1. Overview of extrasolar planet detection methods

LAURANCE R. DOYLE

In this chapter we will describe in a general manner each planet detection method and examine the fundamental astrophysical parameters each technique measures as well as its present measurement limitations for the detection of inner giant planets, jovian outer planets, and Earthlike planets. We then outline several secondary detection methods that may be instituted in the near future with increased detection sensitivity. We then discuss the ranges of each detection method and sketch several cases in which additional parameters may be derived through the acquisition of data from several methods combined. In the final section we discuss habitable zones around M-dwarf systems as potential near-term targets for the detection of life-supporting planets.

1.1. Introduction

In the following sections an overview of the main methods of extrasolar planet detection is presented. This is not a historical review – an excellent review, for example, can be found in Perryman (2000) and the 469 references therein. It is also not an up-to-date listing of extrasolar planet detections or candidates; these can be found at the comprehensive site of the *Extrasolar Planets Encyclopedia* by J. Schneider (www.obspm.fr/encycl/encycl.html). In this chapter we do, however, describe in a general manner each detection method and examine the general astrophysical parameters each technique measures as well as its present measurement limitations. We mention some secondary detection methods that may find application in the near future and what additional parameters may be derived through the acquisition of data from several methods combined. We finally discuss M-dwarf star habitable zones, as these are likely to be the near-term targets for the detection of exobiology on extrasolar planets. This chapter is aimed, in explanatory detail, at the interested college student level.

We note that the detection parameters for the pulsar timing, radial velocity, astrometric imaging, reflected light and eclipsing binary timing methods depend, at any given time, on the orbital phase, $\varphi(t)$, of the extrasolar planet, which is a function of the geometry involved in that detection method. However, detectability depends on the maximum signal produced for a given method, and it is this that we formulate in the equations below. However, we shall point out at which phases this maximum occurs. In keeping with eclipsing binary protocol, the planetary orbital phase $\varphi(t) = 0$ degrees will be when the (darker) planet is in inferior conjunction, that is when it is closest to the observer.

1.2. Pulsar timing

Unexpectedly, the first planetary-mass objects detected around another star were closer to terrestrial-mass than to jovian-mass. The parent star was the pulsar PSR B1257+12, 500 parsecs distant, and the two planetary objects detected around it are a 2.8 Earth (projected) mass ($M_\oplus$) body with a period of 98.22 days and a 3.4 $M_\oplus$ body with a period

Extrasolar Planets, eds. Hans Deeg, Juan Antonio Belmonte and Antonio Aparicio.
Published by Cambridge University Press.

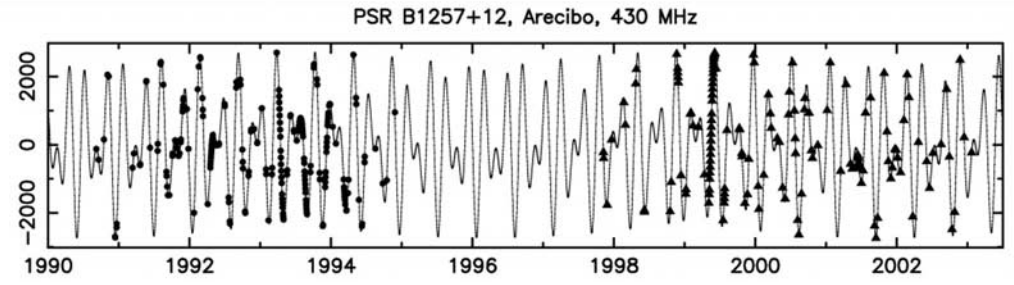

FIGURE 1.1. Time of arrival residuals (in microseconds) of 430 MHz signals from the 6.2-millisecond pulsar PSR B1257 + 12 (from Konacki and Wolszczan 2003) showing that the residuals are dominated by the Keplerian orbits of two planets of actual mass 4.3 (B) and 3.9 (C) Earth-masses (of three in the system, the third being very close to the pulsar). These planets are nearly coplanar (around 50 degrees orbital inclination) and are in actual 3:2 orbital resonance with each other (66.5 and 98.2-day periods; planet A having a period of 25.3 days).

of 66.54 days (Wolszczan 1994; Wolszczan and Frail 1992). The precise radio pulse rates of pulsars (seconds to milliseconds) and their stability as timing 'clocks' (variations in pulse timing on the order of only about a trillionth of a second per year) allow variations in the position of the pulsar to be measured precisely. The variation in timing can occur due to a positional shift in the pulsar around the pulsar–planet barycentre. If such a second mass (planet) is in orbit around the pulsar, the two bodies will orbit around a mutual barycentre, each distance from the barycentre being determined directly by their mass-ratios, where M_* and a_* are the mass and distance (semi-major axis) from the barycentre to the centre of the pulsar and $M_{\rm p}$ and $a_{\rm p}$ are the mass and distance from the barycentre to the planet. The motion of the pulsar around the barycentre causes the addition of (or subtraction of) the light travel time across this distance, which will result in a delay (or early arrival) of the periodic variations in the timing of the pulsar pulses. For a planet in a circular orbit, the maximum amplitude of the delay time will be:

$$\tau = \sin i \left(\frac{a_{\rm p}}{c}\right) \left(\frac{M_{\rm p}}{M_*}\right), \tag{1.1}$$

where i is the inclination of the planet's orbit ($i = 90^\circ$ being edge-on), and c is the speed of light. The pulses will be 'on time' at phases $\varphi(t) = 90$ and 270 degrees, late by an amount τ at $\varphi(t) = 0$ degrees, and early by an amount τ at $\varphi(t) = 180$ degrees orbital phase angle, where zero degrees phase is when the planet is closest to the observer (i.e. inferior conjunction). Note that the sine function in Eq. (1.1) is not negative because it is the pulsar signals that are being measured directly, and the pulsar is at the opposite orbital phase from the planet. Thus, via the foreshortened light travel time across the stellar-barycentre distance, the pulsar timing method can measure the projected planet-to-star mass ratio, the true orbital period of the planet (or planets), and its orbital eccentricity (if the orbit is not circular). General relativistic precessional phase drifts may allow further constraints on the pulsar mass, but only for closer-in planets over longer observing times (see Figure 1.1).

If we define a typical close-in extrasolar giant planet (CEGP) as a 3 jovian-mass planet with a circular orbital semi-major axis (i.e. orbital radius) of 0.05 AU (astronomical unit), and a 'Jupiter' and an 'Earth' as planets with the mass and orbital location (distance from their star) of Jupiter and Earth in the Solar System, respectively, then the half-amplitude timing offsets for such planets around a typical pulsar (assuming the pulsar

to be 1.35 solar masses) would be $\tau = 140$ milliseconds (ms) for a CEGP, 1.65 seconds for a 'Jupiter', and 3 ms for an 'Earth'. That is, these will be the expected maximum delays in the pulse arrival times at a planetary orbital phase $\varphi(t) = 0$ degrees.

1.3. Periodic radial velocity variations

The radial velocity or 'Doppler shift' method has been the most successful extrasolar planet detection method to date, detecting the vast majority of planets as of this writing. The first extrasolar planets around solar-type stars were discovered in this way (Mayor and Queloz 1995; see also Marcy and Butler 1998 and reference therein). Radial velocity variations again cause a wobble in the parent star, but the stellar light flux is generally very constant, so that timing of variations cannot be used to detect this stellar offset around the star–planet barycentre. However, very high precision spectral line measurements (one part in a hundred millionth of a spectral line width) can be performed by superimposing a comparison spectrum with many lines (like an iodine cell in the light path at the observatory) on to the stellar spectrum for a precise measurement of periodic movement in the star's spectral lines.

The stellar spectral lines will move periodically redward or blueward due to the Doppler shift by $\Delta\lambda/\lambda = v/c$, caused by the periodic motion, with a maximum velocity v of the star about the star–planet barycentre. Again, the spectral line variations only measure the component of the motion directly towards or away from the observer, and hence the mass of the body (planet) causing the reflex motion of the star is a minimum mass measurement for the planet, $M_\mathrm{p} \sin i$. The maximum amplitude of this periodic radial velocity variation is given by:

$$K = \left(\frac{2\pi G}{P}\right)^{1/3} \frac{M_\mathrm{p} \sin i}{(M_\mathrm{p} + M_*)^{2/3}} \frac{1}{(1-e^2)^{1/2}}, \tag{1.2}$$

where P is the planetary orbital period, e is the planetary orbital eccentricity, and G is the gravitational constant. K, P and e can be derived from several measurements of the Doppler shift during a planet's orbit. The maxima in K will occur at planetary orbital phases of $\varphi(t) = 90$ degrees (blueshift of stellar spectra lines) and $\varphi(t) = 270$ degrees (redshift of stellar spectral lines). With knowledge of M_* from stellar classification (typically based on low-resolution spectra), the term $M_\mathrm{p} \sin i$ can then be derived. Kepler's third law, $(P/\mathrm{yr}) = (a_\mathrm{p}/1\mathrm{AU})^{3/2}(M_*/M_\odot)^{-1/2}$, where $M_\odot$ is one solar mass, allows then also a derivation of the semi-major axis of the planet.

The precision possible for this detection method is about 1 m/s, this limit imposed by intrinsic stellar surface fluctuations – i.e., variations present in even the most stable solar-type stars (see Figure 1.2). For a CEGP the radial velocity amplitude will be about 56 m/s, for a 'Jupiter' about 13 m/s, and for an 'Earth' about 0.1 m/s. Thus this method may not be expected to detect Earthlike planets around solar-like stars but can, however, detect any jovian-mass bodies within a star's circumstellar habitable zone (CHZ).[1] Hypothetical Earth-sized moons around such bodies have been suggested as being of interest to exobiologists. The detection of Jupiterlike planets are of interest both because of their comparability with our own Solar System as well as the ability of jovian-type planets to remove cometary debris, serving as a possibly necessary 'shield' for any biosystems developing on the inner terrestrial planets of the star system. This method is also, at

[1] The circumstellar habitable zone (CHZ) is defined here as the distance regime around a star where liquid water can persist on the surface of a sufficiently large planet. For a discussion of the CHZ see Chapter 8.

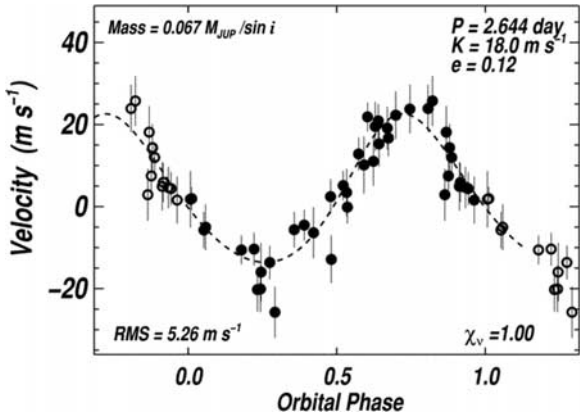

FIGURE 1.2. A Neptune-mass planet orbiting the nearby 0.41 solar-mass M-dwarf GJ 436 (from Butler *et al.* 2004). The 18.1 metre/second variation in the spectral lines of the star with a period of 2.644 days is caused by a planet with a projected mass of about 1.2 Neptune-masses.

present, limited to detection of planets around fairly stable, single star systems as the measurement of these radial velocity variations demands such high spectral line precision measurements.

1.4. Gravitational microlensing

Due to general relativistic effects of bending spacetime, a star moving very close to alignment with a background star will bend – that is, focus – the light of the background star, causing a temporary increase in the combined brightness of the stars by amplifying the light from the background star. The phenomenon, first observed with galaxies, is known as gravitational lensing, A perfect stellar alignment will cause symmetric images around the lensing star; this is known as the 'Einstein ring' (or sometimes an 'Einstein cross'). The Einstein ring radius is given by:

$$R_{\rm E} = \left[\frac{4GM_{*\rm L}}{c^2}\frac{(D_{\rm S}-D_{\rm L})\,D_{\rm L}}{D_{\rm S}}\right]^{1/2}, \tag{1.3}$$

where $M_{*\rm L}$ is the mass of the lensing star, $D_{\rm L}$ is the distance to the lensing star and $D_{\rm S}$ is the distance to the source star. The angle on the sky of the Einstein radius (the Einstein angle) is then given as: $\theta_{\rm E} = R_{\rm E}/D_{\rm L}$. The microlensing magnification, which varies with time, is given by:

$$Q(t) = \frac{u^2(t)+2}{u(t)[u^2(t)+4]^{1/2}}, \tag{1.4}$$

where $u(t)$ is the projected distance between the image of the lensing star and the source star in units of the Einstein radius (Perryman 2000). We can see that for an exact alignment the magnification would become infinite, theoretically. If a planet is in orbit around the lensing star, then observable deviations from the amplification pattern given by Eq. (1.4) may occur, which are caused by a planet-mass distorting the stellar gravitational field.

The probability of alignment among two stars is, even in the Galactic Centre, only about one in 10^6, but once a star is aligned with another star the probability that a planet may also cause an amplification that exceeds 5% of the brightness of the star's amplification itself becomes about one in five (Schneider *et al.* 1999). For this superposition

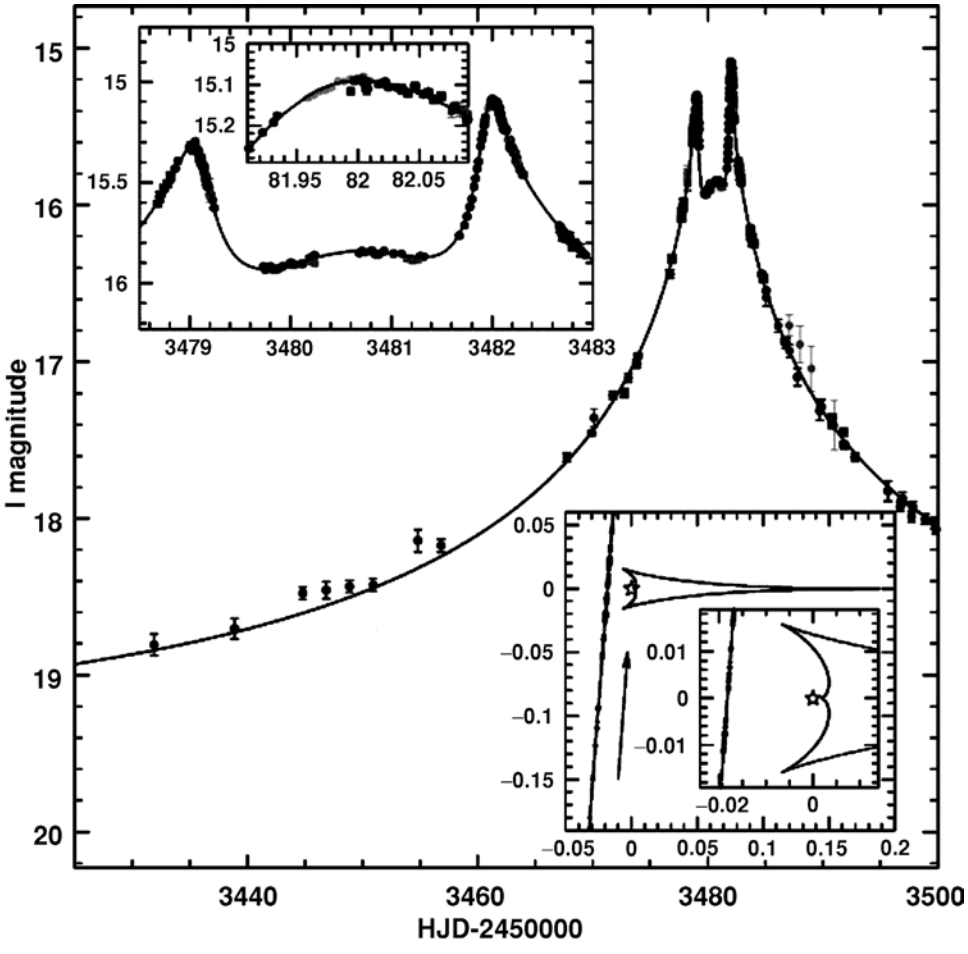

FIGURE 1.3. The light curve of OGLE-2005-BLG-071 (adapted from Udalski *et al.* 2005), showing a binary peak, indicative of the binary nature of this microlensing event, where both a star and a planet move in front of a background star. A magnification of the three peaks (the middle one with a low amplitude) is shown in the inset in the upper left. The bottom right inset shows the caustic surfaces (closed-curve regions of very high magnification) consistent with a binary lens mass ratio of $q = 0.0071$. From an analysis of parallax effects in the wings of the microlensing event, the lensing stellar mass is constrained to be between 0.5 and 0.08 solar masses, at a distance between 1.5 and 5 kiloparsecs, giving a planetary mass from 0.05 to 4 Jupiter masses.

of a brightening due to a planet on top of that due to the amplified star, the term $M_{*\mathrm{L}}$ becomes the mass of the planet, M_{p} in Eq. (1.3).

The duration of a microlensing event is given by:

$$t_{\mathrm{E}} = \sqrt{4GM_{\mathrm{p}}dV/c^2}, \tag{1.5}$$

where d is the distance to the lensing star in parsecs and V is the orbital velocity. If the distance d to the lensing star–planet system is assumed to be 5 kiloparsecs, then for a CEGP the brightening will be about four magnitudes and last for about five days. For a 'Jupiter' the brightening will be about three magnitudes and last about three days. And for an 'Earth' the brightening will be about one magnitude and last about four hours (see

Figure 1.3). Thus we see that these events are quite detectable. The main difficulty of this approach is that the detection of these events requires a large number of stars to be observed that are also pretty far away (a few kiloparsecs), and finally that the detection event cannot be expected to repeat. However, as pointed out in Udalski *et al.* (2005) planetary parameters (such as orbital period) may be determined that are longer than the duration of the observations (note no orbital phase dependence in Eqs. (1.3), (1.4), or (1.5)). In general, however, the planet-to-star mass ratio is what is determined, and since target systems are generally so far away, precision spectra determining the stellar mass are more difficult to obtain than for nearby systems. However, this is at present the planet detection method that works over the longest distance and might give insights into the planetary distribution in quite different stellar populations – for example disc compared to bulge galactic populations.

1.5. Astrometry

Astrometry is perhaps the oldest method to search for extrasolar planets, several having been reported in the mid-twentieth century. This method measures a periodic variation in the position of the star on the 'plane of the sky', subtracting out the star's apparent motion due to the yearly parallax motion and the projection of its real proper motion through space. The motion of a star around the star–planet barycentre thus describes an elliptical motion with semi-major axis (in arcseconds) of:

$$\alpha = \frac{M_{\rm p}}{M_*}\frac{a_{\rm p}}{d}, \tag{1.6}$$

where $a_{\rm p}$ is in AU (astronomical units), and d, the distance to the stellar system, is given in parsecs. For a circular planetary orbit the inclination of the planetary orbit from the observer's line of sight is just $\sin i = b/\alpha$, where b is the semi-minor axis measured. The maximum semi-major axis position of the star occurs at elongation when the planetary orbital phases are $\varphi(t) = 90$ and 270 degrees. Here at maximum elongation $\varphi(t) = 1$. This technique measures the motion of the photometric centroid position of the star in images taken over at least a large fraction of a planet's orbit. It is complementary, for example, to the radial velocity detection method in that it is most sensitive to long-period (large semi-major axis) planets, while the radial velocity method is most sensitive to inner, short-period planets with higher velocity variations. As for the factor d (as indicated in Eq. (1.6)), astrometric detection of extrasolar planets is very sensitive to the distance to the system, and at present is limited to somewhat nearby stellar systems (see Figure 1.4).

For the full amplitude of a CEGP the astrometric offset on a solar-type star at a distance of 5 pc would be about 0.03 milliarcseconds. For a 'Jupiter' it would be about 1 milliarcsecond, and for an 'Earth' the offset would be about 0.6 microarcseconds. This technique can be extended to search for extrasolar planets around radio-emitting stars using very long baseline radio interferometry (see Perryman 2000). Upcoming wide field searches for transiting planets (for example, the NASA *Kepler* mission) may also allow astrometric searches for planets to take place using the same photometric data, since the pointing precision as well as the photometric centroiding of star images should be near the 1 milliarcsecond precision required for astrometry. Near-term spacecraft missions such as *SIM* (*Space Interferometry Mission*) will be specifically designed to optimize astrometric measurements both for stellar parallax determinations and the detection of extrasolar planets in the solar neighbourhood astrometrically. *SIM* should be able to detect nearby extrasolar planets while mapping exact distances to stars by using interferometry to

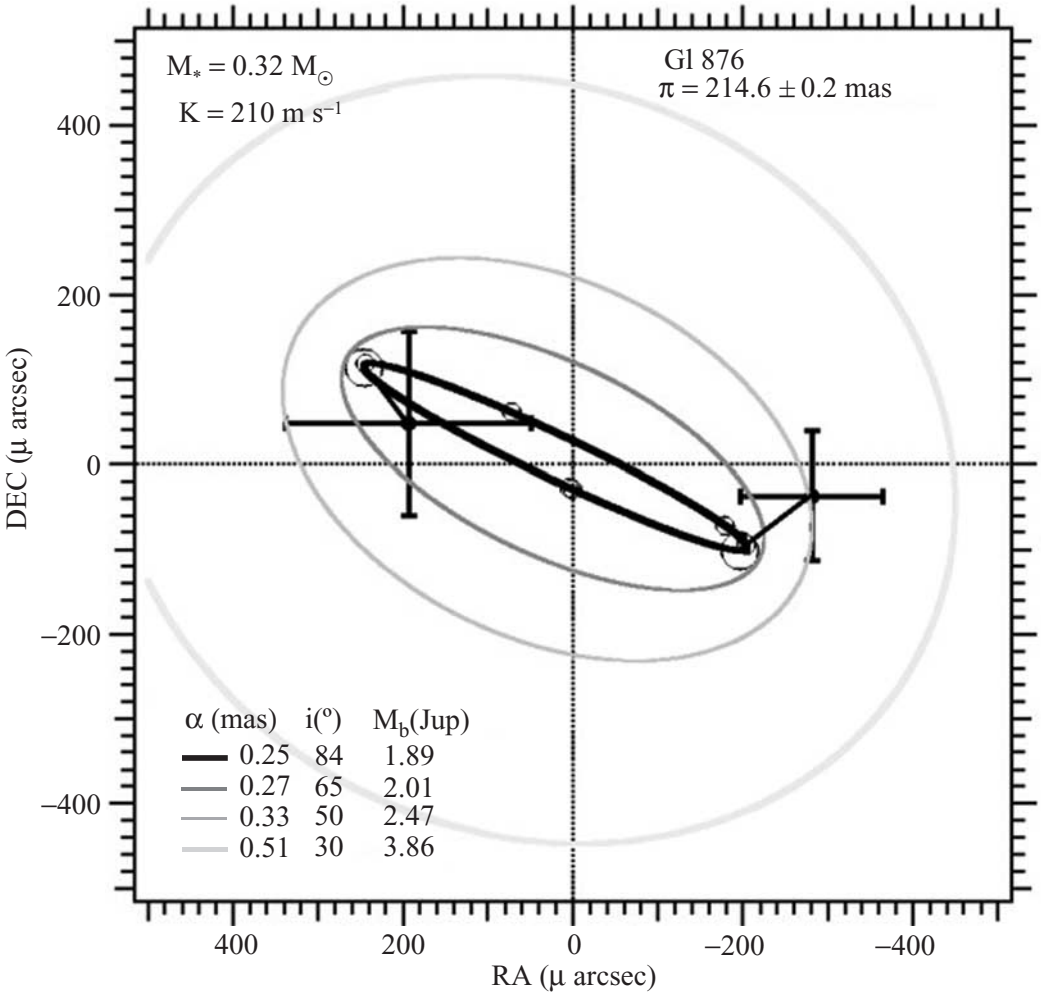

FIGURE 1.4. Astrometric deviations on the plane of the sky measured by the *Hubble Space Telescope* Fine Guidance Sensor in fringe-tracking mode (from Benedict *et al.* 2002). The best fits to the astrometric variations of 0.25 milliarcseconds (the inner ellipsoid, with measurement errors indicated by the crosses) of the star Gliese 876 (with a parallactic mass of about 0.32 solar masses) gives a planetary orbital inclination of about 84 degrees, and thus a planet mass of about 1.89 jovian masses.

accurately measure astrometric wobbles of stars, caused by orbiting planets, to about one microarcsecond in angular resolution.

1.6. Imaging

Direct imaging of an extrasolar planet at visible wavelengths depends on the reflected light from the star that the planet produces which, in turn, depends on its distance from the star, the planet's size, and the nature of its atmosphere (i.e. the product of the geometric albedo, A, and particle phase function $p(\varphi(t))$, which is a measure of the light-scattering nature of the particles in the planetary atmosphere, such as a Lambertian function or, more likely, a steeper function of viewing angle). The ratio of brightness of

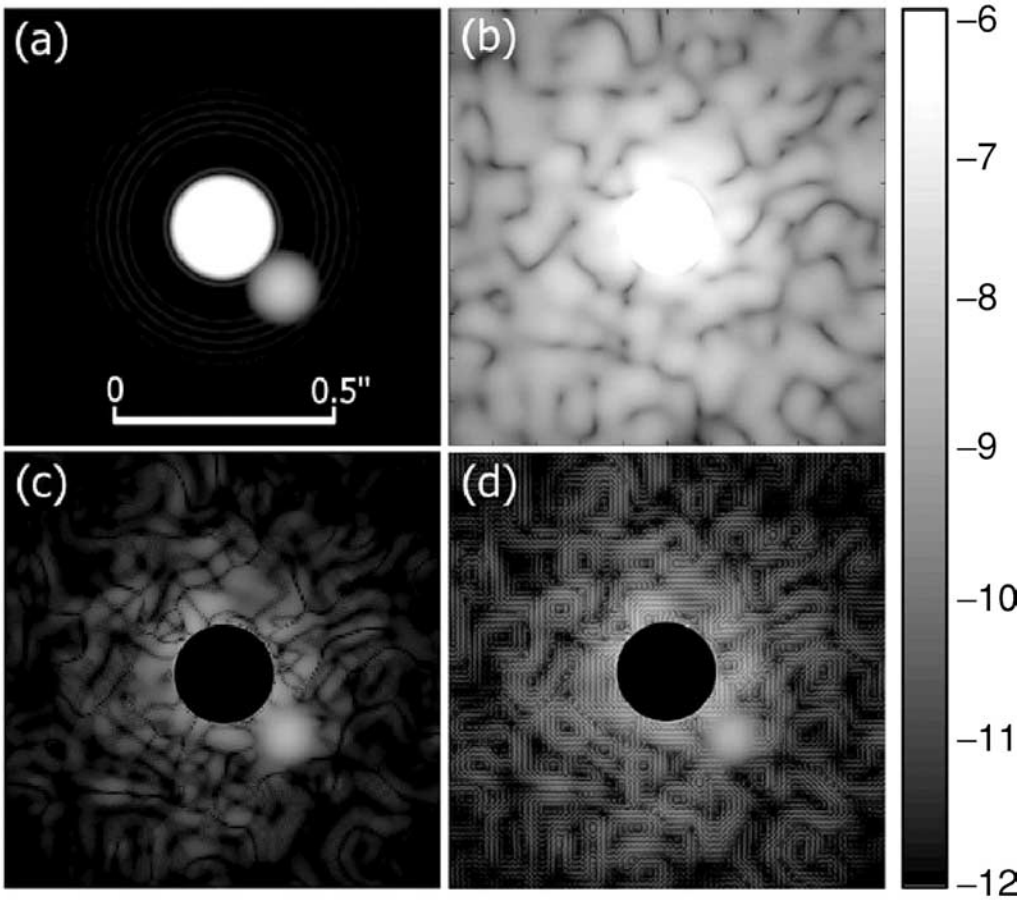

FIGURE 1.5. (a) A star that is ten-thousand times brighter than a companion planet at 0.2 arcsecond separation within the halo of the star (faint rings) with ideal apodization (from Codona and Angel 2004). (b) The system with low-order residuals in the nulling process. (c) Additional destructive interference using an anti-halo technique. (d) Same process as (c) but with spatial light modulation.

the planet to the star is the important factor as the planet, even at a reasonably large angular distance, $a_{\rm p}/d$, from the star, will be 'lost' in the brightness of the diffraction rings of the star as imaged by a telescope (Jupiter viewed from Alpha Centauri would be about 4 arcseconds in angular distance away from the Sun, but typical angular distances of exoplanets are much smaller). The brightness ratio of planet-to-star is:

$$L_{\rm p}/L_* = Ap(\varphi(t)) \left(\frac{R_{\rm p}}{a_{\rm p}}\right)^2, \tag{1.7}$$

where one would want to image a planet at maximum distance from the parent star. This would occur at phases of $\varphi(t) = 90$ and 270 degrees. The maximum separation of the planet and its parent star (see Figure 1.5) is simply (a/d). For $d = 5$ parsecs, at visible wavelengths $L_{\rm p}/L_*$ for a CEGP would be about 16 magnitudes (with a star–planet separation of 0.01 arcseconds), for a 'Jupiter' about 23 magnitudes (separation of about 1 arcsecond), and for an 'Earth' about 25 magnitudes (separation of about 0.2 arcseconds). The contrast between star and planet may improve as one moves into the infrared, if the planet is a warm or hot one. In these cases the planet not only reflects

the stellar light but also emits significant thermal IR radiation. The scattered light from the star can, however, be blocked somewhat by a coronagraphic telescope, and this is the approach being taken by, for example, the first part of the NASA *Terrestrial Planet Finder* (*TPF-C*) mission. The second *TPF* mission (*TPF-I*) is planning to use nulling interferometry, as does the European *Darwin* mission. In that approach the telescope (or telescopes) can be designed such that the light paths from the star will destructively interfere, thus largely cancelling out the stellar flux. The planet, being in a different angular locale, will not undergo destructive interference, and the system may even be designed to allow a planet in a definite position (in the circumstellar habitable zone, for example) to constructively interfere. Significant improvement in ground-based imaging may also be realized with the further development of new techniques. One of several interesting examples is 'dark speckle' imaging in which atmospheric scintillation may cause random destructive interference around a star, allowing the planet image to emerge momentarily, and such images – taken very rapidly – may be summed to give a planetary image (Labeyrie 1995).

1.7. Radio flux

Jovian extrasolar planets with sufficient magnetic field (i.e. rotation rates and metallic cores to produce a dynamo) can emit significant flux at radio wavelengths. The flux produced by a planet can be characterized by:

$$F_{\rm p}(\nu)\left(\frac{a\nu}{cd}\right), \tag{1.8}$$

where ν is the radio frequency being observed, $F_{\rm p}(\nu)$ is the radio flux from the planet, and d is the distance to the star system. Jupiter, for example, would produce a flux density of about 0.3 μJy (micro-Janskys) at a distance of 4 parsecs at a wavelength of 1 mm (synchrotron radiation; Jones 1994). A CEGP at this distance may be expected to produce less than about 0.03 μJy of flux due to being tidally locked in rotation with periods more on the order of several days, thus decreasing the dynamo effect. An Earthlike planet may also be expected to produce a similar strength signature to a CEGP (see Figure 1.6).

For interferometric detection techniques, the flux ratio between that of the planet and the star, $F_*(\nu, t)$, at a given detection frequency and time is the important limiting criterion:

$$F\left(\frac{\text{planet}}{\text{star}}\right) = \left(\frac{F_{\rm p}(\nu,t)}{F_*(\nu,t)}\right), \tag{1.9}$$

where the time dependence refers to both the radio fluctuations in the star and planet, and also to the planet's orbital position for maximum angular separation from the star (i.e. elongation at orbital phases of $\varphi(t) = 90$ and 270 degrees for inclinations not too close to zero degrees).

At 10 MHz this ratio, for a jovian planet around a solar-type star, is about 4×10^{-3} during active phase, but could be as high as 4 during a typical quiet starspot cycle – that is, the planet is brighter than the star at this time! However, interstellar electrons add substantial noise to the detection of flux at this wavelength over any appreciable distance. This is even more the case at lower frequencies. However, at a frequency of about 100 kHz the flux ratio of a jovian-type planet to its star could be as high as 100 during active stellar starspot phases, and as high as 2000 at the quiescent phase of the stellar activity cycle. The proposed square kilometre array (SKA) may have some possibilities of detecting the nearest jovian-type planets in this way. The SKA would be an array of

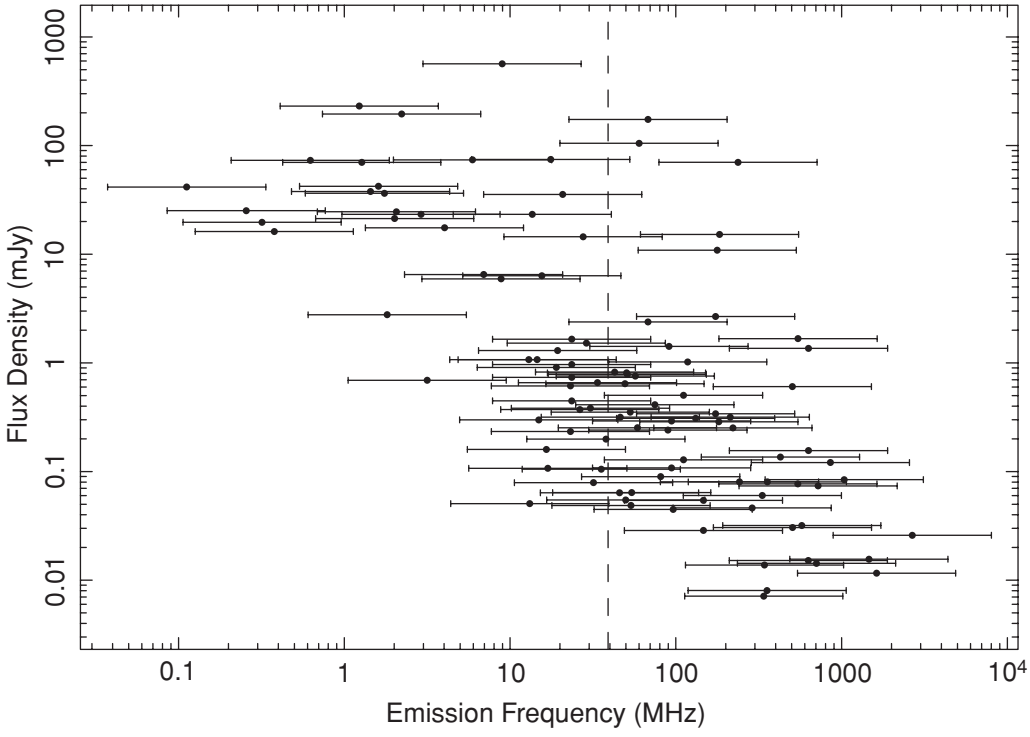

FIGURE 1.6. Theoretical 'burst' flux densities for 106 extrasolar planets (from Lazio *et al.* 2004). The range of expected frequencies is given with their expected flux densities (in milli-Janskys).

centimetre-to-metre wavelength radio telescopes making up a total collecting area of one million square metres with 50% of the collecting area within five square kilometres (for sensitivity), 25% within 150 kilometres (intermediate resolution), and the remaining 25% of the array out to as far as 3000 kilometres for very high angular resolution.

Current radio technology on Earth emits at narrow-band (less than 1 Hertz wide) microwave frequencies in the range 1–10 GHz, and such signals can be many millions of times more powerful than the natural flux from the Sun. This is the basis for the radio searches for extraterrestrial intelligence (SETI) projects (Tarter, 2001). Having but one example of such technological development, constraints on the expected success of detecting such signals are few. One can nevertheless state that at present such SETI projects remain the most unambigous way proposed to detect exobiology since no source in interstellar space is known to produce such narrow-band radio signals (OH masers having a bandpass of several hundred hertz, for example).

1.8. Transit photometry

Although there is a significant coverage in this volume on extrasolar planet detection by transit photometry (see Chapter 3), it is centred on the detection of transits across single stars. Thus, in this section we shall add a discussion of transits across close double star systems – in particular, eclipsing binaries.

1.8.1. *Single-star transits*

Photometric transit detection of extrasolar planets actually detects the shadow of the planet as it crosses in front of the stellar disc. The probability of such an alignment