Integral Field Spectroscopy with the Solar Gravitational Lens

Alexander Madurowicz © and Bruce Macintosh ©
Kavli Institute for Particle Astrophysics and Cosmology, Stanford University, Stanford, California, USA; amaduro@stanford.edu
Received 2021 December 1; revised 2022 March 11; accepted 2022 March 15; published 2022 May 2

Abstract
The prospect of combining integral field spectroscopy with the solar gravitational lens (SGL) to spectrally and spatially resolve the surfaces and atmospheres of extrasolar planets is investigated. The properties of hyperbolic orbits visiting the focal region of the SGL are calculated analytically, demonstrating trade-offs between departure velocity and time of arrival, as well as gravity assist maneuvers and heliocentric angular velocity. Numerical integration of the solar barycentric motion demonstrates that navigational acceleration of \( \frac{dv}{dt} \approx 80 \frac{m}{s} + 6.7 \frac{m}{s^2} \) year is needed to obtain and maintain alignment. Obtaining target ephemerides of sufficient precision is an open problem. The optical properties of an oblate gravitational lens are reviewed, including calculations of the magnification and the point-spread function that forms inside a telescope. Image formation for extended, incoherent sources is discussed when the projected image is smaller than, approximately equal to, and larger than the critical caustic. Sources of contamination that limit observational signal-to-noise ratio \( (S/N) \) are considered in detail, including the Sun, the solar corona, the host star, and potential background objects. A noise mitigation strategy of spectrally and spatially separating the light using integral field spectroscopy is emphasized. A pseudo-inverse-based image reconstruction scheme demonstrates that direct reconstruction of an Earth-like source from single measurements of the Einstein ring is possible when the critical caustic and observed \( S/N \) are sufficiently large. In this arrangement, a mission would not require multiple telescopes or navigational symmetry breaking, enabling continuous monitoring of the atmospheric composition and dynamics on other planets.

1. Introduction
In the wake of modern discoveries of exoplanets around sunlike stars (Mayor & Queloz 1995), advancements in exoplanetary detection have revealed statistics on the vast populations of planets in our galactic neighborhood (Akeson et al. 2013). Continued detection and characterization of these worlds has led others to question the possibility that other lifeforms inhabit these planets, and how we might infer their existence from spectroscopic indicators (Seager 2015). Perhaps the most direct path forward to the detection of biosignatures is the development of exoplanetary direct imaging and spectroscopy (Pueyo 2018), which is the focus of two future flagship NASA mission proposals HabEx (Gaudi et al. 2018) and LUVOIR (LUVOIR Collaboration 2019), as well as the Astro2020 recommendation to combine these efforts into a single 6.5 m class observatory. For these telescope designs, the planets can be resolved spectrally, but spatially appear as a single point, and no current technology has the capability to truly resolve details on the planet’s surface or in its atmosphere. One exception is inferring longitudinal cloud variations from temporal and rotational variability in extremely bright, nearby, substellar-mass brown dwarfs (Crossfield et al. 2014), but this technique cannot be easily extended to faint targets or latitudinal variations. However, an extremely ambitious idea to leverage the high magnification and angular resolution of the solar gravitational lens (SGL) to truly resolve an exoplanet has recently been conceptualized (Turyshhev et al. 2020).

This idea is not entirely new, as it is adapted from similar concepts to use the gravitational lens of the Sun as a link for interstellar communication (Eschleman 1979). However, recent theoretical advancements in the scattering of electromagnetic waves off of arbitrary oblate potentials (Turyshhev & Toth 2021a) has enabled the extremely accurate theoretical modeling of the capabilities of the SGL. The presence of high-order multipole moments in the gravitational potential result in the formation of a critical caustic in the image plane (Loutsenko 2018), which modifies the effective magnification of the lens as well as altering the point-spread function (PSF) formed in the image plane of a telescope observing at the SGL. This effect is critical for reconstruction of extended sources, as we demonstrate in Section 5. The asymmetric PSF formed for points inside the critical caustic enables direct disentanglement of the source intensity distribution from single measurements of the Einstein ring’s azimuthal profile. In this configuration, the reconstruction does not require multiple laterally offset observations of the Einstein ring, which were required using a point-mass or monopole-only model of the gravitational potential (Madurowicz 2020).

The paper is organized into three major Sections 2, 3, and 4. Section 2 considers the orbital mechanics required to send a craft to the SGL as well as maintaining alignment with a target planet. Section 3 broadly covers optical effects related to the gravitational lensing in the presence of an oblate potential. Section 4 discusses the sources of noise that will limit real observations as well as potential mitigation strategies. The paper concludes with Section 5, which demonstrates a simple model of source reconstruction from single measurements of the Einstein ring.
2. Orbital Considerations

2.1. Analytic Hyperbolic Trajectories

There are multiple, equivalent mathematical formulations for unbound hyperbolic trajectories. The formulas that follow are typically found in textbooks on astrodynamics (Schaub & Junkins 2003; McClain & Vallado 2001). The recognizable Cartesian parameterization

\[
\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1
\]

(1)

describes the shape of the hyperbola with two numbers \((a, b)\) named the semimajor and semiminor axes. The graph has two asymptotes at \(y = \pm \frac{b}{a}x\) and two branches opening along the \(x\)-axis that approach the asymptotes as the coordinates approach infinity.

A different formulation uses polar coordinates \((r, \theta)\)

\[
r = \frac{l}{1 - e \cos(\theta)}
\]

(2)

where the shape is described by two new numbers, the semi-latus rectum \(l = -\frac{b^2}{a}\) and eccentricity \(e = \sqrt{1 + \left(\frac{a}{b}\right)^2}\). An important note is that the polar coordinates have their origin as the focus of the right branch, and so to match the Cartesian grid, the \((x, y)\) coordinates can be recovered with the relations \(x = r \cos(\theta) - ae\) and \(y = r \sin(\theta)\).

However, for the purposes of this investigation, we wish to parameterize the hyperbola with two different numbers, the radius at periastron \(r_p\) and the velocity at periastron \(v_p\), so that one may think of the hyperbola shape phase space as a set of initial conditions for a rocket launching from an inner solar system before reaching the SGL. Everywhere along the hyperbola, the velocity associated with the trajectory is governed by the vis-viva equation

\[
v^2 = \mu \left(\frac{2}{r} - \frac{1}{a}\right)
\]

(3)

where \(\mu = GM_\odot\) is the standard gravitational parameter of the Sun. So simply plugging in \(r_p, v_p\), and solving for \(a\) allows determination of the semimajor axis

\[
a = \frac{\mu r_p}{2 \mu - v_p^2 r_p}.
\]

(4)

Additionally, the specific angular momentum \(h = r_p v_p = \sqrt{\mu l}\) of the orbit is conserved and can be used to find the semiminor axis

\[
b = \frac{r_p^{3/2} v_p}{\sqrt{v_p^2 r_p - 2 \mu}}.
\]

(5)

However, this equation will return an imaginary number when \(2 \mu > v_p^2 r_p\), in this case, the chosen \(r_p, v_p\) do not correspond to a hyperbola. An example plot of a basic hyperbola is shown in Figure 1.

Using the \((r_p, v_p)\) basis for hyperbola shapes, we aim to answer a few basic questions about the properties of these orbits to inform a design of a craft that intends to visit the SGL. The questions are: “How long will it take to reach a heliocentric distance of \(r_{\text{SGL}} \approx 600\) au?” “What is the angular velocity of the craft with respect to the Sun at that point in time?” and “How much acceleration is needed to cancel the relative lateral motion to enter a radial hyperbolic trajectory?”

The first question can be answered using the relationship between mean anomaly \(M\), eccentric (or hyperbolic) anomaly \(E\), and time \(t\). Specifically,

\[
E = \text{arcosh} \left(\frac{\cos \theta + e}{1 + e \cos \theta}\right)
\]

(6)

\[
M = e \sinh E - E
\]

(7)

\[
t = \sqrt{-\frac{a^3}{\mu}} M.
\]

(8)

Then it is just matter of finding \(\theta\) corresponding to \(r_{\text{SGL}}\) using

\[
\theta_{\text{SGL}} = \arccos \left(\frac{1}{r_{\text{SGL}}} - \frac{1}{e}\right)
\]

(9)

to compute the time a craft would need to reach such an extreme heliocentric distance. Contours are plotted for a variety of cases in Figure 2(a).

The second question can be analytically derived using the conservation of specific angular momentum. Since \(h = r \times (\frac{\partial \theta}{\partial t})\).

\[
\frac{d\theta}{dt} = \frac{h}{r_{\text{SGL}}^2}
\]

(10)

which is equivalent to Kepler’s second law. We convert these angular velocities into planet-crossing and system-crossing timescales \(T_{\text{planet}} = \theta_{\text{planet}} / \frac{\partial \theta}{\partial t}\) and \(T_{\text{system}} = \theta_{\text{system}} / \frac{\partial \theta}{\partial t}\) using \(\theta_{\text{planet}} = 2R_{\text{Earth}} / d_{\text{s}}\) and \(\theta_{\text{system}} = 2a_{\text{sys}} / d_{\text{s}}\) where \(d_{\text{s}} = 40.54\) ly is the system distance, \(a_{\text{sys}} = 0.062\) au is the target planet semimajor axis, and \(R_{\text{Earth}}\) is the Earth’s radii. These estimates correspond to the TRAPPIST-1 system (Gillon et al. 2017) and are plotted in Figures 2(b) and (c).

The last question can be tackled by means of computing the flight path angle \(\phi_{\text{fp}}\), which gives the angle between the
velocity vector and the direction perpendicular to the radial vector. The flight path angle is computed with

$$\phi_{\text{fpa}} = \arctan \frac{e \sin \theta_{\text{SGL}}}{1 + e \cos \theta_{\text{SGL}}} ,$$

and the magnitude of the velocity vector $v_{\text{SGL}} = \sqrt{\mu(2/r_{\text{SGL}} - 1/a)}$ is computed with the vis-viva equation. Thus the delta-v required to eliminate the transverse motion and enter a radial hyperbolic orbit is

$$dv = v_{\text{SGL}} \sin \left( \frac{\pi}{2} - \phi_{\text{fpa}} \right) .$$

Values of which are plotted in Figure 2(d).

Examining the plots in Figure 2, a few important trade-offs in the orbit parameter space become apparent. Orbits with large velocity at periastron correspondingly have the shortest time to reach great heliocentric distance, while the radius at periastron has a small effect, with a slight preference for starting farther out of the solar system. However, these wide-starting orbits have much greater angular velocity at $r_{\text{SGL}}$, and correspondingly smaller planet-crossing and system-crossing timescales. The orbits with the smallest angular velocity (and thus largest crossing timescales, as well as smallest delta-v to enter the radial trajectory) correspond to orbits with small radius at periastron and large velocity at periastron.

To make a concrete example, suppose a craft is launched in a two per week, enough to resolve effects from Mercury’s short 88 day orbit. The results of the simulation consist of a data cube of size $(N_{\text{particles}}, N_{\text{timesteps}}, 6)$, which contains the six x-, y-, and z- positions and velocities of each of the particles in the integration at every time step. To evaluate the effect of the barycentric motion of the Sun on the pointing of the SGL, we investigate the worst-case scenario. Since the majority of the motion of the Sun occurs in the invariable plane (Souami & Souchay 2012), defined by the plane of the barycenter and perpendicular to the total angular momentum, a target planet that is pointed directly along this angular momentum vector will suffer from the largest perturbations of the position of the Sun along the line of sight. Any other target pointing will be somewhat better from this perspective as the motion of the Sun is projected into the plane of observation. In our coordinate system, this plane is defined by the x- and y-coordinates, and so the telescope craft visiting the SGL has a position of $(0, 0, r_{\text{SGL}})$ once it has reached the observation location and maneuvered to enter a
purely radial orbit trajectory. This next part of the analysis ignores the additional \( z \)-axis motion for simplicity. Although it is a small effect, this could be corrected by adding an additional term for a time-dependent \( z \)-coordinate, \( z = r_{SGL} + (t/T_{SGL})\mu(2/z - 1/a) \), which would require solving the vis-viva equation for \( v(z(t)) \).

In Figure 3(a), the \((x, y)\) position of the Sun relative to the barycenter is plotted in units of solar radii, showing the reflex motion due to perturbations from the planets causing a displacement of around \( \sim 1 \)–\( 2 R_\odot \). These values are used to estimate the angular velocity of the “SGL pointing” vector defined as starting at the stationary craft at \((0, 0, r_{SGL})\) and directed at the position of the Sun \((x_\odot, y_\odot, 0)\). The discretely sampled position vector is the output of the REBOUND simulation. First, we compute the discrete \( x \)- and \( y \)-second-order position differences \( dx = (x_{i+1} - x_{i-1})/2 \) and \( dy = (y_{i+1} - y_{i-1})/2 \), using \texttt{np.gradient}, which handles the edge cases using a first-order finite difference, as well as the discrete time difference \( dt = 100 \) yr/\( 10^4 \) samples. The error on these estimates of the derivative is on the order of \( \mathcal{O}(dt^2) \). Finally, the angular velocity of the pointing vector is computed with

\[
\frac{d\phi}{dt} = \frac{\sqrt{dx^2 + dy^2}}{r_{SGL} dt},
\]

which can be similarly converted into a planet-crossing timescale as before, using \( T_{cross} = \theta_{planet} / \frac{dx}{dt} \). Both the angular velocity and planet-crossing timescale are plotted over time in Figures 3(b) and (c). From these estimates, one can see the average angular velocity is around 2–3 mas per day, and that the corresponding planet-crossing timescales are between 200–300 s. So to design a craft that obtains longer exposure times at ultra-precise alignments, additional navigation acceleration will be needed to counteract this effect.

In order to compute the amount of acceleration necessary to counteract the barycentric motion of the Sun, we first compute the location the craft must station-keep at to maintain the SGL aimed at a fixed point on the sky. If \( d_s \) is the distance between the Sun and source or target planet, and \( d_t = r_{SGL} \) is the distance between the craft and the Sun or lens, a simple similar triangle argument shows the position \((x, y, z = d_t)\) required to station-keep is enhanced by a small factor

\[
x = x_\odot \left( 1 + \frac{d_t}{d_s} \right)
\]

\[
y = y_\odot \left( 1 + \frac{d_t}{d_s} \right)
\]

relative to the position of the Sun due to the additional distance. This enhancement is a small number that depends on the distance to the target being considered. For \( d_s = 40.54 \) ly and \( d_t = 600 \) au, \( \frac{d_t}{d_s} \approx 2 \times 10^{-4} \). This position is shown as a function of time in Figure 4(a) in units of solar radii. Subsequently, the velocity \( v_{x,y} = \frac{dx_{x,y}}{dt} \) and acceleration \( a_{x,y} = \frac{dv_{x,y}}{dt} \) required to maintain this position are computed numerically using the discrete second-order difference using \texttt{np.gradient}, and are plotted in Figures 4(b) and (c).

Examining the figure, it is clear that a small, continuous acceleration of \( \sim 2 \times 10^{-7} \) m s\(^{-2}\) in an oscillatory fashion is necessary to counteract the lateral motion of the Sun around the barycenter to maintain a fixed pointing of the SGL. Considering the possibilities of rocket designs, it seems that ion thrusters may be ideally suited to this task. While the specifics depend on a number of design choices, such a beam divergence, mass of the ion, and the electrical current and voltage used to accelerate the ions (Goebel & Katz 2008), a reasonable estimate of thrust from such an ion thruster is around \( \sim 100 \) mN, which could give accelerations of \( 10^{-4} \) m s\(^{-2}\) for a 1000 kg craft over very long durations, due to the high specific impulse of the ions, which is more than two orders of magnitude needed.

An alternative visualization of this acceleration trajectory is visualized in Figure 5. By converting to polar coordinates, the direction of the acceleration vector in the \( xy \)-plane is plotted as a function of time, which is now the radial coordinate. The log of the magnitude of the acceleration vector is represented by the color of the curve. One can see that acceleration curve spirals outward, as the vector rotates around with an approximate 10 yr period dominated by the effect of Jupiter.
However, the perturbations of the other planets contribute as well, with consistent changes in the acceleration vector’s direction and corresponding changes in magnitude. To continue our analysis, we estimate the total delta-v needed to maintain this acceleration over time as the integral

\[ dv_{\text{total}}(t) = \int_0^t |a_{\text{total}}(t')| \, dt' \]

of the magnitude of the acceleration vector \(|a_{\text{total}}(t)| = \sqrt{a_x(t)^2 + a_y(t)^2}\). This total delta-v is a linear function of the length of the time you would like to maintain the alignment for, with a slope corresponding to the average value \( \langle |a_{\text{total}}| \rangle = \frac{1}{t} \int_0^t |a_{\text{total}}(t')| \, dt' \). As a result, the approximate cost scaling to maintain alignment is

\[ dv_{\text{total}}(t) = \langle |a_{\text{total}}| \rangle t \approx 6.7 \text{ m s}^{-1} \text{ year}, \]

which demonstrates that the station-keeping cost to counteract the solar barycentric motion is sufficiently small as to be reasonably possible.

However, this entire analysis has so far ignored the motion of the target across the sky, composed of the relative proper motion of the system and the orbital motion of the target around its host star. A simple estimate of the acceleration that would be needed to counteract the orbital motion of a hypothetical planet with a face-on, circular orbit of semimajor axis \( a \) and period \( T \) is

\[ a_p = \frac{d_l}{d_s} \frac{4 \pi^2 a}{T^2}, \]

which can dominate over or be comparable to the effect of the solar barycentric motion depending on the assumed planet properties. If \( a = 1 \text{ au}, T = 1 \text{ yr}, d_l = 600 \text{ au}, \) and \( d_s = 10 \text{ pc}, \) then \( a_p \approx 6 \times 10^{-6} \text{ m s}^{-2} \approx 60 \text{ m s}^{-1} \text{ yr}^{-1}, \) but if instead \( d_s = 100 \text{ pc}, \) then \( a_p \approx 0.2 \times 10^{-6} \text{ m s}^{-2} \approx 6 \text{ m s}^{-1} \text{ yr}^{-1}. \) A more complete three-dimensional analysis such as that in Turyshev & Toth (2022) is necessary to evaluate the combined effect, as the orientation of the orbits is important to determine if this acceleration exacerbates or counteracts the solar barycentric motion. Ultimately, precise determination of the target ephemeris will be necessary to locate the exact position needed to send the craft, and predicting this location in advance may be rather difficult.

### 2.3. Constraints on Target Ephemerides

Different exoplanetary detection techniques are capable of constraining different planetary orbital elements. Transit photometry is able to determine the planet’s period and to some extent its orbital inclination, as it must be nearly edge-on...
in order for a transit to occur (Deeg & Alonso 2018). Doppler spectroscopy or the radial velocity method measures the period and velocity semiamplitude, which can be used to infer the orbit’s eccentricity, mass-inclination product $m \sin i$ and, if the star’s mass is known through another technique, the orbital semimajor axis (Eggenberger & Udry 2010). If the planet is amenable to both transit and radial velocity, the combination of both techniques can further constrain the orbit. The Rossiter-McLaughlin effect can be used to measure the relative azimuthal angle on the sky between the planet’s orbit and the stellar spin axis (Ohta et al. 2005; Masuda 2018), using asymmetric distortions in the line profiles of the stellar spectrum as the planet crosses and sequentially obscures different subsets of the stellar disk. However, absolute determination of the planet’s orbital position angle on the sky would still require determination of the orientation of the stellar spin axis, which may be possible using spectro-interferometry to measure the astrometric motion associated with stellar rotation (Kraus et al. 2020), with spectro-astrometric analysis of high-resolution long-slit spectra (Lesage & Wiedemann 2014), or by combining projected spectroscopic rotation rates with photometric rotation periods (Bryan et al. 2020). It may also be possible to make these measurements during the early part of the mission to the SGL, by visiting the image of the star first to aid in navigation before moving to the image of the target planet.

Turyshev & Toth (2022) discusses a strategy that uses astrometric observations of the host star using the SGL to help constrain the planet’s position. However, direct imaging of exoplanets can constrain planetary orbital elements by directly measuring the planet’s astrometric position on the sky. Typical astrometric accuracies obtainable with current instruments on single telescopes are between 3 and 22 mas (Konopacky et al. 2016), while the most precise measurements currently obtained could be around $\sim 100 \mu$as (GRAVITY Collaboration et al. 2019) using multiple telescopes combined into an interferometer. These uncertainties in measured precision are typically relative measurements, comparing the planet’s position to the host star, and an absolute astrometric calibration would require precise determination of the stellar absolute astrometry, which could be measured using GAIA (Brown 2021; Lindegren et al. 2018). The uncertainty in angular position $\theta$ on the sky directly translates into a spatial, lateral uncertainty $dx = d\theta \mu$ of the image of the target in the SGL image plane. The large enhancement due to the great heliocentric distance means that accurate determination of the angular position on the sky of the target is critical to locating the position of the image of the target in the SGL focal plane. An angular uncertainty of $\theta \sim 20$ mas corresponds to an uncertainty in location of $\Delta \chi \sim 8800$ km, while $\theta \sim 100 \mu$as corresponds to $\Delta \chi \sim 44$ km. Since the image of the target planet is rescaled from its original size $R_{\text{image}} = \frac{d}{d_{\text{Earth}}} R_{\text{Earth}}$ by the “plate scale” of the SGL geometry, an Earth-sized planet at a distance of 40 ly has an image only $\sim 3$ km in diameter. With current levels of uncertainty in relative astrometry, it could be missed even with navigation relative to the image of host star. A further caveat is that those measurements are typically for Jupiter-sized planets, since no Earth-like planets have been directly imaged, although mission concepts like HabEx and LUVOIR have been proposed to potentially achieve this (Gaudi et al. 2018; LUVOIR Collaboration 2019).

Additionally, the planet and star are continually in motion, due to the relative orbital motion of the system, and the proper motion and acceleration of the entire system within the galaxy. Typically stars have galactic velocities around $\sim 100 \text{ km s}^{-1}$ and accelerations around $\sim 10^{-10} \text{ m s}^{-2}$ (Silverwood & Easther 2019), which means determining the location needed to visit the image of either the star or planet would need to be predicted in advance, to plan an appropriate trajectory. Given that realistic timescales between launch and arrival at the SGL are of the order of $\sim 100$ yr, and that typical proper motions are of the order of $\sim 10^{-6}$ yr$^{-1}$, the location of the image of a target star now and later would be displaced by millions of kilometers, with uncertainties in current proper motion further worsening the issue. Additionally, predicting the future orbital motion of the relative motion of the planet and star is subject to additional uncertainty.

Software such as orbitize! (Blunt et al. 2020) can be used to fit astrometric observations with three-dimensional Keplerian orbits, allowing estimates of position into the future, although the accuracy of prediction depends strongly on the orbital coverage of the data and its astrometric accuracy. To demonstrate this, Figure 6 was created. First, a series of
simulated astrometric data are generated that correspond to a ground truth Keplerian orbit with a semimajor axis of 1 au, an eccentricity of 0.05, an inclination of 30°, and a solar system 100 pc away. Either 10 or 33 samples were generated over 0.3 or 1.0 yr with an observational uncertainty around ~1 mas, which correspond to an orbital coverage of 33% and 100% for the partially and fully constrained cases respectively.

The simulated data points are fit with the affine-invariant Markov Chain Monte Carlo algorithm developed by Foreman-Mackey et al. (2013), which samples the posterior distributions of all orbital parameters. After a burn-in time of 10⁴ steps and a total number of 10⁶ samples, the posterior distributions have mostly converged, although with less precision on the true orbital elements for the partially constrained orbit. For the partially constrained orbit, attempting to predict the location at a time corresponding to unconstrained and unmeasured locations in the orbit causes the posterior distribution of predicted locations to become large. Even with observational uncertainties on the order of ~1 mas, the future location posterior has uncertainty an order of magnitude larger around ~10 mas. However, a fully constrained orbit predicts the future location with accuracy comparable to the observational uncertainty on the data points themselves, which is a necessary precursor for navigation to the precise location of the planet’s image at the SGL. This demonstrates the importance of complete astrometric orbital coverage on constraining the planetary orbital elements and predicting its future location.

3. Optical Considerations

3.1. Gravitational Field Oblateness

The simplest model for the gravitational potential of the Sun is that of a point mass, \( \Phi = -\frac{GM}{r} \). This model has a very simple formula (Narayan & Bartelmann 1996)

\[
\theta_b = \frac{2R_s}{b} \hat{e}_b
\]  

(18)

for the deflection of a photon as a function of impact parameter \( b \). Here \( R_s = \frac{2GM}{c^2} \) is the Sun’s Schwarzschild radius and \( \hat{e}_b \) is the unit vector in the direction of the impact parameter. Since a photon can travel from the source to the observer around the point mass and receive the same deflection for a path at any azimuthal angle, it is easy to see how gravitational lensing around a point source forms an Einstein Ring symmetrically around the point mass.

However, sophisticated helioseismological analyses (Mecheiri et al. 2004) require the gravitational field to deviate from spherical symmetry. The rotation of the Sun creates an oblate gravitational potential that is described by a zonal spherical harmonic expansion

\[
\Phi = -\frac{GM}{r} \left[ 1 + \sum_{n=2}^{\infty} I_n \left( \frac{R_s}{r} \right)^n P_n(\cos \beta_s) \right],
\]  

(19)

where \( R_s \) is the radius of the Sun, the \( P_n \) are the Legendre Polynomials (Abramowitz & Stegun 1968), \( \beta_s \) is the solar colatitude, and \( I_n \) are the dimensionless multipole moments. Due to the North–South symmetry along the rotation axis, all terms with odd \( n \) are zero. The largest even \( n \) terms are \( I_2 \approx 2 \times 10^{-7} \), \( I_4 \approx -4 \times 10^{-9} \), \( I_6 \approx -3 \times 10^{-10} \), and \( I_8 \approx 1 \times 10^{-11} \) (Roxburgh 2001).

The effect of this deviation from spherical symmetry on the resulting optical properties of the SGL was first investigated by Loutsenko (Loutsenko 2018). They demonstrated that the quadrupole moment \( I_2 \) results in the formation of an astroid-shaped caustic in the image plane with a diameter

\[
D_{\text{astroid}} = \frac{4I_2 R_s^2}{\sqrt{2R_s r}} \sin^2 \beta_s.
\]  

(20)

Turyshev and Toth (Turyshev & Toth 2021a) extended their work and computed the effect of any high-order multipole \( I_n \) on the resulting optical properties of the gravitational lens. In general, the caustic curves are given by higher-order hypocycloids (Turyshev & Toth 2021b), and the general case of photon deflection can be computed with

\[
\theta_s = \frac{2R_s}{b} \left[ \hat{e}_b - \sum_{n=2}^{\infty} I_n \left( \frac{R_s}{b} \right)^n \sin^n(\beta_s) \times (\hat{e}_s \cos[n(\phi_s - \phi)] - \hat{e}_\phi \sin[n(\phi_s - \phi)]) \right],
\]  

(21)

where \( \phi_s \) and \( \phi_\phi \) are the azimuthal orientations of the impact parameter and the solar rotation axis. When \( I_n = 0 \), this formula reduces to the well-known example of a point mass, but includes new terms for the deflection along the \( \hat{e}_\phi \) direction, which is perpendicular to the impact parameter (see Figure 2 in Turyshev & Toth 2021a). This additional deflection gives the SGL entirely new behavior that is not symmetric with respect to azimuthal rotation, and explains geometrically the formation of the Einstein Cross for \( n = 2 \).

Furthermore, they go on to solve Maxwell’s equations by treating the new gravitational potential as a perturbation of the Newtonian metric, and using the angular eikonal method, which takes inspiration from Mie scattering theory. In this formalism, the magnification of the lens \( \mu \), defined as the ratio of the Poynting vectors for a plane wave at the source and resulting image, is (Turyshev & Toth 2021c)

\[
\mu(\mathbf{x}) = 2\pi kR_s |E(\mathbf{x})|^2,
\]  

(22)

where \( k = \frac{2\pi}{\lambda} \) is the wavevector, and the quantity \( E(\mathbf{x}) \) is given by the diffraction integral (Turyshev & Toth 2021c)

\[
E(\mathbf{x}) = \frac{1}{2\pi} \int_0^{2\pi} d\phi_z \exp \left[ -i k \left( \frac{2R_s}{r} \rho \cos(\phi_z - \phi) \right. \right.
\]
\[
+ \left. 2R_s \sum_{n=2}^{\infty} I_n \left( \frac{R_s}{\sqrt{2R_s r}} \right)^n \sin^n(\beta_s) \cos[n(\phi_s - \phi)] \right] \bigg].
\]  

(23)

where \( \mathbf{x} = (\rho, \phi) \) are the polar coordinates in the image plane. If all of the multipole moments \( I_n = 0 \), this diffraction integral can be analytically computed, and the result is that of the monopole-term only, \( E(\mathbf{x}) = J_0(k(\sqrt{2R_s/r}) \rho) \), where \( J_0 \) is the Bessel Function of the First Kind (Abramowitz & Stegun 1968). However, it is not analytically tractable for nonzero \( I_n \), as would be the case for a realistic Sun, and so numerical methods must be relied upon to evaluate Equation (23).

In this paper, we investigate Equation (23) with a discrete approximation to the integral over the azimuthal angles \( \phi_s \). By
it is possible to evaluate the integral numerically. Here \( \phi_{i,j} \in [0, 2\pi] \) is a discrete approximation to the continuous \( \phi_{i,j} \) and \( \Delta_{\xi_{i}} = \frac{2\pi}{N_{p}} \), where \( N_{p} \) is the number of points to use for the azimuthal sampling. \( N_{p} \) must be chosen carefully to be sufficiently large so that the approximation is valid. For “easy” cases where \( k, \beta_{S}, \) and \( \rho \) are relatively small, such as \( \lambda = 4 \mu \text{m} \), \( \beta_{S} = 10^{9}, \) and \( \rho = 0 \text{ m} \), we find that \( N_{p} \gtrsim 10^{4} \) is sufficient. “Hard” cases where \( k, \beta_{S}, \) and \( \rho \) are relatively large such as \( \lambda = 0.5 \mu \text{m}, \beta_{S} = 90^{\circ}, \) and \( \rho = 10^{3} \text{ m} \) can require significantly finer azimuthal sampling, up to \( N_{p} \gtrsim 10^{5} \) to ensure the discrete approximation to the integral does not break down.

3.2. The PSF and the Effects of Telescope Diameter, Wavelength, and Solar Co-latitude

The approximation in Equation (24) depends on many parameters, the location in the image plane \( x = (\rho, \phi) \), the wavelength \( \lambda = 2\pi/\lambda \) of light, the multiple moments \( I_{m} \), the mass \( M_{S} \), and radius \( R_{S} \), of the Sun, and the orientation of the solar rotation axis \( (\beta_{S}, \phi_{S}) \). Furthermore, once one has computed the result of Equation (24), which gives the complex electromagnetic field in the SGL image plane rescaled by the magnification prefactors, it is rather straightforward to find the PSF that would result in the image plane of a telescope using the Fraunhofer approximation and the techniques of Fourier Optics (Hecht 2002; Goodman 1996). If \( P(x) = \rho < D_{\text{tel}}/2 \) is the binary pupil transmission function of a circular aperture telescope, then the PSF in its image plane is

\[
\text{PSF}(\alpha, \beta) = |F^{-1}(\alpha, \beta)[P(x)E(x)]|^2,
\]

where \((\alpha, \beta)\) are the angular coordinates in the conjugate plane, and \( F^{-1}[\ldots] \) is the inverse Fourier Transform. The inverse Fourier Transform, which has a positive sign in the exponent \( e^{+ikx} \) of the exponential basis, is used instead of the Forward Transform to remain consistent with observed orientation of asymmetric PSFs in high-contrast imaging from the ground (see Section 2.4 in Madurowicz et al. 2019). To showcase the effect of these parameters on the resulting images formed, Figure 7 was created. The figure contains three separate cases, with different wavelengths \( \lambda = (1, 4, 16) \mu\text{m} \) and solar co-latitude \( \beta_{S} = (11.25, 22.5, 45)^{\circ} \), with each case labeled in its title. Each case contains three plots, the magnitude of \( |E(x)| \), the phase inside the telescope pupil, and the resulting PSF in the telescope image plane.

A few key observations from this Figure. First, the image formation in the SGL image plane is dominated by the quadrupole moment \( I_{2} \), and the resulting \( |E| \) is shaped primarily like the four-pointed hypocycloid, the asteroid. The higher-order multipole moments have a small effect, but are subdominant for realistic solar values. Second, for larger \( \beta_{S} \), the size of the asteroid caustic grows significantly. For the first case with \( \beta_{S} = 112.25^{\circ}, \) the asteroid diameter is \( \sim 20 \text{ m}, \) but for \( \beta_{S} = 45^{\circ}, \) it is \( \sim 200 \text{ m}. \) At the solar equator, \( \beta_{S} = 90^{\circ}, \) the asteroid diameter reaches its maximum value \( \sim 560 \text{ m}. \) This has the implication that any given target planet, which has a unique location on the sky, will suffer from degradation to the average magnification over the pupil for a set telescope diameter, since the positioning and size of the telescope control the fraction of intercepted light. However, this only is true when the telescope diameter is much smaller than the asteroid diameter. If the telescope is sufficiently large to capture the entire asteroid, then all of the flux can be intercepted, but that would require telescopes much larger than is currently feasible.

Third, the phase of the electromagnetic field appears to have a grid-like pattern that is responsible for the formation of the Einstein Cross. The size scaling of this grid pattern is controlled mostly by the wavelength of the light, with shorter wavelengths having a narrower spacing between the zones, and longer wavelengths having larger zones. It is crucially important that the sampling of the phase inside the pupil is fine enough that the delicate grid pattern is not aliased, and that the pupil area is large enough so that at least one “grid element” is fully contained inside the pupil. If these conditions are not both met, the resulting PSF does not produce the characteristic Einstein Cross, seen on the right. Fourth, the angular size of the four points in the telescope’s PSF grow linearly with wavelength, as is expected. The positioning of the four point images is just outside the angular extent of the solar disk, shown at a heliocentric distance of \( r_{\text{SGL}} = 650 \text{ au}. \) When the telescope is precisely aligned with the optical axis of the point source and the gravitational lens, the locations of the four point sources is symmetric; however, a lateral misalignment of the telescope intercepts an offset phase pattern, which will cause the locations of the points to be perturbed around the edge of the ring; see Figure 8. When the telescope is positioned directly on the caustic edge, the images have moved sufficiently around the edge to merge and form into an arc. When the telescope has moved outside the asteroid caustic, the result is two point sources similar to the monopole-only case. However, the effect of the multipoles is still present, and the intensity of each point is asymmetric.

3.3. The Average and Asymptotic Magnification

To further investigate the effect of the telescope diameter on the intercepted flux, we compute the average magnification across the telescope pupil \( \langle \mu \rangle_{p} \). Equation (24) is computed numerically for points \((\rho, \phi)\) in the image plane, and the average across the pupil is defined as the weighted sum of \( \mu \) in the region inside the telescope pupil \( P \in \{(x, y)\} \sqrt{(x - c_{x})^2 + (y - c_{y})^2} \leq D_{\text{tel}}/2 \).

\[
\langle \mu(x) \rangle_{P} = \frac{\iint_{P} \mu(x) \, dx}{\iint_{P} \, dx} \tag{26}
\]

Here \((c_{x}, c_{y})\) is the location of the center of the telescope in Cartesian coordinates, which has diameter \( D_{\text{tel}} = 2 \text{ m}. \) This metric is used to compare the multiple scenarios plotted in Figure 9.

Examining the behavior of the magnification in Figure 9, it is apparent that the effect of the higher-order multipole moments causes significant alteration to the magnification in the region of
the astroid caustic. This regime is characterized by $\rho \lesssim D_{\text{astroid}}$. Here the maximum of the magnification no longer occurs on the optical axis, but instead in the region of the caustic, and is reduced from the maximum of the monopole by a few orders of magnitude, of which the exact amount depends on the solar colatitude $\beta_S$ as well as the telescope diameter $D_{\text{tel}}$.

Figure 7. Three simulated cases showing from left to right, the magnitude of the electromagnetic field $|E|$ computed numerically with the approximation in Equation (24), the phase of the electromagnetic field inside the pupil of the telescope, centered on the optical axis, and the resulting PSF in the image plane of the telescope. In the left figures, the telescope pupil is plotted with a white circle, and the geometric astroid curve governed by Equation (20) is plotted with a 5% enhancement so as to not cover up the intensity image. In the rightmost images, the angular size of the solar disk is shown with a white circle.
However, far outside the asteroid caustic, the average magnification is consistent with the pure monopole case. In this section, we investigate the asymptotic behavior of these functions and demonstrate an analytic comparison between the wave and geometric theory, which results in a power-law approximation valid in this regime. Using the geometric theory, the magnification of a point-mass lens can be described by Narayan & Bartelmann (1996)
\[ \mu_{\text{geo}} = \frac{1}{1 - \left(\frac{\theta_E}{\beta}\right)^2}, \]  
(27)

where \( \theta_E \) is the angular scale of the Einstein Ring
\[ \theta_E = \sqrt{\frac{2R_s D_L S}{D_s D_L}}, \]  
(28)

and the \( \theta_\perp \) are the angular coordinates of the magnified images
\[ \theta_\perp = \frac{1}{2}(\beta \pm \sqrt{\beta^2 + 4\theta_E}). \]  
(29)

Here \( \beta \) is the angular coordinate of the point source and is geometrically related to the radial coordinate \( \rho \) in the image plane by
\[ \beta = \arctan \frac{D_s}{\rho} - \arctan \frac{D_L}{\rho} \approx \rho \left(\frac{1}{D_L} - \frac{1}{D_s}\right). \]  
(30)

Taking a Taylor expansion of Equation (27) in the regime where \( \beta \ll \theta_E \), results in
\[ \mu_{\text{geo}} \approx \frac{1}{2} + \left(\frac{\theta_E}{2\beta}\right)^2, \]  
(31)

so evidently the dominant scaling of the geometric magnification is \( \mu_{\text{geo}} \propto \frac{1}{\beta} \propto \frac{1}{\rho} \). Taking a Taylor expansion in the regime where \( \beta \gg \theta_E \), results in
\[ \mu_{\text{geo}} \approx 1 + \left(\frac{\theta_E}{\beta}\right)^4, \]  
(32)

and the dominant term is the constant, which is succinctly summarized in the limiting behavior \( \lim_{\rho \to \infty} \mu_{\text{geo}} = 1 \). This limit indicates that far enough away from the lens, the point source simply exists as it would be unperturbed by anything. To compare and contrast this geometric scaling argument with the wave-theoretic treatment described previously, we take an asymptotic expansion of the magnification for a pure monopole SGL. Starting with the expression for the magnification
\[ \mu_{\text{wave}} = 2\pi k R_s \left| J_0\left(k \sqrt{\frac{2R_s}{r}} \rho\right)\right|^2, \]  
(33)

and since the Bessel Functions of the First Kind have asymptotic expansions (Abramowitz & Stegun 1968)
\[ J_0(z) = \frac{2\pi}{z} \left[ \cos \left(z - \frac{n\pi}{2} - \frac{\pi}{2}\right) + e^{i|z|} O\left(\frac{1}{|z|}\right) \right], \]  
(34)

we can conclude, for \( \rho \gtrsim D_{\text{astroid}} \),
\[ \mu_{\text{wave}} \approx 4\pi^2 \sqrt{\frac{r R_s}{2}} \rho^{-1} \cos^2 \left(1 - \frac{\theta_\perp}{\theta_E}\right) \approx \frac{1}{\rho}, \]  
(35)

which shares the same power-law scaling with the geometric theory. Incidentally, the limiting behavior of this function is discrepant as \( \lim_{\rho \to \infty} \mu_{\text{wave}} = 0 \). This likely is a result of the formalism being used to describe the electromagnetic field in the geometric shadow of the Sun, where the direct path component of the field is obscured by the solar disk. So if \( \rho > \rho_{\text{critical}} \), such that \( \beta > \theta_E \), then this description is no longer valid. Indeed, in this case, it would be simpler to imagine that the lens does not exist at all, and to treat the direct path on its own. To summarize, the monopole approximation is valid and consistent with the geometric power-law scaling \( \propto \rho^{-1} \) in the regime where \( D_{\text{astroid}} \lesssim \rho < \rho_{\text{critical}} \). Inside the asteroid, one must consider the perturbations of the higher-order multipoles, and outside the critical distance, one may ignore the lensing entirely. This approximation is useful to quickly evaluate the response of a realistic SGL at large lateral separations, since numerical evaluation of Equation (24) becomes increasingly challenging at large \( \rho \).

3.4. Imaging of Extended Sources

With formalism described in Sections 3.1 and 3.2, it is possible to calculate the PSF of the SGL as well as the PSF as it would appear in the image plane of a telescope in the focal plane of the SGL. In this section, we extend our analysis to the imaging of extended sources under the assumptions of incoherent emission and small angular extent of the source. The former assumption allows one to treat the extended source simply exists as it would be unperturbed by anything. To compare and contrast this geometric scaling argument with a superposition of independent point sources, and the latter allows one to ignore minor perturbations to the orientation of the solar rotation axis (\( \beta_S, \theta_S \)) that would begin to matter if the angular extent of the extended source was sufficiently large. For justification, an Earth-sized planet at a target distance of \( d_S = 100 \) pc occupies a solid angle of \( \Omega_E = \pi \left(R_E/d_S\right)^2 \approx 1.4 \times 10^{-23} \) steradians.

If the Cartesian coordinates in the source plane are \( (x', y') \), then we can approximate the continuous intensity distribution of the target planet \( I(x', y') \) with
\[ I_{n_x, n_y} \approx \left(\hat{\text{III}}_{n_x}(x') \hat{\text{III}}_{n_y}(y') I(x', y')\right), \]  
(36)

where \( \hat{\text{III}}_{n_x}(x') = \delta(x' - \Delta_x (n_x - N_x/2)) \) is a comb function, \( \delta \) is the Dirac delta function, \( n_x \in [0, 1, 2, ..., N_x] \) is the integer index of the pixel, \( \Delta_x \) is the spatial sampling, and the \( N_x/2 \) is a centering offset, so the origin of \( x' = 0 \) occurs at the halfway index \( n_x = N_x/2 \). Additionally, \( (\ldots) \) represents the average of the intensity distribution in the vicinity of the sampling point, and at the end, \( I_{n_x, n_y} \) is a vector (or matrix), indexed by pixel numbers \( (n_x, n_y) \), which represents a discrete approximation to the continuous intensity distribution of the source. In our analysis, we use uniform sampling for both \( x \) and \( y \) and set \( \Delta_x = 2R_{\text{Earth}}/N_x, N_x = 64 \) pixels. As a consequence,
\[ x'_{n_x} = R_{\text{Earth}} \left(\frac{2n_x}{N_x} - 1\right) \]  
(37)
\[ y'_{n_y} = R_{\text{Earth}} \left(\frac{2n_y}{N_y} - 1\right), \]  
(38)

and both \( x', y' \) are \( [-R_{\text{Earth}}, R_{\text{Earth}}] \), since \( \frac{n_x + n_y}{N_x} \in [0, 1] \).

For our analysis, each pixel in this approximation acts as an independent, incoherent source in the imaging problem. Each point on the source object is reimaged by the SGL and appears as an asteroid-shaped PSF in the SGL image plane, with the center of the asteroid offset by its location in the source plane multiplied by the SGL plate scale. If the Cartesian coordinates in the SGL image plane are \( (x_n, y_n) \), then the location of the
The astroid center for a point source located at \((x'_n, y'_n)\) is

\[
x_n = -x'_n \frac{d_l}{d_S},
\]

\[
y_n = -y'_n \frac{d_l}{d_S},
\]

where the negative sign appears due to the inversion of the image. Figure 10 depicts the geometry of the problem for convenience.

In this diagram, which is not to scale, the size of the astroid PSF in the SGL image plane is depicted to be a fraction of the size of the extended object, a diagram of the Earth. One can visualize how a telescope situated directly on the optical axis will see unique PSFs for each of the three points highlighted in the diagram. For the central point on the optical axis, the telescope will sit squarely inside the center of the astroid, where the PSF will form an Einstein cross in the telescope image plane. For the two points that are offset from the optical axis, the telescope centered on the optical axis will instead intercept light from outside the edge of the astroid, which will result in simply two point sources in the telescope’s image plane. It may be helpful to again reference Figure 8, which shows the PSFs that form in the telescope image plane for telescope locations offset relative to the astroid.

To obtain the final intensity distribution in the telescope image plane, we simply sum the contributions of each point in the source plane, using the unique PSF that a telescope would see corresponding to that point. To do this, we compute the PSF over a grid of telescope positions offset relative to the optical axis. It is equivalent to consider a point source located at \((x', y')\) in the source plane and a telescope centered on the optical axis \((x = 0, y = 0)\) as it would be to consider a point source on the optical axis \((x' = 0, y' = 0)\), and have the telescope center offset by negated and plate-scale-reduced coordinates \((c_x = x'_n \frac{d_l}{d_S}, c_y = y'_n \frac{d_l}{d_S})\). Thus the relationship between source plane pixel indices and corresponding telescope centering offsets is

\[
c_x = \frac{d_l}{d_S} R_{Earth} \left(2 \frac{n_x}{N_x} - 1 \right),
\]

\[
c_y = \frac{d_l}{d_S} R_{Earth} \left(2 \frac{n_y}{N_y} - 1 \right).
\]

With the vectors of centering offsets forming a grid of telescope locations \((c_x, c_y)\), we compute the PSF in the telescope image plane using Equation (25), where the pupil transmission function is modified so that \(P(x) \in \{(x, y) | \sqrt{(x - c_x)^2 + (y - c_y)^2} \leq D_{astroid}/2\}\) and the telescope is properly centered around \((c_x, c_y)\). As a computational shortcut, \(E(x)\) in Equation (24) is only computed for points inside the pupil where \(P = 0\) to save on computational cost. Additional computational cost is saved by reverting to the analytic monopole expression for \(E(x) = J_0(k \sqrt{2R_{Earth}/r})\) when the telescope is sufficiently far from the optical axis. Since Figure 9 demonstrates that the monopole approximation starts to be very accurate for \(\rho = \sqrt{c_x^2 + c_y^2} \geq 4D_{astroid}\), this is the condition we check before falling back on the fast analytic computation for \(E(x)\). The final intensity distribution \(I_{extended}\) is computed as a matrix product of the PSF grid \(PSF_{c_x,c_y}\) and the source image vector \(I_{n_x,n_y}\),

\[
I_{extended}(\alpha, \beta) = PSF_{c_x,c_y}(\alpha, \beta) I_{n_x,n_y}.
\]

Thus with the PSF grid previously computed and in memory, the forward model of imaging an extended source can be represented as a linear transformation of the input intensity distribution. However, for this to be achieved with a single matrix operation, the two-dimensional indices of source coordinates indexed by pixel numbers \((n_x, n_y)\) and the point-source offsets corresponding to telescope centering offsets \((c_x, c_y)\) must be unraveled into a single monotonically increasing index. This choice of ordering is arbitrary and can be done in a combinatorially large number of unique ways, although we find unraveling indices one row at a time to be the simplest. Figure 11 was created to demonstrate imaging of extended sources.

In general, the size of the astroid is controlled by the quadrupole moment \(I_2\) and the solar co-latitude \(\beta_S\), see Equation (20), while the size of the image \(D_{image}\) is controlled by the physical dimension of the source, in this case \(R_{Earth}\) as well as the geometry of the problem, including the distances.
between the source and lens $d_S$, and the telescope and the lens $d_L$. The exact relation is

$$D_{\text{image}} = 2R_{\text{Earth}} \frac{d_L}{d_S}. \quad (44)$$

For extended sources of different sizes, there are three regimes that are important to consider. One limit is that of very small sources, or $D_{\text{image}} \ll D_{\text{astroid}}$. In this case, one may think of the extended objects as essentially just a point, and the resulting image formed will be that of the PSF. The other limit is that of very large objects, or $D_{\text{image}} \gg D_{\text{astroid}}$. In this case, the majority of point sources in the SGL image plane will be farther from the optical axis than the dimension of the asteroid, and the resulting PSF in the telescope image plane for each location on the extended object will simply be two points. One can think of this limit as being equivalent to taking $D_{\text{astroid}} \to 0$, or $\beta_S \to 0$, where the telescope is situated directly above or below the solar rotational axis and the projection of the quadrupole moment into the plane of observation becomes negligible. In this limit, the higher-order effects from the multipole moments can be neglected, as nearly all points appear to be reimaged into two points in the telescope image plane, which is identical to the monopole-only analysis described in Madurowicz (2020).

The last regime is that where the dimension of the reimaged target and asteroid have similar sizes, or $D_{\text{image}} \sim D_{\text{astroid}}$. In this regime, one cannot ignore the higher-order effects of the multipoles, as many of the points on the target will be reimaged multiple times, some of the points may sit on the critical caustic, and some of the points may be outside the asteroid entirely. Thus, the final intensity distribution formed in the telescope image plane is a superposition of the multifaceted PSF distribution and intensity distribution from the source.

4. Sources of Contamination

4.1. A Model of the Earth and Sun

Actual physical observations using the SGL will face an extreme difficulty dealing with contamination of source signal with noise from multiple sources along the line of sight, including the Sun and its corona, zodiacal dust in the solar system, and additional nontarget sources in the SGL field of view (FOV). This section investigates the expected brightness from the Sun itself as a primary source of noise, compared to
the expected brightness of a potential, Earth-like, target source by collecting data and models built from observations of the Earth and Sun and scaling them to match the geometry of the SGL problem.

The first Earth data set used in this study is a series of images taken from the Advanced Baseline Imager on board the GOES-R weather satellite (GOES-R Calibration Working Group & GOES-R Series Program 2017). The data includes extremely high-resolution images of the Earth taken from geostationary orbit, across a wide range of wavelengths, from 0.47 μm in the visible light spectrum, to 13.27 μm in the mid-infrared. The specific data set used in this study is the level-1b radiances product, which has been processed and calibrated to represent authentic top-of-the-atmosphere radiances in units of \( \text{W m}^{-2} \text{sr}^{-1} \) or \( \text{mW m}^{-2} \text{sr}^{-1} \) (cm^{-1}) depending on if the channel is dominated by reflection of sunlight or thermal emission from the Earth itself. The GOES-R data is plotted in Figure 12.

While the GOES-R data will be useful to investigate the effect of nonuniform intensity distributions across different wavelengths, the wavelength coverage is not complete. With the goal of investigating higher spectral resolutions, we investigate the Earth model described in Robinson et al. (2011). The Earth spectral model describes disk-integrated specific intensities of the Earth across a massive range in wavelength space, from 0.1 μm to 200 μm, with spectral resolution \( R = \frac{\lambda}{2A} \) ranging from \( \sim 100 \) at the longest wavelengths to \( \sim 10^5 \) at the shortest.

The Earth spectral model also includes variations in brightness due to the inclination of the orbit and the orbital phase. Both the GOES-R data and the Earth spectral model are plotted in Figure 13 in units of spectral flux density or \( \text{W m}^{-2} \mu m^{-1} \) by multiplying by \( \int_0^{2\pi} \int_0^{\pi} \sin \theta \cos \theta d\theta = \pi \text{ sr} \). These units will be convenient for comparison with a model of the spectral flux density of the Sun.

To compare the brightness of the Earth and the Sun across many wavelengths, we use three separate solar models to act as validation and ensure consistency. The first and simplest model is that of a spherical blackbody with radius \( R_\odot \) and temperature \( T = 5778 \text{ K} \), to match the effective temperature of the solar photosphere (Carroll & Ostlie 2007). The blackbody spectrum is governed by the Planck equation

\[
B_\lambda(T) = \frac{2h\nu^2}{\lambda^5} (e^{h\nu/kT} - 1)^{-1},
\]

where \( h \) is Planck’s constant, \( c \) is the speed of light, \( k \) is the Boltzmann constant, and \( T \) is the blackbody temperature. Since the blackbody equation gives specific intensity at the photosphere, to calculate the flux density at the top of Earth’s atmosphere, one must rescale by the solid angle of the Sun as seen from Earth

\[
F_\lambda = B_\lambda \pi \left( \frac{R_\odot}{1\text{au}} \right)^2.
\]
The blackbody model is quite a good description of the radiation from the Sun, due to its physical elegance and analytic form, but is incomplete due to absorption in the solar atmosphere and additional complications. However, the Sun is bright and easily observable, so spectral data at high resolution is plentiful and of high significance. We use two separate solar spectral models that are validated with real observations. The first model is the National Oceanic and Atmospheric Administration (NOAA) Climate Data Record (CDR) solar standard irradiance (Coddington et al., 2015), and the second is the American Society for Testing and Materials (ASTM) E-490 solar standard, which is compiled from multiple unique observations (Kurucz, 1993; Fröhlich & Lean, 1998; Woods et al., 1996; Smith & Gottlieb, 1974) to serve as a reference spectrum across a large range of wavelengths. Both of these models are converted into identical units for comparison in Figure 13.

A few observations regarding the spectral models. The Sun is accurately described by a blackbody spectrum, with small perturbations due to absorption and emission lines that are mostly invisible at the plotted spectral resolution. The Earth model has two peaks, one in the visible spectrum corresponding to the reflection of sunlight, and one in the infrared corresponding to the thermal emission from the planet itself. Additionally, the Earth’s spectrum is significantly altered by absorption in its atmosphere, with many valleys in different wavelength regions indicative of particular molecules such as H2O and CH4.

The previous figure depicts the spectral flux density that an observer would see for the Sun and the Earth if that observer was located at the top of Earth’s atmosphere. However, this arrangement is not equivalent to the geometry of imaging a target planet through the SGL, and so these curves must be rescaled to the appropriate distance using the inverse square law. Additionally, the flux of the planet is magnified by imaging through the SGL, by an amount that depends on its location on sky as well as by the exact alignment between telescope, target, and the Sun. Figure 9 shows that reasonable values of the magnification can range from $\mu = 10^7$ to $10^9$. Additionally, one may obscure the light from the Sun using a starshade (Cash, 2011) or coronagraph (Lyot, 1939), which may need to be specifically designed to obscure the extended disk of the Sun (Ferrari et al., 2009), but could significantly reduced the observed flux from the Sun.

To make a fair comparison of the brightnesses with all of these factors considered, Figure 14 was created. In this Figure, there are three curves for the Sun and the Earth each. Each starts with the Robinson Earth spectral model or the ASTM-E solar spectral model and rescales the flux amplitude depending on different assumptions. From top to bottom, the solar curves are rescaled to $d_L = 650$ au without a coronagraph, $d_L = 650$ au with a coronagraph with a contrast of $10^{-9}$, and $d_L = 10$ pc behind a similar coronagraph. The rescaling factor is simply a multiplication by contrast $\times \left(\frac{d_L}{1\text{au}}\right)^2$. This level of coronagraphic sensitivity has not yet been reached but is projected to be possible by the HABEX mission concept (Gaudi et al., 2018). The three curves for the Earth, from top to bottom, all correspond to a target distance $d_S = 10$ pc, but with $\mu = 10^9$, $10^7$, or 0, corresponding to a possible magnified source behind the SGL and an unmagnified source that one would see without the SGL.

The principal observation this Figure was designed to showcase is that in terms of flux contrast between the target
source and the primary noise source, using the SGL is comparable to direct imaging. The spectral flux density essentially describes the total number of photons that would be captured by the telescope over some bandpass, but is agnostic to the distribution of that light in the focal plane of the telescope, which is discussed in 4.4. The dashed pair of curves correspond to a reasonable expectation of the flux for the standard SGL, while the dotted pair of curves correspond to a reasonable comparison for direct imaging. The ratio of the flux between these two objects gives the contrast as a function of wavelength, which is quite similar. However, the major advantage that the SGL arrangement has over direct imaging is in the absolute brightness of both the target and noise source, being approximately 10^7 times brighter for both the target and the solar/stellar noise contribution. This advantage may prove useful to obtain extremely high-dispersion spectra of an exoplanetary atmosphere, since high dispersion implies correspondingly smaller signal to noise per bin, and high-dispersion instruments typically struggle with low throughput (Vogt et al. 1994).

### 4.2. The Local Astronomical Field

Aside from the light from the target source and the Sun itself, the Einstein ring contains light from all astrophysical sources in the SGL’s FOV. This includes primarily the light from the target planet’s host star, but also includes the probability of faint distant sources that are coincidentally close to alignment with the optical axis. In Figure 9 in Section 3.3, we show the average magnification $\langle \mu \rangle_p$ over the telescope pupil as a function of the telescope’s radial offset in the SGL image plane $\rho = \sqrt{c_x^2 + c_y^2}$. For radial separations significantly greater than the dimension of the astroid, the average magnification is a power law with index $-1$, matching the expectation for a pure monopole SGL.

$$\log(\langle \mu \rangle_p) \approx 8.88 - \log \left( \frac{\rho}{1 \text{m}} \right), \quad \rho \gg D_{\text{astroid}}$$

Since a host star at $d_5 = 10$ pc with a target planet at $a = 1$ au has an angular separation in the source plane $\delta \theta = \frac{a}{d_5} = 0.1$ arcsec, this is equivalent to a telescope radial offset in the SGL image plane of $\rho = d_1 \delta \theta = a \frac{d_1}{d_5} \approx 4.7 \times 10^7$ m and thus an average magnification of $\langle \mu \rangle_p \sim 16$. Figure 15 demonstrates the relative contribution across various wavelengths the host star will have to the observed flux of the target planet and the post-coronagraphic solar contribution.

Examining Figure 15, it is apparent that the target planet’s host star contributes a relatively significant amount of the total flux intercepted. In optical wavelengths, the total contrast between the target planet and its host star is about four orders of magnitude, while at thermal IR wavelengths, it is closer to one order of magnitude. This flux will be spatially separated in the telescope image plane, and as the host star is far off-axis, it will be reimaged into two points on opposite sides of the Einstein ring, similar to in the first panel in Figure 11, which
V magnitude cont.
German, as Doppler broadening of the photosphere absorption lines from the velocity of free electrons creates the appearance of a continuum spectrum. E- is for “emission.” The F- and K-coronal contributions can be reliably separated using the difference in spatial dependence between dust and free coronal contributions can be reliably separated using the posterior probability distributions and covariances. The joint analysis in black constrains the model to the region of overlap between the data sets.

The F- corona is largely considered to be a natural extension of the zodiacal light at smaller angular separations or elongations $\epsilon$ with regard to the Sun, and contains a wavelength dependence that increases toward long wavelengths due to thermal emission from the grains themselves in addition to the reflection of sunlight (Kimura & Mann 1998). However, interpretation of the F- coronal brightness and the corresponding dust distribution is complicated by inversion of the line-of-sight integral. Theoretical predictions for the near solar environment have speculated that the F- corona may decrease strength in the vicinity of the Sun (Lamy 1974; Mukai et al. 1974), due to orbital differentiation from Poynting–Robertson drag, pressure drag, radiation pressure, and sublimation of the grains. Observations from the WISPR instrument on board the Parker Solar Probe (Howard et al. 2019) have provided the first observation of such an F- coronal decrease, inferred from brightness measurements at various elongations at different heliocentric distances between 0.166 and 0.336 au.

Extensive brightness and color measurements of the solar corona have been taken and compiled previously, and for this work, we cross-reference multiple sources in the literature. By combining the data from Figure 2 in Mann (1992), Figure 1 in Kimura & Mann (1998), Figure 1 in Howard et al. (2019), and Figure 9 in Boe et al. (2021), we are able to analyze the solar coronal brightness across a wide range of elongations $\epsilon$. The data are reproduced in Figure 17(a) for convenience. To build a model of the solar coronal brightness, we perform a joint analysis of the various data sets, by fitting for the maximum likelihood power-law model of the data

$$\log \left( \frac{B_{\text{corona}}}{B_\odot} \right) = A + a \log \left( \frac{\epsilon}{R_\odot} \right) - 1,$$

which is linear in log coordinates, and by choosing the base of the exponent as $\frac{\epsilon}{R_\odot} - 1$, the edge of the solar disk at $\epsilon = R_\odot$ corresponds to the origin of the power law. The resulting posterior distributions on the power-law slope $a$ and intercept $A$ are shown in Figure 17(b). The covariances of the inferred parameters show constraints given by different individual data sets, while the joint analysis is restricted to the region of overlap between all data sets. The best-fit parameters as a result of this analysis are $A = -7.836 \pm 0.05$ and $a = -2.071 \pm 0.055$. The resulting power-law curve is shown in black in Figure 17(a).

Most of these measurements are taken with visible light wavelengths, although some effort has been made to make measurements in the near-infrared as well, and the results are compiled into Figures 5 in Boe et al. (2021) and Figure 2 in Kimura & Mann (1998). In Figure 2 of Kimura & Mann (1998), the scatter is large; there is a slight increase at longer wavelengths, going from roughly $10^{-8} B_\odot$ at $\lambda = 0.3 \mu$m to $10^{-8} B_\odot$ at $\lambda = 3 \mu$m for elongations of $\epsilon = 4 R_\odot$. Figure 5 of Boe et al. (2021) reports relative intensities over $\lambda \in [0.5, 0.8] \mu$m, and measures a spectral power-law index of $B_{\text{corona}} \propto \lambda^{-0.91 \pm 0.07}$. This reddening effect could be the result of thermal emission from the particles themselves, as well as preferential directional scattering of thermal emission from the Sun.

Figure 7 in Boe et al. (2021) also demonstrates the azimuthal variation in intensity, with isointensity curves fluctuating from $\epsilon \in [2.0, 2.4] R_\odot$ at $10^{-8} B_\odot$ to $\epsilon \in [1.3, 1.5] R_\odot$ at $10^{-7} B_\odot$. To show the azimuthal variation that can reasonably be expected here, we plot data from the HAO COSMO K-coronagraph (de Wijn et al. 2012; Hou et al. 2013) in Figure 18 with a black-hot color map.

The data in Figure 18 are level-2 PB calibrated data downloaded directly from the HAO website. Some of the many calibration steps include camera nonlinearity, dark, and gain corrections, polarization demodulation, distortion correction, combining images from multiple cameras, removal of sky polarization, and correction for daily variation in sky transmission. From examining the image, it is clear that the observations are roughly in correspondence with the power-law analysis.
described previously, with the coronal brightness ranging from \(10^{-6}B_\odot\) at \(\epsilon \sim R_\odot\) to \(10^{-3}B_\odot\) at \(\epsilon \sim 2R_\odot\), although it is clear that the coronal brightness varies depending on azimuthal angle.

4.4. Implications for Instrumental Design

4.4.1. Coronagraph Design

Assessing the feasibility of imaging the Einstein ring at extremely high contrast is complicated by the redistribution of light in angular coordinates. The Einstein Ring has an approximate radius of \(\sim 1''\), while a host star/target planet separation for an orbit of \(a = 1\) au at \(d_S = 10\) pc is an order of magnitude smaller at 0.1''. Designing a coronagraph to reach high contrast at small inner working angle is challenging (Spaan & Greenaway 2006; Mawet et al. 2012), and is typically investigated in the context of direct imaging of point sources. This is complicated by the extended solar disk and the minimal separation between the edge of the disk and Einstein ring. The angular separation between the edge of the solar disk and the center of the Einstein ring for a source with \(d_S >> d_L\) is (Hippke 2020)

\[
\delta \theta \approx \sqrt{\frac{2R_\odot}{r}} - \frac{R_\odot}{r},
\]

which reaches a maximum value of \(\sim 0.4''\) at a heliocentric distance of \(r \sim 2300\) au independently of wavelength. Specialized coronagraph designs for this particular case have only recently been investigated. An original suggestion from Turyshhev and Toth was the use of an annular coronagraph (Turyshhev & Toth 2020), which would only allow light in from the region of the Einstein ring while blocking light from the solar disk and outer corona. However, as we have seen, the Einstein Ring itself contains contamination from the target planet’s host star, as well as potentially other objects in the field and so this may be insufficient. Loutsenko & Yermolayeva (2021) recently suggested a design that is a combination of an annular mask and a bar mask, with the intention of obstructing light from both the solar disk as well as the magnified off-axis contribution from the target planet’s host star. Their design claims to reach contrasts of \(10^{-8}\) between the solar disk and the region of the Einstein Ring, with similar performance for off-axis sources behind the bar mask. However, neither of these coronagraphs on their own can mitigate contamination from the region of the Einstein ring itself, which includes light from the corona and background sources. A spectroscopic mitigation strategy similar to the method described in Ruffio et al. (2019) could model all sources of contamination in the final image, using spectral differences between noise sources to obtain high contrast in the absence of or in combination with a coronagraph.

4.4.2. Background and Foreground Mitigation

In the process of designing a mission to the SGL, one should be prepared to deal with the possibility of a background object contaminating the Einstein ring. The target planet’s host star will certainly contaminate two points along the Einstein Ring, and a possible background star could contaminate two additional points in a similar fashion. Even contamination from nontarget planets in the target solar system may be relevant if the angular separation between the target and contaminating planet is small enough. Selecting a target planet from a list of multiple potential target planets being considered should choose the safer option, which is farther from the galactic center. This will lower the probability of a background star, but the probability cannot be eliminated entirely. Early and deep observations of the near target region could provide advanced evidence for a clear background, but a robust design should include a contingency plan for mitigating this contamination if the observation is unlucky, as galactic proper motion could place a background object in the FOV sometime between launch and observation. By spatially separating the light in the Einstein ring, either a hardware or a software mask could be applied to the region around the points of contamination; although, scattered light from this unobstructable source will ultimately limit the possible signal-to-noise ratio (S/N) obtainable in unobstructed pixels.

For proper mitigation of the corona and zodiacal light as astrophysical foregrounds, it will be necessary to model the temporal and azimuthal dependence of the coronal brightness to extract the full information contained in the Einstein ring. One observational strategy to achieve this is a differential measurement. A single telescope equipped with an integral field spectrograph could make spatially and spectrally resolved measurements of the Einstein ring and the corona as well as the region of the corona nearby the Einstein ring. By measuring the coronal brightness in the vicinity and extrapolating a model into the region of overlap, a differential measurement may be possible as a mitigation strategy. Another type of instrument design could achieve a similar measurement instead using two separate spacecraft: one being the primary observation telescope at the focal plane of the SGL and the other being an additional companion telescope observing the solar corona from a closer heliocentric distance, where the Einstein ring cannot form at \(d_L \lesssim 550\) au. Simultaneous observations could enable a differential measurement of the brightness of the coronal region with and without the Einstein ring present, and would not rely upon fitting a nearby coronal region and extrapolating. However, this concept would suffer from complexities involved with coordinating observations on different telescopes separated by a great distance.

Another alternative strategy for mitigation of the corona is by sending the observation craft to a significantly larger heliocentric distance of \(d_L = 2300\) au for observing. Both the angular size of the Einstein ring and the solar disk shrink at
larger \( d_L \), but the Einstein ring shrinks proportionally to \( d_L^{-1/2} \) while the solar disk shrinks proportionally to \( d_L^{-1} \). This means at the optimal observation distance \( d_L = 2300 \) au, the angular separation between the edge of the solar disk and the center of the Einstein ring is maximal, and the foreground contamination will be minimal as the Einstein ring overlaps with the much fainter F-corona rather than the bright K-corona at small separations.

The solar foreground curves shown in Figure 19 are a combination of the coronal model fit in Section 4.3 and a fifth-order polynomial limb-darkening model for the brightness of the solar disk described in Hestroffer & Magnan (1998) and using data from Pierce & Slaughter (1977) and Neckel & Labs (1994). The data correspond to measurements at \( \lambda \sim 0.58 \) \( \mu \)m, but additional measurements at wavelengths \( \lambda \in [0.78, 2.44] \) \( \mu \)m are reported in Figure 6 of Pierce (1954), which demonstrates the increasing effects of limb darkening at longer wavelengths. The specific intensity of the Earth \( B_{\text{Earth}} \) is taken from the model in Robinson et al. (2011), and the specific intensity of the Sun is taken from the blackbody model described in Section 4.1. The specific intensity ratio \( B_{\text{Earth}} / B_\odot \sim 10^{-6} \) is not the typical flux contrast ratio of the Earth and Sun \( \sim 10^{-10} \), as it lacks the additional factor of \( (R_{\text{Earth}} / R_\odot)^2 \sim 10^{-4} \) to account for the difference in the solid angle of the bodies. A version of Liouville’s theorem of conservation of phase space volume applied to gravitational lenses (Misner et al. 1973) proves that gravitational lenses preserves surface brightness or specific intensity up to a frequency shift due to gravitational redshift, while the magnification acts to modify the apparent solid angle of the object. This justifies our comparison of the specific intensities of the background and foreground along the line of sight, and demonstrates the advantage of observing at a farther heliocentric distance.

5. Spectral Image Reconstruction

Investigating different instrument designs such as broadband coronagraph masks and multispectral imaging spectrometers will require a full forward model of all sources in the FOV, emission from the solar disk and corona, and light from the target planet, host star, and other contaminating objects in the field, like additional planets or background stars. Such a forward model is within reach by combining all of the models described in this paper, but computationally quite expensive and outside the scope of this paper. An approximation like the quartic polynomial solution described by Turyshev & Toth (2021d) could expedite this computation time at the cost of degraded accuracy, but a complete forward model will be necessary as the mission concept matures and specific parameters of the problem can be further narrowed down to construct instrumental design requirements on the performance of the craft. This complete model could be used to simulate instrument design concepts and test post-processing algorithms to extract information from the Einstein ring in the presence of realistic astrophysical noise sources in a spatially and spectrally resolved manner.

For now, in the absence of such a complete model, we can still test the reconstruction power of such observations in a toy environment. The forward model of the lensing problem in Equation (43) is a linear transformation of the input intensity distribution \( I_{nn} \). The linear transformation is given by the matrix \( \text{PSF}_{cc,cc} \), which describes the impulse response of the SGL to points laterally offset from the optical axis and captures all of the complexity of lensing in the presence of the solar gravitational potential’s zonal spherical harmonics. To test the reconstruction power of an ideal measurement of the extended source’s Einstein ring \( I_{\text{extended}} \), we can pseudo-invert the forward model using the singular value decomposition. The SVD can be represented as a matrix factorization of the form (Hendler & Shrager 1994)

\[
\text{PSF}_{cc,cc} = U \Sigma V^*. 
\]

Where \( U \) and \( V \) are square, real, orthonormal, and unitary matrices, and \( \Sigma \) is a diagonal matrix containing the singular values. This is a useful method to decompose the linear transformation \( \text{PSF}_{cc,cc} \), and because \( U \) and \( V \) are unitary operators, they can be thought of as acting to rotate the basis elements of the space, while \( \Sigma \) acts to stretch the rotated vector along the intermediary axis. This combination of rotate, stretch, derotate naturally allows one to find the pseudo-inverse, by the means of de-rotating, un-stretching, and re-rotating, written...
mathematically as (Gregoric 2001)

\[
\text{PSF}^* = V \Sigma^{-1} U^*.
\]

The inverse of \( \Sigma \) is simply the inverse of each of the diagonal elements, which is why numerically small singular values can cause the pseudo-inverse to be discontinuous, and many routines to find the pseudo-inverse will introduce a cutoff, below which the singular values are set to zero. This cutoff parameter is known as \( r_{\text{cond}} \), and is typically set to values near the machine epsilon, where numerical noise can become problematic.

To investigate the properties of this reconstruction schema across multiple different wavelengths, we compute the PSF forward model matrix at a spatial sampling of \( 256 \times 256 \) pixels across a finite wavelength grid, with \( \lambda \in \{0.5, 1, 2, 4, 8, 16\} \) µm, and interpolate the PSF grids to match the wavelengths of the GOES-R radiances described in Section 4.1, which act as the input intensity distributions and have been downsampled to a resolution of \( 64 \times 64 \) pixels. The GOES-R data ranges in wavelength from \( \lambda \in \{0.47, 13.27\} \) µm, and we interpolate using the following process. For the 0.47 µm channel, we directly use the PSFs computed at 0.5 µm for simplicity and the increased computational cost of integrating Equation (24) at short wavelengths. For the rest of the channels, since each lies between two wavelength grid points, we linearly interpolate the resulting PSF grids with a weighting factor corresponding to the linear distance between the wavelength \( \lambda \), and the two nearest grid points \( \lambda_0, \lambda_1 \). The resulting model is

\[
\text{PSF}_\lambda = \text{PSF}_{\lambda_0} \frac{\lambda - \lambda_1}{|\lambda_1 - \lambda_0|} + \text{PSF}_{\lambda_1} \frac{\lambda - \lambda_0}{|\lambda_1 - \lambda_0|}.
\]

To investigate the properties of this reconstruction schema across multiple different wavelengths, we compute the PSF forward model matrix at a spatial sampling of \( 256 \times 256 \) pixels across a finite wavelength grid, with \( \lambda \in \{0.5, 1, 2, 4, 8, 16\} \) µm, and interpolate the PSF grids to match the wavelengths of the GOES-R radiances described in Section 4.1, which act as the input intensity distributions and have been downsampled to a resolution of \( 64 \times 64 \) pixels. The GOES-R data ranges in wavelength from \( \lambda \in \{0.47, 13.27\} \) µm, and we interpolate using the following process. For the 0.47 µm channel, we directly use the PSFs computed at 0.5 µm for simplicity and the increased computational cost of integrating Equation (24) at short wavelengths. For the rest of the channels, since each lies between two wavelength grid points, we linearly interpolate the resulting PSF grids with a weighting factor corresponding to the linear distance between the wavelength \( \lambda \), and the two nearest grid points \( \lambda_0, \lambda_1 \). The resulting model is

\[
\text{PSF}_\lambda = \text{PSF}_{\lambda_0} \frac{\lambda - \lambda_1}{|\lambda_1 - \lambda_0|} + \text{PSF}_{\lambda_1} \frac{\lambda - \lambda_0}{|\lambda_1 - \lambda_0|}.
\]

The input intensities, forward model of the Einstein ring, and corresponding reconstruction for all of the channels in the GOES-R data are shown in Figure 20.
Examining the results of Figure 20, it is clear that in the limit of large S/N, measurements of the Einstein ring can be used to reconstruct the input intensity distribution. It is important to note that this example relies on superresolution, since the number of reconstructed pixels is larger than the number of independent resolution elements available in the focal plane, which is only possible with very high S/N. For the reflection-dominated channels in the upper half of the figure, the input intensity distribution is a complex function of the distribution of clouds in the planetary atmosphere, due to their high albedo. In the thermal-emission-dominated channels, the surface emits more than the cold clouds at high altitudes, and the intensity distribution is mostly a full disk, with cold pockets due to absorption from the clouds. This difference in input intensity manifests as a unique distribution of light in the Einstein ring, visible most apparently in the reflection-dominated channels, where the bright clouds on the planet’s surface result in a nonuniform azimuthal distribution of intensity in the ring. These azimuthal variations encode the information of the original image. The effect is still relevant for the long wavelength channels but less noticeable by eye. Since the distribution of thermal emission is nearly a uniform disk, the resulting Einstein ring profile is mostly a quadrupolar pattern.

This reconstruction scheme allows one to easily adjust the S/N of the observations and test how the reconstructor fails in the presence of noise. The extremely high S/N = 2 × 10⁷ shown in Figure 20 is unrealistic considering all of the sources of contamination, but represents the limit of an essentially infinite S/N. In this limiting case, only the longest wavelength channels have significant degradation across the entire image, as the PSF grows with wavelength and washes out fine structure hidden within the ring. However, degradation can be seen in shorter wavelength channels as well, specifically around the edge of the caustic. To highlight this effect, another reconstruction was done at S/N = 4000 for λ = 0.5 μm. This reconstruction uses visible color imagery of the Earth as the input intensity distribution, and neglects small discrepancies in wavelengths between the different color channels for simplicity. This reconstruction is shown in Figure 21.

For points on the surface of the planet that are inside the projection of the caustic, the lens reimages those points into four distinct locations around the edge of the Einstein ring in an asymmetric fashion demonstrated in Figure 8, while points outside the caustic are reimaged symmetrically to two points, similarly to the lensing of a point mass or the monopole-term only. The implications of this were understood in Madurowicz (2020), where it was demonstrated that the monopole-only forward model of the lens has a certain antisymmetric null space where it is nearly impossible to tell the difference between an input image and the same image rotated by 180° from just the azimuthal profile of the ring. This antisymmetric null space then requires a certain degree of navigational symmetry breaking, where the craft moves laterally to scan over the image in the SGL focal plane to make inferences about the original intensity distribution. This is problematic for rapidly evolving images such as the distribution of clouds across a planet that is rotating and experience changing weather patterns, and especially difficult if long integration times are required to reach sufficient S/N in the presence of foreground contamination.

However, in the presence of the quadrupole moment, this is not the case. The asymmetric reimaging of the points inside the critical caustic allow for direct disentangling of the original intensity distribution inside the projection of the caustic. This demonstrates a fundamental trade-off for imaging and spectroscopy with the SGL, which must be considered when...
achieve greater departure velocity, such as a solar sail ∼ by a number of engineering challenges that are difficult but not insurmountable. Operating a telescope at a minimum heliocentric distance of 650 au would require extreme patience with conventional and existing rocket technology, with travel times on the order of ~100 yr, or advancements in propulsion to achieve greater departure velocity, such as a solar sail (Turyshev et al. 2020). Furthermore, navigational acceleration of the order of $dv \lesssim 80 \text{m/s} + 6.7 \text{m/s}^2 \text{year}$ is needed to obtain and maintain alignment of the craft, Sun, and a fixed point on the sky in the worst-case scenario. The fixed cost term is due to the relative lateral motion between the craft and solar system barycenter associated with the escaping hyperbolic trajectory’s conservation of specific angular momentum, and the linear cost is associated with perturbations of the solar center of mass with respect to the solar system barycenter due to the gravitational interactions with the other planets in the solar system. Additional acceleration may be needed due to the target planet’s system’s proper acceleration and relative orbital motion; but this is complicated and could potentially reduce the needed acceleration depending on the alignment of the multiple bodies and orbits. Furthermore, no existing exoplanetary characterization methods have obtained sufficient precision to locate the position of the image of the target exoplanet to a comparable degree as the size of the projected image itself, even with GRAVITY-interferometer levels of astrometric precision. Complete astrometric coverage of the target planet will assist in constraining the image location posterior and predicting its future location, which is necessary for navigation.

The optical properties of the SGL depend crucially on the solar photosphere has a specific intensity $B_{\odot}$ while the solar corona is $10^{-6}$ to $10^{-9} B_{\odot}$ depending on sky location. Background stars or other astronomical objects could be reimaged into the Einstein ring, and the probability of observing a background star depends on the target’s galactic coordinates, with the direction of the galactic center being least optimal. A complete strategy to characterize, model, and mitigate all sources of noise has yet to be demonstrated. A pure coronagraphic approach would be limited to a narrow bandpass, and a major advantage of the SGL over direct imaging is the enormous quantity of available photons for high-dispersion spectroscopy. A starshade could be an achromatic solution to obtain high contrast across a wide bandpass. The strategy of spectrally and spatially resolving the light in the vicinity of the Einstein ring potentially has the capability to disentangle all sources of noise, but our results are not conclusive. A complete instrument model is outside the scope of this work. As the desire for knowledge locked behind immense observational barriers increases, further understanding of instrumentation from a universal perspective will be necessary.

6. Conclusion

The prospect of leveraging the SGL to spatially and spectrally resolve the surfaces and atmospheres of extrasolar planets is exciting and fascinating. The combination of extremely high angular resolution and magnification provides unprecedented observational information inaccessible to any other instrument concept. However, the concept is constrained by a number of engineering challenges that are difficult but not insurmountable. Operating a telescope at a minimum heliocentric distance of 650 au would require extreme patience with conventional and existing rocket technology, with travel times on the order of ~100 yr, or advancements in propulsion to achieve greater departure velocity, such as a solar sail (Turyshev et al. 2020). Furthermore, navigational acceleration of the order of $dv \lesssim 80 \text{m/s} + 6.7 \text{m/s}^2 \text{year}$ is needed to obtain and maintain alignment of the craft, Sun, and a fixed point on the sky in the worst-case scenario. The fixed cost term is due to the relative lateral motion between the craft and solar system barycenter associated with the escaping hyperbolic trajectory’s conservation of specific angular momentum, and the linear cost is associated with perturbations of the solar center of mass with respect to the solar system barycenter due to the gravitational interactions with the other planets in the solar system. Additional acceleration may be needed due to the target planet’s system’s proper acceleration and relative orbital motion; but this is complicated and could potentially reduce the needed acceleration depending on the alignment of the multiple bodies and orbits. Furthermore, no existing exoplanetary characterization methods have obtained sufficient precision to locate the position of the image of the target exoplanet to a comparable degree as the size of the projected image itself, even with GRAVITY-interferometer levels of astrometric precision. Complete astrometric coverage of the target planet will assist in constraining the image location posterior and predicting its future location, which is necessary for navigation.

The optical properties of the SGL depend crucially on the target planet’s position on the sky; so ultimately, the target selection and mission design processes should become interconnected. The presence of a significant quadrupole moment due to oblateness from the rotation of the Sun introduces a critical caustic in the lensing problem, which scales proportionally to $\sin^2 \beta_s$, the angle between the solar rotational axis and the line of sight. This effect reduces the magnifying power of the lens and modifies the distribution of light in the Einstein ring. For planets in alignment with the solar rotation axis, the effect disappears, and the lensing problem reduces to a monopole-only gravitational potential, but at the solar equator, the effect is maximal. The antisymmetric imaging of points inside the critical caustic can be exploited to break the rotational symmetry of the monopole-only lensing forward model, allowing for direct disentanglement of the original intensity distribution from single high-S/N measurements of the Einstein ring. However, for this to be possible, the critical caustic must be as large as the projection of the planet, which is only possible for Earth-sized planets at a minimum distance of 100 pc. Furthermore, the large S/N required for superresolution limits the practicality of high-resolution image reconstruction using the quadrupole-based symmetry breaking on its own, especially at long wavelengths. Low-resolution image reconstruction would not require superresolution and would be possible at smaller S/N. However, other strategies exist that can break this degeneracy, which are especially relevant for nearby or polar-aligned targets. Scanning a single telescope across the projected image must account for temporal changes in the target illumination, which could be solved with a multitelescope swarm architecture. For single telescopes, it may just be simpler to observe a target during crescent or quarter-illumination phases where the degeneracy is irrelevant. However, this is only possible for reflection-dominated wavelengths during a fraction of the orbit depending on the target’s orbital inclination. Additionally, the rotation of the planet and atmospheric variability introduce additional complications in reconstructing images from observations taken over time. The combination of all of these effects should be investigated in the design of an optimal reconstruction architecture.

Obtaining high-S/N measurements of the Einstein ring will be difficult, as contamination from other sources along the line of sight obstruct the faint signal. For visible light, the disk of the solar photosphere has a specific intensity $B_{\odot}$ while the solar corona is $10^{-6}$ to $10^{-9} B_{\odot}$ depending on sky location. Background stars or other astronomical objects could be reimaged into the Einstein ring, and the probability of observing a background star depends on the target’s galactic coordinates, with the direction of the galactic center being least optimal. A complete strategy to characterize, model, and mitigate all sources of noise has yet to be demonstrated. A pure coronagraphic approach would be limited to a narrow bandpass, and a major advantage of the SGL over direct imaging is the enormous quantity of available photons for high-dispersion spectroscopy. A starshade could be an achromatic solution to obtain high contrast across a wide bandpass. The strategy of spectrally and spatially resolving the light in the vicinity of the Einstein ring potentially has the capability to disentangle all sources of noise, but our results are not conclusive. A complete instrument model is outside the scope of this work. As the desire for knowledge locked behind immense observational barriers increases, further understanding of instrumentation from a universal perspective will be necessary.
Turyshev, S. G., & Toth, V. T. 2022, *PhRvD*, **105**, 044012
Turyshev, S. G., Shao, M., Toth, V. T., et al. 2020, arXiv:2002.11871
Van Rossum, G., & Drake, F. L. 2009, Python 3 Reference Manual (Scotts Valley, CA: CreateSpace), https://docs.python.org/3/reference/
Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, *NatMe*, **17**, 261

Vogt, S. S., Allen, S. L., Bigelow, B. C., et al. 1994, *Proc. SPIE*, **2198**, 362
Woods, T. N., Prinz, D. K., Rottman, G. J., et al. 1996, *JGRD*, **101**, 9541
Zessewitsch, W., & Nikonow, W. 1929, *Natur*, **123**, 909
Zurlo, A., Vigan, A., Hagelberg, J., et al. 2013, *A&A*, **554**, A21