

LOW-MASS ECLIPSING BINARIES IN THE INITIAL *KEPLER* DATA RELEASE

J. L. COUGHLIN^{1,5}, M. LÓPEZ-MORALES^{2,3,6}, T. E. HARRISON¹, N. ULE¹, AND D. I. HOFFMAN⁴

¹ Department of Astronomy, New Mexico State University, P.O. Box 30001, MSC 4500, Las Cruces, NM 88003-8001, USA; jlcough@nmsu.edu

² Carnegie Institution of Washington, Department of Terrestrial Magnetism, 5241 Broad Branch Road NW, Washington, DC 20015, USA

³ Institut de Ciències de l'Espai (CSIC-IEEC), Campus UAB, Facultat de Ciències, Torre C5, parell, 2a pl, E-08193 Bellaterra, Barcelona, Spain

⁴ California Institute of Technology, MC 249-17, 1200 East California Boulevard, Pasadena, CA 91125, USA

Received 2010 July 24; accepted 2010 December 3; published 2011 February 2

ABSTRACT

We identify 231 objects in the newly released Cycle 0 data set from the *Kepler Mission* as double-eclipse, detached eclipsing binary systems with $T_{\text{eff}} < 5500$ K and orbital periods shorter than ~ 32 days. We model each light curve using the JKTEBOP code with a genetic algorithm to obtain precise values for each system. We identify 95 new systems with both components below $1.0 M_{\odot}$ and eclipses of at least 0.1 mag, suitable for ground-based follow-up. Of these, 14 have periods less than 1.0 day, 52 have periods between 1.0 and 10.0 days, and 29 have periods greater than 10.0 days. This new sample of main-sequence, low-mass, double-eclipse, detached eclipsing binary candidates more than doubles the number of previously known systems and extends the sample into the completely heretofore unexplored $P > 10.0$ day period regime. We find preliminary evidence from these systems that the radii of low-mass stars in binary systems decrease with period. This supports the theory that binary spin-up is the primary cause of inflated radii in low-mass binary systems, although a full analysis of each system with radial-velocity and multi-color light curves is needed to fully explore this hypothesis. Also, we present seven new transiting planet candidates that do not appear among the list of 706 candidates recently released by the *Kepler* team, or in the *Kepler* False Positive Catalog, along with several other new and interesting systems. We also present novel techniques for the identification, period analysis, and modeling of eclipsing binaries.

Key words: binaries: eclipsing – binaries: general – stars: fundamental parameters – stars: late-type – stars: low-mass

Online-only material: color figures, extended figure, figure set, machine-readable and VO tables

1. INTRODUCTION

A double-lined, detached, eclipsing binary (DDEB) is a system that contains two non-interacting, eclipsing stars, in which the spectra of both components can be clearly seen, allowing for the radial velocity (RV) of each component to be obtained. In these systems, the mass and radius of each star can be determined with errors usually less than 1%–2%, thus making DDEBs currently the most accurate method of obtaining masses and radii of stars. Models of main-sequence stars with masses similar to or greater than the Sun have been tested over the years using DDEBs. Popper (1980) compiled available masses and radii of DDEB's with accuracies of $\leq 15\%$, up to that date, and found general agreement with stellar models, though stressed the need for more accurate observations and models. Andersen (1991) provided a compilation of all available DDEB systems up to that date, with accuracies $\leq 2\%$, and showed that the masses and radii of these stars were in general agreement with the current stellar evolution models, with any discrepancies attributable to abundance variations. Torres et al. (2010) recently performed a similar review with nearly double the sample of DDEBs. They were able to show the need to include non-classical effects such as diffusion and convection in stellar models, definitively demonstrate the existence of significant structural differences in magnetically active and fast-rotating stars, test theories of rotational synchronization and orbital circularization, and validate general relativity via apsidal motion rates. However, while observations of DDEBs have enhanced our understanding of stellar structure and evolution for stars

with $M \geq 1.0 M_{\odot}$, low-mass, main-sequence (LMMS) stars ($M < 1.0 M_{\odot}$ and $T_{\text{eff}} < 5800$ K) have not been tested to the same extent.

Although a couple systems with late-G or early-K-type components had been studied prior to 2000 (cf. Popper 1980; Andersen 1991; Torres et al. 2006; Clausen et al. 2009, and references therein), only three LMMS DDEBs with late-K or M-type components were known (Lacy 1977; Leung & Schneider 1978; Delfosse et al. 1999). This number had only increased to nine by the beginning of 2007 (cf. López-Morales 2007, Table 1). Despite the fact that the majority of main-sequence stars are low-mass stars, they are both intrinsically fainter, and physically smaller, than their more massive counterparts. Therefore, they have a lower eclipse probability and are harder to discover and study. As outlined by López-Morales (2007), analysis of these systems showed that the observed radii for these stars are consistently $\sim 10\%$ – 20% larger than predicted by stellar models (Baraffe et al. 1998) for $0.3 M_{\odot} \lesssim M \lesssim 0.8 M_{\odot}$. Fernandez et al. (2009) recently showed that this was also likely the case for five M dwarfs in short-period eclipsing systems with an F-type primary, though since the systems are only single-lined, the masses could not be determined directly. This discrepancy between the radii derived from models and from observations either reveals a flaw in the stellar models for this mass regime, or is due to differences in metallicity, magnetic activity, or interpretation of the light curve data when star spots are present (Morales et al. 2008). As to this last point, Morales et al. (2010) recently noted that improperly taking polar spots into account in the light curve modeling process may possibly cause the derivation of stellar radii a few percent larger than the true values for some of these systems. Of all of these scenarios,

⁵ NSF Graduate Research Fellow.

⁶ Hubble Fellow.

enhanced magnetic activity has been proposed as the principal cause of inflated radii (Chabrier et al. 2007; López-Morales 2007; Morales et al. 2008).

If enhanced magnetic activity is the principal cause of the inflated radii, shorter-period binary systems, with the stellar rotation rate enhanced by the revolution of the system, would be expected to show greater activity and thus larger radii than longer-period systems (Chabrier et al. 2007). Binary systems with component masses of $0.5 M_{\odot}$ are expected to synchronize, and therefore be spun-up, in less than 0.1 Gyr for periods less than 4 days, and in less than 1 Gyr for periods less than 8 days (Zahn 1977, 1994). Thus, the discovery of LMMS DDEBs with $P \gtrsim 10$ days, where the binary components should have natural rotation rates, is crucial to probing if enhanced rotation due to binarity is the underlying cause of this phenomenon. This theory might be supported by measurements of a couple isolated field M and K dwarf stars via very long baseline interferometry, which Demory et al. (2009) found to match stellar models. However, recently a much larger sample of nearly two dozen isolated M and K dwarf stars finds, for $\sim 80\%$ of the sample, larger radii than the model predictions for $0.35 M_{\odot} < M < 0.65 M_{\odot}$ (Boyajian et al. 2010), indicating that there are likely multiple causes of inflation at work, or a remaining flaw in the stellar models.

Though several more LMMS DDEB systems have been found since 2007 (Coughlin & Shaw 2007; Shaw & López-Morales 2007; Becker et al. 2008; Blake et al. 2008; Devor et al. 2008a, 2008b; Shkolnik et al. 2008; Hoffman et al. 2008; Irwin et al. 2009; Dimitrov & Kjurkchieva 2010; Shkolnik et al. 2010), there are to date only seven well-studied systems with $1.0 \text{ day} < P < 3.0 \text{ days}$ (López-Morales 2007, and references therein; Becker et al. 2008; Shkolnik et al. 2008), and only one has a larger period, at $P = 8.4 \text{ days}$ (Devor et al. 2008b). This is mostly due to the fact that ground-based photometric surveys, such as NSVS, TrES, and OGLE, are either cadence, precision, magnitude, or number limited, and thus not sensitive to long periods. The *Kepler Mission*, with three years of constant photometric monitoring of over 150,000 stars with $V \lesssim 17$, at 30 minute cadence and submillimagnitude precision, is the key to discovering a large number of long-period, LMMS DDEBs.

In this paper, we present the results of our search through all the newly available *Kepler* Q0 and Q1 public data for LMMS DDEBs. Section 2 describes the data we use in this paper. Section 3 describes our binary identification technique, and Section 4 describes how we model the light curves. Our selection and list of new LMMS DDEBs is presented in Section 5, and we present new transiting planet candidates in Section 6. In Section 7, we compare the new LMMS DDEBs with theoretical models and conclude with a summary of our results in Section 8. Once accurate mass and radius values exist for a large range of both mass and period, our understanding of these objects should substantially improve, and we will be one step closer to extending to the lower-mass regime the advanced study of stellar structure and evolution that Sun-like and high-mass stars have been a subject of for some time.

2. OBSERVATIONAL DATA

The data used in our analysis consist of the 201,631 light curves made public by the *Kepler Mission*⁷ as of 2010 June 15 from *Kepler* Q0 and Q1 observations. All light curves can be accessed through the Multi-mission Archive at STScI (MAST).⁸

The data consist of 51,366 light curves from *Kepler* Q0 (observed from 2009 May 2 00:54:56 to 2009 May 11 17:51:31 UT) and 150,265 light curves from *Kepler* Q1 (observed from 2009 May 13 00:15:49 to 2009 June 15 11:32:57 UT) each at 29.43 minute cadence. Individual light curves for Q0 contain ~ 470 data points and for Q1 contain ~ 1600 data points. Targets range in *Kepler* magnitude from 17.0 at the faintest, to 5.0 at the brightest.

The *Kepler* team has performed pixel level calibrations (including bias, dark current, flat field, gain, and nonlinearity corrections) identified and cleaned cosmic-ray events, estimated and removed background signal, and then extracted time-series photometry using an optimum photometric aperture. They have also removed systematic trends due to spacecraft pointing, temperature fluctuations, and other sources of systematic error, and corrected for excess flux in the optimal photometric aperture due to crowding (Van Cleve 2010). It is this final, “corrected” photometry that we have downloaded for use in our analysis.

3. ECLIPSING BINARY IDENTIFICATION

Prša et al. (2010) have recently released an initial catalog of eclipsing binary stars they find in the *Kepler* field from the same Q0 and Q1 data we use in this paper. They first identified EB candidates via *Kepler’s* Transit Planet Search algorithm, eliminating those targets already identified as exoplanet candidates. To determine the ephemeris of each candidate, they used Lomb–Scargle, Analysis of Variance, and Box-fitting Least Squares periodogram techniques, combined with manual inspection and modification. They then culled, through manual inspection, non-EB candidates, such as pulsating and heavily spotted stars, as well as duplicates due to contamination from nearby stars, and arrive at their final list of 1832 binaries, which are manually classified as detached, semi-detached, over-contact, ellipsoidal, or unknown. Next, they estimate the principal parameters of each system (temperature ratio, sum of the fractional radii, $e \cdot \cos(\omega)$, $e \cdot \sin(\omega)$, and $\sin(i)$ for detached systems), via a neural network technique called Eclipsing Binaries via Artificial Intelligence (“EBAI”; Prša et al. 2008). For our search, which focuses on the detection of LMMS DDEBs, we have devised our own DDEB identification technique, which we apply to the Q1 data. We do not use the Q0 data in this part of the analysis to avoid discrepant systematics between the two quarters, which complicate the analysis.

Our search consisted of two steps. The first was to identify variable stars, and to do so, we placed a light curve standard deviation limit above which the objects are classified as variables. We first subtracted an error-weighted, linear fit of flux versus time from all data, to remove any remaining linear systematic trends, and then plotted the standard deviation of each light curve versus its average flux and fit a power law. These data are shown in Figure 1, where the black dots correspond to light curves which deviate by less than 1σ from the standard deviation versus average flux fit, and we thus classify as non-variable. The colored dots indicate the variable candidates that deviate by more than 1σ . Next, we used the flux ratio (FR) measurement criterion, which we adapted from the magnitude ratio given in Kinemuchi et al. (2006), and is defined as

$$\text{FR} = \frac{\text{maximum flux} - \text{median flux}}{\text{maximum flux} - \text{minimum flux}} \quad (1)$$

as a measure of whether or not the variable spends most of its time above (low FR value) or below (high FR value) the

⁷ <http://kepler.nasa.gov/>

⁸ <http://archive.stsci.edu/kepler/>

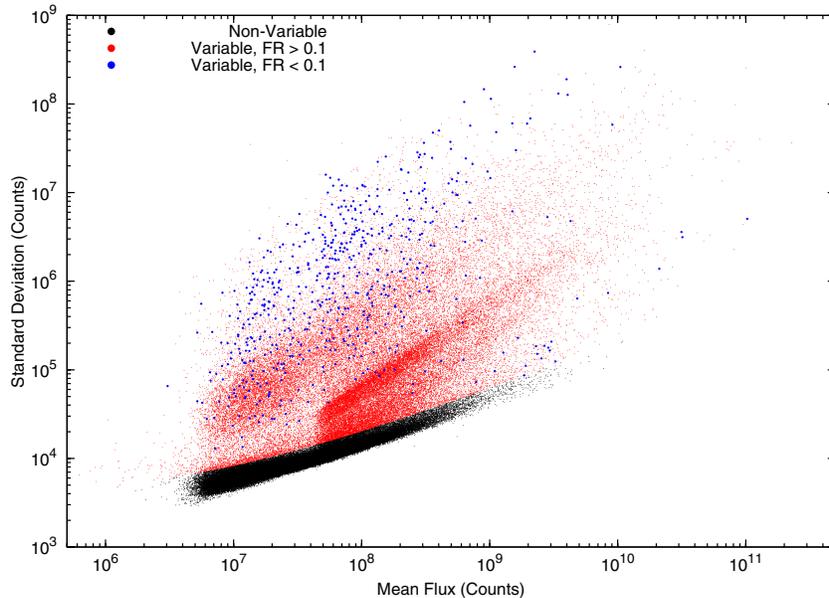


Figure 1. Plot of standard deviation vs. mean flux for the 150,265 stars in Q1. Black dots represent stars that vary by less than 1σ from a best-fit power law to the data, and thus we classify them as non-variables. Red dots represent variables with a flux ratio greater than 0.1. Blue dots represent variables with a flux ratio less than 0.1 and thus are good candidates to be eclipsing binaries.

(A color version of this figure is available in the online journal.)

median flux value. Perfectly sinusoidal variables have $FR = 0.50$, pulsating variables, such as RR Lyrae’s, have $FR > 0.5$, and eclipsing binaries have $FR < 0.5$. As we are principally interested in finding well-detached systems with relatively deep, narrow eclipses, which thus have low FR values, we make a further cut of the systems and only examine those variables with $FR < 0.1$, shown by blue dots in Figure 1.

The second step of the analysis was to determine the orbital period of each candidate. This was done using two independent techniques that are both well suited for detached eclipsing binary systems. The first is phase dispersion minimization (PDM) (Stellingwerf 1978), which attempts to find the period that best minimizes the variance in multiple phase bins of the folded light curve. This technique is not sensitive to the shape of the light curve and thus is ideal for non-sinusoidal variables such as detached eclipsing binaries. The downside of this technique is that if strong periodic features exist in the light curve, which do not correspond to the period of eclipses, such as rapidly varying spots, stellar pulsations, or leftover systematics, they can weaken the signal of the eclipse period. We use the latest implementation given by Stellingwerf (2006) and determine the best three periods via this technique to ensure that the true period is found, and not just an integer multiple, or fraction, thereof.

The second technique we use is one we invented specifically for detached eclipsing binaries, and call eclipse phase dispersion minimization (EPDM). The idea behind EPDM is that we want to automatically identify and align the primary eclipses in an eclipsing binary, thus finding the period of the system. To accomplish this, EPDM finds the period that best minimizes the dispersion of the actual phase values of the faintest N points in a light curve, i.e., the very bottom of the eclipses. Since EPDM only selects the N faintest points in a light curve, it is not affected by systematics or periodic features that do not correspond to the period of eclipses, assuming the systematics do not extend below the depth of the eclipses. The technique works for all binary systems with equal or unequal eclipse depths, and transiting planets, both with either zero or non-

zero eccentricity. Computationally, EPDM is significantly faster than traditional PDM techniques. For a detailed and illustrative explanation of this new technique, please see Appendix A. We use EPDM to find the three best-fit periods for each system as well, for the same reasons as we did with PDM.

We identify 577 EB candidates in the Q1 data. Of these, 486 are listed by Prša et al. (2010) as detached eclipsing binaries, and 20 are identified as semi-detached eclipsing binaries. The 71 remaining candidates were manually inspected by examining both the raw and phased light curves at the six best periods found via PDM and EPDM. Of these 71 remaining candidates, 48 turned out to be false positives with significantly large, sharp systematic features, and one is an apparent red giant (Kepler 010614012, $T_{\text{eff}} = 4859$ K, $\log g = 3.086$, $[M/H] = -0.641$, $R_{\star} = 5.708 R_{\odot}$) with an unusual, asymmetrical, eclipse-like feature that lasts for ~ 3 days with a depth of 1.2%, shown in Figure 2. This does not appear to be a systematic feature due to its very flat out-of-eclipse baseline, contiguous nature, long duration, and the actual time at which the feature occurs, compared to the majority of other objects with strong systematics. The remaining 22 targets are: 2 transiting exoplanet candidates contained in the recently released list of 306 candidates by Borucki & the Kepler team (2010), 3 already published transiting planets (Kepler-5b, Kepler-6b, and TrES-2b), 7 shallow eclipsing systems with primary eclipse depths ranging from 1.4% to 5.7%, visible secondary eclipses ranging from 0.05% to 4.6%, and periods ranging from 4.7 to 45.3 days, the already published transiting hot compact object Kepler 008823868 (Rowe et al. 2010), a 6.4 day eclipsing binary with $T_{\text{eff}} = 5893$ K and eclipse depths of 38.4% and 12.2% (Kepler 006182849), and 8 transiting exoplanet candidates with transit depths ranging from 0.75% to 4.9%, and periods ranging from 2.5 to 24.7 days. For the seven new extremely shallow eclipsing systems, we list their *Kepler* ID numbers, periods, effective temperatures, surface gravities, and primary and secondary eclipse depths in Table 1, and note they could be of interest for follow-up due to the potential to contain brown dwarf or extremely low-mass secondaries, or

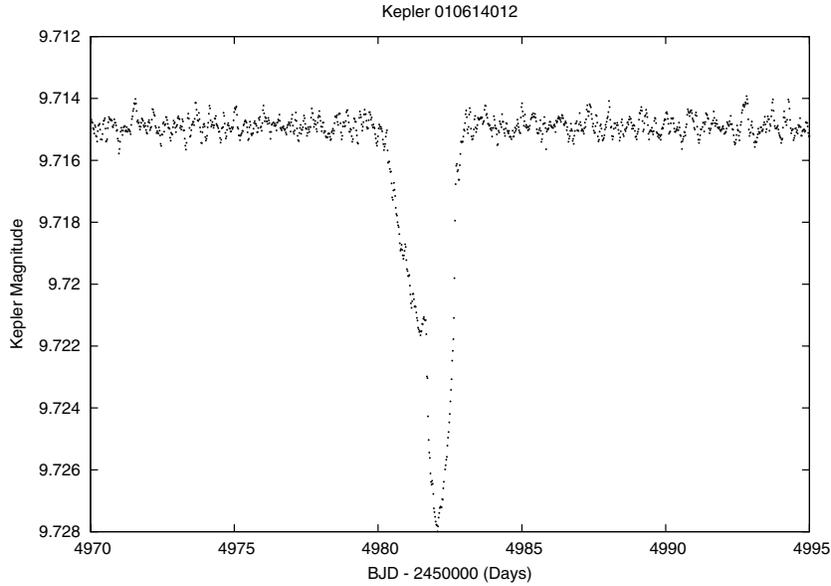


Figure 2. Kepler 010614012, an apparent red giant ($T_{\text{eff}} = 4859$ K, $\log g = 3.086$, $[M/H] = -0.641$, $R_{\star} = 5.708 R_{\odot}$) with a very unusual, shallow, eclipse-like feature.

Table 1

Period, Effective Temperature, Surface Gravity, and Eclipse Depth Estimates for the Seven New Extremely Shallow Eclipsing Systems

Kepler ID	Period (days)	T_{eff} (K)	$\log g$	Primary (%)	Secondary (%)
003098197	38.3840 ^a	5675	4.814	4.9	4.60
004178389 ^b	45.2600 ^a	5645	4.670	3.4	2.80
009016295 ^c	19.9858	5819	4.582	4.1	0.17
009071386 ^b	4.68513	6324	4.267	1.4	0.05
009838975 ^b	18.7000	5018	4.802	5.7	0.21
012017140 ^c	22.8624	6026	4.500	4.7	0.11
012504988 ^b	5.09473	5985	4.464	2.9	0.06

Notes.

^a System is listed in the *Kepler* False Positive Catalog as likely to be an EB.

^b System has non-zero eccentricity.

^c Period derived assuming zero eccentricity.

even anomalously hot exoplanet companions. Of the eight transiting candidates, only one is listed in the *Kepler* false positive catalog,⁹ Kepler 011974540. None of them are in the list of the 306 released candidates by Borucki & the Kepler team (2010), nor are among the 400 planetary candidates currently reserved for follow-up observations (Borucki & the Kepler team 2010). These will be further discussed in Section 6.

4. LIGHT CURVE MODELING

Since the system parameters determined by Prša et al. (2010) are only estimates and do not incorporate spots, and since we seek to obtain as accurate physical parameters as possible, we modeled each system using a robust global minimization scheme with a commonly used, physically detailed eclipsing binary modeling code. We took all 314 detached eclipsing binaries with $T_{\text{eff}} < 5500$ K and that are publicly available (five systems are still proprietary), identified from both our search and the Prša et al. (2010) catalog, combined Q0 and Q1 data if available, and via manual inspection classified systems as double-eclipse

(i.e., contained two visible eclipses), single-eclipse (i.e., only contained one eclipse), or as spurious results that were not recognizable as eclipsing systems. (Given the errors in the KIC temperature determination, and to ensure the primary is below $1.0 M_{\odot}$, we used 5500 K as our cutoff, instead of 5800 K. As well, the definition of a “double-lined” system is one in which the lines of both components are visible in an observed spectrum. Although in general if two eclipses are clearly visible in the photometric light curve, it is likely to be “double-lined,” this cannot be determined without an actual spectrum. Thus, we use the term “double-eclipse” throughout the paper, with the assumption that when observed spectroscopically, the majority of these systems will be observed as “double-lined.”)

We then used the JKTEBOP eclipsing binary modeling program (Southworth et al. 2004a, 2004b) to model every double-eclipse eclipsing binary system, of which there were 231, solving for the period, time of primary minimum, inclination, mass ratio, $e \cdot \cos(\omega)$, $e \cdot \sin(\omega)$, surface brightness ratio, sum of the fractional radii, ratio of the radii, and out-of-eclipse flux. In addition, we also solved for the amplitude and time of minimum of a sinusoidal term imposed on the luminosity of the primary component, with the period fixed to that of the binary, in order to account for spots. Note that in the JKTEBOP model the mass ratio is only used to determine the amount of tidal deformation of the stars from a pure sphere. Thus, it has no effect on the light curve of long-period systems, which due to their large separations are almost perfectly spherical, but must be included to properly model very short period systems, where the tidal deformation can have a significant impact on the light curve. We used the quadratic limb darkening law, which works well for late-type stars (e.g., Manduca et al. 1977; Wade & Rucinski 1985; Claret & Gimenez 1990), with coefficients set to those found by Sing (2010) for the *Kepler* bandpass via interpolation given the systems’ effective temperatures, surface gravities, and metallicities as listed in the Kepler Input Catalog (KIC).¹⁰ We also fixed the gravity-darkening exponent based on the effective temperature as prescribed by Claret (2000). As any contaminating flux from nearby stars in the photometric

⁹ http://archive.stsci.edu/kepler/false_positives.html

¹⁰ http://archive.stsci.edu/kepler/kepler_fov/search.php

aperture has already been compensated for in the *Kepler* pipeline (Van Cleve 2010), we set the amount of third light to 0.0. Note that third light might still exist in some systems if there is a background star or tertiary component that is unidentifiable from ground-based surveys (i.e., less than $\sim 1''$ separation), but since third light is usually unconstrained in a single-color light curve, we do not let it vary. If third light existed in a system and was not accounted for, the solution would result in an inclination determination lower than the true value, and therefore an overestimation of the stellar radius. However, this should only occur in a minority of systems. For a couple binaries in our list, the light curves absolutely could not be modeled without the inclusion of third light (i.e., very sharp eclipses with depths of less than 0.01 mag). For these cases only, we let the third light vary and thus be a non-zero parameter. Additionally, if the effect of spots in a light curve deviates significantly from the adopted sinusoidal shape, it could affect the derived luminosity ratio to a minor extent, but it should not affect the sum of the radii.

In order to model such a large number of systems over such a large solution space, and to ensure we have found the best global solution, we adapted the JKTEBOP code to use a modified version of the asexual genetic algorithm (AGA) described by Cantó et al. (2009), coupled with its standard Levenberg–Marquardt minimization algorithm. Genetic algorithms (GAs) are an extremely efficient method of fitting computationally intensive, multi-parameter models over a large and potentially discontinuous parameter space, and thus ideal for this work. For the details of how GAs work, and the specific changes we made to the Cantó et al. (2009) AGA, please see Appendix B.

We found that our modified AGA does an excellent job of solving well-behaved light curves, simultaneously varying all 12 aforementioned parameters over the entire range of possible solutions. For some of the systems however, strong systematics and/or variable star spots introduced a significant amount of noise, especially in systems with shallow eclipses, for which it was more difficult to arrive at a robust solution. For these systems we had to manually correct the systematics, often by either eliminating the Q0 or Q1 data, equalizing the base flux levels of Q0 and Q1 data, or subtracting out a quasi-sinusoidal variation in the base flux level due to remaining *Kepler* systematics. When possible we attempted to minimize the amount of manual interference. Hopefully this will become much less of a problem with subsequent data releases. We then re-ran the AGA using a larger initial population until a good solution was found. Every light curve in the end was visually inspected to be a good fit compared to the scatter of the data points, and the obtained parameters were confirmed to be reasonable when visually inspecting the light curves.

5. NEW LOW-MASS BINARY CANDIDATES

In order to identify the main-sequence stars from our list of 231 candidates, and determine the best candidates for follow-up, we employ the following technique to estimate the temperature, mass, and radius of each star using the sum of the fractional radii, r_{sum} , and period, P , obtained from our JKTEBOP models, the luminosity ratio, L_r (which is derived from the surface brightness ratio, J , and radii ratio, k , obtained from the models), and the effective temperature of the system, T_{eff} , obtained from the KIC, with an estimated error of ± 200 K.

The value for T_{eff} given in the KIC was determined via interpolation of standard color–magnitude relations as determined by ground-based, multi-wavelength photometry (Van Cleve 2010).

Although in principle one might be able to deconvolve two separate spectral energy distributions from this photometry, in reality given the level of photometric error in the KIC and uncertainty at which binary phase the photometry was obtained, this is untenable. Instead, we assume the stars radiate as blackbodies, and that each star contributes to the determined T_{eff} in proportion to its luminosity. Thus, following our assumption, we obtain the following relation:

$$T_{\text{eff}} = \frac{L_1 T_1 + L_2 T_2}{L_1 + L_2}, \quad (2)$$

where L_1 , L_2 , T_1 , and T_2 are the luminosities and effective temperatures of star 1 and 2, respectively. Still assuming the stars radiate as blackbodies, the luminosity of each star is proportional to its radius squared and temperature to the fourth power, with the temperature proportional to its surface brightness to the one-fourth power. Thus, we find that the luminosity ratio can be expressed as

$$\begin{aligned} L_r &= \frac{L_1}{L_2} = \frac{r_1^2 T_1^4}{r_2^2 T_2^4} = k^2 T_r^4 \\ &= k^2 \left[\left(\frac{\text{SB}_1}{\text{SB}_2} \right)^{1/4} \right]^4 = k^2 (J^{1/4})^4 = k^2 J, \end{aligned} \quad (3)$$

where SB_1 and SB_2 are the surface brightnesses of star 1 and 2, respectively, and r_1 and r_2 are the fractional radii of stars 1 and 2, respectively, defined as R_1/a and R_2/a , where R_1 and R_2 are the physical radius of each star, and a is the semi-major axis of, or separation between, the components. Combining Equations (2) and (3) yields the expression

$$T_{\text{eff}} = \frac{L_r T_1 + T_2}{L_r + 1}, \quad (4)$$

which has two known parameters, T_{eff} and L_r , and two unknown parameters, T_1 and T_2 . To place a further constraint upon the values of T_1 and T_2 , we make the assumption that both stars in the binary are on the main sequence, and employ the mass, temperature, radius, and average of the V -band and R -band luminosity relations given in Baraffe et al. (1998) for $0.075 M_\odot \leq M \leq 1.0 M_\odot$ and in Chabrier et al. (2000) for $M < 0.075 M_\odot$, both assuming an age of 5.0 Gyr and $[M/H] = 0.0$. (We average the V - and R -band luminosities to obtain a very close approximation to the *Kepler* bandpass.) From these models, for a given value of T_1 , there is only one value of T_2 which will reproduce the observed value of L_r . Thus, there only exists one set of unique values for T_1 and T_2 that reproduces both the observed T_{eff} and L_r values for the system.

For each T_1 and T_2 then, we obtain the absolute masses and radii (M_1 , M_2 , R_1 , and R_2) via interpolation from the Baraffe et al. (1998) and Chabrier et al. (2000) models. Then, utilizing Kepler’s third law, given the total mass of the system, we calculate the semi-major axis, a , via

$$a = (GM_{\text{tot}})^{1/3} \left(\frac{P}{2\pi} \right)^{2/3} \quad (5)$$

where M_{tot} is the total mass of the system, $M_1 + M_2$, and G is the gravitational constant. We then multiply each radius determined above by a constant so that the sum of the fractional radii derived from the JKTEBOP model, r_{sum} , is equal to $(R_1 + R_2)/a$,

the sum of the fractional radii when using the physical values of M_1 , M_2 , R_1 , R_2 , and P . This technique is robust because while individual parameters such as i , J , and k can suffer from degeneracies, especially in systems with shallow eclipses, the values of r_{sum} and $L_r = k^2 J$, which we rely on, are firmly set by the width of the eclipses and the difference in their eclipse depths, respectively.

For clarity, we now illustrate the individual steps of this procedure using the example of an actual system, Kepler 002437452. This system was found to have $T_{\text{eff}} = 5398$ K and $L_r = 3.90$ from the KIC and the JKTEBOP modeling, respectively. Now, assuming the stars are main sequence, one could choose values of $T_1 = 4000$ K and $T_2 = 3620$ K, and looking up their luminosities from the Baraffe et al. (1998) models, find that the luminosity ratio between two main-sequence stars with temperatures of 4000 K and 3620 K is 3.90. In this case, the luminosity ratio criterion would be satisfied, but T_{eff} would be ~ 3922 K, nowhere near the measured value of 5398 K. Similarly, one could choose values of $T_1 = 5400$ and $T_2 = 5393$, and this would yield $T_{\text{eff}} = 5398$ K, but L_r would be 1.01, nowhere near the needed value of 3.90. The unique solution that satisfies both the effective temperature and luminosity ratio constraints is that $T_1 = 5591$ and $T_2 = 4647$, which yields both $T_{\text{eff}} = 5398$ and $L_r = 3.90$. Now, given these temperatures, interpolating from the Baraffe et al. (1998) models yields values of $M_1 = 0.963 M_{\odot}$, $R_1 = 0.966 R_{\odot}$, $M_2 = 0.792 M_{\odot}$, and $R_2 = 0.783 R_{\odot}$. Taking the masses, and the period of the system of 14.47184 days, and utilizing Equation (5), we find that the semi-major axis, a , would be $30.1 R_{\odot}$. Dividing the sum of the estimated physical radii by the semi-major axis just calculated, we find a value of 0.058 for the sum of fractional radii. Now, from the JKTEBOP model, this system was measured to have a sum of the fractional radii of 0.084, and so it appears that the current values for the radii are underestimated. Thus, we multiply the radii by a factor of $0.084/0.058 = 1.45$, to obtain our final radii values of $R_1 = 1.40 R_{\odot}$ and $R_2 = 1.13 R_{\odot}$, with, as above, $M_1 = 0.96 M_{\odot}$, $T_1 = 5591$ K, $M_2 = 0.79 M_{\odot}$, and $T_2 = 4647$ K.

Kipping (2010) has recently examined the effects of the long (~ 30 minute) integration time of long-cadence *Kepler* observations on transit light curves, and found that it can significantly alter the morphological shape of a transit curve and result in erroneous parameters if not properly taken into account in the modeling procedure. Certainly, eclipsing binaries are also affected by long integration times, namely by a “smearing” of the eclipses so that they appear to be shallower and have a longer duration. Qualitatively, this would result in a lower inclination and larger sum of the fractional radii, while the luminosity ratio would remain unchanged, since both eclipse depths are equally affected. To quantitatively investigate the extent to which the long integration could affect the derived parameters, we generated model light curves of a typical eclipsing binary, varying its period and the sum of the fractional radii. We then binned these light curves as if they had a 29.43 minute integration time, and the same number of data points as the Q1 *Kepler* light curves. We then resolved the light curves without accounting for the integration time, and compared the computed parameters to those used to generate the original light curve. We found that for the long-cadence *Kepler* integration time of 29.43 minutes, only systems with very low values of r_{sum} and P are significantly affected, as can be seen in Figure 3. These types of systems are less than 2% of our sample. Nevertheless, we modified the JKTEBOP program to perform a numerical integration over a given exposure time, as

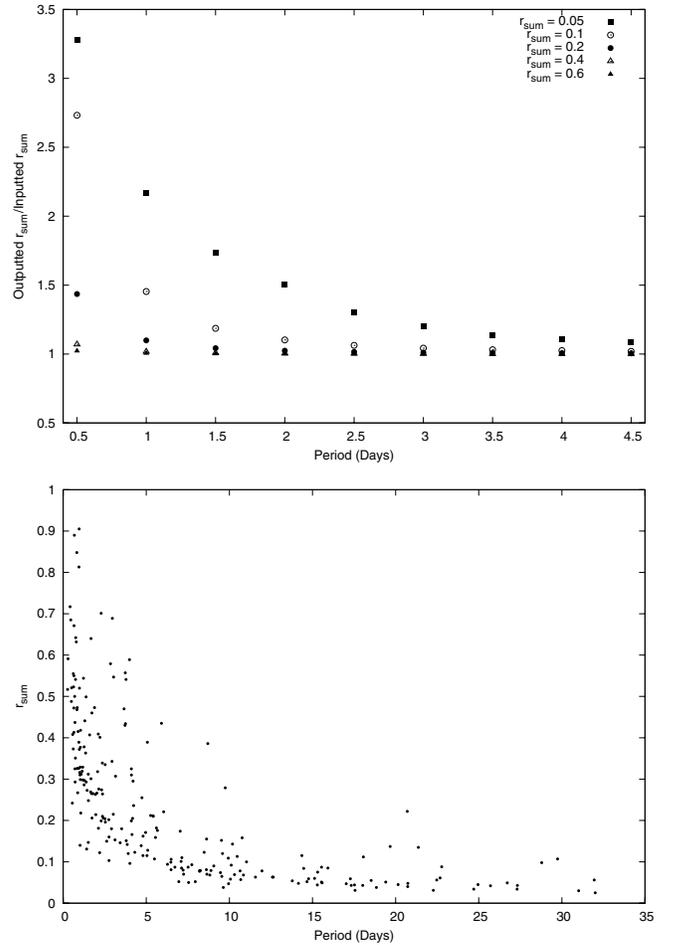


Figure 3. Left: the effect that the 29.43 minute integration time has on the derivation on the sum of the fractional radii, r_{sum} , at a given period. As can be seen, only very small values of r_{sum} and P yield discrepancies $\gtrsim 10\%$, for example, combinations of $P < 3$ days and $r_{\text{sum}} < 0.05$, $P < 1.5$ days and $r_{\text{sum}} < 0.1$, $P < 0.75$ days and $r_{\text{sum}} < 0.2$, etc. Right: the values of r_{sum} vs. period for the binaries we have modeled in this paper, presented in Table 2. Very few of the systems, $\lesssim 2\%$, in our sample lie in a region where they would be significantly affected by the 29.43 minute integration time.

suggested by Kipping (2010). We tested our modifications by solving the aforementioned generated light curves, now taking the integration time into account, and successfully retrieved the inputted parameters.

After estimating the individual mass, radius, and temperature for each component, we re-computed the gravity and limb-darkening coefficients for each individual star, and performed a Levenberg–Marquardt minimization starting from our previously best solutions, taking into account the 29.43 minute integration time. We then repeated the processes of deriving the physical values of the components, interpolating gravity and limb-darkening coefficients, and performing a Levenberg–Marquardt minimization several more times to ensure convergence. The JKTEBOP solutions for all initial 231 candidates are shown in Table 2, including the *Kepler* ID number, effective temperature of the system, apparent *Kepler* magnitude, magnitude range of the light curve, period, time of primary minimum, inclination, eccentricity, longitude of periastron, sum of the fractional radii, surface brightness ratio, radii ratio, luminosity ratio, amplitude of the sine curve applied to the luminosity of the primary star to account for spots, and the amount of third light. Although we list the derived surface brightness and radii

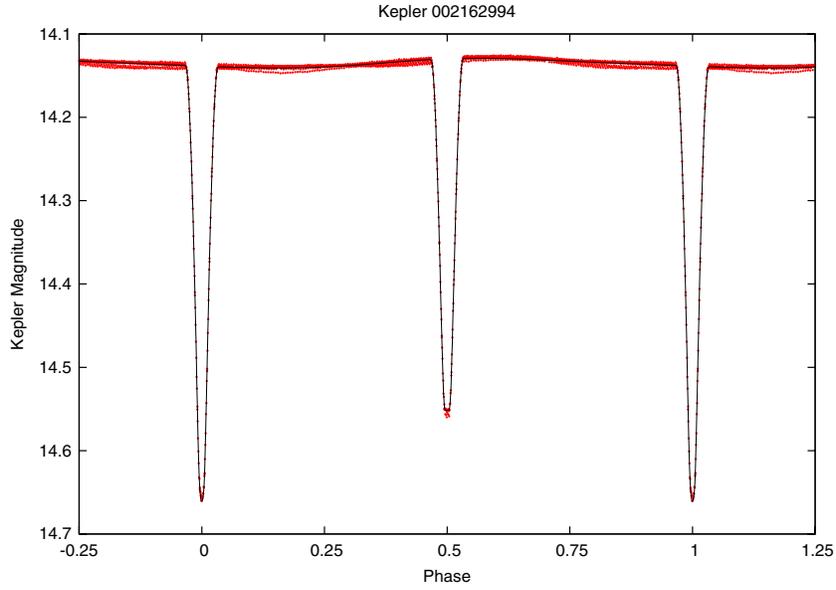


Figure 4. Plots of the light curves of the 231 systems modeled with the JKTEBOP code, presented in Table 2. Only the first plot is shown in the text for guidance. (A color version of this figure and the complete figure set (231 images) are available in the online journal.)

Table 2
Model System Parameters via JKTEBOP for the 231 DDEBs with $T_{\text{eff}} < 5500$ K

Kepler ID	T_{eff} (K)	m_{kep}	Δm_{kep}	Period (days)	T_0 (BJD-2450000)	i ($^\circ$)	e^a	ω^a ($^\circ$)	r_{sum}	J	k	L_r	L_1 Sine Amplitude	L_3
002162994	5410	14.162	0.535	4.101544	5002.545861	89.87	0.01	270	0.199	0.991	0.702	0.4888	0.008	0.00
002437452	5398	16.981	0.256	14.47184	5003.759350	87.46	0.08	90	0.084	0.641	2.47	3.905	0.011	0.00
002580872	5293	14.880	0.374	15.92672	4978.550988	87.95	0.26	102	0.084	1.14	1.25	1.774	0.014	0.00
002719873	5086	15.160	0.235	17.27953	4968.273250	87.76	0.31	90	0.059	0.633	2.64	4.425	0.007	0.00
002852560	5381	15.308	0.460	11.96119	4964.912794	88.06	0.44	41	0.079	1.04	0.986	1.008	0.000	0.00

Notes. ^a Although the values for e and ω are presented in this table for ease of reading, the values of $e \cdot \cos(\omega)$ and $e \cdot \sin(\omega)$ were actually solved for in the analysis.

(This table is available in its entirety in machine-readable and Virtual Observatory (VO) forms in the online journal. A portion is shown here for guidance regarding its form and content.)

ratios here, we note again that they are not always reliable on their own, and thus are combined to obtain the luminosity ratio in our analysis via Equation (3). Plots of each of the eclipsing binaries with their model fit are given in Figure 4.

As a check on the reliability of our analysis technique we took the well-studied low-mass eclipsing binary GU Boo (López-Morales & Ribas 2005), and modeled only the R -band light curve (not using the RV curves) via the exact same procedure as stated above in Sections 4 and 5. The only differences were that we used only the R -band luminosities from the Baraffe et al. (1998) and Chabrier et al. (2000) models, and an integration time of 2 minutes as stated in López-Morales & Ribas (2005). We used only the period, time of primary minimum, and estimated effective temperature of the system from broadband photometry provided in López-Morales & Ribas (2005), as we did for the systems in our main study. We find $T_1 = 3912$ K, $M_1 = 0.61 M_\odot$, $R_1 = 0.62 R_\odot$, $T_2 = 3813$ K, $M_2 = 0.57 M_\odot$, and $R_2 = 0.59$ via our technique. In comparison, López-Morales & Ribas (2005) found with multi-color light curves and RV curves of the system, values of $T_1 = 3920$ K, $M_1 = 0.610 M_\odot$, $R_1 = 0.623 R_\odot$, $T_2 = 3810$ K, $M_2 = 0.599 M_\odot$, and $R_2 = 0.620$. The values derived from our technique using only a single color light curve are accurate to within a few percent of the very precise values

derived from a study using multi-color light and RV curves, thus validating our technique.

As noted above, Prša et al. (2010) estimated the parameters of temperature ratio, sum of the fractional radii, $e \cdot \cos(\omega)$, $e \cdot \sin(\omega)$, and $\sin(i)$ for detached systems, via the EBAI technique (Prša et al. 2008). Before comparing our results to the parameters obtained by Prša et al. (2010), we note that the modeling approach between EBAI and our AGA presented in this paper have some fundamental differences. EBAI is extremely computationally efficient, but relies on a fitted polynomial to the actual data (Prša et al. 2008), which is then compared to a neural network training set of 33,235 light curves generated by the Wilson-Devinney code (Wilson & Devinney 1971; Wilson 1993). Prša et al. (2008, p. 551) note that “...the artificial neural network output is viable for statistical analysis and as input to sophisticated modeling engines for fine-tuning.” In comparison, the use of our AGA coupled with JKTEBOP is computationally slower, but models each actual data point, obtaining an actual best-fit model while varying all physical parameters of interest over the global solution space. As well, our AGA takes into account the 29.43 minute integration time, while EBAI does not. Thus, although the EBAI technique is excellent for mining large databases, identification of light curve morphology, and

Table 3
Temperature, Mass, and Radius Estimates for the 95 New LMMS DDEB Candidates with Amplitudes ≥ 0.1 mag and Both Masses $< 1.0 M_{\odot}$

<i>Kepler</i> ID	Period (days)	T_{eff} (K)	T_1 (K)	T_2 (K)	M_1 (M_{\odot})	M_2 (M_{\odot})	R_1 (R_{\odot})	R_2 (R_{\odot})
002162994	4.102	5410	5593	5038	0.96	0.86	1.39	1.24
002437452	14.47	5398	5591	4647	0.96	0.79	1.40	1.13
002719873	17.28	5086	5246	4382	0.90	0.73	1.08	0.86
002852560	11.96	5381	5385	5378	0.93	0.92	1.06	1.06
003003991	7.245	5366	5554	4598	0.96	0.78	0.83	0.67

(This table is available in its entirety in machine-readable and Virtual Observatory (VO) forms in the online journal. A portion is shown here for guidance regarding its form and content.)

obtaining estimates of parameters for statistical studies, it is not intended to model individual light curves as precisely and accurately as possible. Keeping this in mind, comparing the parameters obtained by Prša et al. (2010) to our solutions for the same systems, we first note a moderate correlation between the sum of radii given by Prša et al. (2010) and our results, with an average discrepancy of $\sim 20\%$. However, some of Prša et al.’s (2010) solutions are unphysical ($r_{\text{sum}} < 0.0$) and visual inspection of the polyfit curves given by Prša et al. (2010) appears to reveal a systematic underestimation of the eclipse depths. With respect to eccentricity, the parameters presented by Prša et al. (2010) reveal an unusually large number of eccentric systems, with only 3% of systems having $e \leq 0.01$, and 11% of systems with $e \leq 0.05$. In contrast, our parameters show 36% of systems with $e \leq 0.01$, and 60% of systems with $e \leq 0.05$, which better matches the large number of systems observed that do not show any offset of secondary eclipse from phase 0.5, and no difference in the eclipse widths, indicative of a circular orbit. There is only a slight correlation between our inclination values and those of Prša et al. (2010), but as we previously noted, their polyfit curves appear not to fit the eclipse depths well. There is practically no correlation between our values for the surface brightness ratio and EBAI’s temperature ratio provided in Prša et al. (2010, p. 35), though they note that for detached systems, the “...eclipse depth ratio is strongly affected by eccentricity and star sizes as well, rendering T_2/T_1 a poor proxy for the surface brightness ratio.”

In Table 3, we list the *Kepler* ID number, orbital period, effective temperature of the system, and the estimated effective temperature, mass, and radius of each stellar component for the 95 systems that contain two main-sequence stars, which we define as having a radius less than 1.5 times the Baraffe et al. (1998) and Chabrier et al. (2000) model relationships, and a light curve amplitude of at least 0.1 mag (suitable for ground-based follow-up and less likely to contain any third light). All of these 95 systems have both stars with masses less than $1.0 M_{\odot}$. Note that we have ordered Table 3 such that Star 1 is always the more massive star, regardless if L_r was greater or less than 1.0 in Table 2. Also note that since we are using V+R-band luminosities, which best correspond to the *Kepler* bandpass, one cannot always use the simple $R^2 \cdot T^4$ relation to derive luminosity ratios from Table 3 to compare to Table 2, since that would correspond to the bolometric luminosity. However, if one takes a system from Table 3, looks up the V+R-band luminosity for each component, based on their mass and temperature, from the Baraffe et al. (1998) and Chabrier et al. (2000) models, and derives a luminosity ratio, this will exactly match the luminosity ratio in Table 2 from the JKTEBOP models, because the technique defines it as such. These results substantially increase the number of LMMS DDEB candidates in general, and provide 29 new LMMS DDEBs with both components

below one solar mass, and at least 0.1 mag eclipse depths, in the heretofore unexplored period range of $P > 10$ days. We further discuss the impact of these systems and comparison to theoretical models in Section 7.

In the left panel of Figure 5 we show an example of a system which did not meet the main-sequence criterion, *Kepler* 004247791, which has $T_{\text{eff}} = 4063$ K and a period of 4.100866 days. If this system were main sequence, via our method, it would have a combined mass of $1.28 M_{\odot}$ and a combined radius of $3.82 R_{\odot}$. This can be seen by the wide, shallow eclipses for a system of this period and effective temperature. Thus, this system contains one or two evolved stars. An additional curiosity of this system is a periodic transit-like feature that is superimposed on the eclipsing binary light curve. The transit feature occurs at just slightly less than half the orbital period of the eclipsing binary, so that it is seen twice per every revolution of the eclipsing binary system, occurring at a slightly earlier phase every revolution. We subtract the model fit from the eclipsing binary and plot the transit feature at its period of 2.02484 days in the right panel of Figure 5. Some possible explanations may include, but are certainly not limited to (1) a background eclipsing binary with no visible secondary eclipse at 0.49376 times the orbital period of the foreground binary, (2) a background eclipsing binary with nearly identical primary and secondary eclipses at 0.98752 times the orbital period of the foreground binary, (3) a circumbinary transiting object, or (4) a transiting object around one of the stars in an almost 2:1 resonant orbit with the binary. Follow-up multi-color light curves, spectra, and radial velocities will be needed to fully characterize this interesting system.

6. NEW TRANSITING PLANET CANDIDATES

For the eight new transiting planet candidates mentioned in Section 3, we combined Q0 and Q1 data, and modeled the transit curves using JKTEBOP, accounting for the 29.43 minute integration time, and using our modified AGA in the same manner described in Section 4. We assumed zero eccentricity and negligible flux from each planet, and interpolated the limb-darkening and gravity-darkening coefficients via the effective temperature, surface gravity, and metallicity from the relations of Sing (2010) and Claret (2000). We then solved for the period, time of primary minimum, inclination, sum of the fractional radii, ratio of the radii, and the out of transit flux level. With this narrowed set of parameters, the AGA proved to be extremely quick and precise, and all fits were confirmed by eye and χ^2 values to accurately fit the data. Plots of the transit light curves with model fits are shown in Figure 6.

To estimate the physical radius of each transiting exoplanet candidate, we took the value for the radius of the host star from the KIC and multiplied by the ratio of the radii, k , found from

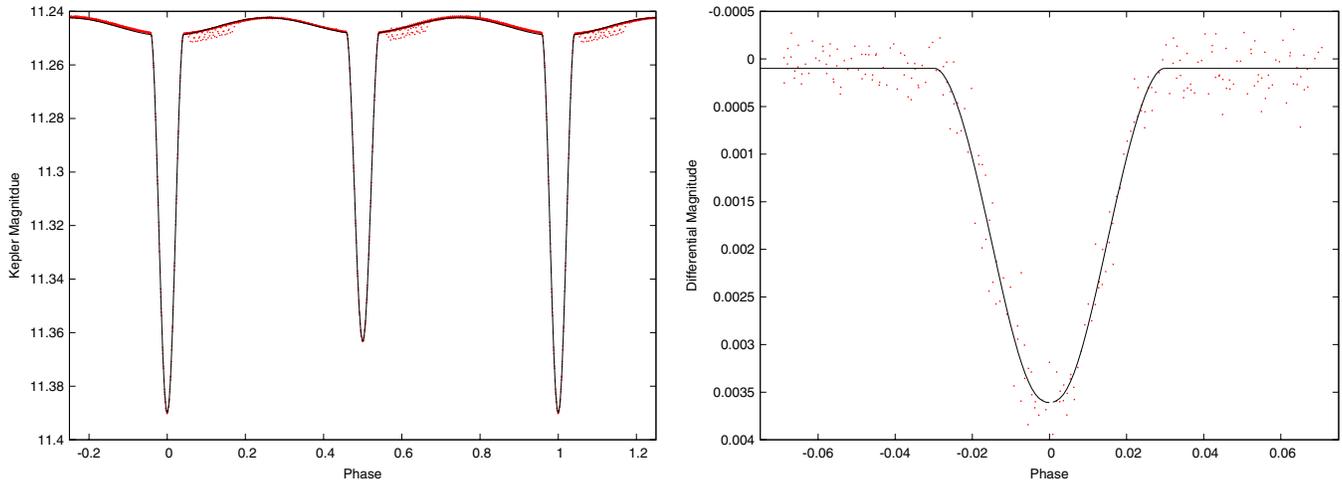


Figure 5. Kepler 004247791, an example of a system which was determined not to be main-sequence in Section 5. Left: the light curve phased at its period of 4.100866 days with our best model fit. Given the shallow, wide eclipses for a ~ 4.1 day period and $T_{\text{eff}} = 4063$ K, if this system were main sequence, it would have a combined mass of $1.28 M_{\odot}$ and a combined radius of $3.82 R_{\odot}$. Thus, this system contains one or more evolved stars. Right: the model-fit subtracted light curve phased at a period of 2.02484 days, showing a transit-like feature imposed on the light curve of the eclipsing binary. Possible explanations may include, but are certainly not limited to a background eclipsing binary with no visible secondary eclipse at 0.49376 times the orbital period of the foreground binary, a background eclipsing binary with nearly identical primary and secondary eclipses at 0.98752 times the orbital period of the foreground binary, a circumbinary transiting object, or a transiting object around one of the stars in an almost 2:1 resonant orbit with the binary.

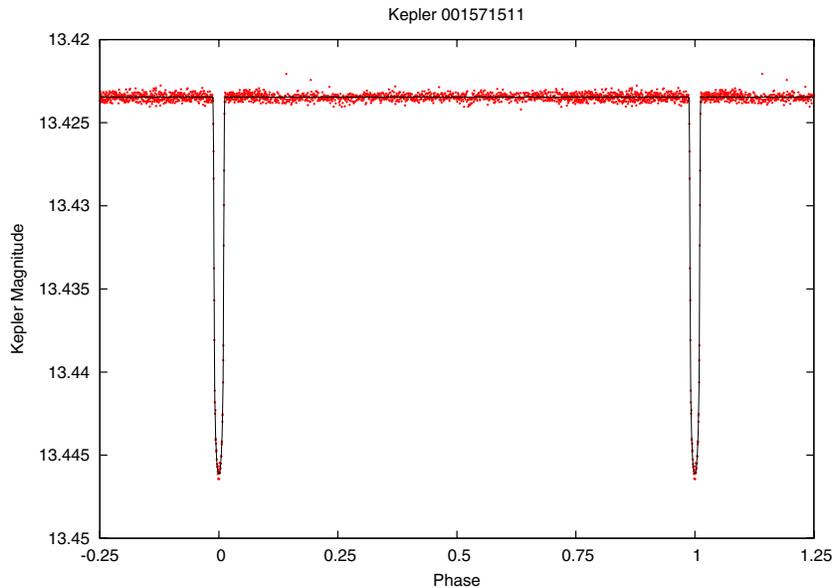


Figure 6. Plots of the light curves of the eight transiting planet candidates modeled with the JKTEBOP code, presented in Table 4. Only the first plot, Figure 6.1, is shown in the text for guidance.

(An extended version of this figure is available in the online journal.)

the model. In Table 4, we list the *Kepler* ID number, apparent *Kepler* magnitude, time of primary minimum, period, effective temperature of the star, inclination, radius of the star, and radius of the exoplanet candidate in both solar radii and Jupiter radii.

As can be seen, the radii for these transiting planet candidates range from 0.56 to $2.1 R_{\text{Jup}}$, with periods between 4.1 and 24.6 days. Only one of these, Kepler 011974540, has been ruled out as a planet from follow-up RV measurements, which are needed for the rest of the candidates to confirm or refute their planetary nature. However, even if these objects turn out not to be planetary mass, they then must be either brown dwarfs or very low mass stars, which still are valuable finds. In the case of brown dwarfs, these targets would be located within the so-called brown dwarf desert (McCarthy & Zuckerman 2004).

7. COMPARISON OF THE NEW LOW-MASS BINARY CANDIDATES TO MODELS

As described in the introduction, one of the current outstanding questions in the study of low-mass stars is whether the inflated radii observed in binaries is caused by their enhanced stellar rotation, and therefore enhanced magnetic activity. We explore this problem in this section using the list of the 95 new LMMS DDEB candidates with estimated individual masses both below $1.0 M_{\odot}$ and light curve amplitudes greater than 0.1 mag, given in Table 3. This sample, for the first time, provides a statistically significant number of systems with orbital periods larger than 10 days.

The left panels of Figure 7 show mass–radius diagrams using the mass and radius of each binary star component estimated in

Table 4
Model Parameters for the Eight Transiting Exoplanet Candidates

<i>Kepler</i> ID	M_{kep}	T_0 (BJD–2454900)	P (days)	$T_{\text{eff},\star}$ (K)	i ($^\circ$)	R_\star (R_\odot)	R_p (R_\odot)	R_p (R_{Jup})
001571511	13.42	68.529019	14.02065	5804	89.28	1.08	0.14	1.43
003342592	14.92	69.190452	17.17864	5717	89.20	0.93	0.14	1.37
005372966	15.37	67.675070	9.286422	5464	88.91	0.92	0.19	1.87
006756669	15.33	65.860125	5.851827	5353	88.34	0.90	0.16	1.59
006805146	13.21	56.568771	13.77974	6214	89.14	1.41	0.21	2.11
008544996	15.20	65.898818	4.081488	5463	87.61	1.00	0.13	1.27
011974540 ^a	13.22	65.862352	24.67058	6507	89.53	0.69	0.06	0.56
012251650	14.76	71.657743	17.76233	4952	88.97	1.00	0.16	1.64

Note. ^a Listed in the *Kepler* False Positive Catalog as “velocity measurements indicate eclipsing binary.”

Section 5. The LMMS DDEB candidates have been separated into three categories, with orbital periods $P < 1.0$ day, $1.0 \text{ day} < P < 10$ days, and $P > 10$ days. Each primary and secondary in a binary pair are traced by a connecting line. We also plot in each panel of Figure 7 the theoretical mass–radius relation predicted by the Baraffe et al. (1998) models for $M \geq 0.075 M_\odot$, and the Chabrier et al. (2000) models for $M < 0.075 M_\odot$, both for $[M/H] = 0.0$, and an age of 5.0 Gyr. We have also defined a main-sequence cutoff as 1.5 times the theoretical mass–radius relation, which is illustrated by the solid line in each diagram. In the models we have used an $\alpha = 1.0$ for $M \leq 0.7 M_\odot$ and interpolated the radius of the models for $0.7 M_\odot < M \leq 1.0 M_\odot$ by fixing the radius of the $1.0 M_\odot$ model to $1.0 R_\odot$, therefore avoiding the dependence of the stellar radius with α between $0.7 M_\odot$ and $1.0 M_\odot$ (Baraffe et al. 1998). We also include in the mass–radius diagrams estimations of the error in our M and R values at several masses, computed by adding and subtracting 200 K (the error in the T_{eff} determinations given by the KIC) from a given temperature and interpolating the mass and radius from the theoretical relations. Note that one of the long-period stars, Kepler 008075618, falls well below the main sequence, with two identical components with $M = 0.91 M_\odot$ and $R = 0.53 R_\odot$. Inspection of this light curve, coupled with the light curve model, reveals that this system could in fact be a single-lined system at half the listed period.

In the figure, many of the stellar radii of binaries with $P < 1.0$ days appear to fall above the model predictions, but as the orbital period increases, a larger fraction of the systems appear to have radii that are either consistent with or fall below the models. There certainly is a fair amount of scatter in these data introduced by the large error in the mass and radius estimations, but a histogram analysis of the radius distributions confirms these apparent trends. In the right panels of Figure 7 we show 5% bin-size histograms representing how many stars have a radius that deviates by a given percentage from the models. The average radius discrepancy is 13.0%, 7.5%, and 2.0% for the short ($P < 1.0$ day), medium ($1.0 \text{ day} < P < 10.0$ days), and long-period ($P > 10.0$ days) systems, respectively. Although a full analysis of each system with multi-color light and RV data is still needed, these preliminary estimates support the hypothesis that binary spin-up is the primary cause of inflated radii in short-period LMMS DDEBs.

8. SUMMARY

We present 231 new double-eclipse, detached eclipsing binary systems with $T_{\text{eff}} < 5500$ K, found in the Cycle 0 data release of the *Kepler Mission*, and provide their *Kepler* ID, estimated effective temperature, *Kepler* magnitude, magnitude range of

the light curve, orbital period, time of primary minimum, inclination, eccentricity, longitude of periastron, sum of the fractional radii, and luminosity ratio. We estimate the masses and radii of the stars in these systems, and find that 95 of them contain two main-sequence stars with both components having $M < 1.0 M_\odot$ and eclipse depths of at least 0.1 mag, and thus are suitable for ground-based follow-up. Of these 95 systems, 14 have periods less than 1.0 day, 52 have periods between 1.0 and 10.0 days, and 29 have periods greater than 10.0 days. This new sample of low-mass, double-eclipse, detached eclipsing binary candidates more than doubles the number of previously known systems and extends the sample into the completely heretofore unexplored $P > 10.0$ day period range for LMMS DDEBs.

Comparison to the theoretical mass–radius relation models for stars below $1.0 M_\odot$ by Baraffe et al. (1998) shows preliminary evidence for better agreement with the models at longer periods, where the rotation rate of the stars is not expected to be spun-up by tidal locking, although, in the absence of RV measurements, the errors on the estimated mass and radius are still quite large. For systems with $P < 1.0$ days, the average radius discrepancy is 13.0%, whereas for $1.0 \text{ day} < P < 10.0$ days and $P > 10.0$ days, the average radius discrepancy is 7.5% and 2.0%, respectively. Ground-based follow-up, in the form of RV and multi-wavelength light curves, is needed to derive the mass and radius of each star in each system to $\sim 1\%$ – 2% , which we have already begun to acquire. With accurate masses and radii for multiple long-period systems, we should be able to definitively test the hypothesis that inflated radii in low-mass binaries are principally due to enhanced rotation rates.

We also present eight new transiting planet candidates. Only one of them is currently listed in the *Kepler* False Positive Catalog. The remaining candidates require RV follow-up to confirm or refute their planetary nature. Even if these systems do not turn out to be planets, they then must be brown dwarf or very low mass, late-type M dwarfs, which would still be a very valuable find. In fact, all false positive planet candidates determined by the *Kepler* team will be of great interest to stellar astrophysics. We also present seven new extremely shallow eclipsing systems, one well-detached binary with deep eclipses, and one apparent red giant with an unusual eclipse-like feature. We also highlight a very unusual eclipsing binary system containing at least one evolved star and an additional transit-like feature in the light curve. Finally, the systems that we determined are not main sequence, and we therefore did not include in the subsequent analysis, should be further studied for valuable science. Accurate mass, radius, and temperature determinations of those systems could yield valuable insights into stellar and binary evolution.

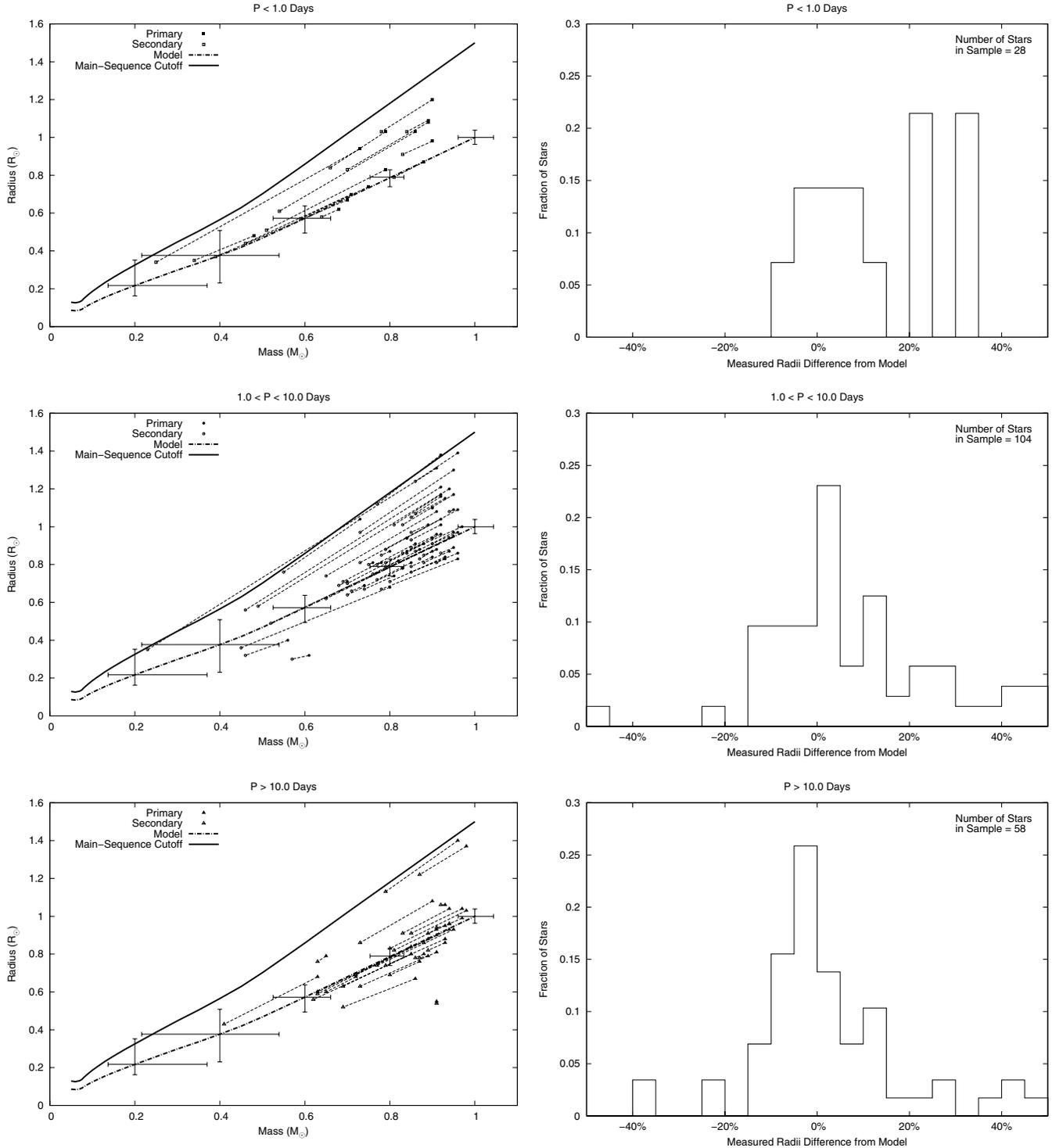


Figure 7. Left: mass–radius diagrams for each binary with both components $< 1.0 M_{\odot}$ and photometric amplitudes greater than 0.1 mag, as given in Table 3, with systems connected by faint lines. The systems are sorted into short-period ($P < 1.0$ days, top panel), medium-period ($1.0 \text{ day} < P < 10.0$ days, middle panel), and long-period groupings ($P > 10.0$ days, bottom panel). The theoretical mass–radius relations of Baraffe et al. (1998) for $0.075 M_{\odot} \leq M \leq 1.0 M_{\odot}$, and of Chabrier et al. (2000) for $M < 0.075 M_{\odot}$, both for $[M/H] = 0.0$ and an age of 5.0 Gyr, are overlaid. The solid line shows the main-sequence cutoff criterion. The error bars indicate the error in mass and radius obtained when interpolating from the mass–temperature–radius relations with an error of 200 K. Right: histograms of the fraction of stars in the sample vs. their deviance from the models for each period grouping. As can be seen by both the mass–radius relation plots and the histograms, shorter period binaries in general appear to exhibit larger radii compared to the models than longer period systems.

We first and foremost thank the entire *Kepler* team and all those who have contributed to the *Kepler Mission*, without which this paper would not be possible. We also thank our second referee for a very thorough and helpful review. J.L.C. acknowledges support from an NSF Graduate Research

Fellowship and the New Mexico Space Grant Consortium. J.L.C., M.L.M., and T.E.H. are grateful for funding from *Kepler's* Guest Observer Program. M.L.M. acknowledges support from NASA through Hubble Fellowship grant HF-01210.01-A/HF-51233.01 awarded by the STScI, which is operated by the

AURA, Inc. for NASA, under contract NAS5-26555. We also thank Robert Edmonds at New Mexico State University for inspiring us to modify the AGA to use a Gaussian distribution to create new individuals. Some/all of the data presented in this paper were obtained from the Multimission Archive at the Space Telescope Science Institute (MAST). STScI is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS5-26555. Support for MAST for non-*HST* data is provided by the NASA Office of Space Science via grant NNX09AF08G and by other grants and contracts. This research has made use of NASA's Astrophysics Data System.

APPENDIX A

ECLIPSE PHASE DISPERSION MINIMIZATION (EPDM)

In this Appendix, we further explain the EPDM technique introduced in Section 3. As mentioned in the text, EPDM finds the period of an eclipsing binary system by seeking the value of the period that best minimizes the dispersion in phase of the faintest N points in a light curve. To illustrate how this method works, we show in Figure 8 the period search analysis of the LMMS DDEB candidate Kepler 006591789, which was found to have a period of 5.088435 days via the JKTEBOP model (see Table 2). The unfolded Q1 light curve is shown in the top left panel of Figure 8. EPDM selects the faintest 20 points of the light curve, which are highlighted by the larger points in that same panel. The number of points should be adjusted based on the quality of the data set. Too few points could result in all the points selected belonging to the same eclipse, if that one eclipse is unusually deep due to systematics or another reason, and thus EPDM will be unable to determine a period. Too many points will cause the results of EPDM to be less precise, as more points are included further away from the center of the eclipses. We have found that 20 points is a good number for *Kepler* data, for which many systems do suffer from moderate systematics, as is evidently the case for Kepler 006591789, as seen by the quasi-sinusoidal variation in the baseline flux.

Having selected the faintest points from the light curve, EPDM then loops over a range of period values. In this case we choose a set of 5000 period values that range from 0.3 to 30 days, evenly distributed in log space, so that shorter periods are as well sampled as longer periods. At each period, the phase of each of the 20 faintest points is calculated via the following standard equation:

$$p = \frac{T}{P} - \text{int}\left(\frac{T}{P}\right), \quad (\text{A1})$$

where p is the phase of a given point, with a time value, T , for a given period, P , and $\text{int}()$ returns the argument rounded down to the nearest integer value. The standard deviation of these 20 phase values is then computed, and we are left with a standard deviation for each trial period. In the bottom left panel of Figure 8, we plot the standard deviation in phase of the 20 points versus each trial period. The lowest values for the standard deviation indicate the best periods, where the eclipses align in phase space, while high values indicate bad periods. As can be seen in the bottom left panel, the standard deviation approaches a value of 0.0 near 10.2 days, 5.1 days, 2.05 days, and decreasing fractions thereof, or period aliases. To determine the three best periods, EPDM first selects the lowest standard deviation, which in this case yields a value of 5.09004 days. It then selects the next lowest value, whose corresponding period

value differs from the first by at least 10%, and yields a value of 10.1747 days. The third period selected via the same method yields a value of 2.54402 days.

To further clarify the technique visually, in the top right panel of Figure 8, we show the same plot as in the bottom left panel, but limited in period range to straddle the best period found, 5.09004 days. At the same period range, in the bottom-right panel, we plot the actual values of the phase for each of the 20 points at each period. (For ease of viewing, we use a lower trial period resolution in the bottom right panel than the top right panel.) As can be seen in the lower-right panel, when the trial period is far from the true period of the system, the actual phase values have a large dispersion, and range completely from 0.0 to 1.0. As the given period gets closer to the true period, the phase values begin to clump, with their dispersion decreasing as the trial period approaches the true value. Indeed, as highlighted by the box in the bottom right panel of Figure 8, at the best period, all the phase values are tightly grouped together at $P = 5.0904$ days, indicating that all the eclipses are extremely well aligned, and the period of the system has been found.

One complication that can arise is if EPDM encounters an eccentric system with two similarly deep eclipses. In this case, when the algorithm selects the N faintest points, it will be selecting points from both eclipses. Since the system is eccentric, there is a phase offset not equal to 0.5 between primary and secondary eclipses, i.e., the two eclipses occur closer to each other in time compared to the period of the system. In this case, if we were to run EPDM as just described, in a plot like the bottom right panel of Figure 8, at the true period of the system there would be two groups of points, each by itself having a very small deviation, but separated from each other in phase by a large amount. Thus, the standard deviation calculation will show a much higher value than it should, and the correct period could not be found. Along similar lines, a problem arises when we consider how to calculate the standard deviation of, for example, the distribution of phase points in the bottom right panel of Figure 8 at a period of 4.9 days, which ranges from 0.8 to 1.0, and then jumps to 0.0–0.05. It is clear this is a continuous group of points, which simply experiences an abrupt jump from phase 1.0 to 0.0. Although they represent a fairly good period, a calculation of their standard deviation would show a high value, and thus indicate a bad period.

To reconcile both these problems, we insert an additional step into the EPDM technique. At each trial period, EPDM searches for a reflection phase, p_r , whose value is between 0.0 and 1.0, that will allow the two distinct phase groupings to align. For each value of p_r , if the phase value of a given point is larger than p_r , a new value for the phase of the point, p , is calculated as

$$p = p - 2.0 \cdot (p - p_r). \quad (\text{A2})$$

The value of p_r which yields the lowest standard deviation for a given trial period is the correct reflection value, and that corresponding lowest standard deviation should be assigned to that trial period. Thus, in the case of an equal depth, eccentric system, where say the N lowest points group around two phases of 0.2 and 0.4, at a value of $p_r = 0.3$, the two distinct groupings would merge into a single group at phase 0.2, with a very small standard deviation at the correct period of the system. As well, in the case where a group of phase points that range from 0.9 to 1.0 and 0.0 to 0.1, p_r allows the points to merge into a single group that only ranges from 0.0 to 0.1. In fact, we have already implemented the use of p_r when generating the bottom left and top right plots of Figure 8.

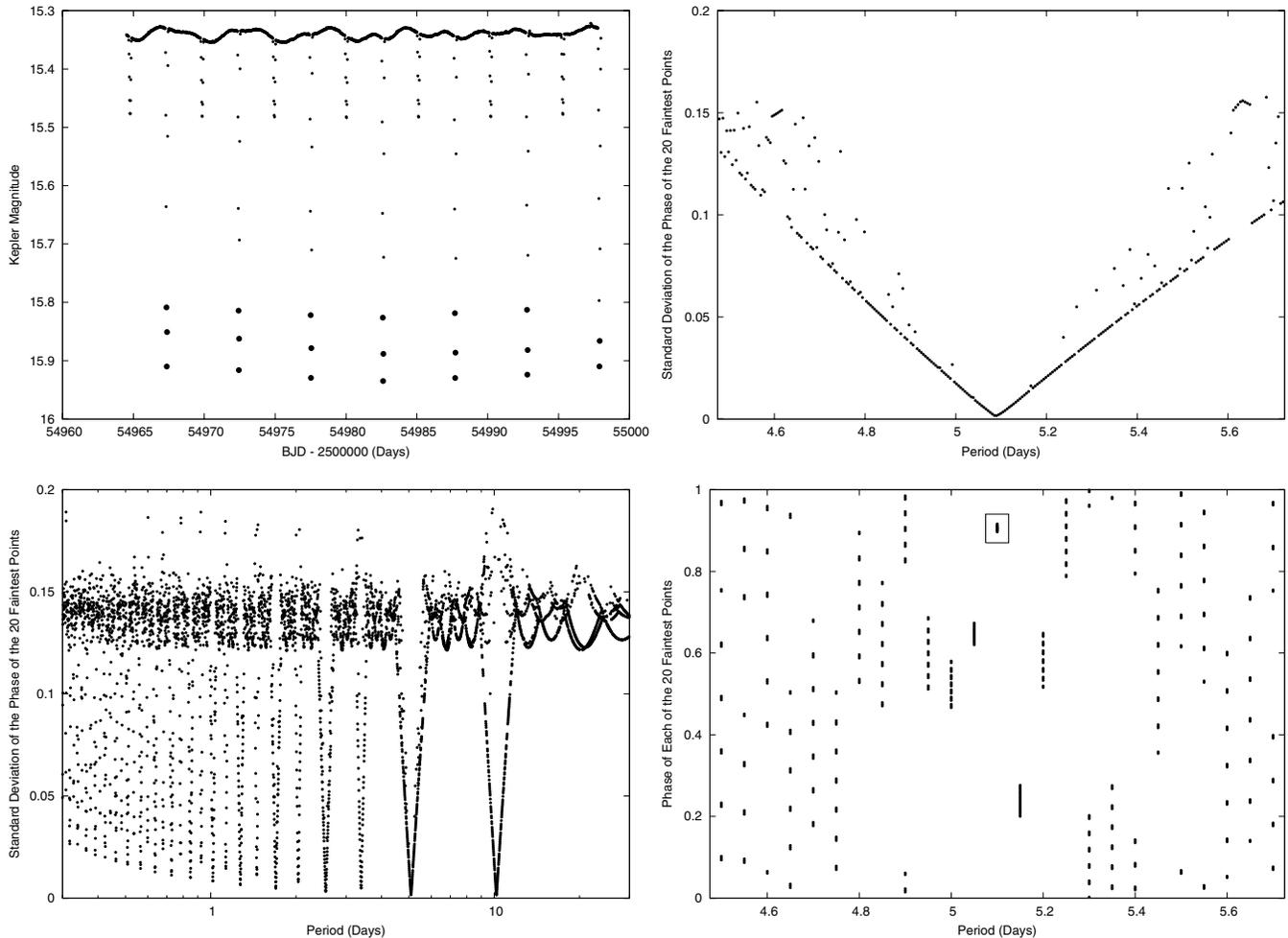


Figure 8. Illustration of the EPDM technique. Top left: the unphased light curve of Kepler 006591789, with the 20 faintest points highlighted by using larger point sizes. Bottom left: standard deviation of the phase values of the 20 faintest points vs. period for this system. As can be seen, the standard deviation approaches 0.0 at ~ 5.1 days, and integer multiples and fractions thereof. Top right: the same plot as in the bottom left panel, but with the period range restricted to show only the period with the lowest standard deviation, and true period of the system. Bottom right: the actual phase values for each of the 20 faintest points at multiple periods, spanning the same period range (but with a lower period resolution, for clarity), as the plot in the top right panel. As can be seen, as the examined period approaches the true period of the system, the phase values of the 20 faintest points strongly clump together, producing a very small standard deviation. The best period is highlighted by a box in the lower-right panel.

In conclusion, because EPDM only utilizes the faintest N points of a light curve, the computations are very quick, especially compared to traditional PDM techniques, which utilize every point in a light curve. This also allows for a more precise determination of the period, as one can apply more computing time toward finer period resolution. As well, for the same reason, EPDM is not affected by systematics or varying star spots, as long as their photometric amplitudes are not on the order of or greater than the amplitude of the eclipses. By selecting the faintest point, or the earliest of the N faintest points, one is also given a good value for the time of primary minimum. We have shown that EPDM can be applied to both eccentric and non-eccentric binaries, and since a transiting planet's light curve is similar to an eclipsing binary with only one visible eclipse, the technique works equally well for transiting exoplanets. In theory, EPDM could also be applied to other variables, such as stars with rotating spots, pulsating variables, and contact binaries, although periods for these systems will be less precise than detached eclipsing binaries, due to the broader minima of those systems. In theory though, one may not have to select the faintest points of a light curve, but possibly a very narrow flux range, and achieve the same result.

APPENDIX B

GENETIC ALGORITHMS FOR ECLIPSING BINARIES

As mentioned in the text, in fitting our sample of eclipsing binaries, we have 12 parameters: period, time of primary minimum, inclination, mass ratio, $e \cdot \cos(\omega)$, $e \cdot \sin(\omega)$, surface brightness ratio, sum of the fractional radii, ratio of the radii, out-of-eclipse flux level, and the amplitude and phase shift of the sinusoid applied to the luminosity of the primary in order to account for spots. We aim to vary these parameters over their entire range of possible solutions, which if left to a grid search for 10^{-3} precision, would require computing on the order of $\sim 10^{36}$ light curves; a computationally prohibitive task. Standard steepest descent minimization schemes such as Levenberg–Marquardt have extreme difficulties in large, multi-parameter solution spaces, especially for eclipsing binaries as the solution space is not at all smooth and has many local minima. Thus, we need a minimization technique that is computationally efficient, not adversely affected by a non-smooth solution space, and able to find the global minimum. These criteria are superbly met by the class of optimization schemes known as GAs.

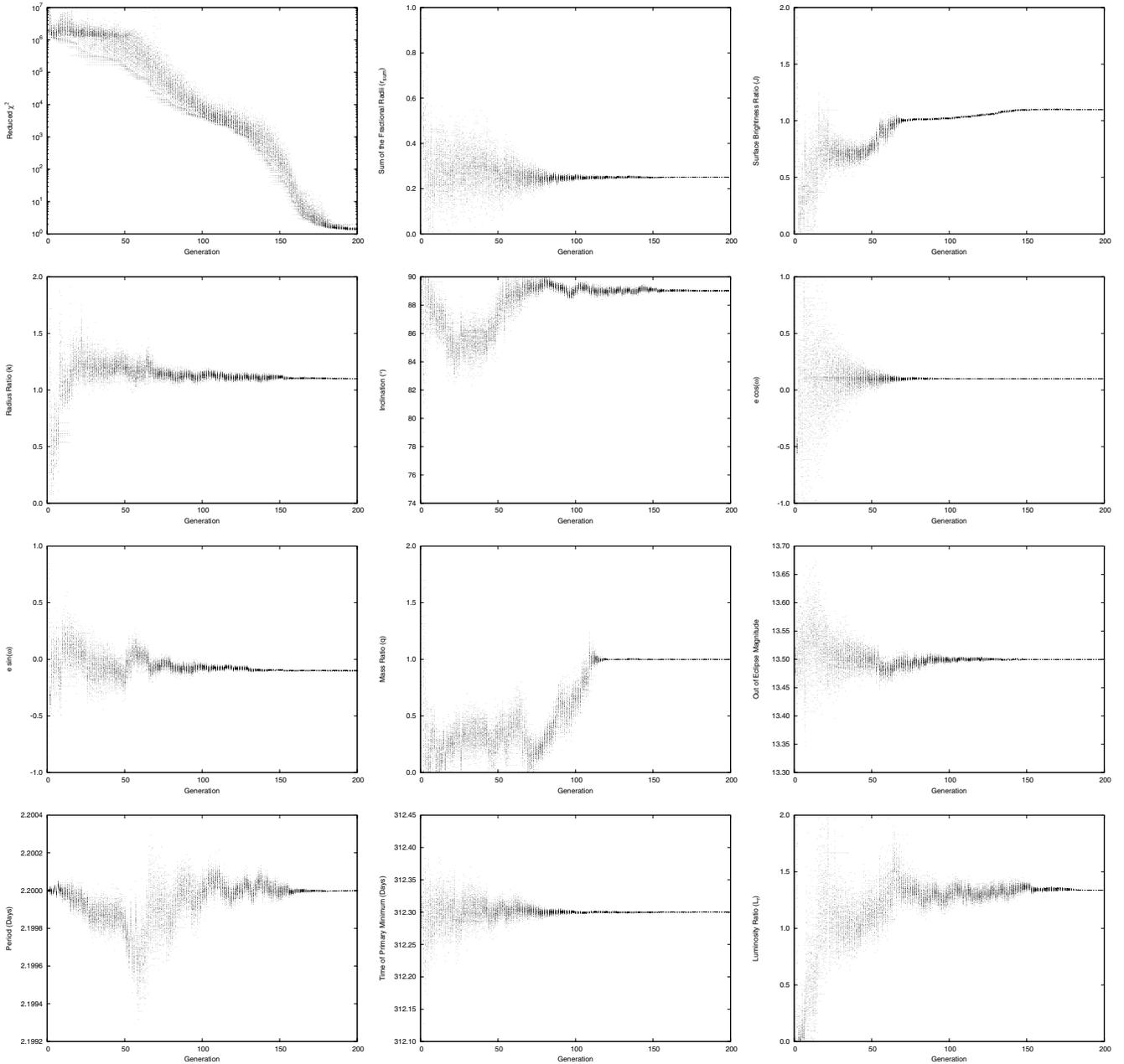


Figure 9. Illustration of how the AGA converges over subsequent generations by solving an artificially generated light curve, re-binned to the number of data points and error typical for a *Kepler* light curve. The parameters of the system are $r_{\text{sum}} = 0.25$, $k = 1.1$, $i = 89^\circ$, $q = 1.2$, $e \cdot \cos(\omega) = 0.1$, $e \cdot \sin(\omega) = -0.1$, $J = 1.1$, $P = 2.20$ days, $T_0 = 312.3$ days, and out-of-eclipse magnitude = 13.5. The derived reduced χ^2 and luminosity ratio are also plotted. The AGA converges rapidly, decreasing the lowest χ^2 value found by an order of magnitude every ~ 20 generations. It can be seen that the parameters that are most significant to the light curve converge the fastest.

In a standard GA (cf. Charbonneau 1995), light curve parameter sets, called individuals, for an initial population of solutions, are randomly generated within a predefined parameter space and compared to the observational light curve. Their corresponding χ^2 value is used as a measure of fitness for natural selection, with parameters from fit individuals bred with each other (subjected to crossover like chromosomes) to create a second generation of new solutions, and parameters from unfit individuals eliminated. After being subject to random mutations, to maintain parameter diversity and ensure discovery of the global minimum, this second generation is compared to the observational data and bred into a third generation of solutions. The process continues for a specified number of generations, until a satisfactorily low

χ^2 is found. Charbonneau (1995) demonstrated the application of GAs to problems in Astronomy and Astrophysics, specifically fitting galactic rotation curves, finding pulsation periods in δ Scuti stars, and fitting magnetodynamical wind models with multiple critical points, showing how the GA quickly finds the global minimum, regardless of the topography of the solution space. It is this type of GA that has been already been incorporated into the ELC eclipsing binary modeling code and used with much success (Orosz & Hauschildt 2000; Orosz et al. 2002).

Cantó et al. (2009) recently proposed a new form of GA called an AGA. In the AGA, instead of breeding new individuals via crossover, individuals are randomly created within a small predefined parameter space, or breeding box, centered on the fittest

members of the previous generation. The size of this breeding box can be shrunk over successive generations to quickly converge to the best-fit solution. As shown by Cantó et al. (2009), the AGA is computationally simpler and more precise since it does not require encoding parameters for crossover, and converges much faster than traditional GAs, without sacrificing any ability to migrate to the global solution, so long as the breeding box size does not decrease too quickly. Cantó et al. (2009) first showed that it far outperformed the standard GA in both computational efficiency and final precision by solving one of the exact same problems presented by Charbonneau (1995). Cantó et al. (2009) additionally demonstrated the application of the AGA to fitting the radial velocities of extrasolar planets and the spectral energy distributions of young stellar objects.

As eclipsing binary solutions have an even larger parameter space with many local minima than most problems, we make a few modifications to the AGA described by Cantó et al. (2009) to ensure discovery of the global minimum. First, while we do exactly copy the fittest 10% of individuals of one generation to the next generation, to ensure forward progress is always made while maintaining parameter diversity, instead of picking the fittest N members of a generation, each of which breeds M offspring, to create a new generation, we randomly select individuals for breeding by weighting them by a factor of $(1/\chi^2)^2$. This ensures that the fittest individuals breed the most offspring, but still allows for a few less fit individuals to breed, maintaining parameter diversity and exploration of the entire parameter space. Second, instead of randomly creating new members within a breeding box of fit individuals, we randomly select a number for each parameter from a Gaussian probability distribution centered on each parameter of a fit individual. Thus, new individuals are not strictly confined to a breeding box, but merely are very likely to be created near a fit individual, and maintain a very small probability that they will be created at many standard deviations away. This mimics mutation in traditional GAs and ensures that the algorithm will not become trapped in a local minimum. Third, as suggested by Cantó et al. (2009), the standard deviation of this normal distribution is chosen for each parameter to be the standard deviation of that parameter in the entire population, times the function $0.1^{(1/\chi_0^2)}$, where χ_0^2 is the χ^2 value of the fittest member of the population. This allows parameters with the greatest impact on the fit, or the smallest range of possible parameters, such as the out-of-eclipse flux level, to converge rapidly, while allowing parameters that are less certain to converge more slowly and thoroughly explore their parameter space. Furthermore, via this method, the standard deviation is shrunk over successive generations, so that the algorithm converges, but only very slowly initially, rapidly increasing as χ_0^2 approaches 1.0, i.e., the global minimum has been found. Finally, we take the fittest 10% of the final generation and perform a standard Levenberg–Marquardt minimization for each member, choosing the member with the resulting lowest χ^2 value as our final solution.

We nominally found, for the eclipsing binaries in our sample, that a population of 100 individuals, bred for 200 generations, does an excellent job of solving the light curves. This only requires the generation of 20,000 light curves, which with the JKTEBOP code only required ~ 3 minutes per light curve to solve on a single 2.0 GHz CPU. Of course, some systems may require a smaller or greater number of individuals and/or generations, but it should not be more than a factor of ~ 2 . One may substantially reduce the number of individuals or

generations required, and thus the run time, if one can limit the range of parameter space. For example, if one knows, or wants to assume, the orbit is circular or nearly circular, one could constrain $|e \cdot \cos(\omega)| < 0.1$ and $|e \cdot \sin(\omega)| < 0.1$. Furthermore, the AGA code is extremely parallelizable, and thus with a multi-core computing cluster one could easily use this technique to model thousands of eclipsing binary light curves, as is to be expected from Pan-STARRS and other large photometric surveys, in a very reasonable time frame.

To visually demonstrate how the AGA works, we have generated a light curve with the following parameters: $r_{\text{sum}} = 0.25$, $k = 1.1$, $i = 89^\circ$, $q = 1.2$, $e \cdot \cos(\omega) = 0.1$, $e \cdot \sin(\omega) = -0.1$, $J = 1.1$, $P = 2.20$ days, $T_0 = 312.3$ days, and out-of-eclipse magnitude = 13.5. We then re-bin these data to match the number of data points in the Kepler Q1 data sets, and add typical Gaussian noise for a bright Kepler star of 0.1 mmag per data point. We then resolve this light curve with the AGA, varying all the aforementioned parameters, and show in Figure 9 the value of each parameter for every individual in each generation, as well as the values for the derived reduced χ^2 and luminosity ratio. One can see how even while searching over the entire global solution space, the AGA rapidly converges to the solution that was used to generate the light curve, with the χ^2 decreasing by a factor of ~ 10 every ~ 20 generations. Even though the best solution of the 200th generation has $\chi^2 \sim 1.5$, if allowed to continue for more generations, this run would eventually converge to $\chi^2 = 1.0$, and performing a simple Levenberg–Marquardt minimization from the best solution quickly produces a $\chi^2 = 1.0$ fit.

REFERENCES

- Andersen, J. 1991, *A&AR*, **3**, 91
 Baraffe, I., Chabrier, G., Allard, F., & Hauschildt, P. H. 1998, *A&A*, **337**, 403
 Becker, A. C., et al. 2008, *MNRAS*, **386**, 416
 Blake, C. H., Torres, G., Bloom, J. S., & Gaudi, B. S. 2008, *ApJ*, **684**, 635
 Borucki, W. J., & the Kepler team. 2010, arXiv:1006.2799
 Boyajian, T. S., et al. 2010, arXiv:1012.0542
 Cantó, J., Curiel, S., & Martínez-Gómez, E. 2009, *A&A*, **501**, 1259
 Chabrier, G., Baraffe, I., Allard, F., & Hauschildt, P. 2000, *ApJ*, **542**, 464
 Chabrier, G., Gallardo, J., & Baraffe, I. 2007, *A&A*, **472**, L17
 Charbonneau, P. 1995, *ApJS*, **101**, 309
 Claret, A. 2000, *A&A*, **359**, 289
 Claret, A., & Gimenez, A. 1990, *A&A*, **230**, 412
 Clausen, J. V., Bruntt, H., Claret, A., Larsen, A., Andersen, J., Nordström, B., & Giménez, A. 2009, *A&A*, **502**, 253
 Coughlin, J. L., & Shaw, J. S. 2007, *J. Southeastern Assoc. Res. Astron.*, **1**, 7
 Delfosse, X., Forveille, T., Mayor, M., Burnet, M., & Perrier, C. 1999, *A&A*, **341**, L63
 Demory, B., et al. 2009, *A&A*, **505**, 205
 Devor, J., Charbonneau, D., O'Donovan, F. T., Mandushev, G., & Torres, G. 2008a, *AJ*, **135**, 850
 Devor, J., et al. 2008b, *ApJ*, **687**, 1253
 Dimitrov, D. P., & Kjurkchieva, D. P. 2010, *MNRAS*, **406**, 2559
 Fernandez, J. M., et al. 2009, *ApJ*, **701**, 764
 Hoffman, D. I., Harrison, T. E., Coughlin, J. L., McNamara, B. J., Holtzman, J. A., Taylor, G. E., & Vestrand, W. T. 2008, *AJ*, **136**, 1067
 Irwin, J., et al. 2009, *ApJ*, **701**, 1436
 Kinemuchi, K., Smith, H. A., Woźniak, P. R., & McKay, T. A. 2006, *AJ*, **132**, 1202
 Kipping, D. M. 2010, *MNRAS*, **408**, 1758
 Lacy, C. H. 1977, *ApJ*, **218**, 444
 Leung, K., & Schneider, D. P. 1978, *AJ*, **83**, 618
 López-Morales, M. 2007, *ApJ*, **660**, 732
 López-Morales, M., & Ribas, I. 2005, *ApJ*, **631**, 1120
 Manduca, A., Bell, R. A., & Gustafsson, B. 1977, *A&A*, **61**, 809
 McCarthy, C., & Zuckerman, B. 2004, *AJ*, **127**, 2871
 Morales, J. C., Gallardo, J., Ribas, I., Jordi, C., Baraffe, I., & Chabrier, G. 2010, *ApJ*, **718**, 502

- Morales, J. C., Ribas, I., & Jordi, C. 2008, *A&A*, **478**, 507
- Orosz, J. A., & Hauschildt, P. H. 2000, *A&A*, **364**, 265
- Orosz, J. A., et al. 2002, *ApJ*, **568**, 845
- Popper, D. M. 1980, *ARA&A*, **18**, 115
- Prša, A., Guinan, E. F., Devinney, E. J., DeGeorge, M., Bradstreet, D. H., Giammarco, J. M., Alcock, C. R., & Engle, S. G. 2008, *ApJ*, **687**, 542
- Prša, A., et al. 2010, arXiv:1006.2815
- Rowe, J. F., et al. 2010, *ApJ*, **713**, L150
- Shaw, J. S., & López-Morales, M. 2007, in ASP Conf. Ser. 362, The Seventh Pacific Rim Conference on Stellar Astrophysics, ed. Y. W. Kang, H.-W. Lee, K.-C. Leung, & K.-S. Cheng (San Francisco, CA: ASP), 15
- Shkolnik, E. L., Hebb, L., Liu, M. C., Reid, I. N., & Cameron, A. C. 2010, *ApJ*, **716**, 1522
- Shkolnik, E. L., Liu, M. C., Reid, I. N., Hebb, L., Cameron, A. C., Torres, C. A., & Wilson, D. M. 2008, *ApJ*, **682**, 1248
- Sing, D. K. 2010, *A&A*, **510**, A21
- Southworth, J., Maxted, P. F. L., & Smalley, B. 2004a, *MNRAS*, **351**, 1277
- Southworth, J., Zucker, S., Maxted, P. F. L., & Smalley, B. 2004b, *MNRAS*, **355**, 986
- Stellingwerf, R. F. 1978, *ApJ*, **224**, 953
- Stellingwerf, R. F. 2006, PDM2 Application, Technical Manual, and Test Data Sets
- Torres, G., Andersen, J., & Giménez, A. 2010, *A&AR*, **18**, 67
- Torres, G., Lacy, C. H., Marschall, L. A., Sheets, H. A., & Mader, J. A. 2006, *ApJ*, **640**, 1018
- Van Cleve, J. 2010, Kepler Data Release 5 Notes (Moffett Field, CA: NASA Ames Research Center)
- Wade, R. A., & Rucinski, S. M. 1985, *A&AS*, **60**, 471
- Wilson, R. E. 1993, in ASP Conf. Ser. 38, New Frontiers in Binary Star Research, ed. K.-C. Leung & I.-S. Nha (San Francisco, CA: ASP), 91
- Wilson, R. E., & Devinney, E. J. 1971, *ApJ*, **166**, 605
- Zahn, J. 1977, *A&A*, **57**, 383
- Zahn, J. 1994, *A&A*, **288**, 829