Doppler searches

In 1952 Otto Struve proposed to use spectroscopic measurements of the radial velocity of stars to search for the signature of the gravitational force of orbiting planets (Struve 1952).¹⁰ As two gravitationally bound objects revolve around a common centre of mass, the existence, and the orbit, of the smaller body can be inferred from observations of the larger body's movement. It is therefore not necessary to detect light from the smaller body (the planet) directly. The radial velocity of the host star can be determined by looking for periodic shifts of the absorption lines in its spectrum, caused by the Doppler effect (Doppler 1842).

The challenge for such an indirect detection based on the gravitational effect of the planet on its host star is the extraordinary precision necessary, as is the case for astrometric measurements. As the masses of star and planet are so grossly different, the effect on the star's motion is extremely small. There are fundamental differences between the Doppler method and the astrometric search: the amplitude of the Doppler signal is independent of the distance to the star. On the other hand, the inclination and the longitude of the ascending node cannot be measured with radial velocities. We only see the radial component of the motion, the projection onto a plane that contains the line of sight.

Mathematical foundations

The relevant mathematical formulation is straight forward, and can be found for example in Quirrenbach 2006.

The projection of the orbit of a body onto the vector describing the line of sight yields:

$$V_0 = K_1 \{ \cos(\nu + \varpi) + e \cos \varpi \}$$

where K_1 is the radial velocity amplitude, and ν is the *true anomaly*, defined as the position angle of the body measured from periastron. V_0 is the radial velocity with respect to the centre of gravity of the system.

The measured velocity V is V_0 corrected for the motion of the system barycentre, γ , and the motion of the observer, β , referred to as 'barycentric correction':

$$V = \gamma + \beta + K_1 \{ \cos(\nu + \varpi) + e \cos \varpi \}$$

The *true anomaly* is related to the *mean anomaly*, M , and the *eccentric anomaly*, E , as follows (T is the time since passage through periastron):

$$M = \frac{2\pi}{P} (t - T)$$

$$M = E - e \sin E$$

and

$$\cos \nu = \frac{\cos E - e}{1 - e \cos E}$$

¹⁰ He also mentioned the possibility of detecting transiting exoplanets with photometry – this was a really far-sighted yet concise paper.

Hence, ω , e , P and T can be derived from the radial velocity time series. The amplitude K_1 holds information about the semi-major axis of the orbit, and thus the mass ratio between both bodies. For $a_1 \sin i$, the projected size of the semi-major axis of the star's orbit, we can write:

$$a_1 \sin i = \frac{P}{2\pi} \sqrt{1 - e^2} K_1$$

The ratio of the semi-major axes of the two bodies is inversely proportional to the ratio of their masses:

$$\frac{m_1}{m_2} = \frac{a_2}{a_1}$$

Kepler's 3rd law relates the sum of the semi-major axis, $a = a_1 + a_2$, to the masses and the period P with the gravitational constant, G :

$$a^3 = \left(\frac{P}{2\pi}\right)^2 G (m_1 + m_2)$$

Together this yields

$$\frac{(m_2 \sin i)^3}{(m_1 + m_2)^2} = \left(\frac{P}{2\pi G}\right) K_1^3 (1 - e^2)^{3/2}$$

The left hand side of this equation is called the mass function of the system. With the assumption $m_1 \gg m_2$ this can be simplified to

$$m_2 \sin i \approx \left(\frac{P}{2\pi G}\right)^{1/3} K_1 m_1^{2/3} \sqrt{1 - e^2}$$

If the host star's mass m_1 is known, $m_2 \sin i$ can be derived.

$m_2 \sin i$ is often called the minimum mass of the planet, as the effect of projection obviously makes the mass appear to be less than its true value. It is interesting to note the distribution of inclination angles, i . The orientation of orbits should be randomly distributed, but the projection, the inclination i , is not. Many three-dimensional orientations lead to a high inclination angle – the orientations with the rotation axis perpendicular to the line of sight, and rotated around the line of sight all have $i = 90^\circ$ - but only exact two orientations lead to a inclination of 0° , namely the orbital axis parallel and antiparallel to the line of sight.

Indeed, what is randomly distributed is $\cos i$, so nearly 90% of the systems have $\sin i \geq 0.5$; their real mass is not more than twice the measured value of $m \sin i$. This is very important for the distinction between real planets and other substellar companions, like brown dwarfs. While we cannot exclude the possibility that a single candidate is above the mass threshold for brown dwarfs, in a statistical sample the values of $m \sin i$ represent the true distribution well.

5.6.1 Instrumentational challenges

As mentioned before, the problem for the Doppler method lies in the small amplitude of the signal due to the huge mass difference between host and planet. To illustrate the magnitude of the signal we look for: Jupiter imposes a reflex motion of 13 m/s on the sun, Saturn 2.7 m/s, and Earth only 0.1 m/s. Only for the Jupiter – Sun system does the barycentre lie (just) outside the sun itself.

Let's assume we want to detect Jupiter-like planets, and therefore need to be able to have a precision of 3 m/s for one integration (which is what can be done with the Hamilton spectrograph we use). If we translate this into spectroscopic resolution: the speed of light divided by 3 m/s equals 10^8 . State-of-the-art high resolution echelle spectrographs in astronomy typically have a resolution between $R = 50.000$ and 120.000 . The necessary precision is at least $1/1000^{\text{th}}$ of a resolution element, normally two pixels of a CCD chip. One pixel corresponds to 1 – 3 km/s. With a standard pixel size of 15 micron, this equals a shift of the spectrum on the CCD of 30 nanometres – an extraordinary precision.

Several requirements must be met to achieve such a low measurement error.

- The host star's spectrum must contain a large number of narrow absorption lines to allow for averaging over many lines and pixels
- The star must be chromospherically quiet - that is, inactive and not pulsating, so the absorption line system is stable over time
- The apparent brightness must be high enough to allow one to reach a substantial (order 100) signal-to-noise ratio during a moderate exposure time.

For the spectrograph, the requirements are

- high resolution to resolve the stellar spectral lines
- large wavelength coverage to record a sufficient number of lines
- very good mechanical and environmental stability on the timescales of exposures and of observing programs
- an extremely precise wavelength calibration method

These points deserve some further comments. Depending on target brightness, overall telescope quantum detection efficiency and desired signal to noise ratio, as well as calibration method, the resolution has an optimal value. Using the iodine cell technique, and assuming a modern low readout noise CCD (<5 e- RMS), simulations show a sweet spot around $R=80.000$ where the detection noise is balanced against resolving power¹¹. Bouchy, Pepe & Queloz (2001) arrive at a similar result with an analytic treatment.

In contrast, larger wavelength coverage is never a disadvantage. However, depending on the calibration method, only a moderate part of the recorded spectrum is actually used to determine the radial velocity shift. Large wavelength coverage can also help to distinguish between genuine planet signals and false signals introduced by atmospheric effects in the host star. For example, the contrast between star spots and the rest of the photosphere decreases towards the near infrared, and so would the spurious Doppler signal that results from the

¹¹ John A. Johnson, private communication.

motion of such spots. Wavelength calibration methods for large parts of the spectrum are consequently the focus of substantial efforts.

The need for mechanical and environmental stability of the spectrograph is clear considering the absolute displacement of the spectrum one wants to measure. Even if the system is mechanically perfect, temperature changes lead to varying dispersion of the gratings used, and changing focal length of the optics. In older spectrographs with glass as the grating substrate material, a temperature difference of 1K can easily lead to a RV shift of several hundred m/s. Changes in outside air pressure of 10% during a night, or even shorter, are not unlikely. The resulting RV shift, primarily because of the change in index of refraction of the medium surrounding the grating, again substantially exceeds the desired precision. Considering this, a spectrograph must either be decoupled from these changes, for example by placing it in the vacuum, or the calibration has to track these changes.

Wavelength calibration methods

Two different methods for wavelength calibration are successfully used: a reference lamp as calibrator, and the absorption cell technique.

Using an emission spectrum recorded simultaneously with the stellar spectrum is the classical way to calibrate a spectrum. In case of a fibre coupled spectrograph, two fibres, the science fibre delivering the starlight and a reference fibre, are mounted on top of each other in the cross-dispersion direction. The amount of cross dispersion permits the spectra of both fibres to be recorded interlaced on the detector. The Swiss planet search team uses this technique on the spectrographs ELODIE, CORALIE and HARPS. A Thorium-Argon lamp is used as the reference lamp; the Thorium spectrum provides stable emission lines to calibrate a wide wavelength range (HARPS: 378 – 691 nm, Mayor et al. 2003). The use of a fibre in this design is crucial to decouple the spectrograph from slit and pupil illumination changes due to guiding errors. The multimode fibre provides a so called ‘scrambling’ of its output by internally mixing the intensity distribution of the excited modes, and effectively provides a stable output light cone geometry relatively independent of the input illumination. As the calibration lamp spectrum cannot track such changes, a slit coupled spectrograph will suffer from uncorrected spectral shifts if one attempts to calibrate the spectrum with this technique.

The great advantage of the iodine cell technique, on the other hand, is that it provides the possibility to track such changes. A gas cell is mounted in front of the entrance slit of the spectrograph such that molecular absorption lines are superimposed on the starlight before entering the spectrograph. This tracks pupil illumination shifts as well as image motion during observation. Any changes in the spectrograph’s dispersion function or its point spread function are also experienced by the absorption spectrum imprinted on top of the stars spectrum, greatly relaxing the necessary intrinsic stability requirements for the spectrograph. This makes it possible to use slit coupled spectrographs for high precision RV work.

An engineering challenge is the selection of a proper absorbing medium. The use of a superimposed absorption spectrum was first proposed by Griffin & Griffin (1973), who suggested the use of the telluric lines in the Earth atmosphere. This technique has recently seen new interest for observations in the infrared (Bean et al. 2010). However, for visible wavelength spectroscopy the technique was altered to use a gas cell. Campbell, Walker & Yang (1988) used a very long cell filled with hydrogen fluoride gas (unfortunately lethal), and achieved a stability of 15 m/s. This led to the discovery of the first extrasolar planet (although it was confirmed only much later). Marcy & Butler (1992) proposed the use of

Iodine instead, which has high absorption coefficients, good wavelength coverage in the centre of the visible spectrum (500 – 630 nm), and is chemically stable. Such a cell was first installed at the Hamilton Spectrograph at Lick observatory; it is the same cell we use for our program.

Iodine produces a very dense forest of lines that are not resolved with the Hamilton spectrograph. It is therefore nearly impossible to separate the star's spectrum and use the iodine lines like the tickmarks on a ruler. Instead, one must use sophisticated forward modelling techniques to simulate the measured spectrum and extract the Doppler shift (Butler et al. 1996). A somewhat famous plot to illustrate this technique is shown in Figure 89. On quiet stars, the achieved long term stability is slightly better than 3 m/s. It must be noted that the Hamilton spectrograph is slit coupled, not fibre fed, and experiences large amounts of guiding jitter, as well as changes in temperature and air pressure. While the precision is often compared to that of HARPS, the instruments are very different, and the competitiveness of the programs observed with the Hamilton spectrograph can be attributed to the ability of the iodine cell to track large instrumental shifts with high precision.

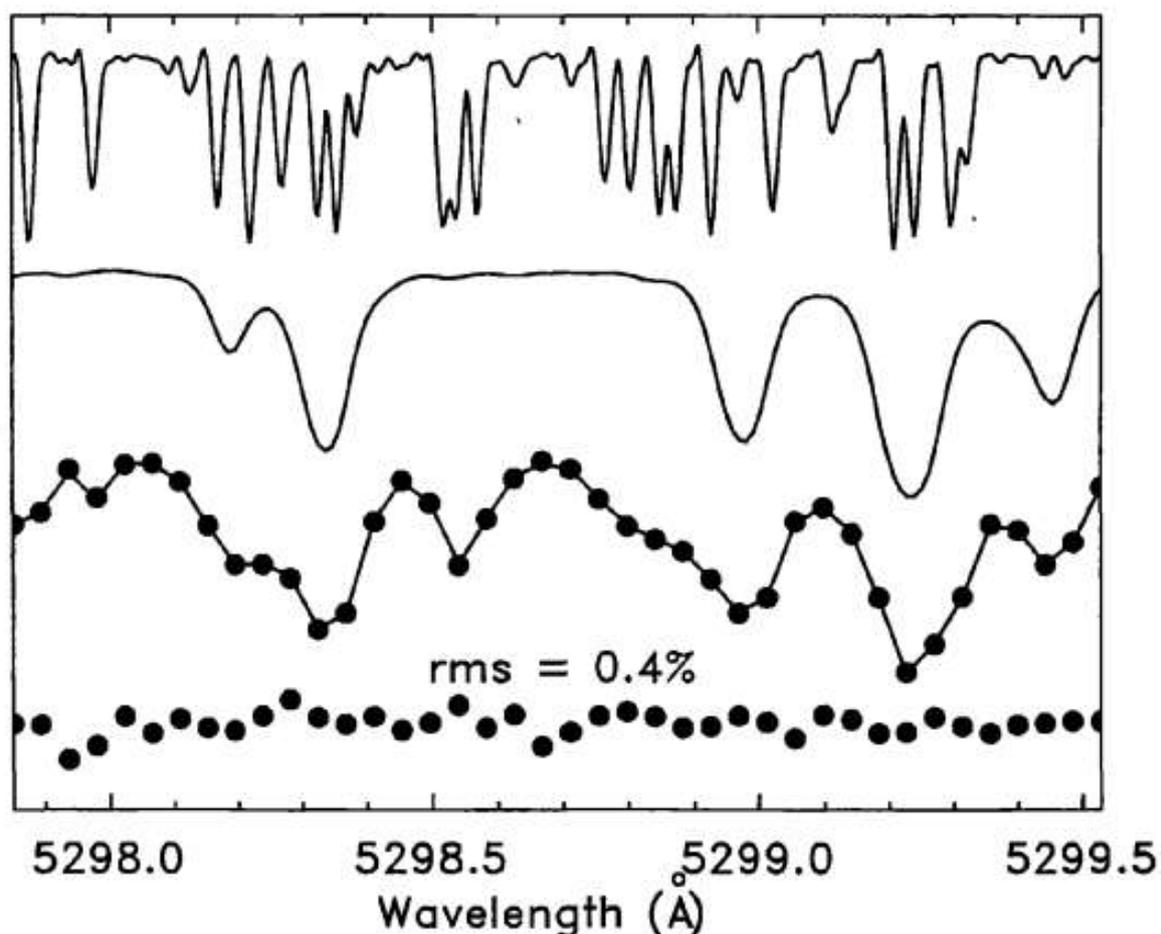


Figure 89: The modelling process. Top: The template iodine cell spectrum. Second: The template spectrum (τ Ceti, G8 V). Third: The points are an observation of τ Ceti made through the iodine absorption. The solid line is a model of the observation. The model is composed of the template iodine and stellar spectra. The free parameters consist of the spectrograph PSF and the Doppler shift of the template star relative to the template iodine. Bottom: The model and observation differ by 0.4% RMS. From: Butler et al. 1996.

6 Data analysis

The following is a description of the data extraction and processing that is done on the raw measurements obtained with the Hamilton spectrograph at Lick observatory.

First, I outline the extraction process from the raw CCD frames, then the principles of the Doppler fitting. Lastly, I discuss the ingredients for the Doppler analysis.

6.1 Data acquisition and extraction

The raw data consists of CCD frames covering the echelle format of the Hamilton spectrograph. On the raw frames, flatfield subtraction and cosmic ray removal are performed with the Hamilton raw data reduction pipeline developed by Jeff Valenti and others in the early 1990s.

Flats are obtained with the help of a quartz lamp and a slit that is longer than the one used for the observations. Because of the change in efficiency of the spectrograph, and mainly the slope in flux from the used quartz lamp as a function of wavelength, filters are used to achieve a sufficiently even illumination of the echelle orders for the flat. Until 2009, two different filters were used, that have a bandpass in the red and blue part of the spectrum, respectively. A flat was a superposition of typically 8 exposures through each of the filters. To avoid unnecessary readout noise and to shorten the time needed for the calibrations, the integration was halted after the exposure time for the blue filter, the filter was changed, and the exposure resumed. Since 2009, after a controller upgrade which sped up the readout considerably (less than 1 minute), a single filter that produces an even response throughout the spectrum is available, and used. Cosmic rays are removed using a standard procedure. The pixels covering the width of each order are then collapsed (added in cross dispersion direction) onto a single measurement. The addition is performed along the CCD columns.

The resulting extracted data is an echelle spectrum covering 39 orders, and is saved in Lick Observatories' .ccd or more recently in the .fits format. The information content between the two formats is equivalent, the data structure only minimally different.

For the RV measurement, only the orders that cover the wavelength region in which the iodine cell imprints the absorption lines onto the spectrum are used, starting typically with the order number 25. Due to a problem with the identification of the first echelle order, the numbering of the orders can shift by plus / minus one. This is taken care of by a bootstrap; we manually put the order number containing the H alpha line into the reduction software. It is desirable to note this directly in the fits header of the reduced data file in the future.

6.2 Doppler shift measurement

The technique to extract the Doppler shift is based on forward modelling of the recorded spectrum and comparing it with a reference observation.

We derive the radial velocity value by measuring the relative shift of the star's spectrum with respect to the iodine spectrum, and comparing it to the shift measured in the first observation of this target. The RV we measure is only relative - we cannot determine the absolute velocity with respect to one coordinate system.

We start with the assumption that any observation can be described with the following model,

$$S(\lambda) = LSF \otimes \{ISS(\lambda + \Delta\lambda) * I2(\lambda)\}$$

where S denotes the measured spectrum (flux as a function of wavelength), LSF the instrumental profile (or the line spread function) of the spectrograph, ISS stands for intrinsic stellar spectrum, and $I2$ is the absorption spectrum of the iodine cell. We generate a model observation using the right hand side of this equation, and then optimize the fit between model and real observation, using the Doppler shift (a wavelength shift $\Delta\lambda$ of the ISS) as a free parameter.

Because of observed changes in the LSF on short time scales, as well as changes in the spectrograph's dispersion function (often called the wavelength solution), there are free parameters allowing for perturbations of these. The code runs two consecutive optimisation computations; during the first one all the parameters are adjusted to minimize the chi square of the model with respect to the observation. Then all the parameters are held fixed, and the Doppler shift is reoptimised in the second pass.

This calculation is not performed globally on the whole spectrum, but on small wavelength chunks spanning 80 pixels each. These are weighted based on their SNR and chi square, combined, and from this a single RV value is derived.

This strategy to extract the RV shift is valid under the following assumptions:

- The ISS remains constant
- The iodine spectrum is correct
- The model for the LSF is correct
- The barycentric correction is right

ISS

The template spectrum (the ISS) is derived from a high signal to noise ratio observation of the star without iodine cell. The knowledge of the ISS is dependent on the correct estimation of the LSF in the reference spectrum. Our template spectra are taken with the same spectrograph. This can lead to a degeneracy between the LSF and remaining signatures of the template's LSF . As no iodine lines are present in the reference spectrum, the LSF can only be measured indirectly. Considerable effort has gone into the necessary deconvolution, and modelling the LSF for this step. Still, the ISS is known only with much less resolution than the $I2$ spectrum.

Iodine spectrum

The $I2$ spectrum of the reference gas cell has been measured using a FTS (Fourier Transform Spectrometer) with a resolution of 300.000. Measurements were taken with the MacMath FTS at Kitt Peak Observatory. The spectrum is not deconvolved with an estimated instrumental profile of the FTS. Efforts are underway to repeat these measurements with higher resolution, the goal is $R=10^6$. It is notable that the quality of the final model fits depends considerably on the used measurement for the iodine cell's spectrum, as I investigated by trying to reduce a spectrum with the use of a FTS scan of a different iodine cell (both FTS scans had the same wavelength resolution).

LSF

A guess of the *LSF* is derived from observations of B stars with the iodine cell. B stars are so hot and rotate so fast that they show no intrinsic stellar features through the iodine range, making them an ideal ‘quartz lamp in the sky’. We believe this method to be superior over using the real quartz lamp, as in the latter case, the light cone illuminating the spectrograph is not the same as for a regular observation, which can lead to differences in the instrumental profile.

Extensive tests have shown that the Hamilton’s *LSF* can change rapidly, between consecutive exposures. It is neither constant over the full CCD detector, nor is it reproducible. The necessary nightly focus adjustment changes the *LSF* considerably, compared to the precision we are looking for. One reason for this is the fact that the electromechanical stages moving the CCD dewar for focussing show hysteresis and tilt. The focus and tilt tolerances are very tight due to the fast focal ratio of the spectrograph camera.

Barycentric correction

We perform a precise correction of the measured radial velocity value to the solar system barycenter, to subtract the effects of the earth’s motion, mainly its rotation and the rotation around the sun, as well as the reflex motions due to the other planets. To calculate these effects to a precision of considerably better than 1 m/s, one must know the exact time when the spectrum was taken to a precision of a few seconds. As the exposures themselves are much longer than this, we determine the photon weighted midpoint, a standard technique for high resolution radial velocity measurements. A photon counter with integrator is fed with 2% of the incoming light by a rotating fan blade mirror behind the spectrograph’s entrance slit, and measures the incoming flux several times a second. From the resulting distribution of flux over the length of the exposure, the midpoint is calculated. The relative motion of the Earth, or more precisely, the location of the spectrograph, including all the mentioned effects, at the midpoint time is then calculated with an IDL program, and the resulting velocity finally subtracted from the measured radial velocity value of the star.

6.3 Doppler code

For many years our group has been using the code developed by P. Butler et al. to do the Doppler analysis for iodine cell spectra taken with the Hamilton spectrograph. The data pipeline has been updated incrementally by many authors. On the quietest stars, the RMS scatter is 3 m/s with the Hamilton spectrograph, an excellent value, considering the spectrograph’s environment is not strictly controlled.

This solution is sometimes impractical, however, as we have to run the reduction on machines at Berkeley. Direct intervention and modification of critical points is complicated. Recently, we started testing the newly written code by D. Fischer from Yale University, largely based on the concepts used in the original code by Butler. During a 3 month visit to Prof. Fischer’s lab at San Francisco State University I had the opportunity to work with her on the code, and she made it available to us to run on our own machines. This gives us the opportunity to work on the code and closely examine the output for complicated targets.

7 A new planet around HIP31592

The vast majority of planetary systems have been found using the radial velocity technique, as discussed above. The largest subset of these systems has a solar-type main sequence star as the primary component. This can be attributed mainly to the early interest in stars similar to our sun combined with the favourable spectral properties of such stars, which enable them to be easily studied with the Doppler method. As the sample sizes continue to increase, we are entering an era where much of the interesting science involves statistical investigation of the planet and host populations, which can inform theories of planet formation. The findings of these investigations are biased by the selection effects. Our knowledge of the planet population is still mostly restricted to main sequence systems. Amongst the trends that have emerged are correlations between the host mass, occurrence of planets and planetary mass. The distributions of orbital parameters have also been investigated. One of the most surprising results is that the distribution of eccentricities spans a wide range, with a mean eccentricity of approximately 0.23 amongst more than 250 well-characterised exoplanet systems (Peek). This is much larger than expected based on the almost circular orbits of gas-giants within our own solar system. Several groups find a positive correlation between the planet occurrence rate and the metallicity of the host stars (Peek etc.). Some exoplanet surveys favour higher metallicity stars as their targets based on this trend (e.g....). It is still questionable whether such a correlation extends to different mass regimes, however.

To improve our understanding of planet formation, it is very important to expand the samples of planetary systems to different types of stars with both higher and lower masses. Several radial velocity surveys have recently focussed on M-dwarfs to probe the lower mass regime. At higher masses, late G and early K-giants provide an evolutionary sweet spot for Doppler searches, as they have numerous metal lines that are important for achieving high precision with the radial velocity method and, at least in some cases, are chromospherically quiet enough for planetary companions to be found. In addition, such stars are very bright and enable a competitive survey to be undertaken with a small telescope. Thus far, 29 planetary companions have been found around giant stars. The discoveries are summarized in Table 14.

A long-term program to search for planetary companions around K-giants has been conducted over the last 11 years, as first described in Frink et al. 2001. The sample now contains 376 stars. The first substellar companion found around one of these stars was presented in Frink et al. 2002. Further exoplanet candidates from the survey were discussed in Mitchell et al. 2007. I contributed substantially to the observational program over the last three years, conducting 15 observation runs of ~6 nights each Lick observatory. These additional data reveal the presence of new planet candidates whose orbits were formerly hidden in the noise. Promising targets are subject to detailed analysis and observed with higher cadence to further constrain the orbits. For at least one of the targets, HIP31592, the orbit is already well constrained. In the following chapter I present my analysis of the radial velocity data for HIP31592 and the derived orbital parameters of its exoplanet companion.

Table 14: Summary of giant stars with exoplanets

Star	HIP	V [mag]	Spectral type	Stellar mass [M_{Sun}]	Planet mass [M_{Jup}]	Period [days]	Ecc.	a2 [AU]	K [m/s]	
ι Dra	75458	3.29	K2 III	1.05	8.8	511	0.71			a
Pollux	37826	1.16	K0 IIIb	1.86	2.9	590	0.06	1.69	46.9	b
HD 11977	8928	4.68	G5 III	1.91	6.54	711	0.4	1.93		c
HD 47536	31688	5.25	K1 III	0.94	5	430			113	d
HD 13189	10085	7.57	K2 II	2-7	8-20	472	0.27	1.5-2.2	173.3	e
ϵ Tau	20889	3.53	K0 III	2.7	7.6	595	0.151	1.93	95.9	f
HD 104985	58952	5.78	G9 III	2.3	8.3	200	0.090	0.95	166.8	g
NGC2423 #3		9.45		2.4	10.6	714	0.21	2.10	137.6	h
NGC4349 #127		10.9		3.9	19.8	678	0.19	2.38	188	h
4 UMa	42527	4.59	K1 III	1.23	7.1	269	0.43		216.8	i
HD 17092		7.74	K0 III	2.3	4.6	360	0.166	1.29	82.4	j
11 Com	60202	4.72	G8 III	2.7	19.4	326	0.231	1.29	302.8	k
18 Del	103527	5.51	G6 III	2.3	10.3	993	0.08	2.6	119.4	l
ξ Aql	97938	4.71	K0 III	2.2	2.8	137	0 (fixed)	0.68	65.4	l
HD 81688	46471	5.40	K0 III-IV	2.1	2.7	184	0 (fixed)	0.81	58.58	l
14 And	116076	5.22	K0 III	2.2	4.8	186	0 (fixed)	0.83	100.0	m
81 Cet	12247	5.65	G5 III	2.4	5.3	953	0.206	2.5	62.8	m
HD 102272 b	57428	8.69	K0 III	1.9	5.9	128	0.05	0.61	155.5	n
HD 102272 c	57428	8.69	K0 III	1.9	2.6	520	0.68	1.57	59	n
HD 240210		8.33	K3 III	0.82	5.21	502	0.15	1.16	161.8	o
BD +20 2457b		9.75	K2 III	2.8	21.42	379	0.15	1.45	322.4	o
BD +20 2457c		9.75	K2 III	2.8	12.47	621	0.18	2.01	160.0	o
HD 173416	91852	6.05	G8 III	2.0	2.7	324	0.21	1.16	51.8	p
42 Dra	90344	4.83	K1.5 III	0.98	3.88	479	0.38	1.19	110.5	q
HD 139357	76311	5.97	K4 III	1.35	9.76	1126	0.10	2.36	161.2	q
11 UMi	74793	-0.37	K4 III	1.80	11.20	516	0.08	1.54	189.7	r
HD 32518	24003	6.43	K1 III	1.13	3.04	158	0.01	0.59	115.8	r
HD 110014	61740	4.65	K2 III	2.17	11.09	835	0.462	2.14	158.2	s
γ^1 Leo	50583	2.01	K0-K4 III	1.23	8.78	426	0.144	1.19	208.3	t

a) Frink et al. 2002, b) Reffert et al. 2006, c) Setiawan et al. 2005, d) Setiawan et al. 2003, e) Hatzes et al. 2005, f) Sato et al. 2007, g) Sato et al. 2003, h) Lovis & Mayor 2007, i) Doellinger et al. 2007, j) Niedzielski et al. 2007, k) Liu et al. 2008, l) Sato et al. 2008a, m) Sato et al. 2008b, n) Niedzielski et al. 2009a, o) Niedzielski et al. 2009b, p) Liu et al. 2009, q) Döllinger et al. 2009a, r) Döllinger et al. 2009b, s) de Medeiros et al. 2009, t) Han et al. 2010

7.1 Known properties of the host star

HIP31592, also known as 7 Canis Majoris, v^2 CMa, HR 2429 and HD 47205, is designated as a giant star of type K1 III or IV (e.g. Hipparcos, SIMBAD, Cenarro et al. 2007). It has a colour index of $B - V = 1.1$ and a visual magnitude of $V = 3.96$ mag. The parallax measured by Hipparcos (Perryman et al. 1997) is 50.4 mas, which translates to a distance of 19.8 pc. Using this distance, an absolute magnitude of $M_V = 2.46$ mag is derived. The metallicity $[\text{Fe}/\text{H}] = 0.18$ dex, $T_{\text{eff}} = 4744$ K, surface gravity $\log g [\text{cm}/\text{s}^2] = 3.11$ and radius of $5.08 R_{\text{sol}}$ were determined by da Silva et al. (2006) using theoretical stellar evolutionary tracks. Cenarro et al. (2007) report values of $[\text{Fe}/\text{H}] = 0.05$, $T_{\text{eff}} = 4751$ K and surface gravity $\log g = 2.93 \text{ cm}/\text{s}^2$ for the Medium-resolution Isaac Newton Telescope library of empirical spectra (MILES). The values of $v \sin i$ found in the literature show a range between about 1 and 3 km/s (see for example Massarotti et al. 2008, Setiawan et al. 2004). It is important to note that the star is unambiguously identified as a slow rotator.

As the star has such bright apparent magnitude, it has been observed and investigated numerous times. It serves as calibrator star for various interferometric programs (Bordé et al. 2002), and for infrared instruments (Blackwell & Lynas-Gray 1998). Its correct parallax was already reported in 1936 (Schlesinger 1936). An early listing in a Colour-Magnitude diagram from Oke (1959) is shown in Figure 90.

7.2 Observations

Fifty spectra of HIP31592 were obtained between 2000 and 2010 using the Hamilton spectrograph (Vogt 1987) on the Coudé Auxiliary Telescope (CAT) at Lick Observatory. The typical exposure time was 600 seconds. The data were reduced using the Doppler code from D. Fischer, using the method described in Butler et al. 1996. A total of six observations had a radial velocity shift that resulted in a wavelength mismatch to the chunks our template is defined on. The code could not handle these data for technical reasons, and so they were excluded from further analysis. The rest of the points are evenly distributed over the periods for which the star was observable. The resulting radial velocity data points and their formal errors are given in Table 15. The observed dispersion is 29 m/s.

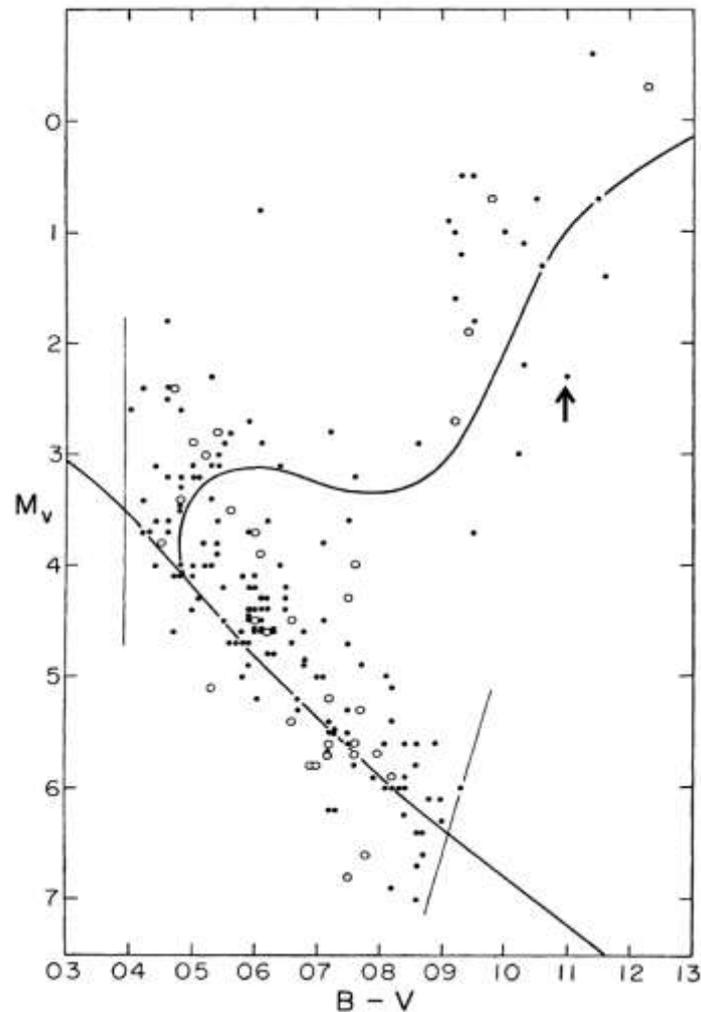


Figure 90: A colour-magnitude diagram for 185 F5-K2 stars, with the colour-magnitude diagram for M67 indicated by the heavy lines, taken from Oke (1959). The position of HIP31592 is indicated by the arrow.

Table 15: Measured radial velocities for HIP31592

Julian date -2440000	v_{rad} (m/s)	$\sigma_{v_{rad}}$ (m/s)
11808.021	18.16	2.43
11853.990	40.39	2.39
11896.855	34.91	2.58
11929.723	6.65	2.67
11992.658	-23.14	3.20
12177.026	-20.18	3.37
12259.801	-21.07	2.92
12297.795	-1.64	3.05
12531.995	41.95	2.96
12543.041	42.81	2.76
12616.817	22.50	2.70
12668.714	34.36	2.77
12901.024	-24.41	3.03
12933.909	-34.31	2.62
12963.961	-26.30	2.73
13025.820	-39.44	4.35
13269.050	36.42	2.53
13288.997	34.03	2.65
13400.787	39.85	2.46
13425.720	36.45	2.41
13442.655	29.03	2.69
13444.635	31.96	2.59
13618.029	-31.46	2.43
13650.022	-49.13	2.28
13656.062	-38.64	2.60
13701.917	-33.61	2.73
13740.972	-35.25	3.50
13788.657	-40.14	2.59
13827.648	-46.98	3.99
13982.026	-0.61	2.12
14054.911	23.53	3.15
14123.821	39.68	2.62
14206.666	23.09	2.65
14418.951	-24.31	2.93
14502.828	-19.59	2.32
14557.664	-22.19	2.50
14712.026	-15.76	2.67
14754.942	5.22	2.45
14777.943	5.42	2.28
14806.842	22.63	2.40
14882.811	15.00	3.39
14948.677	12.50	3.12
15098.048	-8.49	2.63
15121.051	-6.94	2.95
15154.943	-25.88	3.31
15174.901	-29.62	3.20
15241.706	-14.00	2.90

7.3 Orbital solution

A package of IDL programs available in our group was used to investigate a possible orbital solution for the radial velocity data. A Lomb Scargle periodogram (Figure 91) was used to search for significant periodic signals in the data given in Table 15. There is a clear periodicity in the data, with one highly significant period (~800 days) standing out in the periodogram, as expected from a visual examination of the radial velocity data plot shown in Figure 92. The measurements shown in Table 15 are plotted in Figure 92 together with a Keplerian fit. The Kepler orbit was derived by fitting a model to the data points with the parameters given in Table 16. The fitting routine, recently updated by Christoph Bergmann, minimizes the chi squared of the model with respect to the data using the IDL program mpfitfun.pro to drive the chi square optimisation. The residual scatter is 10.6 m/s. The orbital parameters, also shown in Table 16 are derived from the fitted model parameters using the equations outlined in Section 5.6.

Table 16: Model and orbital parameters

Parameter	Value	Error
Period P (days)	785	1.4
Eccentricity e	0.20	0.02
Epoch t_0 (JD)	2451916	10
$a_1 \sin i$ (AU)	0.002467	3.9e-5
Longitude of periastron ω (radians)	1.16	0.09
Radial velocity offset (m/s)	-1.71	0.42
Mass function (solar masses)	3.25e-9	
Primary mass (solar masses)	1.37	0.18
$M_2 \sin i$ (Jupiter masses)	1.92	0.16
a_2 (AU)	1.85	0.08
v_{rad} amplitude (m/s)	34.95	0.55

To see the full phase coverage and the scatter behaviour around the fitted curve, a plot which folds the phases into one is useful. This is shown in Figure 93 with the Keplerian fit overplotted.

I find a best fit orbit with a period of 785 days. The minimum mass of the companion is 1.92 Jupiter masses, assuming a mass of 1.37 solar masses for HIP31592. This mass is derived from modelling stellar evolutionary tracks (Diplomarbeit A. Kuenstler).

The fit matches the radial velocity data very well, as can be seen in Figure 92 and Figure 93. The residuals are shown in the lower panel of Figure 92. The residual scatter is in the normal range for “quiet” stars from our sample, discussed in Hekker et al. 2006. Some of the stars in our sample have large, seemingly uncorrelated RV scatter as well, and the quietest targets in our sample reach values as low as a few m/s RMS.

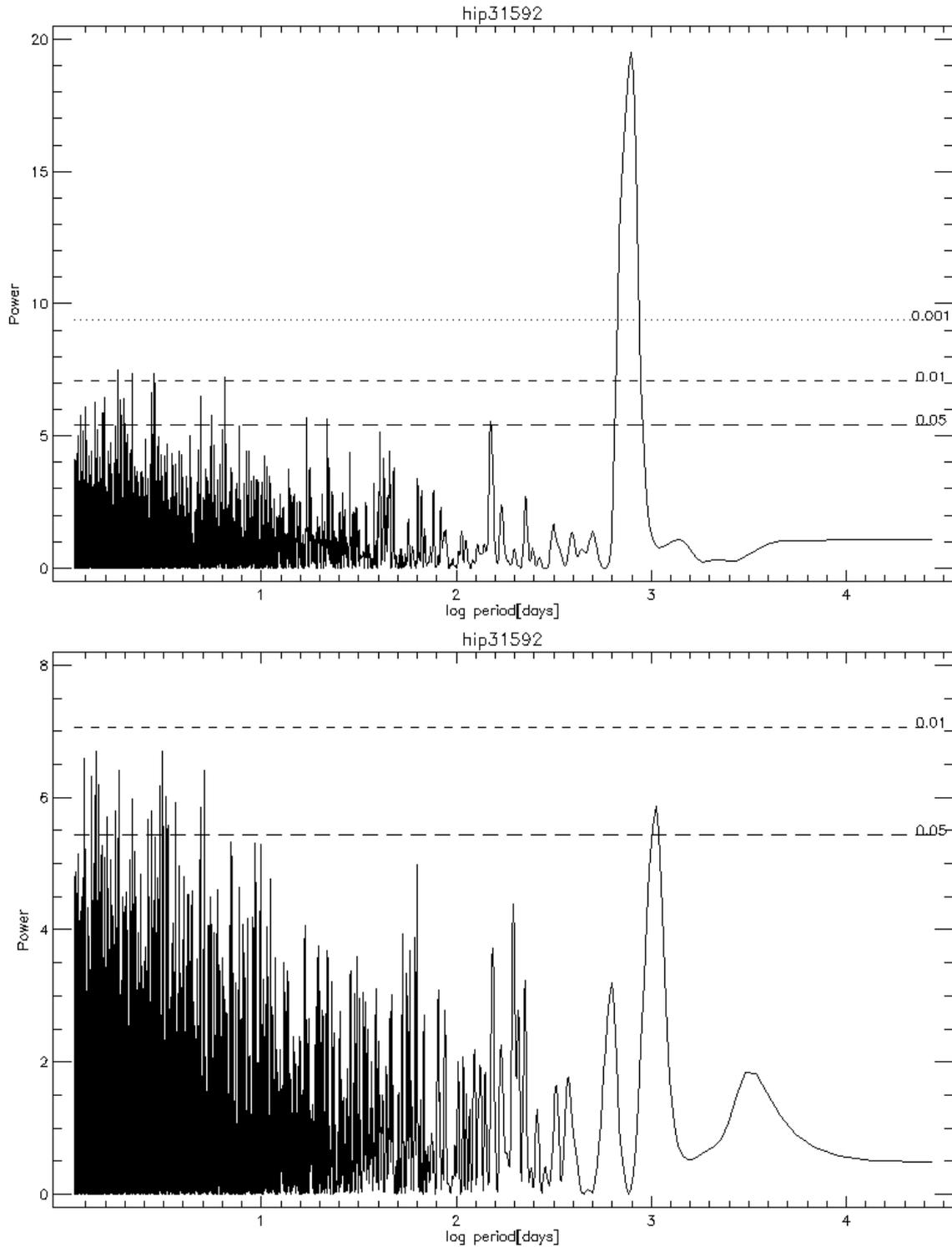


Figure 91: Upper panel: Lomb-Scargle periodogram for the data presented in Table 15. A highly significant peak is visible with a period of ~800 days. This peak has a False Alarm Probability (FAP) of 0%. Lower panel: Lomb-Scargle periodogram of the residuals. No additional highly significant peak is found. A peak at approximately 1200 days has a false alarm probability of ~3%.

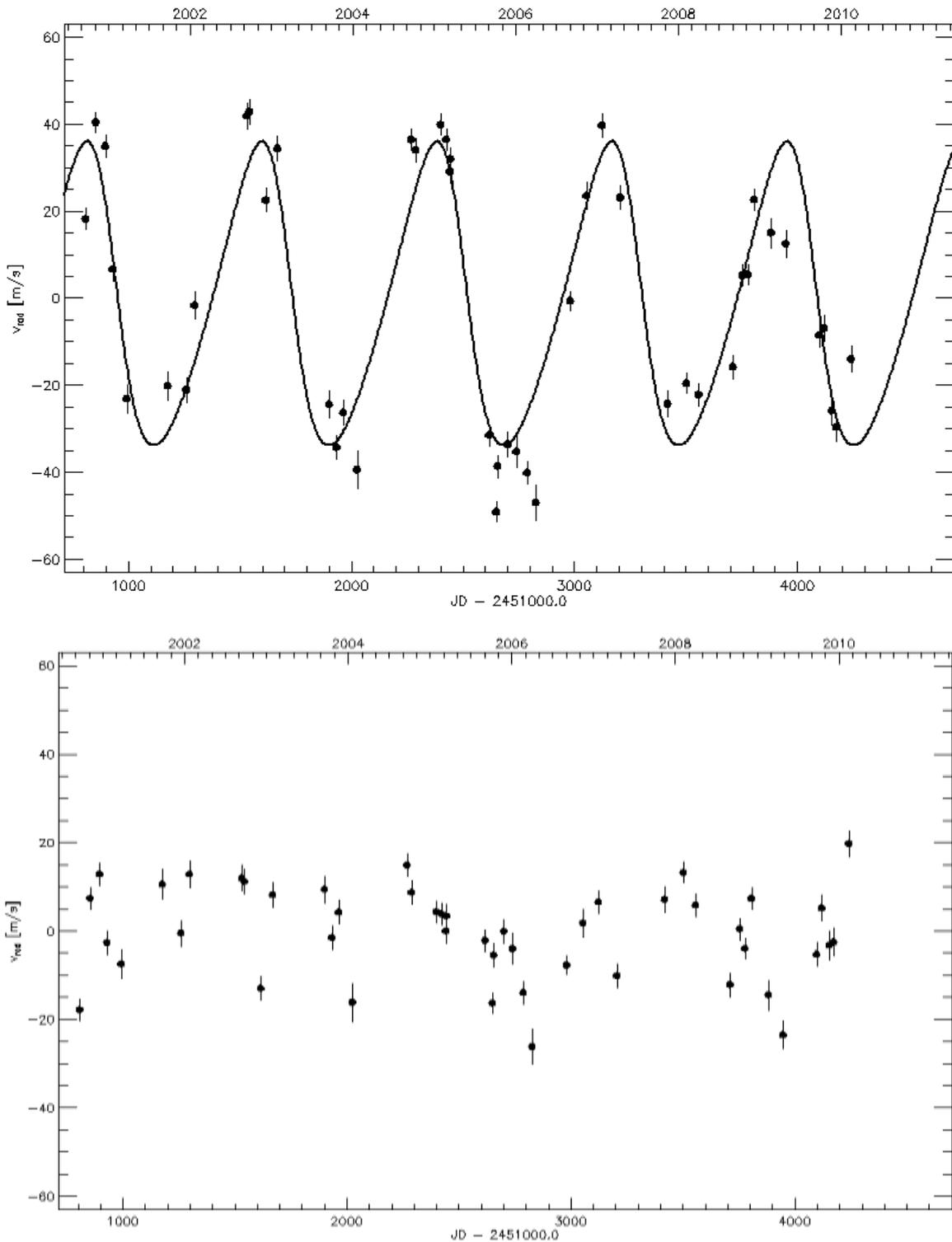


Figure 92. Upper panel: Radial velocity data points obtained between 2000 and 2010 for HIP31592. The best fit Keplerian orbit with a period of 784 days is overplotted. Lower panel: Radial velocity residuals obtained by subtracting the model from the data in the upper panel. The residuals have a scatter of 10.6 m/s and hint at a periodic signal of 1200 days as shown in the periodogram.

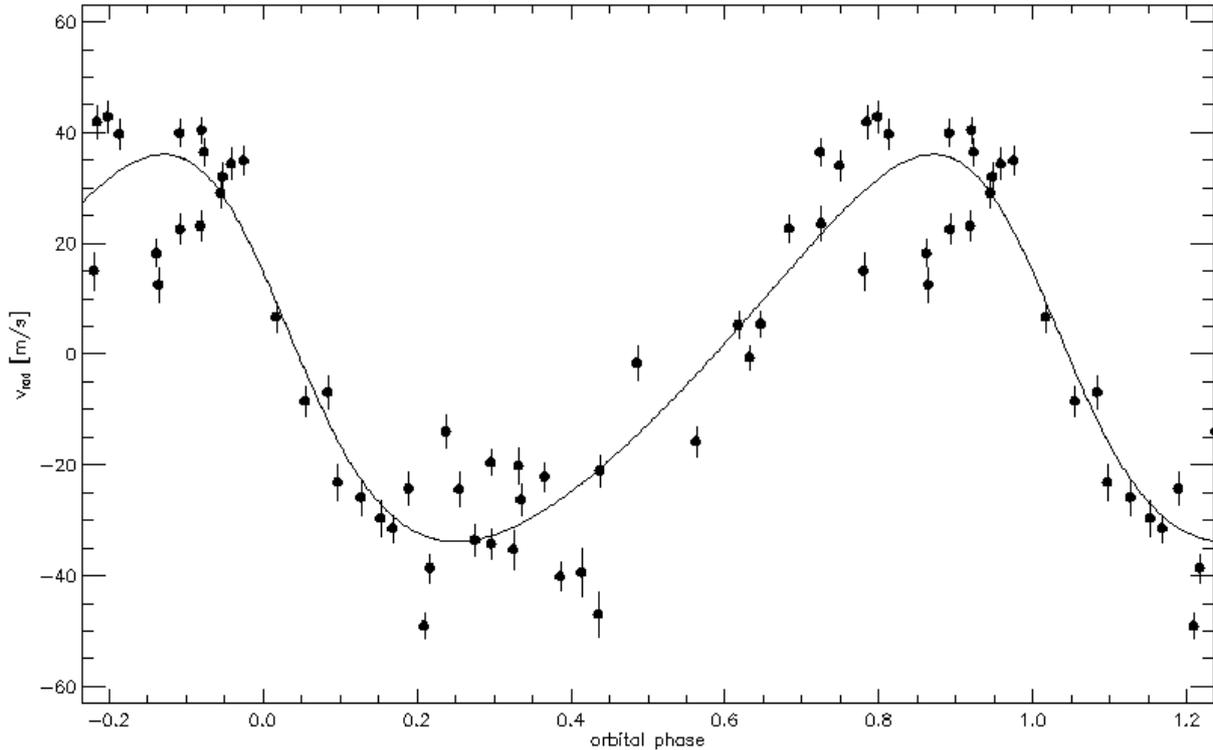


Figure 93: Phase-wrapped plot of the radial velocity data shown in Table 15 and Figure 92, with the Keplerian fit with the parameters given in Table 16 overplotted.

Although the scatter in the residuals is small, I analysed the residuals for further periodic signatures that may indicate another planetary body. The Lomb Scargle periodogram of the residuals (lower panel of Figure 91) shows a smaller peak at a period of 1200 days, as well as several peaks at shorter periods. It is important to note that with the number of data points we have and sampling frequency we are not sensitive to short orbital periods. This can be clearly seen from the random noise at shorter orbital periods in Figure 91. Spurious periods at or below the Nyquist frequency are most likely effects of aliasing, and have no significance.

The secondary peak analysed here occurs in a region of the periodogram that is sampled well above the Nyquist frequency. The FAP for this period is approximately 3%, too high to draw definite conclusions about the existence of another body in the system. On the other hand, due to the inherent scatter in our data originated in the atmospheric jitter of the typical target, a signal at the level I investigate here is almost certainly buried in noise. Identification of likely candidates is important to possibly raise the observing cadence. The detection of long period orbits requires a very long baseline and therefore is very expensive in terms of observing time. To gain the same level of confidence we are used to from typical radial velocity detection in the past, where multiple orbits are well sampled over a short period of time, one has to wait for a very long time, as our knowledge of the orbit increases so very slowly (Wright et al. 2007).

Against this background I further analysed the weak signal, by attempting an orbital fit to the residuals, in this case including a weak linear trend. The result is plotted in Figure 94 with the parameters of the fit and resulting orbital parameters shown in Table 17. A possible companion's orbit is poorly constrained by this data, but the overall fit improves minimally with the inclusion of the second orbit, lowering the residual scatter from 10.6 to 8.5 m/s.

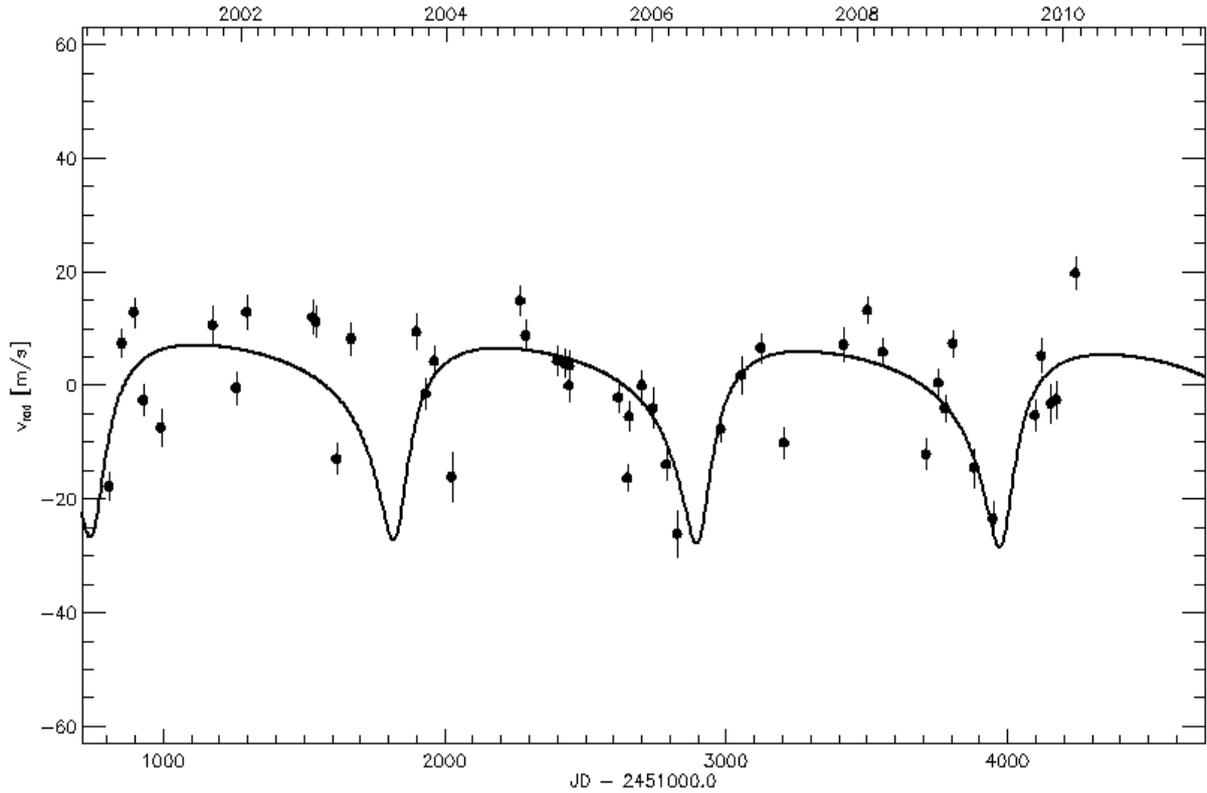


Figure 94: A Keplerian fit to the residuals shown in the lower panel of Figure 92. The eccentricity was very poorly constrained in the fit, and was allowed a maximum value of 0.6 to match the amplitude to the peak to valley RV variation of the data points.

Table 17: Model and orbital parameters for a Keplerian fit to the residuals

Parameter	Value	Error
Period P (days)	1077	5.5
Eccentricity e	0.6	-
Epoch t_0 (JD)	2452830	10
$a_1 \sin i$ (AU)	0.001346	1.2e-4
Longitude of periastron ω (radians)	3.47	0.07
Radial velocity offset (m/s)	-0.12	0.97
Trend	-0.20	0.16
Mass function (solar masses)	2.80e-10	
Primary mass (solar masses)	1.37	0.18
$M_2 \sin i$ (Jupiter masses)	2.28	0.21
a_2 (AU)	2.28	0.10
v_{rad} amplitude (m/s)	17.00	1.49

7.4 Astrometric signature

As outlined before, the Doppler method is not capable of measuring the inclination of an orbit. Hence, with radial velocity data alone we cannot determine the real mass of the planet M_p , but only its value multiplied by the sine of the inclination angle, $M_p \sin(i)$. It is not clear a priori whether the inclination of the system is very low and therefore the companion has a mass very different from its minimal value. From the radial velocity observations of HIP31592 alone the probability that the real mass of the second body is above the threshold for brown dwarfs is very small. The division between brown dwarfs and planets is drawn at about 13 Jupiter masses. With $M_p \sin(i) = 1.92 \pm 0.03 M_J$, the inclination angle would have to be as small as 8.4 degrees at most for it to be a brown dwarf.

The likelihood of a particular inclination angle is determined by the distribution of solid angles, under the assumption the orientation of planetary system is random. Hence the likelihood can be computed as $P(i < i_0) = \cos i_0$. For the upper limit of 8.4 degrees, this gives a probability of 98.9% that the angle is smaller. The companion of HIP31592 is almost certainly a planet.

As another measure to determine an upper limit for the companion's mass, and to rule out the unlikely event that we look almost perfectly face-on at the system, I looked for astrometric signatures of the second body in the Hipparcos data. Although Hipparcos is unlikely to be able to detect a planetary mass companion directly (Pourbaix 2001, Pourbaix & Arenou 2001) a non-detection can place an upper limit on the companion mass. The Hipparcos Intermediate Astrometric Data, published with the Hipparcos Catalogue (Perryman et al. 2007, ESA) is necessary for this kind of analysis. It includes the abscissa residuals from which the astrometric solution was obtained, and allows for the model to be updated based on external data (in this case, the radial velocity data). A suite of IDL programs written by Sabine Reffert was used to fit a model to the residual 1D data from the Hipparcos release, following the procedure outlined in van Leeuwen & Evans (1998) and Reffert & Quirrenbach (2006).

It is necessary to understand that the single measurements only provide information about one dimension, due to the particular way Hipparcos retrieves the data. The model used here includes the five parameters from the original fit by Hipparcos, which are position, proper motion and parallax (RA , DEC , v_{RA} , v_{DE} , and π). These parameters describe the proper motion of the star and its distance. Additionally, the seven parameters describing the orbital motion due to the presence of a planet are used. As the error bars on the parameters retrieved by the radial velocity measurements are much smaller than from the astrometric measurement, they are kept fixed in the model fit. The two additional orbital parameters which are not determined by the radial velocity measurement, the longitude of the ascending node Ω and the inclination i (which we are particularly interested in) are fitted to the Hipparcos data, together with the five original parameters.

The resulting confidence intervals for Ω and i are shown in Figure 95, where the contours indicate the 68.5% (1σ), 90% (2σ) and 98% (3σ) probabilities. We did not find any relevant orbital motion. However, we can rule out inclination angles lower than 9 degrees at the 3σ level, as can be seen in Figure 95. This conveniently rules out the inclination range where $\sin i$ would become small enough to give rise to a companion mass in the Brown dwarf regime, giving a maximum mass of $12 M_J$.

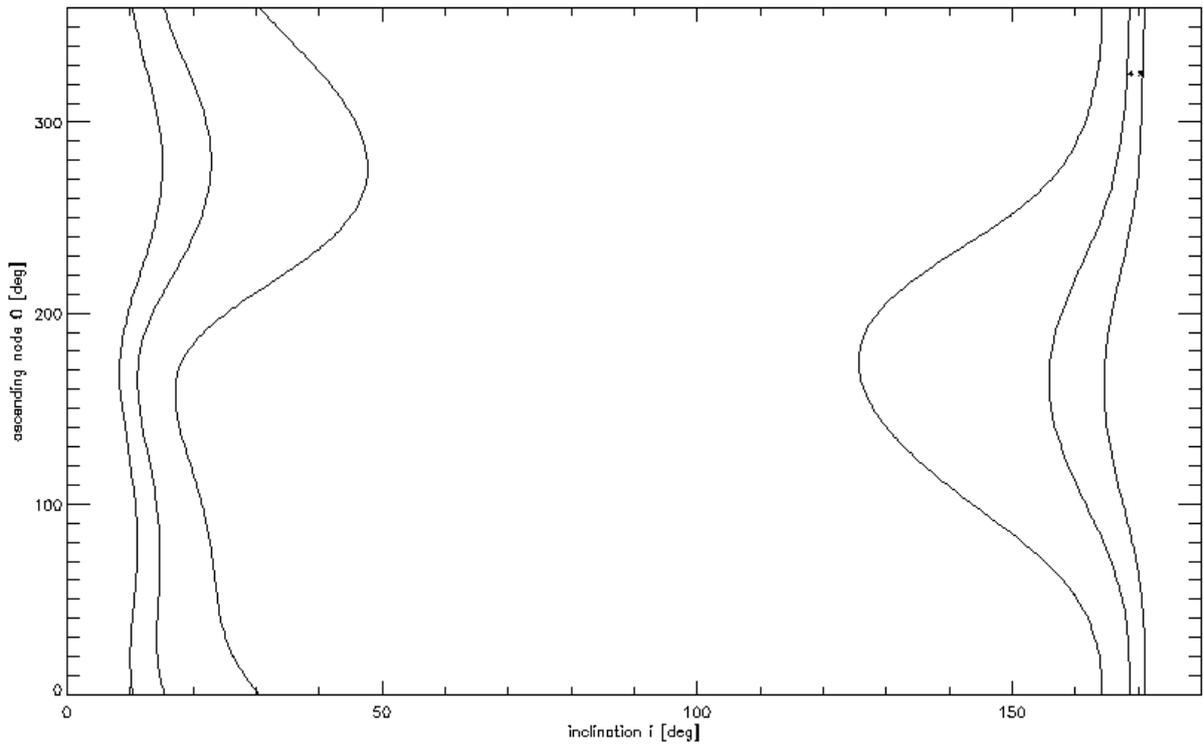


Figure 95: Confidence intervals for Ω and i determined from the combination of Hipparcos and radial velocity data. The lines indicate 1, 2 and 3 sigma contours, from the centre outwards. As no signature could be found, we cannot constrain definite values. However, high and low inclinations can be excluded at the 3 sigma level.

7.5 Comparison to known systems

I compared the found parameters to the systems listed in Table 14. The value of $m \sin(i)$ is slightly lower than other published values, making the companion of HIP31592 the lowest mass planet found around a giant star so far. The extension of the parameter space to lower mass companions is very interesting, of course, to be able to compare the planet population around giant stars to the one found around main-sequence stars, which extends to much lower masses. Figure 96 shows the distribution of companion mass as function of host mass for the 29 known planets around giant stars. HIP31592 is highlighted with a diamond. Similarly, Figure 97 shows the period and planet masses of the known systems. Only four systems with longer periods have been announced. This is not surprising, as orbits with longer periods need a longer baseline to be constrained. The eccentricity we found for the companion of HIP31592 (0.20) is close to the mean of the distribution for giant stars (0.19) shown in Figure 98, as well as for all known exoplanet systems (0.23).

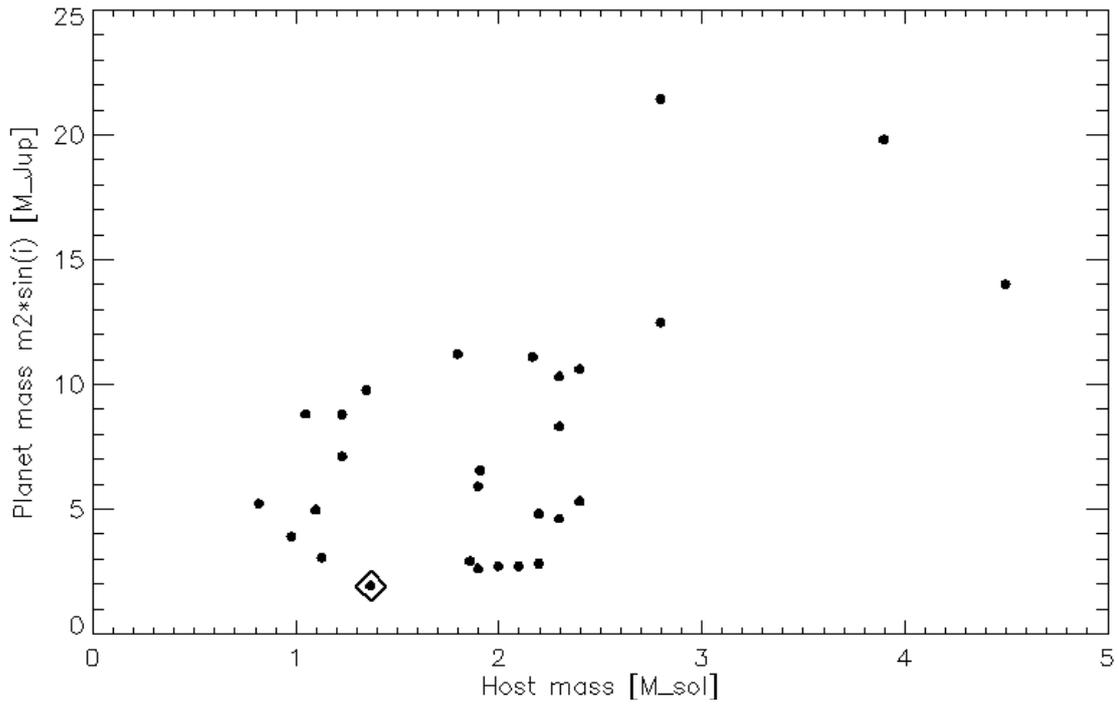


Figure 96: Distribution of minimum planet masses as function of host mass for the 29 announced systems around giant stars, and the location of HIP31592, highlighted by a diamond.

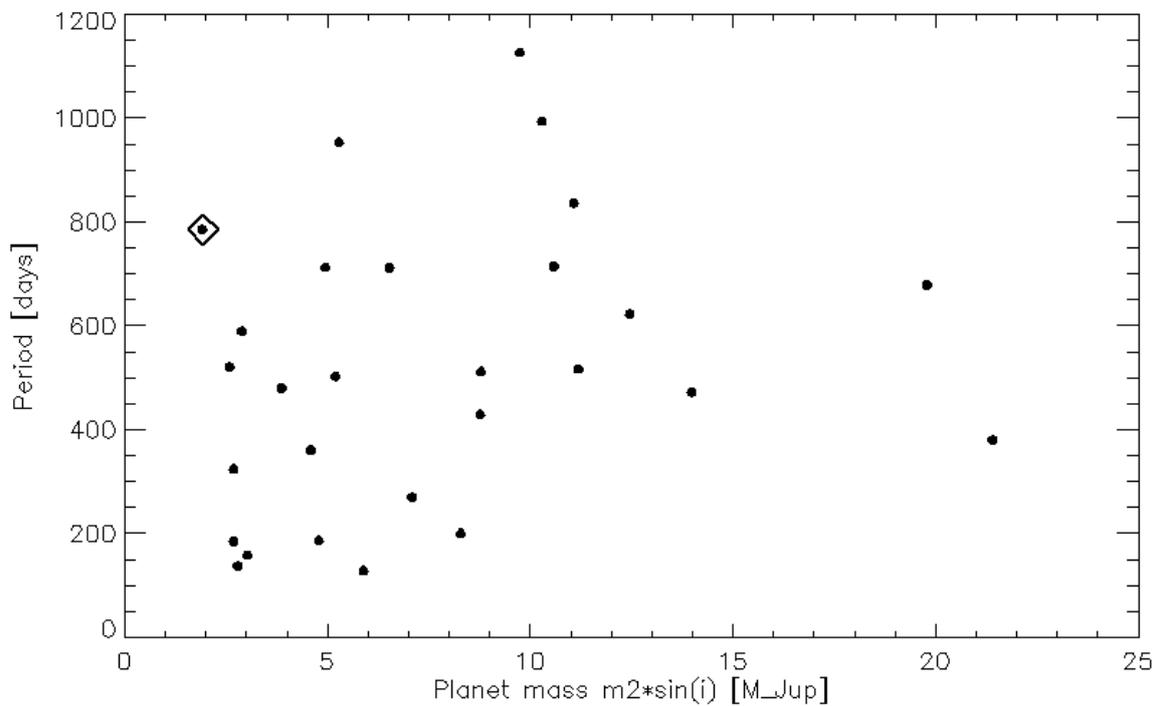


Figure 97: Plot of period versus minimum companion mass for giant star systems with known planets. HIP31592 is indicated as above.

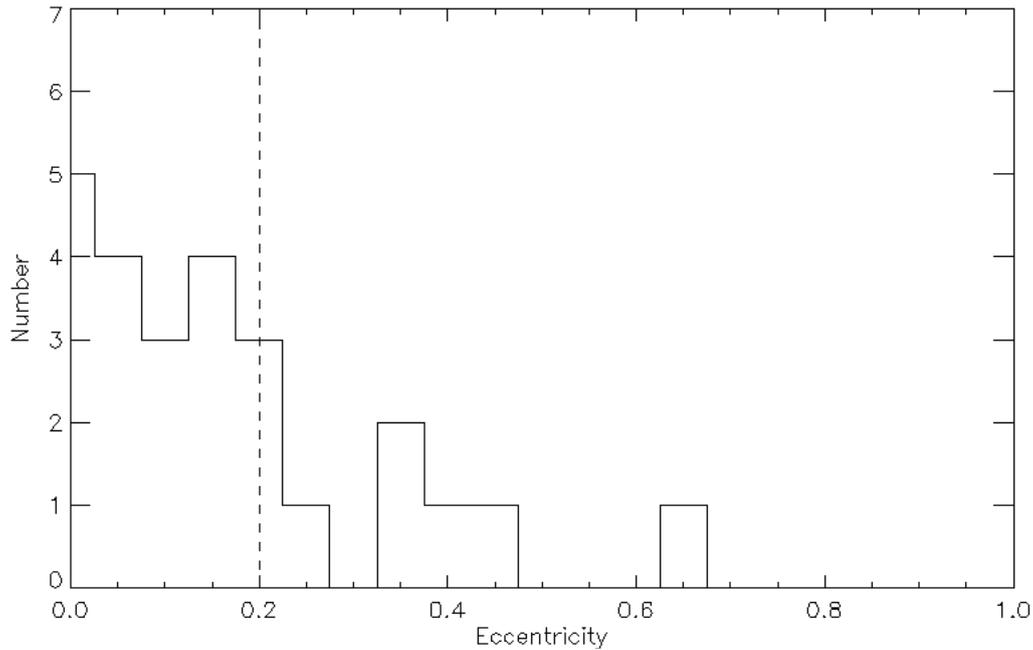


Figure 98: Histogram of eccentricities for planetary systems around giant stars. The value for the planet around HIP31592 is shown with the dashed line.

7.6 Conclusions

A series of spectral observations of HIP31592 were taken at Lick Observatory between 2000 and 2010. The measured radial velocities indicate the presence of a companion with a minimum mass of $1.92 \pm 0.03 M_J$ on an orbit of 785 days. This is the lowest mass companion detected in an orbit around a giant star to date. From the non-detection of this body in the Hipparcos data, a lower limit on the inclination of 9 degrees is found. This gives a maximum mass of $12 M_J$, indicating that it is most likely to be a planet, rather than a brown dwarf. The planet is at a distance of 1.85 AU from the star on an orbit with an eccentricity of 0.2. This is comparable to the values of the other known systems. Possible alternative explanations of the found periodic shift in radial velocity are pulsations of the stellar photosphere or the effect of star spots on the spectrum. These effects have been discussed extensively in this context (Reffert & Quirrenbach 2006, Hekker et al. 2006), and seem to be unlikely, given the large period of more than two years. Follow up observations in a different spectral band (infrared) could rule out these alternative scenarios.

I examined the residuals of this fit for hints of further companions. A harmonic analysis with a Lomb Scargle periodogram shows a second, faint signal at about 1200 days. The best possible fit indeed finds an orbit with 1077 days, lowering the residual scatter minimally. Given the intrinsic jitter of the target sample, no clear conclusion can be derived from this. However, this suggests a more intense follow up observation strategy with a higher observing cadence, ideally evenly spaced at about one data point per week. Such a strategy can help to eventually uncover a faint signal from the noise.

This discovery highlights again the advantages of a long baseline program for detecting exoplanets with the Doppler method.

References

- Arsenault, R., et al. 2006, *The Messenger*, 123, 6
- Arsenault, R., et al. 2008, *Proc. SPIE* 7015, 701524
- Baade, D. 1996, *Reviews in Modern Astronomy*, 9, 95
- Babcock H. W. 1953, *PASP*, 65, 229
- Bean, J. L., Seifahrt, A., Hartman, H., Nilsson, H., Wiedemann, G., Reiners, A., Dreizler, S., Henry, T. J. 2010, *ApJ*, 713, 410
- Beaulieu, J.-P., et al. 2006, *Nature*, 439, L437
- Benedict, G. F., et al. 2002, *ApJ*, 581, L115
- Bennett, D. P., Anderson, J., Bond, I. A., Udalski, A., Gould, A. 2006, *ApJ*, 647, 2, L171 – L174
- Blackwell, D. E., & Lynas-Gray, A. E. 1998, *A&AS*, 129, 505
- Bond, I. A., et al. 2004, *ApJ*, 606, L155 – L158
- Bordé, P., Coudé du Foresto, V., Chagnon, G., Perrin, G. 2002, *A&A*, 393, 183
- Bouchy, F., Pepe, F., & Queloz, D. 2001, *A&A*, 374, 733
- Boyer, C., Ellerbroek, B., Gedig, M., Hileman, E., Joyce, R., & Liang, M. 2008, *Proc. SPIE*, 7015, pp. 70152N-70152N-12
- Butler, R. P., Marcy, G.W., Williams, E., McCarthy, C., Dosanjuh, P., & Vogt, S. S. 1996, *PASP*, 108, 500
- Campbell, B., Walker, G. A. H., Yang, S. 1988, *ApJ*, 331, 902
- Cenarro, A.J., Peletier, R.F., Sanchez-Blazquez, P., et al. 2007, *MNRAS* 374 664
- Chandrasekhar, S. 1950, *Radiative transfer*, Oxford: Clarendon Press
- Charbonneau, D., et al. 2009, *Nature* 462, 891-894
- Cuby, J.-G., et al. 2008, *Proc. SPIE*, 7014, pp. 70141K-70141K-11
- da Silva, L., et al. 2006, *A&A*, 458, 609
- Doppler, C. 1842, *Über das farbige Licht der Doppelsterne und einiger anderer Gestirne des Himmels*, Prag
- ESA, Perryman, M. A. C. 1997, *The Hipparcos and Tycho Catalogues*, ESA SP-1200

- Esposito, S., et al. 2003a, Proc. SPIE, 4839, 164
- Esposito, S., Tozzi, A., Puglisi, A., Fini, L., Stefanini, P., Salinari, P., Gallieni, D., & Storm, J. 2003b, Proc. SPIE, 5169, 149
- Fried, D. L. 1994, in *Adaptive Optics for Astronomy*, D. M. Alloin and J. M. Mariotti eds., pp. 25-57, NATO ASI Vol. 423, Dordrecht: Kluwer
- Fried, D. L. 1965, J. Opt. Soc. Am., 55, 1427–1435
- Frink, S., Quirrenbach, A., Fischer, D., Röser, S., & Schilbach, E. 2001, PASP, 113, 173
- Frink, S., Mitchell, D. S., Quirrenbach, A., Fischer, D. A., Marcy, G. W., & Butler, R. P. 2002, ApJ, 576, 478
- Fugate, R. Q. et al. 1991, Nature, 353, 144-146
- Gould, A., 2000, ApJ, 542, 2, pp. 785-788
- Griffin, R., & Griffin, R. 1973, MNRAS, 162, 255
- Happer, W., MacDonald, G., Max, C., Dyson, F. 1994, J. Opt. Soc. Am., 11, 263-276
- Hardy, R. S. 1998, *Adaptive Optics for Astronomical Telescopes*, pp. 438, New York: Oxford University Press
- Hatzes, A.P. et al. 2003, ApJ, 599, 2, 1383
- Hayano, Y., Iye, M., Takami, H., Takato, N., Gaessler, W., Minowa, Y., Wizinowich, P., & Summers, D. 2003, PASP, 115, 1419
- Hekker, S., Reffert, S., Quirrenbach, A., Mitchell, D. S., Fischer, D. A., Marcy, G. W., & Butler, R. P. 2006, A&A, 454, 943
- Henry, G. W., Marcy, G. W., Butler, R. P., & Vogt, S. S. 2000, ApJ, 529, L41
- Jacob, W. S. 1855, MNRAS 15, 228
- Kolmogorov, A. N. 1941, Comptes rendus (Doklady) de l'Académie des Sciences de l'U.R.S.S., 30, 301–305
- Lenzen, R., et al. 2003, Proc. SPIE, 4841, 944
- Lineweaver, C. H., Fenner, Y., Gibson, B. 2004, 303, 59-62
- Marcy, G. W., & Butler, R. P. 1992, PASP, 104, 270
- Marcy, G. W., Butler, R. P. 1995, Bulletin of the American Astronomical Society, Vol. 27, p.1379
- Massarotti, A., Latham, D. W., Stefanik, R. P., & Fogel, J. 2008, AJ, 135, 209

- Mayor, M. & Queloz, D. 1995, *Nature*, 378, 6555, 355
- Mayor, M., et al. 2003, *The Messenger*, 114, 20
- Meyers Konversations-Lexikon (German encyclopaedia), 1888
- Mitchell, D. S., Frink, S., Quirrenbach, A., Fischer, D. A., Marcy, G. W., Butler, R. P. 2003, *Bulletin of the American Astronomical Society*, 35, 1234
- Oke, J. B. 1959, *ApJ*, 130, 487
- Paczynski, B. 1986, *ApJ*, 304, p. 1-5
- Penndorf, R. 1957, *J. Opt. Soc. Am.*, 47, 2, pp. 176 -182
- Pepe, F. A & Lovis, C. 2008, *Phys. Scr.*, Vol. 130, pp. 014007
- Perelman, Y. 1914, *Distant of Miry*, St. Petersburg: The printing house of Soykina
- Pourbaix, D. 2001, *A&A*, 369, L22
- Pourbaix, D., & Arenou, F. 2001, *A&A*, 372, 935
- Primmerman, C. A., Murphy, D. V., Page, D. A., Zollars, B. G., & Barclay, H. T. 1991, *Nature*, 353, 141
- Queloz, D., et al. 2009, *A&A* 506, 303-319
- Quirrenbach, A., et al. 1998, *Astronomical Interferometry*, Robert D. Reasenberg ed., *Proc. SPIE* Vol. 3350, p. 807-817
- Quirrenbach, A. 2006, in *Extrasolar Planets*, 2005 Saas-Fee Advanced Course 31, eds. D. Queloz, S. Udry, M. Mayor, W. Benz, Springer-Verlag: Berlin Heidelberg
- Radtke, D. & Zeitner, U. D. 2007, *Optics Express*, Vol. 15, Issue 3, pp. 1167-1174
- Reffert, S.; Quirrenbach, A. 2006, *A&A*, Vol. 449, 2, pp.699-702
- Reiners, A., Bean, J. L., Huber, K. F., Dreizler, S., Seifahrt, A., Czesla, S. 2010, *ApJ*, 710, 432
- Reynolds, O. 1883, *An experimental investigation of the circumstances which determine whether the motion of water shall be direct or sinuous, and of the law of resistance in parallel channels*, *Philosophical Transactions of the Royal Society*, 174, 935–982
- Richardson, L. J., Deming, D., Horning, K., Seager, S. and Harrington, J. 2007, *Nature Letters* 445, 892-895
- Rigaut, F. 2002, *European Southern Observatory Conference and Workshop Proceedings*, 58, 11

- Roddier, F. 1981, *Progress in Optics*, Vol.19, 281-376
- Rousset, G., et al. 2003, *Proc. SPIE*, 4839, 140
- Schlesinger, F. 1936, *Catalogue of Stellar Parallaxes*, New Haven: Yale University Press
- Schwab, C., Gaessler, W., Peter, D., Blumchen, T., Aigner, S., & Quirrenbach, A. 2010, in *Adaptative Optics for Extremely Large Telescopes*, eds. Y. Clénet, J.-M. Conan, Th. Fusco, and G. Rousset, EDP Sciences
- See, T. J. J. 1896, *AJ* 16, 17
- Serabyn, E., Mawet, D., & Burruss, R. 2010, *Nature*, 464, 1018-1020
- Setiawan, J., Pasquini, L., da Silva, L., Hatzes, A. P., von der Luehe, O., Girardi, L., de Medeiros, J. R., & Guenther, E. 2004, *A&A*, 421, 241
- Stokes, G. 1851, *On the Effect of the Internal Friction of Fluids on the Motion of Pendulums*, *Transactions of the Cambridge Philosophical Society*, 9, 8–106
- Struve, O. 1952, *The Observatory*, 72, 199
- Swain, M. R., et al. 2010, *Nature*, 463, 7281, pp. 637-639
- Tatarski, V. I. 1961, *Wave Propagation in a Turbulent Medium*, McGraw-Hill Books
- Thomas, S. 2004, in *Astronomy with High Contrast Imaging II*, Vol.12, pp. 177 – 184, EAS Publications Series
- van Dam, M. et al. 2007, *Keck AO Note #489*, 2007
- van Leeuwen, F., & Evans, D. W. 1998, *A&AS*, 130, 157
- van de Hulst, H.C. 1981, *Light Scattering by Small Particles*, New York: Dover
- Vogt, S. S. 1987, *PASP*, Vol. 99, p. 1214-1228
- Ward, P. D., Brownlee, D. 2000, *“Rare Earth: Why Complex Life Is Uncommon in the Universe”*, New York : Copernicus
- Wizinowich, P. L., et al. 2006, *PASP*, 118, 297
- Whewell, W. 1853, *Of the Plurality of Worlds. An Essay*, London: J. W. Parker and son (reissued by Cambridge University Press, 2009)
- Wolszczan, A., Frail, D.A. 1992, *Nature*, 355, 145
- Wright, J. T., et al. 2007, *ApJ*, 657, 533

Acknowledgements

Over the last 4 years, I met many people who made my work, my time and my life interesting, and sometimes fun. To all the people who appreciated my comments, gave me an opportunity to work with them, let me learn, or simply made me feel welcome, a sincere “thank you“.

I want to thank my supervisor Andreas Quirrenbach for letting me work on this topic, for his scientific guidance, his trust in my abilities when I had unusual ideas and his unwavering patience when I wanted things my way. I really appreciate his outstanding competence, as well as his bone dry humour!

My friends and colleagues at LSW, who helped me in many little ways, and made the time we shared enjoyable. Christoph’s and Julian’s office is the most fun place on the hill! Max Camenzind, who made me believe I can and want to do Astronomy in Heidelberg. Most of the actual Astronomy I know I learned from him.

Sabine Reffert for working with me on the K-giants, and her patience with my special preferences concerning computers. I know I’ve often been complicated...

Debra Fischer, who taught me so much, and enabled so much. Thank you for your all your help and confidence in me!

My friend Julien. Thanks for the good times we had, and to the great times ahead!

Sebastian Rabien, Lorenzo Busoni, Richard Green, Michael Lloyd Hart, Jesper Storm, and all the other colleagues in the ARGOS project, for a great working atmosphere.

Diethard Peter and Wolfgang Gässler, for their help with the CalUnit, and their patience with my scheduling, documentation...

Andrew Rakich, for teaching me optics “the right way“, hosting me, sharing many laughs in a short time, and being a friend. Anteaters will never be the same!

My colleagues at MPIA, who always had time, advice and help. The colleagues from the IR Space Astronomy group; Armin, Uli, Ralph and Alexandra, who were great company. Armin Böhm and the staff of the MPIA workshop for many explanations to my persistent self, and quite a number of metal pieces looking more like a swiss cheese. Tom Herbst for his advice and his time as a thesis comitee member. Christian Fendt for his help with all things related to IMPRS or the university, and being my third examiner (and the IMPRS BBQs, too).

To my dearest friends from IMPRS, Marcello, Bagmeet & Swapna, Giovanna, Claudia and Giulia, Cassie and Surhud and Anu, and Kris. You made me feel at home, and when thinking about the time in Heidelberg I will always remember the awesome times we had together! I hope to see you often..

To my friends from home, Steffi, Jens, Daniel, Andreas, Susannah, Maria, Matthias, Annelen, Robin and Simona, Katja and Ali, and Bernardin, Hermann and Werner, and Sarah and Ryan in SF, for being just that – my dear old friends!

To Carola. For so much great music! For the dancing. For your support, and trust, and friendship, and love. For being there whenever I needed someone . Without you I wouldn't have carried on. (Tier!)

Ich möchte meinen Eltern danken, für Ihre immerwährende Unterstützung über so viele Jahre, Verständnis für meine Wünsche, dafür, daß ich ein Zuhause habe, in das ich immer zurück kann. Ich habe das alles wegen Euch geschafft. Und ich danke meiner ganzen Familie, besonders meiner Oma Anneliese und meinem Opa Emil, die immer für mich da waren.

My dearest Ros. Who helped me so much and so often with all this. For your love, and your company. For those great days on the balcony, for trying out german cake, for long nights in front of the computer, for patience and advice and your stubbornness, for putting up pictures and planning travels, for many great dinners shared, for showing me my first real leopard. For being sure. Thank you!