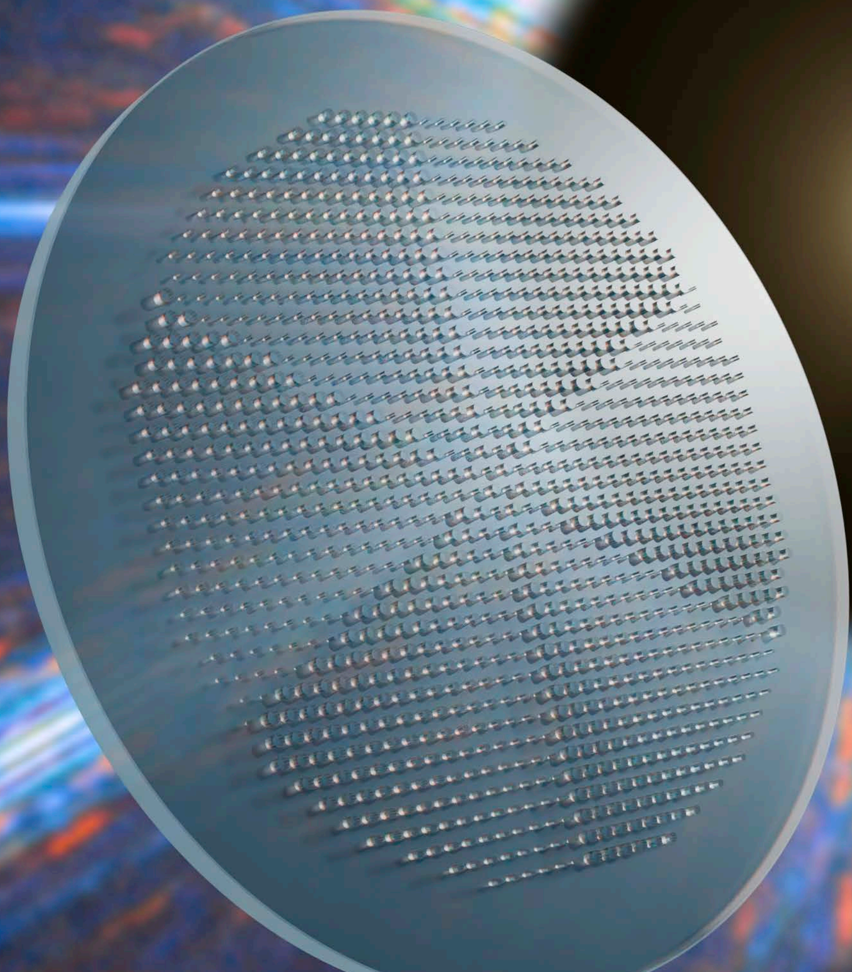
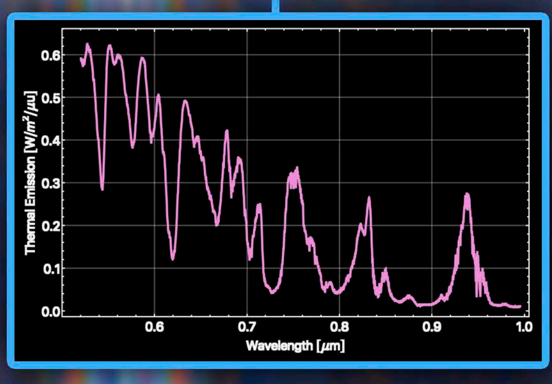
Metasurface Optics for High-Contrast Imaging

Design, Fabrication, and Implementation





Final Report of the
2024 KISS Study

Keck
INSTITUTE FOR SPACE STUDIES

METASURFACE OPTICS FOR HIGH-CONTRAST IMAGING

Date: June 2026

Study Workshop: September 23-27, 2024

Study Leads

J. Kent Wallace, Jet Propulsion Laboratory, California Institute of Technology

james.k.wallace@jpl.nasa.gov

Dimitri Mawet, California Institute of Technology

dmawet@astro.caltech.edu

Tobias Wenger, Jet Propulsion Laboratory, California Institute of Technology

wenger@jpl.nasa.gov

Study Report Prepared for the W. M. Keck Institute for Space Studies (KISS)

The research was carried out in part at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration (80NM0018D0004).

Pre-Decisional Information—For Planning and Discussion Purposes Only

DOI: 10.26206/z9096-t9d30

Recommended citation (long form)

J. K. Wallace, T. Wenger, D. Mawet, O. Absil, V. Chambouleyron, N. Desai, J. Fan, A. Faraon, M. Karlsson, L. Konig, L. Li, M. Millar-Blanchaer, T. Milster, S. Palatnick, J-S. Park, P. Pearson, E. Por, N. Rubin, M. Salama, J. Schuller, E. Serabyn, D. Shanks, Y. Shao, J. Uddin, G. West, F. Zamkotsian (Eds.) 2025. "Metasurface Optics for High-Contrast Imaging." Report prepared for the W. M. Keck Institute of Space Studies (KISS), California Institute of Technology.

Recommended citation (short form)

J. K. Wallace, T. Wenger, D. Mawet (Eds.). 2025. "Metasurface Optics for High-Contrast Imaging." Report prepared for the W. M. Keck Institute of Space Studies (KISS), California Institute of Technology.

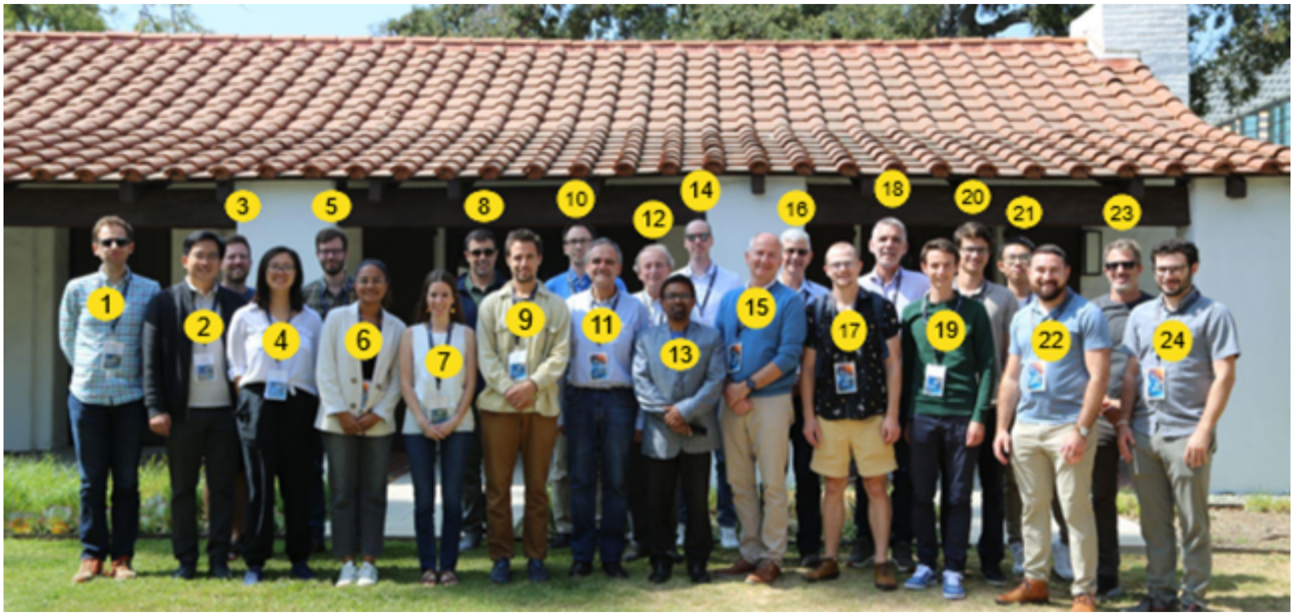
Acknowledgments

Director: Prof. Bethany Ehlmann (at time of KISS study), Prof. Mike Watkins (at time of report publication)

Executive Director: Harriet Brettle

Editing and Formatting: Marcy Harbut

Cover Image: Robert Hurt



Workshop Participants

1	Noah Rubin, University of California, San Diego	13	Jamal Uddin, Coppin State University
2	Joon-Suh Park, Harvard University	14	Olivier Absil, University of Liège
3	Max Millar-Blanchaer, University of California, Santa Barbara	15	J. Kent Wallace, Jet Propulsion Laboratory
4	Lisa Li, University of California, San Diego	16	Thomas Milster, University of Arizona
5	Emiel Por, Space Telescope Science Institute	17	Skyler Palatnick, University of California, Santa Barbara
6	Niyati Desai, Jet Propulsion Laboratory	18	Mikael Karlsson, Uppsala University
7	Maïssa Salama, University of California, Santa Cruz	19	Lorenzo König, Jet Propulsion Laboratory
8	Dimitri Mawet, California Institute of Technology/Jet Propulsion Laboratory	20	Phillippe Pearson, California Institute of Technology
9	Vincent Chambouleyron, University of California, Santa Cruz	21	Yixuan Shao, Stanford University
10	Tobias Wenger, Jet Propulsion Laboratory	22	Garrett West, BAE Systems
11	Frederic Zamkotsian, Laboratoire d'Astrophysique de Marseille	23	Jon Schuller, University of California, Santa Barbara
12	Eugene (Gene) Serabyn, Jet Propulsion Laboratory	24	Daniel Shanks, Jet Propulsion Laboratory

Not pictured: Andrei Faraon, California Institute of Technology; Jonathan Fan, Stanford University

Contents

1 Executive Summary

Pathways to Discovery in Astronomy and Astrophysics for the 2020s (National Academies of Sciences, Engineering, and Medicine, 2021), referred to as Astro2020 or Decadal Survey in this report, underscored the importance of developing a flagship mission to directly image and characterize Earth-like exoplanets orbiting Sun-like stars. This vision has materialized into NASA's Habitable Worlds Observatory (HWO), a mission specifically engineered to identify habitable, Earth-like planets around relatively bright stars like our Sun, utilizing advanced high-contrast imaging technologies: namely coronagraphy and wavefront sensing and control. Achieving HWO's objectives necessitates significant advancements in these starlight suppression technologies. While traditional optical technologies have seen continuous improvements, exploring innovative approaches such as metasurfaces is vital to overcoming existing limitations and expanding design possibilities.

Metasurfaces are engineered two-dimensional structures composed of subwavelength nanostructures that can precisely manipulate the phase, amplitude, and polarization of incident light. Their unique capabilities position them as potential game-changers in high-contrast astronomical imaging.

From September 23 to 27, 2024, the W.M. Keck Institute for Space Studies (KISS) hosted a workshop titled "Metasurface Optics for High-Contrast Imaging," bringing together experts from metasurface technology, astronomy, high-contrast imaging, and micro-nano fabrication. This interdisciplinary gathering included participants from academia, industry, and government agencies, all focused on exploring the integration of metasurface optics into high-contrast imaging systems for exoplanet detection and characterization.

The workshop highlighted several key areas where metasurfaces could significantly impact high-contrast imaging missions:

1. **Coronagraphic Masks:** Metasurfaces offer the potential to create compact and efficient coronagraphs capable of achieving the high-contrast ratios (~10^{-10}) required for imaging Earth-like exoplanets. Their ability to tailor wavefronts at a subwavelength scale allows for precise control over light propagation, which is crucial for effective starlight suppression.
2. **Wavefront Sensing Optics:** The precise manipulation capabilities of metasurfaces synthesize complex optical functions at the core of wavefront sensors and thus enhance wavefront sensing accuracy, leading to improved adaptive optics systems. This enhancement is essential for measuring and correcting optical distortions and

achieving the exquisite wavefront quality necessary for detecting faint exoplanetary signals.

3. **Polarization Aberrations Compensation:** Metasurfaces can be engineered to manage polarization effects effectively, reducing polarization-induced aberrations that can degrade image contrast. This capability is vital for maintaining the integrity of high-contrast imaging systems.

Discussions during the workshop also addressed the challenges associated with metasurface fabrication, particularly in scaling up production while maintaining the precision required for astronomical applications. Advances in nanofabrication techniques, such as innovative etching and lithographic processes, were identified as critical to producing high-quality metasurfaces suitable for space missions.

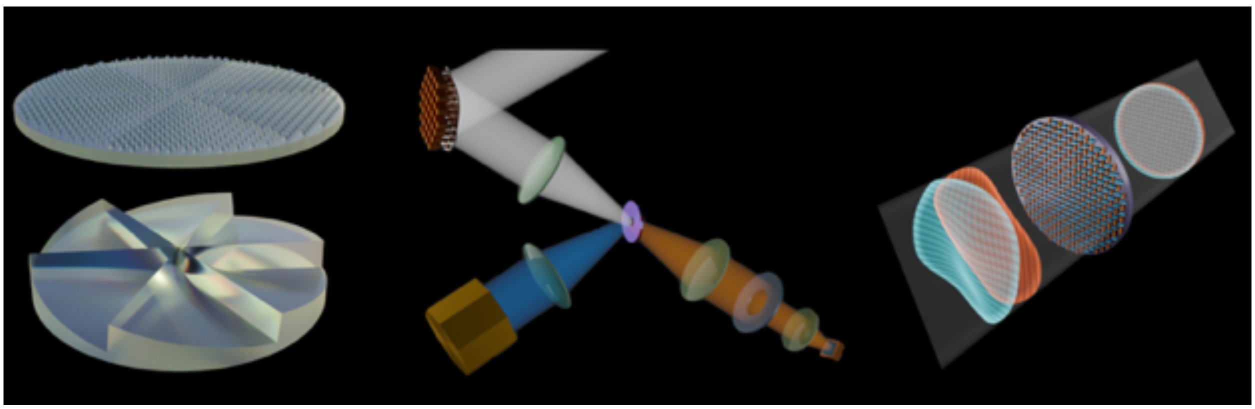


Figure 1.1. This is an illustration of the main applications of metasurfaces for high-contrast imaging that is described in more detail in Section 4 of this report. (Left) Coronagraphic masks; (Center) Wavefront Sensing optics; (Right) Polarization Aberrations Compensation. Image credit: Keck Institute for Space Studies/Robert Hurt (Caltech/IPAC-SELab)

The participants emphasized the necessity for ongoing interdisciplinary collaboration to transition metasurface technology from laboratory research to practical implementation in space-based observatories. Recommendations included increased investment in metasurface research and development, focusing on:

1. **Broadening Operational Bandwidth:** Developing metasurfaces that function effectively across a wide range of wavelengths to accommodate HWO's observational requirements.

2. **Enhancing Environmental Robustness:** Ensuring metasurfaces can withstand the harsh conditions of space, including radiation exposure and temperature fluctuations.

3. **Integrating into System-Level Architectures:** Designing metasurfaces that can be seamlessly incorporated into existing and future high-contrast imaging systems, considering factors like compatibility with other optical components and overall system performance.

In conclusion, metasurfaces present a promising avenue for advancing high-contrast imaging capabilities, potentially enabling the direct observation and characterization of Earth-like exoplanets. Continued research, development, and collaboration are essential to fully realize the benefits of metasurfaces in future space- and ground-based observatories.

2 Introduction

There are times when one senses that a combination of two separate things will work very well if they are brought together. Such a realization is an “a-ha” moment, and when the outcome of the union exceeds expectations, it’s incredible. The combination of metasurface optics and high-contrast imaging happens to be such a special occasion, and it was the genesis of the KISS workshop.

Our work in metasurface optics started modestly enough; we wanted to make a slight change to an existing phase-contrast wavefront sensor. We wanted phase control for two different polarization states simultaneously, and metasurfaces offered a solution. It took a while to mature the design and fabrication process for this metasurface optical element, but within a year we had a well-developed and repeatable method for fabrication. Shortly afterwards, we had this new metasurface device working on-sky and producing impressive results. Then it dawned on us that this metasurface technology could truly shine in the arena of exoplanet detection. Indeed, the technique of starlight suppression to reveal faint companions demands exquisite control of the electric field (amplitude, phase and polarization) at microscopic scales. From our experience with applying metasurfaces to phase-contrast, we knew this was a good match for high-contrast imaging. However, we didn’t know enough detailed information about either topic to know the best way to proceed. We needed to gather a group of the best people working on the leading edge of both fields to help us learn what might be possible to determine the best way to proceed. We therefore proposed a KISS workshop, and this report summarizes its findings.

This workshop aimed to convene a diverse group of researchers, including end users, designers, fabricators, and metrology experts, to explore and advance the application of metasurface optics in the field of astronomy and high-resolution imaging.

The workshop objective was to engage with exoplanet scientists and metasurface researchers and engineers to explore the application of this technology to create solutions for the challenges of direct imaging and spectroscopy of exoplanets. The workshop was able to foster this nascent community through a focused week of technical and social interaction that included:

1. Knowledge exchange and collaboration
2. Understanding metasurface optics in astronomy
3. Showcasing applications and case studies
4. Identifying challenges and exploring solutions

5. Networking and future collaborations

The goal of this workshop was to catalyze advancements in the field of metasurface optics specifically tailored for exoplanet detection and high-contrast imaging, fostering a collaborative environment that drives innovation, knowledge dissemination, and the translation of research findings into practical applications within the realm of astronomy and imaging technology.

The report provides an overview of the current state of exoplanet discovery, introduces metasurface optics, fabrication techniques, characterization, and presents applications within high-contrast imaging. Future opportunities are identified, including concrete recommendations for further work in this important area of research.

3 Exoplanet Discovery: Current State of the Field

3.1 Science motivation: Direct imaging of Earth-like planets

Over the past 30 years, astronomers have discovered an astonishing variety of worlds beyond our solar system. To date, over 6,000 exoplanets have been detected, primarily through indirect methods such as the radial velocity, astrometric, and transit techniques. While these methods have proven highly effective, they have limited remote-sensing capabilities, particularly for stars like our Sun, where Earth-like planets are most likely to be found.

Direct, high-contrast imaging, where photons reflected off a planet's surface or atmosphere are detected, offers our best opportunity to characterize planets similar to Earth. However, imaging such planets around other stars is exceptionally challenging due to the small angular separation and extreme brightness contrast involved. For instance, an Earth-like planet orbiting a Sun-like star at a distance of 10 parsecs would appear at an angular separation of only 0.1 arcseconds and have a flux ratio of 10^{-10} compared to its host star. This makes the imaging and characterization of Earth-like exoplanets in reflected light one of the most formidable technical challenges in modern astronomy.

Over the past two decades, several advanced techniques, inspired by Lyot's pioneering invention of the coronagraph, have been developed to tackle this challenge. However, much work remains to be done to demonstrate our ability to control light diffracted by telescopes and scattered by imperfect optics to the level necessary to achieve this ambitious goal.

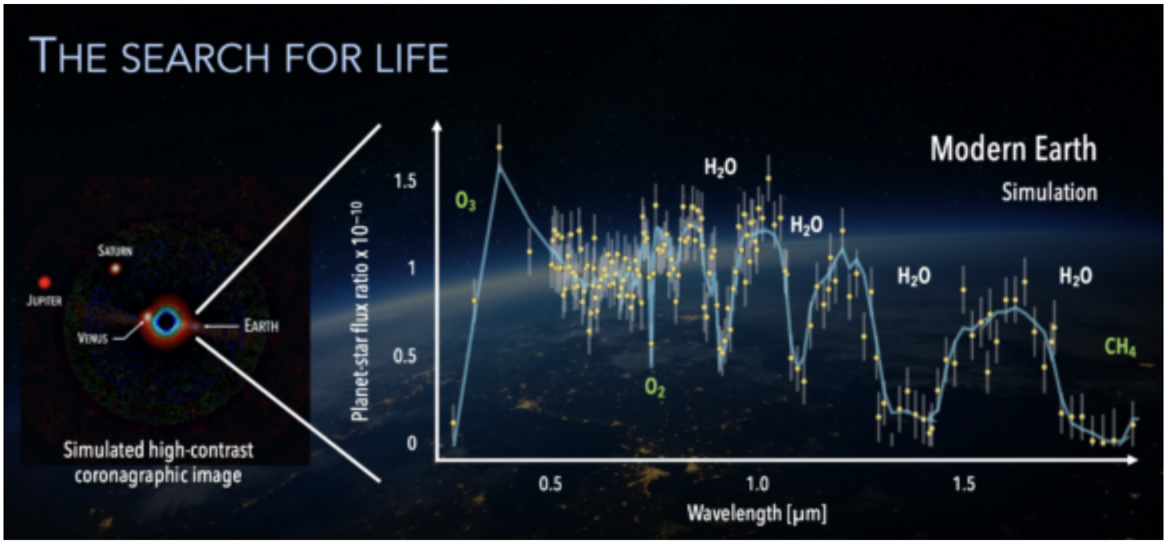


Figure 3.1. Key components of exoplanet science are: a) starlight suppression to enable detection of exoplanet companions, and b) spectroscopy of exoplanet light to look for telltale absorption features. Reproduced from Spohn, T., et al. Exo-Geoscience Perspectives Beyond Habitability. Space Sci Rev 222, 9 (2026).

3.2 Opportunities: HWO, ELTs

The Astro2020 Decadal Survey articulated the impact of advancing high-contrast imaging technologies for the detection and characterization of habitable worlds in the coming decade: "...exploring terrestrial planets outside our Solar System through direct imaging and spectroscopy will advance one of humanity's greatest quests—the search for habitable environments and life outside of the Solar System. This transformative goal is at the scientific forefront, connecting astronomy, astrobiology, planetary and Earth Science, and is one that captures the imagination of all humankind." The Habitable Worlds

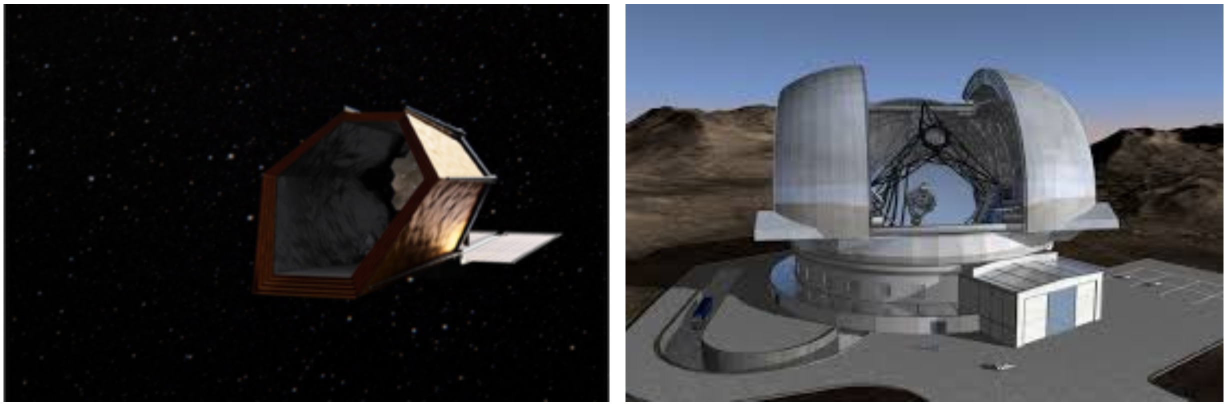


Figure 3.2. The ability to detect exoplanets is limited by the same effect for both ground- and space-based observatories: quasi-static speckles. Therefore, even though the facilities have vastly different operating conditions, the observatory system's stability—and the associated technologies to achieve this—can be shared between the two. Image credit: (Left) NASA's Goddard Space Flight Center Conceptual Image Lab (Right) ESO

Observatory (HWO) is the key resource for meeting the ambitious goal of directly imaging and characterizing planets in the habitable zones of the nearest Sun-like stars.

3.3 Current state of the art

The field of high-contrast imaging has historically been dominated by ground-based observatories, where the most ground-breaking discoveries have been obtained. The largest ground-based optical telescopes, aided by extreme adaptive optics systems and dedicated coronagraphic instruments can nowadays reach contrasts down to a few 10^{-7} at a few arcseconds from sufficiently bright stars, with the contrast performance decreasing for smaller angular separations, typically from a few 10^{-6} at 1 arcsecond to a few 10^{-4} close to the coronagraph's inner working angle. In terms of space-based observatories, coronagraphs on NASA's Hubble and James Webb Space Telescopes have established the current benchmark for high-contrast imaging from space, although they do not employ high-order wavefront control for achieving deeper contrast. Hubble's coronagraph, a part of the Space Telescope Imaging Spectrograph (STIS), has demonstrated a contrast level of 3×10^{-5} at an angular separation of 0.25 arcseconds (approximately 5 λ/D). The JWST NIRCam commissioning results have shown a 5σ detection limit of 1×10^{-5} flux ratios at an angular separation of 0.5 arcseconds (5 λ/D) using the NIRCam bar mask at 3.35 microns. NASA's Roman Space Telescope's Coronagraph Instrument (CGI) is expected to establish a new benchmark for space-based high-contrast imaging after its launch in the mid-2020s.

The common traits of these high-contrast instruments are the need for exquisite wavefront control, and for custom light-suppressing devices, known as coronagraphs.

Both concepts come in many flavors, with various maturity levels and performance potential. One of the main challenges here is to provide wavefront sensing as close as possible to the coronagraphically active element(s) in the optical train. Besides pure imaging, high-contrast imaging instruments are now also aided by integral-field spectrographs, which not only provide an additional boost in detection performance, but also more advanced characterization potential for the detected planets. Finally, these instruments are necessarily paired with custom observing strategies and post-processing algorithms, which can further enhance the dynamic range, but cannot go beyond the fundamental limits set by the photon noise of residual stellar light at the location of the planet. Post-processing algorithms have demonstrated an approximately tenfold improvement, starting from a raw contrast of 10^{-8}, as evidenced by the Roman Space Telescope Exoplanet Imaging Community Data Challenge.

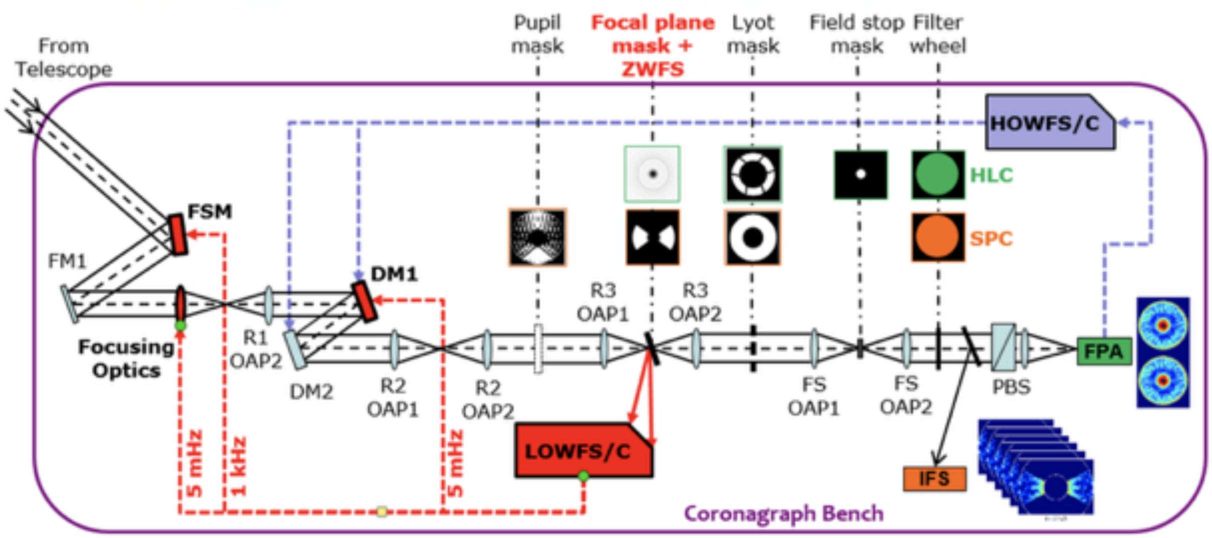


Figure 3.3. This functional block diagram of the high-contrast system illustrates where metasurface and starlight suppression technologies can be applied to exoplanet observing systems. Reproduced from F. Shi, et al., WFIRST Coronagraph Milestone 6 Final Report: Low Order Wavefront Sensing and Control.

In laboratory settings, the deepest in-vacuum contrasts achieved over a 10% bandwidth were 6×10^{-10} over a half-annular dark region from 3–15 λ/D with a Hybrid Lyot Coronagraph at 800 nm, and 4×10^{-10} over a full annulus of 3–8 λ/D with a Lyot Coronagraph at 550 nm. These results were obtained in a static vacuum environment with a clear pupil, representative of a monolithic telescope, in NASA's Exoplanet Exploration Program's High Contrast Imaging Testbed (HCIT). Coronagraph tests with a segmented pupil have been conducted in-air at the High-contrast imager for Complex Aperture Telescopes (HiCAT) facility at the Space Telescope Science Institute. Using a hexagonally segmented deformable mirror to simulate an off-axis segmented telescope pupil, the team has achieved a contrast of 2×10^{-8} in monochromatic 640 nm light from 4.6–12 λ/D with a Lyot Coronagraph in a static environment, and 2×10^{-8} from 2–12 λ/D with a phase-apodized-pupil Lyot coronagraph (PAPLC) with active low-order wavefront control. Additionally, a demonstration of a vortex coronagraph in the HCIT using a static, hexagonally gridded pupil mask showed the capability of using a deformable mirror to suppress diffracted light from the segmentation pattern.

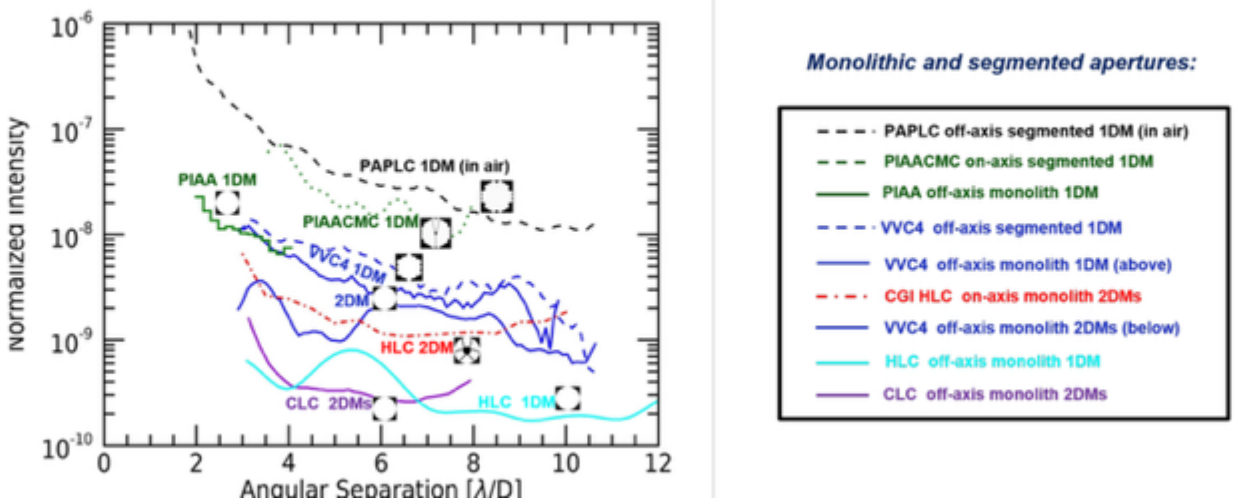


Figure 3. 4. Starlight rejection performance for different coronagraphs are compared in this figure. These results are in vacuum (unless otherwise noted). Metasurface optics can be used for most of these coronagraph architectures. Reproduced from Rus Belikov, et al., "Coronagraph design survey for future exoplanet direct imaging space missions," Proc. SPIE 13092, Space Telescopes and Instrumentation 2024: Optical, Infrared, and Millimeter Wave, 1309266 (23 Aug 2024). Adapted from the original figure by Mennesson, et al., "Current laboratory performance of starlight suppression systems, and potential pathways to desired Habitable Worlds Observatory exoplanet science capabilities", JATIS, (2024).

3.4 Challenges & requirements

Directly imaging an Earth-like planet around a Sun-like star requires reaching 10^{-10} raw contrast levels. A key technical barrier to reaching such deep contrast ratios is the design and fabrication of coronagraphs capable of suppressing starlight to these levels. In addition, for the coronagraph to operate at such levels, a wavefront error and stability on the order of tens of picometers over tens of minutes must be maintained. This will require wavefront sensors capable of sensing to these levels as well as minimal non-common path aberrations (NCPAs) between the wavefront sensor and the coronagraph.

Polarization aberrations arise naturally in optical systems containing reflective surfaces operating at non-normal incidence. Because the phase and amplitude response of optical coatings differ for orthogonal polarization states, reflections from telescope mirrors introduce polarization-dependent wavefront errors that cannot be fully corrected using conventional scalar wavefront control techniques. For future high-contrast imaging systems targeting raw contrasts of 10^{-10}, polarization aberrations become a fundamental limitation rather than a secondary effect. Even when the scalar wavefront is corrected to picometer-level precision, residual polarization-dependent phase and amplitude errors generate leakage terms that can limit achievable contrast and reduce sensitivity to Earth-like exoplanets. Consequently, polarization management must be considered as an integral part of the observatory architecture. Technologies capable of measuring, controlling, or compensating polarization aberrations—including metasurface-based corrective optics—have the potential to significantly relax system-level design constraints while improving overall coronagraph performance.

The inner working angle (IWA) of the coronagraph defines the smallest angular separation at which a planet can be detected relative to its host star. For HWO, achieving an IWA of approximately 2.5 λ/D is highly desirable to maximize access to habitable-zone planets around nearby Sun-like stars. Smaller IWAs increase the number of target systems accessible to the observatory and improve scientific yield. However, reducing the IWA generally increases sensitivity to low-order aberrations, pointing errors, and polarization effects. Coronagraph architectures capable of simultaneously providing high-throughput broadband performance and small inner-working angles therefore represent a major technology development priority.

Any technology developed that meets all the above criteria will also have to be robust enough to survive the observatory launch into space, as well as space conditions once in orbit. The overall requirements are listed in Table 3-1.

Table 3.1. Representative high-level performance requirements for future space-based observatories designed to directly image and characterize Earth-like exoplanets. Values are approximate and intended to illustrate the scale of the technical challenge facing next-generation high-contrast imaging systems such as the Habitable Worlds Observatory.

Requirement	Value	Note
Raw contrast	10^{-10} in visible 10^{-9} in near-IR	
Wavefront error & stability	~10 picometers	Dual polarization wavefront control

Requirement	Value	Note
IWA	2.5 λ/D	
Instantaneous bandwidth	10–20%	
Wavelength coverage	250 nm–1800 nm	UV-Vis-NIR
Spectral resolution	R=70–10,000	
Optical efficiency/component	>90%	TBC, subject to trade-off
Robustness	Resist to launch and radiation	TBD

3.5 The need for metasurfaces

While high-contrast imaging instruments have been reaching up to 10^{-8} in raw contrast, the remaining gap to reach 10^{-10} can hardly be solved based on current technologies without making significant trade-offs. For instance, a classic Lyot coronagraph will not provide the small inner-working angle and high throughput that one would like to reach. The current generation of vortex coronagraphs, on the other hand, is generally based on a vectorial implementation, where the two orthogonal polarizations see different signs for the helical vortex phase ramp. This, in turn, prevents reaching 10^{-10} contrast in the two polarizations at once. The current manufacturing technologies for these vortex phase masks also give a limited control on phase and amplitude at the smallest scales, close to the optical singularity at the center of the vortex.

Non-common path aberrations (NCPAs) arise when the light to the science camera picks up aberrations not seen by the wavefront sensor. Several techniques have been developed to sense these NCPAs (phase diversity/phase retrieval, speckle nulling, electric field conjugation (EFC), pairwise probing, and Zernike wavefront sensing) but the most efficient solution is to do the wavefront sensing at the same location as the starlight suppression.

A particularly important example is the impact of polarization aberrations on coronagraph performance. Unlike conventional scalar wavefront errors, polarization aberrations produce different electric field distributions for orthogonal polarization states. These differences generate residual stellar leakage that cannot be completely removed using a single deformable mirror control solution. Modeling studies for future exoplanet observatories have shown that polarization effects can become a dominant contrast limitation once other error sources are reduced to picometer levels. As a result, technologies capable of directly controlling polarization-dependent phase and amplitude errors may provide substantial gains in achievable contrast and scientific yield.

Table 3.2. Qualitative comparison of conventional dielectric optics, liquid-crystal polymer (LCP) devices, and scalar and vector metasurface implementations for high-contrast imaging applications. Relative strengths and limitations are indicated for key performance categories relevant to future exoplanet observatories.

	Dielectric (bulk optics)	LCP	Metasurface scalar	Metasurface vector
Chromatism	--	+	+	++
Polarization sensitivity	++	-	+	--
Polarization control	--		--	++
Spatial freq. control	-		++	++
Space compatibility	++		+	+
Quality/defects	++		+	+
Large size	++		-	-
Ease of fabrication	++		-	-

Metasurfaces have the potential to contribute to several areas of astronomy related to high-contrast imaging, due to its phase and amplitude control at the micron-scale level, its potential for achromatization, and for its polarization control as well as its ability to be polarization insensitive. Figure 1.1 summarizes the main areas where we see the most potential impact of metasurfaces for high-contrast imaging and which are explained in more detail in Section 8 ("Metasurface Applications within High-Contrast Imaging").

1. **Phase and amplitude control at the micron scale.** Metasurfaces provide spatial control of optical phase and amplitude at dimensions approaching the wavelength of

light, enabling optical functions that are difficult or impossible to realize using conventional manufacturing approaches.

2. **Potential for achromatization.** Through dispersion engineering and advanced meta-atom design, metasurfaces can achieve optical responses that remain effective across broader wavelength ranges than traditional diffractive optics.
3. **Polarization control and polarization-insensitive operation.** Metasurfaces can be designed either to manipulate polarization with high precision or, alternatively, to minimize polarization sensitivity altogether, providing flexibility not available in many conventional optical technologies.

4 Introduction to Metasurface Optics

4.1 Historical primer

Since 2011 (when what is regarded as the initial paper in the field of "metasurfaces" was published) (Yu 2011), the field has exploded, with likely thousands of distinct papers on metasurface devices published each year (Genevet 2017). The range of devices and applications under the heading of metasurfaces has become so diverse that it can occasionally be difficult to discern what is meant by the term in the first place. While "metasurface" is a term broadly applied to a large variety of nanoscale optics (in both guided wave and free-space scenarios), we restrict our definition in this report to devices intended to passively enact an engineered phase, amplitude, and/or polarization profile on a wavefront in free space using structures that are subwavelength in dimension.

It should be said that metasurfaces have a long history, and even those with the definition above are not new, strictly speaking. The ideas underlying metasurface devices are as old as diffraction theory itself. Particularly notable are earlier developments in microwave and radar in the areas of phased arrays and reflect arrays that use extremely similar concepts to contemporary metasurface research. The modern field of metasurfaces can be traced to the genesis of optical holography in the 1960s, and particularly to the development of so-called "computer-generated holography" (CGH) which represent some of the first attempts to synthesize complex wavefronts with sophisticated, human- engineered optics and objects. During the 20th century, advances in lithographic and etching technology owed to the semiconductor industry led directly to the field of *diffractive optics*. Throughout the 1970s, 80s, and particularly the 1990s, extensive efforts were undertaken to realize wavefront-shaping optical elements using both continuous surface relief and multi-level profiles that which involve successive lithographic and etching steps. Particularly notable is work from the 1990s on so-called "blazed binary subwavelength gratings" in France, which directly anticipated the constant-height, modal phase dispersion method of phase imposition employed by current generation "metasurfaces."

Metasurfaces represent a convergence of multiple academic and industrial fields in the early to mid-2010s. In particular, the "meta" prefix is borrowed from the field of metamaterials, an academic field which itself rose to prominence in the early 2000s. The field of metamaterials sought to create artificial dielectric materials with arbitrary control over the constitutive parameters and with three-dimensionally structured media. Numerous theoretical and small-scale experimental studies (primarily at radio and microwave frequencies) showed that if metamaterials could be created with certain exotic material properties (including negative refractive index), several basic laws and limitations of optics could be revisited and perhaps surmounted with new technologies. Popular examples included “perfect” lenses not subject to the traditional diffraction limit, as well as various schemes for optical cloaking of objects. A highly related field known as “plasmonics” was popular in the same time period—plasmonics refers to surface plasmons, guided wave excitations that exist at air and metal interfaces, which were widely touted for a variety of uses in sensing and optics. Despite significant research, both fields struggled to deliver on their initial motivating promise. A variety of issues were at play, but in both cases, material losses and difficulty of fabrication presented substantial challenges (in the case of plasmonics, due to a reliance on metals that are inherently lossy).

Around the year 2010, then, the community in these fields was continually searching for ideas that could be more practical, or devices that offered a more concrete path to real technological impact. It was into this environment the initial 2011 paper appeared. In the 2011 paper, the authors claimed to develop a “generalized law of reflection and refraction”—by patterning a thin layer of gold nanostructures at the interface between silicon and air, the gold nanostructures could shift the phase of incoming light in such a way that it would change direction relative to its predicted shift from Snell’s Law (for mid-infrared wavelengths). This layer of gold nanostructures was billed as a two-dimensional metamaterial (the word “metasurface” was never mentioned in the manuscript). However, perhaps unbeknownst to the authors, the device represented a blazed grating, a well-known device that imparts a linear phase shift on the light and shifts its propagation direction. Nevertheless, this initial paper triggered a flurry of research into what became known as the metasurface field. Soon, intense interest focused on a variety of devices, including metalenses and in general holograms. Suffice it to say, this represented a rediscovery and reinvention of the more established field of diffractive optics by a new community (although in many cases, these researchers were not aware of the connection). Frequently, older ideas from diffractive optics and holography have seen reimplementation more recently under the guise of metasurface research.

However, several new areas of research have emerged under the metasurfaces heading that were not as thoroughly considered with diffractive optical elements of the past. Among these are polarization control, so-called “dispersion engineering,” and more advanced fabrication in new materials for operation in new wavelength regimes. Above all, what the “metasurface moment” in optics research today represents is the accumulation of more people thinking about diffractive optics, with access to more advanced tools, and with more allocated research funding, than ever before.

4.2 Central concepts of metasurface diffractive optics

A general metasurface design/modeling task can be broken into five distinct parts. We describe these as follows:

1. **The illumination:** The light intended for manipulation by a metasurface device. Very commonly, but not exclusively, this is assumed to be a plane wave at a fixed angle, with a fixed polarization state. We label this as $E_{in}(x,y)$.

2. **The metasurface:** The device itself, which manipulates the phase and/or amplitude of the wavefront (often phase-only, as described below). Often, though not exclusively, this is assumed to be fixed in height; this is often a defining feature of “metasurface optics” versus past generations of surface-relief diffractive optical elements.

3. **The “near field”:** Here, near field refers to the field just above the metasurface, after having been manipulated by the metasurface structures themselves (which is not identical to, though not entirely separate from, the term “near field” in optics which often refers to non-propagating evanescent waves just above an interface). We label this as $E_{out}(x,y)$.

4. **The propagator:** The means by which the near field is mapped to a plane located some macroscopic distance away from the metasurface itself. Often, this is as simple as the Fraunhofer propagator, i.e., a Fourier transform.

5. **The “far field”:** Here, we mean the electric field in some plane located an appreciable (macroscopic) distance from the metasurface itself. Often the metasurface itself is designed to enact some behavior in this region, in accordance with its desired role in a larger optical system.

4.2.1. Local approximations in metasurface optics

In general, the picture above requires the solution of Maxwell’s equations over macroscopic regions—in the most general scenario, change of one parameter of the metasurface over a possibly large (centimeter to inch) aperture would require re-simulation of the entire device. This tends to be computationally intractable and unsuitable for design. To simplify the design, modeling, and understanding of metasurfaces, most literature in this domain relies on two foundational assumptions, which we refer to as “local” approximations.

Local approximation #1: The metasurface acts locally on the field. In most metasurface works, it is assumed the metasurface device enacts a transfer function (whether in amplitude, phase, polarization, or all of these) that is strictly local. That is, the metasurface does not move light energy from one point to another, so the effect of the metasurface itself can be written as $E_{out}(x,y) = t(x,y)\, E_{in}(x,y)$— the entire effect of the metasurface is specific to a given point in space (x,y). Another way of saying this is that the metasurface itself is "thin." In this way, the metasurface's action itself can be described by a simple spatially varying phase or amplitude profile as a thin transmission mask—a drastic simplification of what is in general a complicated scattering process.

Local approximation #2: The metasurface is locally periodic. Almost all work on metasurface devices makes use of the so-called "locally periodic approximation." In other words, it is assumed the behavior of each individual structure of the metasurface is independent of that of its neighbors, even its nearest neighbors. This means that in modeling the metasurface, it is sufficient to ascribe to each structure a given amplitude and/or phase transmission that is derived from a simulation of that structure under periodic boundary conditions. This greatly simplifies the process of designing and modeling a metasurface; if, according to local approximation #1, some amplitude and/or phase transmission property is desired, it is sufficient to simulate a *library* of components (using the locally periodic condition under periodic boundary conditions to greatly simplify the simulations) and select components to satisfy the desired phase and/or amplitude at each point across the aperture of the metasurface.

There are many ways of looking at both local approximations stated above, but what these have in common is they are indirect ways of stating the paraxial limit, i.e., these methods of metasurface design are strictly correct only when viewed from small angles in the far field.

Why subwavelength? In this report, and in the literature in general, the subwavelength aspect of metasurface optics is often emphasized. The important point here is that in diffractive optics, if the structures manipulating the wavefront are themselves subwavelength in size and placed at a subwavelength pitch, the first order of diffraction from the inter-structure separation will be at an angle > 90 degrees and, in effect, will be an evanescent diffraction order that cannot propagate to the far field. Subwavelength sampling, then, allows the zero order's bandwidth to extend over the entire half-space of angles above the metasurface. In other words, it avoids diffraction from the individual meta-atoms themselves; light diffracts only from the controlled, supra-wavelength patterns formed by these meta-atoms when assembled into a metasurface.

In earlier work in diffractive optics, especially as applied to the coronagraphy problem that primarily interests us here, such subwavelength diffractive optics were often referred to as "zero order gratings."

Methods of imparting phase with metasurface optical components: In this sub-section, we will discuss methods of controlling the phase of an incoming wavefront with metasurface optical components. First, it merits a brief discussion as to why metasurface work almost always refers to phase control (rather than amplitude control, or mixed phase/amplitude modulation). Phase control is lossless, whereas amplitude control (unless one has the privilege to use gain) is inherently lossy. Moreover, when coupled with propagation (wave interference), phase-only control can implement mixed phase/amplitude control in the far-field. Therefore, it is almost always preferable to control the phase of a wavefront, since with phase-only control, full control over the wavefront's phase and amplitude are possible when combined with propagation without absorbing any of the incoming energy.

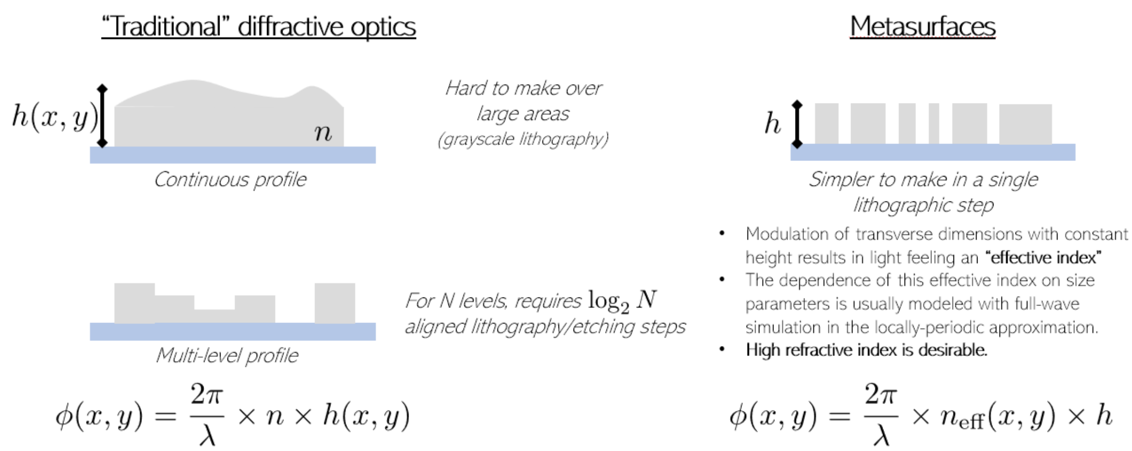


Figure 4.1. This image illustrates the relationship between traditional diffraction optics and metasurface optics. Metasurface optics use a fixed-height nanopillar with a custom-engineered spatial array to create an effective index. Image credit: Noah Rubin

1. Propagation phase: The first, perhaps the most obvious, and the most widespread method of imparting phase with metasurface optics is simply by propagation. This has always been the case in diffractive optics. However, this can take on many different forms. The clearest and most intuitively straightforward way of using propagation to enact a space-variant phase profile is with diffractive optical elements having variable height with (x, y). These are known as surface relief structures, and this was one of the most common approaches of imparting phase in diffractive optics. There are ways to make an arbitrary surface profile (e.g., grayscale lithography), however, this tends to be quite challenging. Multi-level lithographic approaches, in which only a discrete number of heights are accessible, fabricated by a combination of successive lithography and etching steps, have been popular in diffractive optics for decades. Metasurfaces tend to use a third approach altogether - what might be called "effective index" variation. Most metasurfaces use structures of uniform height. By varying their transverse dimensions, one can locally tailor the effective refractive index experienced by the light. This effective index—ranging between that of the constituent material and air—is determined by the spatial confinement of the optical mode within each structure. Being that all variations here are only transverse, these types of structures are dramatically easier to fabricate, with all fabrication typically accomplished in a single step of lithography and etching. As noted previously, however, while thought of as characteristic of metasurface optics, this approach was also pursued in traditional diffractive optics of decades past. It should be noted that all propagation-based phase approaches are necessarily chromatic, since these scale with k1/.

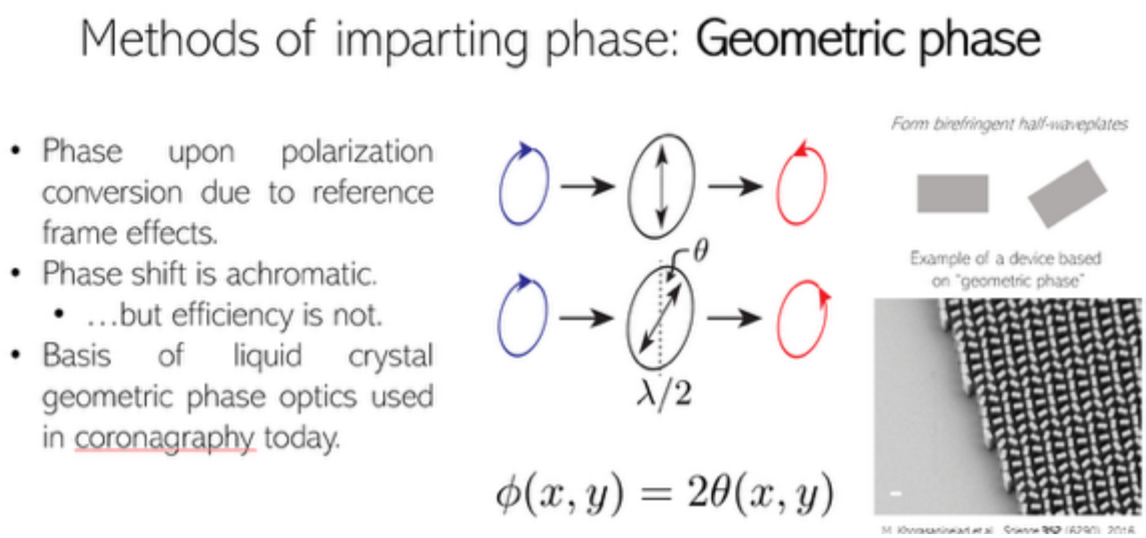


Figure 4.2. Phase can be imparted to the electric field by properly orienting the nanopillars. This is called the geometric phase, and the result is an achromatic phase change. Image credit: Noah Rubin

2. Geometric phase*:* The second most common method of imparting phase with metasurface optical components is based on the polarization of light. This technique uses the "geometric phase," which is a very general concept of physics that goes far beyond diffractive optics. The geometric phase refers to a phase that comes about from a change along a closed path in a higher-dimensional state space—the space of polarization states, in this case. In the most common example, we consider circularly polarized light incident on half-waveplates whose fast axis is oriented at an angle θ. When, for instance, right-circularly polarized light is incident on this half-waveplate, it will convert to left-hand circularly polarized light *for any orientation of the half-waveplate*. However, in the process, the light will pick up a global phase shift that depends on the orientation of the waveplate, in particular, φ = 2 θ. In this way, even though the light has "propagated" through an equal thickness piece of material, it acquires a phase that depends on the geometric parameters of the structure's orientation. This phase is often employed throughout diffractive optics, especially with liquid crystal diffractive optics. Since the phase relies on the orientation of a structure (which is seen as identical by all wavelengths), rather than its shape, per se, it has the advantage of being more achromatic in general (though this is a nuanced statement).

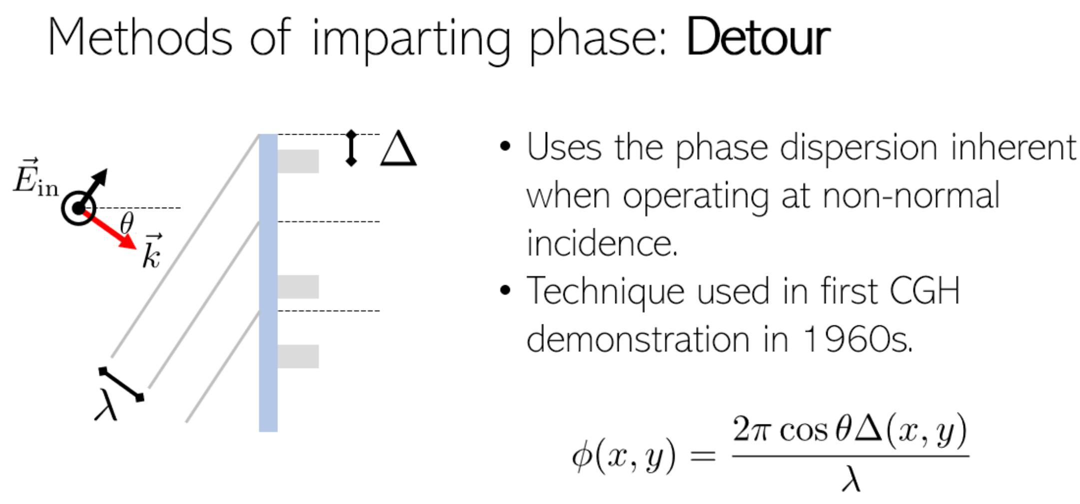


Figure 4.3. Phase can also be imparted by simply changing the transmission of the electric field per each unit cell of an optical element. Image credit: Noah Rubin

3. Detour phase: The detour phase uses the dispersion in phase inherent when operating at non-normal incidence (either by illuminating a hologram at an angle or observing the far field at an angle). In this way, the transverse location of scattering structures changes the phase of the incoming wavefront being sampled (see graphic above). This approach, which was employed to realize some of the first computer generated holograms in the 1960s, is not often used today.

4. Resonance-based phase: A final approach is to use resonant nanostructures. It is well-known that phase undergoes a drastic change when sweeping wavelength across a resonance. Similarly, near a resonance, changing structure size at a fixed wavelength causes a large change in accrued phase. Moreover, this can be done in a way that is ideally free of amplitude modulation if two resonances (in particular, electric and magnetic dipole modes) are arranged to be overlapped by virtue of the structure's design. In past work this has been shown with silicon circular pillar structures with resonances having quality factors of ~10. This approach, while of considerable academic interest, is not often used in contemporary metasurface research due to the fact that its reliance on resonances mandates extremely strict fabrication constraints (it does, however, generally allow for the use of structures with less extreme aspect ratios than the propagation phase approach, which is a notable possible advantage for fabrication). This method of phase control has spawned its own academic field, often referred to as “Huygens metasurfaces” or “Mie-tronics.”

Methods of achromatization: Achromatization is fundamentally challenging without the “true delay” of macroscopic propagation lengths afforded to designers in bulk optics (e.g., achromatic doublets and triplets, etc.).

Methods of achromatization/dispersion engineering:

- Brute force and hope: Creative meta-atom shapes to control group delay and group delay dispersion. Varying periodicity to achromatize.
- Different definitions of what achromatism means.

Inverse design methodologies: Traditional (“forward”) design methods make it difficult to fully explore the design space offered by subwavelength-scale scatterers. By formulating the metasurface design problem as an optimization, it is possible to find structures that can outperform those from library-based designs. Optimizing optical elements is not a new concept, indeed choosing metalens elements from a library can be cast as an optimization. However, these libraries limit themselves to simple structures such as circular or square posts. Topology optimization explores a more general design problem where the structure of the metasurface is itself the optimization variable, allowing for freeform and often unintuitive structures to emerge. In the past decade there have been breakthroughs in efficiently computing gradients that have enabled inverse design, specifically topology optimization, to be applied to many problems in optics and photonics (Molesky et al. 2018). Early work in this area focused on design integrated photonic structures such as Y-splitters, mode converters, grating couplers, and wavelength demultiplexers (Piggott 2015, 2020). These structures achieved high performance within compact footprints, enabling tighter integration of on-chip components. However, over time, topology optimization has been used to demonstrate free-space optical elements, including high-efficiency beam-deflectors (Yang 2018), metalenses (Arabai 2015, Pestourie 2018), color splitters, and photonic crystal resonators (Roberts et al. 2022).

In formulating the design problem, a figure of merit (FoM) is defined that encodes the desired functionality of our metasurface and it is extremized by adjusting the structure of the metasurface. Optimization algorithms can be broadly classified as gradient-based or gradient-free. Gradient-free algorithms do not use the gradient of the FoM and include evolutionary algorithms and particle swarms. These methods typically need a large number of FoM evaluations to converge, which is problematic for many electromagnetic problems that require considerable computational resources. For small problems, these techniques work well and can get close to the global extremum of the FoM. However, for a large portion of photonics design problems, gradient-based optimizers are more suitable since they converge to a solution much faster. Typically, algorithms such as gradient descent, adaptive moment estimation (ADAM), the method of moving asymptotes, and Broyden–Fletcher–Goldfarb–Shanno (BFGS) are used (Pestourie 2018). In most cases, topology optimization is a non-convex problem meaning there is no guarantee that gradient-based optimizers will find the global extremum, only a local one that is dependent on the initial conditions of the optimization. Despite this limitation, there have been many instances where topology-optimized structures have been empirically shown to closely approach absolute performance bounds (Miller 2012). The key to gradient-based approaches is efficiently computing the gradient of the FoM with respect to the optimization variable, which is the refractive index throughout the volume of the metasurface. The adjoint method prescribes an approach to compute these gradients with only two simulations, which allows structures with $>10^4$ parameters to be optimized quickly (Miller 2012). This technique has become ubiquitous in inverse design and has started to be included in commercial simulation packages.

As it relates to coronagraphs, metasurface-based implementations have successfully relied on the library approach with relatively simple structures. Topology optimization could be used in a few different ways to design coronagraph masks. First, similarly to the library approach described above, the metasurface area can be divided into subsections and within each section a specific phase and amplitude over a given spectral bandwidth can be targeted using the adjoint method (Pestourie 2018). These subdivided sections can then be stitched together and can account for more of the longer-range interactions that are fully ignored by the local periodic approximation.

Another approach is to use the adjoint method to optimize the full structure of the metasurface, eschewing the local periodic approximation described in the previous section. This requires simulating the entire metasurface, which can be prohibitive for large-area devices. However, it has been shown that exploiting symmetry and employing approximate solvers can dramatically speed up these simulations (Lin 2022). The geometry of the metasurface can be parameterized into post structures whose dimensions are adjusted based on the gradients calculated via the adjoint method (Arbabi 2020), or the metasurface can be made fully freeform. An advantage of constraining the geometry to simpler structures is that they are more robust to fabrication errors and generally easier to fabricate in terms of minimum feature size.

Finally, the metasurface design procedure can be coupled with system-level modeling such that the metasurface structure can be designed to extremize a FoM that includes quantities of interest such as contrast. Such end-to-end approaches have been used to realize thermal imaging systems, direction of arrival estimation, and full-color metalens imaging (Bayati 2021, Fröch 2025)

Recent work has also begun exploring end-to-end optimization frameworks in which coronagraph performance metrics, such as raw contrast and throughput, are incorporated directly into the metasurface design process.

Considering the vast body of work and variety of open-source codes now available, exploring topology optimization for coronagraph design is a promising avenue for targeting high-performance vector and scalar coronagraphs over a 20% spectral bandwidth.

Key challenges and future direction: While topology optimization and inverse design have demonstrated impressive results for individual meta-atoms and small-scale devices, several open challenges must be addressed before these methods can be routinely applied to the design of coronagraph-grade metasurfaces at the bandwidths and tolerances required for high-contrast imaging.

Creating gratings with achromatized efficiency over bandwidths of at least 20%. Most demonstrated topology-optimized and library-based designs are still optimized at a small number of discrete wavelengths, or over comparatively modest bandwidths, and efficiency often degrades away from the design wavelengths as the structure's effective index and accrued phase both vary with frequency. Reaching HWO-relevant instantaneous bandwidths of 10–20% while simultaneously maintaining high diffraction efficiency, the required vector or scalar polarization response, and a stable center-of-mass for the diffracted beam is therefore a significant open design problem. This likely requires explicitly incorporating multi-wavelength or broadband figures of merit into the optimization from the outset, rather than optimizing at a handful of wavelengths and hoping the response interpolates well in between and may also benefit from combining dispersion-engineered meta-atom libraries with topology-optimized freeform regions to balance bandwidth against computational cost.

Statistical modeling that is aware of fabrication variations. Meta-atom libraries and end-to-end optimized designs are generated under idealized, as-designed conditions, yet as-fabricated devices are subject to line-edge roughness, sidewall angle and verticality errors, critical-dimension bias, and deposition or etch-depth non-uniformity, all of which can shift a structure's phase and amplitude response away from its target value. Because reaching 10^{-10} contrast requires wavefront errors controlled to the picometer level, even small, statistically distributed fabrication deviations can translate into contrast-limiting leakage. Future design frameworks will need to incorporate fabrication-aware, statistical treatments—such as robust or worst-case topology optimization, or Monte Carlo-based tolerancing folded directly into the figure of merit—so that a design's performance is optimized not only for the nominal, as-designed geometry but across the realistic distribution of manufacturing variation it will actually encounter.

Accessibility of design tools to non-experts (MetaGPT). Topology optimization and adjoint-based inverse design currently require specialized expertise spanning electromagnetic simulation, gradient-based optimization, and nanophotonic fabrication constraints, which limits metasurface design to a relatively small number of specialist groups. Broadening participation in the field—bringing in astronomers, instrument scientists, and engineers who understand the desired optical function but not necessarily the underlying meta-atom physics—will require design tools that lower this barrier to entry. Emerging approaches, including higher-level design frameworks that abstract away solver details, pre-packaged and validated meta-atom libraries, and large language model-assisted design assistants such as MetaGPT, could allow a user to specify a desired phase, amplitude, or polarization function and receive a fabricable metasurface layout in return, without first needing to become an expert in inverse design. Simulation tools are discussed further on.

5 Metasurface Fabrication Techniques

Fabricating metasurfaces can be a highly non-trivial endeavor. Small and varying feature sizes, small and varying gaps between features, high aspect ratios, and difficult-to-etch materials are just some of the challenges involved. Despite these challenges, by leveraging microfabrication techniques from the semiconductor industry, metasurfaces fabricated in a variety of materials have been demonstrated to date. Examples of materials used for metasurface fabrication include silicon, titanium dioxide, and silicon dioxide. For a more comprehensive list of materials used in metasurfaces, see later sections in the report.

Depending on material choice and required feature sizes, there are two main fabrication paradigms utilized in the field of metasurfaces, i) etching to define the metasurface pattern, or ii) filling in of features using atomic-layer deposition and etch-back (Damascene process). Regardless of which of these two fabrication paradigms to use, the step that defines the features themselves is arguably the most important step. This is the lithography step and can be done using a variety of different lithography tools of varying complexity. Broadly speaking, lithography can be categorized as either photolithography or electron-beam (e-beam) lithography. Nanoimprint lithography is a technique that replicates a master that has been fabricated using either photo- or e-beam lithography and is closely related to these methods; we devote a section to it here as well even though it also has stark contrasts to the other lithography techniques.

5.1 Lithography techniques

5.1.1 Photolithography

In photolithography (PL), a light-sensitive photoresist (PR) is patterned by exposure to light through a mask. The mask, known as a photomask or reticle, contains the metasurface pattern as either holes (where light can pass through) or opaque marks (where light is blocked). This technique is advantageous because large areas can be patterned at the same time by flood illumination through the mask. Thus, large-scale metasurfaces can be generated. Further, the commonly available PL tools in academic cleanrooms represent tools evolved from mature industrial processes. This means fabrication with such tools is very well understood, consistent, and geared towards high throughput (i.e., many prototypes produced very fast). However, the resolution achievable by PL is limited by the wavelength of light used for exposure. This is roughly 200 nm for widely used 248 nm deep-ultraviolet (DUV) PL machinery. For metasurfaces that can be designed with these resolution constraints, PL can be preferable to e-beam lithography (EBL) because it can be substantially faster, less costly, and allows for large area patterning. For context, the semiconductor industry uses significantly more complex lithography tools, sometimes with immersion liquids to push the resolution limit. Also, in a push towards higher resolution, the industry has been moving from DUV towards Extreme-UV (EUV), and towards high-numerical aperture optical systems. However, these tools are costly investments and very expensive to maintain and will therefore not likely be widely available in academic settings soon.

While the resolution limitations of PL will always be larger than those of EBL, there are methods through which these limitations can be relaxed. One technique is known as optical proximity correction (OPC). OPC involves computationally optimized modifications to the metasurface shapes on a photomask in order to allow them to be more accurately resolved in the PR. For the case of a 248 nm DUV PL tool these modifications can allow for the resolution of structures below 200 nm, and also the more accurate resolution of nanopost cross-sections beyond just circles (e.g., squares, crosses, etc.). However, this comes at the price of added complexity in terms of specialized software and optimizing the mask itself.

5.1.2 Electron-beam lithography

Electron-beam lithography (EBL) is a high-resolution patterning technique that uses a tightly focused beam of high-energy electrons to write intricate structures on a surface coated with an electron-sensitive material, known as an e-beam resist. When exposed to the electron beam, the resist undergoes chemical changes, either breaking molecular bonds, making the exposed regions more soluble (for positive-tone resists), or creating cross-links, making the exposed regions less soluble (for negative-tone resists). During development, the more soluble regions are removed, creating the designed pattern onto the substrate in the form of patterned resist.

Conventional EBL systems operate by steering one or more electron beams to precisely scan the resist-coated substrate, deflecting and blanking (blocking) the beam to form the nanoscale patterns. The resolution of EBL is fundamentally limited by how finely the electron beam can be focused. In practice, resolution is limited by forward scattering of electrons in the resist, backscattered electrons from the substrate, lens aberrations, and Coulomb interactions rather than the wavelength (which is the dominant resolution limitation in optical lithography) of the electron.

Most commercial EBL systems operate at 20–100 kV, with higher voltages reducing electron scattering but requiring careful consideration of material susceptibility to knock-on damage. Pattern fidelity also depends on controlling the electron dose—the total energy delivered to a given area. The time the e-beam spends on each point (dwell time) must be precisely regulated to ensure accurate pattern replication. To speed up the writing process, the beam current density can be increased. However, higher electron flux causes Coulomb interactions (space charge effects), lens aberrations, and increased proximity effects, all of which can degrade resolution. Thus, there is a fundamental trade-off between writing speed and feature sharpness.

One way to mitigate slow writing speeds inherent in point-by-point scanning is through variable-shaped beam (VSB) technology. Unlike traditional approaches that expose patterns pixel by pixel, VSB systems use mechanically controlled apertures to dynamically shape the electron beam into predefined rectangles or trapezoids. This allows larger pattern segments to be written in a single step, significantly increasing throughput without sacrificing precision.

Ultimately, the choice between these techniques depends on the specific requirements of the pattern. While VSB provides a significant speed advantage, point-by-point scanning remains indispensable for applications requiring the highest precision at the smallest feature sizes. By balancing resolution, writing speed, and pattern complexity, EBL continues to play a critical role in advancing nanofabrication across research and industry.

5.1.3 Nanoimprint lithography (NIL)

Another promising fabrication technique is nanoimprint lithography (NIL), which allows for high-throughput, wafer-scale replication of metasurfaces with nanoscale precision. NIL relies on a prefabricated mold to imprint metasurface patterns directly onto a resist layer, which is then used as a template for material deposition or etching. Compared to other methods, NIL offers the advantage of cost-effective large-area fabrication while maintaining nanoscale fidelity. However, factors such as mold degradation, residual layer effects, and defects introduced during demolding must be carefully managed to ensure high-quality metasurface replication. To address these issues, different types of nanoimprint lithography (NIL) processes have been developed (L. J. Guo 2007). The general idea of NIL is to replicate an e-beam written pattern into another material by a molding process. Typically, a stamp is first produced from the e-beam pattern. This stamp will be a negative copy of the original pattern and most commonly consists of a rigid material so that high pressure can be used for the imprint. Two common types of NIL are thermal and ultraviolet (UV). The difference between the two is in the type of resist that is used. Thermal NIL uses a thermoplastic resist, while UV NIL uses a UV-curing resist. The resist is spin coated onto the substrate (typically a silicon wafer). In thermal NIL, the stamp is pressed against the resist at an elevated temperature, and sometimes also under vacuum. After a certain time (normally minutes) the pressure is released, the wafer is cooled, and the stamp can be separated from the resist leaving a more or less exact copy of the original e-beam written pattern in the resist layer on top of the wafer. In the case of UV NIL, the process is at room temperature. The stamp is pressed into the UV-curable resist and the resist is cured by UV light (typically less than a minute) before the pressure is released. This requires the stamp to be UV transparent.

Both of these NIL methods can produce patterns with a feature size less than 10 nm, while the UV NIL method is considered to be more productive since there is no need for elevated temperature. Most often the NIL pattern needs to be transferred into the wafer itself (see next section), or to a thin film on top of the wafer. Here, it is important to have as thin a resist residual layer as possible (see Fig. 5.1). With a thick residual layer, there is a risk the resist structure, now working as an etch mask, will be distorted before the thin film or wafer starts to be etched in the plasma-etching step. To keep the residual layer as thin as possible, the NIL process has to be optimized, mainly by taking account for having an original resist thickness that matches the volume needed to fill up the stamp.

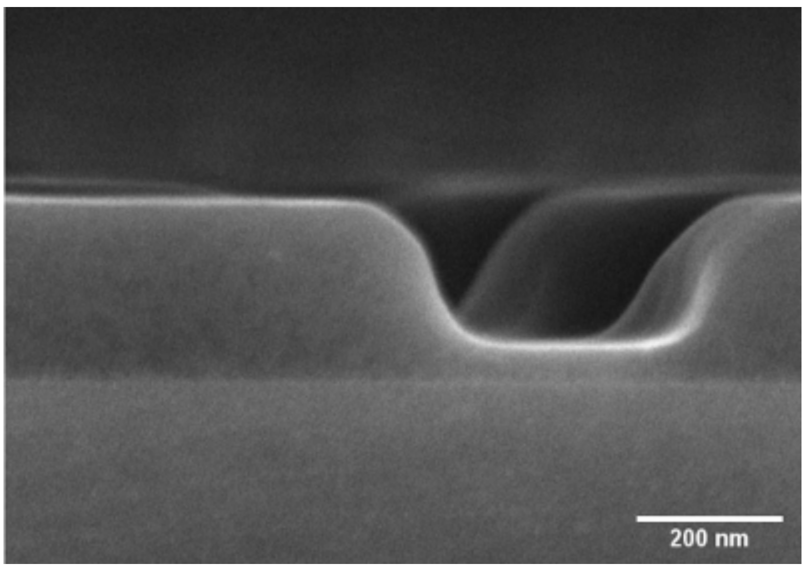


Figure 5.1. Scanning electron microscope picture of a binary grating fabricated in resist on top of a silicon wafer using SAMIM. Opening in the resist is 300 nm and residual thickness is about 50 nm and total resist thickness is 250 nm. Image credit: Pontus Forsberg

A simple variant of NIL, that does not need any specific NIL equipment, is solvent-assisted micro molding (SAMIM), which makes use of polydimethylsiloxane (PDMS) molded to the master pattern (D. R. E. King et al. 1997, E. Vargas Catalán 2016). The PDMS stamp is placed on the substrate with cured photoresist on and subsequently placed in a beaker with ethanol vapor that diffuses through the PDMS stamp. The cured photoresist will dissolve, and due to capillary forces, fill up the nanostructured PDMS stamp. Finally, the substrate is placed on a hot plate, and the PDMS stamp is gently removed, leaving a perfect replica of the e-beam pattern in photoresist on top of the wafer. If the nanostructures need to be in the substrate material, a thin etch mask film is used between the resist and the substrate. The replicated resist structures are then transferred into the masking layer using plasma etching, followed by another plasma etch step to transfer the structures into the substrate (see Fig. 5.2).

NIL can also be used to transfer curved surfaces into a resist, like microlenses, diffractive Fresnel lenses, blazed gratings, etc. NIL can be used to reduce the cost when fabricating metasurfaces, as well as for prototyping. However, great care must be taken when replicating patterns with large variations in the fill factor so that the resist volume needed to fill the pattern varies across the surface. Since the resist is coated with an even thickness, there is a risk of variations in residual layer thickness and/or voids in some areas.

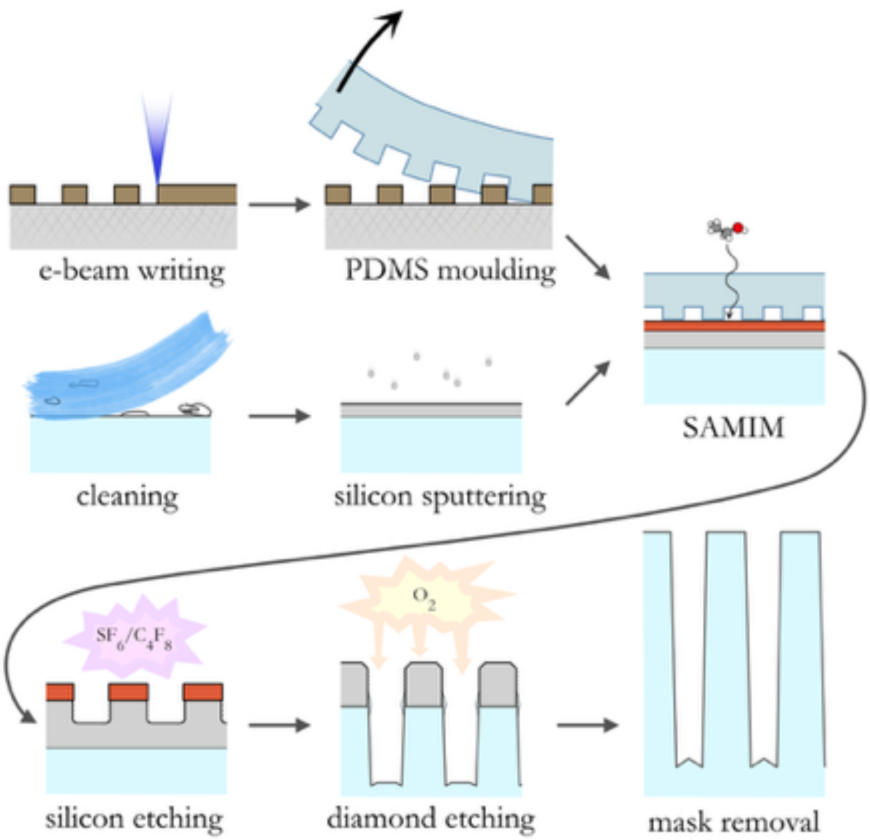


Figure 5.2. Schematic showing a typical process flow on metasurface fabrication in diamond using SAMIM and plasma etching. Image credit: P. Forsberg et al., Diamond & Related Materials 2024, https://doi.org/10.1016/j.diamond.2024.111237]

5.2 Transfer etching

Creating a metasurface pattern is typically done using lithography techniques and the pattern is created in a resist. This pattern can then be transferred into the underlying material(s) that defines the metasurface. Typically, metasurfaces are designed using vertical nanofeatures of varying shapes, and, depending on the type of etch and the gases used, need to be adapted to the metasurface material in question. The achievable verticality and feature sizes depend heavily on the choice of materials.

A material with widespread utility, both in metasurfaces and beyond, is silicon. Silicon wafers are readily available and amorphous silicon (a-Si) can be deposited using, for example, plasma-enhanced chemical vapor deposition (PECVD). The ability to deposit a-Si on a variety of materials opens a wide range of materials to be used for a base substrate of metasurfaces based on a-Si.

Transfer etching of mask materials into a-Si is typically performed using a combination of the gases SF_6 and C_4F_8—the well-known Bosch process in which one cycles through SF_6 (etching) and C_4F_8 (passivating) to achieve a vertical etch. A variant of this process, often referred to as a pseudo-Bosch process, utilizes both gasses simultaneously (no cycling), and can offer greater level of control when large etch depths are not required. The use of other metasurface materials may require the use of etching with other gases such as CHF_3, CF_4, Cl_2, BCl_3.

The mask material itself can be photoresist, in which case only a development of the as-exposed resist is needed before the transfer etch. For a variety of reasons, dielectric hardmasks are often used as an alternative to resist masks. For instance, resist mask erosion is a common issue when using only a resist mask—this leads to roughened features, typically close to the top of the metasurface features. Typically, the use of hardmasks enables better feature control and allows finer metasurface features to be controllably patterned. However, the use of hardmasks brings about the issue of how to pattern the hardmask itself. This can also be done using plasma etching to transfer the resist pattern into the hardmask. On the other hand, hardmask materials that can be evaporated can also be patterned using a liftoff process. In this case, a patterned resist would be covered with evaporated materials, and any area without resist would be coated with the material. Post-evaporation, any resist and hardmask material on top can be dissolved in a solvent such as acetone. This allows patterning of the hardmask without the need to etch the hardmask material, which can be of use for materials that are difficult to etch, such as aluminium oxide. Silicon oxide is another mask material that can be evaporated and therefore useful in a liftoff process. Metals can often also be evaporated and so can also serve as a liftoff-compatible hardmask for etching. However, metals can re-sputter during the plasma etch itself, which can lead to micromasking and result in highly undesirable features introduced during the etch itself.

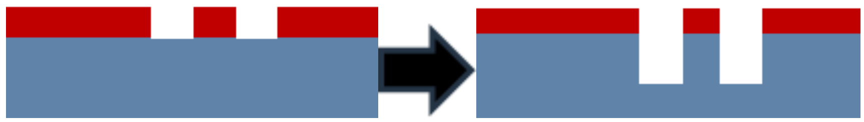


Figure 5.3. An anisotropic etch used to transfer a pattern in resist (red) into the underlying material (blue). Typically, the resist is removed in a solvent or by oxygen plasma towards the very final stages of the fabrication. Image credit: Tobias Wenger

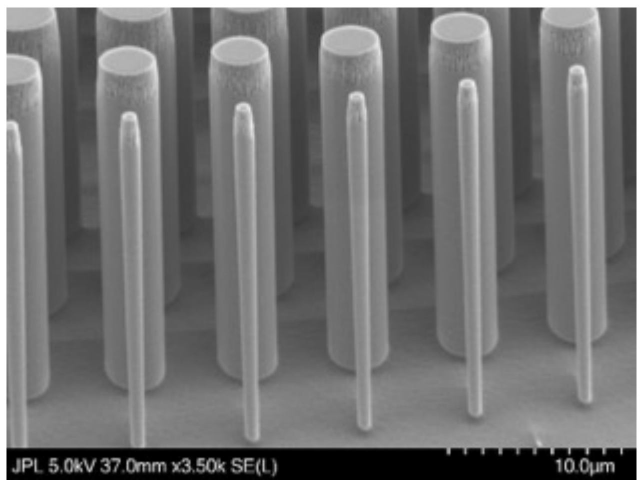


Figure 5.4. Scanning electron microscope (SEM) image of silicon pillars etched using a Bosch process. A photoresist mask was used, and clear signs of resist erosion are visible at the very top of the pillars. Image credit: Tobias Wenger

5.3 Damascene

The damascene process is a widely adopted approach for fabricating metasurfaces, particularly for designs requiring complex geometries and high aspect ratios. This process begins by defining an inverse pattern of the metasurface using sacrificial materials such as e-beam resists or photoresists, which will be removed later. These inverse structures, typically in the form of trenches or cavities, are then conformally filled with the meta-atom material—such as TiO_2—using deposition techniques like atomic-layer deposition (ALD) or chemical vapor deposition (CVD).

If the sacrificial layer consists of thermally sensitive materials, such as resists, deposition must be performed below their glass transition temperature (typically around

100° C for common resists) to prevent deformation. Once deposition is complete and the trenches or cavities are fully filled, excess material is etched back using anisotropic plasma etching to reveal the original surface. Finally, the sacrificial layer (e.g., resist) is removed, leaving behind the deposited metasurface structures. Figure 5.5 illustrates the damascene process used to create 600 nm-tall TiO_2 nanopillar structures.

In a recent development, a damascene process was demonstrated that could create double-layer metasurfaces. In essence, this takes advantage of the planarized stage of the damascene process to create a subsequent layer. This could prove a significant development as it opens up more degrees of freedom in terms of metasurface design that can be used to achieve a target functionality.

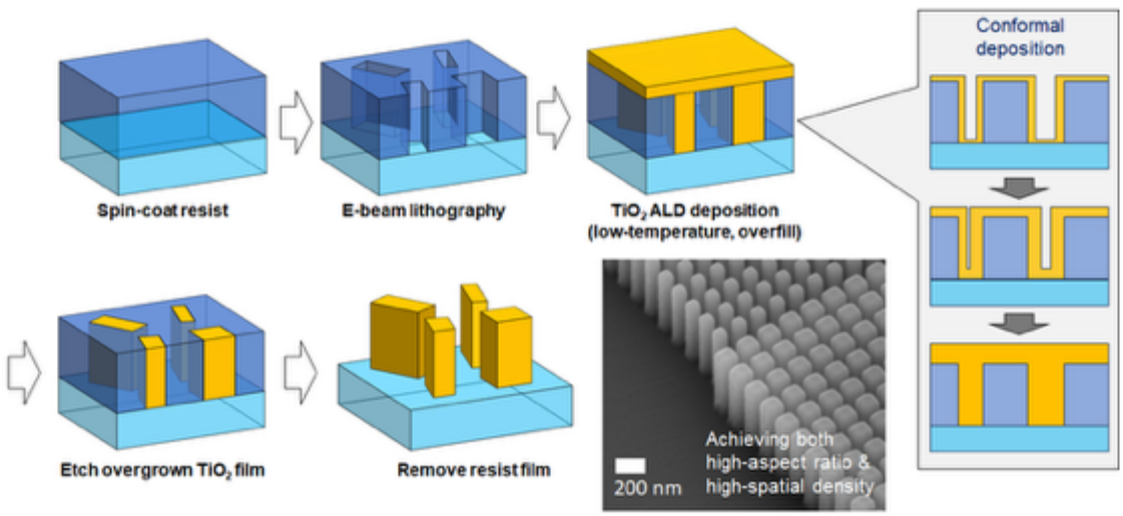


Figure 5.5. The damascene process is illustrated in the multiple steps shown above. Image credit: Joon-Suh Park

5.4 Materials for metasurface fabrication

Metasurfaces can be made from a variety of materials, and perhaps the most important aspect is the optical properties at the wavelength range in question. Low optical absorption is of utmost importance for most applications, along with high refractive index. While a complete list of every material ever used for metasurface fabrication is unfeasible, we here go through and discuss a handful of important metasurface materials and discuss some of their important properties. Metasurface materials in the infrared, visible, and UV parts of the electromagnetic spectrum are discussed separately in the following sections.

5.4.1 Metasurface materials for IR

Silicon is the most ubiquitous material throughout the semiconductor manufacturing world. An appealing property of silicon is the ability to etch very high-aspect ratio structure via the Bosch process (using SF_6 and C_4F_8), which is a widely established etch process throughout industry and academia. Silicon also has a high refractive index, leading to good optical confinement of individual metasurface structures. In addition, a variety of good hardmasks are available such as SiO_2 and Al_2O_3 that are compatible with silicon. Resist is also an option to use for etch mask, mostly if the required etching is not very deep.

Silicon is either deposited using CVD or sputtering, in which case it is amorphous a-Si, or crystalline silicon wafers are used and metasurfaces are fabricated monolithically into the wafer itself. A benefit of using a-Si is it can be used together with highly transparent substrates such as fused silica or sapphire, which makes the entire metasurface device highly transparent. A variety of coating techniques are available such as plasma-enhanced chemical vapor deposition and physical vapor deposition (i.e., sputter). Metasurfaces using this platform have been demonstrated to work for many applications including scalar vortices, gratings, and achromatic lenses.

Metasurfaces utilizing the monolithic approach typically have lower total transmission due to the presence of the substrate with high index. The large surface reflection can be substantially mitigated by depositing an anti-reflective coating on the back surface. Another issue that can arise is absorption inside the substrate itself, which depends on wavelength. A convenient method to avoid this problem is to utilize silicon with a very low amount of trace contaminants—often called float-zone silicon. This type of silicon has a substantially lower absorption than does silicon originating from the more conventional Czochralski method.

Diamond is well known for its extreme hardness and excellent wear properties. Diamond also has other unique properties, like being the material with the broadest optical transmission window. It is transparent from 220 nm up to the mm region, having some small absorption bands resulting from two phonon absorption between 2.5 and 6.5 μm (A. M. Zaitsev 2001). In the infrared, the refractive index goes from 2.39 at a wavelength of 1 μm and then goes down to 2.38 at a wavelength of 2 μm, and then remains more or less constant at 2.38 at the longer wavelengths. The typical absorption coefficient is less than 0.1 cm^{-1}. Moreover, diamond is the best heat conductor of all solids, making it interesting as an optical material in, for instance, high-power laser applications. Synthetic diamond of optical quality, having similar properties as natural diamond, has been commercially available for about 25 years.

For optical purposes diamond comes in two forms: polycrystalline and single crystalline (R. Mildren & J. Rabeau 2013). It should be noted that diamond is quite costly compared to other optical materials, especially in the form of single crystals; therefore, diamond is normally used when the above properties are vital for the application in mind.

Diamond is very difficult and time consuming to machine and the only way to polish diamond is to use diamond particles on top of for instance a silicon carbide wafer. Costs increase when a very smooth surface is needed. However, in the case of texturing, micro- and nanostructures in diamond can be fabricated by using plasma etching in an oxygen or chlorine chemistry (G. S. Sandhu et al. 1989 and E. Vargas Catalán et al. 2016, and ref. A. Toros et al. 2020). Producing optical structures in diamond via inductively coupled plasma etching requires high ion surface bombardment to achieve an etch rate of about 250 nm/min (Nikolajeff and Karlsson 2013). Therefore, it is important to have a hard, thick and well-defined etch mask to accurately etch metasurfaces. Typical etch mask materials are aluminum, silicon oxide, silicon nitride, hydrogen silsesquioxane (HSQ) and silicon. Up to now, the metasurfaces in diamond that have been demonstrated include so-called moth-eye structures to minimize the Fresnel reflection (Fig. 5.6) and the annular groove phase mask (AGPM) (Fig. 5.7) (Karlsson et al. 2003 and C. Delacroix et al. 2013).

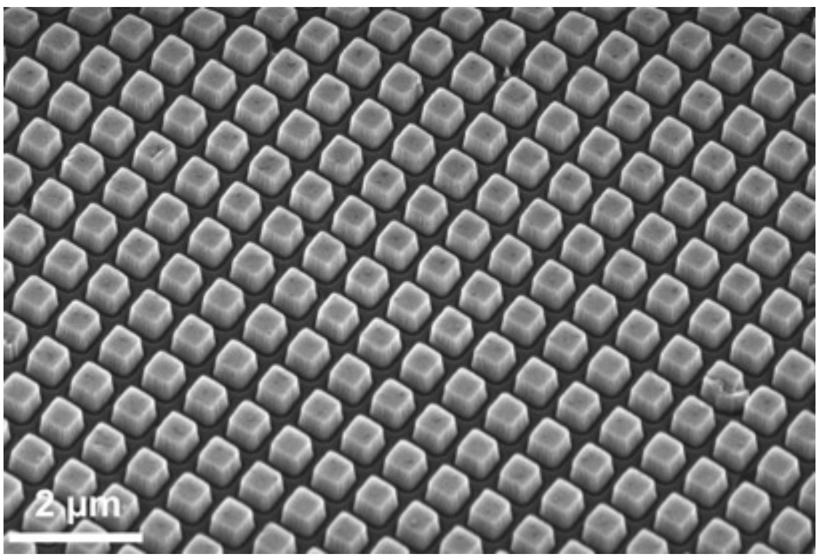


Figure 5.6. SEM picture of meta structures for anti-reflection treatment of diamond. Image credit: Pontus Forsberg

5.4.2 Metasurface materials for visible wavelengths

Titanium dioxide (TiO_2) is one of the first widely adopted material platforms for visible-wavelength metasurfaces, owing to its high refractive index and low absorption. Its high refractive index enables significant phase accumulation within a relatively short optical path length, allowing meta-atoms to be designed with smaller heights. TiO_2 metasurfaces have been realized using various methods, including direct etching techniques, the Damascene process, and nanoimprint lithography (NIL).

One of the earliest demonstrations of TiO_2 metasurfaces utilized direct etching of TiO_2 thin films. The successful fabrication of sub-wavelength features that modulated wavefronts at zeroth-order transmission showed great promise. However, challenges such as lateral etching, microloading effects, non-vertical etching, and sidewall bowing made it difficult to precisely replicate metasurfaces designed with high-aspect-ratio, vertical side-walled structures having dense spatial layouts. These limitations hindered the scalability and uniformity of directly etched TiO_2 metasurfaces.

To address these challenges, the Damascene process emerged as an alternative fabrication approach, enabling the realization of complex designs and high-aspect-ratio structures without the difficulties associated with pattern transfer and etching, such as micro-loading effects and maintaining vertical sidewalls. Due to these advantages, TiO_2 metasurfaces have frequently been demonstrated using this method.

By leveraging these different fabrication techniques, TiO_2 metasurfaces continue to advance, enabling applications in broadband, high-efficiency flat optics, wavefront shaping, and miniaturized optical components for both research and commercial applications.

Silica (SiO_2, fused silica, monolithic) is an optical material with a long heritage and remains a favored choice for space-based optics due to its stability against thermal effects, including refractive index variations and volume changes. With its wide bandgap, silica exhibits high transparency not only in the visible spectrum but also in the ultraviolet and short-wave infrared, making it an ideal material for broadband optical imaging. However, its low refractive index has traditionally limited its effectiveness for metasurface applications. To achieve the same level of transmitted phase control as a 600 nm-tall TiO_2 meta-atom, silica nanostructures must be at least 1.5–2 μm tall, posing fabrication challenges for high-efficiency metasurfaces.

Recent advancements in anisotropic etching techniques have enabled the realization of such structures, leading to several demonstrations of fused silica metasurfaces operating in the visible and vacuum ultraviolet (VUV) regimes. In addition, silica is highly desirable for metasurfaces fabricated using conventional complementary metal oxide semiconductor (CMOS) foundry processes, as it is widely accepted in existing fabrication tools, even under strict material restrictions. A notable example is a monolithic fused silica metalens with a 100 mm diameter, demonstrating the feasibility of scaling to even larger apertures using standard CMOS-compatible tools and techniques.

Beyond its optical advantages, silica metasurfaces exhibit exceptional mechanical and thermal resilience. It has been demonstrated that silica nanopillars can withstand abrupt temperature fluctuations ranging from −197° C to 200° C, as well as strong vibrations, making silica an ideal material for metasurfaces intended for operation in harsh environments, including space. In addition to its mechanical and thermal robustness, silica is also highly resistant to radiation damage, including prolonged exposure to ultraviolet radiation and charged particles, making it a reliable choice for long-duration extraterrestrial missions.

Furthermore, the wide transparency range of silica, extending into the extreme ultraviolet (EUV), makes it a strong candidate for metasurfaces in next-generation semiconductor lithography and high-power laser optics, where low absorption and thermal stability are crucial. While silica's low refractive index presents a design challenge, hybrid metasurfaces that incorporate high-index materials such as TiO_2 or silicon have been explored, offering enhanced phase control while preserving the robustness and CMOS compatibility of silica. These developments position silica as a key material for future metasurface applications across a broad range of wavelengths and operating conditions for optics in space.

Diamond (single crystal) When using diamond for visible light applications, single crystal diamond is preferred since the microcrystalline structure of polycrystalline diamond can scatter light at shorter wavelengths, causing problems in the optical system. The same metasurface fabrication processes as those for polycrystalline diamond can be used for single-crystal diamond. In polycrystalline diamond, there are several grain boundaries containing other forms of bonded carbon (sp^2) that can etch at a different rate than the pure diamond (sp^3). It should be noted this is not a big problem, but at shorter wavelengths, this fact must be taken into consideration. In Fig. 5.7, an AGPM etched in

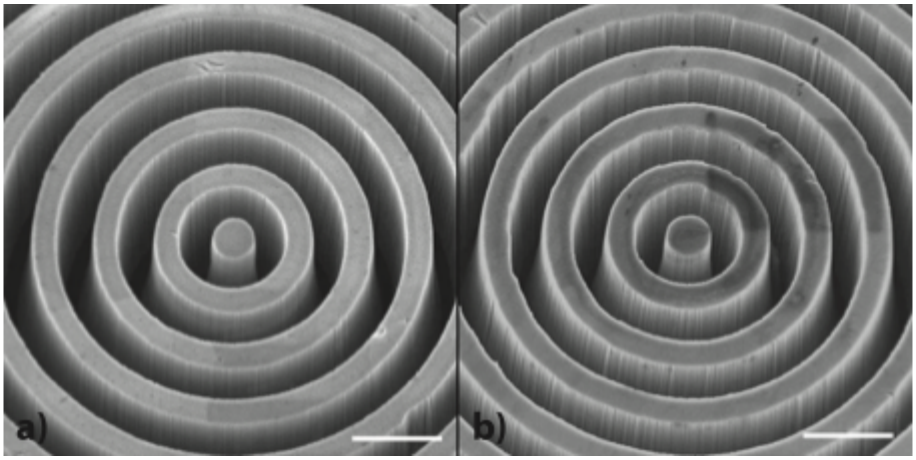


Figure 5.7. SEM pictures show an AGPM in polycrystalline diamond (a) and single-crystal diamond (b). Scale bars are 2 μm long and 0.1 μm wide. Image credit: P. Forsberg et al., Diamond & Related Materials (2013)

single-crystal diamond is shown; for comparison, an AGPM in polycrystalline diamond is also shown.

5.4.3 Metasurface materials for UV

Hafnium oxide (HfO_2) is a wide-gap semiconductor with a high refractive index, with appealing properties for optical metasurfaces. Among its appealing features is the ability to deposit with low-temperature ALD and its transparency from IR into the UV. Hafnium-oxide metasurfaces working in the UV have been demonstrated in the literature (Zhang 2020) and can also be an option for visible-light metasurfaces, although TiO_2 is often the preferred choice in the visible due to a higher refractive index.
HfO_2 is often used for advanced optical coatings and is also a commonly used material in the semiconductor industry and therefore provides a CMOS-compatible route for metasurface manufacturing. In addition, HfO_2 is a robust material with a very high melting point, high dielectric breakdown strength, and exhibits good mechanical properties. Atomic-layer deposition (ALD) of HfO_2 is well understood and there are various precursor choices that can be made.

Tantalum pentoxide (Ta_2O_5) is another wide-gap semiconductor commonly used in both optical coatings as well as in the microelectronics industry. Recent work has demonstrated UV metasurfaces with high efficiencies, establishing Ta_2O_5 as an interesting CMOS compatible metasurface platform. An interesting property of Ta_2O_5 is the ability to etch highly anisotropic features in fluorine plasma chemistry, enabling the use of the direct-etch fabrication method to be used, although Ta_2O_5 can be deposited with ALD and so the damascene method can, in principle, be used as well.

SiO_2 (UV-enhanced fused silica, monolithic) With transparency from 175 nm to the visible spectrum, UV-grade fused silica is an ideal material for ultraviolet optical applications. By etching high-aspect-ratio structures, metasurface-based optical elements can be fabricated from SiO_2. Recent demonstrations have shown that highly efficient metasurfaces that operate in vacuum UV (VUV, 175 nm) can be realized using UV-grade SiO_2, enabling advanced VUV optical components with enhanced performance.

In addition to its optical transparency, fused silica exhibits exceptional resistance to radiation-induced damage, making it an excellent candidate for space-based UV optics and extreme ultraviolet (EUV) lithography applications. Advancements in deep reactive-ion etching (DRIE) and anisotropic plasma etching have enabled the fabrication of high-aspect-ratio SiO_2 nanostructures, allowing metasurfaces to achieve high efficiency across the UV spectrum.

One challenge in fabricating SiO_2 metasurfaces is maintaining smooth sidewalls and high aspect ratios during etching. Recent optimizations in etch chemistry and process control have significantly reduced sidewall roughness, improving optical efficiency. These advancements position monolithic SiO_2 metasurfaces as a promising platform for applications in UV microscopy, high-power laser optics, and biosensing, where high transparency and damage resistance are critical.

6 Metasurface Characterization

To fully understand the performance of a metasurface, it needs to be characterized at both the component level and the system level. Component-level testing is important to save time and energy that would otherwise be spent on incorporating optical elements into a larger system, where efforts to understand performance can be more time consuming and expensive due to the larger engineering scope. Therefore, component-level testing is meant to confirm metasurface properties that are clearly and accurately quantified and compared to device requirements before moving on to the next level of integration. System-level measurements for coronagraphs can be starlight suppression, leakage, and spectral bandpass.

6.1 Component testing

Within component-level testing, we can measure either the physical properties of the meta-atoms, or the optical properties of the metasurface. The physical properties of the metasurface that can be directly measured include post height, lateral dimensions, and refractive index of the substrate or metasurface material. Physical dimensions of the meta-atoms can be measured using scanning electron microscopy (SEM), and atomic force microscopy (AFM). Refractive index properties can be measured using optical characterization such as reflectometry or ellipsometry. Matching the dimension of as-fabricated meta-atoms and as-deposited refractive index to the simulated target design is critical to obtaining the proper optical properties of the metasurface. The optical properties of the fabricated metasurface often include the phase and amplitude impact on incident light, both polarization-dependent and independent.

6.1.1 Structural characterization of individual meta-atoms

Characterization usually begins with understanding the morphology of the elemental nanopillars. Inspecting their geometry is a good starting point for determining if the metasurface fabrication proceeded as expected. Given the small feature size of the nanopillars, this is done with a scanning electron microscope (SEM). Measured properties include factors such as overall physical extent, nanopillar shape and cross-sectional profiles. Instruments that can perform these measurements are scanning electron microscopes and high-resolution optical microscopes.

6.1.1.1 Scanning electron microscope

An SEM uses a focused electron beam that is scanned across the imaged area. By recording various signals arising from the interaction of electron with atoms within the sample, an image is produced. Depending on the operating mode, the SEM records, for example, secondary emission electrons or backscattered/reflected electrons. Different signals have different properties, and this can be utilized to emphasize, for example, topography or material composition within the sample.

An example of a scanning electron microscope is the model JSM-IT210 from JEOL and can be used to measure lateral meta-atom dimensions[1].

Properties of a typical SEM:

- **Magnification:** 100x–1,000,000x

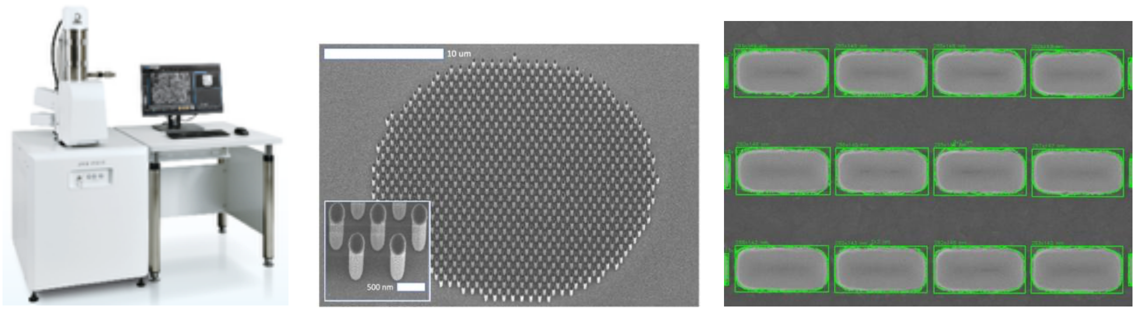


Figure 6.1. (Left) An image of a laboratory scanning electron microscope (SEM) instrument. (Center) An example measurement of a metasurface component. (Right) A sample SEM image analyzed using opencv-python. Image credits: JEOL (left), Kent Wallace (Jet Propulsion Laboratory) (center), Dan Shanks (Jet Propulsion Laboratory) (right)

- **Field of View:** ~1 micron–1 millimeter
- **Spatial Resolution:** ~1 nm
- **Pros:** Ideal for measuring lateral dimensions of individual meta-atoms. Can be integrated with image analysis software for rapid measurement of many meta-atoms simultaneously.
- **Cons:** Non-zero beam diameter of scanning electron beam can blur image, or mis-estimate meta-atom dimensions. Cross-sectional images often require destructive cleaving or ion milling process.

6.1.1.2 Atomic-force microscope

An example of an atomic-force microscope for metrology is the Bruker FastScan that can measure meta-atom heights[2].

[1] https://www.jeol.com/products/scientific/sem/

[2] https://www.bruker.com/en/products-and-solutions/microscopes/materials-afm/dimension-fastscan.html

Figure 6.2. (Left) The fundamental operation of the AFM. (Center) A laboratory instrument. (Right) Sample data of a metasurface. Image credit: Bruker Corp.

Properties of a typical AFM:

- **Field of View:** ~100 um
- **Spatial Resolution:** ~10 nm
- **Pros:** Nanoscale height mapping, most accurate for measuring height of metasurface non-destructively
- **Cons:** AFM tip is often too large to fit in between small meta-atom gaps, requires de-convolution of tip width to obtain accurate measurement of lateral meta-atom dimensions.

6.1.2 Characterizing material properties: complex index of refraction

It is also important to understand the dielectric properties of the as-fabricated nanoposts, as well as the substrate. On one hand, the wavelength-dependent complex index of refraction ($n = n[\lambda] + i\, k[\lambda]$) is a fundamental parameter in the design process, yet, on the other hand, it also varies a lot. That variability can be between deposition chambers but can also vary depending on the deposition process. (The vast majority of nanopillar materials are dielectric with real valued indices of refraction.) It’s common in the design process to start with a default value for the index of refraction that is either intrinsic to the design software, or taken from a textbook or a published paper, but these values reported in the literature are not always sufficient or accurate. Therefore, the best practice is to directly measure the index of refraction of the materials as they are deposited in the chamber that will be used for final metasurface element. This is often done using a small witness or coupon sample on a known substrate and measuring the refractive index of the material using thin film interference.

6.1.2.1 Reflectometry

A reflectometer consists of a broadband white light source output and a spectrometer to measure the wavelength dependence of the reflected intensity of light off a thin film. The light can either be collimated to measure the reflectance off a large area of the thin film sample or coupled to a microscope lens and pinhole to obtain reflectance of a specific point within a small diameter on a sample.

An example of a spectroscope reflectometer: Filmetrics F-10[3]

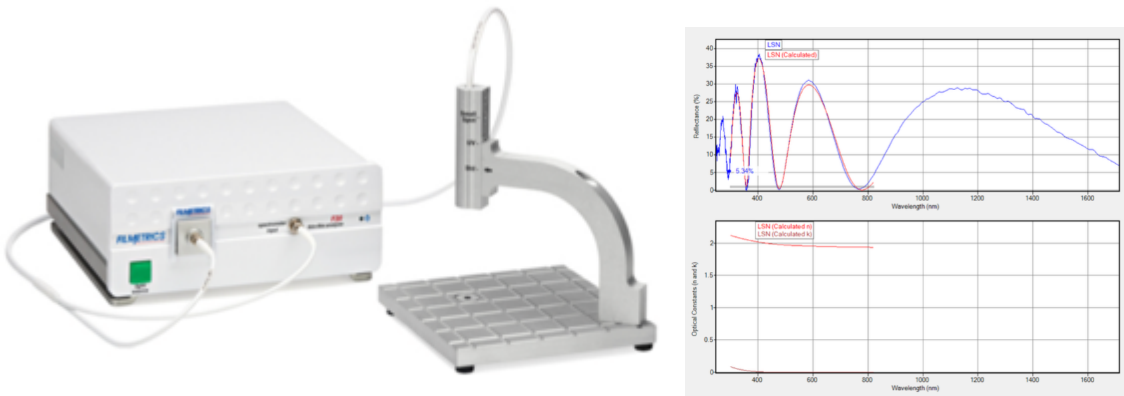


Figure 6.3. (Left) A commercial Filmetrics reflectometer uses fiber-coupled input and output and a collimated lens setup. (Right) Filmetrics software measurement and complex refractive index analysis output. Image credit: KLA Instruments

Properties of a reflectometer:

- **Wavelength Range:** 200 nm to 2000 nm
- Complex index of refraction measurement (n, k)
- **Film thickness measurement:** An ellipsometer similarly uses thin film interference to measure the complex index of refraction of a thin film of sample material. Instead of measuring only the intensity of reflected light off a sample, it uses an actively rotating polarization element to measure the polarization state of generated incident light after reflection at a non-zero angle of incidence off the film. This is often performed broadband at many wavelengths and angles of incidence to construct a more thorough model of the thickness and complex index of the thin film.

[3] https://www.kla.com/products/instruments/thin-film-reflectometers/filmetrics-f10-ar#details

6.1.2.2 Spectroscopic ellipsometry

An example of a spectroscope ellipsometer is the model R2 from J.A. Woollam[4].

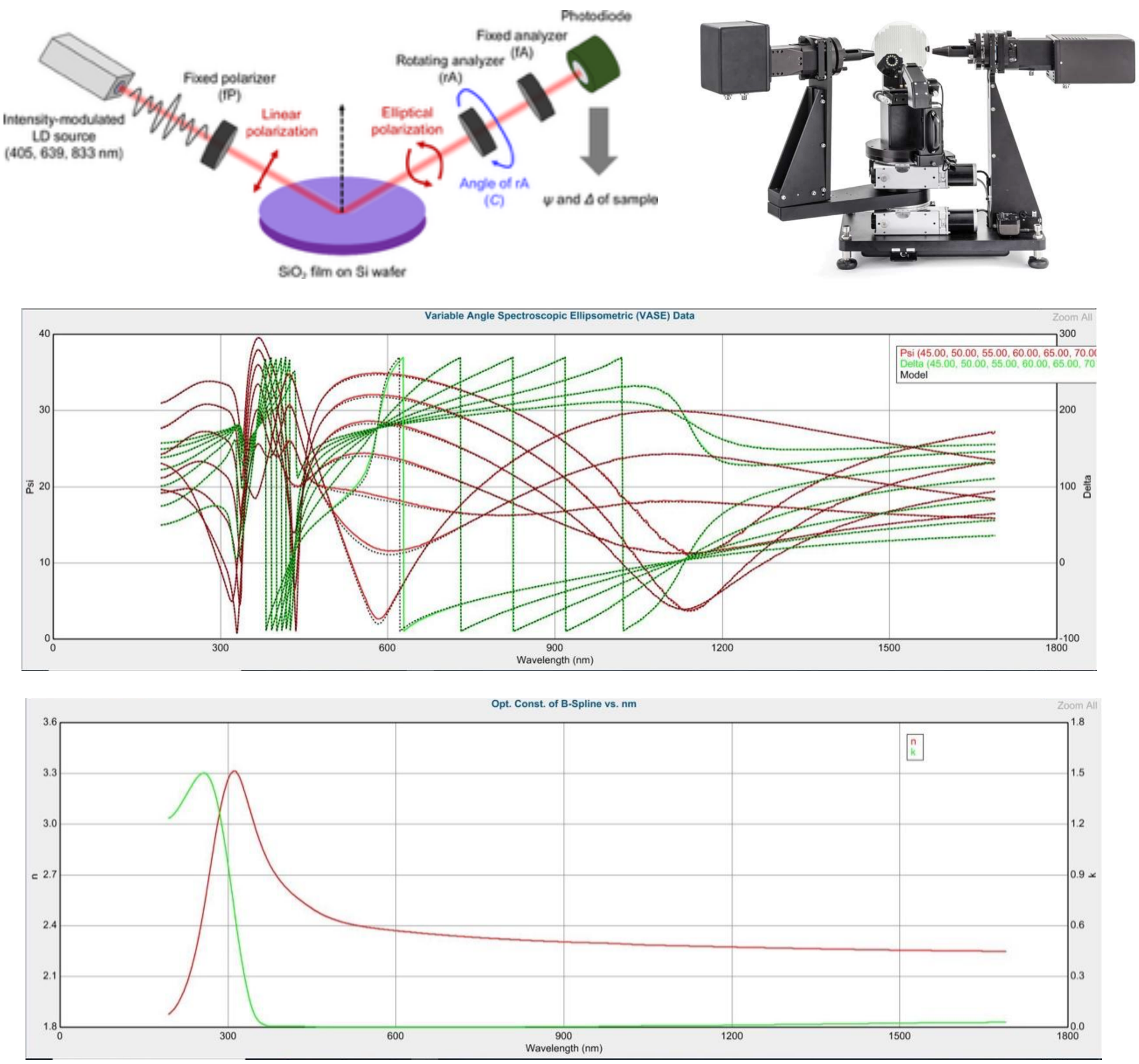


Figure 6.4. (Top Left) Fundamental operation of an ellipsometer . (Top Right) An example of an ellipsometer with robotic arms for automated measurement at multiple angles of incidence. (Center) Sample ellipsometer data taken from a thin TiO_2 film and (Bottom) complex refractive index analysis using J.A. Woollam CompleteEASE software. Image credit: J.A. Woollam

6.1.3 Substrate selection: Ensuring a good foundation

While we recognize the importance of understanding the geometrical and dielectric properties of the nanopillars, of equal importance are the properties of the substrate on which the nanopillars are deposited. Not all optical glass of the same type is made to the

[4] https://www.jawoollam.com/products/rc2-ellipsometer

same level of quality. For instance, fused silica can vary in the homogeneity of index, in the number and sizes of bubbles and inclusions, and in the concentration of trace constituents such as OH-. These properties are related to the glass itself. The substrate must also be prepared (ground, polished, and possibly coated) to be suitable for the metasurface fabrication.

6.1.4 Measuring as-fabricated component optical properties

The methods of characterizing the nano-pillar geometry and dielectric properties, as well as the properties of the substrate, while very important and necessary, still don't inform the ultimate success of the final device. An optical characterization of the device itself is needed to determine if it meets its as-designed properties.

This includes characterization of the transmitted/reflected phase, amplitude, retardance and diattenuation of the output beam as a function of lateral position across the surface. In addition, vacuum and space qualification of the components will be important.

Here, we discuss both commercially available optical metrology equipment, as well as custom-built instruments.

6.1.4.1 Interferometers

A phase-shifting interferometer is used for measuring the phase of an optical component (either in reflection or transmission). Commercial Fizeau interferometers typically work at a single wavelength and polarization and measure the transmitted/reflected phase. There are several companies that make these instruments. An example is the Zygo VeriFire HD.

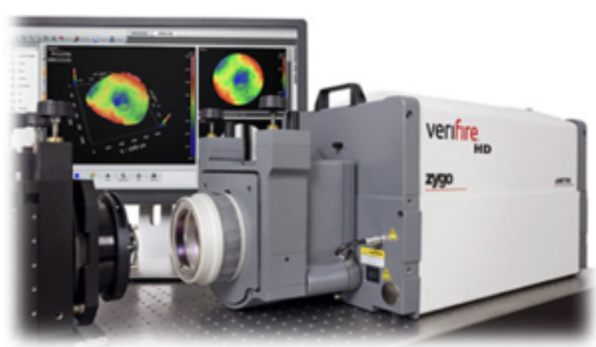


Figure 6.5. An image of a commercially available phase-shifting Fizeau interferometer. This particular model is the Zygo VeriFire HD. Image credit: AMETEK, Inc.

Properties of a typical phase-shifting, Fizeau interferometer:

- **Wavelength (fixed):** 632.8 nm to 1550 nm
- Quantitative phase measurement in either reflection or transmission (double pass)
- **Field of view:** Dependent upon the exact optical configuration.

Custom Built

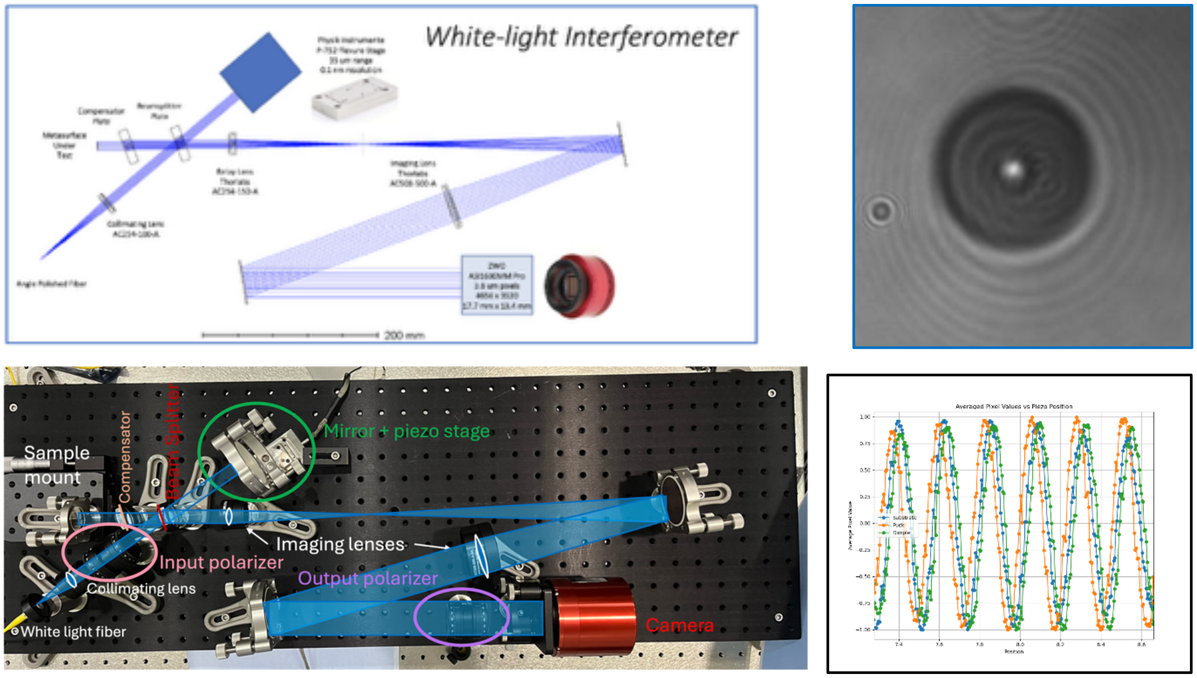


Figure 6.6. These images capture the key aspects of a custom-built Michelson white light interferometer. (Top Left) The optical design of the interferometer; (Bottom Left) The assembled hardware. (Top Right) An image of the device under test.(Lower Right) An example of a white-light fringe scan. Image credit: Kent Wallace

Another option is to build a custom interferometer on a benchtop. The primary advantage of a custom-built system is control over the wavelength, magnification, and accuracy specifications of the system. Additionally, polarization elements can be added to obtain polarization-dependent measurements.

6.1.4.2 Mueller-matrix polarimeter

Commercially Available

Metasurface optical components can be used to modify the polarization state of light. Polarimeters are used to characterize the optical components and provide a quantitative measurement of the full 4x4 Mueller Matrix. An example of this polarimeter is the AxoStep from Axometrics, Inc.

Properties of a typical commercially available Mueller-matrix polarimeter:

- **Wavelength (fixed):** 632.8 nm to 1550 nm
- Quantitative phase measurement in either reflection or transmission (double pass)
- **Field of view:** Dependent upon the exact optical configuration.

Custom Built

Several research groups are also building their own custom benchtop polarimeters. The precision of a custom benchtop rotating-retarder Mueller-matrix-imaging polarimeter is physically constrained by the size and shape of the mechanics that rotate the polarization optics; the specifications around motor step size, backlash, and repeatability determine how finely a retarder can be stepped in the calibration, as well as the accuracy of the polarization analyzing optics.

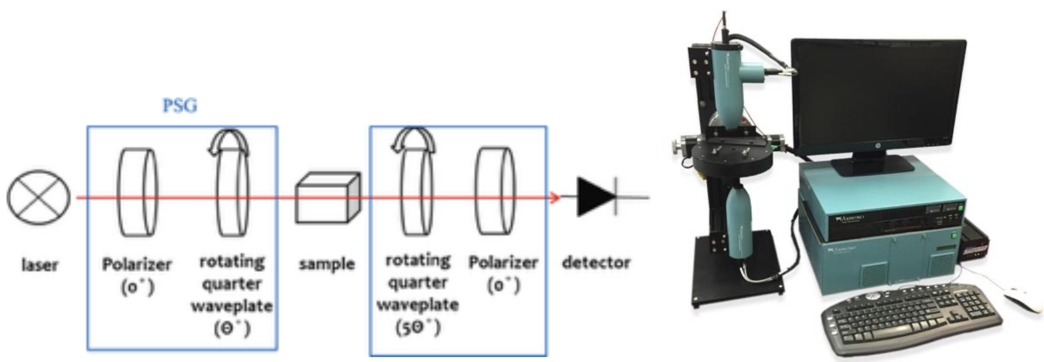


Figure 6.7. (Left) The schematic of the polarimeter instrument. (Right) An image of a commercially available polarimeter from AxoScan. Image credit: Axometrics, Inc.]

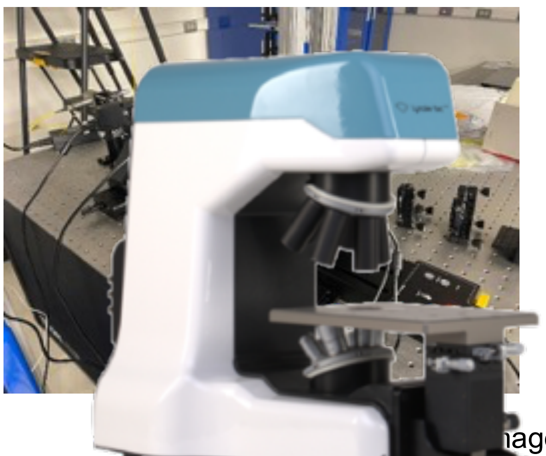
Properties of a custom-built Mueller-matrix polarimeter:

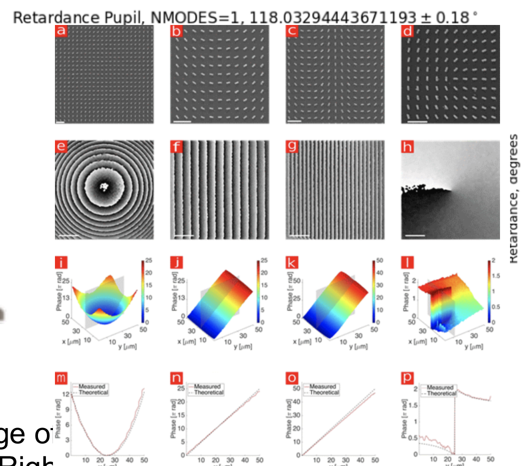
- Wavelength (variable, with a tunable laser source): visible, and near-IR
- Characterization of the full Stokes parameters
- Field of view: Dependent on the exact optical configuration.

6.1.4.3 Digital holographic microscope (DHM)

Commercially Available

Measuring the properties of small optical components is possible using digital holographic microscopes. Such instruments have high spatial resolution and measure transmitted/reflected amplitude and phase. Commercially available systems can be procured from LyncéeTec.





[illegible]age o[illegible] operating in the [illegible] and (Righ[illegible]) an [illegible] of the data acquired from the instrument. Image credit: Max Millar-Blanchaer

Figure 6.9. An image of a commercially available digital holographic microscope is shown above. This particular model is from Lyncée Tec. Image credit: Lyncée Tec

Properties of a typical, commercially available digital holographic microscope:

- **Wavelength (fixed):** 632.8 nm
- Quantitative phase measurement in either reflection or transmission (double pass)
- **Field of view:** Dependent on the exact optical configuration.

Custom Built

Commercially available phase measuring interferometers can be very expensive, prohibitively so in a research department. However, quantitative phase measurements can be performed with a DHM using readily available and affordable off-the-shelf components. Indeed, we have built such instruments for the visible and infrared portion of the spectrum (Wallace 2026).

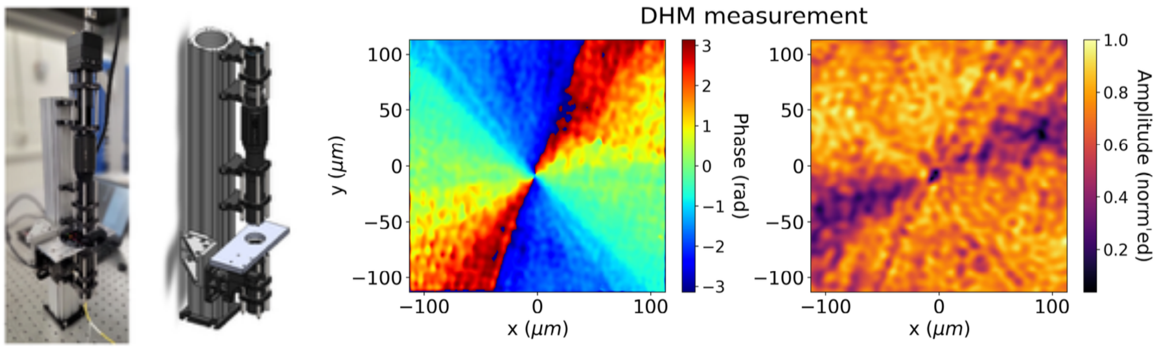


Figure 6.10. An image of a custom -built digital holographic microscope (left). This particular model is for use in the short-wave infrared. The image on the right shows the measured amplitude and phase measurement from this DHM. Image credit: J. Kent Wallace, et al., "Short-wave infrared digital holographic microscope: application to optical characterization of phase and amplitude of metasurface optics," Appl. Opt. 65, 697-704 (2026)

Properties of a custom built available digital holographic microscope:

- **Wavelength:**
    - Fixed: 632.8 nm, 1550 nm
    - Variable (with a tunable laser): Visible and short-wave IR
- Quantitative phase measurement in transmission (reflection is possible)
- **Field of view:** Dependent on the exact optical configuration.

7 Metasurface Applications Within High-Contrast Imaging (HCI)

Metasurfaces have a lot of potential for high-contrast imaging in coronagraphy, wavefront sensing, and polarization aberration correction. Their increased degrees of freedom enable the development of achromatic phase masks for coronagraphy and wavefront sensing, but also polarization correcting optical elements.

7.1 Early successes of metasurfaces in HCI

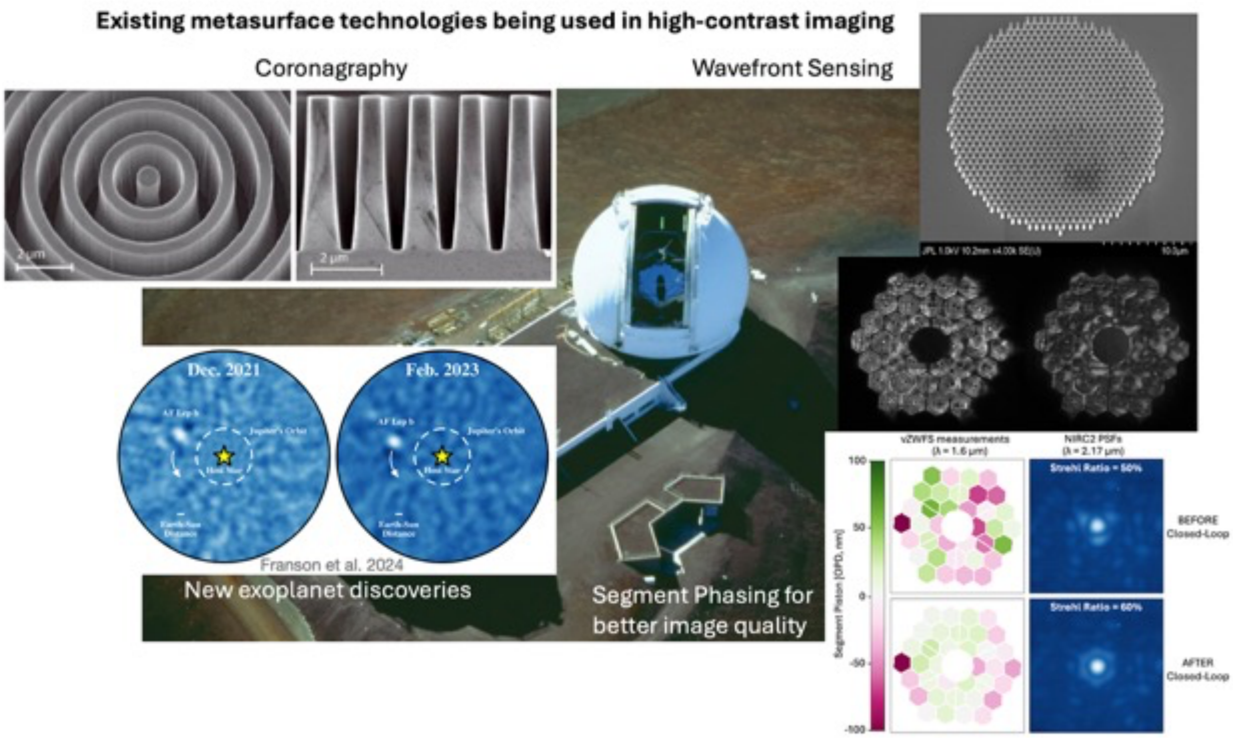


Figure 7.1. Metasurface optics have made advances in high-contrast imaging in two areas: 1) coronagraphic masks (left—annular grove coronagraphs) and 2) wavefront sensing and control (right—metasurface vector Zernike wavefront sensor). Image credits: Pontus Forsberg (annular groove coronagraph, upper left), Kyle Franson, (exoplanet image, lower left), Tobias Wenger (metasurface phase mask, upper right), Kent Wallace (vZWFS images, center right), Maïssa Salama (Keck phasing correction, lower right), W.M. Keck Observatory (Keck telescope, center)

7.1.1 AGPM: Solved chromaticity for coronagraphy

The first use of metasurfaces for high-contrast imaging dates back to the early 2010s, when the first annular groove phase masks (AGPM) were produced for thermal infrared applications. The AGPM concept consists of an optical vortex based on subwavelength gratings (Mawet et al. 2005), which can be locally seen as a 1D version of a metasurface, i.e., a metasurface based on continuous grating lines instead of discrete 2D nanoposts. The main incentive to use subwavelength gratings was (and is still) to provide an achromatic behavior over a wide bandwidth, which can go up to about 50% in the case of the AGPM (e.g., Vargas Catalán et al. 2016).

The manufacturing of the AGPM is based on the same techniques used to produce current mainstream optical metasurfaces, where e-beam lithography is used to produce a master pattern that is subsequently transferred into a dielectric substrate (here diamond) through plasma dry etching (e.g., Forsberg et al. 2024). Manufacturing is followed up by metrology (mostly scanning electron microscopy), to check some of the main grating parameters (period, line width), although the grating depth cannot be measured in a non-destructive way. This metrology step is followed by optical testing on a coronagraphic test bench, where we measure the peak rejection rate provided by the phase mask. Based on a series of measurements at various wavelengths within the wavelength bandpass of interest, one can then infer the actual grating depth and possibly proceed with re-etching to make the grating deeper or shallower, if an appropriate depth was not reached after the initial manufacturing (e.g., Vargas Catalan et al. 2016). It took about three years to go from the first manufacturing tests (started in 2009) to science-grade components (Delacroix et al. 2012), and almost one decade from the inception of the concept (2005) to the first astrophysical on-sky results (Mawet et al. 2013, Absil et al. 2013).

A dozen science-grade AGPMs have been produced since the early 2010s, reaching contrasts up to 10^{-6} at an angular separation of a few λ/D across individual infrared atmospheric windows ranging from the K band (2.0 μm) to the N band (up to 13 μm). They have been retrofitted into high-contrast imaging cameras installed at some of the world's largest ground-based optical telescopes, such as the Very Large Telescope, the W.M. Keck Observatory, and the Large Binocular Telescope. The on-sky exploitation of these coronagraphs has resulted in more than 50 refereed publications over the last 10 years, including the discovery of new exoplanets and circumstellar disks (e.g., Franson et al. 2023).

The AGPM success story is a good example of how astronomers and engineers work together to go from a concept to an actual science-grade device through a series of iterations between design, manufacturing, metrology, and testing. It has also allowed the astronomy community to gain a significant amount of experience in the system-level integration and on-sky operation of optical vortices for high-contrast imaging, as well as in the analysis of the data collected during those observations.

When looking back on how the AGPM was taken from the invention on a drawing board all the way to installation on the telescopes, the most obvious way to be quicker would be to reduce the time from concept (2005) to initial fabrication attempts (2009). A key lesson learned is to align component fabrication with system-level performance simulations at the very beginning of the project. In the case of the AGPM this has allowed for re-optimizing the depth of the grating during fabrication to match actual line width and development of processes to tune the grating depth, both vital to producing high performing diamond AGPMs. Knowing early how much deviation from the optimal design can be tolerated while still having a science-grade performance is also important for setting a clear target for process development. Moreover, it is vital to be able to start testing the performance of the first fabricated prototypes as soon as possible. Pausing fabrication while awaiting testing risks further delays as machines break down or are locked up in other projects. Of course, it is also important to have funding to form groups in the relevant fields to be able to perform the work needed. In the AGPM project, we had groups with competences in design and optical simulations, micro- and nanofabrication and materials science, optics and installation and on-sky measurements in ground-based telescopes. The success in this project can be attributed, to some extent, to having teams with similar aims, but also having good collaborations between the groups in the project. It is of course difficult to know in advance how the different teams will work together, but it is still important to bear this in mind when starting a long-term project.

7.1.2 Vector Zernike Wavefront Sensor

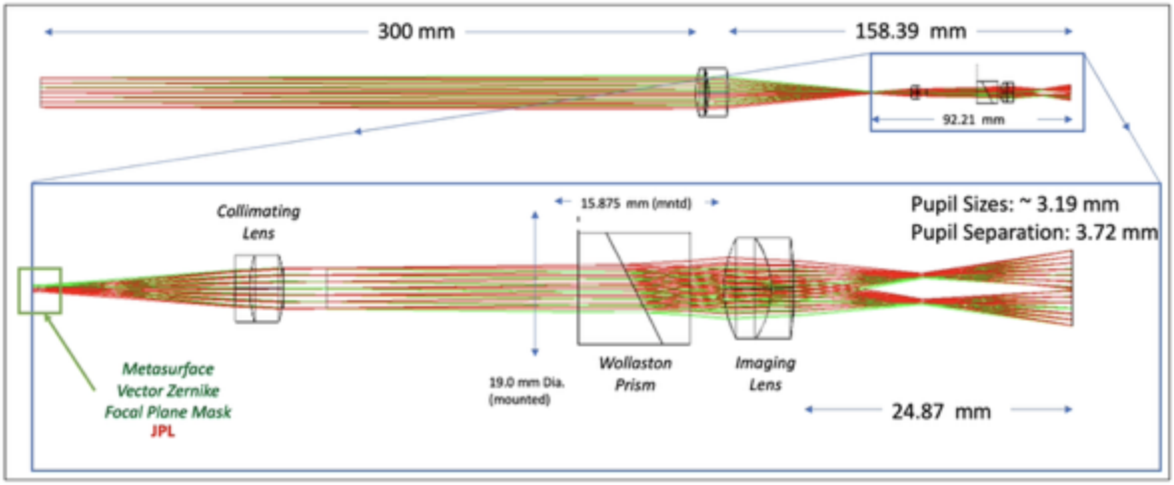


Figure 7.2. The optical design is illustrated above on two different spatial scales. The top image shows the beam path from the input pupil on the left, and the focal plane on the right. The lower image is a zoom-in of the beam path from the metasurface Zernike focal plane mask through the polarization optics and to the detector plane where the pupils are reimaged. Image credit: J.K. Wallace, et al. "Architecting, Implementing and Observing with a Metasurface vector Zernike wavefront sensor on the Keck Telescope," AO4ELT7, ONERA, Jun 2023, Avignon, France. hal-04402835 10.13009/AO4ELT7-2023-088. hal-04402835

The vector Zernike wavefront sensor (vZWFS) installed in 2022 in the Keck Planet Imager and Characterizer (KPIC) on the Keck II telescope is the first on-sky use of a metasurface for wavefront sensing (Salama et al. 2024a). The Zernike wavefront sensor uses the phase contrast technique (Zernike 1934) to convert phase variations into intensity variations that can be imaged. A mask (typically made of glass) is placed in the focal plane, which contains a dimple with diameter matching the size of the point spread function (PSF) core, offsetting its phase by π/2, causing it to act as a reference wavefront interfering with the rest of the PSF, thus encoding the phase variations in the resulting pupil image (N'Diaye et al. 2013). The Zernike wavefront sensor is one of the most sensitive wavefront sensors, but has very limited dynamic range (Guyon 2005, Chambouleyron et al. 2021). The metasurface vector-Zernike mask (Wallace et al. 2023) enables imposing two different phase shifts (e.g. ± π/2), simultaneously, to two orthogonal light polarizations, which are then split by a Wollaston prism, resulting in two pupil images (Doelman et al. 2019). This design doubles the dynamic range compared to the scalar ZWFS (see Figure 4 in Salama et al. 2024a). In addition, the ZWFS is sensitive to phase discontinuities, as opposed to the more commonly used Shack-Hartmann and modulated pyramid wavefront sensors, which measure a wavefront's gradient. The vZWFS on Keck has enabled demonstrations of

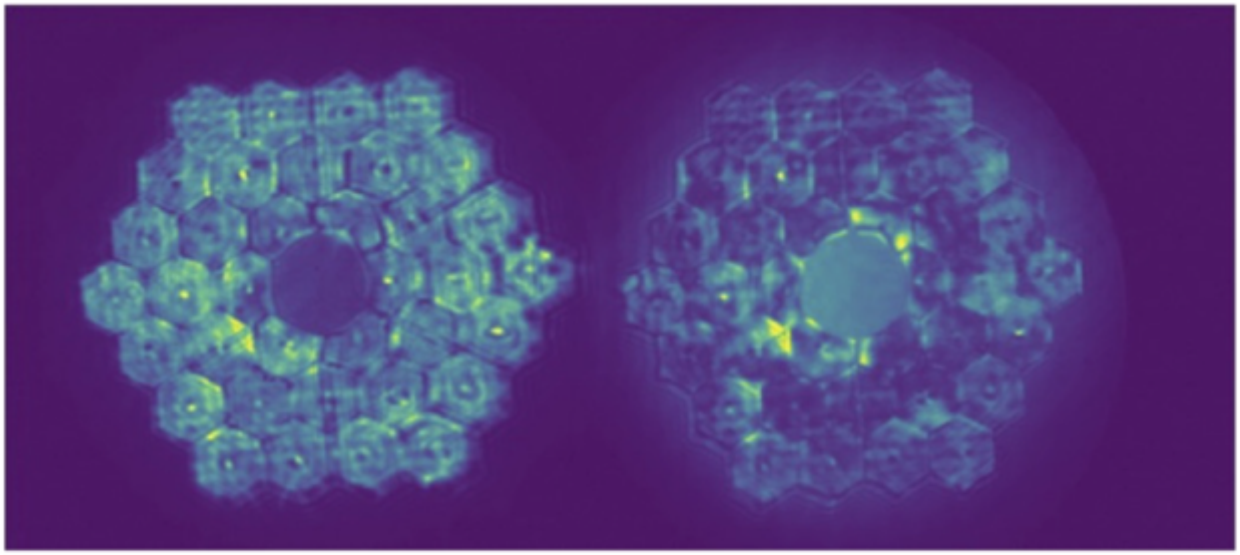


Figure 7.3. A raw image of the pupils viewed with the starlight centered on the focal plane mask is shown. This light passes through the entire optical system, and it senses phase errors immediately before the fiber injection unit of KPIC. Images similar to this were used to make fine adjustments of the primary segment piston in order to improve the system Strehl. Image credit: J. Kent Wallace, et al., AO4ELT7, hal-04402835

closed-loop control of the primary mirror segment pistons, resulting in an improvement in the science image quality (Salama et al. 2024a, 2024b).

This is particularly relevant to next-generation space missions (e.g., HWO) and ground-based extremely large telescopes (ELTs), which will all have segmented primary mirrors. The ZWFS is a key candidate wavefront sensor for HWO, where maintaining closed-loop segment phasing in parallel with science observations will be crucial to mitigate segment co-phasing errors, which are particularly detrimental to reaching 10^{-10} contrasts (Trauger & Traub 2007, Nemati et al. 2020, Pueyo et al. 2022).

7.1.3 Dual-purpose coronagraph: First step integrating starlight suppression and phase sensing

A dual-purpose WFS+coronagraph is currently being tested in the lab (Wallace 2025). The mask consists of a Zernike dimple in reflection and a Lyot coronagraph in transmission. A metasurface mask has been fabricated and implemented in the high-contrast testbed at NASA's Jet Propulsion Laboratory (JPL).

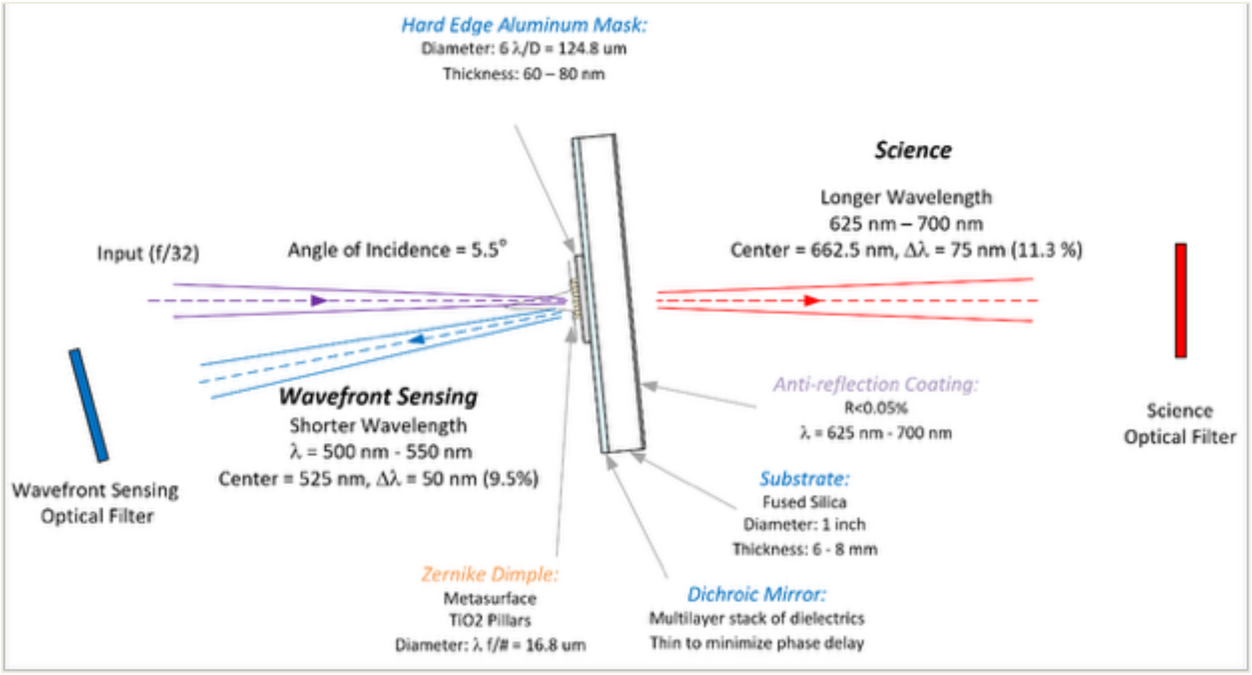


Figure 7.4. The dual-purpose focal plane mask enables wavefront sensing in reflection (blue light), simultaneous with starlight suppression in transmission (red light) with a minimum of non-common path. Image credit: J. Kent Wallace, et al., Proc. SPIE 13627, 1362714 (18 September 2025)

7.2 Coronagraphs

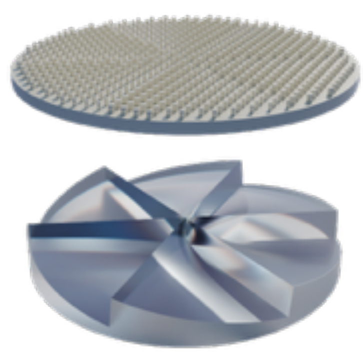


Figure 7.5. The metasurface vortex focal plane mask is at the top of this image, and the equivalent, scalar vortex focal plane mask is shown below. Image credit: W.M. Keck Institute for Space Studies/Robert Hurt (Caltech/IPAC-SELab)

7.2.1 Scalar vortex coronagraph

Science justification:

The future Habitable Worlds Observatory (HWO) requires 10^{-10} raw contrast over 20% bandwidth. A charge-6 vortex coronagraph is one of the best candidates for HWO coronagraphy as it provides at the same time:

- High throughput
- Smaller inner working angle
- Robustness to low-order aberration
- 360° discovery space

Shortfall of current hardware implementations: Scalar versus Vector

While the vector vortex coronagraph (VVC), based on the geometric phase concept, has been used for more than 10 years in ground-based high-contrast imaging applications (Mawet 2010, Serabyn 2017, Kuhn 2016), this implementation has one major shortcoming: the two orthogonal polarizations of the incoming light see different signs for the helical vortex phase ramp. This simple fact presents serious engineering challenges in reaching 10^{-10} contrast in the two polarizations at once. A scalar vortex coronagraph (SVC), that imparts the same phase ramp to both polarizations, can overcome this drawback. This in turns allows one single deformable mirror (DM) to control the two polarizations, and/or avoid losing 50% of the planetary light in the case where the vector concept is used downstream of a polarizer.

For future space-based applications, it is interesting to consider a polarization-independent (scalar) vortex coronagraph instead of the polarization-dependent AGPMs or LCP masks. Currently there are two ways to implement a scalar vortex coronagraph using metasurfaces: Using periodically arranged nanoblocks in a square lattice with small gaps (König 23), or an aperiodic grid relaxing manufacturing constraints (Palatnick 2023). Both frameworks have shown designs down to 10^{-8} contrast and are currently being investigated further. Manufacturing efforts of these designs have been pushed in the past year(s), resulting in several prototypes and ongoing process development. A charge-2 metasurface scalar vortex has been demonstrated on both an all-Si platform and an a-Si on glass platform (Palatnick 2024a, Palatnick 2024b). Using measured phase and amplitude in contrast simulations, the predicted average contrast for these designs is ~10^{-6} across a 200 nm bandwidth in the near infrared (H band). A charge-4 a-Si on glass metasurface scalar vortex is currently in fabrication development, and a TiO_2 on glass metasurface scalar vortex for the visible has been designed for near-term fabrication efforts.

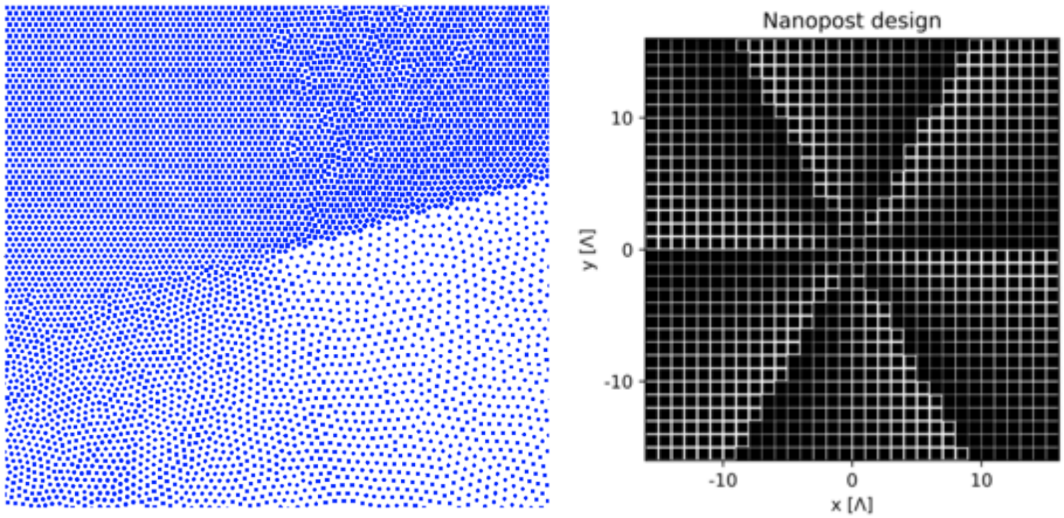


Figure 7.6. (Left) Metasurface scalar vortex phase masks designed for contrasts on the order of 10^{-8} across 20% bandwidth using periodic and (Right) aperiodic metasurface patterns. Image credit: left, König, and right, Palatnick.

Another interesting concept for future high-contrast imaging efforts is the phase knife nuller enabling smaller inner working angles and therefore enlarging the accessible parameter space for future high-contrast imaging efforts. This phase mask consists of two discrete regions with a relative π phase shift and is an excellent concept to demonstrate high-precision metasurface phase masks because it requires a single metasurface region with constant parameters, which is easier to achieve than a spatially varying metasurface. Manufacturing of this kind of mask has been started, resulting in a series of small-scale prototypes (König 2024). These prototypes have been tested for their phase response in a digital holographic microscope (DHM) showing consistent results with simulations. The accuracy of these measurements is limited by the current metrology techniques (see Section 6.1.4.3 Digital Holographic Microscopy). Manufacturing a full-scale phase knife nulling mask is currently underway, after which it will be tested in the Infrared Coronagraphic Testbed (IRCT).

Various other coronagraphic phase mask designs exist, many of which could benefit from a metasurface implementation. Segmented phase masks like the four-quadrant phase mask are an example requiring simple phase patterns similar to the phase knife that could be readily implemented with the same design and manufacturing tools. An azimuthal cosine phase mask could demonstrate smooth phase variations, avoiding the sharp phase jumps of a scalar vortex or segmented phase masks and requiring a smaller phase coverage.

7.2.2 Forked-grating vortex

Current implementations of scalar vortex coronagraphs (SVCs) achieve the desired phase ramp on the wavefront by either varying the surface height of a dielectric substrate or using metasurfaces to vary the effective index of refraction with respect to the azimuth. These approaches are limited by their chromatic leakage or are challenging to manufacture. An alternative approach is spatially separating the leakage from the vortex beam using a concept like the double-grating VVC implementation (Doelman 2020, Doelman 2023). The same concept will be used for the forked grating vortex (FGV) phase mask, except it will utilize a scalar metasurface framework to imprint the desired phase and therefore work for both circular polarizations at the same time. One added benefit of this approach is that the light propagating in the zeroth order could be used for wavefront sensing.

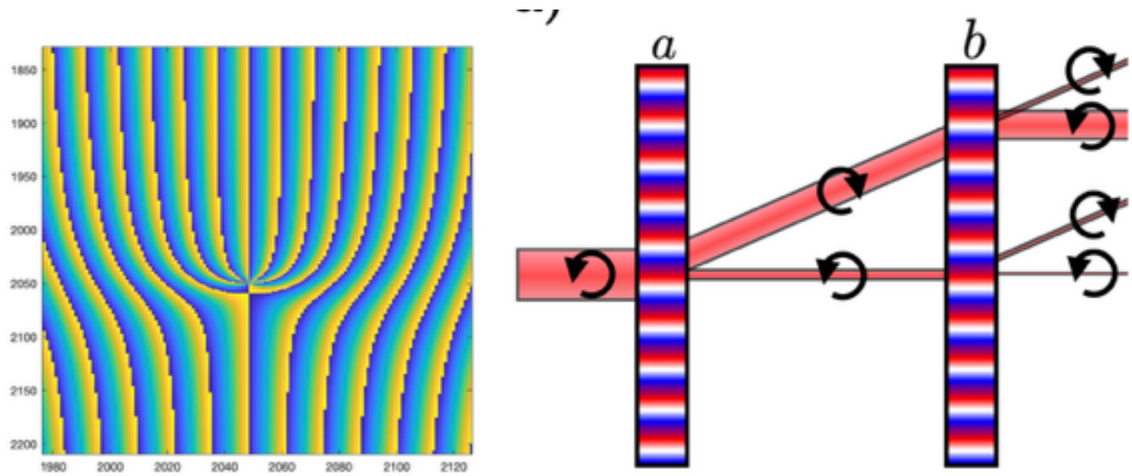


Figure 7.7. Phase profile of the forked grating vortex (left, charge-6 vortex) and working principle of the double-grating vector vortex coronagraph. The same concept will be used for the forked grating vortex (FGV) phase mask, except that it will then work for both circular polarizations simultaneously. Image credit: David Doelman, et al., PASP 132, 2020

The FGV consists of a forked phase grating resulting from the superposition of a phase grating and a vortex phase pattern. The vortex phase ramp and the phase grating should be fabricated together in one single optic. This could be achieved with a dielectric or a scalar metasurface implementation. For an ideal achromatic implementation of the phase profile, the incoming beam is diffracted into the first order and dispersed. A second recombination phase grating or prism is then used to bring the first order back on-axis. The finite distance between the two gratings results in a chromatic displacement of the beam, which is minimized when the distance of the two components is minimized. In case of an imperfect implementation of the forked vortex grating, part of the original beam does not acquire the phase pattern (neither grating nor vortex) and is therefore transmitted unaffected through the FGV. The second phase grating or prism then diffracts most of this leakage off-axis where it is intercepted by the Lyot stop. This allows spatially separating the pure vortex beam from the leakage beam. After the second phase grating, only a highly reduced leakage term is left propagating along the optical axis.

The remaining leakage term is still problematic for contrasts on the order of 10^{-10}. To reduce this leakage, an additional grating could be added (similar to the triple grating vortex coronagraph, Doelman 2023). The alternative to a second (or third) phase grating could be a thin prism, which overcomes the limitation of the grating efficiency, because it would diffract all the remaining leakage off-axis (100% efficiency of the prism). A prism could be implemented as a wedged substrate with the forked grating manufactured on one side of it. An issue with this concept could be the distance between the two interfaces: the forked grating and prism have to be as close as possible, and the thickness of the wedge is determined by the angle of the first diffraction order of the FGV.

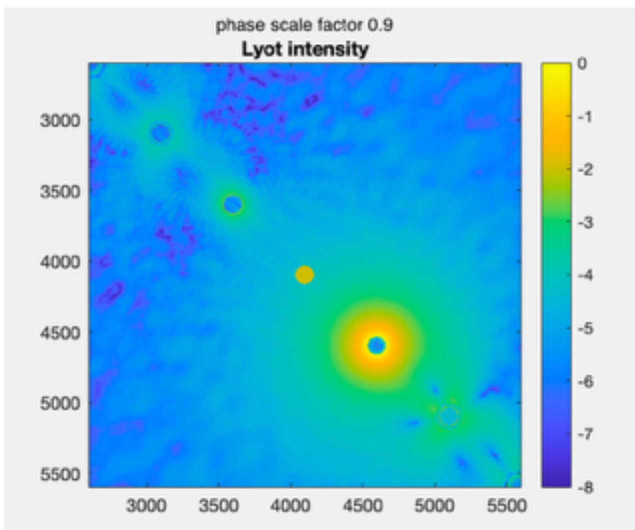


Figure 7.8. Higher diffraction orders carrying vortex modes of higher topological charge appear in the case of imperfect phase implementations. Most of the power is still diffracted into the first order, showing the characteristic ring of fire of a charge-2 vortex. Higher orders are much fainter and show the characteristic ring of fire of higher-order vortex modes. The individual orders carry pure vortex modes limited only by the rings of fire of higher modes leaking into their dark regions and the numerical noise floor of the simulation. Image credit: Lorenzo König

Another limitation of this concept could be the angle of diffraction of the higher vortex orders. Higher diffraction orders appear in the case of imperfect phase implementations, which result in intensity distributions characteristic of higher-topological-charge vortices in higher diffraction orders (Fig. 7.8). These orders leak into the dark region of the first-order vortex. The strength of this effect depends on the angle of the diffraction orders and the topological charge of the vortex (the light in the ring of fire of a vortex is more spread out for higher topological charge), as well as the quality of the phase gradient implementation (i.e., the number of metasurface blocks per grating period). The tradeoff will be choosing a grating period small enough to diffract higher orders sufficiently far away from the first order but providing a smooth enough phase gradient to send as little light as possible into higher diffraction orders.

The steps required to demonstrate this concept include building two components, potentially on the two sides of the same substrate. The first is the focal plane mask featuring the forked grating, and the second is the additional grating or prism required to redirect the spatially separated beam back on-axis, while sending the leakage term off-axis. Steps for achieving the FGV focal plane mask include tolerancing the phase and amplitude requirements for fabrication through simulations, developing an isotropic metasurface design for this FGV focal plane mask, fabricating a prototype within the required tolerances, performing metrology and characterization of the prototypes, and testing its efficiency to spatially separate the leakage term.

The metasurface design could be done with the forward approach or with inverse design. The forward approach would include meta-atom design through either: propagation phase (simple isotropic meta-atoms, monochromatic, polarization-insensitive), geometric phase (simple anisotropic meta-atoms, broadband, polarization-sensitive), or dispersion-engineering (anisotropic, meta-atoms, broadband, polarization-sensitive/insensitive). Figure 7.9 shows an example of a metasurface vortex grating realized with a vectorial metasurface implementation.

The designs can be fabricated with standard metasurface manufacturing approaches. This includes atomic-layer deposition of TiO_2 nanopillars on a substrate, etching of silicon, glass or other materials, or etching of structures in a single substrate, including silicon, diamond, and others.

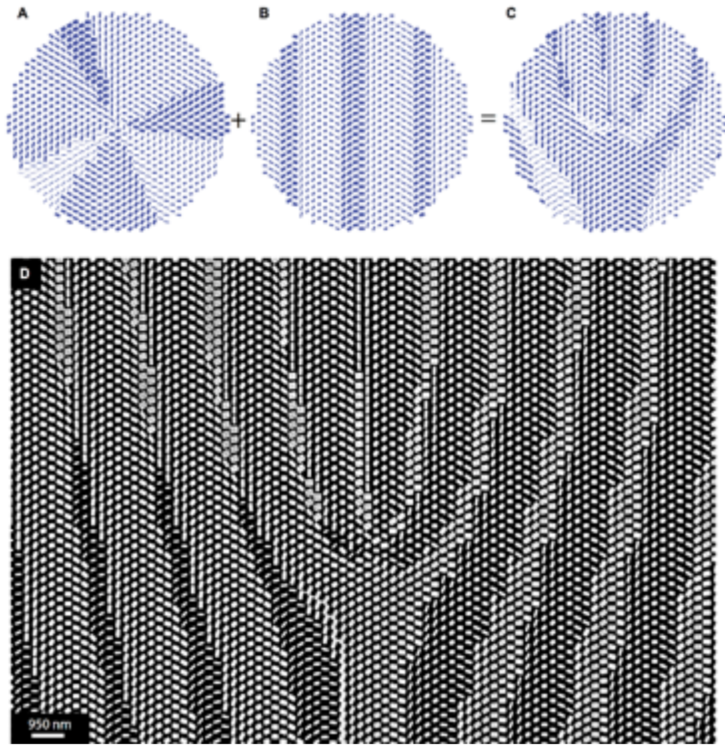


Figure 7.9. Example of metasurface vortex grating, polarization-sensitive. Image credit: Devlin et al., 2017, Science Vol 358, Issue 6365 pp. 896-901: Supplementary Material.

Regarding metrology and testing, the two components should first be measured separately, and possibly in parallel. To evaluate the forked grating alone, a simple imaging setup would be required to measure the off-axis rings of fire (monochromatic and eventually broadband). In addition, standard vortex focal plane mask characterization would be performed including measuring the size of the central defect (where the majority of the PSF lands) and large-scale phase and amplitude characterization and comparison with model or intended design. Another evaluation would be required to measure the recombination grating or prism alone. This would also require a broadband and tunable source that could be set up with a standard vortex to mimic the output of the forked grating focal plane mask.

Ultimately, it’s necessary to evaluate this technology in a relevant environment. High-contrast coronagraphy testbeds would be modified to incorporate both components for testing the final starlight suppression capability of this technology. Starting with in-air testbeds, such as the In-Air Coronagraph Testbed (IACT) at JPL or the High Contrast Spectroscopy Testbed (HCST) at the California Institute of Technology, raw contrast measurements would be conducted in both narrowband and broadband, first without and then with wavefront control. Upon achieving contrasts at the limit of in-air testbeds, dark hole digging experiments in vacuum testbeds would follow.

Some foreseeable challenges include addressing the leakage and separation between high-order terms following the forked grating vortex mask. The leakage from higher-order terms would need to be computed while considering the ring of fire, as explained above. Additionally, internal reflections would need to be managed for the recombination grating/prism.

7.2.3 Phase-apodized-pupil Lyot coronagraph

Another potential future application of metasurfaces is for apodization in pupil-plane coronagraphs. One particularly interesting example here is the phase-apodized-pupil Lyot coronagraph (PAPLC, Por 2020). This coronagraph uses an apodizer in the pupil plane of the telescope, an amplitude knife edge as the focal plane mask, and a Lyot stop to clean up the residual diffraction. This results in a high-throughput coronagraph operating at small angular separations. The original design calls for an achromatic phase apodizer to achieve inherently achromatic coronagraphic performance. Achromatic phase plates can be produced by multi-twist liquid-crystal retarders (Doelman 2017) at the expense of filtering out half the light.

Current experiments are performed by using the two DMs already included in space-based high-contrast imaging systems (Soummer 2022). While DM-based apodization is sufficient for contrasts at the 10^{-10} level in broadband light, the resulting coronagraph will not be robust to low-order aberrations (Por 2023). Future experiments are envisioned to use a combination of a compound prism and DM apodization. While these provide in simulation the required contrast and robustness, in practice, transmissive optics should be avoided as they tend to produce ghosts.

A reflective apodizer made from metasurfaces could solve this issue. Metasurfaces can produce achromatic phase aberrations, albeit with small chromatic deviations. It is an open question whether these chromatic deviations can be mitigated to the level required for 10^{-10} contrast, or if they can be corrected by the DMs in the setup.

7.3 Wavefront sensing

7.3.1 Context: Wavefront sensing in high-contrast imaging

Wavefront sensing and control are foundational technologies for high-contrast imaging. Coronagraphs are designed to suppress diffracted starlight to reveal faint exoplanet companions, but their performance is highly sensitive to optical aberrations. Even extremely small wavefront errors create residual speckles in the science image that can mimic or obscure planetary signals. As a result, achieving and maintaining deep contrast requires continuous measurement and correction of optical aberrations throughout the observing system.

Wavefront sensors provide the measurements needed to drive deformable mirrors, segmented primary mirrors, and other active optical components in a closed loop. In ground-based observatories, wavefront sensing is used primarily to correct atmospheric turbulence and telescope-induced aberrations. In space-based observatories, where atmospheric turbulence is absent, wavefront sensing is instead used to maintain the exceptional optical stability required for coronagraphy, including the correction of thermal drifts, structural disturbances, and segment phasing errors.

Future observatories such as the Habitable Worlds Observatory will require wavefront stability at the level of approximately 10 picometers over timescales of several minutes. Achieving this level of performance will require highly sensitive wavefront sensing architectures with minimal non-common-path aberrations between the sensor and the science channel. Metasurface optics offer unique opportunities in this area because they enable compact optical designs, polarization-dependent functionality, and the possibility of integrating wavefront sensing and coronagraphy into a common optical element. These capabilities have the potential to improve sensitivity, reduce system complexity, and increase overall observatory stability.

7.3.2 Space-based

Achieving 10^{-10} contrast is needed to directly image Earth-like exoplanets over very long science observation timelines. Therefore, actively stabilizing the contrast is essential, and wavefront sensing and control is a key aspect of space-based high-contrast imaging, such as for HWO. The nominal wavefront stability level and time scale required is 10 picometers over 10 minutes. Fortunately, a meta-based Zernike wavefront sensor has been developed and can achieve this level of performance. The goal here is to leverage this early development to reduce system complexity and, with the application of a “vector” version of the technology, dramatically increase sensitivity.

7.3.3 Ground-based

Ground-based, high-contrast instrument fundamental limits are set by capabilities of the associated wavefront sensing. A distinct challenge for ground-based systems is to sense effects of rapidly varying atmospheric turbulence, which are large in magnitude, along with smaller effects due to telescope-induced aberrations. The challenge is to provide enough sensitivity for the measurement over the requisite dynamic range while collecting enough photons during each measurement cycle for adequate signal-to-noise ratio. A meta-based sensor could realize significantly improved photon efficiency over the large dynamic ranges of errors encountered in ground-based systems.

7.3.4 Opportunities for metasurfaces in wavefront sensing

7.3.4.1 State of the art

To our knowledge, there is only one example of a metasurface used for wavefront sensing in the context of high-contrast instruments: the vector Zernike WFS installed on KPIC (Keck II). This sensor is used as a second-stage WFS (fed with light from the Keck-AO system) and is used to perform two tasks:

1. Primary use: segment phasing. A significant Strehl ratio gain was demonstrated thanks to this device (Salama 2024)

2. Non-common-path calibration for the KPIC instrument: This WFS is leveraging the metasurfaces capabilities to introduce differential phase-shifts to the wavefront to increase dynamic range of this sensor.

The vector Zernike WFS on Keck shows that metasurfaces are mature enough to perform useful WFS tasks in high-contrast instruments on-sky. This demonstration represents an important milestone because it moves metasurface technology beyond laboratory validation and into operational astronomical instrumentation. On-sky deployment provides confidence that metasurface optics can survive integration, alignment, calibration, and routine observing conditions while delivering scientifically useful performance.

Dual purpose mask: use of metasurfaces to produce a Zernike WFS in reflexion—being tested in the lab.

7.3.4.2 Metasurfaces to enable new ways to design WFS

Metasurfaces may not yet be ready to be used to build state-of-the-art coronagraph, however, they can be used in the context of wavefront sensing for which constraints are much less stringent.

More sophisticated WFS design allow an increase in dynamic range (e.g., vector ZWFS) or sensitivity (Bi-O edge, c Vérinaud 2024a, 2024b) at the cost of more complex design. Metasurfaces can enable simpler implementation of such concepts by transferring the complexity to the metasurface rather than the optical system. Some concrete examples are given in the next section.

Metasurfaces can be used to build even more complex design in the framework for dual purpose masks.

7.3.5 Challenges

WFS have to work in extended bandwidth (typically ~10/20%), however, metasurfaces suffer from inherent chromaticity because they are grating at first order, which makes it challenging to implement achromatic OPD ramps. This issue can be solved, either by: (i) adding chromatic compensators, or (ii) the more appealing but less mature approach of using multi-layer metasurfaces.

7.3.6 Examples of potential applications

All applications listed below have something in common: they leverage metasurfaces to perform multi-task operation on the EM field depending on polarization states.

7.3.6.1 Space-based and ground-based: Improving the vector Zernike WFS

Efforts are now underway to combine the meta-Zernike wavefront sensor with a focal plane mask to achieve high-contrast imaging contemporaneously with precision wavefront sensing with a minimum of non-common-path errors and reduced system complexity. This so-called "dual-purpose," meta-based focal plane mask has been demonstrated within a scalar Zernike configuration of a classic Lyot coronagraphic mask.

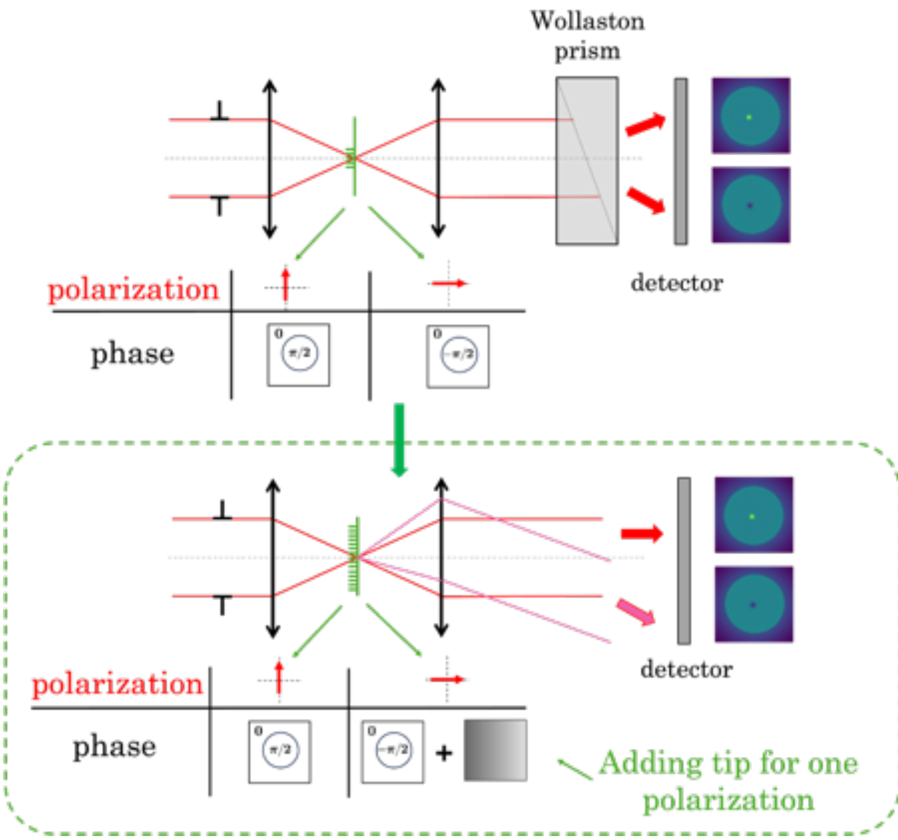


Figure 7.10. Improving the vector-ZWFS by encoding the polarization splitting operation directly in the focal plane mask. Image credit: Vincent Chambouleyron

In parallel, we are making a hybrid design that uses a vector Zernike dimple along with a scalar Lyot focal plane mask. The vector Zernike wavefront sensor has already been demonstrated independently. First, it has been used for performing fine phasing of mirror segments on the Keck II telescope—an on- sky, end-to-end system demonstration. Second, it has been demonstrated in the lab with a segmented deformable mirror. We are in the process of characterizing the performance of this dual-purpose vector Zernike and classic Lyot coronagraph and will write up the results shortly.

Ground-based: meta-pyramid-based WFS and meta-based implementation of the Bi-O edge sensor

A pyramidal WFS is a combination of four meta-enabled polarization linear gratings. However, significant challenges exist in the development of this meta element for expanded optical bandwidth and efficiency.

WFSs used today have high sensitivity that is efficient in terms of encoding phase into intensities, but are not reaching optimum performance. Indeed, they need to cope with large aberrations, which put constraints on their design and reduce their sensitivity. Hence, ground-based WFS and AO system can still be improved to optimize this trade-off between sensitivity and dynamic range.

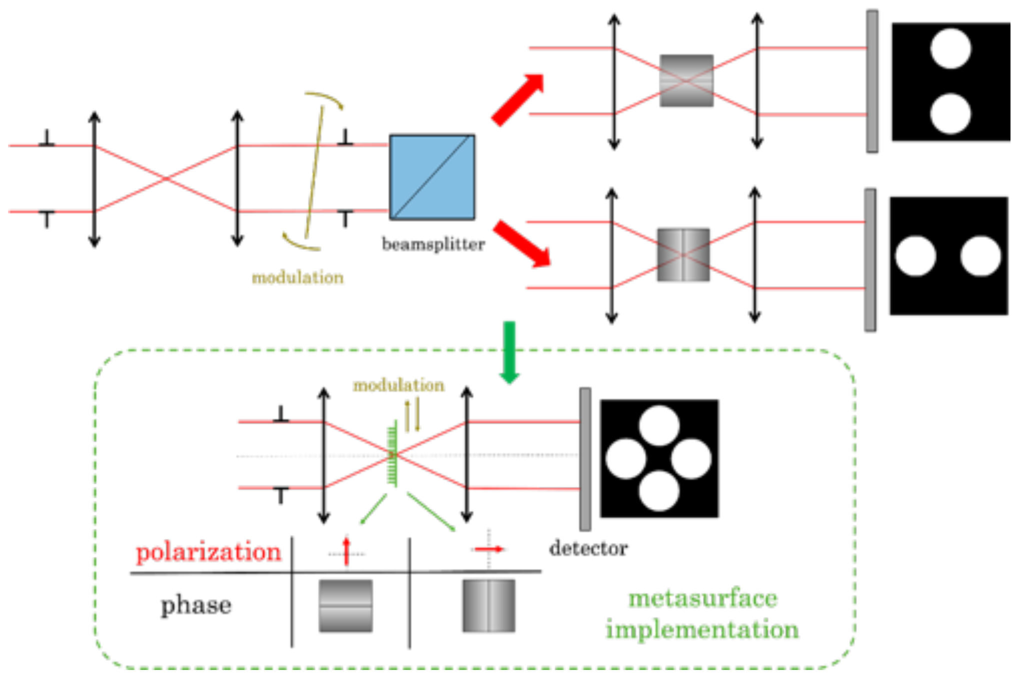


Figure 7.11. A possible implementation of the sharp-biO-Edge with metasurfaces. Image credit: Vincent Chambouleyron, adapted from Vérinaud, C., et al., A&A, vol. 682, 2024. DOI:10.1051/0004-6361/202346660

Measuring for the turbulence but also for the telescope-induced aberrations can be challenging because their natures are different. That point, added to the trade-off between dynamic range and sensitivity and combined with the issue of non-common path aberrations, leads to what may be a new standard for next-generation ground-based high-contrast instruments: multi-WFS systems combining a first-stage WFS with large dynamic range and second-stage WFS that is much more sensitive than the first stage and similar to the one that can be used for space-based applications.

Non-linear curvature wavefront sensor (nlCWFS)

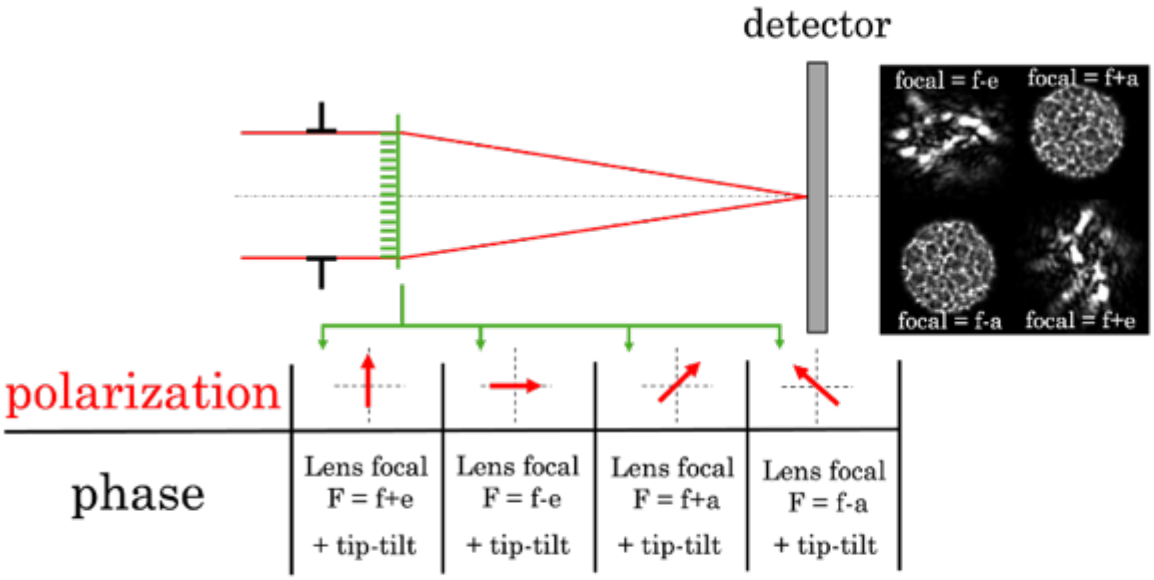


Figure 7.12. Imaging different propagation planes on the same detector with only one metasurface optic application for the nlCWFS. Image credit: Vincent Chambouleyron (center), Olivier Guyon, PASP, 2010 (right)

Space-based and ground-based: Dual-purpose masks

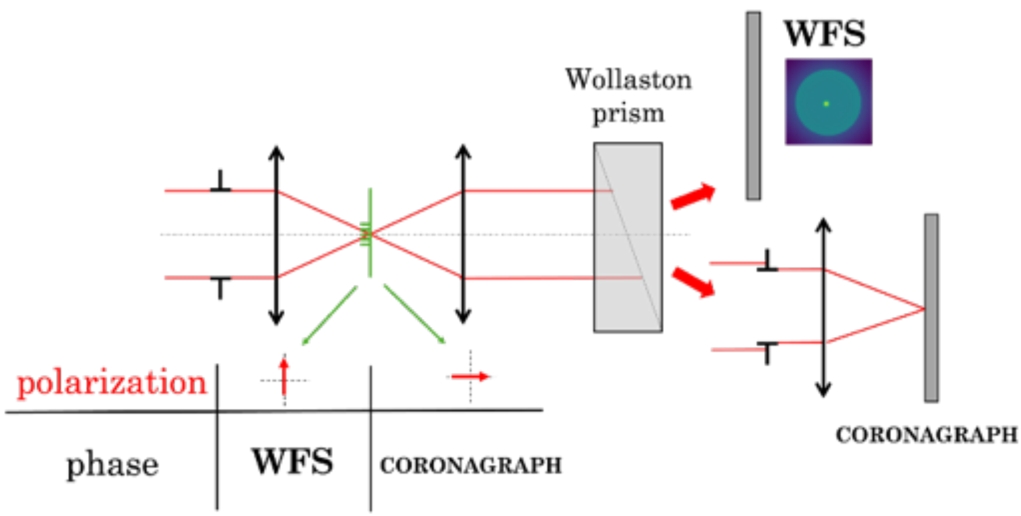


Figure 7.13. A dual-purpose mask made of metasurfaces, one polarization goes through the wavefront sensing function while the orthogonal polarization goes through the coronagraphic function. Image credit: Vincent Chambouleyron

7.4 Polarization correction

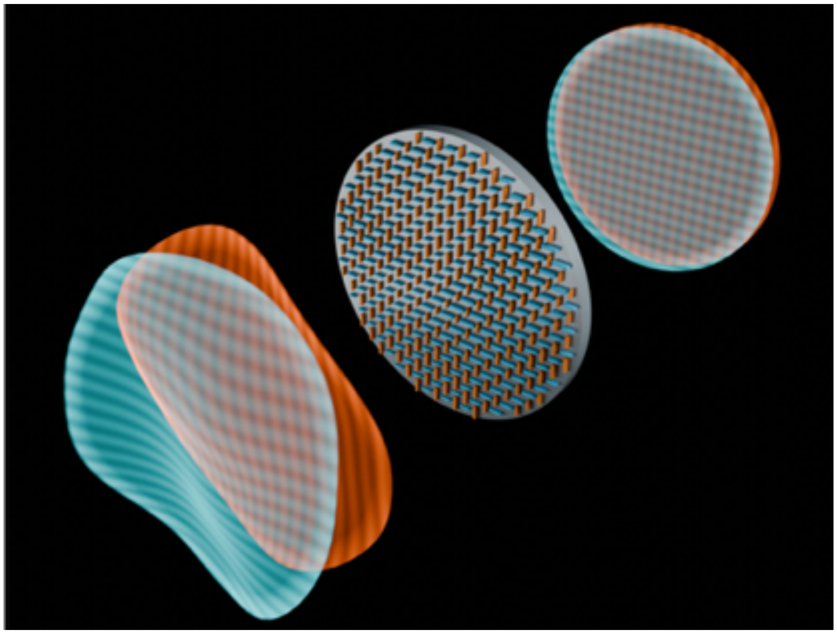


Figure 7.14. This image illustrates the concept of phase aberration correction using metasurface optics. In the lower left, the two orthogonal polarization states are shown in blue and orange. They each have separate and independent wavefront errors. After passing through the metasurface mask, the polarization-dependent aberrations have been corrected. Image credit: W.M. Keck Institute for Space Studies/Robert Hurt (Caltech/IPAC-SELab)

7.4.1 Background and justification

Polarization aberrations are caused by reflections from mirror surfaces at non-normal angles of incidence. The post-reflection wavefront is distorted in a way that is polarization-dependent and cannot be corrected with using a deformable mirror control approach. To achieve 10-10 contrast, we must correct polarization-dependent amplitude and phase errors to the same stringent levels as scalar errors (~10 pm phase, 0.01% amplitude) using deformable mirrors. Managing these polarization errors remains a key barrier.

A metasurface corrector optic can cancel polarization aberrations induced by the telescope, a process analogous to how a deformable mirror cancels scalar wavefront errors. Metasurfaces are ideal for this application because they enable control of polarization aberrations independently at every location within the pupil without significantly affecting the optical path. This corrector can use complex meta-atom shapes to perform correction with chromatic tailoring to match the dispersion of polarization aberrations induced by the telescope. Meta-atom heights can be scaled down to increase the precision of correction within a given fabrication tolerance, at the cost of correction range. For this application, this is an extremely advantageous trade-off, as the range of phase correction needed is on the order of milliradians. Broadband (20% bandwidth) metasurface design has been shown in metagratings using a meta-atom shape optimizer. This metasurface corrector can be placed on an existing fold mirror in the optical design, or on a transparent optic to minimize photon loss from additional reflections.

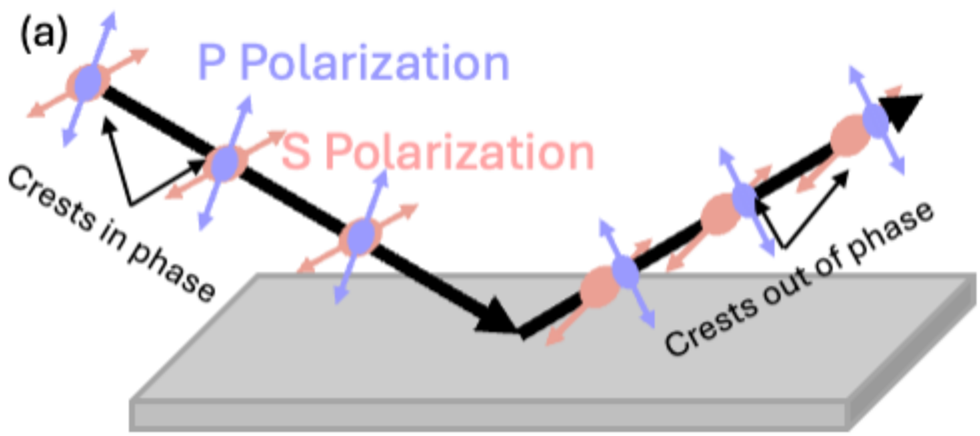


Figure 7.15. (a) A depiction of retardance of light induced by reflection. The black line represents the light ray; blue and red lines represent different polarizations with dots representing the relative locations of the wave crests. Light incident on the mirror with the same phase in the two orthogonal polarization states will exit with different phases. Image credit: Daniel Shanks.

A liquid-crystal-based element is non-ideal for myriad reasons. Liquid crystal dimensions for most applications are fixed to apply a range of 2-pi phase change, using geometric phase. There is no way to scale down the range of geometric phase correction to gain an advantage in precision, as can be done with propagation phase metasurfaces. Additionally, there is no way to align liquid crystals with complex shapes to gain additional degrees of freedom for chromatic control of correction. Correction using spatially varying thin films has also been approached in simulation, but results showed only limited improvements to contrast. This is because these thin films are ideally applied to the first few mirrors on the telescope, which are shared among all wavebands. Optimizing the thin film for such a broadband application is very challenging. To address polarization aberrations (PAs) using metasurfaces, we can design for each wavelength channel of the coronagraph individually.

Certain design choices can be made to minimize the effect of polarization aberrations in the HWO telescope without a metasurface corrector, but each trade-off comes at significant cost. For example, the use of a high-charge (6+) vortex mask makes the coronagraph insensitive to low-order wavefront aberrations, including low-order polarization aberrations, but has lower throughput and a higher inner working angle. The large inner working angle is particularly undesirable in the infrared. Additionally, some coronagraph designs use a polarizer to reduce the effect of polarization aberrations, at the cost of 50% light transmission. Finally, using a slow telescope speed with a long barrel reduces the angle of incidence on the primary mirror, reducing the induced polarization aberrations. This trade-off is the costliest, as a larger telescope is more expensive and challenging for stability.

A precision metasurface polarization aberration corrector could allow for relaxation of these design constraints. This includes relaxing the requirement of the optical design of the system to limit the angle of incidence as much as possible, allowing more flexibility in optimizing the system volume and mass, and reduced packaging constraints on the HWO telescope and coronagraph instrument. The telescope primary mirror would not be constrained to a specific focal ratio (f/#), thereby allowing larger angles of incidence and potentially reducing the distance between the primary and secondary mirrors, improving stability and compactness of the telescope. Additionally, we could remove the need for

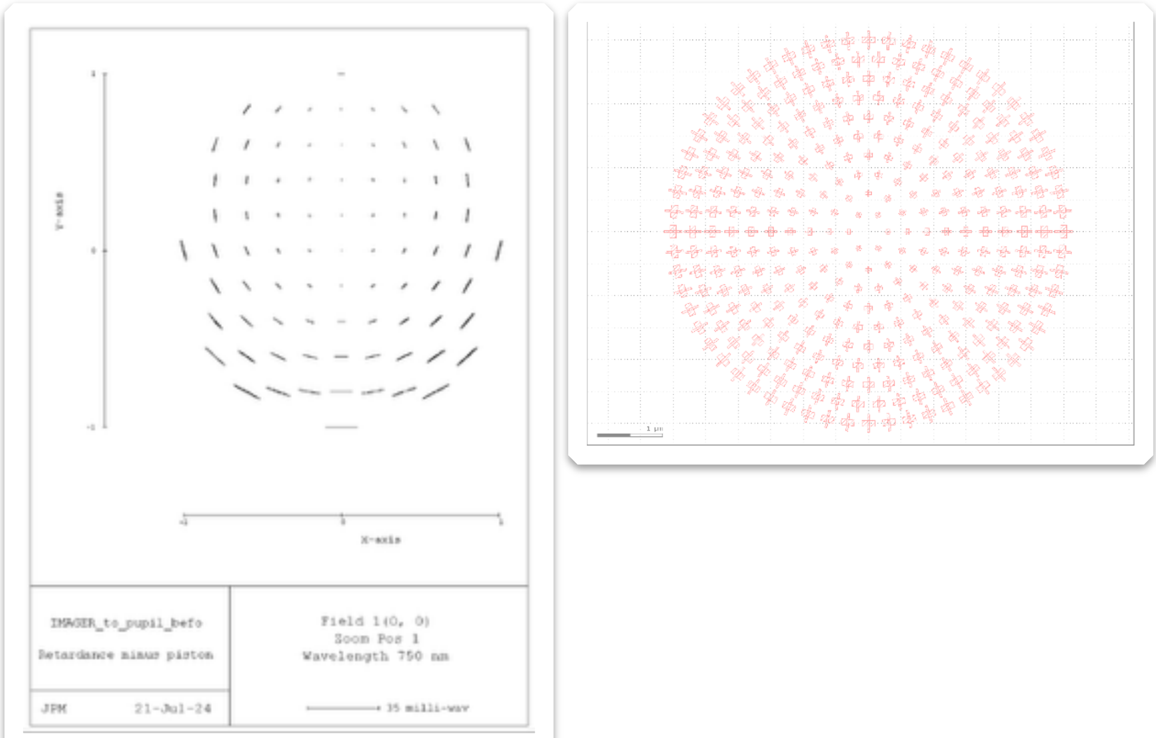


Figure 7.16. (Left) Example target retardance map from the Roman Coronagraph Instrument (CGI) optical model shows the phase retardance distribution across the pupil. Orientations of the lines indicate the eigenpolarization state, or direction of maximum retardance. Ellipticity of these lines represent ellipticity of the eigenpolarization. Length of the lines represents magnitude of retardance. (Image credit: Jim McGuire JPL) (Right) An example layout of a metasurface retardance corrector. Centrally located meta-atoms require less retardance thus are smaller; meta-atoms become physically larger and increase in retardance magnitude imparted as the area of the design approaches the edge of the pupil. Meta atoms are rotated to match the eigenpolarization of the target.

polarization split channels, allowing for increased throughput of the instrument.

Metasurface polarization aberration correctors are currently at technology readiness level 3 (TRL3), having demonstrated that a metasurface can apply an arbitrary polarization-dependent wavefront with some capability of chromatic tuning of the retardance response. The design can be scaled from UV to IR wavelengths by changing meta-atom size and metasurface periodicity.

The ultimate goal of a polarization-correcting metasurface is to design for the 10 pm–10 nm retardance range with arbitrary elliptical eigenpolarization, diattenuation reduction from 0.1% to 0.01% appropriate range of chromatic tunability of these properties, and operation in either reflection or transmission. Coverage of multiples of these categories will likely require multiple metasurface optics. Reduction of retardance by three orders of magnitude alone will likely require multiple cascading metasurfaces, with different designs to reduce the retardance by different orders of magnitude given realistic fabrication tolerances.

While a technology demonstration can be achieved with established metasurface fabrication techniques, novel fabrication will likely be needed to achieve high precision retardance correction over the many desired degrees of control. This could include low-index contrast metasurfaces, stacking multilayer fabrication processes, or multi-height grayscale lithographic designs. Experimentation with low-index contrast platforms translates to a low retardance range of structures which have—in element-wise simulation—achieved high precision and accuracy corrections. Multi-height metasurfaces can allow for increased range of chromatic tunability and retardance correction. Stacked metasurfaces on a transmissive substrate can allow for access to arbitrary elliptical eigenpolarizations.

To achieve TRL4, a demonstration of this polarization correction concept can be performed using a simple curved mirror in an optical benchtop. A basic reflective objective lens applies on the order of 1-10 nm of retardance. This magnitude of retardance is within the sensitivity range of most polarimetric metrology systems, including Mueller matrix imaging polarimeters. Thus, beginning with the large magnitude of retardance correction will provide easy proof of concept, without needing extreme precision of metrology or sensitivity. Then, a series of progressively finer-scaled corrections down to a 0.1 mRad retardance precision can be planned, using more advanced metrology systems.

The TRL5 technology demonstration will include polarization aberration generation and correction in the entrance path of a coronagraph. These polarization aberrations can be generated using a complex optical train of curved mirrors or using a metasurface. We can design such a system to mimic the polarization aberrations induced by the telescope design. The generation of these artificial polarization aberrations in a testbed can be used to confirm the effect of polarization aberrations on coronagraphic contrast for various coronagraph setups, including different mask types and with or without a polarizer in the optical path. These polarization aberrations can then be corrected by a metasurface to show the negation of effect of polarization aberrations directly in terms of contrast.

8 Integrating Metasurface and High -Contrast Imaging Models

Integrating metasurface and high-contrast imaging (HCI) modeling frameworks would be highly beneficial for advancing coronagraph technology. The two domains operate at vastly different scales. Metasurface design requires solving Maxwell's equations in the near-field to capture subwavelength physics. In contrast, HCI modeling uses Fourier optics to simulate system-level performance, propagating wavefronts between pupil and focal planes to predict contrast. This latter step almost always includes the corrective action of deformable mirrors (DMs), which can improve contrast by several orders of magnitude depending on the coronagraph type being modelled.

A unified modeling approach serves several key functions:

- **Performance Tolerancing:** It allows for robustly assessing how fabrication errors in a metasurface impact the final, DM-corrected contrast.
- **DM Interaction:** It helps in understanding how the inherent chromaticity of both the metasurface and the DMs interact, enabling co-design where DMs can more effectively correct for residual metasurface chromatic errors.
- **Rapid Prototyping:** A unified modeling approach that could also be used with measured data enables the evaluation of fabricated metasurfaces by using measured device properties (e.g., phase and amplitude maps) to model expected on-sky contrast with a specific mask. This creates a rapid feedback loop between fabrication and testing, ensuring that only the most promising components are advanced to time-consuming and expensive vacuum testbed experiments.

8.1 Existing frameworks for metasurface modeling

Accurately predicting the optical properties of a metasurface requires near-field electromagnetic modeling of its subwavelength structures. Simulating an entire device at the centimeter scale is computationally prohibitive. Therefore, the standard approach is to use the locally periodic approximation. Here, the optical response of each subwavelength element, or meta-atom, is simulated independently under periodic boundary conditions. This creates a library of meta-atoms, from which a designer can select the appropriate element for each point on the device to build the desired phase and amplitude profile. This method is efficient but assumes the properties of a meta-atom are not influenced by its neighbors. This is usually a highly effective approximation that can accurately model metasurface properties. However, its accuracy is highest where the final metasurface design is not changing much over short-length scales, and when the electromagnetic field is confined strongly within the metasurface pillars themselves. Stated in a different way:

the local approximation is less accurate when the metasurface is varying over short-length scales and when the electromagnetic field is poorly confined. Poor confinement can occur, for instance, when low-index materials are used.

Common numerical methods for these simulations include:

- **Finite-Difference Time-Domain (FDTD):** Solves Maxwell's equations on a discretized spacetime grid.
- **Finite-Element Method (FEM):** Solves the wave equation on a discretized mesh of the simulation volume.
- **Rigorous Coupled-Wave Analysis (RCWA):** Decomposes fields into a basis of Fourier modes.

Common software packages for metasurface modeling are:

- **FDTD:** Ansys Lumerical FDTD, Meep, OptiFDTD, Tidy3D
- **FEM:** Ansys HFSS, COMSOL
- **RCWA:** S^4, FMMAX, RETICOLO

8.2 Existing frameworks for high-contrast imaging modeling

HCI system modeling models propagating light through the entire instrument to predict starlight suppression. Fast Fourier transforms are typically used to propagate between pupil and focal planes. However, to accurately model components with fine features, like the singularity at the center of a vortex coronagraph, more advanced techniques such as matrix Fourier transforms on multiple spatial scales are required.

Furthermore, HCI models must include active wavefront control, typically using electric field conjugation (EFC) algorithms to command DMs. These algorithms create a "dark hole" of high contrast in the image plane by canceling out residual starlight, possibly including imperfections from the metasurface mask itself.

Common software packages for high-contrast imaging modeling:

- PROPER, HCIPy, POPPY, Prysm, dLux, FALCO

8.3 Path forward: An integrated approach

The current state-of-the-art involves a sequential workflow: near-field properties of a metasurface are simulated, and the resulting 2D phase and amplitude maps are imported as a static optical element into an HCI simulation package. For example, in HCIPy, a simulated map for a metasurface vortex, can be implemented using the

MultiScaleCoronagraph class, while a pupil-plane metasurface can be modeled with the Apodizer class. This approach is effective for tolerancing studies and initial performance predictions.

This sequential method, however, has limitations for multi-wavelength simulations. The chromatic response of a metasurface can be complex and non-linear, requiring separate 2D maps to be generated and simulated for each discrete wavelength. A relationship may not be easily parameterized as it can for conventional optics.

The true frontier is end-to-end, integrated modeling. This involves directly interfacing near-field electromagnetic solvers with far-field HCI propagators inside a single optimization loop. Instead of merely optimizing meta-atoms to minimize local phase errors, this approach would optimize the physical metasurface structure to directly maximize the final, post-coronagraph contrast. This would allow for the simultaneous co-design of the metasurface features and the DM control strategy.

Critically, this integrated framework should be agnostic to the source of the metasurface data. This is true even in a sequential workflow—the input 2D phase and amplitude maps can be generated from either simulations during the design phase or from metrology measurements of a fabricated prototype. Incorporating measured data is essential for accurately predicting the performance of real-world hardware, determining precise fabrication tolerances, and providing critical feedback to manufacturers.

9 Future Directions and Opportunities

While the application space for metasurfaces is vast, and the "killer application" for metasurfaces may yet be undiscovered, certain key applications can be foreseen. We see the following as ways to enable a measured, deliberate approach to demonstrating key metasurface technology for HCI soon.

Key Application	Advantages
Polarization aberration compensation	• Directly addresses HWO system-level engineering issues. • There are paths forward for doing this.
Vortex grating coronagraph	• Potential pathway for an achievable 10^{-10} focal plane mask. • Using metasurfaces, this can be made scalar (not polarization-dependent).
Vector dual-knife edge wavefront sensor	• Demonstrates key advantage of metasurfaces.
Higher sensitivity than pyramid WFS.	• Can be done quickly and demonstrated on-sky. • Has immediate benefit to the high-contrast community.
Maturing the software interface between metasurface design and coronagraph design	
Modest-sized, space-based coronagraph system	• Even modestly sized telescopes have apertures large enough to detect—with the right coronagraph—Earth-like planets around relatively close-by stars. • Leveraging recent dramatic reduction in cost and the rise of availability of launch capabilities creates interesting implications for timely and relatively low-cost technology demonstrations. • Such a mission could be seen as a precursor mission for HWO and demonstrate key technologies in coronagraphy and wavefront sensing and control. Small space missions can benefit dramatically from participation of NASA-external entities, where less administrative overhead leads to lower cost and development time, while only marginally increasing the risk profile.

10 Recommendations

The following key recommendations were distilled from the workshop discussions. They represent critical actions required to accelerate the maturation of metasurface optics for high-contrast imaging and ensure their readiness for future flagship missions like the Habitable Worlds Observatory (HWO).

Aggressively Advance Technology Readiness Levels (TRL) for HWO. While the HWO launch is decades away, the timeline for maturing new hardware from concept to flight-readiness is long and demanding. Metasurface optics are currently at a low TRL for space applications. A concerted, community-wide effort must be made to rapidly advance this TRL. While demonstrating the required contrast goals remains the highest priority, this requires more than lab-based contrast demonstrations; it demands a rigorous space-qualification campaign that proves device performance and robustness, as well as long-term stability under realistic operational conditions, including thermal-vacuum cycling and radiation exposure. Systematically retiring these technical risks is a critical step before this technology can be incorporated into the HWO baseline design.

Prioritize and Fund Key Technology Demonstrators. To focus development efforts, the community should prioritize a small number of high-impact technology demonstrations that address critical system-level needs. Based on the workshop findings, we recommend targeting the development of:

- A high-performance vortex coronagraph, which offers a direct and efficient path to achieving the required 10^{-10} raw contrast with a clear inner working angle.
- A polarization aberration corrector, an enabling technology that can mitigate a key source of contrast degradation, thereby relaxing stringent optical design constraints on the telescope and potentially reducing system complexity and mass.
- An advanced metasurface-based wavefront sensor, which promises to improve picometer-level sensing and control while simplifying the optical train, a crucial factor for instrument stability and throughput.

Funding these specific projects will create tangible pathfinders and provide the HWO mission with a toolbox of well-characterized components, increasing design flexibility and performance margin.

Leverage Ground-Based Telescopes for Rapid Maturation. The most effective way to accelerate technology maturation is to gain on-sky heritage. We strongly recommend the technology development plan formally incorporate partnerships with ground-based observatories to serve as proving grounds. These facilities offer the only practical, near-term path for validating metasurface performance against real-world atmospheric and telescope-induced aberrations. Beyond hardware validation, securing on-sky time will provide invaluable operational experience for our scientific and technical teams, test post-processing algorithms on real data, and forge the essential collaborations between the metasurface and astronomy communities that will be vital for HWO's success.

Integrate Science Requirements with Technology Development. To successfully navigate the critical TRL 4-6 maturation phases, technology development must be driven directly by scientific need, not just by its potential. This requires proactive and sustained communication between technologists, the broader exoplanet science community, and programmatic decision-makers. By clearly articulating how metasurface optics can solve specific, high-priority scientific challenges—and by ensuring performance requirements flow directly from the HWO science traceability matrix—we can build the broad support necessary for sustained investment. This ongoing dialogue ensures we are building the right hardware to answer the Decadal Survey's highest-priority science questions, a goal best achieved when scientists, engineers, and project leaders are engaged from the outset.

11 References

11.1 Chapter 1

National Academies of Sciences, Engineering, and Medicine. 2023. Pathways to Discovery in Astronomy and Astrophysics for the 2020s. Washington*, DC: The National Academies Press*. *https://doi.org/10.17226/26141*.

11.2 Chapter 3

Spohn, T., Roberge, A., Way, M.J. et al. Exo-Geoscience Perspectives Beyond Habitability. Space Sci Rev 222, 9 (2026). https://doi.org/10.1007/s11214-026-01265-y

Roman Space Telescope Exoplanet Imaging Community Data Challenge (https://www.exoplanetdatachallenge.com/Roman-dc)

Fang Shi, Kunjithapatham Balasubramanian, Randall Bartos, et al., “WFIRST Coronagraph Milestone 6 Final Report: Low Order Wavefront Sensing and Control”, (https://roman.gsfc.nasa.gov/science/archivedocs/WFIRST_CGI_Milestone6_Final_Report.pdf)

Ruslan Belikov, Christopher Stark, Nick Siegler, et al. "Coronagraph design survey for future exoplanet direct imaging space missions", Proc. SPIE 13092, Space Telescopes and Instrumentation 2024: Optical, Infrared, and Millimeter Wave, 1309266 (23 Aug 2024); https://doi.org/10.1117/12.3020614

Bertrand Mennesson, Ruslan Belikov, Emiel Por, et al. , “Current laboratory performance of starlight suppression systems and potential pathways to desired Habitable Worlds Observatory exoplanet science capabilities”, *Journal of Astronomical Telescopes, Instruments, and Systems*, vol. 10, Art. no. 035004, 2024. doi:10.1117/1.JATIS.10.3.035004.

11.3 Chapter 4

N. Yu, P. Genevet, M.A. Kats, J.-P. Tetienne, F. Capasso, and Z. Gaburro, “Light Propagation with Phase Discontinuities: Generalized Laws of Reflection and Refraction”, Science, vol. 334, no. 6054, p. 333-337, (2011). doi:10.1126/science.1210713.

Patrice Genevet, Federico Capasso, Francesco Aieta, Mohammadreza Khorasaninejad, and Robert Devlin, "Recent advances in planar optics: from plasmonic to dielectric metasurfaces," Optica 4, 139-152 (2017)

Molesky, S., Lin, Z., Piggott, A. Y., Jin, W., Vučković, J., & Rodriguez, A. W. (2018). "Inverse design in nanophotonics." *Nature Photonics*, 12(11), 659–670. https://doi.org/10.1038/s41566-018-0246-9

O. D. Miller, “Photonic Design: From Fundamental Solar Cell Physics to Computational Inverse Design”, PhD Dissertation, University of California, Berkeley. doi:10.48550/arXiv.1308.0212.

Zin Lin, Victor Liu, Raphael Pestourie, and Steven G. Johnson, "Topology optimization of freeform large-area metasurfaces," Opt. Express **27**, 15765-15775 (2019)

Zin Lin, Raphaël Pestourie, Charles Roques-Carmes, Zhaoyi Li, Federico Capasso, Marin Soljačić and Steven G. Johnson, “End-to-end metasurface inverse design for single-shot multi-channel imaging”, Optics Express, vol. 30, no. 16, OSA, p. 28358, (2022). doi:10.1364/OE.449985.

Raphaël Pestourie, Carlos Pérez-Arancibia, Zin Lin, Wonseok Shin, Federico Capasso, and Steven G. Johnson, "Inverse design of large-area metasurfaces," Opt. Express **26**, 33732-33747 (2018)

A. Arbabi, Y. Horie, A. J. Ball, M. Bagheri, and A. Faraon, “Subwavelength-thick lenses with high numerical apertures and large efficiency based on high-contrast transmitarrays”, Nature Communications, vol. 6, Art. no. 7069, 2015. doi:10.1038/ncomms8069.

A. Y. Piggott, J. Lu, K. G. Lagoudakis, et al. & J. Vučković, (2015). "Inverse Design and Demonstration of a Compact and Broadband On-Chip Wavelength Demultiplexer." *Nature Photonics*, 9(6), 374–377.

A. Y. Piggott, E. Y. Ma, L. Su, ... & J. Vučković, (2020). "Inverse-Designed Photonics for Semiconductor Foundries." *ACS Photonics*, 7, 569–575.

R. Trivedi, L. Su, J. Lu, M. F. Schubert, & J. Vučković, (2019). "Data-driven acceleration of photonic simulations." *Scientific Reports*, 9.

D. Sell, J. Yang, S. Doshay, R. Yang, & J.A. Fan, (2017). "Large-Angle, Multifunctional Metagratings Based on Freeform Multimode Geometries." *Nano Letters*, 17, 3752–3757. https://doi.org/10.1021/acs.nanolett.7b01082

E. H. Por, L. Pueyo, and R. Soummer, “Joint optimization of multiple optical planes in a stellar coronagraph”, in Space Telescopes and Instrumentation 2022: Optical, Infrared, and Millimeter Wave, 2022, vol. 12180, Art. no. 121805J. doi:10.1117/12.2630719.

J. A. Roberts, P.-H. Ho, S.-J. Yu, and J. A. Fan, “Electrically Driven Hyperbolic Nanophotonic Resonators as High Speed, Spectrally Selective Thermal Radiators”, Nano Letters, vol. 22, no. 14, pp. 5832–5840, 2022. doi:10.1021/acs.nanolett.2c01579.

J. Yang, D. Sell, and J. A. Fan, “Freeform Metagratings Based on Complex Light Scattering Dynamics for Extreme, High Efficiency Beam Steering”, Annalen der Physik, vol. 530, no. 1, p. 1700302, 2018. doi:10.1002/andp.201700302.

A. Arbabi, E. Arbabi, M. Mansouree, *et al.* Increasing efficiency of high numerical aperture metasurfaces using the grating averaging technique. *Sci Rep* **10**, 7124 (2020). https://doi.org/10.1038/s41598-020-64198-8

E. Bayati, R. Pestourie, S. Colburn, Z. Lin, S.G. Johnson, and A. Majumdar, "Inverse designed extended depth of focus meta-optics for broadband imaging in the visible", Nanophotonics, vol. 11, no. 1–4, Art. no. 431, 2021. doi:10.1515/nanoph-2021-0431.

J.E. Fröch, L. Huang, Z. Zhou, Z. *et al.* Full color visible imaging with crystalline silicon meta-optics. *Light Sci Appl* **14**, 217 (2025). https://doi.org/10.1038/s41377-025-01888-w

11.4 Chapter 5

L. J. Guo, "Nanoimprint Lithography: Methods and Material Requirements", Advanced Materials, 19: 495-513 (2007). https://doi.org/10.1002/adma.200600882

D. R. E. King, et al., Advanced Materials 1997, https://doi.org/10.1002/adma.19970090814

Vargas Catalán, E., Forsberg, P., Absil, O., and Karlsson, M., "Controlling the profile of high aspect ratio gratings in diamond", Diamond and Related Materials, vol. 63, Elsevier, pp. 60–68, 2016. doi:10.1016/j.diamond.2015.08.007.

Pontus Forsberg, Petri Karvinen, Samuel Ronayette, Markku Kuittinen, Olivier Absil, Lorenzo König, Christian Delacroix, Gilles Orban de Xivry, Jean-Christophe Barrière, Eric Pantin, and Mikael Karlsson, "Making the diamond vortex phase masks for the METIS instrument", Diamond and Related Materials, Volume 146, 2024, 111237, ISSN 0925-9635, https://doi.org/10.1016/j.diamond.2024.111237.

Joon-Suh Park. 2023. Nanofabrication platforms for metasurface optics: From prototyping to mass-production. Doctoral dissertation, Harvard University Graduate School of Arts and Sciences.

Optical Engineering of Diamond 2013, eds. R. Mildren and J. Rabeau, Wiley-VHC, https://doi.org/10.1002/9783527648603

A. M. Zaitsev, Optical Properties of Diamond: A Data Handbook; Springer, 2001., https://doi.org/10.1007/978-3-662-04548-0

G. S. Sandhu and W.K. Chu, "Reactive ion etching of diamond", Applied Physics Letters, 55, 437-438, 1989, https://doi.org/10.1063/1.101890

E. Vargas Catalán et al, Astronomy & Astrophysics 2016

A. Toros et al, "Reactive ion etching of single crystal diamond by inductively coupled plasma: State of the art and catalog of recipes", Diamond & Related Materials 2020, https://doi.org/10.1016/j.diamond.2020.107839

Refractive and Diffractive Diamond Optics 2013, F. Nikolajeff and M. Karlsson in Optical Engineering of Diamond, eds. R. Mildren and J. Rabeau, Wiley-VHC, https://doi.org/10.1002/9783527648603.ch4

M. Karlsson and N. Nikolajeff, "Diamond micro-optics: microlenses and antireflection structured surfaces for the infrared spectral region", Optics Express, **11**, 5, pp. 502-507 (2003), https://doi.org/10.1364/OE.11.000502

C. Delacroix, et al, "Laboratory demonstration of a mid-infrared AGPM vector vortex coronagraph", Astronomy & Astrophysics, 553, (2013), https://doi.org/10.1051/0004-6361/201321126

11.5 Chapter 6

J. Kent Wallace, Skyler Palatnick, Max Millar-Blanchaer, Tobias Wenger, Dylan McKeithen, and Eugene Serabyn, "Short-wave infrared digital holographic microscope: application to optical characterization of phase and amplitude of metasurface optics," Appl. Opt. **65**, 697-704 (2026)

11.6 Chapter 7

K. Franson, et al., "Astrometric Accelerations as Dynamical Beacons: A Giant Planet Imaged inside the Debris Disk of the Young Star AF Lep", *The Astrophysical Journal Letters*, vol. 950, no. 2, Art. no. L19, IOP, 2023. doi:10.3847/2041-8213/acd6f6.

J. Kent Wallace, et al., "Architecting, Implementing and Observing with a Metasurface vector Zernike wavefront sensor on the Keck Telescope," AO4ELT7, ONERA, Jun 2023, Avignon, France. hal-04402835 10.13009/AO4ELT7-2023-088. hal-04402835

J. Kent Wallace, et al., "Experimental demonstration of a dual-purpose focal plane mask for active wavefront stabilization during high-contrast imaging," Proc. SPIE 13627, 1362714 (18 September 2025); https://doi.org/10.1117/12.3067404

M. Salama, et al., "Keck Primary Mirror Closed-loop Segment Control Using a Vector-Zernike Wavefront Sensor", ApJ, vol. 967, no. 2, (2024). doi:10.3847/1538-4357

David S. Doelman, Emiel H. Por, Garreth Ruane, Michael J. Escuti, and Frans Snik, "Minimizing the Polarization Leakage of Geometric-Phase Coronagraphs with Multiple Grating Pattern Combinations." *Publications of the Astronomical Society of the Pacific, Issue* 132, No. 1010 (2020): 1–18. https://www.jstor.org/stable/26973265.

Skyler Palatnick, Maxwell A. Millar-Blanchaer, J. Kent Wallace, Demis D. John, Anjali Moore, and Tobias Wenger, "Achromatizing photolithographically patterned metasurfaces with arbitrary, variable unit cell size," Opt. Express **32**, 47057-47075 (2024)

D. Mawet, E. Serabyn, K. Liewer, R. Burruss, J. Hickey and D. Shemo, “The vector vortex coronagraph: laboratory results and first light at Palomar Observatory,” *ApJ*, **709**, 53 (2010). DOI 10.1088/0004-637X/709/1/53.

E. Serabyn, E. Huby, K. Matthews, et al., “The W.M. Keck Observatory infrared vortex coronagraph and a first image of HIP 79124B”, *AJ*, **153**, 42 (2017). DOI 10.3847/1538-3881/153/1/43.

J. Kühn, E. Serabyn, J. Lozi, et al., “An H-band vector vortex coronagraph for the Subaru coronagraphic extreme adaptive optics system”, *PASP*, **130**, 035001. DOI 10.1088/1538-2873/aa9fe5.

Lorenzo König, Skyler Palatnick, Niyati Desai, et al. "Metasurface-based scalar vortex phase mask in pursuit of 1e-10 contrast", Proc. SPIE 12680, Techniques and Instrumentation for Detection of Exoplanets XI, 126800Q (5 Oct 2023); https://doi.org/10.1117/12.2676174

O. Guyon, “High Sensitivity Wavefront Sensing with a Nonlinear Curvature Wavefront Sensor”, PASP, vol. 122, no. 887, p. 49, 2010. doi:10.1086/649646.

Christophe Vérinaud, Cédric Taïssir Héritier, Markus Kasper, et al. "The Bi-O edge sensor: from theory to practice", Proc. SPIE 13097, Adaptive Optics Systems IX, 130971V (27 Aug 2024); https://doi.org/10.1117/12.3019958

C. Vérinaud, et al., “The Bi-O edge wavefront sensor. How Foucault-knife-edge variants can boost extreme adaptive optics”, A&A, vol. 682, 2024. doi:10.1051/0004-6361/202346660.

Emiel H. Por, “Phase-apodized-pupil Lyot Coronagraphs for Arbitrary Telescope Pupils”, *ApJ*, **888** 127 (2020). Doi:10.3847/1538-4357/ab3857

Rémi Soummer, Emiel H. Por, Raphaël Pourcelot, et al. "High-contrast imager for complex aperture telescopes (HiCAT): 8. Dark zone demonstration with simultaneous closed-loop low-order wavefront sensing and control", Proc. SPIE 12180, Space Telescopes and Instrumentation 2022: Optical, Infrared, and Millimeter Wave, 1218026 (27 Aug 2022); https://doi.org/10.1117/12.2630444

David S. Doelman, Frans Snik, Nathaniel Z. Warriner, et al. "Patterned liquid-crystal optics for broadband coronagraphy and wavefront sensing", Proc. SPIE 10400, Techniques and Instrumentation for Detection of Exoplanets VIII, 104000U (1 Sept 2017); https://doi.org/10.1117/12.2273406

Emiel H. Por, Raphaël Pourcelot, Ananya Sahoo, et al. "Broadband performance of the PAPLC coronagraph on the HiCAT testbed", Proc. SPIE 12680, Techniques and

Instrumentation for Detection of Exoplanets XI, 126801D (6 Oct 2023); https://doi.org/10.1117/12.2677757

E. H. Por, "Joint optimization of multiple optical planes in a stellar coronagraph", Proc. SPIE 12180, Space Telescopes and Instrumentation 2022: Optical, Infrared, and Millimeter Wave, 121805J.

Acronyms and Abbreviations

1D	one-dimensional
2D	two-dimensional
a-Si	amorphous silicon
AFM	atomic-force microscope
AGPM	annular groove phase mask
ALD	atomic-layer deposition
AO	adaptive optics
AR	anti-reflective
Astro2020	shorter name for *Pathways to Discovery in Astronomy and Astrophysics for the 2020s*
CGH	computer-generated holography
CGI	Coronagraph Instrument
CMOS	complementary metal oxide semiconductor
CVD	chemical vapor deposition
DHM	digital holographic microscope
DM	deformable mirror
DRIE	deep reactive-ion etching
DUV	deep ultraviolet
EBL	electron-beam (or e-beam) lithography
EFC	electric field conjugation
ELTs	extremely large telescopes
ESO	European Southern Observatory
EUV	extreme ultraviolet
FDTD	finite-difference time-domain

FEM	finite-element method
FGV	forked grating vortex
FoM	figure of merit
HCST	High Contrast Spectroscopy Testbed
HiCAT	High-contrast imager for Complex Aperture Telescopes
HICIT	High Contrast Imaging Testbed
HSQ	hydrogen silsesquioxane
HWO	NASA's Habitable Worlds Observatory
IACT	In-Air Coronagraph Testbed
IR	infrared
IRCT	Infrared Coronagraphic Testbed
IWA	inner working angle
JEOL	Japan Electron Optics Laboratory
JPL	Jet Propulsion Laboratory
JWST	NASA's James Webb Space Telescope
KISS	W.M. Keck Institute for Space Studies
KPIC	Keck Planet Imager and Characterizer
kV	kilovolt
LCP	liquid-crystal polymer
mRAD	milliradian
NASA	National Aeronautics and Space Administration
NCPA	non-common path aberrations
NIL	nanoimprint lithography

NIRCam	Near-Infrared Camera
nlCWFS	nonlinear curvature wavefront sensor
nm	nanometer
OPC	optical proximity correction
PAPLC	phase-apodized-pupil Lyot coronagraph
PAs	polarization aberrations
PDMS	polydimethylsiloxane
PECVD	plasma-enhanced chemical vapor deposition
PL	photolithography
PR	photoresist
RCWA	rigorous coupled-wave analysis
SAMIM	solvent-assisted micro molding
SEM	scanning electron microscope
STIS	Space Telescope Imaging Spectrograph
SVCs	scalar vortex coronagraphs
TRL	technology readiness level
UV	ultraviolet
VSB	variable-shaped beam
VUV	vacuum ultraviolet
vZWFS	vector Zernike wavefront sensor
WFIRST	Wide Field InfraRed Survey Telescope
WFS	wavefront sensing/sensor