SPOTROD: a semi-analytic model for transits of spotted stars

Bence Béky,† David M. Kipping‡ and Matthew J. Holman

Harvard–Smithsonian Center for Astrophysics, 60 Garden St, Cambridge, MA 02138, USA

Accepted 2014 May 25. Received 2014 May 18; in original form 2014 April 1

ABSTRACT

The Hubble Space Telescope and the Kepler space mission observed a large number of planetary transits showing anomalies due to starspot eclipses, with more such observations expected in the near future by the K2 mission and the Transiting Exoplanet Survey Satellite. To facilitate analysis of this phenomenon, we present SPOTROD, a model for planetary transits of stars with an arbitrary limb darkening law and a number of homogeneous, circular spots on their surface. A free, open source implementation written in C, ready to use in PYTHON, is available for download. We analyse Kepler observations of the planetary host star HAT-P-11, and study the size and contrast of more than 200 starspots. We find that the flux ratio of spots ranges at least from 0.6 to 0.9, corresponding to an effective temperature approximately 100–450 K lower than the stellar surface, although it is possible that some spots are darker than 0.5. The largest detected spots have a radius less than approximately 0.2 stellar radii.

Key words: techniques: photometric – stars: individual: HAT-P-11 – starspots.

1 INTRODUCTION

In a transiting planetary system, spots on the face of the host star can result in deviations in the transit light curve from the well-known model described by Mandel & Agol (2002). An unocculted spot – since it is darker than the stellar surface – causes a blend in the opposite sense as a background star, leading to a deeper transit (see, for example, Czesla et al. 2009). On the other hand, a spot causes an anomalous rebrightening when it is eclipsed by the planet, because the planet blocks less flux than it would if the spot was not behind it.

Such spot-induced transit light-curve anomalies were first observed by Silva (2003) in Hubble Space Telescope observations of HD 209458. Other systems exhibiting similar features include HD 189733 (Pont et al. 2007) and TrES-1 (Rabus et al. 2009). Czesla et al. (2009) found a correlation between stellar brightness and transit depth in the system CoRoT-2, which they attribute to varying levels of stellar activity, and show how this effect, when unaccounted for, causes a bias in the planet size estimate.

In the era of the Kepler satellite, a large number of planets transiting active stars have been discovered and observed with high temporal and photometric resolution, providing further examples of transit anomalies. Two such systems are Kepler-17 (Désert et al. 2011) and HAT-P-11 (Bakos et al. 2010). Spots revealed by transit anomalies can be used, for example, to constrain the projected obliquity and the stellar inclination (Deming et al. 2011; Sanchis-Ojeda & Winn 2011). Measuring the spot contrast allows one to constrain the temperature of the spots (Silva 2003; Rabus et al. 2009). However, there is a degeneracy between the spot size and contrast (Pont et al. 2007; Tregloan-Reed, Southworth & Tappert 2013), which makes high-quality data necessary to infer temperatures.

The large number of photometric observations of transit anomalies motivates the development of astrophysical models. Examples include the model by Silva (2003), the one by Wolter et al. (2009), SOAP-T by Oshagh et al. (2013), and PRISh by Tregloan-Reed et al. (2013). These models all assume homogeneous, circular spots, with four input parameters for each spot (two for position, one for size, one for darkness). The first two models simplify the geometry by assuming that the spots are circular in projection, while the other two properly account for the elliptical projected shape given spots that are circular on the stellar surface. These models all define a large-resolution two-dimensional grid either on the stellar surface or in the projection plane, and numerically integrate over two coordinates to calculate the transit light curve.

Integration in two dimensions can be computationally expensive. Kipping (2012) introduced MACULA, an analytic model for a related but different phenomenon: spots on the rotating stellar surface modulating out-of-transit light curves of spotted stars. Its analytic nature makes MACULA faster than numerical models for the same phenomenon, like SOAP (Boisse, Bonfils & Santos 2012).

In this paper, we present SPOTROD, a counterpart of MACULA for transit light curves of spotted stars. We describe the problem as a two-dimensional integral in polar coordinates in the projection plane. Using assumptions similar to those of previous models, we derive an analytic formulation for the integral with respect to the polar angle, so that numerical integration needs to be performed only with respect to the radial coordinate. This semi-analytic nature provides improved speed over previous models requiring

*E-mail: bbeky@cfa.harvard.edu
†Carl Sagan Fellow.
‡Kepler† for 2003 MACULA by Tregloan-Reed et al.
two-dimensional numerical integration. In particular, if the resolution of the integration grid is \( n \) in each dimension, then a double numerical integral takes \( O(n^2) \) time to evaluate, whereas SPOTROD runs in \( O(n) \) time. Typical values are \( n \approx 300 \) for SOAP, \( n \approx 750 \) for the model of Silva (2003), and the grid spacing being one hundredth of the planet diameter for PRISM, resulting in \( n \approx 1000 \) for a typical hot Jupiter or \( n \approx 2000 \) for HAT-P-11b. In this work we use \( n = 1000 \).

We describe the semi-analytic model in Section 2: we state the simplifying assumptions, describe the two-dimensional integral in polar coordinates, and introduce the subroutines of SPOTROD, the free and open source implementation available for the astronomical community. In Section 3, we apply SPOTROD to Kepler observations of HAT-P-11, investigate model artefacts like observational biases and correlations of fit parameters for individual spots, validate SPOTROD on synthetic data generated by PRISM, study the distribution of spot size and contrast on HAT-P-11 that we believe to be physical, and look at the model residuals for validation. Section 4 concludes our findings. Technical details of the model pertaining to calculating angles and handling spots that are partially behind the limb are given in Appendix A.

# 2 SPOT ANOMALY MODEL

## 2.1 Assumptions

Our model has two major assumptions. The first one is that the boundary of each spot is a circle on the surface of the spherical star. We define the radius \( a \) of the spot to be the radius of this circle in three-dimensional space, in units of stellar radius. We assume that \( 0 < a < 1 \). Note that the analytic rotational modulation model MACULA (Kipping 2012) takes as input parameter the half angle \( \alpha \) of the cone with this circle as its directrix and the centre of the star as its apex, which is related to the radius by \( a = \sin \alpha \).

We define the centre of the spot as the intersection point of the surface of the sphere and the axis of this cone. Note that the centre of the spot does not lie in the plane of the boundary. The advantage of this definition is that the centre is on the stellar surface, allowing for easier conversion between input parameters of SPOTROD and MACULA, and easier treatment of stellar rotation. A further advantage of characterizing the spot location with the projection of a point on the stellar surface over the projection of the geometrical centre of the spot boundary in the interior of the star is that its domain does not depend on the spot radius, which also makes it easier to define an isotropic prior for the location of the spot.

The second assumption is that each spot is homogeneous and observes the same limb darkening law as the star. This means that as viewed by the observer, the ratio of the flux from a spot to the flux from the unsotted stellar surface at the same projected distance \( r \) from the centre of the star does not depend on the distance \( r \). We denote this dimensionless flux ratio by \( f \) in accordance with Kipping (2012). Note that the flux ratio is sometimes called contrast, for example, by Tregloan-Reed et al. (2013).

Assuming that a constant flux ratio across the stellar disc is consistent with the findings of Walton, Preminger & Chapman (2003): they study a sample of 18 000 spots on the Sun, and observe no dependence of spot contrast on where the spot is seen. Note that what they call contrast can be expressed as \( f - 1 \) using our notation. As for the homogeneity of spots, we shall see in Section 2.4 how to compose more complicated structures, like a spot with umbra and penumbra, using two homogeneous spots.

## 2.2 Integration

Let \( I(r) \) denote the stellar intensity according to the limb darkening law up to an arbitrary scaling factor, where \( 0 \leq r \leq 1 \) is the projected distance from the centre of the star in units of stellar radius. Let \( C_r \) denote the circle of radius \( r \) in the projection plane concentric with the stellar disc. Then the total out-of-transit flux of the unsotted stellar surface can be calculated as a two-dimensional integral over the polar coordinates \((\theta, r)\), with the inner integral along \( C_r \), and the outer integral with respect to the radial coordinate:

\[
F_0 = \int_0^1 \int_0^{2\pi} I(r) \, d\theta \, r \, dr.
\]

Here \( d\theta \, dr \) is the area of an infinitesimal element in the projection plane. The integrand does not depend on \( \theta \); therefore, the inner integral can be evaluated as the product of the integrand \( I(r) \) and the length \( 2\pi \) of the integration interval.

Now consider the case of a single spot visible on the star. Let \( \gamma(r) \) be half the central angle of the arc of \( C_r \) that overlaps with a spot, as shown in Fig. 1. Then the total flux is

\[
F_1 = \int_0^1 \left( \int_0^{2\gamma(r)} f I(r) \, d\theta + \int_{2\gamma(r)}^{2\pi} I(r) \, d\theta \right) \, r \, dr
\]

\[
= \int_0^1 \left( 2\gamma(r) f I(r) + (2\pi - 2\gamma(r)) I(r) \right) \, r \, dr
\]

\[
= \int_0^1 2\pi (f - 1) \gamma(r) I(r) r \, dr.
\]
multiple spots in a single summation:

\[ F_i = \int_0^1 2 \left( \pi + \sum_{i=1}^n (f_i - 1)\gamma_i(r) \right) I(r)r dr. \]  

(3)

During transit, let \( \delta(r) \) the half central angle of the arc of \( C \), that is obscured by the planet. Let \( \gamma_i(r) \) be the half central angle of the arc on the same circle that overlaps with spot \( i \), but is not obscured by the planet. See spot 2 in Fig. 1 for an example. Then the total flux can be calculated by substituting \( \gamma_i(r) \) for \( \gamma(r) \) in equation (3), and subtracting \( 2\delta(r) \) from the arc length of the unspotted stellar surface:

\[ F_{\text{transit}} = \int_0^1 \left( \sum_{i=1}^n 2\gamma_i(r) f_i I(r) \right) + \left( 2\pi - \sum_{i=1}^n 2\gamma_i(r) - 2\delta(r) \right) I(r) r dr \]

\[ = \int_0^1 2 \left( \pi - \delta(r) + \sum_{i=1}^n (f_i - 1)\gamma_i(r) \right) I(r)r dr. \]  

(4)

The contribution of the planet in this formula is formally equivalent to a spot with flux ratio \( f = 0 \).

The final product of our proposed model is the dimensionless normalized transit light curve

\[ F_{\text{normalized}} = \frac{F_{\text{transit}}}{F_0}. \]  

(5)

At this step, the arbitrary scaling factor in \( I(r) \) cancels out.

Note that we define \( \gamma_i(r) \) and \( \gamma(r) \) to be zero in case the corresponding arcs do not exist, that is, \( C \), does not intersect the spot in projection. Similarly, \( \delta(r) \) is understood to be zero if the planet does not eclipse \( C \). See Appendix A2 on calculating \( \gamma_i(r), \gamma(r), \) and \( \delta \).

In case there are no spots on the stellar surface, even though our model still yields the correct light curve asymptotically for large grid resolution \( n \), we suggest using the fully analytic algorithm of Mandel & Agol (2002) if speed is a consideration.

2.3 Implementation

The model described in this paper is implemented as a software package called SPOTROD. It provides four functions.

1. \texttt{elements} takes the planetary period, semi-major axis, \( k = \cos \sigma, h = \sin \sigma \), and an array of observation times of a transiting planet with respect to the time of mid-transit as input parameters, and calculates the arrays of planar orbital elements \( \xi \) and \( \eta \) using the formalism of Pál (2009).

2. \texttt{circleangle} takes the planetary radius \( R_p \), the distance \( z \) of the centres of the planet and the stellar disc in projection plane, and an array of radii \( r \) as input parameters, and calculates the array \( \delta(r) \).

3. \texttt{ellipseangle} takes the projected spot semi-major axis \( a \), the distance \( z \) of the centres of the projected spot boundary and the stellar disc in projection plane, and an array of \( r \) as input parameters, and calculates the array \( \gamma(r) \). This function is executed internally by \texttt{integratetransit}, so the user does not have to call it directly.

4. \texttt{integratetransit} takes \( R_p \), arrays of the projected coordinates of the planet and the spots, spot radii and flux ratios, an array of the radii \( r \) and weights for numerical integration (the latter are calculated from the integration quadrature and the limb darkening law), and precalculated values of \( \delta(r) \) for each value of \( r \) and each observation time as input parameters. It calculates the normalized light curve \( F_{\text{normalized}} \) using analytic integration with respect to the polar angle and numerical integration with respect to \( r \).

The software can employ any integration quadrature, that is, numerical method that works by evaluating the integrand at given values of \( r \) and summing up using given weights. It evaluates the inner integral analytically in the form of the sum given in equations (3) and (4) on each annulus defined by the input array of \( r \) values. Then it performs numerical integration with respect to \( r \) by adding up the products of these values and the corresponding weights.

Among the simplest integration quadratures are the trapezoidal rule and the mid-point rule. We recommend an integration mesh of \( n \approx 1000 \) values uniformly spaced between 0 and 1. More complicated rules can also be prescribed. For example, since the integrand of the outer integral grows rapidly with \( r \) in equations (1)–(4), one might wish to use a non-uniform mesh that is coarser for small \( r \) and finer for large \( r \).

SPOTROD can also handle arbitrary limb darkening laws, even ones without an analytic formula: it only relies on limb darkening values evaluated at the values of \( r \) used in the integration rule. The limb darkening law and the integration quadrature weights are then multiplied together before they are passed to \texttt{integratetransit}, since it is only their product that is ever used.

Repeated evaluations of \texttt{integratetransit} at the same observation times with fixed planetary orbital parameters are required, for example, for fitting or a Monte Carlo Markov chain (MCMC) exploring spot parameters. The code has been optimized for such use: one needs to calculate the arrays \( \xi \) and \( \eta \), then the projected planetary coordinates at each observation, and finally an array of \( \delta \) only once at the beginning. These values do not depend on spot parameters, and recalculating the same values of \( \delta \) in each iteration would be very costly, since it has to be calculated for each observation time and each \( r \); hundreds of thousands of times in a typical application. Instead, we evaluate \( \delta \) hundreds of thousands of times only once, before we start the fit or MCMC, and then use these precalculated values.

We can also avoid evaluating the function \texttt{ellipseangle} hundreds of thousands of times in each iteration if we neglect the effect of stellar rotation during a single transit. In this case, for a given set of spots, \( \gamma_i(r) \) does not depend on time; therefore, we only need to calculate it \( n \) times: once for each value of \( r \).

It is, of course, also possible to model a transit in the extreme case of a star that rotates so rapidly that spots move substantially during the duration of the planetary transit. In this case, one needs to recalculate the spot positions and call the function \texttt{integratetransit} with a time array of length one for each observation. This method, however, is slower than if we assumed that spots were stationary within the duration of a single transit.

SPOTROD is free and open source software, released under the GNU General Public License. It is implemented in C in the interest of speed, and provides bindings for use in Python. Bindings for different programming languages should be reasonably easy to add.

SPOTROD is publicly available for download at https://github.com/bencec Bey/spotrod, including the C code, Python bindings, two example programs in Python, compilation instructions, and a copy of the license.

2.4 Umbra, penumbra, and faculae

Note that in equations (3) and (4), the contributions of spots add up, regardless of whether they overlap or not. As Kipping (2012) points out in his section 2.4, this feature can be used to build a
composite spot with a central umbra of radius \( a_u \) and flux ratio \( f_u \) and a surrounding penumbra of radius \( a_p \) and flux ratio \( f_p \) by feeding two concentric spots with radius–flux ratio pairs \((a_u, f_u)\) and \((a_p, 1 - f_u + f_p)\) into the model.

Spots should have flux ratio between 0 and 1, \( f = 0 \) for a completely dark spot, and \( f = 1 \) for one indistinguishable from the stellar surface. We note that as the model can handle any value of flux ratio \( f \), faculae and plages (bright areas on the stellar photosphere and chromosphere, respectively) can also be modeled using a flux ratio value exceeding 1, as suggested, for example, by Boisse et al. (2012) and Kipping (2012).

3 APPLICATION TO HAT-P-11

HAT-P-11 is 9.6 visual magnitude K4 dwarf star in the field of the Kepler space telescope (Borucki et al. 2010). It has been known to host HAT-P-11b, a transiting hot Neptune on a 4.9 d orbit (Bakos et al. 2010), before the launch of the Kepler mission, and therefore has been observed with short (1 min) cadence from the beginning, during quarters 0–6, 9–10, 12–14, and 16–17. Missing data in quarters 7, 11, and 15 are due to the failure of a readout module in 2010 January. In quarter 8, only long-cadence observations were taken. We perform our analysis on this data set, using the flux values in the SAP_FLUX column, and dividing each transit by a linear fit to the out-of-transit data within 0.12 d from the mid-transit time to normalize the transit light curves. Visual inspection shows that a linear fit is satisfactory, because the out-of-transit light-curve modulation time-scale is the rotation period of HAT-P-11, which is 29.2 d (Bakos et al. 2010; Béky et al. 2014), much larger than the 0.24 d of the total width of our window.

For our analysis, we adopt the orbital eccentricity and argument of periastron values reported by Bakos et al. (2010) based on radial velocity data and Hipparcos parallax for HAT-P-11. However, we use the revised transit ephemeris, planetary radius and orbital semi-major axis relative to the stellar radius, orbital inclination, and limb darkening parameters of Deming et al. (2011). Their treatment relies on the above eccentricity and argument of periastron values, but accounts for eclipsed and uneclipsed spots, that biased earlier analyses. We number the transits according to this ephemeris, with the mid-transit time of transit 0 being \( T_0 = 2454605.89155 \pm 0.00013 \) (barycentric dynamical time).

Rebrightening events in the transit light curve of HAT-P-11 due to spots were first predicted by Winn et al. (2010), and first reported independently by Sanchis-Ojeda & Winn (2011) and Deming et al. (2011). They are used to constrain the stellar rotational period by Béky et al. (2014), who also compare a model out-of-transit light curve based on the MCMC chains described in this section, and find that it is consistent with the assumption that out-of-transit variation is dominated by rotation of spots.

3.1 Analysis of individual spots

First, we present the analysis of two individual starspots in order to study correlations between spot parameters inherent to the model. Each spot is described by four parameters: \( x \) and \( y \) are the coordinates of the spot centre in stellar radius units, as seen by the observer, in a Cartesian coordinate system whose origin is the centre of the stellar disc, and where the planet is moving approximately in the positive \( x \) direction during transit. (More precisely, for inclined orbits, the \( y \)-axis is defined by projecting the line of sight on the orbital plane, and then projecting that on the sky plane. For an inclined eccentric orbit, the projected velocity of the planet is not exactly parallel to the \( x \)-axis at mid-transit, except for special values of the argument of periastron.) The other two parameters are the spot radius \( a \) and the flux ratio \( f \) described in Section 2.1.

Fig. 2 presents the light curves of transits 74 and 218. For both transits, we identify the deviation from the light-curve model of Mandel & Agol (2002) as an indication for the planet eclipsing a single spot on the surface of the star. The light-curve anomalies are observed about 0.4 h before mid-transit during transit 74, and about half an hour after mid-transit during transit 218. We also plot the best-fitting (least sum of squared residuals) SPOTROD light curves for both transits in red in this figure.

In fact, for transit 218, we find two best-fitting solutions: in the projection plane, the spot can either be situated above or below the transit chord. To illustrate this bimodality, we present in Fig. 3 an observer’s view of the star (large empty circle), the planet (black filled circle), and the best-fitting solutions for the single spot (grey ellipses) during transits 74 and 218. The transit chord is also drawn, as wide as the diameter of the planet. We do not find the same bimodality in the case of transit 74; therefore, only one solution is depicted. Note that in fact we plot two model light curves for transit 218 in Fig. 2, corresponding to the best fits for each mode. However, they are indistinguishable in this figure, which is closely related to our inability to infer which solution describes the spot in reality.

We run parallel tempered MCMC simulations (Earl & Deem 2005) for both transits, using the emcee software package (Foreman-Mackey et al. 2013), at 10 different temperatures, with 100 concurrent chains at each temperature. We first run both simulations for
Figure 3. Projected images of HAT-P-11 during transits 74 (left-hand panel) and 218 (right-hand panel). Large circle is the star, solid black small circle is the planet HAT-P-11b at mid-transit, grey strip is the transit chord, and one or two grey ellipses are best-fitting solutions for the one or two modes of the spot.

1000 steps for burn-in. By inspecting the evolution of the mean and scatter of each parameter, we find that the chain already converges after about half this many steps. Then we run the simulation for another 1000 steps to sample what we believe to be the equilibrium distribution. The lowest, zero temperature chain provides us with the equilibrium distribution, whereas the higher temperature chains guarantee that we explore the entire parameter space and have samples in disconnected modes in numbers proportional to the corresponding posterior probabilities.

Fig. 4 illustrates the results of the MCMC simulation for the single spot models for transits 74 and 218: joint distributions for four pairs of parameters as well as histograms for each parameter are presented. A dashed line is drawn at the $y$ coordinate of the planet at mid-transit, which is very close to the impact parameter $b$ (in fact they would be the same for a circular orbit), to help distinguish the two modes and inspect symmetry with respect to the transit chord. We immediately confirm that there is only one mode for transit 74 and two modes for transit 218. The reason for this is that the anomaly during transit 74 can be well described by a spot that lies under the transit chord; therefore, the two modes overlap. On the other hand, the light curve of transit 218 can only be well modelled if the spot is further away from the transit chord, in which case the two modes are disconnected. We note that we experience bimodality for roughly one quarter of all spots in the entire Kepler data set for HAT-P-11.

Fig. 4 also tells us about the correlations of spot parameters. First, we note that in both cases, $x$ is better constrained than $y$,

Figure 4. Spot parameter distributions of 100,000 MCMC samples for transits 74 (top panels) and 218 (bottom panels) of HAT-P-11b. Four scatter plots show joint distributions of pairs of parameters for each transit: $x$-$y$, $a$-$y$, $f$-$y$, and $a$-$f$. Thin dashed line indicates the $y$ coordinate of planet at mid-transit. The four side panels on top and right present histograms of $x$, $a$, $f$, and $y$. 
even within a single mode. Recall that $x$ is the coordinate (almost exactly) parallel to the transit chord; therefore, it directly relates to when the anomaly is observed, which is well defined by the observations. On the other hand, we shall see that $y$ correlates with other parameters that together shape the transit anomaly, resulting in a larger uncertainty.

We also note that the two solutions of transit 218 have slightly different best-fitting $x$ values, even though $x$ is fairly well constrained. The explanation for this is that since the spot is elliptical in projection, with its semi-major axis not quite parallel to the $y$-axis; therefore, the two spot solutions must have different $x$ coordinates in order to intersect the transit chord at roughly the same $x$ coordinate. This effect can also be observed in Fig. 3.

The joint distribution of $a$ and $y$ for the spot in transit 218 in Fig. 4 tells us that the larger the spot is, the further it has to be from the transit chord. This is expected from geometrical arguments, as the duration of the transit anomaly can be well constrained from the observations. This correlation between spot parameters has been first pointed out by Silva (2003). However, this effect does not show for transit 74, probably due to the spot being close to the transit chord.

If we inspect the joint distribution of $f$ and $y$ for either transit, we notice that if the spot is brighter (has a larger flux ratio), then it is likely to be closer to the transit chord. Such a correlation was first noted by Wolter et al. (2009), and it remains to be explained.

Finally, we notice that there is a strong correlation between $a$ and $f$ for transit 74: the brighter the spot is, the larger it has to be. This correlation has been reported by Pont et al. (2007), Wolter et al. (2009), and Tregloan-Reed et al. (2013). This phenomenon is naively explained by that the rebrightening amplitude is proportional to how much the flux blocked by the planet is less than it would be for the unspotted photosphere. If a spot is not that much darker than the typical stellar surface, a larger area is required to produce the same flux deficit. This argument is expected to hold for spots centred on the transit chord, for which the spot radius directly determines the occulted area. On the other hand, we do not observe the same $a$–$f$ correlation for transit 218, because $a$ correlates strongly with $y$, resulting in a more complex effect on the occulted spot area.

### 3.2 Test on synthetic light curves

In this section, we generate synthetic transit light curves, and apply to them the same analysis as in the last section. We use PRISM (Tregloan-Reed et al. 2013) for generating light curves, and analyse them using SPOTROD. Using different models for data generation and analysis allows us to validate them against each other, that is, make sure that parameters like spot radius and flux ratio are interpreted identically, and they produce the same result.

We take the best-fitting parameters of the spots in transits 74 and 218, convert the projected coordinates to equatorial coordinates as expected by PRISM, and generate two model transits. The difference from the SPOTROD model with the same input parameters has mean $10^{-8}$ and standard deviation $2 \times 10^{-6}$ across all observation times, which indicates good agreement between the two models. For comparison, the mean photon noise is $8 \times 10^{-5}$ for the same data points.

We find that the SPOTROD best fit has residuals with a standard deviation approximately 1.3 times that predicted by the SAP_FLUX_ERR data column of the Kepler data set, and therefore we add independent, normally distributed noise scaled to 1.3 times the corresponding SAP_FLUX_ERR value to each data point calculated by PRISM. We run an MCMC simulation on the resulting synthetic light curves for both transits, in a fashion identical to that explained in the previous section. The resulting chain distribution is presented in Fig. 5.

Comparing Figs 4 and 5, we find that the chains converge to roughly the same parameters, further confirming that PRISM and SPOTROD interpret input parameters in compatible ways. We also find that the extent and shape of the equilibrium distributions, that is, the correlations between spot parameters, are roughly the same. In the case of transit 218, even though the synthetic light curve is generated based on only one of the two modes, we are unable to determine which mode it is from the MCMC analysis: the distribution is bimodal, just like it was for Kepler observations.

The ultimate test to decide whether bimodality and parameter correlations are inherent properties of light-curve models, and not unique to the implementation we use, would be to run the MCMC analysis using PRISM or another model, other than SPOTROD. However, since all previously known models require numerical integration in two dimensions, this would be prohibitively computationally expensive. Instead, we rely on the very small difference of the two light-curve models when run with the same input parameters to conclude that SPOTROD reproduces the results of PRISM.

### 3.3 Distribution of spot parameters

In order to study the spot ensemble distribution, we look for anomalies in 204 transits of HAT-P-11b of which there are complete, high-quality short-cadence Kepler data. First, we identify transit anomalies by visual inspection, fit spot parameters using guesses as initial values, and run MCMC simulations. In a few cases, we find that the chain abandons our initial fit and converges to a solution of a much larger spot with flux ratio close to one, representing a much longer duration and smaller amplitude light-curve anomaly. We attribute this to either noise or the effect of a large number of small spots, neither of which we prefer to incorrectly treat as a single, very large spot. We therefore decrease the number of modelled spots by one for such transits, or omit transits that exhibit such a behaviour with a single spot model, in five cases in total. For another 10 transits, at most 25 per cent samples of the chain are in an isolated mode representing such an unphysical spot; we discard these samples while keeping the rest of the chain.

We end up with 203 spots in 130 transits. For each transit, we independently run the same parallel tempered MCMC simulations as described in Section 3.1, using the EMCEE software package, at 10 different temperatures. For 73 transits with one spot each, we use 100 concurrent chains, for 43 transits with two spots each, 200 concurrent chains, for 12 transits with three spots each, we employ 300 chains, and for 2 transits with four spots each, 400 chains. We have four dimensions of parameter space for each spot: $x$, $y$, $a$, and $f$. We run the simulations for 1000 steps that we discard. Again, inspection of the chain shows that roughly half of this is already enough for convergence. Then we run the chain for another 1000 steps to sample the supposedly equilibrium distribution. This yields the data set that we use for our analysis in this paper, and the same data set is used by Béký et al. (2014).

The next step is to quantify how much better fit these models provide than if we modelled the light curve without the spots. In order to do so, we calculate the Bayesian information criterion (BIC), which is the sum of $\chi^2$ and an additional term penalizing extra model parameters to avoid overfitting (Schwarz 1978). We find that every single one of the 203 spots yields a BIC value at least 25.0 lower than the model without that spot, which is a very
significant improvement, justifying every spot in our analysis. This suggests that probably more spots could be carefully included.

We present the spot radius–flux ratio distribution of all spots in Fig. 6. Note that the distributions of most individual spots overlap, except for a few, which appear as separate clusters in the figure. The distribution is bounded from the side of small radius and large flux ratio by an observational bias: whether we visually inspect the light curves or apply an algorithmic transit anomaly search, small-amplitude anomalies will be lost in photon noise. The maximum eclipsed spot area is the smaller of \( R_p^2 \) and \( a^2 \) relative to the stellar disc; therefore, the maximum transit anomaly amplitude for given \( a \) and \( f \) is \( A_{\text{max}} = \min(R_p^2, a^2)(1 - f) \) (not accounting for limb darkening and projection distortion). The actual amplitude of the transit anomaly can be less, depending on \( y \). Fig. 6 shows the \( A_{\text{max}} = 0.0002 \) curve in red, where 0.0002 is an arbitrarily chosen value, which seems to act as an approximate amplitude threshold for anomalies that we detected.

We note that for small spots, individual spot distributions spread along constant \( \min(R_p^2, a^2)(1 - f) \) curves. The reason is that this quantity describes the total amount of flux missing due to the spot, which is directly related to the shape of the transit light-curve anomaly, and therefore can be well constrained. For small spots, the anomaly does not last long; therefore, there are fewer data points than for large spots. This makes it difficult to resolve the degeneracy between radius and flux ratio. For large spots, however, the transit anomaly has a larger amplitude, which can directly be used to infer the flux ratio. Note that consistently with this argument, the rightmost panels of Fig. 4 show that the flux ratio of the small spot seen in transit 74 has a much larger uncertainty than that of the large spot seen in transit 218, even though they have comparable relative uncertainties of their radii.

This effect results in small spots having a weaker constraint on flux ratio. However, as we will see later in this section, this does not indicate the presence of small, dark spots. Because of this large flux ratio uncertainty of small spots, we are not able to detect a significant correlation between spot radius and flux ratio for the 203 spots studied.

Fig. 6 also presents the spot radius–\( x \) coordinate joint distribution of all identified spots. We find two active latitudes, at \( x \approx -0.2 \) and 0.4, where there are the most spots and where spots seem to be the largest. This is in agreement with the findings of Deming et al. (2011) and Sanchis-Ojeda & Winn (2011). Note how this panel illustrates that the \( x \) coordinate can be very well constrained by the time of transit anomalies.
Figure 6. Spot parameter distributions of the entire MCMC chain of all 203 spots. Three scatter plots present joint distributions of $a$–$f$, $a$–$x$, and $a$–$y$. On the top scatter plot, red curve is the transit anomaly amplitude threshold $A_{\text{max}} = 0.0002$. On the bottom scatter plot, thin dashed line indicates the $y$ coordinate of planet at mid-transit. The four side panels on top and right present histograms of $a$, $f$, $x$, and $y$.

Finally, the bottom-left panel of Fig. 6 shows the joint distribution of spot radius and $y$ coordinate for all spots. Because of the polar orbit of HAT-P-11b, the latter roughly corresponds to the longitude of the spot on the stellar surface. In this case, we can assume that there is no physical correlation between the two parameters; therefore, we have to interpret the joint distribution in terms of observational biases. The observed radius is bounded from below, because very small spots would not cause a detectable signal in the light curve. For values of $y$ further from the transit chord, the smallest detectable radius increases for geometrical reasons: the spot has to overlap with the strip that the planet scans on the stellar surface. Finally, the spot radius is bounded from above by the physical distribution of spots, and we expect this to be independent from $y$. However, this is not reflected in Fig. 6: we see that MCMC states with radius above 0.2 prefer values of $y$ further from $b$. We interpret this as an artefact: we suspect that the light-curve anomaly due to irregularly shaped spots or spot groups, when mistakenly interpreted as a single spot, results in a large spot further from the transit chord. We believe that this radius is unphysical, because we never see transit anomalies that last long enough to require a spot of similar size with $y \approx b$ as an explanation. Therefore, we conclude that the upper limit of physical spot radius distribution is at most somewhat lower than 0.2, the radius of the largest detected spot with $y \approx b$. It is also possible that this is not a single spot either, so the actual upper radius limit might be smaller than this.

The four side panels in Fig. 6 show histograms of spot radius, flux ratio, and $x$ and $y$ coordinates, generated from the MCMC chains for all spots. The same observational biases are reflected here: there are no very small ($a \approx 0$) spots, and no very bright ($f \approx 1$) spots, because these would be undetectable. $x$ has a bimodal distribution according to the two active latitudes, and $y$ is concentrated around the impact parameter, because this is where even small spots are eclipsed by the planet.

To investigate the radius and flux ratio distribution of spots, we plot the median of these parameters in increasing order in Fig. 7. The top panel presents the spot radius and the bottom shows the flux ratio. The horizontal axes indicate the rank of the spot in the order of median values. In addition to the median, we shade the $1\sigma$ and $3\sigma$ intervals of the parameter distribution of each individual spot. The advantage of this presentation over the histograms of Fig. 6 is that we can disentangle the spread of a spot parameter for an individual spot due to uncertainties from the spread of the ensemble distribution due to spots being different.

Together with the histograms of Fig. 6, Fig. 7 helps us confirm the observational biases against small size and large flux ratio. In addition, we note that most spots are smaller than $a \approx 0.15$ (three times the radius of HAT-P-11b, which is approximately $0.059 R_\star$), with very few spots around the size of $a \approx 0.2$. We believe that larger
spots are artefacts, because they are only seen with y different from b.

On the other hand, even though bright spots are more frequent than dark ones, our first impression is that there are a number of almost completely black spots, that is, spots with f close to zero. (Note that a large number density in terms of a spot parameter, that is, large bin count in the histogram, corresponds to a less steep curve when data points are plotted in increasing order of that parameter.) However, when closely inspecting the parameter uncertainties of individual spots, we see that the ones that seem to be consistent with being very dark show a flux ratio distribution that includes much larger flux ratios as well. In fact, only two spots have a 3σ confidence interval that excludes flux ratios above 0.5, even though 30 spots have median flux ratio and 36 spots have best-fitting (least sum of squared residuals) flux ratio lower than 0.5. On the other hand, more than half of all spots studied have a 3σ confidence interval that allows flux ratios below 0.5. Therefore, we cannot either prove or disprove the existence of spots darker than f = 0.5. On the other hand, many spots with flux ratios from approximately 0.6 to 0.9 have small flux ratio uncertainties, suggesting the existence of spots brighter than f = 0.5.

3.4 Spot temperature

Assuming that the stellar photosphere and the spots both radiate as black bodies, one can calculate the spot temperature T_s from the flux ratio f and the effective photospheric temperature T_eff. This method was first applied to infer the spot temperature in the context of spot-induced transit light-curve anomalies by Silva (2003).

We can use the same method to calculate spot temperatures on HAT-P-11 from SPOTROD MCMC results. We use the photospheric effective temperature result T_eff = 4780 ± 50 K of Bakos et al. (2010), and we integrate Planck’s law over the Kepler response function to determine the flux ratio as a function of the spot temperature. The temperature difference ΔT = T_eff − T_s corresponding to certain values of flux ratio f are displayed on the right vertical axis of the bottom panel in Fig. 7.

As we noted in Section 3.3, we cannot infer anything about the existence of spots with flux ratio less than 0.5. Therefore, we cannot determine whether there are spots with flux ratio below the corresponding temperature of T_s = 4180 ± 40 K, or a temperature difference of ΔT = 600 ± 10 K.

On the other hand, Fig. 7 tells us that there are spots with well-constrained flux ratios ranging approximately from f = 0.6 to 0.9. In terms of temperature, this means that there exist spots ranging approximately from T_s = 4330 to 4680 K (ΔT = 450 to 100 K). It is possible that HAT-P-11 also has brighter spots (that are not detected due to observational bias) and darker spots.

For comparison, Walton et al. (2003) show that the Sun exhibits spots ranging from f ≈ 0.15 (with a ≈ 0.03) up to f ≈ 0.7 (with a ≈ 0.01), where the flux ratio is measured at 672.3 nm with 10 nm bandpass. This shows that HAT-P-11 is not the only star where individual spots are thought to exhibit a large range of flux ratios.

An advantage of transit anomalies over spectroscopic methods is the ability to measure temperatures of individual spots. For example, O’Neal et al. (2004) study TiO absorption features to determine the spot temperature to be T_s = 3350 ± 115 K on EQ Vir, a BY Dra-type flare star with T_eff = 4380 K, of spectral type K5 Ve, close to K4 of HAT-P-11. This corresponds to a flux ratio of f = 0.21 ± 0.05 in the Kepler bandpass. We note that this value is lower than the flux ratio of the majority of spots we found on HAT-P-11, though we were not able to either confirm or exclude the existence of spots with such low flux ratios. However, the result of O’Neal et al. (2004) is averaged over all spots visible on the stellar disc, and the temperature range of individual spots cannot be determined by their method.

3.5 Transit anomaly duration and amplitude

In addition to investigating the parameters taken by SPOTROD, x, y, a, and f, it is also interesting to study directly observable properties of spot eclipses: the duration and amplitude of the transit anomaly. Given the results of MCMC calculations, the easiest way to extract these properties is to measure them on model light curves. We therefore draw 1000 states from each chain and generate model light curves with one spot each. We calculate the amplitude of the transit anomaly as the maximum deviation from a spotless model and the duration of the transit anomaly as the length of the time interval on which the model with one spot predicts more flux than the spotless model. The resulting distribution is plotted in Fig. 8.

The transit anomaly duration–amplitude distribution is confined from three sides. The amplitude is bounded from below, as described in Section 3.3. We draw a red line in Fig. 8 at A = 0.0002, the same amplitude value as in Fig. 6. Note, however, that the transit anomaly amplitude is represented directly in Fig. 8, whereas in Fig. 6, we could only calculate the maximum possible value A_max for given values of a and f.

Fig. 8 tells us that visual inspection as performed by the authors results in an amplitude limit that is mostly constant across anomaly durations. Anomalies selected programmatically, however, might have a different boundary, as it is possible to identify an anomaly with a smaller amplitude in noisy data if it lasts sufficiently long.

From the side of short anomalies, the distribution is bounded by geometrical arguments. The amplitude of the anomaly tells us how much flux is missing with respect to a spotless transit. This places a lower limit on the geometrical extent of the spot, which then translates to the duration of the event. To characterize this boundary, we feed 10 000 random black spots to SPOTROD, and plot the envelope of their duration–amplitude distribution in black in Fig. 8. Grey spots (f > 0) would cause anomalies with the same duration but smaller amplitude than if they were black (f = 0); therefore, we only draw black spots to determine this boundary. We notice that the MCMC distribution extends close to this boundary, which means that the chains contain spot states that are almost

Figure 8. Distribution of transit anomaly duration and amplitude in black dots. Each spot is resampled 1000 times from its MCMC chain. The A = 0.0002 constant amplitude line is shown in red. Black curve represents the boundary of parameter space imposed by our model.
entirely black. As discussed in Section 3.3, however, this does not mean that the best fit for those spots is necessarily black.

From the right and above, however, the distribution does not extend to its theoretical limits: the maximum conceivable amplitude would be the transit depth (for a black spot that is larger than the planet), and the maximum conceivable transit duration would be the transit duration. The largest amplitude we observe is roughly half of this, and the longest anomaly lasts only about one quarter of the entire transit. We interpret these limits as indications of the actual spot parameter distribution, namely the lack of large dark spots, and the upper limit on spot size $a \lesssim 0.2$. We also note that in Fig. 8, the longest transit anomalies seem to have small amplitudes, for which we cannot offer an explanation.

While measuring directly observable quantities like transit anomaly duration and amplitude on models regenerated by SPOTROD is very convenient, we note that this inevitably introduces biases. For example, if we observe an anomaly with duration and amplitude that would place it on the left of the black curve in Fig. 8, that could not be reproduced by SPOTROD, and consequently we would measure different duration and amplitude values with our method. However, our MCMC analysis disentangles multiple spots and measures transit anomaly efficiencies, and still yields meaningful conclusions about the darkness of spots and the amplitude threshold for detection.

3.6 Residuals

Finally, we study the distribution of residuals to assess the goodness of fit and to compare our model with 203 spots to the spotless light-curve model. We calculate light curves using SPOTROD and the Mandel–Agol model for the two cases, respectively. We only consider the residuals at observations that took place during transits (between first and fourth contact). We use the SAP_FLUX_ERR column calculated by the Kepler Photometric Analysis module as the error estimate of the light-curve data in the SAP_FLUX column. Fig. 9 displays the normalized residual distribution histograms: in black when calculated using the spotless model, and in red when calculated using our model with spots, with a logarithmic vertical scale. The spotless model residuals have a large excess on the positive side, which we attribute to the presence of spot-induced transit light-curve anomalies. On the other hand, our SPOTROD model accounts for enough spots to make the residual distribution fairly symmetric. We note that this might potentially be used as a detection method to identify targets that exhibit spot-induced transit anomalies.

In terms of the Mandel–Agol fit, it is interesting to note that the largest residual is $25.75\sigma$. Comparing this to the transit depth of roughly $50\sigma$, we see that the largest transit anomaly has an amplitude of half the transit depth, just as we noted in Section 3.5. (The exact transit depth expressed in units of photon noise varies due to the quarterly rotations of the Kepler satellite.)

The relative error of HAT-P-11 photometry is very small ($\lesssim 10^{-4}$). Assuming that it is dominated by photon noise, the error distribution can be approximated with an independent normal distribution for each data point. In this case, it is valid to calculate $\chi^2$ and, from that, reduced $\chi^2$. For the models without and with spots, we get the strikingly different values $\chi^2_{\text{Spotless}} = 1.7 \times 10^2$ and $\chi^2_{\text{Spots}} = 3.3 \times 10^2$, respectively.

The total number of observations taken during the 130 transits is $N = 18,135$. Since we fix orbital parameters, transit ephemeris, and limb darkening, the spotless model has no fit parameters. With the number of data points as the degrees of freedom, we get $\chi^2_{\text{red, spotless}} = 9.2$, which, being much larger than unity, motivates a model with more free parameters.

Our spot model has four fit parameters for each spot, that is $P = 4 \times 203 = 812$ fit parameters in total. Note, however, that Andrae, Schulze-Hartung & Melchior (2010) prove that it is not justified to calculate the degrees of freedom as $K = N - P$ for non-linear models like SPOTROD, and in fact no reliable method is known to calculate $K$ in general. Therefore, we give the value $\chi^2_{\text{red, Spots}} = 1.9$ for reference only, and are left with the symmetry of the histogram as the only way to quantify the goodness of the fit.

4 CONCLUSION

In this paper, we present SPOTROD, a transit light-curve model accounting for both eclipsed and uneclipsed starspots. The advantage of our model over previous methods is that in polar coordinates, we integrate analytically with respect to the polar angle; therefore, time-consuming numerical integration only remains to be performed along a single dimension, the radial coordinate. This feature makes our model fast enough not only for fitting, but also for efficient statistical investigations using, for example, MCMC. A free and open source implementation of our model is publicly available.

The model assumes that spots follow the same limb darkening law as the stellar photosphere, consistent with observations of the Sun (Walton et al. 2003). It also assumes that spots are homogeneous. Umbra–penumbra structure can be mimicked by superimposing two concentric spots, while bright features can also be modelled using a flux ratio exceeding unity.

We apply our model to Kepler data of HAT-P-11 transits. We investigate correlations between fit parameters of individual spots, and confirm findings of previous investigations using similar models. We also study the size and flux ratio distributions. We establish an upper limit of $\lesssim 0.2$ for the spot radius, and find strong indication for the presence of spot with flux ratio ranging from 0.6 to 0.9, corresponding to an effective temperature $100-450$ K lower than that of the spotless photosphere. We cannot prove nor disprove the existence of spots with flux ratio less than 0.5. We do not find a significant correlation between spot size and flux ratio.

Figure 9. Residual histograms of short-cadence Kepler observations that were taken during transits of HAT-P-11b, normalized by SAP_FLUX_ERR. Black histogram corresponds to residuals with respect to the Mandel–Agol light-curve model, red histogram corresponds to residuals with respect to the best-fitting SPOTROD model with a total of 203 spots. Vertical axis (counts) is on a logarithmic scale.
While HAT-P-11 is unique in its brightness and large transit anomaly amplitudes within the Kepler field, SPOTROD can potentially be used to model Kepler observations of other transiting planetary hosts. In addition, comparable quality photometric observations are expected to be taken of a much larger number of stars, for example, by the K2 mission of the Kepler satellite (Haas et al. 2014) and by the Transiting Exoplanet Survey Satellite (Ricker et al. 2010). While our model provides a good fit to observations of HAT-P-11, it is important to keep in mind that our perfectly circular and homogeneous spots are a simplified version of what the stellar surface actually looks like. However, mapping out the projected stellar surface by, for example, two-dimensional deconvolution with the planetary disc, is an underdetermined and potentially numerically unstable inversion problem. The general advantage of model simplifications is that the small degree of freedom makes fitting robust. This is exactly what SPOTROD provides: a simplistic, approximate, but robust and fast way to model transit light curves of spotted stars.

ACKNOWLEDGEMENTS

Work by BB and MJH was supported by NASA under grant NNX09AB28G from the Kepler Participating Scientist Program and grants NNX09AB33G and NNX13A124G under the Origins programme. DMK is funded by the NASA Carl Sagan Fellowships. This paper includes data collected by the Kepler mission. Funding for the Kepler mission is provided by the NASA Science Mission directorate. The MCMC computations in this paper were run on the Odyssey 2.0 cluster supported by the FAS Science Division Research Computing Group at Harvard University. BB is grateful to Robert Noyes for useful comments on the manuscript, and to John A. Johnson for suggesting the name SPOTROD.

REFERENCES

Andrae R., Schulze-Hartung T., Melchior P., 2010, preprint (arXiv:1012.3754)
Bakos G. A. et al., 2010, ApJ, 710, 1724
Béké B., Holman M. J., Kipping D. M., Noyes R. W., 2014, ApJ, 788, 1
Boisse I., Bonfils X., Santos N. C., 2012, A&A, 545, A109
Borucki W. J. et al., 2010, Science, 327, 977
Czesla S., Huber K. F., Wolter U., Schröter S., Schmitt J. H. M. M., 2009, A&A, 505, 1277
Deming D. et al., 2011, ApJ, 740, 33
Désert J.-M. et al., 2011, ApJS, 197, 14
Earl D. J., Deem M. W., 2005, Phys. Chem. Chem. Phys., 7, 3910
Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, PASP, 125, 306
Haas M. R. et al., 2014, Am. Astron. Soc. Meeting Abstr., 223, 228.01
Kipping D. M., 2012, MNRAS, 427, 2487
Mandel K., Agol E., 2002, ApJ, 580, L171
O’Neal D., Neff J. E., Saar S. H., Cuntz M., 2004, AJ, 128, 1802
Oshagh M., Boisse I., Boué G., Montalbo M., Santos N. C., Boufils X., Haghhipour N., 2013, A&A, 549, A35
Pál A., 2009, MNRAS, 396, 1737
Pont F. et al., 2007, A&A, 476, 1347
Rabus M. et al., 2009, A&A, 494, 391
Ricker G. R. et al., 2010, Am. Astron. Soc. Meeting Abstr., 215, 450.06
Sanchis-Ojeda R., Winn J. N., 2011, ApJ, 743, 61
Schwarz G., 1978, Ann. Stat., 6, 461
Silva A. V. R., 2003, ApJ, 585, L147
Tregloan-Reed J., Southworth J., Tappert C., 2013, MNRAS, 428, 3671
Walton S. R., Preminger D. G., Chapman G. A., 2003, Sol. Phys., 213, 301
Winn J. N. et al., 2010, ApJ, 723, L223

APPENDIX A: DERIVATIONS

A1 Geometry

Let $x_1$ and $y_1$ denote the coordinates of the centre of the spot as seen by the observer, in a Cartesian coordinate system with the centre of the stellar disc as the origin, in stellar radius units. Then the angle between the plane of the spot boundary and the line of sight is $\beta = \arccos \sqrt{x_1^2 + y_1^2}$. In projection (as seen by the observer), the spot is an ellipse with semi-major axis $a$ and semi-minor axis $b = a \sin \beta = a \sqrt{1 - (x_1^2 + y_1^2)}$. The centre of this ellipse is the projection of the intersection point of the axis of the cone and the plane in which the spot boundary lies, not the projection of the centre of the spot which is on the stellar surface. The centre of the ellipse lies at a distance of $\sqrt{1 - a^2}$ from the centre of the star; therefore, in projection, the distance between the centre of the stellar disc and the centre of the ellipse is $z = \sqrt{1 - a^2} \cos \beta = \sqrt{1 - a^2} \sqrt{x_1^2 + y_1^2}$. Expressing $x_1^2 + y_1^2$ from equation (A2) and substituting into equation (A1), we can express $b$ in terms of $a$ and $z$: $b = a \sqrt{\frac{1 - z^2}{1 - a^2}}$. The input parameters of the subroutine integrate transit are $x_1$, $y_1$, and $z$; it calculates $z$ from equation (A2), and passes it to ellipse angle, which in turn calculates $b$ using equation (A3).

A2 Calculating $\gamma$, $\gamma^*$, and $\delta$

To calculate $\gamma(r)$, consider a Cartesian coordinate system in the projection plane with the centre of the ellipse as the origin, the $x$-axis parallel to the semi-major axis, and the $y$ to the semi-minor axis. Let the centre of the stellar disc be at $(0, z)$, and let $(x, y)$ denote an intersection point of the ellipse and $C_r$. See Fig. A2. Then $(x, y)$

![Figure A1](https://example.com/figure.png)

**Figure A1.** A cross-section in the plane containing the centre of the star, the centre of the spot, and the observer. The observer is to the right at an infinite distance.
The spot, the centre of the star, and $\gamma_R - \arccos \frac{2 \gamma}{\pi}$ if $-\gamma_2 = C + \gamma$ according to the following cases:

$x + y$ denote the angle between the centre of the spot and 

$r \theta$ denote the radius of the planet and 

$\pi + r y \geq \gamma = y \theta + \gamma$ using (planet arc inside spot arc)

$\gamma = \arccos \frac{z - y}{r}$.

Appendix A3 discusses the case of four intersection points. One or three intersection points are singular cases between other cases, and can be treated along with the case on either side.

Let $R_p$ denote the radius of the planet and $z_p$ its projected separation from the centre of the stellar disc, both in stellar radius units. To calculate $\delta(r)$, we could repeat the above derivation with $a = b = R_p$. Or we can use the law of cosines in the triangle defined by the centre of the star, the centre of the planet, and the intersection point of the edge of the planet with $C_r$:

$$R_p^2 = r^2 + z_p^2 - 2rz_p \cos \delta(r),$$

as seen in Fig. A3. Again, the cases of $C_r$ being disjoint from or entirely occulted by the planet should be tested for separately, yielding $\delta = 0$ and $\delta = \pi$, respectively.

Finally, let $\theta$ denote the angle between the centre of the spot and the centre of the planet as seen from the centre of the star, which we also calculate using the law of cosines. We always choose the angle for which $0 \leq \theta \leq \pi$. Now $\gamma^*(r)$ is determined by $\gamma(r), \delta(r),$ and $\theta$ according to the following cases:

$$\gamma^* = \begin{cases} 
\gamma & \text{if } \theta \geq \gamma + \delta \text{ (arcs disjoint)} \\
\gamma - \delta & \text{if } \gamma > \theta + \delta \text{ (planet arc inside spot arc)} \\
0 & \text{if } \delta \geq \gamma + \delta \text{ (spot arc inside planet arc)} \\
\gamma + \frac{\theta - \delta}{2} & \text{if } \gamma + \delta + \theta \leq 2\pi \text{ (partial overlap)} \\
\pi - \delta & \text{o/w, if } \gamma + \delta + \theta > 2\pi \text{ (circular overlap).} 
\end{cases}$$

Figure A2. The spot, the centre of the star, and $C_r$, as seen from the direction of the observer, in a Cartesian coordinate system centred on the centre of the spot ellipse (not the projection of the spot centre).

Figure A3. The triangle in the sky plane defined by the centre of the star, the centre of the planet, and one intersection point of the edge of the planet with $C_r$. In the first case, the arcs are disjoint; therefore, none of the spot arc is eclipsed by the planet. This happens to spot 1 in Fig. 1. In the second case, the entire planet arc gets subtracted from the spot arc. The part of the spot arc that is not occulted by the planet is now composed of two arcs on either side of the planet, and we define $\gamma^*$ as half the total central angle of them. In the third case, the planet occults the entire spot arc. ‘Otherwise’ for the last two cases means that the triangle inequality holds between $\gamma, \delta,$ and $\theta$. In the fourth case, the planet and spot arcs overlap in a single arc. This happens to spot 2 in Fig. 1. In the last case, they overlap in two arcs. This can happen only if at least one of the planetary disc and the spot ellipse contains the centre of the stellar disc in projection. This situation is illustrated by Fig. A4.

The function circleangle calculates $\beta$ from $r, z_p,$ and $R_p,$ ellipseangle calculates $\gamma$ from $r, z,$ and $a,$ and finally integratetransit calculates $\gamma^*$ based on equation (A6), evaluates the integrals of equations (3) and (4), and calculates $F_{\text{normalized}}$ according to equation (5).

A3 Spots partially behind the limb

The boundary of the spot is assumed to be a circle; therefore, its projection (as seen by the observer) is an ellipse. However, we must investigate whether the boundary of what we see of the spot coincides with the projection of its boundary on the stellar surface. It is easy to see that it is indeed the case as long as no part of the spot
covers another part of it. That is, as long as the entire spot is visible, we will always see it as an ellipse. (Remember the caveat that the centre of this ellipse is not the projection of what we defined as the centre of the spot.)

However, when the spot partially hides behind the stellar limb, we will see it as a shape bounded by an arc of an ellipse (the projection of the spot boundary) and an arc of a circle (the edge of the stellar disc), as illustrated in Fig. A5. We prove in the next section that in this case, the ellipse touches the stellar limb from the inside.

The two cases (the entire spot visible or part of it is behind the limb) are delineated by a critical value of $\beta$ or, equivalently, a critical value of $z$, both of which we expect to depend only on $a$.

From Fig. A1, we can see that in this critical case,

$$
\beta_{\text{crit}} = \alpha
$$

$$
z_{\text{crit}} = \sqrt{1 - a^2} \cos \beta_{\text{crit}} = \sqrt{1 - a^2} \cos \alpha
$$

$$
= \sqrt{1 - a^2} \cos \arcsin a = \sqrt{1 - a^2} \sqrt{1 - a^2}
$$

$$
= 1 - a^2,
$$

using equation (A2) to express $z$ in terms of $a$ and $\beta$.

If $z < z_{\text{crit}}$, then the entire spot is visible. If $z = z_{\text{crit}}$, then the ellipse touches the edge of the stellar limb at one point, as we prove in the next section. If $z > z_{\text{crit}}$, then part of the spot is behind the stellar limb. Note that since we describe the spot with the parameter $z$ (implicitly, through $x_s$ and $y_s$), $z = 1$ corresponding to the spot centre being on the stellar limb, therefore we cannot deal with spots that are partially visible, but their centre is behind the limb. Such spots, however, will have a small contribution to $F_s$ and $F_{\text{transit}}$, because only a very small part of them is visible; furthermore this small part is on the limb, which is usually darker to start with. Furthermore, since $F_s \approx F_{\text{transit}}$ as long as $R_g \ll 1$, omitting such a spot will have a very small effect on $F_{\text{normalized}}$.

### A4 Number of intersection points

In this section, we prove that as long as $z < z_{\text{crit}}$, the spot boundary and $C_r$ can have at most two intersection points. We also explain how to correctly model the ellipse in terms of $y$ if $z > z_{\text{crit}}$ and there are four intersection points. Finally, we prove that if $z = z_{\text{crit}}$, then the spot boundary ellipse touches the stellar limb at one point, and if $z > z_{\text{crit}}$, then at two points.

Let us express $x^2$ from equation (A4) and substitute into equation (A5). This results in a quadratic equation for $y$:

$$
0 = Ay^2 + By + C
$$

$$
A = \frac{a^2}{b^2} - 1
$$

$$
B = 2z
$$

$$
C = r^2 - a^2 - z^2
$$

$$
y_{\pm} = \frac{-z \pm \sqrt{z^2 - \left(\frac{a^2}{b^2} - 1\right) \left(r^2 - a^2 - z^2\right)}}{\frac{a^2}{b^2} - 1}.
$$

Equation (A4) tells us that each solution $y$ represents two intersection points if $|y| < b$, one if $|y| = b$, or none if $|y| > b$. Now let us consider the following inequality:

$$
y_{-} < -b
$$

$$
-\frac{z}{\frac{a^2}{b^2} - 1} < -b.
$$

This inequality is a sufficient condition for that $y_{-}$ does not represent real intersection points, that is, there are at most two intersection points (corresponding to $y_{+}$). Now we increase the left-hand side of inequality (A8). This will make it sharper, leading to a more restrictive, therefore still sufficient (but not necessary) condition. If we find at the end that this still holds whenever $z < z_{\text{crit}}$, that proves our original statement:

$$
-\frac{z}{\frac{a^2}{b^2} - 1} < -b.
$$

Assume that $z > 0$, in which case we also have $a > b$ and thus $A > 0$. Then multiplying by $\frac{a}{b}$ does not change the direction of inequality:

$$
-\frac{b}{a} < -\frac{b^2}{az} \left(\frac{a^2}{b^2} - 1\right)
$$

$$
\frac{a}{z} - \frac{b}{az} < \frac{b}{a}.
$$

Substituting $b$ from equation (A3), we get

$$
\frac{a}{z} - \frac{a (1 - \frac{z^2}{1 - a^2})}{z (1 - \frac{a^2}{1 - a^2})} < \sqrt{1 - \frac{z^2}{1 - a^2}}
$$

$$
\frac{az}{1 - a^2} < \sqrt{1 - \frac{z^2}{1 - a^2}}.
$$

Both sides of inequality (A12) are positive; therefore, we can square them to get an equivalent inequality:

$$
\frac{a^2 z^2}{(1 - a^2)^2} < 1 - \frac{z^2}{1 - a^2}
$$

$$
\frac{a^2 z^2 (1 - a^2)^2}{(1 - a^2)^2} < (1 - a^2)^2 - (1 - a^2) z^2
$$

$$
z^2 < (1 - a^2)^2
$$

$$
z < 1 - a^2
$$

$$
z < z_{\text{crit}}.
$$

(A13)
This inequality is \( \frac{1}{2} \) times inequality (A11) with the inequality sign in the other direction. Inequality (A11), in turn, is equivalent to inequality (A13). That is, the condition on the curvature of the ellipse is equivalent to inequality (A13) with the inequality sign in the other direction: \( z > z_{\text{crit}} \). This is consistent with our previous statements.

Further increasing \( r \) will yield four distinct intersection points, with \(-b < y_- < y_+ < b\). This scenario is also demonstrated in Fig. A5. By continuity, the outside pair of intersection points corresponds to \( y_+ \), because they exist ever since \( r = z - b \). The inside pair only appeared at \( r = z + b \), and therefore corresponds to \( y_- \), which crossed \(-b\) at the same value of \( r \).

Our code always calculates \( \gamma \) based on \( y_+ \). This means that the entire arc between the outside intersection points is considered to be part of the spot, which correctly describes the shape of the spot as the observer sees it. That is, the code gives the correct result even in the case \( z > z_{\text{crit}} \).

Finally, if \( r = 1 \), then the inside and outside pairs of intersection points coincide according to Figs A5 and A6: the ellipse touches the stellar limb from the inside. Another way of saying this is \( y_- = y_+ \), which happens exactly if the discriminant of equation (A7) is zero.

We now prove this statement,

\[
B^2 - 4AC = 0
\]

\[
4z^2 - 4\left(\frac{a^2}{b^2} - 1\right)\left(r^2 - a^2 - z^2\right) = 0
\]

\[
z^2 - \left(1 - \frac{1}{1 - \frac{a^2}{b^2}} - 1\right)\left(1 - a^2 - z^2\right) = 0
\]

\[
z^2 - \left(1 - \frac{1 - a^2}{1 - a^2 - z^2} - 1\right)\left(1 - a^2 - z^2\right) = 0
\]

\[
z^2 = \frac{z^2}{1 - a^2 - z^2} - 1\left(1 - a^2 - z^2\right) = 0.
\]

The last equation trivially holds true, which proves that if \( r = 1 \), then \( y_- = y_+ \). Note that this is true regardless of the value of \( z \): if \( z < z_{\text{crit}} \), then \( y_- < y_+ < b \) (no real solution for \( x \), no intersection points); if \( z = z_{\text{crit}} \), then \( y_- = y_+ = -b \) (touching in a single point with \( x = 0 \), quadruple root); and if \( z > z_{\text{crit}} \), then \( y_- = y_+ > -b \) (touching at two points, multiple roots each, like in Figs A5 and A6), as seen from inequalities (A10)–(A13). This concludes our proof.

We are also allowed to multiply by \((1 - a^2)^2\), which has to be positive. Finally, we arrive exactly at the critical value of \( z \) that we have already established.

This derivation shows that if inequality (A13) holds, then so does inequality (A8). That is, if the spot is entirely visible, then there are at most two intersection points.

To understand the dependence of the number of intersection points on \( r \), we plot \( y_\pm \) as a function of \( r \). In both cases, \( b = 0.12, z = 0.78, r_1 = 0.66, r_2 = 0.90 \).

This paper has been typeset from a \TeX/\LaTeX file prepared by the author.