Design and recording methods for holographic waveguide displays

Zur Erlangung des akademischen Grades eines

DOKTORS DER INGENIEURWISSENSCHAFTEN
(Dr.-Ing.)

von der Fakultät für
Elektrotechnik und Informationstechnik,
des Karlsruher Instituts für Technologie (KIT)

angenommene

DISSERTATION

von

M.Sc. Marvin Gerold Feil
geb. in Laichingen

Tag der mündlichen Prüfung:	16.09.2026
Hauptreferent:	Prof. Dr. rer. nat. Wilhelm Stork
Korreferent:	Prof. Dr. rer. nat. Alois Herkommer

© 2026 Marvin Gerold Feil. All rights reserved.

Parts of this dissertation have been previously published in Optics Express, Applied Optics, and Optics Letters (Optica Publishing Group). The corresponding material is reproduced with appropriate attribution in accordance with the publisher's author reuse policy.

Kurzfassung

Wellenleiterdisplays sind eine wichtige optische Architektur für Displaybrillen, die durch Mehrfachauskopplungen ein großes Sichtfeld und eine große Eyebox bei gleichzeitig kompakter Bauform ermöglichen. Die Integration herkömmlicher planarer Wellenleiterdisplays in Korrektionsgläser mit konvexen Basiskurven führt jedoch zu dickeren und schwereren Brillengläsern. Sphärisch gekrümmte holografische Wellenleiter hingegen ermöglichen dünne und leichte Displaygläser in der Bauform herkömmlicher Korrektionsgläser.

Nach Kenntnis des Autors präsentiert diese Arbeit erstmals ein exaktes geometrisches Designverfahren für sphärische holografische Wellenleiterdisplays. Es wird gezeigt, dass die exakten Eigenmoden, auf denen vorherige Aberrationskorrekturen planarer und zylindrischer Wellenleiterdisplays beruhen, in sphärischen Wellenleitern nicht existieren. Stattdessen werden mithilfe exakter Sprungabbildungen geschlossene geometrische Modenfamilien für planare, zylindrische und sphärische Wellenleiter konstruiert, die im planaren und zylindrischen Grenzfall mit Eigenmoden übereinstimmen. Geschlossene Moden vermeiden Mehrdeutigkeiten in den Strahlfeldern überlappender Generationen von im Wellenleiter geführter Strahlbündel und überführen die Einkoppel- und Expandergitter in analytische Translationen im Modenraum. Verbleibende geometrische Aberrationen entstehen ausschließlich bei der Bildauskopplung. Für sphärische Krümmungen unter $3\,\mathrm{D}$ werden über ein diagonales Sichtfeld von $22°$ für eine $8\,\mathrm{mm}$-Augenpupille monochromatische Winkelfehler von weniger als einer Bogenminute erreicht, vergleichbar mit dem Auflösungsvermögen des menschlichen Auges.

Anschließend befasst sich die Arbeit mit der Belichtung holografischer Wellenleiter. Herkömmliche Verfahren zur Aufzeichnung holografischer Kopplergitter beruhen auf immersionskontaktierter Belichtung, was die Herstellung erschwert. Änderungen der Aufzeichnungswellenlänge und der Aufzeichnungsgeometrie werden als mögliche Ansätze zur Herstellung von Wellenleiterdisplays ohne Immersionskontakt untersucht. Für planare und schwach gekrümmte holografische Gitter werden analytische Grenzen für immersionsfreie Belichtbarkeit hergeleitet. Die Ergebnisse zeigen, dass für typische Photopolymerwellenleiter die betrachteten reflektiven Einkoppler- und Auskopplergitter für die roten, grünen und blauen Farbkanäle ohne Immersionskontakt aufgezeichnet werden können. Praktisch relevante Expandergitter für planare und sphärische Photopolymerwellenleiter erfordern dagegen zwingend eine Immersionsbelichtung.

Die automatisierte Immersionsbelichtung von Wellenleitern erfordert sowohl einen großen Bereich von Strahlwinkeln als auch Strahlaperturen im Zentimeterbereich, um Einkoppel-, Expander- und Auskoppelgitter aufzeichnen zu können. Nach Kenntnis des Autors beschreibt diese Arbeit die erste automatisierte Belichtungsanlage für holografische Wellenleiterdisplays. Das System kombiniert zwei auf Sechsachsrobotern montierte Kollimatoren mit einer eigens entwickelten immersionskontaktierten Prismenkonstruktion. Die Roboterposen werden mithilfe inverser Strahlverfolgung und Roboterkinematik berechnet, wobei Materialschrumpfung und Strahlpolarisation berücksichtigt werden. Die Kalibrierung ergibt einen mittleren absoluten Winkelfehler von weniger als $0.1°$, während interferometrische Messungen über typische Belichtungszeiten eine Phasenstabilität von $\pm\lambda/20$ nachweisen. Die Belichtungsmaschine wird durch die Herstellung eines planaren Wellenleiterdisplays validiert. Die Modulationsübertragungsfunktion des Displays wird bestimmt, und die durch die Pupillenabtastung verursachte Intensitätsmodulation im lasergescannten Bild wird analytisch erklärt.

Diese Ergebnisse belegen die optische Realisierbarkeit sphärischer holografischer Wellenleiterdisplays und identifizieren deren Herstellung als die wesentliche verbleibende Herausforderung. Das entwickelte Designverfahren, die Analyse der

immersionsfreien Belichtbarkeitsgrenzen und die automatisierte Belichtungsanlage für planare Wellenleiter bilden eine Grundlage für die zukünftige Entwicklung sphärischer Wellenleiterdisplays.

Abstract

Waveguide displays are a prominent optical architecture for near-eye displays because exit-pupil expansion enables a large field of view and a large eyebox within a compact form factor. However, embedding conventional planar waveguides into prescription lenses with convex base curves increases lens thickness and weight. Spherically curved holographic waveguides could avoid these penalties, enabling lightweight display glasses with the form factor of conventional eyewear.

To the author's knowledge, this thesis presents the first exact geometric design framework for spherical holographic waveguide displays. It is shown that the exact eigenmodes underlying previous planar and cylindrical waveguide display designs do not exist in spherical waveguides. Instead, closed geometric mode families for planar, cylindrical, and spherical waveguides are constructed using exact jump maps that recover previously known eigenmodes in the planar and cylindrical limit. These modes make overlapping generations of guided beams single-valued and turn incoupler and expander gratings into analytic translations in mode space. Residual geometric aberrations arise only during image extraction at the outcoupler. For spherical curvatures below $3\,\mathrm{D}$, non-sequential ray tracing predicts angular errors below 1 arcmin for monochromatic imaging to an $8\,\mathrm{mm}$ pupil across a 22° diagonal field of view, approaching the visual-acuity limit.

The thesis then addresses the fabrication of holographic waveguides. Conventional recording schemes rely on edge-lit exposure with immersion contact, which complicates both prototyping and fabrication. Shifts in recording wavelength and recording configuration are investigated as potential means of fabricating waveguide displays without immersion contact. Analytical bounds on immersion-free recordability are derived for planar holographic gratings and remain approximately

valid for gratings formed on weakly curved substrates. The results show that the considered red, green, and blue incouplers and outcouplers for typical reflective photopolymer waveguides can be recorded without immersion contact. By contrast, the expander gratings required for practically relevant planar and spherical photopolymer waveguides lie outside the range of immersion-free recordability.

Automated edge-lit waveguide recording requires both a wide range of beam angles and centimeter-scale beam apertures to record the incoupler, expander, and outcoupler gratings. To the author's knowledge, this thesis reports the first automated recording platform for holographic waveguide displays. The system combines two collimators mounted on six-axis robots with a custom edge-lit prism assembly. Exposure poses are computed from the prescribed gratings using inverse ray tracing and robot kinematics, with compensation for material shrinkage and alignment of the recording-beam polarization. Calibration yields a mean absolute angular error below $0.1°$, while interferometric measurements demonstrate relative optical-path stability of order $\pm\lambda/20$ over typical exposure times. The complete workflow is validated through the fabrication of a planar waveguide that transmits readable text. Its modulation transfer function is evaluated, and pupil-sampling-induced intensity modulation in the laser-scanned image is explained analytically.

Together, these results establish the optical feasibility of lightweight, prescription-compatible spherical holographic waveguide displays and identify fabrication as the principal remaining challenge. The developed design framework, recordability analysis, and automated exposure platform for planar waveguides provide a foundation for the future development of spherical waveguide displays.

Contents

Kurzfassung i

Abstract v

1 Introduction 1

2 Fundamentals 9
2.1 Holographic optical elements (HOEs) 9
2.1.1 Local volume grating description 10
2.1.2 Local grating coordinate system 12
2.1.3 Hologram shrinkage 12
2.1.4 Decomposition in recording wavefronts 13
2.1.5 Volume grating diffraction 13
2.1.6 Hologram ray tracing 15
2.2 Planar holographic waveguide displays 16
2.2.1 Planar waveguide eigenmodes 16
2.2.2 Central-mode grating coupler construction 18
2.2.3 Orthogonal waveguides 20
2.2.4 Off-axis modes and achromaticity 21
2.2.5 Achievable display field of view 22
2.3 Waveguide integration in prescription lenses 24
2.3.1 Regular lens geometry 24
2.3.2 Waveguide display prescription correction 26
2.3.3 Lens stack thickness and weight 27

3 Design of curved waveguide displays using closed modes . . 31
3.1 Introduction 31

3.2 Propagation in the waveguide . . . 33
3.2.1 Ray jump maps . . . 33
3.2.2 Beam propagation and rayfield description . . . 35
3.2.3 Waveguide modes and families . . . 35
3.3 Closed and locally consistent modes . . . 36
3.3.1 Removing multi-valuedness: local consistency . . . 36
3.3.2 Supporting finite field of view: closure under tangential grating redirection . . . 37
3.4 Explicit closed mode families . . . 39
3.4.1 Planar family . . . 39
3.4.2 Cylindrical family . . . 39
3.4.3 Spherical family . . . 40
3.4.4 Free-space and guided modes . . . 40
3.4.5 Comparison with previous eigenbeam concepts . . . 41
3.5 Closed mode waveguide design framework . . . 43
3.5.1 Design problem . . . 43
3.5.2 Solution using closed modes . . . 44
3.5.3 Homogeneity and efficiency optimization . . . 47
3.5.4 Polychromatic image transport . . . 47
3.6 Waveguide design studies . . . 48
3.6.1 Exemplary architecture . . . 48
3.6.2 Outcoupler optimization . . . 50
3.6.3 Non-sequential waveguide ray trace . . . 51
3.6.4 Curvature dependence of outcoupling aberrations . . . 53
3.6.5 Color aberrations for finite spectral bandwidth . . . 55
3.6.6 Exemplary light engine design . . . 56
3.7 Wave optical correspondence and scope of the rayfield model . . . 57
3.7.1 Integrability of closed rayfields . . . 57
3.7.2 Integrability of spherical grating couplers . . . 58
3.8 Conclusion . . . 59

4 Limits of immersion-free waveguide recording . . . 61
4.1 Introduction . . . 62
4.2 Immersion-free recording . . . 62
4.2.1 General holographic resolution limit . . . 63

4.2.2 Immersion-free recording condition . . . 64
4.2.3 Incidence angles for immersion-free recording . . . 64
4.2.4 Monochromatic immersion-free recording limit . . . 67
4.2.5 Polychromatic immersion-free recording limit . . . 69
4.2.6 Applicability to spherical waveguide couplers . . . 72
4.3 Recording polarizations . . . 74
4.3.1 Brewster-polarized recording . . . 74
4.3.2 Quasi-Brewster-polarized recording . . . 76
4.4 Conclusion . . . 77

5 Automated planar waveguide recording . . . 79
5.1 Introduction . . . 80
5.2 Recording system design . . . 82
5.2.1 Recording prisms . . . 82
5.2.2 Ray tracing . . . 83
5.2.3 Parasitic hologram suppression . . . 85
5.3 Robotic holography platform . . . 87
5.3.1 System details . . . 88
5.3.2 Robot pose calculation . . . 89
5.3.3 Inverse robot kinematics . . . 91
5.3.4 Robot path planning . . . 93
5.3.5 Angular absolute calibration . . . 93
5.3.6 Interferometric stability . . . 95
5.4 Waveguide sample fabrication . . . 97
5.5 Sample validation . . . 98
5.5.1 Scope of validation sample . . . 98
5.5.2 Imaging test bench . . . 99
5.5.3 Intensity uniformity . . . 100
5.5.4 Image transfer and resolution assessment . . . 103
5.5.5 Optical throughput . . . 105
5.6 Conclusion . . . 106

6 Conclusion and outlook . . . 107
6.1 Conclusion . . . 107
6.2 Outlook . . . 108

A Appendix **111**
A.1 Derivation of the grating vector decomposition 111
A.2 Volume diffraction derivations 113
A.3 Implementation lens-stack weight analysis 114
A.4 Derivation of waveguide jump maps 117
A.5 Mode closure forces affine structure 122
A.6 Properties of the mode families 124
A.7 Closed modes reduce to previous paraxial eigenbeam results . . . 126
A.8 Refraction out of waveguide 128
A.9 Visualizing immersion-free recording 136
A.10 Recording limit analysis for high-refractive-index holograms . . . 139
A.11 Angular sensitivity of the robot optics 141
A.12 Relation between OPW and Stäubli VAL3 robot arm configurations 142
A.13 Fourier phase extraction for recording-stability measurement . . . 144
A.14 Model for sparse-pupil grid modulation 146
A.15 Modulation transfer function measurement 151

Acknowledgements **153**

List of publications **157**
Journal articles 157
Conference contributions 158
Patents 158
Supervised student theses and internships 159

Bibliography **161**

1 Introduction

Augmented reality (AR) overlays digital information, such as text messages, navigation cues, or contextual annotations, onto a user's view of the physical environment. Integrating this functionality into glasses promises everyday wearability, but imposes stringent optical requirements: the display must be compact, lightweight, transparent, and compatible with natural eye motion and realistic variations in spectacle fit.

The near-eye display (NED) is the optical system that forms the virtual image and superimposes it on the world view. Each virtual image point is represented by a ray bundle that enters the eye pupil at a particular chief-ray angle. The bundle is collimated for an image at infinity and divergent for an image at a finite distance s, as shown in Fig. 1.1. The angular extent of all such bundles defines the display field of view. The displayed image must remain visible as the eye rotates, as pupil position varies between users, and as the glasses slip on the face. The optical system must therefore provide not only a numerical aperture sufficient for the specified field of view (FOV), but also an exit pupil diameter D_{EP} large enough to accommodate a specified volumetric eyebox. Within this volume, the eye pupil may shift relative to the exit pupil while the image remains visible without vignetting, see Fig. 1.1. A larger field of view requires a larger system exit pupil to maintain this eyebox. To first order, the available eyebox diameter D_{EYB} is determined by the eye relief EYR, defined as the distance from the eye pupil to the rear surface of the eyeglass, and the system exit pupil distance EPD, defined as the distance from the NED exit pupil to the rear surface of the eyeglass [1]

$$D_{\mathrm{EYB}} = D_{\mathrm{EP}} - 2\,(\mathrm{EYR} + \mathrm{EPD})\tan(\mathrm{FOV}/2). \tag{1.1}$$

The combined requirement for a large field of view and a large exit pupil can be expressed through the one-dimensional system étendue, the product of exit pupil diameter and field of view [2],

$$G_{1\mathrm{D}} = 2D_{\mathrm{EP}} \sin\left(\mathrm{FOV}/2\right). \tag{1.2}$$

This étendue requirement is one of the central constraints in the design of augmented-reality glasses. In a classical lossless optical system without ray splitting, étendue is conserved. Field of view and eyebox can therefore be traded against one another, but achieving both simultaneously requires large étendue and, consequently, large optical apertures, which increase system volume and weight [1–3]. Large system size is incompatible with the glasses-like form factor required for wearable AR glasses.

In *free-space combiner* systems, the virtual-image ray bundles propagate through free-space optics and are superimposed on the world view by a semi-transparent optical element. This principle dates back to Grubb's 1901 reflective gun sight, where a collimated illuminated mask was combined with the scene by a tilted partially reflective plate [4], see Fig. 1.2(a). In modern AR displays, the mask is replaced by a microdisplay or scanned source. The same architecture reappeared in the 1950s and 1960s in aircraft head-up displays (HUDs) and helmet-mounted

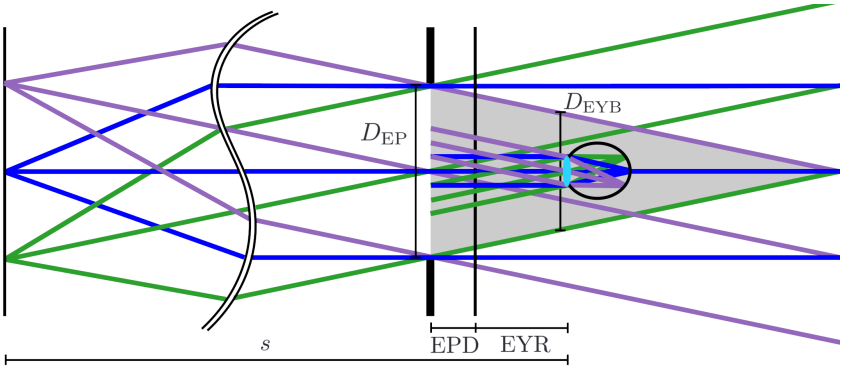


Figure 1.1: Schematic of a near-eye display showing a virtual image at distance s. Exit pupil and eyebox volume are illustrated.

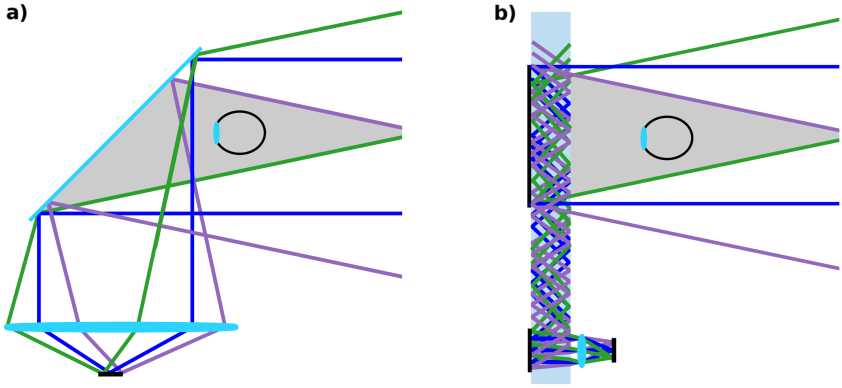


Figure 1.2: Comparison of display architectures. (a) A free-space combiner uses a semi-transparent optical element to overlay the virtual image on the real-world view, requiring a relatively large collimator to support the desired field of view and eyebox. (b) A waveguide combiner guides and repeatedly extracts light to expand the exit pupil within a compact form factor suitable for integration into an eyeglass lens. Actual waveguide substrates are thinner than illustrated, increasing pupil-sampling density.

display systems [5, 6]. However, due to étendue conservation, conventional free-space combiner architectures remain constrained when a large eyebox, wide field of view, and eyeglass-like compactness are required simultaneously.

Waveguide combiners address this constraint by introducing controlled ray splitting. In a 1980 filing, Searle proposed injecting an angular spectrum of collimated ray bundles, corresponding to a virtual image at infinity, into a planar substrate at angles guided by total internal reflection. The ray bundles are repeatedly and partially extracted along the propagation direction [7], see Fig. 1.2(b). This replicates the exit pupil along one dimension and expands the system's eyebox, thereby increasing the effective étendue at the cost of dividing the optical power among the pupil replicas. Amitai and Friesem later extended this principle by partially redirecting the guided light prior to extraction, enabling a matrix-like two-dimensional eyebox expansion (2D EBE) [8, 9]. Eyebox expansion enables highly compact waveguide displays that combine large field of view and large eyebox, making waveguide combiners an attractive architecture for compact AR glasses.

Several waveguide coupler technologies have been developed for light insertion, redirection, and extraction. Holographic waveguides use volume grating couplers, so-called holographic optical elements (HOEs), and build on foundational waveguide work by Upatnieks in the 1980s [10–12], Amitai *et al.* [13, 14] in the 1990s, and early industrial research [15]. This approach was developed into products by Sony [16–18] and BAE Systems [19, 20]. Surface relief grating (SRG) waveguide couplers were developed by Nokia [21] and later used in products such as Microsoft HoloLens [22]. Lumus developed geometric couplers using partially reflective mirrors embedded in the waveguide substrate [23, 24]. These have recently been adopted in the Meta Ray-Ban Display [25].

Holographic waveguide couplers offer two important advantages for consumer AR glasses. First, they can provide high transparency with low levels of visual artifacts. Volume diffraction from HOEs provides higher angular and wavelength selectivity than surface diffraction from SRGs. As a result, most ambient light is far off-Bragg and is not diffracted into the user's or a bystander's eye, reducing rainbow artifacts. Similarly, undesired diffraction orders of the display light are suppressed, reducing active display eyeglow and leakage [22]. Geometric waveguides, by contrast, exhibit several stray-light paths and produce a louvre-blind effect when the partially reflective mirrors are positioned directly in front of the eye [22].

A second, less widely recognized advantage is that holographic waveguides are particularly well suited for integration into thin, lightweight, curved prescription-lens stacks. Regular prescription lenses use a convex front base curvature to conform to face-wrapped eyewear frames and to control off-axis aberrations [26, p. 74], while the back surface provides the required prescription power. Inserting planar optical elements between the front and back prescription-lens surfaces leads to significantly thicker lens stacks than regular eyewear, unlike spherically curved elements that naturally conform to the base curvature [27, 28]. Consequently, existing planar geometric and SRG waveguide displays generally require nearly vanishing base curvature if acceptable thickness and weight are to be maintained, leading to unnaturally flat eyewear with limited face-wrap.

Recent work has demonstrated that spatially varying holographic couplers can enable cylindrically curved waveguides with monochromatic aberration correction [29–34], following earlier conceptual filings [35, 36]. For integration with a nonzero base curvature, the outcoupling HOE must impart the same optical power to the display light as the front surface imparts to the world view. The back surface curvature then brings both the world view and the display image into focus [37].

Holographic waveguides can, in principle, provide both curved-waveguide aberration correction and the required outcoupling power. Their spectral selectivity also provides a path toward full-color curved waveguides by stacking several HOEs dedicated to individual color channels. This combination appears difficult to realize with less selective SRGs, while geometric waveguides lack the required wavefront control. These properties make holographic waveguides particularly promising for lightweight integration into prescription lenses.

Realizing prescription-compatible spherical waveguide displays requires progress in both optical design and holographic fabrication.

From a design perspective, freeform HOEs providing non-paraxial aberration correction have so far been derived only for cylindrical waveguide displays, whereas direct conformity with ophthalmic base curves requires spherical curvature.

From a fabrication perspective, spherical waveguides require spatially varying HOEs recorded by accurately controlled freeform object and reference wavefronts, adjusted for material shrinkage. This exceeds the capabilities of established waveguide recording systems. Currently, even planar holographic waveguide prototyping commonly relies on custom exposure systems rather than dedicated commercial mastering machines. Incoupling, expanding, and outcoupling HOEs are typically recorded in separate purpose-built exposure systems and subsequently relaminated onto a common substrate [38, 39]. Iterative compensation of shrinkage or design changes therefore requires repeated manual realignment of the exposure systems. Conventional waveguide coupler recording additionally relies on immersion contact with edge-lit prisms, which complicates handling and slows fabrication cycles. A reliable and repeatable planar fabrication process

is therefore a necessary experimental foundation for spherical-waveguide development: it enables systematic characterization of material shrinkage, exposure dose, alignment tolerances, immersion-liquid flow, and other process bottlenecks that must be controlled before curved substrates or spatially varying recording wavefronts can be implemented. Recently proposed wavelength-shifted exposure strategies avoid immersion contact and edge-lit exposure for selected waveguide coupler geometries and could substantially simplify fabrication and prototyping if they were applicable to practical planar and spherical waveguide designs [40–42].

To the author's knowledge, this thesis presents the first non-paraxial design framework for spherical holographic waveguides, shows that immersion-free recording is not applicable to practically relevant planar and spherical photopolymer waveguides with two-dimensional eyebox expansion, and introduces an automated robotic exposure platform for the repeatable fabrication of planar holographic waveguide displays using a single custom-designed edge-lit prism assembly. Together, these contributions establish theoretical and experimental foundations for the broader goal of realizing spherical holographic waveguides. The thesis is structured as follows.

Chapter 2 introduces the optical and holographic foundations used throughout the thesis and quantifies the weight advantage of embedding spherical waveguides into prescription lenses. Chapter 3 shows that the eigenmode concept, which has been successfully applied for aberration correction in cylindrical waveguide design, breaks down for spherical waveguides: no such eigenmodes exist. Instead, the chapter introduces closed modes under waveguide jump maps, which coincide with eigenmodes in the cylindrical limit but remain available in the spherical case. Based on this concept, the chapter presents the first non-paraxial designs of spherical waveguide displays. The method is compatible with 2D eyebox expansion, is quasi-analytical, and decomposes the highly non-convex end-to-end design problem into two simplified optimization problems.[1] Chapter 4 examines

[1] Published in: Marvin Feil, Moritz Esslinger, Reinhold Fiess, and Wilhelm Stork, "Closed-mode families for planar, cylindrical, and spherical waveguide display design," *Opt. Express* **34**(14): 26031–26046 (2026). (Editors' pick)

how far immersion-contacted, edge-lit exposure can be avoided by shifting the recording wavelength and configuration. It derives analytical recordability limits and shows that a broad range of relevant incouplers can be recorded from air, whereas typical planar and spherical expander gratings cannot.[2] Chapter 5 introduces an exposure machine based on two six-axis robot arms that automatically adjust the recording-beam positions, directions, and polarization orientations. A custom edge-lit prism system enables edge-lit recording of planar waveguide displays in a single substrate, including incoupler, expander, and outcoupler gratings. The complete process is demonstrated with a planar 2D-EBE validation sample.[3]

The outlook returns to the fabrication of spherical holographic waveguides and outlines two directions for future work. First, it proposes a concept for extending the twin-prism exposure system to curved waveguide displays using refractive freeform optics.[4] Second, it proposes a concept for the spatially resolved scatterometric metrology of waveguide couplers.

[2] Published in: Marvin Feil, Tobias Wilm, Moritz Esslinger, Reinhold Fiess, and Wilhelm Stork, "Limits of immersion-free recording of holographic waveguide displays," *Opt. Lett.* **50**(2): 606–609 (2025).

[3] Submitted for publication: Marvin Feil, Reinhold Fiess, and Wilhelm Stork, "Automated mastering of holographic waveguide displays with robot-mounted optics," *Appl. Opt.* Parts of this work were presented in Marvin Feil, Reinhold Fiess, and Wilhelm Stork, "Automated mastering of holographic waveguide displays," in *Proc. SPIE* **13821**, p. 138210A (2026), and Marvin Feil, Tobias Wilm, Reinhold Fiess, and Wilhelm Stork, "Automated holographic recording using six-axis robot arm pair," in *Proc. SPIE* **13529**, p. 135290J (2025).

[4] Part of this work was awarded first place at the SPIE AR|VR|MR Student Optical Design Challenge 2026: Marvin Feil, Moritz Esslinger, Yuchen Deng, Reinhold Fiess, and Wilhelm Stork, "Freeform Recording of Curved Holographic Waveguides."

2 Fundamentals

In this chapter, we introduce the fundamentals of holographic optical elements (HOEs), discuss the HOE couplers required for conventional planar holographic waveguide displays, and quantify the weight and thickness advantage of curved waveguides over planar ones when integrated into prescription lenses.

2.1 Holographic optical elements (HOEs)

The interference of two coherent optical fields

$$\mathbf{E}_1(\mathbf{r}) = \mathbf{A}_1(\mathbf{r})e^{i\phi_1(\mathbf{r})}, \quad \mathbf{E}_2(\mathbf{r}) = \mathbf{A}_2(\mathbf{r})e^{i\phi_2(\mathbf{r})}, \tag{2.1}$$

produces regions of high and low intensity,

$$I(\mathbf{r}) \propto |\mathbf{E}_1 + \mathbf{E}_2|^2 = |\mathbf{A}_1|^2 + |\mathbf{A}_2|^2 + 2|\mathbf{A}_1 \cdot \mathbf{A}_2^*| \cos(\phi_1 - \phi_2 + \arg(\mathbf{A}_1 \cdot \mathbf{A}_2^*)).$$

Recording this interference pattern in a film of photosensitive material without oversaturation yields, to first order, a spatially varying permittivity structure

$$\epsilon(\mathbf{r}) = \epsilon' + \gamma I(\mathbf{r}). \tag{2.2}$$

Yuri Denisyuk called such a patterned film a wave photograph and showed that it diffracts the original recording field $\mathbf{E}_1$ into $\mathbf{E}_2$, and vice versa [43]. Today we refer to such wave photographs as holograms, a term first introduced by Dennis Gabor [44]. Denisyuk's argument is based on integrating the Fresnel reflection of $\mathbf{E}_1$ over the many constant-permittivity surfaces, or Bragg planes, of the hologram

volume. For holograms in thin films, additional diffraction orders and diffracted fields can occur. These are suppressed in sufficiently thick film layers.

Denisyuk demonstrated this wave reconstruction principle experimentally by recording the interference between a light beam and its reflection from a convex mirror in a silver halide emulsion [43]. For the incident reference wavefront, the hologram reproduces the optical functionality of the convex mirror: the incident field is diffracted into the wavefront that would be produced by reflection from the mirror. We refer to such holograms of optical elements as holographic optical elements (HOEs) [45].

HOEs are strongly wavelength- and angle-selective. A wavefront different from the recording wavefront will generally be diffracted by the HOE differently from how it would be reflected by the recorded convex mirror. HOEs can operate in both transmission and reflection and can be recorded from arbitrary optical elements, such as lenses, mirrors, gratings, or any other element that provides an optical field. The HOE function is independent of the hologram film curvature as long as the geometry is unchanged between recording and replay [45].

In many practical applications, HOEs are used with several different incident beams, for example to image a finite field of view. The HOE can be optimized to diffract these different incident wavefronts approximately into target wavefronts, rather than diffracting one wavefront perfectly. This optimization accounts for the angular and wavelength dispersion of the diffraction physics and typically results in spatially varying permittivity structures that must be recorded using appropriate freeform wavefronts [46, 47].

2.1.1 Local volume grating description

At any position $\mathbf{r}'$, the holographic film is locally planar, irrespective of its overall curvature, with surface normal $\hat{\mathbf{N}}(\mathbf{r}')$. Similarly, the fields in Eq. (2.1) can be approximated

$$\mathbf{E}_i(\mathbf{r}) \approx \mathbf{A}_i(\mathbf{r}')e^{i\left[\phi_i(\mathbf{r}')+\mathbf{k}_i(\mathbf{r}')\cdot(\mathbf{r}-\mathbf{r}')\right]}, \tag{2.3}$$

with the amplitude $\mathbf{A}_i$ varying slowly on the scale of a few wavelengths and the wave vector $\mathbf{k}_i(\mathbf{r}') = \nabla\phi_i(\mathbf{r}')$ being normal to the local phase front. The resulting local interference pattern is a sinusoidal irradiance grating with reciprocal grating vector

$$\mathbf{G}^{\mathrm{irr}}(\mathbf{r}') = \mathbf{k}_2(\mathbf{r}') - \mathbf{k}_1(\mathbf{r}'). \tag{2.4}$$

Ignoring material shrinkage, this pattern is recorded as a permittivity grating with reciprocal vector $\mathbf{G} \approx \mathbf{G}^{\mathrm{irr}}$. The Bragg planes are normal to $\mathbf{G}$, and their spacing is given by the grating pitch $\Lambda = 2\pi/|\mathbf{G}|$. Modern holographic photopolymers are usually highly transmissive, $\mathrm{Im}(\epsilon) \approx 0$, and weakly modulated, $\Delta\epsilon \ll \bar{\epsilon}$. The corresponding refractive-index modulation $n(\mathbf{r}) = \sqrt{\epsilon(\mathbf{r})}$ shown in Fig. 2.1 is therefore also approximately sinusoidal,

$$n(\mathbf{r}) = \bar{n} + \Delta n \cos(\mathbf{G}\cdot\mathbf{r} + \phi_0). \tag{2.5}$$

The achievable index modulation Δn is limited by the finite dynamic range $2\delta n$ of the recording material before saturation, as well as by the visibility V of the recording interference pattern. Low visibility means that the dark fringes are not completely dark, so that the index is also increased in these regions, reducing the attainable modulation. Reaction-diffusion models of photopolymer exposure show that, below saturation, the maximum recorded modulation is [48, 49]

$$\Delta n_{\max} \approx V\delta n = \frac{2\sqrt{I_1 I_2}|\hat{\mathbf{A}}_1 \cdot \hat{\mathbf{A}}_2^*|}{I_1 + I_2}\delta n. \tag{2.6}$$

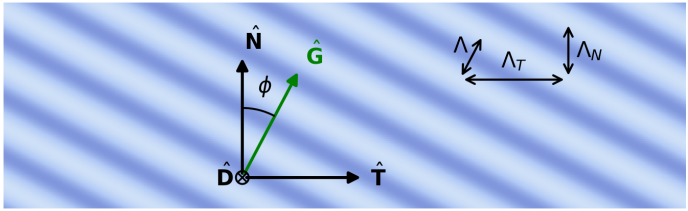


Figure 2.1: Cross-section of a volume grating, visualizing the grating vector and coordinate system. Slant angle ϕ, period Λ, normal period Λ_N and tangential period Λ_T are shown.

2.1.2 Local grating coordinate system

For later analysis, we define the local grating coordinate system $\{\hat{\mathbf{N}}, \hat{\mathbf{T}}, \hat{\mathbf{D}}\}$ attached to the hologram surface, see Fig. 2.1. It is spanned by the surface normal $\hat{\mathbf{N}}$, the tangential grating direction $\hat{\mathbf{T}} = \left(\mathbf{G} - (\mathbf{G}\cdot\hat{\mathbf{N}})\hat{\mathbf{N}}\right) / \left|\mathbf{G} - (\mathbf{G}\cdot\hat{\mathbf{N}})\hat{\mathbf{N}}\right|$, and the binormal $\hat{\mathbf{D}} = \hat{\mathbf{N}} \times \hat{\mathbf{T}}$.

In this system the grating vector is described by only two components,

$$\mathbf{G} = G_T\hat{\mathbf{T}} + G_N\hat{\mathbf{N}}, \tag{2.7}$$

the normal grating component $G_N = \hat{\mathbf{N}}\cdot\mathbf{G}$ and the tangential grating component $G_T = \hat{\mathbf{T}}\cdot\mathbf{G}$. We also define the tangential grating vector $\mathbf{G}_T = \mathbf{G} - G_N\hat{\mathbf{N}}$. The sign of the fringe normal $\mathbf{G}$ is arbitrary. We fix it by requiring $G_N > 0$. By construction of $\hat{\mathbf{T}}$, we have $G_T > 0$. The local pitch and slant of the grating are fully described by G_T and G_N, with pitch $\Lambda = 2\pi/|\mathbf{G}|$, tangential pitch $\Lambda_T = 2\pi/|G_T|$, normal pitch $\Lambda_N = 2\pi/|G_N|$, and slant angle $\phi = \arctan(G_T/G_N)$. In a general coordinate system, the third coordinate of the grating vector specifies the rotation of the grating on the hologram surface.

2.1.3 Hologram shrinkage

Most holographic recording materials shrink during post-exposure curing. Because the holographic film is attached to a carrier substrate, this shrinkage is typically anisotropic, with different shrinkage coefficients in the normal and tangential directions. If the exposure irradiance pattern has periods Λ_N^{irr} and Λ_T^{irr}, the material grating after curing has periods $\Lambda_N = (1-s)\Lambda_N^{\mathrm{irr}}$ and $\Lambda_T = (1-s_T)\Lambda_T^{\mathrm{irr}}$, where s and s_T are the normal and tangential shrinkage coefficients [46, 50]. Equivalently, the irradiance grating vector required to obtain a desired material grating $\mathbf{G} = G_T\hat{\mathbf{T}} + G_N\hat{\mathbf{N}}$ is

$$\mathbf{G}^{\mathrm{irr}}(\mathbf{G}, s, s_T) = G_T^{\mathrm{irr}}\hat{\mathbf{T}} + G_N^{\mathrm{irr}}\hat{\mathbf{N}} = G_T(1-s_T)\,\hat{\mathbf{T}} + G_N(1-s)\,\hat{\mathbf{N}}. \tag{2.8}$$

2.1.4 Decomposition in recording wavefronts

For a desired grating vector $\mathbf{G}$, the required exposure irradiance grating vector $\mathbf{G}^{\mathrm{irr}}$ is specified by Eq. (2.8). Equation (2.4) specifies the irradiance grating vector produced by two recording wave vectors $\mathbf{k}_o$ and $\mathbf{k}_r$. There is an infinite family of recording wave vector pairs satisfying this condition [46, 47]. As derived in Appendix A.1, this family can be parametrized by the recording wavelength λ and the angular recording configuration parameter $\sigma \in [0, 2\pi)$,

$$\begin{aligned}\mathbf{k}_r(\sigma,\lambda) &= -\frac{\mathbf{G}^{\mathrm{irr}}}{2} + \sqrt{k^2 - \frac{\left(G^{\mathrm{irr}}\right)^2}{4}}\,\hat{\mathbf{\Delta}} \\ \mathbf{k}_o(\sigma,\lambda) &= \frac{\mathbf{G}^{\mathrm{irr}}}{2} + \sqrt{k^2 - \frac{\left(G^{\mathrm{irr}}\right)^2}{4}}\,\hat{\mathbf{\Delta}}.\end{aligned} \tag{2.9}$$

Here $k = 2\pi n/\lambda$ is the recording wave vector magnitude inside the recording material with refractive index n and the unit vector $\hat{\mathbf{\Delta}}$ is normal to the grating vector,

$$\hat{\mathbf{\Delta}}(\sigma) = \cos(\sigma)\,\frac{1}{G^{\mathrm{irr}}}\left(G_T^{\mathrm{irr}}\,\hat{\mathbf{N}} - G_N^{\mathrm{irr}}\,\hat{\mathbf{T}}\right) - \sin(\sigma)\left(\hat{\mathbf{N}} \times \hat{\mathbf{T}}\right). \tag{2.10}$$

To achieve maximum recording visibility and, consequently, the maximum hologram modulation in Eq. (2.6), the polarization overlap $|\hat{\mathbf{p}}_r \cdot \hat{\mathbf{p}}_o^*|$ of both recording beams must be unity. Assuming linear polarization, this requires

$$\hat{\mathbf{p}}_r = \hat{\mathbf{p}}_o = \pm\frac{\mathbf{k}_r \times \mathbf{k}_o}{|\mathbf{k}_r \times \mathbf{k}_o|}. \tag{2.11}$$

2.1.5 Volume grating diffraction

Consider a plane wave with wave vector $\mathbf{k}_i$ incident on a volume grating. Due to discrete translation symmetry, as detailed in Appendix A.2.1, the diffracted field

is given by a sum of discrete plane wave diffraction orders m with wave vectors $\mathbf{k}_d^m$. The tangential part of this wave vector $\mathbf{k}_{d,\parallel}^m = \mathbf{k}_d^m - \left(\hat{\mathbf{N}} \cdot \mathbf{k}_d^m\right) \hat{\mathbf{N}}$ is

$$\mathbf{k}_{d,\parallel}^m = \mathbf{k}_{i,\parallel} + mG_T\hat{\mathbf{T}}, \qquad m \in \mathbb{Z}, \tag{2.12}$$

where $\mathbf{k}_{i,\parallel} = \mathbf{k}_i - \left(\hat{\mathbf{N}} \cdot \mathbf{k}_i\right) \hat{\mathbf{N}}$ is the tangential part of the incident wave vector. We call this the tangential momentum matching equation or in angular form the grating equation. The corresponding wave vector normal component $k_{d,N}^m = \hat{\mathbf{N}} \cdot \mathbf{k}_d^m$ is determined by the dispersion relation,

$$k_{d,N}^m = \pm\sqrt{k^2 - \left|\mathbf{k}_{d,\parallel}^m\right|^2} = \pm\sqrt{(2\pi/\lambda)^2\,\bar{\epsilon} - \left|\mathbf{k}_{i,\parallel} + mG_T\hat{\mathbf{T}}\right|^2}. \tag{2.13}$$

The two possible signs correspond to transmitted and reflected diffraction orders.

Efficient diffraction occurs only in the orders that are phase matched through the grating thickness. If there is a nonzero Bragg wave vector mismatch

$$\Delta k(\mathbf{k}_i, \hat{\mathbf{N}}, \mathbf{G}, m) = k_{i,N} + mG_N - k_{d,N}^m \tag{2.14}$$

for a given diffraction order, wavelength and incidence direction, waves generated at different depths interfere destructively, as detailed in Appendix A.2.2. For a thick grating, appreciable diffraction is confined to a narrow neighborhood of the Bragg phase matching condition, $\Delta k = 0$. Combining Eq. (2.14) with Eq. (2.12) gives Laue's formulation of the Bragg condition [51]

$$\mathbf{k}_d^m = \mathbf{k}_i + m\mathbf{G}, \quad m \in \mathbb{Z}, \qquad k = (2\pi/\lambda)\sqrt{\bar{\epsilon}} = |\mathbf{k}_d^m|. \tag{2.15}$$

This condition is best interpreted following Ewald [52]. Efficient diffraction occurs when the sum of the incident wave vector and an integer multiple of the reciprocal grating vector lies on the dispersion relation, or Ewald sphere.

For a given incidence direction and wavelength, sinusoidal volume gratings transfer meaningful energy only in the first order closest to phase matching, either the reflective or transmissive $m = \pm 1$ branch, and not in any higher orders. This

is explained in Appendix A.2.3. This is the origin of the strong angular and wavelength selectivity of thick volume gratings: only one first order is efficiently excited, and only when the incident wave is sufficiently close to phase matching.

Kogelnik's coupled-wave theory treats this two-wave regime [53]. The resulting transmitted and reflected diffracted-field amplitudes are summarized for arbitrary incidence and polarization in [54]. The corresponding diffraction efficiencies for TE and TM polarization are obtained by taking the squared field magnitudes and applying the obliquity factors [53]. For non-absorbing volume phase gratings, the resulting transmission and reflection efficiencies are

$$\eta_{\mathrm{trans}} = \frac{\sin^2\sqrt{\nu_i^2+\xi^2}}{1+\xi^2/\nu_i^2}, \qquad \eta_{\mathrm{reflect}} = \frac{1}{1+\frac{1-\xi^2/\nu_i^2}{\sinh^2\sqrt{\nu_i^2-\xi^2}}}. \tag{2.16}$$

The coefficients ξ and ν_i in Kogelnik's theory differ for reflection and transmission gratings, but may be written in a unified form as shown by Lindlein [46],

$$\nu_i = \frac{d\kappa_i}{\sqrt{|c_S c_R|}}, \qquad \xi = \frac{\vartheta d}{2c_S}. \tag{2.17}$$

Here, d is the grating thickness, $c_R = \hat{\mathbf{N}}\cdot\hat{\mathbf{k}}_{\mathrm{i}}$ and $c_S = \hat{\mathbf{N}}\cdot(\mathbf{k}_{\mathrm{i}}+m\mathbf{G})/k$ are obliquity coefficients, and $\vartheta = \left(k^2-|\mathbf{k}_{\mathrm{i}}+m\mathbf{G}|^2\right)/(2k)$ is the dephasing parameter. The eigenpolarization-dependent coupling coefficient is

$$\kappa_i = \frac{\pi\Delta n}{\lambda}\begin{cases}1 & \text{TE-pol,}\\ \hat{\mathbf{k}}_{\mathrm{i}}\cdot(\mathbf{k}_{\mathrm{i}}+m\mathbf{G})/|\mathbf{k}_{\mathrm{i}}+m\mathbf{G}| & \text{TM-pol.}\end{cases} \tag{2.18}$$

For vanishing Δk, the dephasing parameter ϑ becomes zero.

2.1.6 Hologram ray tracing

We perform geometrical ray tracing through an HOE following Welford [55]. The hologram is represented by an effective diffracting surface carrying the local

grating vector $\mathbf{G}$, index modulation Δn, and surface normal $\hat{\mathbf{N}}$. An incident ray is treated locally as a plane wave with direction $\hat{\mathbf{k}}_{\mathrm{i}}$ inside the holographic substrate. The tangential component of the diffracted ray direction is calculated from the tangential momentum-matching condition in Eq. (2.12), while the normal component follows from the dispersion relation in Eq. (2.13). Only the single non-negligible first diffraction order is traced, with the efficiency given by Eq. (2.16).

Refraction into and out of the holographic substrate is applied at two surfaces surrounding the effective holographic surface. The substrate refractive index is set to the average hologram index [54, 55].

2.2 Planar holographic waveguide displays

This section introduces planar holographic waveguide displays using terminology that connects directly to the mode-based treatment of arbitrarily curved waveguides in Chapter 3.

2.2.1 Planar waveguide eigenmodes

As discussed in Chapter 1, a waveguide display transports image information from the light engine to the eye as an angular spectrum of collimated ray bundles guided by total internal reflection within a transparent substrate. Each bundle is characterized by its wave vector $\mathbf{k}$ and is focused by the eye onto a corresponding image point on the retina. The optical power carried by each angular component determines the brightness of the associated image point.

Propagation is described in the waveguide system $\{\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}}\}$ shown in Fig. 2.2, where $\hat{\mathbf{z}}$ is normal to the substrate and points toward the inner waveguide surface.

At the design wavelength λ_{d}, the free-space wavenumber is $k_0 = 2\pi/\lambda_{\mathrm{d}}$, and the wave vector magnitude in the waveguide core is $k_{\mathrm{d}} = nk_0$, where n is the refractive index of the waveguide core.

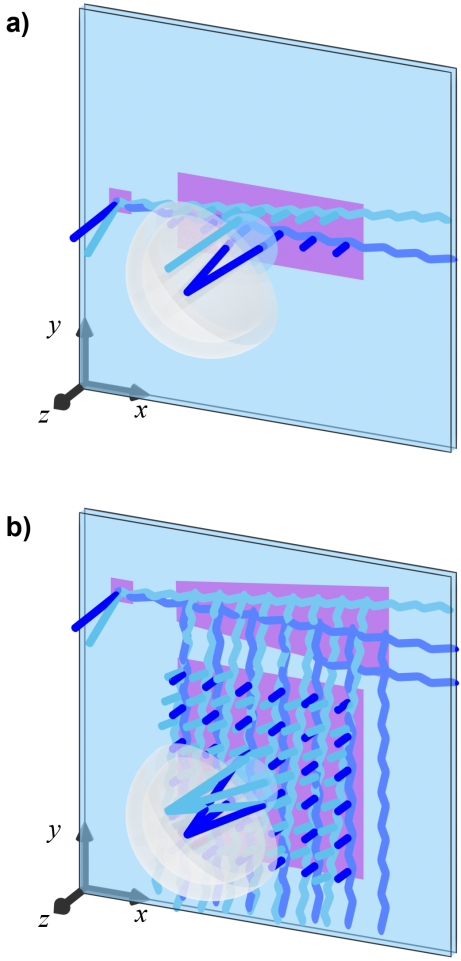


Figure 2.2: Exemplary orthogonal (a) 1D-EBE and (b) 2D-EBE planar waveguide architectures. The central free-space mode is shown in dark blue and incoupled and outcoupled normal to the substrate. In the 2D-EBE architecture, the expander rotates the in-plane propagation direction by $90°$ while preserving the propagation angle $\theta_p = 56°$. One off-axis mode (corresponding to a $(-8°, -8°)$ field point) is shown in each case in light blue.

A ray bundle is guided by total internal reflection when its angle of incidence at both planar interfaces exceeds the critical angle. Equivalently, its tangential wave vector components satisfy

$$k_x^2 + k_y^2 > k_0^2 n_{\mathrm{clad}}^2, \tag{2.19}$$

where n_{clad} is the refractive index of the waveguide cladding. The bundle then propagates between the two interfaces by repeated reflections, as illustrated in Fig. 2.2. We define one *waveguide jump* as a round trip from the outer surface to the inner surface and back. At each reflection, only the normal component of the wave vector changes sign, whereas the in-plane components (k_x, k_y) remain unchanged. These components are therefore invariants of propagation, and the ray bundle recurs after each jump up to a lateral translation. We refer to the set of translated ray bundles generated in this way as a planar waveguide eigenmode, specified by the invariant mode coordinates (k_x, k_y). The two propagation branches of the same mode are

$$\mathbf{k}_{\pm}(k_x, k_y) = k_x \hat{\mathbf{x}} + k_y \hat{\mathbf{y}} \pm k_z \hat{\mathbf{z}}, \qquad k_z = \sqrt{k_{\mathrm{d}}^2 - k_x^2 - k_y^2}. \tag{2.20}$$

Reflections alternate between $\mathbf{k}_+(k_x, k_y)$ and $\mathbf{k}_-(k_x, k_y)$ without changing the mode coordinates.

Modes that do not satisfy Eq. (2.19) are not guided and can cross the waveguide interfaces. In the following, they are referred to as free-space modes. A waveguide image source, such as a laser-beam scanner or a collimated display panel, provides an angular spectrum of such collimated input bundles.

2.2.2 Central-mode grating coupler construction

In a holographic waveguide, volume holographic gratings couple the input free-space modes into guided modes, redirect these modes within the substrate, and finally extract them toward the eye. The coupler HOEs are constructed for the

central free-space mode, which lies at the center of the source angular spectrum and represents the center of the displayed field of view.

Immediately inside the substrate, the central image bundle reaches the incoupler with wave vector $\mathbf{k}^{(\mathrm{i})} = \mathbf{k}_{-}(k_x^{(\mathrm{i})}, k_y^{(\mathrm{i})})$. For normal incidence, $(k_x^{(\mathrm{i})}, k_y^{(\mathrm{i})}) = (0, 0)$ and $\mathbf{k}^{(\mathrm{i})} = -k_{\mathrm{d}}\hat{\mathbf{z}}$.

In a reflective holographic waveguide, an incoupling HOE located at the outer surface diffracts this bundle into the guided mode $\mathbf{k}^{(1)} = \mathbf{k}_{+}(k_x^{(1)}, k_y^{(1)})$. Its propagation angle relative to the substrate normal is $\theta_p = \arccos[|\mathbf{k}^{(1)} \cdot \hat{\mathbf{z}}|/k_{\mathrm{d}}]$.

After one or more jumps, the first guided mode returns to the outer surface with wave vector $\mathbf{k}_{\mathrm{r}}^{(1)} = \mathbf{k}_{-}(k_x^{(1)}, k_y^{(1)})$. In a one-dimensional eyebox-expanding waveguide, this mode is extracted directly by the outcoupling HOE. In a two-dimensional eyebox-expanding waveguide, an expander HOE first redirects it into a second guided mode $\mathbf{k}^{(2)} = \mathbf{k}_{+}(k_x^{(2)}, k_y^{(2)})$, which is subsequently extracted by the outcoupler.

The tangential grating vectors of the reflective incoupler, expander, and outcoupler required for this light redirection follow from the tangential momentum matching condition in Eq. (2.12), whereas the normal grating vector components are specified by Eq. (2.15) such that the central mode is efficiently diffracted:

$$\begin{aligned} \mathbf{G}^{\mathrm{i}} &= \mathbf{k}^{(1)} - \mathbf{k}^{(\mathrm{i})}, \quad \mathbf{G}^{\mathrm{o1D}} = \mathbf{k}^{(\mathrm{out})} - \mathbf{k}_{\mathrm{r}}^{(1)}, \\ \mathbf{G}^{\mathrm{e}} &= \mathbf{k}^{(2)} - \mathbf{k}_{\mathrm{r}}^{(1)}, \quad \mathbf{G}^{\mathrm{o2D}} = \mathbf{k}^{(\mathrm{out})} - \mathbf{k}_{\mathrm{r}}^{(2)}. \end{aligned} \tag{2.21}$$

These expressions also apply to non-normal incoupling and outcoupling caused, for example, by light-engine tilt, face wrap, or pantoscopic tilt. A finite image distance can be introduced by recording the outcoupler with a defocused object wave, thereby adding a spatially varying defocus component to its local grating vector field.

2.2.3 Orthogonal waveguides

The minimal geometry shown in Fig. 2.2 excludes light-engine tilt, face wrap, and non-orthogonal redirection by the expander. We refer to such 1D- and 2D-EBE architectures as *orthogonal* waveguides. Their central mode is incoupled and outcoupled normal to the substrate, $\hat{\mathbf{k}}^{(\mathrm{i})} = -\hat{\mathbf{z}}$ and $\hat{\mathbf{k}}^{(\mathrm{out})} = \hat{\mathbf{z}}$. In the orthogonal 2D-EBE architecture, the expander rotates the in-plane propagation direction by 90° while preserving the propagation angle θ_p. The corresponding guided modes are

$$\mathbf{k}^{(1)} = k_{\mathrm{d}} \left(\sin\theta_p \hat{\mathbf{x}} + \cos\theta_p \hat{\mathbf{z}}\right), \qquad \mathbf{k}^{(2)} = k_{\mathrm{d}} \left(-\sin\theta_p \hat{\mathbf{y}} + \cos\theta_p \hat{\mathbf{z}}\right). \quad (2.22)$$

Equation (2.21) then gives

$$\begin{aligned} \mathbf{G}^{\mathrm{i}} &= k_{\mathrm{d}} \left[\sin\theta_p \hat{\mathbf{x}} + (1+\cos\theta_p)\, \hat{\mathbf{z}}\right], \\ \mathbf{G}^{\mathrm{e}} &= k_{\mathrm{d}} \left[-\sin\theta_p \left(\hat{\mathbf{x}} + \hat{\mathbf{y}}\right) + 2\cos\theta_p \hat{\mathbf{z}}\right], \\ \mathbf{G}^{\mathrm{o1D}} &= k_{\mathrm{d}} \left[-\sin\theta_p \hat{\mathbf{x}} + (1+\cos\theta_p)\, \hat{\mathbf{z}}\right], \\ \mathbf{G}^{\mathrm{o2D}} &= k_{\mathrm{d}} \left[\sin\theta_p \hat{\mathbf{y}} + (1+\cos\theta_p)\, \hat{\mathbf{z}}\right]. \end{aligned} \quad (2.23)$$

The corresponding local tangential and normal grating components are

$$\begin{aligned} G_T^{\mathrm{i}} = G_T^{\mathrm{o1D}} = G_T^{\mathrm{o2D}} &= k_{\mathrm{d}} \sin\theta_p, \\ G_N^{\mathrm{i}} = G_N^{\mathrm{o1D}} = G_N^{\mathrm{o2D}} &= k_{\mathrm{d}} \left(1+\cos\theta_p\right), \\ G_T^{\mathrm{e}} &= \sqrt{2}\, k_{\mathrm{d}} \sin\theta_p, \\ G_N^{\mathrm{e}} &= 2 k_{\mathrm{d}} \cos\theta_p. \end{aligned} \quad (2.24)$$

Thus, the incoupler and outcoupler have identical local components (G_T, G_N) and differ only in their in-plane orientation. The 1D-EBE architecture shown in Fig. 2.3 therefore uses the same local incoupler and outcoupler gratings as the corresponding 2D-EBE architecture.

2.2.4 Off-axis modes and achromaticity

Off-axis image points are represented by incident collimated bundles with tangential wave vector components (k_x, k_y). According to the tangential momentum-matching condition in Eq. (2.12), each planar coupler designed for the central image mode translates the corresponding off-axis image mode in the k_x–k_y plane. For example, the incoupler performs the mapping

$$(k_x, k_y; \lambda) \mapsto (k_x + G_x^{\mathrm{i}}, k_y + G_y^{\mathrm{i}}; \lambda). \tag{2.25}$$

Provided that the translated mode remains inside the guiding region, a 2D-EBE waveguide transports it to the expander, redirects it to $(k_x + G_x^{\mathrm{i}} + G_x^{\mathrm{e}}, k_y + G_y^{\mathrm{i}} + G_y^{\mathrm{e}})$, and transports it to the outcoupler. The extracted free-space mode then has tangential coordinates $(k_x + G_x^{\mathrm{i}} + G_x^{\mathrm{e}} + G_x^{\mathrm{o2D}}, k_y + G_y^{\mathrm{i}} + G_y^{\mathrm{e}} + G_y^{\mathrm{o2D}})$. In a 1D-EBE waveguide, the incoupled mode is extracted directly, without expander redirection, into $(k_x + G_x^{\mathrm{i}} + G_x^{\mathrm{o1D}}, k_y + G_y^{\mathrm{i}} + G_y^{\mathrm{o1D}})$. This mode-translation picture is central to the generalization from planar to curved waveguides in Chapter 3. It

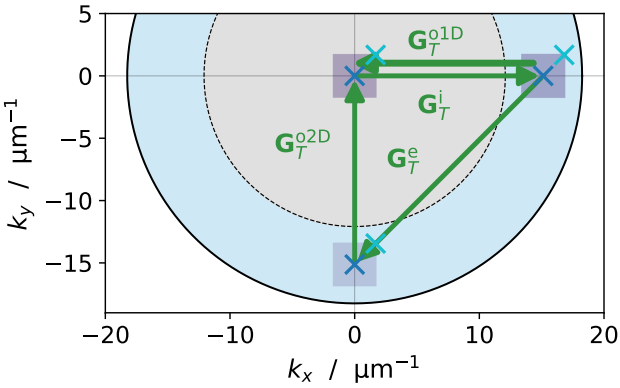


Figure 2.3: k-space diagram illustrating the orthogonal 1D- and 2D-EBE waveguides shown in real space in Fig. 2.2. The marked region corresponds to a $16° \times 16°$ FOV that is translated in k-space by the incoupler, expander and outcoupler. The blue markers correspond to the modes that transport the center image point and the cyan markers to the $(-8°, 8°)$ off-axis image point.

is most transparent in k-space, as illustrated in Fig. 2.3 for orthogonal 1D- and 2D-EBE designs with $\theta_p = 56°$, $\lambda_\mathrm{d} = 520\,\mathrm{nm}$, $n = 1.51$, and $n_\mathrm{clad} = 1$.

If the central output mode is collinear with the incident central mode, the tangential grating vector components satisfy

$$\mathbf{0} = \mathbf{G}_\mathrm{T}^\mathrm{i} + \mathbf{G}_\mathrm{T}^\mathrm{e} + \mathbf{G}_\mathrm{T}^\mathrm{o2D} \quad \text{and} \quad \mathbf{0} = \mathbf{G}_\mathrm{T}^\mathrm{i} + \mathbf{G}_\mathrm{T}^\mathrm{o1D}, \tag{2.26}$$

where $\mathbf{G}_T = G_T\hat{\mathbf{T}} = (\hat{\mathbf{T}} \cdot \mathbf{G})\hat{\mathbf{T}}$. Consequently, every extracted off-axis mode has the same tangential wave vector as its corresponding incident mode. Because this closure condition involves only the fixed grating vectors, it is independent of wavelength. Equation (2.26) therefore defines the directional achromaticity condition of a planar waveguide. This condition does not imply wavelength-independent diffraction efficiency, which remains limited by the spectral and angular selectivity of the volume gratings.

The lateral extent of the outcoupler determines the exit pupil of the system and therefore the eyebox according to Eq. (1.1).

2.2.5 Achievable display field of view

Only modes satisfying the total internal reflection condition in Eq. (2.19) are transported through the waveguide. In the k_x–k_y mode plane, the guided modes occupy the annulus between the cladding-side Ewald circle of radius $k_0 n_\mathrm{clad}$ and the waveguide-side Ewald circle of radius $k_0 n$. In practice, the propagation angle of off-axis modes is additionally limited to θ_max to maintain sufficiently dense output sampling. The usable outer radius is therefore $k_0 n \sin\theta_\mathrm{max}$.

The maximum guidable square field of view is represented by the largest square that fits inside this usable annulus of inner radius $r = k_0 n_\mathrm{clad}$ and outer radius $R = k_0 n \sin\theta_\mathrm{max}$. At the optimum, the nearest side of the square, with edge length $2\Delta k_\mathrm{max}$, is tangent to the inner circle, so that its two farthest vertices are

located at $(r+2\Delta k_{\max}, \pm\Delta k_{\max})$ on the outer circle. The outer-circle constraint $R^2 = (r+2\Delta k_{\max})^2 + \Delta k_{\max}^2$ then gives the positive root

$$\Delta k_{\max} = \frac{k_0}{5}\left(\sqrt{5n^2\sin^2(\theta_{\max}) - n_{\text{clad}}^2} - 2n_{\text{clad}}\right). \tag{2.27}$$

The corresponding guidable horizontal and vertical angular fields of view in air are both

$$\text{HFOV}_{\square}^{\text{air}} = \text{VFOV}_{\square}^{\text{air}} = 2\arcsin(\Delta k_{\max}/k_0). \tag{2.28}$$

A larger rectangular region can, in principle, be fitted into the annulus. In an orthogonal 2D-EBE architecture, however, the expander translates the field in an orthogonal in-plane direction, and the non-square portion of such a rectangle is shifted outside the guided region or into evanescent modes. The square field of view therefore provides the relevant limit for this architecture.

Table 2.1 lists the guidable horizontal angular field of view in air, $\text{HFOV}_{\square}^{\text{air}}$, for several combinations of core and cladding refractive indices. The results show that the guidable field of view is strongly constrained by the available refractive-index contrast. Although high-index substrate materials are available, the average refractive index of typical holographic photopolymers is approximately $n \approx 1.5$ [56]. Holographic materials with $n > 1.5$ are under active investigation [57, 58], as are low-index cladding materials [59].

Table 2.1: Guidable horizontal angular field of view in air, $\text{HFOV}_{\square}^{\text{air}}$, for $\theta_{\max} = 75°$. Values in parentheses correspond to $\theta_{\max} = 90°$.

n_{clad}	$n = 1.5$	$n = 1.7$	$n = 2.0$
1.0	25° (28°)	36° (39°)	52° (56°)
1.2	14° (17°)	25° (28°)	41° (45°)
1.3	8° (11°)	19° (22°)	35° (39°)
1.4	3° (6°)	14° (17°)	30° (34°)

Exploiting the full guidable field of view also requires the holographic couplers to maintain high diffraction efficiency over the corresponding angular bandwidth. Reducing the grating thickness broadens the angular response but, for a fixed refractive-index modulation, reduces the attainable diffraction efficiency, see Eq. (2.16). A sufficiently large modulation Δn is therefore required. Modern high-performance holographic photopolymers provide refractive-index modulations of approximately $\Delta n \approx 0.06$ [56]. Simultaneously achieving a high average refractive index and a large refractive-index modulation remains an active research direction and is important for the broader adoption of holographic waveguide displays [58].

2.3 Waveguide integration in prescription lenses

For consumer adoption, waveguide-display glasses must provide prescription correction while retaining the aesthetics and weight of conventional eyewear. Total device weights above approximately $40\,\mathrm{g}$ are considered uncomfortable for extended wear [60, 61], limiting consumer acceptance.

2.3.1 Regular lens geometry

Prescription frames span a broad range of aperture sizes. Typical examples include the regular *Ray-Ban RX 4340V* glasses and the display-free *Ray-Ban Meta Wayfarer Gen 2* smartglasses, both with an aperture width of $50\,\mathrm{mm}$ and height of $41\,\mathrm{mm}$ [62, 63]. Sunglasses typically have slightly larger apertures. Examples are the *Ray-Ban RB 2184*, with a width of $57\,\mathrm{mm}$ and height of $43\,\mathrm{mm}$ [64], and the *Oakley Flak 2.0 XL*, with a width of $59\,\mathrm{mm}$ and height of $38\,\mathrm{mm}$ [65]. Figure 2.4 compares representative aperture outlines for regular prescription glasses, sunglasses, and sport sunglasses.

The front surface curvature of an eyeglass lens is specified by its base curve, expressed in diopters. For a given base curve, the rear surface is chosen to provide the required prescription power.

Base curve selection affects lens thickness and weight. For myopic prescriptions, stronger base curves generally increase both thickness and weight [26, p. 48]. Stronger base curves can, however, improve off-axis aberration correction [26, p. 74]. Most prescriptions range from $-6\,\mathrm{D}$ to $4\,\mathrm{D}$ [66]. Within this range, base curves of $3\,\mathrm{D}$ to $6\,\mathrm{D}$ are common for myopic prescriptions, whereas $6\,\mathrm{D}$ to $8\,\mathrm{D}$ are common for hyperopic prescriptions [26, p. 76].

In practice, base curve selection is also constrained by frame aesthetics and face-wrap. Sunglasses, and especially sport sunglasses, often wrap more strongly around the face, as illustrated in Fig. 2.5. This typically requires larger base curves, for example approximately $6\,\mathrm{D}$ for standard sunglasses such as the *Ray-Ban RB 2184* and $8.75\,\mathrm{D}$ for sport sunglasses such as the *Oakley Flak 2.0 XL*.

Mechanical stability imposes a minimum lens thickness over the full aperture. Eyeglasses sold in the U.S. market must sustain a ball-drop impact test [67]. In

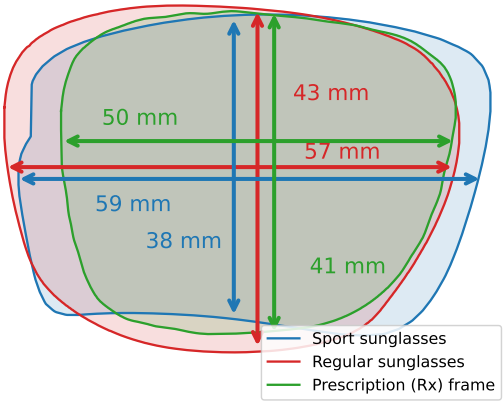


Figure 2.4: Lens aperture outlines, viewed normal to the face-wrap plane, for representative prescription glasses (green), sunglasses (red), and sport sunglasses (blue).

minus lenses, the center is typically the thinnest point. In plus lenses, the edge is typically thinnest. Common minimum thicknesses at these points are $1\,\mathrm{mm}$–$2.2\,\mathrm{mm}$ [26, p. 48]. Safety eyewear requires a minimum lens thickness of $3\,\mathrm{mm}$ [26, p. 48].

2.3.2 Waveguide display prescription correction

Waveguide displays can be prescription-corrected by adding a prescription lens element between the waveguide core and the eye. This so-called pull lens compensates both the real-world view and the display image for the wearer's refractive error. It can provide spherical power only or a combined spherical and cylindrical correction. The same waveguide display can be combined with different pull lenses for user-specific prescriptions.

With existing planar waveguides, finite base curvature requires an additional convex-plano push lens in front of the waveguide. The real-world view then experiences the combined push- and pull-lens powers, whereas the display image experiences only the pull lens. Simultaneous correction of both views therefore requires the waveguide outcoupler to add, in the display path, an optical power

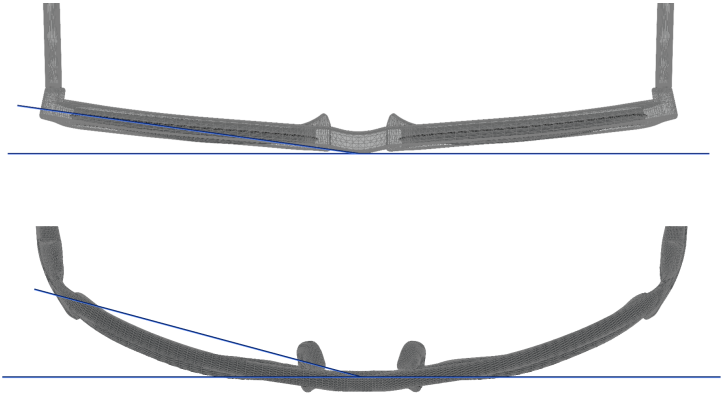


Figure 2.5: Top-view comparison of a standard fashion frame (top) and a high base curve sport frame with large face-wrap angle (bottom).

matching the push lens while leaving the real-world view unaffected [37]. For planar holographic waveguides, this can be implemented by recording the outcoupler with an appropriately focused object beam.

2.3.3 Lens stack thickness and weight

Bonding the push lens, waveguide core, and pull lens without air gaps produces a rigid compound stack that can be assessed as a single element for the U.S. ball-drop impact test [67]. This relaxes the minimum-thickness requirement on the individual push and pull lens elements, enabling a lighter lens stack. We evaluate bonded stacks containing planar, cylindrical, or spherical $0.5\,\mathrm{mm}$-thick waveguide cores for the representative prescription frame at $4\,\mathrm{D}$ base curve ($R_1 \approx 132\,\mathrm{mm}$) and the representative sunglass frame at $6\,\mathrm{D}$ base curve ($R_1 \approx 88\,\mathrm{mm}$) shown in Fig. 2.4, over prescriptions from $-6\,\mathrm{D}$ to $4\,\mathrm{D}$.

The push-lens front radius is fixed by the base curve BC and lens index n via $R_1 = (n-1)/\mathrm{BC}$, with R_1 in meters and BC in diopters. The push-lens rear surface and the pull-lens front surface follow the waveguide-core geometry: planar, cylindrical along the horizontal aperture direction, or spherical. The pull-lens rear radius is determined from the target prescription power using the ABCD matrix of the preceding stack. The derivation is given in Appendix A.3. For cylindrical cores, the buried cylindrical meniscus introduces only a weak astigmatic term, which is neglected in the weight calculation. Eyeglass stack cross-sections are shown in Fig. 2.6 for different waveguide geometries and representative prescriptions of $-6\,\mathrm{D}$ and $4\,\mathrm{D}$.

The push- and pull-lens center thicknesses are chosen as the smallest values that keep every local element thickness above $d_{\min} = 0.5\,\mathrm{mm}$ over the aperture. For the spherical and cylindrical surfaces considered here, the thinnest point lies either at the apex or on the aperture boundary. Details are given in Appendix A.3. Element volumes are obtained by numerical integration of the local thickness over the aperture, also described in Appendix A.3.

The waveguide core is assumed to be polycarbonate, with $n = 1.586$ and $\rho = 1.20\,\mathrm{g\,cm^{-3}}$ and the push and pull lenses are assumed to be Trivex, with $n = 1.53$ and $\rho = 1.11\,\mathrm{g\,cm^{-3}}$ [68]. The resulting weight and minimum stack thickness trends are shown in Fig. 2.7. Representative cross-sections are shown in Fig. 2.6. A single reference lens without a waveguide, with a minimum local thickness of $1\,\mathrm{mm}$, is included for comparison.

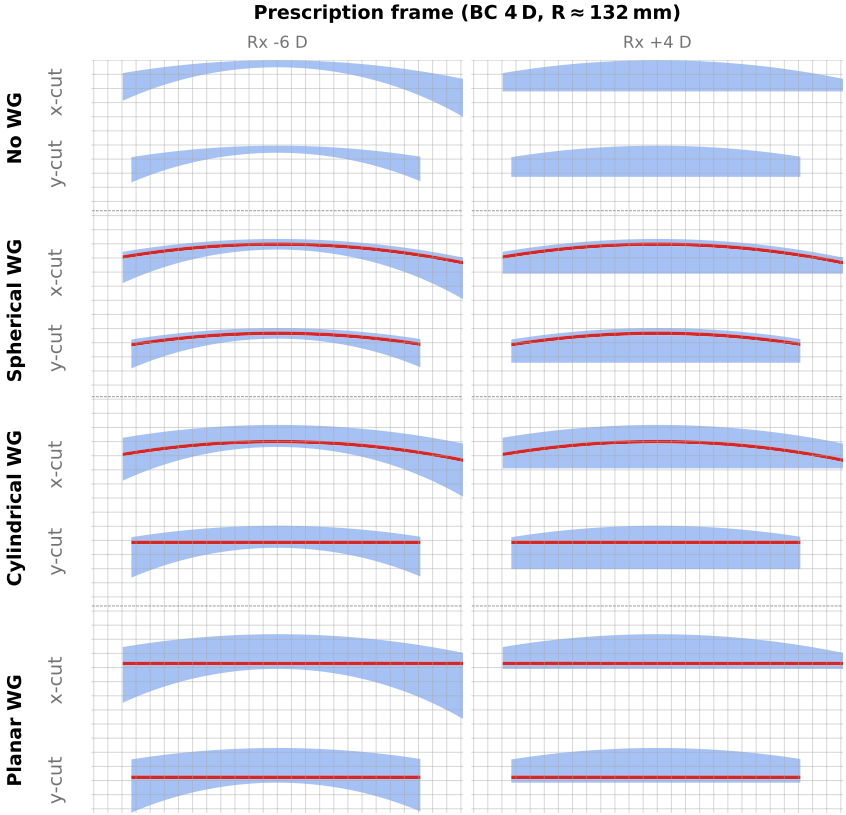


Figure 2.6: Cross-sections (x-cut and y-cut) of lens stacks for a standard prescription frame with base curve $4\,\mathrm{D}$ ($R \approx 132\,\mathrm{mm}$) at prescriptions of $-6\,\mathrm{D}$ (left) and $4\,\mathrm{D}$ (right), without a waveguide and with embedded spherical, cylindrical, and planar waveguides. The waveguide core is shown in red.

Relative to the single-lens reference, the spherical waveguide embedding introduces only a small mass penalty. At the prescription-range endpoints of $-6\,\mathrm{D}$ and $4\,\mathrm{D}$, the mass increase is only about $15\,\%$ to $20\,\%$ for both the regular prescription frame and the sunglass frame. Moreover, this penalty is not intrinsic to the spherical waveguide geometry. The spherical embedding would be essentially weight-neutral relative to a conventional prescription lens if the push and pull elements were made slightly thinner, such that the minimum stack thickness matched that of the single-lens reference.

By contrast, the cylindrical embedding is substantially heavier. In the prescription frame, the mass increases by roughly $80\,\%$ at $-6\,\mathrm{D}$ and $50\,\%$ at $4\,\mathrm{D}$. In the sunglass frame, the corresponding increases are roughly $110\,\%$ and $90\,\%$, respectively. The planar embedding produces the largest penalty for myopic prescriptions, increasing the lens mass by approximately $150\,\%$ in the prescription frame and $200\,\%$ in the sunglass frame at $-6\,\mathrm{D}$. In absolute terms, the planar-waveguide lens

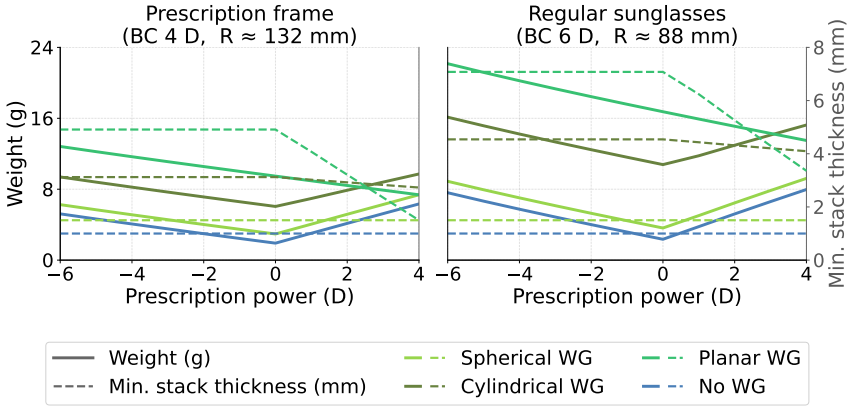


Figure 2.7: Weight (solid lines, left axis) and minimum total stack thickness (dashed lines, right axis) versus prescription power for the prescription lens aperture at BC $4\,\mathrm{D}$ ($R \approx 132\,\mathrm{mm}$, left) and the regular sunglasses lens aperture at BC $6\,\mathrm{D}$ ($R \approx 88\,\mathrm{mm}$, right) shown in Fig. 2.4. Four stack configurations are compared: without waveguide embedding, spherical waveguide embedding, cylindrical waveguide embedding, and planar waveguide embedding.

exceeds $12\,\mathrm{g}$ per lens for $-6\,\mathrm{D}$ myopic users, while the corresponding sunglass lens exceeds $22\,\mathrm{g}$ per lens.

Since total system weight above approximately $40\,\mathrm{g}$ can be uncomfortable during extended wear [60, 61], these results help explain why current waveguide glasses favor flat, low base curvature designs. They also highlight the need for spherical waveguides to enable prescription-compatible AR glasses with the appearance and wearability of regular eyewear.

3 Design of curved waveguide displays using closed modes

Curved waveguide displays could avoid the thickness and weight penalties incurred when planar waveguides are embedded in ophthalmic lenses. However, to our knowledge, no exact design framework has previously been available for spherical waveguides with two-dimensional pupil expansion. We introduce closed geometric mode families for planar, cylindrical, and spherical waveguides based on exact jump maps. These geometric modes make overlapping guided-beam generations single-valued and turn incoupler and expander gratings into analytic translations in mode space. Residual geometric aberrations arise only during image extraction at the outcoupler. For moderate spherical curvatures below $3\,\mathrm{D}$, corresponding to radii above $175\,\mathrm{mm}$, non-sequential ray tracing yields angular errors below 1 arcmin for monochromatic imaging to an $8\,\mathrm{mm}$ pupil over a $22°$ diagonal field, comparable to visual acuity. The framework therefore establishes a route toward thin, prescription-compatible, glasses-like augmented-reality displays.

3.1 Introduction

As discussed in Chapter 1, conventional waveguides are planar, complicating their integration into eyeglass lenses with ophthalmic base curvature. The flat waveguide substrate must be embedded between a convex push lens providing the

Parts of this chapter have been published in: Marvin Feil, Moritz Esslinger, Reinhold Fiess, and Wilhelm Stork, "Closed-mode families for planar, cylindrical, and spherical waveguide display design," *Opt. Express* **34**(14): 26031–26046 (2026). doi: 10.1364/OE.606418 (Editors' pick)

front base curvature and an often concave pull lens providing prescription correction. This increases eyeglass stack thickness and weight as shown in Sec. 2.3.3. Spherically curved waveguides remove this penalty by conforming naturally to ophthalmic base curvature.

However, curvature introduces a fundamental optical difficulty. In a planar waveguide, collimated guided beams remain collimated under repeated total internal reflection and waveguide couplers can be homogeneous gratings. In curved waveguides this translational symmetry is lost: a naively injected collimated beam is generally distorted during propagation, causing imaging aberrations. Blanche, Draper and Zhang *et al.* systematically analyzed curved waveguide displays and proposed paraxial correction schemes for two-dimensional pupil-expanding cylindrical waveguides and one-dimensional spherical waveguides using spatially varying holographic gratings [30, 69, 31, 70]. Their work established the intuition that specially shaped ray bundles can self-reproduce during curved-waveguide propagation within paraxial approximations. Vostrikov *et al.* provided the first rigorous aberration analysis for cylindrical waveguides, deriving the exact cylindrical eigenmodes and coupler expressions [33]. More recently, Weng *et al.* used these eigenmodes for the numerical design of a two-dimensional pupil-expanding cylindrical waveguide [34]. Thus, exact design methods have so far been available only for cylindrical waveguides, not spherical ones.

In this chapter, we show that the eigenbeam concept is too restrictive for spherical waveguides: we prove that no exact, non-paraxial eigenbeams exist for this geometry. Instead, we introduce waveguide jump maps for planar, cylindrical, and spherical geometries and use them to define two mode properties sufficient for waveguide coupler design. The first is *local consistency*, which ensures single-valuedness across overlapping jump generations of one mode. The second is *closure*, which allows grating couplers to operate on a finite field of view of modes. We derive explicit closed and locally consistent mode families for planar, cylindrical, and spherical waveguides. For planar and cylindrical waveguides, they coincide with the eigenmodes known from previous work. We provide a framework for constructing grating couplers from these closed mode expressions, generalizing the planar k-vector diagram to u-mode space. Incouplers and

expanders act as translations in a two-dimensional closed mode space and can therefore be designed analytically for the full field of view. The outcoupler is treated separately because, for curved substrates, the final transformation into collimated air-side beams is not itself a closed operation. We quantify the resulting residual aberrations using non-sequential ray tracing for planar, cylindrical, and spherical waveguides. We demonstrate that spherical waveguides with moderate base curvature can achieve imaging performance comparable to visual acuity. We also investigate light-engine design for closed free-space modes and provide a chromatic-aberration analysis for finite-bandwidth sources.

3.2 Propagation in the waveguide

We describe ray propagation by discrete jump maps and use them to define waveguide modes.

3.2.1 Ray jump maps

Guided rays propagate by consecutive total internal reflections between the inner and outer waveguide surfaces. We describe this propagation on the outer surface S_{out}. A ray is specified by its starting point $\mathbf{p} \in S_{out}$ and outgoing unit direction $\hat{\mathbf{k}}$. One propagation cycle from the outer surface to the inner surface and back defines one *jump*,

$$\tau : (\mathbf{p}, \hat{\mathbf{k}}) \mapsto \big(\tau_p(\mathbf{p}, \hat{\mathbf{k}}), \tau_k(\mathbf{p}, \hat{\mathbf{k}})\big), \tag{3.1}$$

where τ_p gives the next intersection point on S_{out}, and τ_k gives the corresponding outgoing direction after the second reflection. The jump maps are geometry-specific, as shown in Fig. 3.1. For planar, cylindrical, and spherical waveguides, they are a translation, a screw motion, and a rotation.

We attach a surface frame $\{\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}}\}$ to the apex of the outer waveguide shell, with $\hat{\mathbf{z}}$ directed toward the center of curvature. We also use an axis-aligned centered frame $\{\hat{\mathbf{X}}, \hat{\mathbf{Y}}, \hat{\mathbf{Z}}\}$ whose origin is the center of curvature. An outer-surface point

is identified with its apex-frame position vector $\mathbf{p}$. For curved waveguides, the corresponding centered-frame position is $\mathbf{P}(\mathbf{p}) = \mathbf{p} - r\hat{\mathbf{z}}$. Direction vectors have identical components in both frames, whereas position vectors differ by this translation. Unit vectors are denoted by hats, for example, $\hat{\mathbf{P}} = \mathbf{P}/|\mathbf{P}|$. The waveguide thickness is d. For concentric cylindrical and spherical waveguides, the outer radius is r, the inner radius is $r - d = \rho r$, and $\rho = 1 - d/r$. The jump maps are derived in Appendix A.4.

Planar waveguide. With inward surface normal $\hat{\mathbf{z}}$, the jump map is a pure translation,

$$\begin{aligned} \tau_p^{\mathrm{pl}}(\mathbf{p}, \hat{\mathbf{k}}) &= \mathbf{p} + 2d\,\frac{\hat{\mathbf{k}} - (\hat{\mathbf{k}}\cdot\hat{\mathbf{z}})\hat{\mathbf{z}}}{\hat{\mathbf{k}}\cdot\hat{\mathbf{z}}}, \\ \tau_k^{\mathrm{pl}}(\mathbf{p}, \hat{\mathbf{k}}) &= \hat{\mathbf{k}}. \end{aligned} \tag{3.2}$$

Concentric cylindrical waveguide. The jump map is a screw motion,

$$\begin{aligned} \tau_p^{cyl}(\mathbf{p}, \hat{\mathbf{k}}) &= \mathbf{R}_{\hat{\mathbf{y}}}\big(\Gamma(I^{cyl})\big)\mathbf{P}(\mathbf{p}) + r\hat{\mathbf{z}} + \Delta y\,\hat{\mathbf{y}}, \\ \tau_k^{cyl}(\mathbf{p}, \hat{\mathbf{k}}) &= \mathbf{R}_{\hat{\mathbf{y}}}\big(\Gamma(I^{cyl})\big)\hat{\mathbf{k}}, \end{aligned} \tag{3.3}$$

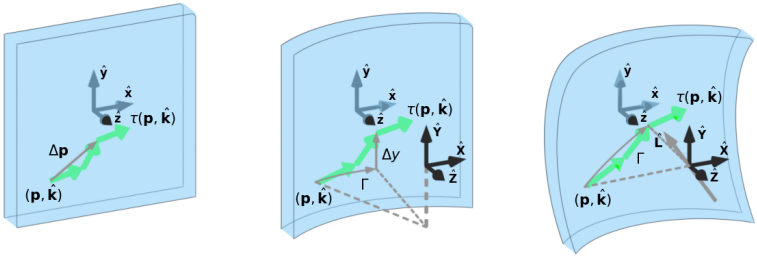


Figure 3.1: The jump map τ transports a ray $(\mathbf{p}, \hat{\mathbf{k}})$ from the outer surface through one round trip back to the outer surface. For planar waveguides, the jump map is a translation. For cylindrical waveguides, it is a screw motion. For spherical waveguides, it is a rotation.

with the rotation matrix $\mathbf{R}$, angular advance $\Gamma(I^{cyl}) = 2\big(\arcsin(\sin I^{cyl}/\rho) - I^{cyl}\big)$, signed meridional angle $I^{cyl} = \operatorname{atan2}(u_\varphi, u_r)$, and axial advance $\Delta y = 2\rho r\,(u_y/u_\varphi)\sin\big(\Gamma(I^{cyl})/2\big)$. Here, $u_y = \hat{\mathbf{y}}\cdot\hat{\mathbf{k}}$, $u_r = \hat{\mathbf{r}}\cdot\hat{\mathbf{k}}$, and $u_\varphi = \hat{\boldsymbol{\varphi}}\cdot\hat{\mathbf{k}}$ with cylinder axis $\hat{\mathbf{y}}$, inward radial normal $\hat{\mathbf{r}}(\mathbf{p})$, and meridional tangent $\hat{\boldsymbol{\varphi}}(\mathbf{p})$, see Appendix A.4.2 Eq. (A.23).

Concentric spherical waveguide. The jump map is a pure rotation about the ray's conserved angular momentum axis $\mathbf{L} = \mathbf{P}(\mathbf{p}) \times \hat{\mathbf{k}}$

$$\begin{aligned}\tau_p^{sph}(\mathbf{p}, \hat{\mathbf{k}}) &= \mathbf{R}_{\hat{\mathbf{L}}}\big(\Gamma(I^{sph})\big)\mathbf{P}(\mathbf{p}) + r\hat{\mathbf{z}},\\ \tau_k^{sph}(\mathbf{p}, \hat{\mathbf{k}}) &= \mathbf{R}_{\hat{\mathbf{L}}}\big(\Gamma(I^{sph})\big)\hat{\mathbf{k}},\end{aligned} \tag{3.4}$$

with the same $\Gamma(I^{sph})$ as for the cylinder but with incidence angle $I^{sph} = \arccos\Big(-\hat{\mathbf{P}}(\mathbf{p})\cdot\hat{\mathbf{k}}\Big)$.

3.2.2 Beam propagation and rayfield description

We describe a beam launched on the outer waveguide surface S_{out} as a rayfield $\hat{\mathbf{k}} : \Omega \to S^2$ that assigns an outgoing geometric ray direction $\mathbf{p} \mapsto \hat{\mathbf{k}}(\mathbf{p})$ to each point $\mathbf{p}$ of its footprint $\Omega \subset S_{out}$. Beam propagation is obtained by applying the jump map τ pointwise to this rayfield. A rayfield is *guided* if its local incidence angle exceeds the critical angle $\arcsin(1/n)$ throughout the footprint, where n is the waveguide refractive index. This condition is preserved in all considered waveguides because $\hat{\mathbf{k}}(\mathbf{p})\cdot\hat{\mathbf{N}}(\mathbf{p}) = \tau_k(\mathbf{p},\hat{\mathbf{k}})\cdot\hat{\mathbf{N}}\big(\tau_p(\mathbf{p},\hat{\mathbf{k}})\big)$ as follows from Eqs. (3.2)–(3.4).

3.2.3 Waveguide modes and families

A guided beam generates a sequence of generations under repeated jumps. We refer to the complete sequence of generations as a *mode*. A mode is fully determined by the rayfield of its launch beam, $\hat{\mathbf{k}}(\mathbf{p})$, since each subsequent rayfield $\hat{\mathbf{k}}^{(n)}(\mathbf{p})$ is obtained by applying the jump map to the previous generation.

In a waveguide display, each image point is encoded by one guided mode. Transporting a finite two-dimensional field of view therefore requires a continuous two-parameter family of modes,

$$\hat{\mathbf{k}}(\mathbf{p}; \mathbf{u}), \tag{3.5}$$

where $\mathbf{u} = (u_1, u_2)^T \in U \subset \mathbb{R}^2$ labels the transported image points.

3.3 Closed and locally consistent modes

We now introduce two mode properties required for grating-coupler design: local consistency and closure. Local consistency removes the multi-valuedness caused by overlapping jump generations, whereas closure ensures that one grating can redirect a finite field of view of modes. The discussion is geometry-independent. Explicit mode families are derived in Sec. 3.4.

3.3.1 Removing multi-valuedness: local consistency

Designing coupler gratings in curved waveguides is nontrivial even for the central image point. Several beam generations of this central mode interact with the same expander and outcoupler grating. In curved waveguides, these generations generally have distinct rayfields because repeated reflections deform the propagating beam. Moreover, in typical waveguide designs the jump length is smaller than the beam diameter to achieve high pupil fill. Therefore, the footprints of the n-th, $(n+1)$-th, and higher generations overlap.

The grating must diffract all overlapping generations as intended. However, if different generations arrive at the same grating position with different local ray directions, a single local tangential grating vector can correctly redirect only one of them. The grating-design problem is well-posed if all overlapping generations of a mode share the same local ray direction. In that case, the overlapping generations can be represented by a single extended rayfield $\hat{\mathbf{k}}(\mathbf{p})$, rather than by separate

rayfields $\hat{\mathbf{k}}^{(n)}(\mathbf{p})$. The grating can then redirect the combined incident beam into any prescribed output rayfield. In holographic grating recording, the combined beam acts as the reference wave. This motivates the following formal definition.

Definition: locally consistent mode. A guided rayfield $\hat{\mathbf{k}} : \Omega \to S^2$ is called *locally consistent* if it can be extended to a larger domain on S_{out} such that

$$\hat{\mathbf{k}}\Big(\tau_p\big(\mathbf{p}, \hat{\mathbf{k}}(\mathbf{p})\big)\Big) = \tau_k\big(\mathbf{p}, \hat{\mathbf{k}}(\mathbf{p})\big), \qquad \forall\, \mathbf{p} \in \Omega. \tag{3.6}$$

Thus, the ray direction predicted by the extended field at the arrival point agrees with the outgoing direction produced by one jump. The same rayfield therefore describes all overlapping jump generations of the mode. Local consistency is precisely the condition that removes the multi-valuedness of the local grating-design problem.

3.3.2 Supporting finite field of view: closure under tangential grating redirection

A locally consistent mode $\hat{\mathbf{k}}(\mathbf{p}; \mathbf{u}^{(1)})$ encoding the central image point must be redirected by the expander into another locally consistent mode $\hat{\mathbf{k}}(\mathbf{p}; \mathbf{u}^{(2)})$, so that the redirected mode remains single-valued at the outcoupler. By tangential wave vector matching, this is achieved by choosing the tangential expander grating vector $\mathbf{G}^{\mathrm{e}}_T = \Big(\hat{\mathbf{T}} \cdot \mathbf{G}^{\mathrm{e}}\Big)\,\hat{\mathbf{T}} = \mathbf{G}^{\mathrm{e}}_{\parallel} = \mathbf{G}^{\mathrm{e}} - \Big(\hat{\mathbf{N}} \cdot \mathbf{G}^{\mathrm{e}}\Big)\,\hat{\mathbf{N}}$ as the tangential difference between the two modes,

$$\mathbf{G}^{\mathrm{e}}_T(\mathbf{p}) = k_{\mathrm{d}} \left(\hat{\mathbf{k}}_{\parallel}(\mathbf{p}, \mathbf{u}^{(2)}) - \hat{\mathbf{k}}_{\parallel}(\mathbf{p}, \mathbf{u}^{(1)})\right), \tag{3.7}$$

where λ_{d} is the mode wavelength for which the grating is designed, $k_{\mathrm{d}} = k(\lambda_{\mathrm{d}})$ is the corresponding wavenumber, and $\hat{\mathbf{k}}_{\parallel} = \hat{\mathbf{k}} - \Big(\hat{\mathbf{N}} \cdot \hat{\mathbf{k}}\Big)\,\hat{\mathbf{N}}$ the tangential rayfield with the local surface normal $\hat{\mathbf{N}}(\mathbf{p})$. For a finite field of view, the same expander must also redirect any other locally consistent incident mode $\hat{\mathbf{k}}(\mathbf{p}; \mathbf{u}^{(3)})$ into another locally consistent mode $\hat{\mathbf{k}}(\mathbf{p}; \mathbf{u}^{(4)})$. Tangential momentum matching

gives $\mathbf{k}_\parallel(\mathbf{p};\mathbf{u}^{(4)}) = \mathbf{k}_\parallel\left(\mathbf{p},\mathbf{u}^{(3)}\right) + \mathbf{G}^{\mathrm{e}}_\parallel(\mathbf{p}) = \mathbf{k}_\parallel\left(\mathbf{p},\mathbf{u}^{(3)}\right) + \mathbf{k}_\parallel\left(\mathbf{p},\mathbf{u}^{(2)}\right) - \mathbf{k}_\parallel\left(\mathbf{p},\mathbf{u}^{(1)}\right)$. For waveguide design we seek a simple mode parametrization in which the diffraction action is additive in mode space $\mathbf{u}^{(4)} = \mathbf{u}^{(3)} + \mathbf{u}^{(2)} - \mathbf{u}^{(1)}$. This requirement motivates closure.

Definition: closed mode family. Let $\mathcal{F} = \{\hat{\mathbf{k}}(\mathbf{p},\mathbf{u})\}$ with $\mathbf{u} \in U \subset \mathbb{R}^2$ be a family of locally consistent rayfields describing waveguide modes. We call $\mathcal{F}$ *closed under tangential grating redirection* if, for any three modes $\hat{\mathbf{k}}(\mathbf{p},\mathbf{u}^{(1)}), \hat{\mathbf{k}}(\mathbf{p},\mathbf{u}^{(2)}), \hat{\mathbf{k}}(\mathbf{p},\mathbf{u}^{(3)}) \in \mathcal{F}$, the tangentially redirected field defined by

$$\begin{aligned}&\hat{\mathbf{k}}_\parallel\left(\mathbf{p},\mathbf{u}^{(3)} + \mathbf{u}^{(2)} - \mathbf{u}^{(1)}\right) = \\ &\hat{\mathbf{k}}_\parallel\left(\mathbf{p},\mathbf{u}^{(3)}\right) + \hat{\mathbf{k}}_\parallel\left(\mathbf{p},\mathbf{u}^{(2)}\right) - \hat{\mathbf{k}}_\parallel\left(\mathbf{p},\mathbf{u}^{(1)}\right)\end{aligned} \tag{3.8}$$

corresponds, after restoring the normal component ($\|\hat{\mathbf{k}}\| = 1$), to another member $\hat{\mathbf{k}}\big(\mathbf{p},\mathbf{u}^{(3)}+\mathbf{u}^{(2)}-\mathbf{u}^{(1)}\big) \in \mathcal{F}$ wherever the resulting field remains non-evanescent ($\|\hat{\mathbf{k}}_\parallel\| < 1$).

A closed mode is specified by its mode coordinate $\mathbf{u}$ and wavelength λ. The wavelength does not affect the normalized rayfield $\hat{\mathbf{k}}_\parallel$, but it scales the wave vector component, $\mathbf{k}_\parallel = k(\lambda)\hat{\mathbf{k}}_\parallel$. Closure implies that *closed gratings* designed from two closed modes $\mathbf{u}^{(1)}, \mathbf{u}^{(2)}$ at wavelength λ_d, with design wavenumber $k_\mathrm{d} = k(\lambda_\mathrm{d})$, act on the full closed family at any wavelength as mode-space translations

$$\begin{aligned}(\mathbf{u},\lambda) &\mapsto (\mathbf{u} + \Delta\mathbf{u}(\lambda),\lambda) \\ &= \left(\mathbf{u} + \frac{k_\mathrm{d}}{k(\lambda)}(\mathbf{u}^{(2)} - \mathbf{u}^{(1)}),\lambda\right).\end{aligned} \tag{3.9}$$

Thus, any incident closed mode $\mathbf{u}$ is mapped to another closed mode $\mathbf{u} + \Delta\mathbf{u}$. This property makes a single expander usable for the entire field of view, rather than only for one image point.

3.4 Explicit closed mode families

As shown in Appendix A.5, a closed mode family is closed if and only if its tangential rayfield is affine in the mode coordinates $\mathbf{u} = (u_1, u_2)^T$. For the planar, cylindrical, and spherical geometries considered here, the relevant families can be written in the linear form

$$\begin{aligned}\hat{\mathbf{k}}_{\parallel}^{g}(\mathbf{p};\mathbf{u}) &= \mathbf{A}_g(\mathbf{p})\mathbf{u},\\ \hat{\mathbf{k}}^{g}(\mathbf{p};\mathbf{u}) &= \mathbf{A}_g(\mathbf{p})\mathbf{u} + \sqrt{1-\left|\mathbf{A}_g(\mathbf{p})\mathbf{u}\right|^2}\,\hat{\mathbf{N}}_g(\mathbf{p}).\end{aligned} \tag{3.10}$$

Here $\hat{\mathbf{N}}_g(\mathbf{p})$ is the inward surface normal specific to geometry g, and $\mathbf{A}_g(\mathbf{p})$ is a 3×2 tangent-valued matrix mapping mode coordinates to local tangential ray components. The linearity of Eq. (3.10) in $\mathbf{u}$ makes closure directly explicit, while local consistency is verified in Appendix A.6.1.

3.4.1 Planar family

For the planar waveguide, $\mathbf{A}_{\mathrm{pl}}$ is the fixed tangent basis $(\hat{\mathbf{x}}, \hat{\mathbf{y}})$, and the inward surface normal is $\hat{\mathbf{z}}$. Equation (3.10) gives the usual collimated planar modes

$$\hat{\mathbf{k}}^{\mathrm{pl}}(\mathbf{p}; u_1, u_2) = u_1\hat{\mathbf{x}} + u_2\hat{\mathbf{y}} + \sqrt{1-u_1^2-u_2^2}\,\hat{\mathbf{z}}. \tag{3.11}$$

3.4.2 Cylindrical family

For the cylindrical waveguide, $\mathbf{A}_{\mathrm{cyl}}$ is formed by the cylindrical unit vectors $(\hat{\boldsymbol{\varphi}}(\varphi), \hat{\mathbf{y}})$, and the inward surface normal is the radial unit vector $\hat{\mathbf{r}}(\varphi)$, see Appendix A.4.2 Eq. (A.23). Equation (3.10) gives

$$\hat{\mathbf{k}}^{\mathrm{cyl}}(\mathbf{p}; u_1, u_2) = u_1\hat{\boldsymbol{\varphi}}(\varphi) + u_2\hat{\mathbf{y}} + \sqrt{1-u_1^2-u_2^2}\,\hat{\mathbf{r}}(\varphi). \tag{3.12}$$

The angle of incidence of these modes is constant because $|\hat{\mathbf{k}}_{\parallel}^{\mathrm{cyl}}| = |\mathbf{u}|$ is independent of position.

3.4.3 Spherical family

For the spherical waveguide, $\mathbf{A}_{\mathrm{sph}}$ is not the local spherical-coordinate tangent basis, because fixed spherical components would not be locally consistent under the spherical jump map. Instead,

$$\mathbf{A}_{\mathrm{sph}}(\mathbf{p}) = \frac{\hat{\mathbf{P}}(\mathbf{p})\hat{\mathbf{P}}(\mathbf{p})^T - \mathbf{I}_3}{\hat{\mathbf{P}}(\mathbf{p}) \cdot \hat{\mathbf{z}}} \big[\hat{\mathbf{x}}\ \hat{\mathbf{y}}\big]. \tag{3.13}$$

The spherical rayfield domain is naturally foliated into great circles, which we call leaves. Each leaf, labeled by the scalar ℓ, is the set of surface points satisfying $\hat{\mathbf{P}} \cdot \mathbf{b}_\ell = 0$, with normal vector $\mathbf{b}_\ell = \mathbf{a} \times \hat{\mathbf{z}} + \ell\hat{\mathbf{z}}$ where $\mathbf{a} = u_1\hat{\mathbf{x}} + u_2\hat{\mathbf{y}}$. On such a leaf, Eqs. (3.10) and (3.13) give $\hat{\mathbf{k}}_{\parallel} = \mathbf{b}_\ell \times \hat{\mathbf{P}}$. Therefore, both the rayfield's angular momentum $\mathbf{L} = \mathbf{P} \times \hat{\mathbf{k}} = r\mathbf{b}_\ell$ and the normal ray direction component are constant on the leaf. By Eq. (3.4), the spherical jump angular advance is also constant. Since the spherical jump rotates each ray about the conserved angular momentum axis, it maps the ray to another point on the same leaf with the direction prescribed by the same rayfield. The spherical family is therefore locally consistent, see Appendix A.6.1 for a detailed proof.

At the apex, $\hat{\mathbf{k}}_{\parallel}^{\mathrm{sph}}(x = 0, y = 0; \mathbf{u}) = u_1\hat{\mathbf{x}} + u_2\hat{\mathbf{y}}$. Thus u_1 and u_2 are the Cartesian tangential direction cosines at the apex, and remain approximately equal to them over a sufficiently small aperture around it.

3.4.4 Free-space and guided modes

Only non-evanescent modes are relevant, requiring $|\hat{\mathbf{k}}_{\parallel}^{g}(\mathbf{p}; \mathbf{u})| = |\mathbf{A}_g(\mathbf{p})\mathbf{u}| < 1$ throughout the aperture. In curved waveguides, rays must also be steep enough to reach the inner shell. For planar and cylindrical families, non-evanescence

is equivalent to $|\mathbf{u}| < 1$. For the cylindrical family, inner-shell intersection is guaranteed by the stricter sufficient bound $|\mathbf{u}| < \rho$. For the spherical family, $|\hat{\mathbf{k}}_{\parallel}^{g}|$ increases away from the apex. Therefore, for a radial aperture of width W centered at the apex, $|\mathbf{u}| < \rho\sqrt{1-(W/2r)^2}$ ensures both inner-shell intersection and non-evanescence.

Guided modes additionally require total internal reflection at every aperture point $\left|\hat{\mathbf{k}}_{\parallel}^{g}(\mathbf{p};\mathbf{u})\right| > 1/n$. For planar and cylindrical waveguides, this reduces to $|\mathbf{u}| > 1/n$. The same threshold applies to the spherical family when the aperture includes the apex. Modes below this threshold are free-space modes. They are not guided and can enter or leave the waveguide. Both mode types are used in the proposed design strategy, see Sec. 3.5.

3.4.5 Comparison with previous eigenbeam concepts

Previous curved-waveguide work introduced self-reproducing ray bundles as building blocks for grating-coupler construction. Blanche and Draper proposed cylindrically focused object beams for cylindrical holographic waveguides using a paraxial Coddington-equation argument [30, 31]. Volkov *et al.* later referred to such self-reproducing bundles as eigenbeams [32], and Vostrikov *et al.* derived exact non-paraxial eigenbeams for cylindrical waveguides [33]. We formalize this concept using the jump-map notation. **Definition: eigenbeams and global rigidity.** A guided beam described by a rayfield $\hat{\mathbf{k}} : \Omega \to S^2$ is called an *eigenbeam*, or *globally rigid* under the jump map, if there exists one rigid motion $g(\mathbf{p}) = \mathbf{R}\mathbf{p} + \mathbf{t}$, with $\mathbf{R} \in SO(3)$ and $\mathbf{t} \in \mathbb{R}^3$, such that

$$\begin{aligned} \tau_p\big(\mathbf{p}, \hat{\mathbf{k}}(\mathbf{p})\big) &= g(\mathbf{p}), \\ \tau_k\big(\mathbf{p}, \hat{\mathbf{k}}(\mathbf{p})\big) &= \mathbf{R}\,\hat{\mathbf{k}}(\mathbf{p}) \end{aligned} \tag{3.14}$$

for all $\mathbf{p} \in \Omega$. Thus, after one jump, the entire beam is recovered up to one global rigid displacement and rotation.

The planar and cylindrical closed modes in Eqs. (3.11) and (3.12) are eigenbeams, as shown in Appendix A.6.3. The cylindrical family coincides with the exact eigenbeams of Vostrikov *et al.* [33], and its small-aperture limit recovers the cylindrically focused beams of Blanche and Draper, as discussed in Appendix A.7. We now show that, for spherical waveguides, no exact guided eigenbeam with finite two-dimensional footprint exists. The paraxial spherical eigenbeams discussed by Zhang and Blanche [70] do not contradict this result. They are small-aperture approximations and coincide in that limit with the closed spherical modes derived here.

Theorem. Let $\Omega \subset S_{out}$ be a finite footprint with nonzero two-dimensional extent on the outer surface S_{out} of a concentric spherical waveguide, and let $\hat{\mathbf{k}} : \Omega \to S^2$ be a smooth guided rayfield. Then $\hat{\mathbf{k}}$ cannot describe a globally rigid eigenbeam in the sense of Eq. (3.14).

Proof. Assume a two-dimensional footprint $\Omega \subset S_{out}$ for the eigenbeam and consider a ray launched at position $\mathbf{p} \in \Omega$, in the centered frame $\mathbf{P} = \mathbf{p} - r\hat{\mathbf{z}}$, in direction $\hat{\mathbf{k}} = \hat{\mathbf{k}}(\mathbf{p})$. Since a jump in the waveguide maps the eigenbeam footprint onto the same sphere, the rigid motion Eq. (3.14) must be a pure rotation: for every launch ray after one jump $\mathbf{P}^{(1)} = \mathbf{R}\mathbf{P}$ and $\hat{\mathbf{k}}^{(1)} = \mathbf{R}\hat{\mathbf{k}}$. The ray's angular momentum after one jump is $\mathbf{L}^{(1)} = \mathbf{P}^{(1)} \times \hat{\mathbf{k}}^{(1)} = (\mathbf{R}\mathbf{P}) \times (\mathbf{R}\hat{\mathbf{k}}) = \mathbf{R}\mathbf{L}$. Spherical propagation conserves $\mathbf{L} = \mathbf{P} \times \hat{\mathbf{k}}$: straight propagation leaves $\mathbf{r} \times \hat{\mathbf{k}}$ unchanged, and reflection at a concentric sphere changes only the radial component of $\hat{\mathbf{k}}$. Hence $\mathbf{R}\mathbf{L} = \mathbf{L}$ for every ray launched in the footprint. $\mathbf{R} = \mathbf{I}_3$ corresponds to zero angular advance. For guided beams $\mathbf{R} \neq \mathbf{I}_3$ and $\mathbf{L} \neq 0$, thus all rays must have angular momentum parallel to the single rotation axis of $\mathbf{R}$. Since $\mathbf{L} \perp \mathbf{P}$, all launch points must lie in the plane through the sphere center perpendicular to this axis. Its intersection with S_{out} is a great circle, not a two-dimensional footprint. Contradiction. ■

As discussed in the next section, waveguide coupler design does not require eigenbeams. Local consistency and closure are sufficient. These weaker conditions coincide with eigenbeams in planar and cylindrical waveguides, but remain available for spherical waveguides.

3.5 Closed mode waveguide design framework

We now formulate the waveguide-design problem and show how closed mode families provide analytical constructions for the incoupler and expander, before discussing the outcoupler.

3.5.1 Design problem

Figure 3.2 illustrates the operating principle of a diffractive waveguide display independent of waveguide geometry. A light engine provides a two-parameter family of free-space input beams. We first assume a monochromatic design wavelength λ_{d}. Finite-bandwidth sources are discussed in Sec. 3.5.4. Each beam represents one image point. Its intensity encodes the pixel value, and its output direction encodes the angular field coordinate. After refraction into the waveguide, the incoupler diffracts each input beam into a guided mode. The guided modes propagate by total internal reflection, are redirected by the expander, and finally reach the outcoupler. The outcoupler should transform all generations of one

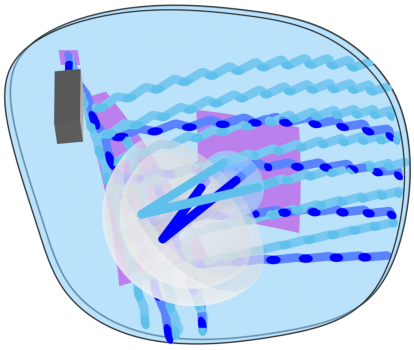


Figure 3.2: The light engine and incoupler encode image points in guided modes. After waveguide transport, the outcoupler collimates these modes, and the eye refocuses them on the retina.

image point into the same collimated air-side beam, so that the eye focuses them onto one retinal point.

Thus, for a prescribed output field of view, the design problem is to determine (i) the free-space input beams generated by the light engine, (ii) the guided modes used for waveguide transport, and (iii) the tangential grating vector fields of the incoupler, expander, and outcoupler, so that each field point exits as a collimated beam with the desired air-side direction.

3.5.2 Solution using closed modes

Closed modes provide both the structured free-space input beams and the guided modes used to encode the image. They also define the incoupler and expander gratings.

We assume reflective diffractive gratings located on the outer waveguide shell, where the closed mode rayfields are defined. This construction can be transferred to transmission gratings.

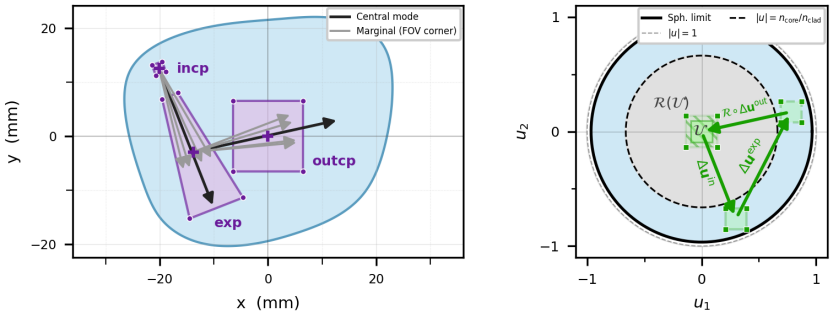


Figure 3.3: Exemplary waveguide cross-section (left) and corresponding $\mathbf{u}$-mode diagram (right). The $\mathbf{u}$-diagram visualizes the grating-induced translations $\Delta\mathbf{u}$.

3.5.2.1 Mode diagrams

Closed mode propagation is fully specified by the single-valued rayfield $\hat{\mathbf{k}}(\mathbf{p};\mathbf{u})$ across the entire eyeglass aperture. Moreover, *closed gratings* constructed from closed modes act in mode space as fixed translations, see Eq. (3.9). This property allows the $\mathbf{k}_{\|}$-vector diagrams used for planar waveguides [22], introduced in Fig. 2.3 of Chapter 2, to be generalized to arbitrary geometry. In planar designs, each guided collimated mode is represented by its in-plane wave vector $\mathbf{k}_{\|}$. Homogeneous gratings translate points in the $\mathbf{k}_{\|}$-plane by their tangential grating vector $\mathbf{G}_T$. The evanescence and total-internal-reflection constraints define the allowed guiding annulus in $\mathbf{k}_{\|}$-space. For curved waveguides, each closed mode is represented by a point in the $\mathbf{u}$-mode diagram and each spatially varying closed grating acts as a fixed translation $\Delta\mathbf{u}(\lambda)$, see Fig. 3.3. The guiding annulus in $\mathbf{u}$-space is derived in Sec. 3.4.4.

3.5.2.2 Incoupler grating

By analogy with the planar waveguide, we propose to refract a spectrum $\mathcal{U}$ of closed free-space modes $\mathbf{u} \in \mathcal{U}$, structured by a suitable light engine, onto the waveguide incoupling grating. We construct the grating from the incident central closed free-space mode and a suitable guided closed mode. The choice of closed free-space mode determines the incidence angle of the light cone. We choose $\mathbf{u} = \mathbf{0}$, fixing it normal to the waveguide surface. The tangential grating vector is

$$\mathbf{G}_T^{\mathrm{i}}(\mathbf{p}) = k_{\mathrm{d}} \left(\hat{\mathbf{k}}_{\|}(\mathbf{p}, \mathbf{u}^{(1)}) - \hat{\mathbf{k}}_{\|}(\mathbf{p}, \mathbf{u} = \mathbf{0}) \right). \tag{3.15}$$

Analogously to planar k-diagrams, this incoupling process can be drawn in the proposed $\mathbf{u}$-diagram (Fig. 3.3). The incoupler shifts the free-space mode spectrum $\mathcal{U}$ at wavelength λ_{d} into the guided spectrum $\mathcal{U} + \Delta\mathbf{u}^{\mathrm{i}} = \mathcal{U} + \mathbf{u}^{(1)}$. The choice of $\mathbf{u}^{(1)}$ fixes the propagation angle and incoupling direction and therefore the expander placement. Figure 3.3 shows a typical waveguide display layout.

3.5.2.3 Expander grating

The expander is a closed grating constructed from the central guided mode $\mathbf{u}^{(1)}$ and the redirected central guided mode $\mathbf{u}^{(2)}$ using Eq. (3.7). The choice of $\mathbf{u}^{(2)}$ sets the central propagation angle and direction after redirection, and therefore the outcoupler placement. Its physical extent must cover the full transported mode spectrum, since off-axis modes separate from the central mode during propagation. In the $\mathbf{u}$-diagram, the expander translates the guided spectrum $\mathcal{U}+\mathbf{u}^{(1)}$ to $\mathcal{U}+\mathbf{u}^{(2)}$ by $\Delta\mathbf{u}^{\mathrm{e}} = \mathbf{u}^{(2)} - \mathbf{u}^{(1)}$, as shown in Fig. 3.3.

3.5.2.4 Outcoupler grating

A natural outcoupler construction is a closed grating that shifts the guided spectrum $\mathcal{U} + \mathbf{u}^{(2)}$ back to a closed free-space spectrum $\mathcal{U}$ around the surface-normal mode. However, this is not generally sufficient for ideal image formation: an eye accommodated to infinity focuses an image point only if all rays associated with that point leave the waveguide as one collimated air-side beam. In curved geometries, closed free-space modes are generally not collimated ray bundles, and they are further deformed by refraction out of the waveguide. Only in the planar case do closed modes coincide with collimated beams, and refraction at the exit surface preserves this collimation. For all other geometries, the outcoupler grating can only be exact for one selected field of view direction.

To construct the exact outcoupler for given air-side direction cosines $\mathbf{c} = (c_x, c_y)$, we assign the corresponding incoupled closed mode coordinate $\mathbf{u}(\mathbf{c}) \in \mathcal{U}$. This arbitrary mapping must be included in the light-engine design to avoid distortion. We choose $\mathbf{u}(\mathbf{c})$ such that this mode's chief ray launched from the outcoupler center refracts into the desired air-side direction $\mathbf{c}$. Placing the outcoupler center at the apex gives, for all geometries, chief-ray launch direction cosines u_1 and u_2. After refraction, the air-side direction cosines satisfy $c_x = \eta_x^g(d, n, r, \mathbf{u})u_1$ and $c_y = \eta_y^g(d, n, r, \mathbf{u})u_2$, with $\mathbf{u} = \mathbf{u}(\mathbf{c})$. The geometry-specific factors η_x^g, η_y^g are derived in Appendix A.8.1 and reduce to n for vanishing curvature.

In Appendix A.8.2 we construct the rayfield $\hat{\mathbf{k}}^{\mathcal{R},g}(\mathbf{p};\mathbf{c},d,r,n)$ by reverse tracing the collimated air-side beam with direction $\mathbf{c}$ to the waveguide outer surface. By reciprocity, the grating

$$\mathbf{G}_T^{\mathrm{o}}(\mathbf{p},\mathbf{c}) = k_{\mathrm{d}}\left(\hat{\mathbf{k}}_{\parallel}^{\mathcal{R},g}(\mathbf{p};\mathbf{c},d,r,n) - \hat{\mathbf{k}}_{\parallel}(\mathbf{p};\mathbf{u}(\mathbf{c})+\mathbf{u}^{(2)})\right), \tag{3.16}$$

perfectly outcouples the incident mode $\mathbf{u}(\mathbf{c})+\mathbf{u}^{(2)}$ into a collimated beam with direction cosines $\mathbf{c}$. For finite-field operation, the implemented outcoupler must compromise between the ideal single-field gratings, as first proposed by Vostrikov *et al.* for cylindrical waveguides [33]. We perform this optimization numerically for the exemplary waveguide in Sec. 3.6.2.

3.5.3 Homogeneity and efficiency optimization

The preceding construction fixes only the tangential grating vector fields required for high-fidelity image transport. These are readily converted to local grating orientation and period. Mode-specific efficiency optimization for eyebox and field of view homogeneity remains an implementation-specific design problem. For volume holographic gratings, the normal component G_N and index modulation can be optimized independently of the prescribed $\mathbf{G}_T$ [47]. For surface-relief gratings, local slant and groove shape provide efficiency degrees of freedom at fixed period and orientation.

3.5.4 Polychromatic image transport

Equation (3.9) shows that closed incoupler and expander gratings remain closed operations for all wavelengths, not only for the design wavelength λ_{d}. Thus, if the light engine provides closed free-space modes $(\mathbf{u},\lambda)\in\mathcal{U}\times\mathcal{S}$, all wavelengths in the spectrum $\mathcal{S}$ still generate single-valued wave vector fields $\mathbf{k}(\mathbf{p},\mathbf{u},\lambda) = k(\lambda)\hat{\mathbf{k}}(\mathbf{p};\mathbf{u})$ at the outcoupler. This requires an achromatic light engine but introduces no additional waveguide-propagation error.

The limiting element is again the outcoupler. In planar waveguides, the dispersive shifts of incoupler, expander, and outcoupler cancel, yielding achromatic image transport, as discussed in Sec. 2.2.4. The same cancellation would occur in $\mathbf{u}$-space if the outcoupler were an ideal closed grating. In curved waveguides, however, the outcoupler requires collimation corrections that are not closed mode translations. These corrections break the wavelength cancellation and introduce chromatic pointing and collimation errors.

In Sec. 3.6.5, we optimize the outcoupler for a finite-bandwidth green source and quantify the residual chromatic aberrations. For strongly curved substrates or broad visible bandwidths, a single outcoupler grating is unlikely to provide good imaging performance across the full spectrum. Wavelength-selective volume holographic gratings may then be used to extract different spectral bands separately.

3.6 Waveguide design studies

We now evaluate the proposed design framework for the two-dimensional pupil-expanding architecture shown in Fig. 3.3. We quantify outcoupler aberrations, system-level imaging performance, curvature dependence, chromatic sensitivity, and light-engine feasibility.

3.6.1 Exemplary architecture

We consider a waveguide of thickness $d = 0.5\,\mathrm{mm}$ and refractive index $n = 1.51$. The incoupled central mode is chosen as $\mathbf{u}^{(1)} = (0.3, -0.7622)$, which corresponds to downward propagation with a slight lateral tilt and places the transported field well inside the guided-mode annulus. The expander redirects the central mode to $\mathbf{u}^{(2)} = (0.78, 0.1715)$. The outcoupler patch is $13 \times 13\,\mathrm{mm}$, and the transported field of view is $16^\circ \times 16^\circ$, represented by $\mathbf{c} \in [-\sin(8^\circ), \sin(8^\circ)]^2$. The design and replay wavelength is $520\,\mathrm{nm}$. Figure 3.3 describes the waveguide

in the apex surface frame $\{\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}}\}$ introduced in Sec. 3.2.1. The outcoupler is centered at $x = y = 0$ as discussed in Sec. 3.5.2.4. We evaluate the same layout for planar, cylindrical, and spherical substrates. Unless curvature is explicitly varied, the curved substrates have a radius $r = 125\,\mathrm{mm}$, corresponding approximately to a $4\,\mathrm{D}$ ophthalmic base curve. All coupler gratings are implemented as volume holographic gratings using the ray-tracing model summarized in Sec. 2.1.6. We assume an index modulation of 0.06, representative of high-performance photopolymers [42], and a grating thickness of $1\,\mu\mathrm{m}$. This implementation affects efficiency but not the imaging errors discussed here. Our non-sequential ray traces use a termination threshold of 10^{-9}, chosen sufficiently low not to affect the reported spot sizes. The displayed spot diagrams are worst-case point-spread estimates that exclude intensity roll-off.

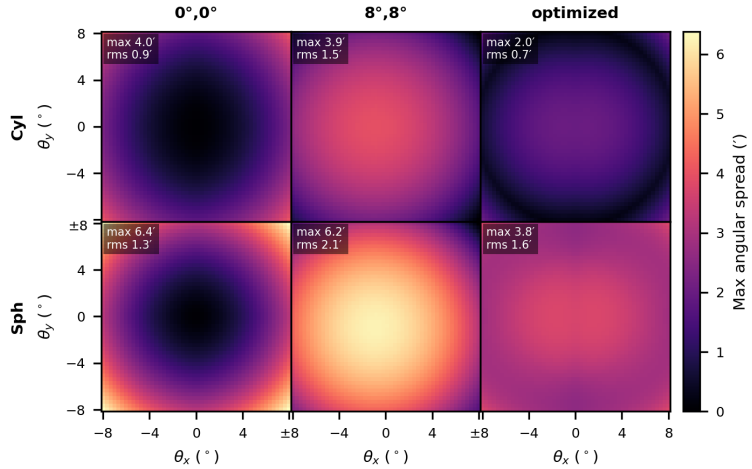


Figure 3.4: Outcoupler optimization for cylindrical (top) and spherical (bottom) waveguides with $r = 125\,\mathrm{mm}$. Panels show the maximum angular collimation error for exact on-axis, exact corner-field, and minimax-optimized gratings.

3.6.2 Outcoupler optimization

The preceding framework shows that, apart from light-engine errors, residual geometric aberrations arise only at the outcoupler. For a pupil-expanding layout with sufficiently small jump length, the closed guided modes fill the full outcoupler area, so the entire outcoupler contributes to each output beam. We therefore evaluate outcoupler performance by sampling the field of view with a 50×50 grid of directions $\mathbf{c}_i$. For each field point, rays are launched from a 25×25 grid of outcoupler positions $\mathbf{p}_j$ according to the incident closed mode rayfield $\hat{\mathbf{k}}_{\parallel}(\mathbf{p}_j; \mathbf{u}(\mathbf{c}_i) + \mathbf{u}^{(2)})$, diffracted by the outcoupler, and refracted out of the waveguide. The maximum angular deviation of any ray from the bundle's centroid defines the collimation error for that field point. This metric is a conservative resolution estimate because a finite pupil samples only part of the beam.

Figure 3.4 compares outcouplers that exactly collimate a single design field point. When the exact grating is constructed for the field center $(0°, 0°)$ or for the corner $(8°, 8°)$, the residual finite-field error reaches about 4 arcmin for the cylindrical waveguide and 6 arcmin for the spherical waveguide. We obtain a finite-field compromise by adding a tangential grating correction,

$$\Delta\mathbf{G}_T(\mathbf{p}) = k_{\mathrm{d}} \Pi^g_{\parallel}(\mathbf{p}) \sum_{1 \le i+j \le N} x^i y^j \left(A_{ij}\hat{\mathbf{x}} + B_{ij}\hat{\mathbf{y}} \right), \tag{3.17}$$

where $\Pi^g_{\parallel}(\mathbf{p}) = \mathbf{I} - \hat{\mathbf{N}}_g(\mathbf{p})\hat{\mathbf{N}}_g(\mathbf{p})^T$ projects onto the local tangent plane, x, y are coordinates of the apex surface frame centered on the outcoupler, and the coefficients A_{ij}, B_{ij} are determined by numerical minimax optimization over the field of view. First-order ($N = 1$) optimization reduces the maximum collimation error to about 2 arcmin for the cylindrical waveguide, for which only A_{10} is nonzero, and 4 arcmin for the spherical waveguide, for which only A_{10} and B_{01} are nonzero, as shown in the right column of Fig. 3.4. Adding higher-order correction terms $N > 1$ did not further reduce the maximum error or introduce additional nonzero coefficients. The first-order grating correction can be interpreted as a phase plate $\phi(x, y) = k_{\mathrm{d}}(A_{10}x^2 + B_{01}y^2)/2$ on the curved surface. This phase plate provides cylindrical optical power in x for the cylindrical

waveguide and astigmatic optical power in x and y for the spherical waveguide. The surface gradient ∇_S of the phase acts as an effective tangential grating vector, $\nabla_S\phi(\mathbf{p}) = \Pi_{\parallel}^{g}(\mathbf{p})\nabla\phi(x, y)$, recovering Eq. (3.17) for the retained first-order terms [71].

3.6.3 Non-sequential waveguide ray trace

We next perform full non-sequential ray traces through the planar, cylindrical, and spherical waveguides with the optimized outcouplers from Sec. 3.6.2. Closed free-space modes for the field center and four field corners are launched immediately before the incoupler at the design wavelength. Each launch beam has a diameter of $1\,\mathrm{mm}$ and is Fibonacci-sampled with 50 rays.

Rays exiting the outcoupler are traced to an eye model centered on the outcoupler with $5\,\mathrm{mm}$ eye relief, see Fig. 3.5. The eye is an ideal sine-condition lens with $f = 19\,\mathrm{mm}$ and an $8\,\mathrm{mm}$ pupil. Similarly, rays are traced to eyes displaced by $\pm 4\,\mathrm{mm}$ to the corners of an $8\,\mathrm{mm} \times 8\,\mathrm{mm}$ eyebox.

We report imaging performance using the conservative maximum geometric angular error, defined as the geometric spot radius, and the corresponding line pairs per degree that are ideally resolved, $1/(4 \cdot \mathrm{spot\ radius})$.

As expected, the planar waveguide gives ray-traced geometric spot radii $\ll 1$ arcsec across the five eyebox and field positions, well below the diffraction-limited Airy radius of the pupil, $1.22 \cdot 520\,\mathrm{nm}/8\,\mathrm{mm} \approx 0.27$ arcmin (56 lp/deg). The cylindrical waveguide retains this precision along the cylinder axis and shows a maximum residual error of 1.3 arcmin ($12\,\mathrm{lp/deg}$) in the curved direction. The spherical waveguide shows a maximum error of 2.4 arcmin ($6\,\mathrm{lp/deg}$) across the field and eyebox. Repeating the analysis with a $4\,\mathrm{mm}$ pupil gives maximum angular errors of 0.7 arcmin ($21\,\mathrm{lp/deg}$) for the cylindrical and 1.3 arcmin ($12\,\mathrm{lp/deg}$) for the spherical waveguide, with the diffraction limit being 0.55 arcmin ($27\,\mathrm{lp/deg}$).

These errors agree with the outcoupler bounds from Sec. 3.6.2, confirming that propagation through the closed mode waveguide itself does not introduce additional geometric aberrations. When the outcoupler is instead designed for one

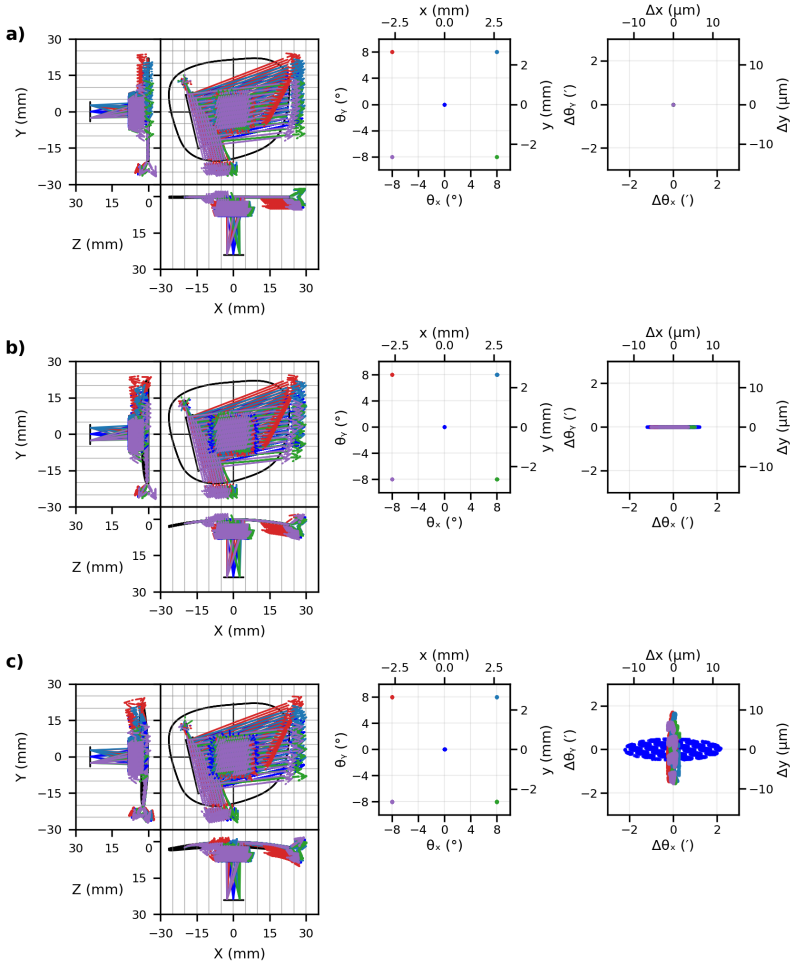


Figure 3.5: End-to-end ray trace for planar, cylindrical, and spherical waveguides. Cross-sections show traced rays. Geometric spot diagrams show absolute and centroid-subtracted retinal distributions after optimized outcoupling.

specific field point, that field point is imaged with a diffraction-limited spot also for cylindrical and spherical substrates.

For comparison, Vostrikov *et al.* reported, for a $4\,\mathrm{mm}$ pupil, a similar maximum outcoupling error of 0.7 arcmin ($21\,\mathrm{lp/deg}$) for a 1D pupil-expanding cylindrical waveguide with a $150\,\mathrm{mm}$ radius of curvature, a $24^\circ \times 24^\circ$ field of view, and an $8\,\mathrm{mm}$-wide eyebox without vertical expansion [33]. Weng *et al.* reported, for a $4\,\mathrm{mm}$ pupil, MTF values above 0.35 at $30\,\mathrm{lp/deg}$ at the field center across a $6\,\mathrm{mm} \times 6\,\mathrm{mm}$ eyebox, for a 2D pupil-expanding cylindrical waveguide designed for a $24^\circ \times 24^\circ$ field of view, a $10\,\mathrm{mm} \times 10\,\mathrm{mm}$ eyebox, and a $100\,\mathrm{mm}$ radius of curvature [34]. Weng *et al.*'s expander grating is recorded from two eigenbeams and therefore coincides for the cylinder with our closed expander construction. The outcouplers of Vostrikov *et al.* and Weng *et al.* are optimized similarly to ours. However, in both works, the incoupler grating performs the mode shaping from collimated free-space beams into guided eigenmodes, whereas in our framework the mode shaping is assigned to the light engine and the incoupler becomes a closed grating. This avoids finite-field incoupling errors, provided that the light engine generates the required closed free-space modes, as discussed in Sec. 3.6.6.

3.6.4 Curvature dependence of outcoupling aberrations

We repeat the monochromatic system ray trace for spherical substrates with radii of curvature between $75\,\mathrm{mm}$ and $225\,\mathrm{mm}$, corresponding to $7\,\mathrm{D}$ to $2\,\mathrm{D}$ base curves, respectively, re-optimizing the outcoupler for each radius. The residual aberration increases monotonically with curvature, as expected from the increasing departure of closed free-space modes from collimated beams. High-base-curve eyewear, such as sports glasses, is therefore more demanding. Figure 3.6 shows that for an $8\,\mathrm{mm}$ pupil maximum angular errors below 1 arcmin (>15 lp/deg) across the field of view are achievable for spherical base curves below approximately $3\,\mathrm{D}$ ($175\,\mathrm{mm}$), whereas stronger curvatures show increasing outcoupling error. For the $175\,\mathrm{mm}$ waveguide and a $4\,\mathrm{mm}$ pupil we find maximum geometric angular errors of 0.5 arcmin (30 lp/deg), below the diffraction limit. These values

exclude light-engine aberrations, treated separately in Sec. 3.6.6 and wave optical considerations discussed in Sec. 3.7. Human visual acuity is commonly associated with a minimum resolvable angle of 1 arcmin, corresponding to a 30 lp/deg vision threshold, with recent work suggesting somewhat higher values [72]. The lp/deg values reported here for the waveguide are conservative resolution estimates derived from the maximum spot radius, not strict cutoff frequencies. Depending on the point-spread function shape and contrast criterion, higher spatial frequencies may still be transmitted with reduced contrast.

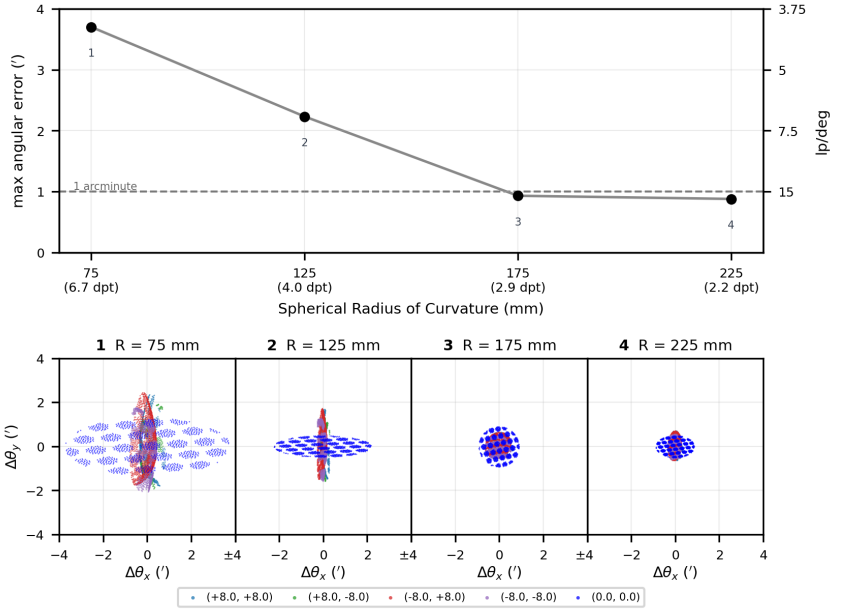


Figure 3.6: Curvature dependence for spherical waveguides: maximum angular deviation (8 mm pupil) versus radius and centroid-subtracted geometric spot diagrams for $r = 75\,\mathrm{mm}$, $125\,\mathrm{mm}$, $175\,\mathrm{mm}$, and $225\,\mathrm{mm}$ after outcoupler optimization. Diopter equivalents calculated for a 1.5 refractive index.

3.6.5 Color aberrations for finite spectral bandwidth

We next quantify the chromatic outcoupler error (8 mm pupil) for a green micro-LED channel centered at $\lambda_d = 520\,\mathrm{nm}$ with $\pm 10\,\mathrm{nm}$ bandwidth, represented by three lines at 510 nm, 520 nm, and 530 nm. As shown in Fig. 3.7, the wavelength-dependent grating shift separates the guided mode spectra in **u**-space. These shifts cancel for ideal closed gratings, but the corrected outcoupler breaks this cancellation and leaves residual wavelength-dependent pointing and collimation errors.

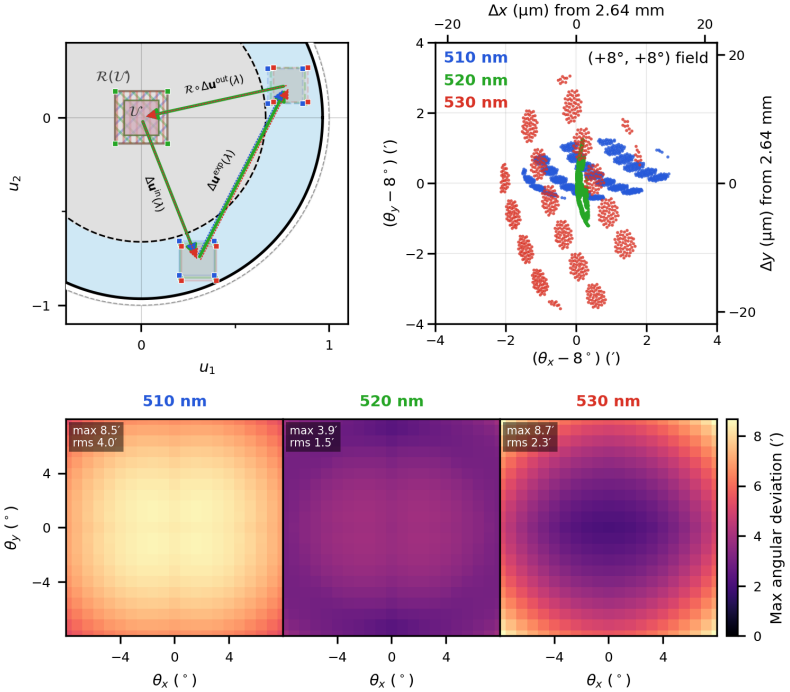


Figure 3.7: Chromatic aberrations for a 20 nm green spectral bandwidth in the spherical waveguide. The plots show wavelength-dependent **u**-space shifts, the corner-field geometric retinal spot, and the absolute angular error across the field of view.

The linear outcoupler correction of Sec. 3.6.2 is re-optimized jointly over all three wavelengths. The bottom row of Fig. 3.7 shows the resulting absolute angular error across the field of view, including both centroid shift and bundle spread. The maximum error increases to about 9 arcmin (2 lp/deg), substantially above the monochromatic case. A full system ray trace for the $(+8°, +8°)$ field point confirms the corresponding wavelength-separated retinal spots. These results illustrate the imaging penalty of finite-bandwidth sources and motivate narrowband laser engines or wavelength-selective outcouplers for curved waveguide designs.

3.6.6 Exemplary light engine design

We now replace the idealized light engine assumed above with an exemplary design for the $r = 125\,\mathrm{mm}$ spherical waveguide. The design transforms the two-dimensional cone of collimated beams from a laser beam scanner (LBS) into

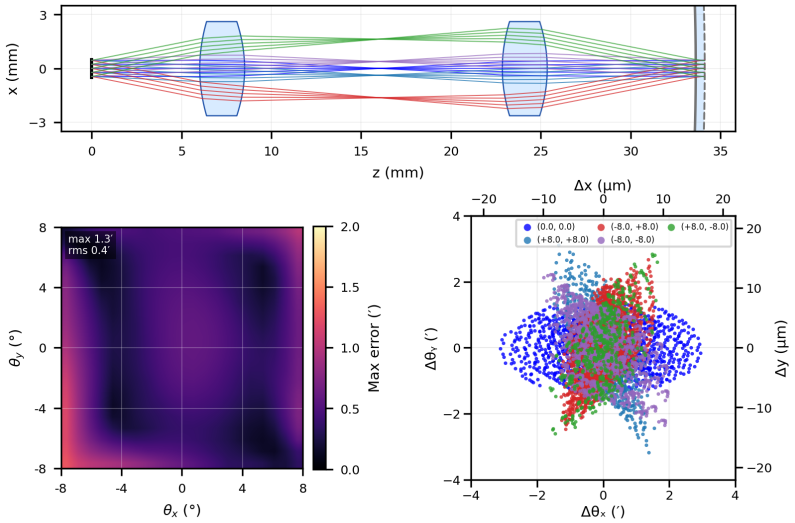


Figure 3.8: Freeform $4f$ beamshaper for the spherical waveguide. The relay transforms collimated scanned beams into closed free-space input modes. The plots show mode error across the field and the resulting geometric retinal spots after full-system tracing.

the structured closed free-space modes required at the incoupler using an approximately unit-magnification $4f$ relay, $f_1 \approx f_2$. The $16° \times 16°$ target field of view is sampled by a 15×15 grid of directions $\mathbf{c}_i$, which are mapped to closed mode coordinates $\mathbf{u}_i = \mathbf{u}(\mathbf{c}_i)$ as described in Sec. 3.5.2.4. The corresponding LBS direction cosines are assigned by the linear relation $\mathbf{s}_i = \mathbf{M}\mathbf{u}_i$, with $\mathbf{M} \in \mathbb{R}^{2\times2}$. For each field sample $\mathbf{c}_i$, the $1\,\mathrm{mm}$ LBS pupil is Fibonacci-sampled with 50 rays launched in direction $\mathbf{s}_i$ and traced sequentially through the relay to the incoupler grating, where we evaluate the maximum angular deviation from the desired closed rayfield $\hat{\mathbf{k}}(\mathbf{p}, \mathbf{u}_i)$. The relay consists of two N-BK7 freeform lenses with two-dimensional Legendre sag profiles of polynomial order 10, corresponding to 66 coefficients per surface. Layout distances, the mapping matrix $\mathbf{M}$, and all surface parameters are jointly optimized to minimize the maximum angular error across the field of view. As shown in Fig. 3.8, the generated input modes deviate from the ideal closed mode rayfields by less than 1.5 arcmin. A full system trace through the relay, incoupler, waveguide, and optimized outcoupler shows that the additional retinal blur beyond the outcoupler-limited case remains of the same order, indicating that closed mode propagation does not substantially amplify light-engine aberrations.

3.7 Wave optical correspondence and scope of the rayfield model

The preceding analysis treats light as geometrical ray bundles, parametrized by rayfields on the outer waveguide surface, as introduced in Sec. 3.2.2. In this section we discuss the limitations of this modeling approach.

3.7.1 Integrability of closed rayfields

In the short-wavelength limit of wave optics, ray directions are normal to the phase fronts, $\hat{\mathbf{k}}(x, y, z) = \nabla\Phi(x, y, z)/k$, where k is the wavenumber. For the

tangential rayfields describing our closed modes on the outer waveguide surface with sag $z_g(x,y)$, this implies $\hat{\mathbf{k}}_\parallel(x,y) = \Pi_g \nabla\phi(x,y)/k$ where $\phi(x,y) = \Phi(x,y,z_g(x,y))$ is the surface phase. Therefore, for a rayfield to admit a wave-optical analogue, its tangential component must be integrable to a scalar phase, i.e., it must be conservative, requiring $0 = \hat{\mathbf{N}}_g \cdot (\nabla \times \hat{\mathbf{k}}_\parallel)$, except at phase singularities [73].[1] Planar and cylindrical closed rayfields satisfy this condition, whereas spherical closed rayfields do not and can therefore only be approximated by conservative rayfields derived from a *regularized* phase. For the $125\,\mathrm{mm}$ spherical waveguide of Sec. 3.6.1, with $16° \times 16°$ field of view and $1\,\mathrm{mm}$ beam diameter, we approximate each closed free-space rayfield incident at the incoupler center (x_i, y_i) by a quadratic phase

$$\phi(x,y,\mathbf{u}) = c_1(\mathbf{u})\bar{x} + c_2(\mathbf{u})\bar{y} + c_3(\mathbf{u})\bar{x}\bar{y} + c_4(\mathbf{u})\bar{x}^2 + c_5(\mathbf{u})\bar{y}^2 \tag{3.18}$$

where $\bar{x} = x - x_i$ and $\bar{y} = y - y_i$. The resulting maximum angular deviation from $\hat{\mathbf{k}}_\parallel^{\mathrm{sph}}(x,y;\mathbf{u})$ is below $0.2'$. Repeating the ray traces of Sec. 3.6.3 with these regularized rayfields increases the maximum spot radius across the five pupil positions and field directions from $2.4'$ to $2.6'$ for the $8\,\mathrm{mm}$ eye pupil. The refractive freeform relay of Sec. 3.6.6 preserves phase integrability of the collimated input by the Malus–Dupin theorem [74], hence, its reported mismatch to the closed rayfields already includes this integrability constraint.

3.7.2 Integrability of spherical grating couplers

To record a holographic grating through interference of two continuous wavefronts, its tangential grating vector field must be the gradient of a scalar Ψ [75]

$$\mathbf{G}_T = \mathbf{k}_\parallel^o - \mathbf{k}_\parallel^r = \Pi_\parallel^g \nabla\left(\phi^o - \phi^r\right) = \Pi_\parallel^g(x,y)\nabla\Psi. \tag{3.19}$$

[1] Since $\nabla \times \nabla\bar{\phi} = 0$ and $\hat{\mathbf{N}}_g \cdot (\nabla \times \hat{\mathbf{N}}_g) = 0$, with $\hat{\mathbf{N}}_g \propto \nabla[z - z_g(x,y)]$, we obtain

$$\hat{\mathbf{N}}_g \cdot (\nabla \times \mathbf{k}_\parallel) = -\hat{\mathbf{N}}_g \cdot \nabla \times [(\hat{\mathbf{N}}_g \cdot \nabla\bar{\phi})\hat{\mathbf{N}}_g] = -(\hat{\mathbf{N}}_g \cdot \nabla\bar{\phi})\hat{\mathbf{N}}_g \cdot (\nabla \times \hat{\mathbf{N}}_g) = 0.$$

The spherical coupler gratings in Eqs. (3.15), (3.7), and (3.16) generally yield non-conservative tangential grating vector fields and therefore cannot be represented exactly by a continuous scalar phase.

Several methods have been developed to realize non-conservative grating vector fields, either by introducing local grating dislocations [76], or by segmenting the grating aperture into several patches with continuous phase but discontinuities at their boundaries [77]. Sufficiently dense segmentation allows accurate approximation of a non-conservative grating vector field. While the thin patch boundaries can be neglected in geometrical optics, diffraction at interfaces and dislocations introduces additional spread beyond the geometric spot radii reported in the previous sections. To avoid dislocations and allow regular recording, non-conservative grating vector fields can be approximated by a regularized phase plate [75]. To each coupler α of the $125\,\mathrm{mm}$ spherical waveguide we fit fifth-order phase plates

$$\Psi^{\alpha}(x,y) = \sum_{\substack{i,j\geq 0 \\ 1\leq i+j\leq 5}} c_{ij}^{\alpha}\,(x-x_{\alpha})^{i}\,(y-y_{\alpha})^{j}\,, \tag{3.20}$$

minimizing $\mathbf{G}_T^{\alpha} - \Pi_{\parallel}^{g}(x,y)\nabla\Psi^{\alpha}$. We then adjust the phase coefficients of all couplers simultaneously by an end-to-end ray-trace optimization. To evaluate the image quality of the regularized spherical couplers, we perform non-sequential ray tracing as described in Sec. 3.6.3. Instead of the exact closed free-space rayfields, we launch the regularized rayfields of Eq. (3.18). We find a maximum geometric spot radius of $2.9'$ across the five field points and five eye positions ($8\,\mathrm{mm}$ pupil), compared to the $2.4'$ achieved with the exact closed gratings. Across 10×10 launched field points over the field of view and 10×10 eye positions distributed uniformly in the eyebox, we find a maximum radius of $3.8'$.

3.8 Conclusion

We have introduced a geometric mode-based design framework for planar, cylindrical, and spherical waveguide displays. The central result is that grating-coupler

design does not require eigenmodes, which do not exist for spherical waveguides. Instead, locally consistent and closed modes are sufficient: consistency removes the multi-valuedness caused by overlapping jump generations at the couplers, while closure makes grating redirection well defined over a finite field of view.

The derived mode families recover the standard planar modes and the known cylindrical eigenmodes, and extend the same analytical construction to spherical waveguides. In closed mode space, incouplers and expanders act as translations, generalizing the planar k-vector diagram to curved substrates. Non-sequential ray tracing confirms that closed mode incoupling, propagation, and expansion do not introduce additional aberrations for the closed target rayfields. The residual imaging error is purely determined by the non-closed outcoupler and by the accuracy with which the light engine generates the target input modes. Both light engine and outcoupler can be optimized independently, without computationally expensive end-to-end waveguide tracing. For moderate ophthalmic spherical base curves, the investigated two-dimensional pupil-expanding architecture achieves geometrical imaging performance close to human visual acuity over a 22° diagonal field of view. We have shown that the geometrically derived spherical closed grating couplers have non-conservative grating-vector fields that are not exactly integrable to scalar phase plates. We presented an approach for approximating these fields with phase plates and discussed the resulting increase in beam spread.

These results provide a route toward thin, prescription-compatible waveguide displays that conform to ophthalmic base curvature without the thickness and weight penalties of embedding planar waveguides in lens stacks. The formalism also provides a basis for more advanced architectures, including crossed gratings, crossed outcouplers, and butterfly expansion schemes. The chromatic analysis indicates a system-level trade-off: narrowband laser engines are naturally compatible with closed mode transport, whereas finite-bandwidth micro-LED engines will require spectral management at the outcoupler, for example through wavelength-selective holographic grating couplers. Very large fields of view may likewise require angle-selective outcouplers. Based on the analytical construction, future work may explore end-to-end regularization strategies for spherical waveguide couplers, which should also discuss wave optical effects.

4 Limits of immersion-free waveguide recording

Recording holographic waveguide gratings in their replay configuration requires at least one recording beam to exceed the critical angle of the hologram–cladding interface. Achieving these angles necessitates edge-lit exposure using prisms and immersion liquids, which slows down display prototyping and is challenging for roll-to-roll mass production. In this chapter, we present analytical equations describing the set of planar holographic gratings that can be recorded without immersion in reflection or transmission using recently proposed methods of wavelength and recording configuration shifting. The limits are formulated in terms of the tangential and normal grating vector components G_T, G_N. We show that, by appropriately choosing the recording polarization, superficial holograms can be mitigated and a large effective index modulation can be achieved. Our results demonstrate that red, green, and blue (RGB) incouplers and outcouplers for typical planar reflective holographic waveguide displays in photopolymer can be recorded without immersion, reducing fabrication complexity for 1D eyebox-expanding waveguides. However, expander grating couplers for practically relevant 2D eyebox-expanding waveguide displays cannot be recorded immersion-free within the sensitivity range of typical photopolymers. We discuss the applicability of these planar limits as local recordability conditions on curved substrates and demonstrate that immersion-free recording is also not applicable to the expander grating of the $125\,\mathrm{mm}$ spherical waveguide display introduced in Chapter 3.

Parts of this chapter have been published in: Marvin Feil, Tobias Wilm, Moritz Esslinger, Reinhold Fiess, and Wilhelm Stork, "Limits of immersion-free recording of holographic waveguide displays," *Opt. Lett.* **50**: 606–609 (2025). doi: 10.1364/OL.545451

4.1 Introduction

Recording waveguide coupler HOEs conventionally in their replay configuration requires at least one recording beam inside the holographic substrate to exceed the critical angle of the waveguide–cladding interface. This necessitates edge-lit recording with prisms optically contacted to the holographic substrate using immersion liquid. However, edge-lit recording complicates prototyping and is challenging for mass production due to the complexities of maintaining a homogeneous immersion film [40] and preventing immersion liquid flow, which can cause phase instabilities [38]. Recently, waveguide gratings have been recorded at wavelengths different from their intended replay wavelength [40, 41], reducing the recording beams' incidence angles and allowing immersion-free recording without edge-lit exposure. Additionally, shifting to a different recording geometry—that is, recording with wave directions different from the desired replay and object waves—has been proposed to reduce the required incidence angles [42]. Immersion-free recording would be of interest for prototyping planar and spherical waveguide displays, simplifying the required exposure system. However, to our knowledge, no systematic investigation has been provided to determine the limits of these immersion-free recording schemes or their applicability to typical waveguide display geometries. In this chapter, we derive analytical expressions to identify which planar holographic gratings are recordable immersion-free using wavelength and recording configuration shifting. We briefly discuss the applicability of these limits to curved substrates.

4.2 Immersion-free recording

Section 2.1 introduced the representation of a prescribed planar holographic grating by its tangential and normal components, $\mathbf{G} = G_T\hat{\mathbf{T}} + G_N\hat{\mathbf{N}}$, with $G_T, G_N > 0$. For this chapter, we consider only the dominant normal hologram shrinkage s and neglect tangential shrinkage s_T. Equation (2.8) gives the irradiance grating components $G_T^{\mathrm{irr}}, G_N^{\mathrm{irr}} > 0$ required to record the prescribed grating.

Equation (2.9) parameterizes the complete family of recording wave vector pairs $\mathbf{k}_r, \mathbf{k}_o$ that produce $\mathbf{G}^{\mathrm{irr}}$ by the recording wavelength λ and configuration parameter $\sigma \in [0, 2\pi)$. Figure 4.1(a) illustrates these configurations for a representative planar waveguide incoupling grating.

4.2.1 General holographic resolution limit

No real-valued wave vector pair may be found from Eq. (2.9) for $G^{\mathrm{irr}} > 2k$, which is a manifestation of the resolution limit $\Lambda^{\mathrm{irr}} \geq \lambda/(2n)$. This constrains the generally recordable grating vectors to

$$\sqrt{G_T^2 + \left(G_N(1-s)\right)^2} \leq \frac{4\pi}{\lambda} n. \tag{4.1}$$

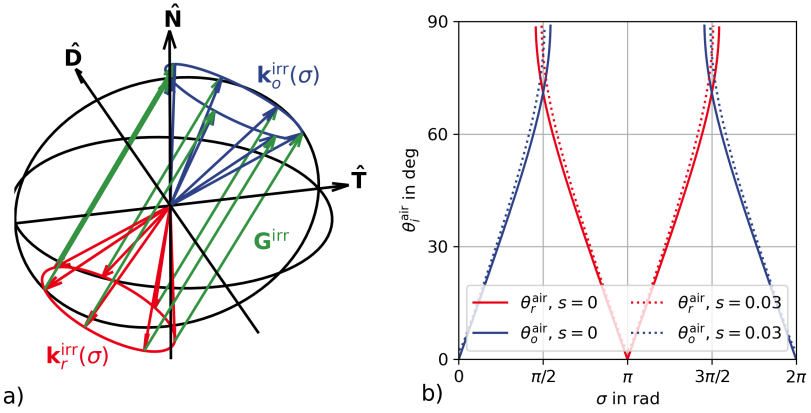


Figure 4.1: (a) Visualization of possible decompositions into reference and object waves (recording configuration σ) for the grating vector of a typical reflective waveguide incoupler (532 nm, propagation angle 56° [38]). (b) Reference and object incidence angles for this grating in air for all possible configurations, with and without typical shrinkage. For $\sigma = \pi/2$ and $3\pi/2$, the smallest limiting incidence angle is achieved with the same incidence angle for both reference and object beams.

4.2.2 Immersion-free recording condition

When recording holograms with edge lighting, a prism is optically attached to the holographic substrate using immersion liquid, allowing arbitrary incidence angles inside the substrate given a suitable prism. For immersion-free recording both recording wavefronts must be transmitted into the substrate from air through the boundary planes normal to $\hat{\mathbf{N}}$. The air-side incidence angle of recording wave $i \in \{r, o\}$ is measured relative to the local surface normal,

$$\theta_i^{\mathrm{air}}(\sigma, \lambda, n; \mathbf{G}^{\mathrm{irr}}) = \arccos\left(|\hat{\mathbf{k}}_i^{\mathrm{air}} \cdot \hat{\mathbf{N}}|\right), \tag{4.2}$$

and must be smaller than $90°$. In practice, only angles smaller than a maximum permissible value $\theta_{\mathrm{max}}^{\mathrm{air}} < 90°$ are allowed due to strongly increasing Fresnel reflection and spot size under grazing incidence. Therefore, a grating $\mathbf{G}^{\mathrm{irr}}$ is recordable immersion-free at a given wavelength λ and configuration σ when it satisfies the resolution limit Eq. (4.1) and

$$\theta_r^{\mathrm{air}}, \theta_o^{\mathrm{air}} < \theta_{\mathrm{max}}^{\mathrm{air}} \Leftrightarrow \bar{\theta}(\sigma, \lambda, n; \mathbf{G}^{\mathrm{irr}}) < \theta_{\mathrm{max}}^{\mathrm{air}} \tag{4.3}$$

with the limiting incidence angle

$$\bar{\theta}(\sigma, \lambda, n; \mathbf{G}^{\mathrm{irr}}) = \max\left(\theta_r^{\mathrm{air}}, \theta_o^{\mathrm{air}}\right). \tag{4.4}$$

4.2.3 Incidence angles for immersion-free recording

The tangential components $k_{i,\parallel}$ of both recording wave vectors $i \in \{r, o\}$ are conserved across the boundary planes due to translational symmetry, $k_{i,\parallel}^2 = k^2 - (\hat{\mathbf{N}} \cdot \mathbf{k}_i)^2$. Inserting Eq. (2.9) gives for recording configuration σ,

$$k_{i,\parallel}^2(\sigma) = k^2 - \left(\sqrt{k^2 - \frac{(G^{\mathrm{irr}})^2}{4}} \frac{G_T^{\mathrm{irr}}}{G^{\mathrm{irr}}} \cos\sigma + \Phi_i \frac{G_N^{\mathrm{irr}}}{2}\right)^2, \tag{4.5}$$

with $\Phi_r = -1$ and $\Phi_o = 1$.

Since the air-side wave-vector magnitude is $2\pi/\lambda$, the incidence angle follows as

$$\theta_i^{\mathrm{air}}(\sigma) = \arccos\sqrt{1 - \left(\frac{\lambda k_{i,\parallel}(\sigma)}{2\pi}\right)^2}, \tag{4.6}$$

provided that $k_{i,\parallel} \leq 2\pi/\lambda$. For $k_{i,\parallel} > 2\pi/\lambda$, no propagating air-side wave exists, we assign the angle the value $+\infty$ so that Eq. (4.3) remains valid. Figure 4.1(b) shows the incidence angles for recording an exemplary waveguide incoupler grating in various configurations σ. Some recording configurations yield lower incidence angles than others.

To determine whether any immersion-free configuration fulfills Eq. (4.3), we seek the configuration σ minimizing the limiting incidence angle in Eq. (4.4):

$$\bar{\theta}_{\min}(\lambda, n; \mathbf{G}^{\mathrm{irr}}) = \min_{\sigma\in[0,2\pi)} \bar{\theta}(\sigma, \lambda, n; \mathbf{G}^{\mathrm{irr}}). \tag{4.7}$$

The angle θ_i^{air} is monotonically increasing in $k_{i,\parallel}^2$. Thus, minimizing $\bar{\theta}(\sigma)$ is equivalent to minimizing $\max(k_{r,\parallel}^2(\sigma), k_{o,\parallel}^2(\sigma)) = k^2 - f(x(\sigma))^2$, where Eq. (4.5) gives

$$f(x) = \min\left(\left|x - \frac{G_N^{\mathrm{irr}}}{2}\right|, \left|x + \frac{G_N^{\mathrm{irr}}}{2}\right|\right) = \left||x| - \frac{G_N^{\mathrm{irr}}}{2}\right|. \tag{4.8}$$

Here $x(\sigma) = \sqrt{k^2 - (G^{\mathrm{irr}})^2/4}G_T^{\mathrm{irr}}/G^{\mathrm{irr}}\cos\sigma$. Consequently, minimizing the limiting incidence angle $\bar{\theta}_{\min}$ with respect to $\sigma \in [0, 2\pi)$ is equivalent to maximizing $f(x)$ over $x(\sigma) \in [-x_l, x_l]$, with $x_l(\lambda, n; \mathbf{G}^{\mathrm{irr}}) = \sqrt{k^2 - (G^{\mathrm{irr}})^2/4}G_T^{\mathrm{irr}}/G^{\mathrm{irr}}$. We consider valid recording wavelengths λ satisfying the resolution limit Eq. (4.1), for which x_l is real and nonnegative.

A recording configuration σ is either transmissive, with both recording waves incident from the same side of the substrate ($k_{r,N}k_{o,N} > 0$), or reflective, with the recording waves incident from opposite sides ($k_{r,N}k_{o,N} < 0$). Grazing configurations with $k_{r,N}k_{o,N} = 0$ are not physical. Equation (2.9) gives $k_{i,N} = x + \Phi_i G_N^{\mathrm{irr}}/2$, showing that a recording configuration is reflective if and only if $|x(\sigma)| < G_N^{\mathrm{irr}}/2$, and transmissive if and only if $|x(\sigma)| > G_N^{\mathrm{irr}}/2$.

On the domain of all reflective recording configurations, $x \in \mathcal{R} = [-x_l, x_l] \cap (-G_N^{\mathrm{irr}}/2, G_N^{\mathrm{irr}}/2)$, we can write $f(x) = G_N^{\mathrm{irr}}/2 - |x|$. If reflective configurations exist, $G_N^{\mathrm{irr}} > 0$, the maximum on the reflective domain is attained at $x = 0$, corresponding to $\sigma \in \{\pi/2, 3\pi/2\}$. Therefore, the smallest limiting incidence angle for any reflective configuration is obtained by inserting Eq. (4.5) into Eqs. (4.6) and (4.4)

$$\bar{\theta}_{\min}^{\mathrm{ref}} = \min_{\substack{\sigma \\ k_{r,N} k_{o,N} < 0}} \bar{\theta}(\sigma, \lambda, n; \mathbf{G}^{\mathrm{irr}}) = \arccos\left(\sqrt{1 - n^2 + \left(\frac{G_N(1-s)\lambda}{4\pi}\right)^2}\right). \tag{4.9}$$

For the $\sigma \in \{\pi/2, 3\pi/2\}$ configurations, the reference and object waves have the same incidence angle, equal to Eq. (4.9).

On the domain of all transmissive recording configurations, $x \in \mathcal{T} = [-x_l, -G_N^{\mathrm{irr}}/2) \cup (G_N^{\mathrm{irr}}/2, x_l]$, we can write $f(x) = |x| - G_N^{\mathrm{irr}}/2$. Transmissive recording configurations exist only if $x_l > G_N^{\mathrm{irr}}/2$. In that case, the maximum on the transmissive domain is attained at the endpoints $x = \pm x_l$, corresponding to $\sigma \in \{0, \pi\}$. Thus, the smallest limiting incidence angle for any transmissive recording configuration follows from Eqs. (4.5), (4.6), and (4.4)

$$\begin{aligned} \bar{\theta}_{\min}^{\mathrm{trans}} &= \min_{\substack{\sigma \\ k_{r,N} k_{o,N} > 0}} \bar{\theta}(\sigma, \lambda, n; \mathbf{G}^{\mathrm{irr}}) \\ &= \arccos\left(\sqrt{1 - n^2 + \left(\sqrt{n^2 - \left(\frac{\lambda G^{\mathrm{irr}}}{4\pi}\right)^2}\frac{G_T^{\mathrm{irr}}}{G^{\mathrm{irr}}} - \frac{\lambda G_N^{\mathrm{irr}}}{4\pi}\right)^2}\right). \end{aligned} \tag{4.10}$$

The maximum value of f on the reflective domain is $f(0) = G_N^{\mathrm{irr}}/2$ and on the transmissive domain $f(\pm x_l) = x_l - G_N^{\mathrm{irr}}/2$. The global minimum of the limiting angle is therefore attained in the reflective configurations $\sigma \in \{\pi/2, 3\pi/2\}$ if $x_l < G_N^{\mathrm{irr}}$ and in the transmissive configurations $\sigma \in \{0, \pi\}$ if $x_l > G_N^{\mathrm{irr}}$. For $x_l = G_N^{\mathrm{irr}}$, all four configurations attain the minimum.

Summarizing, the smallest limiting incidence angle across all configurations σ is

$$\bar{\theta}_{\min}(\lambda, n; \mathbf{G}^{\mathrm{irr}}) = \begin{cases} \bar{\theta}^{\mathrm{ref}}_{\min}, & x_l \leq G^{\mathrm{irr}}_N, \\ \bar{\theta}^{\mathrm{trans}}_{\min}, & x_l > G^{\mathrm{irr}}_N. \end{cases} = \min\left(\bar{\theta}^{\mathrm{ref}}_{\min}, \bar{\theta}^{\mathrm{trans}}_{\min}\right). \tag{4.11}$$

To make the last equality rigorous, if either expression $\bar{\theta}^{\mathrm{ref}}_{\min}$ or $\bar{\theta}^{\mathrm{trans}}_{\min}$ becomes non-real, no propagating air-side solution is available for that branch, and the corresponding angle is assigned the extended-real value $+\infty$. The same convention is used when the reflective or transmissive configuration set is empty.

4.2.4 Monochromatic immersion-free recording limit

A grating with components G_T, G_N is recordable without immersion at a given wavelength λ if and only if it satisfies the resolution limit in Eq. (4.1) and has a limiting incidence angle small enough to fulfill Eq. (4.3) for at least one configuration σ. Equivalently, the minimum limiting incidence angle across all configurations is smaller than the permissible incidence angle, $\bar{\theta}_{\min}(\lambda, n; \mathbf{G}^{\mathrm{irr}}) < \theta^{\mathrm{air}}_{\max}$.

The grating is recordable immersion-free in reflection at a given wavelength λ when the resolution limit Eq. (4.1) is fulfilled and $\bar{\theta}^{\mathrm{ref}}_{\min}(\lambda, n; \mathbf{G}^{\mathrm{irr}}) < \theta^{\mathrm{air}}_{\max}$ or equivalently

$$G_N > \frac{4\pi}{\lambda(1-s)} \sqrt{n^2 - \sin^2(\theta^{\mathrm{air}}_{\max})}. \tag{4.12}$$

Equation (4.12) and the resolution limit Eq. (4.1) together define the gratings G_T, G_N that are recordable immersion-free in reflection at a fixed wavelength λ.

A grating with components G_T, G_N is recordable without immersion in transmission at λ if

$$\left(G_T - \frac{2\pi}{\lambda} \sin\theta^{\mathrm{air}}_{\max}\right)^2 + \left[(1-s)G_N + \frac{2\pi}{\lambda}\sqrt{n^2 - \sin^2\theta^{\mathrm{air}}_{\max}}\right]^2 < \left(\frac{2\pi n}{\lambda}\right)^2. \tag{4.13}$$

This inequality ensures $\bar{\theta}^{\mathrm{trans}}_{\min} < \theta^{\mathrm{air}}_{\max}$ and implies the resolution limit Eq. (4.1).

Figure 4.2 (left) shows the reflection- and transmission-recordable gratings at $532\,\mathrm{nm}$ in a typical photopolymer with $n = 1.51$ and $s = 0.03$ [56]. The color encodes $\bar{\theta}_{\mathrm{min}}(\lambda = 532\,\mathrm{nm}, n = 1.51, s = 0.03; G_T, G_N)$. The lower region corresponds to transmissive recording and the upper region to reflective recording.

For orientation, Fig. 4.2 shows the grating components $G_T(\theta_p, \lambda_\mathrm{d})$ and $G_N(\theta_p, \lambda_\mathrm{d})$ derived in Eq. (2.24) for normal-incidence reflective incouplers and orthogonal expanders commonly used in planar waveguide displays. The marked gratings

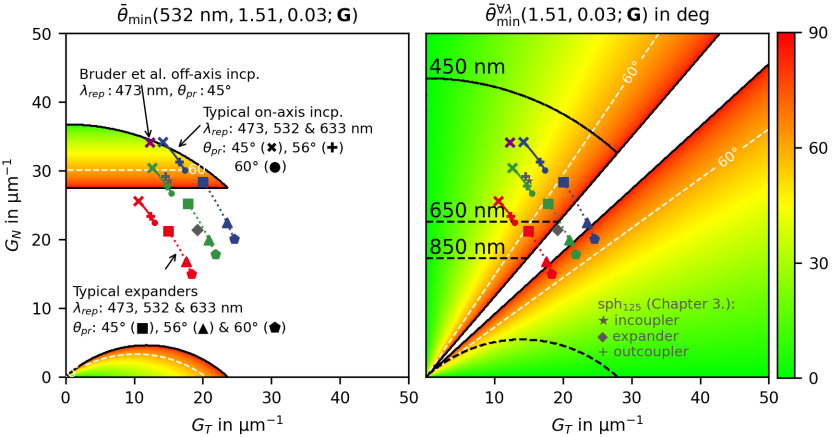


Figure 4.2: Each position (G_T, G_N) represents a holographic grating in a material with $n = 1.51$ and $s = 0.03$. The left plot shows the minimum limiting incidence angle for any recording configuration at $532\,\mathrm{nm}$. The upper region belongs to the reflective branch and the lower region to the transmissive branch. The right plot shows the global infimum over all positive recording wavelengths and configurations, obtained from the reflective and transmissive branch expressions in Eqs. (4.15) and (4.17). The black dotted and dashed outlines indicate the gratings recordable within $450\,\mathrm{nm}$ to $650\,\mathrm{nm}$ and $450\,\mathrm{nm}$ to $850\,\mathrm{nm}$, respectively. White dashed contours mark a limiting incidence angle of $60°$. Colored solid and dotted curves denote on-axis incouplers and orthogonal expanders from Eq. (2.24); the three markers along each curve correspond to propagation angles of $45°$, $56°$, and $60°$. Blue, green, and red denote the replay wavelengths. The off-axis incoupler of [56] is shown in purple, and the $125\,\mathrm{mm}$ spherical-waveguide couplers are shown in gray.

cover the propagation-angle range $\theta_p \in [45°, 60°]$ and representative blue, green, and red display replay wavelengths $\lambda_\mathrm{d} \in \{473\,\mathrm{nm}, 532\,\mathrm{nm}, 633\,\mathrm{nm}\}$.

The figure indicates that green incouplers with moderate propagation angles and blue incouplers with larger propagation angles can, in principle, be recorded immersion-free at a recording wavelength of $532\,\mathrm{nm}$ in the reflective branch. Bruder *et al.* demonstrated immersion-free recording of a blue orthogonal expander and an off-axis incoupler designed for a propagation angle of $\theta_p = 45°$ [56]. Both gratings lie within the recordable region of the reflective branch in Fig. 4.2. For the incoupler, the central image mode is incident at $\theta_\mathrm{i} = 5.5°$ from the surface normal, $\mathbf{k}_\mathrm{B}^{(\mathrm{i})} = k_\mathrm{B}\big[\sin(\theta_\mathrm{i})\hat{\mathbf{T}} - \cos(\theta_\mathrm{i})\hat{\mathbf{N}}\big]$, and is diffracted into a guided mode with $\theta_p = 45°$, where $k_\mathrm{B} = 2\pi n_\mathrm{wg}(\lambda_\mathrm{B})/\lambda_\mathrm{B}$ and $\lambda_\mathrm{B} = 473\,\mathrm{nm}$. Equation (2.21) therefore gives $G_T^\mathrm{B} = \big(\sin(45°) - \sin(5.5°)\big)k_\mathrm{B}$ and $G_N^\mathrm{B} = \big(\cos(5.5°) + \cos(45°)\big)k_\mathrm{B}$. Relative to normal incidence, the off-axis configuration reduces the required angular deflection and hence the grating vector magnitude, corresponding to a larger grating period. This shift is sufficient to make the incoupler recordable immersion-free in reflection at $532\,\mathrm{nm}$.

4.2.5 Polychromatic immersion-free recording limit

A grating G_T, G_N is not recordable immersion-free if no recording wavelength and configuration yield a limiting incidence angle that fulfills Eq. (4.3).

The smallest limiting incidence angle for reflective recording in Eq. (4.9) decreases with increasing recording wavelength. Its smallest value for any wavelength is achieved at the largest wavelength permitted by the resolution limit,

$$\lambda_\mathrm{max}(n, s; G_T, G_N) = \frac{4\pi n}{\sqrt{G_T^2 + \left[G_N(1-s)\right]^2}}. \tag{4.14}$$

Substitution into Eq. (4.9) gives the smallest limiting incidence angle for any reflective configuration and wavelength

$$\bar{\theta}_{\min}^{\mathrm{ref},\forall\lambda}(n,s;\mathbf{G}) = \arcsin\left(\frac{nG_T}{\sqrt{G_T^2+\left[G_N(1-s)\right]^2}}\right). \tag{4.15}$$

Thus a grating is recordable in reflection for some wavelength, $\bar{\theta}_{\min}^{\mathrm{ref},\forall\lambda} < \theta_{\max}^{\mathrm{air}}$, if

$$G_N > \frac{\sqrt{n^2-\sin^2(\theta_{\max}^{\mathrm{air}})}}{(1-s)\sin\theta_{\max}^{\mathrm{air}}}G_T. \tag{4.16}$$

This inequality describes a cone around the G_N axis in G_T–G_N grating space.

In contrast, the smallest limiting incidence angle for transmissive recording in Eq. (4.10) decreases as the recording wavelength decreases. Ignoring dispersion, its minimum is approached for $\lambda \to 0$,

$$\bar{\theta}_{\min}^{\mathrm{trans},\forall\lambda} = \inf_{\lambda>0}\bar{\theta}_{\min}^{\mathrm{trans}}(\lambda,n,s;\mathbf{G}) = \arcsin\left(\frac{n(1-s)G_N}{\sqrt{G_T^2+\left[G_N(1-s)\right]^2}}\right), \tag{4.17}$$

and a grating in the transmissive branch is recordable for some configuration and wavelength, that is $\bar{\theta}_{\min}^{\mathrm{trans},\forall\lambda} < \theta_{\max}^{\mathrm{air}}$, if

$$G_N < \frac{\sin\theta_{\max}^{\mathrm{air}}}{(1-s)\sqrt{n^2-\sin^2(\theta_{\max}^{\mathrm{air}})}}G_T. \tag{4.18}$$

This inequality describes a second cone around the G_T axis. Figure 4.2 (right) visualizes both the reflective and transmissive recordability cones for a typical photopolymer with $n = 1.51$, $s = 0.03$ [56]. The plot colors encode $\bar{\theta}_{\min}^{\forall\lambda}(n,s;\mathbf{G})$.

Gratings outside both cones are not recordable immersion-free at any wavelength or recording configuration. Practically, only a limited part of both cones is recordable because the material is sensitive over a finite wavelength range

$[\lambda_{\mathrm{lower}}, \lambda_{\mathrm{upper}}]$, such as $450\,\mathrm{nm}$ to $850\,\mathrm{nm}$ for state-of-the-art photopolymer [56].

In the reflective branch, the shortest available wavelength λ_{lower} limits the recordable cone at large G_N by an elliptical boundary, which is circular when $s = 0$ through the resolution limit Eq. (4.1). The longest available wavelength limits the reflective region at small G_N through Eq. (4.12) by a horizontal cut-off. Because Eq. (4.9) decreases with increasing wavelength, the smallest reflective-branch angle in $[\lambda_{\mathrm{lower}}, \lambda_{\mathrm{upper}}]$ is attained at $\lambda^{\star}_{\mathrm{ref}} = \min\{\lambda_{\max}(n, s; G_T, G_N), \lambda_{\mathrm{upper}}\}$. If $\lambda_{\max} < \lambda_{\mathrm{lower}}$, the grating violates the resolution limit throughout the available range.

In the transmissive branch, Eq. (4.10) decreases monotonically as the wavelength decreases. Its minimum in the finite range is therefore attained at λ_{lower}. The shortest available wavelength also bounds the transmissive region by the ellipse in Eq. (4.13).

With the branch-specific angles assigned $+\infty$ whenever they are complex and therefore unphysical, the global minimum over the finite wavelength range is

$$\bar{\theta}^{[\lambda_{\mathrm{lower}}, \lambda_{\mathrm{upper}}]}_{\mathrm{min}} = \begin{cases} \min\left[\bar{\theta}^{\mathrm{trans}}_{\mathrm{min}}(\lambda_{\mathrm{lower}}), \bar{\theta}^{\mathrm{ref}}_{\mathrm{min}}(\lambda^{\star}_{\mathrm{ref}})\right], & \lambda_{\max} \geq \lambda_{\mathrm{lower}}, \\ +\infty, & \lambda_{\max} < \lambda_{\mathrm{lower}}. \end{cases} \tag{4.19}$$

Figure 4.3 shows the smallest limiting incidence angle attainable over the practically accessible photopolymer range $450\,\mathrm{nm}$ to $850\,\mathrm{nm}$ and over an extended $100\,\mathrm{nm}$ to $850\,\mathrm{nm}$ range.

By allowing wavelength shifting, a wide range of RGB incoupling gratings can be realized in the considered photopolymer without immersion at recording angles below 60° in the reflective recording configuration. However, immersion-free reflective recording of expanders requires large incidence angles even at moderate waveguide propagation angles and ceases to be possible at larger propagation angles. Intermediate-propagation-angle expanders lie between the reflective and transmissive cones and are not recordable in either branch. At sufficiently large

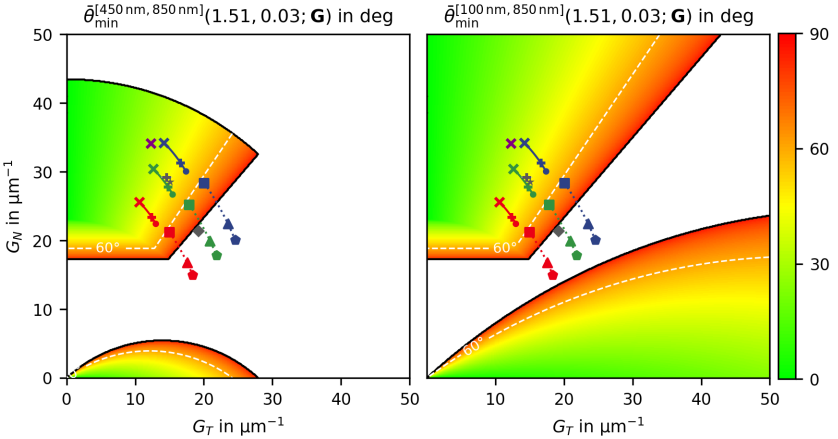


Figure 4.3: Smallest limiting incidence angle attainable over the finite recording-wavelength ranges $450\,\mathrm{nm}$ to $850\,\mathrm{nm}$ (left) and $100\,\mathrm{nm}$ to $850\,\mathrm{nm}$ (right), for $n = 1.51$ and $s = 0.03$. At every grating, the smaller of the reflective- and transmissive-branch angles from Eq. (4.19) is shown. Uncolored regions have no propagating air-side recording solution anywhere in the specified range and are not recordable immersion-free. Grating curves and markers are defined as in Fig. 4.2.

propagation angles, expanders enter the transmissive branch, but only at very short recording wavelengths. For the $60°$ RGB expanders shown here, propagating air-side solutions require recording wavelengths below $100\,\mathrm{nm}$. These wavelengths are far outside the $450\,\mathrm{nm}$ to $850\,\mathrm{nm}$ sensitivity range of state-of-the-art photopolymers [56]. Figure 4.3 shows that extending the lower limit to $100\,\mathrm{nm}$ remains insufficient. Existing holographic photopolymers are not compatible with these extreme-ultraviolet exposure wavelengths. Therefore, the practically relevant expander gratings considered here continue to require edge-lit exposure in photopolymer.

4.2.6 Applicability to spherical waveguide couplers

The limits derived above can also be applied pointwise to freeform gratings with spatially varying $G_T(\mathbf{p})$ and $G_N(\mathbf{p})$. Strictly, however, the derivation applies

only to planar holographic substrates, since the calculation of the air-side incidence angles assumes planar entry surfaces with a locally constant normal $\hat{\mathbf{N}}(\mathbf{p})$. Nevertheless, for HOEs on weakly curved or thin substrates, such as the spherical waveguide couplers of Chapter 3, the limits remain valid to first order.

The grating decomposition in Eq. (2.9) is a purely local relation at $\mathbf{p}$ and is exact irrespective of curvature. The recording air-side angles θ_i^{air}, defined in Eq. (4.2), however, depend on the surface normal at the entry points. Because the substrate has finite thickness d, each recording ray $i \in \{r, o\}$ enters at a point $\mathbf{p}_i$ that is offset from $\mathbf{p}$. To first order, treating the substrate locally as a planar slab, the ray crosses the thickness at internal angle $\theta_{\mathrm{int},i}$ and lands a tangential distance $\Delta s_i \approx d \tan\theta_{\mathrm{int},i}$ away from where it entered. On a spherical substrate, the surface normal is radial, so two surface points separated by an arc length Δs_i subtend a central angle $\Delta s_i/R$; hence, the entry-point normal $\hat{\mathbf{N}}(\mathbf{p}_i)$ is rotated relative to $\hat{\mathbf{N}}(\mathbf{p})$ by $\Delta\alpha_i = \Delta s_i/R$, with R the radius of curvature. In immersion-free recording, all beams enter from air, so the internal angle cannot exceed $\theta_{\mathrm{c}} = \arcsin(1/n)$. The walk-off, and hence the normal tilt, is therefore bounded independently of wavelength and configuration by $|\Delta\alpha_i| \leq d/R \tan\theta_{\mathrm{c}} = d/\left(R\sqrt{n^2-1}\right)$. This tilt bounds the change in the air-side incidence angle by $|\Delta\theta_i^{\mathrm{air}}| \leq \arcsin(n\sin(|\Delta\alpha_i| + \theta_{\mathrm{int},i})) - \arcsin(n\sin(\theta_{\mathrm{int},i}))$ with $\theta_{\mathrm{int},i} = \arcsin(1/n \sin(\theta_i^{\mathrm{air}}))$. Away from grazing incidence the angular error is small relative to the planar prediction in Eqs. (4.9) and (4.10). For the $R = 125\,\mathrm{mm}$ spherical waveguide with $d \leq 0.5\,\mathrm{mm}$ and $n = 1.51$ discussed in Chapter 3, $|\Delta\alpha_i| < 0.21^\circ$ and $|\Delta\theta_i^{\mathrm{air}}| < 2^\circ$ for $\theta_i^{\mathrm{air}} < 80^\circ$. The planar monochromatic limits in Eqs. (4.12) and (4.13), and the arbitrary-wavelength limits in Eqs. (4.16) and (4.18), therefore remain approximately applicable pointwise when the grating lies safely away from a recordability boundary.

The G_T and G_N components of the spherical coupler gratings vary across the coupler aperture. The corresponding G_T–G_N regions are indicated in light gray in Fig. 4.2. Because these regions are small, they are additionally marked by dark gray central markers. Consistent with the planar result, the $125\,\mathrm{mm}$ spherical waveguide expander introduced in Chapter 3 cannot be recorded without immersion for any wavelength or recording configuration.

4.3 Recording polarizations

The inequalities derived in the last section determine the set of gratings that are, in principle, recordable without immersion. However, there are further practical limitations due to polarization effects:

Effective index modulation The effective index modulation Δn of a holographic grating determines the required thickness for high diffraction efficiency and, therefore, indirectly affects the angular and spectral bandwidth of a holographic grating, see Eq. (2.16). As discussed in Sec. 2.1.1, the effective index modulation decreases for visibilities $V < 1$. To achieve unity visibility, both recording waves must exhibit complete polarization overlap. Incomplete polarization overlap reduces the attainable modulation from the material limit δn by the visibility factor V.

Superficial holographic gratings Only part of the recording wave is transmitted out of the holographic substrate back into air due to polarization-dependent Fresnel reflections. At grazing incidence, strong reflections can add coherently to the primary reference and object beams. They can disturb the intended holographic grating and introduce undesired superficial holograms [78, 79]. To prevent this, the reflections must be weak, or the reflected beams must have vanishing polarization overlap with both primary recording waves and among themselves.

In this section, we investigate the recording polarization and wavelength that mitigate these effects for reflective immersion-free recording in the $\sigma \in \{\pi/2, 3\pi/2\}$ configuration.

4.3.1 Brewster-polarized recording

The Fresnel reflectance vanishes for p-polarized light at the Brewster incidence angle $\theta_{\mathrm{Brew}}^{\mathrm{air}} = \arctan(n)$. For reflective recording in the configurations $\sigma \in \{\pi/2, 3\pi/2\}$, both object and reference incidence angles are $\bar{\theta}(\lambda)$, as shown in Eq. (4.9). Choosing the recording wavelength appropriately permits incidence at

the Brewster angle, preventing any reflections. By equating the Brewster angle and Eq. (4.9), the required Brewster wavelength λ_{Brew} is found

$$\lambda_{\mathrm{Brew}} = \frac{n^2}{\sqrt{1+n^2}} \frac{4\pi}{|G_N^{\mathrm{irr}}|}. \tag{4.20}$$

However, the resolution limit Eq. (4.1) must be fulfilled for a grating to be recordable at λ_{Brew}. Inserting Eq. (4.20) into Eq. (4.1) gives

$$G_N \geq \frac{n}{1-s} G_T. \tag{4.21}$$

This defines the cone in G_T–G_N space describing the Brewster-recordable gratings in reflection. It is steeper and encloses fewer gratings than the cone of generally reflective recordable gratings defined by Eq. (4.16). Brewster recording is therefore not applicable to all holographic gratings that may, in principle, be recorded without immersion in reflection. Additionally, while completely removing superficial holograms, Brewster polarization results in only moderate polarization overlap. In general, the s- and p-polarization vectors of either recording wave i are defined as $\hat{\mathbf{p}}_i^{\mathrm{s}} = \mathbf{k}_i \times \hat{\mathbf{N}} / |\mathbf{k}_i \times \hat{\mathbf{N}}|$ and $\hat{\mathbf{p}}_i^{\mathrm{p}} = \hat{\mathbf{p}}_i^{\mathrm{s}} \times \mathbf{k}_i / |\hat{\mathbf{p}}_i^{\mathrm{s}} \times \mathbf{k}_i|$. In the relevant $\sigma \in \{\pi/2, 3\pi/2\}$ recording geometry, the polarization overlap of the p-polarized object and reference waves is

$$|\hat{\mathbf{p}}_{\mathrm{o}}^{\mathrm{p}} \cdot \hat{\mathbf{p}}_{\mathrm{r}}^{\mathrm{p}}| = \left| \frac{\left(\left(G_N^{\mathrm{irr}}\right)^4 + \left(G_N^{\mathrm{irr}}\right)^2 \left(G_T^{\mathrm{irr}}\right)^2 \right) \lambda^4 - 24\pi^2 \left(G_N^{\mathrm{irr}}\right)^2 \lambda^2 n^2 + 128\pi^4 n^4}{8\pi^2 n^2 \left(16\pi^2 n^2 - \left(G_N^{\mathrm{irr}}\right)^2 \lambda^2\right)} \right|. \tag{4.22}$$

The overlap for different gratings at their Brewster wavelength is plotted in Fig. 4.4. All gratings with the same G_N value have the same Brewster wavelength, see Eq. (4.20). The three horizontal lines mark the gratings that can be recorded under Brewster incidence for $450\,\mathrm{nm}$, $650\,\mathrm{nm}$, and $850\,\mathrm{nm}$. The gratings between those lines are recordable with recording wavelengths in this range. Regardless of wavelength, the considered waveguide expanders cannot be recorded in the

Brewster configuration. The considered waveguide incouplers are recorded with reduced polarization overlap and consequently reduced Δn, see Fig. 4.4.

4.3.2 Quasi-Brewster-polarized recording

While vanishing completely at the Brewster angle, the p-polarized reflectance is also very small for surrounding angles. Therefore, to increase polarization overlap, we allow the recording wavelength to differ from the Brewster recording wavelength. For recording in the reflective $\sigma \in \{\pi/2, 3\pi/2\}$ configuration, we choose the wavelength that maximizes the polarization overlap.

At the resolution boundary $\lambda_{\max}$ in Eq. (4.14), the component of the recording vectors orthogonal to $\mathbf{G}^{\mathrm{irr}}$ vanishes, see Eq. (2.9), so the two recording vectors are antiparallel and their p-polarization overlap is one. Because unity is the largest possible overlap, $\lambda_{\max}$ maximizes the overlap for every grating that is

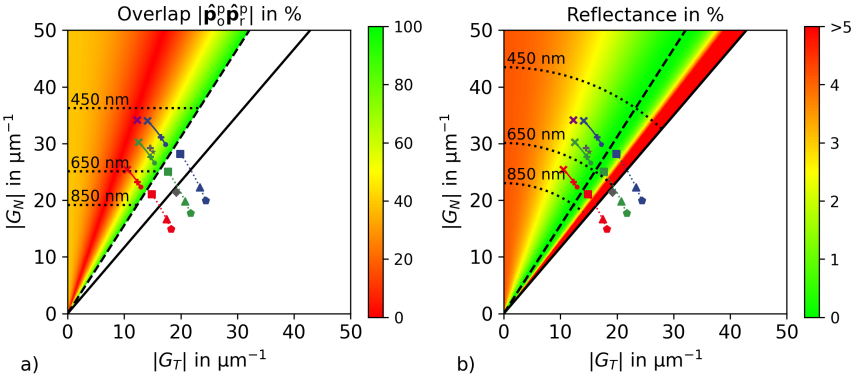


Figure 4.4: (a) Each position (G_T, G_N) in the plots represents a holographic grating. The color shows the polarization overlap achievable in Brewster-polarized recording with vanishing reflectance. Typical incoupling and expanding gratings are marked as defined in Fig. 4.2. (b) The color shows the reflectance for quasi-Brewster-polarized recording with full polarization overlap.

immersion-free recordable in reflection. It also yields the smallest reflective-branch incidence angle, see Eq. (4.9). Figure 4.4(b) shows the corresponding p-polarization reflectance. The considered on-axis RGB incouplers can be recorded this way with perfect polarization overlap and Fresnel reflectance below $2\,\%$. However, for expanding gratings, reflections grow rapidly with propagation angle. To reduce reflections, a multilayer antireflection coating specified for the incidence angle $\bar{\theta}_{\mathrm{min}}(\lambda_{\mathrm{max}})$ can be applied to the holographic film. This coating leaves the previous incidence and overlap analysis intact, as $k_{i,\parallel}$ is conserved in the stack.

4.4 Conclusion

In this chapter, we established analytical limits for immersion-free recording of holographic gratings with wavelength and recording configuration shifting, using the local G_T, G_N formalism. The global minimum limiting incidence angle has a reflective and a transmissive branch. Our analysis demonstrates that all considered RGB incouplers for typical planar reflective holographic waveguide displays in photopolymer can be recorded without immersion, which could simplify prototyping and manufacturing. We also highlighted the importance of recording polarization in achieving high effective index modulation and minimizing superficial holograms, constraining the set of practically recordable gratings. Practically relevant orthogonal expanders and the expander of the spherical waveguide discussed in Chapter 3 remain outside the immersion-free recording range of typical photopolymers and therefore require edge-lit exposure.

5 Automated planar waveguide recording

Automated recording of holographic waveguide displays is limited by the simultaneous need for highly adjustable beam angles exceeding the waveguide–cladding interface critical angle and centimeter-scale recording beams. We report, to our knowledge, the first automated holographic waveguide recording platform, combining two six-axis robot-mounted collimators with a custom edge-lit prism system. Exposure poses are computed from target grating vectors, including shrinkage compensation and polarization alignment, using inverse ray tracing and robot kinematics. Calibration yields mean angular accuracy better than 0.1°, and interferometry shows relative optical-path stability within $\pm\lambda/20$ during recording. A fabricated 2D eyebox-expanding validation sample transmits readable text, and its modulation transfer function is evaluated. We analytically explain a pupil-sampling-induced intensity modulation in the laser-scanned image. The presented fabrication platform supports more rapid and reproducible prototyping of planar holographic waveguide displays.

Parts of this chapter have been published in: Marvin Feil, Reinhold Fiess, and Wilhelm Stork, "Automated mastering of holographic waveguide displays with robot-mounted optics," *Appl. Opt.* **65**(28): 9786–9797 (2026). doi: 10.1364/AO.608922. Parts of this work were presented in Marvin Feil, Reinhold Fiess, and Wilhelm Stork, "Automated mastering of holographic waveguide displays," in *Proc. SPIE* **13821**, p. 138210A (2026), doi: 10.1117/12.3080347; and Marvin Feil, Tobias Wilm, Reinhold Fiess, and Wilhelm Stork, "Automated holographic recording using six-axis robot arm pair," in *Proc. SPIE* **13529**, p. 135290J (2025), doi: 10.1117/12.3055606.

5.1 Introduction

Prototyping holographic waveguide masters remains challenging. In contrast to surface relief gratings (SRGs), which benefit from mature lithography and nanoimprint tools, mastering holographic optical elements (HOEs) for holographic waveguide displays still often relies on purpose-built optical exposure setups and repeated manual alignment [38].

Several automated holographic recording systems have been developed to improve recording flexibility and reproducibility. Early computerized platforms enabled faceted hologram recording by combining substrate translation with adjustable recording beams [80, 81]. Later systems introduced fiber beam delivery, motorized stages, and goniometers for more flexible hologram fabrication [82–86]. Spatial-light-modulator-based holographic printers provide another route to programmable recording wavefronts [87, 47, 88].

These systems are not designed for waveguide holograms and provide either limited recording angle range or small recording field size. These limitations are unsuitable for reflective waveguide display HOEs, which must be centimeter-scale to achieve large eyeboxes while simultaneously requiring recording angles exceeding the critical angle of the hologram-cladding interface [89]. Providing these large recording angles requires immersion-assisted recording, for example using edge-lit prisms or immersion objectives [10, 88]. Therefore, recording full waveguide systems with all coupler HOEs requires an extensive angular range due to the distinct orientation of each grating coupler. Optical wavefront printers, for example, have insufficient étendue to simultaneously provide both centimeter-scale field size and numerical aperture for the angular range [88], requiring faceted printing that impacts hologram fidelity [47]. Existing mechanical systems must similarly trade off beam aperture against angular range.

Consequently, existing waveguide prototyping workflows record the individual HOEs in separate manually aligned recording prism systems and subsequently laminate them onto a common substrate [38, 39]. This requires repeated substrate

handling and immersion contact, introduces lamination tolerances, and slows design iterations. Manual realignment becomes particularly time-consuming when recording angularly multiplexed couplers to increase bandwidth [90].

This chapter introduces, to our knowledge, the first automated recording platform for holographic waveguides, combining two six-axis robot arms with a dedicated dual-prism recording system that we call the twin-prism. The key advantage of our

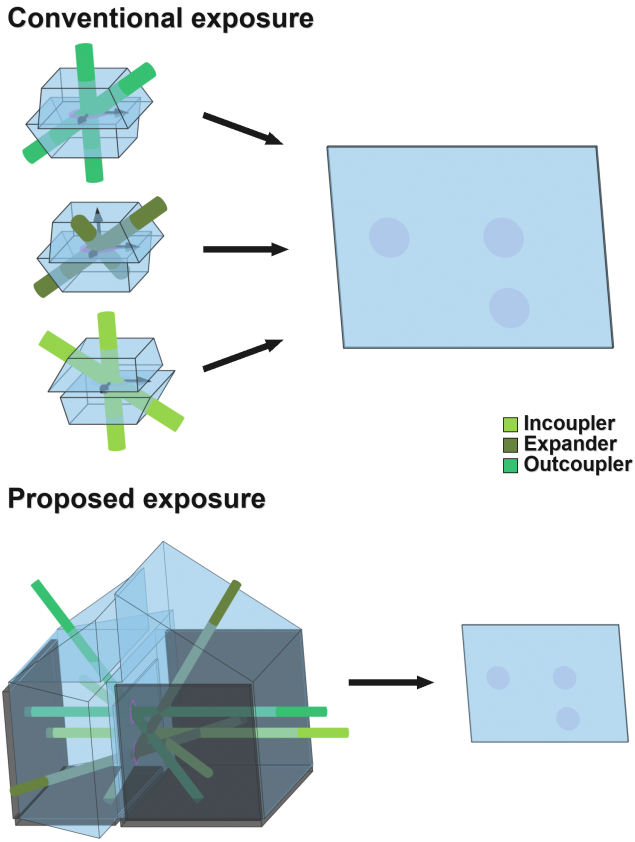


Figure 5.1: Conventional multiple-immersion and relamination workflow and proposed twin-prism waveguide recording.

prism geometry is that all waveguide couplers can be edge-lit using the same prism pair, requiring only a single immersion contacting step, as illustrated in Fig. 5.1. The robots provide control of the beam position, direction, and linear polarization for beam apertures up to $50\,\mathrm{mm}$. Robot recording poses are computed directly from target grating vectors, including shrinkage compensation and polarization alignment, using inverse ray tracing and robot kinematics. The robots sequentially expose the different waveguide couplers in the same substrate, avoiding post-recording relamination and providing intrinsic relative alignment. We investigate platform stability and angular accuracy and demonstrate the full fabrication chain with a validation sample, including laser-scanned image transport. The proposed platform provides a route toward more rapid and reproducible development of holographic waveguide displays.

5.2 Recording system design

In Sec. 2.2.3, we derived the coupler gratings required for planar orthogonal 1D- and 2D-EBE waveguide displays. We now introduce a recording system for these displays. The system is designed for waveguides with propagation angles around $\theta_p = 56^\circ$ and a recording wavelength close to the replay wavelength, while remaining sufficiently flexible to support other propagation angles, recording wavelengths, shrinkage compensation values, and moderately non-orthogonal waveguides with face wrap or light-engine tilt.

5.2.1 Recording prisms

As discussed in Chapter 4, recording waveguide display couplers requires recording wave vectors that are typically inaccessible by illuminating the holographic substrate directly from air. The required internal recording angles in the hologram layer can exceed the critical angle of the air–hologram interface, making edge-lit prism-assisted exposure necessary. In conventional holographic waveguide prototyping, the incoupler, expander, and outcoupler gratings are often recorded

in separate edge-lit prism systems and subsequently laminated onto a common substrate, as illustrated in Fig. 5.1 [38, 39]. This repeated immersion contacting and lamination increases process complexity and introduces relative alignment tolerances between the individual HOEs.

To avoid these repeated handling steps, we use a custom twin-prism system in which the holographic substrate is immersion-contacted once and remains fixed during the complete recording sequence. The assembly supports substrates up to $90\,\mathrm{mm}$ in width. The prism pair, shown schematically in Fig. 5.1, is fabricated from N-BK7, with the edge-lit facets cut at $56°$. It is designed so that the incoupling, expanding, and outcoupling gratings of orthogonal 1D- and 2D-EBE layouts can be recorded without reorienting the substrate. The prism coordinate system $\mathcal{P}$ is centered at the intersection of the prism contact plane, its bottom plane, and its side plane. The assembled twin-prism is shown in Fig. 5.2 (left).

5.2.2 Ray tracing

The accessibility of the required recording angles and the corresponding air-side beams are determined by inverse ray tracing. We consider exemplary orthogonal

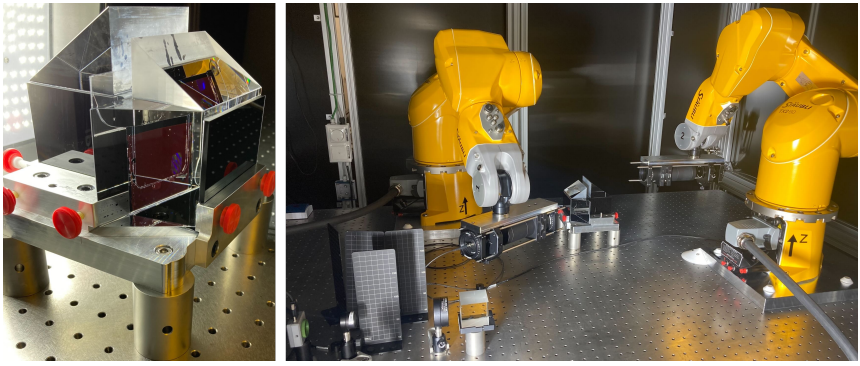


Figure 5.2: Left: 2D-EBE waveguide sample in immersion contact with the twin-prism system after recording. Right: Robotic exposure platform. The bottom-left shows the stability-measurement setup.

1D- and 2D-EBE waveguides on a $60\,\mathrm{mm} \times 80\,\mathrm{mm}$ substrate. The designs use a replay wavelength of $\lambda_\mathrm{d} = 520\,\mathrm{nm}$, a propagation angle of $56°$, and the coupler layout shown in Fig. 5.3, corresponding to the k-vector diagram in Fig. 2.3.

For each coupler $v \in \{\mathrm{i}, \mathrm{o1D}, \mathrm{o2D}, \mathrm{e}\}$ with grating vector $\mathbf{G}^v$ given by Eq. (2.23), Eq. (2.9) gives the object and reference recording wave vectors $\mathbf{k}_\mathrm{o}^v$ and $\mathbf{k}_\mathrm{r}^v$ as functions of the decomposition parameter σ, recording wavelength λ, and material-shrinkage parameters s and s_T. Suitable decomposition parameters for the incoupler, 1D outcoupler, 2D outcoupler, and expander are $\sigma^\mathrm{i} = \pi$, $\sigma^\mathrm{o1D} = \pi$, $\sigma^\mathrm{o2D} = \pi$, and $\sigma^\mathrm{e} = \pi/2$, respectively. The corresponding linear polarization directions $\hat{\mathbf{p}}_\mathrm{o}$ and $\hat{\mathbf{p}}_\mathrm{r}$ for optimal recording visibility are given by Eq. (2.11).

For each beam $b \in \{\mathrm{o}, \mathrm{r}\}$, a ray is launched from the coupler center in the direction opposite its recording wave vector and traced outward through the prism assembly until it refracts into air. The resulting exit position, air-side propagation direction, and polarization in the prism frame $\mathcal{P}$ are denoted by $\mathbf{r}_{b,\mathrm{air}}^\mathcal{P}$, $\hat{\mathbf{s}}_{b,\mathrm{air}}^\mathcal{P}$, and $\hat{\mathbf{p}}_{b,\mathrm{air}}^\mathcal{P}$, respectively.

For visualization, Figs. 5.4 and 5.5 show the corresponding forward trace of the recording beams from the air-side exit positions $\mathbf{r}_{b,\mathrm{air}}^\mathcal{P}$ toward the couplers in directions $-\hat{\mathbf{s}}_{b,\mathrm{air}}^\mathcal{P}$. The traces assume the $14\,\mathrm{mm}$ recording-beam diameter and $\lambda = 532\,\mathrm{nm}$ recording wavelength, used for sample fabrication in Sec. 5.4. The

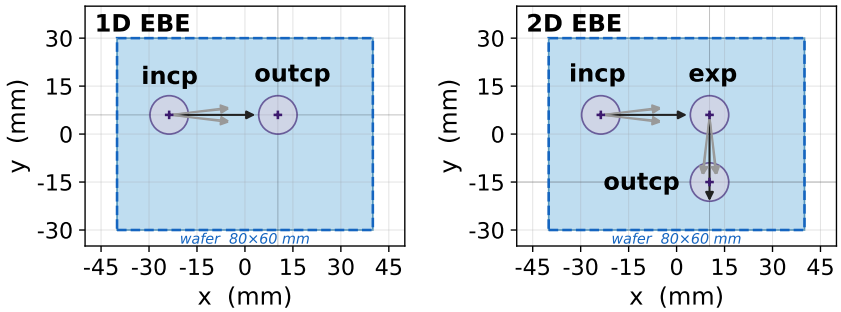


Figure 5.3: Exemplary orthogonal 1D- and 2D-EBE waveguides on a $60\,\mathrm{mm} \times 80\,\mathrm{mm}$ substrate. The 2D-EBE waveguide corresponds to the sample presented in Sec. 5.5.

dashed rays neglect material shrinkage, with $s = s_T = 0$ and $n = 1.51$. The solid rays include shrinkage compensation and are incident normally on the prism facets. They correspond to $s = 2.7\,\%$ and $s_T = 1.6\,\%$ for the incoupler and outcoupler, and $s = 3.6\,\%$ and $s_T = 1.7\,\%$ for the expander.

5.2.3 Parasitic hologram suppression

Fresnel reflections at the immersion-oil, photopolymer, and prism exit interfaces can interfere with the primary recording beams and record parasitic superficial

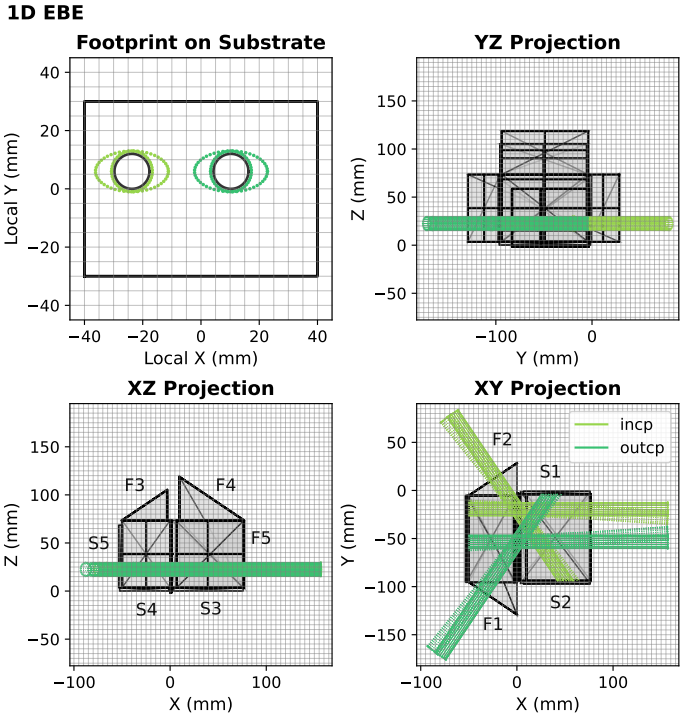


Figure 5.4: Recording-beam ray traces for a 1D-EBE waveguide coupler. Solid rays include shrinkage compensation.

holograms. The twin-prism assembly is therefore designed to suppress unwanted reflected beams before they re-enter the holographic substrate.

The N-BK7 prism material is approximately index-matched to typical Bayfol-type photopolymer films with refractive index near $n = 1.51$ [91]. During exposure, the substrate is contacted to the prisms with an N-BK7-matched immersion oil (Zeiss Immersol-F), reducing Fresnel reflections at the prism–oil–substrate interfaces. After the primary beams pass through the recording region, they are terminated by strategically placed absorbing filter elements bonded to the prism assembly. The filters are made from NG9 absorbing glass and are bonded using

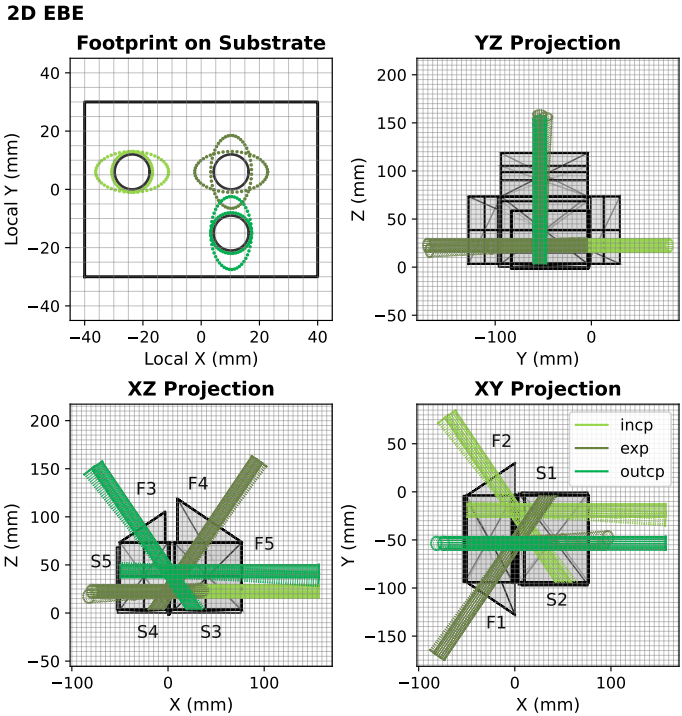


Figure 5.5: Recording-beam ray traces for a 2D-EBE waveguide coupler. Solid rays include shrinkage compensation.

Norland NOA65 cement, both approximately index-matched to the prism material. This approach is inspired by Upatnieks [10].

This absorbing-termination strategy is useful because it allows the recording angles to be varied to account for shrinkage compensation, off-design recording wavelengths, or different waveguide designs. Previous waveguide-recording geometries suppressed reflection-induced stray exposures by choosing prism facets such that the recording beams exit at normal incidence and by reducing the remaining Fresnel reflections with antireflection coatings [38]. Our termination approach suppresses reflections over a broad angular and wavelength range while preserving optical access for the recording beams and simplifies prism mounting through a flat bottom surface.

5.3 Robotic holography platform

During waveguide development, recording angles must be continuously adapted and optimized, for example to account for shrinkage of different holographic composites or new waveguide designs. The twin-prism supports the recording of a wide range of coupler gratings.

The role of the robotic platform is to generate these recording beams reproducibly. Two six-axis industrial robots carry laser collimation heads and allow independent control of the recording position, direction, and polarization roll angle of the object and reference beams. During automated waveguide mastering, the robots move through the exposure poses for the incoupler, expander, and outcoupler and record the HOEs sequentially with computer-controlled laser shutter timing. Changes in waveguide design or shrinkage compensation require only a recalculation of the recording poses, not a manual realignment of the optical setup.

5.3.1 System details

The recording source is a monochromatic $532\,\mathrm{nm}$ laser split into two polarization-maintaining fibers with numerical aperture 0.13. Each fiber feeds an identical robot-mounted collimation head. The fiber output is coupled directly into free space. Its divergence is increased by a $-15\,\mathrm{mm}$ achromat placed $30\,\mathrm{mm}$ from the fiber tip, and the resulting beam is collimated by a $200\,\mathrm{mm}$ achromat. An adjustable iris provides output beam diameters up to $40\,\mathrm{mm}$ with high uniformity. The edge-to-center intensity roll-off is below $50\,\%$. A cross-sectional view of the collimation head is shown in Fig. 5.6.

The collimators are mounted on two Stäubli TX2-60 six-axis robots placed on opposite sides of the prism assembly, as shown in Fig. 5.2. Each robot provides six-degree-of-freedom control of the corresponding recording beam, including rotation about the beam axis for setting the linear polarization direction. The mechanical locking system of the robots is used during exposure to suppress servo-induced motion. All robot poses are calculated with respect to the prism coordinate system $\mathcal{P}$ introduced in Sec. 5.2.1. The Stäubli base frame $\mathcal{B}_b$ of each robot is centered at the first joint, $375\,\mathrm{mm}$ above the robot mounting plate. In the

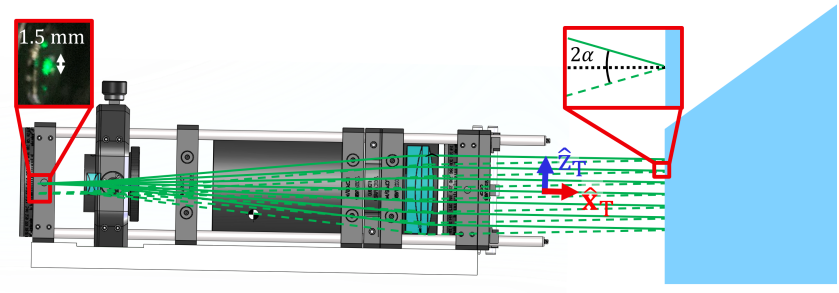


Figure 5.6: Cross section of robot-mounted laser collimators. Illustration of angular calibration procedure described in Sec. 5.3.5.

nominal mechanical layout, the origins of the two robot base frames, expressed in the prism coordinate system, are

$$\mathbf{r}^{\mathcal{P}}_{\mathcal{B}_1} = (-420,\ 166,\ 334)^{\mathsf{T}}\,\mathrm{mm}, \qquad \mathbf{r}^{\mathcal{P}}_{\mathcal{B}_2} = (430,\ -413,\ 334)^{\mathsf{T}}\,\mathrm{mm}.$$

The object recording beam is provided by robot 1 and the reference beam by robot 2, corresponding to the left and right sides of Fig. 5.2 (right), respectively.

5.3.2 Robot pose calculation

We use an inverse algorithm to compute the robot poses required to generate prescribed recording wave vectors at selected coupler positions on the holographic substrate. In contrast to forward-search approaches, in which the hologram position and grating vector are evaluated for many candidate exposure settings [92], our method starts from the desired grating vector $\mathbf{G}^{v}$, recording geometry σ, recording wavelength λ and material shrinkage s, s_T and directly computes the corresponding robot targets.

The ray trace outlined in Sec. 5.2.2 yields the prism exit point $\mathbf{r}^{\mathcal{P}}_{b,\mathrm{air}}$, air-side propagation direction $\hat{\mathbf{s}}^{\mathcal{P}}_{b,\mathrm{air}}$, and polarization $\hat{\mathbf{p}}^{\mathcal{P}}_{b,\mathrm{air}}$ of each recording beam $b \in \{\mathrm{o}, \mathrm{r}\}$ in the prism frame $\mathcal{P}$. For readability, the dependence of these quantities on the selected coupler v and on $(\sigma, \lambda, s, s_T)$ is omitted in the following.

A Cartesian tool frame $\mathcal{T}_b$ is attached to each robot-mounted collimator, as shown in Fig. 5.6. The frame origin is located on the optical axis at the tip of the collimator. The tool-frame axis $\hat{\mathbf{x}}_{\mathcal{T}_b}$ points in the forward laser direction, while $\hat{\mathbf{z}}_{\mathcal{T}_b}$ is normal to the optics mounting plate. The polarization-maintaining fiber is mounted such that the linear polarization axis of the laser coincides with $\hat{\mathbf{z}}_{\mathcal{T}_b}$.

The tool-frame origin is placed a distance d_b from the prism exit point along the back-traced recording chief ray,

$$\mathbf{u}^{\mathcal{P}}_{\mathcal{T}_b} = \mathbf{r}^{\mathcal{P}}_{b,\mathrm{air}} + d_b\,\hat{\mathbf{s}}^{\mathcal{P}}_{b,\mathrm{air}}. \tag{5.1}$$

The tool-frame is oriented such that its optical axis launches the forward ray

$$\hat{\mathbf{x}}^{\mathcal{P}}_{\mathcal{T}_b} = -\hat{\mathbf{s}}^{\mathcal{P}}_{b,\mathrm{air}}, \tag{5.2}$$

and reproduces the required recording polarization

$$\hat{\mathbf{z}}^{\mathcal{P}}_{\mathcal{T}_b} = \chi_b \, \hat{\mathbf{p}}^{\mathcal{P}}_{b,\mathrm{air}}, \tag{5.3}$$

where $\chi_b = \pm 1$ reflects the sign ambiguity of linear polarization. This fully determines the tool orientation in the prism frame,

$$\mathbf{R}_{\mathcal{P}\mathcal{T}_b} = \left[\hat{\mathbf{x}}^{\mathcal{P}}_{\mathcal{T}_b} \; \hat{\mathbf{y}}^{\mathcal{P}}_{\mathcal{T}_b} \; \hat{\mathbf{z}}^{\mathcal{P}}_{\mathcal{T}_b}\right], \qquad \hat{\mathbf{y}}^{\mathcal{P}}_{\mathcal{T}_b} = \hat{\mathbf{z}}^{\mathcal{P}}_{\mathcal{T}_b} \times \hat{\mathbf{x}}^{\mathcal{P}}_{\mathcal{T}_b}. \tag{5.4}$$

Together with the tool-frame origin, this gives the affine transform from the tool frame to the prism frame,

$$\mathbf{T}_{\mathcal{P}\mathcal{T}_b} = \begin{pmatrix} \mathbf{R}_{\mathcal{P}\mathcal{T}_b} & \mathbf{u}^{\mathcal{P}}_{\mathcal{T}_b} \\ \mathbf{0}^{\mathsf{T}} & 1 \end{pmatrix}. \tag{5.5}$$

The six-axis robot servos command the position of the robot's flange frame $\mathcal{F}_b$, centered at the final mechanical mounting plate of the robot wrist, relative to the robot's base frame $\mathcal{B}_b$ according to the target transform $\mathbf{T}_{\mathcal{B}_b\mathcal{F}_b}$. The laser tool is mounted to the flange such that the tool and flange frames are connected by the static transform $\mathbf{T}_{\mathcal{F}_b\mathcal{T}_b}$, with their axes approximately coinciding, up to alignment errors, and their centers mechanically offset by $\mathbf{u}^{\mathcal{F}_b}_{\mathcal{T}_b} = (210, 0, 50)^{\mathsf{T}}\,\mathrm{mm}$. To improve laser pointing accuracy, we refine the laser tool axis $\hat{\mathbf{x}}^{\mathcal{F}_b}_{\mathcal{T}_b}$ orientation in the flange frame using the angular calibration described in Sec. 5.3.5.

The required flange pose in the robot base frame is therefore

$$\mathbf{T}_{\mathcal{B}_b\mathcal{F}_b} = \mathbf{T}_{\mathcal{B}_b\mathcal{P}} \, \mathbf{T}_{\mathcal{P}\mathcal{T}_b} \, \mathbf{T}^{-1}_{\mathcal{F}_b\mathcal{T}_b}. \tag{5.6}$$

Here $\mathbf{T}_{\mathcal{B}_b\mathcal{P}}$ is the fixed transform from the prism frame to the robot base frame. Its translation is obtained from the mechanical offsets described in Sec. 5.3.1, with its orientation defined by the angular calibration described in Sec. 5.3.5.

5.3.3 Inverse robot kinematics

The target transform $\mathbf{T}_{\mathcal{B}_b\mathcal{F}_b}$ does not uniquely determine the physical robot configuration. For each beam, the polarization sign $\chi_b = \pm 1$ gives two equivalent tool orientations, and the distance d_b between the collimator and prism interface can be chosen within the mechanically accessible range. For each resulting flange transform, the Stäubli TX2-60 inverse kinematics can have up to eight joint solutions,

$$\mathbf{j}_b^{(q)} = \mathrm{IK}_q(\mathbf{T}_{\mathcal{B}_b\mathcal{F}_b}(d_b, \chi_b)), \qquad q = 1, \ldots, 8. \tag{5.7}$$

The TX2-60 belongs to the class of six-axis industrial robots with an orthoparallel regional structure and a spherical wrist. We therefore use the analytical OPW inverse-kinematics formulation of Brandstötter *et al.* [93]. For the Stäubli TX2-60, the effective OPW parameters are

$$(a_1, a_2, b, c_1, c_2, c_3, c_4) = (0, 0, 20, 0, 290, 310, 70)\,\mathrm{mm}. \tag{5.8}$$

Here, the OPW parameter b is unrelated to the beam index b used above. The values c_3 and c_4 are the axis distances described in the TX2-60 data sheet [94]. The side projection in the data sheet shows no OPW side offsets, giving $a_1 = a_2 = 0$. The $375\,\mathrm{mm}$ vertical offset from the mounting plate to the first joint is not included in c_1, because the Stäubli base frame is centered on the first joint rather than on the robot bottom plate, as used in the convention of Brandstötter *et al.* The eight analytical branches are labeled using the Stäubli configuration flags "shoulder", "elbow", and "wrist". The relationship between these labels and the Brandstötter *et al.* solutions 1–8 is summarized in Appendix A.12 Tab. A.1.

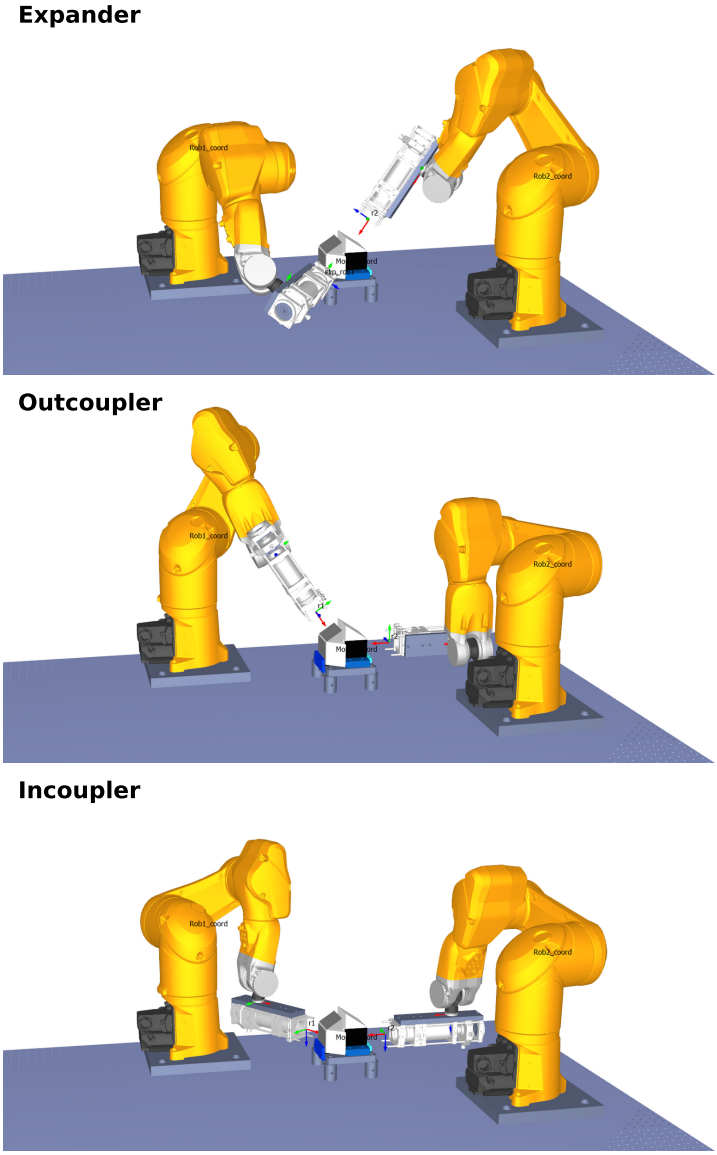


Figure 5.7: Robot poses for exposure of the coupler gratings of the 2D-EBE waveguide.

5.3.4 Robot path planning

The inverse optical pose calculation provides, for each coupler exposure, a family of admissible flange targets $\mathbf{T}_{\mathcal{B}_b\mathcal{F}_b}(d_b, \chi_b)$, parametrized by the tool distance d_b and the polarization sign $\chi_b = \pm 1$. For each target, the analytical inverse kinematics yields up to eight joint configurations. The remaining task is therefore a discrete path-planning problem: a collision-free joint configuration must be selected for each coupler exposure, and the selected configurations must be connected by executable, nonsingular robot trajectories.

This step is performed in a dedicated robot-planning environment [95, 96]. CAD models of the TX2-60 robots, collimation optics, prism mount, substrate holder, optical table, and relevant fixtures are imported into the planning scene. Candidate configurations are checked for collisions between the robots, recording optics, and surrounding hardware. Configurations that produce excessive twisting or unfavorable routing of the polarization-maintaining fibers are also rejected. Among the feasible candidates, we select paths with continuous joint motion and sufficient clearance around the prism system.

Figure 5.7 shows the recording poses for the expander, outcoupler, and incoupler holograms of the 2D-EBE waveguide introduced in Sec. 2.2.3, using the shrinkage-compensated decomposition parameters discussed in Sec. 5.2.2. The tool distance is $d_b = 80\,\mathrm{mm}$. These recording poses are collision-free and can be connected by continuous robot trajectories. The validated trajectories are exported from the planning environment and executed on the TX2-60 controllers using custom drivers.

5.3.5 Angular absolute calibration

For accurate hologram recording, the laser-axis direction in the flange frame, $\hat{\mathbf{x}}_{\mathcal{T}}^{\mathcal{F}}$, and the rotation from the prism frame to the robot base frame, $\mathbf{R}_{\mathcal{BP}}$, must be absolutely calibrated. The calibration is performed independently for the two robot-mounted laser heads.

The twin-prism pair is fabricated with an angular tolerance of one arcminute. The outward normal $\hat{\mathbf{N}}_f^{\mathcal{P}}$ of each facet f is therefore known to within 0.02° and serves as a calibration reference. The prisms are mechanically contacted during calibration. The robot is repeatedly jogged manually to different facets f until the output laser beam is normal to the prism surface, as determined by the laser back-reflection. If the laser is incident at an angle α with respect to the surface normal, the reflected beam re-enters the collimator at an angle 2α with respect to the collimator optical axis, see Fig. 5.6. As derived in Appendix A.11, it is then refocused onto the fiber-output plane with a lateral offset of $h = -\alpha \cdot 1200\,\mathrm{mm}$, with α in radians for the optics described in Sec. 5.3.1. The smallest visually resolvable displacement from the optical axis is estimated to be approximately $0.3\,\mathrm{mm}$, allowing the laser to be adjusted to be normal to the prism surface to within 0.02°.

Since the laser is always directed toward the prism surface, normal incidence requires the laser direction to be antiparallel to the outward normal. For each measured pose, the robot forward kinematics gives the flange rotation $\mathbf{R}_{\mathcal{BF},f}$. The measured laser direction in the robot base frame is therefore

$$\hat{\mathbf{x}}_f^{\mathcal{B}} = \mathbf{R}_{\mathcal{BF},f}\hat{\mathbf{x}}_{\mathcal{T}}^{\mathcal{F}}, \tag{5.9}$$

whereas the corresponding inward prism normal is

$$-\hat{\mathbf{N}}_f^{\mathcal{B}} = -\mathbf{R}_{\mathcal{BP}}\hat{\mathbf{N}}_f^{\mathcal{P}}. \tag{5.10}$$

The angular residual for pose f is

$$\varepsilon_f = \left|\mathrm{atan2}\left(\left\|\hat{\mathbf{x}}_f^{\mathcal{B}} \times \left(-\hat{\mathbf{N}}_f^{\mathcal{B}}\right)\right\|, \hat{\mathbf{x}}_f^{\mathcal{B}} \cdot \left(-\hat{\mathbf{N}}_f^{\mathcal{B}}\right)\right)\right|. \tag{5.11}$$

The calibration parameters are obtained by minimizing the sum of squared angular residuals,

$$\left(\hat{\mathbf{x}}_{\mathcal{T}}^{\mathcal{F}}, \mathbf{R}_{\mathcal{BP}}\right) = \underset{\hat{\mathbf{x}}_{\mathcal{T}}^{\mathcal{F}}, \mathbf{R}_{\mathcal{BP}}}{\mathrm{argmin}} \sum_{f=1}^{N} \varepsilon_f^2. \tag{5.12}$$

Table 5.1: Angular calibration of the two laser tools. Face labels refer to Fig. 5.6, numbers in brackets indicate repeated measurements on the same face in different robot poses.

Arm	Faces used	N	Mean(ε_f)	Max(ε_f)
Robot 1	$F_1(3), F_2(3), F_3(2)$	8	0.07°	0.11°
Robot 2	$F_4(3), F_5(5), S_2(2)$	10	0.09°	0.21°

In the implementation, $\mathbf{R}_{\mathcal{BP}}$ is parametrized by three Euler angles, and $\hat{\mathbf{x}}_{\mathcal{T}}^{\mathcal{F}}$ is represented by a normalized three-vector. The resulting calibration quality is reported by the mean and maximum values of ε_f in Tab. 5.1. The maximum residual angular error, which is in the 0.2° range, is dominated by errors in the robot kinematic chain, such as robot assembly tolerances. The angular precision, or repeatability, of the robots is better than can be quantified with the present back-reflection method.

The proposed calibration determines the absolute laser direction but not the optical-axis roll. The resulting polarization error in the robot pose is set by the mechanical alignment tolerance of the collimator assembly and has only a minor effect on recording fringe contrast.

5.3.6 Interferometric stability

Mechanical jitter of the robot arms or optical-path fluctuations in the fibers can translate the recording interference pattern during exposure. To evaluate the stability of the proposed platform, both robot-mounted collimators are positioned approximately normal to a $50\,\mathrm{mm}$ beamsplitter cube, as shown in Fig. 5.8(top-left). The beamsplitter is near the edge of the reachable workspace of both robot arms and therefore represents a mechanically unfavorable recording configuration with both arms fully extended. The irises of both collimation optics are fully opened to $50\,\mathrm{mm}$. The two beams are manually overlapped at a small mutual angle, producing carrier fringes across the full aperture. The interference pattern is imaged onto a beam profiler (DataRay WinCamD-LCM) after demagnification with a Galilean telescope consisting of $-40\,\mathrm{mm}$ and $200\,\mathrm{mm}$ lenses, as shown in

Fig. 5.8(top-right). The recorded intensity can be written as a slowly modulated carrier-fringe pattern,

$$I(\mathbf{r},t) = I_0(\mathbf{r}) + I_m(\mathbf{r})\cos[\mathbf{k}_0 \cdot \mathbf{r} + \varphi(\mathbf{r},t)]\,, \tag{5.13}$$

where $\mathbf{k}_0$ is the spatial carrier frequency and $\varphi(\mathbf{r},t)$ is the relative phase of the two recording beams. Following the Fourier-transform method of Takeda *et al.* [97], one carrier sideband is isolated in Fourier space, shifted to baseband, and inverse-transformed to obtain a complex fringe field $c(\mathbf{r},t)$. The local fringe phase is then given by $\arg c(\mathbf{r},t)$. Details of the phase-retrieval procedure are given in Appendix A.13. The global phase drift relative to the first recorded frame is calculated from the complex inner product over the beam-overlap region Ω,

$$\Delta\varphi(t_n) = \arg \sum_{\mathbf{r}\in\Omega} c^*(\mathbf{r},0)\, c(\mathbf{r},t_n). \tag{5.14}$$

The unwrapped phase trace in Fig. 5.8(bottom) corresponds to a relative optical-path variation on the order of $\pm\lambda/20$ over a 10 s exposure, corresponding to a

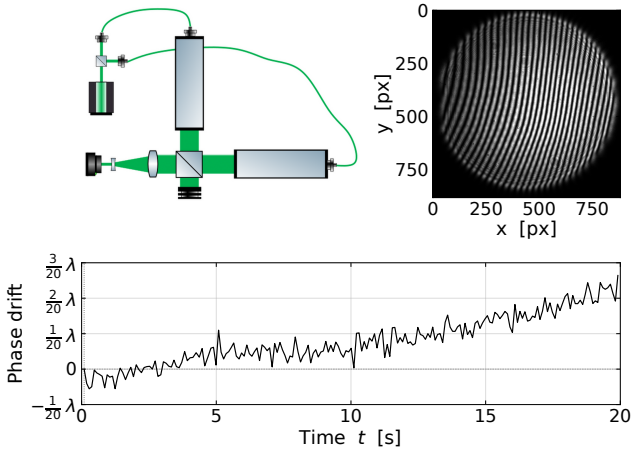


Figure 5.8: Top-left: Interferometer for stability measurement. Top-right: Recorded interference pattern during stability measurement. Bottom: Phase drift extracted over 209 frames.

recording-fringe visibility above $95\,\%$ [98]. The measured stability is therefore sufficient for the exposure times of less than $10\,\mathrm{s}$ used in Sec. 5.4.

5.4 Waveguide sample fabrication

To demonstrate the automated recording workflow, we fabricate a 2D-EBE waveguide with a 56° propagation angle and a 90° expander redirection angle, as introduced in Sec. 2.2.3 and shown in Fig. 2.3.

Development-grade Bayfol photopolymer film [49] is laminated onto $1.1\,\mathrm{mm}$-thick Corning EAGLE XG glass wafers with dimensions of $60\,\mathrm{mm}\times80\,\mathrm{mm}$. The photopolymer is first cut to match the wafer dimensions and then applied in a cleanroom using a double-roller laminator. All handling of the unexposed film is performed under dark-red to near-infrared illumination to minimize unintended pre-exposure.

The laminated substrates are clamped between the twin-prisms using $0.375\,\mathrm{mm}$-thick polycarbonate spacer foils on both sides. The spacers are cut in a U-shape to permit filling of the prism gap with Zeiss Immersol-F immersion fluid from the top. The liquid is injected with a syringe until the full contact region is covered and no visible air inclusions remain. This step is essential because trapped bubbles or inhomogeneous oil films would perturb the recording wavefronts and reduce hologram quality.

No exposure mask or pre-exposure is applied. Both recording beams are adjusted to a diameter of $14\,\mathrm{mm}$ by closing the iris of the collimation optics. The exposure is then executed automatically by the robot controllers, following the recording poses for the expander, outcoupler, and incoupler shown in Fig. 5.7 and the 2D-EBE ray traces in Fig. 5.5. At the recording poses, the $532\,\mathrm{nm}$ laser beams are normal to the prism facets, consistent with the approximately $3\,\%$ normal shrinkage compensation used in Sec. 5.2.2 for typical Bayfol films [56].

The incoupler and outcoupler are exposed for $6\,\mathrm{s}$ with a nominal irradiance of $1.5\,\mathrm{mW\,cm^{-2}}$ per beam, measured in a plane normal to each beam. For a beam of power P, normal cross-sectional area A, and irradiance $I = P/A$ incident at an angle θ_p to the sample normal, the illuminated sample area is $A_p = A/\cos\theta_p$. Equivalently, $A = A_p \cos\theta_p$, and the irradiance in the sample plane is $I_p = P/A_p = I\cos\theta_p$. With one normally incident beam and one beam at $\theta_p = 56°$, the total dose is $tI(1 + \cos\theta_p) \approx 14\,\mathrm{mJ\,cm^{-2}}$. The expander, recorded with two oblique beams, is exposed $30\,\%$ longer, giving a total dose of $1.3\,t\,I(2\cos\theta) \approx 13\,\mathrm{mJ\,cm^{-2}}$. After exposure, the sample remains in the prism assembly and is cured for $15\,\mathrm{min}$ under intense white LED illumination. A second $15\,\mathrm{min}$ curing step is then carried out under UV illumination. Finally, the sample edges are coated with black UV-curing nail polish to suppress parasitic back-reflections from the waveguide facets.

5.5 Sample validation

5.5.1 Scope of validation sample

The fabricated waveguide is used as an end-to-end process validation sample. Its purpose is to close the experimental chain from grating vector design and robot-pose calculation to automated exposure and waveguide display readout. It is not intended as an optimized near-eye-display demonstrator.

For this reason, the sample uses a deliberately simple implementation: circular, unmasked coupler apertures, a thick $1.1\,\mathrm{mm}$ carrier substrate chosen for robust handling during fabrication, and a non-miniaturized laboratory projector and simple relay as the image source. These choices affect the measured image quality. The circular coupler footprints limit the usable field and eyebox, the substrate thickness increases the guided-ray jump length and makes pupil-sampling modulation visible, and the projector/relay optics impose their own resolution limit and do not provide built-in angular efficiency compensation.

Consequently, the MTF, field uniformity, and collected throughput reported below are system-level validation metrics. They demonstrate that the automatically recorded incoupler, expander, and outcoupler operate together as a complete 2D eyebox-expanding waveguide. They should not be compared directly with optimized near-eye-display demonstrators that use designed coupler apertures, appropriately thin substrates, optimized projection engines, and dedicated electronic uniformity correction. The relevant validation criterion here is the functionality of the automated mastering platform.

5.5.2 Imaging test bench

The recorded 2D-EBE waveguide sample is characterized using the imaging test bench shown in Fig. 5.9. An RGB laser projector, originally designed for screen projection, is used as a computer-controlled image source. The projector output is relayed by two achromatic doublets with focal lengths of $40\,\mathrm{mm}$ and $50\,\mathrm{mm}$. The $40\,\mathrm{mm}$ lens is placed as close to the projector as mechanically possible, and the lens separation is adjusted until the output image is approximately afocal.

The relayed image is first inspected directly with a $25\,\mathrm{mm}$ effective-focal-length objective with an $8\,\mathrm{mm}$ entrance pupil (EO Techspec 33892) and imaged onto an 8-bit RGB camera sensor with a $3.45\,\mu\mathrm{m}$ pixel pitch and 2448×2048 pixels

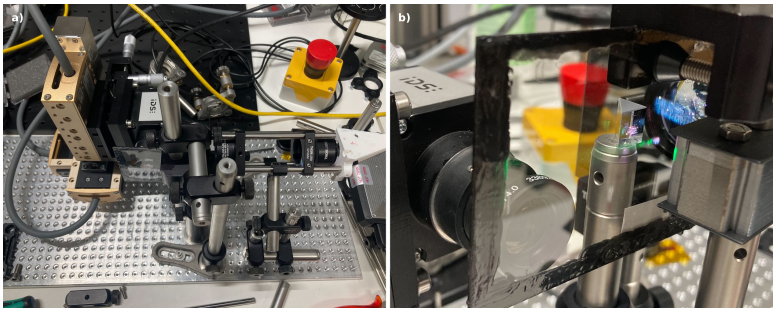


Figure 5.9: Left: Test bench with projector and translatable camera. Right: Active projector relayed through the waveguide.

(IDS U3-3280SE-C-HQ). The relay is adjusted until the projector scan lines are clearly resolved in the center of the image. The relay output is then redirected by a prism mirror and made incident on the waveguide incoupler. The incoupled spot is positioned near the edge of the incoupling hologram adjacent to the expander hologram. This placement avoids a second interaction of the guided angular spectrum with the incoupler.

The same camera and objective are mounted on an automated XYZ translation stage, allowing the output image to be recorded at different pupil positions. For the $25\,\mathrm{mm}$ objective, 1° corresponds to approximately 126 camera pixels. The recorded images are cropped to 2048×2048 pixels, corresponding to an angular field of approximately $16^\circ \times 16^\circ$.Only the green projector channel at $520\,\mathrm{nm}$ is efficiently coupled by the recorded volume holograms. Unless stated otherwise, images are acquired with an ND 0.5 filter placed between the projector and the waveguide, an exposure time of $50\,\mathrm{ms}$, and a camera gain of 1.

5.5.3 Intensity uniformity

We first evaluate the intensity uniformity across the field by displaying a uniform green image and recording the output image at the center of the outcoupler at an eye relief of $5\,\mathrm{mm}$. Two distinct intensity variations are observed. The first is a slowly varying angular envelope caused by the finite Bragg bandwidths of the three recorded holograms. The second is a high-spatial-frequency grid modulation caused by pupil sampling.

5.5.3.1 Hologram bandwidth and display correction

Figure 5.10 (left) shows the corresponding camera capture. The dynamic range of the camera is insufficient to capture the full intensity variation without saturation at the field center. The observed intensity roll-off to the field edges is expected because the total waveguide throughput is determined by the angular

Bragg efficiencies of the coupler gratings. The gratings were recorded for maximum efficiency near the chief ray, so the uncorrected image is brightest near the center of the field.

To compensate this slowly varying intensity envelope, we apply a display-space correction mask to the green projector channel. The mask is constructed from manually adjusted scalar weights that are interpolated to the full projector resolution using a smooth piecewise-cubic interpolation and then multiplied by the displayed 8-bit green-channel image. The minimum value of the correction mask is $5\,\%$. The corrected camera capture is shown in Fig. 5.10(right). Saturation in the image center is removed, while the higher-frequency modulation remains.

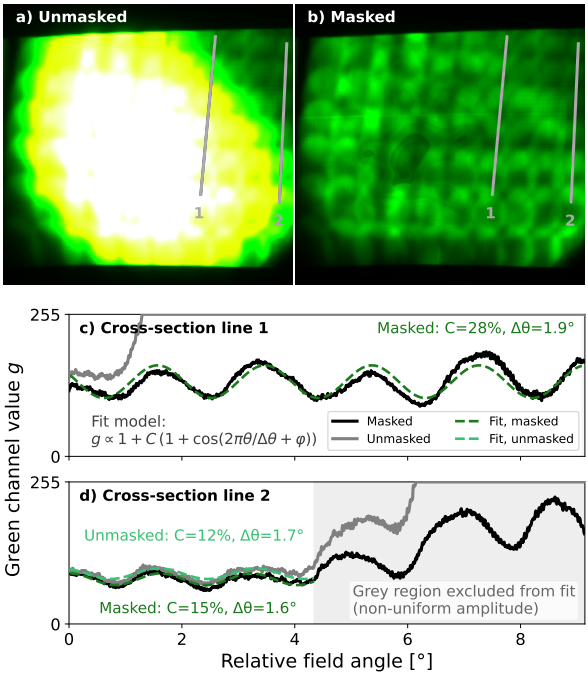


Figure 5.10: Camera captures of a uniform green image (a) before and (b) after display-space intensity correction. The numbered lines indicate the cross sections along which the plots below show the 8-bit green-channel values with fits to the first-order modulation model.

5.5.3.2 Keyhole pupil sampling effect

The residual grid-like intensity modulation in Fig. 5.10(right) shifts when the camera pupil is translated. We attribute this modulation to sparse pupil sampling in the 2D-EBE waveguide. The guided beam generates a discrete two-dimensional lattice of outcoupled replicas, separated by the guided-ray jump length. Each replica acts as a finite "keyhole" through which the observation pupil receives light from the angular component of the image corresponding to the respective guided beam. Because the replica lattice changes with launch angle, neighboring field angles can place different numbers of beam replicas inside the finite pupil. The resulting field-dependent sampling produces the observed grid-like modulation. This artifact has been reported previously [99].

We model this effect analytically in Appendix A.14. The incident beam at the incoupler is modeled as a circular top-hat beam of diameter B, and the camera pupil as a circular aperture of diameter D. Neglecting HOE angular selectivity, guided-mode depletion, and finite coupler apertures, the leading field-center modulation is

$$\begin{aligned} I(\theta_x,\theta_y) &\propto 1 + C\left[\cos\left(\frac{2\pi}{\Delta\theta}\theta_x + \varphi_x\right) + \cos\left(\frac{2\pi}{\Delta\theta}\theta_y + \varphi_y\right)\right], \\ C &= 2\left|\mathrm{jinc}\left(\frac{\pi B}{L_0}\right)\mathrm{jinc}\left(\frac{\pi D}{L_0}\right)\right|, \qquad \Delta\theta = \frac{L_0}{s_{\mathrm{eye}} + L_{\mathrm{eff}}}. \end{aligned} \tag{5.15}$$

Here $L_0 = 2d_{\mathrm{wg}}\tan\theta_p$ is the same-surface jump length of the central guided ray, s_{eye} is the eye relief, and L_{eff} is an effective angular lever arm for the field-angle-dependent displacement of the outcoupling lattice. The contrast is suppressed when the incident beam diameter or the pupil diameter is comparable to L_0.

In the present setup, the relay images the scanner beam onto the incoupler with an estimated diameter of $B \approx 1\,\mathrm{mm}$, while the camera objective has a pupil diameter of $D \approx 8\,\mathrm{mm}$. For $d_{\mathrm{wg}} = 1.1\,\mathrm{mm}$ and $\theta_p = 56°$, the central jump length is $L_0 \approx 3.26\,\mathrm{mm}$, giving a predicted contrast $C \approx 8\,\%$. For the approximately symmetric pupil displacement of the camera from the incoupler in the experiment, with a transverse distance $r_I \approx 30\,\mathrm{mm}$, the lever arm can be estimated from

the refractive index $n = 1.51$ as $L_{\text{eff}} \approx r_I (n \sin\theta_p)^{-1}(1 + \sec^2\theta_p)$. The model predicts $\Delta\theta_{\text{grid}} \approx 1.8°$. To compare these predictions with the camera captures, we extract two cross sections of the green channel along ridges of the grid modulation, such that the orthogonal cosine term in (5.15) is approximately unity, see Fig. 5.10. Assuming that the 8-bit green-channel value g is proportional to the optical intensity within the unsaturated range, we fit (5.15) to the cross sections. For the display-space-corrected capture, the fitted periods are $1.9°$ and $1.6°$, with modulation contrasts of $28\,\%$ and $15\,\%$, respectively. The fits are restricted to regions in which the modulation amplitude is approximately constant, marked with white background in Fig. 5.10, since the first-order model does not account for amplitude variations caused by the finite VPH angular bandwidth. The manually adjusted display-space correction does not perfectly homogenize the modulation amplitude along cross section 2. Although the green-channel signal in the uncorrected capture is saturated over most of the field, the unsaturated field-edge region along cross section 2 exhibits a comparable grid-modulation period and contrast before and after correction, see Fig. 5.10(d). The measured grid periods agree well with the first-order model. We attribute the quantitative discrepancy in modulation contrast to the stated idealized model assumptions and neglected higher-order contributions.

The grid modulation effect is most pronounced for scanned-beam illumination because the incident beam diameter is constrained by the scanner aperture. Larger input beam diameters, as in display-panel illumination, larger observation pupils, and thinner waveguides suppress the modulation contrast. Thinner waveguides also reduce the modulation period.

5.5.4 Image transfer and resolution assessment

Figure 5.11 shows representative camera captures of a checkerboard pattern, a natural image, and a text image displayed through the recorded waveguide with the display-space intensity correction applied. The text is readable and the bright laboratory background remains visible through the transparent waveguide.

To quantify image sharpness, checkerboard targets are recorded in a dark laboratory at the center pupil position and at four displaced pupil positions ($\pm 3\,\mathrm{mm}$). Edge-spread functions are evaluated near the field center and near the four field corners ($\pm 5^\circ, \pm 5^\circ$), using edges parallel and orthogonal to the projector scanline direction. The modulation transfer functions (MTFs) are extracted using the slanted-edge method [100] and shown in Fig. 5.12. The selected edge regions and all extracted MTF_{20} values are given in Table A.2 of Appendix A.15. Spatial frequencies are first calculated in $\mathrm{lp/mm}$ on the camera sensor and then converted to $\mathrm{lp/deg}$ using the known pixel pitch and objective focal length. With the waveguide present, the field-center MTF_{20} remains between 9–11 $\mathrm{lp/deg}$. Across all evaluated field points and pupil positions, the recorded waveguide sample gives a mean value of $8\ \mathrm{lp/deg}$. A direct reference capture without the waveguide gives

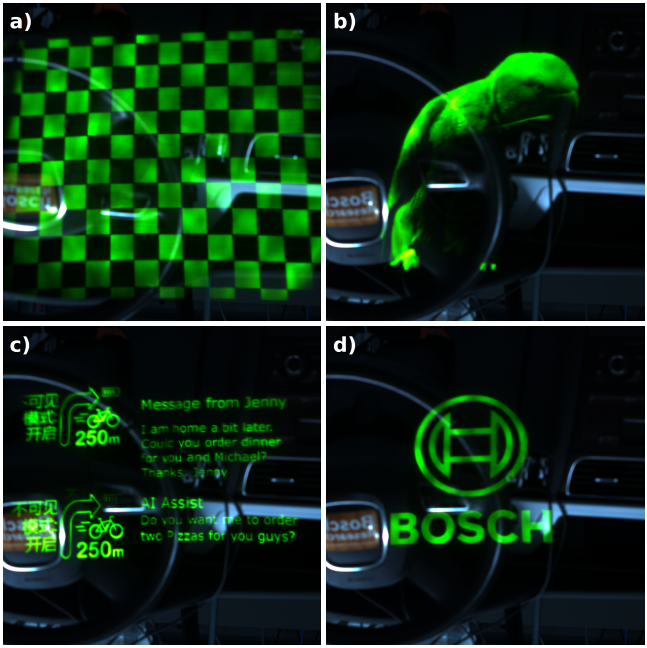


Figure 5.11: Representative images displayed through the waveguide display with projector intensity correction. Captures cover $16^\circ \times 16^\circ$.

8–16 $\mathrm{lp/deg}$ across the evaluated field points, showing that the projector and relay optics alone significantly contribute to the total MTF. The lowest waveguide values occur at outer eyebox and off-axis positions, where the raw captures show strong residual intensity modulation and field-dependent clipping by the circular coupler apertures. The reported MTF_{20} values should therefore not be interpreted as the transfer function intrinsic to the holographic coupler gratings alone, nor as a performance limit of the mastering platform.

5.5.5 Optical throughput

The optical throughput of the validation sample is measured with a power meter using an $18\,\mathrm{mm}\times18\,\mathrm{mm}$ sensor aperture (Thorlabs S170C). For a full green image without (with) the display-space correction mask, the incident power on the

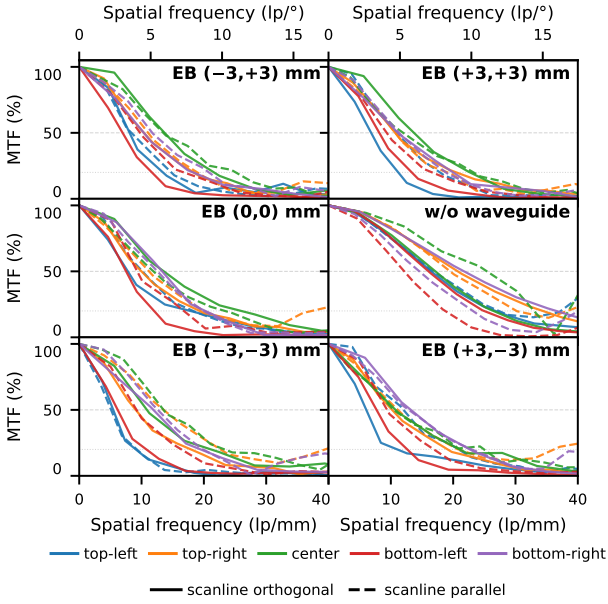


Figure 5.12: MTF results without and with waveguide at different eyebox and field of view positions.

incoupler is $P_A = 4.5\,\mathrm{mW}$ $(1.9\,\mathrm{mW})$. The directly transmitted, non-incoupled power is $P_B = 3.9\,\mathrm{mW}$ $(1.65\,\mathrm{mW})$, giving an angular-spectrum-averaged incoupled fraction $\bar{\eta}_i = (P_A - P_B)/P_A \approx 13\,\%$ in both cases. The total output power across the entire outcoupler is $P_C = 27\,\mu\mathrm{W}$ $(2.5\,\mu\mathrm{W})$, corresponding to a collected total throughput $P_C/P_A = 0.6\,\%$ $(0.13\,\%)$.

5.6 Conclusion

We have demonstrated, to our knowledge, the first automated recording platform for holographic waveguide displays based on two six-axis robot-mounted collimation heads and a custom edge-lit prism system. The platform calculates exposure poses directly from target grating vectors, including shrinkage compensation and polarization alignment, using inverse ray tracing and robot kinematics. This workflow allows incoupler, expander, and outcoupler gratings to be recorded sequentially in a single fixed, immersion-contacted substrate, avoiding repeated prism contacting and post-recording relamination of separate HOEs.

The experimental characterization confirms that robot-mounted optics can provide the angular control and interferometric stability required for holographic exposure. Prism-facet back-reflection calibration gives mean angular absolute accuracy better than 0.1° for both robot arms. Interferometric measurements showed a relative optical-path variation on the order of $\pm\lambda/20$ over typical exposure times, corresponding to high recording stability and integrated fringe contrast exceeding $95\,\%$. The complete fabrication chain was validated by recording and characterizing a 2D eyebox-expanding waveguide sample. The sample transmits images and readable text, and we quantified its intensity uniformity, MTF, and collected throughput. The measured grid-modulation period agrees well with our analytical model for sparse pupil sampling, while the higher measured contrast indicates contributions beyond the first-order prediction.

The chapter establishes and validates a reproducible route from planar waveguide design to automated holographic recording.

6 Conclusion and outlook

6.1 Conclusion

This thesis introduced a concept for lightweight, prescription-compatible AR glasses based on spherical holographic waveguide displays. Section 2.3 quantified the substantial weight reduction achievable by integrating spherical waveguides into prescription-lens stacks, relative to planar and cylindrical waveguides. The work made three contributions toward the broader goal of realizing such spherical waveguide displays.

On the design side, this thesis established the first non-paraxial aberration-correction framework for spherical waveguides. It was shown that the exact eigenmodes underlying previous planar and cylindrical waveguide-display designs do not exist in spherical waveguides. Nevertheless, local consistency and closure were demonstrated to provide weaker mode properties that remain sufficient for coupler construction. A complete design for a spherical waveguide display with 2D eyebox expansion was presented, and its imaging performance was verified using non-sequential ray tracing.

The subsequent analysis investigated whether immersion contacting and edge-lit exposure could be avoided by shifting the recording wavelength, thereby simplifying the prototyping of such waveguide displays. Analytical recordability limits were derived in terms of the local tangential and normal components of the grating vector. Practically relevant expander gratings for waveguides with 2D eyebox expansion were shown to lie outside the immersion-free recording range of typical photopolymers. This included the expander required for the spherical-waveguide design presented in this thesis.

An automated edge-lit recording platform was developed to establish a controlled and reproducible fabrication process for planar holographic waveguides. Two six-axis robotic arms positioned and oriented the recording beams at poses calculated directly from the prescribed waveguide coupler grating vectors, with compensation for material shrinkage. A custom-designed edge-lit prism assembly provided simultaneous access to the recording angles required for the incoupler, expander, and outcoupler gratings. The complete fabrication workflow was demonstrated by producing a functioning planar waveguide display with 2D eyebox expansion. Measurements of the MTF and intensity uniformity confirmed the functionality of the fabricated sample.

Taken together, these results indicate that the principal remaining barriers to spherical holographic waveguide displays lie in fabrication rather than optical design. This thesis established that prescription-compatible spherical waveguides can be designed with imaging performance approaching the visual-acuity limit, while also demonstrating that their spatially varying holographic expander gratings require edge-lit recording in available photopolymers. The automated recording platform developed for planar waveguides provides a controlled basis for investigating critical fabrication parameters, including immersion-liquid contacting, stray-light suppression, material shrinkage, and exposure dose. It therefore constitutes a concrete step toward a reproducible fabrication process for spherical waveguide displays.

6.2 Outlook

Fabricating spherical waveguide displays requires a recording system capable of generating accurate freeform recording wavefronts for the spatially varying coupler HOEs. One possible approach is to extend the proposed edge-lit prism architecture for planar waveguides by cementing two plano-freeform refractive beam-shaping elements for each coupler to the corresponding prism entrance facets. The external recording beams could remain normally incident on the prism facets, while the freeform elements transform them into the spatially varying object and reference

wavefronts required at the curved holographic substrate. This architecture would integrate naturally with the existing robotic exposure platform.

The recording freeforms could be calculated from the prescribed spherical-waveguide coupler grating vector fields using a modified version of the algorithm developed by Jang *et al.* [47], adapted to account for curved hologram surfaces and intermediate prism interfaces. A major challenge will be achieving sufficient accuracy in the freeform optics while accounting for hologram shrinkage. To obtain visual-acuity-limited performance, the outcoupled display rays must have angular errors on the order of 1 arcmin. This corresponds to a required local accuracy on the order of $0.01\,\%$ in the tangential grating vector component and sub-nanometer tangential grating period accuracy.[1] Meeting these tolerances will require a detailed understanding of holographic-material shrinkage, which may be inferred approximately from measurements of planar waveguide gratings.

Achieving this level of accuracy will likely require a closed-loop process comprising exposure, measurement, and adaptation of the recording freeforms. Existing holographic-grating metrology methods are typically scatterometric and rely on measurements of the transmitted or diffracted field over a range of incidence angles and wavelengths [101–105]. Characterizing waveguide coupler gratings is complicated by their tendency to diffract incident probe light into guided modes. First, the diffracted beam is inaccessible to a free-space receiver. Second, the guided light can interact with the grating again, producing secondary outcoupled fields that overlap the transmitted probe field and thereby bias the measured transmission signal. Immersion-based mitigation methods have been proposed [105]. Most existing methods are pointwise, with the notable exception of the imaging-based transmission-measurement approach developed by Steckmann *et al.* for free-space holograms [106].

[1] For normal outcoupling of a ray propagating at $\theta_p = 56^\circ$ in a waveguide with refractive index $n = 1.51$, tangential momentum matching requires $|G_T| = nk_0 \sin\theta_p$, where $k_0 = 2\pi/\lambda$. A relative grating vector error ε produces a residual tangential output wave vector $\Delta k_T = \varepsilon G_T$. Because the tangential component is conserved upon refraction into air, $\sin\Delta\theta_{\mathrm{air}} = \Delta k_T/k_0 = \varepsilon n \sin\theta_p$. For $\varepsilon = 0.01\,\%$, this yields $\Delta\theta_{\mathrm{air}} = \arcsin(10^{-4} \cdot 1.51 \sin 56^\circ) \approx 0.43$ arcmin.

Preliminary results have been obtained for waveguide gratings by extending this approach to immersion-based multispectral probing. A tunable supercontinuum source is expanded and collimated to illuminate a large region of a waveguide sample immersed in a tank with two optical windows. A bi-telecentric imaging system records the transmitted intensity as a function of position on the sample. As the illumination wavelength is varied, regions satisfying the local Bragg condition diffract light out of the transmitted beam and therefore appear darker in the recorded image. Repeating the measurement for several angular orientations of the sample produces a multidimensional transmission data set as a function of spatial position, wavelength, and incidence angle. Fitting the local spectral and angular responses using coupled-wave expressions could then be used to recover the spatially resolved grating vector components, hologram thickness, and local refractive-index modulation. In principle, this method is applicable to both planar and moderately curved waveguide substrates.

Realizing spherical waveguide displays will require further advances in both exposure and metrology methods. In particular, achieving the required exposure accuracy remains a major engineering challenge. The methods and results presented in this thesis provide a foundation for addressing these challenges.

A Appendix

A.1 Derivation of the grating vector decomposition

All decompositions into reference and object wave vectors $\mathbf{k}_r$ and $\mathbf{k}_o$ must satisfy

$$\mathbf{G}^{\mathrm{irr}} = \mathbf{k}_o - \mathbf{k}_r, \qquad |\mathbf{k}_r| = |\mathbf{k}_o| = k = 2\pi n/\lambda. \tag{A.1}$$

The reference wave vector $\mathbf{k}_r$ of any solution may be represented as a linear combination of $\mathbf{G}^{\mathrm{irr}} \neq \mathbf{0}$ and some unit vector $\hat{\mathbf{\Delta}}$ orthogonal to $\mathbf{G}^{\mathrm{irr}}$.

$$\mathbf{k}_r = \mu\mathbf{G}^{\mathrm{irr}} + \Delta\hat{\mathbf{\Delta}} \tag{A.2}$$

At this stage, μ, Δ, and the orientation of $\hat{\mathbf{\Delta}}$ parametrize the possible decompositions. Enforcing equal wave vector magnitudes constrains these quantities. Inserting Eq. (A.2) into Eq. (A.1), one finds the object vector corresponding to this particular solution

$$\mathbf{k}_o = (1+\mu)\mathbf{G}^{\mathrm{irr}} + \Delta\hat{\mathbf{\Delta}} \tag{A.3}$$

Enforcing the dispersion relation Eq. (A.1), one finds

$$\mathbf{k}_r: \quad \left(\frac{2\pi n}{\lambda}\right)^2 = \mu^2\left(G^{\mathrm{irr}}\right)^2 + \Delta^2 \tag{A.4}$$

$$\mathbf{k}_o: \quad \left(\frac{2\pi n}{\lambda}\right)^2 = (1+\mu)^2\left(G^{\mathrm{irr}}\right)^2 + \Delta^2. \tag{A.5}$$

Subtracting Eq. (A.4) from Eq. (A.5) and assuming physical gratings with $G^{\mathrm{irr}} \neq 0$, one finds $\mu = -1/2$. Inserting into Eq. (A.4) and solving the quadratic equation, one finds $\Delta = \pm\sqrt{k^2 - (G^{\mathrm{irr}})^2/4}$. Thus, $|\Delta| = \sqrt{k^2 - (G^{\mathrm{irr}})^2/4}$. This is true for every possible decomposition of the grating vector into reference and object waves. All possible solutions must have the form

$$\begin{aligned} \mathbf{k}_r &= -\frac{\mathbf{G}^{\mathrm{irr}}}{2} + \sqrt{k^2 - \frac{(G^{\mathrm{irr}})^2}{4}}\hat{\mathbf{\Delta}} \\ \mathbf{k}_o &= \frac{\mathbf{G}^{\mathrm{irr}}}{2} + \sqrt{k^2 - \frac{(G^{\mathrm{irr}})^2}{4}}\hat{\mathbf{\Delta}} \end{aligned} \tag{A.6}$$

The unit vector $\hat{\mathbf{\Delta}}$ is arbitrary up to its orthogonality with $\mathbf{G}^{\mathrm{irr}}$. Without loss of generality, it may be parametrized by the angle σ with

$$\hat{\mathbf{\Delta}}(\sigma) = \cos(\sigma)\frac{1}{G^{\mathrm{irr}}}\left(G_T^{\mathrm{irr}}\hat{\mathbf{N}} - G_N^{\mathrm{irr}}\hat{\mathbf{T}}\right) - \sin(\sigma)\hat{\mathbf{N}} \times \hat{\mathbf{T}} \tag{A.7}$$

The entire solution space for the decomposition is therefore parametrized by a single scalar parameter σ. This angle may be restricted to $[0, 2\pi)$. Other values do not add distinct solutions. This proves the completeness of the decomposition given in [47] and in Eq. (A.6).

A.1.1 Symmetry of $[0, \pi)$ and $[\pi, 2\pi)$ configurations

While the angles $\sigma \in [\pi, 2\pi)$ produce different vectors $\hat{\mathbf{\Delta}}$ from those obtained for $\sigma \in [0, \pi)$, they differ only in sign. Inspecting the expressions in Eq. (A.6), one finds

$$\hat{\mathbf{k}}_r(\sigma + \pi) = -\hat{\mathbf{k}}_o(\sigma), \qquad \hat{\mathbf{k}}_o(\sigma + \pi) = -\hat{\mathbf{k}}_r(\sigma). \tag{A.8}$$

Therefore, the $[\pi, 2\pi)$ angles correspond to recording geometries that are merely the recording geometries generated by $[0, \pi)$ mirrored around the origin with switched labeling.

A.2 Volume diffraction derivations

A.2.1 Discrete symmetry

In general, the diffracted field can be written as a Fourier transform, that is, an integral over plane waves with wave vector $\mathbf{k}_d$. The volume grating is invariant under translation by one surface period $\Lambda_T \hat{\mathbf{T}}$ in direction $\hat{\mathbf{T}}$, defined in Sec. 2.1.2, and fully invariant under translations in direction $\hat{\mathbf{D}} = \hat{\mathbf{N}} \times \hat{\mathbf{T}}$. The relative phase of an incident wave with wave vector $\mathbf{k}_i$ and the diffracted field must therefore be invariant modulo 2π under such translations. Therefore, for each Fourier component $\mathbf{k}_d$ of the diffracted field $(\mathbf{k}_d - \mathbf{k}_i) \cdot \hat{\mathbf{T}}\, \Lambda_T = 2\pi m$, or, using $G_T = 2\pi/\Lambda_T$, $\hat{\mathbf{T}} \cdot (\mathbf{k}_d - \mathbf{k}_i) = m G_T$ and therefore $k_{d,T} = k_{i,T} + m G_T$. Similarly, one finds for an arbitrary translation in $\hat{\mathbf{D}}$ that $\hat{\mathbf{D}} \cdot (\mathbf{k}_d - \mathbf{k}_i)$ and therefore $k_{d,D} = k_{i,D}$. The tangential component of the diffracted wave vector is

$$\begin{aligned} \mathbf{k}_{d,\parallel} &= \mathbf{k}_d - \left(\hat{\mathbf{N}} \cdot \mathbf{k}_d\right) \hat{\mathbf{N}} = k_{d,T} \hat{\mathbf{T}} + k_{d,D} \hat{\mathbf{D}} \\ &= \left(k_{i,T} + m G_T\right) \hat{\mathbf{T}} + k_{i,D} \hat{\mathbf{D}} = \mathbf{k}_{i,\parallel} + m G_T \hat{\mathbf{T}} \end{aligned} \tag{A.9}$$

This is the tangential momentum-matching condition Eq. (2.12) and explains the discrete diffraction orders. This derivation is valid also for thin gratings.

A.2.2 Phase matching

This phase matching argument is originally due to Bragg [107]. The incident wave with wave vector $\mathbf{k}_i$ scatters at each point $\mathbf{r}$ on a grating plane, as well as each point $\mathbf{r} + \hat{\mathbf{G}}\Lambda$ on the neighboring grating plane. The incident wave arrives at the second plane with an additional phase $\mathbf{k}_i \cdot \hat{\mathbf{G}}\, \Lambda$ and excites the scattered wave with this phase delay. The scattered wave accumulates an additional phase $-\mathbf{k}_d \cdot \hat{\mathbf{G}}\, \Lambda$ from the second plane to the first plane.

Therefore, to ensure perfectly constructive interference of the scattered fields of both neighboring grating planes,

$$2\pi m \stackrel{!}{=} \mathbf{k}_d \cdot \hat{\mathbf{G}}\,\Lambda - \mathbf{k}_i \cdot \hat{\mathbf{G}}\,\Lambda \Leftrightarrow |\mathbf{G}|^2 m \stackrel{!}{=} \mathbf{G} \cdot (\mathbf{k}_d - \mathbf{k}_i). \tag{A.10}$$

Inserting Eq. (2.12) in Eq. (A.10) one finds that Bragg's phase matching condition is fulfilled if the normal component given by Eq. (2.13) coincides with

$$k_{d,N} \stackrel{!}{=} k_{i,N} + mG_N. \tag{A.11}$$

A.2.3 Sinusoidal volume grating coupling

Because a holographic volume grating is sinusoidal, the m-th diffraction order couples only to its neighboring orders $m \pm 1$ [108]. The incident order, denoted by $m = 0$, can therefore couple directly only to the two first orders $m = \pm 1$ and their respective transmission and reflection branches. In a thick volume grating, efficient energy transfer additionally requires phase matching. For a given incidence wave vector, at most one of the directly coupled first orders can be close to phase matched. Higher orders cannot be phase matched simultaneously with this first order, which suppresses any indirect coupling to higher orders. In contrast, thin volume gratings give rise to multiple orders, see Moharam and Gaylord [108].

A.3 Implementation lens-stack weight analysis

The following details support the lens-stack calculation in Sec. 2.3.3. For each push or pull element, the local thickness is

$$t(x, y) = d_\mathrm{c} + s_2(x, y) - s_1(x, y), \tag{A.12}$$

where s_1 and s_2 are the front and rear surface sags, respectively, and d_c is the center thickness. For a spherical surface of signed radius R, $s_\mathrm{sph}(x, y; R) =$

$R - \operatorname{sgn}(R)\sqrt{R^2 - x^2 - y^2}$, and for a cylindrical surface curved in x only, $s_{\mathrm{cyl}}(x; R) = R - \operatorname{sgn}(R)\sqrt{R^2 - x^2}$. Planar surfaces are treated as $s = 0$.

A.3.1 Center thickness

To ensure a minimum local thickness $d_{\min}$ over the aperture $\mathcal{A}$, each lens element must satisfy

$$d_{\mathrm{c}} + \min_{\mathbf{p}\in\mathcal{A}} \left[s_2(\mathbf{p}; R_2) - s_1(\mathbf{p}; R_1) \right] \geq d_{\min}. \tag{A.13}$$

For the spherical and cylindrical surface combinations used here, the limiting point occurs either at the lens apex or on the aperture boundary. Since the sag difference is zero at the apex, the minimum required center thickness can be obtained by sampling the aperture outline $\mathcal{O}$ and enforcing

$$d_{\mathrm{c}} = \max\left[d_{\min} - \min_{\mathbf{p}\in\mathcal{O}} \left(s_2(\mathbf{p}; R_2) - s_1(\mathbf{p}; R_1) \right),\ d_{\min} \right]. \tag{A.14}$$

A.3.2 Pull-lens rear radius

The push-lens front radius is fixed by the base curve, and its rear radius is fixed by the waveguide geometry. The waveguide thickness and shape also determine the pull-lens front surface. The remaining pull-lens parameters are therefore the rear radius $R_{2,\mathrm{pull}}$ and center thickness $d_{\mathrm{c,pull}}$, which must together provide the target prescription power Φ_{Rx}.

Let

$$\mathbf{M} = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} \tag{A.15}$$

be the cumulative ABCD matrix through all stack elements up to and including refraction at the pull-lens front surface. The output medium of this partial matrix

is the pull-lens material with refractive index n. Propagation through the pull lens and refraction at its rear surface into air give

$$\mathbf{T} = \mathbf{R}_{2,\mathrm{pull}}\mathbf{F}(d_{\mathrm{c,pull}})\mathbf{M}, \tag{A.16}$$

with

$$\mathbf{F}(d) = \begin{pmatrix} 1 & d \\ 0 & 1 \end{pmatrix}, \qquad \mathbf{R}_{2,\mathrm{pull}} = \begin{pmatrix} 1 & 0 \\ (n-1)/R_{2,\mathrm{pull}} & n \end{pmatrix}. \tag{A.17}$$

The required total power is

$$-T_{21} \stackrel{!}{=} \Phi_{\mathrm{Rx}}, \tag{A.18}$$

Solving for the rear radius gives

$$R_{2,\mathrm{pull}} = \frac{(1-n)\,(M_{11} + M_{21}d_{\mathrm{c,pull}})}{nM_{21} + \Phi_{\mathrm{Rx}}}. \tag{A.19}$$

Equations (A.14) and (A.19) are coupled because the required center thickness depends on the rear radius. To avoid an iterative calculation, we first set $d_{\mathrm{c,pull}} \approx d_{\min}$ in Eq. (A.19) and then compute the final center thickness using Eq. (A.14). The resulting focus error is negligible for the present weight analysis. The astigmatic power of the buried cylindrical waveguide is also neglected because its front and rear surface powers nearly compensate.

A.3.3 Volume and weight

The volume of element k is

$$V_k = \int_{x_{\min}}^{x_{\max}} \int_{y_{\min}(x)}^{y_{\max}(x)} t_k(x,y)\,\mathrm{d}y\,\mathrm{d}x. \tag{A.20}$$

The element weight is $m_k = \rho_k V_k$ and the total stack weight is the sum of the push lens, waveguide core, and pull lens.

A.4 Derivation of waveguide jump maps

We derive the jump maps stated in Sec. 3.2.1. As in Sec. 3.2.1, $\mathbf{p}$ denotes the apex-frame position vector of a point on the outer surface and for curved waveguides $\mathbf{P}(\mathbf{p}) = \mathbf{p} - r\hat{\mathbf{z}}$ the corresponding centered-frame position vector. Direction vectors have the same components in both frames. We consider a ray launched from position $\mathbf{p}$ on the outer surface in direction $\hat{\mathbf{k}}$ and derive its position $\mathbf{p}^{(1)}$ and direction $\hat{\mathbf{k}}^{(1)}$ after one jump (propagation to the inner surface, reflection, propagation back to the outer surface, reflection). This yields the jump map

$$\tau : (\mathbf{p}, \hat{\mathbf{k}}) \mapsto (\mathbf{p}^{(1)}, \hat{\mathbf{k}}^{(1)}). \tag{A.21}$$

We write this as $\mathbf{p}^{(1)} = \tau_p(\mathbf{p}, \hat{\mathbf{k}})$ and $\hat{\mathbf{k}}^{(1)} = \tau_k(\mathbf{p}, \hat{\mathbf{k}})$.

A.4.1 Planar waveguide

The ray is launched on the outer plane at point $\mathbf{p}$ in direction

$$\hat{\mathbf{k}} = \hat{k}_x\hat{\mathbf{x}} + \hat{k}_y\hat{\mathbf{y}} + \hat{k}_z\hat{\mathbf{z}}, \qquad \hat{k}_z = \sqrt{1 - \hat{k}_x^2 - \hat{k}_y^2} > 0.$$

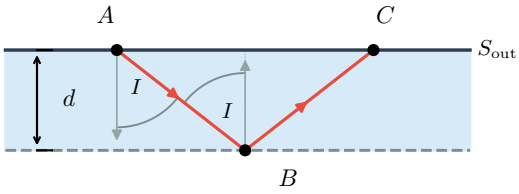


Figure A.1: Cross-section of a planar waveguide along the propagation plane of a single ray. The angle of incidence I of the ray is conserved throughout each jump.

For the geometric interpretation in Fig. A.1, we define the incidence angle I by $\cos I = \hat{k}_z$. The transverse propagation direction is

$$\hat{\mathbf{k}}_T = \frac{\hat{k}_x \hat{\mathbf{x}} + \hat{k}_y \hat{\mathbf{y}}}{\sqrt{\hat{k}_x^2 + \hat{k}_y^2}}.$$

Thus the displacement over one jump is

$$\Delta \mathbf{p} = 2d \tan I \, \hat{\mathbf{k}}_T = 2d \, \frac{\hat{k}_x \hat{\mathbf{x}} + \hat{k}_y \hat{\mathbf{y}}}{\hat{k}_z}.$$

Hence

$$\tau_p^{\mathrm{pl}}(\mathbf{p}, \hat{\mathbf{k}}) = \mathbf{p} + \Delta \mathbf{p}, \qquad \tau_k^{\mathrm{pl}}(\mathbf{p}, \hat{\mathbf{k}}) = \hat{\mathbf{k}}. \tag{A.22}$$

A.4.2 Cylindrical waveguide

In the centered frame we adopt cylindrical coordinates $X = R \sin\varphi$, $Z = -R\cos\varphi$ (with $\varphi = 0$ at the apex), where $R = r$ on the outer surface and $R = \rho r$ on the inner surface. The inward radial and tangential basis vectors are

$$\hat{\mathbf{r}}(\varphi) = -\sin\varphi \, \hat{\mathbf{x}} + \cos\varphi \, \hat{\mathbf{z}}, \qquad \hat{\boldsymbol{\varphi}}(\varphi) = \cos\varphi \, \hat{\mathbf{x}} + \sin\varphi \, \hat{\mathbf{z}}, \tag{A.23}$$

which coincide with the inward surface normal and tangent at the corresponding polar angle. They satisfy the rotation identities

$$\hat{\mathbf{r}}(\varphi + c) = \cos c \, \hat{\mathbf{r}}(\varphi) - \sin c \, \hat{\boldsymbol{\varphi}}(\varphi), \tag{A.24}$$

$$\hat{\boldsymbol{\varphi}}(\varphi + c) = \sin c \, \hat{\mathbf{r}}(\varphi) + \cos c \, \hat{\boldsymbol{\varphi}}(\varphi). \tag{A.25}$$

In the centered frame a point on the outer surface is written as $\mathbf{P} = -r \, \hat{\mathbf{r}}(\varphi) + y \, \hat{\mathbf{y}}$. The ray starting there may propagate in direction

$$\hat{\mathbf{k}} = \hat{\boldsymbol{\varphi}}(\varphi) u_\varphi + \hat{\mathbf{r}}(\varphi) u_r + \hat{\mathbf{y}} u_y \tag{A.26}$$

with some $u_\varphi, u_y \in \mathbb{R}$ such that $u_\varphi^2 + u_y^2 < 1$ and $u_r = \sqrt{1 - u_\varphi^2 - u_y^2}$.

As discussed by Vostrikov *et al.*, there are two special cases: propagation in the sagittal plane ($u_\varphi = 0$) and in the meridional plane ($u_y = 0$) [33].

Sagittal case $u_\varphi = 0$: If the ray propagates in the sagittal plane, its propagation behavior is the same as in the planar waveguide, $\hat{\mathbf{k}}^{(1)}(u_\varphi = 0, u_y) = u_y\hat{\mathbf{y}} + u_r\hat{\mathbf{r}}(\varphi) = \hat{\mathbf{k}}(u_\varphi = 0, u_y)$ and $\mathbf{P}^{(1)} = \mathbf{P} + 2du_y/u_r\hat{\mathbf{y}}$.

Meridional case $u_y = 0$: A ray launched from the outer surface at an angle $I = \arcsin(u_\varphi)$ in the meridional plane remains in that plane because of the symmetry along the cylinder axis. However, propagation in the meridional plane is non-trivial as the surface normal rotates during propagation. As shown by De Hoog *et al.* [109] and visualized in Fig. A.2, applying the sine rule in the triangle OAB gives the angle the ray encloses with the inner surface normal at its intersection point B as

$$\sin(I_i) = \frac{1}{\rho}\sin(I). \tag{A.27}$$

The angle it encloses with the local outer surface normal when returning to the outer surface at C is again I. The total polar angular advance during one jump is therefore $\Gamma = 2(I_i - I) = 2(\arcsin(\sin(I)/\rho) - I)$. Since both the position and direction are rotated by Γ around the cylinder axis, the jump map is a pure rotation for the meridional case. The formula assumes that the ray intersects the inner shell, i.e. $|\sin I| \leq \rho$, otherwise the corresponding inner-surface reflection does not occur.

General skew case ($u_y \neq 0$, $u_\varphi \neq 0$): The helical ray motion makes this case more involved. However, because the cylinder normal has no $\hat{\mathbf{y}}$ component, specular reflection at either surface does not change u_y. The axial motion is therefore a uniform drift superposed on the meridional rotation in the x–z plane, and the ray direction is rotated about the cylinder axis as before.

The *in-plane* (meridional) direction angle $I_\perp$ of the x–z projection of $\hat{\mathbf{k}}$ is given by

$$I_\perp = \operatorname{atan2}(u_\varphi, u_r), \qquad \sin I_\perp = \frac{u_\varphi}{u_\perp}, \qquad u_\perp = \sqrt{u_r^2 + u_\varphi^2}. \tag{A.28}$$

Following the meridional-case argument, the inner-surface in-plane angle I_i satisfies $\sin(I_i) = \sin(I_\perp)/\rho$. The total polar advance is

$$\Gamma = 2(I_i - I_\perp). \tag{A.29}$$

Because u_y is constant during reflections at the cylindrical surfaces, the ray direction is again rotated by the total polar advance after one jump:

$$\hat{\mathbf{k}}^{(1)} = R_{\hat{\mathbf{y}}}(\Gamma)\hat{\mathbf{k}}. \tag{A.30}$$

By symmetry, the total meridional propagation distance is twice the meridional distance from the outer shell to the inner shell s. Applying the sine law yields

$$\frac{s}{\sin(I_i - I_\perp)} = \frac{r}{\sin(\pi - I_i)} = \frac{r}{\sin(I_i)}. \tag{A.31}$$

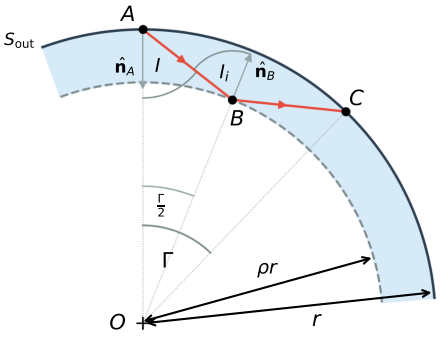


Figure A.2: Meridional cross-section of the jump geometry. A ray launched in the meridional plane from A on the outer surface S_{out} at incidence angle I propagates to B on the inner surface, where it is reflected at incidence angle I_i, and returns to C on the outer surface. The total angular advance seen from the center O is $\Gamma = 2(I_i - I)$.

With each propagation Δs in the meridional plane the ray drifts $\Delta y = \frac{u_y}{u_\perp}\Delta s$ along the cylinder axis.

$$\Delta y = 2s\frac{u_y}{u_\perp} = 2\,\rho r\,\frac{u_y}{u_\perp \sin(I_\perp)}\,\sin(I_i - I_\perp) = 2\,\rho r\,\frac{u_y}{u_\varphi}\,\sin\Big(\frac{\Gamma}{2}\Big). \tag{A.32}$$

Accounting for both polar advance and axial shift, the centered-frame update is

$$\mathbf{P}^{(1)} = R_{\hat{\mathbf{y}}}(\Gamma)\,\mathbf{P} + \Delta y\,\hat{\mathbf{y}} = -r\,\hat{\mathbf{r}}(\varphi+\Gamma) + (y+\Delta y)\hat{\mathbf{y}}. \tag{A.33}$$

Returning to the apex-frame position vector gives

$$\mathbf{p}^{(1)} = \tau_p^{\mathrm{cyl}}(\mathbf{p},\hat{\mathbf{k}}) = \mathbf{P}^{(1)} + r\hat{\mathbf{z}} = R_{\hat{\mathbf{y}}}(\Gamma)\big(\mathbf{p} - r\hat{\mathbf{z}}\big) + r\hat{\mathbf{z}} + \Delta y\,\hat{\mathbf{y}}, \tag{A.34}$$

$$\hat{\mathbf{k}}^{(1)} = \tau_k^{\mathrm{cyl}}(\mathbf{p},\hat{\mathbf{k}}) = R_{\hat{\mathbf{y}}}(\Gamma)\hat{\mathbf{k}}. \tag{A.35}$$

A.4.3 Spherical waveguide

Let $\mathbf{p}$ be the apex-frame position vector of the launch point on the outer sphere, and write $\mathbf{P} = \mathbf{P}(\mathbf{p}) = \mathbf{p} - r\hat{\mathbf{z}}$ for the corresponding centered-frame position vector. The inward unit normal is therefore

$$\hat{\mathbf{r}}_s = -\hat{\mathbf{P}} = -\frac{\mathbf{P}}{r}. \tag{A.36}$$

We decompose the launch direction into normal and tangential parts,

$$\hat{\mathbf{k}} = \hat{\mathbf{k}}_T + u_r\hat{\mathbf{r}}_s, \qquad u_T = |\hat{\mathbf{k}}_T| = \sqrt{1-u_r^2}.$$

The incidence angle in the propagation plane is $I = \arcsin u_T$. Specular reflection at a concentric spherical surface exerts only a radial impulse, no torque about the sphere center. Hence the ray's angular momentum

$$\mathbf{L} = \mathbf{P}\times\hat{\mathbf{k}}, \qquad \hat{\mathbf{L}} = \frac{\mathbf{L}}{|\mathbf{L}|}$$

is conserved during the outer-to-inner-to-outer jump. The ray therefore remains in the fixed plane orthogonal to $\hat{\mathbf{L}}$, spanned by $\mathbf{P}$ and $\hat{\mathbf{k}}$.

In this plane the geometry reduces to the same concentric-circle construction as in the cylindrical meridional case. From Eq. (A.27) we find the inner-surface incidence angle $I_i = \arcsin(u_T/\rho)$, and the total angular advance is given by Eq. (A.29) with $I_\perp$ replaced by I. Thus the net effect of one jump is a rigid rotation in the propagation plane, or equivalently a rotation by Γ about the conserved angular-momentum axis. In centered coordinates,

$$\mathbf{P}^{(1)} = R_{\hat{\mathbf{L}}}(\Gamma)\mathbf{P}, \qquad \hat{\mathbf{k}}^{(1)} = R_{\hat{\mathbf{L}}}(\Gamma)\hat{\mathbf{k}}.$$

Returning to the apex-frame position vector gives

$$\mathbf{p}^{(1)} = R_{\hat{\mathbf{L}}}(\Gamma)\mathbf{P}(\mathbf{p}) + r\hat{\mathbf{z}}. \tag{A.37}$$

A.5 Mode closure forces affine structure

Here we show that a mode family $\{\hat{\mathbf{k}}(\mathbf{p};\mathbf{u})\}_{\mathbf{u}\in\mathcal{U}\subset\mathbb{R}^2}$ is closed if and only if its tangential rayfield is affine in the mode coordinates $\mathbf{u} = (u_1, u_2)^T$.

Closure $\Rightarrow$ affinity: By definition for any three closed modes,

$$\hat{\mathbf{k}}_T(\mathbf{p};\mathbf{u}_3 + \mathbf{u}_2 - \mathbf{u}_1) = \hat{\mathbf{k}}_T(\mathbf{p};\mathbf{u}_3) + \hat{\mathbf{k}}_T(\mathbf{p};\mathbf{u}_2) - \hat{\mathbf{k}}_T(\mathbf{p};\mathbf{u}_1). \tag{A.38}$$

We fix an aperture point $\mathbf{p}$ and write $F_{\mathbf{p}}(\mathbf{u}) := \hat{\mathbf{k}}_T(\mathbf{p};\mathbf{u})$. Using Eq. (A.38) with $(\mathbf{u}_3, \mathbf{u}_2, \mathbf{u}_1) = (\mathbf{u}+\mathbf{z}, \mathbf{v}, \mathbf{u})$ gives

$$F_{\mathbf{p}}(\mathbf{v}+\mathbf{z}) - F_{\mathbf{p}}(\mathbf{v}) = F_{\mathbf{p}}(\mathbf{u}+\mathbf{z}) - F_{\mathbf{p}}(\mathbf{u}). \tag{A.39}$$

Thus, for a given displacement $\mathbf{z}$, the change in the tangential field is independent of the starting mode. We therefore define

$$\Delta_{\mathbf{p}}(\mathbf{z}) := F_{\mathbf{p}}(\mathbf{u}+\mathbf{z}) - F_{\mathbf{p}}(\mathbf{u}), \tag{A.40}$$

with any admissible $\mathbf{u}$. Successive mode displacements then satisfy

$$\Delta_{\mathbf{p}}(\mathbf{z}_1 + \mathbf{z}_2) = F_{\mathbf{p}}(\mathbf{u} + \mathbf{z}_1 + \mathbf{z}_2) - F_{\mathbf{p}}(\mathbf{u}) = \Delta_{\mathbf{p}}(\mathbf{z}_2) + \Delta_{\mathbf{p}}(\mathbf{z}_1), \tag{A.41}$$

where Eq. (A.39) was used in the last step. Hence the response to a combined mode displacement is the sum of the separate responses. Because the rayfield is continuous in $\mathbf{u}$, this additive displacement response is linear and there is a linear map $A(\mathbf{p}) : \mathbb{R}^2 \to T_{\mathbf{p}}S_{out}$ from mode coordinates to tangential displacement such that

$$\Delta_{\mathbf{p}}(\mathbf{z}) = A(\mathbf{p})\mathbf{z}. \tag{A.42}$$

Choosing any reference mode $\mathbf{u}_* \in \mathcal{U}$, we obtain

$$F_{\mathbf{p}}(\mathbf{u}) = F_{\mathbf{p}}(\mathbf{u}_*) + A(\mathbf{p})(\mathbf{u} - \mathbf{u}_*) = A(\mathbf{p})\mathbf{u} + \mathbf{b}(\mathbf{p}), \tag{A.43}$$

with $\mathbf{b}(\mathbf{p}) = F_{\mathbf{p}}(\mathbf{u}_*) - A(\mathbf{p})\mathbf{u}_*$. Thus, closure requires the tangential field to be affine in the mode coordinates:

$$\hat{\mathbf{k}}_T(\mathbf{p}; \mathbf{u}) = A(\mathbf{p})\mathbf{u} + \mathbf{b}(\mathbf{p}). \tag{A.44}$$

Affinity $\Rightarrow$ closure: Assume conversely that the tangential field is affine in the mode coordinate,

$$\hat{\mathbf{k}}_T(\mathbf{p}; \mathbf{u}) = A(\mathbf{p})\mathbf{u} + \mathbf{b}(\mathbf{p}).$$

For any three admissible mode coordinates set $\mathbf{u}_4 = \mathbf{u}_3 + \mathbf{u}_2 - \mathbf{u}_1$. Then

$$\begin{aligned}
\hat{\mathbf{k}}_T(\mathbf{p}; \mathbf{u}_4) &= A(\mathbf{p})(\mathbf{u}_3 + \mathbf{u}_2 - \mathbf{u}_1) + \mathbf{b}(\mathbf{p}) \\
&= \big[A(\mathbf{p})\mathbf{u}_3 + \mathbf{b}(\mathbf{p})\big] + \big[A(\mathbf{p})\mathbf{u}_2 + \mathbf{b}(\mathbf{p})\big] - \big[A(\mathbf{p})\mathbf{u}_1 + \mathbf{b}(\mathbf{p})\big] \\
&= \hat{\mathbf{k}}_T(\mathbf{p}; \mathbf{u}_3) + \hat{\mathbf{k}}_T(\mathbf{p}; \mathbf{u}_2) - \hat{\mathbf{k}}_T(\mathbf{p}; \mathbf{u}_1).
\end{aligned} \tag{A.45}$$

Hence, affine tangential dependence is sufficient for closure.

A.6 Properties of the mode families

We verify the local consistency, closure, and global rigidity properties of the three mode families introduced in Sec. 3.4.

A.6.1 Local consistency

Planar: For the planar family introduced in Sec. 3.4.1, the jump map leaves the direction unchanged and translates the launch point by an aperture-independent displacement, $\mathbf{p}^{(1)} = \mathbf{p} + \Delta\mathbf{p}_{\mathrm{pl}}(\mathbf{u})$. Since the planar rayfield Eq. (3.11) is spatially constant,

$$\hat{\mathbf{k}}^{\mathrm{pl}}(\mathbf{p}^{(1)};\mathbf{u}) = \hat{\mathbf{k}}^{\mathrm{pl}}(\mathbf{p};\mathbf{u}) = \tau_k^{\mathrm{pl}}(\mathbf{p},\hat{\mathbf{k}}^{\mathrm{pl}}),$$

which is the local consistency condition.

Cylindrical: For the cylindrical family introduced in Sec. 3.4.2, the local components (u_1, u_2, u_r) are independent of the aperture coordinates. Hence the cylindrical incidence angle, the polar advance Γ_0, and the axial drift Δy_0 are fixed for a given mode. One jump gives

$$\mathbf{P}^{(1)} = R_{\hat{\mathbf{y}}}(\Gamma_0)\mathbf{P} + \Delta y_0 \hat{\mathbf{y}}, \qquad \hat{\mathbf{k}}^{(1)} = R_{\hat{\mathbf{y}}}(\Gamma_0)\hat{\mathbf{k}}.$$

Since the cylindrical basis co-rotates with the aperture point, $\hat{\boldsymbol{\varphi}}(\varphi + \Gamma_0) = R_{\hat{\mathbf{y}}}(\Gamma_0)\hat{\boldsymbol{\varphi}}(\varphi)$ and $\hat{\mathbf{r}}(\varphi + \Gamma_0) = R_{\hat{\mathbf{y}}}(\Gamma_0)\hat{\mathbf{r}}(\varphi)$. Evaluating the same mode at $\mathbf{p}^{(1)} = \mathbf{P}^{(1)} + r\hat{\mathbf{z}}$, therefore gives

$$\hat{\mathbf{k}}^{\mathrm{cyl}}(\mathbf{p}^{(1)};\mathbf{u}) = R_{\hat{\mathbf{y}}}(\Gamma_0)\hat{\mathbf{k}}^{\mathrm{cyl}}(\mathbf{p};\mathbf{u}) = \tau_k^{\mathrm{cyl}}(\mathbf{p},\hat{\mathbf{k}}^{\mathrm{cyl}}).$$

Spherical: For the spherical family introduced in Sec. 3.4.3, fix one mode $\mathbf{u}$ and set $\mathbf{a} = u_1\hat{\mathbf{x}} + u_2\hat{\mathbf{y}}$. We define the leaf label and leaf vector by

$$\ell(\mathbf{p}) := \frac{(\mathbf{a} \times \hat{\mathbf{P}}) \cdot \hat{\mathbf{z}}}{\hat{\mathbf{P}} \cdot \hat{\mathbf{z}}}, \qquad \mathbf{b}_\ell := \mathbf{a} \times \hat{\mathbf{z}} + \ell\hat{\mathbf{z}}.$$

Equivalently, $\mathbf{b}_\ell = (\mathbf{a} \times \hat{\mathbf{P}})/(\hat{\mathbf{P}} \cdot \hat{\mathbf{z}})$. Hence $\hat{\mathbf{P}} \cdot \mathbf{b}_\ell = 0$, so each leaf is a great circle. Moreover, using Eq. (3.10),

$$\mathbf{b}_\ell \times \hat{\mathbf{P}} = A_{sph}(\mathbf{p})\mathbf{u}.$$

Since the inward surface normal is $-\hat{\mathbf{P}}$, the spherical field can therefore be written on each leaf as

$$\hat{\mathbf{k}}^{sph}(\mathbf{p}; \mathbf{u}) = \mathbf{b}_\ell \times \hat{\mathbf{P}} - \sqrt{1 - |\mathbf{b}_\ell|^2}\, \hat{\mathbf{P}}.$$

Its centered-frame angular momentum is $\mathbf{L}_\ell = \mathbf{P} \times \hat{\mathbf{k}}^{sph} = r\, \mathbf{b}_\ell$, so $\mathbf{b}_\ell$ is parallel to the conserved jump axis. Also $|\hat{\mathbf{k}}_T| = |\mathbf{b}_\ell|$, hence the jump angle $\Gamma_\ell := \Gamma(\arcsin |\mathbf{b}_\ell|)$ is constant on the leaf. The spherical jump rotates all rays on that leaf about $\hat{\mathbf{b}}_\ell$. Since this rotation maps $\hat{\mathbf{P}}$ to $\hat{\mathbf{P}}^{(1)}$ and leaves $\mathbf{b}_\ell$ invariant, evaluating the same field at the jumped point gives

$$\hat{\mathbf{k}}^{sph}(\mathbf{p}^{(1)}; \mathbf{u}) = R_{\hat{\mathbf{b}}_\ell}(\Gamma_\ell)\hat{\mathbf{k}}^{sph}(\mathbf{p}; \mathbf{u}) = \tau_k^{sph}(\mathbf{p}, \hat{\mathbf{k}}^{sph}).$$

A.6.2 Closure

For all three constructed families, the tangential field has the linear form

$$\hat{\mathbf{k}}_T^g(\mathbf{p}; \mathbf{u}) = A_g(\mathbf{p})\mathbf{u}, \qquad g \in \{\mathrm{pl}, \mathrm{cyl}, \mathrm{sph}\}.$$

Therefore, all are closed

$$\hat{\mathbf{k}}_T^g(\mathbf{p}; \mathbf{u}_3) + \hat{\mathbf{k}}_T^g(\mathbf{p}; \mathbf{u}_2) - \hat{\mathbf{k}}_T^g(\mathbf{p}; \mathbf{u}_1) = \hat{\mathbf{k}}_T^g(\mathbf{p}; \mathbf{u}_3 + \mathbf{u}_2 - \mathbf{u}_1).$$

A.6.3 Global rigidity

Planar: The planar family is globally rigid because one jump is the same translation for every aperture point,

$$\mathbf{p}' = \mathbf{p} + \Delta\mathbf{p}_{\mathrm{pl}}(\mathbf{u}), \qquad \hat{\mathbf{k}}' = \hat{\mathbf{k}}.$$

Thus the whole field is reproduced by a single rigid translation.

Cylindrical: The cylindrical family is also globally rigid. For fixed $\mathbf{u}$, the jump angle Γ_0 and drift Δy_0 are aperture-independent, so every ray undergoes the same screw motion,

$$\mathbf{P}' = R_{\hat{\mathbf{y}}}(\Gamma_0)\mathbf{P} + \Delta y_0 \hat{\mathbf{y}}, \qquad \hat{\mathbf{k}}' = R_{\hat{\mathbf{y}}}(\Gamma_0)\hat{\mathbf{k}}.$$

Hence one rigid motion reproduces the full cylindrical field after one jump.

Spherical: The spherical family is not globally rigid on a finite two-dimensional aperture. The local consistency argument shows that each leaf ℓ is rigidly reproduced by a rotation about the leaf angular-momentum axis. However, a finite two-dimensional aperture intersects a continuum of leaves. For $\mathbf{u} \neq \mathbf{0}$, with $\mathbf{a} = u_1\hat{\mathbf{x}} + u_2\hat{\mathbf{y}}$, the vectors $\mathbf{b}_\ell = \mathbf{a} \times \hat{\mathbf{z}} + \ell\hat{\mathbf{z}}$ change direction and magnitude with ℓ, so both the rotation axes and the jump angles $\Gamma(|\mathbf{b}_\ell|)$ are leaf-dependent. A single rigid motion cannot represent these different leaf-wise rotations, agreeing with the non-existence proof for spherical eigenbeams in Sec. 3.4.

A.7 Closed modes reduce to previous paraxial eigenbeam results

A.7.1 Paraxial 1:1 imaging at an off-axis mirror

Following Blanche *et al.* [70, 31], a suitably astigmatically focused beam is an eigenbeam under the off-axis reflection at the outer waveguide surface [70, 31].

Consider a beam whose chief ray is incident on the outer surface at an angle ϕ in the meridional plane. Assume that the sagittal and meridional ray fans of this beam emerge from points at distances s_s and s_m along the chief ray, respectively. The ray fans are then paraxially focused at distances S_s and S_m along the chief ray according to the Coddington equations [110]

$$\frac{1}{s_i} + \frac{1}{S_i} = \frac{1}{F_i}, \tag{A.46}$$

with the sagittal focal length $F_s = r/(2\cos(\phi))$ and meridional focal length $F_m = r\cos(\phi)/2$. A 1:1 eigenbeam satisfies $s_i = S_i = 2F_i$, hence it must be focused by focal lengths

$$f_s = \frac{r}{\cos\phi}, \qquad f_m = r\cos\phi. \tag{A.47}$$

along its chief ray. The corresponding normal distances from the mirror are $d_i = \cos(\phi) f_i$, therefore

$$d_s = r, \qquad d_m = r\cos^2\phi. \tag{A.48}$$

For $\phi = 54°$, this gives $d_m \approx 0.35r$ and $d_s = r$, largely consistent with the reported rule-of-thumb values $d_m \approx 0.3r$ and $d_s \approx 0.9r$ for the sphere [70]. The argument is readily adapted to the cylinder, yielding the same meridional focal length and no sagittal focusing [70, 31].

A.7.2 Paraxial approximation of the closed modes

We now show that, near the apex, the cylindrical and spherical closed modes reduce to the focal distances predicted by Eq. (A.47). The chief ray lies in the meridional x-z plane $\mathbf{Q}_0(t) = t(\sin\phi\,\hat{\mathbf{x}} + \cos\phi\,\hat{\mathbf{z}})$, corresponding to the $u_1 =$

$\sin\phi$, $u_2 = 0$ mode. Around the apex $\hat{\boldsymbol{\varphi}}(x) \approx \hat{\mathbf{x}} + (x/r)\hat{\mathbf{z}}$ and $\hat{\mathbf{r}}(x) \approx \hat{\mathbf{z}} - (x/r)\hat{\mathbf{x}}$, therefore the cylindrical closed modes Eq. (3.12) are approximately

$$\hat{\mathbf{k}}^{\mathrm{cyl}}(x) \approx \left(\sin\phi - \frac{x}{r}\cos\phi\right)\hat{\mathbf{x}} + \left(\cos\phi + \frac{x}{r}\sin\phi\right)\hat{\mathbf{z}}. \tag{A.49}$$

A marginal ray launched at x is $\mathbf{Q}_x(s) \approx x\hat{\mathbf{x}} + s\,\hat{\mathbf{k}}^{\mathrm{cyl}}(x)$. Solving $\mathbf{Q}_0(t) = \mathbf{Q}_x(s)$ gives $t \approx r\cos(\phi) + x\sin(\phi)$ around the apex. Thus the paraxial meridional focus distance is $f_m = r\cos\phi$, in agreement with Eq. (A.47).

For the spherical geometry we have for an outer surface point around the apex $\hat{\mathbf{P}}(x,y) \approx -\hat{\mathbf{z}} + (x/r)\hat{\mathbf{x}} + (y/r)\hat{\mathbf{y}}$, the closed spherical modes rayfield is therefore

$$\hat{\mathbf{k}}^{\mathrm{sph}}(x,y) \approx \left(\sin\phi - \frac{x}{r}\cos\phi\right)\hat{\mathbf{x}} - \frac{y}{r}\cos\phi\,\hat{\mathbf{y}} + \left(\cos\phi + \frac{x}{r}\sin\phi\right)\hat{\mathbf{z}}. \tag{A.50}$$

The meridional fan, obtained by setting $y = 0$, is identical to the cylindrical calculation and therefore focuses at $f_m = r\cos\phi$, with $x_m = r\sin\phi\cos\phi$ and $z_m = r\cos^2\phi$.

The sagittal fan is obtained by setting $x = 0$. A marginal ray launched at y is $\mathbf{Q}_y(s) \approx y\hat{\mathbf{y}} + s\left(\sin\phi\,\hat{\mathbf{x}} - \frac{y}{r}\cos\phi\,\hat{\mathbf{y}} + \cos\phi\,\hat{\mathbf{z}}\right)$. Solving $\mathbf{Q}_0(t) = \mathbf{Q}_y(s)$ gives $t = r/\cos\phi$. Therefore, the sagittal focal distance is $f_s = r/\cos\phi$, in agreement with Eq. (A.47).

A.8 Refraction out of waveguide

In this section, we consider refraction out of spherically and cylindrically curved waveguide substrates. A unit incident ray $\hat{\mathbf{i}}$ at an interface with unit normal $\hat{\mathbf{N}}$ (sign chosen such that $q = \hat{\mathbf{i}} \cdot \hat{\mathbf{N}} > 0$) is refracted into direction

$$\mathcal{S}_{n_1 \to n_2}(\hat{\mathbf{i}}, \hat{\mathbf{N}}) = \mu\left(\hat{\mathbf{i}} - q\hat{\mathbf{N}}\right) + \sqrt{1 - \mu^2(1-q^2)}\,\hat{\mathbf{N}}, \qquad \mu = \frac{n_1}{n_2}. \tag{A.51}$$

A.8.1 Mode coordinate to chief ray direction cosine

First, we derive the mapping between the mode coordinates $\mathbf{u}$ and the air-side direction cosines $\mathbf{c}$ of the mode's chief ray after launch at the surface apex and refraction out of the waveguide. The closed mode chief ray propagates within the waveguide substrate toward the inner surface with direction

$$\hat{\mathbf{k}}_{\mathrm{c}}(\mathbf{u}) = u_1\hat{\mathbf{x}} + u_2\hat{\mathbf{y}} + \sqrt{1-u_1^2-u_2^2}\hat{\mathbf{z}}. \tag{A.52}$$

The ray direction after refraction is

$$\hat{\mathbf{k}}_{\mathrm{c}}^{\mathrm{air}}(\mathbf{c}(\mathbf{u})) = c_x\,\hat{\mathbf{x}} + c_y\,\hat{\mathbf{y}} + \sqrt{1-c_x^2-c_y^2}\,\hat{\mathbf{z}}. \tag{A.53}$$

For a planar waveguide, the direction cosines after refraction are

$$\mathbf{c} = n\mathbf{u}.$$

For curved waveguides, the refracted direction depends on the intersection position of the ray with the inner surface. The inner surface satisfies

$$x^2 + \gamma_g y^2 + (z-r)^2 = (r-d)^2.$$

Here, $\gamma_g = 1$ for the sphere and $\gamma_g = 0$ for the cylinder. Substituting the chief ray $t\hat{\mathbf{k}}_{\mathrm{c}}$ gives the quadratic equation

$$A_i t^2 - 2ru_z t + 2rd - d^2 = 0,$$

with $u_z = \sqrt{1-u_1^2-u_2^2}$ and $A_i = u_1^2 + \gamma_g u_2^2 + u_z^2$. Thus $A_i = 1$ for the sphere and $A_i = 1 - u_2^2$ for the cylinder. Solving for the physical positive intersection distance finds

$$t_i = \frac{ru_z - \Delta_i}{A_i} \quad \text{with} \quad \Delta_i = \sqrt{r^2u_z^2 - A_i(2rd-d^2)}. \tag{A.54}$$

At the intersection point $t_i\hat{\mathbf{k}}_\mathrm{c}$, the inner-surface normal pointing from core to air is

$$\hat{\mathbf{N}}_i^g = \frac{-t_i u_1 \hat{\mathbf{x}} - \gamma_g t_i u_2 \hat{\mathbf{y}} + (r - t_i u_z)\hat{\mathbf{z}}}{r-d}. \tag{A.55}$$

Applying Snell's law, Eq. (A.51), yields the chief ray direction in air $\hat{\mathbf{k}}_\mathrm{c}^\mathrm{air}$, see Eq. (A.53). For the sphere we derive the direction cosines

$$\mathbf{c} = \eta_{sph}(\mathbf{u}; d, n, r)\mathbf{u} \tag{A.56}$$

with

$$\eta_{sph}(\mathbf{u}; d, n, r) = n + \frac{\left(nq_i - \sqrt{1 - n^2\left(1 - q_i^2\right)}\right) t_i}{r-d}, \tag{A.57}$$

where $q_i = \Delta_i/(r-d)$. For the cylinder we have the same expression for c_1 but for $c_2 = nu_2$.

A.8.2 Refraction-compensated output field

We now calculate the rayfield $\hat{\mathbf{k}}^{\mathcal{R},g}(\mathbf{p}; \mathbf{c})$ on the outer waveguide surface that, after refraction out of the waveguide, produces a perfectly collimated wave in the direction $\hat{\mathbf{k}}_\mathrm{c}^\mathrm{air}$, as defined in Eq. (A.53). To determine this field, we trace each collimated beam backward into the waveguide, refract it at the inner surface, and propagate it to the outer surface.

A.8.2.1 Planar waveguide

For the planar waveguide, only refraction at the inner surface must be considered. Propagation does not affect the result because of the translational symmetry of the collimated fields:

$$\hat{\mathbf{k}}_T^{\mathcal{R},\mathrm{pl}} = \frac{c_1}{n}\hat{\mathbf{x}} + \frac{c_2}{n}\hat{\mathbf{y}}. \tag{A.58}$$

A.8.2.2 Curved waveguides

For curved waveguides, each inner-surface point is given in the surface coordinate system by

$$\mathbf{p}_i = x_i\hat{\mathbf{x}} + y_i\hat{\mathbf{y}} + z_{\mathrm{in}}(x_i, y_i)\hat{\mathbf{z}}.$$

The ray in the waveguide is the reciprocal rayfield obtained by applying Snell's law, Eq. (A.51), to the incident collimated ray direction:

$$\hat{\mathbf{k}}_g^-(x_i, y_i; \mathbf{c}) = \mathcal{S}_{1\to n}\left(-\hat{\mathbf{k}}_{\mathrm{c}}^{\mathrm{air}}, -\hat{\mathbf{N}}_i^g(x_i, y_i)\right), \tag{A.59}$$

The ray reaches the outer surface after some distance at

$$\mathbf{p}_o = \mathbf{p}_i(x_i, y_i) + s_o\hat{\mathbf{k}}_g^-(x_i, y_i; \mathbf{c}).$$

The outer point (x_o, y_o) corresponding to the prescribed inner point (x_i, y_i) follows from this intersection equation:

$$(x_i, y_i) \mapsto (x_o, y_o).$$

This map must be inverted to obtain the refraction-compensated forward rayfield on the outer surface. It is the negative of the reciprocal ray field evaluated at the corresponding inner surface point

$$\hat{\mathbf{k}}^{\mathcal{R},g}(x_o, y_o; \mathbf{c}) = -\hat{\mathbf{k}}_{\mathrm{g}}^-(x_i(x_o, y_o), y_i(x_o, y_o); \mathbf{c}). \tag{A.60}$$

In the following paragraphs, we derive this map explicitly for the spherical and cylindrical geometries. We use both surface coordinates and centered coordinates as defined in Sec. 3.2.1, with small and capital letters respectively. Direction vectors are the same across both systems.

Spherical coordinate map: We first derive the inner-to-outer map for the spherical geometry and then invert it. A given ray propagates in the meridional plane spanned by the spherical origin, the surface point, and the air-side direction $\hat{\mathbf{k}}_{\mathrm{c}}^{\mathrm{air}}$.

We choose a primed centered coordinate system with $\hat{\mathbf{Z}}' = \hat{\mathbf{k}}_{\mathrm{c}}^{\mathrm{air}}$, and $\hat{\mathbf{X}}'$ orthogonal to it in this meridional plane. We define the normalized transverse coordinates

$$\xi_i = \frac{X_i'}{r-d} \quad \text{and} \quad \xi_o = \frac{X_o'}{r}.$$

The intersection points of the ray with the inner and outer spherical surfaces are parametrized by these tangential coordinates

$$\mathbf{P}_i'(\xi_i) = r\left[\rho\xi_i\hat{\mathbf{X}}' - \rho\sqrt{1-\xi_i^2}\,\hat{\mathbf{Z}}'\right] \quad \text{and} \quad \mathbf{P}_o'(\xi_o) = r\left[\xi_o\hat{\mathbf{X}}' - \sqrt{1-\xi_o^2}\,\hat{\mathbf{Z}}'\right]. \tag{A.61}$$

Because the reciprocal plane wave is incident along $-\hat{\mathbf{Z}}'$ in the meridional plane, the problem reduces to 2D concentric circles. Applying Snell's law (A.51) at the inner surface gives

$$\begin{aligned}\hat{\mathbf{k}}^{-\prime}(\xi_i) &= \mathcal{S}_{1\to n}\left(-\hat{\mathbf{Z}}', \hat{\mathbf{P}}_i'(\xi_i)\right) \\ &= \frac{\xi_i\left[\sqrt{n^2-\xi_i^2} - \sqrt{1-\xi_i^2}\right]}{n}\hat{\mathbf{X}}' - \frac{\xi_i^2 + \sqrt{n^2-\xi_i^2}\sqrt{1-\xi_i^2}}{n}\hat{\mathbf{Z}}'.\end{aligned} \tag{A.62}$$

The outer point is reached when

$$\mathbf{P}_o'(\xi_o) = \mathbf{P}_i'(\xi_i) + r\ell\,\hat{\mathbf{k}}^{-\prime}(\xi_i; N).$$

Eliminating ℓ gives the scalar inner-to-outer map, with $N = n$ for the sphere:

$$\begin{aligned}\xi_o = \mathcal{A}_N(\xi_i) = &\frac{\rho\xi_i\left[\xi_i^2 + \sqrt{N^2-\xi_i^2}\sqrt{1-\xi_i^2}\right]}{N^2} \\ &+ \frac{\xi_i^2\left[\sqrt{N^2-\xi_i^2} - \sqrt{1-\xi_i^2}\right]}{N^2}\sqrt{\frac{N^2}{\xi_i^2} - \rho^2}.\end{aligned}$$

For the numerical results of Sec. 3.6 we invert the mapping numerically. For a waveguide thin compared to its radius the inverse is approximated by

$$\mathcal{A}_N^{-1}(\xi_o) = \xi_o + (1-\rho)\,\xi_o \frac{\sqrt{1-\xi_o^2}}{\sqrt{N^2-\xi_o^2}} + O\big((1-\rho)^2\big)\,. \tag{A.63}$$

Now consider a point with surface coordinates (x_o, y_o) on the outer surface

$$\mathbf{P}_o = x_o\hat{\mathbf{x}} + y_o\hat{\mathbf{y}} + \Big[z_{out}^{sph}(x_o, y_o) - r\Big]\,\hat{\mathbf{z}}.$$

From this position and the desired ray direction in air we can calculate

$$\hat{\mathbf{X}}' = \frac{\mathbf{P}_o - \left(\mathbf{P}_o \cdot \hat{\mathbf{k}}_{\mathrm{c}}^{\mathrm{air}}\right)\hat{\mathbf{k}}_{\mathrm{c}}^{\mathrm{air}}}{\left|\mathbf{P}_o - \left(\mathbf{P}_o \cdot \hat{\mathbf{k}}_{\mathrm{c}}^{\mathrm{air}}\right)\hat{\mathbf{k}}_{\mathrm{c}}^{\mathrm{air}}\right|}, \qquad \xi_o = \frac{\left|\mathbf{P}_o - \left(\mathbf{P}_o \cdot \hat{\mathbf{k}}_{\mathrm{c}}^{\mathrm{air}}\right)\hat{\mathbf{k}}_{\mathrm{c}}^{\mathrm{air}}\right|}{r}. \tag{A.64}$$

From the inverse map we find the corresponding inner transverse coordinate

$$\xi_i(x_o, y_o) = \mathcal{A}_n^{-1}(\xi_o(x_o, y_o)). \tag{A.65}$$

Therefore, the forward refraction-compensated spherical field can be expressed solely in terms of the surface coordinates by inserting (A.65) and (A.64) into the reciprocal ray direction (A.62) and then reversing it according to (A.60), yielding

$$\begin{aligned}\hat{\mathbf{k}}^{\mathcal{R},sph}(x_o, y_o; \mathbf{c}) &= -\hat{\mathbf{k}}_{sph}^{-}(\xi_i) \qquad\qquad (A.66)\\ &= \frac{\xi_i\left(\sqrt{1-\xi_i^2} - \sqrt{n^2-\xi_i^2}\right)}{n}\hat{\mathbf{X}}' + \frac{\xi_i^2 + \sqrt{1-\xi_i^2}\sqrt{n^2-\xi_i^2}}{n}\hat{\mathbf{k}}_{\mathrm{c}}^{\mathrm{air}}.\end{aligned}$$

Substituting the expressions for $\hat{\mathbf{X}}'$ from Eq. (A.64) and for $\hat{\mathbf{k}}_{\mathrm{c}}^{\mathrm{air}}$ from Eq. (A.53) expresses the ray direction fully in the surface coordinate system defined in Sec. 3.2.1.

Cylindrical coordinate map: For the cylinder, the inner-surface normal has no axial component. We therefore split the desired air-side ray, Eq. (A.53), into its axial and meridional parts

$$\hat{\mathbf{k}}_{\mathrm{c}}^{\mathrm{air}} = c_2\hat{\mathbf{y}} + c_\perp\hat{\mathbf{Z}}'. \tag{A.67}$$

With the meridional magnitude $c_\perp = \sqrt{c_1^2 + c_z^2}$, the unit vector $\hat{\mathbf{Z}}'$ pointing in the direction of the meridional ray projection is

$$\hat{\mathbf{Z}}' = \frac{c_1\hat{\mathbf{x}} + c_z\hat{\mathbf{z}}}{c_\perp}. \tag{A.68}$$

Orthogonal to this unit vector, we define

$$\hat{\mathbf{X}}' = \frac{c_z\hat{\mathbf{x}} - c_1\hat{\mathbf{z}}}{c_\perp}. \tag{A.69}$$

The reciprocal ray refracted into the substrate at an inner cylindrical surface point with inward normal $\hat{\mathbf{N}}_i^{cyl}$ is

$$\hat{\mathbf{k}}_{cyl}^{-} = \mathcal{S}_{1\to n}\left(-\hat{\mathbf{k}}_{\mathrm{c}}^{\mathrm{air}}, -\hat{\mathbf{N}}_i^{cyl}\right).$$

Since $\hat{\mathbf{N}}_i^{cyl}\cdot\hat{\mathbf{y}} = 0$, the axial component follows directly from the tangential part of (A.51),

$$\hat{\mathbf{k}}_{cyl}^{-}\cdot\hat{\mathbf{y}} = -\frac{c_2}{n}.$$

Hence the refracted ray can be written as

$$\hat{\mathbf{k}}_{cyl}^{-} = \kappa\hat{\mathbf{k}}_{\perp}^{-} - \frac{c_2}{n}\hat{\mathbf{y}}, \qquad \kappa = \sqrt{1 - \left(\frac{c_2}{n}\right)^2},$$

where $\hat{\mathbf{k}}_{\perp}^{-}$ is a unit vector in the meridional plane. To determine this direction, let $q' = (-\hat{\mathbf{Z}}')\cdot(-\hat{\mathbf{N}}_i^{cyl})$. Since the meridional component of the incident ray

is $-c_\perp \hat{\mathbf{Z}}'$, the scalar product in (A.51) is $q = c_\perp q'$. Projecting (A.51) onto the meridional plane and dividing by κ gives

$$\hat{\mathbf{k}}_\perp^- = \frac{1}{N_{cyl}} \left[-\hat{\mathbf{Z}}' - q' \left(-\hat{\mathbf{N}}_i^{cyl} \right) \right] + \sqrt{1 - \frac{1}{N_{cyl}^2}(1 - q'^2)} \left(-\hat{\mathbf{N}}_i^{cyl} \right),$$

with effective index $N_{cyl} = n\kappa / c_\perp$. The square-root term follows from the relation $c_\perp^2 + c_2^2 = 1$. Thus

$$\hat{\mathbf{k}}_\perp^- = \mathcal{S}_{1 \to N_{cyl}} \left(-\hat{\mathbf{Z}}', -\hat{\mathbf{N}}_i^{cyl} \right).$$

The remaining construction is therefore identical to the spherical meridional construction. In both cases the reduced problem consists of a unit incident ray $-\hat{\mathbf{Z}}'$, a radial inner-surface normal, and two concentric circular boundaries of radii ρr and r. The only change is the replacement $n \mapsto N_{cyl}$. Consequently, the same normalized transverse coordinates (ξ_i, ξ_o) and the same circle parametrization as in (A.61) apply, and the cylindrical inner-to-outer map is

$$\xi_o = \mathcal{A}_{N_{cyl}}(\xi_i), \qquad \xi_i = \mathcal{A}_{N_{cyl}}^{-1}(\xi_o).$$

For an outer coordinate x_o, the centered meridional point is $\mathbf{P}_{o,\perp} = x_o \hat{\mathbf{x}} + [z_{out}^{cyl}(x_o) - r]\hat{\mathbf{z}}$, and therefore

$$\xi_o(x_o) = \frac{\mathbf{P}_{o,\perp} \cdot \hat{\mathbf{X}}'}{r} = \frac{c_z x_o + c_1 \sqrt{r^2 - x_o^2}}{r c_\perp}, \qquad \xi_i(x_o) = \mathcal{A}_{N_{cyl}}^{-1}(\xi_o(x_o)).$$

Inserting this coordinate into the reciprocal meridional ray expression of (A.62), with n replaced by N_{cyl}, and restoring the axial component gives

$$\hat{\mathbf{k}}^{\mathcal{R},cyl}(x_o;\mathbf{c}) = \frac{\kappa\xi_i\left[\sqrt{1-\xi_i^2}-\sqrt{N_{cyl}^2-\xi_i^2}\right]}{N_{cyl}}\hat{\mathbf{X}}' + \frac{\kappa\left[\xi_i^2+\sqrt{1-\xi_i^2}\sqrt{N_{cyl}^2-\xi_i^2}\right]}{N_{cyl}}\hat{\mathbf{Z}}' + \frac{c_2}{n}\hat{\mathbf{y}}. \tag{A.70}$$

Because $\hat{\mathbf{X}}'$ and $\hat{\mathbf{Z}}'$ are already given in the unprimed basis, (A.70) is expressed in the surface coordinate system. The result is independent of y_o, as required by cylindrical symmetry.

A.9 Visualizing immersion-free recording

In this appendix, we compare edge-lit with immersion-free recording configurations for an exemplary on-axis waveguide incoupler operating at $532\,\mathrm{nm}$ with a $56°$ propagation angle.

A.9.1 Traditional edge-lit in-plane recording

Conventionally, in- and outcouplers are recorded with both recording waves in the $\hat{\mathbf{T}}-\hat{\mathbf{N}}$ plane [38], corresponding to the $\sigma = 0$ or $\sigma = \pi$ recording configurations, see Eq. (A.6). For these configurations, the incidence angle of one of the two recording waves becomes minimal and that of the other becomes maximal, requiring edge-lit exposure. The $\sigma = 0$ configuration is visualized in Fig. A.3.

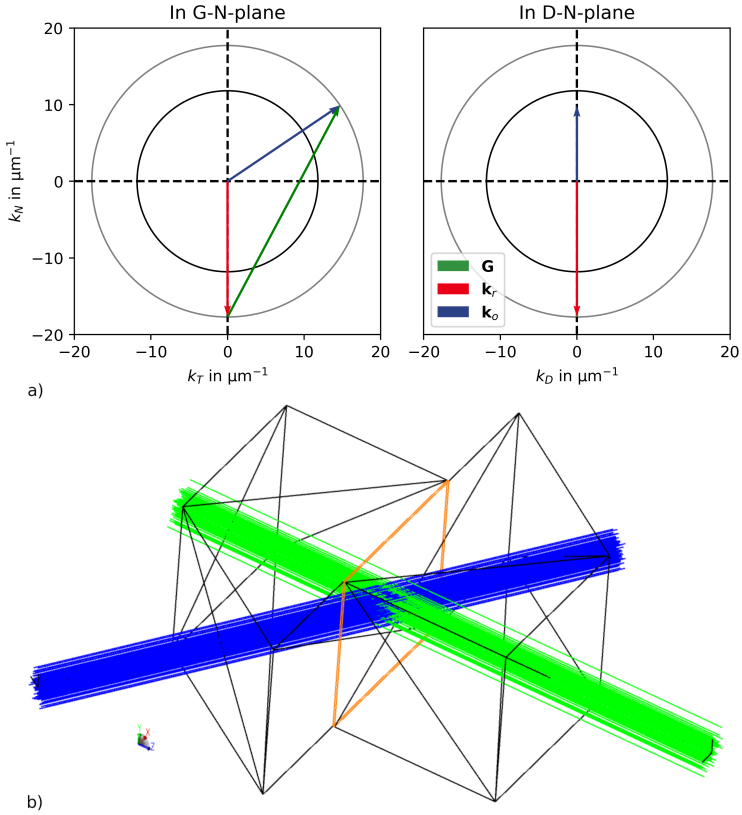


Figure A.3: (a) k-space diagram of the $\sigma = 0$ recording configuration of an on-axis waveguide incoupler operating at $532\,\mathrm{nm}$ with a propagation angle of $56°$. Recording is performed in the same configuration and at the same wavelength as replay in the waveguide display. Both recording wave vectors and the hologram surface normal $\hat{\mathbf{N}}$ lie in one plane, corresponding to in-plane recording. The tangential component of the object wave vector is larger than the radius of the air dispersion sphere, so the configuration cannot be realized without edge-lit exposure. (b) Zemax non-sequential ray-trace model of a practical implementation of this recording configuration. As proposed in [38], two prisms are optically contacted to the substrate using immersion liquid to provide the large object wave incidence angle and reduce multiple reflections.

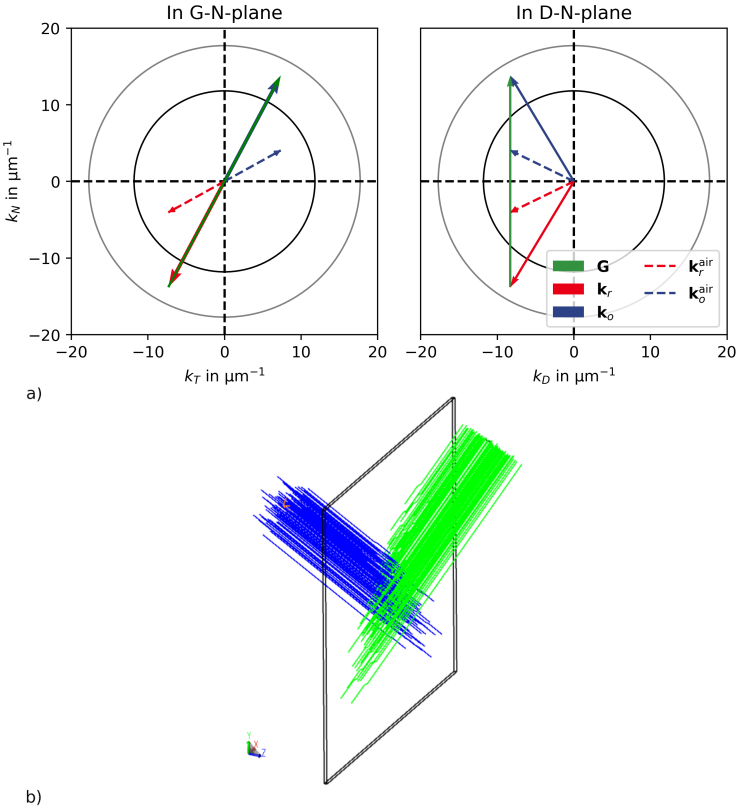


Figure A.4: (a) k-space diagram of the alternative $\sigma = \pi/2$ recording configuration for the incoupler discussed in Fig. A.3. This configuration exhibits the smallest limiting incidence angle and enables immersion-free recording. The tangential components of the recording wave vectors do not exceed the radius of the air dispersion sphere, allowing recording beams incident from air to enter the holographic substrate. The surface normal and the two recording wave vectors no longer lie in one plane. The wave vectors have components orthogonal to the normal vector. In a typical laboratory arrangement, both recording beams therefore point down toward the optical table. Panel (b) shows the corresponding Zemax ray trace.

A.9.2 Immersion-free out-of-plane recording

By shifting the recording configuration of the reflective incoupler from the $\sigma = 0$ recording configuration to the minimum-incidence reflective configuration $\sigma = \pi/2$ or $3\pi/2$, the incidence angles can be reduced sufficiently for light from air to enter the holographic substrate without edge-lit exposure. In this configuration, the two recording wave vectors no longer lie in one plane with the surface normal. They have components orthogonal to the normal vector, corresponding to out-of-plane recording, as shown in Fig. A.4(a). The components of the object and reference recording wave vectors in the $\hat{\mathbf{T}}$–$\hat{\mathbf{N}}$ plane are exactly $\pm\mathbf{G}^{\mathrm{irr}}/2$. They also have equal components in the direction $\hat{\mathbf{D}} = \hat{\mathbf{N}} \times \hat{\mathbf{T}}$, so both vectors satisfy the dispersion relation. This symmetry gives both waves the same incidence angle and yields the smallest limiting incidence angle for reflective recording.

A.10 Recording limit analysis for high-refractive-index holograms

Chapter 4 considers holographic materials with refractive index $n = 1.51$ and normal shrinkage $s = 0.03$, typical for photopolymers [56]. Figure A.5 reproduces Fig. 4.2 for a high refractive index holographic material with $n = 2$ and $s = 0.03$. Waveguide incoupler and expander gratings for these high-refractive-index waveguides exhibit significantly longer grating vectors, corresponding to shorter grating pitches. Additionally, recording wave vectors incident from air experience stronger refraction as they penetrate the holographic material. No relevant expander gratings can be recorded without immersion in such high-refractive-index materials. They lie outside the recordable area in Fig. A.5. Even incoupler recording becomes increasingly challenging: red-channel incouplers that remain recordable require red or near-infrared exposure, while the $60°$ examples have no propagating air-side solution.

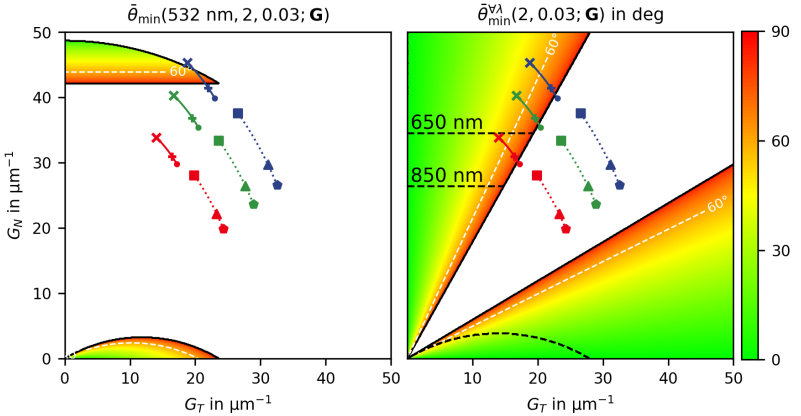


Figure A.5: This figure is a reproduction of Fig. 4.2 from Chapter 4 for $n = 2$. The grating parameters for incouplers and expanders depend on n according to Eq. (2.24).

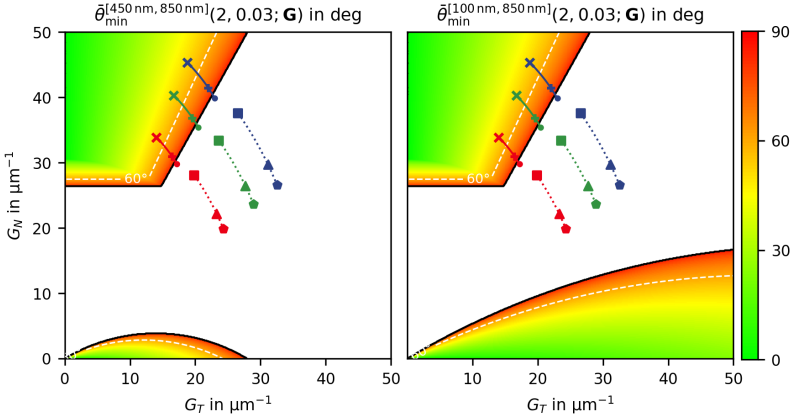


Figure A.6: Smallest limiting incidence angle for $n = 2$ and $s = 0.03$ over the finite recording-wavelength ranges 450 nm to 850 nm and 100 nm to 850 nm. Upper region corresponds to reflective and lower region to transmissive recording.

A.11 Angular sensitivity of the robot optics

The angular sensitivity of the fiber-collimator assembly is obtained with the paraxial ray-transfer matrix

$$M = L(f_2)T(d_2)L(f_1)T(d_1) = \begin{pmatrix} A & B \\ C & D \end{pmatrix}, \tag{A.71}$$

where

$$T(d) = \begin{pmatrix} 1 & d \\ 0 & 1 \end{pmatrix}, \qquad L(f) = \begin{pmatrix} 1 & 0 \\ -1/f & 1 \end{pmatrix}. \tag{A.72}$$

For an input ray $(h, \theta_0)^T$ the output angle is

$$\theta_{\text{out}} = Ch + D\theta_0. \tag{A.73}$$

The lens separation d_2 is chosen such that an on-axis point source is collimated, i.e. $D = 0$. Solving $D = 0$ for d_2 yields

$$d_2 = \frac{d_1(f_1 + f_2) - f_1 f_2}{d_1 - f_1}. \tag{A.74}$$

For the collimation optics discussed in Sec. 5.3.1 with $d_1 = 30\,\text{mm}$, $f_1 = -15\,\text{mm}$, and $f_2 = 200\,\text{mm}$, this gives $d_2 = 190\,\text{mm}$. The angular coefficient is

$$C = \frac{d_2 - f_1 - f_2}{f_1 f_2}. \tag{A.75}$$

Thus the collimated output angle δ generated by a transverse source displacement h is

$$\delta = Ch = -\frac{h}{600\,\text{mm}}. \tag{A.76}$$

Because the back-reflected beam angle is 2α, the refocused displacement is $h = -1200\,\text{mm}\alpha$.

A.12 Relation between OPW and Stäubli VAL3 robot arm configurations

Brandstötter *et al.* enumerate the OPW inverse kinematics as four regional solutions of the first three axes and two spherical-wrist solutions. In their Table II, solutions 1, 2, 5, and 6 use the first shoulder branch $\theta_{1;i}$, whereas solutions 3, 4, 7, and 8 use the second shoulder branch $\theta_{1;ii}$ [93]. We show here how these branches map to the Stäubli VAL3 configuration flags. The VAL3 shoulder criterion for a TX2 arm is "lefty" for $Q_s \geq 0$ and "righty" otherwise, with [111]

$$\begin{aligned} Q_s &= d_1 \sin\theta_2 + d_2 \sin(\theta_2 + \theta_3) + \delta \\ &\overset{\text{TX2}-60}{=} 290\,\text{mm} \sin\theta_2 + 310\,\text{mm} \sin(\theta_2 + \theta_3) \\ &= c_2 \sin\theta_2 + c_3 \sin(\theta_2 + \theta_3). \end{aligned} \tag{A.77}$$

From the TX2-60 data sheet [94] and the Stäubli Robotics Studio software [96] we infer $\delta = 0$, $d_1 = 290\,\text{mm}$. For the TX2-60 OPW parameters $(a_1, a_2, b, c_1, c_2, c_3, c_4) = (0, 0, 20, 0, 290, 310, 70)\,\text{mm}$ the VAL3 shoulder quantity Q_s is identical to the OPW coordinate c_{x_1}. Following Brandstötter *et al.* [93], we denote

$$\psi_1 = \text{atan2}(b, n_{x_1} + a_1)\,. \tag{A.78}$$

For the present case $a_1 = 0$, and the OPW definition $n_{x_1} = \sqrt{c_{x_0}^2 + c_{y_0}^2 - b^2} - a_1$ gives

$$\psi_1 = \text{atan2}\Big(b, \sqrt{c_{x_0}^2 + c_{y_0}^2 - b^2}\Big)\,. \tag{A.79}$$

With $\phi = \text{atan2}(c_{y_0}, c_{x_0})$ and $\rho = \sqrt{c_{x_0}^2 + c_{y_0}^2}$, the two OPW shoulder branches are

$$\theta_{1;i} = \phi - \psi_1, \qquad \theta_{1;ii} = \phi + \psi_1 - \pi. \tag{A.80}$$

To determine the sign of the corresponding OPW c_{x_1} coordinate, we use the inverse base rotation

$$c_{x_1}(\theta_1) = c_{x_0}\cos\theta_1 + c_{y_0}\sin\theta_1. \tag{A.81}$$

Substitution of the first shoulder branch gives

$$\begin{aligned} Q_s(\theta_{1;i}) = c_{x_1}(\theta_{1;i}) &= c_{x_0}\cos(\phi-\psi_1) + c_{y_0}\sin(\phi-\psi_1) \\ &= \rho\cos\phi\cos(\phi-\psi_1) + \rho\sin\phi\sin(\phi-\psi_1) \\ &= \rho\left[\cos\phi\cos(\phi-\psi_1) + \sin\phi\sin(\phi-\psi_1)\right] \\ &= \rho\cos(\phi-(\phi-\psi_1)) = \rho\cos\psi_1 \\ &= \rho\frac{\sqrt{\rho^2-b^2}}{\rho} = \sqrt{\rho^2-b^2} \geq 0. \end{aligned} \tag{A.82}$$

Similarly, substitution of the second shoulder branch gives

$$\begin{aligned} Q_s(\theta_{1;ii}) = c_{x_1}(\theta_{1;ii}) &= c_{x_0}\cos(\phi+\psi_1-\pi) + c_{y_0}\sin(\phi+\psi_1-\pi) \\ &= \rho\cos\phi\cos(\phi+\psi_1-\pi) + \rho\sin\phi\sin(\phi+\psi_1-\pi) \\ &= \rho\left[\cos\phi\cos(\phi+\psi_1-\pi) + \sin\phi\sin(\phi+\psi_1-\pi)\right] \\ &= \rho\cos(\phi-(\phi+\psi_1-\pi)) = \rho\cos(\pi-\psi_1) \\ &= -\rho\cos\psi_1 = -\sqrt{\rho^2-b^2} \leq 0. \end{aligned} \tag{A.83}$$

Therefore OPW branch $\theta_{1;i}$ maps to "lefty", and OPW branch $\theta_{1;ii}$ maps to "righty". VAL3 defines the elbow flags "epositive" and "enegative" by $\theta_3 \geq 0$ and $\theta_3 < 0$, respectively [111]. From the OPW definitions it follows using $a_2 = 0$ that

$$\theta_{3;i}, \theta_{3;iii} \geq 0, \qquad \theta_{3;ii}, \theta_{3;iv} \leq 0.$$

Therefore, $\theta_{3;i}$ and $\theta_{3;iii}$ are "epositive", while $\theta_{3;ii}$ and $\theta_{3;iv}$ are "enegative". VAL3 defines the wrist flags "wpositive" and "wnegative" by $\theta_5 \geq 0$ and $\theta_5 < 0$, respectively. The OPW definitions give $\theta_{5;p} \geq 0$, corresponding to "wpositive",

and $\theta_{5;q} = -\theta_{5;p} \leq 0$, corresponding to "wnegative". The resulting mapping is given in Table A.1, which also contains the numerical encoding $\mathrm{righty} = 1, \mathrm{lefty} = 2, \mathrm{epositive} = 1, \mathrm{enegative} = 2, \mathrm{wpositive} = 1, \mathrm{wnegative} = 2$.

Table A.1: Mapping between the eight OPW inverse-kinematics solutions [93] and the corresponding Stäubli VAL3 configuration flags and their numerical encoding.

OPW solution	OPW regional branch	shoulder	elbow	wrist	(s,e,w)
1	i	lefty	epositive	wpositive	(2,1,1)
2	ii	lefty	enegative	wpositive	(2,2,1)
3	iii	righty	epositive	wpositive	(1,1,1)
4	iv	righty	enegative	wpositive	(1,2,1)
5	i	lefty	epositive	wnegative	(2,1,2)
6	ii	lefty	enegative	wnegative	(2,2,2)
7	iii	righty	epositive	wnegative	(1,1,2)
8	iv	righty	enegative	wnegative	(1,2,2)

A.13 Fourier phase extraction for recording-stability measurement

The interferometric stability measurement in Sec. 5.3.6 evaluates the temporal drift of the recording fringe pattern generated by the two robot-mounted collimators. The two collimated fields are modeled as slowly varying envelopes multiplied by plane-wave carriers,

$$E_i(\mathbf{r}, t) = A_i(\mathbf{r}, t) \exp\{i\,[\mathbf{k}_i \cdot \mathbf{r} + \phi_i(\mathbf{r}, t)]\}\,, \qquad i \in \{1, 2\}. \tag{A.84}$$

Here, A_i describes the beam envelope, $\mathbf{k}_i$ is the carrier wave vector, and ϕ_i contains slowly varying phase contributions from mechanical motion, optical-path drift, and fiber-induced phase changes.

The recorded intensity is

$$I(\mathbf{r},t) = |E_1(\mathbf{r},t) + E_2(\mathbf{r},t)|^2 = A_1^2 + A_2^2 + 2A_1A_2\cos[\mathbf{k}_0\cdot\mathbf{r} + \varphi(\mathbf{r},t)]\,, \tag{A.85}$$

with $\mathbf{k}_0 = \mathbf{k}_1 - \mathbf{k}_2$ and $\varphi(\mathbf{r},t) = \phi_1(\mathbf{r},t) - \phi_2(\mathbf{r},t)$. Equation (A.85) can therefore be written in complex carrier form as

$$I(\mathbf{r},t) = a(\mathbf{r},t) + c(\mathbf{r},t)e^{i\mathbf{k}_0\cdot\mathbf{r}} + c^*(\mathbf{r},t)e^{-i\mathbf{k}_0\cdot\mathbf{r}}, \tag{A.86}$$

where $a(\mathbf{r},t) = A_1^2 + A_2^2$ and $c(\mathbf{r},t) = A_1A_2e^{i\varphi(\mathbf{r},t)}$. The beam envelopes and relative phase are assumed to vary slowly compared with the carrier term $\mathbf{k}_0\cdot\mathbf{r}$. Thus, taking the two-dimensional Fourier transform of Eq. (A.86) gives a central DC lobe and two conjugate sidebands centered at $\pm\mathbf{k}_0$. Following the Fourier-transform fringe-analysis method of Takeda *et al.* [97], one sideband is isolated with a circular window, shifted to the Fourier origin, and inverse Fourier transformed. This yields the complex fringe field $c(\mathbf{r},t)$, from which the local relative phase follows as

$$\varphi(\mathbf{r},t) = \arg c(\mathbf{r},t). \tag{A.87}$$

For the stability measurement, only the spatially averaged temporal drift of the recording pattern is required. We therefore compare each recovered complex field $c(\mathbf{r},t_n)$ to the first recorded frame $c_0(\mathbf{r}) = c(\mathbf{r},0)$. The average phase drift is evaluated over the beam-overlap region Ω as

$$\Delta\varphi(t_n) = \arg\sum_{\mathbf{r}\in\Omega} c_0^*(\mathbf{r})\,c(\mathbf{r},t_n) = \arg\sum_{\mathbf{r}\in\Omega} w(\mathbf{r},t_n)\exp\{i\,[\varphi(\mathbf{r},t_n) - \varphi(\mathbf{r},0)]\}\,. \tag{A.88}$$

where $w(\mathbf{r},t_n)$ is a weight proportional to the local beam amplitudes. For uniform recording beams, it is approximately constant across the overlap region. The resulting temporal phase trace is unwrapped in time.

A.14 Model for sparse-pupil grid modulation

This section derives the analytical model used in Sec. 5.5.3.2 to interpret the grid-like intensity modulation observed in the displayed image. We consider the 90° 2D-EBE waveguide introduced in Sec. 2.2.3 and use the same coordinate convention. The finite scanner beam at the incoupler is modeled as a circular top-hat ray bundle of diameter B, with radius $R_B = B/2$. The coordinate origin is chosen at the center of the incident beam on the incoupler.

The transverse components of the normalized guided wave vectors before and after redirection at the expander are, for the on-axis field,

$$\hat{\mathbf{k}}_{\mathrm{T},0}^{(1)} = \begin{pmatrix} \sin\theta_p \\ 0 \end{pmatrix}, \qquad \hat{\mathbf{k}}_{\mathrm{T},0}^{(2)} = \begin{pmatrix} 0 \\ -\sin\theta_p \end{pmatrix}, \tag{A.89}$$

where θ_p is the central guided propagation angle measured from the substrate normal. For an output field angle (θ_x, θ_y), conservation of tangential wave vector at the outcoupler gives

$$\hat{\mathbf{k}}_{\mathrm{T}}^{(j)}(\theta_x, \theta_y) = \hat{\mathbf{k}}_{\mathrm{T},0}^{(j)} + \frac{1}{n} \begin{pmatrix} \sin\theta_x \\ \sin\theta_y \end{pmatrix}, \qquad j = 1, 2, \tag{A.90}$$

with n the waveguide refractive index. Between two consecutive encounters with the same waveguide interface, a guided ray is laterally displaced by

$$\Delta\mathbf{r}_j(\theta_x, \theta_y) = 2d_{\mathrm{wg}} \frac{\hat{\mathbf{k}}_{\mathrm{T}}^{(j)}(\theta_x, \theta_y)}{\sqrt{1 - \left|\hat{\mathbf{k}}_{\mathrm{T}}^{(j)}(\theta_x, \theta_y)\right|^2}}, \qquad j = 1, 2. \tag{A.91}$$

At the field center this gives

$$\Delta\mathbf{r}_1(0,0) = \begin{pmatrix} L_0 \\ 0 \end{pmatrix}, \qquad \Delta\mathbf{r}_2(0,0) = \begin{pmatrix} 0 \\ -L_0 \end{pmatrix}, \qquad L_0 = 2d_{\mathrm{wg}} \tan\theta_p. \tag{A.92}$$

We assume that the expander and outcoupler are sufficiently large that edge clipping can be neglected locally. A ray launched with offset $\boldsymbol{\xi}$ from the beam center at the incoupler generates an entire 2D lattice of outcoupling-positions due to its guided propagation before and after expander redirection

$$\mathbf{r}_{mn}^{\mathrm{out}} = \boldsymbol{\xi} + m\,\Delta\mathbf{r}_1(\theta_x,\theta_y) + n\,\Delta\mathbf{r}_2(\theta_x,\theta_y), \qquad m,n\in\mathbb{Z}. \tag{A.93}$$

At eye relief s_{eye}, the rays outcoupled at these positions intersect the pupil plane at

$$\mathbf{r}_{mn}^{\mathrm{eye}} = \mathbf{r}_{mn}^{\mathrm{out}} + s_{\mathrm{eye}}\begin{pmatrix}\tan\theta_x\\ \tan\theta_y\end{pmatrix}. \tag{A.94}$$

We model the pupil as a hard aperture of radius R_P and diameter D centered at $\mathbf{r}_{\mathrm{e}} = (x_{\mathrm{e}}, y_{\mathrm{e}})^T$. Ignoring guided-mode depletion, non-uniform coupler efficiency, and angular selectivity, all outcoupled rays are assigned the same power. The detected angular intensity is then proportional to the number of outcoupled rays reaching the pupil, integrated over the incident beam's top-hat irradiance distribution

$$I(\theta_x,\theta_y) \propto \sum_{m,n\in\mathbb{Z}} \int \chi_{R_B}(\boldsymbol{\xi})\chi_{R_P}\left(\mathbf{r}_{mn}^{\mathrm{eye}} - \mathbf{r}_{\mathrm{e}}\right) d^2\boldsymbol{\xi}. \tag{A.95}$$

Here we describe the beam intensity distribution and the pupil by the circular aperture function

$$\chi_R(\mathbf{x}) = \begin{cases} 1, & |\mathbf{x}| < R,\\ 0, & |\mathbf{x}| \geq R. \end{cases} \tag{A.96}$$

With the definitions

$$\begin{aligned} \mathbf{q}(\theta_x,\theta_y) &= s_{\mathrm{eye}}\begin{pmatrix}\tan\theta_x\\ \tan\theta_y\end{pmatrix} - \mathbf{r}_{\mathrm{e}},\\ \mathbf{D}(\theta_x,\theta_y) &= \left(\Delta\mathbf{r}_1(\theta_x,\theta_y), \Delta\mathbf{r}_2(\theta_x,\theta_y)\right). \end{aligned} \tag{A.97}$$

Equation (A.95) becomes

$$I_B(\theta_x,\theta_y) \propto \sum_{m,n\in\mathbb{Z}} \int \chi_{R_B}(\boldsymbol{\xi})\chi_{R_P}\left(\boldsymbol{\xi}+\mathbf{q}+\mathbf{D}\begin{pmatrix}m\\n\end{pmatrix}\right) d^2\boldsymbol{\xi}$$
$$= \sum_{m,n\in\mathbb{Z}} f\left(\mathbf{D}\begin{pmatrix}m\\n\end{pmatrix}\right). \tag{A.98}$$

Where we write the convolution as

$$f(\mathbf{r}) = \int \chi_{R_B}(\boldsymbol{\xi})\chi_{R_P}(\boldsymbol{\xi}+\mathbf{q}+\mathbf{r})d^2\boldsymbol{\xi} = (\chi_{R_B} * \chi_{R_P})(\mathbf{q}+\mathbf{r}). \tag{A.99}$$

The last equality in Eq. (A.98) is a sum over the lattice $\Lambda = \mathbf{D}\mathbb{Z}^2$. We use the lattice Poisson summation formula [112, Sec. 1.3.2.6.7] with the Fourier-transform convention

$$\mathcal{F}[\varphi](\mathbf{Q}) = \int_{\mathbb{R}^2} \varphi(\mathbf{x}) e^{-i\mathbf{Q}\cdot\mathbf{x}} d^2\mathbf{x}, \qquad \varphi(\mathbf{x}) = \frac{1}{(2\pi)^2}\int_{\mathbb{R}^2} \mathcal{F}[\varphi](\mathbf{Q}) e^{i\mathbf{Q}\cdot\mathbf{x}} d^2\mathbf{Q}. \tag{A.100}$$

With this convention the reciprocal lattice is

$$\Lambda^* = 2\pi\mathbf{D}^{-T}\mathbb{Z}^2, \qquad \mathbf{Q}_{\boldsymbol{\ell}} = 2\pi\mathbf{D}^{-T}\boldsymbol{\ell}, \qquad \boldsymbol{\ell} = (\ell_1,\ell_2)^T, \tag{A.101}$$

and the lattice sum

$$\sum_{\mathbf{x}\in\Lambda} f(\mathbf{x}) = \frac{1}{|\det\mathbf{D}|}\sum_{\mathbf{Q}\in\Lambda^*} \mathcal{F}[f](\mathbf{Q}) = \frac{1}{|\det\mathbf{D}|}\sum_{\boldsymbol{\ell}\in\mathbb{Z}^2} \mathcal{F}[f](\mathbf{Q}_{\boldsymbol{\ell}}). \tag{A.102}$$

Using the shift and convolution theorems,

$$\mathcal{F}[f](\mathbf{Q}) = \mathcal{F}[\chi_{R_B}](\mathbf{Q})\mathcal{F}[\chi_{R_P}](\mathbf{Q})e^{i\mathbf{Q}\cdot\mathbf{q}}. \tag{A.103}$$

The Fourier transform of a circular top-hat aperture is the well-known Airy pattern $\mathcal{F}[\chi_R](\mathbf{Q}) = \pi R^2\,\mathrm{jinc}(R|\mathbf{Q}|)$. Absorbing the determinant factor that

slowly varies across the field of view into the proportionality constant gives the intensity

$$I(\theta_x, \theta_y) \propto \sum_{\boldsymbol{\ell} \in \mathbb{Z}^2} \operatorname{jinc}(R_B |\mathbf{Q}_{\boldsymbol{\ell}}|) \operatorname{jinc}(R_P |\mathbf{Q}_{\boldsymbol{\ell}}|) e^{i \mathbf{Q}_{\boldsymbol{\ell}} \cdot \mathbf{q}}. \tag{A.104}$$

Combining reciprocal-lattice terms in pairs $\boldsymbol{\ell}$ and $-\boldsymbol{\ell}$ gives the real cosine series

$$I(\theta_x, \theta_y) \propto 1 + 2 \sum_{\boldsymbol{\ell} \in \mathcal{S}} \operatorname{jinc}(R_B |\mathbf{Q}_{\boldsymbol{\ell}}|) \operatorname{jinc}(R_P |\mathbf{Q}_{\boldsymbol{\ell}}|) \cos\left(\mathbf{Q}_{\boldsymbol{\ell}} \cdot \mathbf{q}\right), \tag{A.105}$$

where $\mathcal{S}$ contains one representative from each nonzero pair $\pm\boldsymbol{\ell}$ and we define the jinc-function as

$$\operatorname{jinc}(x) = \frac{2 J_1(x)}{x} \tag{A.106}$$

where J_1 is the Bessel function of the first kind. We keep only the two fundamental orders $(1, 0)$ and $(0, 1)$. The higher orders are strongly suppressed by the jinc functions. Using that near the field center $|\mathbf{Q}_{10}| \approx |\mathbf{Q}_{01}| \approx \frac{2\pi}{L_0}$ this gives

$$I_B(\theta_x, \theta_y) \propto 1 + 2 \operatorname{jinc}\left(\frac{2\pi R_B}{L_0}\right) \operatorname{jinc}\left(\frac{2\pi R_P}{L_0}\right) \left[\cos(\mathbf{Q}_{10} \cdot \mathbf{q}) + \cos(\mathbf{Q}_{01} \cdot \mathbf{q})\right]. \tag{A.107}$$

Equivalently, using $R_B = B/2$ and $R_P = D/2$, the contrast coefficient used in the main text is

$$C = 2 \left| \operatorname{jinc}\left(\frac{\pi B}{L_0}\right) \operatorname{jinc}\left(\frac{\pi D}{L_0}\right) \right|. \tag{A.108}$$

The absolute value is used because a negative product of jinc factors is equivalent to a phase shift of π. To make explicit that the fundamental orders explain the observed grid-like intensity modulation across the field of view, we linearize the phase terms in the field angles. At the field center,

$$\mathbf{D}_0 = \begin{pmatrix} L_0 & 0 \\ 0 & -L_0 \end{pmatrix}, \qquad \mathbf{Q}_{10}^{(0)} = \begin{pmatrix} 2\pi/L_0 \\ 0 \end{pmatrix}, \qquad \mathbf{Q}_{01}^{(0)} = \begin{pmatrix} 0 \\ -2\pi/L_0 \end{pmatrix}. \tag{A.109}$$

The first-order derivatives of the jump vectors are

$$\left.\frac{\partial \Delta \mathbf{r}_1}{\partial \theta_x}\right|_0 = \frac{L_0}{n \sin \theta_p} \begin{pmatrix} \sec^2 \theta_p \\ 0 \end{pmatrix}, \quad \left.\frac{\partial \Delta \mathbf{r}_2}{\partial \theta_x}\right|_0 = \frac{L_0}{n \sin \theta_p} \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \tag{A.110}$$

$$\left.\frac{\partial \Delta \mathbf{r}_1}{\partial \theta_y}\right|_0 = \frac{L_0}{n \sin \theta_p} \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \qquad \left.\frac{\partial \Delta \mathbf{r}_2}{\partial \theta_y}\right|_0 = \frac{L_0}{n \sin \theta_p} \begin{pmatrix} 0 \\ \sec^2 \theta_p \end{pmatrix}. \tag{A.111}$$

Using $\mathbf{Q}_{\boldsymbol{\ell}} = 2\pi \mathbf{D}^{-T} \boldsymbol{\ell}$ and linearizing the tangents in the definition of $\mathbf{q}$, the two fundamental phases close to the field of view center become

$$\mathbf{Q}_{10} \cdot \mathbf{q} \approx 2\pi \left[-\frac{x_\mathrm{e}}{L_0} + \frac{s_\mathrm{eye} + L_{\mathrm{eff},x}}{L_0} \theta_x \right], \tag{A.112}$$

$$\mathbf{Q}_{01} \cdot \mathbf{q} \approx 2\pi \left[\frac{y_\mathrm{e}}{L_0} - \frac{s_\mathrm{eye} + L_{\mathrm{eff},y}}{L_0} \theta_y \right], \tag{A.113}$$

where

$$L_{\mathrm{eff},x} = \frac{x_\mathrm{e} \sec^2 \theta_p - y_\mathrm{e}}{n \sin \theta_p}, \qquad L_{\mathrm{eff},y} = \frac{x_\mathrm{e} - y_\mathrm{e} \sec^2 \theta_p}{n \sin \theta_p}. \tag{A.114}$$

Thus

$$\Delta\theta_x = \frac{L_0}{s_\mathrm{eye} + L_{\mathrm{eff},x}}, \qquad \Delta\theta_y = \frac{L_0}{s_\mathrm{eye} + L_{\mathrm{eff},y}}. \tag{A.115}$$

In summary, we write

$$I(\theta_x, \theta_y) \propto 1 + C \left[\cos\left(\frac{2\pi}{\Delta\theta_x} \theta_x + \varphi_x \right) + \cos\left(\frac{2\pi}{\Delta\theta_y} \theta_y + \varphi_y \right) \right]. \tag{A.116}$$

Validity The model describes the fundamental grid period and the leading contrast scaling. Away from the field center, the full expression Eq. (A.105) should be used. Higher-order terms in θ_x and θ_y introduce grid curvature through θ_x^2, θ_y^2, and $\theta_x \theta_y$ contributions. In addition, finite expander and outcoupler apertures, guided beam depletion, and Bragg selectivity modify the detailed intensity pattern. Full

finite-aperture ray tracing is therefore required for quantitative prediction over the complete field of view.

A.15 Modulation transfer function measurement

The modulation transfer function is evaluated from checkerboard captures using the slanted-edge method [100]. Figure A.7 shows the captures and selected edge regions. Edges are evaluated near the field center and near the four field corners, with orientations approximately orthogonal and parallel to the projector scan-line direction. Table A.2 reports all extracted MTF_{20} spatial frequencies, in lp/mm and lp/deg, for each evaluated eyebox position, field position, and edge orientation.

Table A.2: Extracted MTF_{20} spatial frequencies (lp/mm / lp/°) for orthogonal (⊥) and parallel (∥) edges, at five field positions, across five eyebox positions and without the waveguide.

Eyebox	Edge	spatial frequency, lp/mm / lp/°				
		top-left	top-right	center	bottom-left	bottom-right
(-3,+3)	⊥	13.4 / 5.9	18.4 / 8.0	21.6 / 9.4	11.7 / 5.1	18.3 / 8.0
	∥	14.8 / 6.5	18.9 / 8.3	25.1 / 11.0	16.8 / 7.3	20.8 / 9.1
(+3,+3)	⊥	11.2 / 4.9	20.5 / 9.0	22.7 / 9.9	13.4 / 5.8	20.0 / 8.7
	∥	17.9 / 7.8	19.1 / 8.3	23.3 / 10.2	16.6 / 7.3	18.8 / 8.2
(0,0)	⊥	17.8 / 7.8	19.3 / 8.4	26.0 / 11.3	12.0 / 5.2	20.2 / 8.8
	∥	18.3 / 8.0	19.9 / 8.7	23.4 / 10.2	16.8 / 7.3	20.0 / 8.7
w/o WG	⊥	26.8 / 11.7	34.4 / 15.0	26.5 / 11.6	25.0 / 10.9	36.5 / 15.9
	∥	27.3 / 11.9	30.2 / 13.2	33.0 / 14.4	18.7 / 8.2	23.2 / 10.1
(-3,-3)	⊥	9.1 / 4.0	17.9 / 7.8	21.2 / 9.2	10.8 / 4.7	19.5 / 8.5
	∥	9.2 / 4.0	22.7 / 9.9	23.3 / 10.2	16.4 / 7.2	19.9 / 8.7
(+3,-3)	⊥	11.0 / 4.8	19.1 / 8.3	21.1 / 9.2	12.6 / 5.5	22.7 / 9.9
	∥	20.9 / 9.1	20.9 / 9.1	25.0 / 10.9	16.9 / 7.4	23.0 / 10.0

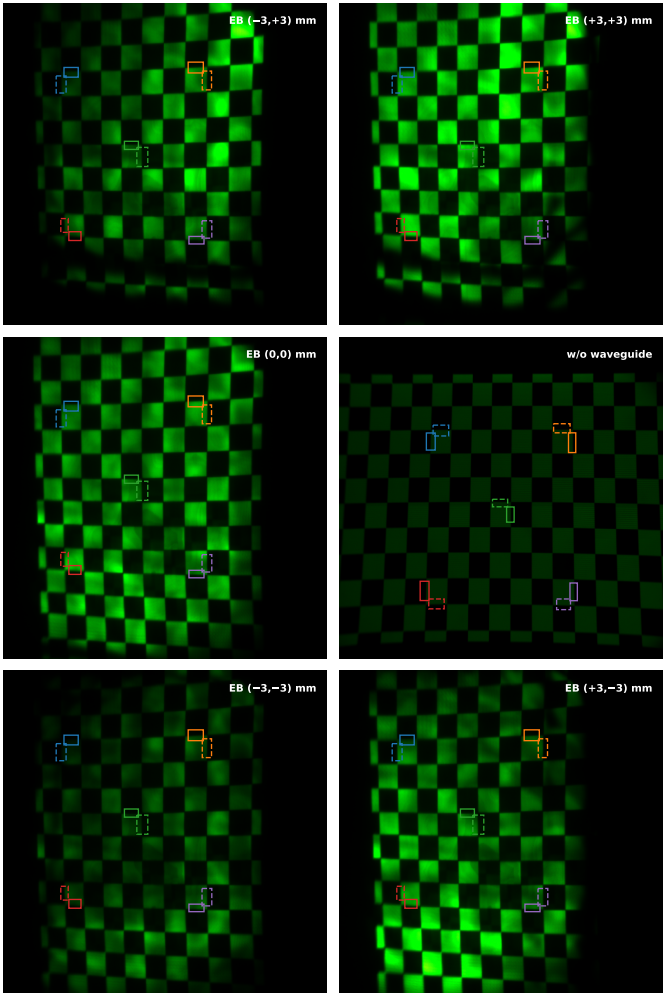


Figure A.7: Checkerboard captures used for resolution evaluation. The color outlines show the edge regions used for the MTF calculation.

Acknowledgements

I would like to acknowledge the many people who contributed to this work and supported me throughout my PhD.

First and foremost, I would like to thank Prof. Dr. Wilhelm Stork for his excellent supervision and guidance throughout this work. I am deeply grateful for our discussions on holography and optics, as well as on research, entrepreneurship, and life more broadly. I particularly appreciate the freedom he gave me to pursue my research in collaboration with Bosch Research and the trust he placed in me to balance the academic and industrial requirements of this work. His approach to research, with its strong emphasis on creating impact beyond the academic world, has had a lasting influence on me.

I would also like to thank Prof. Dr. Alois Herkommer for supporting this work, for his valuable feedback, and for serving as my second reviewer. I am grateful for the opportunity to jointly supervise one master's student with him and to deliver a seminar on holographic waveguide design within his research group. I very much enjoyed our discussions about my work.

My sincere thanks go to Dr. Reinhold Fiess, who supervised this work at Bosch Research. Reinhold is a senior expert in holography and display optics, an exceptional engineer, and, above all, a great mentor and colleague. I learned a great deal from him about navigating a large company, steering a research project, and maintaining a practical, hands-on approach to research and development. When I began this work, I considered myself primarily a scientist. Thanks in large part to Reinhold, I have also become an engineer.

Reinhold proposed the use of six-axis robots for holographic waveguide recording, which formed the initial concept underlying this work. He also taught me a great deal about the formation of spurious holograms caused by stray light, which contributed significantly to the development of the edge-lit prism system presented in Chapter 5. Together, we recorded the holographic waveguide presented in that chapter and optimized the immersion and exposure-process parameters.

I would like to thank my colleague Dr. Tobias Wilm, who has been both a mentor and a valued colleague at Bosch Research. He introduced me to the field of holography, shared his expertise with me, and provided extensive and constructive feedback on my work. I learned a lot from him.

I am equally grateful to Dr. Moritz Esslinger for our many discussions on optical design, from which I learned a great deal. The achievable field-of-view calculation presented in Chapter 2 was originally derived by him. He introduced me to the concept of cylindrical waveguide eigenmodes and provided a proof sketch explaining why corresponding eigenmodes do not exist in spherical waveguides. These insights, together with his numerical studies of curved waveguide designs, were fundamental to my formulation of closed modes for spherical waveguide design.

I would also like to thank Christian Grafenburg for his support in preparing the holographic film used for the sample presented in Chapter 5, as well as for the assembly and testing of the characterization setup briefly described in the outlook of this thesis. I am similarly grateful to my intern, Aurore Bader, for her support with these tasks. The idea of this setup also drew on preliminary work conducted together with my bachelor's student Keno Bürger, whom I would like to thank for his contribution.

My thanks also go to my master's student Yuchen Deng for contributing to the freeform design algorithm briefly described in the outlook, and to Simon Jessen and Jens Hofele for their assistance with robot interfacing and robotics more generally.

I am deeply grateful to my line manager, Hartmut Spennemann, for hiring me at Bosch and for providing continuous managerial support throughout my work. I would also like to thank my project manager, Dr. Thomas Schlebusch, for his guidance and support. Conducting doctoral research in an industrial environment requires balancing the complementary objectives of academic and corporate research and development. I am therefore particularly grateful to Hartmut Spennemann, Dr. Thomas Schlebusch, and Dr. Reinhold Fiess for their exceptional leadership in aligning these objectives and for fostering an environment in which I could produce academically meaningful results while also contributing to Bosch's development goals. I gratefully acknowledge Robert Bosch GmbH for funding this work.

More broadly, I would like to thank the many colleagues at Bosch Research who supported me throughout my PhD. Their expertise, encouragement, and willingness to share their time and knowledge contributed greatly to both this work and my professional development.

During the course of my PhD, the capabilities of large language models increased substantially, and I incorporated these tools into selected aspects of my research workflow. Models provided by OpenAI and Anthropic were used for manuscript review and language polishing, supporting targeted literature and patent searches, discussions during the development of research ideas and mathematical derivations, and drafting specific computer code for laboratory hardware control and optical analysis. All generated content was critically reviewed, and all code and technical outputs were manually verified and tested before use.

Abschließend möchte ich meinen Freunden und meiner Familie von Herzen für ihre große Unterstützung während meiner Promotionszeit danken. Mein ganz besonderer Dank gilt meiner Mutter und meinem Vater, die mich auf diesem Weg stets begleitet, ermutigt und unterstützt haben.

List of publications

Journal articles

Marvin Feil, Moritz Esslinger, Reinhold Fiess, and Wilhelm Stork. Closed-mode families for planar, cylindrical, and spherical waveguide display design. *Optics Express*, 34(14):26031–26046, 2026. DOI: 10.1364/OE.606418. **(Editors' Pick)**.

Marvin Feil, Tobias Wilm, Moritz Esslinger, Reinhold Fiess, and Wilhelm Stork. Limits of immersion-free recording of holographic waveguide displays. *Optics Letters*, 50(2):606–609, 2025. DOI: 10.1364/OL.545451.

Marvin Feil, Reinhold Fiess, and Wilhelm Stork. Automated mastering of holographic waveguide displays with robot-mounted optics. *Appl. Opt.*, 65(28): 9786–9797, 2026. DOI: 10.1364/AO.608922.

Daniel Flamm, Myriam Kaiser, Marvin Feil, Max Kahmann, Michael Lang, Jonas Kleiner, and Tim Hesse. Protecting the edge: Ultrafast laser modified c-shaped glass edges. *Journal of Laser Applications*, 34(1):012014, 2021. DOI: 10.2351/7.0000592. **(JLA Best Paper Award 2023)**.

Zhen Wang, Marvin Feil, Reinhold Fiess, Wilhelm Stork, Ulrich Lemmer, and Dominik Theobald. Compact and high-efficiency multiwavelength beam combiner using transmission volume phase holograms. *Accepted at Optics Express, Issue in Progress*. DOI: https://doi.org/10.1364/OE.609092.

Conference contributions

Marvin Feil, Reinhold Fiess, and Wilhelm Stork. Automated mastering of holographic waveguide displays. In *Optical Architectures for Displays and Sensing in Augmented, Virtual, and Mixed Reality (AR, VR, MR) VII*, volume 13821, page 138210A. SPIE, 2026. DOI: 10.1117/12.3080347.

Marvin Feil, Tobias Wilm, Reinhold Fiess, and Wilhelm Stork. Automated holographic recording using six-axis robot arm pair. In *Holography: Advances and Modern Trends IX*, volume 13529, page 135290J. SPIE, 2025. DOI: 10.1117/12.3055606.

Marvin Feil, Moritz Esslinger, Yuchen Deng, Reinhold Fiess, and Wilhelm Stork. Freeform Recording of Curved Holographic Waveguides, 2026. **1st Place Winner**, *SPIE AR|VR|MR* Student Optical Design Challenge. URL: https://spie.org/conferences-and-exhibitions/photonics-west/program/conferences/ar-vr-mr/student-optical-design-challenge. Accessed 10. July 2026.

Patents

Marvin Feil, Theodorus Thomas Marinus van Schaijk, and Robin Daniel Buijs. *Metrology method for a digital holographic microscope and associated computer program*. Original Assignee: ASML Netherlands B.V., Patent No. WO2025131523A1, pending. Filed 21. November 2024.

Carsten Reichert, Johannes Meyer, and Marvin Feil. *Eye tracking device, duplication optics and data glasses*. Original Assignee: Robert Bosch GmbH, Patent No. DE102024205166A1, pending. Filed 5. June 2024.

Carsten Reichert, Johannes Meyer, and Marvin Feil. *Optical system for a virtual retinal display*. Original Assignee: Robert Bosch GmbH, Patent No. DE102024206046A1, pending. Filed 28. June 2024.

Note: 15 patent filings (10 first authored) are currently under review.

Supervised student theses and internships

2025 **Yuchen Deng**

M.Sc. Thesis: *Design of Freeform Recording Optics for Curved Waveguides*

University of Stuttgart, Institute of Applied Optics (ITO)

2025 **Aurore Elisabeth Bader**

Mandatory Internship (Stage EPFL): *Optimization of an Interference Photolithography Process for AR Applications*

École Polytechnique Fédérale de Lausanne (EPFL)

2024 **Keno Bürger**

B.Eng. Thesis: *Characterization of Large Angle Holograms in Photosensitive Polymers*

Duale Hochschule Baden-Württemberg (DHBW)

Bibliography

[1] Ozan Cakmakci, David M. Hoffman, and Nikhil Balram. 31-4: Invited paper: 3d eyebox in augmented and virtual reality optics. *SID Symposium Digest of Technical Papers*, 50(1):438–441, 2019. DOI: 10.1002/sdtp.12950.

[2] Matthias Hillenbrand, Wolfgang Singer, and Holger Münz. Tackling the challenges of fully immersive head-mounted AR devices. In *International Optical Design Conference 2017*, page 50, Denver, United States, 2017. SPIE. DOI: 10.1117/12.2292077.

[3] Bernard C. Kress. *Optical Architectures for Augmented-, Virtual-, and Mixed-Reality Headsets*. SPIE, 2020. ISBN: 9781510634336.

[4] Howard Grubb. A New Collimating-Telescope Gun-Sight for Large and Small Ordnance. *The Scientific Transactions of the Royal Dublin Society*, 7(10): 321–330, 1901.

[5] Richard L. Newman. *Head-Up Displays: Designing the Way Ahead*, chapter 2. Ashgate Publishing, 1995. ISBN: 13:978-0-291-39811-6.

[6] Charles P. Comeau and James S. Bryan. Headsight television system provides remote surveillance. *Electronics*, pages 86–90, 1961.

[7] Simon Malcolm St. Leger Searle. *Display Apparatus*. Original Assignee: Smiths Industries PLC, Patent No. GB2039383B, 1982. Filed 17. January 1980.

[8] Yaakov Amitai and Asher Friesem. *Holographic Optical Device*. Original Assignee: Yeda Research and Development Co. Ltd., Patent No. US6580529B1, 17. June 2003. Filed 29. March 1999.

[9] Revital Shechter, Yaakov Amitai, and Asher A. Friesem. Compact beam expander with linear gratings. *Applied Optics*, 41(7):1236–1240, 2002. DOI: 10.1364/AO.41.001236.

[10] Juris Upatnieks. Edge-illuminated holograms. *Appl. Opt.*, 31(8):1048–1052, 1992. DOI: 10.1364/AO.31.001048.

[11] Juris Upatnieks. Compact holographic sight. In *Annual Meeting Optical Society of America*, page FF4. Optica Publishing Group, 1985. DOI: 10.1364/OAM.1985.FF4.

[12] Juris Upatnieks. Compact holographic sight. In *Holographic Optics: Design and Applications*, volume 883 of *Proceedings of SPIE*. SPIE, 1988. DOI: 10.1117/12.944140.

[13] Yaakov Amitai, S. Reinhorn, and Asher A. Friesem. Visor-display design based on planar holographic optics. *Applied Optics*, 34(8):1352–1356, 1995. DOI: 10.1364/AO.34.001352.

[14] Yaakov Amitai, I. Shariv, M. Kroch, Asher A. Friesem, and S. Reinhorn. White-light holographic display based on planar optics. *Optics Letters*, 18 (15):1265–1267, 1993. DOI: 10.1364/OL.18.001265.

[15] Ichiro Kasai, Yasushi Tanijiri, Takeshi Endo, and Hiroaki Ueda. A Practical See-Through Head Mounted Display Using a Holographic Optical Element. *Optical Review*, 8(4):241–244, 2001. DOI: 10.1007/s10043-001-0241-z.

[16] Sony Group. Press release: AR HMD—Created from the ground up: AR pioneer Hiroshi Mukawa talks about 20 years of development in optical display, 2024. URL: `https://www.sony.com/en/SonyInfo/technology/stories/entries/interview_de_mukawa/`. Accessed 3. July 2026.

[17] Hiroshi Mukawa, Katsuyuki Akutsu, Ikuo Matsumura, Satoshi Nakano, Takuji Yoshida, Mieko Kuwahara, and Kazuma Aiki. A full-color eyewear display using planar waveguides with reflection volume holograms. *Journal of the Society for Information Display*, 17(3):185–193, 2009. DOI: 10.1889/JSID17.3.185.

[18] Hiroshi Mukawa, Katsuyuki Akutsu, Ikuo Matsumura, Satoshi Nakano, Takuji Yoshida, Mieko Kuwahara, Kazuma Aiki, and Masataka Ogawa. 8.4: Distinguished Paper: A Full Color Eyewear Display Using Holographic Planar Waveguides. *SID Symposium Digest of Technical Papers*, 39(1):89–92, 2008. DOI: 10.1889/1.3069819.

[19] BAE Systems. Press release: Royal navy takes first delivery of q-sight helmet-mounted display for lynx helicopters, 2010. URL: https://www.baesystems.com/en/article/royal-navy-takes-first-delivery-of-q-sight-helmet-mounted-display-for-lynx-helicopters. Accessed 3. July 2026.

[20] Alex Cameron. The application of holographic optical waveguide technology to the Q-Sight family of helmet-mounted displays. In *Head- and Helmet-Mounted Displays XIV: Design and Applications*, volume 7326, page 73260H. International Society for Optics and Photonics, SPIE, 2009. DOI: 10.1117/12.818581.

[21] Tapani Levola. 7.1: Invited paper: Novel diffractive optical components for near to eye displays. *SID Symposium Digest of Technical Papers*, 37(1): 64–67, 2006. DOI: 10.1889/1.2433589.

[22] Yuqiang Ding, Qian Yang, Yannanqi Li, Zhiyong Yang, Zhengyang Wang, Haowen Liang, and Shin-Tson Wu. Waveguide-based augmented reality displays: perspectives and challenges. *eLight*, 3(1):24, 2023. DOI: 10.1186/s43593-023-00057-z.

[23] Yaakov Amitai. P-27: A two-dimensional aperture expander for ultra-compact, high-performance head-worn displays. *SID Symposium Digest of Technical Papers*, 36(1):360–363, 2005. DOI: 10.1889/1.2036446.

[24] Yaakov Amitai. P-21: Extremely Compact High-Performance HMDs Based on Substrate-Guided Optical Element. *SID Symposium Digest of Technical Papers*, 35(1):310–313, 2004. DOI: 10.1889/1.1830976.

[25] Lumus Ltd. Time: Meta Ray-Ban Display Glasses Selected as "Best Inventions of the Year", 2026. URL: https://lumus.com/time-meta-ray-ban-display-glasses-selected-as-best-inventions-of-the-year/. Accessed 3. July 2026.

[26] Darryl Meister and James E Sheedy. *Introduction to Ophthalmic Optics*. Carl Zeiss Vision, 6th edition, 2008.

[27] Tobias Graf. Curved holographic optical elements from a geometric view point. *Optical Engineering*, 60(3):035102, 2021. DOI: 10.1117/1.OE.60.3.035102.

[28] Angeles Marcia Almanza-Workman and Jon H. Appleby. *Ophthalmic Lens Assembly*. Original Assignee: Amazon Technologies, Inc., Patent No. US12468178B2, 11. November 2025. Filed 29. September 2021.

[29] Craig Thomas Draper and Pierre-Alexandre Blanche. *Light Combiner Employing a Curved Waveguide System.* Original Assignee: Arizona Board of Regents on behalf of The University of Arizona, Patent No. US11333893B1, 17. May 2022. Filed 13. May 2020.

[30] Pierre-Alexandre Blanche and Craig T. Draper. Curved waveguide combiner for HUD/AR. In *ODS 2021: Industrial Optical Devices and Systems*, volume 11828, page 1182806. SPIE, 2021. DOI: 10.1117/12.2592767.

[31] Craig T. Draper and Pierre-Alexandre Blanche. Holographic curved waveguide combiner for HUD/AR with 1-D pupil expansion. *Opt. Express*, 30(2): 2503–2516, 2022. DOI: 10.1364/OE.445091.

[32] Andrey Volkov, Ben Sherliker, Helen Smith, Danica Sugic, Sophie Viaene, and Kingsley Hanlon. Curved waveguide with replicated eye-boxes. In *Optical Architectures for Displays and Sensing in Augmented, Virtual, and Mixed Reality (AR, VR, MR) III*, volume 11931, page 119310N. SPIE, 2022. DOI: 10.1117/12.2615024.

[33] Gavril N. Vostrikov, Nikolay V. Muravyev, Aleksandr E. Angervaks, Roman A. Okun, Anastasia S. Perevoznikova, Jaeyeol Ryu, and Andrei N.

Putilin. Method for Compensating Aberrations of a Virtual Image Formed by an Augmented Reality Display Based on a Cylindrical Diffractive Waveguide. *Applied Sciences*, 13(4):2400, 2023. DOI: 10.3390/app13042400.

[34] Jiacheng Weng, Chunyang Pei, Haifeng Li, Rengmao Wu, and Xu Liu. Design and fabrication of curved waveguide display based on freeform polarization volume holograms. *Optics Express*, 33(7):15362–15379, 2025. DOI: 10.1364/OE.557298.

[35] Michael David Simmonds and Mohmed Salim Valera. *Projection Display*. Original Assignee: BAE Systems plc, Patent No. US8810913B2, 19. August 2014. Filed 21. January 2011.

[36] Michael David Simmonds and Mohmed Salim Valera. *Head Mounted Display*. Original Assignee: BAE Systems PLC, Patent No. US8842368B2, 23. September 2014. Filed 27. April 2010.

[37] Oscar A. Martinez and Ozan Cakmakci. *Adding Prescriptive Correction to Eyepieces for See-Through Head Wearable Displays*. Original Assignee: Google LLC, Patent No. US10146054B2, 4. December 2018. Filed 6. July 2015.

[38] Dongwei Ni, Dewen Cheng, Yongdong Wang, Tong Yang, Ximeng Wang, Cheng Chi, and Yongtian Wang. Design and fabrication method of holographic waveguide near-eye display with 2D eye box expansion. *Opt. Express*, 31(7):11019–11040, 2023. DOI: 10.1364/OE.481889.

[39] Takuji Yoshida, Kazutatsu Tokuyama, Yuichi Takai, Daisuke Tsukuda, Tsuyoshi Kaneko, Nobuhiro Suzuki, Takafumi Anzai, Akira Yoshikaie, Katsuyuki Akutsu, and Akio Machida. A plastic holographic waveguide combiner for light-weight and highly-transparent augmented reality glasses. *Journal of the Society for Information Display*, 26(5):280–286, 2018. DOI: 10.1002/jsid.659.

[40] J.J. Sirvent-Verdú, J.C. Bravo, J. Colomina-Martínez, C. Piñol-Galera, G. Nájar, C. Neipp, J. Francés, S. Gallego, and A. Beléndez. Manufacturing reflection holographic couplers for see-through applications recorded

in photopolymers without prisms: An experimental validation. *EPJ Web Conf.*, 287:09046, 2023. DOI: 10.1051/epjconf/202328709046.

[41] Dipanjan Chakraborty, Rosen Georgiev, Vincent Toal, Izabela Naydenova, Dervil Cody, and Suzanne Martin. Comparison of holographic recording options for reflection-format and transmission-format coupler-type diffractive optical elements: theoretical exploration and experimental validation. *Optics Express*, 32(11):20385, 2024. DOI: 10.1364/OE.522326.

[42] Friedrich-Karl Bruder, Johannes Frank, Sven Hansen, Alexander Lorenz, Christel Manecke, Richard Meisenheimer, Jack Mills, Lena Pitzer, Igor Pochorovski, Thomas Roelle, and Brita Wewer. Demonstrations of Bayfol HX vHOEs in see-through display applications. In *Optical Architectures for Displays and Sensing in Augmented, Virtual, and Mixed Reality (AR, VR, MR) III*, volume 11931, page 119310B, San Francisco, United States, 2022. SPIE. DOI: 10.1117/12.2614633.

[43] Yu. N. Denisyuk. On the reproduction of the optical properties of an object by the wave field of its scattered radiation. *Optics and Spectroscopy*, 15: 279–284, 1963.

[44] D. Gabor. A New Microscopic Principle. *Nature*, 161(4098):777–778, 1948. DOI: 10.1038/161777a0.

[45] D. H. Close. Holographic optical elements. *Optical Engineering*, 14(5): 145408, 1975. DOI: 10.1117/12.7971806.

[46] Norbert Lindlein. *Analyse und Optimierung diffraktiver optischer Systeme*. PhD thesis, Friedrich-Alexander-Universität Erlangen-Nürnberg, 1996.

[47] Changwon Jang, Olivier Mercier, Kiseung Bang, Gang Li, Yang Zhao, and Douglas Lanman. Design and Fabrication of Freeform Holographic Optical Elements. *ACM Trans. Graph.*, 39(6), 2020. DOI: 10.1145/3414685.3417762.

[48] Horst Berneth, Friedrich-Karl Bruder, Thomas Fäcke, Rainer Hagen, Dennis Hönel, Thomas Rölle, Günther Walze, and Marc-Stephan Weiser. Holographic recordings with high beam ratios on improved Bayfol® HX photopolymer. In *Holography: Advances and Modern Trends III*, volume 8776, page 877603. International Society for Optics and Photonics, SPIE, 2013. DOI: 10.1117/12.2018618.

[49] Friedrich-Karl Bruder, Johannes Frank, Sven Hansen, Mira Holzheimer, Alexander Lorenz, Christel Manecke, Jack Mills, Lena Nault, Igor Pochorovski, and Thomas Rölle. Generating diffraction efficiency profiles in Bayfol HX vHOEs. In *Holography: Advances and Modern Trends VIII*, volume 12574, page 125740F. SPIE, 2023. DOI: 10.1117/12.2665867.

[50] D. A. Waldman, H.-Y. S. Li, and M. G. Horner. Volume Shrinkage in Slant Fringe Gratings of a Cationic Ring-Opening Holographic Recording Material. *Journal of Imaging Science and Technology*, 41(5):497–514, 1997. DOI: 10.2352/J.ImagingSci.Technol.1997.41.5.art00008.

[51] W. Friedrich, P. Knipping, and M. Laue. Interferenzerscheinungen bei Röntgenstrahlen. *Annalen der Physik*, 346(10):971–988, 1913. DOI: 10.1002/andp.19133461004.

[52] P. P. Ewald. Zur theorie der interferenzen der röntgenstrahlen in kristallen. *Physikalische Zeitschrift*, 14:465–472, 1913.

[53] Herwig Kogelnik. Coupled wave theory for thick hologram gratings. *Bell System Technical Journal*, 48(9):2909–2947, 1969. DOI: 10.1002/j.1538-7305.1969.tb01198.x.

[54] Han-Hsiang Cheng. An accurate volume hologram model for optical design: simulating a multi-color holographic display using ray tracing. In *AOPC 2019: Display Technology and Optical Storage*, volume 11335, page 113350D, Beijing, China, 2019. SPIE. DOI: 10.1117/12.2547729.

[55] Walter T. Welford. A vector raytracing equation for hologram lenses of arbitrary shape. *Optics Communications*, 14(3):322–323, 1975. DOI: 10.1016/0030-4018(75)90327-2.

[56] Friedrich-Karl Bruder, Johannes Frank, Sven Hansen, Mira Holzheimer, Alexander Lorenz, Christel Manecke, Jack Mills, Lena Nault, Igor Pochorovski, Thomas Rölle, and Brita Wewer. Fabricating a complete set of Bayfol HX vHOEs with applications in see through displays using free space recording techniques. In *Optical Architectures for Displays and Sensing in Augmented, Virtual, and Mixed Reality (AR, VR, MR) V*, volume 12913, page 129130U, San Francisco, United States, 2024. SPIE. DOI: 10.1117/12.3001728.

[57] Ao Liu, Yuning Zhang, Zhongwen Shen, Yishi Weng, Jun Xia, and Baoping Wang. 73-3: High Refractive Index Photopolymer Fabricated Holographic Grating used for RGB Waveguide-type Display. *SID Symposium Digest of Technical Papers*, 50(1):1042–1044, 2019. DOI: 10.1002/sdtp.13105.

[58] Alejo Lifschitz. High-refractive index holographic materials for AR waveguide fabrication. In *Optical Architectures for Displays and Sensing in Augmented, Virtual, and Mixed Reality (AR, VR, MR) VI*, volume PC13414, page PC1341407. SPIE, 2025. DOI: 10.1117/12.3041448.

[59] T. Momose, A. Takaki, H. Sato, T. Nakao, and T. Taniguchi. Non-fluorinated low-refractive-index coatings for XR applications. In *Optical Architectures for Displays and Sensing in Augmented, Virtual, and Mixed Reality (AR, VR, MR) VII*, volume 13821, page 138210U. International Society for Optics and Photonics, SPIE, 2026. DOI: 10.1117/12.3079778.

[60] Yujia Du, Haining Wang, Yuxin Ju, and Meng Qiu. The effects of physical load and clamping force on prolonged wearing of AR glasses. In Francisco Rebelo, editor, *Ergonomics in Design*, volume 47 of *AHFE Open Access*, pages 105–112, New York, NY, 2022. AHFE International. DOI: 10.54941/ahfe1001926.

[61] Yong Min Kim, Sangwoo Bahn, and Myung Hwan Yun. Wearing comfort and perceived heaviness of smart glasses. *Human Factors and Ergonomics in Manufacturing & Service Industries*, 31:484 – 495, 2021. DOI: 10.1002/hfm.20895.

[62] Mister Spex. Ray-Ban Wayfarer RX 4340V, n.d. URL: https://www.misterspex.de/p/pg/6680201. Accessed 26. July 2026.

[63] Meta. Ray-Ban Meta Wayfarer (Gen 2), n.d. URL: https://www.meta.com/de/ai-glasses/ray-ban-meta-wayfarer-gen-2/. Accessed 26. July 2026.

[64] Mister Spex. Ray-Ban RB 2184 902/57, n.d. URL: https://www.misterspex.de/p/sg/6749904. Accessed 26. July 2026.

[65] Mister Spex. Oakley Flak 2.0 XL OO 9188 G8, n.d. URL: https://www.misterspex.de/p/sg/6846161. Accessed 26. July 2026.

[66] Katie M. Williams, Virginie J. M. Verhoeven, Phillippa Cumberland, Geir Bertelsen, Christian Wolfram, Gabriélle H. S. Buitendijk, Albert Hofman, Cornelia M. van Duijn, Johannes R. Vingerling, Robert W. A. M. Kuijpers, René Höhn, Alireza Mirshahi, Anthony P. Khawaja, Robert N. Luben, Maja Gran Erke, Therese von Hanno, Omar Mahroo, Ruth Hogg, Christian Gieger, Audrey Cougnard-Grégoire, Eleftherios Anastasopoulos, Alain Bron, Jean-François Dartigues, Jean-François Korobelnik, Catherine Creuzot-Garcher, Fotis Topouzis, Cécile Delcourt, Jugnoo Rahi, Thomas Meitinger, Astrid Fletcher, Paul J. Foster, Norbert Pfeiffer, Caroline C. W. Klaver, and Christopher J. Hammond. Prevalence of refractive error in Europe: the European Eye Epidemiology (E(3)) Consortium. *European journal of epidemiology*, 30(4):305–315, 2015. DOI: 10.1007/s10654-015-0010-0.

[67] United States Code of Federal Regulations, 21 CFR §801.410. Use of impact-resistant lenses in eyeglasses and sunglasses, 2004. URL: https://www.ecfr.gov/current/title-21/chapter-I/subchapter-H/part-801/subpart-H/section-801.410. Accessed 10. July 2026.

[68] VISION EASE (HOYA Vision Care). Polycarbonate vs. Trivex, n.d. URL: https://www.hoyavision.com/globalassets/processing-tips/polycarbonatevstrivextechtips916.pdf. Accessed 26. July 2026.

[69] Pierre-Alexandre Blanche, Tianyao Zhang, and Craig Draper. Holographic curved waveguide combiner for AR/HUD with 2D pupil expansion. In *ODS 2022: Industrial Optical Devices and Systems*, volume 12231, page 1223108. SPIE, 2022. DOI: 10.1117/12.2632864.

[70] Tianyao Zhang, Craig Draper, Pierre Blanche, and Yushi Kaneda. 2D curved holographic waveguide combiner for augmented reality with pupil expansion. In *Optical Architectures for Displays and Sensing in Augmented, Virtual, and Mixed Reality (AR, VR, MR) IV*, volume 12449, page 124491Q. SPIE, 2023. DOI: 10.1117/12.2648597.

[71] Francesco Aieta, Ali Kabiri, Patrice Genevet, Nanfang Yu, Mikhail A. Kats, Zeno Gaburro, and Federico Capasso. Reflection and refraction of light from metasurfaces with phase discontinuities. *Journal of Nanophotonics*, 6: 063532, 2012. DOI: 10.1117/1.JNP.6.063532.

[72] Maliha Ashraf, Alexandre Chapiro, and Rafał K. Mantiuk. Resolution limit of the eye – how many pixels can we see? *Nature Communications*, 16(1): 9086, 2025. DOI: 10.1038/s41467-025-64679-2.

[73] Alasdair C Hamilton and Johannes Courtial. Metamaterials for light rays: ray optics without wave-optical analog in the ray-optics limit. *New Journal of Physics*, 11(1):013042, 2009. DOI: 10.1088/1367-2630/11/1/013042.

[74] Charles-Michel Marle. The works of william rowan hamilton in geometrical optics and the malus–dupin theorem. *Banach Center Publications*, 110: 177–191, 2016. DOI: 10.4064/bc110-0-12.

[75] N. Davidson, A. A. Friesem, and E. Hasman. Optical coordinate transformations. *Appl. Opt.*, 31(8):1067–1073, 1992. DOI: 10.1364/AO.31.001067.

[76] C. Paterson. Diffractive optical elements with spiral phase dislocations. *Journal of Modern Optics*, 41(4):757–765, 1994. DOI: 10.1080/09500349414550771.

[77] S. K. Case, P. R. Haugen, and O. J. Løkberg. Multifacet holographic optical elements for wave front transformations. *Appl. Opt.*, 20(15):2670–2675, 1981. DOI: 10.1364/AO.20.002670.

[78] A. Scharmann and J. Ziller. Influence of light polarization on the diffraction efficiency of TIR holograms. *Optics Communications*, 24(2):181–184, 1978. DOI: 10.1016/0030-4018(78)90114-1.

[79] Craig T. Draper, Colton M. Bigler, Micah S. Mann, Kalluri Sarma, and Pierre-Alexandre Blanche. Holographic waveguide head-up display with 2-D pupil expansion and longitudinal image magnification. *Appl. Opt.*, 58(5): A251–A257, 2019. DOI: 10.1364/AO.58.00A251.

[80] Steven K. Case and Volker Gerbig. Laser Beam Scanners Constructed From Volume Holograms. *Optical Engineering*, 19(5):195711, 1980. DOI: 10.1117/12.7972592.

[81] H. Bartelt and S. K. Case. Coordinate Transformations Via Multifacet Holographic Optical Elements. *Optical Engineering*, 22(4):224497, 1983. DOI: 10.1117/12.7973150.

[82] Max Monti. Robot for recording analog multifacet holograms. In *Practical Holography VII: Imaging and Materials*, volume 1914, pages 34 – 40. SPIE, 1993. DOI: 10.1117/12.155028.

[83] Max Monti. Large-area laser scanner with holographic detector optics for real-time recognition of cracks in road surfaces. *Optical Engineering*, 34(7): 2017 – 2023, 1995. DOI: 10.1117/12.204793.

[84] Hosung Jeon, Minwoo Jung, Moonsung Park, Daerak Heo, Bonggu Shim, and Joonku Hahn. Hologram printer for freeform holographic optical elements. In *Digital Holography and 3-D Imaging 2022*, page Tu2A.8. Optica Publishing Group, 2022. DOI: 10.1364/DH.2022.Tu2A.8.

[85] Hosung Jeon, Sungjin Lim, Youngjin Jeon, Woongjae Baek, Daerak Heo, Youngmin Kim, Hwi Kim, and Joonku Hahn. Holographic printing for

generating large-angle freeform holographic optical elements. *Optics Letters*, 47(2):257–260, 2022. DOI: 10.1364/OL.448480.

[86] Hosung Jeon, Youngmin Kim, and Joonku Hahn. Aberration-free warp projection on a horopter screen using freeform holographic optical elements. *Optics Express*, 31(5):7466–7479, 2023. DOI: 10.1364/OE.484960.

[87] Johannes Hofmann, Anna-Katharina Friedel, Reinhold Fiess, and Wilhelm Stork. Angle-compensated holographic wave front printing for the fabrication of holographic optical elements operating in the infrared. *Optical Engineering*, 59(10):102405, 2019. DOI: 10.1117/1.OE.59.10.102405.

[88] Tobias Wilm, Johannes Hofmann, Reinhold Fiess, Simone Höckh, and Wilhelm Stork. Immersion-based holographic wave front printer setup for volume holographic retinal projection elements. In *Holography: Advances and Modern Trends VII*, volume 11774, page 117740F. International Society for Optics and Photonics, SPIE, 2021. DOI: 10.1117/12.2589102.

[89] Marvin Feil, Tobias Wilm, Moritz Esslinger, Reinhold Fiess, and Wilhelm Stork. Limits of immersion-free recording of holographic waveguide displays. *Optics Letters*, 50(2):606–609, 2025. DOI: 10.1364/OL.545451.

[90] Suyeon Choi, Changwon Jang, Douglas Lanman, and Gordon Wetzstein. Synthetic aperture waveguide holography for compact mixed-reality displays with large étendue. *Nature Photonics*, 19(8):854–863, 2025. DOI: 10.1038/s41566-025-01718-w.

[91] Covestro AG. Bayfol HX120 Preliminary Data Sheet, n.d. URL: https://solutions.covestro.com/-/media/covestro/solution-center/products/datasheets/imported/bayfol/bayfol-hx120_en_86220288-21894048-21894045.pdf. Accessed 3. October 2025.

[92] Robert R. McLeod and Andrew N. Sias. Inverse design of two-beam holographic exposure systems. In *Photosensitive Materials and their Applications II*, volume PC12151, page PC1215105. International Society for Optics and Photonics, SPIE, 2022. DOI: 10.1117/12.2623937.

[93] Mathias Brandstötter, Arthur Angerer, and Michael Hofbaur. An analytical solution of the inverse kinematics problem of industrial serial manipulators with an ortho-parallel basis and a spherical wrist. In *Proceedings of the Austrian Robotics Workshop*, pages 7–11, 2014.

[94] Stäubli International AG. Stäubli TX2-60 Product Datasheet, n.d. URL: https://www.staubli.com/content/dam/robotics/products/robots/tx2/TX2-60-6-axis-product-data-sheet-EN.pdf. Accessed 28. May 2026.

[95] RoboDK Inc. RoboDK: Robot simulation and offline programming software, n.d. URL: https://robodk.com/. Accessed 28. May 2026.

[96] Stäubli International AG. Stäubli Robotics Suite, n.d. URL: https://www.staubli.com/us/en/robotics/products/robot-software/staubli-robotics-suite.html. Accessed 28. May 2026.

[97] Mitsuo Takeda, Hideki Ina, and Seiji Kobayashi. Fourier-transform method of fringe-pattern analysis for computer-based topography and interferometry. *Journal of the Optical Society of America*, 72(1):156–160, 1982. DOI: 10.1364/JOSA.72.000156.

[98] Daniel B. Ott, Ivan B. Divliansky, Marc A. SeGall, and Leonid B. Glebov. Stabilization system for holographic recording of volume Bragg gratings using a corner cube retroreflector. *Applied Optics*, 53(6):1039–1046, 2014. DOI: 10.1364/AO.53.001039.

[99] Takuma Kuno, Toshiteru Nakamura, Takahiro Matsuda, Shinsuke Onoe, Yoshiho Seo, and Satoshi Ouchi. Fiber-scanning based ultra-compact near-eye display with a narrow-beam waveguide. *Optical Review*, 28(6):716–722, 2021. DOI: 10.1007/s10043-021-00702-3.

[100] International Organization for Standardization. ISO 12233:2024 Digital cameras – Resolution and spatial frequency responses, 2024. URL: https://www.iso.org/standard/88626.html. Accessed 10. July 2026.

[101] Yueming Han, Chengzhe Chai, Rui Xiao, Haolei Mao, Xin Li, Lei Zhao, Chaohao Wang, and Jianyu Wang. K-vector microscope: high-resolution testing technique for freeform holographic optical elements. *Optics Express*, 33(17):35798–35814, 2025. DOI: 10.1364/OE.569045.

[102] Chengzhe Chai, Rui Xiao, Yueming Han, Shujin Zheng, Haolei Mao, Xin Li, and Lei Zhao. 49-1: Invited paper: Prism-free fabrication and high-spatial-resolution characterization of holographic optical waveguide. *SID Symposium Digest of Technical Papers*, 56(S1):416–419, 2025. DOI: 10.1002/sdtp.18824.

[103] Tina Sabel. Spatially resolved analysis of bragg selectivity. *Applied Sciences*, 5(4):1064–1075, 2015. DOI: 10.3390/app5041064.

[104] J Breitenbach and J.L.J Rosenfeld. Goniospectrometer measurements of the optical performance of a holographic optical element. *Solar Energy*, 68 (5):427–437, 2000. DOI: 10.1016/S0038-092X(00)00011-6.

[105] Wan-Pin Tsai, Yang-Kuan Tseng, Tsung-Xian Lee, and Ching-Cherng Sun. Volume holographic waveguide combiner design for AR glasses. In Naamah Argaman, Hong Hua, and Daniel K. Nikolov, editors, *Optical Architectures for Displays and Sensing in Augmented, Virtual, and Mixed Reality (AR, VR, MR) V*, volume PC12913, page PC1291303. SPIE, 2024. DOI: 10.1117/12.3002656.

[106] Gregory J. Steckman and Frank Havermeyer. High spatial resolution measurement of volume holographic gratings. In *Practical Holography XX: Materials and Applications*, volume 6136, page 613602. SPIE, 2006. DOI: 10.1117/12.646413.

[107] W. L. Bragg. The diffraction of short electromagnetic waves by a crystal. In *X-ray and Neutron Diffraction*, The Commonwealth and International Library: Selected Readings in Physics, pages 109–125. Pergamon, 1966. DOI: 10.1016/B978-0-08-011999-1.50015-8. Originally published in Proceedings of the Cambridge Philosophical Society, 17, 43–57 (1913).

[108] M. G. Moharam and T. K. Gaylord. Rigorous coupled-wave analysis of planar-grating diffraction. *J. Opt. Soc. Am.*, 71(7):811–818, 1981. DOI: 10.1364/JOSA.71.000811.

[109] Edward DeHoog, Jason Holmstedt, and Tin Aye. Field of view limitations in see-through HMD using geometric waveguides. *Applied Optics*, 55(22): 5924–5930, 2016. DOI: 10.1364/AO.55.005924.

[110] E. A. Igel and R. L. Hughes. Optical analysis of solar facility heliostats. *Solar Energy*, 22(3):283–295, 1979. DOI: 10.1016/0038-092X(79)90143-9.

[111] Stäubli International AG. VAL 3-Handbuch Version 9, D.280.935.02-H - 12/01/2022, 2022.

[112] G. Bricogne. Fourier transforms in crystallography: Theory, algorithms and applications. In *International Tables for Crystallography*, chapter 1.3, pages 25–98. International Union of Crystallography, 2006. DOI: 10.1107/97809553602060000551.