

The *Kepler Mission* and Eclipsing Binaries

David Koch¹, William Borucki¹, Gibor Basri², Timothy Brown³,
Douglas Caldwell⁴, Jorgen Christensen-Dalsgaard⁵, William
Cochran⁶, Edna DeVore⁴, Edward Dunham⁷, Thomas N. Gautier⁸,
John Geary⁹, Ronald Gilliland¹⁰, Alan Gould¹¹, Jon Jenkins⁴, Yoji
Kondo¹², David Latham⁹, Jack Lissauer¹, and David Monet¹³

¹NASA Ames Research Center, Moffett Field, CA, USA

²University of California-Berkeley, Berkeley, CA, USA

³Las Cumbres Observatory Global Telescope, Golenta, CA USA

⁴SETI Institute, Mountain View, CA, USA

⁵Aarhus University, Denmark

⁶University of Texas at Austin, Austin, TX, USA

⁷Lowell Observatory, Flagstaff, AZ, USA

⁸Jet Propulsion Laboratory, Pasadena, CA, USA

⁹Smithsonian Astrophysical Observatory, Cambridge, MA, USA

¹⁰Space Telescope Science Institute, Baltimore, MD, USA

¹¹Lawrence Hall of Science, UC-Berkeley, Berkeley, CA, USA

¹²NASA Goddard Space Flight Center, Greenbelt, MD, USA

¹³United States Naval Observatory, Flagstaff, AZ, USA

Abstract. The *Kepler Mission* is a space-based photometric mission with a differential photometric precision of 14 ppm (at $V = 12$ for a 6.5 hour transit). It is designed to continuously observe a single field of view (FOV) of greater than 100 square degrees in the Cygnus-Lyra region for four or more years. The primary goal of the mission is to monitor more than one-hundred thousand stars for transits of Earth-size and smaller planets in the habitable zone of solar-like stars. In the process, many eclipsing binaries (EB) will also be detected and light curves produced. To enhance and optimize the mission results, the stellar characteristics for all the stars in the *Kepler* FOV with $V < 16$ will have been determined prior to launch. As part of the verification process, stars with transit candidates will have radial-velocity follow-up observations performed to determine the component masses and thereby separate eclipses caused by stellar companions from transits caused by planets. The result will be a rich database on EBs. The community will have access to the archive for further analysis, such as, for EB modeling of the high-precision light curves. A guest observer program is also planned to allow for photometric observations of objects not on the target list but within the FOV.

Keywords. Planet detection, exoplanets, differential photometry, eclipsing binaries

1. Introduction

The *Kepler Mission* is NASA's first mission capable of detecting Earth-size and smaller planets in orbit around solar-like stars. It is a photometric space-based mission designed specifically to search for habitable planets orbiting in or near the habitable zone (HZ) of solar-like stars by detecting sequences of planetary transits, a method first described by Borucki & Summers (1984). A habitable planet is taken to be from 0.8 to $2.2 R_{\oplus}$ or, if one assumes an Earth-like density, from about 0.5 to $10 M_{\oplus}$. For planets less than about $0.5 M_{\oplus}$ the surface gravity is too small to retain a life sustaining atmosphere. If M_{\oplus} is greater than about 10 , then there is sufficient gravity for a planet to hold onto the lightest and most abundant gases, H and He, and become a gas-giant. The HZ is taken to be the range of distances from a star where liquid water can exist on the surface of a

planet (Kasting *et al.* 1993). Specifically the *Kepler Mission* is designed to be capable of detecting a 1.0 Earth-size planet at 1 AU from a solar-like (G2V) star in four years (four transits) with a single transit signal-to-noise ratio of four for a grazing transit of just 6.5 hours duration. An Earth-size transit of a solar-like star produces a relative change in brightness of 84 parts per million and lasts for 13 hours when it crosses the center of the star. Given this detection sensitivity, *Kepler* has a broad detection capability from a Mars-size planet in the HZ of a $V = 9$ G8-dwarf or $V = 13$ M2-dwarf to a two-Earth radii planet in the HZ of a $V = 13$ F0-dwarf to planets significantly smaller than the Earth for orbital periods on the order of a few days. The latter will produce hundreds of transits during the four year mission lifetime.

The mission will produce an unprecedented photometric database in terms of the continuity, duration, photometric precision and number of stars observed. This unique database should provide a wealth of information for other astrophysical uses. Examples are given below.

2. Mission Overview

The *Kepler Mission* is based upon a classical Schmidt telescope design with a 95-cm aperture and more than one-hundred square degree field of view (FOV). The FOV is equivalent to about six Palomar Schmidt plates. The FOV is located in the Cygnus-Lyra region centered on RA = 19^h 22^m 40^s, Declination = 44° 30', just above the galactic plane and looking down the Orion arm of the Galaxy. This provides a rich star field that is continuously viewable throughout the year as the spacecraft orbits the Sun and a sample of stars similar to our local neighborhood. The typical stellar distances for most of the usable stars are from a few hundred parsecs to about 1 kpc. Shot noise limits the distance of usable stars. The apparent magnitude range is from about $V = 9$ to 15.

The FOV is oriented such that all but twelve stars brighter than $V = 6$ are placed in the gaps between the CCDs. The focal plane consists of 42 CCDs, which are 2200 columns by 1024 rows of 27 micron pixels, thinned-back illuminated and anti-reflection coated. Each CCD has two output amplifiers and a pair of CCDs is mounted together to form a square module. The modules are mounted in a 5×5 array with each of the four corner modules having instead a 512×512 pixel CCD. These are used as fine guidance sensors.

The single FOV will be viewed for the entire mission. However, every three months the photometer-spacecraft must be rotated 90° about the optical axis to keep the fixed-solar array pointed to the Sun and the focal-plane radiator pointed to deep space. The positions and orientations of the CCD modules have been chosen so that the focal plane is four-fold symmetric. Thus after a 90° rotation all of the selected target stars remain on active pixels. The photometer-spacecraft will be launched into an Earth-trailing heliocentric orbit, similar to Spitzer. It is expected to drift away from the Earth at the rate of about 0.1 AU per year. The mission is scheduled for launch on a Delta II in November 2008.

The mission has been described in a number of papers (Borucki *et al.* 2005, Koch *et al.* 2004). The basic photometric capability and mission design have remained nearly the same since selection by NASA in December 2001 as the tenth Discovery mission. The detailed design has matured, and the three fundamental design requirements have not been descoped:

1. Detection of an Earth-size transit at 4σ for a $V = 12$ solar-like star in 6.5 hours;
2. The capacity to observe 170,000 stars for at least the first year and 100,000 stars at the end of the mission; and
3. A mission duration of at least four years.

A recent significant change that was made to the mission design was to go from an articulated high-gain antenna (HGA), which allowed for downlinking of the science data every four days without interrupting the observing and having a fifteen-minute integration time, to a body-mounted HGA, which requires a one-day interruption in observing every 31 days for data downlink and required the integration time to double to thirty minutes in order to halve the amount of downlink data.

3. Photometric Characteristics

To understand both how planet detection is performed and how the data might be used for other astrophysical purposes we describe here the basic characteristics of the *Kepler* photometer and the data processing.

We describe the photometric precision of *Kepler* in terms of all the possible sources of noise that have to be taken into account when looking for transits in the light curves. We refer to this as the Combined Differential Photometric Precision (CDPP) and we use the units of parts per million (ppm). The CDPP includes the photon shot noise, variability of the source and the measurement noise. The latter includes not only the CCD read noise and electronics noise, but also the shot noise from the sky due to background stars, zodiacal light, stray light in the photometer, ghosting in the optics, etc. The design requirement is a CDPP ≤ 20 ppm for a $V = 12$ G2V star and 6.5-hour integration. For design purposes we have assumed stellar variability, which is red noise, to be 10 ppm for a solar-like star for a 6.5-hour integration. For a 6.5 hour integration, a $V = 12$ G2V star produces 5×10^9 electrons in the *Kepler* photometer, yielding a shot noise of 14 ppm. Thus the measurement noise by design has to be ≤ 10 ppm at $V = 12$ for the CDPP to meet the requirement. Design and testing indicate that this will be achieved. The dynamical range of the photometer for planet detection is roughly from $V = 9$ to $V = 15$. This depends on the stellar type and apparent brightness, the planetary size and orbit and the mission duration. At $V = 9$ the required CDPP ≤ 12 ppm and at $V = 14$ the required CDPP ≤ 41 ppm for a 6.5-hour integration.

Unlike most conventional telescopes, the *Kepler* photometer has a rather broad point spread function (psf) of about six arcsec and large pixels. One pixel is $3''.98$. This is necessary due to the very large FOV, the need for large well depth to accommodate the shot noise requirement and the desire to minimize the photometric sensitivity to intrapixel quantum efficiency variations.

The system has a capacity at the beginning of the mission for 170,000 stars. Sometime prior to the fourth year this needs to be cut back to 100,000 stars, simply being limited by the telecom system usage. All of these stars will have a sampling resolution of thirty minutes. A subset of 512 stars will have a sampling resolution of one minute. Initially this will be used to measure p-mode oscillations in the brighter stars. Once transits are detected, some portion of these short-cadence targets will be used to observe high signal-to-noise ratio (SNR) transits.

To minimize the shot noise, the photometer has a single broad bandpass from 430 to 890 nm FWHM. The short wavelength cutoff was selected to avoid the Ca II H&K lines. For the Sun, 60% of the irradiance variation is at wavelengths less than 400 nm, but only accounts for about 12% of the total flux (Krivova *et al.* 2006). Hence the 430 nm cutoff helps to improve the system SNR. The red cutoff was chosen to avoid fringing within the CCDs and to minimize the potential for false positives produced by faint reddened background EB.

4. Target Selection

The primary goal of *Kepler* is to detect terrestrial planets around solar-like stars, that is, late-dwarf (F, G, K and M) stars. The magnitude range is roughly $V = 9$ to 15. Since there did not exist a catalog to this depth near the galactic plane with the necessary information on which to base the selection, the project undertook the process of creating the *Kepler Input Catalog* (KIC) led by a team from SAO. The KIC is based on new multi-band photometric observations using the same $g - r - i - z$ filter sequence as the Sloan survey (Abazajian *et al.* 2003) plus an additional filter for the Mg b lines at 516.7, 517.3 and 518.4 nm. The net result will be a catalog with classification information for the two million stars in the *Kepler* FOV to $K < 14$. Using model fitting and a newly expanded Kurucz library of spectra (Castelli & Kurucz 2003), the catalog will contain the effective temperature, $\log(g)$, metallicity [Fe/H], reddening-extinction, mass and radius. The catalog will be federated at the USNO Flagstaff Station with other catalogs, including 2MASS and USNO-B for cross reference. Each star will be ranked for its potential for terrestrial planet detection using the KIC and a merit function to determine the minimum detectable planet size in or near the HZ of each star. This process will be reapplied post-launch incorporating the measured CDPP for each star to re-rank the stars on the target list as part of a necessary on-going down-selection process.

5. Data Processing

On-board the integration time for the CCDs may be set from 2.5 to 8.0 sec. The longer integration time helps to improve the CDPP for the fainter objects, but also results in saturation of more stars. With an integration time of 5 sec, stars brighter than about $V = 12$ saturate. However, we have demonstrated in laboratory testing that precision photometry can still be achieved with bloomed pixels provided the rail voltages are properly set on the CCD and that the full scale on the analog-to-digital converter (ADC) is greater than full well. The CCDs are read out in half a second without a shutter. Individual integrations are co-added on-board for thirty minutes, although for a subset of 512 target stars, one-minute co-additions are preserved. Once a thirty-minute co-add is accumulated in computer memory for all the 95 megapixels in the focal plane, the pixels of interest for the target stars are read out, about 3% of the pixels. There are additional collateral pixels from over-clocking and from masked regions. These are used for removing the bias, smear and determining the dark level. The values are re-quantized to account for the larger value of noise on the high end of the scale. The data are then Huffman encoded (compressed) and stored on the solid-state recorder for later transmission to the ground.

Smear is a result of clocking out the data without a shutter. Every pixel in a column passes under every piece of sky in a column during a read out. This produces a column unique constant offset. And it also produces an optical fat-zero that helps to keep the traps full.

On the ground, the raw data are unpacked, decompressed and archived. The bias and smear are removed. Cosmic ray hits are removed and the background flux is estimated and removed. Then two parallel paths are followed to obtain stellar flux time series, one using aperture photometry and the other using difference image analysis. Ensemble normalization is applied to the raw flux time series to remove common-mode noise at each cadence in time. These data are then used to produce de-trended relative flux time series for each target. These light curves are then archived and will be made available for others to use, for example for modeling eclipsing binary systems. To each time series

a wavelet transform is applied and conditioned with a whitening matched filter. The time series are then folded modulo all possible periods and searched for a multiple-event-detection statistic above a threshold of seven sigma for planetary transits, yielding an 84% detection rate for an eight sigma folded transit signal. A similar process using a Fourier transform rather than a wavelet transform is used to conduct the reflected light search for short period non-transiting giant planets.

6. Eclipsing Binary Detection

From the processing described above, we expect to detect all the eclipsing binaries from within our target set with periods shorter than the mission duration as well as a significant fraction of those with even longer periods that exhibit only a single eclipse. Note that a single grazing Earth-like transit of a $V = 12$ solar-like star has an $\text{SNR} = 4$ and a gas giant like Jupiter produces a 1% transit depth which for *Kepler* will have an $\text{SNR} \approx 400$. So even a single grazing eclipse of a long period binary will have a recognizable character in the data with a significant SNR. Recall that an eclipsing binary signature of a V-shape signal will be significantly different from the U-shape character of a planetary transit. The eclipsing binary events will be cataloged for the astronomical community to further analyze as part of the Data Analysis Program (see below), perhaps using automated processes such as eBAI (Guinan *et al.*, 2007).

For a typical terrestrial planetary transit, the SNR will be too low to distinguish the event from a grazing eclipsing binary (a transiting white dwarf will actually cause a brightening (Sahu & Gilliland 2003)). Stellar companions will induce radial velocity variations that are typically $\gg 1 \text{ km s}^{-1}$. The *Kepler* project will conduct moderate resolution spectroscopy to identify these situations using, for example, the SAO Digital Speedometer (Latham 1992). The net result is that the *Kepler Mission* will produce high precision light curves for more than a thousand eclipsing binaries that are expected to be detected by the mission and available for use by the astrophysical community.

What about non-eclipsing binaries that are on the *Kepler* target list? Many of these can also be identified in a rather circuitous fashion. The most accurate way (to a few percent) to obtain the stellar size is to use the distance to the star, the measured flux, the extinction-reddening, the effective temperature from the KIC and to apply the Stefan-Boltzmann law to derive the stellar area. If the area appears to be something like twice what one would expect for a dwarf with the given effective temperature, based on the modeling that went into deriving $\log(g)$ in the KIC, then the target must be either a binary or perhaps a chance optical double star. Spectroscopic observations will in most cases lead to recognition of its true nature. With regard to the distance, we are expecting that this can be derived from astrometric parallaxes out to 1 kpc using the *Kepler* data. Based upon laboratory measurements (Koch, *et al.* 2000), simulations and the high SNR of the data, this does appear to be plausible. The GAIA mission (Niarchos, Munari and Zwitter 2006) results will also be very useful in refining and confirming some of these cases. Getting the stellar areas correct, especially eliminating a factor of two ambiguity, is necessary for correctly determining the planetary areas.

7. Additional Astrophysical Uses for the Data

Given the uniqueness in precision, completeness, duration and number of stars in the archived data there are potentially a host of other astrophysical uses for the *Kepler* data. These include such things as; measuring p-mode oscillations (Brown & Gilliland 1994), which yield the density and age of stars; analysis of stellar activity, which can yield star

spot cycles (especially if the mission is extended beyond about half of a solar cycle) and white light flaring; the frequency of Maunder minimums for solar-like stars, which has implications for paleoclimatology and perhaps the future of our Earth's climate; stellar rotation rates; cataclysmic variables, providing pre-outburst activity and mass transfer rates; and active galactic nuclei, providing a measure of the "engine" size in BL Lacs, quasars and blazars. Performing p-mode observations of all of the usable stars in the FOV (about 5000 are bright enough) has already been planned as part of the *Kepler* science team effort.

8. Community Participation and Data Access

The community can participate in the *Kepler* results in several ways: a guest observer (GO) program, a data analysis program (DAP) and a participating scientist program (PSP), all of which will be competed by NASA Headquarters through the annual Research Opportunities in Space and Earth Sciences (ROSES). Although the program is open to scientists worldwide, only United States proposals may receive funding.

8.1. Guest Observer Program

Within the GO program, scientists may propose to view objects within the *Kepler* FOV which are not already on the planet detection target list, whether galactic or extragalactic. These objects may be intrinsically variable stars, such as, pulsating (Cepheids, RR Lyrae, Mira, etc.), rotating (ellipsoidal, etc.), eruptive (T Tauri, Wolf-Rayet, etc.) and explosive (novae, super-novae, cataclysmic variables). In general one may assume that all F-, G-, and K-dwarf stars to $V = 14\text{--}15$ and M-dwarf stars to $K = 14.5$ are already on the list and any other object is not. Proposed objects will typically be observed for a minimum of three months and to as long as the mission duration based on justification. We will try to schedule the observations to coincide with any other coordinated observing. Capacity for three thousand objects has been set aside for this program. These will be at thirty-minute cadence except that twenty-five can be allocated at one-minute cadence. The data will be processed using the standard *Kepler* pipeline to produce de-trended light curves. The raw data will also be archived and available if the investigators choose to perform their own processing, for example, if one wants to carry out psf fitting to the pixel level data.

8.2. Data Analysis Program

All the *Kepler* data will be placed in the Multi-mission Archive at the STScI (MAST) and open to the public once the data have been validated and release approved by the principal investigator and NASA Headquarters, based on a pre-arranged schedule. There will be a latency of about one year due to the one-month intervals between data downlinks, transmission gaps which may not be recovered for another month, the need to de-trend the data in blocks of three months (each roll orientation), etc. Keep in mind that the nature and usefulness of the data has to do with the greatest extent of the time series for each object, not what is contained in a single snapshot of the FOV. This is not an imaging mission. Thus there will probably be more value and emphasis on DAP investigations that make use of the full mission length rather than snippets of early data. The archived data will consist of the calibrated and de-trended flux times series (light curves) for all target stars observed. GOs, however, should receive their data much sooner. We expect to reprocess the data several times based on a better understanding of the nature of the photometry as the mission progresses, including a final reprocessing

of the entire data set following the end of flight operations. The data will be maintained at the MAST for up to ten years after the end of the mission.

8.3. Participating Scientist Program

The PSP is separate from the GO and DAP programs in that the PSP is reserved for proposals that complement and enhance the core planet detection scientific program of the *Kepler Mission*. These programs might consist of supportive ground based follow-up observing, additional or improved analysis methods for planet detection, etc.

9. Status and Summary

The *Kepler Mission* is progressing through development toward a launch at the end of 2008. The optics have been delivered and tested. All of the flight CCDs have been in hand for over a year. The photometer assembly, integration and test will be completed within the year and delivered for spacecraft integration. The mission is designed to be capable of detecting hundreds of terrestrial planets, if they are common, or provide a significant null result if they are not. Either result would be profound. In addition, the photometric data base will be unique in precision, completeness, duration and number of objects and serve the community in many areas of astrophysical research for years to come.

Acknowledgements

Funding for this mission is provided by NASA's Discovery Program Office.

References

- Abazajian, Kevork *et al.* 2003, *ApJ*, The First Data Release Of The Sloan Digital Sky Survey, 126, 2081–2086
- Borucki, W. J. & Summers, A.L. 1984, *Icarus* 58, 121
- Borucki, W. J., Koch, D., Basri, G., Brown, T., Caldwell, D., DeVore, E., Dunham, E., Gautier, T., Geary, J., Gilliland, R., Gould, A., Howell, S., Jenkins, J., & Latham, D. *Kepler Mission: Design, Expected Science Results, Opportunities to Participate*, 2005, in *A Decade of Extrasolar Planets around Normal Stars* (ed. M. Livio), Cambridge: Cambridge University Press, in preparation
- Brown, Timothy M. and Gilliland, Ronald L. *Asterioseismology*, 1994, *ARAA*, 32, 37–82
- Castelli, F. & Kurucz, R. L. 2003, New grids of ATLAS9 model atmospheres, in *IAU Symposium 210, Modelling of Stellar Atmospheres*, (eds. N.E. Piskunov, W.W. Weiss. and D.F. Gray) 2003.
- Guinan, E. F., Engle, S. G., & Devinney, E. 2007, in *Solar and Stellar Physics through Eclipses* (eds. O. Demircan, S. O. Selam and B. Albayrak), ASP Conf. Series, in press
- Latham, D. 1992, in *IAU Coll. ASPCS*, McAlister and Hartkopf (eds.), 32:135, 110
- Kasting, J.F., Whitmire, D.P., & Reynolds, R.T. 1993, *Icarus* 101, 108
- Koch, D. G., Borucki, W., Dunham, E., Jenkins, J., Webster, L., & Witteborn, F. 2000, CCD Photometry Tests for a Mission to Detect Earth-Size Planets in the Extended Solar Neighborhood, in *SPIE Conference 4013, UV, Optical and IR Space Telescopes and Instruments*, Munich, Germany
- Koch, D., Borucki, W., Dunham, E., Geary, J., Gilliland, R., Jenkins, J., Latham, D., Bachtell, E., Berry, D., Deininger, W., Duren, R., Gautier, T. N., Gillis, L., Mayer, D., Miller, C., Shafer, D., Sobek, C., Stewart, C., & Weiss, M. 2004, Overview and status of the *Kepler Mission*, in *SPIE Conf 5487, Optical, Infrared, and Millimeter Space Telescopes*, Glasgow, Scotland
- Krivova, N. A., Solanki, S. K. & Floyd, L. 2006, *A&A*, 452, 631
- Niarchos, P.G., Munari, U., & Zwitter, T. 2007, these proceedings, 244

Sahu, K. C. & Gilliland, R. L. 2003, *ApJ*, 584, 1042

Discussion

RALPH GAUME: Given the reprioritizations going on now at NASA Headquarters and the fact that *Kepler* is getting close to launch, do you feel that *Kepler* is safe from budget reprogramming?

KOCH: I don't know of any program that is "safe". We have had to change our baseline about a dozen times, mostly due to external influences. I think all programs are on thin ice and very often one does not know how thin the ice may be at any one time.

CARLSON CHAMBLISS: Low-metallicity stars are probably less likely than others to have Earth-like planets. Are these going to be systematically removed from the 100,000-star sample?

KOCH: No. The *Kepler Mission* is designed to conduct an unbiased search. The mission has enough capacity to include every star in the field of view for which an Earth-size planet can be detected.