DIRECT DETECTION OF TERRESTRIAL EXOPLANETS:
COMPARING THE POTENTIAL FOR SPACE AND GROUND TELESCOPES

Roger Angel

Center for Astronomical Adaptive Optics, Steward Observatory, The University of Arizona,
Tucson, AZ 85721, USA, rangel@as.arizona.edu

1. INTRODUCTION

Direct detection of terrestrial exoplanets presents a challenge that lies beyond the scope of any current telescope. In this paper we consider the performance of different new telescopes with this potential. As our yardstick for sensitivity estimates, we take the earth in a solar system twin at 10 pc. At this range a dozen or so single stars like the sun would be accessible. Terrestrial planets could be detected through optical reflected starlight or thermal emission at around 10 µm. Both spectral regions carry spectroscopic information about the chemistry and biochemistry of a planet.

For all schemes under consideration and analyzed here, the detected signal will be dominated by sources other than the planet. These include zodiacal light, starlight that is imperfectly rejected, and thermal emission from the atmosphere, telescope and zodiacal dust. A fundamental limit to the sensitivity will be set by the photon noise arising from these sources.

The stellar, planetary and zodiacal fluxes of sun and earth seen from 10 pc are listed in table 1 in Janskys. Photon fluxes (which determine photon noise limits) are in direct proportion, given in photons/m²/sec by $1.51 \times 10^7 F(Jy) \Delta \lambda / \lambda$. Note that the flux ratio at 11 µm of $2 \times 10^7$ is a thousand times more favorable than in the optical.

In order to detect a planet with a maximum angular separation of only 0.1 arcsec, very efficient rejection of starlight is necessary. Stellar diffraction can be reduced with the aid of an apodized pupil or coronagraph, even to the $10^{-10}$ contrast level needed in the optical. [1,2].

Table 1. Fluxes from the solar system as seen from 10 pc. The Earth’s optical flux at maximum extension is taken to be $1.5 \times 10^{10}$ of the sun’s. The thermal flux is from ref [3]. The zodiacal brightness given is for 1 AU radius, as seen with the ecliptic plane tilted 30° to the line of sight [4].

<table>
<thead>
<tr>
<th>$\lambda$ (µm)</th>
<th>Star (Jy)</th>
<th>Planet (µJy)</th>
<th>Zodiacal dust (µJy/arcsec^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>40</td>
<td>6.0</td>
<td>5.2</td>
</tr>
<tr>
<td>0.8</td>
<td>67</td>
<td>10.1</td>
<td>9.6</td>
</tr>
<tr>
<td>2.2</td>
<td>30</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>11</td>
<td>1.64</td>
<td>340</td>
<td>1050</td>
</tr>
</tbody>
</table>
At 11 µm, even though the contrast is less taxing, very large filled aperture (~100 m) would be needed for coronagraph suppression. A useful alternative in this wavelength region is the nulling interferometer [5]. In space, one can build with relatively small cryogenic telescopes. The beams are combined so the starlight interferes destructively, while light from a close-by planet (in the right position) will interfere constructively. Very high suppression of the stellar disc can be achieved with multiple elements [6].

Even with the star strongly suppressed, zodiacal light or thermal emission may be problematic. The density of zodiacal dust around nearby stars is currently unknown, except for very dense discs like that around like β Pic. Zodiacal clouds down to solar levels will be detectable at ~1 AU from nearby candidate stars with the LBT interferometer [7]. Table 1 shows the surface brightness of a system like our own, twice that of solar system dust seen from earth, as the line of sight passes right through the zodiacal disc of the other system. The listed values are for a system viewed at 1 AU and at a typical angle of 30° to its ecliptic.

 Unless a system is viewed with a beam width much less than 1 AU, the dust will be brighter than the planet at optical as well as infrared wavelengths. Cloud structure could thus mask a real planet unless images are made with high resolution, for example by interferometric synthesis imaging in the infrared. This can be accomplished by a multi-element nulling interferometer, with resolution limited by the maximum baseline [8]. Even a low resolution interferometer with an asymmetric beam could distinguish a planet from a symmetric dust cloud.

The best observing conditions for new telescopes are found in space, where there is no atmospheric aberration and cryogenic optics can be used to eliminate telescope thermal background. However, much larger telescopes and interferometer elements can be built on the ground, where adaptive optics can obtain high-Strehl, diffraction-limited images using the bright star as reference. The limiting sensitivity for optical detection of terrestrial exoplanets from the ground was previously analyzed by Angel [9]. Thermal detection as well as optical may be possible, given a large enough aperture and favorable location. The best ground-based sites are on the Antarctic plateau, such as Dome C [10].

The goal of this paper is to compare on the same basis all the options for terrestrial planet detection, ground and space, optical and thermal. Section 2 gives an analysis of the photon noise sensitivity limits set by background light and the residual stellar halo after optimal correction with adaptive optics. Section 3 gives the strength of background sources and atmospheric properties at different sites. In section 4 we derive the limiting sensitivities for specific telescope and interferometer concepts, compiling all the results in a single table. In section 5 we consider the extent to which the required tolerances and stability can be met in practice.

2. ANALYSIS OF SENSITIVITY LIMITS SET BY PHOTON NOISE

In this section we develop a consistent treatment of sensitivity limits set by photon noise for space and ground systems at optical and thermal wavelengths.

2.1 Diffuse Background

In the simplest case, the exoplanet will be superposed on a brighter uniform background. The signal/noise for detection will be determined by photon shot noise of this background, and thus by the width of the telescope beam. As we discuss below, in every case we can expect images of high Strehl ratio. Though not currently obtained with ground-based adaptive optics at near-infrared wavelengths, high Strehl should be achievable with systems tailored to take advantage of the bright reference star flux. The PSF in general will not be an Airy function, because apodizing or coronagraphic systems are used. We will make the simplifying assumptions that the PSF is Gaussian with FWHM \( \lambda/D_{\text{eff}} \), and Strehl=1, and also that the image better than critically sampled with a noise free detector. The optimum signal/noise ratio for a weak signal against a stronger background, obtained by PSF fitting, is then:

\[
\frac{S}{N} = 0.66 \frac{C_{p}}{\sqrt{C_{b}}}
\]

where \( C_{p} \) is the total planet count and \( C_{b} \) is the total background count in solid angle \( (\lambda/D_{\text{eff}})^2 \) [11].

In general, \( D_{\text{eff}} \) will be less than the physical diameter of the telescope, because of pupil reduction at an apodizing or Lyot stop. The effective area will also be less than the primary mirror diameter. For simplicity we will assume the same effective diameter for flux collection as for resolution, i.e. effective area for flux collection = \( \pi D_{\text{eff}}^2/4 \). For fixed total observation time \( T \), the effective integration time \( t_{\text{eff}} \) may be reduced for some coronagraphs, because only part of the focal plane is free from diffracted light. We will assume that for very high rejection ratios half the search area is lost, effectively halving the on-source integration time.

For multi-element interferometers, the beam width is set by the diameter of the individual apertures. If the combining optics are effectively lossless, the total sky flux at the nulled output will be the same as that received by a single aperture [7]. For the planet, if it is located at the peak of constructive interference, the flux
will be that collected by the whole array. In this circumstance, the signal/noise ratio for the interferometer is the same as for a filled-aperture telescope of diameter $D_{\text{eff}}$ that has the same total collecting area as the array. But for a total observation time $T$ with a multi-element nulling interferometer, the effective on-source integration time $t_{\text{eff}} \sim T/4$, because at any moment only about $1/4$ of the search region around the star shows constructive interference.

With these assumptions, any single filled-aperture telescope or interferometer can be characterized by the parameters $D_{\text{eff}}$ and $t_{\text{eff}}$, and equation (1) can be rewritten for the signal/noise in the limit of photon noise in the background as

$$\left( \frac{S}{N} \right)_{\text{background}} = 0.011 \times 10^4 \frac{F_p}{\sqrt{I_b \lambda}} \frac{D_{\text{eff}}^2}{\lambda} \sqrt{q t_{\text{eff}}/\lambda} \quad (2)$$

where $F_p$ is the planet flux in Jy and $I_b$ the sky intensity in Jy/arcsec$^2$. $D_{\text{eff}}$ and $\lambda$ are in meters and $q$ is the effective system quantum efficiency, $t_{\text{eff}}$ the integration time and $\Delta \lambda$ the bandwidth.

In some cases, particularly optical detection on the ground, the dominant background will be from the stellar halo, rather than an independent diffuse source. It is convenient in this case to characterize the halo strength at the planet in terms of ratio $R_{\text{psf}}$ of PSF intensity at the central star to that at the position of the planet. In order for equation (2) to hold, $R_{\text{psf}}$ must exceed the ratio $R_{\text{bgr}}$ of star to background flux in $(\lambda/D)^2$, given by

$$R_{\text{bgr}} = 2.35 \times 10^{-11} \frac{F^*}{I_b} \left( \frac{D}{\lambda} \right)^2 \quad (3)$$

If the PSF background dominates, i.e. $R_{\text{psf}} < R_{\text{bgr}}$, and the fluctuations in this background are given by photon noise statistics, the signal-to-noise S/N is given by:

$$\left( \frac{S}{N} \right)_{\text{halo}} = 2270 \frac{F_p}{\sqrt{F^* D_{\text{eff}}}} \frac{q}{\sqrt{\lambda/\lambda}} R_{\text{psf}} t_{\text{eff}} \quad (4)$$

where $F^*$ is the stellar flux in Jy.

### 2.2 Halo intensity after correction of Kolmogorov turbulence

Adaptive optics for exoplanet detection, and the limit to sensitivity set by photon noise in the wavefront sensor, were explored by Angel [12]. A more recent paper describes a way to augment a standard AO system by focal plane measurements of the complex amplitude of the residual halo [9]. The temporal evolution of the halo PSF is analyzed in terms of the errors that develop as the turbulent aberration evolves. These are most directly measured by interferometry in the focal plane, rather than the pupil plane as is conventional. The theoretical limit to accuracy is the same in both cases, but the focal plane method has the big advantage of eliminating calibration errors. The measurement is null, and there are no non-common path errors because the science imager and error-sensing detector are one and the same. The complex amplitude of the halo speckles is measured via interferometric elements integrated into the coronagraph, with the blocked core light serving as a reference beam. The optimum correction interval is the time required to record just a few photons per speckle. Even though the resulting amplitude and phase measurements are rather inaccurate, they are direct and good enough to yield a correction that will reduce the halo. In general it will be necessary to correct both corrugations and intensity variations of the wavefront. One way to do this is with a pair of deformable mirrors conjugated to the primary mirror and to the high turbulent layers responsible for scintillation.

The results given by Angel [9] for AO correction of Kolmogorov turbulence lead to an optimum correction interval independent of telescope diameter:

$$\Delta t = 0.053 \frac{5 \alpha_{\text{lim}}}{\lambda_0} \frac{\theta^2}{\lambda_0^2} \frac{\theta^2}{(F^* v^2 \cos^2 \alpha q \Delta \lambda / \lambda)^{1/3}} \quad (5)$$

Here $\lambda$ is the wavelength of observation and $\theta$ the angular radius from the central star. The modeled wavefront aberration is a phase screen characterized by Fried’s length $r_0$ at reference wavelength $\lambda_0$, in motion with speed $v$ at angle $\alpha$ to the planet direction. Typically, $\cos \alpha = 1/\sqrt{2}$.

Broad photometric bands will be necessary to increase signal strength. In order to eliminate the radial blurring of the speckles, the focal plane must be imaged simultaneously in multiple narrow spectral bands. For $R_{\text{psf}} < 10^6$ the chromatic wavefront errors due to dispersion of air can typically be neglected, but to achieve higher values of $R_{\text{psf}}$ it will be necessary to make small differential wavefront corrections in the separate narrow-band wavelength channels.

When corrections are made at the interval $\Delta t$ given above, we obtain

$$R_{\text{psf}} = \frac{I(0)}{I(\theta)} = 4.0 \times 10^6 F^* D^2 q \Delta \lambda / \lambda \Delta t \quad (6)$$

This follows from equation (7) of ref [9], bearing in mind that $I(0)$ is twice the direct stellar PSF because an equal amplitude reference beam has been added. Even
so, the total photon count in \((\lambda/D)^2\) in \(\Delta t\) still averages only 3, and the variance will be dominated by photon rather than speckle noise. If, as we argue in section 5 below, the incipient speckle structure can be decorrelated from one correction cycle to the next, it follows that when measurements are averaged over many cycles, the averaged halo noise will be given approximately by photon noise. The signal/noise ratio for planet detection can then be obtained from equation (4) with \(R_{psf}\) from equation (6).

Note that from equations (4), (5) & (6), \(R_{psf}\) and signal/noise ratio depend only weakly on wavelength, \((R_{psf} \sim \lambda^{0.99}\) and \(S/N \sim \lambda^{1.16}\)). In practice, wavelengths <1 \(\mu\)m will be preferred, because imaging array detectors with high quantum efficiency and single photon counting capability are available.

3. OBSERVING CONDITIONS FROM SPACE AND GROUND

3.1 Diffuse and Zodiacal backgrounds

At a typical 30° ecliptic latitude and 90° longitude, the thermal zodiacal background seen from space (at \(\sim 11\) \(\mu\)m) is ~ 350 mJy/arcsec\(^2\). The solar system zodiacal light is 2.6 mJy/arcsec\(^2\) at 500 nm and 4.8 mJy/arcsec\(^2\) at 800 nm [4]. From the ground the thermal sky emission is lowest in the region 10-12 \(\mu\)m. Under clear conditions at the South Pole, thought to be representative of Dome C, sky brightness is \(\sim 25\) Jy/arcsec\(^2\), while at Mauna Kea under good conditions it is \(\sim 150\) Jy/arcsec\(^2\) [13]. These values for sky background for ground and space are compiled in table 2. We list also as the total background appropriate for resolved exosystems the sum of local and exosystem fluxes from table 1.

Table 3. Telescope thermal emission

<table>
<thead>
<tr>
<th>Temperature and emissivity</th>
<th>210K</th>
<th>273 K</th>
<th>273 K</th>
</tr>
</thead>
<tbody>
<tr>
<td>(I_0) in Jy/arcsec(^2)</td>
<td>42</td>
<td>470</td>
<td>1180</td>
</tr>
<tr>
<td>(e=0.03)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(e=0.08)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(e=0.2)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

When sky and telescope backgrounds are combined, the ratio of ground based to space background is \(1.9\times10^5\) for Antarctica and \(3.8\times10^6\) for OWL. To achieve the same sensitivity as a space interferometer the effective aperture must be increased (from equation (2)) by \((B_{ground}/B_{space})^{1/2}\), i.e. 21 times for Antarctic and 44 times for OWL backgrounds.

3.2 Character of atmospheric aberration

Winter seeing in Antarctica is projected from turbulence measurements to be similar to current good sites [10]. Thus for all the ground telescopes we adopt the same typical strength of atmospheric turbulence, given by \(r_0 = 0.25\) m at reference wavelength \(\lambda_0=800\) nm. For the wind speed we take \(v = 2\) m/sec for Antarctica and 15 m/sec averaged over the turbulent layers for normal sites. Lower turbulence speeds make for improved adaptive correction, because the atmospheric aberration changes relatively slowly.

4. PHOTON NOISE LIMITED SENSITIVITIES OF DIFFERENT TELESCOPES

In this section, the photon-noise-limited signal-to-noise ratio for detection of an earth-like planet is determined for various telescope configurations. Many details that alter the limiting sensitivity lie beyond the scope of this overview. We have, though, made explicit all the assumptions, and have tried to be consistent, aiming in each case to find the highest sensitivity that could be achieved with an efficient optical system. Our estimates are compared with previous estimates when available, or with existing observational data.

Equations (3), (4) & (5) are used as appropriate with the planet and star fluxes given in table 1, and backgrounds from tables 2 and 3. All cases are calculated for the same fractional bandwidth, \(\Delta \lambda/\lambda = 0.2\), system quantum efficiency, \(q = 0.2\) and total integration time \(T = 24\) hr to search all round the star at \(\sim 0.1\) arcsec radius. For those cases where only part of the search space is accessible, the effective integration time \(t_{eff} < T\). The results are compiled for comparison in table 4.
4.1 Nulling interferometer in space

Many concepts have been described and analyzed since Bracewell’s original paper [5]. Here we estimate sensitivity for an interferometer representative of TPF or Darwin, with four elements of diameter d=2 m yielding D_{\text{eff}} = 4 m. With element resolutions \( \lambda/d = 1.1 \) arcsec, the background will be dominated by local zodiacal emission, with exo-zodiacal light at solar system level adding negligible photon noise. The projected sensitivities for \( t_{\text{eff}}=24/4 = 6 \) hr from equation (3) is \( 8.4\sigma \). This estimate is consistent with detailed TPF projections which, scaled for effective aperture derived from total area, yield \( 9.1\sigma \) for \( 24 \) hr exposure [3].

4.2 3.5 and 7 m space coronagraphs

New coronagraph designs show that suppression of diffraction such that \( \text{R}_{\text{psf}} \sim 10^{10} \) is in principle possible at a radius of a few diffraction widths [1] or even closer [2]. Extremely high stability and accurate optics are required, with monolithic primary mirrors preferred. For present rocket fairings, the largest such mirror is limited to 3.5 m. Given an expanded fairing on an EELV, or a cargo modification of the Shuttle STS system, a primary mirror as large as 7 m could be launched. These are the two diameters we will analyze.

The price of apodization to the \( 10^{-10} \) level will be a substantial reduction in the effective diameter \( D_{\text{eff}} \). For some designs \( D_{\text{eff}} \) is no more than half the physical diameter. We will assume a ratio of 2/3, so a 3.5 m telescope \( D_{\text{eff}} = 2.33 \) m, about the same as HST, and for a 7 m mirror \( D_{\text{eff}} = 4.7 \) m. For the best coronagraphs, an earth at 0.1 arcsec should be just resolved at wavelengths of 500 nm for the smaller telescope, \( (\lambda/D_{\text{eff}} = 0.044 \) arcsec) and 800 nm for the larger \( (\lambda/D_{\text{eff}} = 0.035 \) arcsec).

Using the resolved, total background fluxes from table 2, \( R_{\text{bg}}=2.7\times10^6 \) for \( D_{\text{eff}} = 2.33 \) m at \( \lambda=500 \) nm and \( R_{\text{bg}}=3.8\times10^8 \) for \( D_{\text{eff}} = 4.7 \) m at \( \lambda=800 \) nm. Provided the \( R_{\text{psf}} \) is significantly higher, the signal/noise ratio will reach the limits set by zodiacal background of \( 11\sigma \) and \( 34\sigma \) respectively. Here we have taken \( t_{\text{eff}} = 12 \) hrs, consistent with 50% focal plane coverage for the coronagraph.

There is little current understanding now as how such high values for \( R_{\text{psf}} \) could be realized in practice and held stable for long periods. As we show in section 5, residual wavefront Fourier components at \(-1\) m wavelength would need to have amplitude of no more than 1-2 picometers. If in reality the halo flux were dominant, for example \( R_{\text{psf}}=10^3 \), the photon noise alone from the halo would reduce the S/N ratio, now given by equation (5), to \( 2.1\sigma \) and \( 5.5\sigma \) respectively for the cases considered.

As a check on the validity of our sensitivity analysis, we compare our prediction with measurements made with HST ACS WFC, also limited by zodiacal background photon noise. Ford et al. [14] report 5\( \sigma \) detection of a star with \( m_{\text{HDFS}}=27.45 \) (i.e. \( 38 \) nJy, \( 3.8\times\)brighter than our standard earth) in a broad filter at 814 nm in 2600 sec. For \( t=2600 \) sec, \( D_{\text{eff}}=2.4 \) m, \( q=0.4 \) (as measured for the ACS WFC) and using local zodiacal background only, equation (2) yields a S/N ratio of 20. The reason the actual performance is lower is that HST+ACS does not yield the very high Strehl we have assumed, placing 50% of the stellar flux in an area \( 4(\lambda/D)^2 \) compared to effectively 66% in \( (\lambda/D)^2 \) for our ideal Gaussian image. When this reduction is included, we project S/N = 6.6, close to the measured value.

4.3 A 21 m Antarctic ELT

An Antarctic telescope should use 2-mirror optics to minimize emissivity, and be built to minimize construction and on-site engineering costs. One option would be a duplicate of the Giant Magellan Telescope, (GMT), a 21 m telescope planned for Chile [15]. Its primary mirror will be made of 8.4 m segments built like the 6.5 m Magellan mirrors (figure 1). For minimum thermal background, wavefront correction for adaptive optics will be built into the deformable Gregorian secondary, and instruments will mount directly at the Gregorian focus.

The telescope is well suited for terrestrial exoplanet detection by optical imaging. The full aperture of 25 m edge-to-edge yields a resolution of \( \lambda/D=6.7 \) mas at \( \lambda=0.8 \) \( \mu \)m and the individual unobscured 8.4 m elements have resolution 20 mas. By individually apodizing the outer 6 elements at a telescope pupil, a full-telescope beam profile with \( >10^{-6} \) star suppression in a full ring at 0.1 arcsec radius can be realized, free of diffraction spikes so \( t_{\text{eff}}=24 \) hr [16]. We will assume the flux loss is half that of the space systems, i.e. 28%, for \( D_{\text{eff}}=17.5 \) m.

Figure 1. Concept for the 21 m GMT (Giant Magellan Telescope) with seven 8.4 m segments and deformable Gregorian secondary.
The low Antarctic wind speed is favorable. Taking the star flux from section 2 and atmospheric properties for section 3, we obtain an optimum AO correction interval $\Delta t=1.0$ msec, and corresponding halo intensity limited by photon noise in the wavefront sensor at $R_{\text{psf}}=3.8 \times 10^6$ (equations (5) & (6)). The signal/noise ratio for a 24 hr exposure from equation (4) is a useful 5.9. At this level, several dozens of the best candidates could be surveyed during a year of study.

For thermal searches, the diffraction-limited GMT resolution will not adequately resolve a planet at $\lambda=11$ $\mu$m ($\lambda/D=0.09$ arcsec). But high contrast could be achieved by combining light from the six 8.4 m clear apertures as elements of a nulling interferometer. The planet's angular separation for constructive interference is $\lambda/2b$, where $b$ is the element separation. With a maximum $b=17$ m, a planet could be discriminated as close as 0.07 arcsec from the star.

Taking $D_{\text{eff}}=20$ m and summing the background of 25 Jy/arcsec$^2$ for sky and 42 Jy/arcsec$^2$ for the telescope (tables 2 and 3), and with $t_{\text{eff}}=6$ hr, as for the space interferometers, we obtain from equation (3) a sensitivity to an earth at 10 pc of $0.5 \times 10^7$ for $\lambda=1$ $\mu$m and $\theta=0.1$ arcsec. We will assume that suppression to $R_{\text{psf}} > 10^6$ can be accomplished with a coronagraph over half the field (between diffraction spikes), i.e. $t_{\text{eff}}=12$ hr, and that the resulting reduction in resolution and effective area is the same as adopted for the 21 m telescope, i.e. $D_{\text{eff}}=25.5$. The optimum $\Delta t$ is then 0.28 msec and $R_{\text{psf}}=1.9 \times 10^8$, for a S/N ratio from equation (3) of 4.1. This is lower than for an Antarctic GMT, because the gain of increased aperture is more than offset by faster turbulence and less-efficient apodization because of small segments. [Our earlier estimate for 30 m aperture scales to 6 $\times$ for $t_{\text{eff}}=12$ hr [8]. This is consistent when we allow for the apodization loss in $D_{\text{eff}}$, not previously included].

### 4.4 30m ELT at a conventional site

30 m aperture is still too small to resolve planet and star at 11 $\mu$m with a coronagraph. The telescope pupil could be masked off to form nulling interferometer elements, as for the 21 m above. However, the higher thermal background at a normal site results in still lower sensitivity, a S/N ratio of 0.34 in 24 s.

For optical detection at 800 nm, a coronagraph will be necessary. We will assume that the 30 m (and 100 m ELTs below) will have primary mirrors synthesized from 1–2 m size hexagonal segments. Effects such pupil blocking by secondary supports and segment tip/tilt errors cause diffraction spikes and increased halo. For the CELT, Troy and Chanan [17] estimate $R_{\text{psf}} \sim 3 \times 10^5$ for $\lambda=1$ $\mu$m and $\theta=0.1$ arcsec. We will assume that suppression to $R_{\text{psf}} > 10^6$ can be accomplished with a coronagraph over half the field (between diffraction spikes), i.e. $t_{\text{eff}}=12$ hr, and that the resulting reduction in resolution and effective area is the same as adopted for the 21 m telescope, i.e. $D_{\text{eff}}=25.5$. The optimum $\Delta t$ is then 0.28 msec and $R_{\text{psf}}=1.9 \times 10^8$, for a S/N ratio from equation (3) of 4.1. This is lower than for an Antarctic GMT, because the gain of increased aperture is more than offset by faster turbulence and less-efficient apodization because of small segments. [Our earlier estimate for 30 m aperture scales to 6 $\times$ for $t_{\text{eff}}=12$ hr [8]. This is consistent when we allow for the apodization loss in $D_{\text{eff}}$, not previously included].

### 4.5 100 m ELTs

The most ambitious and most powerful current ground-based concept for exoplanet detection is the 100 m filled-aperture OWL [18]. Its resolution at 11 $\mu$m is 0.023 arcsec, high enough for direct diffraction-limited imaging at 0.1 arcsec with an efficient infrared coronagraph, and thus avoiding the interferometric nulling needed for smaller filled apertures. There will thus be no need for multiple exposures to search all around a

<table>
<thead>
<tr>
<th>Table 4. Signal-to-noise ratios projected for detection of an Earth twin at 10 pc.</th>
</tr>
</thead>
<tbody>
<tr>
<td>In every case, the exposure $t=24$ hr, the effective quantum efficiency $q=0.2$ and the bandwidth $\Delta \lambda/\lambda=0.2$.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>telescope</th>
<th>$\lambda$ (µm)</th>
<th>mode</th>
<th>$D_{\text{eff}}$ (m)</th>
<th>$t_{\text{eff}}$ (hr)</th>
<th>$\Delta t$ (sec)</th>
<th>$R_{\text{bgr}}$</th>
<th>$R_{\text{psf}}$</th>
<th>$I_{\text{psf}}$ (nm)</th>
<th>S/N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Space interferometer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4x2 m</td>
<td>11</td>
<td>nulling</td>
<td>4</td>
<td>6</td>
<td>14,400</td>
<td>$&gt; R_{\text{bgr}}$</td>
<td>14.5</td>
<td>8.4</td>
<td></td>
</tr>
<tr>
<td>Space filled aperture</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.5 m</td>
<td>0.5</td>
<td>coronograph</td>
<td>2.3</td>
<td>12</td>
<td>75</td>
<td>$&gt; R_{\text{bgr}}$</td>
<td>0.0015</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>7 m</td>
<td>0.8</td>
<td>coronograph</td>
<td>4.7</td>
<td>12</td>
<td>16</td>
<td>$&gt; R_{\text{bgr}}$</td>
<td>0.0021</td>
<td>34</td>
<td></td>
</tr>
<tr>
<td>7 m</td>
<td>0.8</td>
<td>coronograph</td>
<td>4.7</td>
<td>12</td>
<td>0.42</td>
<td>$3.8 \times 10^7$</td>
<td>1$\times$10$^3$</td>
<td>0.013</td>
<td>5.5</td>
</tr>
<tr>
<td>ELTs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21 m Antarctic</td>
<td>11</td>
<td>nulling</td>
<td>21</td>
<td>6</td>
<td>24</td>
<td>$0.0010$</td>
<td>2.1</td>
<td>2$\times$10$^6$</td>
<td>1200</td>
</tr>
<tr>
<td>0.8</td>
<td>coronograph</td>
<td>17.5</td>
<td>6</td>
<td>24</td>
<td>0.00028</td>
<td>1.9$\times$10$^6$</td>
<td>0.34</td>
<td>4.1</td>
<td></td>
</tr>
<tr>
<td>30 m</td>
<td>11</td>
<td>Nulling</td>
<td>30</td>
<td>6</td>
<td>24</td>
<td>$0.00028$</td>
<td>1.7</td>
<td>6$\times$10$^6$</td>
<td>1340</td>
</tr>
<tr>
<td>0.8</td>
<td>coronograph</td>
<td>25.5</td>
<td>6</td>
<td>12</td>
<td>0.00028</td>
<td>2.1$\times$10$^7$</td>
<td>1340</td>
<td>0.023</td>
<td>4.0</td>
</tr>
<tr>
<td>100 m</td>
<td>11</td>
<td>coronograph</td>
<td>85</td>
<td>24</td>
<td>12</td>
<td>$34$</td>
<td>6$\times$10$^6$</td>
<td>8$\times$10$^7$</td>
<td>17</td>
</tr>
<tr>
<td>0.8</td>
<td>coronograph</td>
<td>85</td>
<td>100</td>
<td>24</td>
<td>0.0010</td>
<td>34</td>
<td>6$\times$10$^7$</td>
<td>8$\times$10$^7$</td>
<td>17</td>
</tr>
<tr>
<td>100 m Antarctic</td>
<td>11</td>
<td>coronograph</td>
<td>85</td>
<td>24</td>
<td>12</td>
<td>$34$</td>
<td>6$\times$10$^6$</td>
<td>8$\times$10$^7$</td>
<td>17</td>
</tr>
<tr>
<td>0.8</td>
<td>coronograph</td>
<td>85</td>
<td>100</td>
<td>24</td>
<td>0.0010</td>
<td>34</td>
<td>6$\times$10$^7$</td>
<td>8$\times$10$^7$</td>
<td>17</td>
</tr>
</tbody>
</table>
star, and \( t_{\text{eff}} = 24 \text{ hr} \). To allow for apodization loss, we take \( D_{\text{eff}}=85 \text{ m} \). The OWL as currently envisaged uses segmented spherical primary and flat secondary mirrors, with a total of 7 warm surfaces reflecting the light to an adaptively corrected focus. The thermal emissivity is thus likely to be relatively high, and we adopt the \( e=20\% \) telescope background given in table 1. With these parameters, the S/N from equation (3) is 4.0\( \sigma \) for 24 hr in the 11 \( \mu \)m band. This is high enough to make practical a thermal survey of many candidates.

Our thermal sensitivity method gives a result consistent with the projection by Gillett and Mountain [19] of 5\( \sigma \) as the sky background limited sensitivity for a 100 \( \mu \)Jy source seen by an 8 m ground based telescope in a 3-hr integration. Scaling from this to a 24 hr integration on a 0.35 \( \mu \)Jy source with 85 m aperture yields 3.4\( \sigma \).

In the optical region at \( \lambda=0.8 \) \( \mu \)m, 100 m aperture yields a diffraction-limited resolution of 1.65 mas, placing the 0.1 arcsec search radius at 60 \( \lambda/D \). Coronagraphic losses will likely be less than for smaller telescopes and longer wavelengths, but we will take a conservative \( D_{\text{eff}} = 85 \text{ m} \) and \( t_{\text{eff}} = 12 \text{ hr} \) to allow for loss due to diffraction spikes. From equations (4), (5) & (6) we obtain \( \Delta t=0.28 \text{ msec} \) and \( R_{\text{psf}}=2.1\times10^7 \) for a signal-to-noise ratio of 46 in 24 hr. This is the highest sensitivity of any system analyzed, 1.5 times higher than the 7 m space coronagraph. Hawarden et al [18] estimate a 1 \( \mu \)m S/N ratio for OWL of 30 in 24 hr, a “conservative minimum”. Our higher somewhat higher value may reflect the reduced halo of the interferometric coronagraph.

We complete our list with the largest and most powerful telescope that could reasonably be built on Earth, a Better OWL (BOWL) in the form of a 100 m, infrared-optimized ELT, located in Antarctica. The 24 hr limits then become 17\( \sigma \) in the 11 \( \mu \)m band and 90\( \sigma \) in the 0.8 \( \mu \)m band. The thermal sensitivity is thus comparable to a full scale TPF or Darwin with four 3 m elements, though from the ground there would not be access to the broad spectral range needed to search for the strong absorption features of CO\(_2\), water and ozone. The optical sensitivity is so high that the telescope could be used to search terrestrial exoplanets for spectral features seen in the “blue dot” spectrum of Earth [20]. The water bands at 1.1 and 1.4 microns would be detectable through the very dry Antarctic atmosphere. The telescope could thus be used to make a serious start on the search for extra-solar life. The BOWL would be designed for low emissivity, with 2-mirror optics, a fast aspheric primary and deformable secondary. As we have suggested for the GMT, it would make sense to build the first BOWL at an accessible site, followed by a second in Antarctica.

5. OPTICAL TOLERANCES AND METROLOGY

The analyses above deal with the random errors caused by photon noise. But the photon noise limit will not be realized unless optical tolerances are understood and met. We list here briefly some of the challenges and approaches.

Optical errors must be controlled to meet two criteria. First, if the projected photon noise levels are to be realized, the instantaneous wavefront or path length must be precise enough to keep the residual stellar flux at or below the background levels derived in section 4. Second, the long-term temporal and spatial statistical properties of the stellar halo must allow for subtraction or averaging to the photon noise limit. We deal with these in turn.

5.1 Instantaneous accuracy

Column 8 in table 4 lists the \( R_{\text{psf}} \) values that must be maintained to obtain the photon noise limiting the S/N values of the last column. With each there is an associated length tolerance \( l_{\text{tol}} \). For interferometers this is the rms path length error that will result in star flux leaking through at a level higher than the assumed background. For direct imaging, it is the maximum amplitude for Fourier components of wavelength \( \sim 1 \) m that will diffract light as bright as the background into the planet search region. \( l_{\text{tol}} \) is given by

\[
l_{\text{tol}} = \frac{\lambda}{2\pi \sqrt{R_{\text{psf}}}} \tag{7}
\]

and is listed in column 9 of table 4. It ranges from \( \sim 1 \) \( \mu \)m for the ground-based interferometers to 1-2 pm for the space coronographs. It will be necessary to meet this tolerance initially and to maintain it during operation.

In the case of the ground-based coronagraphs, the nature of the atmospheric optical disturbance is well understood. The listed tolerances of 20-60 pm for the Fourier components are consistent with the active servo control of the adaptive optics system, with corrections at \( \Delta t\#1 \) msec.

Space coronagraphs will need optics made or adjusted to the extraordinary tolerances of 1-2 pm, and likely a system of continuous active control. The required accuracy represents a large improvement over the HST. For the NICMOS coronagraph, \( R_{\text{psf}} \) (0.5 arcsec, 1.65 \( \mu \)m) \( \sim 4\times10^7 \) [21]. The amplitude of the corresponding residual wavefront Fourier components of wavelength 0.7 m must be \( \sim 8 \) nm, nearly 4 orders of magnitude larger than we require. Measurement to picometer...
accuracy with starlight can be made only on long time scales, because of the limits set by photon noise. The measurement times are shown in column 6 of Table 4, derived by equating $R_{\text{pref}} = R_{\text{br}}$ (set by zodiacal light) and inverting equation (6). They range from less than a second for the halo dominated case ($R_{\text{pref}} = 10^5$) to a minute for the 3.5 m coronagraph at the zodiacal limit.

In the latter case, disturbances with frequencies $> 0.01$ Hz and amplitude $> 1$ pm would require active suppression based on non-stellar, non-common path metrology. This task is beyond currently understood technology. Modeling, experimentation development and tests in space will be needed to understand if such stability is achievable.

5.2 **Long term halo subtraction or averaging**

**Optical coronagraphs**

Long integrations will be needed to average out photon noise from background or halo photons. We can hope to see the planet if the result is a very smooth halo, or if the halo structure remains constant over time, when subtraction methods can be used.

Subtraction of exposures before and after spacecraft roll is the strategy used to search for faint companions with NICMOS. But the changes in the halo, probably due to focus change, set a limit to companion detection at $\sim 3 \times 10^5$ star [21]. Remembering that speckle intensity is determined by amplitude squared, the corresponding slow wavefront changes must be at the 1 nm level. Thus a factor 1000 improvement in long-term wavefront stability would be required for this method to work for terrestrial exoplanet detection with space coronagraphs. Note that this stability improvement is necessary in addition to coronagraph improvement, since a coronagraph does not suppress the effect of wavefront errors.

For ground-based optical coronagraphs, the constantly changing turbulence rules out the subtraction method, and speckle pattern averaging must be used. For this to work, it is crucial that the speckles be decorrelated between successive correction cycles. If so, the central limit theorem applies, and the speckle noise will average out as the root of the number of independent images. Since we record only about 3 photons per correction cycle per beam width, with half of these coming from the uniform reference wavefront, the averaged speckle noise adds little to the photon noise.

In current AO systems it is observed that atmospherically induced halo speckles are not de-correlated on the correction time scale, especially those that are closest to the star [22]. This is because low order errors that determine the inner speckles evolve slowly, changing on a time scale $\lambda/\theta v$. The simple wavefront reconstruction algorithms normally used for adaptive optics result in the same error developing after each of many successive corrections. As noted by Sivaramakrishnan et al [23], simply adding random noise to a wavefront will not result in decorrelated speckles. But as discussed by Stahl & Sandler [24] and Angel [9], decorrelation can be accomplished with a tracking algorithm. Past evolution of the wavefront (as evidenced by the measured history of the speckle complex amplitudes) may be used to anticipate future evolution. Given a good algorithm, the actual future path should differ by a random amount from the prediction, and the speckle patterns will change randomly, as desired.

Long-term changes in the halo pattern will arise because ground apertures will be synthesized from multiple segments. The amplitude and phase terms in the wavefront due to segments misalignments and gaps, will change with thermal, gravitational and wind perturbations. Very small underlying quasi-fixed aberrations will show up after some integration as diffraction peaks in the focal planes, and may be difficult to distinguish from planets. The degree to which photon noise can be reached in long integrations will have to be measured with AO systems at the telescope, where different tracking algorithms and spatial, temporal and chromatic differencing techniques can be developed and tested. One recourse we note is that peaks of stellar origin will show definite phase in the interferometric measurement, while true planetary companions will be incoherent with no definite phase. This discriminator between stellar halo peaks and planetary signals is an advantage for the interferometric coronagraph.

**Thermal detection**

At infrared wavelengths, stellar wavefront and path-length errors can be actively measured and controlled to bring the halo level not only below the diffuse background, but directly to the a level fainter than a terrestrial planet. The situation is much more favorable than for coronography because

1) the wavelength is much longer
2) the contrast is much more favorable
3) a strong flux of shorter wavelength stellar photons is available for interferometric metrology on short time scales.

The characteristic tolerance length to control stellar flux to the level of the planet is given by

$$l_{\text{planet}} = \frac{\lambda}{2\pi} \sqrt{\frac{F_p}{F^*}}$$

At 11 $\mu$m wavelength, $l_{\text{planet}} = 0.8$ nm. (At 0.5 $\mu$m it is 1 pm).
Optical starlight is commonly used as a surrogate to measure and correct the wavefront for infrared imaging. This is the basis of most current adaptive optics [25]. In the same way, interferometric pathlength can also be stabilized, as is being implemented at the LBT nulling interferometer, [7]. The 2.2 µm K band is chosen for ground measurements, to minimize chromatic differences in atmospheric pathlength. To eliminate non common-path errors, the same semi-transparent beam-splitter will be used to interfere the two telescope beams at 11 and 2.2 µm.

In general, the characteristic time for path length measurement to accuracy ∆t_{planet} with starlight of wavelength λ_m and flux F*(λ_m) is given by an extended form of equation (6):

\[ \Delta t = \frac{F^*(\lambda)}{F^*(\lambda_m)} \left( \frac{\lambda_m}{\lambda} \right)^2 \frac{1}{4 \times 10^6 F_p(\lambda) D^2 \Delta \lambda_m / \lambda_m} \]  

(9)

For the space infrared interferometers at 11 µm with D_eff=2 and 6 m, ∆t = 1 and 10 sec, while for the ground telescopes ∆t = 0.1 sec for the GMT and 0.01 sec for the 100 m. In space, laser metrology can be used to maintain 1 nm accuracy on intervals shorter than a few seconds. On the ground, correction of the of the 100 m telescope’s image halo to the planet level on 10 msec intervals should ensure freedom from systematic errors while the long integration proceeds to overcome sky noise.

6. CONCLUSIONS

Based on photon noise considerations, all of the telescopes examined in section 4, both ground- and space-based, have the potential to detect terrestrial exoplanets, through thermal emission or reflected light. On balance, thermal detection with the space interferometers or a 100 m ground telescope seems more secure, because we can see how to control systematic errors. Wavefront and path errors 500 times larger than for optical wavelengths can be tolerated, and these tolerances can be met by active servo systems, with measurements made of shorter wavelength starlight in common-path optics. Space-based interferometers with multiple 3 m elements should have the sensitivity over and wide enough waveband to go beyond simple detection, to make a search for greenhouse gas spectral features.

All the optical coronagraph options studied have potentially good sensitivity, particularly the 100 m ground-based telescope, but the control of systematic errors presents an enormous challenge for both ground and space. It is not clear at this time whether any fundamental advantage is held in space or on the ground. Similar strategies for sensing and controlling long term errors may lead to similar sensitivities.

Much experimentation with more advanced active correction techniques and coronographs is needed. Many of the techniques needed for terrestrial planet detection from ground and space can be developed and tested at ground telescopes and interferometers, with the advantage of ready access to make experiments and modify equipment. In the process, it is likely that exo-zodiacal emissions and brighter giant exoplanets will be detected with existing facilities. The key ground coronagraph issue, namely the extent to which the residual atmospheric halo can be smoothed out by averaging, should be resolved by these experiments. At this point we should remain open all the possibilities.

ACKNOWLEDGEMENTS

I am grateful to Johanan Codona, Phil Hinz and Nick Woolf for helpful comments on draft versions of the paper. This work is supported by the NSF under grant AST-0138347 and by AFOSR under contract F49620-01-1-0383.

REFERENCES


