

Optical design and quantitative perturbation analysis of a nulling interferometer

AE4020

H. Vlot

Delft University of Technology

Optical design and quantitative perturbation analysis of a nulling interferometer

by

H. Vlot

to obtain the degree of Master of Science

at the Delft University of Technology,

to be defended publicly on Thursday May 23, 2024 at 14:00.

Student number: 4363523
Project duration: May, 2023 – May, 2024
Thesis committee: Dr. J. J. D. Loicq, TU Delft, Supervisor
Dr. ir. R. Saathof, TU Delft, Chair
Dr. A. J. L. Adam, TU Delft, External

Cover: Artist rendition of ESA's Darwin nulling interferometer
Style: TU Delft Report Style, with modifications by Daan Zwan-
eveld

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

Preface

Dear reader,

It has been quite the journey, this master thesis. When I first met my supervisor, Jérôme Loicq, I knew I was in for a formidable engineering challenge with many interesting aspects. This project brought together a deeply rooted fascination for exoplanets, physics and engineering. Yet, with many interesting aspects come many intricate complications, some of which I have encountered on the way and kept my head spinning due to all the complex considerations to be made.

Needless to say, I would not have been able to do this on my own and thus would like to express my gratitude towards a number of people. First and foremost, I want to thank Jérôme, for being an excellent supervisor and guide through this sometimes confusing 'proof of competency' stage of my student career. I remember many a meeting where I went in pessimistic about my own capabilities and methodology, yet after each meeting with Jérôme, I came out inspired and motivated to do the best I can to move forward in the process. Second, I would like to thank Thomas van der Sijs and Luuk Zonneveld, for being tireless in listening to my many questions and providing me with much needed outside perspectives. Additionally, I want to express my appreciation for Jan de Graaff's upbeat spirit and memorable one-liners ("Een PhD is geen sprint of een marathon, Henri, het is een loopgravenoorlog. "). Last but not least, I am sincerely grateful to Lidija Nikolic and Roland Horsten for providing me with great facilities to study at the physics department.

As for anyone who is tasked to read through my thesis, I would like to end with the following quote:

I have made this longer than usual because I have not had time to make it shorter.
- Blaise Pascal, 1657

H. Vlot
Delft, May 2024

Summary

Finding signs of life on planets outside of Earth has been the holy grail of astronomy. Answering this question of 'are we alone in the universe' requires not only detection of exoplanets, but also characterising their atmospheres. Detections of exoplanets has grown exponentially since the first discovery in 1995, which can be attributed to the refinement of the radial velocity and transit method. These methods are currently biased towards massive (super-Earths and heavier) planets that orbit close to their host. In order to find an Earth-twin in the habitable zone and analyse the composition of its atmosphere, different detection methods are needed. One of the most promising methods is nulling interferometry, which can obtain both a high angular resolution, as well as a large enough rejection to reduce the bright starlight that outshines the planet by a factor 10^{10} in the visible spectrum. This contrast is reduced in the mid-infrared (MIR) to 10^7 , as the planet thermal radiation has a lower equilibrium temperature than the reflected solar radiation. Another benefit of observing a potential Earth-twin in the MIR is the abundance of molecular biosignatures in the absorption spectrum of the observed atmosphere. As the Earth's atmosphere absorbs partly or completely most of the MIR spectrum, a space based nulling interferometer is required to fully utilise the potential of this technique. Although this technique has been successfully demonstrated on ground, atmospheric turbulence limits the coherent observation time, as well as the attainable null depth. Several space based nulling interferometers were proposed in the late 1990s, by both ESA and NASA. However, due to budgetary limitations, none of these missions ever materialised. During the height of this period, many works were published on possible configurations for a formation flying interferometer, the effect of noise sources on the instrument and potential science return of a nulling interferometer. In all of these works, the objective was to understand the nulling interferometer on a system level, or test out specific components, however, considerations on the telescope design are lacking in the literature. This brings up the research question of the thesis: "To which extent does the choice of telescope design influence the stability and performance of a space based, mid-infrared, nulling interferometer?"

To address the research question, nulling interferometers with both on-axis and off-axis telescopes were designed in Code V and analysed in Matlab. For each of these models, the telescope optics were perturbed and the effect on the null depth analysed. The off-axis Cassegrain was shown to be more sensitive under perturbation, compared to the nulling interferometer with an on-axis Cassegrain design. Lastly, three telescope designs were compared under perturbation: the Cassegrain, Ritchey-Chrétien (RC) and Gregorian. The Gregorian was shown to be the most sensitive, with a null depth loss of one order of magnitude more in some cases, compared to the Cassegrain and RC. The RC had the best performance and stability in most perturbation cases, making it a promising candidate for a space-based, mid-infrared, nulling interferometer.

Contents

Preface	i
Summary	ii
Nomenclature	xviii
1 Introduction	1
1.1 Introduction to detecting exoplanets	1
1.2 Nulling interferometry	2
1.3 Towards a space based nulling interferometer	3
1.4 Research question	5
1.5 Thesis outline	5
2 Theoretical background - Optics	6
2.1 Electromagnetic waves	6
2.2 Polarised light	7
2.3 Fresnel equations	8
2.4 Zernike polynomials	10
2.5 Aberrations	11
2.5.1 Spherical aberration	12
2.5.2 Astigmatism	12
2.5.3 Coma	13
2.5.4 Field curvature	13
2.5.5 Distortion	14
2.6 Chapter summary	14
3 Theoretical background - Interferometry	15
3.1 Young's experiment	15
3.2 Stellar interferometry	17
3.3 Nulling interferometry	18
3.3.1 General principle	18

3.3.2	Aperture configurations	19
3.3.3	Practical implementation	21
3.4	Transmission map	22
3.5	Chapter summary	23
4	Optical design	25
4.1	Basic properties of optical systems	25
4.2	Requirements	25
4.3	On axis telescope designs	27
4.3.1	Newtonian telescope	27
4.3.2	Cassegrain	27
4.3.3	Gregorian	28
4.4	Off-axis telescope designs	29
4.5	Beam collimation	32
4.6	Delay line	34
4.7	Beam combiner	35
4.7.1	Beam combiner types	35
4.7.2	Bulk optics beam combiners	36
4.7.3	Modified Mach-Zehnder beam combiner	36
4.8	Chapter summary	37
5	Methodology	38
5.1	Modelling split apertures and beam combination	38
5.2	Beam propagation	38
5.2.1	FFT Beam Propagation	39
5.2.2	Beam Synthesis Propagation (BSP)	40
5.3	Deriving transmission map	41
5.4	Sensitivity analysis	42
5.4.1	Tilt	42
5.4.2	Translation	43
5.4.3	Surface shape	44
5.4.4	Null depth using finite sampling	44

5.5	Performance analysis	45
5.5.1	Spot diagram	46
5.5.2	Exit pupil map	47
5.5.3	Modulation Transfer Function	47
5.5.4	Diffraction limit	47
5.6	Chapter summary	48
6	Results	50
6.1	Optical model	50
6.1.1	Aperture	50
6.1.2	Collimator	52
6.1.3	Beam combiner	55
6.2	Nulling interferometer optical model	58
6.2.1	Bracewell interferometer	58
6.2.2	Double Bracewell interferometer	59
6.3	Simulating transmission maps	59
6.4	Intermezzo: Nulling without an optical model	61
6.5	Perturbation analysis - Off axis designs	65
6.5.1	Axial translation perturbation	66
6.5.2	Lateral translation perturbation	67
6.5.3	Collimator focal length dependence	69
6.6	On axis design	76
6.7	Classical telescope design comparison	80
7	Recommendations	84
8	Conclusion	86
	References	88
A	Appendix A: Zernike polynomials	94
B	Appendix B: Maxwell and the wave equation	95
C	Appendix C: Modified Mach Zehnder interferometer	96
D	Appendix D: Unfinished or unverified results	100
D.1	Three aperture design	100

D.2 Monte Carlo simulation 100

E Appendix E: Null depth calculation with finite sampling 103

F Appendix F: Field of View determination 105

G Appendix G: Classical telescope design null depth comparison 107

List of Figures

1.1	Exoplanet census of current confirmed discoveries, separated into different detection methods.	2
1.2	A simulated emission spectrum of a star and orbiting planet. Due to thermal radiation of the planet, an extra signature shows up in the radiation spectrum. Adapted from [14].	3
1.3	The transmissive windows of the Earth's atmosphere. Although there are some small transmissive windows between 1 and 10 μm , they are relatively opaque and need long integration times or extreme adaptive optics to be usable for astronomical observations.	4
2.1	Definitions of the s- and p- components of the reflected and transmitted field.	9
2.2	(a) Amplitude coefficients for the reflected and transmitted wave, separated into the perpendicular (s) and parallel (p) component, as function of the angle of incidence. (b) Phase of the complex amplitude coefficients.	10
2.3	Visual representation of the first 10 Zernike polynomials. First row: piston. Second row (from left to right): vertical tilt, horizontal tilt. Third row: vertical astigmatism, defocus, oblique astigmatism. Fourth row: vertical trefoil, vertical coma, horizontal coma, oblique trefoil.	11
2.4	Spherical aberration resulting from a flat field incident on a biconvex lens. Due to the difference in angle of incidence along the lens height, the focal plane shifts as function of the incident ray height. The resulting spot on the detector plane has an elliptical shape.	12
2.5	A lens with astigmatism. The rays in the meridional plane are colored green and come to a focus at the image plane. The rays in the sagittal plane are colored blue and come to a focus before the image plane.	13
2.6	A singlet which shows field curvature. As the field angle increases, the focal length decreases.	13
2.7	(a) Barrel distortion. (b) Pincushion distortion	14

3.1	Schematic representation of Young's double slit experiment. On the left, there is a screen with two slits, spaced by a distance d . The considered rays from the centre of the slits are angled with an angle θ , which results in a path length difference ΔR . The intensity, and thus the locations of the dark and bright fringes on the screen, depends on the angle θ	16
3.2	(a) A Michelson stellar interferometer with baseline B and aperture diameter D . (b) The fringe visibility as function of angular separation α , shown with the red line. The blue line indicates the envelope due to the single aperture with diameter D . Adapted from [32].	17
3.3	General schematic representation of a Bracewell nulling interferometer. The beam combiner in the middle is preceded by delay lines, which are used to equalize the path length between both apertures. Adapted from [34]	19
3.4	Left: Transmission map with a star in the centre and three planets with their corresponding trace through the fringes. (green / red / blue dots and circles). Right: The resulting modulation from rotating the transmission map. Adapted from [35]	19
3.5	Selection of architectures considered in [19], grouped by their null order p	20
3.6	A central slice of the transmission map of a nulling interferometer. The effect of a higher null order is shown on the right, where the dark fringe is wider compared to the one on the left. Adapted from [36].	21
3.7	Functional diagram of a potential nulling interferometer with two apertures.	22
3.8	Simulated monochromatic transmission maps, based on Equation 3.14, of (a) 2 apertures (b) 3 apertures and (c) 4 apertures, using a baseline of 10 meters and wavelength of $10 \mu\text{m}$	24
4.1	Comparison between the fundamental mode of a single mode fibre (red) and the electric field at the focus of a telescope without a central obscuration. Adapted from [53].	29
4.2	A conventional off-axis telescope system, using a eccentric section of the conventional on-axis system, such that the optical axis of both parent mirrors co-align. Adapted from [30].	30
4.3	Distances and angles relevant to the off-axis mirror system. l'_1 is the primary mirror (M1) focal length, l_2 is the secondary mirror (M2) focal length, l'_2 is the telescope back focal length, i_1 is the primary mirror tilt angle and i_2 is the secondary mirror tilt angle. Adapted from [30].	31
4.4	Schematic of the secondary mirror and collimator. β is the angle of incidence/reflection. γ is the angle suspended by the optical axis ray after reflection off the secondary mirror and the global horizontal. The tilt angle α_c is the angle between the global horizontal and the vertex.	33

4.5	Dependencies of the collimator surface parameters on the primary mirror tilt angle. (a) Collimator tilt angle, as function of primary mirror tilt angle. (b) Collimator off axis angle, as function of primary mirror tilt angle. (c) Collimator radius of curvature as function of primary mirror tilt angle.	34
4.6	(a) Rooftop delay line. (b) Catseye delay line, consisting of a parabolic primary mirror and a fold flat.	34
4.7	Symmetric beam combiner designs, based on classic interferometers. The bold arrows indicate the balanced output, while the dotted arrows indicate the unbalanced outputs. Schematic taken from [65].	36
4.8	A classic (left) and modified MZ (right) beam combiner, with two input beams, A and B. The dashed line indicates the half transmissive coating on the beam splitter.	37
5.1	Example of a zoomed system. (a) Zoom 1, where the second surface of the beamsplitter is reflective. (b) Zoom 2, where the second surface is transmissive. (c) A render of both zooms overlaid.	39
5.2	A schematic representation to show the difference between a non-zoomed interferometer model and a zoomed interferometer model. On the left, a single wavefront, with wave vector \vec{k}_1 is incident on the interferometer model, consisting of two apertures. Since the model is not separated into zooms, there is only one entrance pupil, with an Entrance Pupil Diameter (EPD) as large as the baseline. Because of this, $\phi_1 = \phi_2$. On the right, a zoomed interferometer model is shown. Due to the 'zooming', the entrance pupil is split into two, and thus two wavefronts are incident, with wavevector k_1 and k_2 . To make sure both wavefronts have the same phase, a phase shift $\delta\phi$ has to be applied.	40
5.3	A screenshot of the BSP window in Code V. At the input beam tab, a choice can be made between a spherical or Gaussian beam, or a field based on external data. Under the propagation controls, the number of beamlet rings can be set, as well as settings for resampling the field at a certain surface.	41
5.4	Base case of the telescope configuration, without any perturbation. The blue structures represent the (orientation of the) telescope assembly, consisting of the primary mirror, secondary mirror and collimator. The dashed dotted line denotes the line of sight with the target, whereas the dashed line denotes the line of sight of a given surface. The solid line is perpendicular to the line of sight of the spacecraft.	42
5.5	Four cases of tilt perturbations can be considered. (a) One arm of the interferometer is tilted by an angle α . (b) Both arms are tilted by an angle α , but the spacecraft is still pointed correctly towards the target. (c) Both arms are tilted by an angle α and the spacecraft is not pointed correctly. (d) Both arms are tilted, but not necessarily by the same angle α	43

5.6	Left: translation perturbation along the global x, y and z axis, denoted by the Code V commands (XDE, YDE, ZDE). Centre: translation perturbation along axes local to the optical surface. Right: Perturbation along optical axis ray (OAR).	44
5.7	Estimated dependence of axial mirror displacement on the image blur size, β .	45
5.8	Three surface shape parameters will be considered for the perturbation analysis. Left: Radius of curvature (RDY), centre: Conic constant (SCO C1), right: Off axis angle (SCO C2).	45
5.9	Spot diagrams of a lens with astigmatism, simulated with Code V.	46
5.10	Spot diagrams of a lens with coma, simulated with Code V.	46
5.11	Spot diagrams of a lens with spherical aberration, simulated by Code V. . . .	46
5.12	Top: an object is imaged with 100% contrast in the object plane. After the object is imaged, the contrast is reduced to 90%. Bottom: When the spatial frequency is increased, the contrast in the image plane is reduced further, due to the finite extent of the PSF.	48
6.1	(a) Diagram of an off-axis Cassegrain telescope, with a back focal length of 2000 mm, a primary mirror focal length of 2000 mm and tilt angle of 15° . (b) To collimate the off-axis Cassegrain, a single off-axis parabola is used. The collimated beam is then re-imaged with a perfect lens, so that the optical performance can be analysed.	51
6.2	Spot diagrams superimposed on the field of view, which is demarcated by the dashed square. The blue circles demarcate the Airy disk, which has a radius of $39.04 \mu m$ at a wavelength of $4 \mu m$. Note that the size of the Airy disk circles are not to scale, compared to the FoV, for the sake of visualisation.	51
6.3	Diffraction based MTF for the off-axis Cassegrain, evaluated from 4 to $8 \mu m$. The red dashed graphs show the diffraction limited performance, whereas the solid lines show the simulated performance for the off axis diffraction MTF at a field angle of $(3', 3')$	52
6.4	(a) Spot diagrams for the off-axis Cassegrain with collimator, simulated with a wavelength of $4 \mu m$. A perfect lens was used to image the pupil plane. The top spot diagram, showing the distribution for the $(3', 3')$ field angle, could be interpreted as defocus, or astigmatism. Note that the Airy disk is inside the spot diagram. (b) The spot diagrams, evaluated outside of the focal plane. . .	53
6.5	The diffraction based MTF for the off-axis Cassegrain with collimator, evaluated from 4 to $8 \mu m$, at a field angle of $(3', 3')$. When comparing the diffraction limited MTF and the one at $(3', 3')$, the degradation in performance becomes apparent, as no modulation is left after 6 cycles/mm.	54

6.6	(a) Simulated intensity, using BSP, at the detector plane of the off-axis Cassegrain and collimator, for the on axis field. (b) Simulated intensity for the off axis field.	54
6.7	The modelled beam splitter, with one compensator plate, to compensate for the optical path. The reflected path reflects off the backside of the beam splitter. Both the beam splitter and compensator plate are made of ZnSe. In both renders, two zooms are overlaid for visualisation purposes. The grey ring in the right figure is the aperture stop, while the dark grey circular surfaces represent the detector plane. Blue surfaces are refractive surfaces and light grey surfaces are reflective.	55
6.8	Phase and amplitude of the transmitted and reflected complex field, as modelled by BSP, at the detector plane with a beam linearly polarised in the x direction (perpendicular to plane of incidence).	56
6.9	Phase and amplitude of the transmitted and reflected complex field, as modelled by BSP, at the detector plane with a beam linearly polarised in the y (parallel to plane of incidence) direction.	56
6.10	Schematic of a modified Mach-Zehnder beam combiner. It consists of 8 zooms, which are overlaid to make it appear as one component. The top inset shows the destructive output and the right inset shows the constructive output, both simulated with BSP. The input wave is a monochromatic plane wave at $10 \mu m$.	57
6.11	Side view of the 2 aperture nulling interferometer with a baseline of 15 meters. The destructive outputs are located on the top side of the MMZ BC, while the constructive outputs are located on the right side of the BC.	58
6.12	Intensity fields at the detector plane, as simulated by BSP. Upon close inspection of the fields, it can be seen that they are mirrored with respect to the horizontal axis.	59
6.13	(a) Destructive field of the nulling interferometer in the pupil plane, where neither of the aperture beams was flipped. (b) Destructive field in the pupil plane, resulting from rotating one of the beams by 180 degrees.	59
6.14	(a) Front view of the 4 aperture, off axis Cassegrain nulling interferometer. (b) Perspective view.	60
6.15	(a) Transmission map for the nulling interferometer with a baseline of 4.1 m and monochromatic incident field, showing the angular separation of 251 mas between the dark and bright fringe. (b) When using polychromatic light, the fringe visibility is reduced near the edges of the field of view.	60
6.16	(a) Monochromatic transmission map of a 4 aperture, double Bracewell nulling interferometer (b) Polychromatic transmission map, with a $2 \mu m$ bandwidth.	61

- 6.17 Several ways of degrading the null are shown here, as described by the model in Equation 6.1 and Equation 6.3. (a) Destructive field, after a certain translation δx has been applied. (b) Destructive field, showing a ring of non-zero intensity, due to a radius mismatch ($r_1 \neq r_2$). (c) Destructive field which has been degraded uniformly, due to an amplitude mismatch ($\delta A \neq 0$). (d) Destructive field, after astigmatism has been applied to one of the complex amplitudes ($\delta\varphi = Z_2^2$, see Table 2.2). 62
- 6.18 (a) Null depth loss as a result of piston in one of the beams. At an OPD of 1/1000th of a wave, the null depth loss exceeds the threshold of $1 \cdot 10^{-5}$. (b) The null depth loss due to a mismatch in amplitude between the two interfering beams. 63
- 6.19 (a) Null depth loss due to beam shear, plotted on a log-log scale. (b) 63
- 6.20 Null depth loss due a selection of aberrations, as function of the (a) P-V aberration magnitude and (b) RMS aberration magnitude. 64
- 6.21 Comparison of the null depth loss due to beam shear, as derived by Hyde et al. and simulated by Code V. For small displacements, the Code V model underestimates the null depth loss. 65
- 6.22 (a) Axial translation of the focal length for different values of the collimator focal length. The smaller focal length collimator clearly shows a higher sensitivity to axial perturbations, compared to the longest focal length. A tilt angle of -40.34 was used on the collimator. (b) Axial translation of the focal length for different values of the collimator focal length, using a collimator tilt angle of -39.1. . . . 66
- 6.23 Null depth as function of lateral translation perturbation, applied to each telescope surface, along the x, y and z axis. Δx and Δy show similar behaviour, with the primary mirror being the most sensitive, then the secondary mirror and lastly the collimator. For Δz perturbations, the collimator is again the least sensitive and the primary and secondary mirror have similar sensitivities. 67
- 6.24 Three graphs showing the effect of perturbations applied to the collimator. For both XDE and YDE perturbations, the null depth change does not depend significantly on neither the primary mirror tilt angle, nor the collimator focal length. For ZDE misalignment, there is some dependence on the primary mirror tilt angle, with a lower tilt angle causing the least change in the null depth. . . 69
- 6.25 (a) Nominal exit pupil map of the nulling interferometer. (b) Exit pupil map resulting from a 20 μm perturbation along the y axis, with a P-V WFE of $5.38 \cdot 10^{-3}$ waves and RMS WFE of $1 \cdot 10^{-3}$ waves. (c) Exit pupil map resulting from a 20 μm perturbation along the z axis, with a P-V WFE of $2.83 \cdot 10^{-3}$ waves and RMS WFE of $7 \cdot 10^{-4}$ waves. 70

6.26	(a) Geometrical representation of a perturbation δz along the z axis, as viewed in the y - z plane. l_F is the nominal distance to the focus and l'_F is the distance to the focus after misalignment. θ_0 (θ'_0) is the angle subtended by the surface vertex and the focus. (b) The geometry of a δx perturbation, as seen from the x - z plane. For both schematics, the solid surface indicates the unperturbed position, whereas the dashed surface is the perturbed position.	71
6.27	Apparent field angle as function of δx or δz . For the given range, $\delta\theta_x$ is always larger than $\delta\theta_z$	72
6.28	Left: Change of the null depth, compared to the nominal case, due to rotations around the x axis. Centre: Null depth change due to rotations around the y axis. Right: Null depth change resulting from perturbations of the off axis angle	73
6.29	Comparison of null depth change due to RDY (radius of curvature) and SCO C1 (conic constant) perturbations.	74
6.30	Side view schematic of the on-axis Cassegrain nulling interferometer model. . .	76
6.31	Left: null depth as a result of XDE perturbation applied to all three telescope surfaces. Middle: null depth as a result of YDE perturbation applied to the telescope surfaces. Right: null depths due to ZDE perturbation.	77
6.32	Left: ADE (alpha tilt) perturbation comparison between the on axis and off axis Cassegrain telescope design, shown for all three telescope optical surfaces. Middle: BDE (beta tilt) perturbation comparison. Right: SCO C2 (off axis angle) perturbation comparison.	77
6.33	Left: Null depth change due to RDY (radius of curvature) perturbation, shown for the primary, secondary and collimator surface. Right: Null depth change due to SCO C1 (conic constant) perturbation.	78
6.34	(a) Exit pupil map of the on-axis Cassegrain, with an SCO C1 perturbation of $1 \cdot 10^{-4}$ applied to the primary mirror, having a P-V WFE of 6.6E-4 waves. (b) Exit pupil map of the off-axis Cassegrain, with an SCO C1 perturbation of $1 \cdot 10^{-4}$ applied to the primary mirror, with a P-V WFE of 4.8E-2 waves. . . .	79
6.35	(a) Exit pupil map of the on-axis Cassegrain, as a result of an SCO C1 perturbation of $1E-4$ on the secondary mirror, showing a P-V WFE of 8E-5 waves. (b) Exit pupil map of the off-axis Cassegrain, as a result of an SCO C1 perturbation of $1E-4$ on the secondary mirror, with a P-V WFE of 5.7E-3 waves. The exotic shape of the exit pupil map can be attributed to the large astigmatism and trefoil component of the wavefront aberration.	79
6.36	The radar plot consists of three axes, the data points of which are taken from the above plot, as indicated.	81
6.37	Translation perturbation of the secondary mirror. The comparison is made between an off-axis Cassegrain, RC and Gregorian design. All null depths calculated with a $10 \mu m$ central wavelength and bandwidth of $1 \mu m$	82

6.38	Rotation perturbation of the secondary mirror. The comparison is made between an off-axis Cassegrain, RC and Gregorian design.	82
6.39	Surface shape perturbation of the secondary mirror. The comparison is made between an off-axis Cassegrain, RC and Gregorian design.	83
C.1	Schematic of all zooms used to make a 2 input, 4 output MMZ.	97
C.2	Phase for each beam interacting with the MMZ beam combiner	98
C.3	Intensity for each beam interacting with the MMZ beam combiner	99
D.1	Side and top view of a three aperture telescope design.	101
D.2	Flow chart of the Monte Carlo simulation process.	101
D.3	Comparison of the off-axis Cassegrain and Gregorian telescope design.	102
E.1	Left: Transmission map of a 2 aperture Bracewell interferometer, where the minimum null depth is shifted to the right of the 0 field angle line, due to a translation in the primary mirror. Right: Horizontal slice of the transmission map, where the red line indicates the null depth without any perturbations applied to the optical surfaces. The blue line corresponds to the horizontal slice of the shifted transmission map.	104
E.2	The blue line shows the calculated null depth when a relatively low sampling rate is used for the transmission map, as function of the axial displacement. When the transmission map is resampled around the minimum, which is shown in red, the null depth shows a less erratic trend.	104
F.1	(a) Spot diagrams for the on-axis and off-axis field, evaluated at $4 \mu m$. (b) Diffraction based MTF of the considered telescope design, evaluated from 4 to $8 \mu m$ and a field angle of ($18''$, $18''$).	105
F.2	Exit pupil map of the off-axis Cassegrain simulated at a field angle of ($18''$, $18''$) and wavelength of $4 \mu m$, showing oblique astigmatism. Given that the P-V is larger than 0.25 waves, it is not considered to be diffraction limited.	106
G.1	Translation perturbation of the primary mirror. The comparison is made between an off-axis Cassegrain, RC and Gregorian design. All null depths calculated with a $10 \mu m$ central wavelength and bandwidth of $1 \mu m$	108
G.2	Rotation perturbation of the primary mirror. The comparison is made between an off-axis Cassegrain, RC and Gregorian design.	108
G.3	Surface shape perturbation of the primary mirror. The comparison is made between an off-axis Cassegrain, RC and Gregorian design.	109

G.4	Translation perturbation of the collimator. The comparison is made between an off-axis Cassegrain, RC and Gregorian design. All null depths calculated with a $10\ \mu m$ central wavelength and bandwidth of $1\ \mu m$	109
G.5	Rotation perturbation of the collimator. The comparison is made between an off-axis Cassegrain, RC and Gregorian design.	110
G.6	Surface shape perturbation of the collimator. The comparison is made between an off-axis Cassegrain, RC and Gregorian design.	110

List of Tables

2.1	Amplitude coefficients for an electric field incident on a slab of ZnSe at 45°	9
2.2	Table of the first ten Zernike polynomials and the associated aberration name.	11
4.1	Conic constants and their corresponding conic sections.	26
4.2	Scaling laws of the primary aberrations, as function of the f/# and field angle θ . Adapted from [50, p. 111].	27
4.3	System requirements for the MIR nulling interferometer.	27
6.1	Design parameters for an off axis Cassegrain. The parameters R_1 , R_2 , K_2 and i_2 are calculated from Equation 4.9, Equation 4.13, Equation 4.14 and Equation 4.12.	50
6.2	Design parameters for an off axis parabola to collimate the telescope beam.	52
6.3	Comparison of the null depth loss, due to aberrations, derived by Hyde et al. and the null depth loss as simulated with Code V.	64
6.4	The considered contributions to null depth loss and their corresponding requirements.	65
6.5	Code V commands of the considered perturbation parameters.	65
6.6	Table containing the Zernike polynomial expansion coefficients for the exit pupil map. The coefficients are determined by fitting the corresponding Zernike polynomial to the exit pupil sphere. Each surface was perturbed by a z decenter of 100 μm . Zernike polynomials that have $<1 \cdot 10^{-4}$ RMS wavefront errors are denoted by "-".	68
6.7	Comparison of Zernike decomposition for ADE and BDE perturbation of 0.1°. Zernike terms 1 to 3 are omitted, as they have no contribution to the RMS WFE.	73
6.8	Beam shear, spot size difference and RMS WFE due to translation of the collimator. The third, fifth and seventh column indicate the expected null depth, based on the model from Hyde et al. The beam shear was determined with Code V's "Real Ray Trace" function, the spot diameter was determined with the "Spot diagram function" and the RMS WFE was determined with the "Wavefront analysis" function.	75

6.9 Expected null depth loss, based on the Hyde et al. model, for the considered rotational perturbation parameters. 75

6.10 Expected null depth loss, based on the Hyde et al. model, for the considered rotational perturbation parameters. 75

6.11 Design parameters for the off-axis RC and Gregorian. 80

Nomenclature

Abbreviations

Abbreviation	Definition
BC	Beam Combiner
BSP	Beam Synthesis Propagation
FFT	Fast Fourier Transform
FKSI	Fourier Kelvin Stellar Interferometer
FoV	Field of View
LBTI	Large Binocular Telescope Interferometer
MTF	Modulation Transfer Function
MMZ	Modified Mach Zehnder
OPD	Optical Path Difference
OPL	Optical Path Length
pc	Parsec
P-V	Peak-to-Valley
RC	Ritchey-Chretien
RMS	Root Mean Square
WFE	Wavefront Error

Symbols

Symbol	Definition	Unit
B	Baseline	[m]
D	Aperture diameter	[m]
Z_n^m	Zernike polynomial	[-]
R_n^m	Radial polynomial	[-]
$f/\#$	Focal ratio / f-number	[-]
K	Conic constant	[-]
i	Tilt angle	[rad]
k	Wave number	[m ⁻¹]
N	Null depth	[-]
λ	Wavelength	[m]
θ	Spherical coordinate, angle with z axis	[rad]
ϕ	Spherical coordinate, angle with x axis	[rad]
φ	Phase, used in OPD	[rad]
α, β	Angle on sky	[rad]

1

Introduction

1.1. Introduction to detecting exoplanets

Since centuries past, mankind has looked up to the stars and dreamt of the worlds that might be out there. Coming one step closer to answering questions yet unknown, astronomers started building telescopes, to peer into the sky and observe its riches. Galileo, Huygens, Newton and Cassegrain pioneered the astronomical telescope in the early to mid 1600s, improving the effectiveness step by step. Since then, a multitude of designs and variations have been realised. With each improvement, astronomers could peer further into the sky, going beyond the planets of our solar system. Observing the unfathomable extent of the universe, sooner or later the question arises: are we alone in the universe? To answer this, it is necessary to find ways to directly observe planets outside our solar system and characterise their atmospheres.

Yet however much the classical telescope has improved, imaging planets outside of the solar system remains a formidable challenge. To this day, direct imaging of exoplanets can only be done in a few exceptional cases. Of the more than 5500 confirmed exoplanets, only 1% has been observed with direct imaging¹. The vast majority of exoplanets has been identified indirectly, deriving its existence from other observables, such as the radial velocity of the star, or the reduction in intensity of the star over time when the planet transits the star in the line of sight. Although these methods are effective at identifying exoplanets, they are biased towards a specific type of planet, usually large (super-Earth sized) planets close to the star and in the line of sight. In Figure 1.1, a current census is shown of the confirmed exoplanets at the time of writing. Notice the gap in discoveries at an orbit period of one Earth year and planet mass of one Earth mass.

The difficulty of directly observing terrestrial planets in the habitable zone is twofold: First, a main sequence star is 10^{10} times brighter than a terrestrial planet in the habitable zone, when observed in the visible spectrum. Second, the angular resolution needed to find such a planet is on the order of $0.1''^2$, assuming a planet orbiting at 1 AU around a star 10 pc away. To illustrate the difficulty of achieving such a resolution, consider the James Webb Space Telescope, launched in December 2021. This state of the art space based observatory,

¹Source: <https://exoplanets.nasa.gov/>

²1 arcsecond, denoted by $''$, is 1/60 of a degree.

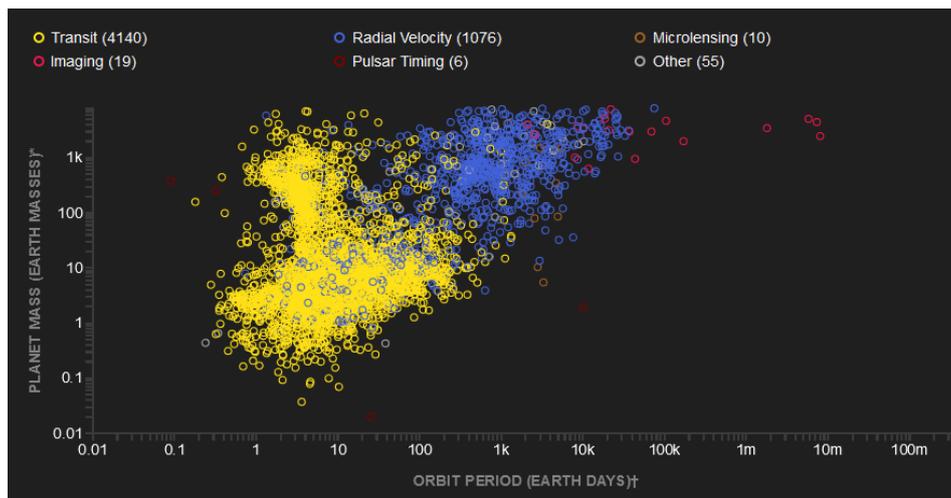


Figure 1.1: Exoplanet census of current confirmed discoveries, separated into different detection methods.

with its 6.5 m diameter aperture, has an angular resolution of approximately $0.3''$. To obtain even higher angular resolutions, interferometry can be utilised, where instead of the individual aperture size, the angular resolution depends on the separation between the apertures. This technique has virtually no limit on angular resolution, with the most extreme example coming from the Event Horizon Telescope (EHT), which has a baseline on the order of the size of the Earth. EHT was used to generate the first image of a black hole [1], approximately 40 microarcseconds across.

As for detecting exoplanets, interferometry thus addresses the issue of angular resolution, however it does not necessarily account for the contrast between the star and planet. Coronagraphy is a technique to suppress the starlight sufficiently to observe its nearby surroundings. To achieve this, a mask is placed in the line of sight to reject the starlight. An upcoming space based coronagraph, slated to launch by May 2027, is the Nancy Grace Roman Space Telescope (formerly known as the Wide Field InfraRed Survey Telescope), which has a target contrast of 10^{-9} , but an inner working angle of 250 milliarcseconds at $1 \mu m$ [2].

1.2. Nulling interferometry

To obtain both the angular resolution, as well as the required rejection for direct detection of Earth-like exoplanets in the habitable zone, nulling interferometry was proposed by R.N. Bracewell in 1978 [3]. A nulling interferometer uses optical phase shifting to achieve destructive interference on the line of sight of the instrument, such that the star light can be rejected over a wide spectrum. To find the planet, the interferometer is rotated around the line of sight, so that the planet signal is modulated onto the bright and dark fringes. This technique has been demonstrated successfully, both in the lab [4, 5] as well as on ground based observatories [6–9]. The main limitation on ground based observatories comes from atmospheric turbulence, which limits both the coherent observation time, as well as the null depth that can be achieved, even with adaptive optics systems [10].

To come back to the question of 'are we alone in the universe?', a nulling interferometer has the potential to answer this question, when it is used in conjunction with a spectrometer. Without touching upon the philosophical question of what it means to be alive, astrophysicists and as-

tronomers have found some consensus of what life could look like outside of our familiar climate³. The indicators of any form of life are formalised in the notion of biosignatures. To follow the NASA astrobiology roadmap [12], published in 2008, *"A biosignature is an object, substance and/or pattern whose origin specifically requires a biological agent"*. These biosignatures are best observed in the mid-infrared spectrum, due to the abundance of absorption features in that part of the electromagnetic spectrum. It was in part for this reason that Quanz et al. labelled it as *"the mid-infrared opportunity"* [13]. Furthermore, the contrast between a star and planet is reduced by at least two orders of magnitude in the mid-infrared spectrum, as the thermal radiation from the planet is colder and thus less bright compared to the reflected solar radiation. To illustrate this, a hypothetical solar system was modelled by Des Marais et al. [14], shown in Figure 1.2, where the reduced contrast is highlighted. Thus, ideally a nulling interferometer observes in the mid-infrared spectrum.

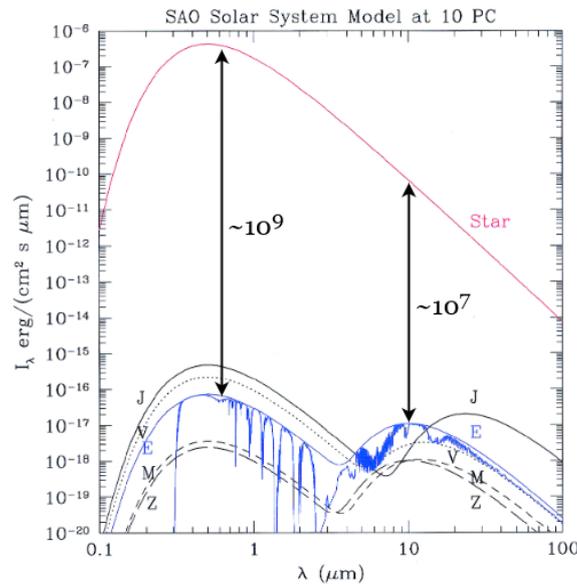


Figure 1.2: A simulated emission spectrum of a star and orbiting planet. Due to thermal radiation of the planet, an extra signature shows up in the radiation spectrum. Adapted from [14].

1.3. Towards a space based nulling interferometer

Due to atmospheric attenuation at the mid-infrared, there is only a limited amount of windows that are feasible for nulling interferometry on the ground. As is shown in Figure 1.3, between 1 and 10 μm , there are only a handful of usable windows, with the widest window around 10 μm . To avoid the atmospheric attenuation, as well as turbulence, space based missions have been proposed in the late 1990s by both ESA and NASA. ESA's Darwin mission [15] and NASA's Terrestrial Planet Finder - Interferometer (TPF-I)[16] were meant to be formation flying interferometers, but were cancelled due to budgetary constraints.

During the early design phase of Darwin and TPF-I, many works were published on high level modelling and theory development of the predicted performance of nulling interferometers,

³The question of what life would look like on other planets is a scientific interdisciplinary effort that has been discussed extensively, and will continue to evolve, as there are still many unanswered underlying questions. Even the perception of the concept of a biosignature is still being debated. For a recent discussion on this, see for example *"Is there such a thing as a biosignature?"* [11]

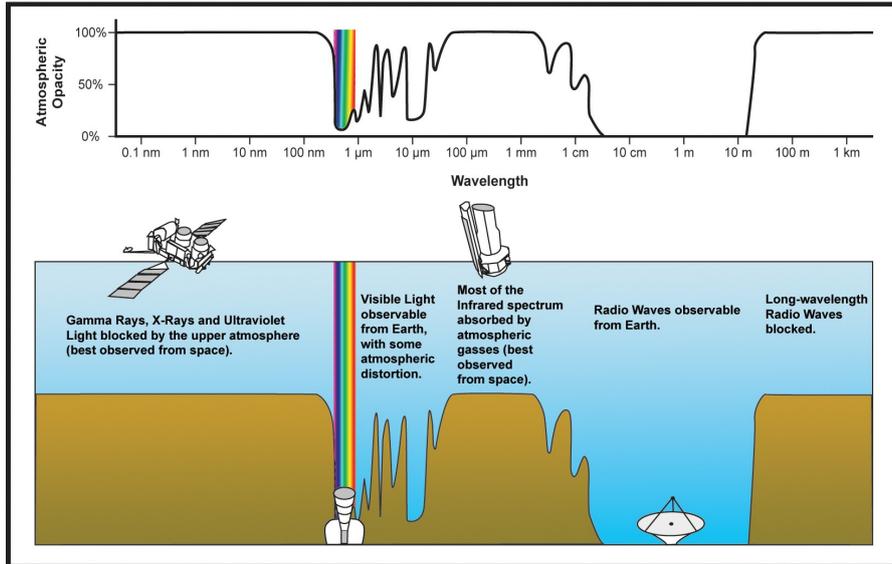


Figure 1.3: The transmissive windows of the Earth’s atmosphere. Although there are some small transmissive windows between 1 and 10 μm , they are relatively opaque and need long integration times or extreme adaptive optics to be usable for astronomical observations.

based on number of apertures, configuration and diameter of the apertures and baseline. At the time, it was unknown what an optimal aperture configuration would be, so many different configurations have been proposed and analysed [17–20]. This led the teams of Darwin and TPF-I to converge to the same design: the so-called Emma X array, which is a 4 aperture rectangular configuration [21].

To show that the large class missions would be feasible, precursor missions were proposed for both Darwin and TPF-I. PEGASE, the precursor to Darwin, was meant to be a formation flying instrument, where the flat siderostats would be separated from the central beam combiner spacecraft [22]. The Fourier-Kelvin Stellar Interferometer (FKSI), on the other hand, was supposed to be a structurally connected 2 aperture nulling interferometer [23]. Neither PEGASE nor FKSI ever materialised, again due to budgetary constraints, as well as lack of technology readiness.

A more recent example of a proposed space based nulling interferometer is the Large Interferometer for Exoplanets (LIFE). LIFE aims to be a revival of the forlorn era of Darwin and TPF-I. The concept shows similarities with its predecessor, but it also builds on (relatively) new technology developments, such as adaptive nulling [24] and kernel nulling [25]. Whereas it was widely accepted in the early 2000s that an Emma X-array configuration was optimal, a new study has shown that a 5 aperture pentagonal kernel nuller yields the highest number of planet detections [26].

However much the nulling interferometer was studied on a system level, very little, if any work has been published detailing the optical design aspect of a nulling interferometer. To the author’s knowledge, no work has been published on the choice of telescope design and its potential influence on the stability and performance of a nulling interferometer. For ground based nulling interferometers, the infrastructure already present was used, whereas the proposed space based missions never made it past the early phases of the design process.

1.4. Research question

The previous discussion leads us to the research question to be answered:

To which extent does the choice of telescope design influence the stability and performance of a space based, mid-infrared, nulling interferometer?

As there are many considerations to make for an optical design of a telescope for nulling interferometry, the research question will be divided into subquestions:

1. Can an optical model accurately predict the performance and stability of a nulling interferometer under perturbation?
2. How does an on axis telescope design compare to an off axis telescope design in terms of the stability of the nulling interferometer?
3. How do the archetypal Cassegrain, Ritchey-Chrétien and Gregorian designs compare in terms of the performance and stability of the nulling interferometer?

1.5. Thesis outline

To address the research question, an optical model of a nulling interferometer will be made and analysed. Before this model can be discussed, a basic understanding of optics and interferometry is needed, which are discussed in chapter 2 and chapter 3, respectively. With the theoretical basis established, a rationale behind the optical design of the nulling interferometer is given in chapter 4. This acts as a starting point for chapter 5, which elaborates on the methodology of the sensitivity analysis. Then, verification of the optical model and results of the sensitivity analysis will be discussed in chapter 6. In chapter 7, recommendations for future research are given, which finally leads to the conclusion, given in chapter 8.

2

Theoretical background - Optics

In this chapter a handful of concepts concerning optics are worked out in order to provide the reader with enough context to follow the discussion on nulling interferometry. In the first section, the wave nature of light and wave properties are treated. In section 2.2, the polarisation of light is discussed, which will be used in chapter 6 when the beam splitter model is evaluated. Additionally, the Fresnel equations are given in section 2.3, which are also needed to correctly validate the beam splitter model.

2.1. Electromagnetic waves

One of the most fiercely debated topics in the history of physics has been the matter of describing light as a wave or a particle. Several experiments have been performed to establish whether light should be considered as a wave or particle, including Young's double slit experiment, which will be elaborated upon in section 3.1. Although by now it is established that light behaves as both a wave and particle, it is most convenient for the discussion of interferometry to consider the wave nature of light. A widely used set of equations to describe the behaviour of electromagnetic waves is the Maxwell equations.

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0} \quad (2.1)$$

$$\nabla \cdot \mathbf{B} = 0 \quad (2.2)$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \quad (2.3)$$

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J} + \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \quad (2.4)$$

These equations are coupled first order partial differential equations. With some derivation, worked out in Appendix B, they can be decoupled, such that two separate second order partial differential equations can be defined for both \mathbf{E} and \mathbf{B} .

$$\nabla^2 \mathbf{E} = \mu_0 \epsilon_0 \frac{\partial^2 \mathbf{E}}{\partial t^2} \quad (2.5)$$

$$\nabla^2 \mathbf{B} = \mu_0 \epsilon_0 \frac{\partial^2 \mathbf{B}}{\partial t^2} \quad (2.6)$$

From this, it becomes apparent that for each Cartesian coordinate, the above equations satisfy the wave equation:

$$\nabla^2 f = \frac{1}{\nu^2} \frac{\partial^2 f}{\partial t^2} \quad (2.7)$$

The benefit of expressing the Maxwell equations in this way is that known solutions of the wave equation can be used to describe the evolution of an electromagnetic wave in time and space. One of those solutions is a plane wave, which is a field that has constant phase in the plane perpendicular to the direction of travel. Suppose the plane wave travels in the +z direction (such that $k_x = k_y = 0$), the plane wave is then described by:

$$\mathbf{U}(x, y, z, t) = \mathcal{A}(\mathbf{r}) e^{i(kz - \omega t)}. \quad (2.8)$$

Alternatively, it can be separated into a space dependent part and a time dependent part:

$$U(\mathbf{r}, t) = U(\mathbf{r}) e^{-i\omega t}, \quad (2.9)$$

where $U(\mathbf{r})$ is the complex amplitude:

$$U(\mathbf{r}) = \mathcal{A}(\mathbf{r}) e^{i\phi(\mathbf{r})}, \quad (2.10)$$

such that the modulus of $U(\mathbf{r})$ is $\mathcal{A}(\mathbf{r})$ and the argument of $U(\mathbf{r})$ is the phase $\phi(\mathbf{r})$.

Although stars emit spherical waves, when the distance is large enough, the incoming wavefront can be approximated by a plane wave. A perfectly collimated beam is by definition a plane wave, as it has a focus at infinity. As such, Equation 2.8 can be used to describe the complex amplitude in a pupil plane beam combiner, as well as the incident wave on the nulling interferometer.

2.2. Polarised light

When light is described as an electromagnetic wave, it can be specified by its polarisation. The polarisation specifies in which direction the light is oscillating. To exemplify, let us again consider the plane wave travelling in the +z direction. In that case, the electric field has no z component:

$$\mathcal{E}(z, t) = \begin{pmatrix} \mathcal{A}_x \cos(kz - \omega t + \phi_x) \\ \mathcal{A}_y \cos(kz - \omega t + \phi_y) \\ 0 \end{pmatrix} \quad (2.11)$$

When the x and y components are in phase (i.e. $\phi_x - \phi_y = 0$ or $\phi_x - \phi_y = \pi$), the light is linearly polarised. When the fields are out of phase by $\pi/2$ radians and have the same magnitude, the light is circularly polarised. When the x and y components of the magnitude are different and out of phase by $\pi/2$, the field is said to be elliptically polarised. To quantify this polarisation state, the Jones vector can be defined:

$$J = \begin{pmatrix} \mathcal{A}_x e^{i\phi_x} \\ \mathcal{A}_y e^{i\phi_y} \end{pmatrix}. \quad (2.12)$$

Where $E_{0x/y}$ is the electric field magnitude in the x and y direction, respectively and $\phi_{x/y}$ is the phase in the x or y direction.

An incident field can behave differently when interacting with optical elements when it is polarised, compared to when it is not. For example, in a non-polarisation-maintaining optical fibre, polarised light at the input can change the polarisation state at the output of the fibre, whereas non-polarised light would be unaffected. In this case, having polarised light might be unwanted, in which case it can be made unpolarised with polarisation scramblers. Conversely, when polarised light is desired, polarizers or wave plates can be used. Another notable effect is when polarised light is incident with a particular angle on a partly transmissive medium, there exists an angle where only one polarisation state is reflected, known as the Brewster angle.

As for nulling interferometry, the polarisation state can have a detrimental effect on the null when the polarisation states are not matched correctly between the combined beams [27, 28]. Properly addressing this in the optical model would require extensive analytical understanding of the propagation of polarisation states throughout the interferometer, which would be a complete study on its own. For this reason, it is assumed that the observed sources are unpolarised and that the optics introduce no polarisation dependence.

2.3. Fresnel equations

In order to quantify how different components of an incoming field are transmitted and reflected on an incident medium, the Fresnel equations can be used. In the following section, we will see that when a linearly polarised beam is incident on a beam splitter the phase of the beam can change, depending on how the beam travels through the material. The theory presented here will be used to verify the outputs of the Modified Mach Zehnder beam combiner, which will be discussed in more detail in subsection 4.7.3 and subsection 6.1.3. In those sections, a linearly polarised wave will be propagated through (a collection of) beam splitters and the output complex amplitudes will be compared to the relations presented in the subsequent section.

Consider a surface with an electric field incident at an angle θ to the surface normal. The surface divides two media with refractive index n_i and n_t . The plane of incidence is parallel to the \vec{k} vector and perpendicular to the surface. Now, the field can be decomposed into two components: \vec{E}_\perp and \vec{E}_\parallel ¹, as shown in Figure 2.1.

Fresnel's equations are defined by the fraction of the parallel (or perpendicular) component of the reflected and incident field.

¹Alternatively, the fields can be denoted \vec{E}_s and \vec{E}_p , for perpendicular and parallel, respectively. The 's' comes from the German word senkrecht.

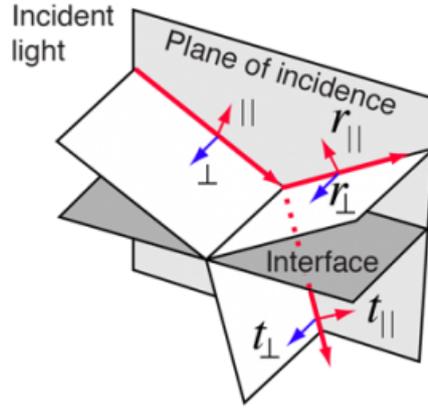


Figure 2.1: Definitions of the s- and p- components of the reflected and transmitted field.

$$r_{\perp} \equiv \left(\frac{E_{0r}}{E_{0i}} \right)_{\perp} = \frac{n_i \cos \theta_i - n_t \cos \theta_t}{n_i \cos \theta_i + n_t \cos \theta_t}, \quad (2.13)$$

$$t_{\perp} \equiv \left(\frac{E_{0t}}{E_{0i}} \right)_{\perp} = \frac{2n_i \cos \theta_i}{n_i \cos \theta_i + n_t \cos \theta_t} \quad (2.14)$$

$$r_{\parallel} \equiv \left(\frac{E_{0r}}{E_{0i}} \right)_{\parallel} = \frac{n_t \cos \theta_i - n_i \cos \theta_t}{n_t \cos \theta_i + n_i \cos \theta_t} \quad (2.15)$$

$$t_{\parallel} \equiv \left(\frac{E_{0t}}{E_{0i}} \right)_{\parallel} = \frac{2n_i \cos \theta_i}{n_i \cos \theta_t + n_t \cos \theta_i}. \quad (2.16)$$

Whenever this fraction is positive, both amplitudes are positive and thus they are in phase, while they are out of phase when the fraction is negative. From this, we can see that the exact point of phase reversal depends on the index of refraction of the incident and transmitted media, as well as the angle of incidence. The angle of refraction (θ_t is a dependent parameter, defined by the others through Snell's law). This principle is of particular importance in understanding the Mach-Zehnder interferometer. To get a better understanding of these equations, let us consider a practical example. Let us assume an incident plane wave in air on a surface made of ZnSe, with a refractive index of 2.39279 (assuming a wavelength of $10 \mu m$) and angled at 45° . Using the above equations, we find the following values for r_{\perp} , r_{\parallel} , t_{\perp} and t_{\parallel} :

Amplitude Coefficient	Value
r_{\perp}	-0.5277
r_{\parallel}	0.27843
t_{\perp}	0.47234
t_{\parallel}	0.53406

Table 2.1: Amplitude coefficients for an electric field incident on a slab of ZnSe at 45° .

So, for this specific case, only the perpendicular component is phase shifted, relative to the incident wave. If the angle of incidence would have been past the so called polarisation angle, both the parallel and perpendicular component of the reflected wave would have been phase

shifted by π radians. When extending this calculation to all incident angles between 0 and 90 degree, we obtain the figure given below.

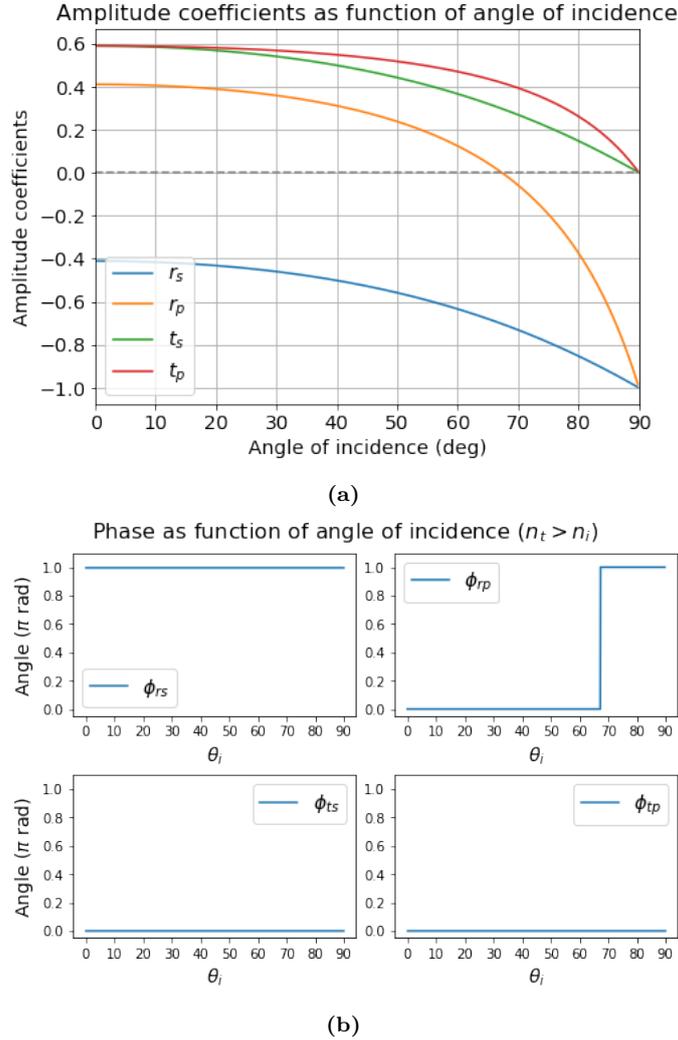


Figure 2.2: (a) Amplitude coefficients for the reflected and transmitted wave, separated into the perpendicular (s) and parallel (p) component, as function of the angle of incidence. (b) Phase of the complex amplitude coefficients.

2.4. Zernike polynomials

Zernike polynomials are a set of polynomials that are orthogonal on the unit disk. Similar to how the Fourier series can be used to approximate any periodic function, the Zernike polynomials can be used to approximate any wavefront that is sufficiently smooth. Such a 'Zernike decomposition' of the wavefront will be used extensively in section 6.5. The benefit of this particular description of wavefront disturbances is that the polynomial terms correspond to physical aberrations. Each polynomial has two components: a radial one and an azimuthal one:

$$Z_n^m(\rho, \phi) = \begin{cases} R_n^m(\rho) \cos(m\phi) & \text{if } m > 0, \\ R_n^m(\rho) \sin(-m\phi) & \text{if } m < 0, \end{cases} \quad (2.17)$$

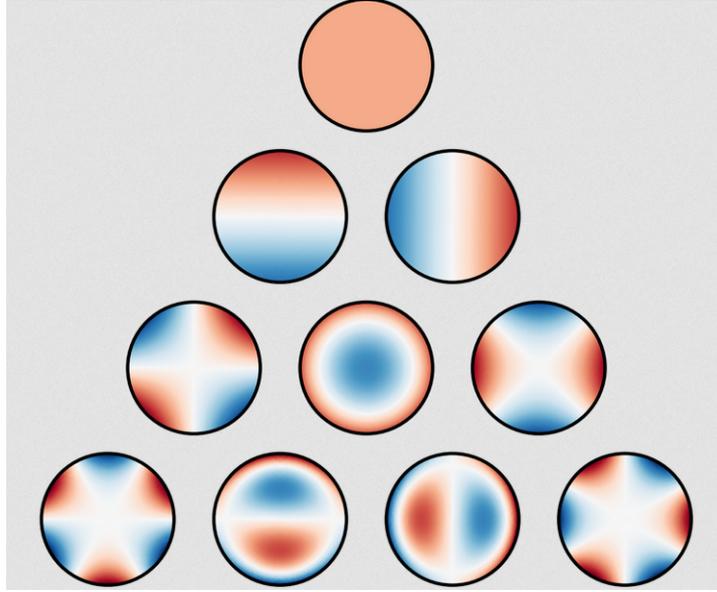


Figure 2.3: Visual representation of the first 10 Zernike polynomials. First row: piston. Second row (from left to right): vertical tilt, horizontal tilt. Third row: vertical astigmatism, defocus, oblique astigmatism. Fourth row: vertical trefoil, vertical coma, horizontal coma, oblique trefoil.

$$R_n^m(\rho) = \sum_{k=0}^{\frac{n-m}{2}} \frac{(-1)^k (n-k)!}{k! \left(\frac{n+m}{2} - k\right)! \left(\frac{n-m}{2} - k\right)!} \rho^{n-2k}, \quad (2.18)$$

where $R_n^m(\rho)$ is the radial component and $\cos m\phi$ or $\sin(-m\phi)$ is the azimuthal component. Hence, the summation variable n indicates the radial degree, while m is referred to as the azimuthal degree. The first 10 Zernike polynomials (up until the third radial degree) are given in Table 2.2 and derived in Appendix A.

Z_n^m	Z_j	Aberration name
Z_0^0	1	Piston
Z_1^{-1}	$\rho \sin(\phi)$	Vertical tilt
Z_1^1	$\rho \cos(\phi)$	Horizontal tilt
Z_2^{-2}	$\rho^2 \sin(2\phi)$	Oblique astigmatism
Z_2^0	$(\rho^2 - 1)$	Defocus
Z_2^2	$\rho^2 \cos(2\phi)$	Vertical astigmatism
Z_3^{-3}	$\rho^3 \sin(3\phi)$	Vertical trefoil
Z_3^{-1}	$(3\rho^3 - 2\rho) \sin(\phi)$	Vertical coma
Z_3^1	$(3\rho^3 - 2\rho) \cos(\phi)$	Horizontal coma
Z_3^3	$\rho^3 \cos(3\phi)$	Oblique trefoil

Table 2.2: Table of the first ten Zernike polynomials and the associated aberration name.

2.5. Aberrations

All practical optical systems suffer from some form of aberration. An aberration is a deviation from a perfect spot or wavefront. They can thus be expressed in terms of ray aberrations, or wavefront aberrations. In this section, the most commonly occurring aberrations will be

discussed qualitatively. When one of these aberrations is present in one of the arms of the interferometer, the null will be degraded, due to either wavefront errors, or mismatches in the beam shape.

2.5.1. Spherical aberration

When a ray is incident on a spherical surface close to the optical axis, the angle of incidence is close to zero degrees and is not refracted strongly. To calculate the angle of refraction, Snell's law can be used in the paraxial approximation. When the rays are closer to the edge of the lens, the angle of incidence becomes larger and thus the angle of refraction. Because of this, rays further from the optical axis come to a focus in a shorter distance from the lens. To reduce spherical aberration, the radius of curvature of the second surface should be minimised, such that the angle of incidence is approximately equal over the lens surface. Parabolic mirrors have the special property that they do not suffer from spherical aberration. This is due to the property of parabolas that any ray parallel to the axis of symmetry ends up in the focal point of the parabola.

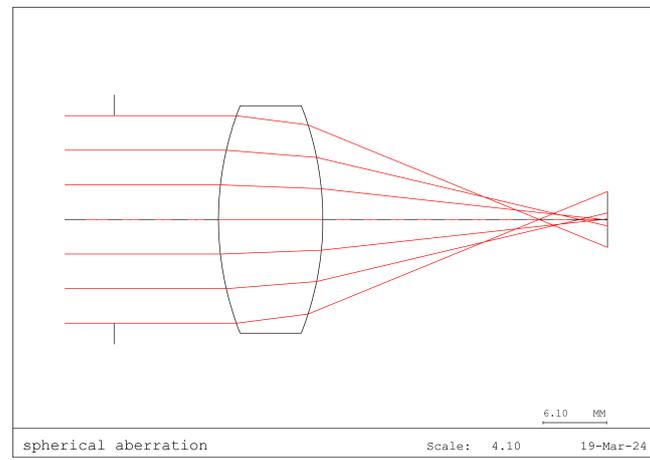


Figure 2.4: Spherical aberration resulting from a flat field incident on a biconvex lens. Due to the difference in angle of incidence along the lens height, the focal plane shifts as function of the incident ray height. The resulting spot on the detector plane has an elliptical shape.

2.5.2. Astigmatism

Astigmatism is an aberration where the meridional and sagittal rays have a different focal plane, illustrated by the Code V model given in Figure 2.5. It is caused by a difference in radius of curvature along the meridional and sagittal plane. The result of this is that point-like objects are defocused in two perpendicular axes, appearing elongated along that direction. It is known from geometrical optics theory that astigmatism is the dominant aberration in off-axis optical systems [29]. However, with specific orientation of the optical elements, it is possible to eliminate astigmatism in a two element off axis telescope [30]. A diffractive element suffering from astigmatism can be corrected by a cylindrical lens, such that the focal length in one axis is corrected, but not the other axis.

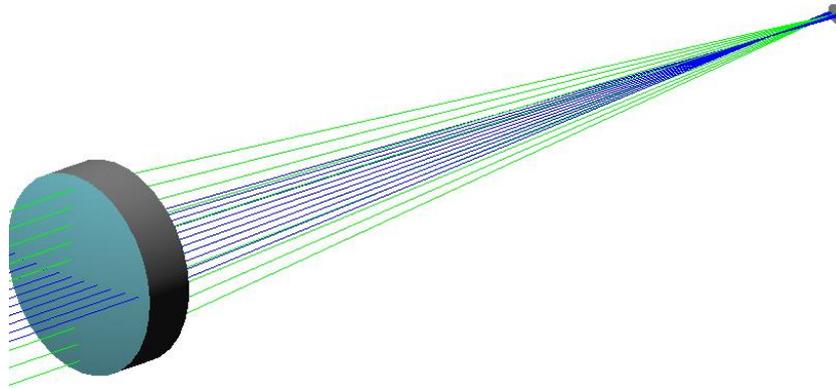


Figure 2.5: A lens with astigmatism. The rays in the meridional plane are colored green and come to a focus at the image plane. The rays in the sagittal plane are colored blue and come to a focus before the image plane.

2.5.3. Coma

Unlike spherical aberration and astigmatism, coma is an aberration that depends on the field angle and does not occur when the field is parallel to the optical axis. This aberration is especially prevalent in telescopes and causes off-axis point sources to appear like a comet. Coma can be minimised by shifting the aperture stop relative to the entrance pupil.

2.5.4. Field curvature

Field curvature, sometimes referred to as Petzval field curvature, is an aberration that depends on the field angle. Whereas coma spreads out a point source, creating a tail, with field curvature a point source is merely defocused. The focal length thus depends on the field angle, as can be seen from Figure 2.6. Field curvature can be corrected either by a curved image surface, or a field corrector.

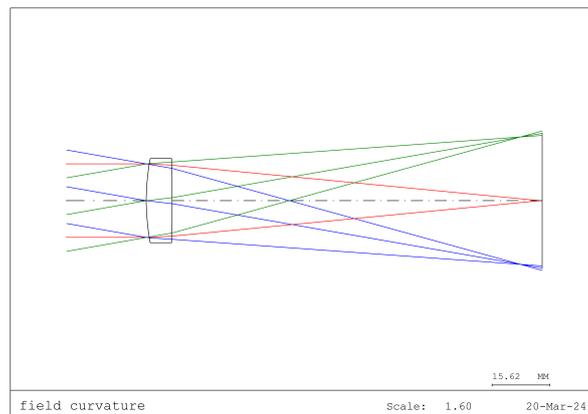


Figure 2.6: A singlet which shows field curvature. As the field angle increases, the focal length decreases.

2.5.5. Distortion

When distortion aberration is present in a system, there is no image blur, but the magnification is not uniform across the image. Generally speaking there are two textbook examples of distortion: barrel and pincushion distortion. This distinction is not exhaustive, as distortion can also appear in a combination of these two examples. For the former, the sides of the image are magnified more, compared to the center of the image, while for the latter, the sides of the image are magnified less. The effect of this is that for barrel distortion, objects near the edge of the image appear larger, compared to when that object would be at the centre of the image.

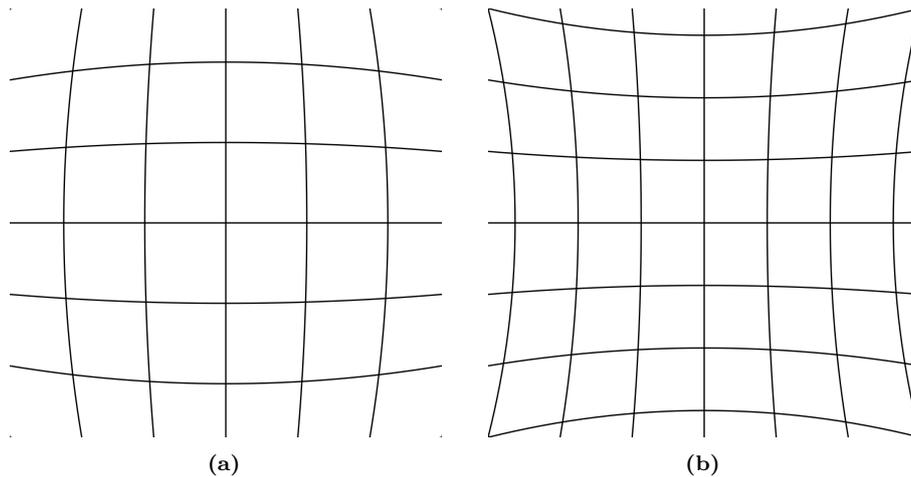


Figure 2.7: (a) Barrel distortion. (b) Pincushion distortion

2.6. Chapter summary

In this chapter, several theoretical concepts from optics have been discussed. First the origins of the plane wave was considered through the Maxwell equations. The humble plane wave serves as a starting point for propagating a wavefront through the interferometer. Second, polarisation of light and its effect on optical components was discussed in section 2.2. Although polarisation states do influence the performance of a nulling interferometer, a simplification is made that the polarisation states are matched perfectly between the interferometer arms. Third, the Fresnel equations were analysed in section 2.3. An example was given for a ZnSe plate in air, angled at 45° . It showed that in this case, only the perpendicular component of the reflected field is phase shifted relative to the incident field. This example will be revisited in chapter 6, to validate the beam splitter model. Fourth, the Zernike polynomials were discussed in section 2.4. These polynomials are a representation of physical aberrations in optical systems and will be used extensively in section 6.5. Finally, some aberrations and their origin are discussed in section 2.5

3

Theoretical background - Interferometry

In this chapter, a theoretical framework will be given for interferometry and later in the chapter, nulling interferometry. In the first section, Young's experiment will be used as a stepping stone for the theory of interferometry and nulling interferometry. From that, the discussion will be extended briefly to stellar interferometry in section 3.2. Using that as a basis, nulling interferometry is introduced in section 3.3. Finally, the response of a nulling interferometer will be derived in section 3.4.

3.1. Young's experiment

Young's double slit experiment is a classic physics experiment showing the wave nature of light, performed by Thomas Young in 1803. A schematic representation of the experiment is shown in Figure 3.1. It consists of a screen with two slits and behind that, another screen for observation. In an actual setup, a light source and a single pinhole would be used to create approximate plane waves at the double slit screen.

When the slits are close enough and smaller than the wavelength of the incoming field, a fringe pattern will emerge on the observation plane. The term 'close enough' is not exact, but can be formalised in the notion of coherence length, which will be discussed later in this section. For now, let us assume a plane wave is incident on the two slits. By the Huygens-Fresnel principle, two spherical waves will emanate from the slits. These spherical waves will fall on the observation screen and add constructively or destructively, depending on the difference in distance travelled to the screen. As indicated in Figure 3.1, the intensity on the screen can also be expressed in terms of the difference in time of arrival, τ , as they are related by $\Delta R = c\tau$. Let us now assume that $U(\mathbf{r}_1, t)$ is a monochromatic plane wave, such that it can be expressed as:

$$U(\mathbf{r}_1, t) = A(\mathbf{r}_1)e^{-i\omega t}, \quad (3.1)$$

where $A(\mathbf{r}_1)$ is the complex amplitude, ω the frequency of the time harmonic plane wave and

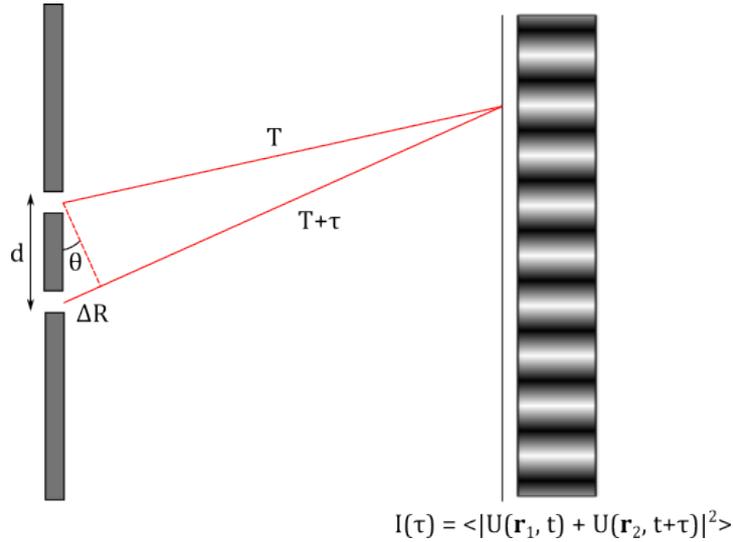


Figure 3.1: Schematic representation of Young's double slit experiment. On the left, there is a screen with two slits, spaced by a distance d . The considered rays from the centre of the slits are angled with an angle θ , which results in a path length difference ΔR . The intensity, and thus the locations of the dark and bright fringes on the screen, depends on the angle θ .

at the time. With some algebra, it can be shown that the intensity at the screen is given by:

$$I(\tau) = |A(\mathbf{r}_1)|^2 + |A(\mathbf{r}_2)|^2 + 2|A(\mathbf{r}_1)||A(\mathbf{r}_2)|\cos(\omega\tau + \phi), \quad (3.2)$$

where $\omega = c\frac{2\pi}{\lambda}$ and ϕ is a phase difference between $A(\mathbf{r}_1)$ and $A(\mathbf{r}_2)$. Note that this means that the fringes are visible on the screen, which need not always be the case, as we will see for pupil plane combination. For now, we will assume that $\phi = 0$ ¹, in which case the maximum of the intensity occurs when $\omega\tau = m \cdot 2\pi$, with m integer. Using the definitions for ΔR and ω , this can be written as:

$$\Delta R = m\lambda. \quad (3.3)$$

This makes intuitive sense, as the above equation simply states that constructive interference occurs whenever the path length difference is exactly equal to a multiple of the wavelength, in other words, when the two waves are in phase. Finally, for large distances between the slits and the observation screen, the angle of the maxima can be expressed as:

$$\theta_{max} = \frac{\Delta R}{d} = \frac{m\lambda}{d}. \quad (3.4)$$

Similarly, the angle of the minima is given by:

$$\theta_{min} = \frac{(m + \frac{1}{2})\lambda}{d}. \quad (3.5)$$

So the fringe spacing increases for increasing wavelength or decreasing distance between the slits. However, as mentioned, whether fringes are formed on the screen depends on the characteristics of the source, as well as the slit spacing. A monochromatic point source is perfectly

¹In the case of $\phi \neq 0$, the fringe pattern simply shifts up or down.

coherent, so that regardless of the slit spacing, fringes can be formed on the screen. In practice, no source of light is perfectly monochromatic and only unresolved sources on sky appear as point sources. To quantify how far slits can be positioned apart while still creating fringes, the coherence length can be defined. It depends on the width of the spectrum emitted by the source, $\Delta\lambda$, as well as the centre wavelength of the spectrum, λ_c :

$$\Delta l_c = \frac{\lambda_c^2}{\Delta\lambda}. \quad (3.6)$$

Thus, when the bandwidth is small, for example with a laser source, the coherence length can become large. When the distance between the slits is within the coherence length, fringes can appear on the observation plane.

Young's experiment is a useful academic tool to introduce the topic of interferometry and coherence, but has limited use outside of that context. Almost a century after Young first performed his experiment, Albert A. Michelson utilised the concept of interference to measure the diameter of stellar objects.

3.2. Stellar interferometry

Although stellar interferometry is a complex topic in its own right, it is useful for the treatment of nulling interferometry to briefly discuss it here. The Michelson stellar interferometer is in principle very comparable to Young's double slit experiment. However, instead of using narrow slits, Michelson used two flat mirrors to create a baseline, which were then imaged by a single concave reflector [31]. Both beams are focused on the image plane, where the interference fringes appear. A schematic representation of the Michelson stellar interferometer is shown in Figure 3.2a. In this particular case, the fringes are visible on the screen, as indicated in red in Figure 3.2b. The envelope of the fringes is determined by the PSF of the single aperture.

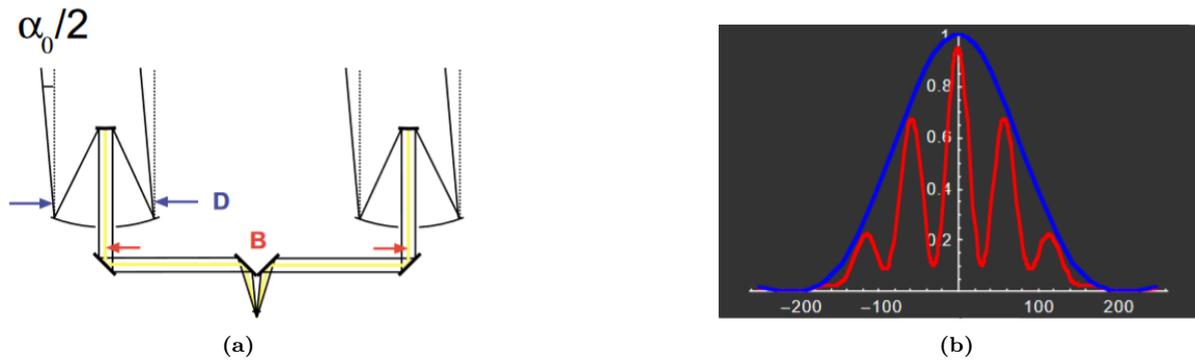


Figure 3.2: (a) A Michelson stellar interferometer with baseline B and aperture diameter D . (b) The fringe visibility as function of angular separation α , shown with the red line. The blue line indicates the envelope due to the single aperture with diameter D . Adapted from [32].

Just as with Young's experiment, the distance between the maximum and minimum of the fringes is equal to $\lambda/(2 \cdot B)$. One might wonder how it is possible for a stellar interferometer with a baseline many orders of magnitude larger than the coherence length is able to make fringes from a stellar, polychromatic source. This can be explained with the notion of the fringe visibility:

$$V = \frac{I_{max} - I_{min}}{I_{max} + I_{min}}. \quad (3.7)$$

The fringe visibility depends on the angular size of the source, as well as the baseline. The visibility is a measure for coherence, as when the measured wavefront is less coherent, the fringe visibility reduces. A stellar interferometer thus measures the degree of coherence through the fringe visibility. By changing the baseline, the visibility will vanish at some given separation, which then determines the angular size of the source. This is how Michelson measured the size of Betelgeuse, with the 100 inch (2.54 m) Hooker telescope [33]. It should be noted that a stellar interferometer can only form fringes when the path length between the apertures is equal, such that both beams are coherent. If a target is at zenith, or exactly on the line of sight of the instrument, the path length difference is zero when the beam combiner is placed in the middle of the baseline. However, when the target moves away from zenith, for example when the Earth rotates, the path length difference is equal to $B \sin \theta$, with B the baseline and θ the angle between zenith and the target. To compensate for this, a delay line is used on both arms, such that the path length can be controlled for each individual arm of the interferometer.

3.3. Nulling interferometry

In this section, the principle of nulling interferometry will be introduced, starting with the general principle in subsection 3.3.1. To extend the general principle, the effect of increasing the number of apertures and their spatial configuration will be considered in subsection 3.3.2. Finally, some practical implications of a potential nulling interferometer will be discussed in subsection 3.3.3.

3.3.1. General principle

In the previous section, the stellar interferometer was introduced, which allows us to see inside the PSF of the single aperture telescope. However, in order to observe exoplanets directly, the starlight that outshines the planet by a factor of one billion needs to be reduced sufficiently. In 1978, R.N. Bracewell proposed a novel idea: if one of the arms of a stellar interferometer could be delayed with a π phase, the dark fringe would appear in the centre of the baseline [3]. If this instrument then would be pointed towards a star, the light would be extinguished. On the other hand, light from the planet, when located inside the bright fringe, would reach the detector, as the path length difference ensures constructive interference. A basic schematic of a two aperture nulling interferometer is shown in Figure 3.3. It consists of two apertures, two delay lines, and a beam combiner. There are several ways to combine beams for interferometry, some of which will be mentioned in section 4.7. The bright, or constructive output can be discarded, or used for fringe tracking.

The fringes formed by the instrument are referred to as the transmission map, which will be discussed in more detail in section 3.4. It is a mapping that shows which part of the sky is transmitted and which part is extinguished. To know where the planet is, the interferometer needs to be rotated around the line of sight. In this way, the light from the planet is modulated onto the fringes. Depending on the modulation frequency, the location of the planet can be inferred, as is shown in Figure 3.4.

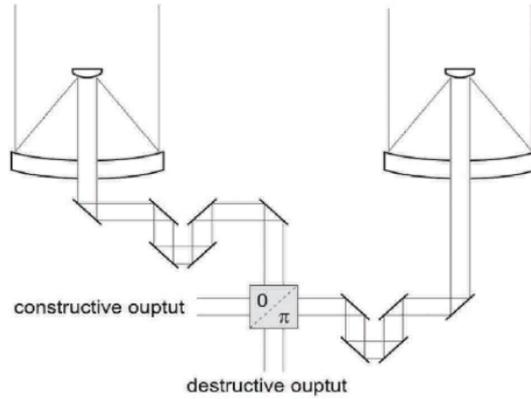


Figure 3.3: General schematic representation of a Bracewell nulling interferometer. The beam combiner in the middle is preceded by delay lines, which are used to equalize the path length between both apertures. Adapted from [34]

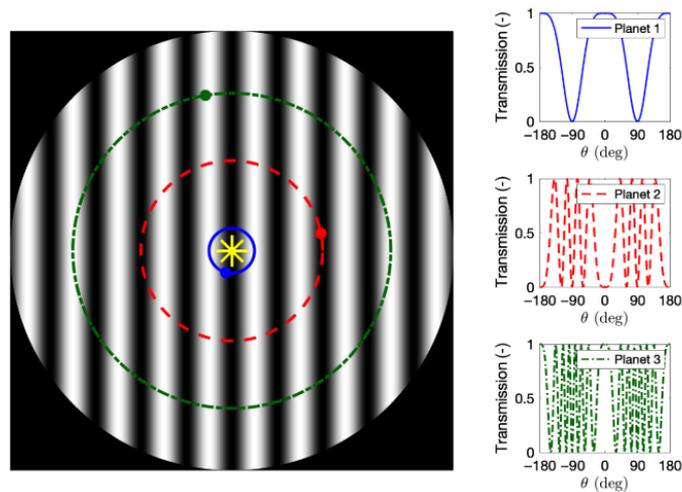


Figure 3.4: Left: Transmission map with a star in the centre and three planets with their corresponding trace through the fringes. (green / red / blue dots and circles). Right: The resulting modulation from rotating the transmission map. Adapted from [35]

3.3.2. Aperture configurations

The Bracewell interferometer consists of two apertures, however there is in principle no limit on the number of apertures that could be used for nulling interferometry. Naturally, a compromise has to be made between performance, cost and complexity of the overall system. Even so, a higher number of apertures does not necessarily mean a better performance of the nulling interferometer. To assess the effect of the number of apertures, several considerations have to be made. First, the response of the interferometer is proportional to:

$$R \propto (L\theta/\lambda)^p, \quad (3.8)$$

with L the length of the longest baseline between the apertures, θ the angle on sky, λ the wavelength and p the null order, corresponding to a specific aperture configuration [19]. The effect of the null order is that the null becomes wider at higher values of p and thus reduces the stellar leakage, as less starlight is transmitted to the detector (see Figure 3.6). Stellar leakage

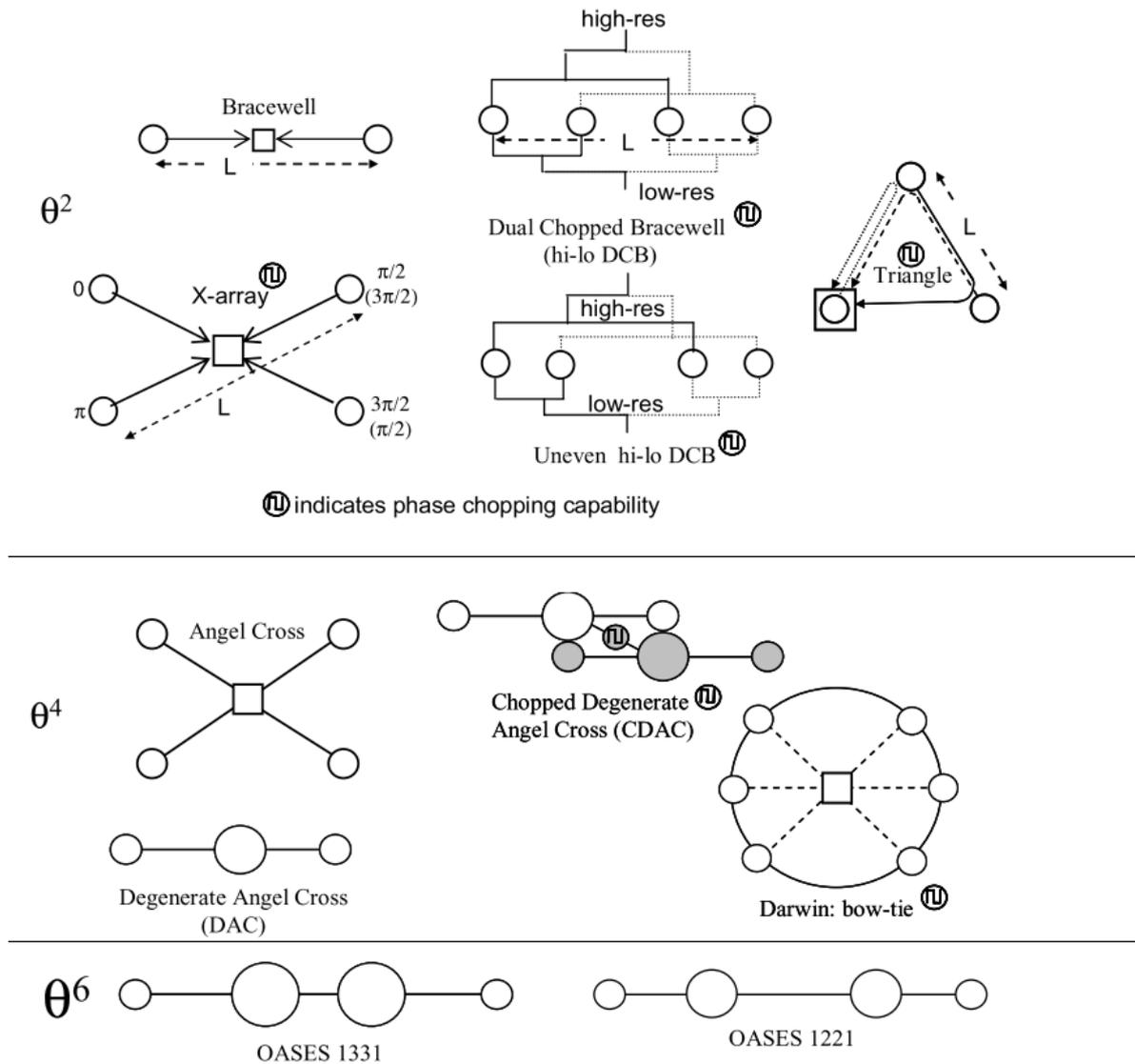


Figure 3.5: Selection of architectures considered in [19], grouped by their null order p .

occurs because a star has a finite extent and thus not all starlight is rejected with the same amount.

A selection of configurations was considered in [19], shown in Figure 3.5. In this schematic, the configurations are grouped by their null order p . The θ^2 configurations can consist of two, three or four apertures, whereas the θ^4 and higher null order configurations need at least three apertures. In context of the DARWIN mission, it was predicted that a four aperture X-array configuration would have the highest science return [37]. This result refuted the assumption that reducing the stellar leakage would directly translate to better performance.

Another important parameter in Equation 3.8 is the baseline length L . With a longer baseline, the fringes become narrower, increasing the stellar leakage. However, as shown by R. Norbruis through numerical modelling, having a longer baseline increases the planetary yield, despite the increased stellar leakage.

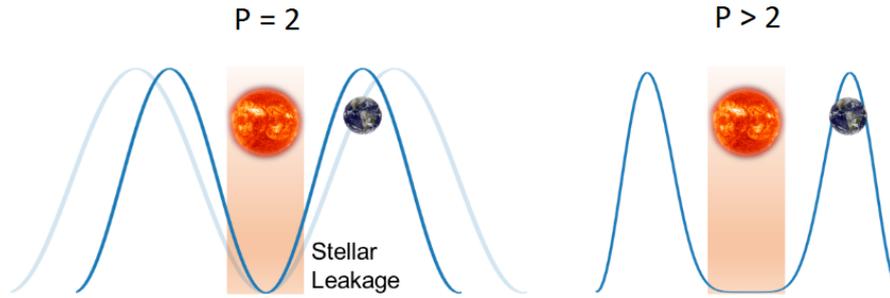


Figure 3.6: A central slice of the transmission map of a nulling interferometer. The effect of a higher null order is shown on the right, where the dark fringe is wider compared to the one on the left. Adapted from [36].

3.3.3. Practical implementation

When realising an actual nulling interferometer, the diagram in Figure 3.3 does not suffice. To discuss the practical implications of designing a nulling interferometer, consider the functional diagram given in Figure 3.7. For context, a brief discussion of each component will be given here, or in the referred sections. For a more in-depth treatment, one can refer to the literature or literature review.

- Optics - The main subject of this study, which will be discussed more extensively in chapter 4.
- Tip-tilt sensor / mirror - The tip-tilt mirror corrects tip-tilt errors in the optical path. The sensor can be as simple as a quadrant sensor to measure tip-tilt, but if wavefront errors also need to be measured, more elaborate sensors can be used, such as the Shack-Hartmann sensor [38], Pyramid wavefront sensor [39], or curvature wavefront sensor [40]. Although tip-tilt and wavefront errors could be in principle be corrected with a single deformable mirror, tip-tilt and wavefront errors are typically corrected separately, due to the limited stroke in a deformable mirror.
- ODL - Used to equalise the path lengths between the apertures. On ground, the largest contribution to OPD is due to the target being off-zenith. In space, this contribution is less significant, although pointing errors can still contribute to OPD. Optical delay lines will be discussed more extensively in section 4.6.
- Single mode fibre injection - To correct for higher order aberrations in the optics, as well as pointing errors, fibre injection can be used for spatial filtering.
- π Phase shift - The achromatic phase shifter is a critical component that shifts one of the input beams by π in the 2 aperture configuration, or fractions of π otherwise. There are several options available discussed in the literature:
 - Dispersive plate [41, 42]
 - Focus crossing mirrors [42, 43]
 - Mirror Periscope [42, 43]
 - Fresnel Rhomb [44, 45]

All these methods have been tested on testbenches [42, 46], which showed that the mirror periscope approach is the most promising.

- Beam combiner - A component that combines the beams for interferometry. See section 4.7.

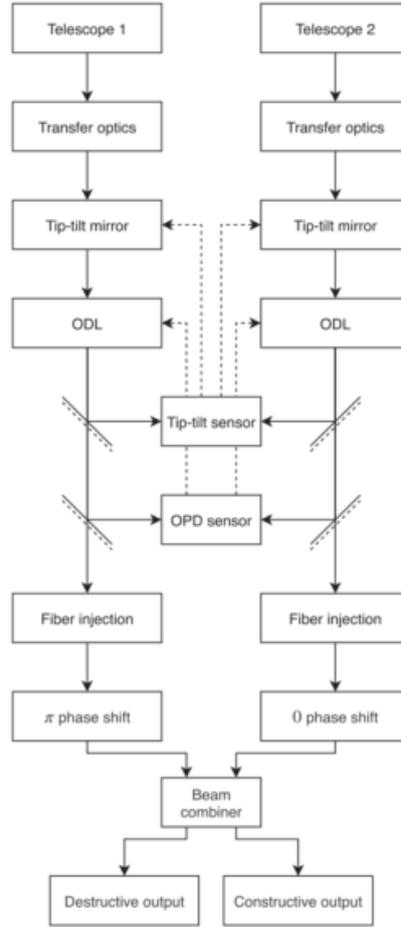


Figure 3.7: Functional diagram of a potential nulling interferometer with two apertures.

3.4. Transmission map

The transmission map shows which part of the field of view is nulled and which part of the field of view is transmitted to the detector. It can be seen as a sieve that only transmits light for certain angles on sky. The shape of the transmission map depends on the configuration of the apertures.

For a given baseline \vec{B} , and pointing vector \vec{s} , the OPD is given by $\vec{B} \cdot \vec{s}$. Since the target is on sky and the interferometer remains in the xy plane, it is most convenient to work with spherical coordinates. A given pointing vector expressed in terms of the angles on sky $\vec{s}(\alpha, \beta)$, can be expressed as a pointing vector in spherical coordinates $\vec{s}(\theta, \phi)$ as follows: A given angle on sky can be converted to spherical coordinates in the following way:

$$\theta = \cos^{-1}(s_z) \quad (3.9)$$

$$\phi = \tan^{-1}(\beta/\alpha), \quad (3.10)$$

where $s_z = \sqrt{1 - \alpha^2 - \beta^2}$ (assuming $\|\vec{s}\| = 1$). Then, to express the pointing vector in terms of spherical coordinates, the following transformation can be used:

$$\vec{s}(\theta, \phi) = \begin{pmatrix} s_x \\ s_y \\ s_z \end{pmatrix} = \begin{pmatrix} \sin(\theta) \cos(\phi) \\ \sin(\theta) \sin(\phi) \\ \cos(\theta) \end{pmatrix}. \quad (3.11)$$

With the pointing vector defined, the OPD can be calculated for each field angle in the FoV:

$$OPD(\theta, \phi) = 2\pi \frac{\vec{B} \cdot \vec{s}(\theta, \phi)}{\lambda} \quad (3.12)$$

As the baseline is defined on a plane, assuming the telescopes are well aligned in space, there is no z component in the OPD:

$$\varphi_{OPD}(\theta, \phi) = \frac{2\pi}{\lambda} \begin{pmatrix} B_x \\ B_y \end{pmatrix} \cdot \begin{pmatrix} \sin(\theta) \cos(\phi) \\ \sin(\theta) \sin(\phi) \end{pmatrix} \quad (3.13)$$

Each field coming into the aperture receives an optical path difference, compared to the chosen reference (one of the apertures, or the centre of the array), depending on the angle on sky. The optical path difference results in a piston, or constant phase over the detector plane, which is added to the complex amplitude. The resulting transmission for a given field angle (α, β) , is then calculated by integrating the square of the complex amplitude at the detector plane:

$$T(\alpha, \beta) = \left| \sum_n A_n \cdot e^{i(\varphi_{OPD}(\alpha, \beta) + \Delta\varphi)} \right|^2, \quad (3.14)$$

where the sum goes over each aperture. The intensity on the detector depends on the brightness distribution on sky and the transmission:

$$I_{detector} = \iint B(\alpha, \beta) T(\alpha, \beta) d\alpha d\beta \quad (3.15)$$

In the case of polychromatic light, each wavelength is added incoherently:

$$I_{total} = I_{\lambda_1} + I_{\lambda_2} + \dots + I_{\lambda_n} \quad (3.16)$$

This gives the intensity in Watts for each angle on sky, however, the transmission map ought to be normalised, so that it is defined in terms of the null depth. The transmission maps, as calculated from Equation 3.14, are shown in figure Figure 3.8 for 2, 3 and 4 apertures, with a baseline of 10 meters and using a wavelength of 10 μm .

3.5. Chapter summary

In this chapter, the theoretical foundation was laid for nulling interferometry. Through Young's double slit experiment, the concept of constructive and destructive interference was introduced. Second, by means of stellar interferometry, it was shown that it is possible to "see" inside the PSF of a single aperture, when two apertures are separated by a baseline B . Nulling

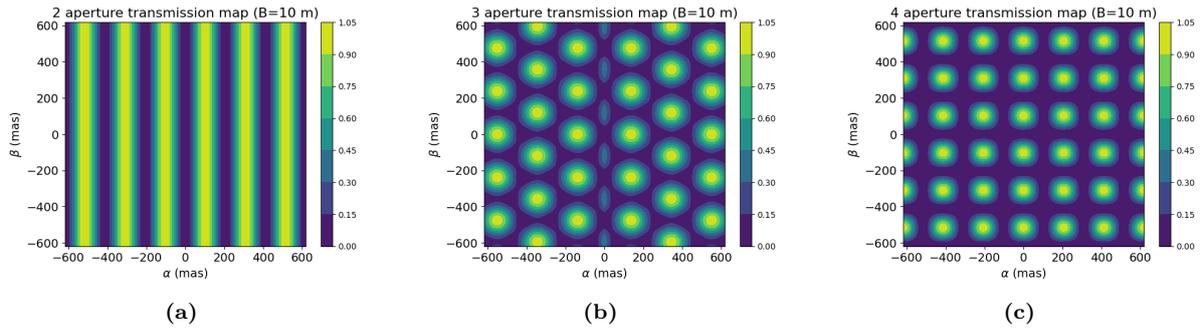


Figure 3.8: Simulated monochromatic transmission maps, based on Equation 3.14, of (a) 2 apertures (b) 3 apertures and (c) 4 apertures, using a baseline of 10 meters and wavelength of $10 \mu\text{m}$.

interferometry, which was discussed in section 3.3, relies on the same principle, but with an achromatic π phase shift, such that the destructive fringe appears on axis, which allows for extinction of the starlight, while the bright fringes can be used to detect the exoplanet. In this section, the effect of different aperture configurations was briefly discussed, as well as some practical aspects of an operational nulling interferometer. Finally, in section 3.4, the theory of deriving a transmission map was given.

4

Optical design

This chapter is concerned with the optical design of the nulling interferometer. First, in section 4.1, the basic properties of an optical system are stated. Then, the requirements driving the optical model are discussed in section 4.2. In section 4.3, the design parameters of several on-axis telescope designs are considered. The theory needed to design an off-axis telescope that performs the same as its on-axis counterpart is provided in section 4.4. To collimate the beam, an off-axis parabola will be used, which is discussed in section 4.5. Finally, the choice of delay line and beam combiner will be considered in section 4.6 and section 4.7, respectively.

4.1. Basic properties of optical systems

Before telescope designs are considered, it is useful to briefly review the basic properties of optical systems and curved surfaces.

First, any optical system can be characterised by its entrance and exit pupil. Second, the f-number ($f/\#$) of a system is defined as the focal length divided by the aperture diameter. A 'fast' optical system has a lower $f/\#$, compared to a 'slow' optical system. For a given aperture diameter, a fast optical system is thus more compact compared to a slow optical system.

A curved mirror is characterised by two parameters: its radius of curvature and its conic constant. The radius of curvature (RoC) determines the focal length through the relation $f = R/2$ ¹. When the RoC is negative, the mirror surface is concave and forms a real inverted image. When the RoC is positive, the mirror surface is convex and forms a virtual image. The surface shape is determined by the conic constant, related to the conic sections:

4.2. Requirements

The optical design is driven by several requirements:

¹For a tilted surface, the focal length is given by $f = R/(1 + \cos \theta)$

Conic constant	Conic section
$K < 1$	Hyperbola
$K = -1$	Parabola
$-1 < K < 0$	Prolate ellipse
$K = 0$	Circle
$K > 1$	Oblate ellipse

Table 4.1: Conic constants and their corresponding conic sections.

1. The spectral bandwidth of the nulling interferometer will range from 4-18.5 μm , which is similar to the proposed bandwidth of LIFE [47]. This is seemingly a very broad bandwidth, but is not impossible to attain. For example, the MIRI instrument on the JWST has a spectral bandwidth of 5 to 28 μm [48, 49]. Although a single focal plane could capture the complete spectrum, anti-reflection coatings are optimised for a target wavelength. Furthermore, the modulation efficiency [47] of a nulling interferometer decreases with increasing bandwidth, such that it would be favourable for that metric to divide the spectral bandwidth into several channels, for example 4-8 μm , 8-12 μm , 12-16 μm and 16-18.5 μm .
2. The Inner Working Angle (IWA) of an interferometer is defined as the angular separation between the first dark and bright fringe, given by $\lambda/(2B)$. In order to resolve an Earth-like planet in the HZ of a main-sequence star, the IWA should be on the order of 100 mas, given an orbital radius of 1 AU around a star at a distance of 10 pc. As the IWA scales linearly with the wavelength, the longest wavelength in the spectrum drives the baseline requirement. However, since the modulation efficiency is non-zero for angular separations smaller than the IWA, it does not necessarily pose a lower limit on the baseline. For example, for the longest wavelength of 18.5 μm , the baseline should be 19 m to obtain an IWA of 100 mas. Despite this, the baseline requirement is limited to 15 m for the proposed LIFE mission [47], which puts the IWA at 127 mas for a wavelength of 18.5 μm . For the purposes of this study, a baseline of 15 m will be used. It should be noted that, although a longer baseline does increase the geometrical stellar leakage, it is still favourable in terms of the number of planets detected, as was pointed out by R. Norbruis [36].
3. The Outer Working Angle (OWA) is given by λ/D and is the angle beyond which there is no transmission. Increasing the single aperture area increases the signal to noise ratio and thus the planetary yield [36] as long as all observed planets are within the OWA. There are several factors that limit the aperture size. First, the size of the launcher that transports the telescope to space should be considered. For example Ariane 6 has a fairing diameter of 5.4 m². Using the full fairing diameter would be infeasible, considering multiple telescopes would be needed for the interferometer, as well as a beam combiner spacecraft. Second, the cost of manufacturing mirrors scales with the aperture diameter: $\text{Cost} \propto D^{1.6}$ [50, p. 97]. Third, for a given focal ratio, aberrations increase with increasing aperture diameter, according to the scaling laws given in Table 4.2.

Taking these points into consideration, a starting point of a 1 m aperture is chosen, which showed promising results for several aperture configurations [36]. Given that the OWA scales linearly with wavelength, the FoV of the telescope should be at least 825 mas for the shortest wavelength of 4 μm and 3.7 arcseconds at a wavelength of 18.5 μm .

²<https://www.arianespace.com/vehicle/ariane-6/>

Aberration	Scaling law
Spherical	$(f/\#)^{-3}$
Coma	$\theta(f/\#)^2$
Astigmatism	$\theta^2(f/\#)^{-1}$
Field curvature	$\theta^2(f/\#)^{-1}$
Distortion	θ^3

Table 4.2: Scaling laws of the primary aberrations, as function of the $f/\#$ and field angle θ . Adapted from [50, p. 111].

The requirements are thus stated as follows:

Parameter	Requirement
Baseline	15 m
Spectral bandwidth	3-20 μm
IWA	127 mas @18.5 μm
OWA	825 mas @4 μm 3.7 arcseconds @18.5 μm
Aperture diameter	≥ 1 m

Table 4.3: System requirements for the MIR nulling interferometer.

4.3. On axis telescope designs

In this section, an overview will be given of the three classic reflective telescope designs: Newtonian (subsection 4.3.1), Cassegrain (subsection 4.3.2) and Gregorian (subsection 4.3.3).

4.3.1. Newtonian telescope

The Newtonian telescope is the most simple of all reflective telescope designs. It consists of a parabolic primary and a secondary flat to angle the focus away from the optical axis. The main benefit of this design is its simplicity, as only one curved surface has to be manufactured. However, this simplicity is also a disadvantage, as the size of the telescope is on the same order as the focal length of the primary mirror. Additionally, due to the fact there is only a single parabolic mirror, a Newtonian telescope is only free of spherical aberration

4.3.2. Cassegrain

The Cassegrain is a widely used telescope design, due to its relative simplicity and good image quality. It consists of a parabolic primary and hyperbolic secondary.

The design parameters of a Cassegrain are the effective focal (f'), the back focal length (L) and the distance between the primary and secondary mirror (d_1). From those parameters, the surfaces can be defined. For the primary mirror, the focal length f'_1 is given by (following the notation from Wilson [51]):

$$f'_1 = \frac{f'}{m_2}, \quad (4.1)$$

and the conic constant K_1 is chosen to be -1. The magnification of the secondary mirror, m_2 , can be calculated as follows:

$$m_2 = \frac{f' - L}{d_1}, \quad (4.2)$$

from which the secondary mirror focal length can be calculated:

$$f'_2 = \frac{L}{m_2 + 1}. \quad (4.3)$$

Finally, the conic constant of the secondary mirror can be expressed as function of m_2 :

$$K_2 = 1/(b_{s2}) = - \left(\frac{m_2 + 1}{m_2 - 1} \right)^2. \quad (4.4)$$

A more modern variant of the Cassegrain is the Ritchey-Chrétien telescope, which has a hyperbolic primary, instead of a parabolic primary. In that case, the conic constant of the primary mirror is given by:

$$K_1 = -1 - \frac{2L}{d_1 m_2^3}, \quad (4.5)$$

and the conic constant of the secondary mirror:

$$K_2 = 1/(b_{s2}) = - \left[\left(\frac{m_2 - 1}{m_2 + 1} \right)^2 + \frac{2f'}{d_1(m_2 + 1)^3} \right]^{-1} \quad (4.6)$$

4.3.3. Gregorian

The Gregorian telescope is a two mirror telescope with an intermediate image between the primary and secondary mirror. It is made up of a parabolic primary and elliptical secondary. To calculate the primary and secondary mirror focal length and secondary mirror conic constant, Equation 4.1 to Equation 4.4 can still be used, with the only alteration that the effective focal length is taken to be negative. Although the Cassegrain is generally favoured over the Gregorian design, it does have the particular advantage of having an intermediate focus. This is especially useful for solar observatories, as a field stop can be placed at the intermediate focus, to limit the amount of heat transported to the secondary mirror. For nulling interferometry, a field stop could also be useful, either to limit the amount of stray light coupling into the beam combiner, or for spatial filtering.

4.4. Off-axis telescope designs

on-axis telescope designs have been the archetypal form ever since the inception of the telescope. As this type of telescope is a very mature instrument and relatively cheap to manufacture, compared to more complex designs, it has been used in many space based observatories.

However, there are multiple drawbacks to using an on-axis design, which degrade the performance of an interferometer. First, the central obscuration inherent to such a design degrades the fibre coupling efficiency. A single mode fibre, as the name suggests, propagates the fundamental mode through the fibre core, without significant loss of energy. The coupling efficiency can be expressed as an overlap integral between the incoming electric field and the fundamental mode of a single mode fibre [52]. When the incoming light matches the envelope of this fundamental mode, the light is coupled and transmitted optimally into the fibre. If there is a mismatch in the profile of the incoming radiation and the fundamental mode, the coupling efficiency is reduced. Such a difference in profile is illustrated in Figure 4.1. A central obscuration changes the shape of the electric field and thus worsens the mismatch between the fundamental mode and incident electric field.

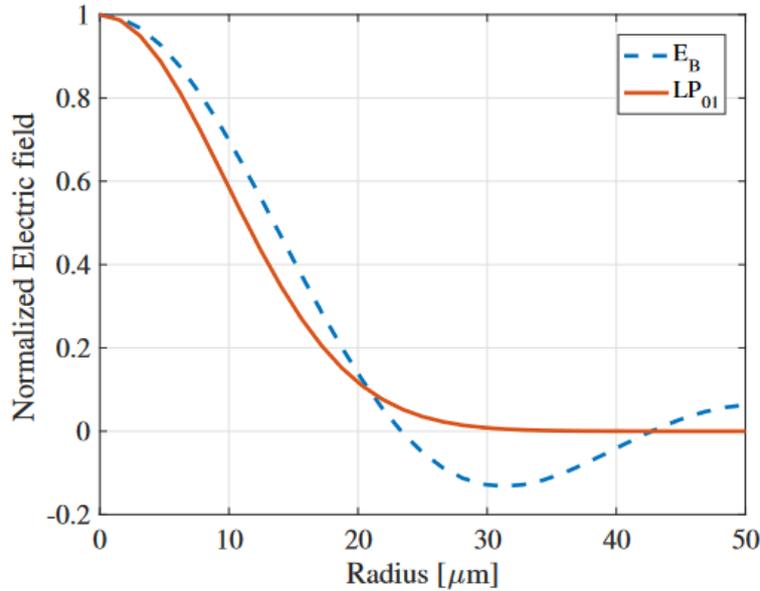


Figure 4.1: Comparison between the fundamental mode of a single mode fibre (red) and the electric field at the focus of a telescope without a central obscuration. Adapted from [53].

Spatial filtering can reduce the effect of higher order aberrations and thus ease the stability requirements on the spacecraft, which improve the stability of the null [54]. Because of this, virtually every nulling interferometer uses some form of spatial filtering. As was shown by Ait Hocine [53] there exists an analytical relation between the obscuration ratio α and the coupling efficiency ρ :

$$\rho(\alpha, \beta) = 2 \left[\frac{e^{-\beta^2} - e^{-(\beta^2\alpha^2)}}{\beta\sqrt{(1-\alpha^2)}} \right]^2, \quad (4.7)$$

where β is a constant factor depending on single mode fibre parameters and the wavelength of the incoming light. From this, it is straight forward to deduce that as α goes to 1, the coupling

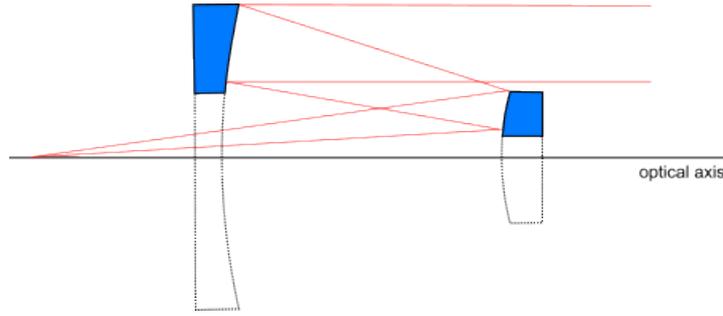


Figure 4.2: A conventional off-axis telescope system, using an eccentric section of the conventional on-axis system, such that the optical axis of both parent mirrors co-align. Adapted from [30].

efficiency goes to 0.

Another significant effect with on-axis designs is the loss of PSF quality, due to the presence of the secondary mirror, as well as the support structure to hold up the secondary mirror. The support structure creates a particular "spider-like" pattern on the PSF and reduces the encircled energy within a given radius, although the largest effect comes from the central obscuration. So, in order to avoid a loss of coupling efficiency and to maximise the throughput, off-axis designs have to be considered.

There are many different ways to design an off-axis telescope. To list some high level design choices; free-form surfaces can be used, or conventional conic sections, a telescope can consist of two, three or more mirrors and in the case of conic sections, the parent mirror axes might be co-aligned or not. A frequently used three mirror design is the Three Mirror Anastigmat. Because it has three surfaces, it corrects for spherical aberration, coma as well as field curvature. This allows for a wide field of view, compared to two mirror telescopes. However, for the purpose of nulling interferometry, a wide field of view is not needed, as the outer working angle typically is only a few arcseconds. The added mirror adds an extra loss term in the transmission, as well as adding degrees of freedom which perturb the null. For this reason, only two mirror designs are considered.

The most basic off-axis system uses an eccentric section of a classic on-axis design, which means that the parent mirror axes are coinciding. However, it is not immediately clear whether such a system is optimal or not, in terms of optical performance. In order to establish this, Chang et al. formalised a geometrical aberration theory for two mirror off-axis designs [29]. The result of this theory is a closed-form equation which shows how linear astigmatism can be eliminated for such a system. The theory has shown that using a confocal system, where the focus of the primary and secondary mirror coincide, instead of using a co-aligned optical axis, is the most optimal design.

As this theory requires a considerable amount of geometry and algebra, the most important equations for designing the system are stated here. A schematic of an example off-axis confocal system is given in Figure 4.3. There are two angles of interest to the current analysis: i_2 and θ_0 , as those can be calculated from the other design parameters (l'_1 , l_2 , l'_2 and i_1). i is the angle between the reference ray and the axis normal to the surface vertex, which can be thought of as the tilt angle. The second angle, θ_0 is the angle between the parent mirror axis and the axis normal to the surface vertex.

Note that when $\theta_{0,1} = \theta_{0,2}$, the parent mirror axes are co-aligned, in which case the system

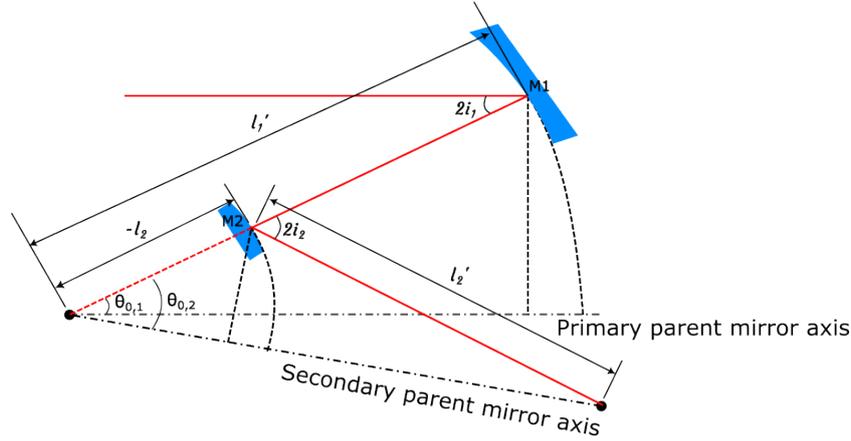


Figure 4.3: Distances and angles relevant to the off-axis mirror system. l_1' is the primary mirror (M1) focal length, l_2 is the secondary mirror (M2) focal length, l_2' is the telescope back focal length, i_1 is the primary mirror tilt angle and i_2 is the secondary mirror tilt angle. Adapted from [30].

is referred to as a common off-axis system. When the angles are not equal and the focal point of M1 and M2 coincide, the system is referred to as a confocal off-axis system. To fully characterise the system, four lengths are needed:

- l_1' - The focal length of the primary mirror.
- $-l_2$ - The focal length of the secondary mirror.
- l_2' - The back focal length of the mirror system.
- l_1 - The distance from the primary mirror to the object plane. Taken to be ∞ in the rest of the text.

These three lengths, together with the angle i_1 are design parameters, while the other angles and radii of the mirror surfaces can be calculated. As an example, let us consider an off-axis Cassegrain telescope. For such a system, the primary mirror is parabolic ($K_1 = -1$), while the secondary mirror is hyperbolic ($K_2 < -1$). To calculate R_1 , equation 3 from [30] can be used:

$$\frac{1}{l_1'} = \frac{2 \cos^2 i_1}{R_1} - \frac{1}{l_1} \quad (4.8)$$

$$R_1 = 2l_1' \cos^2 i_1. \quad (4.9)$$

To calculate the secondary mirror tilt angle i_2 , relative to the optical axis ray, equation 38 from [30] can be used:

$$\frac{l_1 + l_1'}{l_1} \tan i_1 = \frac{l_2 + l_2'}{l_2} \tan i_2 \quad (4.10)$$

$$i_2 = \tan^{-1} \left(\frac{l_2' \cdot (l_1 + l_1')}{l_1 \cdot (l_2 + l_2')} \tan(i_1) \right). \quad (4.11)$$

Given that l_1 is infinite, Equation 4.11 can be further simplified to:

$$i_2 = \tan^{-1} \left(\frac{l'_2}{l_2 + l'_2} \tan(i_1) \right). \quad (4.12)$$

Note: in Code V, there are two methods of decentering and tilting surfaces. The method 'Decenter and bend', bends the optical axis after reflection, such that the optical axis of the next element co-aligns with the reflected ray. The method 'Decenter and return' returns the reference frame to that of the previous surface. In that case, the tilt angle should be defined relative to the optical axis of the previous surface, prior to reflection, which is simply $\alpha_2 = 2i_1 - i_2$

With this angle determined, the secondary mirror radius follows from equation 23 of [30]:

$$R_2 = \frac{2 \cos^2 i_2}{\frac{1}{l_2} + \frac{1}{l'_2}}. \quad (4.13)$$

The conic constant of the secondary mirror, K_2 can be determined with trigonometry and the definition of conic sections, although this method is relatively cumbersome. Alternatively, K_2 can be calculated from conic relations:

$$K_2 = \frac{R_2}{a} - 1, \quad (4.14)$$

with a the semi major axis of the ellipse subtended by M2 and its focii, given by $(l_2 + l'_2)/2$. Finally, the axis tilt angle of the secondary mirror, $\theta_{0,2}$, can be calculated using equation 30 from [29]:

$$K_2 \sin^2 \theta_{0,2} = -\sin^2 i_2. \quad (4.15)$$

Using trigonometric identities, $\theta_{0,2}$ can be isolated:

$$\frac{1 - \cos(2\theta_{0,2})}{2} = -\frac{\sin^2 i_2}{K_2}, \quad (4.16)$$

which can be rearranged to the final form:

$$\theta_{0,2} = \cos^{-1} \left[\frac{1}{2} \left(1 + 2 \frac{\sin^2 i_2}{K_2} \right) \right]. \quad (4.17)$$

With these equations defined, it is very straight forward to design an off-axis telescope that rejects linear astigmatism, as well as spherical aberration, in case a non-spherical primary mirror is used. Some Code V models will be shown in subsection 6.1.1. In the following section, the beam combiner will be discussed

4.5. Beam collimation

To transport the optical beam through the interferometer effectively, it needs to be collimated. This can be achieved with a single off axis parabola, or more complex systems, like

the Schwarzschild objective [55, 56], as used in the Laser Interferometer Space Antenna (LISA) [57]. In the case of the collimator, it is desirable to have more degrees of freedom, such that the sub-assembly can be used to compensate for translational and rotational perturbations of the primary and secondary mirror. However, due to the complexity of designing such a system, the single off axis parabola was chosen to collimate the beam.

To calculate the tilt and off axis angle for the collimator, equations and angles from the previous section can be used. A schematic is given in Figure 4.4. In this figure, the optical path from the secondary mirror to the collimator and corresponding reflection is shown. Here, γ is the angle between the global horizontal, defined by the plane perpendicular to the object plane, and the OAR before reflection. α_c is the collimator tilt angle, defined relative to the global horizontal. β is the angle of incidence, which in this case is equal to the off axis angle.

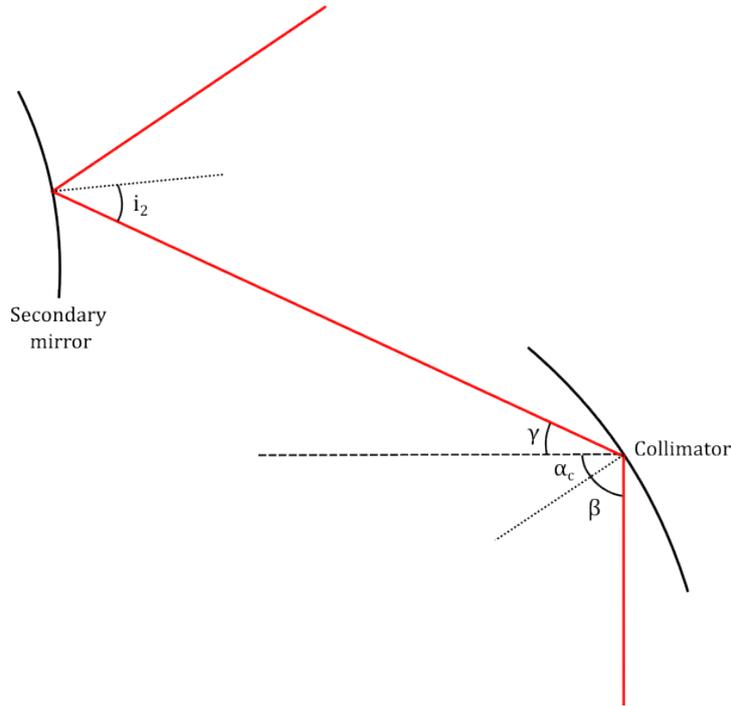


Figure 4.4: Schematic of the secondary mirror and collimator. β is the angle of incidence/reflection. γ is the angle suspended by the optical axis ray after reflection off the secondary mirror and the global horizontal. The tilt angle α_c is the angle between the global horizontal and the vertex.

The collimator is designed in such a way that the ray after reflection makes a 90° angle with the global horizontal. γ can be determined using basic geometry, and is related only to i_1 and i_2 : $\gamma = 2i_2 - 2i_1$. The angle of reflection, β can be expressed in terms of α_c and γ : $\beta = \alpha_c + \gamma$. Then, using the condition that $\alpha_c + \beta = 90$, the collimator tilt angle and off axis angle can be calculated:

$$\alpha_c = 45 - i_2 + i_1 \quad (4.18)$$

$$\beta = 45 + i_2 - i_1 \quad (4.19)$$

From Equation 4.18 and Equation 4.19, it can be seen that both the collimator tilt angle and off axis angle depend on the primary mirror tilt angle and are only equal when $i_1 = i_2 = 0$.

The focal length will be a design parameter, such that the radius of curvature depends on the focal length, as well as the off axis angle:

$$R_c = f_c(1 + \cos(2\beta)) \quad (4.20)$$

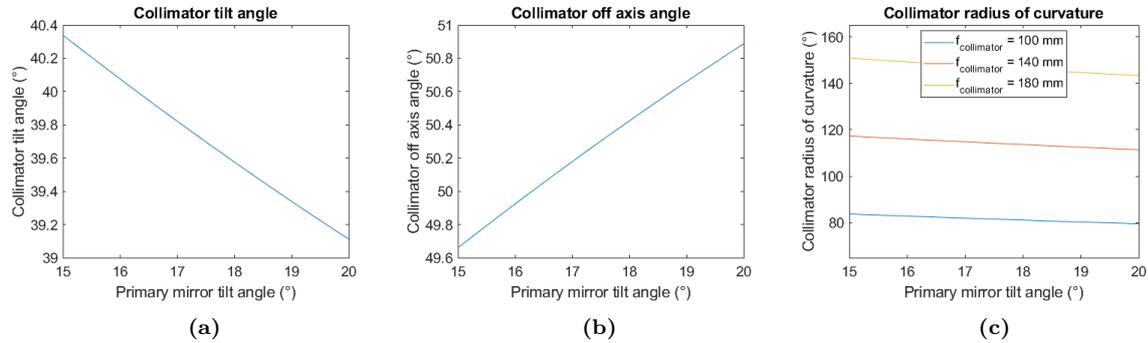


Figure 4.5: Dependencies of the collimator surface parameters on the primary mirror tilt angle. (a) Collimator tilt angle, as function of primary mirror tilt angle. (b) Collimator off axis angle, as function of primary mirror tilt angle. (c) Collimator radius of curvature as function of primary mirror tilt angle.

4.6. Delay line

Delay lines generally come in two forms, shown schematically in Figure 4.6:

1. A simple "rooftop" orientation, where two mirrors are angled at 90 degrees with respect to each other and both tilted by 45 degrees to the optical axis. Such a design is used for the 85 m long delay line in the W.M. Keck observatory [58], allowing the optical path to be delayed by up to 170 m.
2. A cat's eye reflector consists of a primary parabolic mirror and a fold flat. This setup is used for the VLTI long delay lines, as well as the proposed precursor of TPF-I: FKSI.

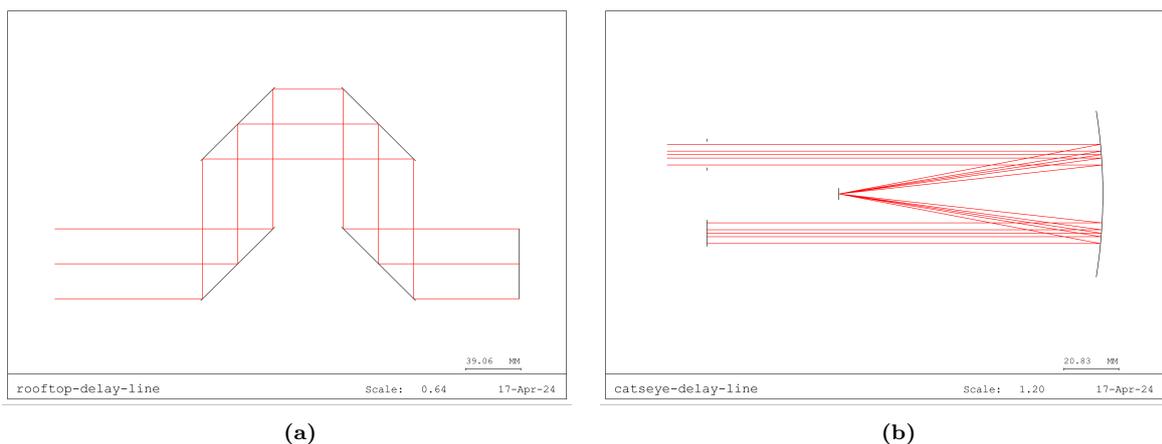


Figure 4.6: (a) Rooftop delay line. (b) Catseye delay line, consisting of a parabolic primary mirror and a fold flat.

The cat's eye design uses only two reflective surfaces and is thus more compact and lightweight compared to the rooftop design, while it also has less degrees of freedom. However, given that

a cat's eye reflector could introduce additional aberrations, the rooftop design is chosen, such that the only sources of aberrations come from the telescope optics.

4.7. Beam combiner

4.7.1. Beam combiner types

Generally speaking, there are three ways to do beam combination:

- Bulk optics, using beam splitters, dichroics and compensator plates³. Examples of nulling interferometers that have used this type of beam combiner is the Large Binocular Telescope Interferometer (LBTI) [41], the Keck Interferometer Nuller (KIN) [8, 59], PEGASE [22] and the Fourier Kelvin Stellar Interferometer (FKSI) [23]. It should be noted that although PEGASE and FSKI never materialised, their preliminary design did include a Mach Zehnder beam combiner.
- Using fibre optics. Fibre based splitters, or couplers, use the phenomenon of evanescent waves, which are waves that exponentially decay, as opposed to for example a plane wave, that can in principle propagate indefinitely. In an optical fibre, the coupled field propagates through the core. However, there is also a part of the field that dissipates through the cladding as an evanescent wave. When two fibre cores are brought together, the evanescent wave can couple to the neighbouring fibre core and continues propagating through the core. One notable implementation of this method was the FLUOR fibered beam combiner for interferometry at the IOTA interferometer [60] and later at the CHARA array [61], combining beams in the K band (2.0-2.4 μm) and L band (3.4-4.1 μm). Despite the simplicity of this method, it never gained popularity as a beam combination scheme for nulling interferometry. This is due to two reasons: First, the bandwidth of fibre couplers is relatively limited, typically on the order of 10-20%. Second, despite the limited bandwidth, the splitting ratio is chromatic, varying on the order of 10% (splitting ratio of 40/60 instead of 50/50).
- Integrated Optics / photonics. This technique works on the same principle as the previously mentioned fibre optics beam combination. Instead of using optical fibres, single mode waveguides are etched into a substrate, which reduces the size of the beam combiner significantly. In addition to the reduced size, an advantage of integrated optics is the lack of need for alignment. Notwithstanding these advantages, integrated optics has only limited use for the purpose of nulling interferometry in the mid-infrared, since an achromatic, broadband mid-infrared coupler is yet to be developed. It should be noted that in recent years, several projects using integrated optics on nulling interferometers have been proposed and demonstrated, including NOTT (formerly known as Hi-5) [9], GLINT [62] and a 4 input photonic kernel nuller, demonstrated on the VLTI [63].

The beam combiner of choice for the studied nulling interferometer will be based on bulk optics. Given the large bandwidth of the nulling interferometer of 4-20 μm and the relatively small bandwidth of integrated optics and optical fibres, many separate channels would be needed to reach the required spectral bandwidth. The chromaticity, meaning differences in splitting ratio

³A compensator plate is a (sometimes wedged) piece of glass to compensate for the optical path difference that occurs after reflection off a plate beamsplitter.

(photonics) or coupling efficiency (fibre optics beam combiner), is detrimental to the attainable null, as will be analysed in section 6.4.

4.7.2. Bulk optics beam combiners

When considering bulk optics beam combiners, using beamsplitters, there are still several options available, including a single beam splitter, a Michelson interferometer⁴, a Sagnac and Mach-Zehnder interferometer. In 2001, Serabyn published a paper detailing the requirements and need for a fully symmetric beam combiner. In this publication, three configurations were given for beam combiners derived from classical interferometers (Michelson, Sagnac and Mach-Zehnder), shown in Figure 4.7. Each beam combiner has two inputs and four outputs, two of which are balanced and two of which are unbalanced. As can be explained with the right-and-left incidence theorem [64], the complex transmissive coefficient t and its reciprocal⁵ t' are equal, regardless of internal absorption, or other properties of the beam splitter or input beam. Because of this, the difference $r(t' - t)$ (balanced) is zero, while the difference $t(r' - r)$ (unbalanced) depends on the properties of the input beam, wavelength, polarisation state and angle of incidence. Of the three configurations presented in Figure 4.7, the Mach-Zehnder (MZ) beam combiner has the least amount of reflective surfaces, which gives it a few distinct advantages over the other two options: First, the reflective loss is minimised. Second, the interferometer is more compact, lightweight and cheaper. Lastly, due to cosmic radiation, reflective surfaces and coatings can degrade over time due to contamination. Reducing the amount of reflective surfaces minimises the loss in system transmission.

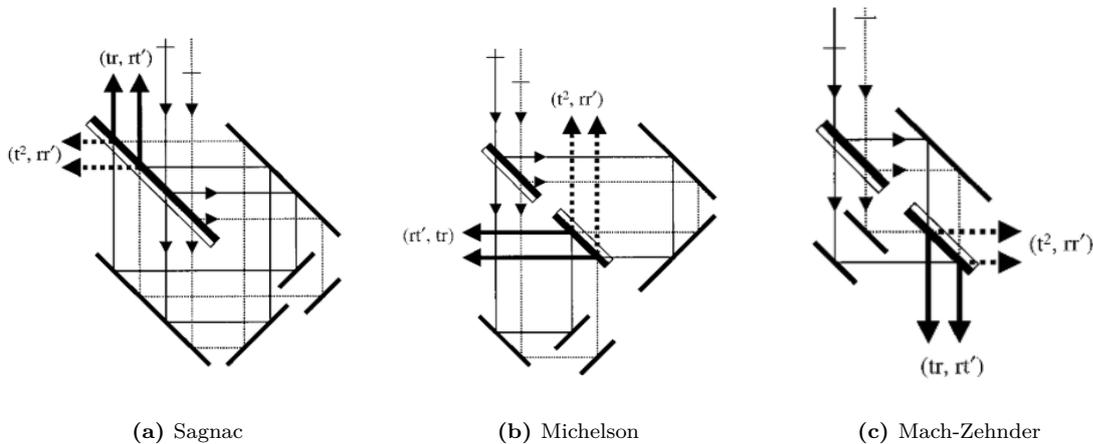


Figure 4.7: Symmetric beam combiner designs, based on classic interferometers. The bold arrows indicate the balanced output, while the dotted arrows indicate the unbalanced outputs. Schematic taken from [65].

4.7.3. Modified Mach-Zehnder beam combiner

Given the arguments in the previous section, a Modified Mach Zehnder (MMZ) beam combiner was designed in Code V as beam combiner for the nulling interferometer. With a classical MZ

⁴Although the Michelson interferometer also uses a single beam splitter, the distinction is made here between the single beam splitter scheme, where the input beams originate from two different sides of the beam splitter and no reflective optics are needed to combine the beams, in contrast to the Michelson interferometer, where both beams originate from the same side.

⁵ t is the transmissive coefficient corresponding to an input beam incident from the front of the beamsplitter (black shaded area), whereas t' corresponds to an input beam incident from the backside of the beamsplitter.

beam combiner, a single beam is split into two beams by the first beam combiner, after which both beams are recombined into a constructive and destructive output. The constructive output has no relative phase shift between the beams, while the destructive output has a relative π phase shift. The reason between the difference of the constructive and destructive output is due to the optical path taken through the beam combiner and the difference in reflections encountered in the optical path. In an MZ beam combiner, the two beam splitters are facing in opposite directions, meaning that their half-reflective surfaces are on opposite sides, one facing up and one facing down, as shown in Figure 4.8. This orientation is to make sure the optical path length between both paths is the same. The relative π phase shift can be explained with the Fresnel coefficients. When a field is reflected off a mirror in air, the phase is shifted by π . However, when a field is inside a refractive medium and splits into a transmitted and reflected beam, the latter of which goes back into the refractive medium, the beam is not phase shifted after reflection.

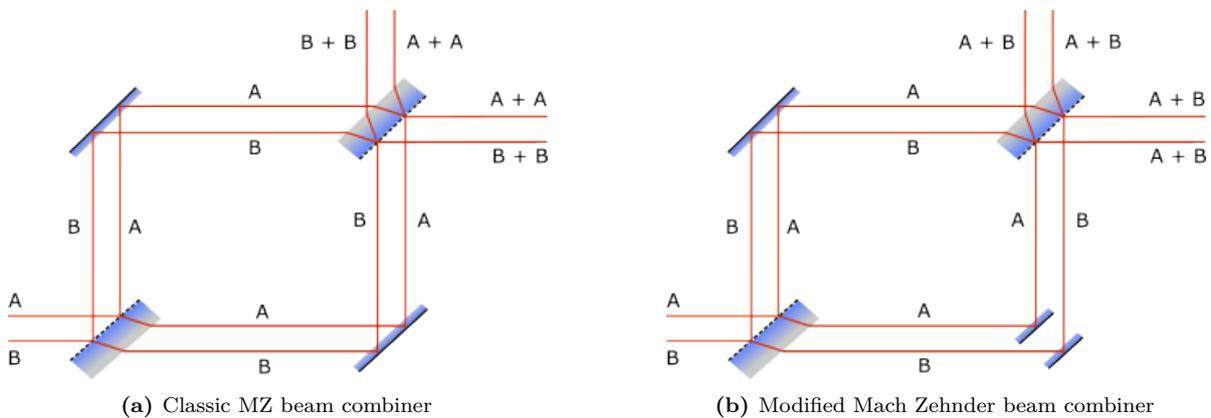


Figure 4.8: A classic (left) and modified MZ (right) beam combiner, with two input beams, A and B. The dashed line indicates the half transmissive coating on the beam splitter.

The need for the MMZ can be seen from Figure 4.8a. In this configuration, beams A and B are not cross-combined, which means that there is no interferometry between beam A and B. A simple solution is to split the right bottom mirror into two smaller mirrors (see Figure 4.8b), such that beams A and B are combined after the second beam splitter.

4.8. Chapter summary

In this chapter, several considerations were made concerning the optical design of a nulling interferometer. First, the basic properties of optical systems were summarised. In section 4.2, the requirements for the nulling interferometer were given, which are summarised in Table 4.3. In section 4.4, it was argued that an off-axis telescope has several advantages over an on-axis telescope design, including increased throughput and fibre coupling efficiency. The remainder of the section was concerned with the specific design parameters of an off-axis telescope design. As pupil plane beam combination will be used, the telescope needs to be afocal, which is achieved with a single off-axis parabola. For the delay lines, a simple rooftop design will be used. As for bulk optics pupil plane beam combiners, there are several options available, including the classical Sagnac, Michelson and Mach-zehnder (MZ) beam combiner. Of those three options, the MZ beam combiner is the most compact and lightweight.

5

Methodology

5.1. Modelling split apertures and beam combination

An individual aperture of a (nulling) interferometer can be seen as part of a larger aperture that is subtended by the individual apertures. An incoming wavefront is thus 'sampled' by N apertures. There is no proper way to model this in Code V, so a trick has to be applied. Instead of sampling a larger wavefront with individual apertures, a single, smaller wavefront is propagated through each arm of the interferometer. Each arm is modelled individually by what is called a 'zoom' in Code V, which is equivalent to 'multi-configuration' in OpticStudio (formerly known as Zemax)¹. The same problem arises when modelling a beam splitter in Code V. As there is no way to split a wavefront into two parts, a beam splitter has to be modelled by two zooms that each model either the reflected or transmitted path. An example is given in Figure 5.1. In this example, there are two zooms, one for the reflected path and one for the transmitted path. Figure 5.1c shows a render of both zooms overlaid. In a similar fashion, the arms of an interferometer will be modelled by individual zooms. The consequence of this method is that the individually propagated wavefronts through the interferometer arms have to be combined in Matlab. Furthermore, when a tilted wavefront is incident on the instrument, a phase delay has to be applied in Matlab, to make sure both wavefronts are coherent. This effect is shown schematically in Figure 5.2.

5.2. Beam propagation

In order to determine the complex amplitude at the detector and take diffraction effects into account, diffraction propagation is needed. This is a critical part of the process, as it should be done well to accurately determine the system's response to perturbations. In Code V, two methods are available: FFT beam propagation and Beam Synthesis Propagation (BSP).

¹See for example <https://support.zemax.com/hc/en-us/articles/1500005489141-Using-Multi-Configuration-Operands-to-control-parameters-in-a-single-configuration-system>

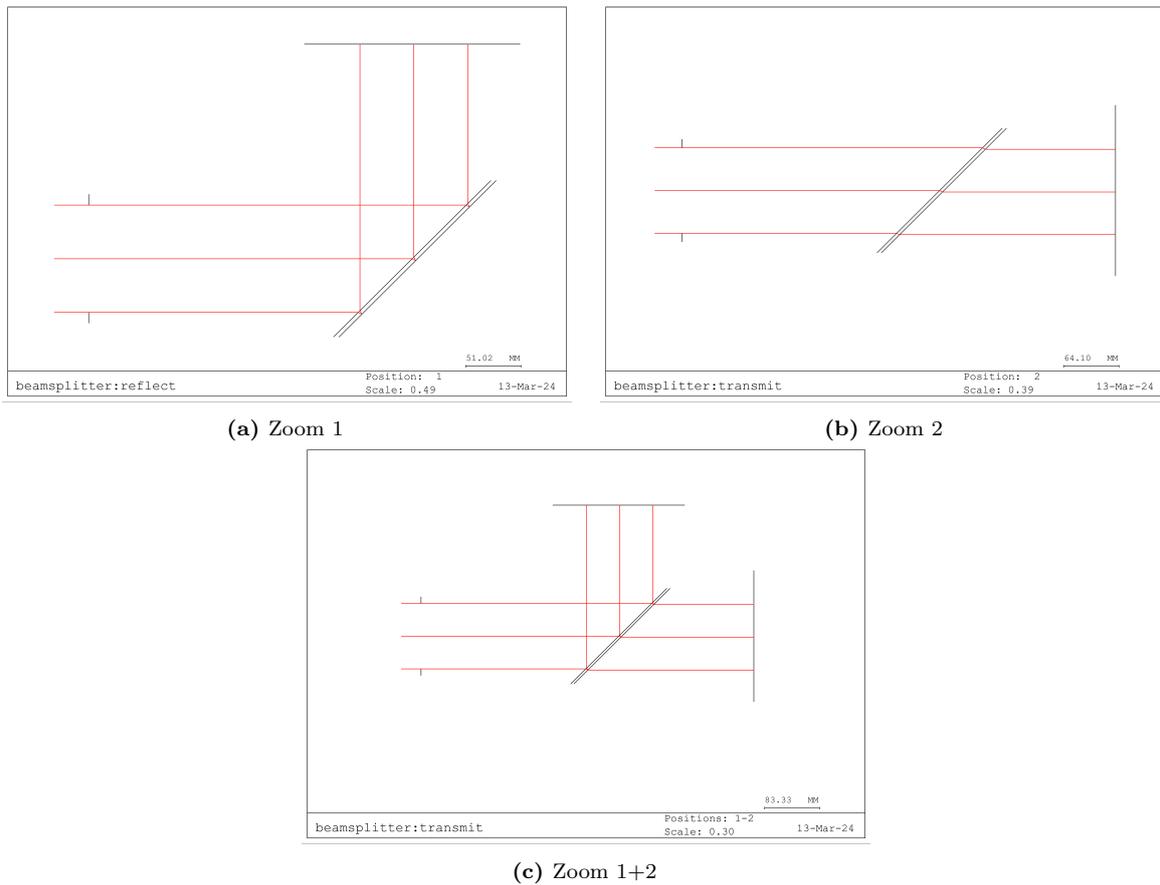


Figure 5.1: Example of a zoomed system. (a) Zoom 1, where the second surface of the beamsplitter is reflective. (b) Zoom 2, where the second surface is transmissive. (c) A render of both zooms overlaid.

5.2.1. FFT Beam Propagation

The FFT beam propagation function uses a combination of diffraction and ray tracing methods to model the field at a given plane. For each surface, or range of surfaces a specific propagation method can be chosen:

- Automatic - Let Code V choose one of the below methods automatically, based on the Fresnel number ($r^2/(\lambda \cdot D)$, where r is the beam radius, λ the wavelength and D the propagation distance).
- Far field - Also referred to as the Fraunhofer propagator, this method uses the Fraunhofer diffraction integral, propagated from the current position to infinity.
- Near field - Also referred to as the Fresnel propagator, which uses the Fresnel diffraction integral.
- Sphere to sphere - Propagates the wavefront from an input reference sphere to an output sphere.

Although FFT beam propagation yields more accurate results, it is slower and more complex to set up appropriately. Because of this, the documentation recommends using the easier to use Beam Synthesis propagation, which will be discussed in the next section.

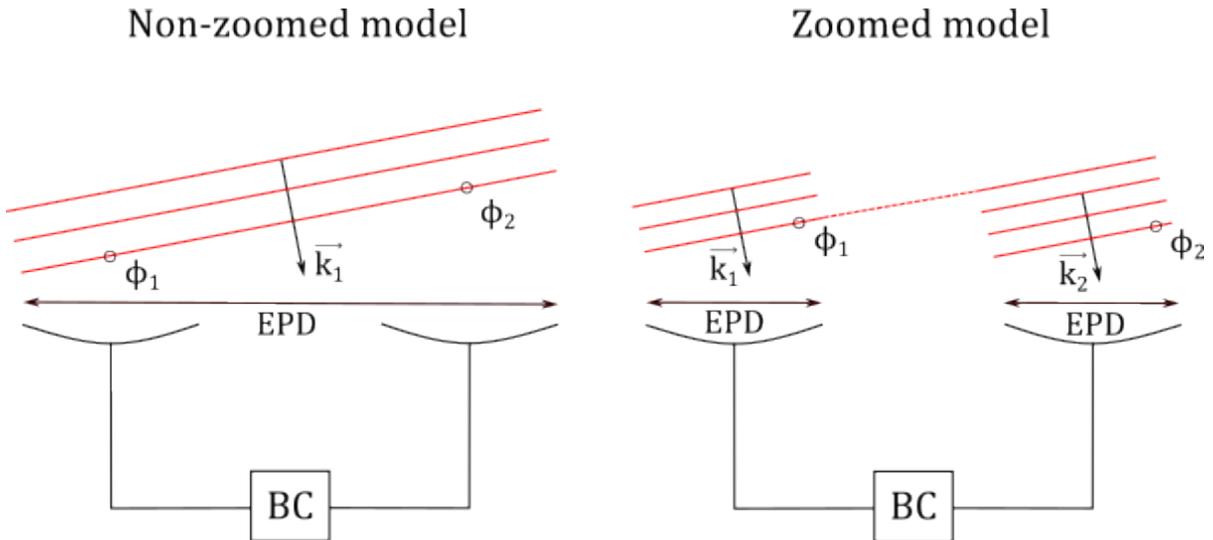


Figure 5.2: A schematic representation to show the difference between a non-zoomed interferometer model and a zoomed interferometer model. On the left, a single wavefront, with wave vector \vec{k}_1 is incident on the interferometer model, consisting of two apertures. Since the model is not separated into zooms, there is only one entrance pupil, with an Entrance Pupil Diameter (EPD) as large as the baseline. Because of this, $\phi_1 = \phi_2$. On the right, a zoomed interferometer model is shown. Due to the 'zooming', the entrance pupil is split into two, and thus two wavefronts are incident, with wavevector k_1 and k_2 . To make sure both wavefronts have the same phase, a phase shift $\delta\phi$ has to be applied.

5.2.2. Beam Synthesis Propagation (BSP)

Beam Synthesis Propagation (BSP) is an advanced, Code V proprietary algorithm for propagating an incident field through the optical system. Normally, Code V would ray trace through the system up until the exit pupil, after which the last part is propagated by taking the wave nature of light into account. With BSP, an input wave is propagated through the entire optical system, making use of so-called "beamlets", which are small beams that make up the whole beam. The beamlets are propagated individually and are added together at the chosen output plane. This is a valid approximation, assuming the linearity of the wave equation. Several input beam types can be chosen, as is shown in Figure 5.3: A spherical, Gaussian, or a custom wavefront can be selected. When the wavefront radius is not specified, it will default to infinity, making it into a plane wave.

The ease of use of BSP comes in part from the pre-analysis option. With this, an algorithm automatically determines settings appropriate for the system. These settings include:

- NRI - The number of rings of beamlets to propagate
- GPX/GPY - The number of points in the output grid
- RSF - Resampling of the grid at a given surface

These settings can be considered a starting point and can always be altered for better sampling of the system, if needed.

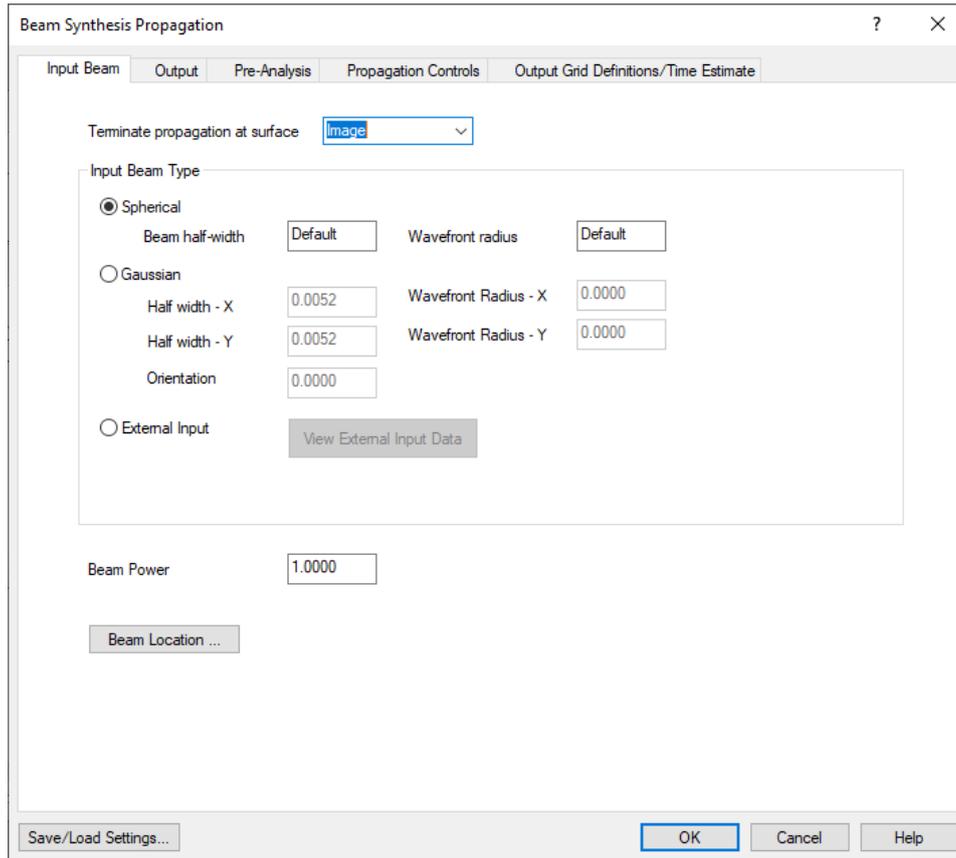


Figure 5.3: A screenshot of the BSP window in Code V. At the input beam tab, a choice can be made between a spherical or Gaussian beam, or a field based on external data. Under the propagation controls, the number of beamlet rings can be set, as well as settings for resampling the field at a certain surface.

5.3. Deriving transmission map

In order to get the transmission map from the optical model, the following steps are taken:

1. Propagate a plane wave through each arm of the interferometer for each wavelength in the spectrum, using Code V's BSP. The wave vector of the incident field is parallel with the optical axis of the telescope.
2. Retrieve the complex amplitudes in Matlab, using Code V's COM API, for each aperture and wavelength.
3. Calculate the optical path difference for each field angle in the field of view, and each aperture and wavelength, as described in section 3.4. The OPD is applied to the complex amplitude as a constant phase difference across the detector plane. The π phase shift is done digitally, for the sake of simplicity in the optical model.
4. Combine the complex amplitudes and take the magnitude squared to get the intensity at the detector plane. Integrate over the detector plane to get the power impinging on the detector, for one specific field angle.
5. To get the transmission map, the power per field angle is normalised.

It should be noted here that in reality, not only piston, but also a tilt contribution would be present in the phase term of the complex amplitude at the detector plane. The tilt is neglected

for two reasons: First, given the small field of view, the beam shear resulting from the tilt is negligible due to the small field of view of the interferometer. Second, taking into account the tilt would require a new diffraction propagation for each point in the field of view, which would be very expensive computationally.

5.4. Sensitivity analysis

The main objective of the thesis is to perform a sensitivity analysis. The goal of this analysis is to understand how a nulling interferometer behaves under physical perturbations. These perturbations can be caused by vibrations in the satellite, thermal expansion or gradients, manufacturing errors or temporal degradation of optical surfaces.

The physical system perturbations will be done according to the following steps:

1. Propagate a plane wave through the system with the Code V BSP functionality, without any perturbations.
2. Retrieve the complex amplitude data with Matlab, via the Code V API and calculate the null depth.
3. Apply perturbation (tilt, decenter, radius of curvature, ...) to a surface.
4. Propagate again and calculate null depth.

The number of systems, surfaces and parameters to be perturbed create a large parameter space, too large to all be evaluated properly. As such a selection of systems and parameters will be made to be evaluated, which will be discussed in the following sections.

5.4.1. Tilt

When tilt is concerned, a single element can be tilted independently, or the entire telescope assembly can be tilted with respect to the rest of the optical train. When perturbing the tilt angle of the entire telescope assembly, there are several possible cases. First, consider the base case, given two telescope assemblies, without any perturbation, as depicted in Figure 5.4.

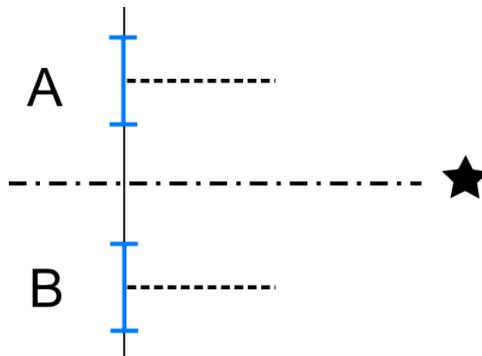


Figure 5.4: Base case of the telescope configuration, without any perturbation. The blue structures represent the (orientation of the) telescope assembly, consisting of the primary mirror, secondary mirror and collimator. The dashed dotted line denotes the line of sight with the target, whereas the dashed line denotes the line of sight of a given surface. The solid line is perpendicular to the line of sight of the spacecraft.

The four perturbation cases are:

- A. Telescope A is tilted by an angle α , relative to the axis perpendicular to the line of sight of the spacecraft.
- B. Both telescope A and B are tilted by an angle α , but both are still aligned with the line of sight of the spacecraft.
- C. The line of sight of the spacecraft is tilted with an angle α
- D. Telescope A is tilted with angle α_1 , whereas telescope B is tilted with an angle α_2 .

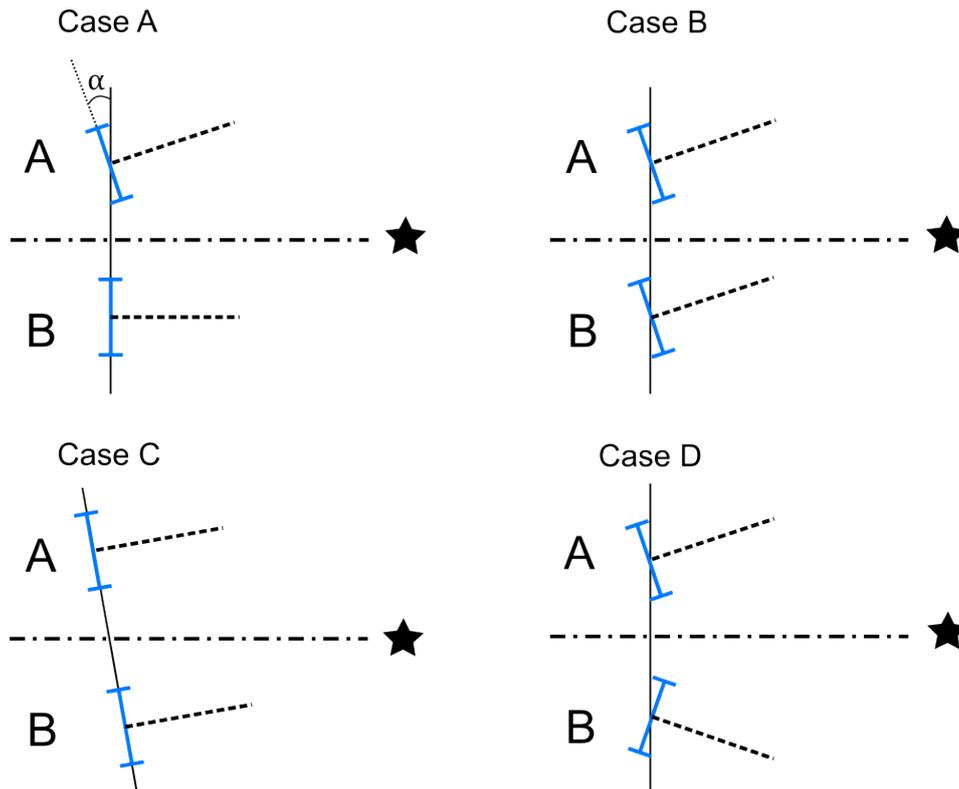


Figure 5.5: Four cases of tilt perturbations can be considered. (a) One arm of the interferometer is tilted by an angle α . (b) Both arms are tilted by an angle α , but the spacecraft is still pointed correctly towards the target. (c) Both arms are tilted by an angle α and the spacecraft is not pointed correctly. (d) Both arms are tilted, but not necessarily by the same angle α .

5.4.2. Translation

Two types of translation can be defined: lateral translation, which is the translation along the x , y and z axis and axial translation, which translates the surface along the optical axis. The latter is a composite of the former, but the effect on the null depth is not a linear addition of the separate lateral translations.

Lateral translation can be considered in two coordinate systems: the global coordinate system and a coordinate system local to the perturbed surface, depicted schematically in Figure 5.6. Both have their own merit: a global coordinate system is the most general case for considering perturbation modes. A local coordinate system allows us to express perturbation modes in

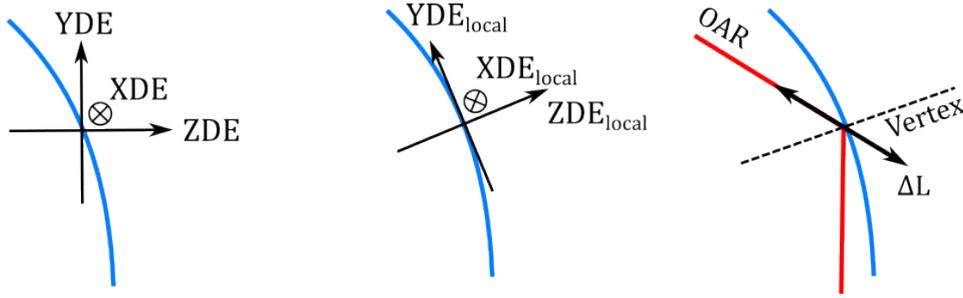


Figure 5.6: Left: translation perturbation along the global x, y and z axis, denoted by the Code V commands (XDE, YDE, ZDE). Centre: translation perturbation along axes local to the optical surface. Right: Perturbation along optical axis ray (OAR).

terms of aberrations, such as defocus and astigmatism. Axial translation will only be considered for the off axis parabola that collimates the beam, depicted schematically in Figure 5.6.

One design parameter is the radius of curvature of the off axis parabola, which acts as a collimator, and thus the distance to the focal point of the telescope. According to the Space Instrumentation Engineering lecture notes [66], the image blur size of the spot can be estimated for different types of mirrors and aberrations. For the (on axis) parabolic mirror, the blur size due to astigmatism is given by the following equation:

$$\beta = \frac{(L - f)^2 U^2}{2f^2 f\#}, \quad (5.1)$$

where L is the distance between the optical surface and aperture stop, f is the effective focal length, U is the half field angle in radians and $f\#$ is the f-number, given by f/D . From this it can be seen that there is no astigmatism when the focal length matches exactly the distance to the aperture stop. However, when this distance is increased or decreased, for example due to manufacturing errors or vibrations, a quadratic dependence emerges. This effect is shown in Figure 5.7 for different values of f , keeping U and the $f\#$ the same. From this, it can be seen that a mirror with a larger focal length is less sensitive to deviations in L than a smaller focal length mirror.

5.4.3. Surface shape

The previously considered perturbations only affect the positioning and orientation relative to other optical surfaces. To also take into account surface deformations, three more surface parameters are considered, shown in Figure 5.8

5.4.4. Null depth using finite sampling

The null depth is defined by the ratio between the minimum and maximum intensity of the transmission map:

$$N = \frac{I_{min}}{I_{max}}. \quad (5.2)$$

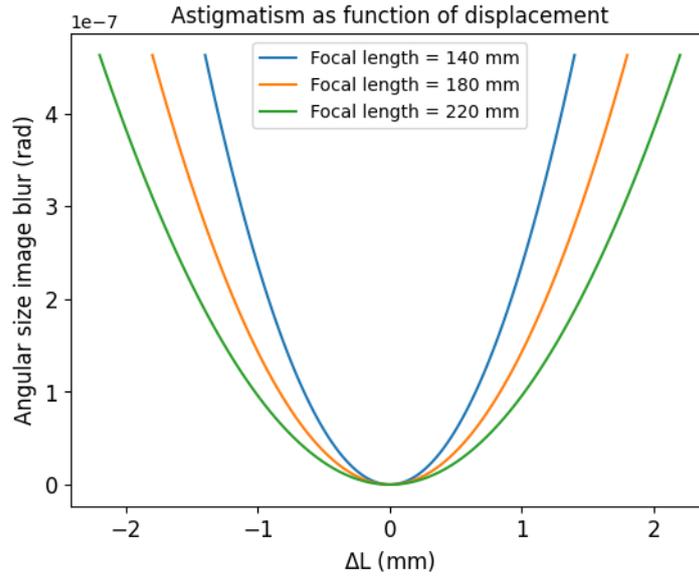


Figure 5.7: Estimated dependence of axial mirror displacement on the image blur size, β .

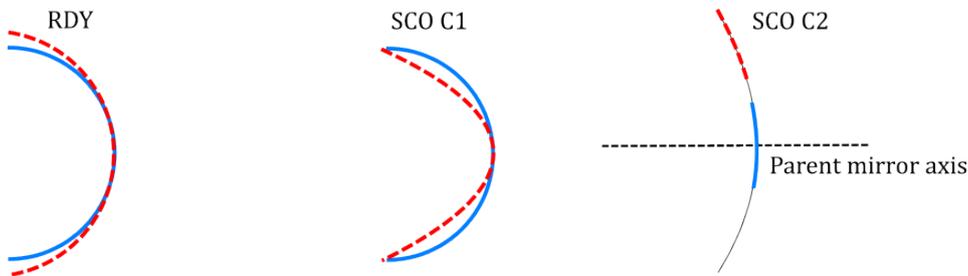


Figure 5.8: Three surface shape parameters will be considered for the perturbation analysis. Left: Radius of curvature (RDY), centre: Conic constant (SCO C1), right: Off axis angle (SCO C2).

Theoretically, this is a straight forward calculation. However, when simulating the transmission map, there is a complication in determining the minimum intensity, which is addressed in Appendix E.

5.5. Performance analysis

To compare different designs, the null depth is derived as function of the perturbation magnitude, for each surface and each parameter. Although this gives an indication of which design is more sensitive to certain perturbation, it does not provide an explanation of the result. To gain more insight in the optical performance of the nulling interferometer, whether or not with perturbations applied, several tools can be used, such as the spot diagram (subsection 5.5.1), exit pupil map (subsection 5.5.2) and modulation transfer function (subsection 5.5.3). All of these methods can be used to determine whether or not an optical design is diffraction limited, which will be addressed in subsection 5.5.4. Another, often used metric is the Strehl ratio, which is related to the wavefront error. As the exit pupil map gives both the shape of the wavefront error, as well as a metric to determine diffraction limited performance, it is favoured over the Strehl ratio. Nonetheless, it comes down to preference which combination of methods is used to determine diffraction limited performance.

5.5.1. Spot diagram

The spot diagram is a simple tool to assess the quality of the image plane spot. A grid of rays is traced through the entrance pupil and through the optical system. The distribution of rays on the image plane is then plotted in the spot diagram. The spot diagram is thus derived using geometrical optics. The benefit of the spot diagram is that it is fast to simulate and that one can quickly determine which aberration is dominant across the FoV. However, since it is based on geometrical optics, diffraction effects are not taken into account.

The spot diagrams of astigmatism, coma and spherical aberration are shown in Figure 5.9 to Figure 5.11. All spot diagrams were simulated by Code V, using the "SPO" function. First, a ZnSe lens with a high $f\#$ is optimised at $10\ \mu\text{m}$ to be diffraction limited, such that it has no significant spherical aberration. Then, an interferogram is applied to the first surface, to force a certain aberration.

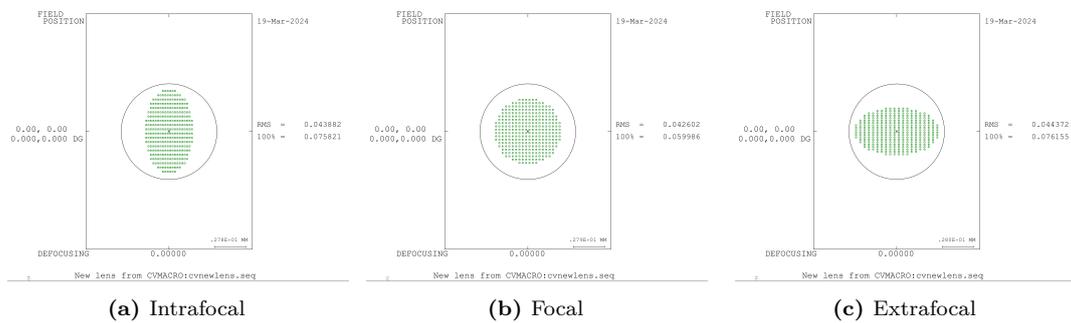


Figure 5.9: Spot diagrams of a lens with astigmatism, simulated with Code V.

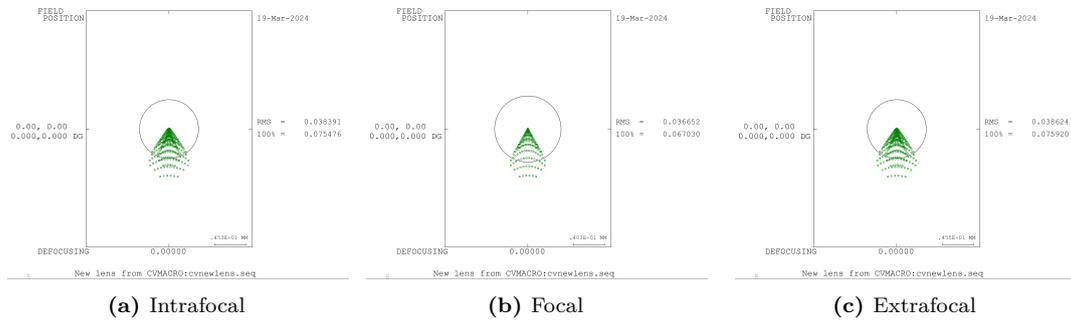


Figure 5.10: Spot diagrams of a lens with coma, simulated with Code V.

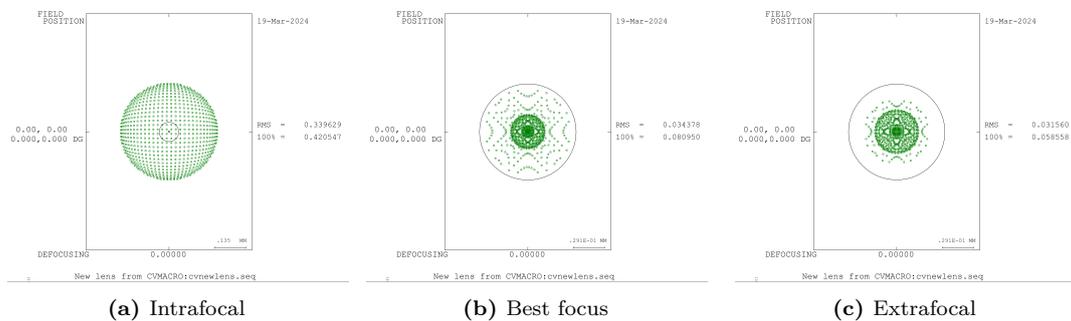


Figure 5.11: Spot diagrams of a lens with spherical aberration, simulated by Code V.

5.5.2. Exit pupil map

The exit pupil map shows the wavefront aberrations in the exit pupil. Similar to the spot diagram, an exit pupil map can be used to evaluate which wavefront aberration is dominant in the exit pupil. The exit pupil map has several benefits over the spot diagram. First, the exit pupil map can directly quantify how far away an optical system is from diffraction limited performance, by means of the P-V Wavefront Error (WFE). Second, the exit pupil map can be decomposed in terms of Zernike polynomials (section 2.4), quantifying the contribution of each aberration in the exit pupil. Zernike decomposition will be used extensively in section 6.5, to quantify the different contributions to the wavefront error after a perturbation has been applied to the optical model. Despite the advantages, it is still useful to consider the spot diagram, in order to assess how well the optical system performs from a geometrical optics perspective.

5.5.3. Modulation Transfer Function

Before the Modulation Transfer Function (MTF) can be discussed, the concepts of contrast and spatial frequency need to be considered. In imaging, any object can be expressed in terms of the spatial frequency distribution, where fine details correspond to high spatial frequency, while coarse features correspond to low spatial frequencies. Instead of considering features on an object to be imaged, the problem is simplified to a grating like object, that only consists of dark and bright lines, spaced apart equally, corresponding to a single spatial frequency. Spatial frequency is thus expressed in lines/mm. The contrast is the difference between the brightest and darkest signal in the imaged grating. When the spatial frequency is low, a given optical system can adequately image the object, as shown in the top of Figure 5.12, such that the imaged contrast is high. When the spatial frequency of the block wave increases, the contrast is reduced, as can be seen from the bottom part of Figure 5.12. This has to do with the finite extent of the point spread function. The point spread function (PSF) can be thought of as the impulse response of an imaging system. When an object is imaged, the image that appears is a convolution between the PSF and the object. In this light (no pun intended), it becomes apparent why the contrast drops when the grating frequency is increased. As the distance between the bright lines is reduced, the imaged objects will start to overlap more, reducing the contrast. At some spatial frequency, there will be no more contrast, which is referred to as the cutoff frequency. The MTF is a measure of the amount of contrast at a given spatial frequency, relative to the maximum contrast at a spatial frequency of 0 lines/mm.

Although the MTF cannot be used to quantify how close an optical system is to diffraction limited performance, it is the only method that takes spatial frequencies into consideration and is thus a valuable tool to determine how well an optical design performs in the real world.

5.5.4. Diffraction limit

Now that the tools have been discussed that will be used to assess the optical system, the diffraction limited performance can be expressed in terms of performance parameters derived from those tools. The diffraction limit is the smallest possible separation that can still be resolved by an imaging system, which can be calculated in terms of the wavelength and diameter of the entrance pupil:

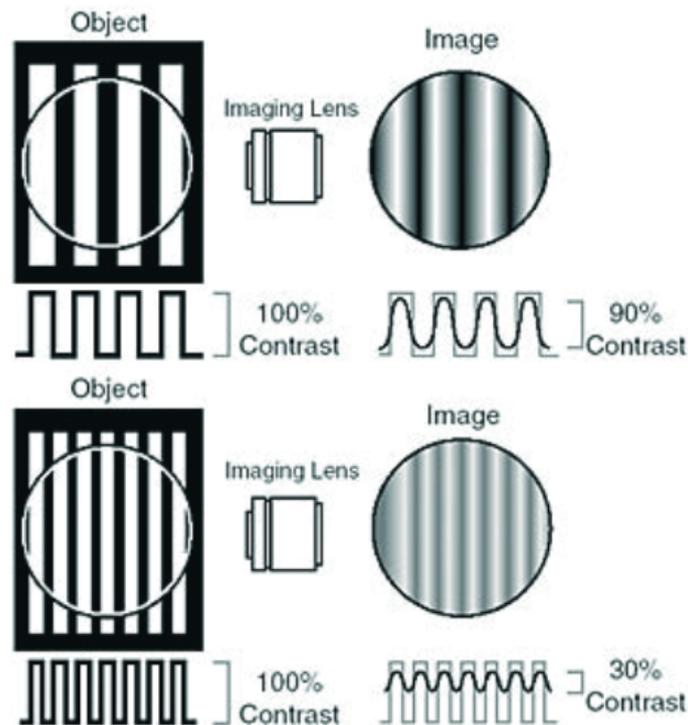


Figure 5.12: Top: an object is imaged with 100% contrast in the object plane. After the object is imaged, the contrast is reduced to 90%. Bottom: When the spatial frequency is increased, the contrast in the image plane is reduced further, due to the finite extent of the PSF.

$$\theta = 1.22 \frac{\lambda}{D} \quad (5.3)$$

A system is said to be diffraction limited when the following statements are true²:

1. The RMS spot diagram size is within the Airy Disk diameter.
2. The RMS wavefront error (WFE) is less than 0.075 waves, while the P-V WFE is less than 0.25. This is an equivalent to stating that the Strehl ratio should be higher than 0.8.
3. The MTF is close to the diffraction limited MTF.

5.6. Chapter summary

In this chapter, several aspects concerning the methodology of this study were discussed. First, because of the fact Code V is a sequential ray tracing software so-called 'zooms' need to be used to model an interferometer, specifically the separated apertures and the beam combiner optics. Second, two methods of diffraction propagation were discussed: Fast Fourier Transform beam propagation and Code V's proprietary Beam Synthesis Propagation algorithm. The latter method will be used to determine the complex amplitude in the pupil plane, which in turn is used to calculate the null depth. Third, the perturbation analysis was discussed in section 5.4.

²The source of these requirements is the FORTA workshop that was given for the course Space Instrumentation

The goal of this analysis is to quantify how much the null depth is degraded for a small set of perturbation parameters, and to which extent the telescope design influences this result. To this end 8 perturbation parameters were chosen: three translational (x/y/z decenter), two rotational (rotation around x/y axis) and three surface shape parameters (conic constant, radius of curvature and off-axis angle). Lastly, the notion of diffraction limited performance was considered. To assess and quantify this, several methods are available, including a spot diagram, the exit pupil map and modulation transfer function. Diffraction limited performance is then reached when the following is true: the RMS spot diagram size is within the Airy disk diameter, the RMS wavefront error is less than 0.075 waves, the P-V WFE is less than 0.25, and the MTF is close to the diffraction limited MTF.

6

Results

In this chapter, the results of the study will be discussed. First, a baseline optical model for the telescope design and beam combiner is shown and analysed in section 6.1. In this section, it will be argued that the off-axis telescope design is diffraction limited within the OWA. Additionally, it will be shown that the beam combiner behaves according to what is expected from theory. The telescope and beam combiner are then combined into a two and four aperture nuller, discussed in section 6.2. Then, a simulation of the transmission map is shown in section 6.3. Before the sensitivity analysis is considered in section 6.5, an exposition is given in section 6.4 on the different contributions that can degrade the null. This provides the necessary insight to interpret the results of the sensitivity analysis.

6.1. Optical model

6.1.1. Aperture

As a starting point, a confocal off-axis Cassegrain telescope was designed, following the theory discussed in section 4.4. The design parameters are listed in Table 6.1 and a side view is shown in Figure 6.1a.

Parameter	Value	Unit
K_1	-1	-
l'_1	2000	mm
R_1	3732.0508	mm
i_1	15	°
K_2	-2.5765	-
i_2	19.6601	°
R_2	-1182.4118	mm
l_2	500	mm
D	1	m

Table 6.1: Design parameters for an off axis Cassegrain. The parameters R_1 , R_2 , K_2 and i_2 are calculated from Equation 4.9, Equation 4.13, Equation 4.14 and Equation 4.12.

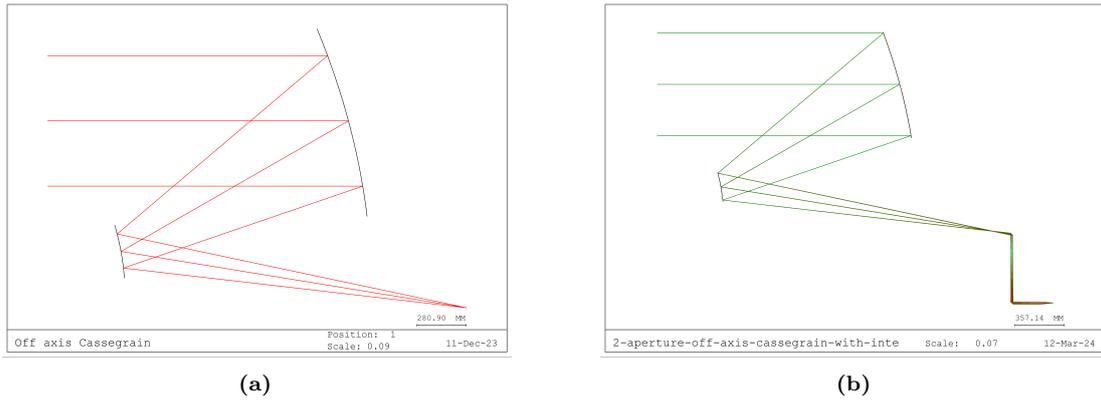


Figure 6.1: (a) Diagram of an off-axis Cassegrain telescope, with a back focal length of 2000 mm, a primary mirror focal length of 2000 mm and tilt angle of 15° . (b) To collimate the off-axis Cassegrain, a single off-axis parabola is used. The collimated beam is then re-imaged with a perfect lens, so that the optical performance can be analysed.

To know whether or not this design is diffraction limited and what the field of view (FoV) is, the spot diagrams can be plotted for different field angles. The FoV of a telescope is determined by the field angles within which the telescope is still diffraction limited. Figure 6.2 shows nine spot diagrams superimposed on the FoV, in such a way that the location of the spot diagram corresponds to the field angle with which the spot diagram was simulated. The FoV is highlighted by the dashed square. The blue circle indicates the size of the Airy disk, determined by the size of the aperture and wavelength. The central spot diagram, corresponding to a field angle of $(0', 0')$ ¹, shows a perfect spot, whereas the bottom right spot diagram, corresponding to a field angle of $(3', 3')$, shows coma. Given that all spot diagrams fit inside their airy disks for the considered field angles, it can be said that the telescope is diffraction limited over a FoV of 6×6 arcminutes. It should be noted that this is much larger than the OWA of the interferometer, which is on the order of a an arcsecond.

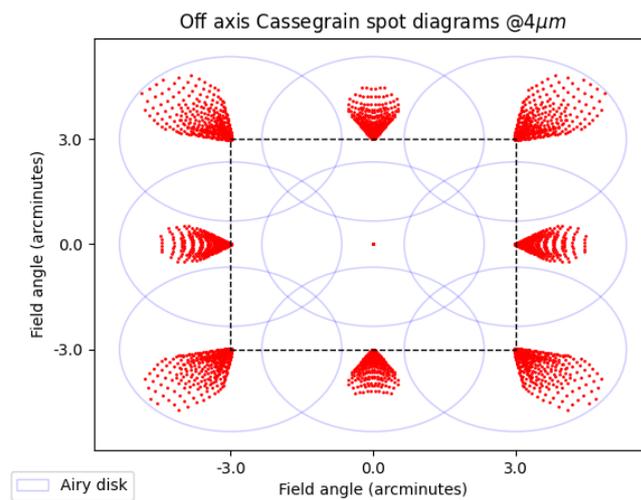


Figure 6.2: Spot diagrams superimposed on the field of view, which is demarcated by the dashed square. The blue circles demarcate the Airy disk, which has a radius of $39.04 \mu\text{m}$ at a wavelength of $4 \mu\text{m}$. Note that the size of the Airy disk circles are not to scale, compared to the FoV, for the sake of visualisation.

¹The arcminute is one sixtieth of a degree, denoted by the single quote ($'$)

The spot diagram is derived with geometrical optics, and although useful in its own right, it does not provide a complete performance evaluation. Another requirement for diffraction limited performance is that the diffraction based MTF of the instrument is close to the ideal case. Figure 6.3 shows 4 MTF curves, two for the ideal, diffraction limited case and two for the MTF at the corner of the FoV ($3'$, $3'$). Because the MTF curves overlap for the largest field angle in the FoV, it can be said that the telescope is diffraction limited for the total FoV, from a diffraction optics point of view.

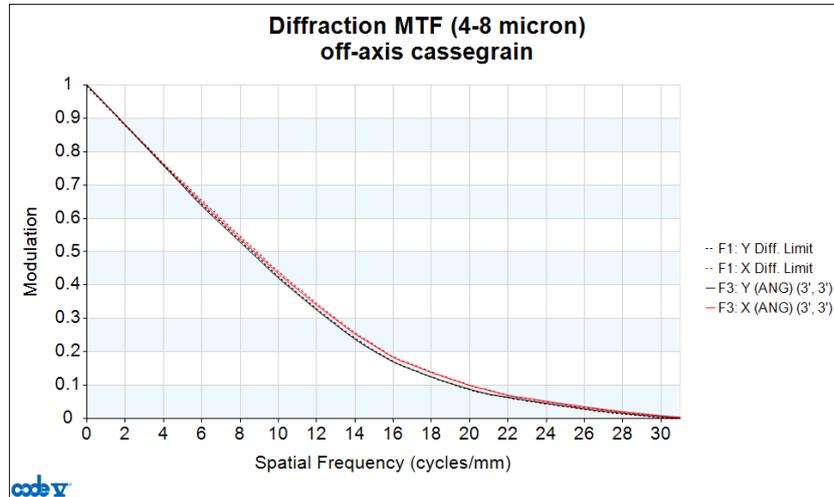


Figure 6.3: Diffraction based MTF for the off-axis Cassegrain, evaluated from 4 to 8 μm . The red dashed graphs show the diffraction limited performance, whereas the solid lines show the simulated performance for the off axis diffraction MTF at a field angle of ($3'$, $3'$).

6.1.2. Collimator

After the telescope optics, an objective needs to be placed, such that the beam is collimated for beam combination. As was discussed in section 4.5, a single off-axis parabola (OAP) is used for this purpose. The design parameters are given in Table 6.2 and a side view of the telescope and objective is shown in Figure 6.1b. Given an entrance pupil diameter of 1 m and spot size after collimation of 13.7 mm, the expansion ratio is 0.0137.

Parameter	Value	Unit
i	40.3399	$^\circ$
θ_0	49.6601	$^\circ$
R	-117.3269	mm
K	-1	-

Table 6.2: Design parameters for an off axis parabola to collimate the telescope beam.

With a collimated beam, some functions of Code V that rely on having a finite caustic do not work properly anymore. To still analyse the optical performance, a perfect lens model can be used. This perfect lens introduces no additional aberrations, while still bringing the collimated beam to a focus. In this case, a perfect lens with a focal length of 100 mm is used. The spot diagrams in the focal plane are shown in Figure 6.4a for the on axis field and the off axis field of ($3'$, $3'$). The spot diagrams just outside the focal plane are shown in Figure 6.4b.

From these figures, two observations can be made. First, the dominant aberration in the off-

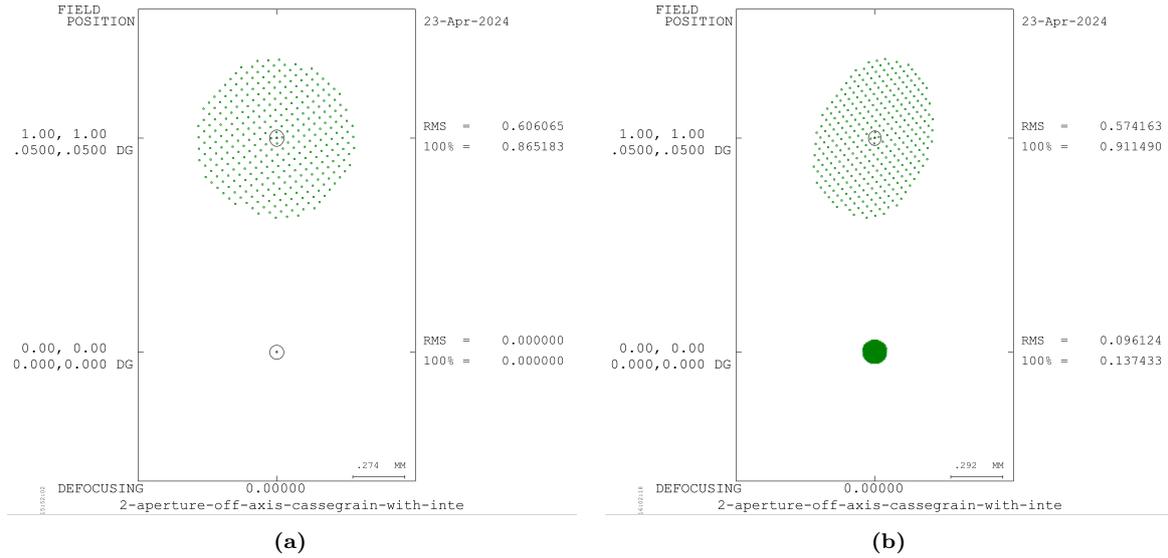


Figure 6.4: (a) Spot diagrams for the off-axis Cassegrain with collimator, simulated with a wavelength of $4 \mu\text{m}$. A perfect lens was used to image the pupil plane. The top spot diagram, showing the distribution for the $(3', 3')$ field angle, could be interpreted as defocus, or astigmatism. Note that the Airy disk is inside the spot diagram. (b) The spot diagrams, evaluated outside of the focal plane.

axis field is now astigmatism, as opposed to coma (compare to Figure 5.9b). For astigmatism, the spot appears circular in the focal plane, such that it is difficult to distinguish from pure defocus. To confirm which of the two it is, a small defocus can be introduced by translating the image plane. The effect of this is shown in Figure 6.4b, which shows the tilted, oval shaped spot diagram, corresponding to oblique astigmatism. This is the expected aberration for an OAP [29].

The second observation is that, with the addition of the OAP, the telescope is not diffraction limited anymore within the FoV, given that the size of the spot diagram is much larger than the Airy disk. This is also shown in Figure 6.5, as now the simulated MTF curves do not overlap anymore with the diffraction limited MTF curves. To make sure the telescope is still diffraction limited, there are two possible solutions: either optimise the design, or reduce the FoV with an aperture stop. As a nulling interferometer is a narrow field instrument, restricted by the OWA, and given that the previously considered FoV is much larger than the OWA of $\sim 1''$, optimising the design is not necessary. To find the new FoV, the field angles are reduced to the point where all requirements for diffraction limited performance are met. Some notes and details on this process are given in Appendix F. The new FoV with the addition of the collimator was determined to be $13'' \times 13''$, which fulfils the requirement of being diffraction limited within the OWA (850 mas at a wavelength of $4 \mu\text{m}$).

As a last step of verification, it is useful to compare the output of the BSP with the exit pupil map and spot diagrams. Figure 6.6 shows the simulated near-field intensities for both the on-axis and the off-axis fields. A plane wave was used as an input field, with a wavelength of $10 \mu\text{m}$. Thirty five rings of beamlets make up the incident field, resulting in 3787 beamlets propagating through the system. The off-axis field shows oblique astigmatism, which agrees with the exit pupil map and spot diagram. Now that the optical performance of the off axis Cassegrain has been analysed and the FoV is determined, the beam combiner design will be discussed in the next section.

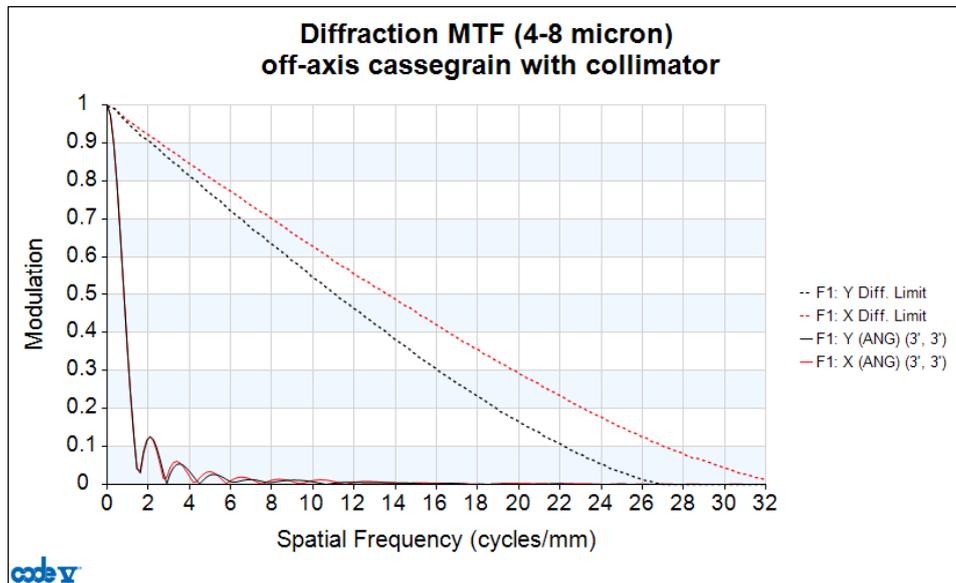


Figure 6.5: The diffraction based MTF for the off-axis Cassegrain with collimator, evaluated from 4 to 8 μm , at a field angle of (3', 3'). When comparing the diffraction limited MTF and the one at (3', 3'), the degradation in performance becomes apparent, as no modulation is left after 6 cycles/mm.

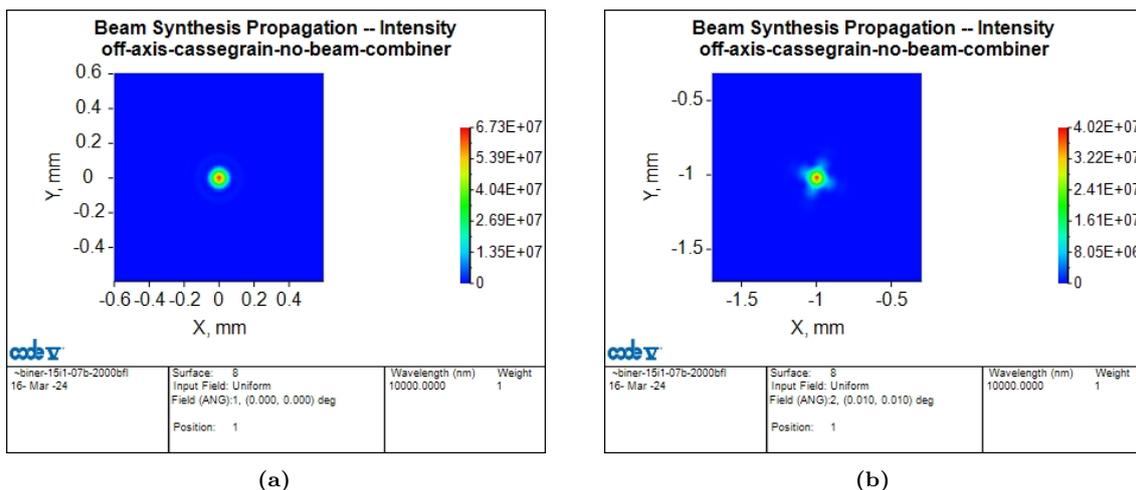


Figure 6.6: (a) Simulated intensity, using BSP, at the detector plane of the off-axis Cassegrain and collimator, for the on axis field. (b) Simulated intensity for the off axis field.

6.1.3. Beam combiner

As a first step towards modelling a beam combiner, a simple plate beam splitter model was made in Code V, which is shown in Figure 6.7. The beam splitter is made of ZnSe, which is transmissive in the MIR, with relatively little chromaticity.² To model the beam splitter, two separate configurations, so-called 'zooms' in Code V, are made. Most parameters in Code V can be zoomed, which means that their value can be changed for each separate zoom. In the case of the beam splitter, the back surface is made reflective in one zoom, while for the other zoom, the surface is made refractive.

In order to make sure the optical path length between both outputs of the beam splitter is equal, a compensator plate of ZnSe was placed after the transmitted output of the beam splitter.³

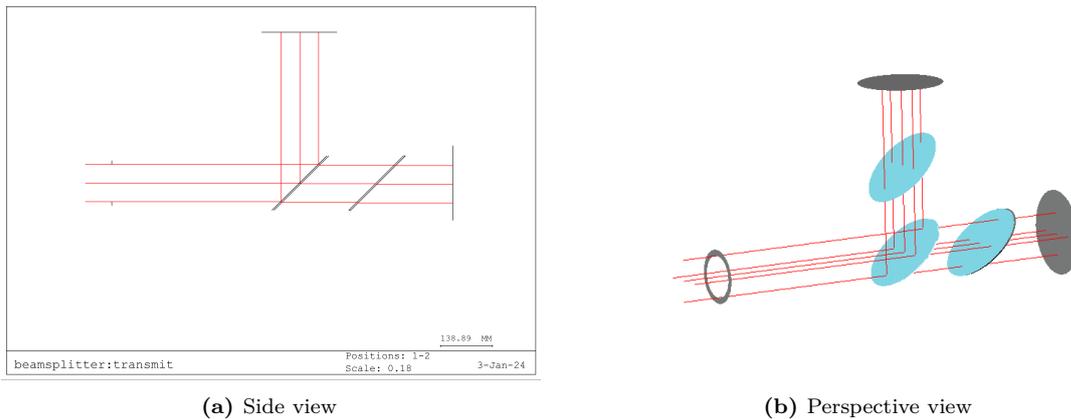


Figure 6.7: The modelled beam splitter, with one compensator plate, to compensate for the optical path. The reflected path reflects off the backside of the beam splitter. Both the beam splitter and compensator plate are made of ZnSe. In both renders, two zooms are overlaid for visualisation purposes. The grey ring in the right figure is the aperture stop, while the dark grey circular surfaces represent the detector plane. Blue surfaces are refractive surfaces and light grey surfaces are reflective.

To verify whether this beam splitter model is physically correct, the phase of the reflected and transmitted fields can be compared to the expected outcome of the Fresnel equations discussed in section 2.3. From that discussion, it was found that for the case of a 45 degree angled plate of ZnSe, only the perpendicular component was phase shifted (see Table 2.1).

In this model, the y axis is parallel to the plane of incidence, while the x axis is perpendicular. This means that for linearly polarised light in the x direction, there should be a relative π phase shift between the transmitted and reflected field⁴. To verify this, BSP is again used, by simulating the propagation through the beam splitter of a plane wave linearly polarised in the x direction. The x component of the phase and amplitude at the detector plane is shown in Figure 6.8 for both the reflected and transmitted field. The reflected field has a phase of π radians at the detector, while the transmitted field has a phase of 0π radians, resulting in a relative π phase shift. Similarly, when a beam polarised in the y direction is incident on the beam splitter, there should be no phase shift. When comparing the phase of the transmitted

²The refractive index drops by only 2.3% between 3 and 14 μm

³In Figure 6.7b, a refractive surface is shown after the reflective output of the beamsplitter. This is a zoomed surface, without any optical power or refractive index. It is an artefact of the zooming, as the compensator plate surfaces for the transmissive zoom need to end up somewhere in the optical path of the reflected zoom.

⁴By default, Code V adds an anti-reflective coating to each air-glass interface. These have to be removed, otherwise the results do not agree in this particular case, due to the asymmetry in anti-reflective coating passes.

field with the reflected field, shown in Figure 6.9, it becomes apparent that there is no relative phase shift between the two fields.

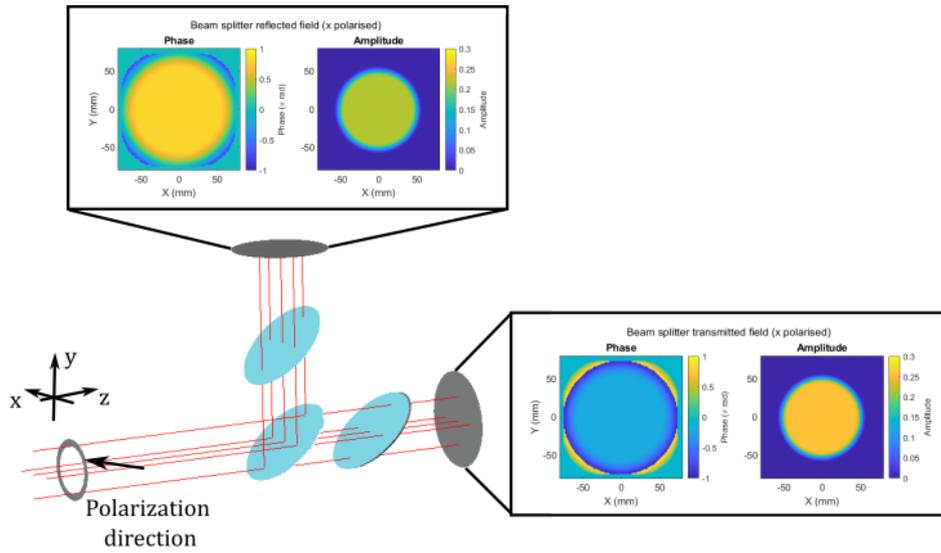


Figure 6.8: Phase and amplitude of the transmitted and reflected complex field, as modelled by BSP, at the detector plane with a beam linearly polarised in the x direction (perpendicular to plane of incidence).

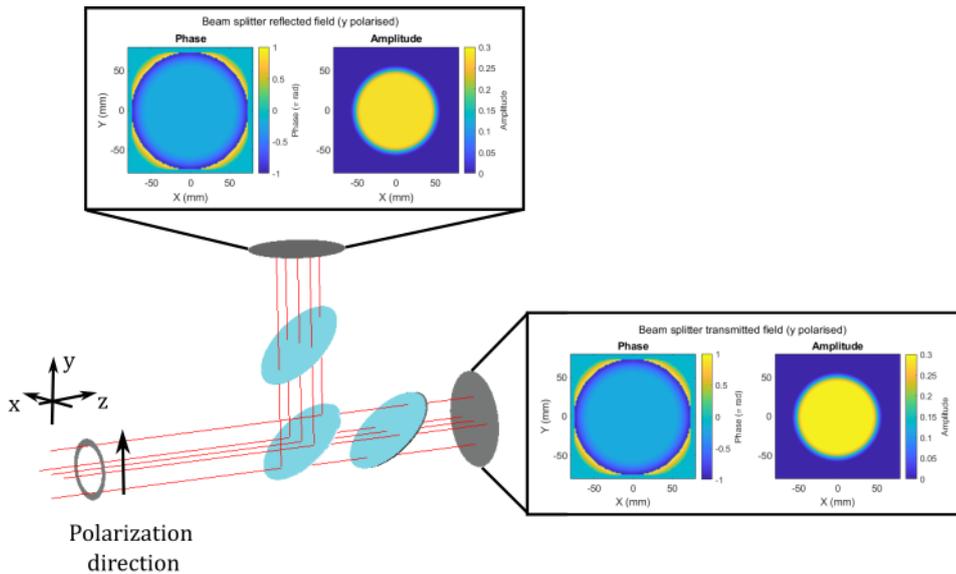


Figure 6.9: Phase and amplitude of the transmitted and reflected complex field, as modelled by BSP, at the detector plane with a beam linearly polarised in the y (parallel to plane of incidence) direction.

Now that it is determined that the beam splitter can be correctly modelled with BSP, the extension is made to a modified MZ beam combiner (MMZ). To do this, two beam splitters are placed between three reflective surfaces, such that the two input beams are split and recombined into two constructive and two destructive outputs, as shown in Figure 4.8b. All eight zooms that make up the schematic shown in Figure 6.10 are discussed in Appendix C.

Due to the vertical separation of the input beams, a path length difference is introduced between the beams that are combined, which will be corrected by the delay lines. Given that

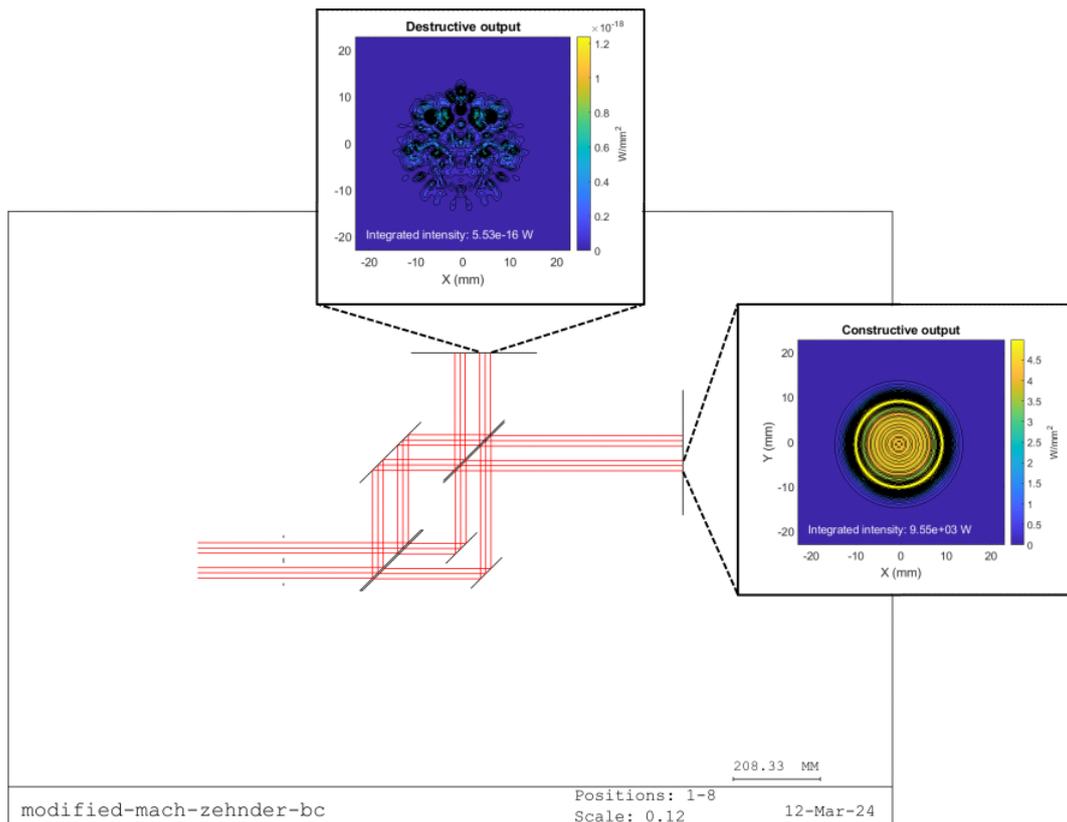


Figure 6.10: Schematic of a modified Mach-Zehnder beam combiner. It consists of 8 zooms, which are overlaid to make it appear as one component. The top inset shows the destructive output and the right inset shows the constructive output, both simulated with BSP. The input wave is a monochromatic plane wave at $10 \mu m$.

there are no delay lines yet, an artificial path length difference is introduced to equalise the phase. With the path length equalised, it is expected that the intensity of the destructive output is close to 0. In Figure 6.10, the constructive and destructive output of the MMZ beam combiner are shown in the insets. As a reminder of the methodology, the individual waves are simulated with BSP and combined in Matlab. The destructive output has an integrated intensity on the order of 10^{-16} W, which results in a hypothetical null depth of $\sim 10^{-19}$. Needless to say, this is a highly idealised result, as only flat surfaces are used with perfect reflective coatings.

To summarise this section, an off-axis Cassegrain was modelled in Code V, showing diffraction limited performance at a wavelength of $4 \mu m$ over a FoV of $13'' \times 13''$. Additionally, a simple beam splitter model was made and analysed, showing that the model agrees with what is expected from the Fresnel equations. Finally, the beam splitter model was used in a MMZ beam combiner, showing constructive and destructive interference in the outputs. In the next section, the telescope and MMZ model will be combined into a nulling interferometer model.

6.2. Nulling interferometer optical model

6.2.1. Bracewell interferometer

The nulling interferometer optical model is shown in Figure 6.11. In addition to the Cassegrain telescope and MMZ beam combiner, a delay line is placed, consisting of four flat mirrors. The distance between the collimator and following fold flat can be adjusted to increase or decrease the baseline.

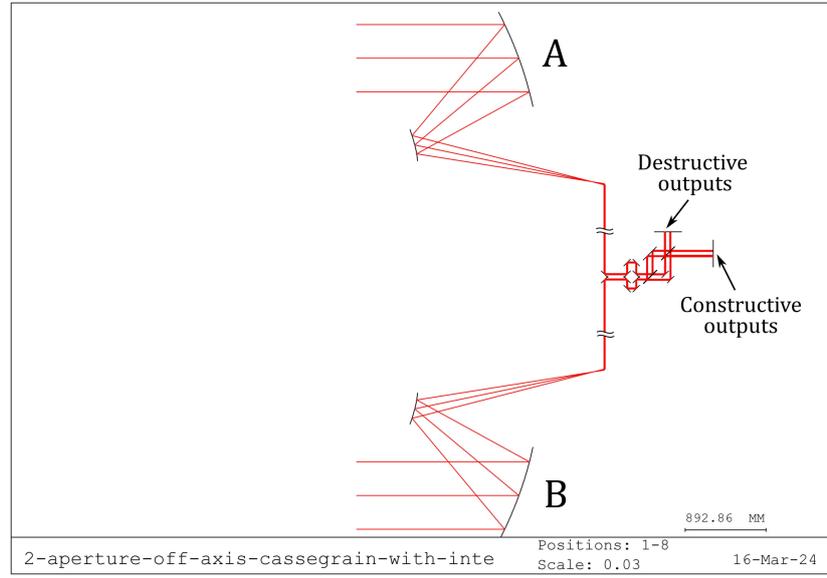


Figure 6.11: Side view of the 2 aperture nulling interferometer with a baseline of 15 meters. The destructive outputs are located on the top side of the MMZ BC, while the constructive outputs are located on the right side of the BC.

Before the destructive fields of the interferometer are considered, a plane wave will be propagated through aperture A and B individually, in order to evaluate the properties of the field falling on the detector. This could be achieved in practice by placing a shutter in one of the apertures, blocking off its optical path. To this end, a polychromatic plane wave with a centre wavelength of $10\ \mu\text{m}$ and bandwidth of $1000\ \text{nm}$ is propagated with BSP⁵, the result of which is shown in Figure 6.12. Upon close inspection, it can be seen that the fields are rotated by 180 degrees, relative to each other. The off axis Cassegrain apertures are orientated such that they are rotationally symmetric around the line of sight of the interferometer. The result of this is that the pupil planes are mirrored around the horizontal axis. Because of this, one of the pupil planes is flipped digitally, to optimise the null depth.

To quantify the difference in null depth, the on axis (no field angle) destructive fields are plotted in Figure 6.13. The null depth of the destructive field where neither of the fields is flipped is $1.39 \cdot 10^{-3}$, whereas the null depth corresponding to the case where one field is flipped, is $2.26 \cdot 10^{-7}$.

⁵In Code V, it is not possible to define a single polychromatic wave. Instead, several monochromatic waves are propagated and combined in Matlab, using equal weights on the monochromatic waves, such that the spectrum is a top hat distribution.

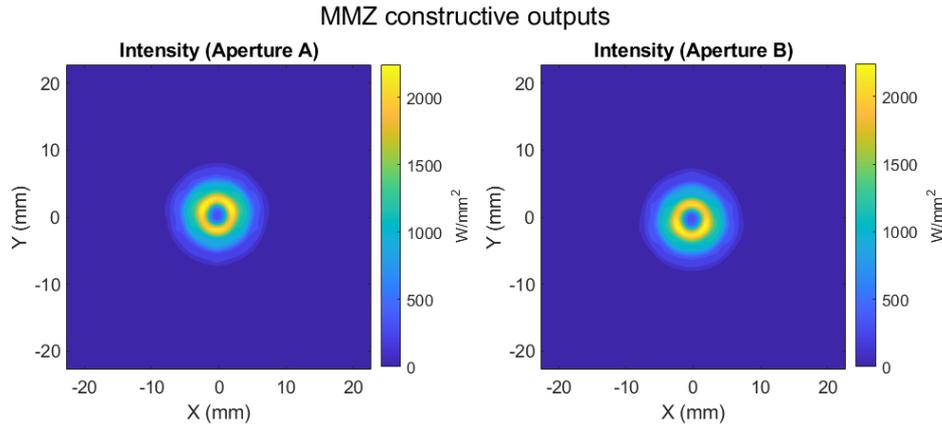


Figure 6.12: Intensity fields at the detector plane, as simulated by BSP. Upon close inspection of the fields, it can be seen that they are mirrored with respect to the horizontal axis.

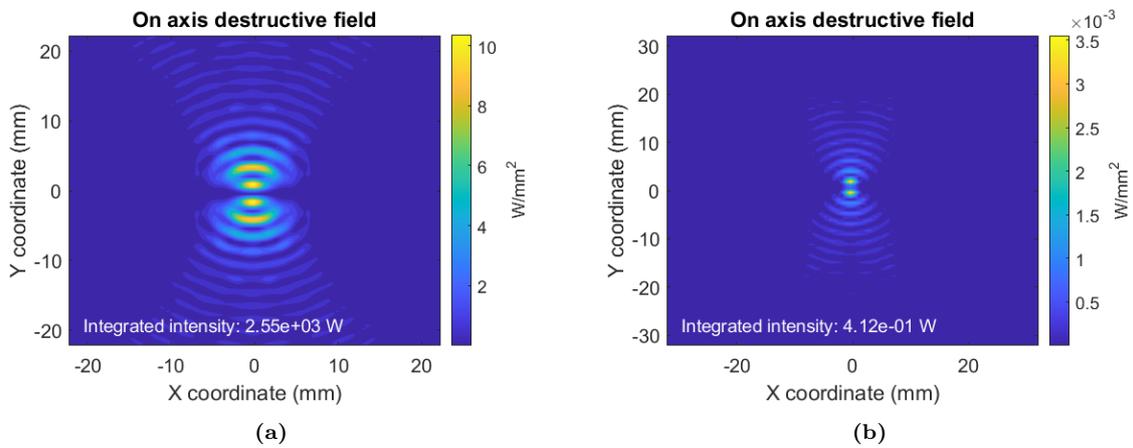


Figure 6.13: (a) Destructive field of the nulling interferometer in the pupil plane, where neither of the aperture beams was flipped. (b) Destructive field in the pupil plane, resulting from rotating one of the beams by 180 degrees.

6.2.2. Double Bracewell interferometer

As was discussed in subsection 3.3.2, a four aperture nulling interferometer is predicted to perform better than a two aperture interferometer. To analyse its performance compared to a Bracewell interferometer, a double Bracewell interferometer was designed. As the name suggests, the design consists of two Bracewell interferometers, where each pair is combined with an MMZ beam combiner. The combined beams are then cross-combined with a single beamsplitter and compensator plate. The beam combination strategy is comparable to the KIN beam combiner [67].

6.3. Simulating transmission maps

The transmission map shows where light is passed through and where it is blocked by destructive interference. For a 2 aperture interferometer, the transmission map is expected to have only horizontal or vertical fringes, depending on the orientation of the baseline. Following the methodology explained in section 3.4, the transmission map for the two aperture interfer-

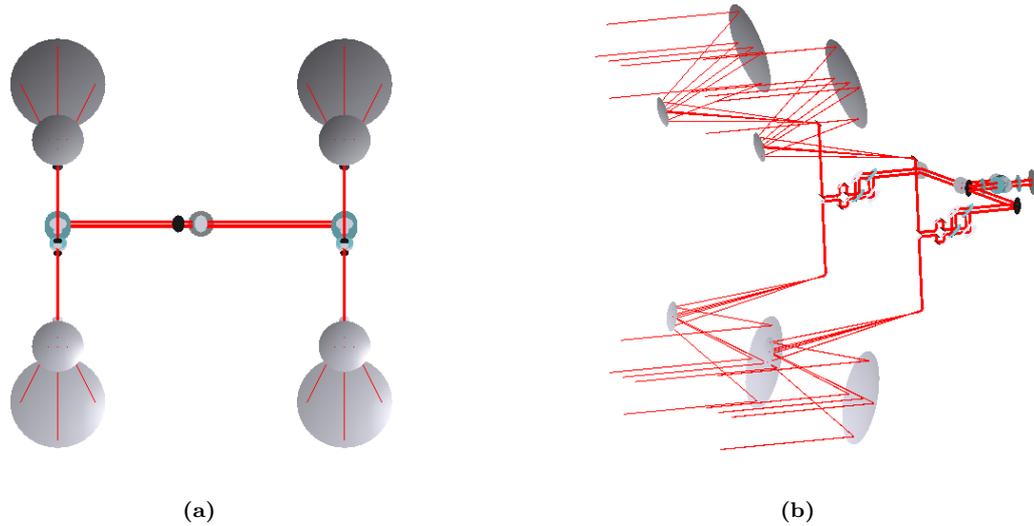


Figure 6.14: (a) Front view of the 4 aperture, off axis Cassegrain nulling interferometer. (b) Perspective view.

ometer was derived, using a baseline of 4.1 m and a monochromatic field at $10\ \mu\text{m}$. As can be seen from Figure 6.15a, the angular separation between the null and first bright fringe is 251 milliarcseconds, which equals the known theoretical value of $\lambda/(2B)$. As the input field is monochromatic, the fringe visibility is not reduced at the edges of the field of view. In the case of a polychromatic incident field, $2\ \mu\text{m}$ in the case of Figure 6.15b, the fringe visibility reduces when the angular separation increases, which agrees with the theory. The minimum null depth for the monochromatic transmission map is $7.5\text{E-}8$, while the null depth is higher for the polychromatic case, with a null depth of $3.0\text{E-}7$. The increase in null depth can be explained mainly from the path length difference between wavelengths, due to the chromaticity in the refractive index of the beam splitters.

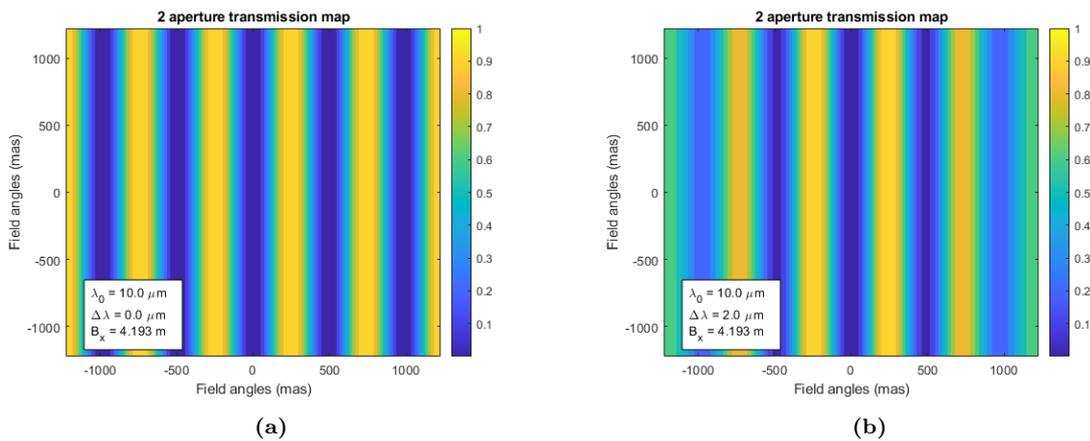


Figure 6.15: (a) Transmission map for the nulling interferometer with a baseline of 4.1 m and monochromatic incident field, showing the angular separation of 251 mas between the dark and bright fringe. (b) When using polychromatic light, the fringe visibility is reduced near the edges of the field of view.

To further test the methodology presented in section 3.4, again a transmission map is calculated, now for the 4 aperture nuller. As expected from theory, the transmission map for a 4 aperture nuller has a rectangular pattern, instead of the vertical fringes. For a monochromatic incident

field, the fringes do not lose visibility near the edges of the field of view. The minimum null depth for the transmission map shown in Figure 6.16a is $8.95 \cdot 10^{-8}$. For a polychromatic incident wave, $2 \mu\text{m}$ in the case of Figure 6.16b, the fringe visibility reduces radially outward from the centre of the field of view. The minimum null depth has increased by approximately 5%, to $9.39 \cdot 10^{-8}$.

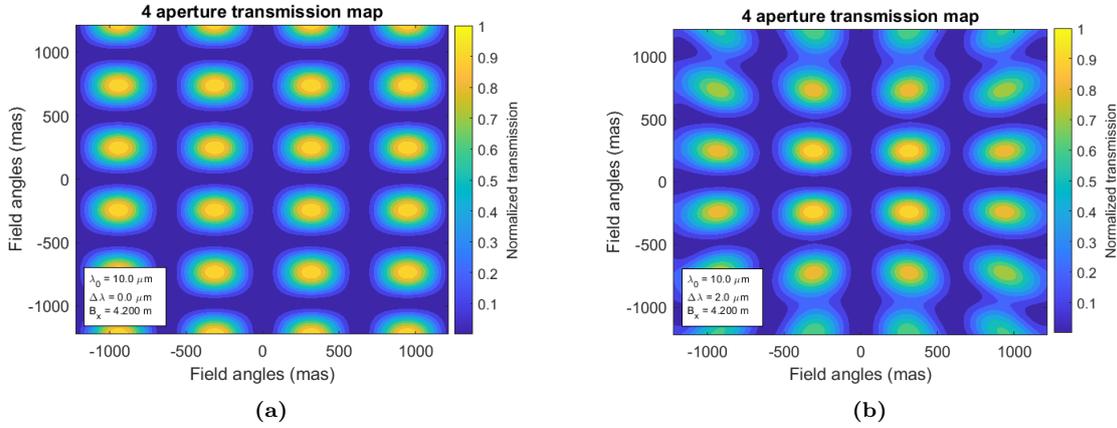


Figure 6.16: (a) Monochromatic transmission map of a 4 aperture, double Bracewell nulling interferometer
(b) Polychromatic transmission map, with a $2 \mu\text{m}$ bandwidth.

6.4. Intermezzo: Nulling without an optical model

At this point, before the perturbation analysis will be discussed, it is useful to take a step back and consider the different ways to perturb the null. There are several contributions to degrading the null, each with their own means of correction:

1. Difference in optical path length. Can be corrected with delay lines.
2. Beam shear, can be corrected with a set of periscope.
3. Difference in spot size. Can be corrected with afocal that has variable expansion/compression ratio.
4. Difference in amplitude. Can be corrected with a knife edge.
5. Even when both beams overlap perfectly, differences in the phase distribution of one of the spots can degrade the null. Higher order aberrations can be corrected with deformable mirrors or spatial filtering. Tilt errors can be corrected with a tip tilt mirror.

The question that arises is, what is the effect of each contribution on the null and are all methods of correction needed? Most of these contributions have been estimated by Hyde et al. [68], with the goal of providing a nulling budget for the FKSI. To verify the estimates given in Hyde et al., a simple model is made in Python. In this model, two circular spots with variable amplitude and phase are combined with and without an achromatic π phase shift:

$$I_{\text{destructive}} = \iint |A_1 + A_2 e^{i(\delta\varphi + \pi)}|^2 dy dx, \quad (6.1)$$

$$I_{\text{constructive}} = \iint |A_1 + A_2|^2 dy dx, \quad (6.2)$$

where

$$A_1 = \begin{cases} 1, & \text{if } x^2 + y^2 < r_1^2 \\ 0, & \text{otherwise} \end{cases} \quad A_2 = \begin{cases} 1 + \delta A, & \text{if } (x - \delta x)^2 + y^2 < r_2^2 \\ 0, & \text{otherwise} \end{cases} \quad (6.3)$$

from which the null depth is calculated as $I_{\text{destructive}}/I_{\text{constructive}}$. Polarisation effects are neglected, to simplify the model. In the ideal case, $\delta A = 0$, $\delta x = 0$, $r_1 = r_2$ and $\delta\varphi = 0$. For the following discussion, all parameters will be assumed ideal, unless mentioned otherwise.

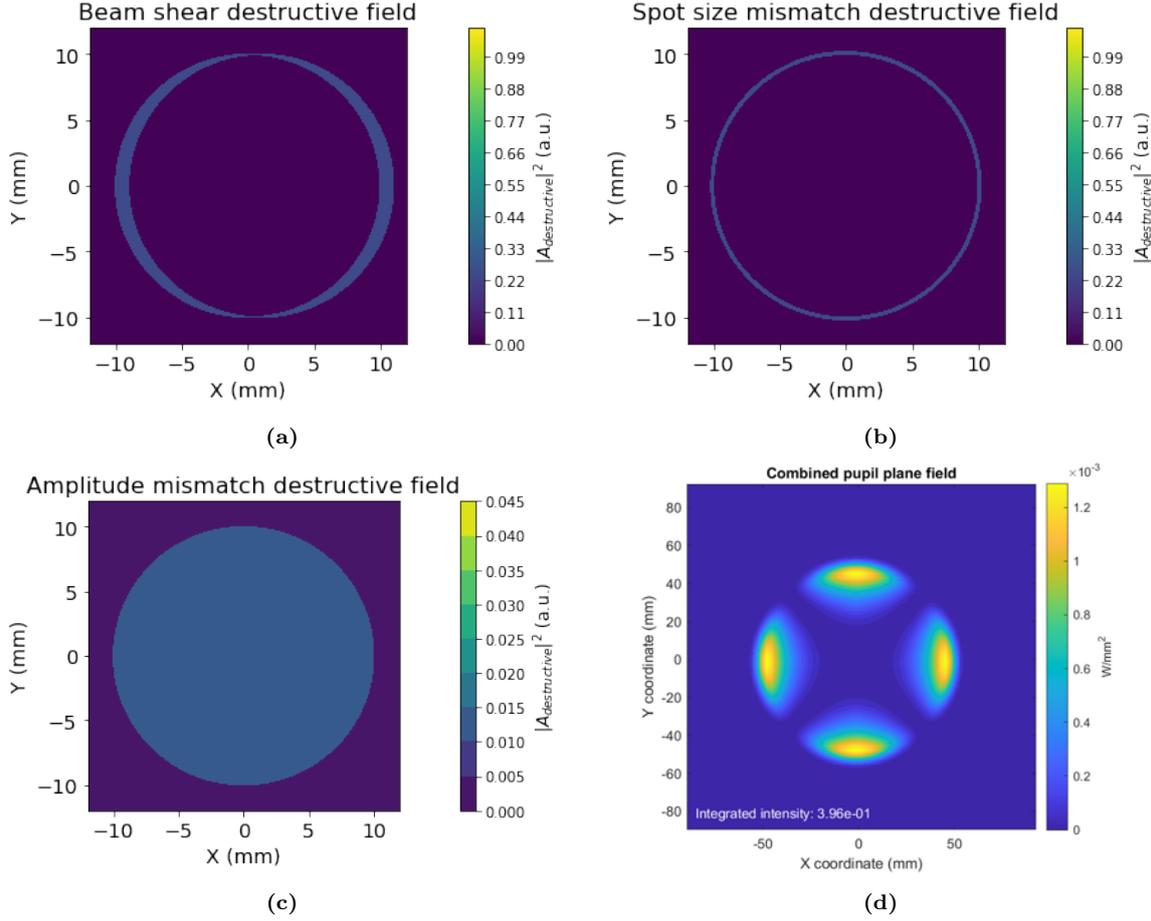


Figure 6.17: Several ways of degrading the null are shown here, as described by the model in Equation 6.1 and Equation 6.3. (a) Destructive field, after a certain translation δx has been applied. (b) Destructive field, showing a ring of non-zero intensity, due to a radius mismatch ($r_1 \neq r_2$). (c) Destructive field which has been degraded uniformly, due to an amplitude mismatch ($\delta A \neq 0$). (d) Destructive field, after astigmatism has been applied to one of the complex amplitudes ($\delta\varphi = Z_2^2$, see Table 2.2).

For the first case, a uniform $\delta\varphi$ is used: $\delta\varphi = 2\pi/\lambda \cdot z$. In Figure 6.18a, the above model is compared to the prediction given in Hyde et al., showing that both models agree. In red, a threshold is given for a null depth loss of 10^{-5} . To achieve this, the OPD should not exceed 1/1000th of a wave, or no more than 10 nm for a wavelength of 10 μm .

For the second case, δx is taken to be non-zero. In Hyde et al., the beam shear null depth loss is estimated as $1 - f$, with f the pupil overlap. With some basic geometry and algebra, it can be shown that the overlap area is given as:

$$A_{\text{overlap}} = 2r^2 \cos^{-1}\left(\frac{d}{2r}\right) - \frac{d}{2} \sqrt{4r^2 - d^2}, \quad (6.4)$$

with r the beam radius and d the linear displacement, such that the pupil overlap $f = A_{\text{overlap}}/(2\pi r^2)$. The comparison with the model and Hyde et al. is shown in Figure 6.19a, from which it can be seen that the beam displacement should be no more than $\sim 3 \cdot 10^{-5}$ beam radii, or 30 nm for a beam radius of 1 cm. In their paper, Hyde et al. make the comment on page 554 that, *"The use of the fibre also assures 100% beam overlap"*. Although this is indeed the case after fibre coupling, the coupling efficiency with the fibre is reduced when the beam overlap is non-ideal.

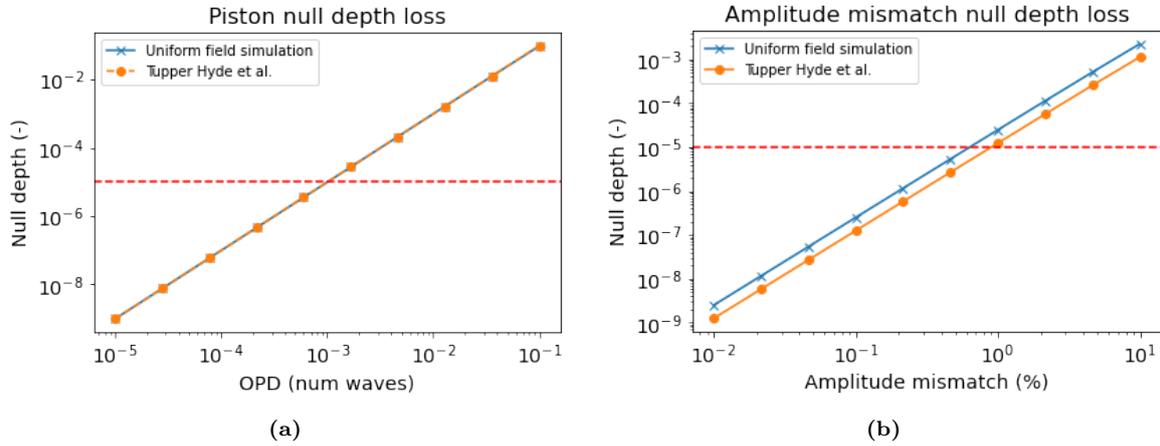


Figure 6.18: (a) Null depth loss as a result of piston in one of the beams. At an OPD of 1/1000th of a wave, the null depth loss exceeds the threshold of $1 \cdot 10^{-5}$. (b) The null depth loss due to a mismatch in amplitude between the two interfering beams.

The third case is not mentioned in Hyde et al., so it cannot be compared. However, it is expected that a mismatch in spot size is on the same order of magnitude as beam shear. In fact, this requirement is marginally more stringent compared to the beam shear requirement. Indeed, as is shown in Figure 6.19b, the spot size mismatch should be no more than $\sim 2 \cdot 10^{-5}$ beam radii.

For the fourth case, an amplitude mismatch between the interfering beams was modelled. They are compared in Figure 6.18b, showing a factor two difference between the estimate and model. Nevertheless, the amplitude mismatch should be no more than about a percent.

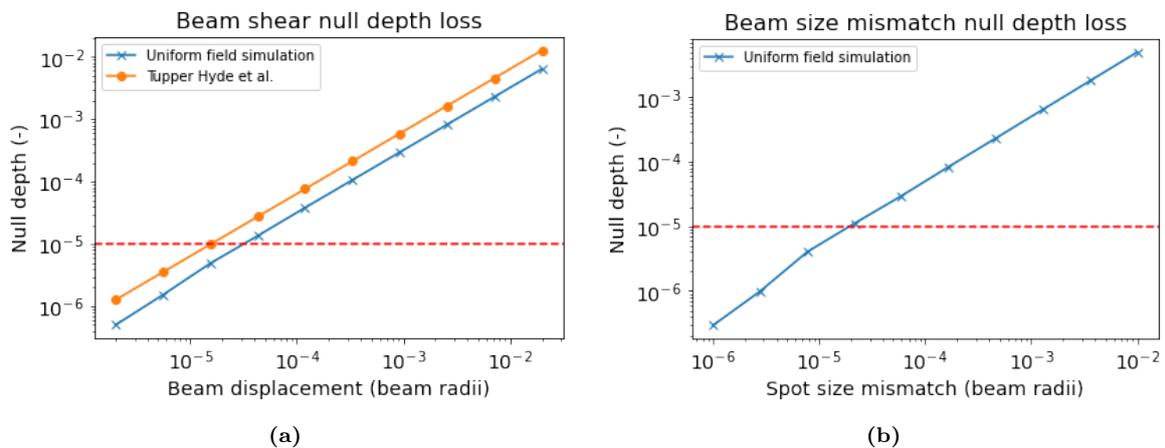


Figure 6.19: (a) Null depth loss due to beam shear, plotted on a log-log scale. (b)

Finally, a simple pupil plane combiner was made in Code V, in order to verify the null depth

calculation under perturbations. The model consists of two clear apertures (10 mm diameter) and an image plane, such that the effect of optics does not have to be considered. To calculate the null depth dependence, a Zernike polynomial is applied to the aperture with a certain aberration magnitude, while the other aperture remains unaberrated. The results are shown in Figure 6.20, expressed both in P-V magnitude of the aberration, as well as the RMS magnitude. As derived by Hyde et al. [68], the loss of null depth does not depend on the particular shape of the aberration, but only on the RMS magnitude, according to the following relation:

$$\delta N_{Aberration} = \frac{1}{4} \left(\frac{2\pi\sigma}{\lambda} \right)^2 \quad (6.5)$$

Some values are compared in Table 6.3, showing the agreement between both models. The requirement on the RMS WFE can then be calculated, filling in $1 \cdot 10^{-5}$ on the left side of Equation 6.5. This leads to a WFE of no more than 10 nm at $10 \mu m$, which is a very stringent requirement, provided that the state-of-the-art JWST had a requirement of an RMS WFE < 110 nm [69, p. 571]

RMS magnitude (waves)	Null depth loss (Hyde et al.)	Null depth loss (Code V)
$1 \cdot 10^{-1}$	$9.87 \cdot 10^{-2}$	$9 \cdot 10^{-2}$
$1 \cdot 10^{-2}$	$9.87 \cdot 10^{-4}$	$1 \cdot 10^{-3}$
$2 \cdot 10^{-3}$	$3.95 \cdot 10^{-5}$	$3 \cdot 10^{-5}$

Table 6.3: Comparison of the null depth loss, due to aberrations, derived by Hyde et al. and the null depth loss as simulated with Code V.

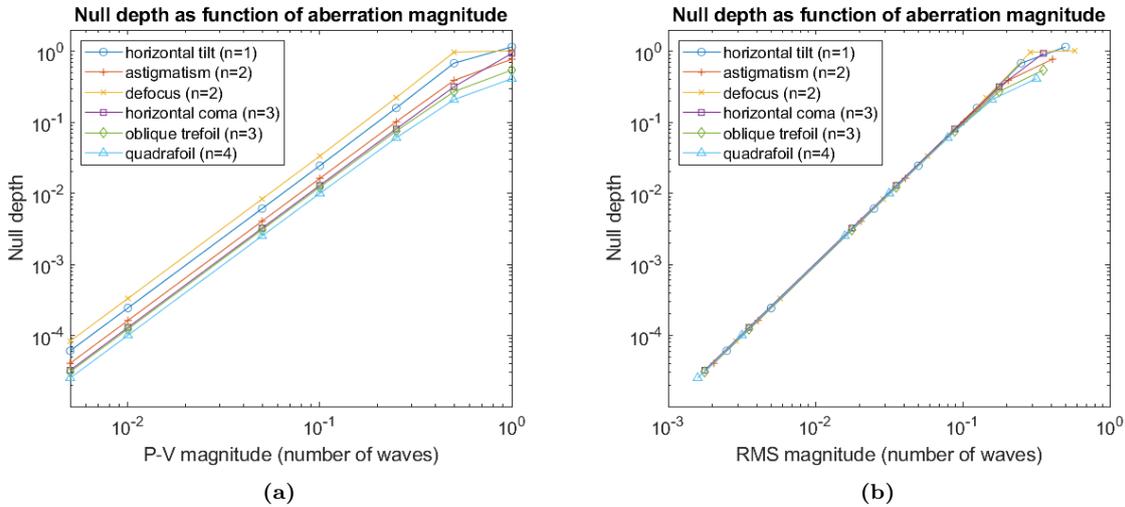


Figure 6.20: Null depth loss due a selection of aberrations, as function of the (a) P-V aberration magnitude and (b) RMS aberration magnitude.

The same pupil plane combiner was used to simulate null depth loss due to beam shear. The result is presented in Figure 6.21. For small amounts of beam shear, the Code V model underestimates the amount of null depth loss, which can be attributed to the sampling of the pupil plane spot. A sampling of $9 \mu m$ was used, so that a beam displacement of $1 \cdot 10^{-3}$ is already below the sampling rate. It should be noted that in practice, a spot is never a perfect top-hat as the one used in the pupil plane combiner. In a realistic spot, there are amplitude as well as phase variations across the pupil, such that undersampling becomes less of an issue.

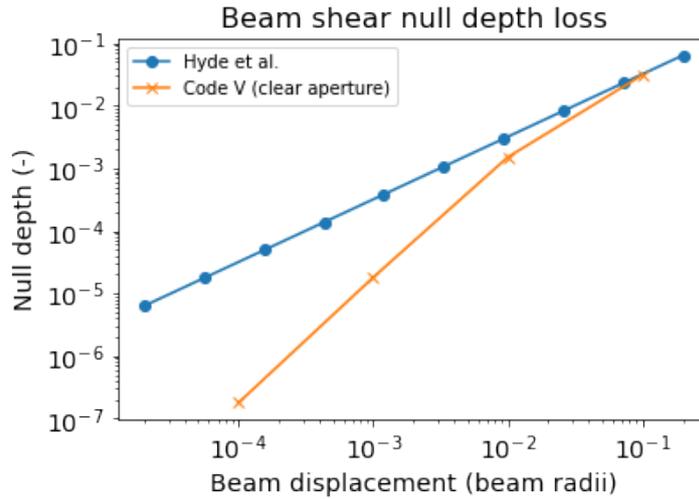


Figure 6.21: Comparison of the null depth loss due to beam shear, as derived by Hyde et al. and simulated by Code V. For small displacements, the Code V model underestimates the null depth loss.

The results of this section can be summarised in Table 6.4.

Null depth loss contribution	Requirement	Unit
Piston	1E-3	Waves
Amplitude mismatch	1	%
Beam shear	3E-5	Beam radii
Beam size mismatch	2E-5	Beam radii
RMS wavefront error	1E-3	Waves

Table 6.4: The considered contributions to null depth loss and their corresponding requirements.

6.5. Perturbation analysis - Off axis designs

In this section, the results of the perturbation analysis will be discussed. The goal of this analysis is to find an optical design that is least sensitive to perturbations in one of the surfaces. First, to analyse the stability of the null depth as function of the collimator focal length, an axial translation perturbation is done in subsection 6.5.1. Then, the perturbation analysis is extended to other parameters, to determine which telescope design performs best. As the perturbation parameter space is too large to analyse extensively, a small selection is made. An overview of the parameters was discussed in section 5.4 and a summary is given in Table 6.5.

Code V command	Description
XDE/YDE/ZDE	Lateral translation along the x/y/z axis
ADE/BDE	Rotation around the local x/y axis
RDY	Radius of curvature
SCO C1	Conic constant
SCO C2	Off-axis angle

Table 6.5: Code V commands of the considered perturbation parameters.

6.5.1. Axial translation perturbation

For this perturbation mode, discussed in subsection 5.4.2, the collimator of one of the interferometer arms is displaced along the optical axis. For the following simulations, a monochromatic field was used, with a wavelength $10\ \mu\text{m}$.

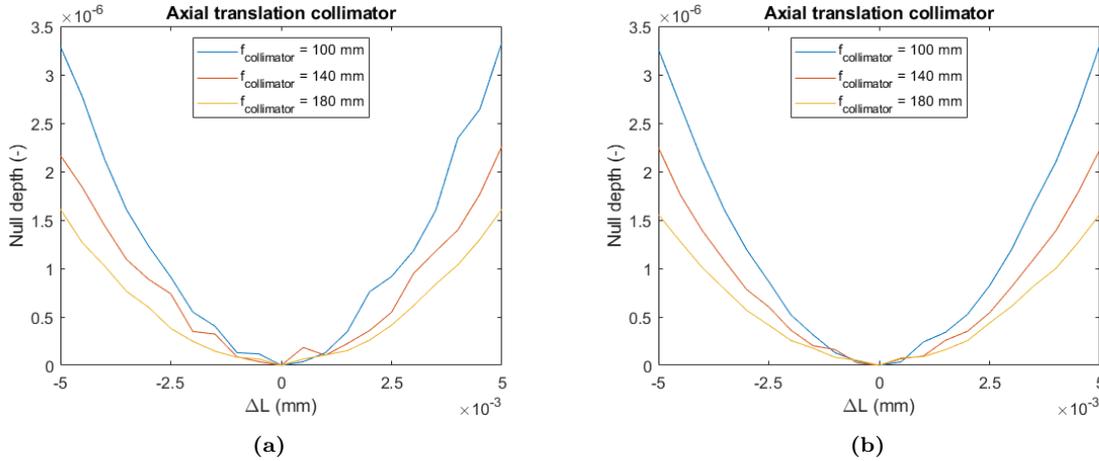


Figure 6.22: (a) Axial translation of the focal length for different values of the collimator focal length. The smaller focal length collimator clearly shows a higher sensitivity to axial perturbations, compared to the longest focal length. A tilt angle of -40.34 was used on the collimator. (b) Axial translation of the focal length for different values of the collimator focal length, using a collimator tilt angle of -39.1 .

The axial perturbation reduces the null depth due to two contributions: First, the defocus causes coma aberration in the reflected field. Second, due to the defocus, the spot size is reduced or increased, causing a mismatch with the unperturbed spot of the other aperture. Figure 6.22a shows that decreasing the focal length, and thus decreasing the radius of curvature of the collimator, the interferometer becomes more sensitive to axial translation. This result agrees with the approximation given in Equation 5.1. It also reveals a trade-off to be made, for this specific design of collimator. Although maximising the radius of curvature would reduce the sensitivity to axial perturbation, it does come at the cost of having a larger spot size on the detector and thus the need for a larger active area detector, increasing the cost of the instrument. However, there is still other perturbation modes to consider, so we are not yet in a position to assess this trade-off.

To see if the result is consistent across designs, the simulation is repeated with a primary mirror tilt of 15° , which decreases the collimator tilt from -40.34° to -39.1° . Again, the collimator is translated along the optical axis and the null depth is determined as function of the translation. The result is shown in Figure 6.22b and agrees well with Figure 6.22a.

6.5.2. Lateral translation perturbation

From the previous section, it was concluded that for axial perturbation, the change in null depth is minimised when the collimator focal length is maximised, regardless of the tilt angle. To continue the sensitivity analysis, lateral translation will be considered in this section. First, perturbation along the global x, y and z axis will be considered for the three mirror surfaces making up the telescope. The result is shown in Figure 6.23.

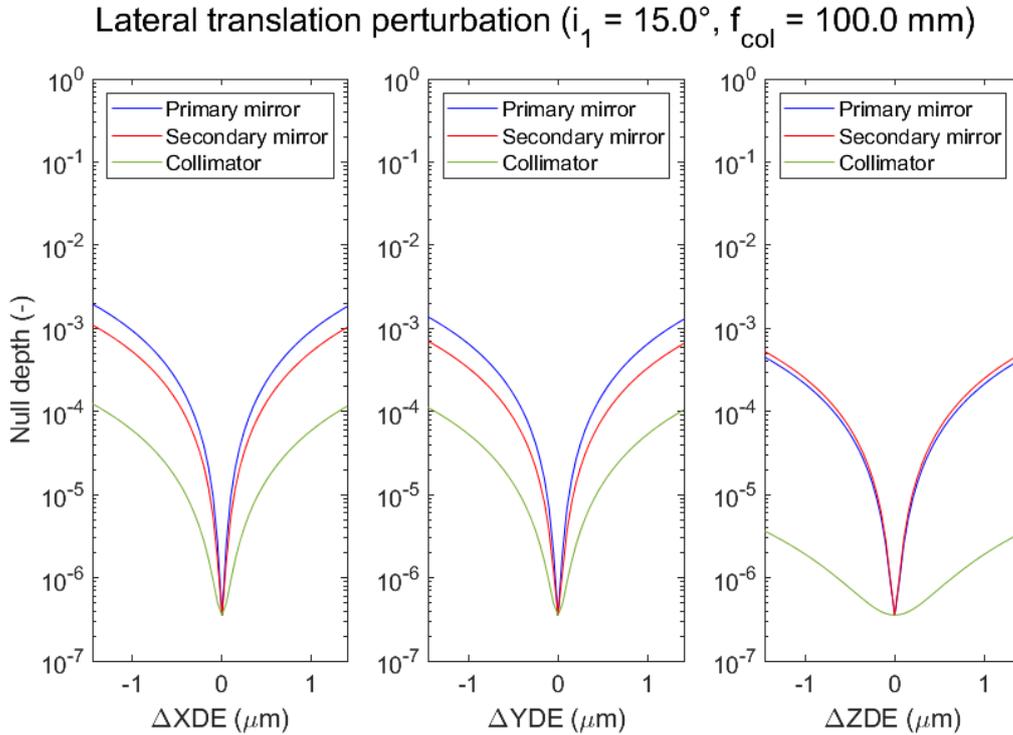


Figure 6.23: Null depth as function of lateral translation perturbation, applied to each telescope surface, along the x, y and z axis. Δx and Δy show similar behaviour, with the primary mirror being the most sensitive, then the secondary mirror and lastly the collimator. For Δz perturbations, the collimator is again the least sensitive and the primary and secondary mirror have similar sensitivities.

Several observations can be made from this. First, for all three perturbation modes, the collimator is the least sensitive. This can be explained from the fact that the collimator is the last curved optical surface of the telescope. Generally speaking, curved surfaces add aberrations to a wavefront, whereas flat surfaces have minimal effect on the wavefront. It is then to be expected that aberrations due to a perturbation of the primary mirror are amplified by the secondary mirror and collimator. On the other hand, aberrations caused by a decenter of the collimator propagate largely unaltered to the pupil plane. To further illustrate this point, Table 6.6 shows the exit pupil fit to each Zernike polynomial for perturbations along the z axis. The RMS wavefront error for each surface perturbation corresponds to the behaviour seen in the null depth. In other words, the primary and secondary mirror have comparable null depths for a given ΔZDE translation, while the RMS wavefront errors are on the same order. Second, given that the collimator is least sensitive to perturbations in the z direction, it would be a suitable component for correcting defocus in the focal plane. From Table 6.6, we see that defocus is the most significant term in the Zernike decomposition (ignoring piston, as that is corrected by the delay lines) for a z decenter. It is in part for this reason that a two element collimator would be favourable over a one element collimator. Third, when comparing

perturbations along the x and y axis for a given surface, they are within the same order of magnitude. This is somewhat curious, given the asymmetry of the collimator. It is at this point useful to compare the results from section 6.4 with the null depths found in Figure 6.23. In table Table 6.8, the beam shear and RMS WFE are given after the collimator is translated by $1 \mu m$. Comparing the numbers shows that the Code V model underestimates the null depth by one to two orders of magnitude. This could be due to the sampling rate in the detector plane, which is large compared to the beam shear. Nonetheless, the difference between x, y and z perturbation is comparable to what is expected from the Hyde et al. model. It also shows that the largest contribution to degradation of the null is due to beam shear.

Perturbed surface	Zernike number	Aberration name	RMS wavefront error
Primary mirror	1	Piston	0.0798
	2	Horizontal Tilt	-
	3	Vertical Tilt	0.0198
	4	Vertical Astigmatism	0.0188
	5	Defocus	0.0460
	6	Oblique Astigmatism	-
	7	Horizontal Trefoil	-
	8	Horizontal Coma	-
	9	Vertical Coma	0.0069
	10	Vertical Trefoil	0.0025
		RMS of polynomial	0.0540
Secondary mirror	1	Piston	0.0856
	2	Horizontal Tilt	-
	3	Vertical Tilt	0.0205
	4	Vertical Astigmatism	0.0202
	5	Defocus	0.0494
	6	Oblique Astigmatism	-
	7	Horizontal Trefoil	-
	8	Horizontal Coma	-
	9	Vertical Coma	0.0072
	10	Vertical Trefoil	0.0028
		RMS of polynomial	0.0577
Collimator	1	Piston	0.0057
	2	Horizontal Tilt	-
	3	Vertical Tilt	0.0005
	4	Vertical Astigmatism	0.0009
	5	Defocus	0.0033
	6	Oblique Astigmatism	-
	7	Horizontal Trefoil	-
	8	Horizontal Coma	-
	9	Vertical Coma	0.0002
	10	Vertical Trefoil	-
		RMS of polynomial	0.0034

Table 6.6: Table containing the Zernike polynomial expansion coefficients for the exit pupil map. The coefficients are determined by fitting the corresponding Zernike polynomial to the exit pupil sphere. Each surface was perturbed by a z decenter of $100 \mu m$. Zernike polynomials that have $<1 \cdot 10^{-4}$ RMS wavefront errors are denoted by "-".

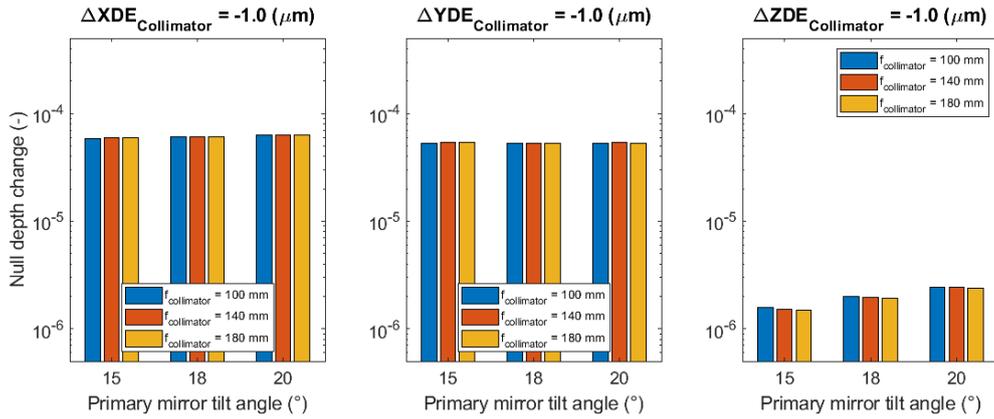


Figure 6.24: Three graphs showing the effect of perturbations applied to the collimator. For both XDE and YDE perturbations, the null depth change does not depend significantly on neither the primary mirror tilt angle, nor the collimator focal length. For ZDE misalignment, there is some dependence on the primary mirror tilt angle, with a lower tilt angle causing the least change in the null depth.

6.5.3. Collimator focal length dependence

To be able to compare different design variations, the data shown in Figure 6.23 can be reduced to several data points instead of traces for each perturbation and each surface. To this end, a vertical line is drawn at a given perturbation magnitude. The null depth at that perturbation magnitude is then taken as a single data point. In that case, Figure 6.23 would be reduced to nine data points, three for each subplot. This allows us to quantify the dependence of both the tilt angle of the primary mirror and the collimator focal length on the null depth. For this part of the analysis, nine variations on the off-axis Cassegrain were made, varying the primary mirror tilt angle from 15 to 20 degrees in three steps and varying the collimator focal length from 100 to 180 mm in three steps.

First, the XDE, YDE and ZDE parameters are compared in Figure 6.24. For both XDE and YDE perturbations, neither the primary mirror tilt angle nor the collimator focal length changes the sensitivity significantly, whereas there is some dependence on the collimator focal length, as well as the primary mirror tilt angle for ZDE perturbations. It should be noted here again that the primary mirror tilt angle influences the collimator tilt angle, as discussed in section 4.5. From that discussion it became clear that the dependence of the collimator tilt angle on the primary mirror tilt angle is relatively weak.

The comparison between XDE and YDE on one the hand and ZDE on the other, raises two questions:

1. Why is it that XDE and YDE perturbations result in a degradation of the null by almost two orders of magnitude more compared to ZDE perturbation?
2. Why is there some dependence on the tilt angle for ZDE perturbation, while this dependence is negligible for XDE and YDE perturbations?

When the collimator is perturbed along the x and y axis, astigmatism, coma and defocus are present in the aberrated wavefront. From Figure 6.25b, it can be seen that of those three contributions, astigmatism is the dominant aberration. The same aberrations are present when the collimator is decentered along the z axis, however, defocus is the most significant

contribution for this perturbation, which becomes apparent from the exit pupil map plotted in Figure 6.25c. In addition to wavefront errors, a path length difference is introduced, for all three cases, which is corrected by the delay lines. Nonetheless, the path length difference is more significant when the collimator is decentered along the z axis.

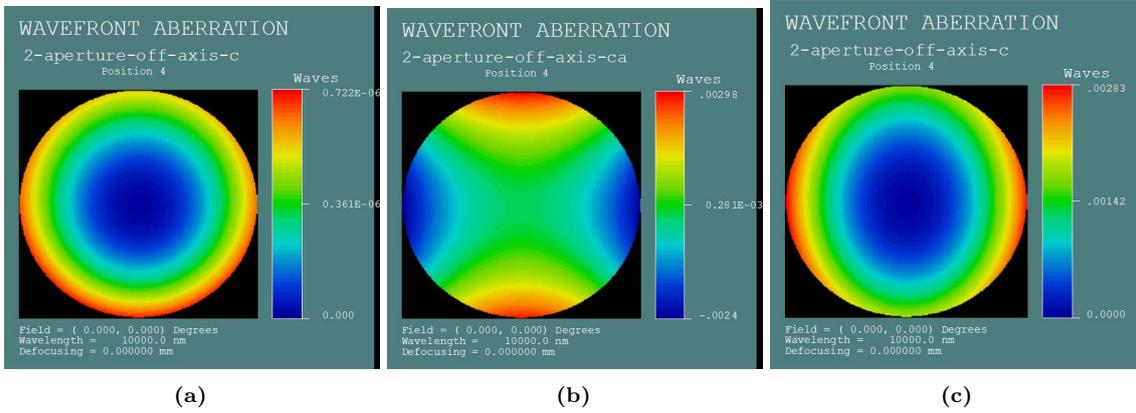


Figure 6.25: (a) Nominal exit pupil map of the nulling interferometer. (b) Exit pupil map resulting from a $20\ \mu\text{m}$ perturbation along the y axis, with a P-V WFE of $5.38 \cdot 10^{-3}$ waves and RMS WFE of $1 \cdot 10^{-3}$ waves. (c) Exit pupil map resulting from a $20\ \mu\text{m}$ perturbation along the z axis, with a P-V WFE of $2.83 \cdot 10^{-3}$ waves and RMS WFE of $7 \cdot 10^{-4}$ waves.

The exit pupil map gives part of the explanation why a y decenter degrades the null more significantly compared to a z decenter. In section 6.4, it was shown that the RMS WFE can be used to estimate the effect on the null depth. In this case, the RMS WFE due to decenter along the x axis is about 50% higher, compared to the RMS WFE due to a decenter along the z axis, resulting in a null depth difference of about a factor of two. A much larger contribution comes from the beam shear. A Δz perturbation results in defocus as well as beam shear, although the beam shear is not as significant as when the collimator is perturbed along the x axis.

Geometrical derivation

Instead of drawing conclusions from just the wavefront maps and Zernike decompositions, it is insightful to consider the problem geometrically. To this end, a geometrical derivation has been done to quantify the contribution of defocus and astigmatism. The hypothesis is that both perturbations (Δx , Δz) can be decomposed in two components and that those components manifest differently for the Δx and Δz misalignments.

The geometry of a Δz perturbation is presented in Figure 6.26a. In this schematic, l_F is the distance to the focus, l'_F is the new distance to the focus after perturbation, θ_0 is the angle subtended by the surface vertex and the focus and $\theta'_{0,z}$ the new angle. Note that before and after the Δz decenter, the optical axis ray angle does not change, such that $\theta_0 - \theta'_{0,z}$ can be interpreted as a field angle. The apparent field angle resulting from a perturbation Δz is given by:

$$\delta\theta(\delta z) = \theta_0 - \theta'_0 = \theta_0 - \tan^{-1} \left(\frac{l_F \sin(\theta_0)}{l_{F,z} + \delta z} \right). \quad (6.6)$$

To verify if this relation makes sense, let us consider two edge cases. When $\delta z = 0$, $\delta\theta$ should be 0. Indeed,

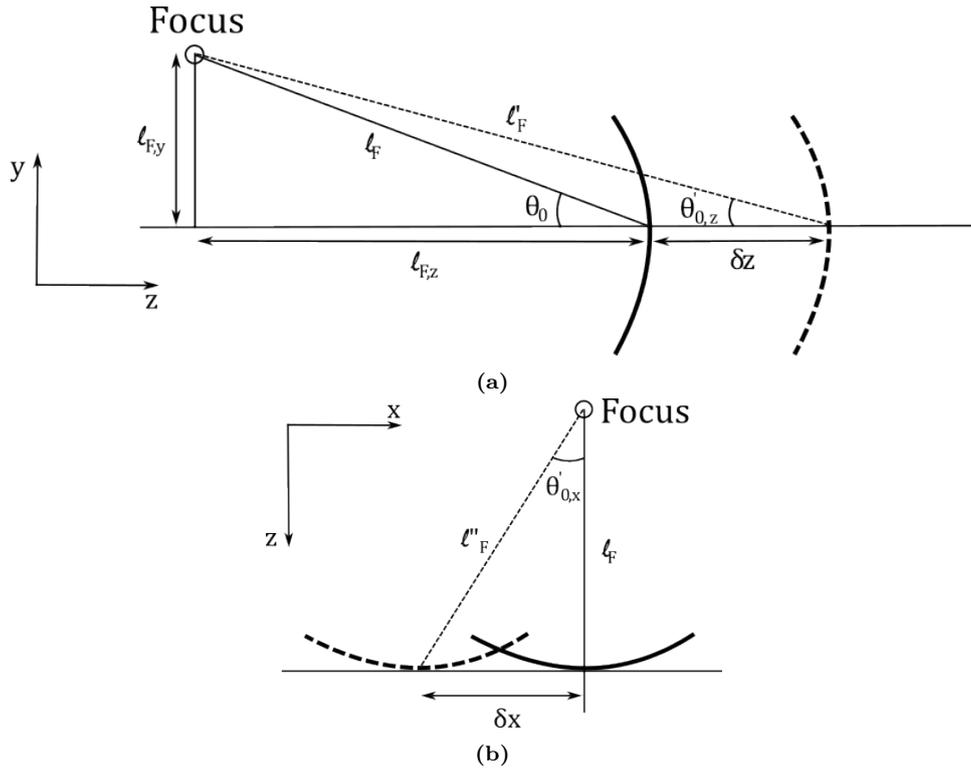


Figure 6.26: (a) Geometrical representation of a perturbation δz along the z axis, as viewed in the y-z plane. l_F is the nominal distance to the focus and l'_F is the distance to the focus after misalignment. θ_0 (θ'_0) is the angle subtended by the surface vertex and the focus. (b) The geometry of a δx perturbation, as seen from the x-z plane. For both schematics, the solid surface indicates the unperturbed position, whereas the dashed surface is the perturbed position.

$$\delta\theta(\delta z = 0) = \theta_0 - \tan^{-1}\left(\frac{l_F \sin(\theta_0)}{l_{F,z}}\right) = \theta_0 - \tan^{-1}\left(\frac{l_F \sin(\theta_0)}{l_F \cos(\theta_0)}\right) = 0. \quad (6.7)$$

The second edge case is when $\theta_0 = 0$, in other words, the parabola is on-axis. In that case, there is no astigmatism for $\delta z \neq 0$ and again $\delta\theta = 0$.

A schematic representation of a Δx misalignment is shown in Figure 6.26b. The distance to the focus is again l_F and l''_F after a Δx misalignment. The field angle due to a perturbation in the x direction is given by:

$$\delta\theta(\delta x) = \theta'_{0,x} = \tan^{-1}\left(\frac{\delta x}{l_F}\right). \quad (6.8)$$

To show which of these two are more significant for a given value of δz or δx , the dependence is compared in Figure 6.27. This shows that the field angle in the z-x plane grows faster for a given perturbation magnitude, compared to the z-y plane.

The deviation from the nominal focal length due to the perturbations are:

$$\delta l_F(\delta z) = l'_F - l_F = \sqrt{l_{F,y}^2 + (l_{F,z} + \delta z)^2} - l_F, \quad (6.9)$$

$$\delta l_F(\delta x) = l''_F - l_F = \sqrt{l_F^2 + \delta x^2} - l_F. \quad (6.10)$$

With a Taylor expansion it can be shown that $\delta l_F(\delta x) \propto \delta x^2$, whereas $\delta l_F(\delta z) \propto \delta z$. Given that δx and δz are close to 0, this means that $|\delta l_F(\delta z)| > |\delta l_F(\delta x)|$.

To conclude the geometrical derivation, both δx and δz perturbations were decomposed into two components: a field angle ($\delta\theta$) and a defocus (δl_F) component. Given that $\delta\theta(\delta x) > \delta\theta(\delta z)$ and $|\delta l_F(\delta z)| > |\delta l_F(\delta x)|$, it can be assumed that for a δx perturbation the field angle, and thus astigmatism, is more significant, whereas for a δz perturbation, the defocus component is more significant. This result agrees with what was shown in Figure 6.25c. Furthermore, Equation 6.8 and Equation 6.10 show that there is no dependence on the collimator tilt angle for a δx perturbation, whereas there is some dependence on the tilt angle for a δz decenter.

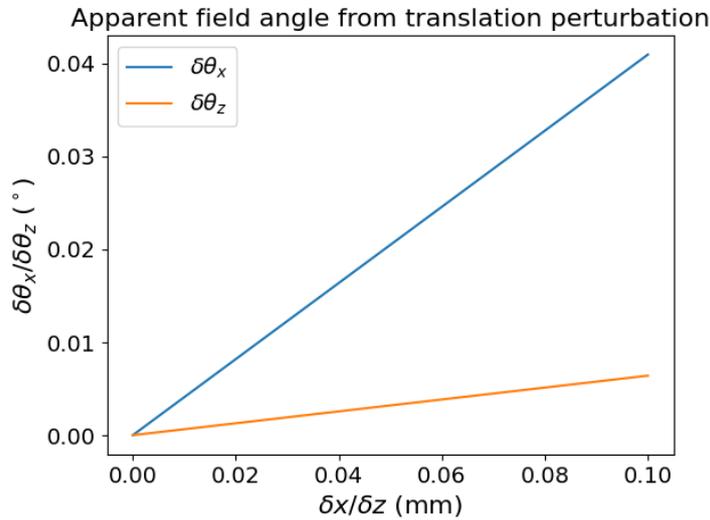


Figure 6.27: Apparent field angle as function of δx or δz . For the given range, $\delta\theta_x$ is always larger than $\delta\theta_z$.

Rotational perturbation

To continue the discussion, let us consider the results shown in Figure 6.28. The first observation is that for ADE, BDE as well as SCO C2 perturbation, an increase in collimator focal length results in a higher null depth change. This is exactly opposite of axial perturbation of the collimator, in which case a longer focal length reduces the effect on the null. The main contribution for this case is astigmatism, which scales with the focal length, such that a longer focal length results in more astigmatism for a given field angle. Again, the primary mirror tilt angle has relatively little effect on the sensitivity. The second observation is that ADE perturbations have more effect on the null depth, compared to BDE perturbations. This can be expected from the asymmetry in the system, where an excessive ADE tilt mainly causes vertical astigmatism, whereas an excessive BDE tilt causes oblique astigmatism, with a marginally higher RMS wavefront error, as shown in Table 6.7. Just as before, the exit pupil maps were simulated for both an ADE and BDE perturbation of 0.1° , after which the RMS WFE is derived by fitting the wavefront error to the given Zernike polynomials. The third

observation is that off axis angle perturbations are much less sensitive, compared to ADE and BDE tilts. Whereas ADE and BDE tilts of 18 mas cause a null depth change on the order of 10^{-8} , for off axis angle perturbations, it is 30 times larger. To give a sense of whether or not it is possible to keep the pointing within such a small angle, it can be compared to the optical telescope element of the JWST, which was required to reach a line of sight stabilization within 7.3 milliarcseconds [69, p. 561]. Table 6.9 gives a breakdown of the three major contributions to null depth loss for the three perturbation parameters discussed in this section. Given that a change in the off-axis angle only results in wavefront errors, it is the least sensitive of the three. For ADE and BDE, beam shear and spot size difference dominate the null depth loss.

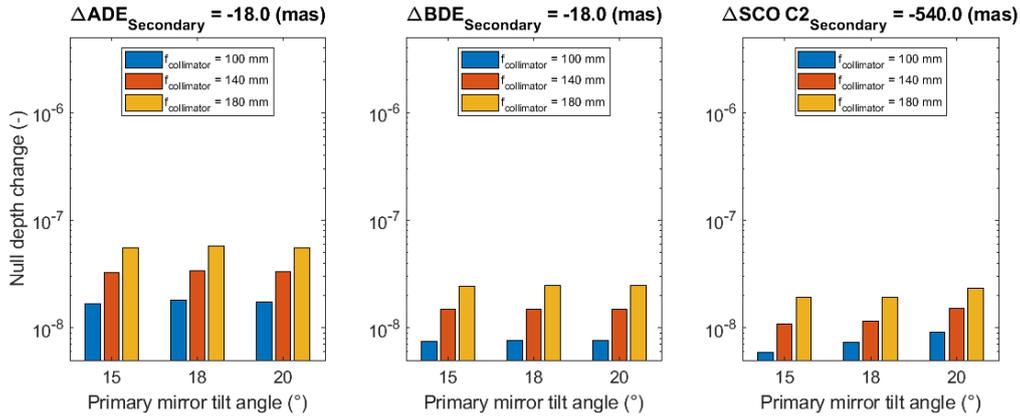


Figure 6.28: Left: Change of the null depth, compared to the nominal case, due to rotations around the x axis. Centre: Null depth change due to rotations around the y axis. Right: Null depth change resulting from perturbations of the off axis angle

Perturbation	Zernike Number	Aberration Name	RMS WFE
ADE	4	Vertical Astigmatism	0.0146
	5	Defocus	0.0006
	6	Oblique Astigmatism	0
BDE	4	Vertical Astigmatism	0
	5	Defocus	0
	6	Oblique Astigmatism	0.0095

Table 6.7: Comparison of Zernike decomposition for ADE and BDE perturbation of 0.1° . Zernike terms 1 to 3 are omitted, as they have no contribution to the RMS WFE.

Surface shape perturbation

Finally, perturbations in the radius of curvature and conic constant are shown in Figure 6.29. As can be seen from Figure 6.29, changes in the radius of curvature degrade the null equally, independent of tilt angle and only weakly dependent on the collimator focal length. The resulting aberration is dominantly defocus, which is independent of tilt angle. When the conic constant is perturbed, there are two effects on the beam: defocus and astigmatism. Whereas RDY perturbations are not dependent on the collimator focal length, the conic constant perturbations do show a dependence on the focal length. This can be attributed to astigmatism, which worsens with increasing focal length. Once more, the beam shear, spot diameter difference and RMS WFE due to the perturbations are given in Table 6.10. Contrary to the previously considered perturbation parameters, only the spot diameter difference contributes significantly to the degradation of the null.

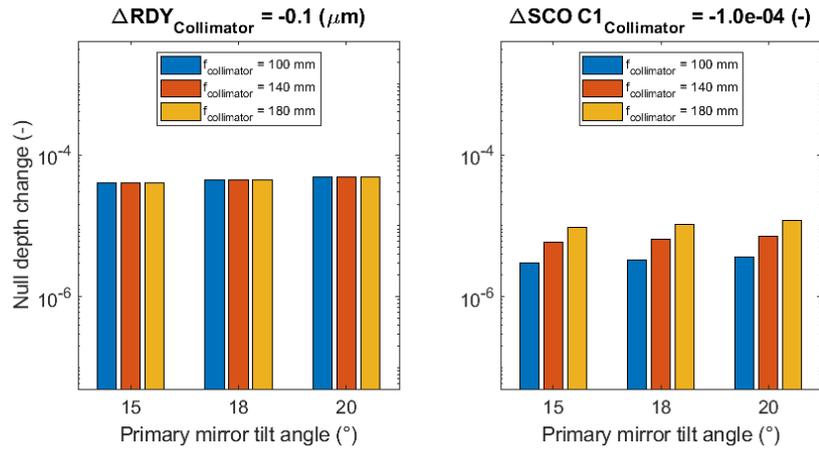


Figure 6.29: Comparison of null depth change due to RDY (radius of curvature) and SCO C1 (conic constant) perturbations.

To summarise this section, a perturbation analysis was done on the off-axis Cassegrain design. First, the collimator was perturbed along the optical axis, showing that increasing the focal length of the collimator minimises the null depth change for that particular parameter. Second, all surfaces were displaced laterally in a global coordinate system, showing that displacements along the x and y axis cause the same change in the null depth, where the largest contribution comes from beam shear. For decenters along the z axis, both beam shear and spot size differences perturb the null. Third, the perturbation parameters ADE, BDE and SCO C2 were applied to the telescope surfaces. This showed that minimising the collimator focal length makes the telescope least sensitive to these particular perturbations. The largest contribution to degrading the null for ADE and BDE perturbations is again beam shear, whereas for SCO C2 perturbations, only wavefront errors degrade the null. Finally, changes in the radius of curvature and conic constant were considered. Perturbing these parameters do not result in any beam shear, but have a large effect on the spot size in the pupil plane, degrading the null.

Perturbation	Beam shear (μm)	Expected null depth	Spot diameter dif- ference (μm)	Expected null depth	RMS WFE (waves)	Expected null depth
1 μm XDE	19.42	9.02E-4	1.69	1.0E-4	5.6E-5	3.1E-8
1 μm YDE	20.14	9.36E-4	5.37	2.8E-4	5.5E-5	3.0E-8
1 μm ZDE	4.31	2.00E-4	6.74	3.4E-4	3.5E-5	1.2E-8

Table 6.8: Beam shear, spot size difference and RMS WFE due to translation of the collimator. The third, fifth and seventh column indicate the expected null depth, based on the model from Hyde et al. The beam shear was determined with Code V's "Real Ray Trace" function, the spot diameter was determined with the "Spot diagram function" and the RMS WFE was determined with the "Wavefront analysis" function.

Perturbation	Beam shear (μm)	Expected null depth	Spot diameter dif- ference (μm)	Expected null depth	RMS WFE (waves)	Expected null depth
1" ADE	26.35	1.29E-3	1.42	8.74E-5	3.8E-5	1.43E-8
1" BDE	17.06	8.33E-4	0.25	1.89E-5	2.5E-5	6.17E-9
1" SCO C2	0	-	0	-	6.4E-5	4.04E-8

Table 6.9: Expected null depth loss, based on the Hyde et al. model, for the considered rotational perturbation parameters.

Perturbation	Beam shear (μm)	Expected null depth	Spot diameter dif- ference (μm)	Expected null depth	RMS WFE (waves)	Expected null depth
50 μm RDY	0	-	112.91	4.12E-3	2.0E-3	3.95E-5
1E-4 SCO C1	0	-	19.61	8.82E-4	7.8E-4	6.01E-6

Table 6.10: Expected null depth loss, based on the Hyde et al. model, for the considered rotational perturbation parameters.

6.6. On axis design

In section 4.4 it was stated that it is known that an off-axis design has better throughput compared to an on-axis design and suffers less from diffractive effects of the central obscuration, as well as the suspension structure of the secondary mirror. To assess whether or not this notion affects the nuller performance, an on-axis Cassegrain design was made and compared to the off-axis design. Both telescopes have an effective focal length of 8 m and entrance pupil diameter of 1 m. The on-axis Cassegrain, shown schematically in Figure 6.30 has a central obscuration ratio of 0.33.

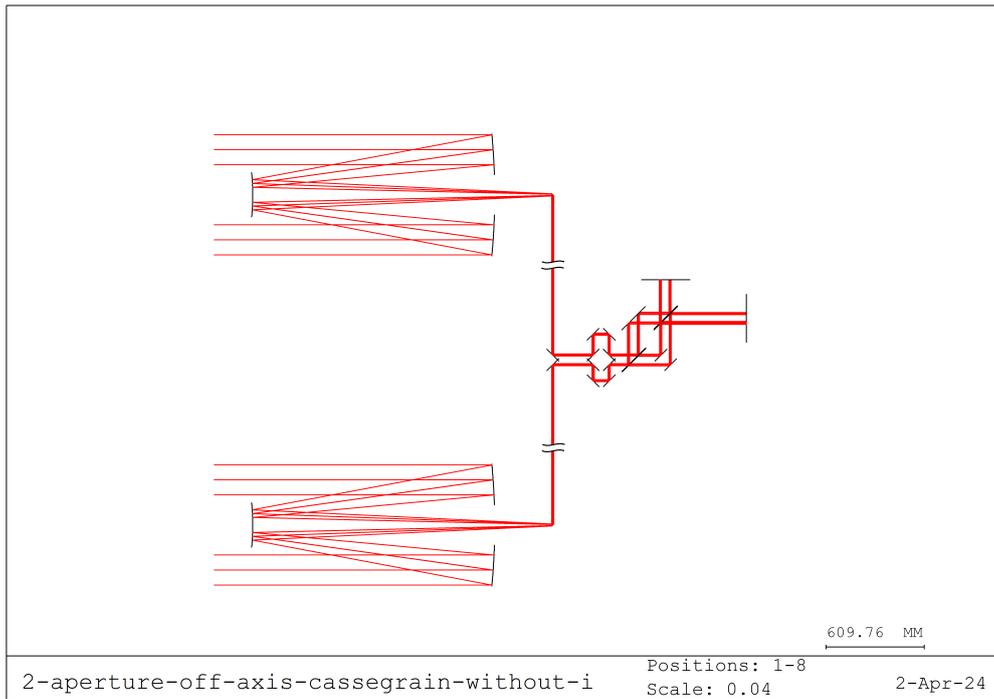


Figure 6.30: Side view schematic of the on-axis Cassegrain nulling interferometer model.

Following the same methodology as for the off-axis Cassegrain, the null depths at a certain perturbation magnitude can be plotted and compared to the off-axis case, as shown in Figure 6.31 to Figure 6.33.

First, the XDE, YDE and ZDE misalignment are compared in Figure 6.31, for the global coordinate case. For the XDE and YDE misalignments, there is a clear downwards trend between the consecutive surfaces, such that the primary mirror perturbation degrades the null the most, whereas the collimator perturbation degrades the null the least. Additionally, the difference between the on-axis and off-axis design for the XDE and YDE misalignment are at most a factor of two, while for the ZDE case, the difference between the on-axis and off-axis is between one and two orders of magnitude. As was established in subsection 6.5.3, the main contributor to null depth change for the XDE and YDE perturbation, when uncorrected, is beam shear, which is on the same order of magnitude for the on- and off-axis Cassegrain. On the other hand, for the ZDE misalignment, there can be a defocus component, as well as beam shear, depending on which surface is considered. For the on-axis Cassegrain, the beam stays aligned with the optical axis, such that misalignment along the z axis of the primary and secondary mirror only cause an enlargement of one of the spots. The clear downwards trend that is present for the XDE and YDE perturbations is not seen for ZDE decenters, neither for

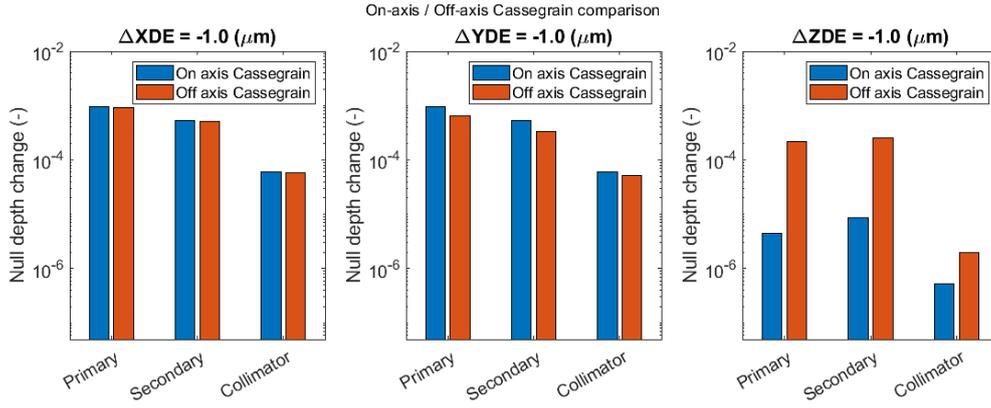


Figure 6.31: Left: null depth as a result of XDE perturbation applied to all three telescope surfaces. Middle: null depth as a result of YDE perturbation applied to the telescope surfaces. Right: null depths due to ZDE perturbation.

the on-axis nor the off-axis Cassegrain.

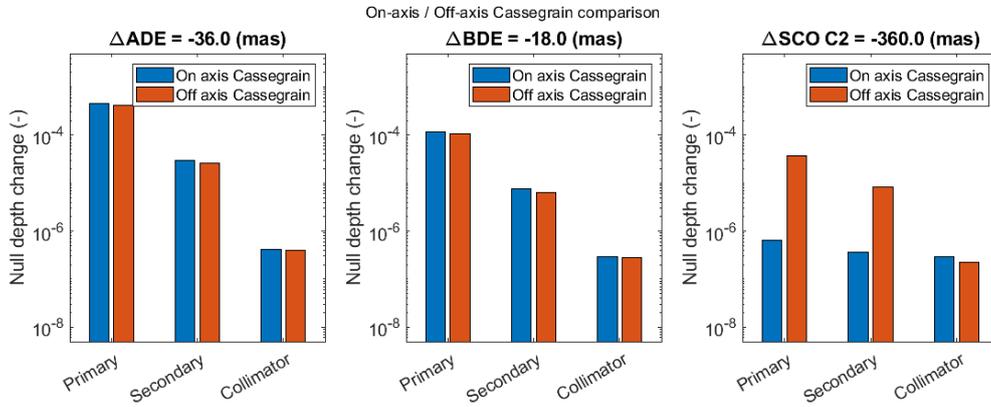


Figure 6.32: Left: ADE (alpha tilt) perturbation comparison between the on axis and off axis Cassegrain telescope design, shown for all three telescope optical surfaces. Middle: BDE (beta tilt) perturbation comparison. Right: SCO C2 (off axis angle) perturbation comparison.

Second, the null depth change due to excessive tilt around the x axis, y axis and perturbation of the off axis angle are shown in Figure 6.32. The results are similar to the previously considered parameters, as also here there is a clear downwards trend in null depth change for consecutive surfaces and there is only a marginal difference between the on- and off-axis Cassegrain for ADE and BDE perturbations. A notable difference is that for off axis angle perturbations on the collimator, both the on- and off-axis Cassegrain show a null depth change of around 10^{-7} , whereas for the same misalignment on the primary and secondary mirror, the difference in null depth change between the on-axis and off-axis design is almost two orders of magnitude. This can be partly explained by the similarity in geometry for the collimator in both designs. For the on-axis Cassegrain, a tilt angle and off-axis angle of 45 degrees is used. For the off-axis Cassegrain, it is -40.34° and -49.66° , respectively. Compare this to for example the primary mirror, with a tilt angle and off-axis angle of 0 degrees in the on-axis design, whereas for the off-axis Cassegrain, the tilt angle and off-axis angle is 15 and -15 degrees.

Lastly, the null depth change due to RDY and SCO C1 perturbation are shown in Figure 6.33. As was already discussed in subsection 6.5.3, RDY perturbations are only weakly dependent

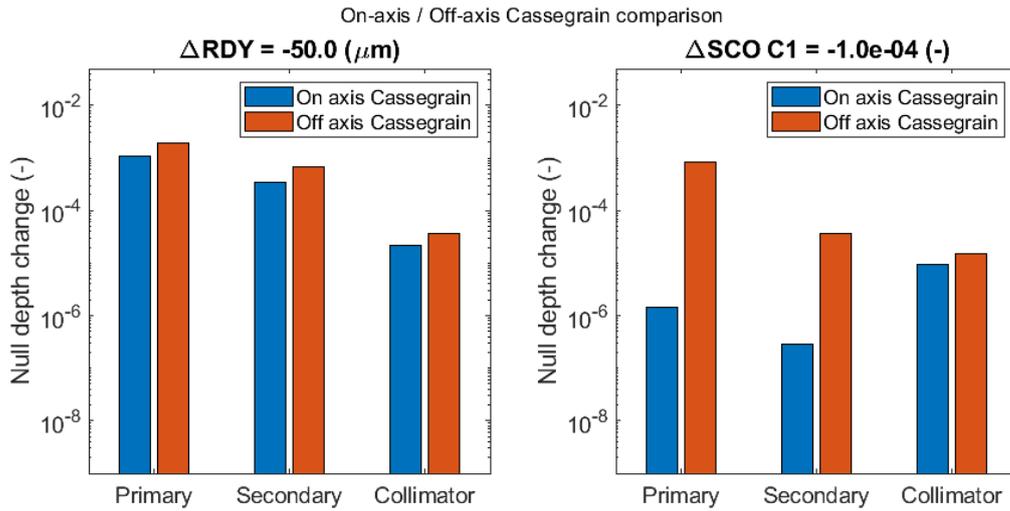


Figure 6.33: Left: Null depth change due to RDY (radius of curvature) perturbation, shown for the primary, secondary and collimator surface. Right: Null depth change due to SCO C1 (conic constant) perturbation.

on the primary mirror tilt angle, which can also be seen from the figure. The differences in null depth change due to conic constant perturbations are quite distinct, compared to the RDY results. As these results are difficult to interpret geometrically (with the exception of the collimator, due to the similarity in geometry), the exit pupil maps can again be considered, given in Figure 6.35 for the secondary mirror and Figure 6.34 for the primary mirror. From this, the difference in aberrations becomes clear. For the on-axis Cassegrain and perturbations of the secondary mirror conic constant, the exit pupil map shows spherical aberration, with a P-V value of $6.6E-4$ waves. The off-axis Cassegrain suffers from defocus, astigmatism, as well as trefoil and a negligible coma component, resulting in a P-V wavefront aberration of $4.8E-2$ waves.

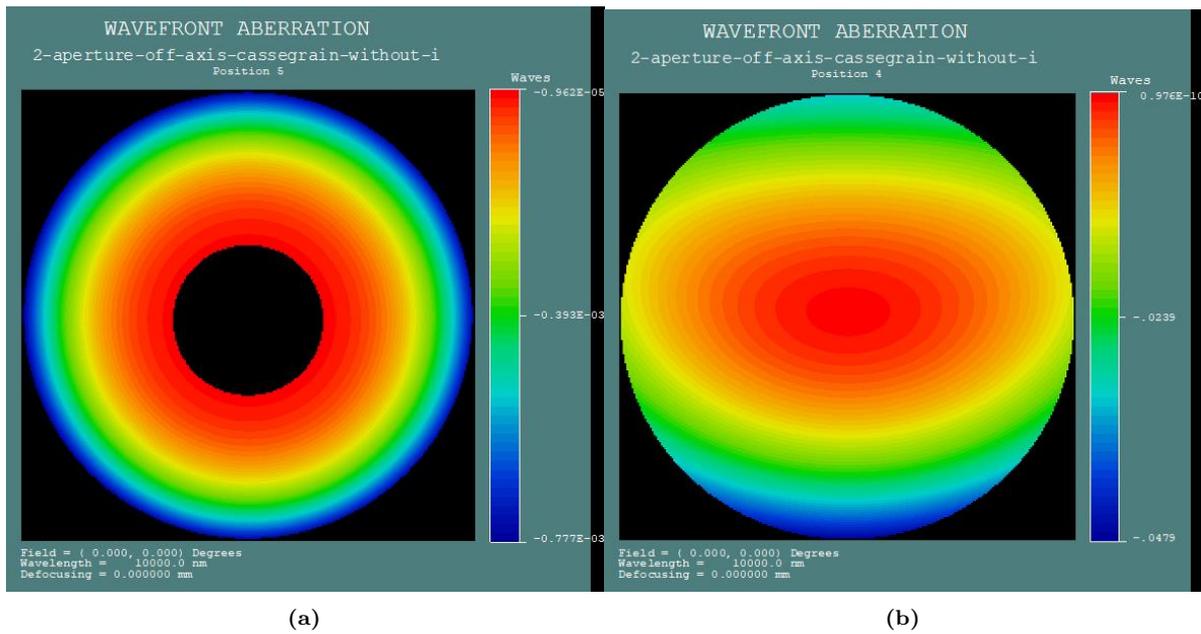


Figure 6.34: (a) Exit pupil map of the on-axis Cassegrain, with an SCO C1 perturbation of $1 \cdot 10^{-4}$ applied to the primary mirror, having a P-V WFE of $6.6E-4$ waves. (b) Exit pupil map of the off-axis Cassegrain, with an SCO C1 perturbation of $1 \cdot 10^{-4}$ applied to the primary mirror, with a P-V WFE of $4.8E-2$ waves.

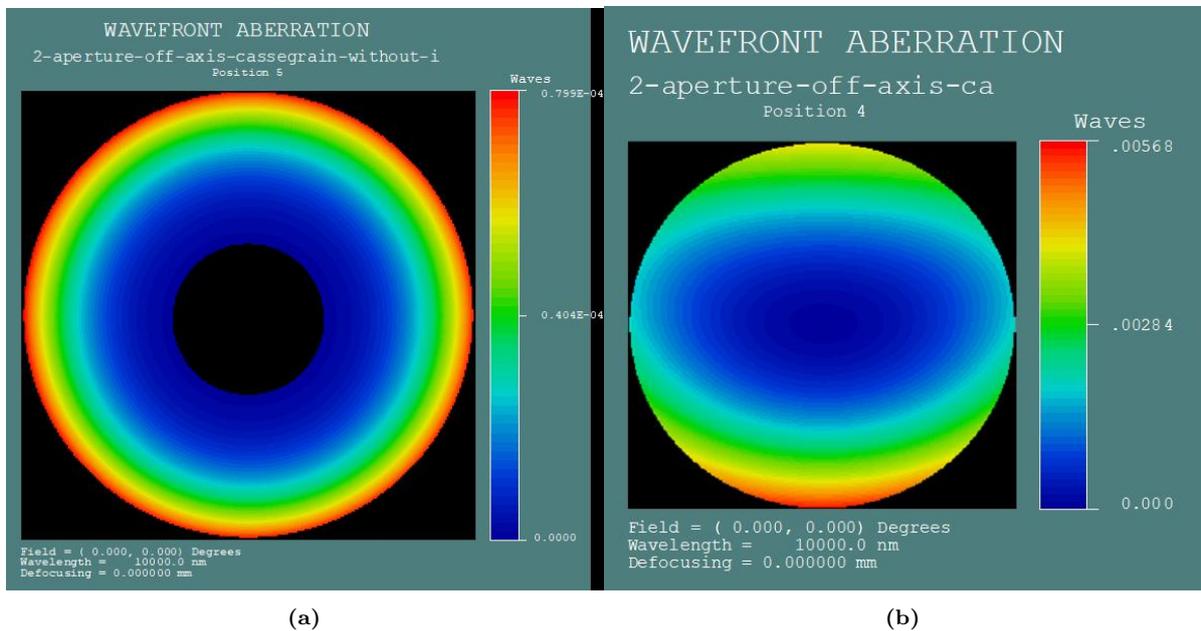


Figure 6.35: (a) Exit pupil map of the on-axis Cassegrain, as a result of an SCO C1 perturbation of $1E-4$ on the secondary mirror, showing a P-V WFE of $8E-5$ waves. (b) Exit pupil map of the off-axis Cassegrain, as a result of an SCO C1 perturbation of $1E-4$ on the secondary mirror, with a P-V WFE of $5.7E-3$ waves. The exotic shape of the exit pupil map can be attributed to the large astigmatism and trefoil component of the wavefront aberration.

6.7. Classical telescope design comparison

For this last section, three telescope designs will be compared in terms of their stability and performance. They are based off of the classical Cassegrain, Ritchey-Chretien (RC) and Gregorian, discussed in section 4.3. The design parameters are given in Table 6.11. One notable difference between the RC and Gregorian is the presence of an intermediate focus in the latter. Because of this, the radius of curvature is rather small on the Gregorian secondary mirror. Furthermore, the secondary mirror is elliptical, instead of hyperbolic.

Surface	Parameter	Ritchey-Chrétien	Gregorian
Primary mirror	R_1 (mm)	-3909.8096	-3732.0508
	K_1 (-)	-1.0310	-1.0000
	i_1 ($^\circ$)	15.0000	15.0000
	$\theta_{o,1}$ ($^\circ$)	-15.0000	-15.0000
Secondary mirror	R_2 (mm)	-1502.57162	764.8550
	K_2 (-)	-3.4426	-0.3881
	i_2 ($^\circ$)	10.3399	17.9012
	$\theta_{o,2}$ ($^\circ$)	-10.8873	-19.6601
Collimator	R_c (mm)	118.8482	154.1538
	K_c (-)	-1.0055	-1.0000
	i_c ($^\circ$)	40.3399	47.9012
	$\theta_{o,c}$ ($^\circ$)	49.6601	42.0988

Table 6.11: Design parameters for the off-axis RC and Gregorian.

Just like before, a perturbation is applied to each surface and the null depth is calculated. To get a more complete look compared to the one-dimensional representation of the null depth dependence given before, a radar plot is made with three axes: The minimum null depth, the null depth at a certain perturbation magnitude and the Full-Width at Half Maximum (FWHM) of the distribution, indicated in Figure 6.36. Note that the axes on the radar plot are chosen such that a larger area indicates better stability and performance.

Perturbations applied to the secondary mirror are shown in Figure 6.37 to Figure 6.39. Starting with the translation perturbation parameters, the Gregorian is the most sensitive, where the null is degraded by one order of magnitude more compared to the Cassegrain and RC. For the rotational parameters ADE and BDE, the Gregorian is marginally better compared to the Cassegrain and RC. The off-axis angle (SCO C2) is the only parameter where the Gregorian is the most stable of all three designs, outperforming the Cassegrain and RC by one order of magnitude. For radius of curvature and conic constant perturbations, the Gregorian is again surpassed by the Cassegrain and RC, both in terms of the stability, as well as minimum null depth. In most cases, the RC is marginally more stable under perturbation compared to the Cassegrain telescope design.

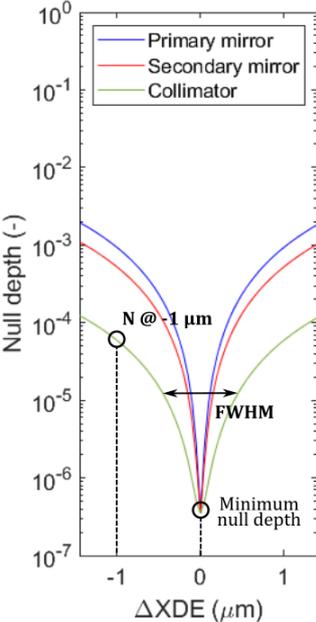


Figure 6.36: The radar plot consists of three axes, the data points of which are taken from the above plot, as indicated.

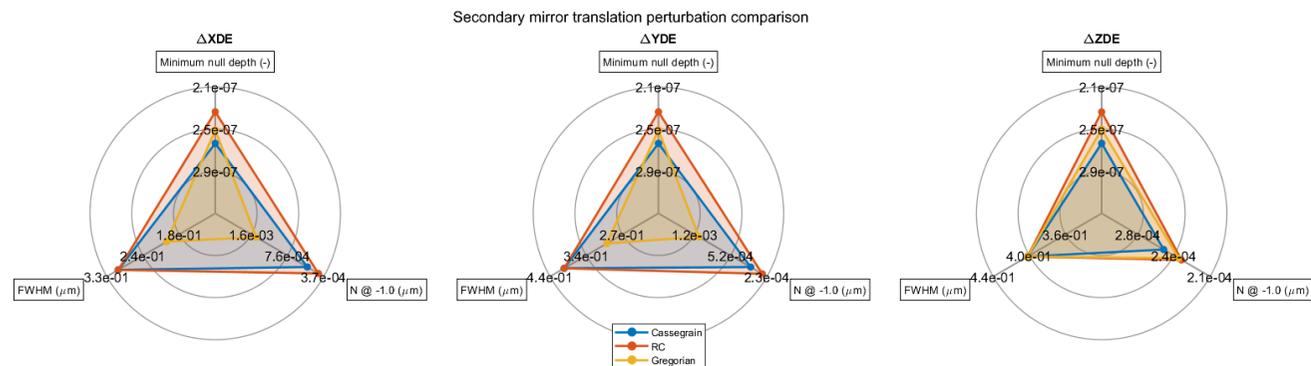


Figure 6.37: Translation perturbation of the secondary mirror. The comparison is made between an off-axis Cassegrain, RC and Gregorian design. All null depths calculated with a $10 \mu\text{m}$ central wavelength and bandwidth of $1 \mu\text{m}$.

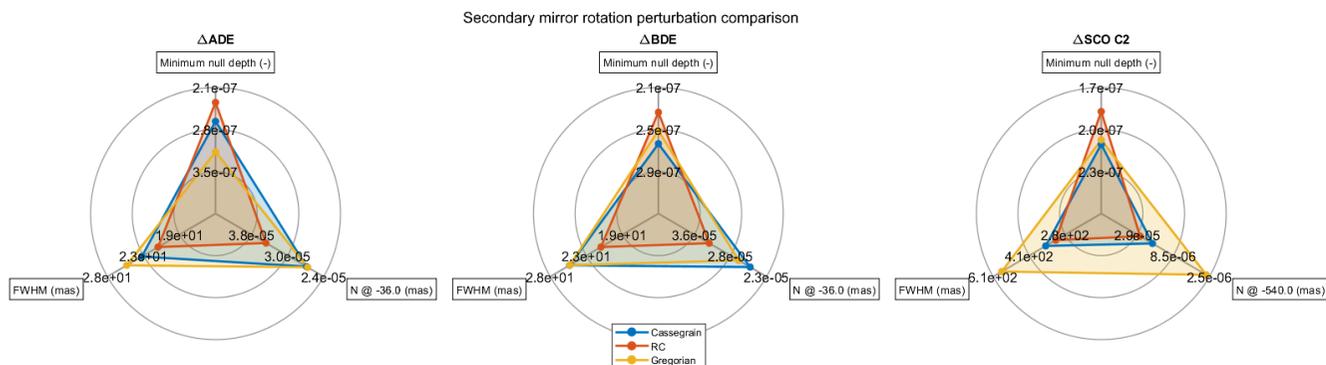


Figure 6.38: Rotation perturbation of the secondary mirror. The comparison is made between an off-axis Cassegrain, RC and Gregorian design.

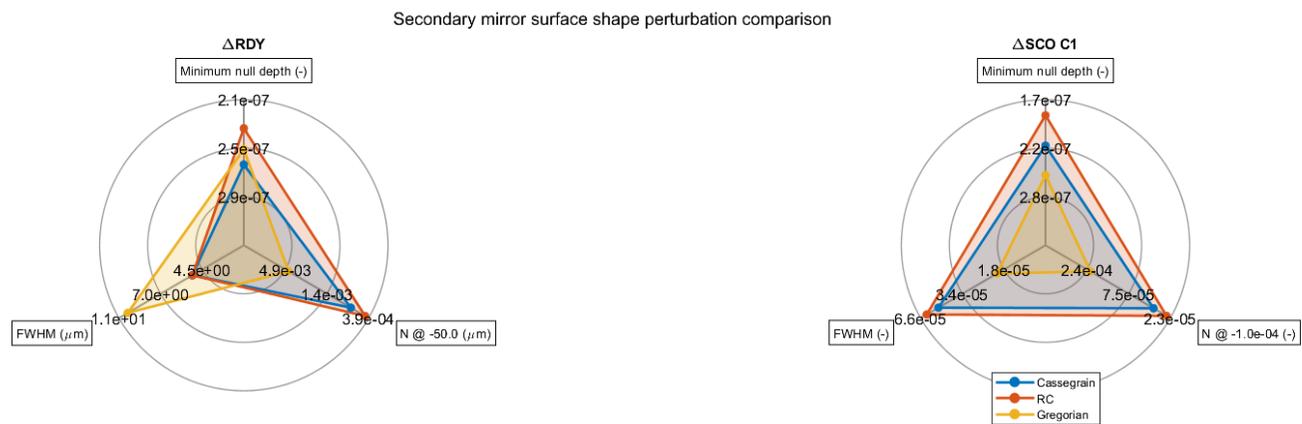


Figure 6.39: Surface shape perturbation of the secondary mirror. The comparison is made between an off-axis Cassegrain, RC and Gregorian design.

7

Recommendations

The discussed results have been obtained with some simplifications, which should be addressed in future research.

One of the idealisations that has been made was the achromatic phase shift. In the work presented, a digital π phase shift has been introduced, however, in reality this could be quite complicated, depending on the aperture configuration. For a 2 aperture interferometer, a π phase shift suffices, which can be introduced by either a focus crossing, or a periscope, given that those options are inherently achromatic. For a 3 or 5 aperture configuration, the phase shifts become fractions of π , such that the reflective solutions are incapable of providing the correct phase shift. For such a configuration, a dispersive plate would be needed, introducing a chromatic dependence.

Furthermore, In all the presented results, polarisation dependencies have been neglected, or simplified to linear polarisation, as this would complicate the analysis of the results. As the Fresnel equations show, reflection and transmission coefficients depend on the polarisation state of the incoming field, as well as the angle of incidence. Given that the pupil plane beam combination relies on the equivalence of the individual fields, polarisation dependencies makes the null more susceptible to perturbations in the optical surfaces, as these can cause a mismatch between the polarisation components.

Additionally, in this study, only one dimensional perturbations were considered, which is a major simplification. In an actual space based nulling interferometer, all perturbations should be considered simultaneously. This points towards a Monte Carlo study of the nulling interferometer, where many random samples of the perturbation parameter space are taken and applied. This way, a statistical analysis can be done to understand how likely it will be that a space based nulling interferometer will operate within the specifications, given the tolerances of the optical components. A Monte Carlo simulation was done for this study, however, it was not reported on in the main document, as the results remained unverified. For a proposal of the setup and preliminary results, the reader is referred to Appendix D.

The number of parameters tested is relatively small, compared to the parameter space of the optical model. Additional parameters that could be considered are for example gravity sag, thermal gradients, element or group wedge, or even polarisation related perturbations.

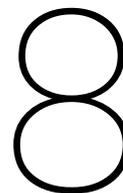
Especially if a Monte Carlo simulation would be performed, it is important to carefully consider the possible parameters and their magnitude, possibly with the addition of compensators.

Aside from the simplifications, future research could go in several directions, to make the model more representative of an actual implementation of a space-based nulling interferometer. One of those steps forward would be to make a full end-to-end simulation, in which planet yield would be taken into account. This would be the most representative metric, as null depth itself does not translate directly to the planet yield. To achieve such an end-to-end model, an interface could be implemented which works in conjunction with planetary population models such as P-POP [35, 70], or LIFEsim [47].

The most significant step forward that could be made is to build in correction capabilities into the model, specifically tip-tilt correction and lateral beam translation. As was discussed several times in chapter 6, beam shear is a major contribution to the loss of null depth. If corrected, it could reveal the fundamental limitations that would favour one design over the other.

As was shown in chapter 6, aberrations are a major contribution to degradation of the null. Although these could be corrected actively by a deformable mirror, that would complicate the design further. It could prove to be more worthwhile to passively compensate as many aberrations as possible, for example by exploring other telescope designs, such as the three mirror anastigmat, or Yolo telescope (named after the Yolo county in California) [71], or introducing aspheric terms to the discussed optical designs.

Lastly, spatial filtering is extensively discussed in the literature, as it would be beneficial to the stability of the null, but was not taken into account for this study. Modelling the coupling to a single mode fibre in Code V would have taken a considerable amount of time, if at all possible. Given that BSP can take an external file as input field, it would theoretically be possible to propagate up until a certain surface, use the complex amplitude at that surface as input for a separate model that calculates the field after spatial filtering and then use the filtered field as an input to further propagate through the model.



Conclusion

To conclude, several optical models of nulling interferometers were designed for this study. Both on-axis and off-axis telescope designs were considered, while for the beam combiner, a modified Mach Zehnder (MMZ) beam combiner was designed. The MMZ beam combiner optical model was verified by comparing its output to the Fresnel equations and showing constructive and destructive outputs. Before the sensitivity analysis was performed, the telescope design was determined to be diffraction limited over a wide field of view. Apart from the delay line, no correction capabilities were present in the optical model.

Comparing the results of the perturbation analysis with theoretical predictions, it was shown that the optical model underestimates the null depth loss in an absolute sense, although relatively, the optical model agrees with the theoretical derivations. Through the numerical perturbation analysis, it was shown that beam shear and spot size differences are the main contribution to degrading the null, when no tip-tilt mirrors, periscopes and adjustable afocals are present. The benefit of this is that correcting beam shear and spot size differences are less complex compared to an adaptive optics system. Nonetheless, this study underscores the need for some form of wavefront correction, as a RMS WFE of $\lambda/1000$ already degrades the null beyond the 10^{-5} requirement. The perturbation analysis also showed that the most significant wavefront errors are low-order, such as tip/tilt, astigmatism, defocus and coma, which can be corrected by means of actuators, possibly avoiding the need for deformable mirrors.

To maximise the throughput and fibre coupling efficiency, an off-axis telescope design would be ideal, but comes at the expense of increased complexity and thus cost of manufacturing, as well as increased sensitivity to perturbations. This last point was confirmed with a perturbation analysis, showing that the off-axis Cassegrain has comparable or worse sensitivity to the on-axis Cassegrain.

Lastly, three off-axis telescope designs were compared, which were the Cassegrain, Ritchey-Chrétien (RC) and Gregorian. The RC, due to its hyperbolic primary and secondary, is known to perform better than the Cassegrain. The Gregorian on the other hand is not favoured due to its performance, but is still used in some observatories because of the intermediate focus, which allows for placing a field stop. The perturbation analysis showed that perturbations applied to the secondary mirror of the Gregorian causes the largest null depth loss, which

can be attributed to its relatively tight radius of curvature. The RC showed slightly better performance and stability compared to the Cassegrain, making it a promising candidate for nulling interferometry.

References

- [1] Event Horizon Telescope Collaboration et al. “First M87 event horizon telescope results. IV. Imaging the central supermassive black hole”. In: *arXiv preprint arXiv:1906.11241* (2019).
- [2] D Spergel et al. “Wide-field infrared survey telescope-astrophysics focused telescope assets WFIRST-AFTA final report”. In: *arXiv preprint arXiv:1305.5422* (2013).
- [3] R. N. Bracewell. “Detecting nonsolar planets by spinning infrared interferometer”. In: *Nature* 274.5673 (1978), pp. 780–781. ISSN: 1476-4687. DOI: 10.1038/274780a0. URL: <https://doi.org/10.1038/274780a0>.
- [4] Robert D. Peters, Oliver P. Lay, and Peter R. Lawson. “Mid-Infrared Adaptive Nulling for the Detection of Earthlike Exoplanets”. In: *Publications of the Astronomical Society of the Pacific* 122.887 (2010), p. 85. ISSN: 1538-3873 0004-6280. DOI: 10.1086/649850. URL: <https://dx.doi.org/10.1086/649850%20https://iopscience.iop.org/article/10.1086/649850/pdf>.
- [5] Stefan Martin et al. “High performance testbed for four-beam infrared interferometric nulling and exoplanet detection”. In: *Applied optics* 51.17 (2012), pp. 3907–3921. ISSN: 2155-3165.
- [6] D. Defrère et al. “NULLING DATA REDUCTION AND ON-SKY PERFORMANCE OF THE LARGE BINOCULAR TELESCOPE INTERFEROMETER”. In: *The Astrophysical Journal* 824.2 (2016), p. 66. ISSN: 0004-637X. DOI: 10.3847/0004-637X/824/2/66. URL: <https://dx.doi.org/10.3847/0004-637X/824/2/66%20https://iopscience.iop.org/article/10.3847/0004-637X/824/2/66/pdf>.
- [7] W. Danchi et al. “Enabling the direct detection of earth-sized exoplanets with the LBTI HOSTS project: a progress report”. In: *Proc.SPIE*. Vol. 9907, p. 990713. DOI: 10.1117/12.2233397. URL: <https://doi.org/10.1117/12.2233397>.
- [8] M. M. Colavita et al. “The Keck Interferometer”. In: *Publications of the Astronomical Society of the Pacific* 125.932 (2013), p. 1226. ISSN: 1538-3873 0004-6280. DOI: 10.1086/673475. URL: <https://dx.doi.org/10.1086/673475>.
- [9] Romain Laugier et al. “Asgard/NOTT: L-band nulling interferometry at the VLTI”. In: (2023).
- [10] Olivier Guyon. “Limits of Adaptive Optics for High-Contrast Imaging”. In: *The Astrophysical Journal* 629.1 (2005), p. 592. ISSN: 0004-637X. DOI: 10.1086/431209. URL: <https://dx.doi.org/10.1086/431209>.
- [11] Christophe Malaterre et al. “Is there such a thing as a biosignature?” In: *Astrobiology* 23.11 (2023), pp. 1213–1227.
- [12] David J. Des Marais et al. “The NASA Astrobiology Roadmap”. In: *Astrobiology* 8.4 (2008). doi: 10.1089/ast.2008.0819, pp. 715–730. ISSN: 1531-1074. DOI: 10.1089/ast.2008.0819. URL: <https://doi.org/10.1089/ast.2008.0819>.

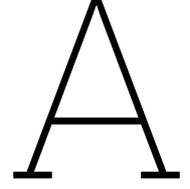
- [13] Sascha P. Quanz et al. “Atmospheric characterization of terrestrial exoplanets in the mid-infrared: biosignatures, habitability, and diversity”. In: *Experimental Astronomy* (2021). ISSN: 1572-9508. DOI: 10.1007/s10686-021-09791-z. URL: <https://doi.org/10.1007/s10686-021-09791-z><https://link.springer.com/content/pdf/10.1007/s10686-021-09791-z.pdf>.
- [14] David J. Des Marais et al. “Remote Sensing of Planetary Properties and Biosignatures on Extrasolar Terrestrial Planets”. In: *Astrobiology* 2.2 (2002). PMID: 12469366, pp. 153–181. DOI: 10.1089/15311070260192246. eprint: <https://doi.org/10.1089/15311070260192246>. URL: <https://doi.org/10.1089/15311070260192246>.
- [15] Alan J. Penny et al. “Darwin interferometer”. In: *Astronomical Interferometry*. Ed. by Robert D. Reasenberg. Vol. 3350. International Society for Optics and Photonics. SPIE, 1998, pp. 666–671. DOI: 10.1117/12.317130. URL: <https://doi.org/10.1117/12.317130>.
- [16] R. Martin Stefan et al. “TPF-Emma: concept study of a planet finding space interferometer”. In: *Proc.SPIE*. Vol. 6693, p. 669309. DOI: 10.1117/12.734835. URL: <https://doi.org/10.1117/12.734835>.
- [17] Bertrand Mennesson and Jean Marie Mariotti. “Array Configurations for a Space Infrared Nulling Interferometer Dedicated to the Search for Earthlike Extrasolar Planets”. In: *Icarus* 128.1 (1997), pp. 202–212. ISSN: 0019-1035. DOI: <https://doi.org/10.1006/icar.1997.5731>. URL: <https://www.sciencedirect.com/science/article/pii/S0019103597957316>.
- [18] Charley Noecker and Roger P Linfield. “Error analysis for a compound nulling interferometer”. In: *Interferometry in space*. Vol. 4852. SPIE, pp. 412–422.
- [19] Serge Dubovitsky and Oliver P Lay. “Architecture selection and optimization for planet-finding interferometers”. In: *New Frontiers in Stellar Interferometry*. Vol. 5491. SPIE, pp. 284–295.
- [20] Oliver P. Lay. “Imaging properties of rotating nulling interferometers”. In: *Applied Optics* 44.28 (2005), pp. 5859–5871. DOI: 10.1364/AO.44.005859. URL: <https://opg.optica.org/ao/abstract.cfm?URI=ao-44-28-5859>.
- [21] C. Ruilier et al. “System assessment study of the ESA Darwin Mission: concepts trade-off and first iteration design on novel Emma arrangement”. In: *Proc.SPIE*. Vol. 6693, 66930T. DOI: 10.1117/12.733679. URL: <https://doi.org/10.1117/12.733679>.
- [22] M Ollivier et al. “PEGASE, an infrared interferometer to study stellar environments and low mass companions around nearby stars”. In: *Experimental Astronomy* 23 (2009), pp. 403–434.
- [23] W. C. Danchi et al. “The Fourier-Kelvin Stellar Interferometer (FKSI): a review, progress report, and update”. In: *Proc.SPIE*. Vol. 7013, 70132Q. DOI: 10.1117/12.790649. URL: <https://doi.org/10.1117/12.790649>.
- [24] P. Lay Oliver, Jeganathan Muthu, and Peters Robert. “Adaptive nulling: a new enabling technology for interferometric exoplanet detection”. In: *Proc.SPIE*. Vol. 5170, pp. 103–112. DOI: 10.1117/12.503759. URL: <https://doi.org/10.1117/12.503759>.
- [25] Frantz Martinache and Michael J. Ireland. “Kernel-nulling for a robust direct interferometric detection of extrasolar planets”. In: *A&A* 619 (2018). URL: <https://doi.org/10.1051/0004-6361/201832847>.

- [26] Jonah T. Hansen, Michael J. Ireland, and Life Collaboration the. “Large Interferometer For Exoplanets (LIFE) IV. Ideal kernel-nulling array architectures for a space-based mid-infrared nulling interferometer”. In: *A&A* 664 (2022). URL: <https://doi.org/10.1051/0004-6361/202243107>.
- [27] Eugene Serabyn. “Nulling interferometry: symmetry requirements and experimental results”. In: *Interferometry in Optical Astronomy*. Ed. by Pierre J. Lena and Andreas Quirrenbach. Vol. 4006. International Society for Optics and Photonics. SPIE, 2000, pp. 328–339. DOI: 10.1117/12.390223. URL: <https://doi.org/10.1117/12.390223>.
- [28] Nicholas M. Elias II, David W. Draper, and M. Charles Noecker. “Polarization and planet-detecting nulling interferometers”. In: *Instruments, Methods, and Missions for Astrobiology VIII*. Ed. by Richard B. Hoover, Gilbert V. Levin, and Alexei Y. Rozanov. Vol. 5555. International Society for Optics and Photonics. SPIE, 2004, pp. 215–230. DOI: 10.1117/12.564219. URL: <https://doi.org/10.1117/12.564219>.
- [29] Seunghyuk Chang and Aluizio Prata. “Geometrical theory of aberrations near the axis in classical off-axis reflecting telescopes”. In: *Journal of the Optical Society of America A* 22.11 (2005), pp. 2454–2464. DOI: 10.1364/JOSAA.22.002454. URL: <https://opg.optica.org/josaa/abstract.cfm?URI=josaa-22-11-2454>.
- [30] Seunghyuk Chang et al. “Linear astigmatism of confocal off-axis reflective imaging systems and its elimination”. In: *Applied Optics* 45.3 (2006), pp. 484–488. DOI: 10.1364/AO.45.000484. URL: <https://opg.optica.org/ao/abstract.cfm?URI=ao-45-3-484>.
- [31] A. A. Michelson. “Light Waves in Astronomy”. In: *Science* 57.1486 (1923), pp. 703–707. DOI: 10.1126/science.57.1486.703. eprint: <https://www.science.org/doi/pdf/10.1126/science.57.1486.703>. URL: <https://www.science.org/doi/abs/10.1126/science.57.1486.703>.
- [32] Andreas Glindemann. “VLTI tutorial - Introduction to Stellar Interferometry”. Published on website of European Southern Observatory. URL: https://www.eso.org/sci/facilities/paranal/telescopes/vlti/tuto/tutorial_introduction_to_stellar_interf.pdf.
- [33] A. A. Michelson and F. G. Pease. “Measurement of the Diameter of Alpha-Orionis by the Interferometer”. In: *Proceedings of the National Academy of Sciences* 7.5 (1921), pp. 143–146. DOI: 10.1073/pnas.7.5.143. eprint: <https://www.pnas.org/doi/pdf/10.1073/pnas.7.5.143>. URL: <https://www.pnas.org/doi/abs/10.1073/pnas.7.5.143>.
- [34] C. S. Cockell et al. “Darwin—A Mission to Detect and Search for Life on Extrasolar Planets”. In: *Astrobiology* 9.1 (2009). doi: 10.1089/ast.2007.0227, pp. 1–22. ISSN: 1531-1074. DOI: 10.1089/ast.2007.0227. URL: <https://doi.org/10.1089/ast.2007.0227>.
- [35] Dandumont Colin et al. “Exoplanet detection yield of a space-based Bracewell interferometer from small to medium satellites”. In: *Journal of Astronomical Telescopes, Instruments, and Systems* 6.3 (2020), p. 035004. DOI: 10.1117/1.JATIS.6.3.035004. URL: <https://doi.org/10.1117/1.JATIS.6.3.035004><https://www.spiedigitallibrary.org/journals/Journal-of-Astronomical-Telescopes-Instruments-and-Systems/volume-6/issue-3/035004/Exoplanet-detection-yield-of-a-space-based-Bracewell-interferometer-from/10.1117/1.JATIS.6.3.035004.pdf>.
- [36] R. Norbruis. “Detecting exoplanets with a smallsat-based multi-aperture nulling interferometry telescope”. MA thesis. University of Technology Delft, 2023.
- [37] Wallner Oswald et al. “DARWIN mission and configuration trade-off”. In: *Proc.SPIE*. Vol. 6268, p. 626827. DOI: 10.1117/12.671658. URL: <https://doi.org/10.1117/12.671658>.

- [38] L. L. A. Vosteen et al. “Wavefront sensor for the ESA-GAIA mission”. In: *Proc.SPIE*. Vol. 7439, p. 743914. DOI: 10.1117/12.825240. URL: <https://doi.org/10.1117/12.825240>.
- [39] Roberto Ragazzoni. “Pupil plane wavefront sensing with an oscillating prism”. In: *Journal of modern optics* 43.2 (1996), pp. 289–293. ISSN: 0950-0340.
- [40] Bo Xin et al. “Curvature wavefront sensing for the large synoptic survey telescope”. In: *Applied Optics* 54.30 (2015), pp. 9045–9054. DOI: 10.1364/AO.54.009045. URL: <https://opg.optica.org/ao/abstract.cfm?URI=ao-54-30-9045>.
- [41] Hinz Phil et al. “Commissioning the LBTI for use as a nulling interferometer and coherent imager”. In: *Proc.SPIE*. Vol. 9146, 91460T. DOI: 10.1117/12.2057340. URL: <https://doi.org/10.1117/12.2057340%20https://www.spiedigitallibrary.org/conference-proceedings-of-spie/9146/1/Commissioning-the-LBTI-for-use-as-a-nulling-interferometer-and/10.1117/12.2057340.short>.
- [42] A. Schuller Peter et al. “The NULLTIMATE test bench: achromatic phase shifters for nulling interferometry”. In: *Proc.SPIE*. Vol. 7734. Includes temperature dependent refractive index equations, 77342E. DOI: 10.1117/12.857136. URL: <https://doi.org/10.1117/12.857136>.
- [43] Yves Rabbia et al. “Achromatic Phase Shifters: The ”Mirror” Approaches”. In: *GENIE-DARWIN Workshop-Hunting for Planets*. Vol. 522. 2003.
- [44] Dimitri Mawet et al. “Fresnel rhombs as achromatic phase shifters for infrared nulling interferometry”. In: *Optics Express* 15.20 (2007), pp. 12850–12865. DOI: 10.1364/OE.15.012850. URL: <https://opg.optica.org/oe/abstract.cfm?URI=oe-15-20-12850>.
- [45] C. Hanot et al. “Fresnel rhombs as achromatic phase shifters for infrared nulling interferometry: first experimental results”. In: *Proc.SPIE*. Vol. 6693, p. 66931L. DOI: 10.1117/12.758405. URL: <https://doi.org/10.1117/12.758405>.
- [46] Robert O. Gappinger et al. “Experimental evaluation of achromatic phase shifters for mid-infrared starlight suppression”. In: *Applied Optics* 48.5 (2009), pp. 868–880. DOI: 10.1364/AO.48.000868. URL: <https://opg.optica.org/ao/abstract.cfm?URI=ao-48-5-868>.
- [47] Felix Dannert et al. “Large Interferometer For Exoplanets (LIFE): II. Signal simulation, signal extraction, and fundamental exoplanet parameters”. In: (2022).
- [48] Michael E. Ressler et al. “Performance of the JWST/MIRI Si:As detectors”. In: *High Energy, Optical, and Infrared Detectors for Astronomy III*. Ed. by David A. Dorn and Andrew D. Holland. Vol. 7021. International Society for Optics and Photonics. SPIE, 2008, 70210O. DOI: 10.1117/12.789606. URL: <https://doi.org/10.1117/12.789606>.
- [49] G. H. Rieke et al. “The Mid-Infrared Instrument for the James Webb Space Telescope, VII: The MIRI Detectors”. In: *Publications of the Astronomical Society of the Pacific* 127.953 (July 2015), p. 665. DOI: 10.1086/682257. URL: <https://dx.doi.org/10.1086/682257>.
- [50] Pierre Bely. *The design and construction of large optical telescopes*. Springer, 2003. ISBN: 0387955127.
- [51] Ray N Wilson. *Reflecting Telescope Optics: 240 Fig., 39 Tab.(Astronomy and Astrophysics Library)*. Springer, 1999.
- [52] Ernst-Georg Neumann. *Single-mode fibers: fundamentals*. Vol. 57. Springer, 2013.

- [53] Art Hocine Farouk. “Coupling of a single-mode fiber with a deformable mirror in a space-based nulling interferometer”. Thesis. 2020.
- [54] Bertrand Mennesson, Marc Ollivier, and Cyril Ruilier. “Use of single-mode waveguides to correct the optical defects of a nulling interferometer”. In: *Journal of the Optical Society of America A* 19.3 (2002), pp. 596–602. DOI: 10.1364/JOSAA.19.000596. URL: <https://opg.optica.org/josaa/abstract.cfm?URI=josaa-19-3-596>.
- [55] Antonio Budano, Francesco Flora, and Luca Mezi. “Analytical design method for a modified Schwarzschild optics”. In: *Applied Optics* 45.18 (2006), pp. 4254–4262. DOI: 10.1364/AO.45.004254. URL: <https://opg.optica.org/ao/abstract.cfm?URI=ao-45-18-4254>.
- [56] M. Dolores Mouriz et al. “Schwarzschild spectrometer”. In: *Applied Optics* 50.16 (2011), pp. 2418–2424. DOI: 10.1364/AO.50.002418. URL: <https://opg.optica.org/ao/abstract.cfm?URI=ao-50-16-2418>.
- [57] Shannon R Sankar and Jeffrey C Livas. “Optical telescope design for a space-based gravitational-wave mission”. In: *Space Telescopes and Instrumentation 2014: Optical, Infrared, and Millimeter Wave*. Vol. 9143. SPIE, pp. 283–289.
- [58] M. Mark Colavita, Peter L. Wizinowich, and Rachel L. Akeson. “Keck Interferometer status and plans”. In: *New Frontiers in Stellar Interferometry*. Ed. by Wesley A. Traub. Vol. 5491. International Society for Optics and Photonics. SPIE, 2004, pp. 454–463. DOI: 10.1117/12.552149. URL: <https://doi.org/10.1117/12.552149>.
- [59] S. L. Crawford et al. “Final laboratory integration and test of the Keck Interferometer nuller”. In: *Proc.SPIE*. Vol. 5905, 59050U. DOI: 10.1117/12.619131. URL: <https://doi.org/10.1117/12.619131>.
- [60] Foresto Vincent Coude du et al. “FLUOR fibered instrument at the IOTA interferometer”. In: *Proc.SPIE*. Vol. 3350, pp. 856–863. DOI: 10.1117/12.317153. URL: <https://doi.org/10.1117/12.317153>.
- [61] Foresto Vincent Coude du et al. “FLUOR fibered beam combiner at the CHARA array”. In: *Proc.SPIE*. Vol. 4838, pp. 280–285. DOI: 10.1117/12.459942. URL: <https://doi.org/10.1117/12.459942>.
- [62] Marc-Antoine Martinod et al. *Achromatic nulling interferometry and fringe tracking with 3D-photonics tricouplers with GLINT*. Vol. 12183. SPIE Astronomical Telescopes + Instrumentation. SPIE, 2022. URL: <https://doi.org/10.1117/12.2627978>.
- [63] Peter Marley Chingaïpe, Frantz Martinache, and Nick Cvetojevic. “4-input photonic kernel-nulling for the VLTI”. In: *Optical and Infrared Interferometry and Imaging VIII*. Vol. 12183. SPIE, pp. 448–457.
- [64] P Yeh. “Optical Waves in Layered Media”. In: (2006).
- [65] Eugene Serabyn and Marc M Colavita. “Fully symmetric nulling beam combiners”. In: *Applied Optics* 40.10 (2001), pp. 1668–1671. ISSN: 2155-3165.
- [66] Ir. A. Kamp. *Space Instrumentation Engineering*. University of Technology Delft, 2007.
- [67] E. Serabyn et al. “THE KECK INTERFEROMETER NULLER”. In: *The Astrophysical Journal* 748.1 (2012), p. 55. ISSN: 0004-637X. DOI: 10.1088/0004-637X/748/1/55. URL: <https://dx.doi.org/10.1088/0004-637X/748/1/55>.
- [68] T. Tupper Hyde et al. “Requirements formulation and dynamic jitter analysis on Fourier-Kelvin stellar interferometer”. In: *Proc.SPIE*. Vol. 5497. Contains contributions to nulling budget, pp. 553–564. DOI: 10.1117/12.552186. URL: <https://doi.org/10.1117/12.552186>.

-
- [69] Jonathan P Gardner et al. “The james webb space telescope”. In: *Space Science Reviews* 123 (2006), pp. 485–606.
- [70] Jens Kammerer and Sascha P. Quanz. “Simulating the exoplanet yield of a space-based mid-infrared interferometer based on Kepler statistics”. In: *A&A* 609 (2018), A4. URL: <https://doi.org/10.1051/0004-6361/201731254>.
- [71] R. A. Buchroeder and A. S. Leonard. “Wide-Field Tilted-Component Telescope: a Leonard Extended Yolo All-Reflecting System”. In: *Appl. Opt.* 11.7 (July 1972), pp. 1649–1651. DOI: 10.1364/AO.11.001649. URL: <https://opg.optica.org/ao/abstract.cfm?URI=ao-11-7-1649>.



Appendix A: Zernike polynomials

In this appendix, the first 10 zernike polynomials are worked out.

$$R_n^m(\rho) = \sum_{k=0}^{\frac{n-m}{2}} \frac{(-1)^k (n-k)!}{k! \left(\frac{n+m}{2} - k\right)! \left(\frac{n-m}{2} - k\right)!} \rho^{n-2k} \quad (\text{A.1})$$

$$R_0^0(\rho) = \frac{(-1)^0 \cdot (0-0)!}{0! \cdot (0-0)! \cdot (0-0)!} \rho^{(0-0)} = 1 \quad (\text{A.2})$$

$$R_1^1(\rho) = \frac{(-1)^0 \cdot (1-0)!}{0! \cdot (1-0)! \cdot (0-0)!} \rho^{(1-0)} = \rho \quad (\text{A.3})$$

$$R_2^0(\rho) = \sum_{k=0}^1 \frac{(-1)^k (2-k)!}{k! (1-k)! (1-k)!} \rho^{2-2k} = \frac{(-1) \cdot 2!}{0! \cdot 1! \cdot 1!} \rho^{2-0} + \frac{(-1) \cdot 1!}{1! \cdot 0! \cdot 0!} \rho^0 = \rho^2 - 1 \quad (\text{A.4})$$

$$R_2^2(\rho) = \frac{(-1)^0 \cdot (2-0)!}{0! \cdot (2-0)! \cdot 0!} \rho^{2-0} = \rho^2 \quad (\text{A.5})$$

$$R_3^1(\rho) = \sum_{k=0}^1 \frac{(-1)^k (3-k)!}{k! (2-k)! (1-k)!} \rho^{3-2k} = \frac{1 \cdot 3!}{0! \cdot 2! \cdot 1!} \rho^3 + \frac{(-1) \cdot 2!}{1! \cdot 1! \cdot 0!} \rho^1 = 3\rho^3 - 2\rho \quad (\text{A.6})$$

$$R_3^3(\rho) = \frac{(-1)^0 \cdot (3-0)!}{0! \cdot (3-0)! \cdot 0!} \rho^3 = \rho^3 \quad (\text{A.7})$$

B

Appendix B: Maxwell and the wave equation

The Maxwell equations are coupled first order, partial differential equations. With a clever manipulation, they can be uncoupled, at the cost of becoming second order. The Maxwell equations in the absence of any charge or current read as follows:

$$\nabla \cdot \mathbf{E} = 0, \tag{B.1}$$

$$\nabla \cdot \mathbf{B} = 0, \tag{B.2}$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \tag{B.3}$$

$$\nabla \times \mathbf{B} = \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t}. \tag{B.4}$$

The trick is to apply the curl to Equation B.3 and Equation B.4. Since both manipulations are identical, only the E field will be worked out here. To start, the triple product vector identity is used on the left side:

$$\nabla \times (\nabla \times \mathbf{E}) = \nabla(\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E} = \nabla \times \left(-\frac{\partial \mathbf{B}}{\partial t}\right). \tag{B.5}$$

To simplify, Equation B.1 can be used:

$$\nabla^2 \mathbf{E} = -\frac{\partial}{\partial t}(\nabla \times \mathbf{B}). \tag{B.6}$$

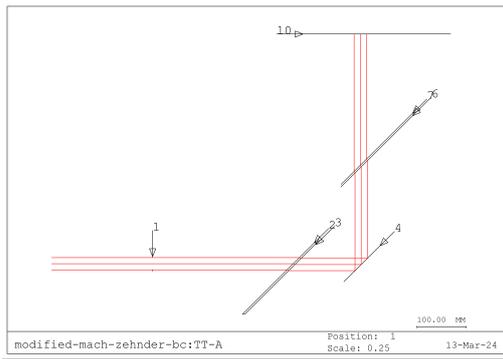
In the above equation, the order of ∇ and $\frac{\partial}{\partial t}$ has been swapped. This is valid, as both of them are linear operators, and the second partial derivative of \mathbf{B} is continuous. Then, finally Equation B.4 can be used to uncouple the differential equation:

$$\nabla^2 \mathbf{E} = \mu_0 \epsilon_0 \frac{\partial^2 \mathbf{E}}{\partial t^2}. \tag{B.7}$$

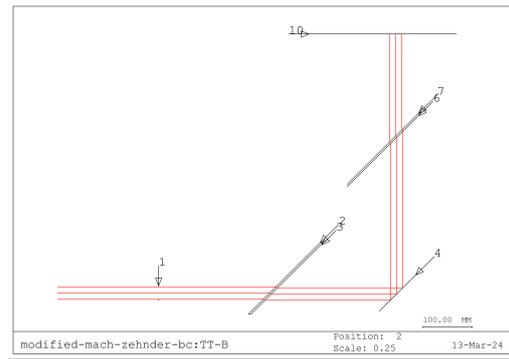
C

Appendix C: Modified Mach Zehnder interferometer

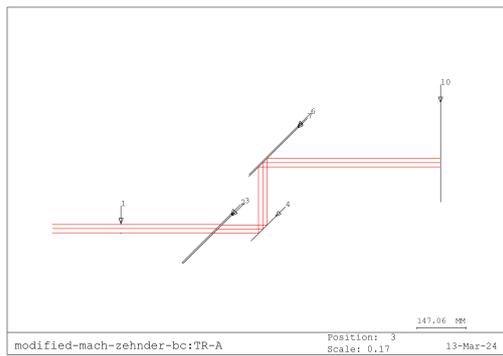
In Figure C.1 schematics are given for each zoom making up the MMZ beam combiner. In Figure C.2 and Figure C.3, the phase and intensity fields are given, respectively. Zooms 1, 2, 7 and 8 are not used, as those combine into the destructive outputs, which are unbalanced, as was explained in section 4.7. Of the two constructive outputs, only one is used to derive the transmission map.



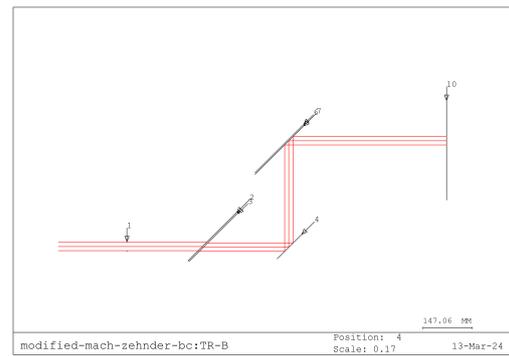
(a) Zoom 1



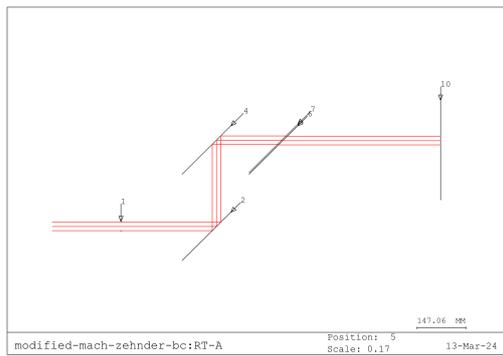
(b) Zoom 2



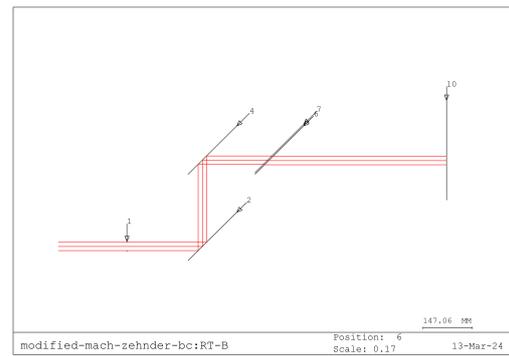
(c) Zoom 3



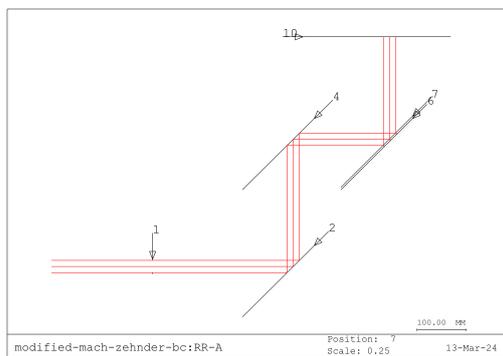
(d) Zoom 4



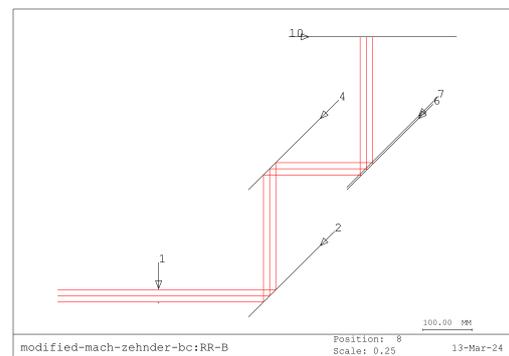
(e) Zoom 5



(f) Zoom 6



(g) Zoom 7



(h) Zoom 8

Figure C.1: Schematic of all zooms used to make a 2 input, 4 output MMZ.

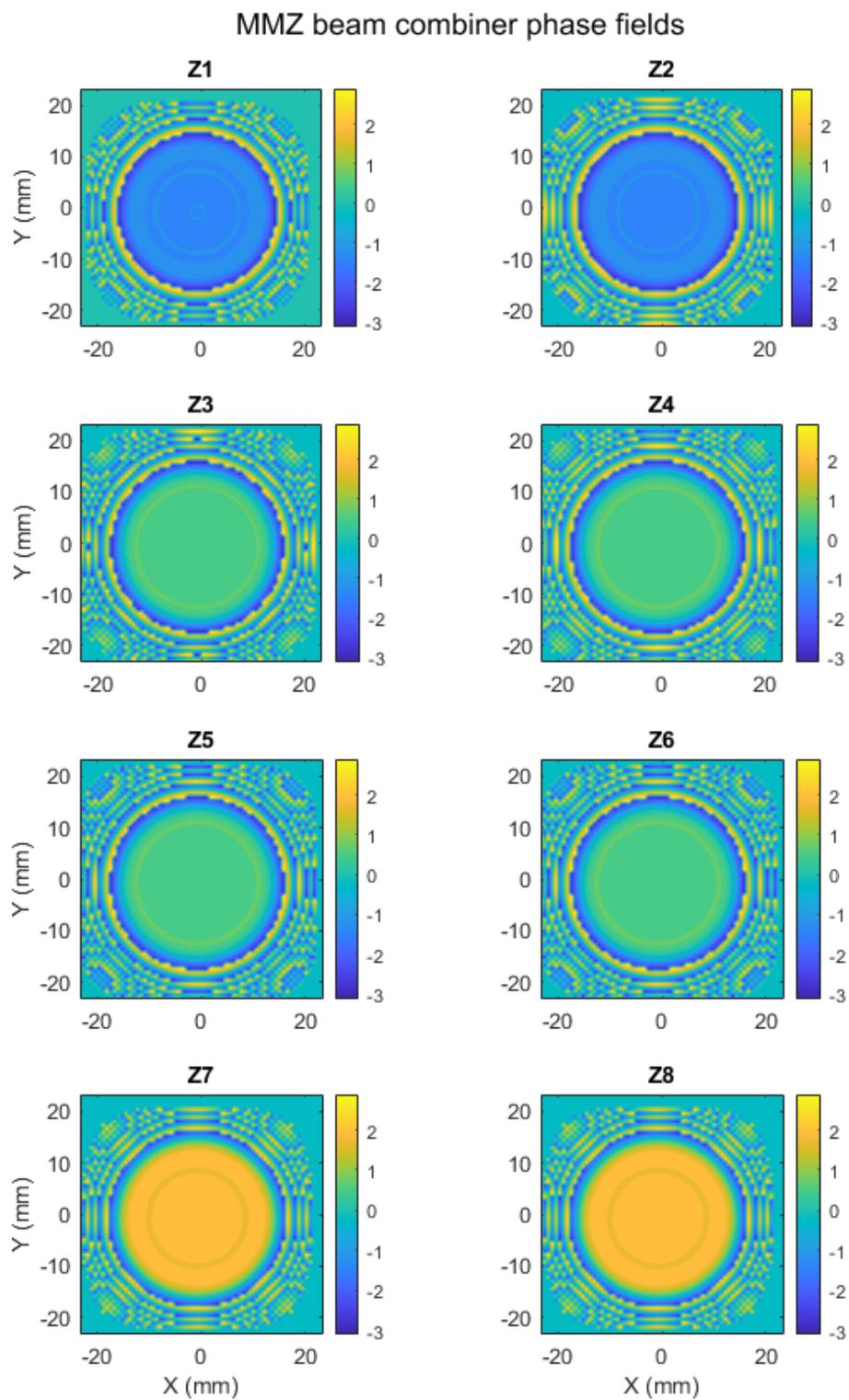


Figure C.2: Phase for each beam interacting with the MMZ beam combiner

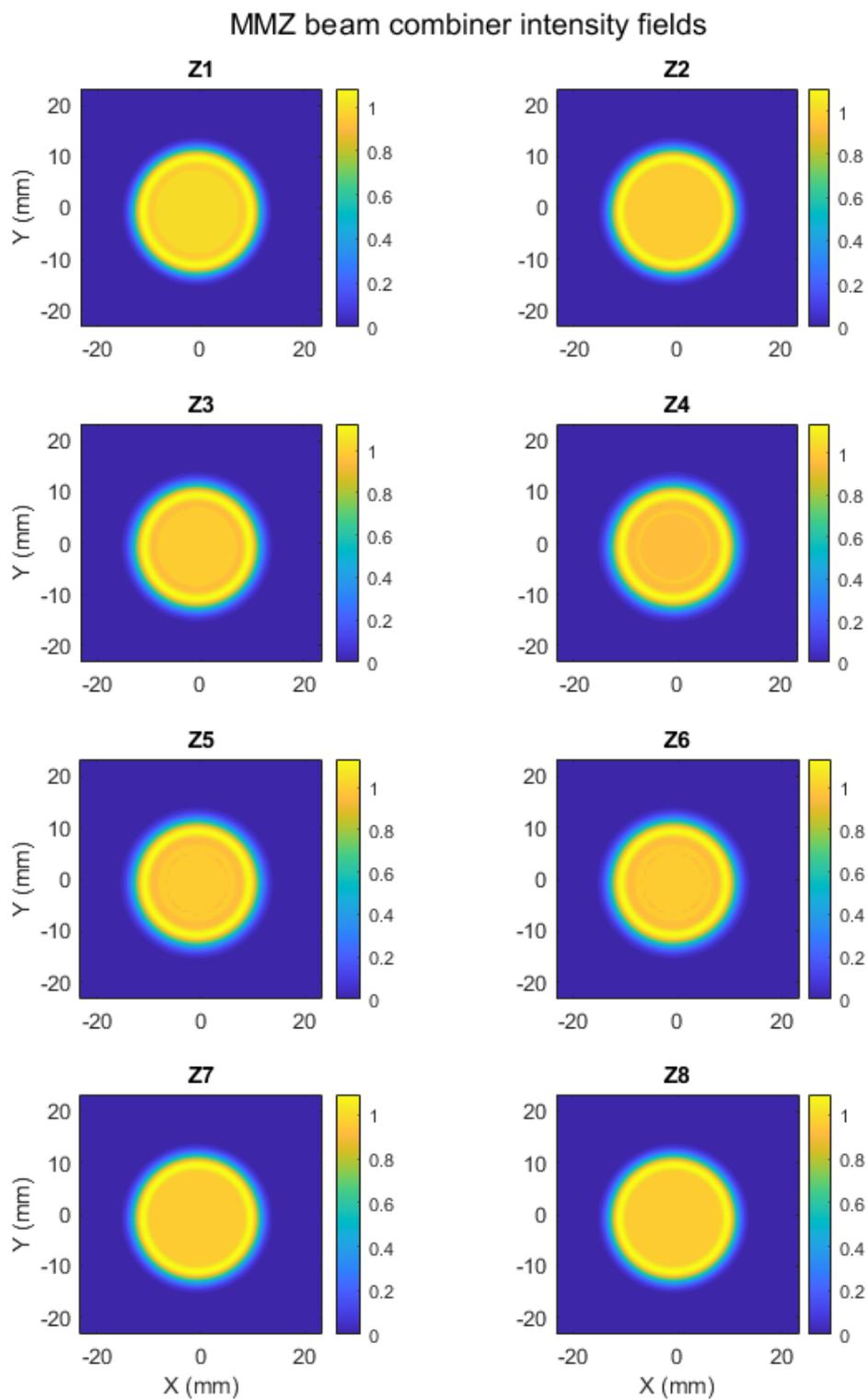


Figure C.3: Intensity for each beam interacting with the MMZ beam combiner

D

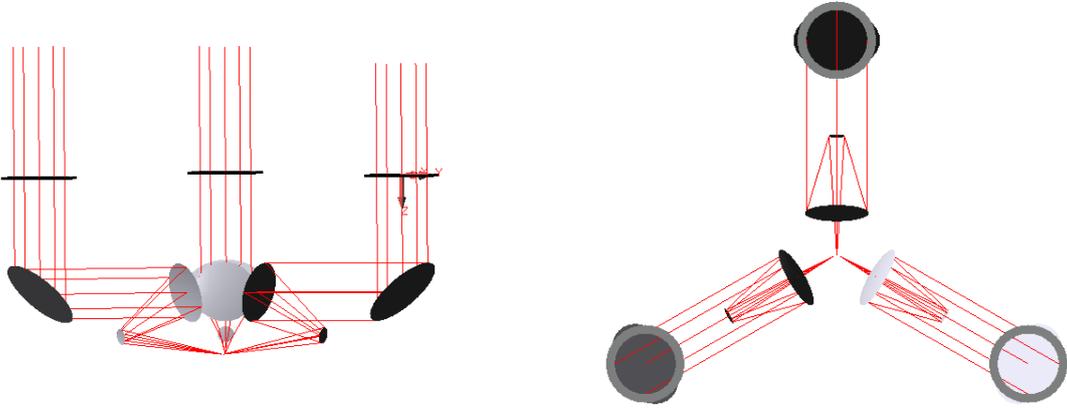
Appendix D: Unfinished or unverified results

D.1. Three aperture design

The initial scope of the project included testing a two, three and 4 aperture nuller and comparing the performances. The scope was made smaller once it became clear how challenging it is to model a three aperture nuller. The telescope design is shown in Figure D.1. The difficulty in making this configuration of telescopes can be explained with the fact that Code V deals with rotations intrinsically. This means that first a rotation around the x axis is performed and after that, a rotation around the local y axis. This way, rotations are not commutative. In terms of external rotations (meaning, with respect to a global coordinate system), the order of operations is trivial: first, rotate by 15 or 20 degrees around the x axis then rotate the primary mirror with 120 degrees around the y axis. To calculate the needed intrinsic rotation angles, the Python module SciPy was used, specifically the Rotation module. Modelling a three aperture beam combiner proved to be cumbersome, so the project focus was shifted towards the four aperture nuller, which relies mostly on MMZ beam combiners.

D.2. Monte Carlo simulation

After the parameter ranges have been determined, a Monte Carlo simulation can be set up. The flow of the simulation is shown in Figure D.2. First, all current values are acquired from the Code V model with the COM API. Second, also using the COM API, each surface parameter in the parameter space is perturbed by a (uniformly distributed) random value. Third, a plane wave is propagated through the system using BSP. From that result, the null depth is calculated, as discussed in chapter 3, and finally the results are aggregated by null depth into an accumulative distribution. With this distribution, it can be determined in how many percentage of the cases a target null depth was reached. A preliminary Monte Carlo simulation was ran with 1000 trials, the result of which is shown in Figure D.3



(a) Side view

(b) Top view

Figure D.1: Side and top view of a three aperture telescope design.

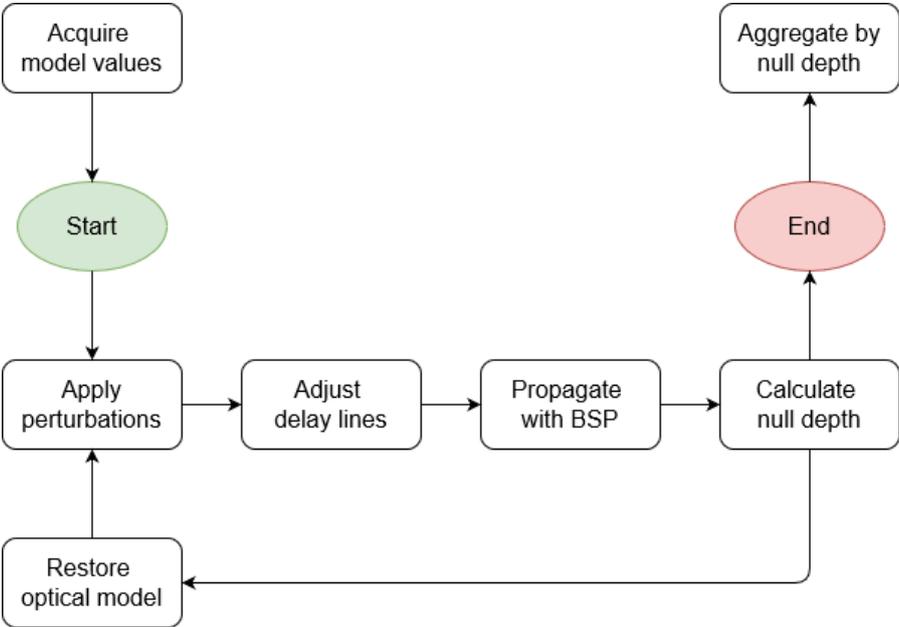


Figure D.2: Flow chart of the Monte Carlo simulation process.

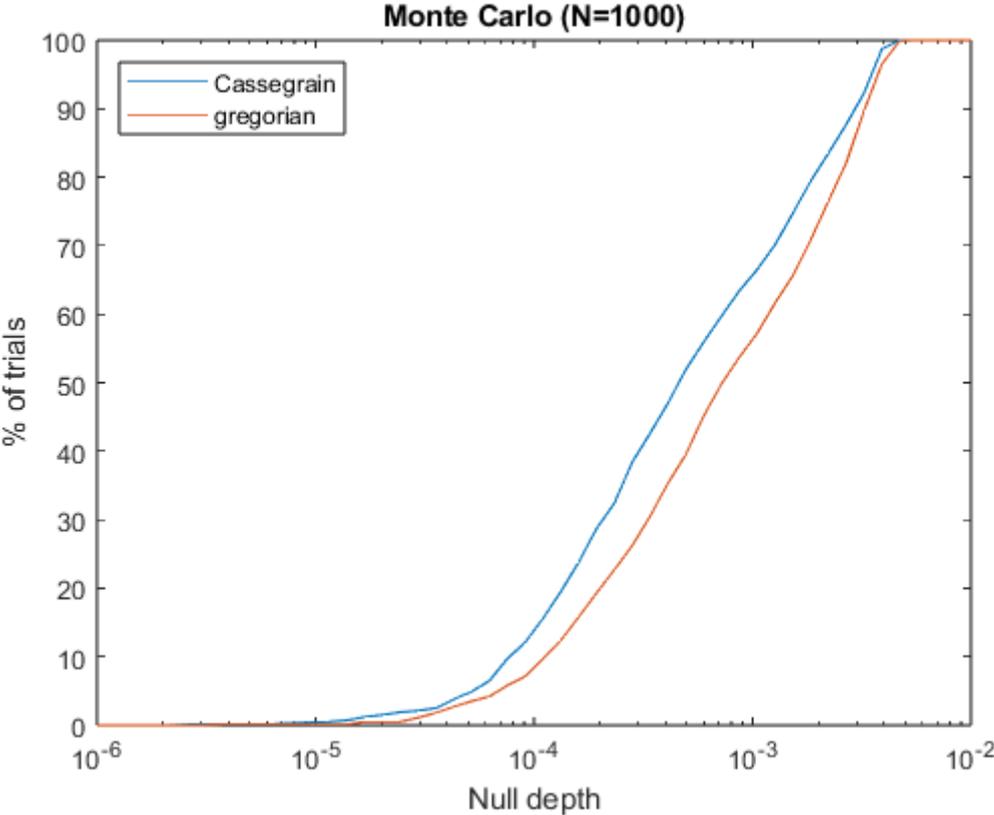


Figure D.3: Comparison of the off-axis Cassegrain and Gregorian telescope design.

E

Appendix E: Null depth calculation with finite sampling

When perturbing a surface of the interferometer, the transmission map can shift, such that 0 OPD does not correspond to the minimum null depth anymore. This is indicated on the left of Figure E.1, where the solid black line indicates the 0 OPD line and the dashed line is the line where the transmission map is minimum. Since the transmission map is determined using a finite sampling at regular intervals, this can result in a higher minimum null depth than should be the case, as is shown on the right of Figure E.1. The solution to this is to resample the transmission map around the minimum.

From Figure E.2, it can be seen that due to the finite sampling and shifting of the transmission map, the null depth changes erratically, up to several orders of magnitude per step of ΔL . To mitigate this, the transmission map is resampled with a much higher sampling rate around the minimum, after which a more accurate null depth is found.

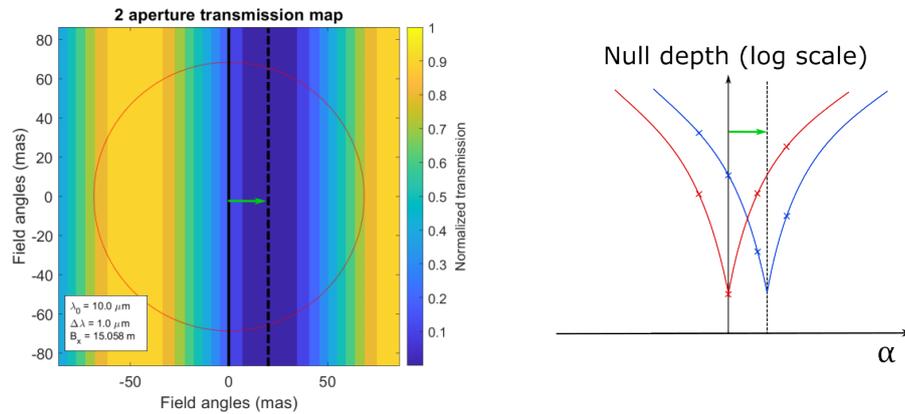


Figure E.1: Left: Transmission map of a 2 aperture Bracewell interferometer, where the minimum null depth is shifted to the right of the 0 field angle line, due to a translation in the primary mirror. Right: Horizontal slice of the transmission map, where the red line indicates the null depth without any perturbations applied to the optical surfaces. The blue line corresponds to the horizontal slice of the shifted transmission map.

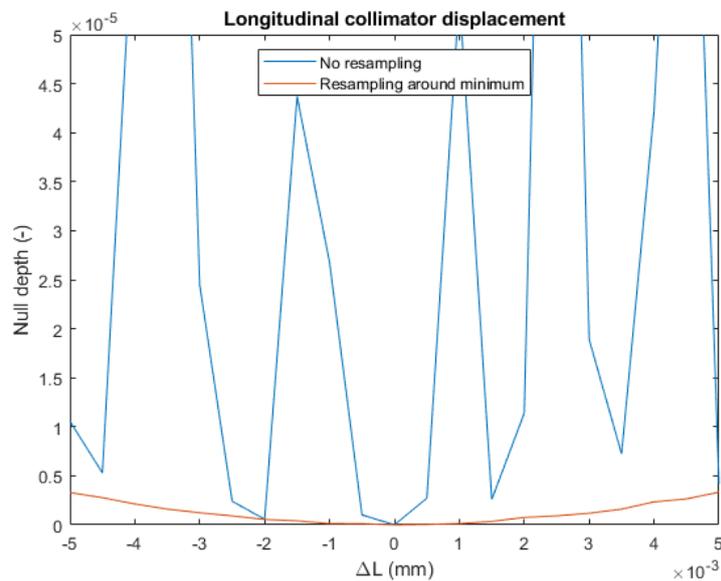


Figure E.2: The blue line shows the calculated null depth when a relatively low sampling rate is used for the transmission map, as function of the axial displacement. When the transmission map is resampled around the minimum, which is shown in red, the null depth shows a less erratic trend.

F

Appendix F: Field of View determination

When the largest field angle is reduced to (18",18") (a factor of 10 compared to (3', 3')), the spot diagram fits inside the Airy disk, as shown in Figure F.1a. However, the diffraction based MTF, plotted in Figure F.1b shows that the MTF curves still do not overlap. Both the requirement for the spot diagram to be within the Airy disk and that the MTF curves should be 'close to the diffraction limited MTF curves' are a qualitative requirement for diffraction limited performance. Alternatively, the requirement on the P-V WFE can be considered, which should not exceed 0.25 waves, as was mentioned in subsection 5.5.4. A simulated exit pupil map for the (18", 18") is shown in Figure F.2, which exceeds the P-V WFE of 0.25 waves.

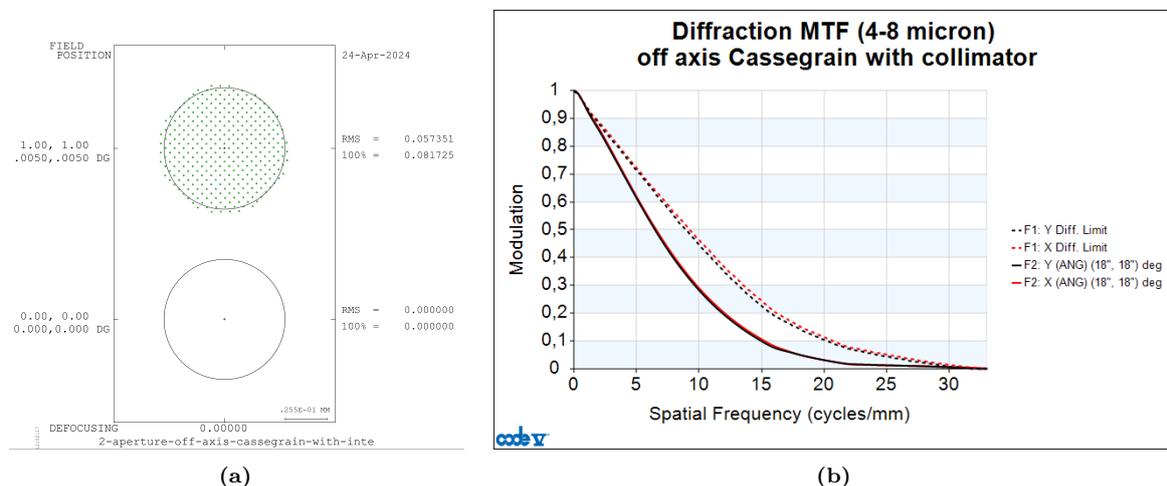


Figure F.1: (a) Spot diagrams for the on-axis and off-axis field, evaluated at 4 μm . (b) Diffraction based MTF of the considered telescope design, evaluated from 4 to 8 μm and a field angle of (18", 18").

It should be noted here that Code V gives the following warning, for both the pupil map option (PMA), as well as the diffraction based MTF option (MTF):

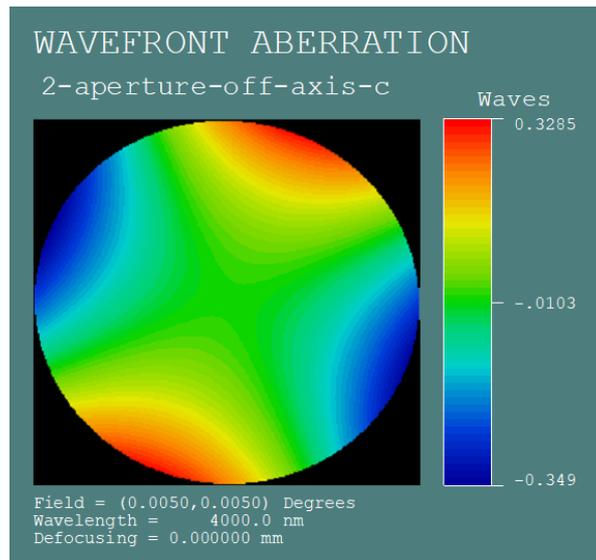


Figure F.2: Exit pupil map of the off-axis Cassegrain simulated at a field angle of (18", 18") and wavelength of $4\ \mu\text{m}$, showing oblique astigmatism. Given that the P-V is larger than 0.25 waves, it is not considered to be diffraction limited.

Warning: This system is not isoplanatic at field(s) 1, 2. Any PMA results are of questionable value for assessing the image quality, except in those cases where the user is interested only in the image of a point-like object. The image quality for the unacceptable fields varies significantly over a lateral region of 12 Airy disk diameters.

This shows how non-trivial it is to properly assess the optical performance of an off-axis/non-isoplanatic system. In part for this reason, BSP will be used for evaluating the interferometer performance under perturbations in section 6.5. As for diffraction limited performance, the field of view can be reduced further to ensure ideal performance, as the new field of view (13x13 arcseconds) would still be much larger than the outer working angle of about 1 arcsecond.

G

Appendix G: Classical telescope design null depth comparison

In this appendix, the radar plots comparing the stability of the three telescope designs are given. The plots are shown on the next page onwards. Perturbations applied to the primary mirror give similar results across the three designs. For perturbations applied to the collimator, the results are more distinct. Note the difference in axes limits on all the plots.

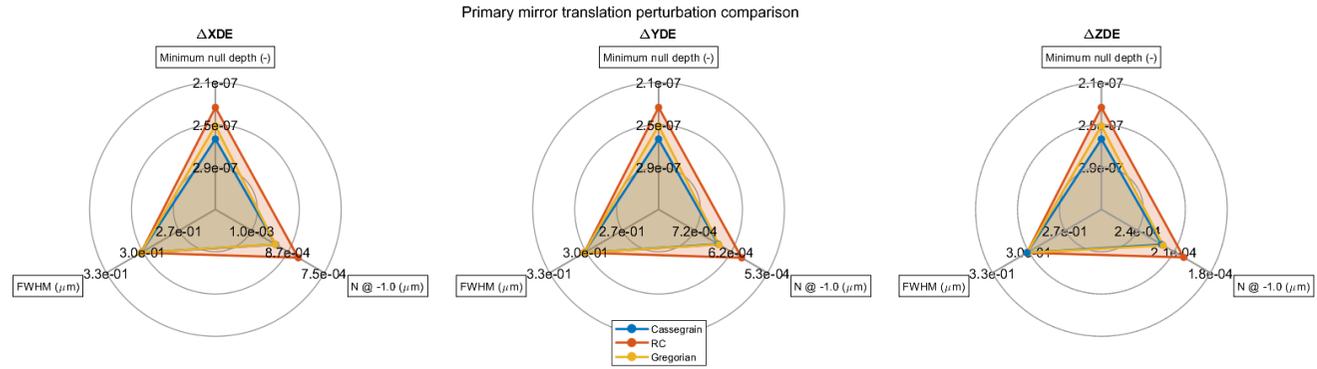


Figure G.1: Translation perturbation of the primary mirror. The comparison is made between an off-axis Cassegrain, RC and Gregorian design. All null depths calculated with a $10 \mu\text{m}$ central wavelength and bandwidth of $1 \mu\text{m}$.

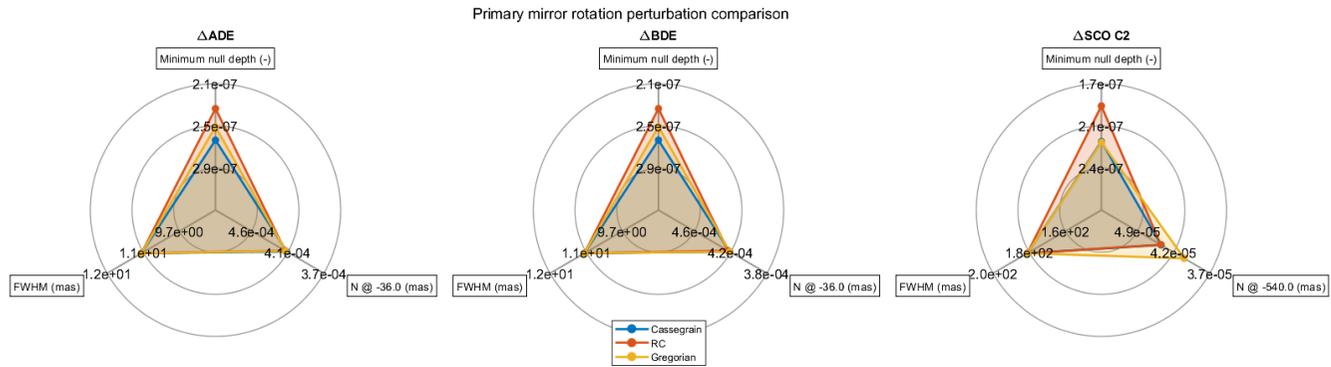


Figure G.2: Rotation perturbation of the primary mirror. The comparison is made between an off-axis Cassegrain, RC and Gregorian design.

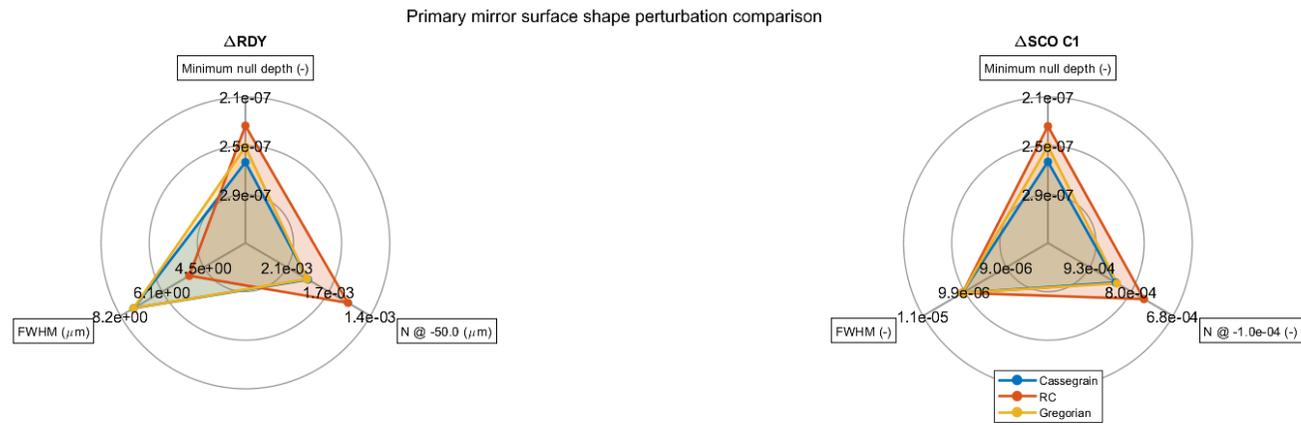


Figure G.3: Surface shape perturbation of the primary mirror. The comparison is made between an off-axis Cassegrain, RC and Gregorian design.

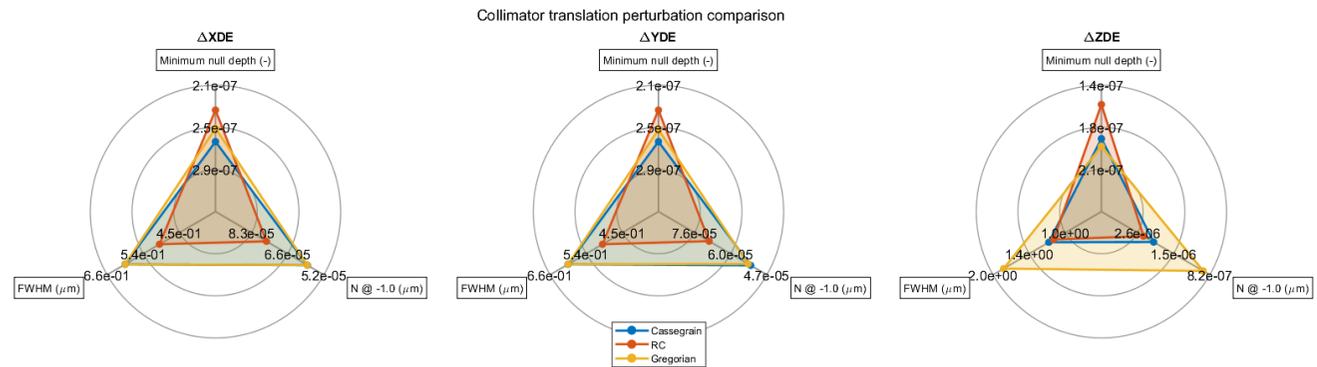


Figure G.4: Translation perturbation of the collimator. The comparison is made between an off-axis Cassegrain, RC and Gregorian design. All null depths calculated with a $10 \mu\text{m}$ central wavelength and bandwidth of $1 \mu\text{m}$.

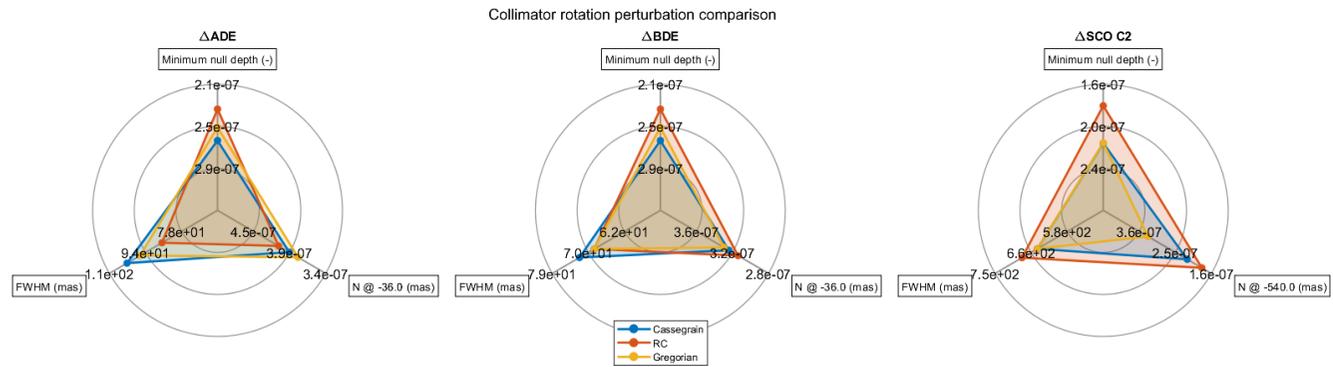


Figure G.5: Rotation perturbation of the collimator. The comparison is made between an off-axis Cassegrain, RC and Gregorian design.

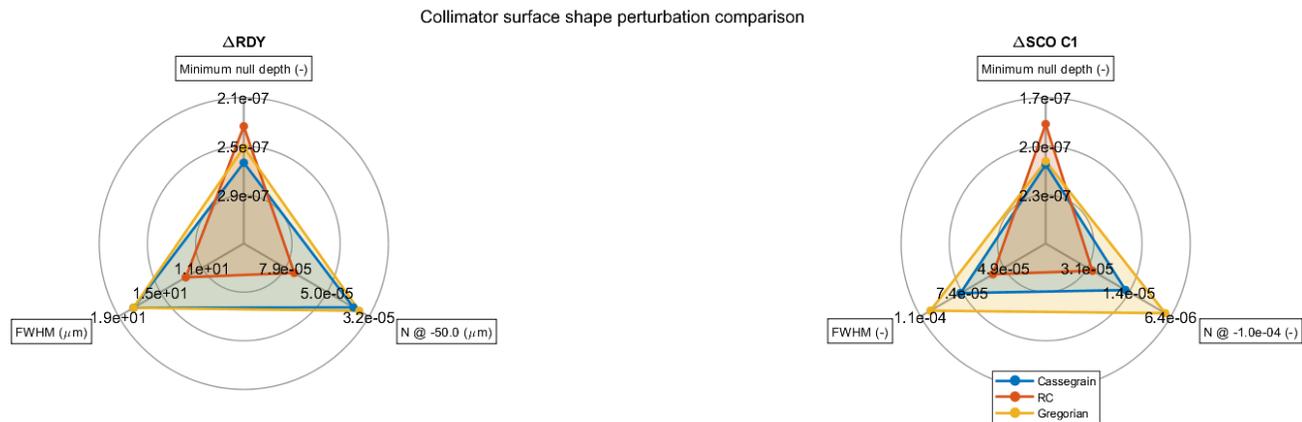


Figure G.6: Surface shape perturbation of the collimator. The comparison is made between an off-axis Cassegrain, RC and Gregorian design.