doi: 10.1038/nature05636

Supplemental Information: Methodology
A Spectrum of an Extrasolar Planet

L. Jeremy Richardson[1], Drake Deming[2], Karen Horning[3], Sara Seager[4,5] & Joseph Harrington[6]

[1]*Exoplanets and Stellar Astrophysics Laboratory, NASA Goddard Space Flight Center, Mail Code 667, Greenbelt, MD 20771*

[2]*Planetary Systems Laboratory, NASA Goddard Space Flight Center, Mail Code 693, Greenbelt, MD 20771*

[3]*Department of Physics and Space Sciences, Florida Institute of Technology, 150 W. University Blvd., Melbourne, FL 32901*

[4]*Department of Terrestrial Magnetism, Carnegie Institution of Washington, 5241 Broad Branch Rd., NW, Washington, DC 20015*

[5]*Department of Earth, Atmospheric, and Planetary Sciences, Massachusetts Institute of Technology, 77 Massachusetts Ave., Cambridge, MA 02139*

[6]*Department of Physics, University of Central Florida, Orlando, FL 32816*

Here we describe the details of our methods and analysis to derive the planetary spectrum from our IRS observations of HD 209458 b. The main result is summarized in Supplemental Figure 1, which shows the average spectrum of the planet for the two observed eclipse events (compare to Figure 2b in main text). In order to allow other researchers to compare models to our observed spectrum, we provide the final planetary spectra in contrast units (planetary to stellar flux) for both events as a text file (also available for download as Supplementary

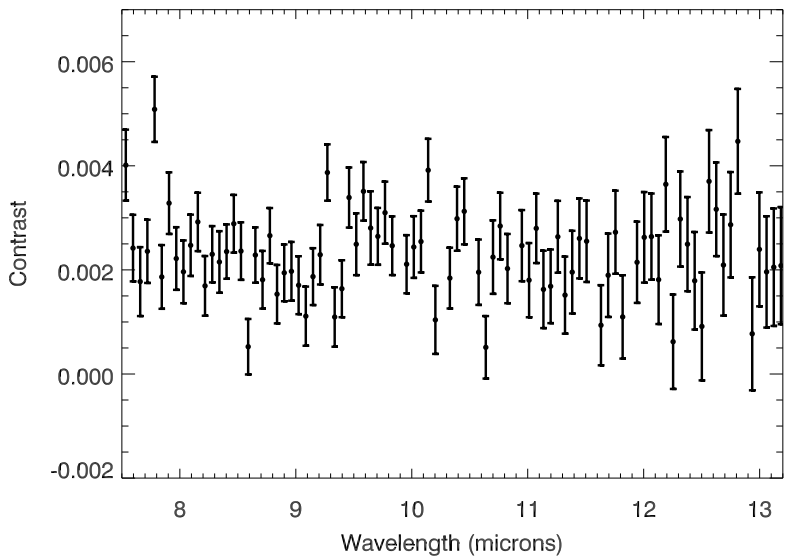

Figure 1: Average planetary spectrum from the two eclipse events.

Information).

1 Spectral Extraction

We created a custom procedure for extracting the combined light spectra (star+planet) from the IRS images. We read in the Basic Calibrated Data (BCD) images, created from raw data by the reduction pipeline at the Spitzer Science Center (version S13.2.0). Each BCD is a 2-D image representing a single 60.95-sec integration on the source. The dimensions of the image represent wavelength and spatial distance parallel to the spectrograph slit. The standard procedure for IRS Staring Mode is to observe half the images with the star at position A on the slit, nod the telescope, and then record the other half of the observations at position B. We found this procedure to be optimal for the purpose of extracting the planet spectrum.

The 560 BCD images are separated into four groups (of 140 images each) based on eclipse event and nod position. For each group, we perform the following steps:

1. *Identify and correct bad pixels.* We employed the `IRSCLEAN_MASK` (created by J. Ingalls and the IRS Instrument Support Team) routine downloaded from the Spitzer website. It accounts for known bad pixels on the array and allows the user to identify and correct other "rogue pixels" by inspecting the 2-D data.

2. *Median filter the images.* The median image is derived by calculating the median value of each pixel from the stack of images. For the A position, we exclude the first 50 images from the median calculation to avoid bias due to the systematic ramp that occurs at the beginning of each observing sequence (see Figure 1 in the main text). The median image is subtracted from each individual image, and we then apply the `SIGMA_FILTER` routine (see the IDL Astronomy Library) to the difference image; we reject and correct a given pixel (centered on a box of width 5 pixels) if it exceeds 10σ of the pixels in the box. This serves to correct any rogue pixels not identified in Step 1.

3. *Create the background-subtracted images* by subtracting from each individual image the median image from the *opposite* nod position. For wavelengths near the center of our band-pass, the background is about 2 percent of the stellar intensity. We checked the hypothetical possibility that background fluctuations, not removed by the nod, might affect our results. This was checked two ways; first, we performed the entire analysis without subtracting the background, finding essentially the same results, but with higher noise. Second, we pro-

duced a set of background spectra, by extracting the background in each image as if the star were present, but using the opposite nod position from the star. Analyzing these spectra in lieu of the real data, we see no effects in the background that would contaminate our planet spectrum.

4. *Extract the spectrum from each image.* We find the maximum value (peak) in the spectrum at each row (which represents wavelength), and we add the 4 pixels on either side of the peak, orthogonal to the dispersion direction (for a total of 9 pixels) to obtain the flux at this wavelength. We ignore the curvature of the spectrum on the array. We assign the wavelength of each point by averaging the corresponding pixels in the calibration file `b0_wavsamp_wave.fits` from each event. Two members of our team extracted the spectra using separate analysis routines based on this method, and we ran both sets of spectra through our entire analysis, obtaining virtually identical results. We also used the SPICE software from the SSC to extract spectra, and we verified that these spectra are also consistent with the conclusions of this paper, again by running them through the entire analysis. However, the SPICE spectra do not produce as high a signal-to-noise ratio in the planet spectrum for this very specialized problem. We also varied the width of the window to 6 pixels on either side of the peak (13 pixels total), and we ran these spectra through our entire analysis. In deciding which version of the spectra to use for our final analysis, we computed the chi-squared statistic for the difference between the planet spectra derived for the two eclipse events, and we use the spectra (9-pixel width, custom extraction) that produce the minimum chi-squared in the difference.

5. *Apply a multiplicative factor*. This factor varies with wavelength, and it essentially corrects for the discontinuity caused by the telescope nod and imperfect flat-fielding of the detector array. In order to avoid bias in this correction, the factor is calculated using only the data with the planet in eclipse (i.e., behind the star), and separate factors are calculated and applied for each wavelength. At this point, the A and B spectra are recombined, and from here onward, we consider only two groups of spectra, corresponding to the two eclipse events. Note, however, that our analysis is done independently at each wavelength using the time series of intensity. Thus, combining the spectra specifically means that we are adjusting the time series at each wavelength to fix the discontinuity caused by the nod.

6. *Normalize the spectra* by dividing each spectrum by the average spectrum, shown in Supplemental Figure 2a, in that eclipse event (to convert the spectra to "contrast" units). This step essentially just normalizes the intensity in each time series to unity. We then subtract the average time series, shown in Supplemental Figure 2b, from the individual time series at every wavelength point. This serves to remove the first two systematic effects in the data (see Section 4). Supplemental Figure 3 shows the "stacks" of spectra (wavelength on the x-axis and phase on the y-axis) before and after the normalization.

7. *Calculate the errors* by computing the standard deviation in each time series of normalized contrast values at each wavelength. The standard deviation is calculated by shifting the time series by one step, subtracting the shifted series, and computing the standard deviation of the difference, with the end points omitted. We divide this precision by the square-root of 2, and we assign it to all contrast values in the time series at that wavelength. The error ranges from

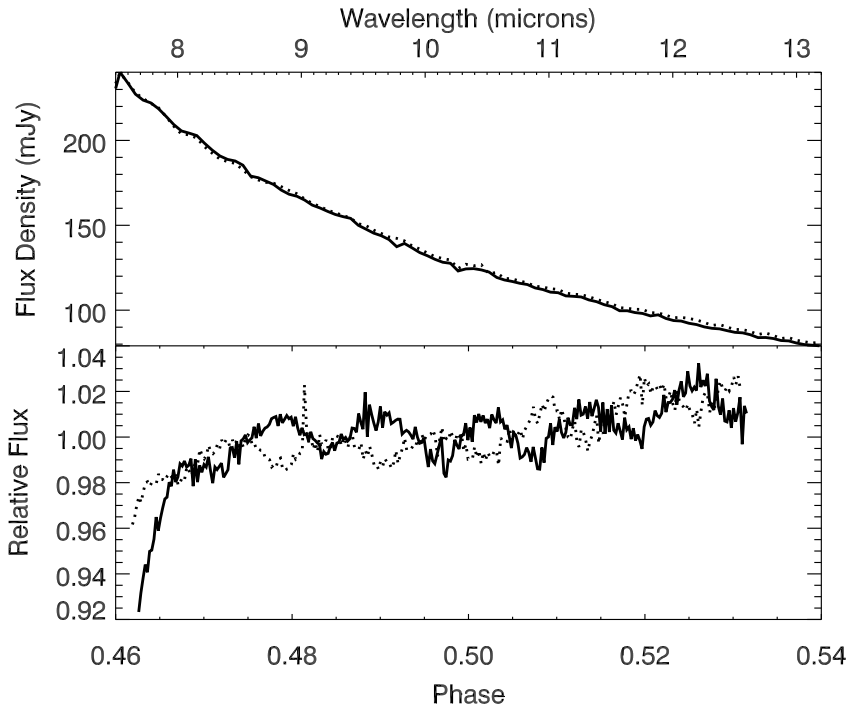

Figure 2: Normalizations used in the analysis. Upper panel shows the average spectrum for each eclipse event (solid line for first event, dotted line for second event), and the flux density is approximate because we have not accounted for slit losses. Lower panel similarly shows the average time series for both events.

0.0055–0.029 (noise is higher at longer wavelengths). These values are within 50 to 70% of the fractional statistical fluctuations in the number of photons (electrons) detected. These per-point precisions are propagated through the linear regressions (see below), resulting in errors assigned to the planet spectrum at each wavelength.

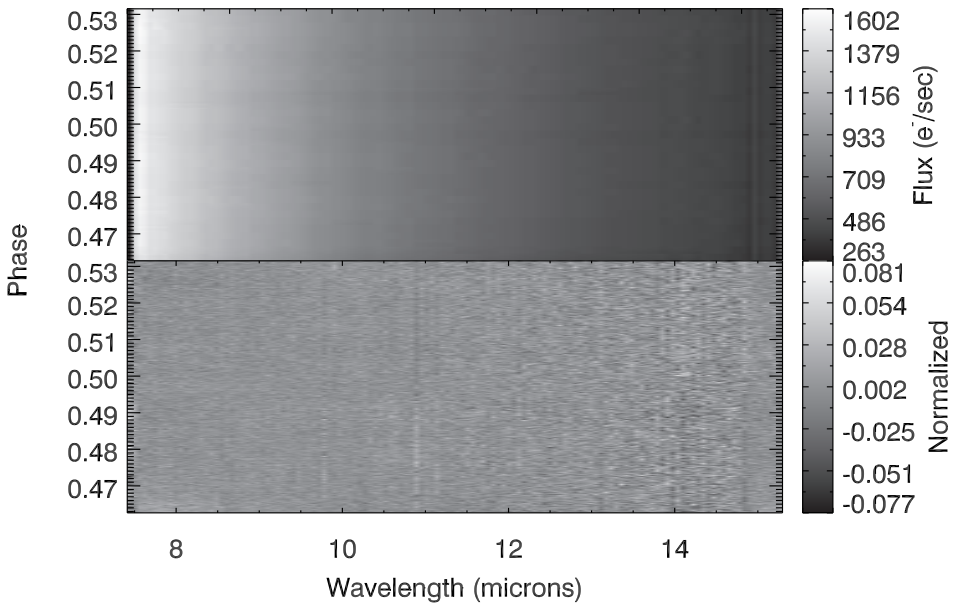

Figure 3: Spectra obtained for first eclipse event. Upper panel shows the uncorrected spectra (after executing Step 5). Lower panel shows the normalized residual spectra (after executing Step 6).

2 Fitting the Eclipse Curve

We fit a model eclipse curve to the time series at each wavelength. The eclipse amplitude (duration and central phase held fixed) is estimated using multiple linear regression, simultaneously with a residual linear ramp and a residual periodic oscillation. The periodic oscillation used as the independent variable in the regression is obtained by subtracting from the average time series a fourth-order polynomial fit to the average time series; the result is the oscillation alone with the correct period. The amplitude of the eclipse fit at each wavelength gives the planetary spectrum, and is nearly equivalent to subtracting the in-eclipse combined light spectrum from the out-of-

eclipse combined light spectrum, which we verified by actually subtracting those spectra. By allowing for a residual linear ramp in the linear regressions, we are drawing on our experience[1] that the differences in the ramp from pixel to pixel are linear to a good approximation, except for the first $\sim$30 integrations, where higher order effects can sometimes remain. We experimented with omitting the first 30 integrations from the regressions, but this had little effect on our results. By fitting for a residual 1.02 hr oscillation in the regressions, we allow for the possibility that the oscillation may not be perfectly subtracted by our procedure. However, we found that our eclipse amplitudes are remarkably insensitive to fitting (or not) for a residual oscillation. We attribute this to the fact that the time scale of the oscillation is several times shorter than the eclipse duration, and this bandwidth difference mitigates significant interaction with the fitted eclipse amplitude. Errors on the fitted eclipse amplitudes are returned by the regression routine, based on the per-point precisions described below.

Reject wavelengths where the reduced chi-squared of the eclipse fit is greater than a cutoff value. We found that the linear regressions usually fit the contrast time series data to within the noise. Accordingly, the distribution of reduced chi-squared values is centered closely on unity, and we reject any wavelength where the reduced chi-squared exceeds 1.2. This stringent criterion eliminates false-positive detections of residual eclipses, so they are not manifest as spurious features in the planet spectrum. Of 94 wavelength bins shortward of 13.2 μm, we reject 9 bins from the first eclipse event and 8 from the second event.

3 Calibration and Consistency Checks

Calibrate the results to contrast units. The resulting spectra (one for each eclipse event) are averaged together (Supplemental Figure 1). We checked to ensure that we obtain the same result regardless of whether the binning or averaging was performed first. Prior to averaging the two eclipses, a calibration is applied to the planet spectrum from each event to adjust for the wavelength variation of the slit losses from slow image drift, and to place the spectra on an absolute contrast scale. The measured preliminary depth of the eclipse (0.25%) in the IRAC 8 μm band (kindly communicated by D. Charbonneau in advance of publication) is used as a calibration for the effect of slit losses. We weight the wavelengths in our spectra so as to simulate the IRAC 8-micron bandpass as closely as possible within the limit imposed by the incomplete wavelength overlap. We removed the telescope oscillation using a Fourier notch filter. This produces a synthetic IRAC eclipse from our IRS data (as in Figure 1 of the main text), and we average it over the two eclipses. This average eclipse is too deep, due to a component of slow telescope drift perpendicular to the slit. We scale the extra depth with wavelength, and subtract these corrections from the contrast values in our planet spectrum. The primary effect of this procedure is a zero-point correction to the contrast; the variations with wavelength are smaller (< 0.1%), and vary gradually over our bandpass. The scaling with wavelength is based on the measured wavelength dependence of the intensity fluctuations created by the 1.02-hour telescope oscillation, since the wavelength scaling is independent of temporal frequency.

Apply a high-pass filter analysis, as an alternate procedure for removing the telescope and

instrument systematics to check the reality of the sharp spectral feature at 7.78 μm. This procedure begins by fitting each spectrum with a high-order polynominal, and removing this fit to yield residual intensities. Both fourth and sixth order polynomials were used, and produced similar results. The principle of this analysis is that most telescope and instrument systematics vary slowly with wavelength (see Section 4) and are removed by the polynomial fit. However, sharp spectral features will remain, and their differential eclipses will be detectable by linear regression, as described above. As for the main analysis, we rejected poor fits to avoid false-positive detection of residual eclipses; for this analysis we tightened our limit in reduced chi-squared to 1.15. Of the 94 wavelength bins shortward of 13.2 μm, we reject 7 bins from the first eclipse event and 13 bins from the second event. We find eclipses in the 7.78 μm feature with correct ingress and egress times that repeated for both eclipse events.

Determine the width of the 7.78 μm feature. Another key test for the reality of the 7.78 μm feature is to determine whether it is consistent with the 2-pixel spectral resolution of IRS. A feature occupying a single pixel, for example, is not likely to be real, since even an instrinsically sharp feature will be broadened to the 2-pixel resolution of the instrument. Therefore, we measured the width of this emission by fitting a Gaussian profile to the data. The strong-filtering analysis has the property that it suppresses broad spectral structure such as the 9.65 μm feature, and it also attenuates the wings of the 7.78 μm feature. We fit to the width of the 7.78 μm feature in the data from our primary analysis (Figure 2 of the main text), and this fit is shown in Supplemental Figure 4. The result indicates a width of 2.2 pixels, in agreement with the spectral resolution of the observations.

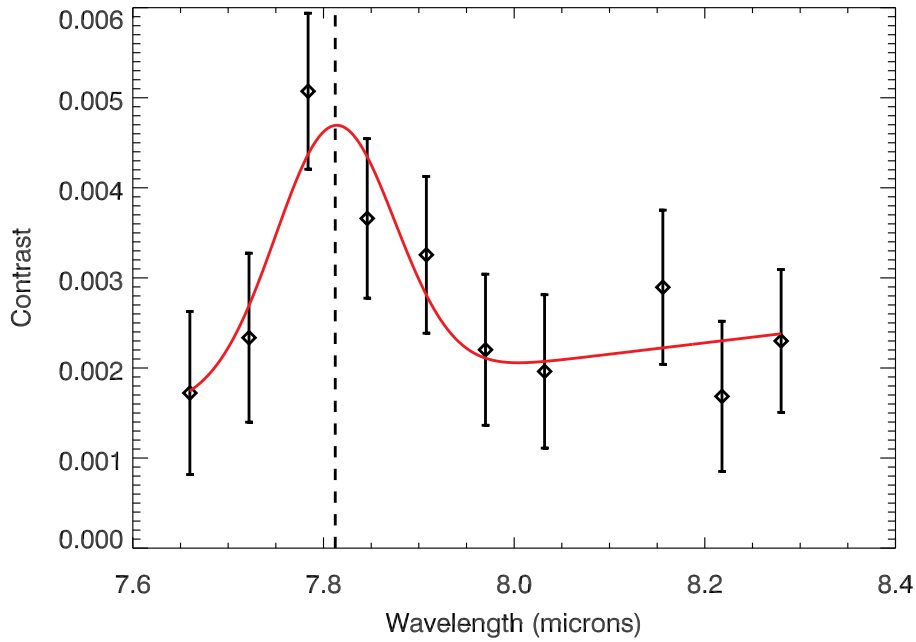

Figure 4: Fit of Gaussian plus linear baseline to the unidentified feature near 7.8 μm. Data points represent the average of the two eclipse events, with one exception: the point at 8.25 μm is the value at that wavelength from the first eclipse only, since the point from the second event is clearly discrepant (see Figure 2a of the main text).

4 Systematic Effects

We now discuss the three systematic effects in the data and their removal.

1. The first systematic effect is a gradual increase in intensity over the 6 hours of each eclipse, that we denote as the "ramp". The ramp[1] is shown in Supplemental Figure 2b (but is seen more clearly in Figure 1 of the main text). The cause of the ramp is not completely under-

stood, but is likely related to charge trapping in the Si:As detector material (H. Knutson and D. Charbonneau, private communication). The other systematic effects are due to telescope pointing errors. These pointing errors are comprised of a 1.02-hour oscillation, and a drift on longer time scales, both described below. To first order, both the ramp and the 1.02-hour oscillation are removed by subtracting the average time series as described above. (This subtraction also removes the average eclipse, but still allows us to derive the planet spectrum by finding the differential eclipse depths as a function of wavelength.)

2. The 1.02 hr intensity oscillation is due to a periodic telescope pointing error, shown in Supplemental Figure 2b. This effect is well known to the SSC, and we used our data to verify that the telescope pointing oscillation is indeed the cause of the intensity oscillation seen in our spectra. We measured the spatial position of the star along the slit, by fitting to the centroid of the spatial intensity distribution at each wavelength. We find that these positions show the same 1.02 hr oscillation, and are strongly correlated with the intensity oscillation seen in our data. We also verified that the phase of the intensity oscillation is independent of wavelength, and its amplitude is weakly dependent on wavelength. The wavelength dependence is measured and found to be consistent with expectations based on diffraction of the PSF. By subtracting the average time series from the individual time series at each wavelength, we very effectively remove the oscillation. However, our analysis also includes residual oscillation that remain at some wavelengths (as mentioned in Section 2).

3. The third systematic effect is a slow telescope drift that causes the depth of the eclipse (Figure 1 of the main text) to appear deeper than it actually is. The existence of this slow drift

was indicated by the position of the star parallel to the slit, that was derived as part of our analysis. We have no direct information on stellar motion perpendicular to the slit for our observations, but D. Charbonneau and H. Knutson showed us position information from their 30-hour sequence of IRAC photometry on HD 189733, and this revealed significant drift in both orthogonal coordinates, on long time scales. Hence we developed a calibration procedure that removes the effect of slow drift from our observations in a very general manner, as described in Section 3.

1. Deming, D., Harrington, J., Seager, S. & Richardson, L. J. Strong Infrared Emission from the Extrasolar Planet HD 189733b. *ApJ* **644**, 560–564 (2006).