

Finding Earth-size planets in the habitable zone: the *Kepler Mission*

William Borucki¹, David Koch¹, Gibor Basri², Natalie Batalha³,
Timothy Brown⁴, Douglas Caldwell⁵, Jørgen Christensen-Dalsgaard⁶,
William Cochran⁷, Edward Dunham⁸, Thomas N. Gautier⁹,
John Geary¹⁰, Ronald Gilliland¹¹, Jon Jenkins⁵, Yoji Kondo¹²,
David Latham¹⁰, Jack J. Lissauer¹ and David Monet¹³

¹NASA Ames Research Center, Moffett Field, CA 94035, USA

²University of California-Berkeley, Berkeley, CA, 94720, USA

³San Jose State University, San Jose, CA, 95192, USA

⁴Las Cumbres Observatory Global Telescope, Golenta, CA 93117, USA

⁵SETI Institute, Mountain View, CA, 94043, USA

⁶University of Aarhus, Denmark

⁷University of Texas at Austin, Austin, TX, 78712, USA

⁸Lowell Observatory, Flagstaff, AZ, 86001, USA

⁹Jet Propulsion Laboratory, Pasadena, CA, 91109, USA

¹⁰Smithsonian Astrophysical Observatory, Cambridge, MA, 02138, USA

¹¹Space Telescope Science Institute, Baltimore, MD, 21218, USA

¹²NASA Goddard Space Flight Center, Greenbelt, MD, 20771, USA

¹³United States Naval Observatory, Flagstaff, AZ, 86002, USA

Abstract. The *Kepler Mission* is a space-based mission whose primary goal is to detect Earth-size and smaller planets in the habitable zone of solar-like stars. The mission will monitor more than 100,000 stars for transits with a differential photometric precision of 20 ppm at V=12 for a 6.5 hour transit. It will also provide asteroseismic results on several thousand dwarf stars. It is specifically designed to continuously observe a single field of view of greater than 100 square degrees for 3.5 or more years.

This overview describes the mission design, its goals and capabilities, the measured performance for those photometer components that have now been tested, the Kepler Input Catalog, an overview of the analysis pipeline, the plans for the Follow-up Observing Program to validate the detections and characterize the parent stars, and finally, the plans for the Guest Observer and Astrophysical Data Program.

Keywords. Planet detection, exoplanets, differential photometry, space-based telescope

1. Introduction

Over 250 exoplanets have been detected as of the time of this symposium (Marcy, report in IAUS 249, 2007). Most of these are gas giants. A few approaching super-Earths in short period orbits have also been found (Rivera *et al.*, 2005). However, the holy grail is to find habitable planets, that is, those in the habitable zone (HZ) (Kasting *et al.*, 1993), where liquid water can exist on their surfaces, and with a size and density such that they can have a life-sustaining atmosphere, that is, from about 0.8 to 2.2 R_{\oplus} or, if one assumes an Earth-like density, from about 0.5 to 10 M_{\oplus} . Finding planets more than three hundred times less massive than Jupiter is not trivial. Even finding extra-solar Jupiters took until 1995 (Mayor & Queloz, 1995). However, the transit method proposed by Borucki and Summers (1984) can detect Earth-size planets in the HZ. Thus, it was necessary to show that: 1) the variability of the Sun and presumably most stars similar to

the Sun on the time scale of a transit (on the order of 10-12 hours) is substantially smaller in amplitude than that of a Sun-Earth transit analog (84 ppm) (Jenkins 2002) and 2) to demonstrate that a space-based photometer with all the known forms of realistic noise has a combined differential photometric precision ≤ 20 ppm, one-sigma on the time scale of a transit (Koch, *et al.* 2000). The concept was proposed as a Discovery mission four times before finally being selected in December 2001 as NASA's tenth Discovery mission.

2. Mission Overview

Various aspects of the mission have been described in a number of papers (Borucki, *et al.* 2005, Koch, *et al.* 2004, Koch, *et al.* 2006). The top level scientific goals are to:

1. Determine the frequency of terrestrial and larger planets in or near the habitable zone of a wide variety of spectral types of stars;
2. Determine the distributions of sizes and orbital semi-major axes of these planets;
3. Estimate the frequency of planets and orbital distribution of planets in multiple-stellar systems;
4. Determine the distributions of semi-major axis, albedo, size, mass and density of short-period giant planets;
5. Identify additional members of each photometrically discovered planetary system using complementary techniques; and,
6. Determine the properties of those stars that harbor planetary systems.

To achieve these goals, three fundamental design requirements were established: the photometric precision, the mission life time and the number of stars observed. The photometric requirement is to detect individual Earth-size ($R=1.0 R_{\oplus}$) transits of 6.5 hrs (half of a central transit duration for a planet at 1 AU) of a twelfth magnitude solar-like star with an SNR of greater than or equal to four, when all sources of noise (stellar variability, shot and instrument) are included. Since a periodic sequence of at least three transits is required, the mission must last three or more years to detect planets in the HZ. Finally, the photometer is required to observe enough solar-like stars to produce a statistically meaningful result. Hence, the aperture size, field of view and location on the sky have been chosen to provide at least 100,000 dwarf stars.

The photometer design is based upon a classical Schmidt telescope with a 95 cm diameter aperture and more than one-hundred square degree field of view (FOV). The FOV is equivalent to about six Palomar Schmidt plates. The completed flight focal plane is shown in Figure 1.

The position of the FOV on the sky is 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 sample of stars similar to our local neighborhood and is a rich star field that is continuously viewable throughout the year as the spacecraft drifts away from the Earth as it orbits the Sun. The typical stellar distances for most of the usable stars are from a few hundred parsecs to about 1 kpc. Shot noise limits the magnitude of usable stars to about V=15-16 for F-, G- and K-dwarfs and about V=16.5 for M-dwarfs.

The single FOV will be viewed for the entire mission including any extended mission life. But to keep the fixed-solar array pointed toward the Sun and the focal-plane radiator pointed to deep space, the photometer-spacecraft must be rotated 90° about the optical axis every 93 days. The orientations of the CCD modules have been chosen so that the focal plane is four-fold symmetric (Figure 1). Therefore, after a 90° rotation all of the selected target stars remain on active pixels and the orientation of the columns remains the same on the sky except for the central module which is only two-fold symmetric.



Figure 1. Completed flight focal plane for *Kepler* with 42 science CCDs and 4 fine guidance sensor CCDs in the corners. The science CCDs are thinned, back-illuminated, and AR-coated with 1024 by 2200 pixels each with two readouts per CCD for a total of 84 science outputs. The pixels are 27 micrometers with a plate scale of 3.98 arcsec/pixel. The FOV is 13.9° side to side.

3. Measured Photometric Performance

The ability to detect transits depends on the photometric precision of the entire system on the time scale of transits, that is, from a few hours to about a half of a day. Absolute photometric accuracy is not necessary, since one is only looking for temporal variations. Common-mode noise can be removed by performing ensemble normalization within each of the 84 data channels, thereby taking out effects, e.g., gain variations and DC offsets.

The total noise we describe as the Combined Differential Photometric Precision (CDPP) in units of parts per million (ppm). The CDPP includes variability of the source, the photon shot noise and the measurement noise. Stellar variability is inherent in the source and uncontrollable. From the ACRIM-SMM (Willson *et al.* 1991) and DIARAD-SOHO (Froelich, 1997) measurements show that on the time scale of a transit, the solar variability is ≤ 10 ppm during solar-max. Note that solar variability is red noise and does not scale with $1/t^{1/2}$. Shot noise is determined by the brightness and spectral type of the star and the photometer design (aperture, obscuration, transmissions, reflectivity, band-pass filtering, quantum efficiency, etc.) To minimize the shot noise, the photometer has a single broad bandpass. Figure 2 shows the measured spectral response for the *Kepler* photometer. When convolved with the spectrum of a G2V (solar-like) star with $V=12$, the photometer will measure 4.7×10^9 photoelectrons in 6.5 hrs (91% duty cycle). This results in a shot noise contribution of 14.6 ppm.

The final contribution to CDPP is the measurement noise. This noise includes not only that from the CCD, but also electronic and optical cross talk, pointing jitter and any other noise introduced by the methodology used to make the measurements and reduce the data. At least fourteen individual terms are tracked and measured.

The Kepler Technology Demonstration (Koch, *et al.* 2000) was used once again in 2007 to test a single string engineering model of the flight detector system design. These tests demonstrated that the instrument noise requirement was met (after making modifications

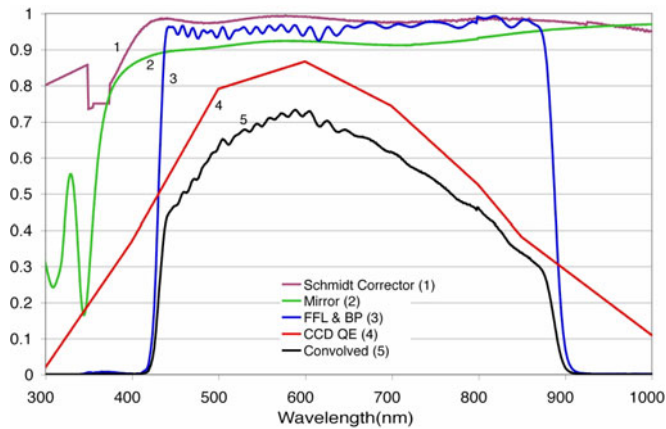


Figure 2. Spectral response curves. The FFL & BP is the average for the field flattener lenses on the CCD modules, which also have the bandpass filters. The CCD QE is the average from all 42 CCDs. The convolved also includes a 3-4.5% loss for contamination. The red cutoff is to avoid fringing in the CCDs and to minimize the potential for false positives produced by faint reddened background eclipsing binaries. The blue cutoff is designed to avoid the UV and Ca II H&K lines. For the Sun, 60% of the irradiance variation is at wavelengths less than 400 nm, but it only accounts for 12% of the total flux (Krivova, *et al.*, 2006).

in the design to eliminate under/overshoot due to an analog bandwidth issue) and that Earth-size transit signals in individual stars could be detected. For bloomed star images, it was again shown that photometric precision was preserved and transits were detected. A comprehensive test program of the components and full up focal plane has been performed at operating temperatures to measure and characterize such parameters as pixel response non-uniformity (PRNU), 2-D biases, under/overshoot, clocking cross talk (due to the fine guidance sensors which are read out at 10 Hz) and smear (due to reading out the CCD without having a shutter). All of this characterization information is being incorporated into the data reduction pipeline. The net result of the measurement noise contribution is 7.7 ppm. Figure 3 is a histogram of the individual CDPP values. The design requirement is a $\text{CDPP} \leq 20$ ppm for a $V=12$ G2V star and 6.5 hour integration.

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. Since *Kepler* does not send down data from every pixel in the focal plane, but rather only approximately 3-5% of the pixels of interest that are used for each star of interest, one has to preselect the stars to observe. 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 in 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 is 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 and Kurucz, 2003), the catalog contains the effective temperature, $\log(g)$, metallicity [Fe/H], reddening-extinction, mass and radius. The catalog was federated at the USNO Flagstaff Station with other catalogs, including 2MASS and USNO-B for cross reference and contains about 15 million objects.

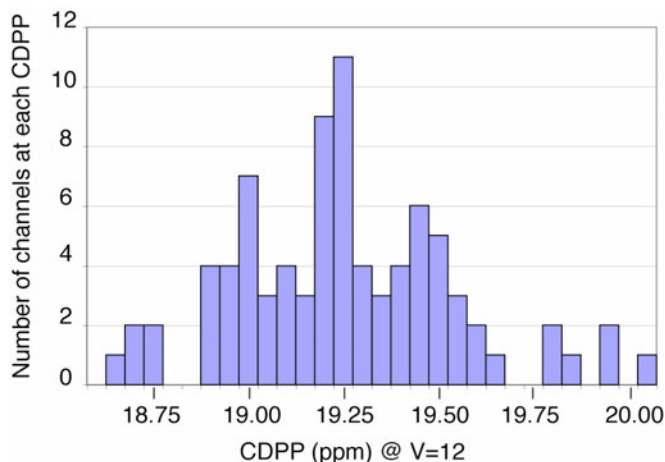


Figure 3. Combined Differential Photometric Precision (CDPP) for each of the 84 data channels for a $V=12$ solar-like star taking into account the measured noise from each channel, the effects of the predicted pointing stability, shot noise and the canonical 10 ppm for stellar variability.

Prior to launch, each star will be ranked for its potential for terrestrial planet detection using 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 of the brighter 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 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. 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. First a 2-D bias correction is made. Then the correction for undershoot is made, followed by the gain and non-linearity corrections. Cosmic ray hits are identified and removed. This is followed by the smear, dark current and flat-field corrections. Then two parallel paths

are followed to obtain stellar flux time series, one using optimal 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 as well, for example, for asteroseismic analysis and eclipsing binary modeling. 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 orbital 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 (Jenkins 2004).

All threshold crossing events then go through a validation process before they are considered viable candidates. This includes ensuring that each individual transit detected is consistent within the statistical limits to the character of the folded event. The individual pixels that make up each star are also examined to ensure that they do not exhibit any anomalous behavior. Finally, the results from the difference image analysis (DIA) method are examined to determine if there was any centroid shift during the event, which would be indicative of either a background eclipsing binary or a background transiting giant planet that is slightly off center from the target star and within the unresolved image.

6. Follow-up Observing Program

First, any transiting planet candidate with sufficient SNR to warrant it will be switched from the long cadence (thirty-minute) target list to the short cadence (one-minute) target list, to provide better resolution of the transit time, duration and shape.

Considerable ground-based observing resources will be utilized to eliminate false positives and to characterize both the host star and the planet itself where possible. This will be a step wise process, which will begin with using moderate precision radial velocity measurements to eliminate binary systems, such as with the Tillinghast Reflector Echelle Spectrograph on Mt. Hopkins and using high spatial resolution images from WIYN at Kitt Peak to eliminate any background object not detected with DIA. Once the viability of the candidate has been established, then more precious resources will be utilized such as the HET at McDonald Observatory, Keck and ultimately HARPS-North. The latter will have the capability of confirming the detection of planets as small as 2 Earth-masses out to 0.04 AU with 20% uncertainty with 50 hrs of observing of a $V=12$ G2V star (Lovis, *et al.* 2007). All of these spectra will also be used to improve our understanding of the parent star and to search for any companion non-transiting planets in the system.

7. Additional Planned Analyses

In addition to using the data to search for sequences of transits and reflected light from (not necessarily transiting) close-in giant planets, the data will also be used for: obtaining parallax measurements, thereby obtaining the stellar size from the luminosity and distance, which is essential to knowing the size of the transiting planet; determining stellar rotation rates; and performing asteroseismic analysis (Christensen-Dalsgaard *et al.*, 2007). The latter analysis will be performed by the Kepler Asteroseismic Science Consortium (KASC) to measure p-mode oscillations (Brown and Gilliland, 1994) of stars in the FOV brighter than about $V=11.5$, which have been observed with a one-minute cadence. This will yield the mass, radius, density and age of those stars.

8. Community Participation and Data Access

The community can participate in the *Kepler Mission* in several ways: a participating scientist program (PSP), a guest observer (GO) program, and an Astrophysical Data Analysis Program (ADP), 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 US proposals may receive funding.

8.1. Participating Scientist Program

The PSP is conceived to solicit proposals from scientists in the community to provide for potential direct enhancements to achieving the primary goals of the *Kepler Mission*. As described in the NASA ROSES-2007 section D.10, this includes such things as, detection of non-transiting planets using timing variations, improvements in determining the size of stars, and performing ground-based observing to aid in elimination of false-positives. Eight proposals have been selected for addition to the *Kepler* science team.

8.2. Guest Observer Program

Within the GO program, scientist may propose to view objects within the *Kepler* FOV which are not already on the planet detection target list, whether galactic or extragalactic. The *Kepler Target Catalog* (KTC) is expected to become publicly available shortly before launch. In general one may assume that all F-, G-, and K-dwarf stars to $V=14-15$ and M-dwarf stars to $V=16.5$ 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. Capacity for three thousand objects has been set aside for this program. The data will be processed using the standard *Kepler* pipeline to produce de-trended light curves. Solicitation for the GO is expected prior to launch with observing to begin shortly after commissioning.

8.3. Astrophysical Data Analysis Program

There are potentially a host of other astrophysical uses for the *Kepler* data (Granados & Borucki, 1994), given the uniqueness in precision, completeness, duration and number of stars in the archived data. Potential uses include such things as; analysis of white light flaring and stellar activity, which can yield star spot cycles, especially if the mission is extended beyond about half of a solar cycle; the frequency of Maunder minimums for solar-like stars, which has implications for paleoclimatology and perhaps the future of our Earth's climate; 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.

9. Status and Summary

Kepler is NASA's first mission capable of detecting Earth-size and smaller exoplanets in the HZ. The hardware and software are progressing through development toward a launch in February of 2009. The photometer assembly, integration and test will be completed in early 2008 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.

Acknowledgements

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

References

- Abazajian, Kevork *et al.* 2003, The First Data Release Of The Sloan Digital Sky Survey, *ApJ*, 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. 2005, *Kepler Mission: Design, Expected Science Results, Opportunities to Participate*, in (ed. M. Livio), *A Decade of Extrasolar Planets around Normal Stars* (Cambridge: Cambridge University Press) in preparation
- Brown, Timothy, M. & Gilliland, Ronald L. 1994, *Asteroseismology*, *ARAA*, 32, 37–82
- Castelli, F. & Kurucz, R. L. 2003, New grids of ATLAS9 model atmospheres, in (eds. N.E. Piskunov, W. W. Weiss. and D. F. Gray) *IAU Symposium 210, Modelling of Stellar Atmospheres*
- Christensen-Dalsgaard, J., Arentoft, T., Brown, T. M., Gilliland, R. L., Kjeldsen, H., Borucki, W. J. & Koch, D., 2007, The *Kepler* Asteroseismic Investigation. in (eds L. Gizon & M. Roth) *Proc. HELAS II International Conference: Helioseismology, Asteroseismology and the MHD Connections*, (*Göttingen, J. Phys.: Conf. Ser.*), in the press.
- Frohlich, C. 1987, *JGR* 92, 796
- Granados, A. F. & Borucki, W. J., (eds.) 1994 *Astrophysical Science with a Spaceborne Photometric Telescope NASA Conf. Publ. 10148* (Mt View)
- Jenkins, Jon M, 2002, The Impact of Solar-like Variability on the Detectability of Transiting Terrestrial Planets, *ApJ* 575, 493
- Jenkins, Jon M. & Doyle, Laurance R. ,2003, Detecting Reflected Light from Close-in Extrasolar Giant Planets with the Kepler Photometer, *ApJ*, 595, 429
- 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)
- Koch, D. G., Borucki, W., Basri, G., Brown, T., Caldwell, D., Christensen-Dalsgaard, J., Cochran, W., Dunham, E., Gautier, T., Geary, J., Gilliland, R., Jenkins, J., Kondo, Y., Latham, D., Lissauer, J., & Monet., D. 2006, The *Kepler Mission: Astrophysics and Eclipsing Binaries*, *ApSS* 304, 389
- Krivova, N. A., Solanki, S. K., & Floyd, L. 2006, *A&A*, 452, 631
- Lovis. C. *et al.*, 2007, in (eds. Debra Fischer, Fred Rasio, Steve Thorsett & Alex Wolszczan) *ASP Conference Series, Proceedings of the Santorini conference on Extreme Solar Systems*, in preparation.
- Mayor, M. & Queloz, D. 1995, A Jupiter-Mass Companion to a Solar-Type Star, *Nature*, 378, 355
- Rivera, Eugenio J., Lissauer, Jack J., Butler, R. Paul, Marcy, Geoffrey W., Vogt, Steven S., Fischer, Debra A., Brown, Timothy M., Laughlin, Gregory, & Henry, Gregory W. 2005, A 7.5 M_⊕ Planet Orbiting the Nearby Star, GJ 876, *ApJ*, 634, 625
- Sahu, K. C. & Gilliland, R. L. 2003, *ApJ*, 584, 1042
- Willson, R. C., Gulakis, S., Janssen, M., Hudson, H. S. & Chapman, G. A. 1981, *Science* 211, 700