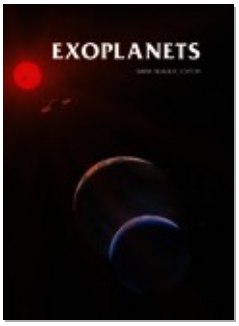
# Exoplanets

Sara Seager

Published by University of Arizona Press



# Statistical Distribution of Exoplanets

**Andrew Cumming**
*McGill University*

This chapter discusses the current statistical sample of exoplanets. We discuss the selection effects in radial velocity and transit surveys, followed by a brief introduction to statistical techniques for characterizing the orbital properties of planets and how to include completeness corrections in population studies. We then highlight the major features of the planet population discovered so far, discuss some of the implications for planet formation theories, and the future prospects for increasing the sample of known planets.

## 1. INTRODUCTION

The last 15 years have seen a tremendous rate of discovery of exoplanetary systems. At the time of writing in September 2010, the Extrasolar Planets Encyclopedia website reports 490 planets in 413 planetary systems (*http://exoplanet.eu,* maintained by J. Schneider, Paris Observatory, accessed on September 1, 2010). These planets have masses that range from 3 $M_\oplus$ and up, orbital periods from close to one day to several years, and a wide range of eccentricities. The number of confirmed multiple planet systems is 49, with many other single planet systems showing evidence for an additional companion at long orbital periods. This large number of planets has allowed the first studies of the distribution of orbital periods, eccentricities, and planet masses, as well as the dependence on host star properties. For example, it is now well established that the planet occurrence rate increases strongly with the metallicity of the host star. The number of transiting planets is over 100, allowing studies of the mass-radius relation for hot Jupiters, and the first glimpses of exoplanet atmospheres.

One reason to be interested in the statistical properties of exoplanets is to answer the question of how common planetary systems are, particularly those with habitable planets. We now know that at least 10% of solar type stars, perhaps up to 20%, harbor gas giant planets, and information about lower masses will follow. Already, lower-mass planets with masses ~10 $M_\oplus$ comparable to Neptune are being studied in close orbits. Systems similar to our own solar system, dominated in mass by Jupiter at 5 AU from the Sun, are beginning to be found, including a Jupiter/Saturn analog recently detected in a microlensing event. The frequency of planets determined from current data is an important input for future surveys, such as astrometric and direct searches.

Another reason for studying the statistical distributions of planet properties is that they offer us a tremendous amount of information about the process of planet formation. Planets likely form in the outer regions of protoplanetary disks and migrate inward. The radial distribution of observed planets offers clues to the physics of migration and the process by which it stops close to the star. The increasing planet occurrence rate with stellar metallicity and stellar mass is likely directly related to the amount of planet-building materials in the protoplanetary disk. The nonzero eccentricities of most exoplanets perhaps point to planet-planet scattering or other gravitational interactions as an important process early in the life of most planetary systems. Part of the puzzle is to understand how our solar system, in which the planets have nearly circular orbits and the gas giants orbit at larger distances than most of the exoplanets so far discovered, fits into this picture.

When studying the statistical distributions of planet properties, we must be careful to account for selection effects. For example, in radial velocity searches, the velocity precision (typically $\sigma \approx$ 1–3 m/s) sets a lower limit on the amplitude of velocity variations that can be detected. The result is that low-mass planets in wide orbits are much harder to detect than massive planets in close orbits. We must include this fact when trying to estimate the fraction of stars with planets, and when interpreting the distribution of planetary masses and orbital periods.

This chapter is an overview of the statistical properties of the known planets as of September 2010. We focus on the planets detected by radial velocity, transit searches, and microlensing techniques. Radial velocity and transit searches have both led to large samples of planets. Although microlensing has so far uncovered only a handful of planets, the results already constrain the occurrence rates of planets with different properties than those detected by the radial velocity and transit methods. We start in section 2 by describing the selection effects in each of these search techniques, and how to understand them. We review the main statistical techniques that are used to characterize the properties of the planet from the data, and infer properties of the population from large surveys. In section 3, we present the observations and highlight some of the lessons so far about planet formation from the distribution of planet properties. We conclude in section 4 with a discussion of future prospects for this field. The number of planets promises to increase dramatically in future years, making this a particularly exciting time for the study of exoplanet statistics.

## 2. STATISTICAL TECHNIQUES AND SELECTION EFFECTS

In this section, we first discuss how to set the detection threshold for a radial velocity or transit observation (section 2.1), and then how to understand the resulting selection effects in radial velocity (section 2.2), transit (section 2.3), and microlensing surveys (section 2.4). In section 2.5, we discuss the importance of understanding the properties of the host star and population of stars in the survey. In section 2.6, we briefly introduce some of the techniques that have been developed to characterize the orbital properties of a planet from the data. Finally, in section 2.7, we discuss how to use our knowledge of the selection effects to infer properties of the planet population.

### 2.1. Setting a Detection Threshold

A typical set of radial velocity measurements is shown in Fig. 1, taken from the paper by *Robinson et al.* (2007) that announced the discovery of a planet with minimum mass 1.9 $M_{Jup}$ in a 675-d low-eccentricity orbit around the star HD 5319. The velocity variations due to the planet can be clearly seen in this dataset, as they have a much larger amplitude than the measurement errors. But it is interesting to consider what would happen for smaller-mass planets, which induce a smaller amplitude of velocity variations. When this amplitude becomes comparable to the measurement errors, there must be a point at which we can no longer say with confidence that we are seeing an orbiting planet in the data. How to calculate this detection threshold is the subject of this section.

We start by considering radial velocity data, but similar ideas apply to analysis of lightcurves for transits. Imagine we are given a set of N measured velocities $v_i$ where i labels each individual measurement, i = 1 to N, along with the observation time $t_i$ and the measurement error $\sigma_i$. The simplest approach is to carry out a $\chi^2$ fit of a Keplerian orbit to the data. For each trial orbit, we calculate

$$\chi^2 = \sum_{i=1}^{N} \left( \frac{v_i - V(t_i)}{e_i} \right)^2 \qquad (1)$$

where $V(t_i)$ is the predicted velocity at time $t_i$ for the orbital parameters being considered, and the estimated error $e_i$ for each measurement includes the Doppler measurement errors $\sigma_i$, but could also be augmented by other sources of noise such as intrinsic stellar variability. The best fitting orbit is the one that minimizes the value of $\chi^2$. [We show how to calculate $V(t_i)$ in section 2.2. We need five parameters to specify $V(t_i)$ for each planet in the model: the orbital period P, the planet minimum mass $M_p$ sin i, the eccentricity e, longitude of periastron $\omega$, and the time of periastron $t_p$. In addition, we need to include the systemic velocity, and in some cases (such as in Fig. 1) add a long-term velocity trend.]

Minimizing $\chi^2$ in this way tells us the set of best-fitting orbital parameters, i.e., characterizes the orbit. (In addition, how quickly $\chi^2$ grows as the parameters are changed from their best-fitting values gives a measure of the error in each parameter. We will say more about this in section 2.5.) But the question of whether we have a detection is a different one. A model that includes a planet will always fit the data better, in the sense that it will always lower the minimum $\chi^2$, because we have more free parameters to adjust. The important question to ask is whether the reduction in $\chi^2$ is significant. If we were to model the velocities in Fig. 1 without the planet, the value of $\chi^2$ would be significantly larger. Because we can lower $\chi^2$ by such a large amount with only five additional parameters, the planet model is strongly preferred. The principle at work here is similar to Occam's razor; we want to keep the model as simple as possible while adequately describing the data.

This idea underlies the Lomb-Scargle (LS) periodogram, a commonly used tool for searching for periodicities in measurements that are unevenly sampled in time (*Lomb,* 1976; *Scargle,* 1982). The LS periodogram power is calculated for each trial period P as

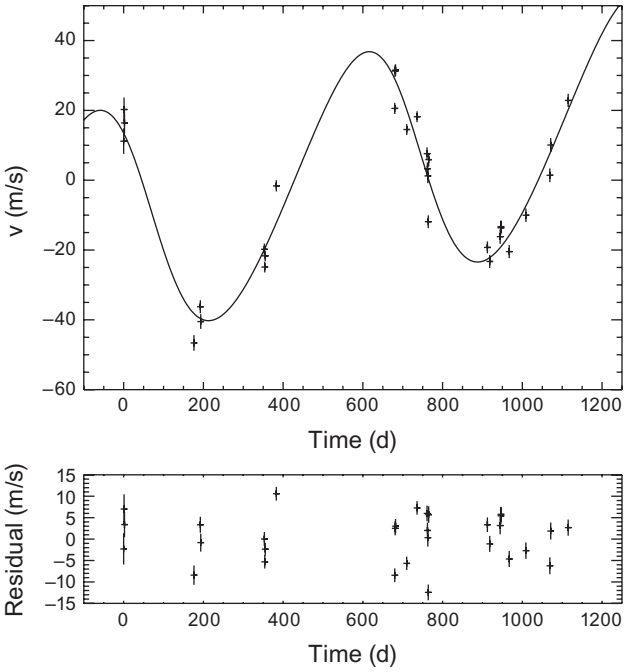



**Fig. 1.** An example radial velocity dataset for the star HD 5319 (*Robinson et al.,* 2007). The error bars correspond to the measurement errors. The best fit model is shown, which has a planet in a 1.84 yr orbit with K = 33.7 m/s and e = 0.12. A long term acceleration is included in the fit indicating a massive long period companion is also present in the system. The lower panel shows the measured velocities after subtracting the best fit Keplerian orbit. The scatter is significantly larger than the measurement errors, but can be accounted for by intrinsic stellar jitter.

$$z(P) = \frac{\Delta\chi^2}{\chi^2} = \frac{\chi_0^2 - \chi^2(P)}{\chi^2(P)} \quad (2)$$

where $\chi_0^2$ is the minimum value of $\chi^2$ from a fit of a constant to the data, and $\chi^2$ (P) is the minimum $\chi^2$ from a fit of a sinusoid with period P plus constant. If including the sinusoid reduces $\chi^2$ significantly, the power z will be large. The procedure is then to calculate z(P) for a number of different trial periods and look for the period that gives the largest power, $z_{max}$. Figure 2 shows the LS periodogram for the data shown in Fig. 1. There is a large peak in the power at a period just under 700 days, which matches nicely the final orbital period determined by *Robinson et al.* (2007).

The LS periodogram is often used in planet searches to quickly identify likely periods for planets. For circular orbits, the velocity variations are sinusoidal, and so the LS periodogram fits the correct model, but even for eccentric orbits, where the velocity curve becomes significantly non-sinusoidal, the LS peridogram often gives a good initial estimate for the orbital period. (To do better for noncircular orbits, *Cumming* (2004) defined a "Keplerian periodogram" in an analogous way to equation (2) but fitting a Keplerian orbit at each period.)

We can use the values of $z_{max}$ to decide whether we have detected a planet: Only if the power exceeds a threshold value, $z_{max} > z_{th}$, do we count the fit as significant and say that we have a detection. The detection threshold $z_{th}$ can be determined with Monte Carlo simulations. By generating many synthetic datasets consisting of a constant velocity plus random noise chosen to represent the measurement errors and stellar variability, the distribution of $z_{max}$ that arises due to noise fluctuations can be determined. [In the Monte Carlo simulations, the noise can be generated by assuming a Gaussian distribution, or by using the measured velocities themselves to estimate the noise distribution (so-called "bootstrapping") (see *Press et al.,* 1992). For example, *Marcy et al.* (2005a) discuss the calculation of the false alarm probability by scrambling the order of the detected velocities.] The threshold $z_{th}$ is then set to be the value of $z_{max}$ that is exceeded in some small fraction F of the Monte Carlo simulations. If we observe $z_{max} > z_{th}$, then we know that we have a detection, with a small false alarm probability F that the signal is due to noise fluctuations. The choice of F is determined by our willingness to tolerate false alarms. The higher we set $z_{th}$, the lower the false alarm rate F, but we also lose the ability to detect low-amplitude signals. This trade off is something that always has to be dealt with when deciding where to set the detection threshold (e.g., *Wainstein and Zubakov,* 1962). Figure 2 of *Cumming* (2004) gives the false alarm probability for Keplerian orbit fits for different values of N and $\Delta\chi^2/\chi^2$.

An important quantity is the number of statistical trials that we make in our search of parameter space. The more different combinations of parameters we try, the more likely it is that noise fluctuations will mimic a planetary signal and give a large value of z. Therefore the value of $z_{th}$ increases as we search a larger range of parameters. An example is the number of independent frequencies $N_f$ (or periods) that are searched. Monte Carlo simulations must be used to determine $N_f$ , but a simple analytic estimate gives $N_f \approx 3000$ for typical radial velocity datasets (*Cumming,* 2004). This means that to detect a signal with a false alarm probability of $10^{-3}$, the periodogram power for that signal must be large enough, or the Keplerian fit good enough, that the false alarm probability for a search at a single frequency would be $\sim 10^{-6}$.

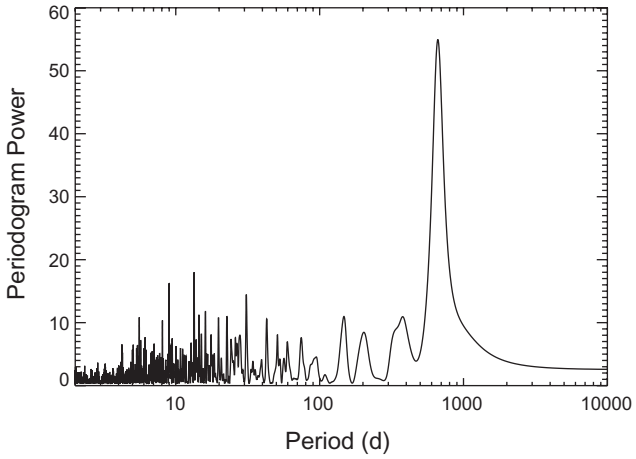



**Fig. 2.** Lomb-Scargle periodogram of the data for HD 5319 shown in Fig. 1, which measures the improvement in $\chi^2$ when the data are modeled by a sinusoidal velocity variation plus constant rather than a constant alone. There is a peak at 1.8 yr, which matches well with the final derived orbital period when a long-term trend and nonzero eccentricity are included. The periodogram shows many peaks, spaced equally in frequency with separation ~1/T where T is the timespan of the dataset. The peaks are equally spaced in frequency, so that on the log scale used here, the density of peaks increases toward shorter periods.

For transits, the same approach applies, of defining a detection statistic, and setting a threshold that takes into account the noise level and number of trials. There are a variety of detection statistics that have been proposed for detecting the signature of a transiting planet. These detection statistics are essentially based on $\chi^2$ fitting of a box-shaped dip to the stellar lightcurve. *Moutou et al.* (2005) carry out a blind test comparison of five different detection methods on simulated lightcurves, which include a matched filter method (*Jenkins et al.,* 1996), box-fitting least squares (*Kovács et al.,* 2002), and a Bayesian approach (*Defaÿ et al.,* 2001), finding that no one technique offers an advantage in all situations. Indeed, *Schwarzenberg-Czerny and Beaulieu* (2006) point out the equivalence of some of these methods for Gaussian noise, but this may not be the case if the data have a red noise component coming from systematic errors that introduce correlations over time (*Pont et al.,* 2006).

It is important to stress that we have been discussing the statistical detection threshold, but that may in fact not be the

criterion used to determine detections, especially early in a survey when systematic effects are not fully understood, and the same detection threshold may not be applied uniformly in a given survey. For example, the first planets discovered by Doppler surveys had large signal-to-noise ratios, and are essentially detectable by eye. *Marcy and Butler* (1998) note that "experience shows that a confident detection requires that the amplitude be ~4 times the Doppler error." Recent radial velocity detections have velocity amplitudes much closer to the noise level, and statistical analysis has become more important, with the planet search teams often reporting false alarm probabilities for their detections (*Marcy et al.,* 2005a). Another example is the OGLE-III transit survey, in which the transit candidates were selected by eye from a list of "pre-candidates." *Gould et al.* (2006a) pass a number of simulated transits through the same by-eye selection procedure in an attempt to objectively quantify the corresponding detection threshold as a function of the signal-to-noise ratio.

The other point to note is that the methods discussed here identify significant periodic signals in radial velocity data or dips in stellar lightcurves, but it may be the case that the detected signal is not due to a planet. For example, in radial velocity searches, a periodic signal could be caused by stellar magnetic activity, and be related to the rotation period of the star. In transit surveys, typically only a small fraction of identified candidates turn out to be transiting planets; most are other kinds of systems such as transits by low-mass stars.

### 2.2. Selection Effects in Radial Velocity Searches

Having set a detection threshold, we can ask what orbital parameters lead to signals that are detectable. The radial velocity of a star with an orbiting planet is

$$V(t) = V_Z + K\left[\cos\left(f(t)+\omega\right)+e\cos\omega\right] \tag{3}$$

where $V_Z$ is the systemic velocity, K is the semiamplitude, and the remaining factor in square brackets sets the shape of the velocity curve. The amplitude is

$$K = \frac{28.4\text{m s}^{-1}}{\sqrt{1-e^2}}\left(\frac{M_P \sin i}{M_J}\right)\left(\frac{P}{1\text{ yr}}\right)^{-1/3}\left(\frac{M_\star}{M_\odot}\right)^{-2/3} \tag{4}$$

where the mass of the star is $M_\star$, the planet mass is $M_P$, the orbital period is P, and the inclination of the orbit to the line of sight is i (where i = 90° if we are looking at the orbit edge-on). The shape of the velocity curve depends on the longitude of periastron ω and how the true anomaly f varies with time, given by the relations

$$\tan\left(\frac{f(t)}{2}\right) = \left(\frac{1+e}{1-e}\right)^{1/2}\tan\left(\frac{E(t)}{2}\right)$$
$$E(t) - e\sin E(t) = M(t) \qquad M(t) = \frac{2\pi}{P}(t-t_P) \tag{5}$$

where E(t) is the eccentric anomaly, and M(t) is the mean anomaly. For a circular orbit, f(t) = E(t) = M(t), and the velocity variations are sinusoidal.

There are three main factors that determine whether the velocity variations induced by a planet will be detectable in a radial velocity dataset. The first is the signal-to-noise ratio. A larger-amplitude K leads to a larger-power $z_{max}$ on average, and so increases the likelihood of detection. Using analytic results for Gaussian noise, the signal-to-noise ratio required for a 50% detection rate is

$$\frac{K_0}{\sqrt{2}s} \approx \frac{1}{\sqrt{N}}\left[2\ln\left(\frac{N_f}{F}\right)\right]^{1/2} \tag{6}$$

(*Cumming,* 2004) (here we give the large N limit), where $N_f$ is the number of independent frequencies that were searched, and F is the desired false alarm probability. The noise level s includes contributions from measurement errors and stellar jitter. The $1/\sqrt{N}$ factor should be no surprise: If we take N measurements each with an error s, then the uncertainty in the mean of those measurements we expect to be $s/\sqrt{N}$, and this gives a measure of what velocity amplitude could be detected. For example, for typical values $N_f \approx 1000$ and $F \approx 10^{-3}$ (so that on the order of one false alarm would be expected in a survey of 1000 stars), we find $K_0 \approx 7\ s/\sqrt{N}$. Using equation (4), the planet mass that can be detected 50% of the time is

$$M_{p,50} = \frac{0.14\ M_J}{\sin i}\left(\frac{P}{1\text{ yr}}\right)^{1/3}\left(\frac{M_\star}{M_\odot}\right)^{2/3}\left(\frac{s}{3\text{ m/s}}\right) \tag{7}$$

or

$$M_{p,50} = \frac{1.0\ M_\oplus}{\sin i}\left(\frac{P}{1\text{ d}}\right)^{1/3}\left(\frac{M_\star}{0.3\ M_\odot}\right)^{2/3}\left(\frac{s}{1\text{ m/s}}\right) \tag{8}$$

where we assume N = 30 (with $M_p \propto 1/\sqrt{N}$ for different N).

The second factor is the time span of the observations, T. For long orbital periods P > T, we see only part of the orbit, and the velocity variation is smaller than the full velocity amplitude K. Therefore the velocity amplitude needed for detection $K_{det}$ increases with period for P > T. The rate at which it increases depends on the phase of the orbit being detected: When the orbit is observed near a zero-crossing rather than a maximum or minimum, the velocity variation is larger. Averaging over phase, *Cumming* (2004) shows that for a 50% detection threshold (in which case we can rely on the zero-crossing phases only), $K_{det} \propto P$ whereas for a larger detection efficiency, the scaling is $\propto P^2$. Of course, even if a long-period orbit is detected, it is not possible to determine whether the companion is of planetary mass: Characterization of the orbit is possible only after a whole orbit has been observed. For example, the velocities in Fig. 1 show a clear upward trend over the time span of the data, presumably due to an additional companion with a long-period orbit. However, only after continued monitoring will its orbital period and mass be securely identified.

The third factor is the shape of the velocity curve, which depends on the eccentricity and the orientation of the orbit. The orbit shown in Fig. 1 has a low eccentricity, giving a radial velocity curve that is close to sinusoidal. A planet on a very eccentric orbit on the other hand has a much more distorted velocity curve (Fig. 3), with large velocities during periastron passage, but smaller velocities during most of the orbit. Unless we are lucky enough to observe the star while the planet is close to periastron, the orbit could go undetected. The two velocity curves in Fig. 3 are drawn from a large number of simulated observations using the same observation times and velocity amplitude, but random phases, of an orbit with e = 0.5. The upper panel in the figure shows a case that was detected by the detection algorithm, whereas the lower panel was not. The effect of eccentricity on detectability has been calculated by several authors (*Endl et al.,* 2002; *Cumming,* 2004; *Wittenmyer et al.,* 2006). There is good agreement that the detectability falls off for e ≥ 0.5–0.6.

In equations (7) and (8) above, we use values s = 1 and 3 m/s, corresponding to the long-term precision currently available. For example, the error bars in Fig. 1 show the measurement errors from the HIRES spectrometer at the Keck telescope, and range from 1.5 to 3.6 m/s. A long-term accuracy of $\sigma \approx 1$ m/s has been achieved at the HARPS spectrograph at La Silla Observatory, Chile (*Lovis et al.,* 2006). The ability to track stellar radial velocities at the ~1 m/s level on timescales of 10 yr is a remarkable achievement. In fact, the precision is good enough that the scatter in the residuals to orbital solutions usually has a significant component from stellar variability or "jitter" (added in quadrature to the measurement errors), which we discuss in section 2.4.

The observed detection threshold of radial velocity surveys matches these expectations quite well. For example, *Cumming et al.* (2008) show that equation (6) provides a good description of the K-N relation for announced planets from the Keck Planet Search, if the noise s for each star is estimated from the residuals to the best-fitting orbital solution. However, for N ≥ 60 there is an apparent floor in the signal-to-noise ratio of announced planets of K/s ≈ 2. The likely cause of this is that at small signal-to-noise ratio it is more difficult to rule out other explanations such as stellar variability and thereby confirm a statistically significant signal is indeed due to a planet.

### 2.3. Selection Effects in Transit Searches

Searches for transiting planets look for dips in the stellar lightcurve as the planet repeatedly transits its host star. *Mandel and Agol* (2002) and *Seager and Mallén-Ornelas* (2003) give analytic expressions for transit lightcurves. Four main parameters can be determined from the lightcurve. The time between transits gives the orbital period. The depth of the transit depends on the ratio of the areas of the disk of the planet and star, $\delta \approx (R_p/R_\star)^2$, or $\delta \approx 0.01$ for Jupiter. The other two observables are the duration of the transit and the time for ingress or egress, which allow the orbital radius and inclination to be determined. The detailed shape of the transit lightcurve also depends on the stellar limb darkening.

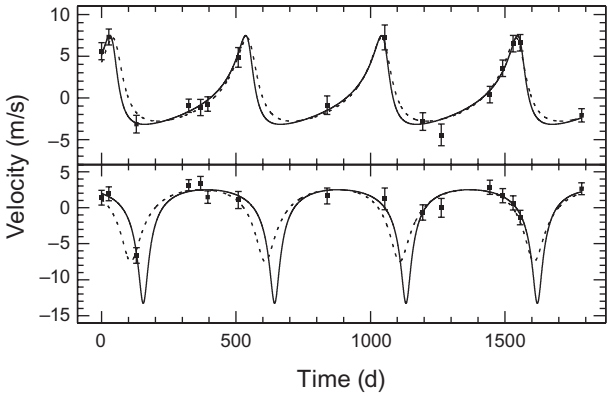



**Fig. 3.** Examples of velocity curves with e = 0.5 that are (top panel) and are not (bottom panel) detected. The dotted line in each case shows the true orbit, the points are the observed velocities, and the solid curve shows the best-fitting orbit. In both cases, the solid curve gives a lower $\chi^2$ than the dotted curve. The lower panel has only a single measurement during the periastron passage, and is not a significant detection.

The observed transiting planets have radii ranging from ≈0.15 to 1.8 $R_{Jup}$, masses ≈5 $M_\oplus$ to >10 $M_{Jup}$ (radial velocity follow-up is required to determine the mass of a transiting planet), and short orbital periods, concentrated at ~1 to several days. These properties are consistent with the fact that detectability falls off strongly with increasing orbital period and decreasing planet radius. To see this (e.g., *Horne,* 2003; *Gaudi et al.,* 2005; *Gaudi,* 2005), the first step is to write down the expected signal-to noise ratio. If there are $N_{obs}$ observations made of a given star, the number of observations during transit will be on average $N_t \approx N_{obs}(R_\star/\pi a)$, since the duration of the transit is roughly a fraction $\approx R_\star/\pi a$ of the orbital period (a planet orbiting farther out spends a smaller fraction of its time in transit than one orbiting close in). The signal-to-noise ratio will then be

$$S/N \approx \sqrt{N_t}\,\frac{\delta}{\sigma} \tag{9}$$

where $\sigma$ is the photometric precision per observation. By analogy with equation (6) for radial velocities, we expect that there will be a critical value of S/N required for a detection. For example, *Tingley* (2003) finds that S/N ≈ 16 for a 50% detection rate with 1% false alarm probability, the exact number depending on the particular detection statistic adopted (although note that a FAP of much less than 1% will be typically necessary given the large numbers of stars surveyed).

The sensitivity to orbital period and planet radius comes about because the search volume, or the number of stars surveyed, is different for different planet parameters (*Gaudi,* 2005). If we assume that the noise level is set by source dominated photon noise, then $\sigma \propto 1/\sqrt{N_\gamma} \propto d$ where d is the distance to the star, since the stellar flux falls off as $1/d^2$. Therefore, if we fix the signal-to-noise threshold and stellar properties, equation (9) gives a search volume $\propto d^3 \propto N_t^{3/2}\, \delta^3 \propto R_p^6/P$. The probability that a planetary orbit has an inclination leading to a transit is $\approx R_\star/a$. Multiplying the transit probability by the search volume, and assuming a constant space density of stars, we find that the number of detections is $\propto R_p^6/P^{5/3}$ (*Gaudi,* 2005). Therefore, the number of planets discovered should increase strongly with planet radius, and fall off strongly with orbital period. Based on this simple formula, we see that going from an orbital period of 1 to 3 d gives a factor of 6 suppression in detectability, and going from a radius of 1 $R_{Jup}$ to 1.3 $R_{Jup}$ increases detectability be a factor of 5.

This estimate predicts the general fall off with period well, but in detail the detection probability as a function of period shows a series of sharp dips around this trend set by the spacing of the observations [see Fig. 1 of *Beatty and Gaudi* (2008) for an example of the detectability as a function of orbital period, known as the window function]. This aliasing arises because the observations are taken during the same ≈8-h window each day. Depending on the phase of the orbit at the time of observation, this can make orbits with periods close to an integer number of days easier or harder to detect. For example, if the phase is such that the transit occurs in every observing window, then the detectability is greater than a noninteger period, but the opposite could also be true, that the phase is such that transits happen outside the observing window! *Gould et al.* (2006a) show that for a signal-to-noise limited survey the first effect wins out, leading to a net enhancement of detectability at integer periods. The finite spacing of the observations has other consequences. For example, scheduling radial velocity follow up to confirm the mass of a candidate transiting planet requires that two transits be observed, giving the orbital period. For a given observing strategy, this becomes less likely for longer orbital periods (*Gaudi et al.,* 2005).

Another important complication is the presence of red noise (*Pont et al.,* 2006). Equation (9) for the signal-to-noise ratio assumes that the photometric errors are uncorrelated, i.e., that the noise was white. In fact, systematic effects such as changing atmospheric or telescope conditions over time lead to correlations in the errors in transit surveys, so that equation (9) overestimates the effective signal-to-noise. Several recent papers attempt to untangle the effects of red noise on detection thresholds in transit surveys; see, e.g., *Pont et al.* (2006) and *von Braun et al.* (2009).

Understanding the selection effects in transit surveys in detail has proved to be complex. Although there are now more than 100 transiting planets, the discoveries at first did not come as quickly as expected based on the known occurrence rate of hot Jupiters from radial velocity surveys. It became apparent that the yield of transit surveys was initially overestimated by large factors (e.g., *Horne,* 2003). The yields are now much better understood. *Beatty and Gaudi* (2008) consider a number of different transit surveys that have adopted different observing strategies, and carefully estimate the predicted yields, finding numbers that are close to the actual number of reported detections.

### 2.4. Selection Effects in Microlensing Surveys

Ten planets discovered using microlensing had been published at the time of writing, with masses in the range 3 $M_\oplus$ to 3.5 $M_{Jup}$, and orbital radii from 0.6 to 5.1 AU. This range of planet mass and orbital radius can be understood by considering that the planet must cause a significant deviation from the microlensing lightcurve that would otherwise result from lensing of the background star by the lens star. This requires that the orbital radius of the planet lies close to the Einstein radius

$$R_E \approx 2.9\,\mathrm{AU}\left(M_L/M_\odot\right)^{1/2}\left(D/\mathrm{kpc}\right)^{1/2} \qquad (10)$$

where $R_E^2 = (4\,GM_L/c^2)D$, $M_L$ is the mass of the lens (planet host star), and D is the reduced distance ($1/D = 1/D_{LS} + 1/D_{OL}$ with $D_{LS}$ being the lens source distance, and $D_{OL}$ the observer lens distance).

Given a planet in the lensing zone (typically a within about a factor of 2 of $R_E$), the probability of detection can be close to 100% for high-magnification central caustic events, in which the source image sweeps around the Einstein ring (*Griest and Safizadeh,* 1998). This is a powerful way to detect multiple planet systems, since there is a high probability of seeing all detectable planets within the lensing zone. For lower-magnification planetary caustic events, the detection probability depends on the relative path of the source and lens, but is typically tens of percent for Jupiter-mass planets to a few percent for Earth-mass planets (*Gould and Loeb,* 1992; *Bennett and Rhie,* 1996).

A limit on the detectable planet mass is set by the finite angular size of the source star, which washes out the perturbation due to low-mass planets if the angular size of the source star is greater than the Einstein radius of the planet. (However, note that detection of finite size effects in the lightcurve is important, as it enables a measurement of the Einstein radius and therefore the lens mass.) The detectable mass can therefore be roughly estimated by setting the angular size of the planet's Einstein radius $\theta_E = R_{E,p}/D_L = q^{1/2}R_E/D_L$, where $q = M_p/M_\star$ is the mass ratio, to the angular radius of the source $\theta_\star = R_\star/D_S$. This gives

$$M_P \gtrsim 1M_\oplus\left(\frac{R_\star}{R_\odot}\right)^2\left(\frac{D_L}{D_S}\right)^2\left(\frac{D}{1\,\mathrm{kpc}}\right)^{-1} \qquad (11)$$

where the stellar radius depends on the type of star, $R_\star \sim R_\odot$ for a main-sequence star or could be ~10 $R_\odot$ for a giant. Such low planet masses at orbital radii of several AU are

not possible to detect with other techniques, so microlensing provides a unique and important probe of the planet distribution in this region.

### 2.5. Stellar Properties

Alongside the rapid growth in the number of known exoplanets has been a large amount of work to understand the properties of their host stars. Understanding stellar properties is important for detecting and characterizing the planets in a given system, and for interpreting the statistical properties of the planet sample.

For transiting planets, knowledge of the planet properties is limited by how well the properties of the host star can be determined. For example, the geometric measurements of the planet radius $R_P$ and orbital radius a are in units of the stellar radius. Differences in the way stellar properties are determined from star to star are important to take into account when looking at the sample of transiting planets as a whole. *Torres et al.* (2008) and *Southworth* (2008) present subsamples of transiting planets for which the stellar parameters have been obtained in a uniform way.

For radial velocity searches, the intrinsic stellar radial velocity variability, or stellar jitter, provides an important source of noise. Stellar jitter is believed to correlate with magnetic activity, and can arise from a few different sources. For example, a magnetic spot that covers a fraction $f_{spot}$ of the area of a rotating star would give a shift in the observed line centroids of roughly $f_{spot}$ v sin i $\approx$ 10 m/s ($f_{spot}$/0.01) (v sin i/1 km/s) (*Saar and Donahue,* 1997) (to set the scale, note that $2\pi R_\odot$/10 d = 5 km/s). An old inactive star like the Sun has $f_{spot} \sim 10^{-3}$; active stars can have $f_{spot}$ of several percent. Other sources of stellar jitter are spatial variations in convective velocities across the surface, changes in line profiles with time, or stellar oscillations. The effect of stellar jitter can be seen in the lower panel of Fig. 1, in which the residuals have a scatter $\approx$6 m $s^{-1}$, greater than the measurement errors, but consistent with the predicted jitter for this star. [For F, G, and K dwarfs in radial velocity searches, the level of jitter has been empirically calibrated (*Wright et al.,* 2004; *Wright,* 2005; see also *Saar and Donahue,* 1997; *Saar et al.,* 1998) in terms of observables absolute magnitude $M_V$, color B–V, and activity level S. The quantity S measures the flux in Ca H and K emission lines (an indicator of magnetic activity) relative to the neighboring continuum.] As well as hindering detection as an extra noise source, time-dependent jitter can also mimic planetary orbits. Timescales range from the stellar rotation period (~10 d) to magnetic cycle timescales of ~10 yr. Monitoring magnetic activity indicators over time is important to check that a planetary signal is in fact real [e.g., see *Queloz* (2001) for a study of the magnetically active star HD 166435].

As well as individual stellar properties, the properties of the sample of stars surveyed is important to understand when trying to draw conclusions about the planet population. For example, the selection of stars in the Keck Planet Search is described by *Wright et al.* (2004) and *Marcy et al.* (2005a). The $\approx$975 stars are selected to be chromospherically quiet, to lie at most 3 magnitudes above the main sequence (thereby excluding giant stars), to have no companion within 2", a color selection B–V > 0.55, and are magnitude limited to V = 8. As discussed by *Marcy et al.* (2005a), the fact that the survey is magnitude limited introduces some interesting biases. The first is a Malmquist bias, that brighter stars such as subgiants can be seen to a larger distance, and so they are overrepresented in the sample. Second, there is a metallicity bias introduced by the fact that at a fixed B–V color, metal-rich stars are brighter than metal-poor stars (e.g., see *Santos et al.,* 2004; also see discussion in *Gould et al.,* 2006a). This comes about because metal-rich stars are redder than metal-poor stars of the same mass, due to increased line blanketing. Therefore in a given B–V bin, there are more-massive, brighter, metal-rich stars, and less-massive, fainter, metal-poor stars. Because the sample is magnitude-limited, metal-rich stars are seen out to a larger volume and therefore overrepresented. The same effect leads to a correlation between stellar metallicity and mass for stars in the sample (*Fischer and Valenti,* 2005).

A second example is the CORALIE survey of 1650 dwarfs (*Udry et al.,* 2000), which is selected using Hipparcos parallaxes to be volume limited (d < 50 pc) rather than magnitude limited, in which case the Malmquist biases are absent. However, the fainter cool dwarfs are removed from the sample by implementing a color-dependent distance cutoff for K dwarfs, and so the search volume depends on the B–V bin. As a final example, the N2K consortium deliberately targets metal-rich stars in a search for closely orbiting (P $\lesssim$ 14 d) planets (*Fischer et al.,* 2005). It is important to keep these different sample biases in mind when comparing the results of different surveys, or looking at differences between the planet populations around stars with different metallicities or masses.

### 2.6. Exploring Parameter Space: Fitting and Uncertainties

We now turn to techniques for exploring parameter space to find the best-fitting orbital solution, and to assess the errors in the fitted parameters. Consider fitting Keplerian curves (equation (3)) to radial velocity measurements. A common approach to finding the best-fitting solution is to use a Levenberg-Marquardt algorithm (e.g., *Press et al.,* 1992). This is an efficient algorithm that marches downhill from a starting guess to the $\chi^2$ minimum. The difficulty is that in general there isn't a single $\chi^2$ minimum in parameter space, but rather a complicated $\chi^2$ surface with many local minima. For example, the many peaks in periodogram power in Fig. 2 reflect the many local minima in $\chi^2$ as a function of period for a circular orbit fit. For nonzero eccentricities, there are also many peaks as a function of the phase (time of periastron $t_P$), since there can be many values of $t_P$ for which the radial velocity peak at periastron passage intercepts the sparsely spaced data points. The problem becomes more complicated for multiple planet systems since

each additional planet adds an extra five parameters to the model. A grid search through parameter space with a resolution of 100 in each direction would require an additional factor of $10^{10}$ evaluations for each additional planet added.

Often, Levenberg-Marquardt is used with many different starting points in parameter space to determine which leads to the lowest value of $\chi^2$. The LS periodogram can be used to give initial guesses for the orbital period. For multiple planet systems, the data can be analyzed by subtracting successive planet orbits from the data and analyzing the residuals for further companions, before carrying out a full multiple-planet fit.

A related question is determining the uncertainties in fitted parameters. The probability distribution of certain parameters such as eccentricity is not necessarily Gaussian (e.g., *Shen and Turner,* 2008), so that the usual estimate of uncertainty from the covariance matrix, roughly $\delta_a \approx (\partial^2\chi^2/\partial a^2)^{-1/2}$ for parameter a (e.g., *Press et al.,* 1992), does not give a good measure of the error.

A good framework for systematically calculating the probability distributions of the different parameters of the model is a Bayesian approach. [We only have room here to give the basic idea. For a short but thorough introduction to the subject that will get you up and calculating quickly, we recommend the book by *Sivia* (1996).] The goal is to calculate the probability distribution of model parameters given the data, Prob($\vec{a}$|data), where, for example, $\vec{a} = (V_0, P, K, e, \omega, t_P)$ for a single Keplerian orbit. Bayes' theorem

$$\mathrm{Prob}\left(\vec{a}\,\middle|\,\mathrm{data}\right) \propto \mathrm{Prob}\left(\vec{a}\right)\mathrm{Prob}\left(\mathrm{data}\,\middle|\,\vec{a}\right) \tag{12}$$

gives this probability in terms of something we can calculate directly, the probability of obtaining the data given the parameters $\vec{a}$, Prob(data|$\vec{a}$). If the noise is Gaussian and uncorrelated from observation to observation, the probability of obtaining a set of measurements $\{t_i, v_i, \sigma_i\}$ if the underlying velocities are $V(t_i)$ is Prob(data|$\vec{a}$) =

$$\begin{aligned}\mathrm{Prob}\left(\mathrm{data}\,\middle|\,\vec{a}\right) &= \prod_{i=1}^{N}\frac{\exp\left(-\left(v_i - V\left(t_i\right)\right)^2\Big/\sigma_i^2\right)}{\sqrt{2\pi\sigma_i^2}} \\ &= \frac{1}{\sqrt{2\pi}\prod\sigma_i}\exp\left(-\frac{1}{2}\sum_{i=1}^{N}\left(\frac{v_i - V\left(t_i\right)}{\sigma_i}\right)^2\right) \\ &= \frac{1}{\sqrt{2\pi}\prod\sigma_i}\exp\left(-\frac{\chi^2\left(\vec{a}\right)}{2}\right)\end{aligned} \tag{13}$$

where in the last step, we see that the probability is related in a simple way to the $\chi^2$ for that choice of parameters. The remaining factor in equation (12) is the prior probability of the parameters $\vec{a}$, Prob($\vec{a}$), which allows us to specify any prior information about the parameters.

Equations (12) and (13) give the joint probability densities for all parameters $\vec{a}$. However, often we are interested in the probability distribution of some subset of the parameters of the model. For example, consider circular orbits, for which the model has four parameters $\vec{a} = (P, K, \phi, V_0)$,

$$V\left(t_i\right) = V_0 + K\sin\left(\frac{2\pi t_i}{P} + \phi\right) \tag{14}$$

Often we are interested in the period P and velocity amplitude K of the orbit (since they relate to physical properties of the system), but not $\phi$ and $V_0$. The probability distribution of P and K can be obtained by marginalizing, or integrating over the uninteresting parameters

$$\mathrm{Prob}\left(K, P\,\middle|\,\mathrm{data}\right) = \int dV_0 d\phi\ \mathrm{Prob}\left(K, P, \phi, V_0\,\middle|\,\mathrm{data}\right) \tag{15}$$

A good way to think of this integral is as a weighted average of the probability over the parameters $V_0$ and $\phi$. Therefore, all possible values of $\phi$ and $V_0$ are taken into account, in contrast to traditional $\chi^2$ minimization, in which one would find the best fit $V_0$ and $\phi$ at each K and P. For eccentric orbits, similar integrals are needed, but over different choices of the Keplerian parameters.

There has been a lot of work recently on Markov Chain Monte Carlo (MCMC) techniques as a method for evaluating marginalization integrals for Keplerian fits to radial velocity data (e.g., *Ford,* 2006; *Gregory,* 2007). The basic algorithm is very simple. One starts off at some point in the parameter space $\vec{a}$, and then proposes jumping to a new location $\vec{a}' = \vec{a} + \delta\vec{a}$. If $\chi^2$ for the new set of parameters at $\vec{a}'$ is lower than at $\vec{a}$, then we accept the jump. In this way, we always accept a move to a better-fitting model. If the $\chi^2$ is larger at the new location, we accept the jump with a probability $\exp(-\Delta\chi^2/2)$. Every jump that is accepted generates a new "link" in the chain. After following this simple algorithm many times (typically at least several thousands of accepted jumps are required), we end up with a sample of points in parameter space. The amazing property of this algorithm is that the density of points in parameter space is proportional to the probability density in equation (12). [For this to be the case, certain rules must be followed, for example, the jumps we make in parameter space $\delta\vec{a}$ must have certain properties, e.g., the probability to jump from point A to point B in parameter space should be the same as B to A, and the jumps should be uniform in the prior distribution. The size of the jumps is generally adjusted so that the jump is accepted ≈25% of the time. The choice of parameters in which to make jumps can be tuned to make this efficient (*Ford,* 2006).] Marginalization is then trivial. Say we generate a series of points in the parameter space (P, K, $\phi$, $V_0$) for a circular orbit; to find Prob(P), we simply plot a histogram of the P values of all the points that were visited.

As an example of the kind of constraints that are obtained, Fig. 4b shows the joint probability distribution of P and e derived from the radial velocity data for the star

HD 72659 from *Butler et al.* (2003). The radial velocity data for this star (Fig. 4a) show clear long-period variations. However, because the orbit has not yet closed, a range of orbital periods are possible, with longer periods corresponding to a larger eccentricity. This plot is based on Fig. 4 of *Ford* (2005), but calculated by direct integration (*Cumming and Dragomir,* 2010) rather than using a MCMC technique.

There are alternative methods for investigating the uncertainties in derived parameters. *Marcy et al.* (2005a) calculate uncertainties by a Monte Carlo method in which they make a large number of radial velocity curves consisting of the best-fitting model plus "noise" drawn from the residuals to the best fit. The resulting distribution of measured parameters gives an estimate of the uncertainties.

There are also many other methods for searching complicated parameter spaces that have been applied in the literature. For example, *Lovis et al.* (2006) use a genetic algorithm in which sets of parameter values are bred (along with occasional mutations) such that only the fittest (the best-fitting models) survive. *Gregory* (2007) has developed a parallel tempering technique in which MCMC chains with different temperatures are run simultaneously, and exchange information. A "hot" chain, which has artificially increased error bars, explores the broad parameter space, jumping large distances and exploring many local minima; a "cold" chain takes small steps locally and explores the local minima. This technique has been applied with success to a number of multiple-planet systems.

In some cases, more than one method may be applied. A good example is fitting microlensing lightcurves. The $\chi^2$ surface in that problem is extremely complex. Small changes in parameters that affect the location of caustics can lead to dramatic changes in the lightcurve and therefore $\chi^2$. This makes it nontrivial to explore the parameter space and find the best-fitting solutions. *Gould et al.* (2006b) adopt a brute force scan of parameter space in some nonlinear parameters, and use a minimization algorithm for others. *Bennett et al.* (2008) use MCMC to explore the complex parameter space.

A different application of Bayesian techniques is in microlensing and transit searches, when a statistical model of the stellar population must be considered. For example, in microlensing events, constraints on the lens mass and distance can be derived using as input priors on the distances and masses of lenses and sources [see *Beaulieu et al.* (2006) for an application of this approach to one of the detected microlensing planets].

### 2.7. Completeness Corrections: Determining Planet Occurrence Rates

Understanding selection effects is crucial for population studies, as it allows us to correct for incompleteness, i.e., to determine the effective size of the stellar sample. For each set of planet parameters, e.g., mass, orbital period, and eccentricity, the idea is to work out which stars in the survey have data that would allow a planet with those parameters to be detected. Only those stars are useful in constraining the population. To see this, consider a simple example. Imagine surveying 100 stars, detecting 10 planets, and being able to rule out the presence of planets around the remaining 90 stars. Then the best estimate of the fraction of stars with planets is 10%, since we detected 10 planets out of 100 stars surveyed. However, it could be that the data for 80 of those stars are not good enough to say whether a planet is present or not. In that case, only 20 stars have data good enough to constrain the presence of a planet, and since planets were detected around 10 of these, the best estimate of the planet fraction is 50%. In practice, the ability to detect or rule out the presence of a planet depends on the orbital parameters, and so when doing this, we must account for every combination of orbital parameters of interest.

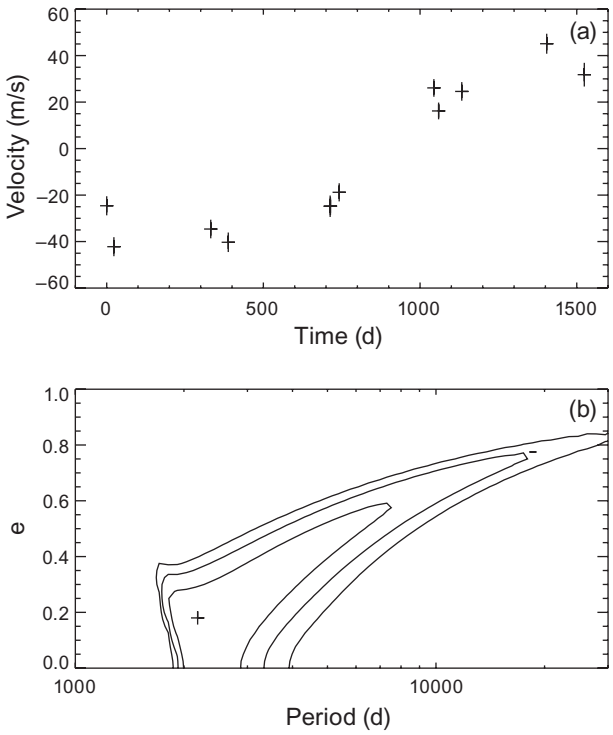



**Fig. 4.** Constraints on the period and eccentricity of the companion to HD 72659 (data from *Butler et al.,* 2003), for a companion with mass less than 10 $M_{Jup}$. The best fit values given in that paper are e = 0.18 and P = 5.98 yr, shown by the cross. From *Cumming and Dragomir* (2010), based on Fig. 4 of *Ford* (2005).

We can write this argument down mathematically using a maximum likelihood method. Assume that a survey of $N_\star$ stars detects $N_p$ planets, and rules out planetary companions for the remaining $N_\star$–$N_p$ stars. If the fraction of stars with planets is f, then the probability that a given star has zero or one planet can be written $e^{-f}$ or $f e^{-f}$ (considering each star as an independent Poisson trial). The total likelihood is the product of these probabilities for all stars

$$L = \left(fe^{-f}\right)^{N_d} \left(e^{-f}\right)^{N_\star - N_d} = f^{N_d} e^{-N_\star f} \quad (16)$$

Maximizing the likelihood with respect to f by setting $\partial L/\partial f = 0$ gives $f = N_d/N_\star$, exactly what we expect: The fraction of stars with planets is given by the number of detections divided by the total number of stars. We can introduce selection effects into this calculation by including the fact that for each star i there is some probability $p_i$ of detecting a planet. Then

$$\begin{aligned} L &= \prod_{i=1}^{N_d} p_i f e^{-p_i f} \prod_{j=1}^{N_\star - N_d} e^{-p_j f} \\ &= f^{N_d} \left(\prod_{i=1}^{N_d} p_i\right) \exp\left(-f \sum_{k=1}^{N_\star} p_k\right) \end{aligned} \quad (17)$$

where the sum labeled i is over stars with a detected planet, the sum labeled j is over stars with no detection, and the sum labeled k is over all stars. The value of f that maximizes the likelihood L is now $f = N_d/N_{eff}$ where the effective number of stars $N_{eff} = \sum_{i=1}^{N_\star} p_i$. This is the mathematical statement of our previous argument. We are guaranteed to detect a planet around stars with $p_i = 1$, and so these stars contribute fully to $N_{eff}$; stars with lower quality or fewer data points that have small $p_i$ do not contribute as much to $N_{eff}$. In practice, the detection probability depends on the orbital parameters of the planet, and varies from star to star depending on the number of data points, measurement errors, and stellar properties.

There have been several detailed calculations of detection probabilities and thresholds for radial velocity surveys (*Walker et al.,* 1995; *Cumming et al.,* 1999, 2008; *Endl et al.,* 2002; *Naef et al.,* 2005; *Wittenmyer et al.,* 2006; *O'Toole et al.,* 2009). These papers essentially use the same method for determining detection probabilities or upper limits on the planet mass at a given orbital period. The idea is to use a Monte Carlo method to find the value of velocity amplitude K as a function of period that results in a particular detection efficiency. The technique for modeling the noise differs in different studies, ranging from assuming Gaussian noise with amplitude set by the observed variability, to sampling directly from the observed residuals to the best-fit orbit.

Early investigations of the planet mass and orbital period distributions did not take these detailed star-by-star mass limits into account (*Tabachnik and Tremaine,* 2002; *Udry et al.,* 2003; *Lineweaver and Grether,* 2003), but recent work has included mass limits calculated for each star. *Naef et al.* (2005) estimate planet occurrence rates in the ELODIE Planet Search, *Cumming et al.* (2008) use their results from the Keck Planet Search to constrain the minimum mass and orbital period distribution of planets, and *O'Toole et al.* (2009) use observations from the Anglo-Australian Planet Search to focus on the frequency and mass distribution of low-mass planets. *Cumming et al.* (2008) use a method similar to *Tabachnik and Tremaine* (2002), but including the upper limits derived separately for each star, rather than assuming a constant velocity limit for the whole survey. The method is a generalization of equations (16) and (17) to bins in the mass-period plane [based on the method of *Avni et al.* (1980) for one-dimensional distributions with upper limits].

Calculations of the incidence of planets or limits on the planet fraction in transit surveys include *Gilliland et al.* (2000), *Brown* (2003), *Mochejska et al.* (2005), *Bramich and Horne* (2006), *Gould et al.* (2006a), *Beatty and Gaudi* (2008), and *Weldrake et al.* (2008). Monte Carlo methods are used in which fake planetary transits are injected into the observations, and subject to the same analysis procedure as the real data. As discussed in section 2.3, the effective size of the stellar sample in a transit survey depends sensitively on the planet radius and orbit. Determining this number accurately requires a model of galactic structure, the stellar mass function, and accounting for the effects of extinction. For example, *Gould et al.* (2006a) give a detailed discussion of the selection effects and stellar sample size for the OGLE-III transit surveys, and derive the frequencies of hot and very hot Jupiters.

Examples of calculations of detection sensitivities and occurrence rates in microlensing surveys are *Gaudi et al.* (2002) and *Snodgrass et al.* (2004). Exactly as we discuss above, the challenge is to identify a subset of microlensing events for which planetary systems could have been identified, e.g., free from nonplanetary features and with good enough signal-to-noise data. Detection sensitivities are then calculated as a function of planet-lens mass ratio and separation [see Fig. 8 of *Gaudi et al.* (2002) for examples] and used to infer occurrence rates of planets.

## 3. PROPERTIES OF OBSERVED PLANETS

In this section, we discuss the statistical properties of observed planets, highlighting the major features of the population discovered so far, and some of the implications for planet formation theories. As we go, we will point out the places where the selection effects discussed in section 2 play a role.

### 3.1. Catalogs of Exoplanet Properties

There are a few different catalogs of exoplanet properties. The Extrasolar Planets Encyclopedia website (*http://exoplanet.eu*), maintained by J. Schneider of the Paris Observatory, gives properties of all planets announced to date, organized by discovery method. In this section, we use data from the Extrasolar Planets Encyclopedia for planets discovered by radial velocity measurements, transits, or microlensing, accessed on September 1, 2010, a total of 469 planets with 101 transiting planets and 10 microlensing planets. *Butler et al.* (2006a) published a catalog of 172 exoplanets within 200 pc with $M \sin i < 24\ M_{Jup}$. An updated version of this catalog is available (*http://exoplanets.org/exotable/exoTable.html*), which at the time of writing lists 372 planets. This catalog includes

radial velocity data for those stars with planets observed at the Keck, Lick, or Anglo-Australian telescopes. When showing plots of planet properties in this section, we use data from the Extrasolar Planets Encyclopedia, but our conclusions would not be changed had we used the *Butler et al.* (2006a) sample.

For transiting planets, there are several other sources of information. A website maintained by F. Pont (*http://www.inscience.ch/transits/*) gives properties and references for all known transiting planets. As we mentioned in section 2.4, *Torres et al.* (2008) and *Southworth* (2008) have published data for subsamples of the transiting planets in which the data have been analyzed in a uniform way.

There is also a lot of information available about planet host stars. For example, the Spectroscopic Properties of Cool Stars (SPOCS) survey (*Valenti and Fischer,* 2005; *Takeda et al.,* 2007) provides a catalog of stellar properties for 1040 F, G, and K stars that have radial velocity measurements as part of the Keck, Lick, or Anglo-Australian Planet Search programs.

A new resource is the NASA/IPAC/NExScI Star and Exoplanet Database (NStED) available at *http://nsted.ipac.caltech.edu/.* Designed to support NASA's planet-finding and characterization activities, this is a comprehensive database of stellar and planet properties, including radial velocity data and lightcurves.

### 3.2. Fraction of Stars with Planets

The fraction of F, G, and K stars with giant planets is fairly well understood, out to orbital periods of several years that can be probed with the current data, with good agreement between different surveys. The simplest estimate that can be made is to divide the number of detected planets by the number of stars surveyed. *Marcy et al.* (2005b) did this for the Lick, Keck, and Anglo-Australian planet surveys. They found that 16/1330 = 1.2 ± 0.3% of stars have a hot Jupiter ($a < 0.1$ AU), 88/1330 = 6.6% of stars have a gas giant within 5 AU, and extrapolating the observed orbital period distribution to longer periods, 12% of F, G, and K stars have a gas giant within 20 AU. *Udry and Santos* (2007) give the corresponding numbers for the CORALIE Planet Search, finding 0.8% of stars with hot Jupiters, and 63/1120 = 5.6% of stars having giant planets within 5 AU.

These estimates are lower limits to the true planet fraction for these stellar samples, because they do not account for selection effects. Including completeness corrections for the Keck Planet Search, *Cumming et al.* (2008) concluded that 11 ± 1.7% of stars harbor gas giants within 5 AU, or 17% within 20 AU (Saturn mass and larger), using a flat extrapolation of the period distribution observed for $a < 3$ AU. *Naef et al.* (2005) find 7.3 ± 1.5% of stars have gas giants ($M > 0.5\ M_{Jup}$) within 5 AU from the ELODIE survey. (Taking into account the different lower mass limits, these numbers are consistent at the $1\sigma$ level.)

It is important to be aware of the selection effects in the stellar sample when discussing the planet occurrence rate. The planet fractions we mentioned previously are for the particular samples of F, G, and K stars studied, which are biased selections of stars. For example, *Beatty and Gaudi* (2008) argue that a comparison between the 0.8 ± 0.3% hot Jupiter occurrence rate reported by *Udry and Santos* (2007) for the volume-limited CORALIE survey and the 1.5 ± 0.6% reported by *Cumming et al.* (2008) for the magnitude-limited Keck sample should take into account metallicity bias in the magnitude-limited sample. Accurately knowing the hot Jupiter occurrence rate is necessary to understand the yields of transit searches (*Beatty and Gaudi,* 2008). *Gould et al.* (2006a) calculated a hot Jupiter frequency of $0.3^{+0.4}_{-0.2}\%$ ($P < 5$ days) from the OGLE-III transit survey results, which appears consistent with the radial velocity numbers (see also *Gaudi et al.,* 2005).

At larger separations, beyond the reach of transit and current radial velocity surveys, there are indications from microlensing that low-mass planets are common. *Gould et al.* (2006b) detected a Neptune-mass 13 $M_\oplus$ planetary companion to a 0.5 $M_\odot$ star with orbital separation 2.7 AU. They were able to infer that such cool Neptune planets (~10 $M_\oplus$ in orbits beyond the ice line, ~1–4 AU) are common based on the detection of two planets from four events sensitive to them, with a derived frequency and 90% upper and lower limits of $0.38^{+0.31}_{-0.22}$ (see also *Beaulieu et al.,* 2006). This is significantly larger than the frequency of Jupiter-mass planets at similar distances from solar-mass stars, found to be 2–5% (for Jupiter mass and above between 2–5 AU) by *Cumming et al.* (2008).

### 3.3. Mass and Orbital Period Distributions

The mass-orbital period distribution is shown in Fig. 5. The two different symbols distinguish the transiting and radial velocity planets. The gray solid curves show the expected 50% and 99% detection thresholds for a radial velocity survey of eight years with $\sigma = 3$ m/s for solar mass stars. These curves match the lower envelope of detected planets quite well. Figure 6 shows the mass-semimajor axis distribution, now including the planets discovered by microlensing.

Perhaps the most striking feature is that the distribution of orbital periods is bimodal, with a population of "hot" planets with orbital periods close to 3 d, and a population of long-period planets with periods of ≥200 d. Figure 7 shows the distribution of semimajor axes. *Udry et al.* (2003) and *Jones et al.* (2003) emphasized this "period valley" in the period distribution.

The rising number of gas giants with orbital period has been fit with a power-law distribution. *Cumming et al.* (2008) fit the orbital period and M sin i distribution from the Keck Planet Search jointly with a power law, and found that $dn \propto M^\alpha P^\beta d\ln M\, d\ln P$ with $\alpha = -0.31 \pm 0.2$ and $\beta = 0.26 \pm 0.1$ for a range of masses and periods M sin i = 0.3–10 $M_{Jup}$ and P = 2–2000 d [similar values for $\alpha$ and $\beta$ were found by *Tabachnik and Tremaine* (2002), who included detections from several surveys but with a simplified treatment of detection thresholds]. In terms of semimajor axis, this corresponds to $dn/d\ln a \propto a^{0.39}$, assuming equal

stellar masses. *Cumming et al.* (2008) pointed out that an equally good description of the distribution is a step in the number of planets per decade beyond orbital periods of ≈300 d or semimajor axis ≈ 1 AU. For the Keck sample of F, G, and K dwarfs, they found the fraction of stars with planets per decade in orbital period dn/d $\log_{10}$ P = 6.5 ± 1.4% at long periods, compared with dn/d $\log_{10}$ P = 1.3 ± 0.4% at short periods.

The detection curves in Fig. 5 rise steeply at long periods, as discussed in section 2.2, but the data show a much sharper cutoff at P ≈ 2000–3000 d. This is because even though a companion can be detected statistically, it cannot be confirmed as a planet, as opposed to a low-mass star, for example, until a full orbit has been observed. There is a lot of information in the radial velocity datasets about orbits at long periods. Analysis of these long-period candidate planets is beginning to appear in the literature (e.g., *Fischer et al.,* 2001; *Wright et al.,* 2007; *Patel et al.,* 2007).

The group of hot Jupiters at short orbital periods shows a "pileup" in their orbital periods at P ≈ 3 d, with a tail that stretches out to longer periods. Figure 8 shows the distribution of radial velocity and transiting planets at short periods. There are planets with orbits inside 3 d, particularly those found by transit surveys, which survey a larger number of stars and whose sensitivity depends strongly on orbital period. Initially there were questions about why the "very hot" planets found by transit surveys were not being found by radial velocity surveys, but in fact the likely explanation is that they are rare (*Gaudi et al.,* 2005; *Gould et al.,* 2006a). Interestingly, the distribution of transiting planets is showing increasing numbers at 3–4-d orbital periods.

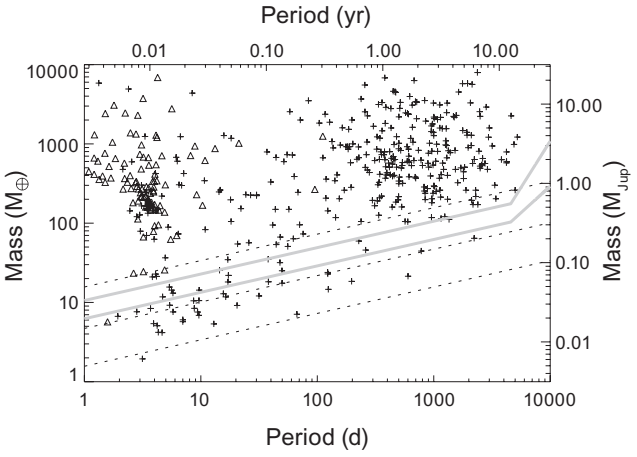



**Fig. 5.** Mass and orbital periods of known exoplanets. Crosses show planets detected by radial velocity only; triangles show transiting planets. The dotted lines correspond to stellar radial velocity amplitudes of K = 1, 3, and 10 m/s for a solar mass star. The gray curves give the 50% and 99% detection threshold for a survey of 8 yr with σ = 3 m/s, assuming solar-mass stars (see equation (7)).

*Marcy et al.* (2005b, following *Marcy and Butler,* 2000) fit a power law to the observed minimum mass distribution, finding dn/d$M_p$ sin i ∝ 1/($M_p$ sin i)$^{1.05}$. In other words, the planets are distributed almost uniformly in log mass, rising slightly more steeply than that toward lower masses. The power law fits of *Tabachnik and Tremaine* (2002) and *Cumming et al.* (2008) given earlier rise more quickly to lower masses than the *Marcy et al.* (2005b) fit, which is as expected since the completeness corrections account for undetected low-mass planets.

The observed $M_p$ sin i distribution of planets is shown in Fig. 9, for all observed planets and for those in close orbits with P < 100 d. For orbital periods beyond P ≈ 1 yr, the low-velocity amplitude makes it difficult to constrain the mass distribution below a Saturn mass. However, in close orbits, particularly the "hot" orbits within 10-d orbital periods, Fig. 5 shows that there are now many examples of Neptune-mass planets and below (mass range ~10 $M_\oplus$). Indeed, *Lovis et al.* (2006) discovered a three-Neptune planetary system around HD 69830. The long-term radial velocity precision of only ≈1 m/s of the HARPS spectrometer at La Silla Observatory allowed the detection of these signals from this 0.86 $M_\odot$ star, each of which has an amplitude of less than 3 m/s. Many, but not all, of the known hot Neptunes orbit M dwarfs, which helps with their detection since a less-massive star has a larger-velocity amplitude for a given planet mass (in equation (8) we scale to a typical $M_{dwarf}$ mass).

The distribution of planet periods and masses has been contrasted with that of low-mass stars and stellar binaries. The roughly dn/dM ∝ 1/M mass distribution is significant because it clearly separates the distribution of planet masses from that of low-mass stars. These objects are not an extension of the low-mass stars/brown dwarfs to lower masses. [Indeed, there is a paucity of brown dwarfs at semimajor axes <3 AU, the so-called "brown dwarf desert" (*Marcy and Benitz,* 1989)]. This has been used to argue for differ-

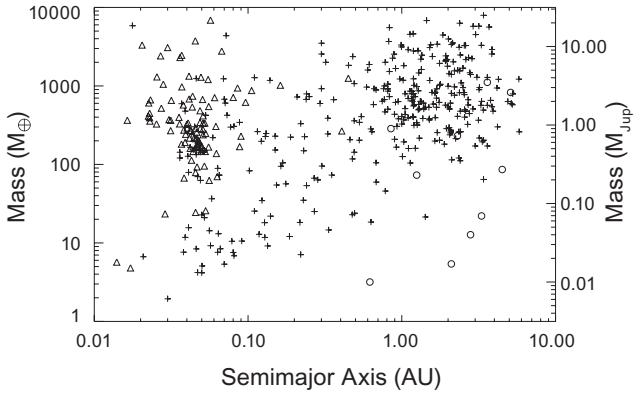



**Fig. 6.** Mass and semimajor axis of known exoplanets. Crosses show planets detected by radial velocity only; triangles show transiting planets; circles show planets discovered by microlensing.

ent formation mechanisms for brown dwarfs and planets. *Jones et al.* (2003) contrasted the distribution of planetary orbital periods with the stellar binary orbital period distribution, which shows no dip at intermediate orbital periods, being well fitted by a log-Gaussian (*Duquennoy and Mayor,* 1991), again consistent with the different formation mechanism of giant planets in a gas disk through which migration occurs.

In general, the distribution of planet masses and orbital periods can be described by core accretion models of planet formation including inward migration (e.g., *Ida and Lin,* 2004a,b, 2005, 2008a,b; *Armitage,* 2007; *Mordasini et al.,* 2009a,b). Models based on type II migration of gas giants predict a smooth increase in the number of gas giants with orbital radius, because of the increasing migration rate as the planet moves inward. In principle, therefore, how quickly the observed distribution increases to long periods can be used to determine the ratio of migration to formation/disk-depletion timescales. The most recent models by *Ida and Lin* (2008a,b) include type I migration, and propose that the upturn in planet frequency at ~1 AU corresponds to retention of solids near the ice line, which gives a preferred radius in the protoplanetary disk. *Mordasini et al.* (2009a,b) also find an upturn, although further out than observed by about a factor of 2, coming from two different evolutionary tracks: one in which cores migrate inward before reaching the condition for runaway accretion of gas, the other from cores that run away *in situ*. In addition, they predict a distribution of $\approx$20 $M_\oplus$ planets ranging in semimajor axis from ~0.1 to 5 AU, which is consistent so far with the Neptunes being found within a wide range of semimajor axis both by radial velocity surveys and microlensing.

### 3.4. Eccentricity Distribution

Exoplanets have a large range of eccentricities. This goes against what we might have expected in comparison with the solar system, in which the orbits are almost circular (e.g., Jupiter has an eccentricity of 0.05; Mercury's eccentricity is significantly larger than the other planets, with e = 0.21). The eccentricity distribution is shown in Fig. 10 for planets with orbital periods greater than 10 d, for which tidal damping of eccentricity is not expected to have occurred. The distribution is consistent with being flat at low eccentricity, continuing up to eccentricities beyond the median value of 0.25–0.3 and then drops sharply above e = 0.5. The reason for the drop at high eccentricity may be partly due to selection effects. Figure 11 illustrates this by comparing the K–e distribution against the detection curves calculated by *Cumming* (2004) (this is an updated version of Fig. 13 in that paper). The planets with large eccentricities all have large K values, K > 20 m/s for e $\geq$ 0.7. This suggests that the tail of eccentricities in Fig. 10 may be underestimated due to the finite sampling of radial velocity surveys, although a detailed analysis taking into account the selection effects on a star-by-star basis in different surveys is needed.

*Marcy et al.* (2005b) point out that the most massive planets have nonzero eccentricities. This is not due to a selection effect, since more massive planets should be easier to detect at all eccentricities than low-mass planets, yet only lower-mass planets are found with circular orbits. Figure 12 shows the eccentricity against mass for planets with P > 10 d (to avoid circularization effects). The median eccentricity for $M_P \sin i \geq 5\ M_{Jup}$ is clearly greater than the median eccentricity for $M < 5\ M_{Jup}$. This could indicate that the most massive planets have a different formation mechanism (*Ribas and Miralda-Escudé,* 2007; *Ford and Rasio,* 2008).

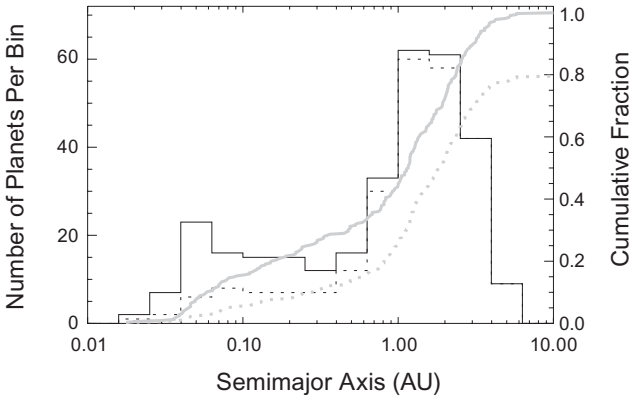



**Fig. 7.** The distribution of semimajor axes for nontransiting planets (radial velocity detection only). The solid curve is all planets; the dotted curves are for $M_P \sin i > 0.5\ M_{Jup}$. The gray curves show the cumulative distribution in each case. The cutoff at a $\approx$ 5 AU is due to the finite length of the radial velocity surveys (orbital periods ~10 yr and greater).

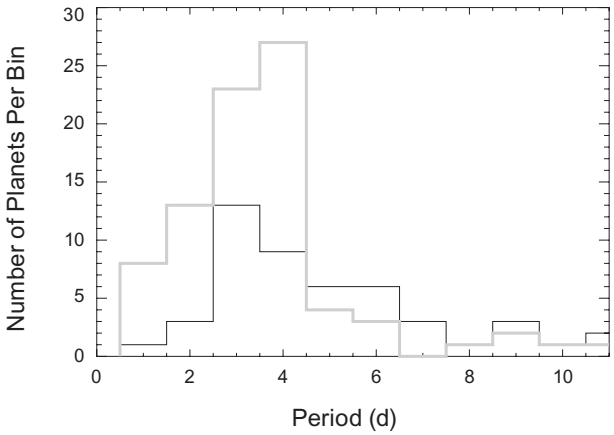



**Fig. 8.** Orbital periods of "hot" planets. The solid curve is for radial velocity detected planets; the gray curve is for transiting planets.

The eccentricity-semimajor axis distribution is shown in Fig. 13. The orbits of short-period planets are expected to be circularized by tides, and the eccentricity-semimajor axis distribution provides evidence for this. The circularization timescale for small e due to tides raised on the planet is

$$\tau_e = \frac{4Q}{63\Omega}\left(\frac{M_P}{M_\star}\right)\left(\frac{a}{R_P}\right)^5 \quad (18)$$

where the parameter Q describes the tidal dissipation inside the planet. For a Jupiter-mass and radius planet orbiting a solar-mass star, this is $\tau_e \approx 1$ G.y. $(Q/10^6)(a/0.05\ \mathrm{AU})^{13/2}$. The value of Q is uncertain; values somewhere between ~$10^5$ to $10^6$ are likely (*Wu,* 2003). Therefore we see that within a periastron distance on the order of 0.05 AU, the circularization time is smaller than 1 G.y., and it is reasonable to expect that the planet's orbit will have circularized over the lifetime of the star. In Fig. 13 the dotted curve shows a constant value of a(1−e) = 0.05 AU. Indeed, this roughly separates the occupied and unoccupied regions, suggesting that tidal dissipation sets the upper envelope of the distribution. This cannot be the only factor, however, as there is a significant group of planets with nonzero eccentricity at short orbital periods. In some cases, a yet-undetected additional companion may excite the eccentricity of the short-period planet through gravitational interactions.

Earlier, we discussed the differences in the mass and orbital period distributions of exoplanets compared to stellar binaries, and how these differences have been used to argue that the planets are a distinct population with a different formation mechanism. For the eccentricity distribution, the opposite is the case. *Stepinski and Black* (2001) pointed out that the period-eccentricity distribution for planets is very similar to that for binary stars, and suggested that this may result from a similar formation mechanism. More recently, *Halbwachs et al.* (2005) point out that there are differences, e.g., there are several long-period planets with almost circular orbits (analogous to our solar system), which is not the case in the stellar binaries.

The origin of the observed exoplanet eccentricities is still not understood. Since gas giants are believed to form in close to circular orbits in a gas disk, most ideas rely on dynamical interactions following formation of the planet to increase its eccentricity. Building on their earlier work, *Ford and Rasio* (2008) find that two-body scattering with a range of planet mass ratios can explain the observed eccentricity distribution of most planets, but that there should be a well-defined maximum eccentricity of ≈0.8 (for initially circular orbits of both planets), so that the most eccentric planets require some other mechanism to excite their eccentricity. *Jurić and Tremaine* (2008) simulate systems of several planets and find that the eccentricity distribution for a wide range of initial conditions relaxes to the form $dn \propto e \exp\left[-\frac{1}{2}(e/0.3)^2\right] de$. This distribution is shown as the dotted curve in Fig. 10, and matches the shape of the observed distribution well at larger eccentricities, but not for small eccentricities. Therefore, an extra mechanism is needed to make the population of planets on circular orbits (perhaps these are planets that formed in a disk and did not undergo gravitational scattering).

One problem with these models is that the distribution of semimajor axis tends to be smooth with increasing a, and does not match the factor of 5 observed increase in the number of planets orbiting beyond ~1 AU compared to within 1 AU (see Fig. 7 of *Jurić and Tremaine,* 2008). This increase is present even when planets with circular orbits are excluded (see Fig. 13). Other proposals for producing the observed eccentricities are due to the interaction of a

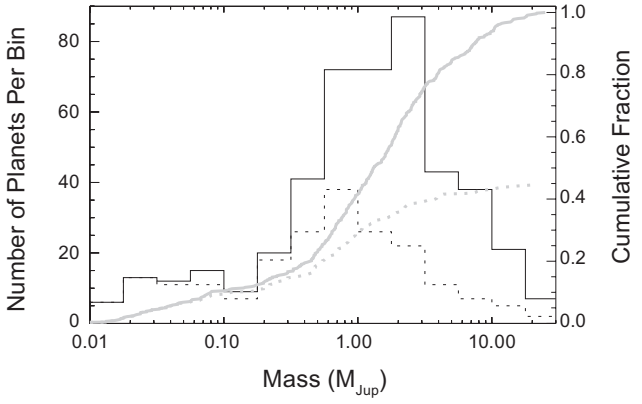



**Fig. 9.** Observed mass distribution of exoplanets. The gray curve shows the cumulative fraction. The dotted curves are for planets with P < 100 d.

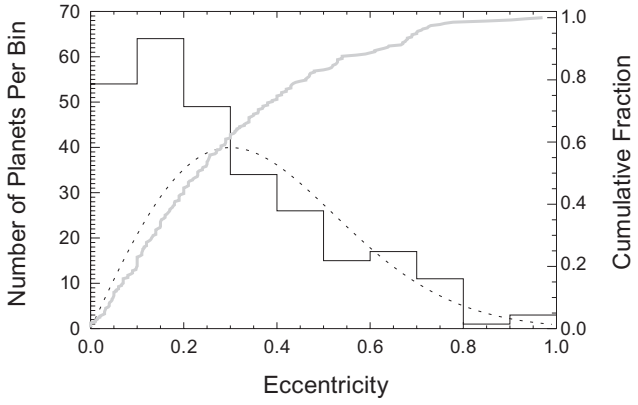



**Fig. 10.** Eccentricity distribution for planets with orbital periods greater than 10 d. The gray curve shows the cumulative distribution. The dotted curve shows the theoretical distribution from planet-planet scattering derived by *Jurić and Tremaine* (2008). The normalization has been adjusted to roughly fit the tail of the observed distribution.

migrating planet with the disk itself (*Goldreich and Sari,* 2003), and interactions with a passing star (*Zakamska and Tremaine,* 2004), or companion star (*Holman et al.,* 1997; *Wu and Murray,* 2003; *Fabrycky and Tremaine,* 2007; *Wu et al.,* 2007). The observed eccentricity distribution is likely due to a combination of these mechanisms.

### 3.5. Planet Incidence Versus Stellar Metallicity

It is now well established that the occurrence rate of giant planets increases dramatically with increasing metallicity of the host star. Evidence that this is the case emerged from the first planet discoveries. *Gonzalez* (1997) performed abundance analyses of four of the early planet discoveries, and noted that they were all metal rich relative to the Sun. This trend has been confirmed with the many planet detections since then: There are many more planets known around metal-rich than metal-poor stars. One might worry that this could be an effect of the stellar selection. We have already discussed the possible metallicity biases possible in stellar samples; e.g., metal-rich stars are overrepresented in a magnitude-limited sample with a cutoff in B–V (section 2.4). To eliminate these selection effects, *Santos et al.* (2004) and *Fischer and Valenti* (2005) measured metallicities in a uniform way for samples of stars with and without planets, and so could directly address the likelihood that a star in a particularly metallicity bin would have a planet.

Figures 14 and 15 show the results from *Fischer and Valenti* (2005), who determined metallicities and other stellar parameters for more than 1000 stars observed as part of the Keck, Lick, and Anglo-Australian Planet Searches. They identified a subset of 850 stars that they estimated had similar radial velocity detectability (more than 10 observations spanning four or more years). [You may also wonder whether the typical velocity precision would change with metallicity and could introduce a bias. In fact, the mean velocity precision seems to be constant for [Fe/H] in the range –0.5 to 0.5 (see *Valenti and Fischer,* 2008).] Figure 14 compares the metallicity distributions for the stars with and without planets, and Fig. 15 shows the fraction of stars with planets in each metallicity bin. *Fischer and Valenti* (2005) find that the probability that a star has a planet is well fit by the expression

$$\begin{aligned}\mathrm{Prob(planet)} &= 0.03\times 10^{2.0[\mathrm{Fe/H}]} \\ &= 0.03\left[\frac{N_{Fe}/N_H}{(N_{Fe}/N_H)_\odot}\right]^2 \end{aligned} \quad (19)$$

(shown by the gray curve in Fig. 15).

For hot Jupiters, another way to address the incidence of planets at low metallicities is transit searches of globular clusters. *Gilliland et al.* (2000) found that the frequency of hot Jupiters in the globular cluster 47 Tuc is at least an order of magnitude below that found in radial velocity surveys of stars in the solar neighborhood. They found no transiting hot Jupiters in their sample of ~34,000 main-sequence stars, whereas if the occurrence rate is 1%, 17 detections would have been expected. More recently, *Weldrake et al.* (2008) placed limits on the occurrence rate of planets at periods between 1 and 5 d in the globular cluster ω Cen. The majority of stars studied are members of the low-metallicity population of the cluster, which has [Fe/H] = −1.7. The combined rate of hot and very hot Jupiters that they find is approximately <1/600. We should note that both of these studies are for planets with radii larger than $R_p$ = 1.3 $R_{Jup}$ or 1.5 $R_{Jup}$ respectively, which decreases the expected rate compared to that for all hot Jupiters (see the distribution of transiting planet radii in Fig. 18).

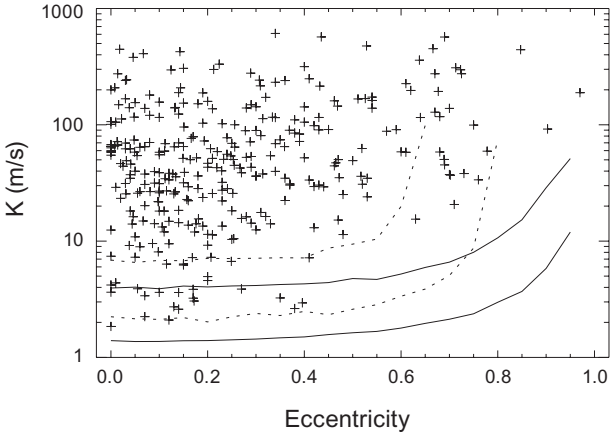



**Fig. 11.** Radial velocity amplitude against eccentricity. The curves show 50% (solid) and 99% (dotted) detection thresholds calculated by *Cumming* (2004), assuming 16 (top curve in each pair) and 39 (bottom curve in each pair) measurements, and σ = 5 m/s.

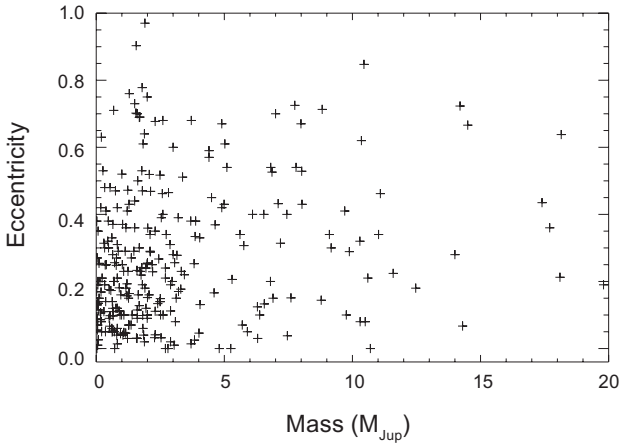



**Fig. 12.** Planet mass against eccentricity. More massive planets have a larger median eccentricity than low-mass planets.

The origin of the metallicity correlation could be either that planet formation or migration are enhanced around metal-rich stars, or that the metallicity of the star is enhanced by planet formation itself. Following the suggestion of *Lin et al.* (1996) that the hot Jupiters migrated inward along with the remains of the protoplanetary disk, *Gonzalez* (1997) proposed that the addition of heavy elements to the outer convection zone of the star would be enough to raise the star's metallicity by a significant amount. The thickness of the convection zone at the time of the accretion event, which depends on the mass of the star (more massive stars have thinner outer convective envelopes) and the timing of the accretion event (the mass of the convection zone decreases as the star settles onto the main sequence), is critical since the added material is mixed throughout the convection zone and therefore diluted. However, studies of the metallicities of stars as a function of their effective temperature, or in subgiants compared to main-sequence stars, do not show any clear evidence that the accretion of solids gives rise to the observed metallicity dependence of the planet occurrence rate (*Pinsonneault et al.,* 2001).

The alternative is that metallicity affects the migration process, which brings planets closer to their host star where they are easier to detect, or that metallicity directly affects planet formation. In the context of the core-accretion model for forming Jupiters, this would not be surprising. We might naively expect the probability of forming a planet to scale with $N^2$ (equation (19)) if planet formation involved binary collisions of solids. Core-accretion models by *Ida and Lin* (2004b) or *Mordasini et al.* (2009b) do indeed show an increasing abundance of planets with metallicity.

The behavior at low metallicity is still being debated. The planet abundance against [Fe/H] appears consistent with going to a constant at low values of [Fe/H] (Fig. 15), as suggested by *Santos et al.* (2004). However, *Cochran et al.* (2008) find that many of the stars with low [Fe/H] are members of the thick disk population rather than the thin disk, and have a larger fraction of α-elements than thin disk stars. This means that [Fe/H] actually underestimates their total metallicity if measured as the total mass fraction of elements heavier than helium, $Z/Z_\odot$. It may be that the planet fraction does indeed go to zero below some value of $Z/Z_\odot$, as expected as the raw materials for planet building become rarer. There is clearly a lot of interesting work to do studying the relationships between stellar abundances and planet occurrence rates and properties.

### 3.6. Planet Incidence Versus Stellar Mass

Stellar mass is another parameter that could affect the likelihood of forming planets, and their masses and orbital properties. The distribution of masses and metallicities of known planet-bearing stars is shown in Fig. 16. The range of stellar masses goes from ≈0.3 $M_\odot$ (main-sequence M dwarfs) to almost 3 $M_\odot$ (evolved stars). However, most planet detections are concentrated in the stellar mass range close to 1 $M_\odot$ (mostly G and K spectral types). This is because radial velocity surveys have traditionally focused on this narrow range of stellar masses, where the stellar spectra are most amenable to radial velocity measurements. At higher stellar masses, main-sequence stars earlier than late F-type are generally not included in radial velocity surveys, because of the low radial velocity precision that can be achieved (~50–100 m/s). The reason for this is that these stars have hot atmospheres giving fewer spectral lines, and are rapid rotators giving high levels of jitter and rotationally broadened lines. At lower masses, because M dwarfs are faint they require more observing time to achieve the same precision as brighter solar-type stars.

Observational programs have begun to enlarge the sample

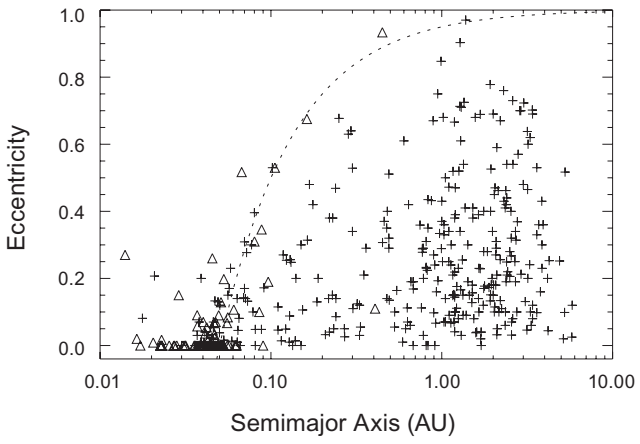



**Fig. 13.** Semimajor axis-eccentricity distribution. The dotted curve shows the curve of constant pericentric distance a(1−e) = 0.05 AU. Crosses show planets detected by radial velocity only; triangles show transiting planets.

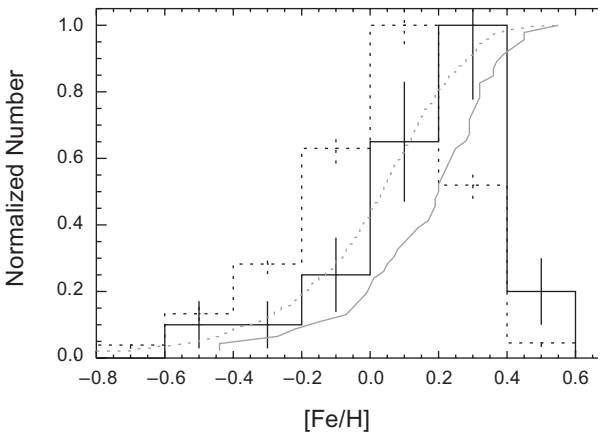



**Fig. 14.** Distribution of stellar metallicities of stars with uniform radial velocity detectability (dotted curves), and those with planets (solid curves). The histograms have been normalized so that the peak value is 1. Data are from *Fischer and Valenti* (2005) and *Valenti and Fischer* (2005). There are 850 stars in total, of which 47 have planets.

of low- and high-mass stars surveyed for planets, and evidence is emerging that the planet occurrence rate and even the distribution of orbital properties depends on stellar mass. One potential selection effect that must be considered carefully is the mass-metallicity correlation introduced into a stellar sample by making a cut in B–V (section 2.4).

At the low-mass end, there is now good evidence for a low occurrence rate of giant planets around M dwarfs. For example, *Butler et al.* (2006b) estimated the occurrence rate of giant planets within 2.5 AU as 2/114 = 1.8 ± 1.2% for planet masses ≥0.4 $M_{Jup}$. The equivalent occurrence rate for F, G, and K stars is between 6% and 7% (*Cumming et al.,* 2008). Taking into account differences in selection effects between the F, G, and K stars and the M dwarfs, *Cumming et al.* (2008) found that M dwarfs were 3–10 times less likely to harbor a gas giant planet with an orbital period less than 2000 d (set by the duration of the Keck survey).

A low occurrence rate is predicted by core accretion models for planet formation, which find that Jupiter-mass planets should be rare around M dwarfs, with the mass function of planets shifted toward lower masses (*Laughlin et al.,* 2004; *Ida and Lin,* 2005; *Kennedy and Kenyon,* 2008). *Kennedy and Kenyon* (2008) assume that the mass of the protoplanetary disk scales $\propto M_\star$, and include a detailed calculation of the position of the snow line. Their Fig. 7 shows that the probability of having at least one giant planet is six times lower for 0.4 $M_\odot$ star than a 1 $M_\odot$ star, in good agreement with the observed fraction from radial velocity surveys. A challenge for these models is to still be able to produce massive planets, e.g., the recent microlensing discovery of a 3.8 $M_{Jup}$ planet at 3.6 AU from a 0.46 $M_\odot$ M dwarf (*Dong et al.,* 2009).

At higher stellar masses, the evidence is that giant planets are more common. The most massive stars in Fig. 16 are evolved stars, either subgiants or giants. Whereas it is not possible to achieve high-precision velocities for massive stars on the main sequence, once these stars evolve off the main sequence and move to later spectral types, it again becomes possible to look for planets. *Johnson et al.* (2007, 2008) describe the first results from survey of 159 subgiants with Lick and Keck, for which they can achieve a precision of ≈2–5 m/s. They use these results to investigate the stellar mass dependence of the planet occurrence rate. To ensure uniform detectability between stars of different masses, they select stars that have at least eight observations with an observing time necessary to detect a companion at a = 2.5 AU, and consider only $M_p \sin i > 0.8\ M_{Jup}$ (this mass limit is set by the need to be able to detect companion around subgiants, which have lower radial velocity precision, and being more massive, lower stellar velocity amplitudes for a given planet mass). They find planet occurrence rates of 1.8 ± 1.0% for $M_\star < 0.7\ M_\odot$, 4.2 ± 0.7% for $0.7\ M_\odot < M_\star < 1.3\ M_\odot$, and 8.9 ± 2.9% for $M_\star > 1.3\ M_\odot$. The increasing trend of planet occurrence rate with mass remains after correcting for the mean metallicity of each mass bin.

This trend is consistent with the core accretion models of *Kennedy and Kenyon* (2008). They find that the planet occurrence rate is sensitive to the assumed dependence of snow line location with stellar mass. For example, *Ida and Lin* (2005), with a steeper dependence of snow line position on mass, predict a decrease in the planet fraction for $M_\star \geq 1\ M_\odot$.

*Johnson et al.* (2007, 2008) emphasize the different semimajor axis distribution of the planets discovered around subgiants and giants compared to the sample of F, G, and K stars. This is illustrated in Fig. 17, which shows the mass semimajor axis distribution of known planets, with stellar masses <0.5 $M_\odot$ (corresponding to M dwarfs) and >1.5 $M_\odot$ (evolved stars) highlighted. [Note that the ≈3 $M_\oplus$ microlensing planet MOA-2007-BLG-192-L b may orbit a substellar object. The best-fit mass for the host star is well

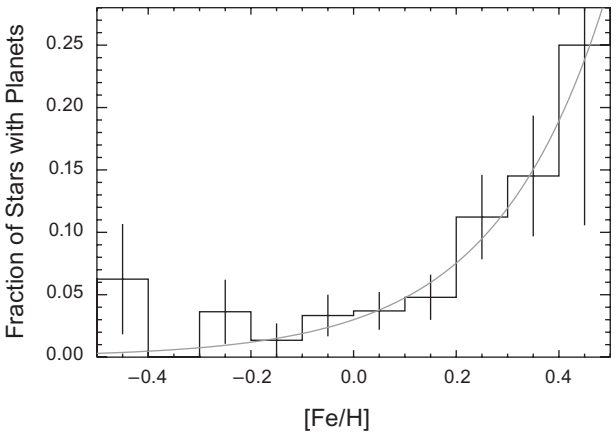



**Fig. 15.** Fraction of stars with planets as a function of [Fe/H], calculated using the same data as in Fig. 14. The gray curve shows the relation $0.03 \times 10^{2[Fe/H]}$ from *Fischer and Valenti* (2005).

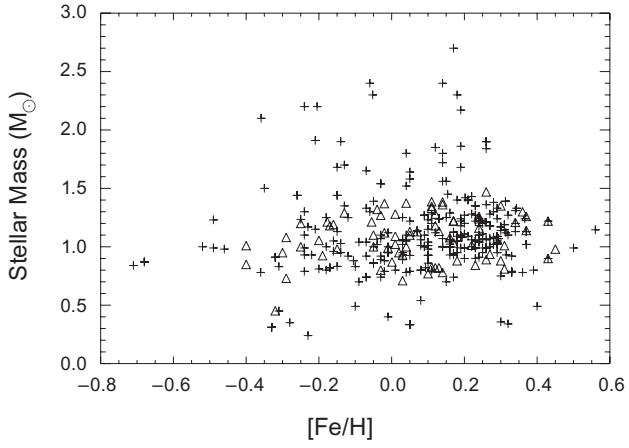



**Fig. 16.** Distribution of stellar metallicities and masses for transiting planet host stars (triangles) and stars with RV detected planets (crosses).

below an M dwarf mass, $M_\star = 0.06^{+0.028}_{-0.021}\ M_\odot$ (*Bennett et al.,* 2008).] The planets orbiting evolved stars are concentrated in long-period orbits, out beyond 0.6 AU [the exception in Fig. 17 is the hot Jupiter orbiting HD 102956 (*Johnson et al.,* 2010)]. This cannot be due to selection effects, since shorter orbital periods would be easier to detect. Similarly, no planet with mass ~1 $M_{Jup}$ has been detected within 0.1 AU of an M dwarf. Continued observations of evolved stars and M dwarfs will clarify this trend. Particularly for giants, tidal effects on close orbits must be included when trying to infer how much of this trend is due to differences in planet formation or migration between stars of different masses (*Sato et al.,* 2008).

### 3.7. Mass-Radius Relation of Transiting Planets

The statistical study of the properties of transiting planets is just beginning. Transiting planets are particularly interesting because we are not restricted to measuring the orbital elements, but can study planet properties such as radius and atmospheres. Some interesting trends in the population are already emerging.

The masses and radii of the transiting planets are shown in Fig. 18. At a given mass, the transiting planets show a range of radii, or alternatively a range of densities. Given the uncertainties in the physics of giant planet interiors (e.g., *Saumon and Guillot*, 2004), the study of a large sample of gas giant planets is of great interest. The first transiting planet HD 209458 had a larger radius than expected (1.3 $M_{Jup}$) (*Charbonneau et al.,* 2000), which has been attributed to tidal heating caused by damping of eccentricity driven by gravitational interactions with a second planet in the system (*Bodenheimer et al.,* 2003, *Laughlin et al.,* 2005b), or the insulating effect of stellar irradiation as the planet migrates inward, which keeps the internal entropy of the planet at a larger value than a planet cooling in isolation (*Burrows et al.,* 2003). A smaller radius than expected can be explained by the presence of a rocky core. For example, the Saturn-mass planet orbiting HD 149026 has a radius of only 0.73 $R_{Jup}$, and is inferred to have a core of mass ≈70 $M_\oplus$ (*Sato et al.,* 2005). *Guillot et al.* (2006) propose that the radii of all transiting planets can be accommodated by having different size cores, with the core mass increasing with stellar metallicity (see also *Burrows et al.,* 2007).

Figure 19 shows the masses and orbital periods of closely orbiting planets. Inspection of this figure shows that the planet mass is anticorrelated with the orbital period (for planet masses ≥0.1 $M_{Jup}$). One proposal is that this is evidence that planet-planet scattering is responsible for producing hot Jupiters. Some fraction of planets will scatter into orbits in which they approach very close to the star, and could tidally circularize there. In this picture, there should be a limit on the orbital radius of hot Jupiters of twice the Roche limit (*Rasio and Ford,* 1996; *Ford and Rasio,* 2006), since a highly eccentric orbit circularizes to twice its periastron distance. [To see this, conserve angular momentum $\propto\sqrt{GMa(1-e^2)}$ during the circularization, so that $a_{final} = a_{initial}(1-e^2) \approx 2a_{periastron}$ for $e \approx 1$.] Writing the Roche limit $a_R$ as $R_p = 0.462\ a_R(M_P/M_\star)^{1/3}$, then $a_{circ} = 2a_R$ gives $M_p = 1.1\ M_{Jup}\ (R_p/R_{Jup})3(P/1\ d)^{-2}$, which is plotted as a dotted line in Fig. 19. This is quite close to the lower envelope of the planets with $M_p > 0.1\ M_{Jup}$.

Other correlations between parameters have been discussed in the literature. *Hansen and Barman* (2007) point out that the transiting planets appear to fall into two classes based on Safronov number, $\Theta = (1/2)(v_{esc}/v_{orb})^2 = (a/R_p)(M_p/M_\star)$, where $v_{esc}$ is the escape speed from the planet, and $v_{orb}$ is the planet's orbital velocity. The size of the Safronov

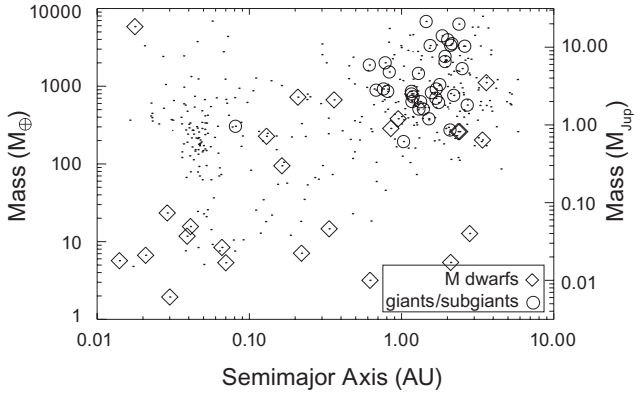



**Fig. 17.** Mass and orbital period distribution as a function of stellar mass. The dots show all the planets plotted in Fig. 6. The circled points are planets around stars more massive than 1.5 $M_\odot$ (evolved stars: giants and subgiants); the diamonds show stars less massive than 0.5 $M_\odot$ (M dwarfs). The dotted lines correspond to stellar radial velocity amplitudes of K = 1, 3, and 10 m/s for a solar-mass star.

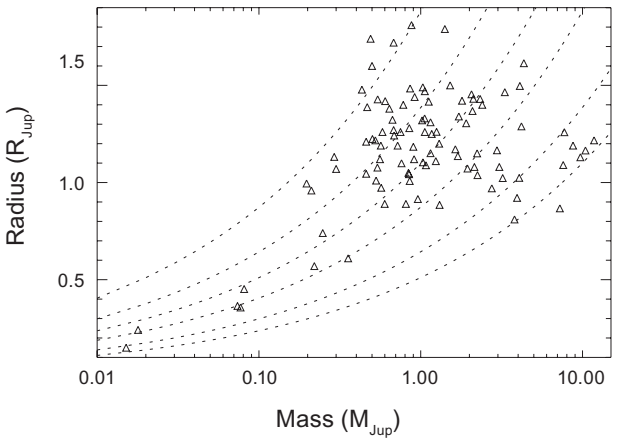



**Fig. 18.** Mass-radius relation for observed transiting planets. The dotted curves show constant mean density $\langle\rho\rangle = 3\ M_P/4\pi R_p^3 = 0.2$, 0.5, 1.0, 2.0, 5.0, and 10.0 g cm$^{-3}$.

determines whether the planet captures or scatters other objects in neighboring orbits, and so this difference could point to differences in migration or stopping mechanisms. However, the statistical significance of the correlation has been questioned (*Southworth,* 2008). As the number of transiting planets increases further, we are sure to learn a tremendous amount about their internal structure and formation and evolution histories.

### 3.8. Multiple Planet Systems

There are 46 multiple-planet systems in the set of planets considered here, about 10% of the 397 planetary systems. Of these, 15 have three or more planets and 5 have four or more planets, including the star 55 Cnc, which has five companions (*Fischer et al.*, 2008). Note that the chance of having additional planets is much greater for a two-planet system (15 out of 46) than for a single-planet system (46 out of 397). This may be partially due to a selection effect, in that two-planet systems will have been observed more carefully than many single planet systems, making it more likely to detect additional companions. Analysis of the multiplicity of planets taking into account selection effects has not been carried out.

Several multiple-planet systems show orbital period ratios that are integers or close to integers. For example, the two planets in GJ 876 have P ≈ 30 and 60 d, and are believed to be in a 2:1 mean-motion resonance. A natural way to form such an arrangement is by migration. As planets migrate inward, they can be trapped in resonant configurations (e.g., *Kley et al.,* 2005). The occurrence of resonant configurations is a strong argument that migration does take place.

Planets in resonant configurations are particularly interesting since over time they will show correlated changes in orbital elements. An example is the GJ 876 system (*Laughlin et al.,* 2005a; *Rivera et al.,* 2005; *Bean and Seifahrt,* 2009). The fit to the radial velocity measurements for this system is significantly improved if the gravitational interactions between the planets is included. Instead of fitting Keplerians, the equations of motion for the multiple body system are integrated directly to predict the expected radial velocity evolution. Because the dynamical interactions depend on the planet masses, whereas the Keplerian fits determine $M_P$ sin i, the interactions can be used to determine the inclination of the orbits, and whether the orbits are coplanar. Another example is the three-planet system orbiting υ And (the first multiple-planet system discovered orbiting a main-sequence star). The outer two planets orbit with periods of 241 d and 1301 d (the third planet is a hot Jupiter), and are expected to show evolution of their eccentricities and longitudes of periastron on long timescales of thousands of years (*Ford et al.,* 2005).

In Fig. 20, we show the known planets in the mass orbital period plane, highlighting those planets that are known to be in multiple systems. Note that this does not include stars with single planets that have an underlying trend in their radial velocity, many of which could have a planet at long orbital periods (*Fischer et al.,* 2001; *Wright et al.,* 2007). Quite striking in this figure is that the majority of the super Earths and Neptunes ($M_P$ sin i ≤ 30 $M_\oplus$) are in multiple-planet systems (see, e.g., *Mayor et al.,* 2009). However, one could imagine that this is a selection effect. Neptune-mass planets are detected with small signal-to-noise values, requiring many observations. Multiple-planet systems are likely to have been subject to intense observational scrutiny, accumulating many observations that could make them amenable to detecting planets with low masses.

Microlensing allows planetary systems at large orbital separations to be probed. There has been one detection so far; *Gaudi et al.* (2008) report the detection of a planetary

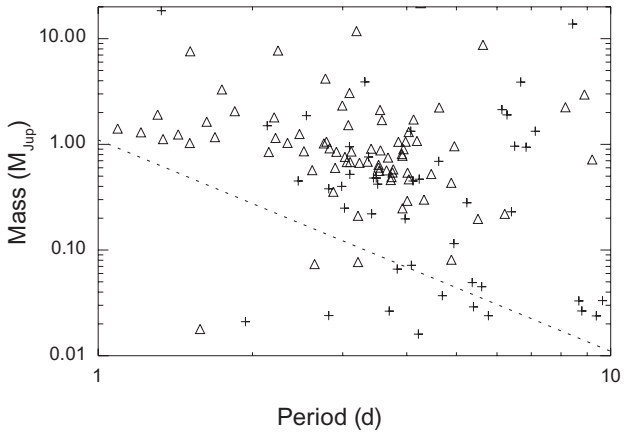



**Fig. 19.** Mass and orbital periods of closely orbiting exoplanets. Crosses show planets detected by radial velocity only; triangles show transiting planets. The dotted line shows twice the Roche limit, assuming $R_p = R_{Jup}$.

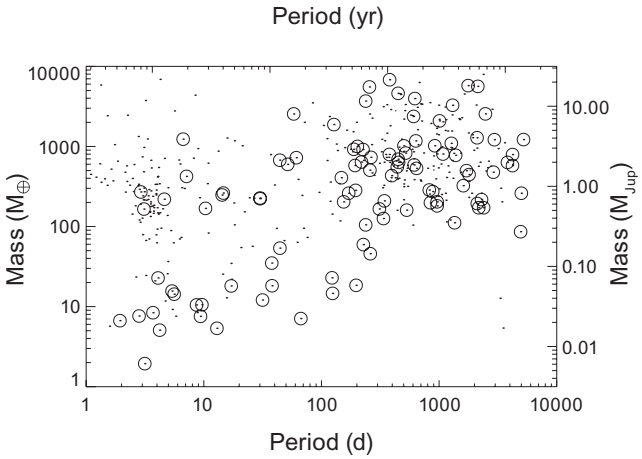



**Fig. 20.** Orbital period and mass distribution of planets, with members of multiple planet systems circled. Announced multiple planet systems only. This does not include those planets that show a long-term trend indicating a planet further out.

system that is a scaled version of our Jupiter/Saturn system, with 0.71 $M_{Jup}$ and 0.27 $M_{Jup}$ planets with orbital separations of 2.3 and 4.6 AU from a 0.5 $M_\odot$ host star. Based on the fact that the two other Jupiter-mass planets detected by microlensing were not very sensitive to multiple-planet systems, they suggested that the occurrence rate of multiple Jupiters is likely to be high, which promises more of these systems will be discovered in the future.

## 4. FUTURE PROSPECTS

This is an exciting time in exoplanet statistics. The current sample of planets is now large enough to yield statistically meaningful constraints on orbital properties, giving input for planet formation theories. In the next few years, we should learn much more as the sample increases further and new regions of the mass-orbital period plane and a wider range of stellar properties are explored.

Doppler surveys are moving forward in a few different ways. Doppler precision continues to improve beyond the 1 m $s^{-1}$ level. *Pepe and Lovis* (2008) discuss the prospects for lowering the Doppler precision from the current ≈1 m $s^{-1}$ level to the cm $s^{-1}$ level. They conclude that $\sigma \approx 10$ cm $s^{-1}$ should be possible for some stars that are "quiet" enough, although it is not yet clear for how long such a precision could be maintained. Precise radial velocity measurements so far have been made at optical wavelengths, but *Ramsey et al.* (2008) demonstrate the first measurements in the near-infrared at the 10 m $s^{-1}$ level. This has the potential to allow studies of M dwarfs that are not currently possible, moving the mass function of hot planets to smaller masses. As part of the Sloan Digital Sky Survey III, the Multi-Object APO Radial Velocity Exoplanet Large-Area Survey (MARVELS) will monitor 11,000 bright stars with a precision ~10 m $s^{-1}$. This survey makes use of a dispersed fixed-delay interferometer to make simultaneous observations of ≈60 stars in a given exposure, and ≈500 stars per night (*Mahadevan et al.,* 2008). *Johnson et al.* (2008) have added 300 new subgiants to their program to look for planets around massive stars, which should enlarge the sample by 20–30 more planets in the next few years. Continued monitoring of stars included in current surveys will reveal planets in long-period orbits. *Wright et al.* (2008) announced the discovery of a Jupiter twin around the star HD 154345. This planet has a mass of 0.95 $M_{Jup}$, and a 9.2-yr circular orbit, with no evidence so far for other planets in the system.

For transits, the number of planets has grown tremendously in the last few years, and will continue to increase in the near future with observations from the ground and from space (the CoRoT and Kepler missions). Whereas photometric precision from the ground is limited to ≈0.1%, spacebased observations allow a photometric precision adequate to detect Earth-sized planets [recall that $\delta \approx (R_P/R_\star)^2 \approx 10^{-4}$ for $R_P = R_\oplus$]. The NASA Kepler mission, launched in 2009, will continuously monitor $\approx 10^5$ stars for transits over the 3.5-yr mission lifetime. It should find hundreds of terrestrial-mass planets if they are common around solar-type stars (scaling from the ≈10% probability of observing a transit for a hot Jupiter at 0.05 AU, we see that the probability is ≈0.5% at 1 AU for a solar-type star), out to orbits of 1 AU and beyond, and increase the sample of hot Jupiters by tens (*Basri et al.,* 2005). This will provide a census of Earths and super Earths in close orbits.

Even with only a handful of detections, microlensing has already provided interesting constraints on the planet population beyond the snow line, complementing the discoveries made by radial velocity and transit surveys, e.g., indicating that cold Neptunes are common at orbital radii of a few AU. Microlensing offers some unique views of the planet population: It is potentially sensitive to Earth-mass planets beyond the snow line; it can detect old, free-floating planets; it surveys a different population of stars, being sensitive to planets in the bulge and disk of the Milky Way at large distances, and probes a wide range of host stars. Next-generation microlensing searches are currently being put together or planned (including the MOA-II and OGLE-IV upgrades, and the South Korean KMTNet project undergoing construction over the next few years), involving continuous wide-field large area surveys that would survey ~$10^7$ stars on ~10-min timescales, giving thousands of microlensing events per year that would all be monitored for planetary perturbations (rather than a small subsample as in current alert-follow-up surveys). Simulations indicate a detection rate of ≈10 super Earths per year and ≈1 Earth-mass planet per year (e.g., *Gaudi,* 2008), an order of magnitude improvement on the current detection rate. A spacebased mission such as the Microlensing Planet Finder [a 1-m-class telescope, imaging the galactic bulge continuously for several months (*Bennett et al.,* 2007)] would be 100 times as effective at finding Earth-mass planets (*Bennett,* 2004).

Direct searches are beginning to detect planetary companions, e.g., the HR 8799 system with three massive gas giants at ≈20–70 AU (*Marois et al.,* 2008), and had already placed interesting statistical constraints on giant planet occurrence rates at large orbital radii. The Gemini Deep Planet Survey (*Lafrenière et al.,* 2007), a near-infrared survey of 85 nearby young stars using adaptive optics at the Gemini North telescope, was able to place 95% upper limits on the fraction of stars with a gas giant between 10 and 25 AU of 0.28, 25 and 50 AU of 0.13, and 50 and 250 AU of 0.09. This compares to the fraction 0.17–0.2 for gas giants within 20 AU from the Keck radial velocity survey (*Cumming et al.,* 2008). Follow-on projects in the next few years such as the Gemini Planet Imager (*Macintosh et al.,* 2006) should increase the detection rate of gas giants at separations ≳5 AU, informing our understanding of the early evolution of gas giants, and when and how planets migrate to large orbital radii.

Future astrometry space missions should allow detection of a significant sample of planets at AU separations. Depending on the bright star single measurement precision achievable (they assume 8 μas), *Casertano et al.* (2008) find that Gaia could potentially find several thousands of gas giant planets at orbital radii 1–4 AU around stars within 200 pc, with hundreds of multiple-planet systems. The flux-limited

stellar sample would consist of a range of spectral types, ages, and metallicities. The Space Interferometry Mission, with a single measurement accuracy of 1 μas, could uncover tens of terrestrial planets (~3 $M_\oplus$) in habitable zones around stars within 30 pc (*Ford and Tremaine,* 2003; *Catanzarite et al.,* 2006).

Theoretical models of core accretion can reproduce many of the statistical properties of exoplanets; recent papers include *Ida and Lin* (2008b), *Kennedy and Kenyon* (2008), and *Mordasini et al.* (2009b). As the observational sample grows to include a wider range of orbital, planet, and stellar properties, a challenge for theorists is to identify which properties of the observed distributions constrain key aspects of the theoretical models, and can distinguish between alternative scenarios. In this way, the full potential of the upcoming observational discoveries can be realized.

***Acknowledgments.*** I thank the referees for a careful reading of this chapter and constructive suggestions that improved the text. I am grateful for support from the National Sciences and Engineering Research Council of Canada (NSERC) and the Canadian Institute for Advanced Research (CIFAR).

## REFERENCES

Armitage P. J. (2007) Massive planet migration: Theoretical predictions and comparison with observations. *Astrophys. J., 665,* 1381–1390.

Avni Y., Soltan A., Tananbaum H., and Zamorani G. (1980) A method for determining luminosity functions incorporating both flux measurements and flux upper limits, with applications to the average X-ray to optical luminosity ratio for quasars. *Astrophys. J., 238,* 800–807.

Basri G., Borucki W. J., and Koch D. (2005) The Kepler mission: A wide-field transit search for terrestrial planets. *New Astron. Rev., 49,* 478–485.

Bean J. L. and Seifahrt A. (2009) The architecture of the GJ 876 planetary system. Masses and orbital coplanarity for planets b and c. *Astron. Astrophys., 496,* 249.

Beatty T. G. and Gaudi B. S. (2008) Predicting the yields of photometric surveys for transiting extrasolar planets. *Astrophys. J., 686,* 1302–1330.

Beaulieu J.-P. et al. (2006) Discovery of a cool planet of 5.5 Earth masses through gravitational microlensing. *Nature, 439,* 437.

Bennett D. P. (2004) The detection of terrestrial planets via gravitational microlensing: Space vs. ground-based surveys. In Extrasolar Planets: Today and Tomorrow (J.-P. Beaulieu et al., eds.), p. 59. ASP Conf. Proc. 321, San Francisco.

Bennett D. P. and Rhie S. H. (1996) Detecting Earth-mass planets with gravitational microlensing. *Astrophys J., 472,* 660.

Bennett D. P. et al. (2007) An extrasolar planet census with a space-based microlensing survey. White paper submitted to the NASA/NSF ExoPlanet Task Force. *ArXiv e-prints, arXiv:0704.0454.*

Bennett D. P. et al. (2008) A low-mass planet with a possible sub-stellar-mass host in microlensing event MOA-2007-BLG-192. *Astrophys. J., 684,* 663.

Bodenheimer P., Laughlin G., and Lin D. N. C. (2003) On the radii of extrasolar giant planets. *Astrophys. J., 592,* 555.

Bramich D. M. and Horne K. (2006). Upper limits on the hot Jupiter fraction in the field of NGC 7789. *Mon. Not. R. Astron. Soc., 367,* 1677.

Brown T. M. (2003) Expected detection and false alarm rates for transiting jovian planets. *Astrophys J. Lett., 593,* L125.

Burrows A., Sudarsky D., and Hubbard W. B. (2003) A theory for the radius of the transiting giant planet HD 209458b. *Astrophys. J., 594,* 545–551.

Burrows A., Hubeny I., Budaj J., and Hubbard W. B. (2007) Possible solutions to the radius anomalies of transiting giant planets. *Astrophys. J., 661,* 502–514.

Butler R. P., Marcy G. W., Vogt S. S., Fischer D. A., Henry G. W., Laughlin G., and Wright J. T. (2003) Seven new Keck planets orbiting G and K dwarfs. *Astrophys. J., 582,* 455.

Butler R. P., Wright J. T., Marcy G. W., Fischer D. A., Vogt S. S., Tinney C. G., Jones H. R. A., Carter B. D., Johnson J. A., McCarthy C., and Penny A. J. (2006a) Catalog of nearby exoplanets. *Astrophys. J., 646,* 505–522.

Butler R. P., Johnson J. A., Marcy G. W., Wright J. T., Vogt S. S., and Fischer D. A. (2006b) A long-period Jupiter-mass planet orbiting the nearby M dwarf GJ 849. *Publ. Astron. Soc. Pacific, 118,* 1685–1689.

Casertano S. et al. (2008) Double-blind test program for astrometric planet detection with Gaia. *Astron. Astrophys., 482,* 699–729.

Catanzarite J., Shao M., Tanner A., Unwin S., and Yu J. (2006) Astrometric detection of terrestrial planets in the habitable zones of nearby stars with SIM PlanetQuest. *Publ.. Astron. Soc. Pacific, 118,* 1319.

Charbonneau D., Brown T. M., Latham D. W., and Mayor M. (2000) Detection of planetary transits across a Sun-like star. *Astrophys. J. Lett., 529,* L45–L48.

Cochran W. D., Endl M., MacQueen P. J., and Barnes S. (2008) The planet-metallicity correlation. *American Astronomical Society, DPS meeting #40,* #4.05.

Cumming A. (2004) Detectability of extrasolar planets in radial velocity surveys. *Mon. Not. R. Astron. Soc., 354,* 1165.

Cumming A. and Dragomir D. (2010) An integrated analysis of radial velocities in planet searches. *Mon. Not. R. Astron. Soc., 401,* 1029.

Cumming A., Marcy G. W., and Butler R. P. (1999) The Lick Planet Search: Detectability and mass thresholds. *Astrophys. J., 526,* 890.

Cumming A, Butler R. P., Marcy G. W., Vogt S. S., Wright J. T., and Fischer D. A. (2008) The Keck Planet Search: Detectability and the minimum mass and orbital period distribution of extrasolar planets. *Publ. Astron. Soc. Pacific, 120,* 531.

Defäy C., Deleuil M., and Barge P. (2001) A Bayesian method for the detection of planetary transits. *Astron. Astrophys., 365,* 330–340.

Dong S., Gould A., Udalski A., Anderson J., Christie G. W., et al. (2009) OGLE-2005-BLG-071Lb, the most massive M dwarf planetary companion? *Astrophys. J., 695,* 970.

Duquennoy A. and Mayor M. (1991) Multiplicity among solartype stars in the solar neighbourhood. II — Distribution of the orbital elements in an unbiased sample. *Astron. Astrophys., 248,* 485–524.

Endl M., Krster M., Els S., Hatzes A. P., Cochran W. D., Dennerl K., and Dbereiner S. (2002) The Planet Search Program at the ESO Coud Echelle spectrometer. III. The complete Long Camera survey results. *Astron. Astrophys., 392,* 671–690.

Fabrycky D. and Tremaine S. (2007) Shrinking binary and planetary orbits by Kozai cycles with tidal friction. *Astrophys. J., 669,* 1298.

Fischer D. A. and Valenti J. (2005) The planet-metallicity correlation. *Astrophys. J., 622,* 1102–1117.

Fischer D. A., Marcy G. W., Butler R. P., Vogt S. S., Frink S., and Apps K. (2001) Planetary companions to HD 12661, HD 92788, and HD 38529 and variations in Keplerian residuals of extrasolar planets. *Astrophys. J., 551,* 1107–1118.

Fischer D. A. et al. (2005) The N2K consortium. I. A hot Saturn planet orbiting HD 88133. *Astrophys. J., 620,* 481–486.

Fischer D. A. et al. (2008) Five planets orbiting 55 Cancri. *Astrophys. J., 675,* 790–801.

Ford E. B. (2005) Quantifying the uncertainty in the orbits of extrasolar planets. *Astron. J., 129,* 1706–1717.

Ford E. B. (2006) Improving the efficiency of Markov Chain Monte Carlo for analyzing the orbits of extrasolar planets. *Astrophys. J., 642,* 505–522.

Ford E. B. and Rasio F. A. (2006) On the relation between hot Jupiters and the Roche limit. *Astrophys. J. Lett., 638,* L45–L48.

Ford E. B. and Rasio F. A. (2008) Origins of eccentric extrasolar planets: Testing the planet-planet scattering model. *Astrophys. J., 686,* 621–636.

Ford E. B. and Tremaine S. (2003) Planet-finding prospects for the Space Interferometry mission. *Publ. Astron. Soc. Pacific, 115,* 1171–1186.

Ford E. B., Lystad V., and Rasio F. A. (2005) Planet-planet scattering in the upsilon Andromedae system. *Nature, 434,* 873–876.

Gaudi B. S. (2005) On the size distribution of close-in extrasolar giant planets. *Astrophys. J. Lett., 628,* L73–L76.

Gaudi B. S. (2008) Microlensing searches for planets: Results and future prospects. In *Extreme Solar Systems* (D. Fischer et al., eds.), p. 479. ASP Conf. Ser. 398, San Francisco.

Gaudi B. S. et al. (2002) Microlensing constraints on the frequency of Jupiter-mass companions: Analysis of 5 years of PLANET photometry. *Astrophys. J., 566,* 463.

Gaudi B. S., Seager S., and Mallen-Ornelas G. (2005) On the period distribution of close-in extrasolar giant planets. *Astrophys. J., 623,* 472–481.

Gaudi B. S. et al. (2008) Discovery of a Jupiter/Saturn analog with gravitational microlensing. *Science, 319,* 927.

Gilliland R. L. et al. (2000) A lack of planets in 47 Tucanae from a Hubble Space Telescope search. *Astrophys. J. Lett., 545,* L47.

Goldreich P. and Sari R. (2003) Eccentricity evolution for planets in gaseous disks. *Astrophys. J., 585,* 1024–1037.

Gonzalez G. (1997) The stellar metallicity-giant planet connection. *Mon. Not. R. Astron. Soc., 285,* 403–412.

Gould A. and Loeb A. (1992) Discovering planetary systems through gravitational microlenses. *Astrophys J., 396,* 104.

Gould A., Dorsher S., Gaudi B. S., and Udalski A. (2006a) Frequency of hot Jupiters and very hot Jupiters from the OGLEIII transit surveys toward the galactic bulge and Carina. *Acta Astron., 56,* 1–50.

Gould A., Udalski A., An D., Bennett D. P., Zhou A.-Y., et al. (2006b) Microlens OGLE-2005-BLG-169 implies that cool Neptune-like planets are common. *Astrophys. J. Lett., 644,* L37–L40.

Gregory P. C. (2007) A Bayesian Kepler periodogram detects a second planet in HD 208487. *Mon. Not. R. Astron. Soc., 374,* 1321–1333.

Griest K. and Safizadeh N. (1998) The use of high-magnification microlensing events in discovering extrasolar planets. *Astrophys. J., 500,* 37.

Guillot S., Guillot T., Santos N. C., Pont F., Iro N., Melo C., and Ribas I. (2006) A correlation between the heavy element content of transiting extrasolar planets and the metallicity of their parent stars. *Astron. Astrophys., 453,* L21–L24.

Halbwachs J. L., Mayor M., and Udry S. (2005) Statistical properties of exoplanets. IV. The period-eccentricity relations of exoplanets and of binary stars. *Astron. Astrophys., 431,* 1129–1137.

Hansen B. M. S. and Barman T. (2007) Two classes of hot Jupiters. *Astrophys. J., 671,* 861-871.

Holman M., Touma J., and Tremaine S. (1997) Chaotic variations in the eccentricity of the planet orbiting 16 Cygni B. *Nature, 386,* 254.

Horne K. (2003) Status and prospects of planetary transit searches: Hot Jupiters galore. In *Scientific Frontiers in Research on Extrasolar Planets* (D. Deming and S. Seager, eds.), pp. 361–369. *ASP Conf. Ser. 294,* San Francisco.

Ida S. and Lin D. N. C. (2004a) Toward a deterministic model of planetary formation. I. A desert in the mass and semimajor axis distributions of extrasolar planets. *Astrophys. J., 604,* 388–413.

Ida S. and Lin D. N. C. (2004b) Toward a deterministic model of planetary formation. II. The formation and retention of gas giant planets around stars with a range of metallicities. *Astrophys. J., 616,* 567–572.

Ida S. and Lin D. N. C. (2005) Toward a deterministic model of planetary formation. III. Mass distribution of short-period planets around stars of various masses. *Astrophys. J., 626,* 1045–1060.

Ida S. and Lin D. N. C. (2008a) Toward a deterministic model of planetary formation. IV. Effects of type I migration. *Astrophys. J., 673,* 487–501.

Ida S. and Lin D. N. C. (2008b) Toward a deterministic model of planetary formation. V. Accumulation near the ice line and super-Earths. *Astrophys. J., 685,* 584–595.

Jenkins J. M., Doyle L. R., and Cullers D. K. (1996) A matched filter method for ground-based sub-noise detection of terrestrial extrasolar planets in eclipsing binaries: Application to CM Draconis. *Icarus, 119,* 244–260.

Johnson J. A., Fischer D. A., Marcy G. W., Wright J. T., Driscoll P., et al. (2007) Retired A stars and their companions: Exoplanets orbiting three intermediate-mass subgiants. *Astrophys. J., 665,* 785–793.

Johnson J. A., Marcy G. W., Fischer D. A., Wright J. T., Reffert S., Kregenow J. M., Williams P. K. G. and Peek K. M. G. (2008) Retired A stars and their companions. II. Jovian planets orbiting κ CrB and HD 167042. *Astrophys. J., 675,* 784–789.

Johnson J. A., Bowler B. P., Howard A. W., Henry G. W., Marcy G. W., et al. (2010) Retired A stars and their companions V. A hot Jupiter orbiting the 1.7 Msun subgiant HD 102956. *ArXiv e-prints*, arXiv:1007.4555.

Jones H. R. A., Butler R. P., Tinney C. G., Marcy G. W., Penny A. J., McCarthy C., and Carter B. D. (2003) An exoplanet in orbit around $\tau^1$ Gruis. *Mon. Not. R. Astron. Soc., 341,* 948–952.

Jurić M. and Tremaine S. (2008) Dynamical origin of extrasolar planet eccentricity distribution. *Astrophys J., 686,* 603–620.

Kennedy G. M. and Kenyon S. J. (2008) Planet formation around stars of various masses: The snow line and the frequency of giant planets. *Astrophys. J., 673,* 502–512.

Kley W., Lee M. H., Murray N., and Peale S. J. (2005) Modeling the resonant planetary system GJ 876. *Astron. Astrophys., 437,* 727–742.

Kovács G., Zucker S., and Mazeh T. (2002) A box-fitting algorithm in the search for periodic transits. *Astron. Astrophys., 391,* 369–377.

Lafrenière D., Doyon R., Marois C., Nadeau D., Oppenheimer B. R., et al. (2007) The Gemini Deep Planet Survey — GDPS. *Astrophys. J., 670,* 1367–1390.

Laughlin G., Bodenheimer P., and Adams F. C. (2004) The core accretion model predicts few jovian-mass planets orbiting red dwarfs. *Astrophys. J. Lett., 612,* L73–L76.

Laughlin G., Butler R. P., Fischer D. A., Marcy G. W., Vogt S. S., and Wolf A. S. (2005a) The GJ 876 planetary system: A progress report. *Astrophys. J., 622,* 1182–1190.

Laughlin G., Wolf A., Vanmunster T., Bodenheimer P., Fischer D., Marcy G., Butler P., and Vogt S. (2005b) A comparison of observationally determined radii with theoretical radius predictions for short-period transiting extrasolar planets. *Astrophys. J., 621,* 1072–1078.

Lin D. N. C., Bodenheimer P., and Richardson D. C. (1996) Orbital migration of the planetary companion of 51 Pegasi to its present location. *Nature, 380,* 606–607.

Lineweaver C. H. and Grether D. (2003) What fraction of Sun-like stars have planets? *Astrophys. J., 598,* 1350–1360.

Lomb N. R. (1976) Least-squares frequency analysis of unequally spaced data. *Astrophys. Space Sci., 39,* 447–462.

Lovis C. et al. (2006) An extrasolar planetary system with three Neptune-mass planets. *Nature, 441,* 305–309.

Macintosh B., Graham J., Palmer D., Doyon R., Gavel D., et al. (2006) The Gemini Planet Imager. In *Advances in Adaptive Optics II* (B. L. Ellerbroek and D. B. Calia, eds.), 62720. Proc. SPIE, Vol. 6272, DOI: 10.1117/12.672430.

Mahadevan S., van Eyken J., Ge J., DeWitt C., Fleming S. W., Cohen R., Crepp J., and Vanden Heuvel A. (2008) Measuring stellar radial velocities with a dispersed fixed-delay interferometer. *Astrophys. J., 678,* 1505–1510.

Marcy G. W. and Benitz K. J. (1989) A search for substellar companions to low-mass stars. *Astrophys. J., 344,* 441–453.

Marcy G. W. and Butler R. P. (1998) Detection of extrasolar giant planets. *Annu. Rev. Astron. Astrophys., 36,* 57–97.

Marcy G. W., Butler R. P., Vogt S. S., Fischer D. A., Henry G. W., Laughlin G., Wright J. T., and Johnson J. A. (2005a) Five new extrasolar planets. *Astrophys. J., 619,* 570–584.

Marcy G., Butler R. P., Fischer D., Vogt S., Wright J. T., Tinney C. G., and Jones H. R. A. (2005b) Observed properties of exoplanets: Masses, orbits and metallicities. *Progr. Theor. Phys. Suppl., 158,* 24–42.

Mandel K. and Agol E. (2002) Analytic light curves for planetary transit searches. *Astrophys. J. Lett., 580,* L171–L175.

Marois C., Macintosh B., Barman T., Zuckerman B., Song I., Patience J., Lafrenière D., and Doyon R. (2008) Direct imaging of multiple planets orbiting the star HR 8799. *Science, 322,* 1348.

Mayor M., Bonfils X., Forveille T., Delfosse X., Udry S., Bertaux J.-L., Beust H., Bouchy F., Lovis C., Pepe F., Perrier C., Queloz D., and Santos N. C. (2009) The HARPS search for southern extra-solar planets. XVIII. An Earth-mass planet in the GJ 581 planetary system. *Astron. Astrophys., 507,* 487.

Mochejska B. J. et al. (2005) Planets in stellar clusters extensive search. III. A search for transiting planets in the metal-rich open cluster NGC 6791. *Astron. J., 129,* 2856.

Mordasini C., Alibert Y., and Benz W. (2009a) Extrasolar planet population synthesis. I. Method, formation tracks, and mass-distance distribution. *Astron. Astrophys., 501,* 1139.

Mordasini C., Alibert Y., Benz W., and Naef D. (2009b) Extrasolar planet population synthesis. II. Statistical comparison with observations. *Astron. Astrophys., 501,* 1161.

Moutou C., Pont F., Barge P., Aigrain S., Auvergne M., et al. (2005) Comparative blind test of five planetary transit detection algorithms on realistic synthetic light curves. *Astron. Astrophys., 437,* 355–368.

Naef D., Mayor M., Beuzit J.-L., Perrier C., Queloz,D., Sivan J.-P., and Udry S. (2005) The ELODIE Planet Search: Synthetic view of the survey and its global detection threshold. In *Proc. 13th Cambridge Workshop on Cool Stars, Stellar Systems and the Sun* (F. Favata et al., eds.), pp. 833–836. ESA SP-560, Noordwijk, The Netherlands.

O'Toole S. J., Jones H. R. A., Tinney C. G., Butler R. P., Marcy G. W., Carter B., Bailey J., and Wittenmyer R. A. (2009) The frequency of low-mass exoplanets. *Astrophys. J., 701,* 1732.

Patel S. G., Vogt S. S., Marcy G. W., Johnson J. A., Fischer D. A., Wright J. T., and Butler R. P. (2007) Fourteen new companions from the Keck and Lick radial velocity survey including five brown dwarf candidates. *Astrophys. J., 665,* 744–753.

Pepe F. A. and Lovis C. (2008) From HARPS to CODEX: Exploring the limits of Doppler measurements. *Phys. Scripta T, 130,* 014007.

Pinsonneault M. H., DePoy D. L. and Coffee M. (2001) The mass of the convective zone in FGK main-sequence stars and the effect of accreted planetary material on apparent metallicity determinations. *Astrophys. J. Lett., 556,* L59.

Pont F., Zucker S., and Queloz D. (2006) The effect of red noise on planetary transit detection. *Mon. Not. R. Astron. Soc., 373,* 231–242.

Press W. H., Teukolsky S. A., Vetterling W. T., and Flannery B. P. (1992) *Numerical Recipes: The Art of Scientic Computing,* 2nd edition, Cambridge Univ., Cambridge.

Queloz D. et al. (2001) No planet for HD 166435. *Astron. Astrophys., 379,* 279–287.

Ramsey L. W., Barnes J., Redman S. L., Jones H. R. A., Wolszczan A., Bongiorno S., Engel L., and Jenkins J. (2008) A Pathfinder instrument for precision radial velocities in the near-infrared. *Publ. Astron. Soc. Pacific, 120,* 887–894.

Rasio F. A. and Ford E. B. (1996) Dynamical instabilities and the formation of extrasolar planetary systems. *Science, 274,* 954–956.

Ribas I. and Miralda-Escudé J. (2007) The eccentricity-mass distribution of exoplanets: Signatures of different formation mechanisms? *Astron. Astrophys., 464,* 779–785.

Rivera E. J., Lissauer J. J., Butler R. P., Marcy G. W., Vogt S. S., Fischer D. A., Brown T. M., Laughlin G., and Henry G. W. (2005) A 7.5 Earth mass planet orbiting the nearby star, GJ 876. *Astrophys. J., 634,* 625.

Robinson S. E., Laughlin G., Vogt S. S., Fischer D. A., Butler R. P., Marcy G. W., Henry G. W., Driscoll P., Takeda G., and Johnson J. A. (2007) Two jovian-mass planets in Earthlike orbits. *Astrophys. J., 670,* 1391.

Saar S. H. and Donahue R. A. (1997) Activity-related radial velocity variation in cool stars. *Astrophys. J., 485,* 319–327.

Saar S. H., Butler R. P., and Marcy G. W. (1998) Magnetic activity-related radial velocity variations in cool stars: First results from the Lick extrasolar planet survey. *Astrophys. J. Lett., 498,* L153–L157.

Santos N. C., Israelian G., and Mayor M. (2004) Spectroscopic [Fe/H] for 98 extra-solar planet-host stars: Exploring the probability of planet formation. *Astron. Astrophys., 415,* 1153.

Sato B., Fischer D. A., Henry G. W., Laughlin G., Butler R. P., et al. (2005) The N2K consortium. II. A Transiting hot Saturn around HD 149026 with a large dense core. *Astrophys. J., 633,* 465–473.

Sato B., Izumiura H., Toyota E., Kambe E., Ikoma M., et al. (2008) Planetary companions around three intermediate-mass G and K giants: 18 Delphini, ξ Aquilae, and HD 81688. *Proc. Astron. Soc. Japan, 60,* 539.

Saumon D. and Guillot T. (2004) Shock compression of deuterium and the interiors of Jupiter and Saturn. *Astrophys. J., 609,* 1170–1180.

Scargle J. D. (1982) Studies in astronomical time series analysis. II — Statistical aspects of spectral analysis of unevenly spaced data. *Astrophys. J., 263,* 835–853.

Schwarzenberg-Czerny A. and Beaulieu J.-P. (2006) Efficient analysis in planet transit surveys. *Mon. Not. R. Astron. Soc., 365,* 165–170.

Seager S. and Mallén-Ornelas G. (2003) A unique solution of planet and star parameters from an extrasolar planet transit light curve. *Astrophys. J., 585,* 1038–1055.

Shen Y. and Turner E. L. (2008) On the eccentricity distribution of exoplanets from radial velocity surveys. *Astrophys. J., 685,* 553–559.

Sivia D. S. (1996) *Data Analysis: A Bayesian Tutorial.* Oxford Univ., New York.

Snodgrass C., Horne K., and Tsapras Y. (2004) The abundance of galactic planets from OGLE-III 2002 microlensing data. *Mon. Not. R. Astron. Soc., 351,* 967.

Southworth J. (2008) Homogeneous studies of transiting extrasolar planets. II. Physical properties. *Mon. Not. R. Astron. Soc., 394,* 272.

Stepinski T. F. and Black D. C. (2001) On orbital elements of extrasolar planetary candidates and spectroscopic binaries. *Astron. Astrophys., 371,* 250–259.

Tabachnik S. and Tremaine S. (2002) Maximum-likelihood method for estimating the mass and period distributions of extrasolar planets. *Mon. Not. R. Astron. Soc., 335,* 151–158.

Takeda G., Ford E. B., Sills A., Rasio F. A., Fischer D. A., and Valenti Jeff A. (2007) Structure and evolution of nearby stars with planets. II. Physical properties of 1000 cool stars from the SPOCS catalog. *Astrophys. J. Suppl., 168,* 297.

Tingley B. (2003) Improvements to existing transit detection algorithms and their comparison. *Astron. Astrophys., 408,* L5–L7.

Torres G., Winn J. N., and Holman M. J. (2008) Improved parameters for extrasolar transiting planets. *Astrophys. J., 677,* 1324–1342.

Udry S. and Santos N. C. (2007) Statistical properties of exoplanets. *Annu. Rev. Astron. Astrophys., 45,* 397–439.

Udry S., Mayor M., Naef D., Pepe F., Queloz D., Santos N. C., Burnet M., Conno B., and Melo C. (2000) The CORALIE survey for southern extra-solar planets II. The short-period planetary companions to HD 75289 and HD 130322. *Astron. Astrophys., 356,* 590–598.

Udry S., Mayor M., and Santos N. C. (2003) Statistical properties of exoplanets I. The period distribution: Constraints for the migration scenario. *Astron. Astrophys., 407,* 369–376.

Valenti J. A. and Fischer D. A. (2005) Spectroscopic properties of cool stars (SPOCS). I. 1040 F, G, and K dwarfs from Keck, Lick, and AAT Planet Search programs. *Astrophys. J. Suppl., 159,* 141–166.

von Braun K., Kane S. R. and Ciardi D. R. (2009) Observational window functions in planet transit surveys. *Astrophys. J., 702,* 779.

Wainstein L. A. and Zubakov V. D. (1962) *Extraction of Signals from Noise.* Prentice-Hall, New Jersey.

Walker G. A. H., Walker A. R., Irwin A. W., Larson A. M., Yang S. L. S., and Richardson D. C. (1995) A search for Jupiter-mass companions to nearby stars. *Icarus, 116,* 359–375.

Weldrake D. T. F., Sackett P. D., and Bridges T. J. (2008) The frequency of large-radius hot and very hot Jupiters in ω Centauri. *Astrophys. J., 674,* 1117.

Wittenmyer R. A., Endl M., Cochran W. D., Hatzes A. P., Walker G. A. H., Yang S. L. S., and Paulson D. B. (2006) Detection limits from the McDonald Observatory Planet Search program. *Astron. J., 132,* 177–188.

Wright J. T. (2005) Radial velocity jitter in stars from the California and Carnegie Planet Search at Keck observatory. *Publ. Astron. Soc. Pacific, 117,* 657–664.

Wright J. T., Marcy G. W., Butler R. P., and Vogt S. S. (2004) Chromospheric Ca II emission in nearby F, G, K, and M stars. *Astrophys. J. Suppl., 152,* 261–295.

Wright J. T., Marcy G. W., Fischer D. A., Butler R. P., Vogt S. S., et al. (2007) Four new exoplanets and hints of additional substellar companions to exoplanet host stars. *Astrophys. J., 657,* 533–545.

Wright J. T., Marcy G. W., Butler R. P., Vogt S. S., Henry G. W., Isaacson H., and Howard A. W. (2008) The Jupiter twin HD 154345b. *Astrophys. J. Lett., 683,* L63–L66.

Wu Y. (2003) Tidal circularization and Kozai migration. In *Scientific Frontiers in Research on Extrasolar Planets* (D. Deming and S. Seager, eds.), pp. 213–216. *ASP Conf. Ser. 294,* San Francisco.

Wu Y. and Murray N. (2003) Planet migration and binary companions: The case of HD 80606b. *Astrophys. J., 589,* 605–614.

Wu Y., Murray N. W., and Ramsahai J. M. (2007) Hot Jupiters in binary star systems. *Astrophys. J., 670,* 820–825.

Zakamska N. L. and Tremaine S. (2004) Excitation and propagation of eccentricity disturbances in planetary systems. *Astron. J., 128,* 869–877.