

Radial velocities

THE MOTION OF A SINGLE PLANET in orbit around a star causes the star to undergo a reflex motion about the star–planet *barycentre* (centre of mass). This results in the periodic perturbation of three observable properties of the star, all of which have been detected (albeit typically in different systems): in radial velocity, in angular (or astrometric) position on the sky (Chapter 3), and in the time of arrival of some periodic reference signal (Chapter 4).

The first convincing exoplanet detection was made by radial velocity measurements in 1995. By the end of 2017, 662 planets in 504 systems (102 of which are multiple) have been discovered by this technique, with a dozen or so which also transit their host star.

2.1 Orbits and orbit fitting

2.1.1 Description of orbits

As in all orbiting systems, both star and planet orbit the common system barycentre. Under the inverse square law of gravity¹, each moves in a closed elliptical orbit in inertial space, with the centre of mass at one focus (Figure 2.1). Such an ellipse is described in polar coordinates (with respect to a focus) by

$$r = \frac{a(1 - e^2)}{1 + e \cos v}, \quad (2.1)$$

or in Cartesian coordinates (with respect to the centre)

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1, \quad (2.2)$$

with the semi-major axis a and the semi-minor axis b related to the *eccentricity* e by

$$b^2 = a^2(1 - e^2). \quad (2.3)$$

¹The relations given in this section, their derivations, and more extensive dynamical considerations, can be found in various recent texts (e.g., Murray & Dermott 2000, Chapter 2; Hilditch 2001; Cole & Woolfson 2002, Section M) as well as in a number of earlier authoritative treatments of orbits (e.g. Aitken, 1918; Binnendijk, 1960; Batten, 1973; Heintz, 1978a; Roy, 1978).

[The *ellipticity*, $\eta = (a - b)/a = 1 - \sqrt{(1 - e^2)}$, is an alternative measure of non-circularity, not used further.]

The *pericentre* distance q and *apocentre* distance Q are given by²

$$\begin{aligned} q &= a(1 - e), \\ Q &= a(1 + e). \end{aligned} \quad (2.4)$$

A line through a focus and parallel to the minor axis defines the *latus rectum*, with the *semi-latus rectum* (related to the planet's angular momentum) of length

$$p = a(1 - e^2). \quad (2.5)$$

True, eccentric, and mean anomaly Various angles in the orbit plane, referred to as ‘anomalies’ (the word dating from the time that planetary motions were considered anomalous), are used to describe the position of a planet along its orbit at a particular time (Figure 2.1).

The *true anomaly*, $v(t)$, also frequently denoted $f(t)$, is the angle between the direction of pericentre and the current position of the body measured from the barycentric focus of the ellipse. It is the angle normally used to characterise an observational orbit.

The *eccentric anomaly*, $E(t)$, is a corresponding angle which is referred to the *auxiliary circle* of the ellipse. The true and eccentric anomalies are geometrically related by

$$\cos v(t) = \frac{\cos E(t) - e}{1 - e \cos E(t)}, \quad (2.6)$$

or, equivalently,

$$\tan \frac{v(t)}{2} = \left(\frac{1+e}{1-e} \right)^{1/2} \tan \frac{E(t)}{2}. \quad (2.7)$$

²An *apsis* (plural *apsides*) is the point on the orbit of minimum distance (*pericentre*, or *periapsis*) or maximum distance (*apocentre*, or *apoapsis*) from the barycentre. The line connecting the two is the *line of apsides*, which defines the orbit's major axis. Derivative terms refer to the body being orbited: *perigee/apogee* around the Earth, *perihelion/aphelion* around the Sun, *periastron/apastron* around a star, *peribac/apobac* around a barycentre, and *perigalacticon/apogalacticon* around a galaxy. The general terms *pericentre/apocentre* are used preferentially here.

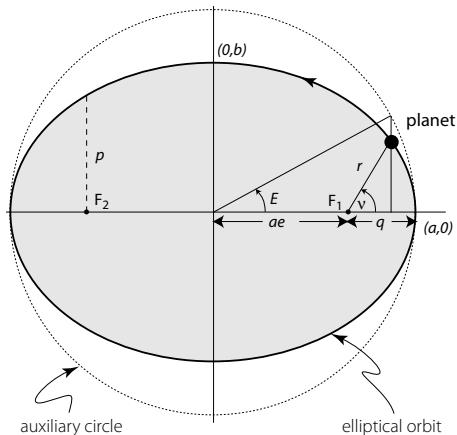


Figure 2.1: Geometry of an elliptical orbit. Points on the orbit can be described in terms of the true anomaly v (with respect to the ellipse) or the eccentric anomaly E (with respect to an auxiliary circle of radius equal to the semi-major axis a). Focus F_1 is the system barycentre, F_2 is the ‘empty’ focus.

The *mean anomaly*, $M(t)$, is an angle related to a fictitious mean motion around the orbit, used in calculating the true anomaly. Over a complete orbit, during which the real planet (or the real star) does not move at a constant angular rate, an average angular rate can nevertheless be specified in terms of the *mean motion*

$$n \equiv 2\pi/P, \quad (2.8)$$

where P is the orbital period. The mean anomaly at time $t - t_p$ after pericentre passage is then defined as

$$M(t) = \frac{2\pi}{P}(t - t_p) \equiv n(t - t_p). \quad (2.9)$$

The relation between the mean anomaly, $M(t)$, and the eccentric anomaly, $E(t)$, can be derived from orbital dynamics. This relation, *Kepler’s equation*, is given by

$$M(t) = E(t) - e \sin E(t). \quad (2.10)$$

The position of an object along its orbit at any time can then be obtained by calculating the mean anomaly M at that time from Equation 2.9, (iteratively) solving the transcendental Equation 2.10 for E , and then using the geometrical identity Equation 2.6 to obtain v .

An analytical approach, which yields partial derivatives which can be used in various fitting and data analysis methods (including Fisher analysis of co-variances, uncertainties and correlations) is given by Pál (2009).

Orbit specification A Keplerian orbit in three dimensions (Figure 2.2) is described by seven parameters: $a, e, P, t_p, i, \Omega, \omega$. The first two, a and e , specify the size and shape of the elliptical orbit. P is related to a and the component masses through Kepler’s third law (see below), while t_p corresponds to the position of the object

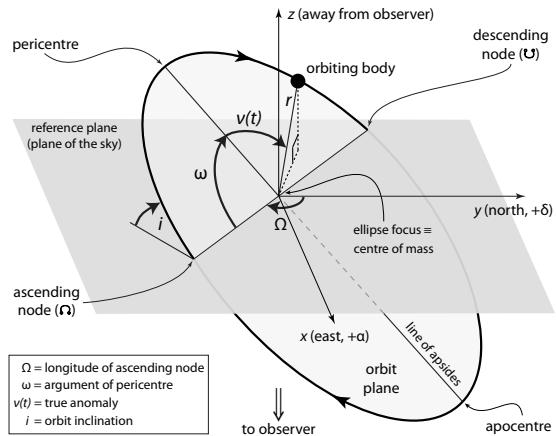


Figure 2.2: An elliptical orbit in three dimensions. The reference plane is tangent to the celestial sphere, i is the inclination of the orbit plane, and the nodes define the intersection of the orbit and reference planes. Ω is the longitude of the ascending (receding) node, measured in the reference plane. ω is the fixed angle defining the object’s argument of pericentre relative to the ascending node. The true anomaly, $v(t)$, is the time-dependent angle characterising the object’s position along the orbit. The right-handed xyz coordinate system has x towards east (increasing a), y towards north (increasing δ), and z away from the observer, consistent with the Thiele–Innes constants (Equations 3.18–3.20). The reference axis, y , contrasts with the use of x , the vernal equinox, as adopted for the solar system.

along its orbit at a particular reference time, generally with respect to a specified pericentre passage.³

The three angles (i, Ω, ω) represent the projection of the true orbit into the observed (apparent) orbit; they depend solely on the orientation of the observer with respect to the orbit. With reference to Figure 2.2:

- i specifies the *orbit inclination* with respect to the reference plane, $0 \leq i < 180^\circ$. $i = 0^\circ$ corresponds to a face-on orbit. For binary star orbits, motion is referred to as prograde (in the direction of increasing position angle on the sky, irrespective of the relation between the rotation and orbit vectors) if $i < 90^\circ$, retrograde if $i > 90^\circ$, and projected onto the line of nodes if $i = 90^\circ$.
- Ω specifies the *longitude of the ascending node*, mea-

³A few remarks are in order: (i) some texts state that just six parameters are required, and omit P , implicitly invoking the relation between a and P (and the component masses) as given by Kepler’s third law; (ii) a is the semi-major axis of the orbiting body with respect to the system barycentre, assumed here to be in linear measure unless otherwise noted. If a is determined in angular measure, as in the relative astrometry of binary stars, the system distance d (equivalently the parallax ω) is required to establish the linear scale; (iii) the parameters of the two co-orbiting bodies (e.g. a star and planet) with respect to the barycentre are identical, with the exception of their a which differ by a factor M_p/M_\star , and their ω which differ by 180° ; (iv) in general, the semi-major axis of the true orbit does not project into the semi-major axis of the apparent orbit.

sured in the reference plane. It is the node where the object moves away from the observer through the plane of reference, i.e. where z goes from negative to positive. [For solar system objects, it is the node where an orbiting object moves north through the plane of reference.]

- ω specifies the *argument of pericentre*. It is the angular coordinate of the object's pericentre relative to its ascending node, measured in the orbital plane and in the direction of motion. [For $e = 0$, where pericentre is undefined, $\omega = 0$ can be chosen such that t_p gives the time of nodal passage.]

Three other angles with respect to the adopted reference direction are used in the specification of orbits:

$$\tilde{\omega} = \Omega + \omega \quad \text{the } \textit{longitude of pericentre} \quad (2.11)$$

$$\theta = \tilde{\omega} + \nu \quad \text{the } \textit{true longitude} \quad (2.12)$$

$$\lambda = \tilde{\omega} + M \quad \text{the } \textit{mean longitude} \quad (2.13)$$

Since Ω and ω are measured in different planes, the longitude of pericentre, $\tilde{\omega}$, is a 'dog-leg' angle. The true longitude and mean longitude are correspondingly offset with respect to the true anomaly and mean anomaly, respectively. Despite its name, the mean longitude is again a linear function of time and, as for the mean anomaly, has only an auxiliary geometrical interpretation.⁴

Kepler's laws Kepler's three laws of planetary motion are: (1) the orbit of a planet is an ellipse with the Sun at one focus; (2) the line joining a planet and the Sun sweeps out equal areas in equal intervals of time; (3) the squares of the orbital periods of the planets are proportional to the cubes of their semi-major axes.

The first and third laws are consequences of the inverse square law of gravity, while the second follows from conservation of angular momentum (and is true for any radial law of attraction). Kepler's laws originally

referred to relative orbits with respect to the Sun, but corresponding formulations apply also to 'absolute orbits' defined with respect to the barycentre.

For the general two-body problem where the mass of the secondary is not neglected, both orbits are ellipses with their foci at their common barycentre (Figure 2.3). Kepler's third law takes the general form

$$P^2 = \frac{4\pi^2}{GM} a^3, \quad (2.14)$$

with M and a taking different values according to the type of orbit being measured:

(a) relative orbits: the motion of the planet, now relative to the star rather than the barycentre (Figure 2.3), can be found by applying an acceleration to the system which cancels that of the star, viz. GM_p/r^2 where r is the instantaneous star–planet separation

$$r = \frac{a(1-e^2)}{1+e\cos\nu}. \quad (2.15)$$

Then

$$P^2 = \frac{4\pi^2}{G(M_\star + M_p)} a_{\text{rel}}^3, \quad (2.16)$$

where the coordinate origin is now the star, not the barycentre, and a_{rel} is the semi-major axis of the relative orbit, i.e. of the planet around the star.⁵

For $M_p \ll M_\star$ and in units of Earth's orbit of 1 au

$$P \approx 1 \text{ yr} \left(\frac{a_{\text{rel}}}{\text{au}} \right)^{3/2} \left(\frac{M_\star}{M_\odot} \right)^{-1/2}. \quad (2.17)$$

(b) absolute orbits: the orbit of the star around the system (star–planet) barycentre is given by

$$P^2 = \frac{4\pi^2}{GM'} a_\star^3, \quad (2.18)$$

where

$$M' \equiv \frac{M_p^3}{(M_\star + M_p)^2}, \quad (2.19)$$

and a_\star is the semi-major axis of the stellar orbit around the system barycentre. An equivalent expression gives the orbit of the planet around the system barycentre in

⁴ Various forms of Figure 2.2 are given in the literature. Indeed, this version differs from that in the first edition which, following Batten (1973), erroneously showed Ω measured in the direction north through west. With reference to the following standard texts (Lennart Lindegren, priv. comm.): Aitken (1918) provided no figure, but his text is consistent with Figure 2.2. Binnendijk (1960), Figure 25, is drawn for negative i , i.e. with Ω showing the position angle of the *descending* rather than the *ascending* node, following the convention in visual binary work of choosing the node that gives $\Omega < 180^\circ$, and indicating the nodal ambiguity by assigning $i = \pm X^\circ$ (with $X > 90^\circ$ for a retrograde orbit). The positive sign would then mean that the node is ascending, negative that it is descending. If the ambiguity is subsequently resolved by radial velocity data, and it turned out that Ω referred to the descending node, one could either (1) retain Ω and assign $i < 0^\circ$, or (2) increase Ω by 180° , increase ω by 180° , and assign $i > 0^\circ$ (and so preserving the 6 Thiele–Innes constants, §3.4.5). Batten (1973), Figure 1.3, erroneously shows Ω drawn from north through west, whereas it should be measured from north through east, as stated in his text (pp 9–10). Heintz (1978a), Figure 7, is wrong in the same way as Batten (1973).

⁵This measurement of *relative* separation does not arise for exoplanet orbits when the planet is unseen. It is, in contrast, a situation relevant for the relative astrometry of binary stars, where an orbit is measured as a separation and position angle of one star with respect to another; then, the combined system mass can be determined if P and a_{rel} are measurable, while individual masses can only be determined if the mass ratio can be established, either from the ratio of the distances from the barycentre, or the ratio of their speeds around it.

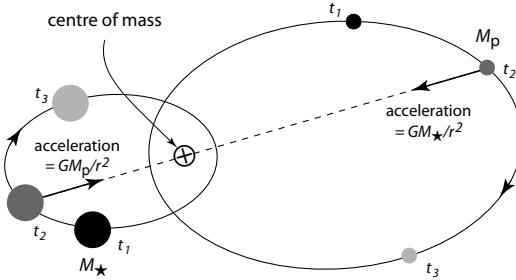


Figure 2.3: Two orbiting bodies, shown at times t_1, t_2, t_3 , move about their common barycentre. Both bodies follow orbits having the same shape and period, but of different sizes and with ω differing by 180° .

terms of the semi-major axis of the planet around the system barycentre a_p .

It follows that the sizes of the three orbits are in proportion $a_\star : a_p : a_{\text{rel}} = M_p : M_\star : (M_\star + M_p)$, with $a_{\text{rel}} = a_\star + a_p$. Furthermore, $e_{\text{rel}} = e_\star = e_p$, $P_{\text{rel}} = P_\star = P_p$, the three orbits are co-planar, and the orientations of the two barycentric orbits (ω) differ by 180° .

Since the planet is assumed to be invisible (the case for essentially all radial velocity observations at the present time, although see §2.5.4), the orbital motion of the star around the barycentre is only correctly determined by astrometry if its position is measured with respect to an ‘absolute’ (quasi-inertial) reference frame.

That all seven orbital elements are accessible to astrometric measurements for an arbitrary projection geometry essentially follows from two principal considerations: the star position versus time allows the maximum and minimum angular rates to be determined, and hence the position of the line of apsides. With the orientation of the major axis so established, appeal to Kepler’s second law fixes the orbit inclination.

Radial velocity measurements of the host star also give information on its barycentric orbital motion, although not all seven Keplerian orbital elements are accessible from the line-of-sight velocity variations alone.

2.1.2 Orbits from radial velocities

Radial velocity (Doppler) measurements describe the projected motion, along the line-of-sight, of the primary star as it orbits the system barycentre.

Radial velocity semi-amplitude With Figure 2.2 now considered as representing the orbit of the star around the barycentre, the star’s z -coordinate along the line-of-sight can be derived from trigonometry

$$z = r(t) \sin i \sin(\omega + \nu), \quad (2.20)$$

where $r(t)$ is the distance from the barycentre. Then

$$v_r \equiv \dot{z} = \sin i [\dot{r} \sin(\omega + \nu) + r \dot{\nu} \cos(\omega + \nu)]. \quad (2.21)$$

Some algebraic substitutions for r and \dot{r} lead to

$$v_r = K [\cos(\omega + \nu) + e \cos \omega], \quad (2.22)$$

where the *radial velocity semi-amplitude* is given by⁶

$$K \equiv \frac{2\pi}{P} \frac{a_\star \sin i}{(1 - e^2)^{1/2}}. \quad (2.23)$$

From Equation 2.22, it can be seen that v_r varies around the orbit between limits of $K(1 + e \cos \omega)$ and $K(-1 + e \cos \omega)$. This expression for v_r as a function of the true anomaly $\nu(t)$ can be transformed into an expression for v_r as a function of time through Equations 2.6–2.10.

The shape of the radial velocity curve is determined by e and ω (Figure 2.4). Together with P , their combination constrains the value of $a_\star \sin i$ (Equation 2.23). But neither a_\star nor $\sin i$ can be determined separately.

Two alternative expressions for K are instructive. Substituting Equations 2.18 and 2.19 into 2.23 gives

$$K^2 = \frac{G}{(1 - e^2)} \frac{1}{a_\star \sin i} \frac{M_p^3 \sin^3 i}{(M_\star + M_p)^2}. \quad (2.24)$$

With the product $a_\star \sin i$ determined as above, it follows that radial velocity measurements provide a value for the *mass function* (i.e. Equation 2.19 modified by a dependency on the orbit inclination)

$$\mathcal{M} \equiv \frac{M_p^3 \sin^3 i}{(M_\star + M_p)^2}. \quad (2.25)$$

For $M_p \ll M_\star$ this reduces to

$$\mathcal{M} \approx \frac{M_p^3 \sin^3 i}{M_\star^2}. \quad (2.26)$$

Further, if M_\star can be estimated from its spectral type and luminosity class (or otherwise), then $M_p \sin i$ can be determined. The mass of the planet nevertheless remains uncertain by the unknown factor $\sin i$.

Equations 2.18, 2.19 and 2.23 can be combined to give an alternative expression for K without the explicit appearance of a_\star (e.g. Cumming et al., 1999, eqn 1)

$$K = \left(\frac{2\pi G}{P} \right)^{1/3} \frac{M_p \sin i}{(M_\star + M_p)^{2/3}} \frac{1}{(1 - e^2)^{1/2}}, \quad (2.27)$$

which can also be written (Torres et al., 2008, eqn 1)

$$\frac{M_p \sin i}{M_J} = 4.919 \times 10^{-3} \left(\frac{K}{\text{m s}^{-1}} \right) (1 - e^2)^{1/2} \left(\frac{P}{\text{days}} \right)^{1/3} \left(\frac{M_\star + M_p}{M_\odot} \right)^{2/3}. \quad (2.28)$$

⁶Often denoted K_1 in binary star work to emphasise the motion of the primary, this amplitude of radial motion is analogous to the projected semi-major axis measured astrometrically.

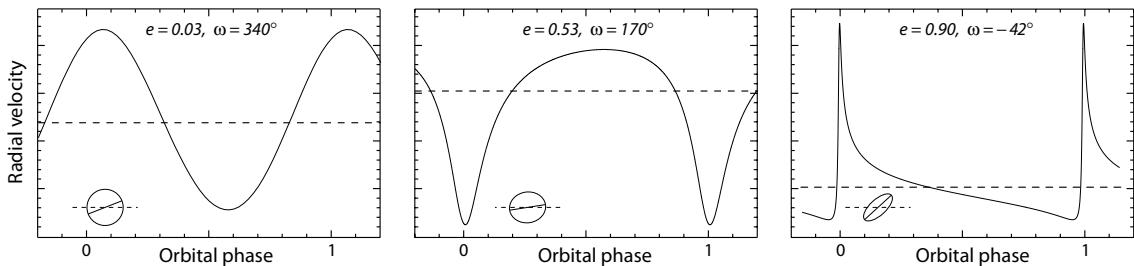


Figure 2.4: Example stellar radial velocity curves, illustrating their dependence on e and ω , for HD 73256 (Udry et al., 2003a, Figure 2), HD 142022 (Eggenberger et al., 2006, Figure 4), and HD 4113 (Tamuz et al., 2008, Figure 1). Horizontal dashed lines show the systemic velocity (viz. the radial velocity of the barycentre). The ellipses at lower left show the viewing geometries.

For a circular orbit with $M_p \ll M_\star$, for example, the stellar velocity variations are sinusoidal with amplitude

$$K = 28.4 \text{ m s}^{-1} \left(\frac{P}{1 \text{ yr}} \right)^{-1/3} \left(\frac{M_p \sin i}{M_J} \right) \left(\frac{M_\star}{M_\odot} \right)^{-2/3}. \quad (2.29)$$

For Jupiter around the Sun ($a = 5.2$ au, $P = 11.9$ yr) $K_J = 12.5 \text{ m s}^{-1}$. For Earth, $K_\oplus = 0.09 \text{ m s}^{-1}$.

Keplerian observables Of the 7 elements describing an orbit in three-dimensions ($a, e, P, t_p, i, \Omega, \omega$), Ω cannot be determined from radial velocity measurements (cf. Figure 2.2). Furthermore, only the combination $a_\star \sin i$ is determined, with neither a_\star nor $\sin i$ individually.

Radial velocity measurements alone do not provide individual planetary masses, their sum, or their ratio.⁷ Taken alone, small periodic radial velocity modulation is insufficient to infer, unambiguously, the presence of a companion of planetary mass. Without additional information, such as an astrometric orbit, a constraint on the orbit inclination i from photometric transits, spectroscopic line profiles, or statistical deconvolution (e.g. Jorissen et al., 2001), small values of K could indicate either a low-mass planet, or an object of significantly higher mass with small orbital inclination, i.e. with the plane of the orbit almost face-on to the line-of-sight.

Periodogram analysis In single- or multiple-planet fitting, preliminary period estimates are often used to simplify the search, for example using the Lomb–Scargle algorithm (Lomb, 1976; Scargle, 1982). This modified periodogram analysis, equivalent to least-squares fitting of sine waves, targets the efficient and reliable detection of a periodic signal in the case of unevenly-spaced observation times and in the presence of noise.

⁷For double-lined spectroscopic binaries with two distinct spectra (not the case for typical exoplanet orbits, but see §2.5.4), the mass function for each component can be established separately. Then the component mass ratio can also be estimated, but still neither the separate masses, nor the orbit inclination. Further inferences can be made if the system is eclipsing (such that $i \approx 0$), or if the system can also be resolved astrometrically.

Although not generally as robust as Bayesian methods in selecting and quantifying signals (e.g. through MCMC, §2.1.5), they are computationally efficient and are good at signal visualisation (Figure 2.5).

Variations Variations of the Lomb–Scargle periodogram have been developed to variously account for measurement errors (e.g. Horne & Baliunas, 1986; Gilliland & Baliunas, 1987; Irwin et al., 1989; Baluev, 2008a; Vio et al., 2013); a frequency-dependent mean (e.g. Cumming et al., 1999; Zechmeister et al., 2009); non-sinusoidal functions (e.g. Breithorst, 2001; Cumming, 2004); multiple periodic signals (e.g. Anglada-Escudé & Tuomi, 2012; Baluev, 2013a; Jenkins et al., 2014; Hara et al., 2017); Bayesian formalism using marginalised likelihoods, BGLS (Mortier et al., 2015), and for identifying periodicities caused by stellar activity (Mortier & Collier Cameron, 2017).

Most searches account for white noise through weighting by measurement errors. Algorithms have also been developed to treat correlated ‘red noise’ (e.g. in time or wavelength), such as RedFit for paleoclimate data (Schulz & Mudelsee, 2002), and Agatha for radial velocities (Feng et al., 2017a).

A comparative evaluation of the the generalised Lomb–Scargle periodogram (GLS), the modified Bayesian formalism (BGLS), and the multifrequency periodogram scheme (FREDEC, for frequency decomposer; Baluev, 2013a) was carried out by Pinamonti et al. (2017). For more complex signals, the residual analyses of GLS and BGLS yielded 30% false alarms, compared to some 10% for FREDEC.

2.1.3 Single planet fitting

There are five observables related to the star’s Keplerian orbit which can be estimated for each planet on the basis of radial velocity measurements alone: e , P , t_p , and ω , and the combination $K = f(a, e, P, i)$.

Two further terms are usually taken into account: a *systemic velocity*, γ , describing the constant component of the radial velocity of the system’s centre of mass relative to the solar system barycentre, which may also include an instrument-dependent radial velocity offset; and a linear trend parameter, d , which may accommodate instrumental drifts as well as unidentified contributions from massive, long-period companions.

From Equation 2.22, the radial velocity of a star with an orbiting planet can then be expressed as a function of

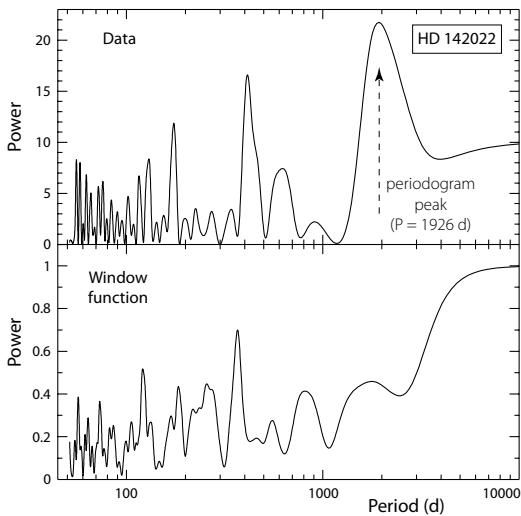


Figure 2.5: Lomb–Scargle periodogram analysis applied to radial velocities for HD 142022 from CORALIE. The power indicates the signal’s statistical significance, not its true amplitude. The window function, determined by the observation epochs, is shown below. The highest peak corresponds to $P = 1926$ d, while the best-fit solution gives $P = 1928$ d. From Eggenberger et al. (2006, Figure 2), reproduced with permission © ESO.

the true anomaly as (e.g. Wright & Howard, 2009, eqn 1)

$$v_r(t) = K [\cos(\omega + \nu(t)) + e \cos \omega] + \gamma + d(t - t_0). \quad (2.30)$$

Establishing the astrometric or radial velocity parameters for a single orbiting planet is typically based on the search of parameter space using χ^2 minimisation.

Zechmeister & Kürster (2009) give an analytic solution for the generalisation to a full sine wave fit, taking account of measurement errors and an offset. They include a specific algorithm to search for the period of the best-fit Keplerian orbit to radial velocity data. Other analytical methods are described by Delisle et al. (2016).

2.1.4 Multiple planet fitting

Kinematic fitting For a system of n_p planets, the total radial velocity signal can be approximated as a linear sum over n_p terms of the form contained in Equation 2.30, giving a total of $5n_p + 1$ Keplerian parameters to be fit, including γ (and optionally d).

This first-order approach to fitting multiple systems considers that the reflex motions of the star caused by each planet are independent, i.e. ignoring the effects of planet–planet gravitational interactions. In this kinematic (or Keplerian) fitting, the dominant planet signal is identified, its Keplerian contribution subtracted from the observational data, and the process repeated until all significant planet signals have been accounted for. A more rigorous multi-planet χ^2 fit to the original data can then be made using these results as starting values.

Such an approach was adopted for the five-planet fit to the 18 years of Doppler data for 55 Cnc, with weights assigned to account for signal-to-noise, photospheric motions, and instrument errors (Fischer et al., 2008).

Linearisation An alternative to least-squares fitting of Keplerian motions to the radial velocity measurements is to seek linear parameters which can be constructed from certain of the orbital elements.

Konacki & Maciejewski (1999) described a harmonic component analysis which they applied to 16 Cyg. This is based on a Fourier expansion of the Keplerian motion in which the harmonic coefficients are functions of all orbital elements. The coefficients, obtained by linear least-squares, are then used as starting estimates for a local minimisation of the non-linear problem.

Wright & Howard (2009) formulated a linearisation which reduces the search space for n_p planets from the $5n_p + 1$ Keplerian parameters to $3n_p$ non-linear variables corresponding to (e, P, t_p) for each planet. In this approach, Equation 2.30, generalised to the case of n_p planets, is re-cast as (Wright & Howard, 2009, eqn 4–7)

$$v_r(t) = \sum_{j=1}^{n_p} [h_j \cos \nu_j(t) + c_j \sin \nu_j(t)] + v_0 + d(t - t_0), \quad (2.31)$$

where, for each planet,

$$h_j = K_j \cos \omega_j, \quad c_j = -K_j \sin \omega_j, \quad v_0 = \gamma + \sum_{j=1}^n K_j e_j \cos \omega_j.$$

The non-linear terms (e, P, t_p) are then searched for algorithmically (e.g. using the Levenberg–Marquardt or Markov Chain Monte Carlo methods), with the linear parameters (h, c, v_0) , which are transformable back to (ω, K, γ) , solved for analytically at each search step.

Compressed sensing techniques In an extension of the more advanced periodogram-type analyses described above, new techniques from the field of signal processing are also being applied. Exploiting developments in the analysis of sparse signals, often referred to as *compressed sensing* or *compressive sampling*, Hara et al. (2017) detail a technique which circumvents ‘brute-force’ minimisation (cf. Equations 2.30–2.31) and searches for the presence of all planets simultaneously and rapidly, essentially exploiting the *a priori* information that the representation of an exoplanet signal in the Fourier domain is sparse. They demonstrated the method with published radial velocity data for 55 Cnc, GJ 876, HD 10180 and HD 69830.

Dynamical fitting In numerous multi-planet systems, gravitational interaction between two or more planets further modifies the total radial velocity or ‘astrometric signature’ as a function of time. Such interaction can result in detectable variations in the planetary orbits even

over short intervals measured in years. A more complete dynamical (or Newtonian) fit involving N-body integrations must then be used.

Various self-consistent algorithms, which incorporate mutual perturbations in fitting the radial velocity data, were developed for the first resonantly interacting system GJ 876, resulting in substantially improved fits to the radial velocity data (Laughlin & Chambers, 2001; Rivera & Lissauer, 2001; Nauenberg, 2002a; Lee & Peale, 2002). The latter was verified against full hydrodynamic evolution of embedded planets by Kley et al. (2004).

Even good dynamical fits can result in orbital parameter solutions which might be stable over years or decades, but unstable on time scales comparable to the age of the planetary system. Since such formal solutions may be considered as implausible, short-term and long-term dynamical stability is a further constraint that must be satisfied by multiple planet orbit fitting. The issue is considered further in §2.12.

Degeneracies In inadequately-sampled data, model ambiguities can frequently arise (e.g. Morais & Correia, 2011). Ambiguities between high eccentricity and multiple planet fits can also arise from various degeneracies (Wittenmyer et al., 2013c): (a) between a single planet on an eccentric orbit and two planets on circular orbits in a 2:1 configuration (Anglada-Escudé et al., 2010a); (b) between an eccentric planet and two co-orbital planets, i.e., in a 1:1 resonance or Trojan pair (Laughlin & Chambers, 2002; Giuppone et al., 2012a); (c) between a single eccentric planet and a circular planet with a long-period companion (Rodigas & Hinz, 2009); and (d) between a single eccentric planet and two near-circular planets with poorly sampled periods (e.g. HD 159868, Wittenmyer et al., 2012c).

2.1.5 Bayesian methods

As described in Sections 2.1.3 and 2.1.4, orbit parameters and their uncertainties are typically obtained by period searches using methods such as the Lomb–Scargle periodogram to establish the orbital period(s), and then estimating other parameters using minimisation algorithms. In practice, there are significant statistical challenges, which are complicated in the case of multiple planets by various degeneracies between orbital parameters. Determining the *number* of planets favoured by a particular data set turns out to be a non-trivial problem, posing an even greater computational challenge.

Bayesian methods, with their associated prior probabilities (see box, page 24), aim for a more rigorous approach, with advantages over more traditional methods if, for example, the data do not cover a complete orbit (e.g. Anglada-Escudé & Tuomi, 2012). They also allow a more rigorous model selection, most notably the number of planets favoured by the data (e.g. Brown, 2004;

Clyde et al., 2007; Ford & Gregory, 2007; Cumming & Dragomir, 2010; Faria et al., 2016a).

Further background to the use of Bayesian conditional probabilities for exoplanet orbit modeling, example reference priors, and the sensitivity of the detection criteria to prior choice, can be found in various studies (e.g. Tuomi, 2012; Tuomi & Anglada-Escudé, 2013).

Priors and search space Tuomi & Anglada-Escudé (2013), for example, take their reference set of priors to be the Keplerian parameters (radial velocity amplitude K , longitude of pericentre ω , eccentricity e , mean anomaly M_0 , and $\log P$), along with a constant reference velocity γ , a radial velocity jitter σ_j (viz., the excess white noise on top of the estimated instrument uncertainties), and a correlation coefficient between the noise of subsequent epochs ϕ (as defined by Tuomi et al., 2013a,b). Typically, but not exclusively, these may be uniform over a specific interval. The planets are, in this case, assumed to be gravitationally interacting. The choice of $\log P$ improves the sampling efficiency of parameter space, valid because of its scale invariance (Ford, 2006b).

A key problem in the application of Bayesian methods to exoplanet searches is the extensive parameter space that must be searched, and the resulting computational cost, especially in calculating the Bayesian ‘evidence’. The development of Markov Chain Monte Carlo techniques (MCMC, box, page 25) has made Bayesian methods, and probabilistic methods in general, more tractable (e.g. Gregory, 2005; Ford, 2005, 2006b; Balan & Lahav, 2009; Gregory, 2011a; Hou et al., 2012; Foreman-Mackey et al., 2013).

Nested sampling (Skilling, 2004) is a Monte Carlo method for efficient calculation of the Bayesian evidence, along with posterior inferences. It is the basis of the MultiNest algorithm (Feroz & Hobson, 2008; Feroz et al., 2009a,b), which provides efficient sampling in problems that contain multiple modes and/or large degeneracies. Details of this approach in the context of exoplanet detection are given by Feroz et al. (2011b). It has been applied, for example, to 47 UMa, HD 10180, and HD 37124 (Feroz et al., 2011b), HIP 5158 (Feroz et al., 2011a), and CoRoT-7 (Faria et al., 2016a).

Acceptance criteria In the comparison of models with different numbers of planetary companions, Tuomi (2011) proposed that the probability density of the k -th planet should have a unique maximum that can be interpreted as a Keplerian signal (and not caused by data gaps or noise). Furthermore, the probability of confidently finding a k -th planetary signal requires that $P(z_k) \gg P(z_{k-1})$, where z represents the k Keplerian signals. In practice, to claim that there are k planets orbiting the target star, Tuomi (2011) has proposed that the probability of finding k signals should formally be at least 150 times greater than that of finding $k-1$ signals.

Bayesian inference: Bayesian inference (or analysis) is a method for summarising uncertainty and making estimates and predictions using probability statements conditional on observed data and an assumed model. Qualitatively, it asserts that the posterior probability of a model fit is proportional to the ‘likelihood’ times the ‘prior’.

More precisely (e.g. Feroz et al., 2011b), for a set of parameters Θ , under a model (or hypothesis) H , and for a data set \mathbf{D} , Bayes’ theorem states that the posterior probability distribution of the parameters, $\text{Pr}(\Theta|\mathbf{D}, H) \equiv P(\Theta)$, is given by

$$\text{Pr}(\Theta|\mathbf{D}, H) = \frac{\text{Pr}(\mathbf{D}|\Theta, H)\text{Pr}(\Theta|H)}{\text{Pr}(\mathbf{D}|H)}, \quad (2.32)$$

where $\text{Pr}(\Theta|H) \equiv \pi(\Theta)$ is the prior distribution of parameters, $\text{Pr}(\mathbf{D}|\Theta, H) \equiv \mathcal{L}(\Theta)$ is the sampling distribution, or ‘likelihood’, of the observed data conditional on its parameters, and $\text{Pr}(\mathbf{D}|H) \equiv \mathcal{Z}$ is the distribution of the observed data marginalised (in the probabilistic sense) over the parameter(s), referred to as the ‘marginal likelihood’ or ‘Bayesian evidence’.

In situations of relevance here, such an approach represents an extensive search across a wide parameter space, thus carrying a significant computational penalty.

For *parameter estimation*, the normalising evidence factor can usually be ignored, since it is independent of the parameters Θ . Inferences can be obtained by taking samples from the (non-normalised) posterior using standard MCMC sampling methods where, at equilibrium, the chain contains a set of samples from the parameter space distributed according to the posterior. This posterior constitutes the complete Bayesian inference of the parameter values, and can be marginalised over each parameter to obtain individual parameter constraints.

For *model selection*, the evidence factor is crucial. Numerically, it is the factor normalising the posterior over the parameter set Θ (Feroz et al., 2011b, eqn 2)

$$\mathcal{Z} = \int \mathcal{L}(\Theta) \pi(\Theta) d^D \Theta, \quad (2.33)$$

where D is the dimensionality of parameter space. Evaluation of this multi-dimensional integral is numerically challenging, with evidence evaluation being at least an order of magnitude more costly than parameter estimation alone (Feroz et al., 2011b).

There is an extensive literature on the advantages and disadvantages of Bayesian approaches (reference texts include Box & Tiao, 1992; Berger, 1985). Although less contentious when prior distributions have a physical basis or a plausible scientific model, Bayesian methods nevertheless have their critics (e.g. Gelman, 2008).

Examples Tuomi et al. (2013b) performed Bayesian comparisons of stellar jitter models for τ Cet, favouring one with Gaussian noise superimposed on moving average components with exponential decay over hours to days. This resulted in significant improvements of the statistical models, and enabled the detection of signals with amplitudes below 1 m s^{-1} .

Numerous Bayesian planet searches have been reported (e.g. Gregory, 2005; Balan & Lahav, 2009; Tuomi, 2011; Nelson et al., 2014b; Tuomi et al., 2014; Brewer

& Donovan, 2015; Feng et al., 2016a; Faria et al., 2016a), leading to various candidates, amongst them HD 208487 c (Gregory, 2007a), HD 11964 b–d (Gregory, 2007b), and GJ 676 d (Anglada-Escudé & Tuomi, 2012).

Suggestions have been made for up to 4 planets around GJ 163 (Tuomi & Anglada-Escudé, 2013), up to 6 around HD 40307 (Tuomi et al., 2013a), up to 7 around GJ 667 (Anglada-Escudé et al., 2013c), up to 9 around HD 10180 (Tuomi, 2012), and up to 15 around GJ 581, although not accounting for stellar jitter (Brewer & Donovan, 2015). Many of these have not yet been independently confirmed; those that have can be assessed from the online catalogue compilations.

2.1.6 Algorithmic implementation

Detailed descriptions of the analysis of radial velocity data in practice is widely reported. Examples include procedures for HARPS data (e.g. Doyle et al., 2013; Triaud et al., 2013a); for Lick data (Valenti & Fischer, 2005; Wright & Howard, 2009; Howard et al., 2010a); and for CARMENES data (Trifonov et al., 2018).

Radial velocities and orbits Various algorithms for the determination of radial velocities, or orbits derived from them, have been described, and in several cases routines provided (for some considerations of the effects they consider, see e.g. Wright & Eastman 2014, §5). Some include a dynamical treatment. Routines include:

EXOFAST is a suite of IDL routines designed to fit transits and radial velocity variations simultaneously or separately, and to characterise parameter uncertainties and covariances with a differential evolution Markov Chain Monte Carlo method (Eastman et al., 2013).

ExoSOFT, the Exoplanet Simple Orbit Fitting Toolbox, with routines for fitting any combination of radial velocity and astrometric data (Mede & Brandt, 2017).

Frequency Decomposer is a parallelised algorithm for the decomposition of a noisy time series into a number of sinusoidal components (Balucinska-Church, 2013b).

Keplerian periodogram, in which the signal is modeled by a highly non-linear Keplerian radial velocity function (Balucinska-Church, 2015b), a formulation claiming significant efficiency gains with respect to the Lomb-Scargle periodogram, especially for high-eccentricity orbits.

PlanetPack for the analysis of radial velocity data with unknown jitter, employing multiple Keplerian orbits or dynamical N-body simulations (Balucinska-Church, 2013d,e, 2014).

PyORBIT for radial velocity time series and ancillary data (such as photometry and activity indicators) to simultaneously characterise the orbital parameters and the host star activity (Malavolta et al., 2016).

rvfit for fitting radial velocities of stellar binaries and exoplanets using ‘adaptive annealing’ global minimisation methods (Iglesias-Marzoa et al., 2015b,a).

RVLIN (IDL) for fitting an arbitrary number of Keplerian curves (Wright & Howard, 2012), and BOOTTTRAN for the corresponding errors (Wang & Wright, 2012).

SERVAL, Spectrum Radial Velocity Analyser, is a least-squares algorithm to derive radial velocities and additional spectral diagnostics (Zechmeister et al., 2018).

Spectroscopy Made Easy for synthetic spectral fitting (e.g. Valenti & Piskunov, 1996; Kane et al., 2011c).

Stacked Bayesian Lomb-Scargle periodogram for identifying periodicities caused by stellar activity (Mortier & Collier Cameron, 2017).

Systemic Console for computing Keplerian fits (e.g. Meschiari et al., 2009; Meschiari & Laughlin, 2010; Jones et al., 2013b; Vogt et al., 2014b).

Advanced search schemes Because the equations describing an astrometric or radial velocity orbit are non-linear, unstructured searches of parameter space for multiple planets may be computationally prohibitive, and with many false local χ^2 minima. More sophisticated search schemes are therefore desirable. Amongst those widely used are numerical schemes based on genetic algorithms, Levenberg–Marquardt minimisation, and Markov Chain Monte Carlo methods.

Some early searches used *genetic algorithms* to explore global parameter space (e.g. Goldberg, 1989; Charbonneau, 1995), e.g. for v And (Butler et al., 1999; Stepinski et al., 2000), GJ 876 (Laughlin & Chambers, 2001), 55 Cnc (Marcy et al., 2002), HD 12661 (Goździewski & Maciejewski, 2003, PIKAIA), and μ Ara (Pepe et al., 2007, Stakanof). According to Stepinski et al. (2000) and Goździewski & Maciejewski (2003), the method is inefficient in identifying very accurate best-fit solutions, but provides good starting points for more precise gradient methods such as Levenberg–Marquardt.

The *Levenberg–Marquardt method* is an efficient algorithm for finding a local χ^2 minimum for non-linear models. It varies smoothly between ‘steepest descent’ far from the minimum, and the inverse-Hessian method as a minimum is approached. It has been widely used for both radial velocity and astrometric orbit fitting. Implementations include `mrqmin` (Press et al., 2007).

Due to sparse sampling, measurement errors, parameter degeneracy and model limitations, there are frequently no unique values of the basic model parameters such as period and eccentricity, and a Bayesian approach (§2.1.5) can provide more robust estimates of parameter uncertainties. The *Markov Chain Monte Carlo method* (MCMC) has been widely used for orbit fitting for radial velocity observations, as well as for photometric transit light curves (box, page 25).

Algorithms Amongst available algorithms:

EXOFIT is a Bayesian tool employing a Markov Chain Monte Carlo method, and specifically tailored to exoplanet radial velocity fitting for a one- or two-planet sys-

Markov Chain Monte Carlo, MCMC: Probabilistic data analysis, including Bayesian inference (§2.1.5), has transformed many areas of research in the past decade. Significant gains have come from numerical methods for approximate inference, especially Markov Chain Monte Carlo (MCMC; Metropolis et al., 1953; Hastings, 1953).

Probabilistic data analysis procedures involve computing and using the posterior probability density function for the parameters of the model or the likelihood function. In some cases it is sufficient to find the maximum of one of these, but it is often necessary to understand the posterior probability density function in detail. MCMC methods are designed to sample from, and thereby provide sampling approximations to, the posterior probability density function efficiently, even in parameter spaces with large numbers of dimensions. MCMC links two concepts: the Markov Chain, being a random process with the property that the future depends only on the current state of the process and not the past (i.e. it has no memory); and Monte Carlo methods, which randomly sample from some underlying process to provide an estimate of something too complex or time consuming to find deterministically.

The two concepts are put together to construct a Markov Chain which converges to the desired probability distribution after a number of steps. The state of the chain is used as a sample from the desired distribution, and the process is repeated. MCMC algorithms use different techniques for generating/sampling the Markov Chain (e.g. those of Metropolis–Hastings; Gibbs; or Goodman–Weare).

Numerous MCMC algorithms are used in the exoplanet community, including `emcee` (Foreman-Mackey et al., 2013); MC3 (MC3Cubed), Multi-Core Markov-Chain Monte Carlo (Cubillos et al., 2016b, 2017c); and GEMC, based on DE-MC (Tregloan-Reed et al., 2018). Specific optimisations for exoplanet studies have been developed (e.g. Ford, 2004b; Gregory, 2005; Ford, 2006b; Gregory, 2007a,b; Carter et al., 2011c). Of many example applications, Gregory & Fischer (2010) describe parallel tempering, simulated annealing, and genetic crossover applied to 47 UMa; Gregory (2011a) describe efficient sampling in highly correlated parameter spaces; and Triaud et al. (2011) describe fitting to WASP light curves.

More complete discussion of MCMC methods can be found in, e.g., Mackay (2003), while key concepts in this context are summarised by Foreman-Mackey et al. (2013).

tem (Balan & Lahav, 2009, 2011). A uniform analysis of 94 systems was carried out by Hollis et al. (2012).

Joker: a Monte Carlo sampler for sparse or noisy radial velocity measurements of a 2-body system (Price-Whelan et al., 2017).

MPFIT: an IDL implementation of Levenberg–Marquardt minimisation (Markwardt, 2009). One of many applications to radial velocity fitting is described, for example, by Wright & Howard (2009).

RUN DMC is a parallel code for analysing radial velocity observations using N-body integrations and differential evolution MCMC (Nelson et al., 2014a). Application to 55 Cnc involving both radial velocity and transit observations is described by Nelson et al. (2014b).

2.1.7 Detectability and selection effects

Various studies have been made of planet detectability from radial velocity data as a function of period and signal amplitude (e.g. Nelson & Angel, 1998; Eisner & Kulkarni, 2001b; Cumming, 2004; Narayan et al., 2005).

Cumming (2004) provides analytic expressions for planet detectability as a function of period and eccentricity, and some insight into the detection limits. For short orbital periods, $P \lesssim T$ where T is the duration of the observations, the radial velocity semi-amplitude threshold alone characterises detectability. At long periods, $P \gtrsim T$, the observations cover only part of the orbit by definition, and detectability depends on which part of the orbit is being sampled. If the orbit is close to a velocity maximum/minimum, or to a zero crossing, velocity variations are ‘sine-like’ or ‘cosine-like’ respectively, and the velocity variations are

$$\Delta v = K \sin\left(\frac{2\pi T}{P}\right) \approx K\left(\frac{2\pi T}{P}\right), \quad \text{or} \quad (2.34)$$

$$\Delta v = K \cos\left(\frac{2\pi T}{P}\right) \approx \frac{K}{2} \left(\frac{2\pi T}{P}\right)^2 \quad (2.35)$$

respectively. Averaging over phase introduces a dependency on the adopted detection efficiency, ϵ_D : for $\epsilon_D = 0.5$, the amplitude must be large enough that sine-like phases are detected, but cosine-like phases do not have to be, and the velocity threshold scales as $K \propto P$. For $\epsilon_D = 0.99$, almost all phases must be detected, requiring a large amplitude, which then scales as $K \propto P^2$.

While the dependence of K on e given by Equation 2.27 results in large K for highly eccentric orbits, in practice eccentricity acts to make detection more difficult at short periods, where an uneven sampling often results in poor phase coverage during rapid pericentre passages. At longer periods, the increased velocity amplitude and acceleration near the pericentre increase detectability. The transition to the long-period regime occurs for orbital periods $P \approx T/(1-e)^2$. The analysis also allows the completeness of existing surveys to be assessed (Cumming et al., 2008). The results emphasise that there remains a significant selection effect against detecting eccentric orbits for $e \gtrsim 0.6$.

Shen & Turner (2008) found that, once a planet is detected, the eccentricities derived from Keplerian fitting are biased upwards for low signal-to-noise and moderate numbers of observations. They suggest that the numbers of exoplanets with low-eccentricity may be underestimated in current samples. The effect may be evident in results from Keck, for example, where for $P > 10$ d the observed eccentricity distribution is nearly flat for large-amplitude systems, $K > 80 \text{ m s}^{-1}$, but rises linearly towards low eccentricity at lower amplitudes, $K > 20 \text{ m s}^{-1}$ (Valenti et al., 2009).

Rodigas & Hinz (2009) showed that there is an additional bias due to the presence of an undetected outer

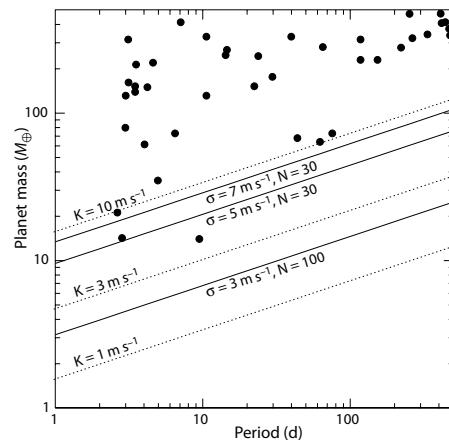


Figure 2.6: Minimum mass limits for a 50% detection threshold as a function of the number of observations, N , and the combined error (measurement and stellar jitter) σ , for $M_{\star} = 1 M_{\odot}$ (Equation 2.36). The observation duration is assumed longer than P_{orb} . Circles are planets known in 2005. Dotted lines show various radial velocity semi-amplitudes, K . From Narayan et al. (2005, Figure 7), by permission of IOP Publishing/AAS.

companion. For moderate eccentricity $0.1 < e < 0.3$, for example, there is a 13% probability that a modeled eccentric orbit is in fact circular, with the model fit confused by the undetected outer companion.

Narayan et al. (2005) evaluated the detection probability for low mass planets, either in isolation or near mean motion resonance with a hot Jupiter. For a 50% detection rate their expression for the minimum detectable mass is

$$M_{\text{p,min}} \sim 4M_{\oplus} \left(\frac{N}{20}\right)^{-\frac{1}{2}} \left(\frac{\sigma}{1 \text{ m s}^{-1}}\right) \left(\frac{P}{1 \text{ d}}\right)^{\frac{1}{3}} \left(\frac{M_{\star}}{M_{\odot}}\right)^{\frac{2}{3}}, \quad (2.36)$$

for $N \gtrsim 20$, where N is the number of observations, and σ is the quadratic sum of the Doppler velocity measurement error and stellar jitter (Figure 2.6).

2.1.8 Scheduling

In a study for the Space Interferometric Mission SIM, Ford (2004a) considered several non-adaptive observing schedules (i.e. schedules fully defined *a priori*) for a targeted astrometric planet search. These included time intervals which are regular periodic (at constant spacing); Golomb ruler⁸; regular power law; regular logarithmic; regular geometric; and random uniform. The efficiency for planet searches was found to be relatively insensitive to the actual observing schedule.

⁸A Golomb ruler has unique integer intervals between each pair of marks, thus providing the maximum number of unique baselines. The number of marks defines its order; it is ‘perfect’ if it includes all distances up to its length, and ‘optimal’ if no shorter ruler of that exists. Thus the (optimal and perfect) Golomb ruler of order 4 and length 6 has marks at 0, 1, 4, 6.

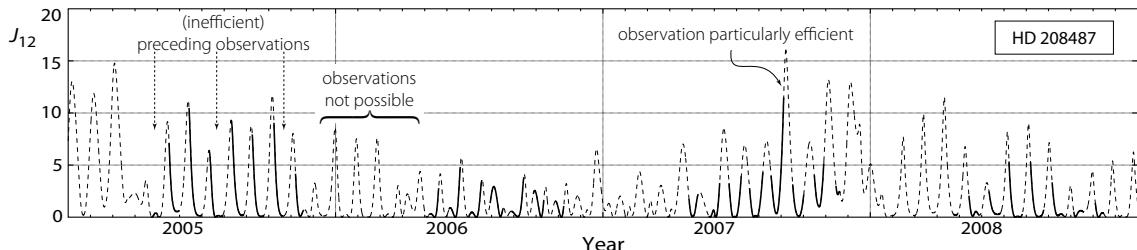


Figure 2.7: Scheduling future radial velocity observations to discriminate between two orbit models for HD 208487. The parts of the curve shown dashed correspond to times when observations are impossible. Observations are optimal when scheduled to coincide with the information peaks J_{12} . Arrows mark the three preceding radial velocity measurements, which in retrospect are seen to have been of little additional information value. From Baluev (2008b, Figure 2), © Oxford University Press.

Adaptive scheduling, in which information from previous observations is used to plan future ones most efficiently, can optimise their information content, and can perform significantly better. The objective is to predict future epochs which yield, for example, maximum improvement in orbit parameters and planet masses, the most favourable epochs for distinguishing between alternative models, or an optimum series of future observations. For multiple planet systems, and in particular for resonant orbits, such optimal timings may be concentrated in rather narrow time intervals.

Examples Loredo (2004) outlined a maximum entropy sampling strategy for determining an exoplanet orbit based on Bayesian inference, and Ford (2008) developed a more rigorous adaptive scheduling algorithm exploiting information theory. Based on a small number of initial observations, this proceeds by assuming a prior for the distribution of orbital periods and masses, using Bayesian inference to calculate the posterior probability distribution for all model parameters, using this to calculate the predictive distribution for the radial velocities at some future time, and finally choosing observing times at which additional observations would be most valuable. Monte Carlo simulations demonstrated that such searches are more efficient in terms of detection, and can measure orbital parameters more accurately, than non-adaptive algorithms.

Baluev (2008b, 2010) used criteria from optimal design theory (based on Fisher information, Shannon information, and Kullback–Leibler divergence) to design similar general scheduling formalisms. As an example, Figure 2.7 illustrates the problem of deciding between two orbital solutions based on the 35 measurements of HD 208487 from Butler et al. (2006b). Wright et al. (2007) identified periodicity attributable to an additional planet at either 28.6 or 900 d, but their data was inadequate to rule between them due to aliasing.

If the probability densities $p_{1,2}(v)$ describe the distribution of the radial velocity predictions for two possible orbit fits, the expectation of the likelihood ratio statistic considering the first or the second model as true

are given by (Baluev, 2008b, eqn 3)

$$I_{2|1} = \int p_1(v) \ln \frac{p_1(v)}{p_2(v)} dv, \quad I_{1|2} = \int p_2(v) \ln \frac{p_2(v)}{p_1(v)} dv \quad (2.37)$$

The maximum information is given by (Baluev, 2008b, eqn 11)

$$\begin{aligned} J_{12} &= I_{2|1} + I_{1|2} \\ &= -1 + \frac{1}{2} \left(\frac{\sigma_1^2}{\sigma_2^2} + \frac{\sigma_2^2}{\sigma_1^2} \right) + \left(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2} \right) \frac{(v_1 - v_2)^2}{2}, \end{aligned} \quad (2.38)$$

under the assumption that the distributions of v_i are close to Gaussian. The largest values of J_{12} then correspond to the most promising times for future observations to rule out one of the alternative models. Although Equation 2.38 is derived by Baluev (2008b) from a formal consideration of information content, its interpretation is straightforward: observations should be made when the two models imply the largest predicted differences in the measured radial velocity, while the uncertainties of these predictions, σ , should be small enough to avoid statistically insignificant differences.

For HD 208487 (Figure 2.7) the radial velocity predictions for the two different orbit models differ by up to 20 m s^{-1} . The function J_{12} identifies the epochs which maximise discrimination between them. It turns out that a single observation at one of the peaks during the 2005 observing season could have decided between competing models at the $\gtrsim 2\sigma$ level, while the actual observations (indicated by arrows) fell at epochs of low information content, and added little new to the contested two-planet solution. The subsequent season (2006) offered limited opportunity for further discrimination.

Further considerations in optimal scheduling are given by Pál (2009). The scheduling of CARMENES observations is described by Garcia-Piquer et al. (2017).

Survey optimisation Bottom et al. (2013) examined statistical and instrumental effects with the goal of maximising planet detection yields for a fixed observing time. They showed that G and K dwarfs observed at 400–600 nm are optimum for surveys targeting completeness to a given M_p and P_{orb} , while M dwarfs observed at 700–800 nm are optimum for habitable-zone planets.

2.2 Measurement principles

2.2.1 Doppler shifts

An instantaneous measurement of the stellar radial velocity about the star–planet barycentre is given by the small, systematic Doppler shift of the many absorption lines that make up the stellar spectrum.

If, in the observer's reference frame, the source is receding with velocity v at an angle θ relative to the observer–source direction, the change in wavelength

$$\Delta\lambda = \lambda_{\text{obs}} - \lambda_{\text{em}}, \quad (2.39)$$

is related to the velocity by the expression for the relativistic Doppler shift (e.g. Lang, 1980, eqn 2–226)

$$\lambda_{\text{obs}} = \lambda_{\text{em}} \frac{(1 + \beta \cos\theta)}{(1 - \beta^2)^{1/2}}, \quad (2.40)$$

where λ_{obs} , λ_{em} are observed and emitted wavelengths, and $\beta = (v/c)$. For $v \ll c$ and $\theta \ll \pi/2$, the expression reduces to the classical form

$$v_r = v \cos\theta \approx \left(\frac{\Delta\lambda}{\lambda_{\text{em}}} \right) c, \quad (2.41)$$

where, conventionally, positive values indicate recession. Special relativistic terms correspond to changes in v_r of several m s^{-1} , and are therefore significant. Equation 2.40 omits the effect of the refractive index of air at the spectrograph, n_{air} (1.000 277 at standard temperature and pressure, STP), which introduces errors of $\lesssim 1 \text{ m s}^{-1}$ (Marcy & Butler, 1992, eqn 3).

With the radial velocity semi-amplitude values of $K_J \simeq 12.5 \text{ m s}^{-1}$ and $K_\oplus \simeq 0.09 \text{ m s}^{-1}$ as indicative goals, the detection of planets around solar-type stars has demanded long-term radial velocity accuracies of some 15 m s^{-1} or preferably significantly better, corresponding to an accuracy of a few parts in 10^8 in wavelength. This must be maintained over months or years.

High-accuracy radial velocities for exoplanet detection are typically acquired using échelle spectrographs with high spectral resolving power (typically $R \equiv \lambda/\Delta\lambda \sim 50000 - 100000$), and operated in the optical region (450–700 nm). Many diffraction orders are cross-dispersed, and recorded simultaneously on rectangular format CCDs providing large numbers of resolved absorption lines (Figure 2.8). The principles of an échelle spectrograph are described by, e.g., Vogt (1987) in the case of the Lick Observatory Hamilton spectrometer, by Vogt et al. (1994) in the case of Keck-HIRES, and by Baranne et al. (1996) in the case of ELODIE.

High instrumental stability and accurate wavelength calibration is demanded to minimise effects of gravitational and thermal telescope flexure, and other instrument drifts. Large telescopes and long integration times are still required to achieve the necessary high signal-to-noise, and corresponding sub-pixel accuracies.

2.2.2 Spectral resolution

For unresolved spectral lines, measurement precision increases with improved spectral resolution (Bouchy et al., 2001). In the photon-noise limited regime the error ϵ on the line-centre measurement can be approximated by (Pepe et al., 2014a)

$$\epsilon = \frac{\sigma^{1.5}}{(2I_0)^{0.5} \cdot \sigma \mathcal{C}} \left(1 - \frac{\mathcal{C}}{2} \right)^{\frac{1}{2}}, \quad (2.42)$$

where σ is the spectral line width observed through the spectrograph, $\mathcal{C} = (I_{\text{min}} - I_0)/I_0$ is the line contrast (where I_0 and I_{min} are the photoelectron counts per resolution element in the continuum and line minimum respectively), and $\sigma \mathcal{C}$ is the equivalent width.

A spectral resolution of at least $R \equiv \lambda/\Delta\lambda = 100000$ has been found to provide the best precision on slowly-rotating, quiet, solar-type stars. High spectral resolution and adequate line sampling provide better signal-to-noise per spectral line, and also lead to smaller instrumental errors in both the radial velocity measurement and the calibration process (Bouchy et al., 2001).

Some recent instruments for 8-m class telescopes or larger have adopted significantly higher resolutions of $R = 200000 - 300000$, notably PEPSI (optical) and iLocater (near infrared), both for the LBT, and complementary in their wavelength coverage (Table 2.4).

Ultra-high spectral resolution Throughput is an important requirement, and ultra-high resolution spectrographs (e.g. the $R = 10^6$ AAT-UHRS, Diego et al. 1995; or the $R = 600000$ échelle at Steward Observatory, Ge et al. 2002a) do not provide the efficiencies of $\gtrsim 10\%$ needed for competitive planet searches.

Extremely large telescopes To first order, instrument errors scale with the resolution element expressed in wavelength units (Pepe et al., 2014a). However, with increasing telescope size, spectral resolution becomes a cost driver. For seeing-limited instruments, the optical étendue (i.e. beam cross section area times solid angle, $E = A\Omega$) increases with telescope size, and so does the instrument size for a given spectral resolution. For 8-m and extremely-large telescopes, this aspect has become a technical and managerial challenge requiring novel, compact designs (Pasquini et al., 2010; Pepe et al., 2014b; Szentgyorgyi et al., 2012; Marconi et al., 2016).

2.2.3 Cross-correlation spectroscopy

Information about the instantaneous Doppler shift is contained in the many thousands of absorption lines in the high-resolution optical spectrum of solar-type stars. This information can be concentrated into a few parameters by cross-correlation, even at low signal-to-noise (Figure 2.9). This involves multiplying the stellar spectrum by a weighted (binary) mask, and finding the minimum of the product as a function of the Doppler shift.

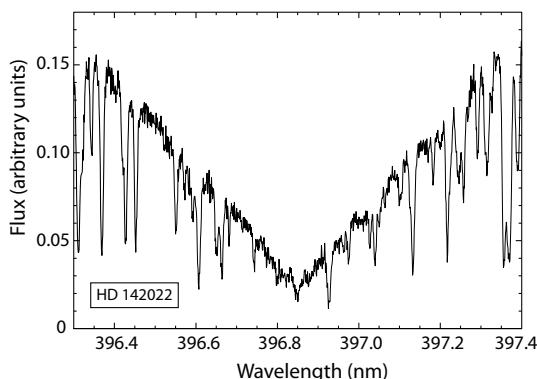


Figure 2.8: The Ca II H absorption-line region of the HARPS échelle spectrum for HD 142022, illustrating the density of lines in 1 nm of the 378–691 nm spectrum. From Eggenberger et al. (2006, Figure 1), reproduced with permission © ESO.

The technique, originally implemented using a physical mask, was proposed by Fellgett (1955), first demonstrated by Griffin (1967), and extended to échelle spectroscopy using CORAVEL by Baranne et al. (1979). As Fellgett phrased it, ‘It is uneconomical, both in telescope time and in labour of reduction, to observe the details of a stellar spectrum if the sole object of the observation is to measure a radial velocity.’

Limitations of a fixed physical template, in accuracy and adaptability to spectral type, led to the cross-correlation subsequently implemented numerically as a box-shaped template (Queloz, 1995; Baranne et al., 1996). Finer details of the method have been developed progressively (e.g. Simkin, 1974; Sargent et al., 1977; Tonry & Davis, 1979; Bender, 1990). The essentials are to determine the value of ϵ minimising (Queloz, 1995)

$$C(\epsilon) \propto \int_{-\infty}^{+\infty} S(v) M(v - \epsilon) dv, \quad (2.43)$$

where S is the spectrum and M is the mask, both expressed in velocity space v . Associated errors are established from Monte Carlo modeling. Weighting according to the relative line depths further optimises the signal-to-noise, and can also reduce the perturbing effects of telluric lines (Pepe et al., 2002).

The precise shape of the resulting *cross-correlation function* depends on the intrinsic spectral line shapes and on the template line widths, and overall represents a mean profile of all lines in the template. Accordingly, in addition to the radial velocity, the width yields the stellar rotational velocity $v \sin i$, while the equivalent width provides a metallicity estimate if T_{eff} is known approximately (Mayor, 1980; Benz & Mayor, 1981; Queloz, 1995). In the absence of systematic line asymmetry, the underlying shape of the cross-correlation function is well approximated by a Gaussian, with asymmetry reflecting systematic structure in the individual spectral lines.

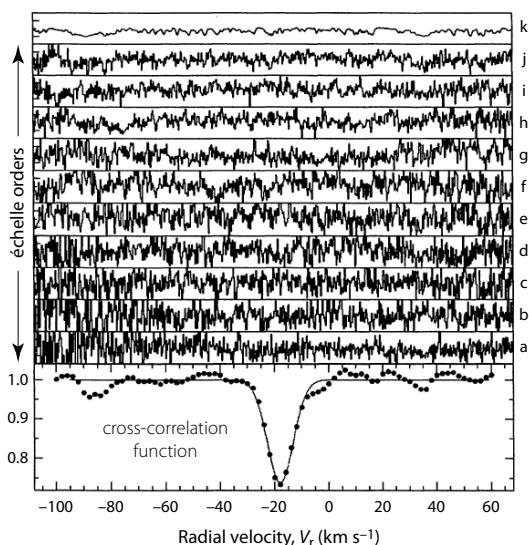


Figure 2.9: Example cross-correlation function for a K0 III star with $S/N \approx 1$. Observations at $R = 40\,000$ span 411–444 nm in ten orders, each covering 4 nm. Some 1000 lines match the template. The top trace is the same order as that immediately below, but at $S/N = 40$. The cross-correlation function is shown at bottom. From Queloz (1995, Figure 2).

For HARPS data, an alternative to cross-correlation is based on least-squares matching of each observed spectrum to a high signal-to-noise ratio template derived from the same observations. Thus, HARPS TERRA (Template-Enhanced Radial velocity Re-analysis Application) reports a significant improvement in accuracy, especially for M dwarfs (Anglada-Escudé & Butler, 2012), where a mean activity jitter of 2.3 m s^{-1} has been reported for a sample of 78 stars (Perger et al., 2017a).

2.2.4 Determination of barycentric velocities

The measured Doppler shift, Equation 2.39, in practice includes effects other than the line-of-sight velocity of the target star’s centre of mass. Contributions from the motion of the observer around the solar system barycentre, due to Earth rotation and orbital motion, are time varying, significant at levels of up to 0.5 and 30 km s^{-1} respectively, and must be accounted for.

Higher-order relativistic and secular projection effects, and additional spectroscopic line shifts introduced by gravitational redshift and stellar surface effects such as convective flow, pulsations, and star spots (Dravins, 1975; Nidever et al., 2002; Lindegren & Dravins, 2003) may be less critical in determining the relative radial velocities central to exoplanet detection. But there are situations where their contributions must be considered, and all must be accounted for in any *absolute* determination of radial space motion at the m s^{-1} level. The magnitude of these effects is as follows.

Barycentric correction To detect changes in a star's radial velocity due to an orbiting planet, measured velocities must be referred to a non-moving frame, or at least one with a constant rectilinear space motion. The *solar system barycentre* (i.e. the solar system's centre of mass, as opposed to the Sun's centre of mass or *heliocentre*) is consequently adopted as reference. For details see, e.g., Wright & Eastman (2014, §3.2).

The time-varying motion of the Earth around the barycentre, which includes gravitational perturbations from the other planets, is described by the *solar system ephemerides* provided by JPL (e.g. Konopliv et al., 2006) and IMCCE (e.g. Fienga et al., 2008, 2009, 2014). By adjusting for known effects, residual velocity terms were brought below 1 m s^{-1} (cf. Stumpff, 1980), and now routinely to below 0.01 m s^{-1} (Wright & Eastman, 2014).

Nonetheless a spurious one-year signal in the HARPS data, at an amplitude of a few m s^{-1} , was only recently traced to the deformation of spectral lines crossing block stitchings of the CCD as the Earth orbits the barycentre (Dumusque et al., 2015b).

Earth rotation, precession, nutation These effects are important at levels of $\sim 0.01 \text{ m s}^{-1}$. They are discussed, for example, by Wright & Eastman (2014, §6.2.1).

Stellar space motion Effects of a star's radial motion on the secular evolution of its parallax and proper motion are treated by Dravins et al. (1999), and encapsulated in Equations 3.13–3.14. Equivalently, a constant space motion results in a changing radial velocity with time due to the changing projection geometry.

Knowledge of the star's parallax ϖ , proper motion μ , and systemic radial velocity γ , allows prediction of this contribution to any secular evolution, which might otherwise be attributed to long-period planetary orbits.

For Barnard's star, observed with VLT-UVES over five years (Figure 2.10), Kürster et al. (2006) measured a secular radial velocity acceleration consistent with the predicted value of $4.50 \text{ m s}^{-1} \text{ yr}^{-1}$ based on the Hipparcos proper motion and parallax, combined with the absolute radial velocity of -110.5 km s^{-1} (Nidever et al., 2002). An increase of $4.515 \pm 0.002 \text{ m s}^{-1} \text{ yr}^{-1}$ was determined from Lick and Keck observations over 25 years between 1987 and 2012 (Choi et al., 2013a).

Gravitational redshift The contribution from the star's *gravitational redshift* is given by (e.g. Misner et al. 1973, eqn 25.26N; Lang 1980, eqn 2–234)

$$v_r \simeq \frac{GM_\star}{R_\star c}. \quad (2.44)$$

This is valid in the Newtonian limit, $R_\star \gg R_S$, where the Schwarzschild radius $R_S \equiv 2GM_\star/c^2$. The contribution amounts to 636 m s^{-1} for the Sun, and ranges from $\sim 680 \text{ m s}^{-1}$ at F5V to $\sim 500 \text{ m s}^{-1}$ at M5V, and several tens of km s^{-1} for white dwarfs. Effects due to *changes* in gravitational redshift are discussed in Section 2.5.1.

Stellar rotation Stellar rotation imposes small radial velocity effects (Gray, 1999), while the contribution of variable meridional flows may be comparable (Beckers, 2007). More critical at high accuracies are the effects of surface features such as star spots, which can induce a radial velocity amplitude of a few m s^{-1} (e.g. Saar & Donahue, 1997; Hatzes, 2002), resulting in periodic modulation at the stellar rotation period (§2.4.5).

Stellar convection Convective motion in the photospheres of cool stars leads to spectral line asymmetries (§2.4.6). A net blueshift results from the contribution of photons from the larger and brighter photospheric granules compared with the downward motion in the darker and cooler intergranular lanes.

The effect is of order -0.5 km s^{-1} for the Sun and other solar-type stars, and ranges from -1000 m s^{-1} at F5V to about -200 m s^{-1} at K0V (Dravins et al., 1981). Effects due to pressure broadening are $\lesssim 100 \text{ m s}^{-1}$ for main sequence stars (Dravins, 1999). Corrective approaches are detailed by Meunier et al. (2017).

Stellar activity This is considered in Section 2.4.5.

Observatory-related effects Also relevant are potential timing errors (box, page 104) and, at the highest target accuracies, effects such as atmospheric chromatic aberration (Blackman et al., 2017), and errors associated with imperfect knowledge of the telescope position.

Higher-order source effects Various higher-order effects, including gravitational redshift variations due to change in stellar radius, Zeeman effect, and planet-induced tides, are considered in Section 2.5.3.

Radial velocity zero point As a result of these astrophysical and instrumental effects, establishing the zero point for absolute radial motions at $\lesssim 50 \text{ m s}^{-1}$ remains difficult. Nidever et al. (2002) determined barycentric radial velocities for 889 late-type stars observed at Keck and Lick with typical accuracies of 0.3 km s^{-1} , and found a difference in zero point of 53 m s^{-1} compared with the system of the 38 stable FGK stars adopted as standards by the Geneva group (Udry et al., 1999a,b).

Practical implementation For the detection of Earth analogues, with reflex motion $\sim 0.09 \text{ m s}^{-1}$, collective correction of the above effects to $\lesssim 0.01 \text{ m s}^{-1}$ is implied. Such accuracies are targeted with E-ELT-HIRES, at least in the context of the 'Sandage test' (§2.6.6).

Wright & Eastman (2014) describe the theory and implementation of many of the above terms at levels of $\sim 0.01 \text{ m s}^{-1}$, including those arising from the Earth's rotation and orbital motion, the relative motion of the solar system with respect to target star, relativistic effects, atmospheric chromatic aberration, and errors associated with imperfect knowledge of the telescope position, timing, stellar position and space motion.

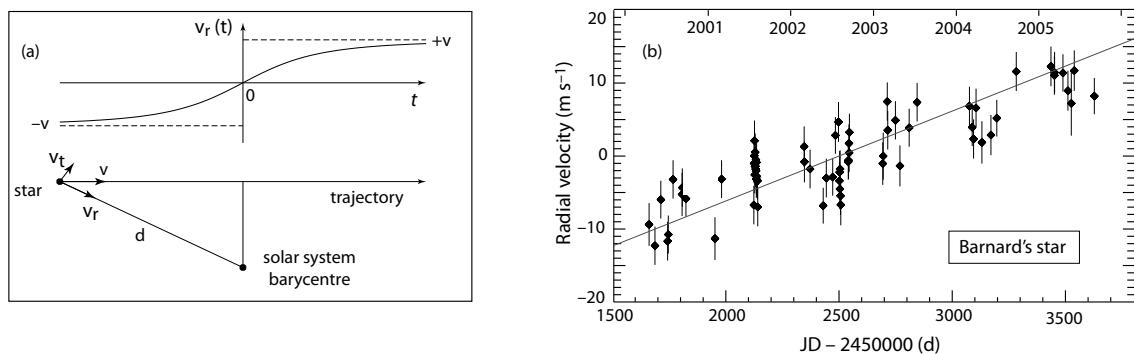


Figure 2.10: Geometry of the space motion of a nearby star. In (a), bottom panel, $g v_t$ and v_r are, respectively, the transverse and radial components of the space velocity v , and d is the current distance; (a, top) absolute radial velocity as a function of time and its asymptotic limits $\pm v$. (b) barycentric radial velocity measurements of Barnard's star over five years, compared with the predicted secular acceleration (solid line). Adapted from Kürster et al. (2003, Figure 1) and Kürster et al. (2006, Figure 3).

2.3 Wavelength calibration

Accurate wavelength calibration is a prerequisite for reaching high radial velocity accuracy. In the early 1970s, accuracies of around 1 km s^{-1} were limited by photographic plate technology, and by guiding errors at the spectrograph slit which introduced shifts in the stellar spectrum relative to comparison arc lines.

2.3.1 Telluric lines

Wavelength calibration was significantly improved by the use of telluric (atmospheric) water vapour lines (Griffin, 1973; Griffin & Griffin, 1973; Walker et al., 1973; Gray & Brown, 2006), largely eliminating errors caused by the different optical paths of the stellar beam and the calibration lamp. Disadvantages include limited spectral ranges where telluric lines of suitable strength are found, in addition to various systematic errors arising from variable path length as a function of zenith distance, producing variations of 20 m s^{-1} per air mass, and from significant winds, producing systematic shifts of 20 m s^{-1} (Gray & Brown, 2006).

Analysis of HARPS archival data over 6 years for the bright stars τ Cet, μ Ara, and ϵ Eri suggested a long-term stability of atmospheric lines of $\sim 10 \text{ m s}^{-1}$. A simple atmospheric model yielded short-term stability of $\sim 2 \text{ m s}^{-1}$ (Figueira et al., 2010b).

2.3.2 Gas cells

The use of captive gases to provide a dense and accurate wavelength reference, superimposed on the stellar spectral lines, started with the use of hydrogen fluoride (HF). Although toxic and corrosive, its 3–0 vibration band gave a well-spaced line distribution, with no isotopic confusion, and of similar natural width to those in typical stellar spectra (Campbell & Walker, 1979).

An alternative, iodine (I_2), also mononuclidic, was used by Beckers (1976) and Koch & Woehl (1984) for solar observations, and later by Marcy & Butler (1992) for their radial velocity programme at the Lick Observatory 3-m telescope (a retrospective is given by Beckers, 2005). It has a strong line absorption coefficient, and requires a path length of only a few cm. Accuracies improved accordingly, to around 25 m s^{-1} by the early 1990s, and to some 3 m s^{-1} just a few years later (Butler et al., 1996).

The gas cell is placed in the telescope light path, just before the spectrograph slit. Sharp absorption lines of known wavelength are superimposed on the stellar spectrum, thereby providing calibration of both wavelength and the spectrograph point-spread function. For an échelle, the latter is complex with many degrees of freedom (Valenti et al., 1995).

The spectrum through the absorption cell is then modeled as (Marcy & Butler, 1992, and their figure 1)

$$I_{\text{obs}}(\lambda) \propto [I_\star(\lambda + \Delta\lambda_\star) T_{I_2}(\lambda + \Delta\lambda_{I_2})] \otimes \text{PSF}, \quad (2.45)$$

where $\Delta\lambda_\star$ and $\Delta\lambda_{I_2}$ are the shifts of the star spectrum and iodine transmission function, determined by least-squares fitting to the composite spectrum, I_{obs} ; and $\otimes \text{PSF}$ represents convolution with the spectrograph point-spread function. I_\star , via $I_\star \otimes \text{PSF}$, is obtained *a priori* by observing each star without the iodine cell in place. The iodine cell transmission function, T_{I_2} , is obtained from external measurements using a high-resolution Fourier transform spectrometer. The final corrected Doppler shift is then given by $\Delta\lambda = \Delta\lambda_\star - \Delta\lambda_{I_2}$.

Modified calibration sequences have been developed. Iterating from an existing spectral template of a similar star can eliminate the need for an observed stellar template spectrum (Johnson et al., 2006c). Modified calibration sequences to measure binary stars, and in particular double-lined spectroscopic binaries, are described by Konacki (2005b).

The iodine gas cell remains the standard in many instruments (including AAT-UCLES, HET-HRS, Keck-HIRES, Lick 3 m, Magellan-MIKE, OAO-HIDES, Subaru-HDS, Tautenburg, TNG-SARG, and VLT-UVES), being of particular relevance for spectrographs which are not intrinsically highly stabilised. Alternative implementation strategies also exist (Johnson et al., 2006c).

The advantages of the iodine absorption cell are the large number (several thousand) of absorption lines, and the common path of the starlight and the iodine absorption. The measured iodine lines therefore simultaneously track changes in the instrument point-spread function. The disadvantages are the 20–30% loss of light, and the clustering of the absorption-line bands in the visible (500–620 nm), making wavelength calibration for redder M dwarfs more problematic.

2.3.3 Emission lamps

ELODIE (Baranne et al., 1996), and its successor HARPS (Mayor et al., 2003), used a thorium–argon (Th–Ar) emission lamp as the reference wavelength spectrum. In practice, two optical fibers are used to transfer light to the spectrograph, one collecting the stellar light, the other simultaneously recording either a Th–Ar reference spectrum, or the background sky.

The advantages of the Th–Ar lamp for wavelength calibration are the large numbers of strong emission lines over a wide optical to infrared range (Palmer & Engleman, 1983; Hinkle et al., 2001; Lovis & Pepe, 2007; Kerber et al., 2008), and the improvement in throughput due to the absence of the iodine absorption cell.

In addition to the use of the Th–Ar lamps, CARMENES is also using U–Ne lamps, which provide a higher line density in the near infrared (Redman et al., 2011, 2012), and suffer less from excessively bright (overexposed) noble gas lines (Quirrenbach et al., 2012).

Given the significant ageing effects of emission line lamps (and in particular the Ar lines), a common practice is to install two such lamps, with one used every night, and the other as master calibrator used only rarely (Quirrenbach et al., 2012).

2.3.4 Infrared calibration

One issue that has restricted high-accuracy radial velocities in the near-infrared has been the lack of a suitable wavelength calibrator. I₂ and Th–Ar lines used in the visible, for example, are sparse in the near infrared.

Infrared calibration using imprinted atmospheric lines reaches 10–20 m s⁻¹. For the Kitt Peak McMath Fourier transform spectrometer, Deming et al. (1987) found variations of ~20 m s⁻¹ based on telluric CH₄. For the $R = 100\,000$ Nasmyth-mounted cryogenic échelle spectrograph CRIRES at the VLT, designed for high-resolution spectroscopy between 0.96–5.2 μm, Seifahrt

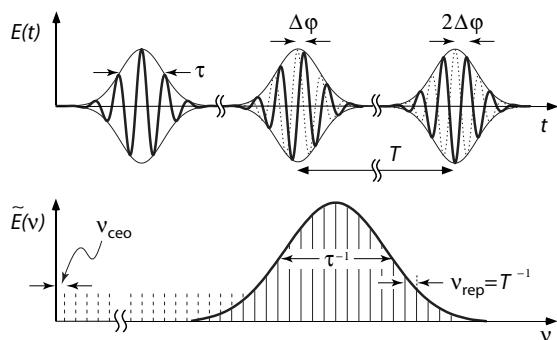


Figure 2.11: Principle of the laser frequency comb. A pulse train (top) produces the optical frequency ‘comb’ in Fourier space (bottom), with shorter pulse envelopes τ producing broader frequency combs. Within the laser cavity, the round-trip pulse time, T , determines the repetition frequency, $v_{\text{rep}} = 1/T$. Dispersive elements cause a difference in the group and phase velocities, shifting the carrier wave with respect to the envelope (top) by $\Delta\phi$ per pulse. In the frequency domain, the comb shifts by $v_{\text{ceo}} = \Delta\phi/2\pi T$. From Murphy et al. (2007, Figure 1), © Oxford University Press.

& Käfl (2008) reported telluric N₂O lines near 4.1 μm stable to ~10–20 m s⁻¹. Both studies used external N₂O gas cells for wavelength calibration.

The development and optimisation of near-infrared gas cells is ongoing. Mahadevan & Ge (2009) concluded that H¹³C¹⁴N, ¹²C₂H₂, ¹²CO, and ¹³CO together could provide useful calibration in the 1.65 μm H band. Valdivielso et al. (2010) describe a gas mixture (including C₂H₂, N₂O, NH₃, and various chloromethanes and hydrocarbons) with absorption lines spanning a wide region in the near infrared. Experiments using an NH₃ cell for VLT-CRIRES in the 2.2 μm K band reached a precision of 3–5 m s⁻¹ over weeks or months (Bean et al., 2010b). Anglada-Escudé et al. (2012c) evaluated the use of CH₄ cells, and performed a comparison with the NH₃ cell adopted for VLT-CRIRES.

For CARMENES (\$2.6.5), U–Ne emission lamps are in use for the near-infrared arm.

2.3.5 Laser frequency combs

Gas cells and emission lamps depart somewhat from an ideal radial velocity standard. A more robust calibrator would cover the entire optical/infrared range with individually unresolved lines, of uniform spacing and intensity, and with accurately known wavelengths determined by fundamental physics. *Laser frequency combs*, or *astro combs*, generated from mode-locked femtosecond-pulsed lasers (Reichert et al., 1999; Jones et al., 2000a; Udem et al., 2002), offer this possibility (Figure 2.11).

The application of laser combs to spectrograph calibration for exoplanet detection, at ~0.01 m s⁻¹, was evaluated by Murphy et al. (2007). A number of laboratory prototypes have subsequently been developed and

demonstrated (e.g. Osterman et al., 2007; Li et al., 2008a; Steinmetz et al., 2008; Schettino et al., 2010; Phillips et al., 2012a,b; Johnson et al., 2014a; Probst et al., 2015; Del'Haye et al., 2016; Kashiwagi et al., 2016; McCracken et al., 2017b; Xu et al., 2017a).

Principles Laser frequency combs rely on the storage of a single laser pulse, maintained on a repetitive path, and circulating in a cavity as a carrier wave. After each circuit, a copy of the pulse is emitted through an output mirror, resulting in an indefinite pulse train; the energy lost is replenished by stimulated emission in the lasing medium. Absolute line frequencies are

$$\nu = \nu_{\text{ceo}} + n \nu_{\text{rep}}, \quad (2.46)$$

where $\nu_{\text{rep}} = 1/T$ is the laser repetition rate (with T being the round-trip travel time), ν_{ceo} is a carrier-envelope offset frequency which must be calibrated (resulting from non-commensurability of the carrier-wave and repetition frequencies), and n is an integer, with modes given by $n \sim 10^5 - 10^6$. Both ν_{rep} and ν_{ceo} can be synchronised to atomic clocks.

Using GPS time referencing, the comb frequencies are expected to have long-term fractional stability and accuracy better than 10^{-12} , corresponding to a velocity variation below 0.01 m s^{-1} . By providing a common absolute frequency standard, external time referencing should also allow the direct comparison of measurements made at different observatories.

In practice, various architectures are used to extend the wavelength coverage (including frequency doubling or spectral broadening), or to filter the dense mode spacing to one matched to the spectrograph resolving power, e.g. using a Fabry-Pérot étalon (McCracken et al., 2017a).

Practical implementation Early astronomical demonstrations/calibrations included use on the solar spectrograph at the Vacuum Tower Telescope (VTT, Tenerife Steinmetz et al., 2008). A laser frequency comb developed at CfA Harvard was used to calibrate the Whipple Observatory's TRES spectrograph (Li et al., 2008a; Latham, 2008b). The prototype used a model-locked titanium-sapphire laser operating at $\sim 800 \text{ nm}$, providing $\nu_{\text{rep}} \sim 1 \text{ GHz}$. A Fabry-Pérot cavity increased the intrinsic line spacing to a more usable 40 GHz . One comb was operated over a 100 nm band around 800 nm , and a second over a 20 nm band around 400 nm , yielding calibration and stability below 1 m s^{-1} (Phillips et al., 2012b). Subsequent deployment was made at HARPS-N (TNG, La Palma), where it is centred at 550 nm and generates 6000 lines equally spaced by 0.015 nm over a range of 100 nm (Li et al., 2012a).

A Yb fiber-fed laser comb was evaluated for HARPS in November 2010 and January 2011, and yielded calibration lines over 450 – 590 nm (Wilken et al., 2012). Applied to HD 75289, they demonstrated short-term repeatability of 0.025 m s^{-1} (Figure 2.12). Other laser combs have since been implemented and tested, or are planned, amongst them TNG-GIANO (Schettino et al., 2011); VTT (Doerr et al., 2012); AAT-UHFR (Murphy et al., 2012); HET-Pathfinder (Ycas et al., 2012); Wendelstein-FOCES (Brucalassi et al., 2016a); NAO-Xinglon 2.16-m (Ye et al., 2016); and SALT-HRS (McCracken et al., 2017b).

Future challenges include broader wavelength coverage (e.g. 0.37 – $2.5 \mu\text{m}$ for E-ELT-HIRES), as well issues of operational lifetime and automation (McCracken et al., 2017a).

2.3.6 Fabry-Pérot étalons

Fabry-Pérot interferometers or étalons also provide a quasi-periodic and dense grid of lines and homoge-

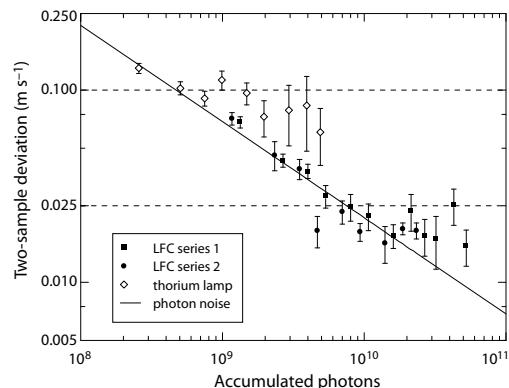


Figure 2.12: Calibration of the laser frequency comb for ESO-HARPS, showing the two-sample deviation of the two long calibration runs. The smallest number of accumulated photons per series, N , represents one acquisition, and several consecutive acquisitions can be added to obtain larger values of N . For $N < 5 \times 10^9$, the uncertainty decreases as $N^{-0.5}$, due to photon noise (solid line). For $N > 5 \times 10^9$ there is no further reduction, which flattens at $\sim 0.025 \text{ m s}^{-1}$ (lower dashed line). This represents the minimum calibration uncertainty achieved, and is reached after 4 min. The two-sample deviation of 116 Th lamp calibrations is also shown (open diamonds). One acquisition with the Th lamp has a much larger photon noise, and a minimum uncertainty of $\sim 0.1 \text{ m s}^{-1}$. Four data points of 5×10^9 photons require 2 h. From Wilken et al. (2012, Figure 3), by permission from Nature/Springer/Macmillan Ltd, ©2012.

neous amplitudes, and can provide economical alternatives to the laser frequency comb (Bauer et al., 2015). Operated with a hollow-cathode lamp, they can map the lamp's spectral features onto an absolute scale, for example calibrating large distortions (of some 50 m s^{-1}) in the wavelength solution of the HARPS data reduction software, and yielding differences of $\leq 10 \text{ m s}^{-1}$ with results from a laser frequency comb (Bauer et al., 2015).

They have been used or considered for various state-of-the-art instruments, including HARPS (Wildi et al., 2011), CARMENES (Schäfer & Reiners, 2012; Reiners et al., 2014), APOGEE and HPF (Halverson et al., 2014), and SPIRou (Cersullo et al., 2017). Actively stabilised étalons are also used (e.g. Halverson et al., 2014; Schwab et al., 2015), and high-accuracy (0.1 m s^{-1}) dual cavity systems have been considered (Banyai & Reiners, 2017).

2.3.7 Radial velocity standards

Verification of global error control is provided by continual monitoring of stars suspected to be true null standards, such as HD 185144 (σ Dra), a bright intrinsically quiet star with no discernible variations, or with simple, well-characterised planetary systems. These include (e.g. Vogt et al., 2014b) HD 168746 (with a 6.4-d Saturn-mass), HD 185269 (with an eccentric 6.8-d Jupiter), HD 187123 (with a 3.1-d hot Jupiter), and HD 209458 (with its 3.5-d hot Jupiter).

2.3.8 Fiber coupling

Various state-of-the-art instruments employ fiber feed to couple the telescope focus to the spectrograph. This offers two advantages. Firstly, it overcomes variable light illumination of the spectrograph slit, and the associated complications of wavelength calibration. This is because multiple internal reflections within the fiber cause scrambling of the image structure between input and output, decoupling the illumination of the spectrograph optics and detector from guiding errors and calibration source misalignments (Hubbard et al., 1979; Barden et al., 1981; Heacox & Connes, 1992; Heacox, 1986, 1988; Hunter & Ramsey, 1992; Connes et al., 1996; Walker et al., 2003b). In addition, the spectrograph can be distanced from the telescope, in a more mechanically and thermally stabilised environment, also minimising the effects of gravitational flexure.

Fiber-feed was first adopted for ELODIE at the OHP 1.9-m telescope (Baranne et al., 1996), and used for the detection of the first exoplanet discovered around a main sequence star, 51 Peg (Mayor & Queloz, 1995). The same concept has subsequently been used for HARPS, CARMENES, HPF and others. In CARMENES, for example, each (optical or near infrared) spectrograph is fed by two optical fibers with $100\text{ }\mu\text{m}$ core diameter (1.5 arcsec on the sky), with one fiber carrying the starlight, and the second used for simultaneous emission line lamp calibration or sky subtraction (Quirrenbach et al., 2012). A core diameter of $300\text{ }\mu\text{m}$ is used for HPF. These *multi-mode fibers* contrast with the *single-mode fibers*, for which a much smaller diameter core (typically $10\text{ }\mu\text{m}$) allows only one mode of light to propagate (with a corresponding reduction in internal reflections and therefore reduced attenuation).

Modal noise and fiber shaking The $100\text{--}300\text{ }\mu\text{m}$ core diameters in typical use imply the propagation of some $10^2\text{--}10^4$ electromagnetic modes (Mahadevan et al., 2014a). This finite number of transmission modes introduces an error term in high-resolution spectra which cannot be eliminated by flat-fielding, because the measured intensity distribution is sensitive to small changes in fiber position or illumination. This *modal noise*, rather than photon shot noise, limits the achievable signal-to-noise ratio in fiber-fed spectrographs, to $\text{S/N} \lesssim 500$ at $R \sim 150\,000$ in the experiments of Baudrand & Walker (2001). They observed an (unexplained) empirical relation between the number of monochromatic modes and the level of modal noise, and found that continuous, low-amplitude agitation of the fiber eliminated the problem. Mechanical fiber agitation, or ‘shaking’, is now a standard component of fiber-fed systems (e.g. Sutherland et al., 2016). An alternative uses an integrating sphere with a diffuser that is moved rapidly using electrostrictive polymers (Mahadevan et al., 2014a).

Scrambling and fiber cross section Despite the intrinsic light-scrambling properties of optical fibers (Hunter & Ramsey, 1992; Avila et al., 1998; Avila & Singh, 2008) the illumination produced by a circular optical fiber still depends on how light is fed into the fiber, i.e. motion of the stellar image at the fiber en-

trance can produce a change in the illumination of the spectrograph, and hence mimic a radial velocity effect. Double scramblers were developed to provide further improvements (Hunter & Ramsey, 1992; Baranne et al., 1996).

Non-circular fibers, and specifically fibers with a square or octagonal cross section, can provide better scrambling as they lack the symmetry that leads to incomplete radial scrambling in circular fibers (Perruchot et al., 2011; Chazelas et al., 2012a; Plavchan et al., 2013a; Sutherland et al., 2016). Better performance with octagonal fibers has been demonstrated for both HARPS-N (Cosentino et al., 2012) and SOPHIE+ (Bouchy et al., 2013), resulting in an improvement from $5\text{--}6\text{ m s}^{-1}$ for SOPHIE to around $1\text{--}2\text{ m s}^{-1}$ for SOPHIE+. Use of octagonal fibers is now widespread, while D-shaped fibers offering ‘chaotic scrambling’ are also under development (Stürmer et al., 2016).

Adaptive optics and photonic lanterns Future instruments, notably for very large telescopes, are likely to benefit from the use of adaptive optics (to reduce the size of the seeing-limited stellar image) coupled to a *photonic lantern* (or photonic reformatter), which converts a multimode fiber input into a small number of single mode fiber outputs (Leon-Saval et al., 2010).

Such an approach, already proposed by Connes et al. (1996), offers prospects of reducing instrument size, allowing for more compact instrument designs with good optomechanical stability, while also eliminating modal noise (Ihle et al., 2010; Schwab et al., 2014; Feger et al., 2014; Crepp, 2014; Harris et al., 2016). The first such operational instrument is likely to be iLocater, under development for the $2 \times 8.4\text{-m}$ LBT, and targeting operations in 2018 (Crepp et al., 2016).

2.4 Accuracy limits and error sources

Improvements in detector technology, wavelength calibration, and various other aspects of instrument design brought the state-of-the-art radial velocity accuracy to some $3\text{--}5\text{ m s}^{-1}$ by the mid-1990s (e.g. Butler et al., 1996), and to around $0.3\text{--}0.5\text{ m s}^{-1}$ today (Pepe & Lovis, 2008). Such accuracy represents a displacement of a few nm at the CCD detector, demanding a combination of optical and algorithmic techniques to maintain the required metrology over several years.

Although the radial velocity amplitude is independent of the distance to the star, signal-to-noise considerations limit observations to the brighter stars, typically $V \lesssim 8\text{--}10$ mag. Equation 2.27 indicates that radial velocity measurements implicitly favour the detection of massive planets, those with small a (i.e. small P) and, with adequate temporal sampling, large e (but see §2.1.2 for detectability in practice).

Error sources potentially imposing practical limits on achievable accuracy include instrumental errors, photon noise, and stellar ‘noise’, or stellar ‘jitter’ at various time scales and amplitudes (§2.4.5).

Other effects include spectral contamination (e.g. Pepe et al., 2004b; Pepe & Lovis, 2008; Cunha et al., 2013), including the contribution of micro-telluric lines, i.e. telluric lines that are not masked out for the radial velocity computation (Cunha et al., 2014), as well as contributions due to unrecognised planetary companions.

2.4.1 Photon noise

Achievable radial velocity accuracies scale with signal-to-noise. Estimates for ELODIE, and comparisons with real data, were detailed by Connes et al. (1996). For the HARPS instrument at the ESO 3.6-m telescope, a photon-limited precision of $\sim 1 \text{ m s}^{-1}$ is reached for a $V = 7.5$ mag G dwarf in around 60 s. In the absence of other noise sources, a single measurement accuracy of 0.1 m s^{-1} requires exposure times $\gtrsim 100$ min. Photon-noise limits for CORALIE and HARPS are given by Bouchy et al. (2001).

2.4.2 Detection versus signal-to-noise

The signal-to-noise ratio at which a radial velocity survey can detect planets with a given period depends on the strength of the signal K , the magnitude of the measurement uncertainties σ , the survey duration T , and the total number of observations N . The time-dependent radial velocity of a star orbited by a planet on a circular orbit with period P has the form $v(t) = K \sin(2\pi t/P - \phi_0)$, where ϕ_0 is an arbitrary phase. For a uniform and continuous sampling of the radial velocity curve, such that observations at N epochs are taken over a survey duration T , each with a measurement uncertainty σ , then the detection S/N is

$$\text{S/N}(\phi_0) = \left\{ \frac{N}{\sigma^2} \frac{1}{T} \int_{-T/2}^{T/2} [v(t) - \langle v(t) \rangle]^2 dt \right\}^{\frac{1}{2}}, \quad (2.47)$$

where $\langle v(t) \rangle$ is the time average of $v(t)$ and is a function of ϕ_0 . Averaging over ϕ_0 , gives a ‘phase-averaged’ S/N for a radial velocity detection

$$\langle \text{SNR} \rangle_{\phi_0} = \left(\frac{N}{2} \right)^{\frac{1}{2}} \left(\frac{K}{\sigma} \right) \left\{ 1 - \frac{1}{\pi^2} \left(\frac{P}{T} \right)^2 \sin^2 \left(\frac{\pi T}{P} \right) \right\}^{\frac{1}{2}}. \quad (2.48)$$

For $P/T \ll 1$, the S/N for a detection is very nearly independent of period

$$\langle \text{SNR} \rangle_{\phi_0} \sim \left(\frac{N}{2} \right)^{\frac{1}{2}} \left(\frac{K}{\sigma} \right). \quad (2.49)$$

As expected, more epochs, improved measurement errors, and longer survey durations increase the detection S/N. Even for $P \sim T$, Equation 2.49 remains a good approximation. Although this assumes no detection dependence on eccentricity, in practice planets with $e \gtrsim 0.6$ are more difficult to detect (Cumming, 2004)

Radial velocity fitting challenge To establish further confidence in claimed detections given the various analysis methods used by different groups and based on different instruments, a ‘fitting challenge’ has evaluated the consequences of the use of different activity indicators, red-noise models, and a Bayesian analysis framework (Dumusque, 2016; Dumusque et al., 2017).

Results were formulated in terms of the ratio $K/N = (K_{\text{pl}}/\text{RV}_{\text{rms}}) \sqrt{N_{\text{obs}}}$, where K_{pl} is the semi-amplitude of each planetary signal, N_{obs} is the number of observation in each system, and RV_{rms} is the radial velocity rms of each system once the best-fit of a model (consisting of a linear correlation with $\log(R'_{\text{HK}})$ plus a second-order polynomial as a function of time) was removed.

Their findings are that planets can be found down to $K/N = 5$, with a threshold of $K/N = 7.5$ at the level of 80–90% recovery rate found for a number of methods. For $K/N < 7.5$ recovery rates drop dramatically, while false positives increase, suggesting that $K/N = 7.5$ provides an appropriate threshold for planet discovery.

2.4.3 Exposure metering

Starting with ELODIE (Baranne et al., 1996), many instruments including Keck-HIRES, the Lick Observatory Hamilton Spectrometer, and the Lick Automated Planet Finder (APF), employ an integral exposure meter to monitor ongoing exposures. This is used to optimise exposure times for a required signal-to-noise, and to calculate the photon-weighted midpoint of each exposure for barycentric correction (Kibrick et al., 2006).

HIRES, for example, uses a ‘propellor’ mirror behind the spectrometer slit to direct a few percent of the light to a dedicated photomultiplier. APF employs a semi-transparent ‘pellicle’ mirror, operating at 1 Hz, for auto-guiding, exposure metering, and barycentric weighting. A multiple-channel exposure meter for EXPRES at the Lowell 4.3-m Discovery Channel Telescope, designed to take account of chromatic atmospheric effects, is described by Blackman et al. (2017).

2.4.4 Instrument errors

Potential instrument-related error sources include the effects of mechanical and thermal stability, wavelength calibration (§2.3), spectrograph illumination (§2.3.8), and timing and coordinate systems (§2.2.4).

Self-calibration versus simultaneous reference Any instrumental effect that produces a distortion or a shift of the spectral line in the detector-pixel space will be interpreted, if not detected and recognised, as a wavelength change and thus a Doppler shift. Broadly, two methods of tracking changes in the instrumental profile changes have been used (Pepe et al., 2014a).

The first superimposes an absorption spectrum of a reference gas cell on the stellar spectrum, such that the instrument profile is continuously measured (§2.3.2). This *self-calibration technique* is particularly effective in spectrographs with varying instrument profiles, notably for slit spectrographs. Disadvantages are the restricted bandwidth of the gas-cell spectrum, the loss of efficiency due to absorption in the light path, and the need for a sophisticated deconvolution process to recover the stellar spectrum and thus the radial velocity.

Table 2.1: Estimated radial velocity jitter, σ'_{rv} , for various spectral types and luminosity classes.

Type	Jitter (m s^{-1})	Comment	Reference
FV	7–30	$v_{\text{rot}} = 8 - 10 \text{ km s}^{-1}$	Saar et al. (1998)
"	2–7	inactive stars	Wright (2005)
GV	20–45	$v_{\text{rot}} = 8 - 10 \text{ km s}^{-1}$	Saar et al. (1998)
GKV	4.6 ± 1.8	$v_{\text{rot}} \leq 2 \text{ km s}^{-1}$	Saar et al. (1998)
"	2–5	inactive stars	Wright (2005)
MV	2–7	inactive stars	Wright (2005)
GKIV	5–10	subgiants	Wright (2005)
KIII	20	giants	Hekker et al. (2006b)

The second method, the *simultaneous reference technique*, as employed for ELODIE, HARPS and others, assumes a stabilised instrumental profile that does not change between two wavelength calibrations of the spectrograph, such that the relationship between detector pixel and wavelength remains valid over time scales of typically one night. A second channel carrying a spectral reference is continuously fed to the spectrograph to monitor and correct for potential instrumental drifts or instrument profile changes. It is then essential that the changes that the scientific and the reference channels are subject to are identical over the time scale of one observing night. Therefore, the whole design of the instrument must be optimised for stability of the instrumental profile, demanding fiber feed and light scrambling, as well as pressure, mechanical, thermal and optical stability. The effort is compensated for by an unrestricted spectral bandwidth and the acquisition of an ‘uncontaminated’ scientific spectrum.

2.4.5 Stellar activity

Activity in the stellar atmosphere (due to spots and/or plages), as well as stellar oscillations, and surface granulation, may all contribute astrophysical noise or ‘jitter’ to the radial velocity measurements, and become a prominent issue at the m s^{-1} level. These phenomena alter the shape of the stellar absorption lines by injecting spurious or systematic radial velocity signals that may mask or mimic planetary signals, and therefore complicate determination of the underlying centre-of-mass Keplerian motion. The effects are one of the dominant issues faced by future spectrographs targeting 0.1 m s^{-1} accuracies (e.g. Fischer et al., 2016a).

There are many examples of these effects reported in the literature, for example in the cases of HD 166435 (Queloz et al., 2001), HD 219542 (Desidera et al., 2003, 2004a), TW Hya (Huélamo et al., 2008), and BD+20 1790 (Figueira et al., 2010a).

Jitter due to inhomogeneities of the stellar atmosphere (spots, plages) is often significant, frequently imposes a limit to the accuracy of radial velocity measurements, and is expected to vary on time scales comparable to the stellar rotation period. It is correlated

with stellar chromospheric activity (e.g. Saar & Donahue, 1997; Saar et al., 1998; Butler & Marcy, 1998; Saar & Cuntz, 2001; Tinney et al., 2002b; Shkolnik et al., 2005; Jenkins et al., 2006; Desort et al., 2007; Saar, 2009). Emission in the core of the Ca H/K lines is a useful proxy, although active stars seen pole-on can have a high activity index but low radial velocity jitter.

The effects of stellar oscillations on Doppler measurements are typically smaller than those produced by activity, and are most significant for giants and subgiants. O’Toole et al. (2008) made an analysis using data from asteroseismological investigations using AAT-UCLES. They found a power-law relation between measured noise and both the luminosity-to-mass ratio of the target stars and integration times (Figure 2.13), showing that it is advantageous to average over several p-mode oscillation frequencies (O’Toole et al., 2009b). HARPS results have also shown that 15-min integrations are sufficient to damp resulting radial velocity variations to below 0.2 m s^{-1} (Udry & Mayor, 2008).

Surface granulation can induce variability of order 1 m s^{-1} for solar type stars (Kjeldsen et al., 2005), which may demand measurements over several hours to damp.

The overall effects of jitter can be characterised as an excess in the radial velocity standard error as (Saar et al., 1998; Wright, 2005)

$$\sigma'_{\text{rv}} = (\sigma_{\text{rv}}^2 - \sigma_{\text{meas}}^2)^{1/2}, \quad (2.50)$$

where σ_{meas} is the contribution of the measurement error to the radial velocity standard error, σ_{rv} , for a given star, and σ'_{rv} includes contributions from all of the effects noted above.

Astrophysical contributions are dependent on a number of variables, including stellar rotational velocity and age; some indicative estimates are given in Table 2.1 and Figure 2.14. The activity–rotation–age relationship for solar-type stars also applies to early M dwarfs (e.g. Maldonado et al., 2017; Scandariato et al., 2017).

Various recent models of the resulting contributions to radial velocity time series have been developed (e.g. Korhonen et al., 2015; Rajpaul et al., 2015; Gregory, 2016; Herrero et al., 2016), including specific models for M dwarfs (Andersen & Korhonen, 2015).

Summary for GK dwarfs For GK dwarfs, stellar signals relevant at $\sim 1 \text{ m s}^{-1}$ can be decomposed in four components (Dumusque et al., 2017):

- (a) solar-type oscillations (e.g. Kjeldsen et al., 2005; Arentoft et al., 2008; O’Toole et al., 2008; Dumusque et al., 2011a);
- (b) granulation phenomena (e.g. Dravins, 1982; Lindegren & Dravins, 2003; Del Moro, 2004; Del Moro et al., 2004; Dumusque et al., 2011a; Gray & Oostra, 2018).
- (c) short-term activity signals on the stellar rotation period time scale (e.g. Saar & Donahue, 1997; Saar, 2009;

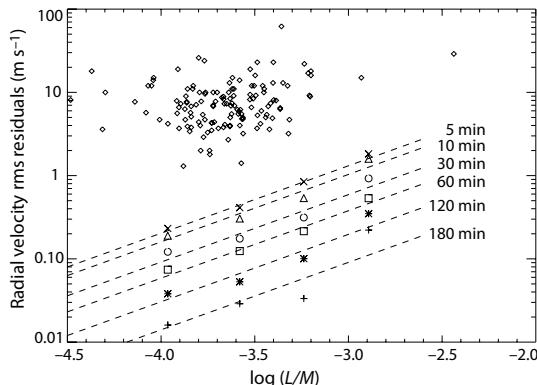


Figure 2.13: Oscillation jitter of AAT–UCLES asteroseismology targets as a function of $\log_{10}(L/M)$ (luminosity–mass ratio) for various simulated integration times. Power laws derived by O’Toole et al. (2008, their eqn 2) are shown as dashed lines. Residual rms values for the sample of known planets from Butler et al. (2006b) are shown as small diamonds. From O’Toole et al. (2008, Figure 4), © Oxford University Press.

Meunier et al., 2010a; Boisse et al., 2012a; Robertson et al., 2014; Borgniet et al., 2015; Robertson et al., 2015a; Haywood et al., 2016);

(d) long-term activity signals on the magnetic cycle period time scale (e.g. Makarov, 2010; Dumusque et al., 2011c; Meunier & Lagrange, 2013a; Díaz et al., 2016b; Lanza et al., 2016).

Treatment Regarding short-term stellar activity, the most difficult signals to treat are non-periodic stochastic terms arising from the evolution and decay of active regions. Numerous correction techniques have been investigated (Dumusque et al., 2017): • fitting sine wave and harmonics at the rotation period of the star (Boisse et al., 2011); • use of red-noise models (Gregory, 2011b; Tuomi et al., 2013a; Feroz & Hobson, 2014); • exploiting contemporaneous photometry (Aigrain et al., 2012; Haywood et al., 2014b; Dumusque et al., 2015a); • modeling activity-induced signals with Gaussian process regression, whose covariance properties are shared either with the stellar photometric variations (Haywood et al., 2014b; Grunblatt et al., 2015), or a combination of several spectroscopic indicators (Rajpaul et al., 2015), or determined from the radial velocities themselves (Faria et al., 2016a); • using linear correlations between the different observables, i.e., radial velocity, bisector span, and width of the cross correlation function (Baranne et al., 1996; Pepe et al., 2002), photometry (Queloz et al., 2001; Boisse et al., 2009; Robertson et al., 2014, 2015a), and magnetic field strength (Hébrard et al., 2014a); • checking for season per season phase incoherence of signals (Dumusque et al., 2012, 2014a; Santos et al., 2014); and • avoiding the impact of activity by using wavelength-dependent criteria (Anglada-Escudé & Butler, 2012; Tuomi et al., 2013a).

Earth-mass planets in the habitable zone The sun spot and plage properties of the Sun over one solar cycle, between 1993–2003, have been used to infer the radial velocity curve of a solar-type star exhibiting such features (Lagrange et al., 2010b; Meunier et al., 2010a). The radial velocity amplitude varies, in a complex way, from a few tenths up to $\sim 5 \text{ m s}^{-1}$. Assuming radial velocity accuracies in the range $0.01\text{--}0.1 \text{ m s}^{-1}$ for a Sun-

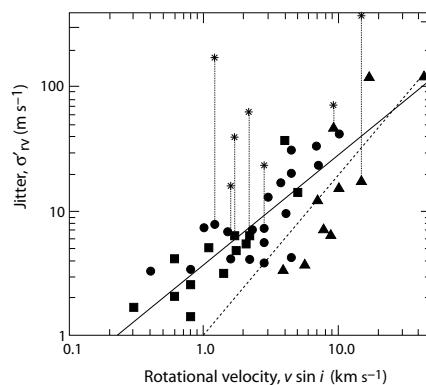


Figure 2.14: Early results from the Lick survey, showing the dependency of astrophysical jitter, σ'_{rv} (Equation 2.50) on stellar rotational velocity, $v \sin i$ for main sequence stars, indicated as F (triangles), G (circles) and K (squares). Stars with planets are plotted before (asterisks) and after (connected by vertical dashed lines) accounting for the planet’s contribution. The power-law fits have exponent 0.9 for G and K stars (solid line), and 1.3 for F stars (dashed). From Saar et al. (1998, Figure 2a), by permission of IOP Publishing/AAS.

like star at 10 pc, detection of a $1M_{\oplus}$ planet in the habitable zone, between 0.8–1.2 au, would require weekly sampling over several years, with frequent temporal sampling being crucial. Variations in the Sun’s total solar irradiance during cycle 23, as measured by SOHO–VIRGO, is reported by Lanza et al. (2003).

Activity indicators Chromospheric activity indicators include Ca II H and K, H α , He I D3, and Na I D1 and D2 (Figure 2.15). The frequently used Mt. Wilson S-value scale is defined as the ratio of the sum of the flux in the cores of the Ca II H and K lines to the sum of two continuum bands, redward and blueward of the H and K lines (Wilson, 1968). In the near-infrared, the Ca II triplet provides a proxy (Robertson et al., 2016).

The R'_{HK} index is derived from the S-value by converting into a standard scale and correcting for the residual photospheric emission, leading to a non-dimensional measure of chromospheric emission. This effectively removes any dependencies on the bolometric flux, allowing accurate comparison of the chromospheric emission across different spectral types (e.g. Knutson et al., 2010; Isaacson & Fischer, 2010; Astudillo-Defru et al., 2017a; Suárez Mascareño et al., 2017c).

Detailed studies Many detailed studies of the origin and effects of stellar noise on radial velocity measurements have been reported, including for M dwarfs (e.g. Saar & Fischer, 2000; Santos et al., 2000c; Narayan et al., 2005; Wright, 2005; Reiners, 2009; Santos et al., 2010a; Martínez-Arnáiz et al., 2010; Lagrange et al., 2011; Dumusque et al., 2011c,b; Boisse et al., 2011; Lanza et al., 2011a; Barnes et al., 2011a; Gomes da Silva et al., 2011, 2012; Boisse et al., 2012a; Aigrain et al., 2012; Cegla et al., 2012, 2014; Bastien et al., 2014b; Barnes et al., 2014; Astudillo-Defru et al., 2017a).

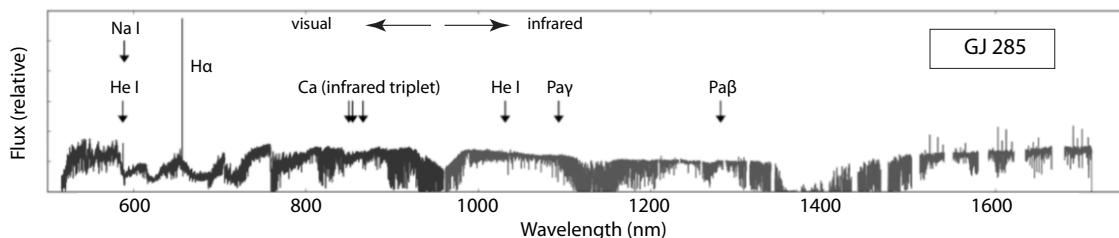


Figure 2.15: Optical–infrared spectrum of the M4.5V dwarf GJ 285 observed with CARMENES. Chromospherically sensitive lines (indicated) are distributed across the spectrum. In the optical, the H α line is particularly prominent and several molecular band heads can be distinguished. The infrared spectrum shows the atmospheric J- and H-band ‘windows’ around 1.2–1.6 μm , with a number of prominent airglow emission lines beyond 1.5 μm . From (Quirrenbach et al., 2016, Figure 4), with permission ©SPIE.

The Sun as a template Monitoring of the Sun has also been used to characterise solar-type stellar activity and its impact on exoplanet detectability (e.g. Lagrange et al., 2010b; Meunier et al., 2010a; Lagrange et al., 2011; Meunier & Lagrange, 2013a,b; Borgniet et al., 2015; Marchwinski et al., 2015; Meunier et al., 2015; Haywood et al., 2016). The Sun’s radial velocity signal is also measurable in reflected light from minor bodies of the solar system, including the Moon, the Galilean satellites, and several asteroids, providing further insight into the relation between radial velocity and stellar activity (Lanza & Morlino, 2015; Lanza et al., 2016).

A dedicated solar telescope feeding the HARPS-N spectrograph (equipped with a laser frequency comb) is being used for intensive long-term monitoring of the Sun, and currently reaches weekly rms noise levels of 0.6 m s^{-1} (Dumusque et al., 2015a).

Accounting for stellar jitter Star spots and plages can induce radial velocity signals of $1\text{--}100 \text{ m s}^{-1}$ (Saar & Donahue, 1997). There have been many attempts to remove this large-scale jitter, including the use of observational strategies, spot simulations, Keplerian models, bisector analysis, harmonic decomposition, Fourier component analysis/pre-whitening, and principal component analysis (Saar & Donahue, 1997; Hatzes, 2002; Desort et al., 2007; Melo et al., 2007; Bonfils et al., 2007; Boisse et al., 2009; Hatzes et al., 2010; Dumusque et al., 2011c; Boisse et al., 2011; Hatzes, 2013b; Gettel et al., 2013; Hatzes, 2013b; Moulds et al., 2013; Petit et al., 2015; Robertson et al., 2015a; Vanderburg et al., 2016d; Barnes et al., 2017; Davis et al., 2017).

Notwithstanding these removal techniques, active stars are still not ideal candidates for radial velocity follow-up and are thus often left out of planet surveys.

Even ‘quiet’ stars (those with little or no activity) exhibit jitter due to granulation and stellar oscillations, which can introduce radial velocity signals of 1 m s^{-1} or more (e.g. Schrijver & Zwaan, 2000). Observing strategies can be adapted to average out this small-scale jitter (Pepe et al., 2011; Dumusque et al., 2011c).

Correlation with stellar rotation Active areas moving across the stellar limb can lead to detection bias at the stellar rotation period, a specific problem if the planet is tidally-locked close to co-rotation (e.g. Petit et al., 2015; Vanderburg et al., 2016d).

Correlation with magnetic cycles From long-term monitoring, Santos et al. (2010a) found that the stellar activity indicators Ca II, H α , and He I trace the magnetic cycle in solar type stars, as do the different parameters of the spectroscopic cross-correlation function such as the bisector and full width half maximum. At least for early K dwarfs, radial velocity amplitudes are perturbed by $\lesssim 1 \text{ m s}^{-1}$ over the magnetic cycle.

For GJ 328 observed over 10 yr, the long-period magnetic activity cycle introduces an additional signal in the stellar radial velocity curve of amplitude $6\text{--}10 \text{ m s}^{-1}$ (Robertson et al., 2013).

Algorithmic implementation Routines for the computation of photometry and radial velocity terms induced by star spots include SOAP (Boisse et al., 2012a) and SOAP 2.0 (Dumusque et al., 2014a).

2.4.6 Excluding other sources of periodicity

Effects other than an orbiting exoplanet can result in periodic stellar radial velocity variations, including most of the noise terms considered in the previous section.

Star effects Surface activity such as spots and plages, as well as inhomogeneous convection, can lead to periodic radial velocity variations: a spot with a filling factor of a few percent can induce a radial velocity amplitude of a few m s^{-1} (Saar & Donahue, 1997). Activity may be discounted if the radial velocity period is distinct from the stellar rotation period, or if the star shows only very low level photometric variability at the radial velocity period (although spots are a time-dependent phenomenon). Spectroscopic measurements of the Ca II H and K lines provide a proxy for surface (magnetic) activity (Baliunas et al., 1995), and may reveal rotational

modulation attributable to stellar rotation. For close-in planets, magnetic interactions may cause additional stellar activity, as in CoRoT-7 (Lanza et al., 2010).

Stellar pulsations may also cause radial velocity variations, but tend to be discounted if their period and amplitude do not correspond to known excitation mechanisms for the relevant spectral type.

Binary companions Stellar companions can mimic exoplanet signatures, and can be difficult to rule out (see, e.g., HD 41004; Santos et al., 2002b; Zucker et al., 2003, 2004). Hipparcos catalogue astrometry, both the double star annex and the goodness-of-fit statistics for the single-star model (catalogue fields H29–H30), provides a first level of discrimination.

The presence of a lower-level stellar spectrum can also be investigated using TODCOR, a cross-correlation algorithm developed to detect and measure radial velocities of two components of a spectroscopic binary (Mazeh & Zucker, 1994). Its application to HD 41004 is described by Zucker et al. (2003). An extension to triple stellar systems TRICOR/TRIMOR is described, and applied to HD 188753, by Mazeh et al. (2009b).

A particularly confusing configuration may arise when a tight binary system orbits the target star. Schneider & Cabrera (2006) suggested that the binary wobble around its own centre of mass can result in the same modulation of the three basic observables (radial velocity, astrometric position, and time of arrival of a periodic signal) as for an orbiting planet. A more detailed analysis by Morais & Correia (2008) showed that the binary actually mimics two planets, rather than a single orbiting planet, but with very similar orbital periods.

Period aliases In radial velocity work, as in other areas of astronomy, observation sequences are frequently interrupted by semi-regular natural phenomena including the diurnal and lunar cycles, resulting in aliases due to stroboscopic effects.

Baluev (2012) examined the problem of distinguishing between two possible models, specifically for period alias ambiguities. Considering the statistical problem of selecting the best-fitting model (and as an alternative to the computationally-intensive Bayesian approach), he argued that the (non-Bayesian) *Vuong test* (Vuong, 1989) performs particularly well, and applied it to 55 Cnc, GJ 876, GJ 3634, HD 75898, and HD 156668.

Retracted discoveries Various retracted discoveries serve as illustrations of the complexities of orbit fitting in practice:

MARVELS-1 (TYC 1240–945–1) was reported as a $28M_J$ companion in a $P = 5.9$ d, $e \sim 0$ orbit (Lee et al., 2011b); further radial velocities, imaging data, and bisector analysis, subsequently reclassified it as a face-on stellar binary in a hierarchical 3–4 star system (Wright et al., 2013).

HIP 11952 was reported as a 2-planet system from FEROS (Setiawan et al., 2012). Unconfirmed by HARPS (Desidera et al., 2013), re-analysis of the FEROS data attributed the claim to an error in barycentric correction (Müller et al., 2013a).

HD 41248 was reported as a 2-planet system in a possible

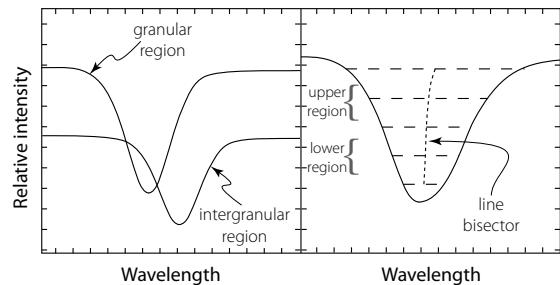


Figure 2.16: Schematic of line profile asymmetry caused by convection, and the resulting line bisector. Left: line profiles resolving individual convection cells would show blueshifted lines from the upflow regions, with redshifted lines from the intergranular downflow lanes. Right: the resulting line profile at low spatial resolution, with a consequently asymmetric line bisector. Adapted from Dravins et al. (1981, Figure 1), reproduced with permission © ESO.

7:5 resonance (Jenkins et al., 2013c), but subsequently attributed to stellar activity (Santos et al., 2014).

2.4.7 Bisector analysis

The *spectral line bisector* is the locus of median points midway between equal intensities on either side of a spectral line, thereby dividing it into two halves of equal equivalent width (Figure 2.16). Line bisectors are used to quantify stellar line profile asymmetries, and are important in identifying the underlying cause of certain types of radial velocity variation.

Since the work of Voigt (1956), line bisectors have been used to describe the nature and strength of convection in the Sun, and in stellar photospheres on the cool side of the granulation boundary where deep convective motion is significant (Dravins, 1975; Dravins et al., 1981; Gray, 1982, 1983; Nowak & Niedzielski, 2008). In such photospheres, local upward motion in the brighter (hotter) convective cells is balanced by a downflow in the darker (cooler) intergranular lanes. In the absence of spatial resolution, averaging over many granules, an overall convective blueshift, of order -0.5 km s^{-1} , results from the dominant contribution of blueshifted photons originating from the larger and brighter granules.

Bisectors due to solar-like granulation are shaped like a distorted 'C', and like the top-half of the letter 'C' for stars somewhat hotter or more luminous than the Sun. They show reversed curvature on the hot side of the granulation boundary, due to the structure of the surface convection zone (Gray, 1989; Gray et al., 2008).

Bisector metrics The shape of the spectral line bisector can be quantified in various ways: (a) the *bisector velocity span* (or 7% span) is constructed as the difference in bisector velocity between upper and lower regions of the line (Figure 2.16), avoiding the wings and cores (Toner & Gray, 1988; Hatzes, 1996; Queloz et al., 2001); (b) the *bisector inverse slope* (Queloz et al., 2001; Santos et al., 2002b) is $v_t - v_b$, where v_t is the mean bisector velocity between 10–40% of the line depth (top), and v_b is

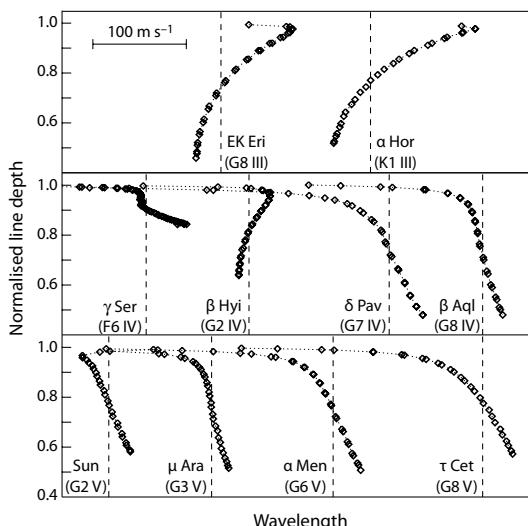


Figure 2.17: Mean bisectors, constructed from the cross-correlation function, of various bright stars according to luminosity class, III (top row) to V (bottom). Vertical dashed lines indicate the radial velocity. Error bars are the size of the plot symbols or smaller. From Dall et al. (2006, Figure 2), reproduced with permission © ESO.

that between 55–90% (bottom); (c) the *bisector curvature* is the difference in velocity span between the upper and lower halves of the bisector (e.g. Hatzes, 1996; Nowak & Niedzielski, 2008), for example $(v_3 - v_2) - (v_2 - v_1)$, where v_1, v_2, v_3 are mean velocities at 20–30%, 40–55%, and 75–100% of the line depth (Dall et al., 2006). Various other forms are also used (e.g. Dall et al., 2006; Baştürk et al., 2011; Figueira et al., 2013).

Bisector analysis is usually made directly on the mean line profile given by the cross-correlation function, and used as an aid to distinguish planetary signatures from other types of radial velocity modulation (Povich et al., 2001; Martínez Fiorenzano et al., 2005). For a planetary signal, the bisector span is expected to be independent of radial velocity, whereas a distinct correlation arises for blended systems (e.g. for HD 41004: Santos et al., 2002b, Figure 3), or for periodic variations due to star spots (e.g. for HD 166435: Queloz et al., 2001, Figure 7).

As one of many examples, Dall et al. (2006) used HARPS observations of bright stars, both active and inactive, to derive the mean bisectors from individual lines, as well as from the cross-correlation function (Figure 2.17). They showed that different spectral lines can show different bisector shapes, even between lines of the same element, calling for caution in deriving global stellar properties from the bisector derived from the cross-correlation function alone.

2.5 Higher-order radial velocity effects

2.5.1 Gravitational redshift variations

The first-order effect of gravitational redshift is described in Section 2.2.4. Fluctuations in stellar radius would lead to variations in an escaping photon's gravitational redshift, introducing noise into radial velocity measurements by shifting the *centroid* of the observed

spectral lines. The effect is in contrast with other types of stellar jitter which result in line asymmetries (Cegla et al., 2012). From the velocity shift for a photon detected at large distance from the stellar surface (Equation 2.44), a radius change of δR_\star would result (at constant mass) in a velocity variation $\delta v_r = (\delta R_\star / R_\star) v_r$.

Solar radius variations For the Sun, $v_r = 636 \text{ m s}^{-1}$, such that $\delta R_\odot \sim 0.01\%$ would result in $\delta\lambda \sim 0.06 \text{ m s}^{-1}$. Since the radial velocity signal due to the Earth orbiting the Sun is $\sim 0.09 \text{ m s}^{-1}$, a radius change in a host star of $\geq 0.01\%$ could mimic, or mask, the radial velocity signal of an Earth-like planet.

Cegla et al. (2012) tabulate various recent estimates of solar radius *variations* (see also §12.1.7). While some studies report solar radius changes of 0.01% or more observed from ground (e.g. Penna et al., 2010, and references), space-based observations from SOHO–MDI suggest that R_\odot is stable to $\sim 10^{-6}$ (Bush et al., 2010).

Stellar radius variations In addition to the widely-observed pulsation-driven variations, a number of other mechanisms may give rise to stellar radius variations. Cegla et al. (2012) assessed the effects of the Applegate mechanism (box, page 114), magnetic-field induced convective inhibition, and the Wilson depression of star spots (§4.4.3). They found that the latter two could induce gravitational redshift variations up to 0.04 m s^{-1} and 0.03 m s^{-1} respectively (Figure 2.18). They suggest that for transiting systems, transit duration variations could allow changes in stellar radius to be monitored and perhaps corrected.

2.5.2 Zeeman effect

Models of activity-induced radial velocity variations (§2.4.5) typically focus on the impact of temperature contrast in spots, according to which the effects should diminish toward longer wavelengths. The *Zeeman effect* (splitting of a spectral line into several components in the presence of a static magnetic field) also has an effect on radial velocity measurements, but in the opposite sense: its relative importance should grow with wavelength because the Zeeman displacement itself grows with λ , and because a magnetic and cool spot contributes more to the total flux at longer wavelengths.

Reiners et al. (2013) modeled the effects of active regions, including both temperature contrast in spots and Zeeman broadening, calculating stellar line profiles using polarised radiative transfer models including atomic and molecular Zeeman splitting over the range $0.5\text{--}2.3 \mu\text{m}$. The amplitude of the radial velocity signal caused by the Zeeman effect can be comparable to that caused by temperature contrast; a spot magnetic field of 0.1 T can produce a similar radial velocity amplitude as a spot temperature contrast of 1000 K.

Table 2.2: Estimates of the tidal velocity amplitude for the 17 systems (from all known in 2012) with $K_{\text{tide}} > 1 \text{ m s}^{-1}$, as determined by Arras et al. (2012, Table 1). For the eccentric orbit cases, the $k = 2$ orbital velocity amplitude, eK_{orb} , is given for comparison. NT under inclination indicates the planet is non-transiting.

Planet	$M_p \sin i$ (M_J)	P_{orb} (d)	e	ω (°)	i (°)	M_\star (M_\odot)	R_\star (R_\odot)	K_{tide} (m s^{-1})	eK_{orb} (m s^{-1})
CoRoT-14 b	7.6	1.51			79.6	1.13	1.21	4.1	
HAT-P-2 b	8.7	5.63	0.517	185	90.0	1.36	1.64	5.0	483
HAT-P-7 b	1.8	2.20			84.1	1.47	1.84	1.0	
HAT-P-23 b	2.1	1.21	0.106	118	85.1	1.13	1.20	3.6	39
HD 41004 B b	18.4	1.33	0.081	178	NT	0.40	0.48	3.6	517
HIP 13044 b	1.3	16.20	0.250	219	NT	0.80	6.70	2.7	30
OGLE-TR-56 b	1.3	1.21			78.8	1.17	1.32	2.1	
OGLE2-TR-L9 b	4.3	2.49			79.8	1.52	1.53	1.9	
SWEEPS-11	9.7	1.80			84.0	1.10	1.45	7.0	
WASP-4 b	1.1	1.34			88.8	0.93	1.15	1.1	
WASP-12 b	1.4	1.09			86.0	1.35	1.60	4.8	
WASP-14 b	7.7	2.24	0.090	255	84.8	1.32	1.30	1.6	91
WASP-18 b	10.1	0.94	0.009	-92	86.6	1.24	1.36	31.9	15
WASP-19 b	1.2	0.79	0.005	3	79.4	0.97	0.99	2.8	1
WASP-33 b	4.6	1.22	0.174	-89	87.7	1.50	1.44	5.9	120
WASP-43 b	1.8	0.81			82.6	0.58	0.93	8.9	
XO-3 b	11.8	3.19	0.260	346	84.2	1.21	1.38	3.1	385

2.5.3 Planet-induced tides

Close-in massive exoplanets raise significant tides in their stellar hosts (§10.11.2). Arras et al. (2012) computed the radial velocity signal due to this fluid motion in the ‘equilibrium tide approximation’, in which the fluid motion is incompressible and follows gravitational equipotentials (§10.11.6).

In the particular case of a circular orbit, the general expression for the tidal amplitude simplifies to (Arras et al., 2012, eqn. 15)

$$\nu_{\text{tide}} = \frac{3}{2} n R_\star \left(\frac{M_p}{M_\star} \right)^3 f_2 \sin^2 \theta_0 \sin [2(nt - \phi_0)] \\ \simeq 1.13 \text{ m s}^{-1} \left(\frac{M_p}{M_J} \right) \left(\frac{M_\odot^2}{M_\star (M_\star + M_p)} \right) \quad (2.51)$$

$$\times \left(\frac{R_\star}{R_\odot} \right)^4 \left(\frac{1 \text{ d}}{P_{\text{orb}}} \right)^3 \sin^2 \theta_0 \sin [2(nt - \phi_0)], \quad (2.52)$$

where n is the orbital frequency, θ_0 and ϕ_0 are defined in terms of the star–planet–observer coordinates, and f_2 contains information on the limb darkening (the second numerical expression uses Eddington limb darkening with $f_2 \simeq 1.1$). The tidal velocity term is largest for massive short-period planets around stars with large radii.

Their study shows that the orbital and tidally-induced radial velocity terms are distinct in phase for large eccentricities. For more circular orbits, and because both the tidal fluid flow and the epicyclic motion of a slightly eccentric orbit produce a radial velocity signature at twice the orbital frequency, the tidally-induced term can mimic a finite orbital eccentric.

The 17 systems for which tidal contributions are predicted to exceed 1 m s^{-1} are given in Table 2.2, the largest

being $\sim 30 \text{ m s}^{-1}$ for WASP-18 b. Interestingly, the WASP-18 planet has a small but finite $e = 0.007 - 0.009$ (Hellier et al., 2009a; Triaud et al., 2010; Miller et al., 2012; Knutson et al., 2014c), and a pericentre longitude $\omega \sim \pi/2$, implying a coincidental orbit alignment with the line-of-sight. That this result is explained, in phase and amplitude, by the tidal velocity signature of an $e = 0$ orbit, suggests that the reported eccentricity is a signature of the tidally induced radial velocity in the stellar host.

As a corollary, the measurement of both the orbital and tidal velocities for non-transiting planets may allow planet mass and inclination to be separately determined solely from radial velocity data.

2.5.4 Planet radial velocity signals

In almost all discussions of exoplanet radial velocities to date, information on the exoplanet orbit is (indirectly) derived from spectroscopy of (and therefore motion of) the host star, and any possible contributory signal from the planet itself is either irrelevant or ignored. Radial velocity measurements accordingly only yield a value of the *mass function* (§2.1.2), with the consequence that (in the absence of further constraints) the planet mass remains uncertainty by the unknown factor $\sin i$.

Direct measurement of the *planet’s* radial velocity, in addition to the star’s barycentric motion, would eliminate the $\sin i$ uncertainty in the orbital orientation, and provide the planet’s true mass (and orbit inclination) directly, even for non-transiting system (§2.1.2).

In principle, such measurements could be made through high-resolution spectroscopy of the light from the host star *reflected* by the planet, or from the (infra-red) thermal *emission* of the planet itself. As detailed

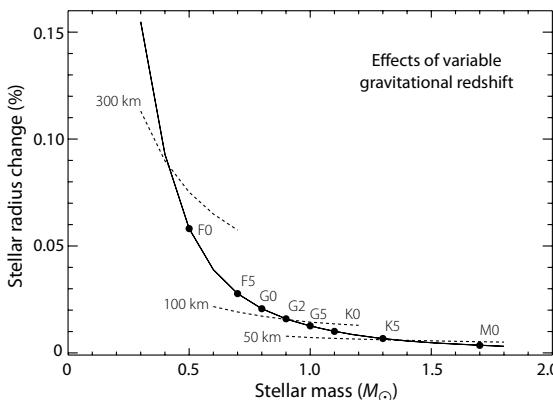


Figure 2.18: The effects of variable gravitational redshift due to a change in stellar radius. The solid curve shows the percentage change in stellar radius required to induce a δV_{grav} equivalent to an Earth-twin radial velocity signal. Solid circles indicate specific spectral types. Dashed curves represent stellar radius variations of 50, 100 and 300 km. From Cegla et al. (2012, Figure 1), © Oxford University Press.

in Chapter 6, such thermal emission has been detected for several transiting planets, either from Spitzer or low-resolution infrared spectroscopy from ground.

Ground-based high-resolution spectroscopy of the reflected or emitted light from the planet is particularly challenging because of the very low planet/star flux ratio ($\sim 10^{-4}$ in the K-band), combined with contamination due to the Earth's atmospheric (telluric) absorption. The method is rendered feasible because the planet signal, unlike that of the telluric lines, is subject to a changing Doppler shift arising from the changing radial component of the planet's orbital velocity (some $100\text{--}150 \text{ km s}^{-1}$ for hot Jupiters). The telluric absorption can therefore be removed while preserving the planet signal, which is then analysed by cross correlation with a range of model planet atmosphere spectra.

Observational results: infrared Since 2012, a small number of planets have been detected at high spectral resolution in the near infrared, notably using VLT-CRIRES, resulting in direct signatures of the planet's orbital motion (Table 2.3).

τ Boo: the first exoplanet measured in this way, *τ Boo b*, had previously been subject to several attempts to measure its orbit inclination through reflected light, without (convincing) success (§6.15.1). Brogi et al. (2012b) reported the detection of CO absorption in the thermal day-side spectrum of *τ Boo b*, from VLT-CRIRES observations (in 2011 April). At $R \sim 100\,000$, they could trace the planet's radial velocity over a large range in phase, yielding $M_p = 5.95 \pm 0.28 M_J$ and $i = 44.5 \pm 1.5^\circ$, thus confirming the planetary (mass) nature of the companion, and demonstrating that atmospheric characterisation is also possible for non-transiting planets. Similar obser-

Relevance of carbon monoxide: CO exhibits a dense sequence of deep absorption lines at around $2.3 \mu\text{m}$. Models predict CO to be one of the most abundant molecules in hot gas giants (Cooper & Showman, 2006; Sharp & Burrows, 2007). The possible detection of CO in the atmosphere of some transiting hot Jupiters has been reported (§11.6.2), mostly from low-resolution ($R < 40$) HST-NICMOS near-infrared spectra or broad Spitzer photometry between $3.6\text{--}24 \mu\text{m}$, either via transmission (primary transit) and emission (secondary eclipse) observations.

Snellen et al. (2010a) detected CO in the transmission spectrum of HD 209458 b using $2.30\text{--}2.33 \mu\text{m}$ high-resolution spectroscopy, providing the first direct measurement of the radial velocity of a transiting hot Jupiter. Brogi et al. (2012b) and Rodler et al. (2012) reported the first detection of an atmospheric chemical species, and associated radial velocity shift, for a *non-transiting* planet.

Strong CO *absorption* indicates an atmosphere with decreasing temperature towards higher altitudes. This is in contrast to the temperature inversion surmised for some other highly irradiated planets (Burrows et al., 2008a; Fortney et al., 2008a), and supports models in which the absorbing compounds believed to cause such atmospheric inversions are destroyed by the ultraviolet emission from the active host star (Knutson et al., 2010).

vations with VLT-CRIRES (in 2011 June) by Rodler et al. (2012) gave $M_p = 5.6 \pm 0.7 M_J$ and $i = 47 \pm 6^\circ$. This high-resolution spectroscopy provided the first 'direct' determination of the mass of a non-transiting planet.

HD 189733: this transiting planet was similarly observed with VLT-CRIRES at 2.0 and $2.3 \mu\text{m}$ (de Kok et al., 2013). No absorption from H_2O , CO_2 , or CH_4 was detected, but a 5σ CO absorption, at a contrast of 4.5×10^{-4} , revealed the planet's orbital radial velocity of $154^{+4}_{-3} \text{ km s}^{-1}$, yielding $M_\star = 0.846 \pm 0.06 M_\odot$ and $M_p = 1.16 \pm 0.06 M_J$. Detection of Na was subsequently reported, with VLT-UVES, by Khalafinejad et al. (2017).

51 Peg: VLT-CRIRES observations gave $M_p = 0.46 \pm 0.02 M_J$ and $i = 79.6 \text{--} 82.2^\circ$ (Brogi et al., 2013).

HD 179949: VLT-CRIRES observations of HD 179949 b (Brogi et al., 2014) revealed CO and H_2O absorption with a combined S/N of 6.3, at a projected planet orbital velocity $K_p = 142.8 \pm 3.4 \text{ km s}^{-1}$, yielding $M_p = 0.98 \pm 0.04 M_J$ and $i = 67.7 \pm 4.3^\circ$.

v And: Keck-NIRSPEC observations detected the planet signal in H_2O for the non-transiting *v And b*, yielding $M_p = 1.7^{+0.33}_{-0.24} M_J$ and $i = 24 \pm 4^\circ$ (Piskorz et al., 2017).

For the future, simulations by de Kok et al. (2014) suggest that there are several (narrow) wavelength regions in the L-band that could yield cross-correlation signals for hot Jupiters (in H_2O , CH_4 , CO_2 , C_2H_2 , and HCN) that might exceed these initial detections by a factor 2–3 for the same integration time.

Simulations for the IGRINS near-infrared spectrograph at McDonald Observatory's 2.7-m telescope (Park et al., 2014), employing similar multiple-line cross-

Table 2.3: Planetary masses based on direct planetary radial velocity determinations at high spectral resolution in the near infrared. All are based on CO absorption at around $2\mu\text{m}$ with VLT-CRIRES. For the only transiting planet, HD 189733 b, the analysis assumes $i = 85.5 \pm 0.1^\circ$, as determined from transit and radial velocity measurements (Triaud et al., 2009).

Planet	$M_p (M_J)$	$i (\circ)$	Reference
τ Boo b	5.95 ± 0.28	44.5 ± 1.5	Brogi et al. (2012b)
"	5.6 ± 0.7	47 ± 6	Rodler et al. (2012)
HD 189733 b	1.16 ± 0.06	85.5 ± 0.1	de Kok et al. (2013)
51 Peg b	0.46 ± 0.02	80.9 ± 1.3	Brogi et al. (2013)
HD 179949 b	0.98 ± 0.04	67.7 ± 4.3	Brogi et al. (2014)

correlation techniques, suggest that comparable measurements might be possible with IGRINS in some dozen cases (Gullikson & Endl, 2013, Table 3), with single epoch observation times between 1–100 min.

Observational results: optical Similar techniques have been developed in an attempt to detect *reflected* light following similar principles (Martins et al., 2013). Evidence for such a spectroscopic detection of reflected light from 51 Peg b, with HARPS, was reported by Martins et al. (2015). A wider treatment of reflected light, and phase curves, is given in Section 6.15.1.

Coronagraphy Riaud & Schneider (2007) proposed extracting the planet's radial velocity signal from the residual stellar flux halo remaining in coronagraphic adaptive optics imaging in combination with the planned extremely large telescopes. The radial velocity signal is estimated to be detectable within the residual contrast ratio of $10^3 - 10^4$, compared to the contrast ratio of 10^{10} in the absence of a coronagraph.

Kawahara et al. (2014) considered the application of coronagraphy to the direct measurement of planetary molecular lines. Even for an unresolved close-in system, the combination of a visible nuller and extreme adaptive optics could reduce the stellar photon noise, and increase the S/N of the planetary signal. For a 30-m telescope observing 55 Cnc b (at an angular separation of $0.6\lambda/D$ in the K band) such *spectroscopic coronagraphy* could increase contrast by a factor 50–130, resulting in a S/N enhanced by a factor 3–6 for warm Jupiters and Neptunes at 10 pc. Tip–tilt error is the most crucial factor, with low-order speckles also contributing. If tip–tilt errors can be reduced to $\lesssim 0.3$ mas, the S/N gain reaches 10–30, allowing the detection of warm super-Earths.

Planet rotation The effect of planet rotation (or spin) on the planetary radial velocity in the day-side spectra of exoplanets was considered by Kawahara (2012), who provides an analytic formula for the intensity-weighted radial velocity from the planet surface, showing that the planet radial velocity signal is distorted by its spin, and that this anomaly can be characterised by the spin radial velocity at the equator combined with a projected

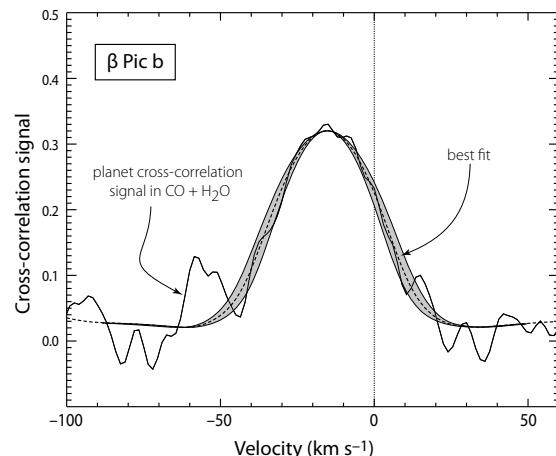


Figure 2.19: Spin of β Pic b: the solid line is the cross-correlation signal of the planet in the lines of CO + H₂O. The dashed line indicates the best fit, rotationally broadened by 25 km s^{-1} . The grey area indicates the 1σ uncertainty in the level of broadening of $\pm 3\text{ km s}^{-1}$. From Snellen et al. (2014, Figure 2b), by permission from Nature/Springer/Macmillan Ltd, ©2014.

angle on a celestial plane between the spin axis and the axis of orbital motion λ_p (in a manner analogous to the Rossiter–McLaughlin effect).

Snellen et al. (2014) reported near-infrared spectroscopic observations, at $R = 100\,000$, of the young gas giant β Pic b. CO absorption in the planet's thermal spectrum is blueshifted with respect to that from the host star by approximately 15 km s^{-1} , consistent with a circular orbit. The combined line profile exhibits rotational broadening of $\sim 25\text{ km s}^{-1}$ (Figure 2.19), meaning that β Pic b spins significantly faster than any planet in the solar system, although it is in line with the extrapolation of the (general) trend in spin velocity with planet mass in the solar system.

Similarly-derived spin measurements have been reported for GQ Lup, also with VLT-CRIRES (from both CO and H₂O), giving $v \sin i = 5.3_{-1.0}^{+0.9}\text{ km s}^{-1}$ (Schwarz et al., 2016a); and for PSO J318–22, with Gemini–N–GNIRS, of $v \sin i = 17.5_{-2.8}^{+2.3}\text{ km s}^{-1}$ (Allers et al., 2016). PSO J318–22 is a free-floating planetary mass member of the β Pic moving group, with an estimated age of 23 ± 3 Myr.

A further rotation rate, for 2M J1207 b, derived from rotational photometric modulation observed with HST–WFC3 (§9.6.5) was determined by Zhou et al. (2016d).

An overview of planet rotation in the solar system is given in Section 12.4.5, and these various spin/rotation rates (solar system planets, and exoplanet spins determined from radial velocity and photometric modulation measurements) are collected in Figure 12.15.

The relatively slow rotation of GQ Lup, for example, is most likely due to its young age, < 5 Myr, viz. when it is still in the process of accreting material and angular momentum (Schwarz et al., 2016a).

Doppler imaging The technique of *Doppler imaging* produces two-dimensional maps of rotating objects using high-resolution spectroscopy, exploiting the varying Doppler shifts across a rotating object. The principles, and application to the surface structure of brown dwarfs and exoplanets, are described elsewhere (box, page 440).

Application to hot Jupiters Complementing thermal (infrared) measurements (§11.4.2), Doppler measurements can also place constraints on the ‘meteorology’ of hot Jupiters. Showman et al. (2013a) have shown that the atmospheric circulation, and Doppler signature, of hot Jupiters splits into two regimes. Under weak stellar insolation, day–night thermal forcing generates fast zonal jet streams from the interaction of atmospheric waves with the mean flow. Air along the terminator (as seen during transit) then flows toward Earth in some regions and away from Earth in others, leading to a Doppler signature exhibiting superposed blueshifted and redshifted components.

Under intense stellar insolation, however, the strong thermal forcing damps these planetary-scale waves, inhibiting their ability to generate jets. Strong frictional drag likewise damps these waves and inhibits jet formation. As a result, this second regime exhibits a circulation dominated by high-altitude, day-to-night airflow, leading to a predominantly blueshifted Doppler signature during transit. Their atmospheric circulation models suggest that cool planets like GJ 436 b lie in the first regime, HD 189733 b is transitional, while planets hotter than HD 209458 b lie in the second regime (Figure 2.20).

Additionally, the amplitude of the Doppler shifts constrains the strength of frictional drag in their upper atmospheres. Under relatively weak drag, the wind speeds at the terminator of their models of HD 209458 b reach $4\text{--}6 \text{ km s}^{-1}$, depending on altitude and forcing conditions. Under strong drag, the wind speeds are slower.

For HD 209458 b, Snellen et al. (2010a) tentatively inferred wind speeds toward Earth of $2 \pm 1 \text{ km s}^{-1}$, consistent with models including atmospheric drag through interaction with a planet magnetic field (Miller-Ricci Kempton & Rauscher, 2012).

2.5.5 Determination of inclination

The unknown inclination of most radial velocity detected systems results in the well-known $\sin i$ dependent uncertainty on M_p . Although most of these topics are covered elsewhere, this section summarises conditions favourable for the determination of orbit inclinations.

Astrometric constraints Constraints are currently provided in a very few cases by high-accuracy astrometry, notably from HST (§3.7). More can be expected from Gaia (§3.9). Prospects from WFIRST and EXO are considered by Brown (2015b).

Transit constraints Direct inclination constraints are available for the dozen planets discovered from radial velocity observations but subsequently found to transit (see Appendix D).

Dynamical constraints For some multi-planet systems, mutual inclinations can be constrained by considerations of long-term dynamical stability (§2.12).

Circumbinary systems For planets orbiting binary stars, the stellar binary orbit may be more easily determined by astrometry, such that dynamical constraints on the inclination of the planetary orbit with respect to the plane of the sky and the

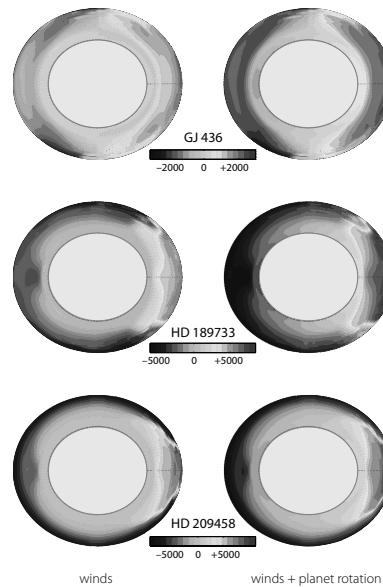


Figure 2.20: Line of sight wind velocities (m s^{-1}) along the terminator at mid-transit, for three hot Jupiters. The radial coordinate is log pressure, from $2 \times 10^7 \text{ Pa}$ (inside) to 0.2 Pa (outside). Left: winds alone; right: wind plus planet rotation. Top to bottom: models transition from high-altitude velocities with both blue and redshifted components, to velocities that are entirely blueshifted. From Showman et al. (2013a, Figure 8), by permission of IOP Publishing/AAS.

orbit’s ascending node may also be possible. For the circumbinary HD 196885 b, the most likely configurations for the planet are either nearly co-planar (prograde or retrograde), or highly inclined orbits near the Lidov–Kozai equilibrium points at $i = 44^\circ$ or $i = 137^\circ$ (Giuppone et al., 2012b).

Resonant systems In well-observable resonant systems, monitoring of the resonant argument and the precession rates may lead to determination of the true masses. For GJ 876, for example, Rivera et al. (2005) used a 3-planet Newtonian fit to derive a minimum mass for planet d of $M_p \sin i = 5.89 \pm 0.54 M_\oplus$ and, assuming co-planar orbits, an inclination of the GJ 876 planetary system to the plane of the sky of $\sim 50^\circ$.

Tidal fixed points Long-term orbital evolution of multi-planet systems under tidal dissipation may converge to a stationary state, the *tidal fixed point*. This is characterised by a lack of oscillations in the eccentricities and apsidal alignment among the orbits, with its precise nature dictated by mutual interactions among the planets as well as non-Keplerian effects. Under these circumstances, the $\sin i$ degeneracy can be resolved, and true masses obtained. Details, and applicability to 61 Vir, are described by Batygin & Laughlin (2011).

Statistical and Bayesian constraints Statistical decoupling of the effects of unknown inclination angle on the derived mass distribution can be made if the true underlying mass distribution is known or assumed, e.g. if the true mass function is described by a power law (Brown, 2011; Lopez & Jenkins, 2012). Ho & Turner (2011) argued that the value of $\sin i$ for a system with any particular observed value of $M_p \sin i$ cannot be assumed to be drawn randomly from an isotropic distribution,

but rather the posterior distribution from which it is drawn depends on the intrinsic distribution of M_p for the exoplanet population being studied. They applied a Bayesian derivation to several ‘toy models’ for the intrinsic distribution of M_p .

Effect on phase curves Constraints on i may in principle be available as potentially detectable phase variations along the orbit (§6.15.3), which are dependent on the companion mass (Kane & Gelino, 2012a). Effects to be considered include the reflected light from the companion, which depends on its radius (itself somewhat dependent on mass), semi-major axis, eccentricity, orbital inclination and geometric albedo (determined by its atmospheric properties). As the companion mass increases, ellipsoidal variations may contribute (§6.15.3). Self-luminosity may contribute for massive planets and brown dwarfs of young age, $\lesssim 1$ Gyr. Kane & Gelino (2012a) give predicted examples for the high-mass companions to HD 114762 and HD 162020.

Radial velocity tidal signatures Measurement of both orbital and tidal velocities for non-transiting planets may allow planet mass and inclination to be separately determined solely from radial velocity data (§2.5.3).

Radial velocities of planets Direct measurement of the radial velocity of a *planet*, in addition to that of the barycentric motion of the host star, eliminates the $\sin i$ uncertainty in the orbital orientation, and provides M_p and i directly, even for non-transiting systems (§2.5.4).

2.6 Radial velocity instruments

2.6.1 Overview

A list of radial velocity instruments applied to exoplanet detection and characterisation is given in Table 2.4. It includes both large successful survey instruments, which now typically make use of many tens of nights on each of many telescopes throughout the world, as well as an incomplete list of more modest and less certain plans; it is intended to give a flavour of the activity in this field.

Most instruments are échelle spectrographs; specific ‘names’ are given when commonly used. The division between ‘early’ and ‘ongoing’ is a little arbitrary, but intended to recognise the radial velocity surveys that were being undertaken in advance of the first more secure exoplanet detection announced in 1995.

For some of the dedicated spectrographs installed on (or planned for) the largest telescopes, Figure 2.21 shows the collecting area versus spectral resolution, for both optical and near-infrared instruments.

From an optical design perspective, most modern high-resolution échelle spectrometers belong broadly to one of two families (Malacara & Thompson, 2004, Section 9.11): the double-pass Littrow design as used in SOPHIE (Perruchot et al., 2008), MIKE (Bernstein et al., 2003), PFS (Crane et al., 2010) and others; and the white-pupil design pioneered by ELODIE (Baranne et al., 1996) and used by FEROS (Stahl et al., 1999), UVES (Dekker et al., 2000), HARPS (Pepe et al., 2000), HERMES (Raskin et al., 2011) and others.

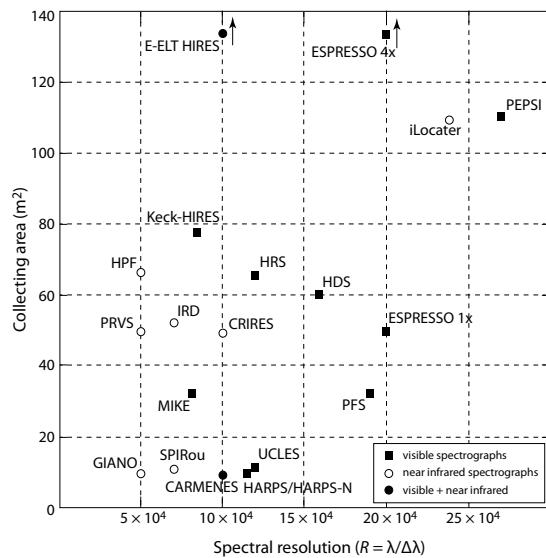


Figure 2.21: Collecting area versus resolution for existing and planned visible and near-infrared radial velocity spectrographs on large telescopes. From Crepp et al. (2016, Figure 1), with permission ©SPIE.

2.6.2 State-of-the-art in échelle spectroscopy

Of the numerous high-performance radial velocity instruments now operational, this section gives a short description of just two established exemplars: HARPS at the ESO 3.6-m telescope at La Silla, and HIRES at the Keck I 10-m telescope in Hawaii.

HARPS HARPS (High Accuracy Radial Velocity Planet Searcher) built on experience gained with ELODIE and CORALIE, and was designed to reach accuracies of around 1 m s^{-1} (Mayor et al., 2003). It has been operating at the ESO 3.6-m telescope, La Silla, since 2003 (Rupprecht et al., 2004). It is an échelle spectrograph with $R = 115\,000$, and operates over a spectral range 378–691 nm distributed over échelle orders $N = 89$ –161. The detector consists of two CCDs (totaling $4\text{k}\times 4\text{k}$, $15 \mu\text{m}$ pixels), with one spectral order ($N = 115$, 530–533 nm) lost in the gap between the two.

The spectrograph is housed in a vacuum vessel, thermally controlled to a few mK to minimise spectral drifts due to temperature and air pressure variations. It is fed by two fibers, one collecting the stellar light, the other simultaneously recording either a Th–Ar reference spectrum, or the background sky. The fibers have an aperture on the sky of 1 arcsec, with image scrambling providing a uniform spectrograph pupil illumination independent of decentering.

In 60 s integration it produces a signal-to-noise ratio of 110 per pixel at 550 nm for a $V = 6$ mag G2V star, and a photon noise error of about 0.9 m s^{-1} . Errors introduced by guiding, focus, and instrumental uncertainties lead to typical long-term radial velocity accuracies of about 1 m s^{-1} rms for spectral types later than G, and for non-rotating stars with $v \sin i < 2 \text{ km s}^{-1}$. Short-term precision of 0.2 m s^{-1} and long-term precision of 0.3 – 0.6 m s^{-1} have been achieved (Pepe & Lovis, 2008).

Table 2.4: Overview of spectroscopic instruments applied to radial velocity exoplanet searches. Resolving power, $R = \lambda/\Delta\lambda$, and accuracy σ (in $m s^{-1}$), are indicative, the former tabulating the maximum value claimed, and the latter particularly dependent on magnitude ($\dots \rightarrow \dots$ indicates that accuracies have improved accordingly over the project's development life). References cover a mix of early and more recent reports, and instrumental descriptions, intended to provide an entry to the more detailed literature.

Telescope–instrument	Lead institute(s)		R (max)	σ	Start	Reference
Early:						
CFHT 3.6-m	échelle	DAO/UBC	100 000	15	1980	Campbell et al. (1988), Walker et al. (1995)
KPNO 0.9-m	échelle	Kitt Peak	74 000	4	1987	McMillan et al. (1990), McMillan et al. (1994)
HJS 2.7-m	échelle	U. Texas/McDonald	60 000	4–7	1988	Cochran & Hatzes (1994), Endl et al. (2008a)
Lick/Shane 3-m	HAMILTON	Lick Observatory	60 000	10 → 3	1986	Vogt (1987), Cumming et al. (1999)
ESO CAT 1.4-m	CES	ESO planet search	100 000	20	1992	Hatzes et al. (1996), Kürster et al. (1999)
OHP 1.9-m	ELODIE	France–Switzerland	42 000	15	1993	Mayor & Queloz (1995), Baranne et al. (1996)
Whipple 1.5-m	AFOE	SAO-HAO	70 000	5	1994	Brown et al. (1994), Noyes et al. (1997a)
Ongoing:						
Keck-I 10-m	HIERES	Keck	85 000	3 → 1	1993	Vogt et al. (1994), Vogt et al. (2000)
AAT 3.9-m	UCLES (CYCLOPS2)	Anglo-Australian	100 000	3	1998	Tinney et al. (2001), Horton et al. (2012)
ESO Euler 1.2-m	CORALIE	Geneva	50 000	7 → 3	1998	Queloz et al. (2000b), Tamuz et al. (2008)
McDonald HET 9.2-m	HRS	U. Texas/McDonald	120 000	3	1998	Tull (1998), Cochran et al. (2007)
ESO Paranal VLT 8-m	UVES	ESO	110 000	30 → 2	1999	Dekker et al. (2000), Joergens (2008)
TNG 3.6-m	SARG	Italy/INAF	46 000	3	2001	Gratton et al. (2004)
Tautenburg 2-m	échelle	Thüringer LS	67 000	3–10	2001	Hatzes et al. (2003b)
OAO 1.9-m	HIDES	Japan	70 000	6	2001	Sato et al. (2005b), Sato et al. (2008a)
ESO La Silla 3.6-m	HARPS	Geneva	115 000	0.3–1	2003	Mayor et al. (2003), Mayor et al. (2009a)
MPG 2.2-m	FEROS/FEROS-II	Max Planck–ESO	48 000	10	2003	Kaufer et al. (1999), Setiawan et al. (2008b)
Magellan II 6.5-m	MIKE	Magellan	83 000	5	2004	López-Morales et al. (2008), Minniti et al. (2009)
Mauna Kea/Subaru 8.2-m	HDS	NAOJ	160 000	4–5	2001	Noguchi et al. (2002), Aoki (2014)
BOAO 1.8-m	BOES	Korea	44 000	6	2004	Izumiura (2005)
Xinglong 2.2-m	CES	China	44 000	30	2005	Izumiura (2005)
Sloan 2.5-m	interferometry	Exoplanet tracker	6 700	20	2006	Ge (2002), van Eyken et al. (2007)
OHP 1.9-m	SOPHIE/SOPHIE+	OHP, replaces ELODIE	75 000	6 → 3	2006	Perruchot et al. (2008), Bouchy et al. (2013)
Sloan 2.5-m	MARVELS*	SDSS	11 000	3–20	2008	Ge et al. (2008), Ge & Eisenstein (2009)
Magellan II 6.5-m	PFS	Carnegie	190 000	1	2009	Crane et al. (2008), Crane et al. (2010)
Mercator 1.2-m	HERMES	Belgium	85 000	2.5	2009	Raskin et al. (2011)
Mt Abu 1.2-m	PARAS	India	70 000	3–5	2010	Chakraborty (2008), Chakraborty et al. (2010)
Calar Alto 2.2-m	CAFE	CAHA (Germany/Spain)	67 000	20	2011	Aceituno et al. (2013)
CTIO 1.5-m	CHIRON	Heidelberg/Yale	80 000	<1	2011	Schwab et al. (2010)
TNG 3.6-m	HARPS-N	SAO–Geneva	115 000	0.3–1	2012	Latham (2008b), Cosentino et al. (2012)
Lick/Levy 2.4-m	APF	Lick Observatory	150 000	<1	2013	Vogt et al. (2014b)
Fairborn 2-m AST TOU (EXPERT-III)		Florida	110 000	–	2013	Ge et al. (2012), Ge et al. (2014)
Tubitak Turkey 1.5-m	RTT150	Ankara	55 000	10	2013	Yilmaz et al. (2013), Yilmaz et al. (2017)
SAAO–SALT 10-m	SALT–HRS	SAAO	67 000	3–4	2014	Bramall et al. (2012)
LCOGT 6 × 1-m	NRES	LCOGT (Las Cumbres)	53 000	3	2015	Eastman et al. (2014)
Mt. Hopkins/4 × 1-m	MINERVA	US (Harvard, etc)	80 000	1	2015	Swift et al. (2015a)
Mt. Graham LBT 2 × 8.4-m	PEPSI	AIP Potsdam	270 000	1	2017	Strassmeier et al. (2015)
Near infrared:						
ESO Paranal VLT 8-m	CRIRES	ESO (1–5 μm)	100 000	10 → 5	2007	Käufl et al. (2004), Bean et al. (2010b)
Palomar 5-m	TEDI*	Berkely (0.9–2.4 μm)	20 000	5	2007	Edelstein et al. (2007)
IRTF 3-m	CSHELL	Lowell (1.1–5.5 μm)	46 000	15–50	2011	Crockett et al. (2011), Gagné et al. (2016)
Mauna Kea/Subaru 8.2-m	IRD	NAOJ (0.97–1.75 μm)	70 000	1	2014	Tamura et al. (2012), Kotani et al. (2014)
TNG 3.6-m	GIANO	Italy/INAF (1–2.5 μm)	50 000	7	2014	Oliva et al. (2012c)
Fairborn 2-m AST	FIRST	Florida (0.8–1.8 μm)	72 000	–	2014	Ge et al. (2007)
McDonald 2.7-m	IGRINS	Texas/Korea (1.5–2.5 μm)	40 000	–	2014	Gullikson & Endl (2013), Park et al. (2014)
McDonald HET 9.2-m	HFF	Penn State (0.9–1.7 μm)	50 000	1–3	2017	Mahadevan et al. (2014b)
Mauna Kea CFHT 3.6m	SPIRou	FRA/CAN (0.9–2.4 μm)	70 000	1	2017	Delfosse et al. (2013b)
IRTF 3-m	iSHELL	Lowell (1.1–5.2 μm)	80 000	<3	2017	Rayner et al. (2016)
Mt. Graham LBT 2 × 8.4-m	iLocater	USA/ITA/DEU (0.9–1.3 μm)	240 000	0.1	2018	Crepp et al. (2016)
Magellan 6.5-m	Maroon-X	Chicago (0.7–0.9 μm)	80 000	0.7	2018	Seifahrt et al. (2016)
Gemini 8.1-m	PRVS	Gemini (0.9–1.7 μm)	50 000	7–10	study	Ramsey et al. (2008), Jones et al. (2008c)
Near infrared+optical:						
Calar Alto/CAHA 3.5-m	CARMENES	DEU/ESP (0.55–1.7 μm)	82 000	1	2016	Quirrenbach et al. (2012), Reiners et al. (2018)
TNG 3.6-m	GIARPS	GIANO+HARPS (0.38–2.45 μm)	–	–	2018	Claudi et al. (2016)
Future (optical and/or infrared):						
LAMOST	EDI	China	10 000	1	study	Zhang (2011)
Small telescopes	RHEA	Macquarie, Australia	75 000	–	2017	Feger et al. (2014), Feger et al. (2016)
ESO VLT 4 × 8-m	ESPRESSO	ESO (Geneva, Bern)	200 000	0.1	2017	Liske et al. (2009), Pepe et al. (2014b)
Lowell DCT 4.3-m	EXPRES	Yale (0.7–0.9 μm)	150 000	<0.3	2017	Jurgenson et al. (2016)
ESO Paranal VLT 8-m	CRIRES+	ESO	100 000	<5	2018	Oliva et al. (2012b)
ESO La Silla 3.6-m	NIRPS	Montreal/Geneva (1–1.8 μm)	100 000	–	2019	Conod et al. (2016)
LCO GMT 24.5-m	G-CLEF	Magellan	120 000	0.1	2022	Ben-Ami et al. (2016)
E-ELT 42-m	HIERES	ESO/INAF (0.37–2.5 μm)	100 000	0.01	2024	Oliva et al. (2015a), Marconi et al. (2016)

Lovis et al. (2008) summarised results from the first three years, and detail the global error budget taking account guiding accuracy and photon noise. The HARPS guaranteed time high-precision sample is presented in Sousa et al. (2008).

A polarimetric unit, HARSPol (Piskunov et al., 2011), allows a full polarisation spectral analysis. This has been applied to the discovery of various massive magnetic stars (Alecian et al., 2011, 2014), observations of the planet-host τ Boo (Borsa et al., 2015; Mengel et al., 2016), and others.

Keck–HIRES A second example of high-accuracy radial velocity instrumentation is the Keck I 10-m telescope on Mauna Kea, Hawaii, equipped with the HIRES échelle spectrometer. HIRES was one of the five first-light instruments for the Keck I telescope, designed and built at UCO/Lick Observatory (Vogt et al., 1994). The project started in 1988, and achieved first light in 1993. While described as a conventional échelle spectrometer, it was designed to have a relatively large order separation, at the expense of overall wavelength coverage, to allow both for accurate sky subtraction (for faint objects), and for image slicing (for bright objects).⁹

In the configuration described by Vogt et al. (2000), the resolving power is $R = 80000$, and the spectra span a wavelength range from 390–620 nm. Wavelength calibration uses an iodine absorption cell.

2.6.3 Other optical spectrographs

This section gives a summary of some of the more ambitious or pioneering instruments recently undertaken, ordered alphabetically by instrument name. Others are listed in Table 2.4.

APF (Lick) The Lick Observatory Automated Planet Finder (APF) uses the Levy spectrometer at the Mt Hamilton 2.4-m telescope (Vogt et al., 2014b). Wavelength calibration is provided by an iodine cell. First light was in June 2013, with robotic operation since January 2014. First discoveries, based on Keck–HIRES candidates, were GJ 687 (Burt et al., 2014) and the 4-planet system HD 141399 (Vogt et al., 2014a).

CHIRON (CTIO) CHIRON is a high-resolution échelle spectrograph at the 1.5-m CTIO telescope since 2011, covering 415–880 nm, and yielding $R = 79000 – 136000$ (Tokovinin et al., 2013). The échelle grating is housed in a vacuum enclosure, with temperature held to $\pm 0.2^\circ$. Stable illumination is provided by light-scrambling octagonal multimode fibers, with an iodine cell for wavelength calibration.

EPDS (WIYN) The NASA-funded Extreme Precision Doppler Spectrometer (EPDS) will be operational on the 3.5-m WIYN telescope in 2018, timed with the data flow from the Transiting Exoplanet Survey Satellite (TESS) mission. It is a $R > 70000$ fiber-fed system targeting an accuracy of $0.1\text{--}0.2 \text{ m s}^{-1}$, led by Pennsylvania State University.

⁹Given the impact of discoveries by Keck–HIRES (Table 2.5), it is noteworthy that amongst the first-light science originally foreseen (including quasar absorption lines, Be in the early Universe, Li abundances, and asteroseismology), exoplanet detection and characterisation did not figure (Vogt et al., 1994).

HARPS–N (TNG) The visible spectrograph HARPS–North has been deployed at the Telescopio Nazionale Galileo (TNG, La Palma) since 2012. Based on HARPS, it exploits an ultra-stable environment, includes better scrambling using octagonal fibers and a tip-tilt system to increase image sharpness, and targets measurements under 1 m s^{-1} (Cosentino et al., 2012). Goals include Kepler field observations, with significant TNG time devoted to it (Hébrard et al., 2013a; Berdiñas et al., 2016).

HDS (Subaru) The High Dispersion Spectrograph (HDS) is the échelle spectrograph for the Subaru Telescope (Noguchi et al., 2002). It provides $R = 90\,000 – 160\,000$, sensitivity from 300 nm to $1 \mu\text{m}$, and single exposures covering a range of 150–250 nm. New image slicers that improve the efficiency at very high resolution were installed in 2011 (Aoki, 2014).

PFS (Magellan) The Carnegie Planet Finder Spectrograph (PFS) operates at the Magellan Clay 6.5-m telescope at Las Campanas Observatory, Chile (Crane et al., 2010). It is optimised for high precision and high stability, covers 388–668 nm across 64 orders, provides $R = 38000$, and uses an iodine cell as wavelength reference, achieving accuracies $\sim 1 \text{ m s}^{-1}$.

Meter-class telescopes Instrument advances (e.g. fiber feeds) have enabled lower-cost instruments on meter-class telescopes. These include: in India, the Mt Abu 1.2-m (Chakraborty et al., 2008); in Hungary, the Piszkestető 1.0-m and Szombathely 0.5-m (Csák et al., 2014); in Germany, the Großschwabhausen 0.9-m (Mugrauer et al., 2014a); and in Slovakia, the Stará Lesná 0.6-m (Pribulla et al., 2015; Garai et al., 2017).

2.6.4 Infrared spectrographs

Rationale As radial velocity work in the optical has advanced over the past two decades, interest in high-resolution spectroscopy in the near infrared has intensified, confronting a new range of technical challenges posed by the detectors, cryogenic operation, atmospheric transmission, and wavelength calibration.

To a large extent this interest has been motivated by the search for planets around late M dwarfs. These are brighter in the infrared, less massive (accentuating the radial velocity effect for a given planet mass) and, being cooler, their habitable zones lie closer to the host star. Habitable-zone planets are therefore more easily detectable, and over shorter time spans (Reiners et al., 2010; Rodler et al., 2011).

Simulations based on stellar atmospheric models suggest that the highest precision is achieved around $1 \mu\text{m}$ for cool M dwarfs and around $1.25 \mu\text{m}$ for L dwarfs (Rodler et al., 2011). Line intensity and continuum air-glow emission in the H-band, important for instrument design, is described by Oliva et al. (2013, 2015b). They identified some 1500 emission lines over the range 0.97– $2.4 \mu\text{m}$, of which some 80% are OH transitions.

The influence of star spots is also reduced in the near-infrared compared with the visible (Desort et al., 2007; Barnes et al., 2011a), such that a comparison at optical and infrared wavelengths can help to discriminate planet-induced velocity changes from stellar jitter.

For these reasons, various new instruments are becoming operational in the infrared. This following summarises a few representative instruments, ordered alphabetically, with a more complete listing in Table 2.4.

CRIRES+ (VLT) Upgrade of VLT-CRIRES (Figueira et al., 2010c) to the cross-dispersed CRIRES+ aims for a significant improvement in sensitivity and wavelength coverage (0.9–5.4 μm) as a result of its improved detectors (Oliva et al., 2012b). First light is due in 2018 Q2.

GIANO (TNG) GIANO is a high-resolution ($R = 50\,000$) infrared spectrograph at the 3.6-m Telescopio Nazionale Galileo, La Palma, covering 0.95–2.5 μm in a single exposure (Oliva et al., 2012a,c). It has been operational since 2014. The accuracy is some 7 m s^{-1} with respect to telluric lines, with an rms calibration accuracy of 300 m s^{-1} as determined from U–Ne lamps.

HPF (HET) The Habitable zone Planet Finder (HPF) is a stabilised fiber-fed near-infrared spectrograph for the Hobby-Eberly 10-m telescope (Mahadevan et al., 2012; Stefansson et al., 2016). It covers the Y and J bands (0.9–1.7 μm) to provide precise radial velocities optimised for mid-M dwarfs. It includes a cryostat cooled to 200 K, dual fiber-feed, a gold-coated échelle grating, and a Hawaii-2RG (H2RG) detector. Wavelength calibration uses a U–Ne lamp, with tests of laser frequency combs ongoing. First light, on GJ 3470, was reported in 2017 November.

IGRINS (McDonald) The Immersion Grating Infrared Spectrometer (IGRINS) is a compact high-resolution near-infrared échelle spectrograph for the 2.7-m Harlan J. Smith telescope at the McDonald Observatory, covering 1.45–2.45 μm in a single exposure with $R = 40\,000$ (Gullikson & Endl, 2013; Park et al., 2014). The primary disperser is a silicon immersion grating. Individual volume phase holographic gratings serve as cross-dispersers for separate spectrographs covering the H and K bands, with two 2048 \times 2048 Teledyne HAWAII-2RG detectors.

IRD (Subaru) The Infrared Doppler Instrument (IRD) at the Subaru 8.2-m also aims at detecting Earth-like planets around nearby M dwarfs (Tamura et al., 2012; Kotani et al., 2014). It is a fiber-fed, near-infrared spectrometer operating over 0.97–1.75 μm at $R = 70\,000$. New technologies to achieve 1 m s^{-1} include very low thermal expansion ceramic for most of the optical components including the optical bench, and use of a laser frequency comb for wavelength calibration.

SPIRou (CFHT) SPIRou (SpectroPolarimètre Infra-Rouge) at the CFHT 3.6-m is a near-infrared spectropolarimeter covering the range 0.9–2.4 μm , providing unpolarised and polarised spectra at $R = 73\,500$, with accuracies of 1 m s^{-1} (Delfosse et al., 2013b; Moutou et al., 2015a). First light is due in early 2018.

2.6.5 Optical-infrared spectrographs

Rationale Simultaneous observations in the optical and near infrared (Table 2.4, see also Figure 2.15), are likely to be particularly important in extending the M dwarf exoplanet surveys, while at the same time providing strong constraints on wavelength-dependent noise sources such as star spots (e.g. Reiners et al., 2010; Ma & Ge, 2012; Anglada-Escudé et al., 2013a).

This approach has been demonstrated, for example, with two of the stable M dwarfs observed with HARPS,

Barnard's star (GJ 699) and GJ 588. Anglada-Escudé et al. (2013a) showed that Barnard's star is stable to 0.009 m s^{-1} over 4 yr, and that the radial velocity signals correlated with activity disappear when using the reddest part of the HARPS wavelength range. Similar effects are seen in the observations of T Tauri stars at 670 nm and 2.3 μm (Crockett et al., 2012).

CARMENES, with its dual optical-infrared spectrograph arms, has been operational since the start of 2016. E-ELT-HIRES is adopting a similar approach. GIARPS is an ongoing project to operate the HARPS-N (optical) and GIANO (near infrared) instruments simultaneously at the TNG (Claudi et al., 2016). NIRPS is an infrared complement to the HARPS instrument at the ESO 3.6-m telescope, due for operation in 2019.

CARMENES (CAHA) CARMENES (Calar Alto high-Resolution search for M dwarfs with Exoearths with Near-infrared and optical Echelle Spectrographs) is a German–Spanish project, operational at the CAHA 3.5-m telescope in southern Spain since the start of 2016 (Quirrenbach et al., 2011a, 2012, 2014).

It consists of two temperature-stabilised échelle spectrographs together covering 0.55–1.7 μm , fed by (agitated octagonal) fibers from the Cassegrain focus, and employing simultaneous calibration with emission-line lamps (Th–Ne in the optical, ^{235}U –Ne in the infrared) or with a Fabry–Pérot étalon. The visible-light arm (0.52–0.96 μm , $R = 93\,000$, employing a 4096 \times 4096 e2v CCD) is operated at 12 C, while the near-infrared arm (0.96–1.71 μm , $R = 82\,000$, employing two 2048 \times 2048 HAWAII-2RG infrared detectors) is cooled to 140 K.

Of relevance for mid- to late-M stars, the efficiency has been optimised around 1.0 μm (Y band). A survey of 300 M dwarfs has an assignment of 600 nights between 2016–18, with ~60 observations per star. It emphasises very cool stars with spectral type M4V or later, but also includes earlier and brighter M dwarfs (Reiners et al., 2017). With a long-term radial velocity precision target of 1 m s^{-1} per measurement, the goal is to detect a 2 M_{\oplus} planet in the habitable zone of an M5V star, and super-Earths of $\lesssim 5 M_{\oplus}$ across the habitable zones.

First results include follow-up observations in the visual channel of 9 known M dwarfs, with the confirmation of all except one, and the discovery of a second planet in GJ 1148 (Trifonov et al., 2018). Estimated (V-channel) formal rms uncertainties are around 1–2 m s^{-1} , comparable to those for HARPS and Keck–HIRES (their figure 14). The first discovery, a planet around the 10.7 pc M dwarf GJ 617A (HD 14739), for which they estimate $\sigma_{\text{VIS}} = 1.7 \text{ m s}^{-1}$ and $\sigma_{\text{NIR}} = 8.6 \text{ m s}^{-1}$ also made use only of the visual channel (Reiners et al., 2018).

NIRPS (La Silla) NIRPS (Near Infra Red Planet Searcher) is an $R = 100\,000$ infrared spectrograph to complement the HARPS instrument at the ESO 3.6-m telescope, under development by a consortium led by the University of Montreal and the Geneva Observatory. It comprises an adaptive optics and fiber feed system with a 0.4 arcsec field, employing a Shack–Hartmann wavefront sensor with 14 \times 14 sub-apertures, a Hawaii 4RG 4k \times 4k detector, and operating between 0.98–1.8 μm (Conod et al., 2016). Operation is scheduled for 2019 August.

GIARPS (TNG, La Palma) GIARPS targets the simultaneous use of the HARPS-N (optical) and GIANO-B (infrared) spectrographs to achieve high-resolution spectroscopy over 0.38–2.45 μm in a single exposure (Claudi et al., 2016).

2.6.6 Future instrument plans

Future instruments aim for significant improvements in radial velocity accuracy over current levels of $0.3\text{--}0.5\,\text{m s}^{-1}$, with target accuracies $\sim 0.1\,\text{m s}^{-1}$ or better.

ESPRESSO (VLT) ESPRESSO (Echelle Spectrograph for PREcision Super Stable Observations, Pasquini et al., 2009; Pepe et al., 2013b, 2014b) will be installed in the Combined Coudé Laboratory of the ESO–VLT, and linked to the four Unit Telescopes through optical coudé trains. It can be operated either with a single Unit Telescopes, or with up to four combined incoherently for a gain of some 1.5 mag. It targets a radial velocity precision at the $0.1\,\text{m s}^{-1}$ level, with a limiting magnitude 2 mag fainter than HARPS (Dumusque et al., 2011b). It will test a number of new subsystems foreseen for HIRES for E–ELT, including fiber scrambling, CCD thermal stabilisation, and the use of laser frequency combs for wavelength calibration. It saw first light in 2017 November.

iLocater (LBT) iLocater is a cross-dispersed échelle spectrograph covering the Y and J-bands ($0.97\text{--}1.30\,\mu\text{m}$), with $R = 150\,000\text{--}240\,000$, and targeting $0.1\,\text{m s}^{-1}$ accuracies (Crepp et al., 2016). It aims to operate at the diffraction limit of the individual 8.4-m diameter telescopes (each using extreme adaptive optics), with the use of single mode optical fibers to eliminate modal noise (\$2.3.8). Priority will be assigned to follow-up measurements for the TESS mission. In contrast with (all other) seeing-limited spectrometers, iLocater will also reduce contamination from neighbouring stars permitting Doppler studies of close-separation binaries. Operation is targeted for 2018.

HIRES (E–ELT) E–ELT HIRES is a merger, dating from 2012, of two previous high spectral resolution instrument concepts, CODEX (COsmic Dynamics EXperiment, an ultra-stable optical, $0.37\text{--}0.71\,\mu\text{m}$, $R = 120\,000$ spectrograph) and SIMPLE (a near-infrared, $0.84\text{--}2.5\,\mu\text{m}$, $R = 130\,000$ adaptive optics-assisted spectrograph; Origlia et al. 2010). A goal of CODEX was to detect the expansion of the Universe directly, by measuring the Doppler shift of high-redshift quasar Ly- α absorption lines as a function of time (the ‘Sandage test’). CODEX incorporated high-stability slanted volume phase holographic gratings, fiber-optic scrambling to reduce the impact of guiding errors, and wavelength calibration using laser frequency combs (Pepe & Lovis, 2008; Pasquini et al., 2008a,b; Liske et al., 2009).

HIRES is currently foreseen as a dual-channel (optical-infrared) spectrograph, providing $R \sim 100\,000$ over $0.37\text{--}2.5\,\mu\text{m}$ (Marconi et al., 2016). Exoplanet requirements foresee a radial velocity accuracy of $\sim 0.1\,\text{m s}^{-1}$, while the Sandage test requires a night time stability, and absolute wavelength calibration, maintained over decades, of $0.02\,\text{m s}^{-1}$. The latter are considered as goals rather than design drivers. The HIRES Phase A study should be completed in March 2018. First light is tied to that of the E–ELT, currently 2024.

2.6.7 Externally dispersed interferometry

Background Various interferometric-based techniques have been applied to spectroscopy and radial velocity measurements in the past, although not for exoplanets. In *Fourier transform spectroscopy*, for example, light enters a Michelson interferometer through a beam-splitter. The resulting interference pattern, in frequency space, depends on the path delay within the

interferometer, and the recorded intensity is a measure of the light’s temporal coherence. One mirror is moved to introduce a variable path length, the temporal coherence is measured as a function of this path delay, $I(\delta)$, and the spectrum is recovered by Fourier transform of $I(\delta)$. Fourier transform spectroscopy provides for very high spectral resolution, but only for very bright sources due to the inefficient use of photons which are sampled only sequentially at each path delay setting.

Fixed delay interferometers with narrow band passes isolating a single spectral line have used this principle for solar measurements since the 1980s, reaching accuracies of around $3\,\text{m s}^{-1}$ (e.g. Kozhevaton et al., 1995, and references). Other approaches have employed a more extended optical delay path range within the interferometer, without resorting to physical scanning, for example using grisms or holographic gratings (*holographic heterodyne spectroscopy*) in place of the mirrors (e.g. Connes, 1985; Frandsen et al., 1993; Douglas, 1997).

Use of a post-disperser The combination of a Michelson interferometer with post-disperser was suggested for spectroscopy already by Edser & Butler (1898). It entered the exoplanet literature with the work of Erskine & Ge (2000), described as a *fringing spectrometer*. Subsequent developments have been described under the epithets of a *dispersed fixed-delay interferometer*, DFDI (Wang et al., 2011, 2012a), or an *externally dispersed interferometer*, EDI (Ge, 2002; Ge et al., 2002b; Erskine, 2003; Erskine et al., 2003; van Eyken et al., 2004; Mahadevan et al., 2008b; van Eyken et al., 2010; Wan et al., 2011; Erskine et al., 2016).

A prototype instrument ('exoplanet tracker') was used to confirm the known exoplanet, 51 Peg b, at the KPNO 2.1-m by van Eyken et al. (2004), and as the basis of the W.M. Keck Exoplanet Tracker instrument (Mahadevan et al., 2008a,b).

The first planet discovered using this technique was HD 102195 (ET1), made at the KPNO 0.9-m, and confirmed with the KPNO 2.1-m and the HET–HRS (Ge et al., 2006), and subsequently by HARPS (Melo et al., 2007). The planet has $P = 4.1\,\text{d}$ and $M_p \sin i = 0.49 M_J$.

MARVELS The concept also underlies the MARVELS survey at Apache Point Observatory (Multi-object APO Radial-Velocity Exoplanet Large-area Survey). MARVELS formed part of the six-year Sloan Digital Sky Survey SDSS III extension, operating between 2008–14. Its scientific goal was to monitor 11 000 $V = 8\text{--}12$ stars over 800 sq. deg. for six years. The survey targeted 90% F8 and later main sequence stars and subgiants, and 10% G and K giants with $V = 7.6\text{--}12$ (van Eyken et al., 2007; Ge & Eisenstein, 2009; Ge et al., 2009). The instrument acquired radial velocities, at $3\text{--}20\,\text{m s}^{-1}$ depending on V magnitude, for 60 objects simultaneously in the 3° field, by lining up the distinct low-resolution spectra across the CCD (Wan et al., 2006). A trial survey of 420 $V = 8\text{--}12$ mag solar-type stars in 2006 was reported by Ge et al. (2007).

The survey discovered a number of low-mass (sub-)stellar companions, and a few candidate brown dwarfs, designated

MARVELS–n (Fleming et al., 2010; De Lee et al., 2013; Jiang et al., 2013b; Grieves et al., 2017). Of these, MARVELS–1 (TYC 1240–945–1) was subsequently re-classified as a face-on binary (Lee et al., 2011b; Wright et al., 2013), while a further candidate, TYC 3010–1494–1, proved to be a long-period high-eccentricity spectroscopic binary (Mack et al., 2013). A giant planet (MARVELS–7 b) and a brown dwarf candidate (MARVELS–7 c) were discovered orbiting the close binary HD 87646, representing the first close binary with more than one substellar circum-primary companion (Ma et al., 2016a).

Including 10 new MARVELS candidates in the ‘brown dwarf desert’ (\$2.10.5), Grieves et al. (2017) derived a brown dwarf occurrence rate around solar-type stars with $P < 300$ d of 0.56%.

Related initiatives

An All-Sky Extrasolar Planet Survey (ASEPS), using wide-field telescopes to monitor stars to $V = 12$ –13 mag, targeted increasing the number of exoplanet detections by a factor of 10–100 (Ge, 2007; Ge et al., 2007). Observing strategies for large-scale multi-object surveys were simulated by Kane et al. (2007). A similar instrument concept has been considered for LAMOST (Zhang, 2011). A prototype at the BTA 6-m telescope is described by Panchuk et al. (2009, 2010).

Broadly similar principles are used in the *dispersed Fourier transform spectrometer* (dispersed FTS) considered by the USNO (Hajian et al., 2007). But rather than operating at a fixed delay position, the interferogram of the Fourier transform spectrometer is coarsely sampled over a wide range of delay positions, so that a complete high-resolution broad-band spectrum can be reconstructed. The addition of a post-disperser essentially creates several thousand separate narrow-band Fourier transform spectrometers operating in parallel.

Extension to the infrared

The same technique can be used at longer wavelengths. The TEDI instrument (TripleSpec Exoplanet Discovery Instrument) extended the concept to the near infrared (0.9–2.4 μm) by coupling a wide-angle Michelson interferometer to the existing $R = 3000$ TripleSpec échelle spectrometer at the Palomar 5-m telescope, whose spectral resolution alone is insufficient for exoplanet radial velocity measurement. The instrument is used in phase uniform mode, and both interferometer outputs are relayed to the spectrograph (Edelstein et al., 2007, 2008; Muirhead et al., 2011; Erskine et al., 2014).

2.6.8 Absolute accelerometry

A conceptually different approach to determining radial velocity *variations* is to measure accelerations.

Connes (1985, 1994) proposed a system in which the stellar light, and light from a white light source, are passed alternately through a Fabry–Pérot étalon. A feedback loop adjusts the étalon spacing to achieve coincidence. A tunable laser tracks the Fabry–Pérot spacing, and a measurable beat signal is generated from a stabilised laser. The objective is to lock onto the stellar lines, and to monitor fluctuation using terrestrial frequency standards. The radial velocity of the source can be assigned some arbitrary value at some reference time and, accounting for the (known) heliocentric correction terms, subsequent observations provide the absolute acceleration with respect to the observer.

Connes (1985) estimated that for a $P = 10$ yr orbit, a 1 m telescope observing a 10 mag star for 10 yr and an

observing time of 10×1 h per year, would reach an acceleration amplitude of $2 \times 10^{-5} \text{ m s}^{-2}$, corresponding to a radial velocity error of 0.1 m s^{-1} . A laboratory system was built and tested (Schmitt, 1997; Bouchy et al., 1999).

A more recent assessment of the merits of this approach are given by Panchuk et al. (2010).

2.7 Introduction to the radial velocity results

2.7.1 The first radial velocity exoplanets

Early radial velocity surveys (Table 2.4), on a relatively small number of stars, were primarily aimed at characterising the substellar/brown dwarf mass function by searching for binary companions of main-sequence stars with masses below $1 M_\odot$ (Campbell et al., 1988; Marcy & Moore, 1989; Marcy & Benitz, 1989; McMillan et al., 1990; Duquennoy & Mayor, 1991; Tokovinin, 1992). Some were part of programmes to establish improved IAU radial velocity standards (Latham et al., 1989).

As accuracies improved towards plausible planetary signals of 10 – 20 m s^{-1} , existing groups intensified their efforts, and others started new programmes, leading to the monitoring of many more stars over several years.¹⁰

The first radial velocity detections were announced cautiously, and only substantively confirmed some years later. Thus Campbell et al. (1988) identified a possible $P = 2.7$ yr, $1.7 M_J$ object around γ Cep, parameters which were subsequently questioned (Walker et al., 1992), but which were eventually confirmed by the 1981–2002 study of Hatzes et al. (2003a). Latham et al. (1989) reported a $P = 84$ d, $11 M_J$ companion to HD 114762, which they suggested was a probable brown dwarf. These values were confirmed by Cochran et al. (1991), and further refined by Butler et al. (2006b). Hatzes & Cochran (1993) reported a possible $P = 558$ d, $2.9 M_J$ companion to the K giant HD 62509 (β Gem), parameters again substantially confirmed in the 25-year baseline study by Hatzes et al. (2006).

The discovery of 51 Peg b The discovery of a very short-period $P = 4.2$ d ($a = 0.05$ au) $0.47 M_J$ planet surrounding the star 51 Peg, was announced by Mayor & Queloz (1995). The discovery was promptly confirmed by the Lick Observatory group, who were also quickly able to report two new planets around stars that they had been monitoring: 70 Vir (Marcy & Butler, 1996) and 47 UMa (Butler & Marcy, 1996). The compelling realisation that planetary mass objects existed around main

¹⁰Walker (2012) provides a recent perspective on the radial velocity exoplanet searches carried out in the 1970s–1980s, and states ‘It is quite hard nowadays to realise the atmosphere of scepticism and indifference in the 1980s to proposed searches for extrasolar planets. Some people felt that such an undertaking was not even a legitimate part of astronomy’.

sequence stars marked the start of a substantive and world-wide acceleration in exoplanet research.

The reality of 51 Peg b was the subject of some early and intense controversy. In part this was motivated by its unexpectedly short orbital period and close proximity to the parent star (although see footnote, page 62). But an alternative explanation – that the radial velocity shifts arose from non-radial oscillations – was also put forward to explain possible distortions in the absorption line bisector. Studies that followed (Gray, 1997; Hatzes et al., 1997; Marcy et al., 1997; Gray & Hatzes, 1997; Willems et al., 1997; Brown et al., 1998a,b; Gray, 1998; Hatzes et al., 1998a,b) finally resulted in a consensus that the planet hypothesis was the most reasonable.

Recent perspectives have been given by Mayor & Queloz (2012) and Cenadelli & Bernagozzi (2015).

2.7.2 Example radial velocity curves

Out of nearly 700 planets now discovered from radial velocity measurements, a small selection of graphical results is shown in Figures 2.22–2.25 as examples of the variety, measurement quality, and orbit reconstruction possibilities. These figures relate to the following systems (and where all radial velocity derived mass estimates are implicitly uncertain by the factor $\sin i$):

HD 4113 (G5V): Tamuz et al. (2008) used CORALIE to discover a $1.63M_J$ planet in a highly eccentric ($e = 0.90$) $a = 1.28\text{au}$, $P = 526\text{-d}$ orbit. The radial velocity curve shows a sharp, strongly asymmetric form arising from the eccentricity (Figure 2.22), and an underlying linear trend attributable to an outer, longer-period planet.

55 Cnc (G8V): Fischer et al. (2008) used Lick and Keck observations (1989.1–2007.4), to characterise five orbiting planets. Planets were identified by periodogram analysis of residuals to successive Keplerian fits (Figure 2.23). $\{M_p \sin i, P\}$ (M_J, d) are $\{0.024, 2.8\}$, $\{0.83, 14.6\}$, $\{0.17, 44.4\}$, $\{0.15, 261\}$, $\{3.9, 5400\}$.

HD 40307 (K2.5V): Mayor et al. (2009b) used HARPS to characterise the orbits of three ‘super-Earth’ planets (Figure 2.24; the total radial velocity semi-amplitude of the combined three-planet signal is only $\sim 6 - 7\text{ m s}^{-1}$). $\{M_p \sin i, P\}$ (M_\oplus, d) are $\{4.2, 4.3\}$, $\{6.8, 9.6\}$, $\{9.2, 20.5\}$.

61 Vir (G5V): Vogt et al. (2010b) used 4.6 years of Keck–HIRES and AAT data to characterise this ‘super-Earth and two Neptune’ planetary system (Figure 2.25). The planets are all in low-eccentricity orbits; their $\{M_p \sin i, P\}$ (M_\oplus, d) are $\{5.1, 4.2\}$, $\{18.2, 38\}$, $\{24.0, 124\}$.

2.7.3 Present radial velocity census

Apart from the extreme planetary system around the pulsar PSR B1257+12, and until the first transiting planet discovered by photometry in 2003 (OGLE–TR–56: Konacki et al., 2003a), almost the first hundred planets around normal main-sequence stars were discovered

Table 2.5: Number of planets discovered by the 20 most prolific radial velocity instruments, based on the NASA Exoplanet Archive compilation as of the end of 2017. This listed 662 planets discovered by radial velocity measurements. A planet discovered by observations with N instruments, is counted for $1/N$ discoveries by each. A total of 17 radial velocity discoveries are accounted for by other observatories/instruments.

Observatory/instrument	Number
Keck–HIRES	162
ESO–HARPS	154
Euler–CORALIE	51
AAT–UCLES	40
Lick–Hamilton	32
HET–HRS	28
OHP–SOPHIE	27
Okayama–HIDES	26
TNG–HARPS–N	23
OHP–ELODIE	18
Bohyunsan–BOES	15
Magellan–MIKE	13
Subaru–HDS	10
ESO–FEROS	10
Tautenburg	8
APF–Levy	8
McDonald–HJS	6
ESO–CHIRON	5
Whipple–TRES	5
Magellan–PFS	4

using radial velocity techniques. The number of radial velocity discoveries continues to rise, as enlarged surveys, higher measurement accuracies, and longer temporal baselines take effect (Table 2.5).

As of 2010 November 1 (the first edition of this handbook), there were 494 known exoplanets, of which 461 in 390 systems (45 multiple) had measured radial velocity orbits, and 358 had been discovered by radial velocity measurements.

As of the end of 2017 there were 3572 planets (some 2500 from Kepler/K2) of which 662 planets, including 102 multiple systems, had been discovered by radial velocity measurements. A concise bibliography of the radial velocity planets is given in Appendix C.

Figure 2.26 shows the planets discovered from radial velocity measurements by year, as a function both of M_p (ordinate) and semi-major axis a (circle size). The decreasing lower mass bound with time reflects the improving instrumental accuracy leading to smaller detectable values of K . For a given planet mass, $K \propto P^{-1/3}$ (Equation 2.27), which explains their preferentially smaller a .

Figure 2.27 illustrates the region probed by the radial velocity discoveries, compared with transit and astrometric detections at current and projected accuracies.

Figure 2.28 shows the number of exoplanets discovered to date by radial velocity monitoring as a function of spectral type.

As discoveries proceed, more significant statistical

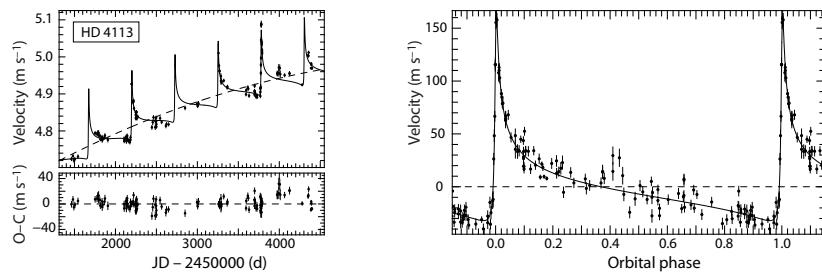


Figure 2.22: Left: radial velocity measurements for HD 4113, observed with CORALIE, with the best-fit Keplerian solution including a linear drift (residuals below). Right: the same measurements folded at $P = 526.58$ d, yielding $K = 97.7 \text{ m s}^{-1}$, $a = 1.28 \text{ au}$, $e = 0.90$, and $M_p \sin i = 1.63 M_J$. The planet-star separation varies between 0.12–2.4 au. The long-term drift, and its curvature, may imply a brown dwarf companion at 8–20 au, with $P \sim 20$ –90 yr. From Tamuz et al. (2008, Figure 1), reproduced with permission © ESO.

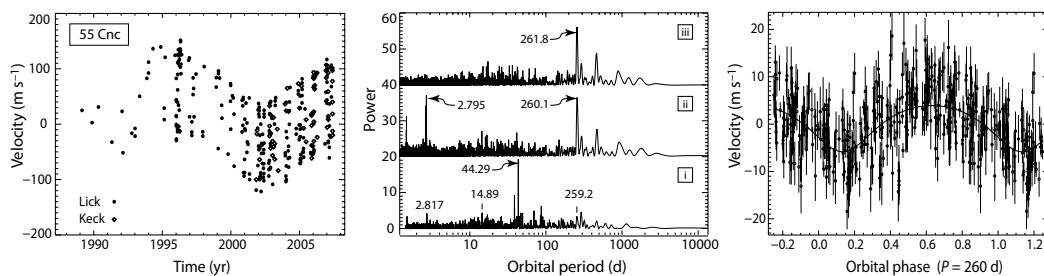


Figure 2.23: The five-planet system 55 Cnc. Left: velocities from Lick/Keck (1989.1–2007.4). Middle: periodograms of residuals to successive Keplerian models: (i) with respect to the 2-planet (14.65 and 5200 d) model; the peak at 44.3 d is due to the third planet; (ii) with respect to the resulting 3-planet model (other major peaks are aliases); (iii) with respect to the 4-planet model (2.8, 14.7, 44.3, and 5200 d); the peak marked is due to the fifth planet (the other major peak is an alias). Right: periodicity of the fifth planet, after subtracting the other planets. From Fischer et al. (2008, Figures 1–4, 8), by permission of IOP Publishing/AAS.

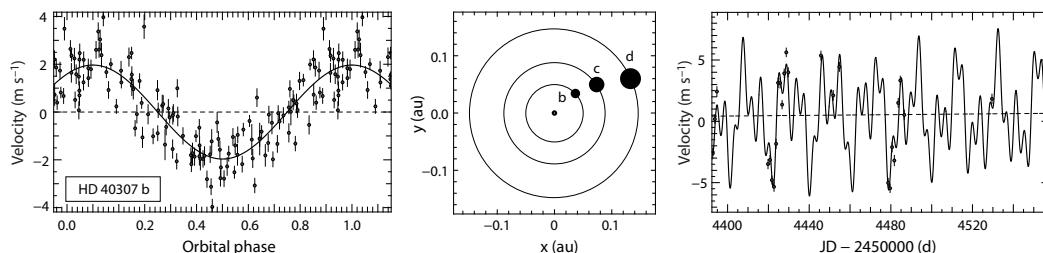


Figure 2.24: The ‘three super-Earth’ system HD 40307, observed with HARPS. Left: phase-folded velocities and Keplerian curve for the lowest mass planet b, after correction for the other two planets. Middle: pole-on view of the orbits; planet size is proportional to mass. Right: measured velocities with the three Keplerian model superimposed. From Mayor et al. (2009b, Figures 2–4), reproduced with permission © ESO.

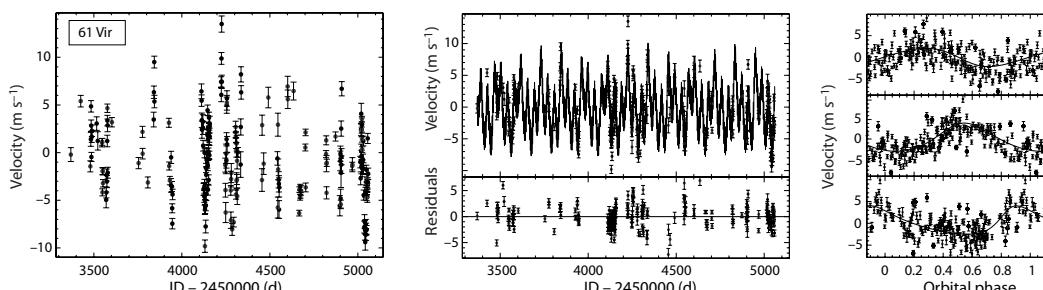


Figure 2.25: The ‘super-Earth and two Neptune’ system 61 Vir, observed with Keck and AAT. Left: relative radial velocities. Middle: observed velocities, and the three-planet model fit (residuals below). Right: modeled contributions from the three planets b–d (top to bottom; $e = 0.1, 0.14, 0.35$ respectively). From Vogt et al. (2010b, Figures 2, 6–7), by permission of IOP Publishing/AAS.

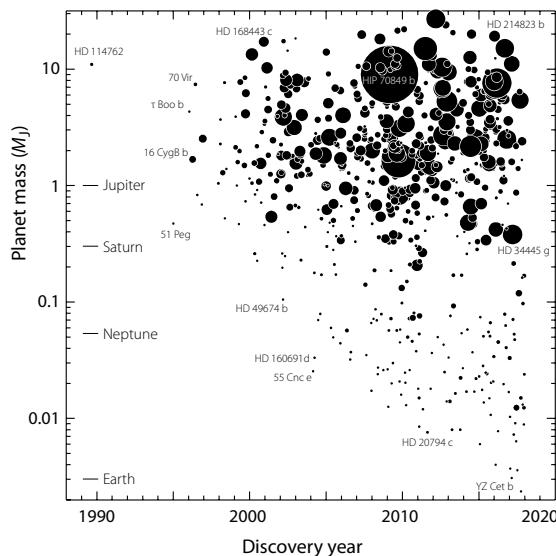


Figure 2.26: Planets discovered by radial velocity measurements, according to mass versus year of discovery. Data are for the 654 planets from the NASA Exoplanet Archive, 2017 December 31. Circle sizes are proportional to semi-major axis. The figure shows the decreasing mass detection limit with time, and the increase in semi-major axis orbits as the temporal baseline increases. Within a calendar year, discovery dates are random.

analyses of the planetary population are becoming possible, with statistically distinct properties and correlations becoming evident through principal component or hierarchical clustering analyses (e.g. Marchi, 2007).

2.7.4 Reviews

Various reviews of radial velocity instrumentation, modeling, noise sources, discoveries, and findings have been made as the field has advanced (e.g. Marcy et al., 2006; Udry & Santos, 2007; Marcy et al., 2008; Santos, 2008; Johnson, 2009; Mayor et al., 2014; Pepe et al., 2014a; Panchuk et al., 2015; Hatzes, 2016a,b; Díaz, 2018; Figueira, 2018; Oshagh, 2018).

2.7.5 On-line compilations

There are numerous on-line resources and compilations related to exoplanet surveys and results. The following (cf. Table 1.4) are particularly useful in the context of radial velocity exoplanet detections, since they maintain up-to-date records of new discoveries, compilations of related parameters, and a variety of tools for their inspection and analysis: (a) the NASA Exoplanet Archive; (b) the Extrasolar Planets Encyclopaedia, which is a compilation of exoplanet announcements, data, and bibliography, maintained by Jean Schneider at the Observatoire de Paris–Meudon; (c) the Exoplanet Orbit Database, which includes results from radial velocity

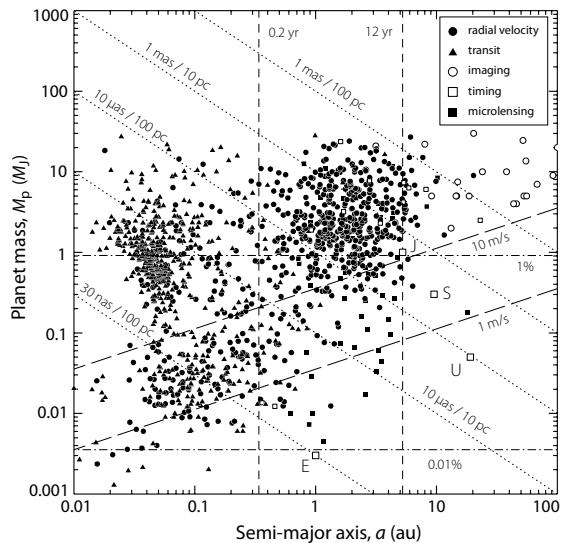


Figure 2.27: Detection domains based on orbital motion, with imaging and microlensing discoveries included. Masses are $M_p \sin i$ for radial velocity discoveries. Lines from top left to bottom right show astrometric signatures of 1 mas and 10 mas at distances of 10 and 100 pc (Equation 3.1, assuming $M_\star = M_\odot$). Short- and long-periods cannot be detected by Gaia astrometry: vertical lines show limits at $P = 0.2$ and 12 yr. Lines from top right to bottom left show radial velocity semi-amplitudes of $K = 10$ and $K = 1 \text{ m s}^{-1}$. Horizontal lines are transit thresholds of 1% and 0.01%, corresponding to $\sim 1R_J$ and $\sim 1R_\oplus$ respectively. Positions of E(arth), J(upiter), S(aturn) and U(ranus) are shown. Data from the NASA Exoplanet Archive, 2017 December 31.

and transit surveys for stars within 200 pc, aiming to incorporate the most secure spectroscopically-measured orbital parameters. It updates the Catalogue of Nearby Exoplanets (Butler et al., 2006b; Jones et al., 2008a), and is maintained by Jason Wright.

2.8 Surveys according to stellar type

For the reasons noted below, most early Doppler planet searches concentrated on G and K main sequence stars (dwarfs), in the mass range $\sim 0.7 - 1.3 M_\odot$. Observing programmes progressively expanded to cover most late-type main sequence stars brighter than $V \sim 7.5 - 8.5$ mag in a systematic manner, with others subsequently focusing on giants, M dwarfs, young stars, stars in clusters, and others. A compilation is given in Table 2.6.

An overview of the various surveys as a function of stellar class or category, and present results, are given in this and subsequent sections. Multiple planet systems are considered in Section 2.12, and planets around binary and multiple stars in Section 2.13.

2.8.1 Main sequence stars

G and K main sequence stars were the primary targets for the early radial velocity surveys. They have nu-

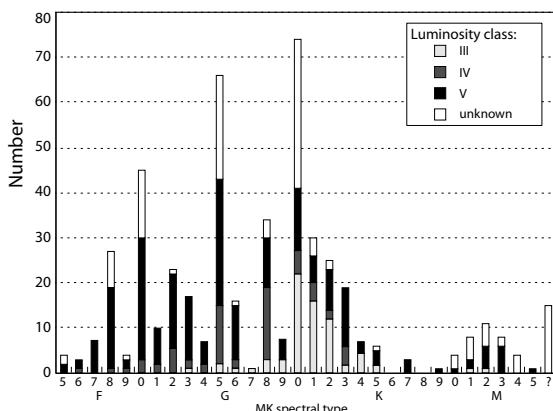


Figure 2.28: Radial velocity discoveries as a function of spectral type and luminosity class. Data, for 481 planet hosts, are from the NASA Exoplanet Archive, 2017 December 31. For 15 hosts, no spectral type is specified (box at far right). For more than 100 additional stars, no luminosity class is specified (unshaded).

merous absorption lines, are relatively bright, are relatively slow rotators with low rotational broadening, and have relatively stable atmospheres, with photometric jitter for inactive G dwarfs extending down to $2\text{--}3 \text{ m s}^{-1}$ (Table 2.1). Early reviews of these radial velocity searches were given by Latham (1997); Butler & Marcy (1998); Latham et al. (1998); Marcy & Butler (1998b); Nelson & Angel (1998); and Marcy & Butler (2000).

Of the larger survey programmes, CORALIE and HARPS are now surveying about 1600 stars in the southern hemisphere (Tamuz et al., 2008). The Anglo-Australian Planet Search has been in operation since 1998 (Tinney et al., 2001).

In the northern hemisphere, the ELODIE survey of more than 1000 targets originally focused on metal-rich stars (e.g. da Silva et al., 2006, 2007). ELODIE was replaced by SOPHIE in 2006, and the same programme has continued under the SOPHIE exoplanet consortium (e.g. Santos et al., 2008).

The Lick Planet Search programme began in 1987 when the first spectrum of τ Cet was taken with an iodine cell and the Hamilton Spectrograph. Upgrades to the instrument improved the Doppler precision from about 10 m s^{-1} in 1992 to about 3 m s^{-1} in 1995. The project detected dozens of exoplanets with orbital periods ranging from a few days to several years. The Lick survey identified the first planet in an eccentric orbit (70 Vir), and the first multi-planet system around a normal main sequence star (ν And). The Lick Planet Search with the Hamilton Spectrograph effectively ended when a heater malfunction compromised the integrity of the iodine cell. More than 14 000 velocities for 386 stars were obtained between 1987–2011 (Fischer et al., 2014).

The California & Carnegie search program uses the Keck, Lick, and AAT telescopes to survey about 1000

stars in the north and south (Valenti & Fischer, 2005).

The N2K consortium is using the Keck, Magellan, and Subaru telescopes to survey the ‘next 2000’ stars (Fischer et al., 2005; López-Morales et al., 2008). Estimates of T_{eff} , [Fe/H], and binarity were constructed from a starting list of more than 100 000 FGK dwarfs (Ammons et al., 2006; Robinson et al., 2006b). Their primary target list was then constructed from the more than 14 000 main sequence and subgiant stars with $d < 110 \text{ pc}$, $V < 10.5$ and $0.4 < B - V < 1.2$. Their final selection of 2000 stars was also biased towards higher metallicity targets, using broad-band photometry to define a subset with $[\text{Fe}/\text{H}] > 0.1$.

A large survey, of $\sim 10 000$ F8 and later dwarfs, started with the Sloan SDSS III MARVELS in 2008 (§2.6.7).

The Lick–Carnegie survey denotes the 20-year survey of FGKM dwarfs with Keck–HIRES. Butler et al. (2017) provide 60 949 precision velocities for 1624 stars, tabulating 357 significant periodic signals, of which 225 have been published as planet claims, 60 are classified as significant candidates awaiting photometric follow-up, and a further 54 requiring further confirmation. Candidates include a $P = 9.9\text{-d}$ companion to the nearby La-lande 21185.

Kinematics and chemistry Adibekyan et al. (2013a) analysed the kinematics and chemical properties of 850 HARPS-observed FGK dwarfs in the solar neighbourhood ($\log g \geq 4 \text{ dex}$; $5000 < T_{\text{eff}} < 6500 \text{ K}$; $-1.39 < [\text{Fe}/\text{H}] < 0.55 \text{ dex}$; $d \leq 45 \text{ pc}$). Stars were assigned to the thin disk, thick disk and high- α metal-rich populations based on $[\alpha/\text{Fe}]$ versus $[\text{Fe}/\text{H}]$.

They found a gradient of the rotational velocity of the thin disk stars with $[\text{Fe}/\text{H}]$ of $-17 \text{ km s}^{-1} \text{ dex}^{-1}$, and a steep positive gradient for both thick disk and high- α metal-rich stars of $+42 \text{ km s}^{-1} \text{ dex}^{-1}$. There was no correlation between orbital eccentricities and metallicity for thin disk stars, but a steep negative gradient for both the thick disk and high- α metal-rich stars of -0.18 dex^{-1} , in reasonable agreement with more distant SDSS stars (Lee et al., 2011d).

The results suggest that radial migration played an important role in the formation and evolution of the thin disk. For the thick disk stars, heating and migration (rather than accretion) could explain the steep positive gradient of V_{ϕ} with $[\text{Fe}/\text{H}]$. The high- α metal-rich population shares properties of both the thin and thick disk populations, and comparison with stars from simulations including radial migration suggests that they may have originated from the inner Galaxy.

2.8.2 Early-type dwarfs

Observationally, radial velocity measures for higher mass main sequence stars (A–F spectral type) are complicated by three effects: a smaller number of spectral

Table 2.6: Radial velocity surveys focusing on particular stellar types. Surveys, instruments, and descriptions evolve with time, so the table is indicative rather than definitive. References describe the sample or more recent global results. Listings are chronological in survey start date within each category, and there is often overlap between some categories.

Selection	Telescope/instrument	N(stars)	Start	Reference
Main sequence:				
F8–M5	ESO–CES [ended 1998]	37	1992	Endl et al. (2002); Zechmeister et al. (2013)
GK, $V < 7.65$, $v \sin i < 5 \text{ km s}^{-1}$	OHP–ELODIE	324	1994	Queloz et al. (2001)
FGKM (California–Carnegie)	Keck+Lick+AAT	1040	1995	Valenti & Fischer (2005)
FGK	Lick–Hamilton	107	1995	Marcy & Butler (1998a)
F7–M5, $< 50 \text{ pc}$	Keck–HIRES	530	1996	Vogt et al. (2000)
FGKM	Lick+Keck	844	1998	Nidever et al. (2002)
FGK IV–V, $V < 8$, $\delta < -20^\circ$	AAT–UCLES	300	1998	Jones et al. (2002a)
F–M, metal-rich	AAT–UCLES	20	1999	Tinney et al. (2003b)
metal-poor	Keck–HIRES	200	2003	Sozzetti et al. (2006a, 2009b)
metal-poor	HARPS	104	2003	Santos et al. (2007, 2011)
F–M (N2K), metal-rich	Keck+Magellan+Subaru	2000	2004	Fischer et al. (2005, 2007a)
early-type, B8–F7	ELODIE/SOPHIE+HARPS	185	2004	Desort et al. (2009b)
metal-rich, Hipparcos-based	OHP–ELODIE/SOPHIE	1061	2004	da Silva et al. (2006)
F stars	Tautenburg	–	2005	Hartmann et al. (2010)
low activity (Rocky Planet)	AAT–AAPPS	55	2007	O’Toole et al. (2009b,a)
metal-rich (Calan–Hertfordshire)	ESO–HARPS (CHEPS)	350	2008	Jenkins et al. (2009a, 2013a)
F8 and later	SDSS III MARVELS	10 000	2008	Ge et al. (2009)
metal-poor	FEROS	70	2008	Setiawan et al. (2010)
FGK, metal poor	ESO–HARPS	–	2010	Santos et al. (2014)
Distance-limited samples:				
distance-limited, $d < 50 \text{ pc}$	La Silla–CORALIE	1650	1998	Udry et al. (2000); Tamuz et al. (2008)
FGK distance-limited, $d < 57.5 \text{ pc}$	ESO–HARPS	850	2008	Lo Curto et al. (2010); Sousa et al. (2011b)
GKM (Eta–Earth)	Keck–HIRES	230	2008	Howard et al. (2009, 2011a,b, 2014)
Solar twins and Jupiter analogues:				
solar twins	ESO–HARPS	88	2010	Ramírez et al. (2014c)
Jupiter analogues	AAT	–	2011	Wittenmyer et al. (2011c, 2013b, 2014a, 2016a)
Jupiter analogues	SOPHIE	–	2012	Boisse et al. (2012b)
Jupiter analogues	HARPS	–	2012	Zechmeister et al. (2013)
M dwarfs:				
M2–M5, $V < 11.5$	Mt Wilson 2.5-m	65	1984	Marcy et al. (1986)
Lick/Keck	Lick+Keck	24	1994	Marcy et al. (1998)
$V < 7.5$	AAT–UCLES	7	1998	Jones et al. (2002a)
CNS3, $d < 9 \text{ pc}$, $V < 15$	OHP–ELODIE/SOPHIE (N)	127	1995	Delfosse et al. (1999a)
" $d < 12 \text{ pc}$	La Silla–FEROS (S)	200	2000	Bonfils et al. (2004)
ESO survey	VLT–UVES	40	2000	Zechmeister et al. (2009)
"	VLT–UVES	26	2002	Guenther & Wuchterl (2003)
McDonald survey	HET–HRS+HJS	100	2002	Endl et al. (2003)
$d < 11 \text{ pc}$	HARPS	200	2003	Mayor et al. (2009a); Bonfils et al. (2013a)
ultracool (late M and L)	Keck-II NIRSPEC	59	2004	Blake et al. (2010)
Palomar	Palomar–TEDI	100	2007	Edelstein et al. (2007)
K–M (M2K)	Keck	1600	2009	Apps et al. (2010); Fischer et al. (2012a)
M5–M9 (ROPS)	Magellan–MIKE	7	2012	Barnes et al. (2012)
M dwarfs (optical+infrared)	CAHA–CARMENES	300	2016	Quirrenbach et al. (2014)
Subgiants:				
subgiants	Lick+Keck	159	2006	Johnson et al. (2006b, 2010c, 2011a)
subgiants/giants	AAT–UCLES (Pan-Pacific)	170	2011	Wittenmyer et al. (2011a, 2016c,d)
Giants:				
G/K/clump giants	Lick+Keck	45	1999	Nidever et al. (2002)
"	Lick–CAT	179/373	1999	Hekker et al. (2006b); Quirrenbach et al. (2011c)
"	Okayama–HIDES/Subaru	300	2001	Sato et al. (2008a,b, 2010)
"	La Silla–FEROS	83	2001	Setiawan et al. (2003a,b)
" (in 13 open clusters)	HARPS	115	2003	Lovis & Mayor (2007)
" (Penn State/Torun)	HET–HRS	1000	2004	Gettel et al. (2012a); Niedzielski et al. (2016a)
"	Tautenburg 2-m	62	2004	Döllinger et al. (2007, 2009a)
"	BOA–Okayama	190	2005	Omiya et al. (2011)
"	SDSS III–MARVELS	1000	2008	Ge et al. (2009)
"	BOA–BOES	55	2010	Han et al. (2010)
"	La Silla–FEROS	166	2010	Jones et al. (2013b)

(continued over)

Table 2.6, continued.

Selection	Telescope/instrument	N(stars)	Start	Reference
Open clusters:				
Hyades giants	Okayama-HIDES	4	2001	Sato et al. (2007)
Hyades dwarfs	Keck+others	94	2002	Cochran et al. (2002)
13 clusters	CORALIE+HARPS	115	2003	Lovis & Mayor (2007)
M67	HARPS	88	2003	Pasquini et al. (2012); Brucalassi et al. (2014)
Praesep	TRES	53	2012	Quinn et al. (2012b)
Young stars:				
young stars, Cha I cloud	VLT-UVEES	12	2000	Joergens (2006)
young stars, 1–100 Myr	FEROS+HARPS	200	2003	Setiawan et al. (2008a,b)
β Pic/TW Hya associations	Keck+NIRSPEC	20	2010	Bailey et al. (2012)
young associations (8–300 Myr)	HARPS	26	2010	Lagrange et al. (2013)
T Tauri	IRTF-CSHELL	9	2010	Crockett et al. (2012)
Binary and multiple stars:				
CORAVEL single-lined (SB1)	ELODIE+CORALIE	101	2001	Eggenberger (2010)
wide binaries, $a = 100 - 1000$ au	TNG-SARG	100	2003	Gratton et al. (2003)
binary/multiple	Keck-HIRES/HET-HRS	450	2003	Konacki (2005a,b)
visual binaries	Okayama-HIDES	9	2004	Toyota et al. (2009)

lines due to higher surface temperatures; line broadening due to higher rotational velocities, of order $v \sin i \approx 100 - 200$ km s $^{-1}$ for A-type stars (Galland et al., 2005a); and large atmospheric ‘jitter’ due to higher surface inhomogeneities and pulsation, of as much as 50 m s $^{-1}$ or more for mid-F stars (Saar et al., 1998, see also Table 2.1). Together, these limit the applicability of the basic cross-correlation method (Griffin et al., 2000).

Chelli (2000) considered cross-correlation in Fourier space, which allows enhanced suppression of high-frequency noise and low-frequency continuum variations. The correlation is made between each spectrum of the target star and a reference spectrum specific to that star constructed from the sum of all the available spectra. When applied to early-type stars (Galland et al., 2005a), it suggested improved detection possibilities: for example, reaching the planetary domain with ELODIE for A-type main sequence stars with $v \sin i$ up to 100 km s $^{-1}$ for $P < 10$ d, and for all A and F-type stars with HARPS, even for large $v \sin i$ (Figure 2.29).

The resulting search for a sample of A–F main sequence stars, with ELODIE/SOPHIE in the north, and HARPS in the south, led to the first detection of a $9.1 M_J$ companion around the F6V star HD 33564 (Galland et al., 2005b), a $25 M_J$ planet/brown dwarf in orbit around HD 180777, a pulsating A9 dwarf with a high rotation of $v \sin i = 50$ km s $^{-1}$ (Galland et al., 2006b), and a series of subsequent discoveries from the southern survey (Desort et al., 2009b, 2010, and references).

2.8.3 Evolved stars: subgiants and giants

Observational complications inherent to high-mass dwarfs can be somewhat circumvented by observing their counterparts which have evolved away from the main sequence. After main sequence hydrogen burning, radii expand and atmospheres cool, leading to an

increase in metal lines. Angular momentum loss as stars cross the subgiant branch results from a coupling of stellar winds to rotationally-generated magnetic fields (Gray & Nagar, 1985; Schrijver & Pols, 1993; do Nascimento et al., 2000). The combination of cooler atmospheres and smaller rotational broadening makes evolved stars well suited for precision radial velocity surveys.

Subgiants Subgiants offer a reasonably favourable region of the HR diagram for radial velocity searches, with low rotational velocities of order $v \sin i \lesssim 5$ km s $^{-1}$, and reasonably low photometric jitter of around 5–10 m s $^{-1}$ (Wright, 2005, see also Table 2.1). Surveys at Keck and Lick have been made around 150 subgiants, selected from Hipparcos according to $V < 7.6$ mag, $2 < M_V < 3.5$, and $0.55 < B - V < 1.0$ (Johnson et al., 2006b), and numerous planets have been reported (Johnson et al., 2007b; Johnson, 2008; Johnson et al., 2010b,c, 2011b; Harakawa et al., 2010). Some 70 radial velocity detected systems around luminosity class IV hosts were listed in the NASA Exoplanet Archive at the end of 2017.

G and K giants Various radial velocity surveys of G and/or K giants have been made (Table 2.6): at Lick (e.g. Frink et al., 2002; Nidever et al., 2002; Hekker et al., 2006b, 2008; Reffert et al., 2015; Trifonov et al., 2015); Okayama-HIDES and Subaru-HDS (e.g. Sato et al., 2003, 2005b, 2008a,b, 2010); La Silla-FEROS (Setiawan et al., 2003a,b, 2004; Jones et al., 2013b); CORALIE and HARPS (Lovis & Mayor, 2007); Tautenburg (Hatzes et al., 2005; Döllinger et al., 2007, 2009a); HET-HRS (Niedzielski et al., 2007; Niedzielski & Wolszczan, 2008; Nowak & Niedzielski, 2008; Niedzielski et al., 2009a,b); the Sloan SDSS III MARVELS survey of some 1000 GK giants (Ge et al., 2009); BOAO-BOES (Han et al., 2010; Lee et al., 2012b); and Penn State-Torün (Gettel et al., 2012a).

By 2008, nine (radial velocity detected) planets around stars with $M \gtrsim 1.5 M_\odot$ had been reported (John-

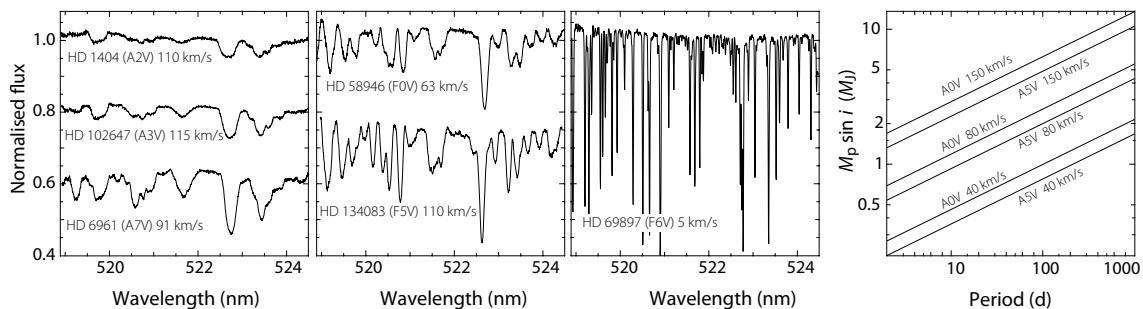


Figure 2.29: Left three panels: spectra of early-type stars acquired with ELODIE, showing the progression in both the number and the broadening of spectral lines for representative examples of A2V, A3V, A7V, F0V, F5V and F6V spectral types. Right: simulated mass detection limits for A-type stars observed with HARPS. From Galland et al. (2005a, Figures 1 and 10), reproduced with permission © ESO.

son, 2008), more than 40 by 2012 (Jones et al., 2013b; Refert et al., 2015), and some 90 by the end of 2017.

In contrast to main sequence stars, there is a lack of planets orbiting interior to ~ 0.6 au (Figure 2.30). This may be the result of tidal interaction between the planet and the stellar envelope, such that planets orbiting interior to a certain distance would spiral inward to be engulfed by the host star (e.g. Siess & Livio, 1999b; Sato et al., 2008a; Villaver & Livio, 2009; Kunitomo et al., 2011). However, subgiants (of relatively smaller radii) show a similar trend, meaning that formation or migration effects, in addition to stellar evolution, may also be present (Johnson et al., 2007b; Bowler et al., 2010a).

2.8.4 M dwarfs

Motivation Towards lower stellar masses, late K and M dwarfs are faint in the optical, such that high signal-to-noise spectra tend to be limited by telescope size. But they are important targets for exoplanet surveys because of their prevalence in the Galaxy (and therefore the large numbers which exist close to the Sun), and because differences in the number and type of exoplanets orbiting them reflect different conditions in the protoplanetary disks which provide additional constraints for planet formation theories.

Since they are brighter in the infrared, the development of near-infrared spectroscopy (§2.6.4) has intensified the interest in M dwarfs. Being less massive than GK dwarfs, the radial velocity amplitude is accentuated for a given planet mass (Equation 2.27). And being cooler, their habitable zones lie closer to the host star, such that habitable-zone planets are more easily detectable, and over shorter time spans (Reiners et al., 2010; Rodler et al., 2011).

Surveys As a consequence, numerous M dwarf radial velocity surveys have been initiated (Table 2.6). These started with the pioneering survey of 65 M dwarfs at the Mount Wilson 2.5-m telescope in 1984 (Marcy et al.,

1986). Radial velocity instruments incorporating programmes with a substantial focus on M dwarfs in the infrared are listed in Sections 2.6.4, 2.6.5, and 2.6.6.

The HARPS survey of 102 southern M dwarfs was carried out between 2003–2009 with radial velocities at the level of $1\text{--}3\text{ m s}^{-1}$ (Bonfils et al., 2013a). The sample corresponds to a volume-limited list of M dwarfs with $d < 11$ pc, $\delta < +20^\circ$, $V < 14$ mag, and $v \sin i \lesssim 6.5\text{ km s}^{-1}$, and is sensitive to a few Earth-mass planets for periods up to several hundred days.

The M2K programme (Apps et al., 2010; Fischer et al., 2012a) is a Doppler survey of M and K dwarfs selected from the SUPERBLINK proper motion survey (Lépine & Shara, 2005; Lépine & Gaidos, 2011), and set up to better understand the frequency and architectures of planetary systems around low mass stars.

The Red Optical Planet Survey was initiated with the Magellan 6.5-m and MIKE spectrograph, operating at $0.5\text{--}0.9\text{ }\mu\text{m}$ (Barnes et al., 2012). For 7 late-type M dwarfs (M5.5–M9), radial velocities at $10\text{--}30\text{ m s}^{-1}$ accuracy suggest a habitable zone detection limit $\sim 6M_\oplus$.

The CARMENES survey of 300 M dwarfs has been assigned 600 nights at the CAHA 3.5-m telescope between 2016–18 (Quirrenbach et al., 2014; Alonso-Floriano et al., 2015). At 1 m s^{-1} precision, goals are to detect $2M_\oplus$ planets in the habitable zone at M5V, and $\lesssim 5M_\oplus$ more generally. Simultaneous optical–infrared measurements should constrain effects of stellar activity.

Massive planets Microlensing surveys suggest that both ice and gas giant planets are common beyond the ice line (Gould et al., 2010b). However, the fraction of gas giant planets detected inside the ice line by Doppler surveys is relatively low for late K and early M dwarfs.

At the time when only 11 planets were known, Marcy et al. (1986) already concluded that gas giants occur less frequently around M dwarfs, a conclusion supported by later studies (Endl et al., 2003, 2006a; Butler et al., 2006a; Bonfils et al., 2007; Cumming et al., 2008).

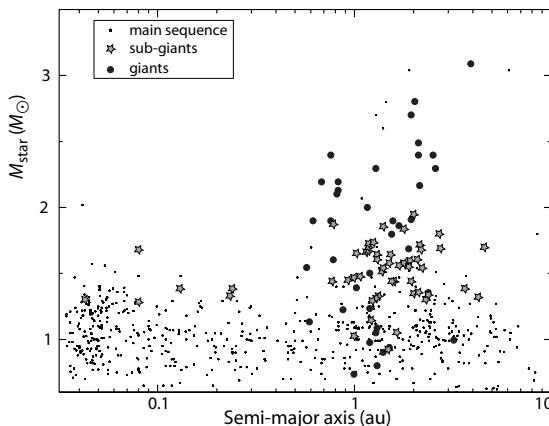


Figure 2.30: Semi-major axis distribution for planets around evolved stars, as of early 2013. Giants are indicated by filled circles, and subgiants by filled stars. For comparison, planets around main sequence stars are shown as small black circles. The giant star at $a \sim 0.1$ au in the original figure corresponded to HIP 13044, which has since been retracted and therefore removed from this figure. From Jones et al. (2013b, Figure 1), reproduced with permission © ESO.

By late 2010, only seven Doppler-detected giant planets ($M > 0.3M_J$) were known around six M dwarfs (Johnson et al., 2010d). These included GJ 317 of $1.2M_J$ (Johnson et al., 2007a, Figure 2.31), and one of the most massive at $2.1M_J$ around HIP 79431 (Apps et al., 2010).

Quantitatively, Endl et al. (2008a) estimated a 1.3% detection rate for Jovian-type planets at $a < 1$ au earlier than M5V. Johnson et al. (2010d) estimated that $3.4^{+2.2}_{-0.9}$ % of stars with $M_\star < 0.6M_\odot$ host planets with $M \sin i > 0.3M_J$ and $a < 2.5$ au. Restricted to metal-rich stars with $[Fe/H] > +0.2$, the occurrence rate rises to $10.7^{+5.9}_{-4.2}\%$. Bonfils et al. (2013a) found that giant planets ($M_p \sin i = 100 - 1000M_\oplus$) have a low frequency ($f \lesssim 0.01$ for $P = 1 - 10$ d and $f = 0.02^{+0.03}$ for $P = 10 - 100$ d). Santerne et al. (2016b) found an occurrence rate of giant planets with $P < 400$ d of $4.6 \pm 0.6\%$.

Lower-mass planets At lower planetary masses, planets are common around M dwarfs. Bonfils et al. (2013a) found that super-Earths ($M_p \sin i = 1 - 10M_\oplus$) appear to be very abundant, with $f = 0.36^{+0.25}_{-0.10}$ for $P = 1 - 10$ d and $f = 0.35^{+0.45}_{-0.11}$ for $P = 10 - 100$ d.

Kepler provides more robust statistics for lower mass planets, and also suggests that a reduced planet occurrence with later spectral type only applies to gas giants (Dressing & Charbonneau, 2013). Correcting for lower transit detection around higher mass stars with larger radii, Howard et al. (2012b) found that 20–30% of low-mass stars have planet candidates with Neptune-like radii between $2 - 4R_\oplus$, while the fraction of more easily detected Jupiter-radius planets remains at a few percent. Howard et al. (2012b) also observed a rising occurrence of small-radius planets among cooler, less massive stars.

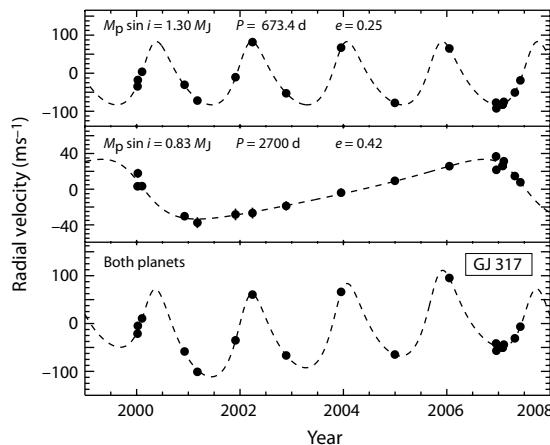


Figure 2.31: The M dwarf GJ 317, and the best-fitting two-planet model. The top two panels show the single-planet fits with the other component removed. Adding an outer planet with $P \approx 2700$ d decreases the scatter of the residuals from 12.5 to 6.32 m s^{-1} , and the reduced χ^2 from 2.02 to 1.23 . From Johnson et al. (2007a, Figure 4), by permission of IOP Publishing/AAS.

Multi-planet systems Latham et al. (2011) analysed multi-planet transiting systems detected with Kepler, and found that solar type and hotter stars are more common hosts of single transiting planets, while multi-planet systems are more often detected around cooler stars. Among the 170 multi-planet systems detected by Kepler, 78% contain planets no larger than Neptune, while close-in gas giant planets are far less common in multi-planet systems (Latham et al., 2011).

Further evidence of a high multiplicity of planets around low-mass stars was given by Anglada-Escudé et al. (2013a), who noted that there were three known systems with 3–4 planets around M dwarfs at the time (GJ 581, GJ 676A, and GJ 876), while only 12 M dwarfs were reported to have a single orbiting planet.

Of particular dynamical interest is GJ 876 (\$2.12.4), with its two resonant Jovian-mass planets and an inner super-Earth (Rivera et al., 2005), and a fourth outer Uranus-mass planet in a Laplace resonance, with periods 30.4, 61.1, and 126 d (Rivera et al., 2010b).

Metallicity effects While the planet–metallicity correlation among solar-type stars is most evident for those hosting large-radius planets, a similar correlation for low-mass stars is significant even among hosts of small-radius planets (Schlaufman & Laughlin, 2011). Neves et al. (2013) used improved M dwarf metallicity estimates from HARPS ($\sigma \sim 0.08$ dex) to confirm the correlation for giant planets, while for Neptune-mass and smaller hosts no significant correlation was found.

Habitable zone planets Over the planetary mass range $1 \leq M \sin i \leq 10M_\oplus$, the frequency of habitable-zone planets is $0.41^{+0.54}_{-0.13}$ (Bonfils et al., 2013a).

2.9 Surveys according to other criteria

2.9.1 Nearby stars and volume-limited samples

Eta–Earth survey The Eta–Earth survey is a NASA–UC programme undertaken by the California Planet Search group with Keck–HIRES and targeting the nearest stars (the survey name reflects the notation assigned to the fraction of (Sun-like) stars that have at least one planet in the habitable zone, η_{\oplus} , §11.7.4). The target list is composed of 232 G, K, and M dwarfs (of which 66 are M dwarfs), being nearby ($d < 25$ pc), bright ($V < 11$), and with low chromospheric activity ($R'_{HK} < -4.7$).

Eta–Earth has detected several low-mass planets (Howard et al., 2009, 2011a,b, 2014). It has also demonstrated that for G and K dwarfs the planet mass function rises steeply with decreasing planet mass, i.e. that small planets are common (Howard et al., 2010b).

Volume-limited surveys Volume-limited samples include the $d < 50$ pc sample surveyed by Keck–HIRES (Vogt et al., 2000), and the $d < 57.5$ pc sample surveyed by HARPS (Lo Curto et al., 2010; Sousa et al., 2011a). The latter includes a range of spectral types from M0–F2, with target stars known to have low levels of activity, low rotation rates, and to be free of known binaries and variable stars (Lo Curto et al., 2010; Sousa et al., 2011b). Samples have been selected on the basis of Hipparcos parallaxes, and will presumably be refined and extended with those from Gaia.

2.9.2 Specific nearby stars

Various nearby stars are now known to host planets, showing that even the immediate solar neighbourhood is rich in planetary systems.

Within 7 pc, radial velocity surveys have discovered gas giants orbiting ϵ Eri (Hatzes et al., 2000), GJ 876 (Delfosse et al., 1998; Marcy et al., 1998), and GJ 832 (Bai-ley et al., 2009); intermediate mass planets (Neptunes and super-Earths) orbiting GJ 674 (Bonfils et al., 2007), GJ 876 (Rivera et al., 2010b), HD 20794 (Pepe et al., 2011), and GJ 581 (Mayor et al., 2009a), an approximately $1M_{\oplus}$ planet orbiting α Cen B (Dumusque et al., 2012), and a $5M_{\oplus}$ planet orbiting GJ 15A (Howard et al., 2014).

Amongst the radial velocity detected planets, a number have attracted particular attention in view of their proximity to Earth, or their historical interest:

- α Cen B: an Earth-mass planet orbiting α Cen B, a member of the closest stellar system to the Sun, with $P = 3.236$ d and $a = 0.04$ au, was reported from HARPS observations by Dumusque et al. (2012). This was subsequently considered controversial because of the presence of strong stellar activity (Hatzes, 2013b), and attributed to the artefacts of the window function (time sampling) in the original data (Rajpaul et al., 2016).
- *Barnard's star*: the long controversy surrounding the possible existence of a planet around the M dwarf Barnard's star (GJ 699), the second nearest star system to Earth (box, page 83), unless

essentially face-on, was emphatically concluded on the basis of HARPS observations by Anglada-Escudé et al. (2013a), and even more so with the 248 precise Doppler measurements obtained from Lick and Keck between 1987–2012 by Choi et al. (2013a). The latter found no significant periodic amplitudes above 2 m s^{-1} , setting upper limits on the minimum mass of any planets with orbital periods from 0.1–1000 d, excluding $M_p \sin i > 2M_{\oplus}$ and $P < 10$ d, and $M_p \sin i > 10M_{\oplus}$ and $P < 2$ yr.

- *Kapteyn's star*: Kapteyn's star is a low-metallicity M subdwarf, and the nearest halo object to the Sun. Its low metallicity and halo kinematics suggest an ancient origin, consistent with its low activity and slow rotation. Bayesian analysis of HARPS, HIRES, and PFS observations led to the announcement of two planets with periods of 48 and 120 d (Anglada-Escudé et al., 2014, the innermost within the habitable zone), raising questions about the system's existence and survival given its halo origin and kinematic history. Robertson et al. (2015c) found that while the star is photometrically stable, spectral activity indices reveal a rotation $P_{\text{rot}} = 143$ d, suggesting that the claimed planetary signal at an integer fraction (1/3) of the stellar rotation period is likely to be an artefact of stellar activity.

- *Proxima Cen*: at $d = 1.295$ pc, the red dwarf Proxima Cen (α Cen C) is the Sun's closest stellar neighbour. From HARPS and VLT–UVES observations, Anglada-Escudé et al. (2016a) reported a terrestrial mass planet with $M_p \sin i = 1.27M_{\oplus}$ and $P = 11.2$ d ($a = 0.05$ au), and with an equilibrium temperature within the range where water could be liquid on its surface.

2.9.3 Solar twins and Jupiter analogues

Solar twins A survey of 88 stars considered as solar twins (§8.5.2) is being conducted with HARPS (Ramírez et al., 2014c), assisted by spectroscopy from Magellan–MIKE (Ramírez et al., 2014c).

Jupiter analogues In identifying systems which are most like the solar system, the presence of a Jupiter analogue, viz. a planet of Jupiter mass in a Jupiter-like orbit, may be a key ingredient (but see §12.2.2).

The high frequency of eccentric orbits amongst giant planets with semi-major axis $a \gtrsim 1 – 2$ au suggests that common perturbation mechanisms may cause non-circular orbits both for the giant planets themselves, and for any lower mass rocky planets closer in. The eccentric giant planets tend to eject lower-mass objects, or enhance orbital eccentricities. Around solar-mass stars, analogues to Jupiter, both in terms of low eccentricity and comparable orbital semi-major axis, may therefore offer the best prospects for harbouring Earth-mass planets moving in circular orbits within the habitable zone.

One of the first such Jupiter analogues reported orbits the $0.88M_{\odot}$ star HD 154345 (Wright et al., 2007). With $M_p \sin i = 0.92M_J$, $P = 9$ yr, $a = 4.2$ au, and $e = 0.16$, the planet moves in a nearly circular orbit with no other giant planets lying inward of it. There may be additional planets with $P = 45 – 60$ d, while the star itself has a heavy element abundance close to solar, with $[\text{Fe}/\text{H}] = -0.10 \pm 0.04$ (Valenti & Fischer, 2005).

In around 2011, the Anglo–Australian Planet Search, in operation since 1998 (Tinney et al., 2001), shifted

its priority to the detection of Jupiter analogues. The observing strategy (Wittenmyer et al., 2013b) and target list were modified, with the aim of producing an accurate determination of their frequency (Wittenmyer et al., 2011c, 2013b, 2014a, 2016a). The modified target list includes stars with long-term velocity stability such that Jupiter analogues can be excluded (e.g. Wittenmyer et al., 2006), as well as those with incomplete orbits suggestive of long-period giants. Various candidates have been reported (Wittenmyer et al., 2011c, 2014a). Wittenmyer et al. (2011c) found that 3.3% of stars in their sample host Jupiter analogues. For 202 solar-type stars, Wittenmyer et al. (2016a) gave a frequency of $6.8^{+2.8}_{-1.6}\%$ for $a = 3 - 7$ au.

Other similar searches are also ongoing, with SOPHIE (Boisse et al., 2012b), with HARPS (Zechmeister et al., 2013), and as part of the Lick–Carnegie programme (Rowan et al., 2016).

Sample results By the end of 2012 this population of long-period Jupiter-mass planets numbered only around 20 candidates (Boisse et al., 2012b). Amongst their properties, there appears to be a fall-off in the mass–period distribution beyond ~ 4 au (e.g. Boisse et al., 2012b, Figure 11), perhaps reflecting the decrease in frequency of giant planets at larger distance ($\gtrsim 5$ au) predicted by core accretion theories (Mordasini et al., 2012a). Those with a complete orbit are mostly of low eccentricity ($e \lesssim 0.25$), while those with incomplete coverages are almost entirely eccentric ($e \gtrsim 0.25$), reflecting the expectation that eccentric orbits are more easily detected for periods longer than the observation times (Cumming, 2004).

Bryan et al. (2016b) used a Doppler survey at Keck combined with NIRC2 K-band adaptive optics imaging to search for massive, long-period companions to 123 systems with one or two planets previously detected from radial velocities. Sensitive to Jupiter-mass planets out to 20 au, they discovered eight new long-period planets, and 20 systems with statistically significant trends indicating the presence of an outer companion beyond 5 au. They reported a total occurrence rate of companions as $52 \pm 5\%$ in the range $1 - 20 M_J$ and 5–20 au, a declining frequency for gas giant planets beyond 3–10 au, that hot gas giants are more likely to have an outer companion than cold gas giants, and that planets with an outer companion have higher than average eccentricities than their single counterparts, suggesting that dynamical interactions between planets may play an important role.

Establishing whether these ‘Jupiter analogues’ represent planets which formed beyond the snow line and migrated inward as the protoplanetary disk dissipated, or whether their present orbits result from some contribution from planet–planet scattering, will require further study. Enlarged samples are expected as the duration of high-accuracy Doppler surveys continues, with a substantial harvest expected from the later Gaia astrometry data releases around 2020 (§3.9).

Examples In terms of its measured period and eccentricity (if not its mass, at $M_p \sin i = 3.9 M_J$), 55 Cnc d is one of the closest Jupiter analogues discovered to date (Nelson et al., 2014b). Estimates of $P = 4909 \pm 30$ d and $e = 0.02 \pm 0.008$ by Endl et al. (2012), were revised to $P = 4867$ d and $e = 0.0269$ by Nelson et al. (2014b), compared to that of Jupiter, viz. $P_J = 4333$ d and $e_J = 0.049$. A Jupiter twin around the solar twin HIP 11915 was reported by Bedell et al. (2015).

2.9.4 Effects of metallicity

Host star metallicity A correlation between the presence of gas giant planets and high metallicity of the host star was noted in the early years of exoplanet discoveries (Gonzalez, 1997). More homogeneous studies have confirmed the trend (e.g. Fischer & Valenti, 2005; Sousa et al., 2011b), with investigations extending to the abundances of numerous elements in addition to iron. The consensus is that the effect is attributable to primordial effects rather than to self-enrichment by accretion, and is considered further in §8.4. The correlation has been taken to imply that core accretion is a plausible mechanism for the formation of giant planets (§10.13.3), while it may also be consistent the formation by the tidal downsizing of giant planet embryos (§10.5.2).

The correlation between occurrence and metallicity may not extend to the lowest mass planets (Udry et al., 2006). A correlation for planets orbiting giant stars was not found by Pasquini et al. (2007), but was present in the sample of Reffert et al. (2015).

From their analysis of planet frequency as a function of [Fe/H], Sousa et al. (2011b) derive a power-law fit, for masses in the range $0.1 - 25 M_J$, given by

$$P_{\text{planet}} = 0.038 \left[\frac{(N_{\text{Fe}}/N_{\text{H}})}{\langle N_{\text{Fe}}/N_{\text{H}} \rangle} \right]^{2.58}. \quad (2.53)$$

Their large volume-limited sample with spectroscopic metallicities again confirms the correlation between planet frequency and host star metallicity, with stars of lower metallicity also hosting less massive planets.

Metal-poor stars The observed correlation between exoplanet occurrence and host star metallicity suggests that radial velocity searches around metal-poor stars should have relatively low success. Nonetheless, specific searches around low-metallicity stars have been targeted since 2003.

A HARPS survey down to $[\text{Fe}/\text{H}] = -1.4$ found only three moderately metal-poor stars hosting giant planets with $P > 700$ d (Santos et al., 2007, 2011). The frequency of super-Earths and Neptunes is $\sim 13\%$, comparable to that for solar-metallicity hosts (Faria et al., 2016b).

Keck–HIRES carried out a 3-year survey around 160 stars down to $[\text{Fe}/\text{H}] = -2.0$, finding only one (with $[\text{Fe}/\text{H}] = -0.49$), and no evidence for giant planets within 2 au (Sozzetti et al., 2006a, 2009b).

FEROS observed 96 A and F stars in the range $-4 < [\text{Fe}/\text{H}] < -0.5$ (Setiawan et al., 2010, 2012). Reports of the very low metallicity HIP 13044, a Galactic halo stream star with $[\text{Fe}/\text{H}] = -2.0$ (Setiawan et al., 2010), and HIP 11952 with $[\text{Fe}/\text{H}] = -1.8$ (Setiawan et al., 2012), have not been confirmed (see Appendix C).

As of the end of 2017, there are still only ~ 40 planet hosts with $[\text{Fe}/\text{H}] < -0.5$, of which only 10 are radial velocity discoveries. The most extreme are the 2-

planet system BD+20 2457 with $[Fe/H] = -1.00$ (Niedzielski et al., 2009b), Kapteyn's star (§2.9.2) with $[Fe/H] = -0.89$ (Anglada-Escudé et al., 2014), and HD 114762 with $[Fe/H] = -0.77$ (Latham et al., 1989; Kane et al., 2011c).

2.9.5 Open clusters

The determination of stellar masses for giant stars is complicated by the fact that evolutionary tracks for stars covering a wide range of masses converge to the same region of the Hertzsprung–Russell diagram. This consideration motivated Sato et al. (2007) and Lovis & Mayor (2007, 115 stars in 13 clusters) to target giants in intermediate-age (0.2–2 Gyr) open clusters, where cluster membership combined with the estimated cluster age provides a more secure estimate of the host star masses compared to that of field giants.

A number of planets orbiting stars in open clusters have now been discovered from radial velocity surveys (see Appendix C, where NGC and Pr(aesepe) members are classified explicitly), including:

Hyades: early observations of 94 main sequence stars yielded no close-in giant planets (Cochran et al., 2002; Paulson et al., 2002, 2003, 2004a,b). The first discoveries came with the $7.6M_J$ planet orbiting the giant ϵ Tau (Sato et al., 2007), and the cluster's first hot Jupiter, HD 285507 (Quinn et al., 2014).

M67 (NGC 2682): Pasquini et al. (2012) reported 11 stars with variability which may indicate planet-hosting stars. A 7-yr programme to measure 88 main sequence and evolved stars in M67 (Pasquini et al., 2012; Brucalassi et al., 2016b, 2017) has discovered five 1-planet systems. Results yield a giant planet occurrence of $18^{+12}_{-8}\%$ slightly higher but consistent with that of field stars, but with a substantially higher rate of cluster hot Jupiters, $5.7^{+5.5}_{-3.0}\%$ (Brucalassi et al., 2017).

Praesepe: Quinn et al. (2012b) reported the discovery of two giant planets amongst 53 single members surveyed in this metal-rich open cluster, Pr 201 b and Pr 211 b (the latter now known to be multiple, Malavolta et al. 2016). These represented the first known hot Jupiters in an open cluster, and the first known to orbit Sun-like, main-sequence cluster stars, providing a lower limit of $3.8^{+5.0}_{-2.4}\%$ on the hot Jupiter frequency in this cluster.

Others: two massive planets (10.6 and $19.8M_J$) were reported around giants in the open clusters NGC 2423 and NGC 4349 respectively (Lovis & Mayor, 2007).

2.9.6 Young stars and associations

While there is a consensus that planets form from the gas and dust disks around young stars, details of their early formation, including the time scale of planet formation, remain unclear. Observationally, radial velocity detection around young stars is hampered by increased host star activity, and planet discoveries in systems with disks have typically derived from direct imaging.

Young stars Searches for planets around young stars with protoplanetary disks, of age 1–100 Myr, have been targeted by FEROS/HARPS (Setiawan et al., 2008a). Discoveries include a $6.1M_J$ planet around HD 70573, a 100 Myr GV star (Setiawan et al., 2007), and a $9.8M_J$ planet around TW Hya, a young (8–10 Myr) star surrounded by a prominent circumstellar disk (Setiawan et al., 2008a). The planet orbits the star with $P = 3.56$ d ($a = 0.04$ au), inside the inner rim of the disk, perhaps demonstrating that planets can form before the disk has been dissipated by stellar winds and radiation. CI Tau is a (candidate) $P = 9$ -d $10M_J$ planet orbiting a 2 Myr old classical T Tauri star still possessing a protoplanetary disk (Johns-Krull et al., 2016a; Rosotti et al., 2017).

Jeffers et al. (2014) estimated detection prospects around young active G and K dwarfs. Models of stellar activity were based on tomographic imaging of the G dwarf HD 141943 and the K1 dwarf AB Dor, including contributions from plages and star spots. They showed that increased stellar activity in young solar-type stars strongly impacts the detection of Earth- and Jupiter-mass planets, and that the degree of activity jitter is directly correlated with stellar $v \sin i$. They concluded that close-in Jupiter-mass planets can be detected around fast-rotating young active stars, Neptune-mass planets around moderate rotators, while super-Earths are only detectable around very slowly rotating stars.

Associations In the β Pic and TW Hya associations, monitoring by Keck–NIRSPEC of 20 young stars excluded the presence of any hot ($P < 3$ d) companions more massive than $8M_J$ and any ‘warm’ ($P < 30$ d) companions more massive than $17M_J$ (Bailey et al., 2012).

A survey of pre-main sequence (T Tauri) stars in Taurus–Auriga was reported by Crockett et al. (2012). Strong magnetic fields give rise to large cool star spots, which can mimic a radial velocity modulation, with simultaneous observations at 670 nm and $2.3\mu\text{m}$ used to constrain the nature of the signals.

A HARPS survey of 26 A–K stars in nearby young (8–300 Myr) associations yielded no planets (Lagrange et al., 2013). Their technique should allow the detection of planets less massive than Jupiter, reaching a few Neptune masses around young stars with ages $\gtrsim 30$ Myr.

2.9.7 Follow-up of transit candidates

Verification of candidate transiting planets typically involves follow-up radial velocity observations, although less direct validation techniques are adopted for host stars too faint for such observations (§6.12.5).

Amongst numerous such efforts required in vetting the large numbers of accessible (brighter) Kepler candidates, and by way of illustration only, many have been confirmed with OHP–SOPHIE (e.g. Ehrenreich et al., 2011a; Santerne et al., 2011b; Bouchy et al., 2011a; Santerne et al., 2011a; Bonomo et al., 2012b; Santerne et al.,

2012b,a; Díaz et al., 2013; Moutou et al., 2013a; Barros et al., 2014b; Deleuil et al., 2014; Santerne et al., 2014; Díaz et al., 2014b; Bruno et al., 2015; Almenara et al., 2015b; Bourrier et al., 2015a). Orbital parameters for 231 transiting giant planets in 45 systems have been refined by HARPS-N over 3 years (Bonomo et al., 2017a).

False-positive rate for Kepler close-in giants From observations with OHP–SOPHIE, Santerne et al. (2012a) derived a false-positive rate for Kepler candidates with transit depth $> 0.4\%$, $P < 25$ d and brighter than 14.7 mag. Out of 46 candidates, they found five undiluted eclipsing binaries, two brown dwarfs, six diluted eclipsing binaries, and nine new transiting planets that complemented 11 already published. The remaining 13 candidates were not followed-up or remain unsolved due to photon noise limitation or lack of observations. From these results they computed the false-positive probability for Kepler close-in giant candidates to be $35 \pm 6\%$, compared with previous estimates of below 5%.

2.10 Masses and orbits

2.10.1 Mass distribution

Various distributions of the radial velocity planet population are shown in Figure 2.32. In the following discussions, it is recalled that for any individual object the measured mass is uncertain by the factor $\sin i$. Similarly, the underlying distribution of M_p (rather than that of the minimum mass $M_p \sin i$) cannot be derived unambiguously (cf. Jorissen et al., 2001). For a random orientation of orbital planes, $\langle M_p \rangle = (\pi/2) M_p \sin i$.

2.10.2 Mass of host star

The results of the surveys of FGK main sequence stars in the mass range $0.3 - 1.4 M_\odot$, along with the few, typically low-mass planets detected around M dwarfs, and the more massive planets found from the evolved star surveys, suggest that the material trapped in the form of planets is strongly correlated with the mass of the primary star (e.g. Lovis & Mayor, 2007, their Figure 11).

If the surface density of disk material furthermore scales with the central star mass, there should be an additional correlation between planet occurrence and stellar mass (Laughlin et al., 2004a; Ida & Lin, 2005b; Kennedy & Kenyon, 2008b).

At higher stellar mass, $M_\star \gtrsim 1 M_\odot$, the occurrence rate of giant planets for 373 G and K-type giant stars, measured at Lick over 12 years, was assessed by Reffert et al. (2015). They found strong evidence for a planet-metallicity correlation, in agreement with the findings for main-sequence stars. Planet occurrence increases in the interval $1 - 1.9 M_\odot$, with a maximum at $\sim 1.9 M_\odot$, and a rapidly falling occurrence beyond $2.5 - 3.0 M_\odot$, with no planets around stars $\gtrsim 2.7 M_\odot$, albeit with 113 stars with $2.7 - 5 M_\odot$ in their sample.

2.10.3 Period distribution

The significant numbers of giant planets with $P \sim 3 - 10$ d ($a \sim 0.05 - 0.1$ au), orbiting close to their host stars, was essentially unexpected before the early exoplanet discoveries.¹¹ The early trend has continued: some 30 radial velocity discoveries with $M_p \sin i > 0.5 M_J$ have $P < 10$ d. As discussed elsewhere, these ‘hot Jupiters’ are believed to have formed much further out, before being scattered or migrating inwards to their present location, with some mechanism halting the migration before the planets fall onto their host stars.

The observed masses and orbits of giant planets within 5 au are reproduced rather well by current theories of their formation and migration (Kley et al., 2005), including their dependence on stellar metallicity and mass (Ida & Lin, 2004b, 2005b; Armitage, 2007b).

The period distribution for $P \gtrsim 1000$ d ($a \gtrsim 2$ au) is less certain, with the apparent decline beyond 3 au probably largely attributable to the limited duration, of around 10 years, of the high-precision Doppler surveys (Kholshevnikov & Kuznetsov, 2002). Some of the best constraints, out to some 8–10 au, come from the almost 30-yr baseline for 17 objects included in the early CFHT and McDonald Observatory observations (Wittenmyer et al., 2006, 2007). Extrapolation out to 20 au, comparable to the orbit of Uranus, is correspondingly uncertain, but even a flat extrapolation would approximately double the known rate (Marcy et al., 2005a). This suggests that a large population of still undetected Jupiter-mass planets may exist between 3–20 au, perhaps implying that as many as 20% of Sun-like stars have a giant planet within 20 au (Cumming et al., 2008).

Adaptive optics imaging targeting the detection of giant planets around young stars from their thermal infrared emission using Gemini–Altair (Lafrenière et al., 2007a) and VLT–NACO (Apai et al., 2008) has yielded independent limits on their occurrence, also suggesting a decreasing number beyond 20–30 au.

Current theories support the idea that gas giants may be rare beyond 20 au. The declining densities of protoplanetary disks with increasing radius, combined with the longer dynamical time scales, together imply a longer time scale for planetary growth compared with protoplanetary disk lifetimes of order 3 Myr (Alibert et al., 2005a; Hubickyj et al., 2004).

Mass–period relation A correlation is seen between M_p and a (Figure 2.32b). There is a paucity of massive planets on short-period orbits (Zucker & Mazeh, 2002; Udry et al., 2002; Jiang et al., 2007). There is also a rise in

¹¹Although Struve (1952), in considering the timeliness of radial velocity searches, commented that ‘*It is not unreasonable that a planet might exist at a distance of 1/50 au... Its period around a star of solar mass would then be about 1 day.*’

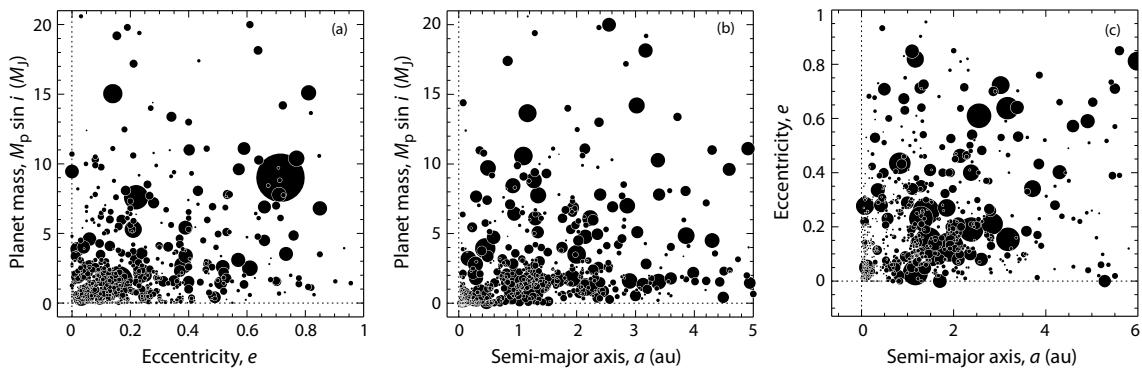


Figure 2.32: Properties of the radial velocity planets: (a) planet mass $M_p \sin i$ versus eccentricity e , with circle sizes proportional to a ; (b) planet mass $M_p \sin i$ versus semi-major axis a , with circle sizes proportional to e ; (c) eccentricity e versus semi-major axis a , with circle sizes proportional to $M_p \sin i$. Data are from the NASA Data Archive, 2017 December 31.

the maximum planet mass with increasing a (Udry et al., 2003c). Since more massive planets are presumably formed further out in the protoplanetary disk, where accretion material is abundant and orbital paths longer, the larger masses are perhaps less easily displaced by whatever migration mechanism forces them inward.

2.10.4 Eccentricities

Most pre-discovery theories of planetary formation suggested that exoplanets would be in circular orbits similar to those in the solar system (Boss, 1995; Lissauer, 1995). In practice (Figure 2.32), there is a significant correlation between a (or P) and e (Stepinski & Black, 2000). Close-in planets are in preferentially low eccentricity orbits, while exoplanets with $P \geq 6$ d have e spanning the range 0–0.93, with a median $e \sim 0.3$. A few, long-period, low-eccentricity orbits are found, representing a small sample of solar system analogues.

As discussed elsewhere, the origin of the eccentric orbits probably arises from several mechanisms: gravitational interaction between multiple giant planets; interaction between the giant planets and planetesimals in the early stages of planet formation; or the secular influence of an additional passing mass, either unbound or bound (with the four most extreme eccentricities being found orbiting components of binary systems). Furthermore, for small pericentre distances, tidal circularisation appears to be an important damping mechanism.

With significant numbers of exoplanets now characterised, observed eccentricities can impose an informative prior on their distribution (e.g. Hogg et al., 2010; Savransky et al., 2011; Wang & Ford, 2011; Zakamska et al., 2011; Kipping, 2013b; Lucy, 2013; Wittenmyer et al., 2013c; Limbach & Turner, 2015). There are situations, as in the simulations of astrometric orbit detection with Gaia (Perryman et al., 2014a, see also §3.9), where this knowledge can provide a useful constraint,

rather than assuming (say) a uniform prior over the range $0 \leq e < 1$.

Some authors (e.g. Steffen et al., 2010; Wang & Ford, 2011; Kipping et al., 2012b) have adopted a combination of a Rayleigh distribution (to reflect the effects of planet–planet scattering) and an exponential distribution (to reflect the effects of tidal dissipation).

Zakamska et al. (2011, their Figure 12) determined an intrinsic (de-biased) eccentricity distribution from the radial velocity detection sample of Butler et al. (2006b), representing the sum of two populations: one on circular orbits (38% of all planets) and the remaining a ‘dynamically active’ distribution with eccentricities following the form of Jurić & Tremaine (2008)

$$\frac{dN_{\text{pl}}}{de} \propto e \exp \left[-\frac{1}{2} \left(\frac{e}{0.3} \right)^2 \right]. \quad (2.54)$$

Kipping (2013b) proposed the Beta distribution

$$P_\beta(e; a, b) \propto e^{a-1} (1-e)^{b-1}, \quad (2.55)$$

which can replicate a wide and diverse range of probability distributions. Regression of the cumulative density function of e from radial velocity detected planets, suggests that $a = 0.867 \pm 0.044$ and $b = 3.03 \pm 0.17$ provide a good fit to the observed distribution (Figure 2.33). Furthermore, short-period planets have a higher proportion of low- e orbits than long-period planets (Figure 2.34), consistent with tidal circularisation.

Correlation of eccentricity with multiplicity The known population of exoplanets exhibits a much wider range of orbital eccentricities than the planets of the solar system, with a much higher average eccentricity.

From radial velocity surveys, Jurić & Tremaine (2008) and Rodigas & Hinz (2009) noted that eccentricities in 2-planet systems tend to be lower than in single-planet systems. From 403 (of 441 known) radial velocity planets, including 127 in multi-planet systems with multiplicities up to six, Limbach & Turner (2015) noted an

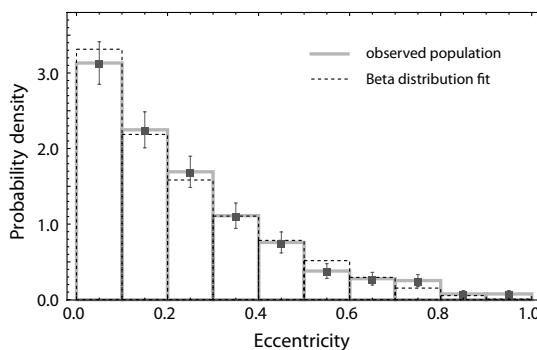


Figure 2.33: Probability density distribution of e for 396 exoplanets (black bars), from www.exoplanets.org. The solid histogram is the observed sample, with errors from Poisson statistics. The dashed histogram is the probability density function of a synthetic population generated using the maximum likelihood parameters of a Beta distribution regressed to the observed sample. Using just two shape parameters, the fitted Beta distribution is consistent with the observed distribution. From Kipping (2013b, Figure 3), © Oxford University Press.

anti-correlation of eccentricity with multiplicity. Specifically, as the multiplicity increases, the eccentricity decreases as a power law of index -1.20 (Figure 2.35). Mean and medians of the solar system planets fits an extrapolation to the 8-planet case rather precisely. Similar results are found for transiting systems (§6.30.3).

The model of Limbach & Turner (2015) implies that $\sim 80\%$ of their 1-planet and 25% of their 2-planet systems have additional, as yet undiscovered, members. They also concluded that the low solar system eccentricities are not anomalous when accounting for multiplicity (as confirmed by Xie et al., 2016) and that, if low eccentricities favour high multiplicities, habitability may be more common in higher-multiplicity systems.

2.10.5 Brown dwarf desert

Due to the large reflex motion they induce in their host stars, early radial velocity surveys were expected to discover brown dwarf companions to solar-type stars with relative ease. Their frequency and orbits would provide observational constraints on theories of formation and evolution of both planets and brown dwarfs.

A prominent feature of the early radial velocity exoplanet discoveries, in contrast, was the general absence of close-in ($a < 3 - 4$ au) substellar objects with masses in the range $10 - 80 M_J$, a paucity referred to as the *brown dwarf desert* (Marcy & Butler, 2000).

Subsequent radial velocity surveys of several thousand stars (including those of SDSS APOGEE and MARVELS, §9.2.2) have confirmed this pattern, finding brown dwarf companions out to $P \sim 10$ yr, but yielding only a small number in the mass and separation range characterising the ‘desert’ (e.g. Grether & Lineweaver,

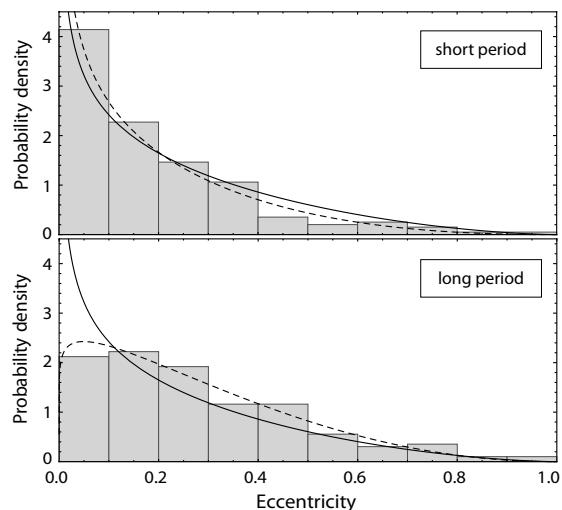


Figure 2.34: Probability density function for short-period (top) and long-period planets (bottom) from www.exoplanets.org, divided at the median period of 382.3 d. The solid line is the regression of a single Beta distribution to both sets. The dashed line is the regression of two independent Beta distributions. From Kipping (2013b, Figure 4), © Oxford University Press.

2006; Sozzetti & Desidera, 2010; Sahlmann et al., 2011b; Ma & Ge, 2014; Bouchy et al., 2016; Santerne et al., 2016b; Troup et al., 2016; Wilson et al., 2016a; Borgniet et al., 2017; Grieves et al., 2017).

Of these, Sahlmann et al. (2011b) placed an upper limit of close ($P \lesssim 300$ d) brown dwarf companions around solar-type stars of $0.3 - 0.6\%$. Grether & Lineweaver (2006) found that $\sim 16\%$ of solar-type stars have close ($P < 5$ yr) companions with $M_p > 1 M_J$, but only $< 1\%$ are brown dwarfs. The minimum number of companions per unit interval in log mass, or the ‘desert’ part of the desert, is at $\sim 30 - 55 M_J$ (Grether & Lineweaver, 2006; Ma & Ge, 2014). The ‘desert’ may be particularly pronounced for solar-mass stars (Duchêne & Kraus, 2013), and more populated for more massive hosts (Guillot et al., 2014b; Troup et al., 2016).

Díaz et al. (2012) aimed to improve statistical knowledge of their frequency in a sample of northern hemisphere stars observed with OHP-SOPHIE. Their improved masses increased the number of substellar objects or candidates with $M_p > 10 M_J$ orbiting solar-type stars in relatively short orbits. Mass and eccentricity distributions of the resulting 60 or so objects are shown in Figure 2.36. The full results of the programme, including occurrence rates, are pending (Wilson et al., 2016a).

Borgniet et al. (2017) derived consistent occurrence rates for AF-type stars: brown dwarf rates for $P = 1 - 1000$ d of 2% and 3% for $M_\star = 1.1 - 1.5 M_\odot$ and $M_\star = 1.5 - 3 M_\odot$ respectively, and corresponding Jupiter-mass ($1 - 13 M_J$) rates of 4% and 6% for $P = 1 - 1000$ d, and 2% and 4% for $P = 1 - 100$ d.

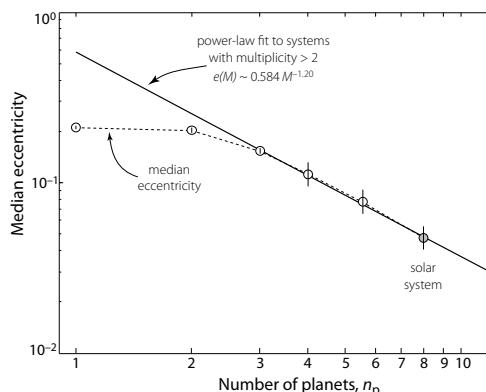


Figure 2.35: Power-law fit to median eccentricity in radial velocity systems with $n_p > 2$. The fit suggests that the 1- and 2-planet data are higher-multiplicity systems with undiscovered members. From Limbach & Turner (2015, Figure 5), with permission ©(2015) National Academy of Sciences.

Imaging and astrometry Similar statistics follow from infrared coronagraphic searches at close separation (McCarthy & Zuckerman, 2004), and other imaging programmes (Kraus et al., 2011; Evans et al., 2012). The majority of those detected may be H-burning stars with low orbit inclination (Halbwachs et al., 2000; Udry et al., 2000). Other mass constraints have come from Hipparcos astrometry (Halbwachs et al., 2000). Greatly improved statistics can be expected from Gaia (§3.9), for which the brown dwarf desert occupies a region of pronounced astrometric sensitivity, and many thousands of desert occupants might be expected.

HST-NICMOS imaging of 255 stars in 201 systems within 10 pc, including 138 M dwarfs in 126 systems, suggests that the desert extends to binaries with low-mass primaries, and is largely independent of primary mass, mass ratio, and separation (Dieterich et al., 2012). At the same time, wide L and T dwarf companions from 2MASS and other near infrared surveys (§9.2.2) suggest that brown dwarfs may not be unusually rare as wide (> 1000 au) companions to F–M0 main sequence stars (Gizis et al., 2001; Grether & Lineweaver, 2006).

Similar low occurrence rates across a range of discovery methods underlines that the absence is not some artefact of the radial velocity method. As a corollary, the clear mass separation between stellar and planetary companions to solar-type stars renders exoplanets distinguishable by their high occurrence at low masses, irrespective of their actual values of $\sin i$.

Examples As of end 2017, there were some 60–70 known brown dwarf companions around solar-type stars (Ma & Ge, 2014; Grieves et al., 2017). Objects occupying the brown dwarf desert include the radial velocity discovery HD 137510, with $M_p \sin i = 26 M_J$, $e = 0.4$, $P = 798$ d (Endl et al., 2004); the transiting CoRoT-3 b (Deleuil et al., 2008) and CoRoT-33 b (Csizmadia et al., 2015); the gravitational lens companion MOA-2007-BLG-197Lb (Ranc et al., 2015); and various SDSS-III APOGEE (Troup et al., 2016), and MARVELS (Grieves et al., 2017) candidates.

Period valley at lower companion mass In their samples of radial velocity detected planets, Jones et al. (2003) and Udry et al. (2003c) noted a significant *period valley* (viz., a lack of planets) in the interval $P = 10 – 100$ d. Udry et al. (2003c) showed that this is attributable to a lack of massive planets ($M_p \sin i > 2 M_J$). These results were further confirmed, at least

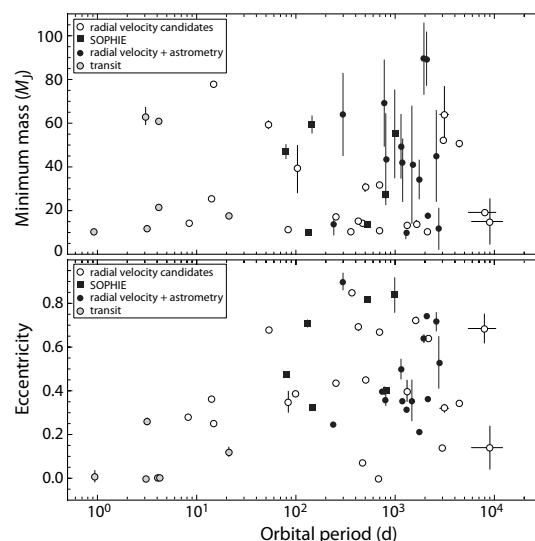


Figure 2.36: Radial velocity discoveries with (minimum) masses between $9 – 90 M_J$ that orbit solar-type stars with $P < 10^4$ d, showing the mass-period (top) and eccentricity-period relations. Objects classified through SOPHIE by Diaz et al. (2012) are shown as solid squares. Other symbols indicate companions for which the true mass is known from astrometry (black circles) or transits (grey circles). Open circles are radial velocity discoveries with no other mass information, with indicated mass interpreted as a ($\sin i$ dependent) lower limit. From Diaz et al. (2012, Figure 13), reproduced with permission ©ESO.

for more massive planets ($M_p > 100 M_\oplus$), from four years of data on 24 bright stable stars from the Anglo-Australian Planet Search (Wittenmyer et al., 2010), which allowed rigorous detection limits to be placed on planetary companions in the range $P = 2 – 300$ d. A possibly related paucity of close-in low-mass planets is seen in transiting planet samples, referred to as the ‘sub-Jovian desert’ (§6.27).

Constraints on formation models The brown dwarf desert has no unambiguous origin. It has been attributed to inward migration within an evolving protoplanetary disk (Armitage & Bonnell, 2002), or as signifying a different formation mechanism for the two (planetary and brown dwarf) populations (Matzner & Levin, 2005). It may emerge as a natural consequence of formation by core accretion, in which unimpeded dynamical accretion of gas is a runaway process that is terminated only when the residual gas is depleted either globally or locally in the form of a gap in the vicinity of their orbits. Since planet masses grow rapidly from 10 to $100 M_\oplus$, gas giants rarely form with asymptotic masses in this intermediate range. The model of Ida & Lin (2004a), for example, specifically predicts few objects in the range $10 – 100 M_\oplus$ and $a < 3$ au. The mass distribution in the desert region described by current population synthesis models is described further in Section 10.13.3.

Duchêne & Kraus (2013) suggested that tidal interaction with host stars may further shape the brown dwarf desert. Damiani & Diaz (2016) found that F-type stars may host massive companions for a significantly longer time than G-type stars for $P < 5$ d; however, they conclude that brown dwarf occurrence rates should be largely unaffected by tidal decay for $P \gtrsim 10$ d, independent of host star mass.

Ma & Ge (2014) found that brown dwarf companions $\lesssim 40M_J$ have an eccentricity distribution consistent with that of massive planets, while those of larger mass have an eccentricity distribution consistent with that of binaries. They also found that host stars of brown dwarf companions are not metal rich, and have a significantly different metallicity distribution compared to host stars of giant planets. They concluded that brown dwarfs $\lesssim 40M_J$ likely form in a protoplanetary disk through gravitational disk instability, and their eccentricity is excited through scattering, while those of higher mass likely form in the same way as stars, through molecular cloud fragmentation (§9.7). Troup et al. (2016) found 14 brown dwarf candidates around metal-poor stars ($[Fe/H] < -0.5$), perhaps consistent with the finding that lower metallicity populations generally have a higher number of binaries (Carney et al., 2003).

2.11 Results according to planet type

Based on 8 yr of HARPS data, Mayor et al. (2011) already found that more than 50% of solar-type stars harbour at least one planet of any mass and with $P \lesssim 100$ d. The occurrence rate of gaseous giant planets grows with $\log P$, strongly increasing with host star metallicity, and with eccentricities up to 0.9 or more. Some 14% of solar-type stars have a planetary companion more massive than $50M_{\oplus}$ with $P < 10$ yr.

2.11.1 Low-mass planets

Udry & Mayor (2008) had already noted that, below $\sim 0.1M_J$, and despite the progressively smaller radial velocity variations and observational incompleteness, which together rendered the detection of lower mass planets more difficult, the frequency of low-mass planets was observed to rise. As of late 2010, eight planets were known with $M \sin i < 0.02M_J$, of which two were transit discoveries (CoRoT-7 b and GJ 1214 b).

By the end of 2017, some 45 radial velocity discoveries had $M \sin i < 0.02M_J$ ($\sim 6M_{\oplus}$) and some 20 below $0.01M_J$ ($\sim 3M_{\oplus}$), mostly in multiple systems with previously-detected planets.

Amongst the early low-mass discoveries were the $7.5M_{\oplus}$ GJ 876 d in the 4-planet system (Rivera et al., 2005), the Neptune-mass 3-planet system around HD 69830 (Lovis et al., 2006; Ji et al., 2007), the 3-planet super-Earth system HD 40307 (Mayor et al., 2009b), and the 2-planet HD 215497 (Lo Curto et al., 2010).

The lowest mass planets are being discovered around the lowest mass stars. The three lowest mass discoveries, at $0.002 - 0.003M_J$ ($0.6 - 0.9M_{\oplus}$, $P_{\text{orb}} = 2 - 5$ d) orbit the $0.13M_{\odot}$ ($d = 3.6$ pc) M dwarf YZ Cet (Astudillo-Defru et al., 2017b).

Low-mass planets are frequently found as members of multi-planet systems, and may indeed be found preferentially in multiple systems (Lo Curto et al., 2010).

Wittenmyer et al. (2011b) presented an analysis of 67 solar-type stars from the Anglo-Australian Planet

Search specifically targeted for high-precision observations. They found that planet occurrence increases strongly with decreasing planet mass. Their results are consistent with those from other surveys: for periods shorter than 50 d, they found that 3% of stars host a giant planet ($M \sin i > 100M_{\oplus}$), and that 17% of stars host a planet with $M \sin i < 10M_{\oplus}$.

The discovery of so many low-mass planets, close to the detection threshold and over a relatively short period of the high-precision Doppler surveys, suggests the existence of a large population of low-mass planets, perhaps reaching 30% for G and K dwarfs (Mayor et al., 2009b). Some of the scatter in radial velocity measurements for individual objects is also likely to be attributable to undetected, low-amplitude, multi-planet systems.

Eta Earth The subject is covered in Chapter 6.

2.11.2 Super-Earths and Neptunes

Based on 8 yr of HARPS data, Mayor et al. (2011) found that the mass distribution for planets with $M_p \lesssim 30M_{\oplus}$ (i.e. super-Earths and Neptunes) behaves differently from gaseous giant planets, increasing strongly between 30 and $15M_{\oplus}$, with $e \lesssim 0.45$, no preference for metal rich stars, and most belonging to multi-planetary systems.

Wolfgang & Laughlin (2012) reported that 30–50% of G and K dwarfs in the solar neighbourhood host planets with $M_p < M_{\text{Neptune}}$ ($17M_{\oplus}$) in orbits of $P \lesssim 50$ d. To reconcile these numbers with the early (Q0–Q2) findings of the Kepler mission, that some 15% of main-sequence dwarfs harbour a short-period planet with $R < 4R_{\oplus}$, they hypothesised either a mass–density relationship extrapolated from the solar system, a population consisting of both dense silicate–iron planets and low-density gaseous planets, or a multi-valued mass–radius relationship, which allows planets of similar mass to have significantly different radii. In such a scenario, HARPS would be detecting a large population of dense low-mass planets, while Kepler detects a large population of gaseous sub-Neptunes.

2.11.3 High-mass planets

For the 1200 FGKM dwarfs in the solar neighbourhood monitored by the California and Carnegie Planet Search programme (Wright et al., 2004a), Marcy et al. (2008) report that 87% of stars observed for more than a decade show no Doppler variations at a 3σ limit of 10 m s^{-1} . These limits largely exclude the existence of a significant number of undiscovered Jupiter-mass planets within 3 au, and Saturn-mass planets within 1 au.

At least 6–7% have giant planets, with $M_p > 0.5M_J$ and $a < 5$ au, a similar fraction also being reported by the Geneva group (Udry & Mayor, 2008). Some 15% of these fall into the category of ‘hot Jupiters’, with $P \lesssim 10$ d, and orbiting very close to their host stars.

The remaining 6% of stars show long-term radial velocity trends, often with significant curvature, indicating substellar, brown dwarf, or planetary companions with orbital periods of a decade or more (Patel et al., 2007).

2.11.4 Hot Jupiters

Wright et al. (2012) determined a fraction $1.2 \pm 0.38\%$ of solar-type stars (FGK dwarfs) in the solar neighbourhood hosting hot Jupiters as measured by the California Planet Survey from the Lick and Keck planet searches. This is consistent with the fraction of $0.89 \pm 0.36\%$ with $M \sin i > 50M_{\oplus}$ and $P < 11$ d reported from HARPS and CORALIE (Mayor et al., 2011). These numbers are more than double those reported by Howard et al. (2012b) for Kepler stars, and the rate of Gould et al. (2006a) from the OGLE-III transit search, although of only marginal statistical significance due to the small sample size.

Further details on hot Jupiters are given elsewhere, including their overall occurrence rates in radial velocity and transit surveys summarised (§1.6.2), discovery and associated properties from transit observations (§6.28), and current understanding of their formation (§10.7.1).

2.12 Multi-planet systems

A number of general dynamical considerations relevant for multi-planet systems, which apply equally to systems found from transit experiments (notably Kepler), are considered in Chapter 10: including a description of resonances (§10.8), stability (§10.9), and chaotic orbits (§10.9.4).

2.12.1 General considerations

Frequency of multiple systems The first radial velocity target known to comprise multiple planets was the triple planetary system *v* And (Butler et al., 1999). This was followed by the discovery of a second planet orbiting 47 UMa (Fischer et al., 2002b), and a resonant pair orbiting GJ 876 (Marcy et al., 2001a). The tenth triple system, HIP 14810, was announced by Wright et al. (2009a).

Some 10–15% of systems discovered from radial velocity measurements are known to be multiple (Figure 2.38), with a further comparable fraction showing evidence for multiplicity in the form of long-term radial velocity trends. An increasing number continue to be discovered, partly as a result of improved accuracies, and partly as a result of longer measurement baselines.

Highest confirmed multiplicities are the 6-planet systems HD 10180, HD 34445, and HD 219134, and the 5-planet systems 55 Cnc, GJ 667C, and HD 40307. There are seven 4-planet systems.

Longer temporal coverage of known systems is leading to the discovery of outer companions with long orbital periods. More than 30 radial velocity discoveries

have $P \gtrsim 10$ yr, with a few of 20–30 yr. Improved accuracy is leading to the discovery of numerous low-mass planets in multiple systems, with around 40 now known below $0.02M_{\mathrm{J}}$ ($6M_{\oplus}$), and 20 below $0.01M_{\mathrm{J}}$ ($3M_{\oplus}$).

Co-planarity The $\sin i$ uncertainty for each planet means that the extent to which multiple systems have co-planar orbits is not well constrained by present observations. Numerical simulations, such as those by Thommes & Lissauer (2003), Adams & Laughlin (2003), and others, suggest that a significant fraction of planetary systems involving giant planets may be substantially non-coplanar. Dynamical mechanisms that lead to fast amplification of the relative inclination are especially effective in the first-order resonance configurations (Thommes & Lissauer, 2003). Also, dynamical relaxation and collisional scattering of protoplanets may favour large relative inclinations, even if they initially emerge in a flat protoplanetary disk.

Constraints on the relative orbital inclinations of individual systems are frequently presented on the basis of long-term numerical orbit integrations, which may reveal islands of stability or instability for certain hypothesised relative inclinations. Direct observations, from astrometric measurements, are restricted to systems with large astrometric signatures for each individual planet (§3.1). The 3-planet system *v* And, observed with HST-FGS, currently provides the most convincing example (§3.7), with the large relative inclination between planets c and d ($\Delta i_{cd} = 29.9 \pm 1^\circ$) lending preliminary support to the simulation results.

Statistics of multi-planet systems A difference between single and multiple planetary systems noted early on in the discovery of radial velocity planets is that the pile-up of hot Jupiters between 0.03–0.07 au, and the discontinuity at ~ 1 au observed in the distribution of single systems, are both absent from the more uniform distribution seen in multi-planet systems (Figure 2.37). Systems with multiple giant planets appear to lead to a suppressed occurrence of close-in planets. Both characteristics presumably reflect details of migration in the two types of system. Various other differences in the properties of single and multiple systems, in $M_p \sin i$, e , and M_* , the implications for migration models, and the possibilities of selection effects in the present samples, are considered further by Wright et al. (2009b).

Dynamical modeling In the basic kinematic (or ‘Keplerian’) fitting approach (§2.1.2), the total radial velocity signal due to n_p planets is assumed to result from the independent reflex motions due to each planet separately.

Many multi-planet systems may, however, show effects of gravitational planet–planet interactions, even over relatively short time scales. These can lead to observable evolution of the orbital parameters over periods of years, and to radial velocity variations of the star

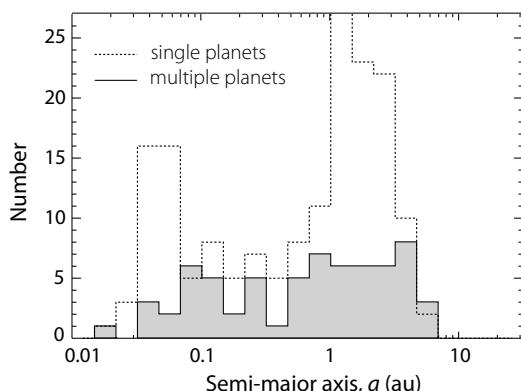


Figure 2.37: Distribution of semi-major axes for multiple planet systems (solid) and apparently single systems (dashed). The pile-up of hot Jupiters, and the jump in abundance beyond 1 au seen in the single planet systems are not evident in the multiple planet systems. From Wright et al. (2009b, Figure 9), by permission of IOP Publishing/AAS.

that differ substantially from those derived assuming the planets are executing independent Keplerian motions.

Dynamical analyses of these more complex systems are made using either of two approaches. N-body numerical integration methods are used both for dynamical (or ‘Newtonian’) orbit fitting (see also §2.1.2), and for studying the detailed planetary motions over relatively short time intervals in the future, including testing the validity of analytical results. N-body integration rests on the basic laws of gravity and motion. Results may yield ‘deprojected’ planetary masses (without the $\sin i$ ambiguity) and, in favourable cases based on stability arguments, relative orbital inclinations.

Analytical methods make use of the *disturbing function*, the difference in gravitational potential due to a star alone, and that due to a star and other perturbing planets. Analytical theory successfully describes two principal phenomena seen in multiple systems: secular (non-periodic) evolution, and resonances.

In *secular theory*, terms that depend on the planet’s mean motion, n , as well as other higher-order orbit terms, are ignored, and the theory describes the system’s *secular evolution*, essentially predicting how the shape of an average orbit evolves with time. In most two planet systems, secular theory predicts that their eccentricities oscillate, with an increasing eccentricity of one planet (corresponding to an increase of orbital angular momentum) being accompanied by decreasing eccentricity of the other through conservation of angular momentum. Secular theory also predicts the general behaviour of the difference in the two longitudes of pericentre, $\Delta\tilde{\omega}$: depending on initial conditions, $\Delta\tilde{\omega}$ may oscillate around 0 (aligned libration), 2π (anti-aligned libration), or circulate through 2π .

In *resonant theory*, terms that depend on the mean

motions are included, but only those related to the resonance under study are considered.

Radial velocity systems in resonance A number of multi-planet systems discovered by radial measurements (as well as many discovered through transit measurements) are found to be in orbital resonances (§10.8). Examples, and inferences that can be made from them, are considered in §2.12.4. From the complexity of the phenomenon it is apparent that very diverse systems, and complex behaviour in the long-term orbit integrations, can be expected. Table 2.7 lists the apparent period commensurabilities for well-characterised radial velocity detected multi-planetary systems.

Multiple systems and theories of formation As developed in detail in Chapter 10, the core accretion model provides a compelling scenario for giant planet formation. In brief, from an embryonic disk of dust and gas, dust particles collide and grow to form progressively larger planetary cores. If a significant amount of gas remains in the disk, a sufficiently massive core can gravitationally accrete more gas, rapidly growing in mass. Such giant planets are likely to form beyond the snow line, at around 3 au for solar-type stars, where ices can participate in the initial planetary cores. The fact that some 20% of known exoplanets orbit within 0.1 au, where little ice is available, leads to the hypothesis that short-period planets formed far out and migrated inwards to their present locations.

The discovery of numerous systems in or near mean motion resonance lends support to migration, with hydrodynamical and N-body simulations with externally applied damping suggesting that the 2:1 mean motion resonance (and others) appear naturally as the planets migrate inwards at different rates. Various mechanisms may be responsible for the wide distribution of observed eccentricities, some driving eccentricities to larger values while damping others to low values. An early era of strong planet–planet scattering producing large values may be followed by a damping phase as a result of subsequent interactions with remaining planetesimals. Evidence for these various mechanisms are detailed in §2.12.3 and §2.12.4.

2.12.2 Architectures and classification

The multi-planet systems discovered from radial velocity measurements span a range of architectures. Wright (2010) classified them into five broad and non-exclusive categories: systems with three or more giants (§2.12.3), systems with two giants in mean motion resonance (§2.12.4), other systems in which planet–planet interactions are non-negligible (§2.12.5), non-interacting systems (§2.12.6), and systems containing only lower mass planets with $M_p \sin i < 20 M_{\oplus}$ (§2.12.7).

This somewhat arbitrary classification nevertheless serves as a guide for the following discussions of various dynamical effects observed in multi-planet systems.

2.12.3 Systems with three or more giant planets

A few multiple radial velocity systems are considered here in some detail, in order to illustrate the kind of analysis and insight that is possible for these systems. A number of other such systems are now known.

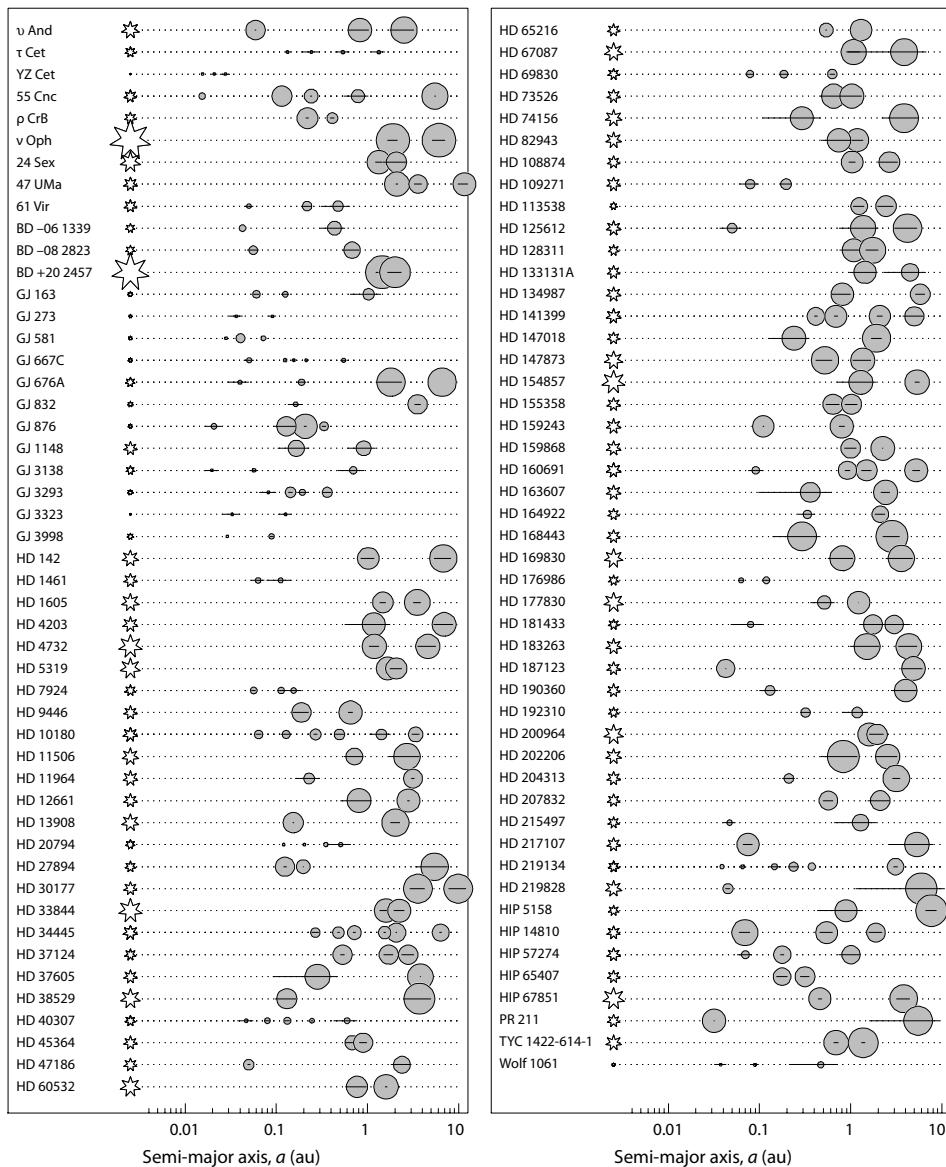


Figure 2.38: Radial velocity discovered systems of multiplicity 2 or higher, from the NASA Exoplanet Archive, 2017 December 31 (245 planets in 97 $n \geq 2$ systems). The host star mass is represented at left with size proportional to M_* (ranging from $0.13 M_\odot$ for YZ Cet to $3.0 M_\odot$ for v Oph). Each planet in the system is shown to the right, with size $\propto \log M_p$ (from $0.002 M_J$ for YZ Cet b to $27 M_J$ for v Oph c). Solid horizontal lines indicate maximum and minimum star–planet distance from their eccentricities. Highest multiplicities (6) are for HD 10180, HD 34445, and HD 219134. A similar plot for transiting systems is given in Figure 6.114.

v And The first system discovered to be multiple was the 3-planet system *v* And. In addition to the $0.6 M_J$ object in a low-eccentricity 4.6-d orbit originally detected by Butler et al. (1997), two more distant planets were identified from subsequent radial velocity observations, with $M_p \sin i$ of 2.0 and $4.1 M_J$, a of 0.82 and 2.5 au, and large e of 0.23 and 0.36 respectively (Butler et al., 1999). Evidence of a fourth outer planet from the Keplerian fits ($M_p \sin i = 1.06 M_J$) was reported by Curiel et al. (2011).

Various work has addressed its stability (Holman et al., 1997; Krymolowski & Mazeh, 1999; Laughlin & Adams, 1999; Lissauer, 1999; Rivera & Lissauer, 2000; Stepinski et al., 2000; Barnes & Quinn, 2001; Chiang et al., 2001; Jiang & Ip, 2001; Lissauer & Rivera, 2001; Chiang & Murray, 2002; Michtchenko & Malhotra, 2004; Michtchenko et al., 2006; Libert & Henrard, 2007; Rivera & Haghighipour, 2007; Libert & Tsiganis, 2009a; McArthur et al., 2010; Libert & Santos, 2013; Deitrick et al., 2015).

Table 2.7: Apparent period commensurabilities for well-characterised radial velocity detected multi-planetary systems (similar commensurabilities occur for transiting systems), based on Wright et al. (2011c, Table 3).

System	Planets	Period ratio	Reference
μ Ara	b, e	2:1	Pepe et al. (2007); Goździewski et al. (2007)
24 Sex	c, b	2:1	Johnson et al. (2011b); Wittenmyer et al. (2012b)
GJ 876	e, c, b	4:2:1	Marcy et al. (1998); Rivera & Lissauer (2001)
55 Cnc	c, b	3:1	Fischer et al. (2008); Zhou et al. (2008)
HD 10180	d, e	3:1	Lovis et al. (2011)
HD 10180	e, f	5:2	Lovis et al. (2011); Tuomi (2012)
HD 37124	c, d	2:1	Vogt et al. (2005); Wright et al. (2011c)
HD 45364	c, b	3:2	Correia et al. (2009); Rein et al. (2010); Correia-Otto et al. (2013)
HD 60532	c, b	3:1	Desort et al. (2008); Laskar & Correia (2009); Desort et al. (2009a); Alves et al. (2016)
HD 73526	c, b	2:1	Tinney et al. (2006); Sándor et al. (2007a); Zhang et al. (2010); Wittenmyer et al. (2014b)
HD 82943	b, c	2:1	Ji et al. (2003a); Mayor et al. (2004); Lee et al. (2006); Baluev & Beaugé (2014)
HD 108874	c, b	4:1	Vogt et al. (2005); Goździewski et al. (2006)
HD 128311	c, b	2:1	Vogt et al. (2005); Sándor & Kley (2006); Zhang et al. (2010)
HD 200964	c, b	4:3	Johnson et al. (2011b); Wittenmyer et al. (2012b); Tadeu dos Santos et al. (2015)
HD 202206	c, b	5:1	Correia et al. (2005); Goździewski et al. (2006)

Dynamical insights The innermost planet, with $e \sim 0$, significantly exceeds the minimum stability requirement given by the Hill radius criterion, suggesting little interaction with the outer two companions. For the two planets c and d in contrast, the system stability depends strongly on the planet masses, and hence their relative orbital inclinations, with certain combinations implying chaotic or unstable orbits.

Mazeh et al. (1999) derived a mass for planet d of $10.1 \pm 4.7 M_J$ using Hipparcos astrometry, implying $i = 156^\circ$, and masses of the inner two planets of 1.8 ± 0.8 and $4.9 \pm 2.3 M_J$ if the orbits are co-planar (see also Reffert & Quirrenbach 2011). Astrometry from HST has established the mutual inclination of planets c and d as $\sim 30^\circ$ (discussed further in §3.7).

Goździewski et al. (2001) used the fast orbit indicator MEGNO to derive $\langle \mathcal{J} \rangle(t)$ as a function of the inclination of planet c, and the relative inclination of planets c and d (Equation 3.17). Quasi-periodic and chaotic zones are evident in various regions of the i_c, i_{c-d} parameter space (Figure 10.32).

From numerical integration using Mercury, Ford et al. (2005) found that for a co-planar configuration the system becomes dynamically unstable when $\sin i < 0.5$ while, if the coplanarity condition is relaxed, dynamical instability resulted from relative inclinations $\gtrsim 40^\circ$. They found that the long-term secular evolution of the orbits (Figure 10.33) was best modeled as resulting from an impulsive perturbation to planet d, in which a sudden change in its eccentricity was produced by a close encounter with another planet. After a brief period of chaotic evolution lasting $\sim 10^3$ yr, the perturbing planet would have been ejected, leaving the remaining planets (c and d) in a stable configuration resembling that observed today.

With the discovery of the fourth planet, v And e, numerical integration of the orbital solution for all four planets (Currie et al., 2011) showed that the system is stable for at least 10 Myr, with the orbit of the fourth planet coinciding with an island of stability reported by Rivera & Haghighipour (2007). All four planets have very strong mutual interactions, with planets b and c in apsidal alignment, and the orbit of planet e being close to an external 3:1 resonance with planet c.

With no observational constraints on the orbital plane of planet b, Deitrick et al. (2015) used N-body simulations to search for stable three-planet configurations that are consistent with the combined radial velocity and astrometric solu-

tion, finding that the orbit of planet b must lie near the invariant plane of planets c and d, but can be either prograde or retrograde. Their solutions predict that the mass of planet b is in the range $2 - 9 M_J$, with an inclination angle from the sky plane of $< 25^\circ$. Combined with brightness variations in the combined star/planet light curve (or phase curve) determined by Harrington et al. (2006) and Crossfield et al. (2010), their results imply that $R_b = 1.8 R_J$, relatively large for a planet of its age. But with $e_b > 0.1$ in several of their stable solutions, generating upward of 10^{19} W in the interior of the planet via tidal dissipation (as also inferred for HD 209458 b by Ibgui & Burrows 2009), the radius could be inflated by an amount consistent with the phase curve observations.

HD 37124 HD 37124 is an example of dynamical stability modeling leading to the discovery of additional planets. Vogt et al. (2000) detected a 150-d Jupiter-mass planet from Keck-HIRES observations. Butler et al. (2003) identified a second planet with $P \sim 6$ -yr. The two-planet solution was shown to be unstable by Goździewski (2003a), with further Keck data identifying a third planet (Vogt et al., 2005). $\{M_p \sin i / (M_J), P \text{ (d)}, e\}$ for planets b, c, d are respectively $\{0.638, 154, 0.055\}$, $\{0.697, 2300, 0.2\}$, $\{0.624, 844, 0.14\}$. Further dynamical studies have been undertaken (Goździewski et al., 2006, 2008a; Wright et al., 2011c; Elser et al., 2013).

HD 74156 The two planet system was originally reported by Naef et al. (2004) from ELODIE data. A third planet was detected from a combination of CORALIE, ELODIE and HET-HRS data by Bean et al. (2008b), with $M_p \sin i = 0.40 M_J$, $P = 347$ d, and $e = 0.25 \pm 0.11$.

Predicting planet d Raymond & Barnes (2005) had predicted the parameter space that such a third planet could have occupied (also in three other systems), by identifying stability regions for massless test particles defined by the previously known planets. For HD 74156 they had shown that a Saturn mass planet could exist within a broad region of $a = 0.9 - 1.4$ au and $e \leq 0.15$. Barnes et al. (2008b) gave an alternative slightly improved orbit, and suggested that the confirmed prediction of

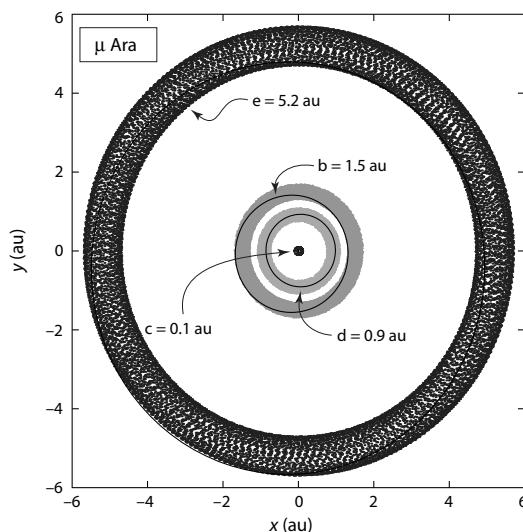


Figure 2.39: Face-on view of μ Ara (HD 160691), illustrating orbital evolution of the four planets over 1 Myr. Present orbits are shown as solid lines, while dots (indistinguishable here for the inner three planets) correspond to positions every 50 yr. The values of a are constant, while the eccentricities undergo small variations in the range 0.09–0.13, 0.16–0.21, 0–0.19, and 0.08–0.11 for planets b–e respectively. From Pepe et al. (2007, Figure 6), reproduced with permission © ESO.

the third planet points to planet formation being an efficient process, with systems typically containing many planets.

μ Ara (HD 160691) The properties of this four-planet system were progressively revealed by AAT–UCLES (Butler et al., 2001; McCarthy et al., 2004) and HARPS (Santos et al., 2004a; Pepe et al., 2007) data. The fourth planet was also tentatively announced by Goździewski et al. (2007). Simulations demonstrating the system stability are shown in Figure 2.39. The semi-major axes remain almost constant, while the eccentricities of all four planets undergo small variations. Various other dynamical analyses have also been made (Bois et al., 2003; Goździewski et al., 2003, 2005a; Short et al., 2008).

55 Cnc The 55 Cnc system was the first known planetary system comprising five planets (Table 2.8). The first, 55 Cnc b, was reported by Butler et al. (1997), and the second (and possibly a third) also from Lick data by Marcy et al. (2002). 55 Cnc c was confirmed by McArthur et al. (2004) based on Lick, ELODIE and HET Doppler measurements and HST–FGS astrometry. They also announced 55 Cnc e with $P = 2.8$ d.

Dynamical insights Fischer et al. (2008) used 18 years of Lick and Keck data to confirm the four proposed planets, and to identify a fifth, 55 Cnc f, moving in the large empty zone between two other planets. All five reside in low-eccentricity orbits, four having $e < 0.1$. They adopted a multi-planet Keplerian fitting procedure (§2.1.2), using a numerical N-body simulation to show that the system is dynamically stable.

Table 2.8: The planets of 55 Cnc, ordered by increasing semi-major axis. Data are from Fischer et al. (2008), and do not include possible updates suggested by Nelson et al. (2014b).

Planet	$M_p \sin i$ (M_J)	a (au)	P (d)	e
e	0.02	0.038	2.80	0.26
b	0.83	0.115	14.65	0.02
c	0.17	0.240	44.38	0.05
f	0.15	0.789	260.67	0.00
d	3.90	5.888	5371.82	0.06

55 Cnc has some basic structural attributes found in our solar system: the orbits are rather circular and nearly co-planar, and a dominant gas giant lies at a distance of about 6 au. Although planets b and c have a period ratio of 3.027:1.000, hinting at a possible mean motion resonance (Marcy et al., 2002; Ji et al., 2003b; Voyatzis & Hadjidemetriou, 2006), the 3:1 mean motion resonance was excluded by the N-body model of Fischer et al. (2008), as later confirmed by Nelson et al. (2014b), as none of the relevant resonant arguments librate.

Other dynamical simulations were made by Gayon et al. (2008) using MEGNO, who found that about 15% of the systems resulting from the nominal orbital elements of the system are highly chaotic. Raymond et al. (2008b) evaluated the stability of the large region between planets f and d using N-body integrations that included an additional, yet-to-be-discovered planet g with a radial velocity amplitude of $\sim 5 \text{ m s}^{-1}$, i.e. $M_p \approx 0.5 - 1.2 M_{\text{Saturn}}$. They found a large stable zone extending from 0.9–3.8 au with $e < 0.4$, which could contain 2–3 additional planets each of $M_p \sim 50 M_{\oplus}$. Any planets exterior to planet d must reside beyond 10 au.

N-body simulations using 1418 high-precision radial velocity observations from four observatories (Lick, Keck, HET, and HJS), along with transit times (and durations) for the innermost planet, 55 Cnc e, was made by Nelson et al. (2014b). Amongst their findings were that dynamical stability dictates that the orbital plane of planet e must be aligned to within 60° of the orbital plane of the outer planets, assumed to be co-planar.

2.12.4 Systems in mean motion resonance

The 2:1 resonance system GJ 876 GJ 876 was the first known M-dwarf host. The orbit elements are particularly well determined due to their short periods, and orbit modeling shows that the two most massive planets b and c are locked in a 2:1 resonance, the first to be discovered amongst exoplanet systems.

2:1 mean motion resonance Marcy et al. (1998) using Keck, and Delfosse et al. (1998) using ELODIE/CORALIE, both reported a 61-d, $2M_J$ planet. From further Keck observations, Marcy et al. (2001a) showed that the radial velocity signal was actually the combination of two planets in a 2:1 mean motion resonance, with the inner having $P = 30.1$ d and $0.6M_J$. In the original data, the inner companion was indistinguishable from an additional orbital eccentricity of the outer. The axes of the two orbits are nearly aligned, and the orbital elements evolve significantly with time. A third, non-interacting, $7.5M_{\oplus}$ mass planet in a 1.9 d orbit was detected by Rivera et al. (2005).

Laughlin & Chambers (2001), Rivera & Lissauer (2001), and Nauenberg (2002a) independently developed fitting procedures to account for the gravitational interactions between the

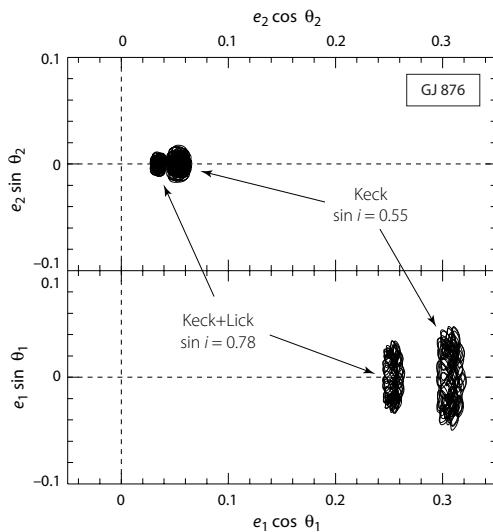


Figure 2.40: Small-amplitude librations, about 0° , of the two 2:1 mean motion resonance variables for GJ 876. Trajectories for 3100 d (the average pericentre precession period for both planets) are shown as $e_j \sin j$ versus $e_j \cos j$ ($j = 1, 2$) for the best-fit solutions to the Keck data alone, and for the combined Keck and Lick data, both from Laughlin & Chambers (2001). Values of $\sin i$ are obtained in the best-fit solution. From Lee & Peale (2002, Figure 1), by permission of IOP Publishing/AAS.

two outer planets. Rivera & Lissauer (2001) found that most of their solutions are stable for at least 10^8 yr, while test particles orbiting between the two planets are lost in $\lesssim 300$ yr.

Lee & Peale (2002) showed that both mean motion resonance variables, $\phi_1 = 2\lambda_2 - \lambda_1 - \dot{\omega}_1$ and $\phi_2 = 2\lambda_2 - \lambda_1 - \dot{\omega}_2$ (cf. Equation 10.33), perform very small amplitude librations around 0° (Figure 2.40). Their simultaneous libration about 0° implies that the variable $\phi_3 = \phi_1 - \phi_2 = \dot{\omega}_1 - \dot{\omega}_2$ also librates around 0° , and the planets are consequently in three resonances at the 2:1 mean motion commensurability. The small libration of ϕ_3 means that the lines of apsides of the two orbits are therefore nearly aligned, and conjunctions of the two planets occur very close to their longitudes of pericentre (Beaugé & Michtchenko, 2003).¹² The existence of the mean motion resonance is taken as confirmation that the two orbits are essentially co-planar, with the deep resonance (i.e. with small libration amplitudes) implying that the system is stable indefinitely.

Laughlin et al. (2005a) incorporated additional data obtained over 16 yr, and confirmed the small libration amplitudes (Figure 2.41). Such a configuration can be explained by a slow differential migration resulting from an interaction of the planets with a protoplanetary disk (Snellgrove et al., 2001; Lee & Peale, 2002; Kley et al., 2005; Beaugé et al., 2006; Crida et al., 2010b). They also showed that configurations with modest mutual inclinations are possible, which has an interesting corollary: for non-coplanar configurations the line of nodes of the inner planet precesses at about -4° yr^{-1} .

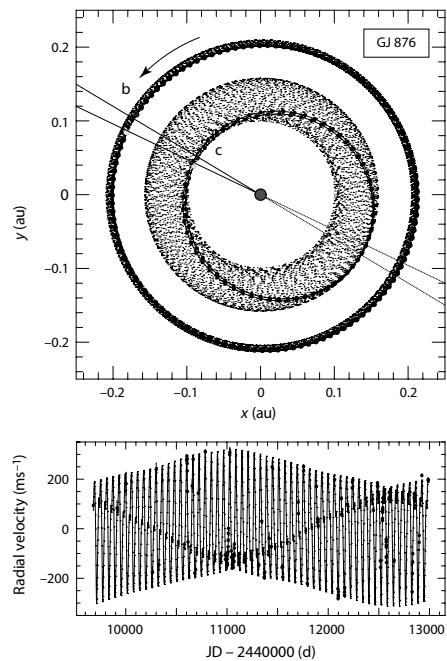


Figure 2.41: The 2:1 mean motion resonance in GJ 876. Top: the two clouds of dots show the planet positions every 0.5 d over 10 yr, illustrating precession in the line of apsides of $-41^\circ \text{ yr}^{-1}$. Connected circles show the planet positions every 0.5 d for 60 d. The two lines radiating from the central star mark the longitudes of pericentre, ω_b and ω_c , at JD 2 449 710. They oscillate about alignment with a libration amplitude of 34° . Bottom: the stellar reflex velocities (circles), and the solution from the self-consistent, co-planar, 3-body integration. From Laughlin et al. (2005a, Figures 1–2), by permission of IOP Publishing/AAS.

The inner planet may therefore be observed to transit the host star at some time in the relatively near future, i.e. when either the ascending or descending node precesses through the line-of-sight, even though it is not transiting at the present time. This general phenomenon of ‘transitional transits’ is described further in Section 6.19.8.

Laplace resonance A fourth planet, GJ 876 e, was discovered from continued radial velocity monitoring by Rivera et al. (2010b). N-body fits show that the four-planet system has an invariable plane with an inclination of $59^\circ 5$, and is stable for more than 1 Gyr. Their model places the fourth planet in a three-body ‘Laplace’ resonance (§10.8.3) with the two giant planets ($P_c = 30.4$ d, $P_b = 61.1$ d, $P_e = 126.6$ d). Unlike the case of the Galilean satellites of Jupiter (Figure 2.42), the three planets come close to a triple conjunction once per orbit of planet e. The critical argument for the Laplace resonance, $\phi = \lambda_c - 3\lambda_b + 2\lambda_e$, librates with an amplitude of $40^\circ \pm 13^\circ$.

The 2:1 resonance system HD 128311 The two-planet system HD 128311 offers an interesting insight into planetary migration along similar lines. The first planet was reported by Butler et al. (2003), and the second by Vogt et al. (2005). A Trojan 1:1 solution was suggested by Goździewski & Konacki (2006), while the most recent orbit solution is given by Wittenmyer et al. (2009).

¹²The nature of the resonance contrasts with the case of the Io-Europa system, where the lines of apsides are anti-aligned, and where conjunctions therefore occur when Io is near its pericentre, and Europa is near its apocentre.

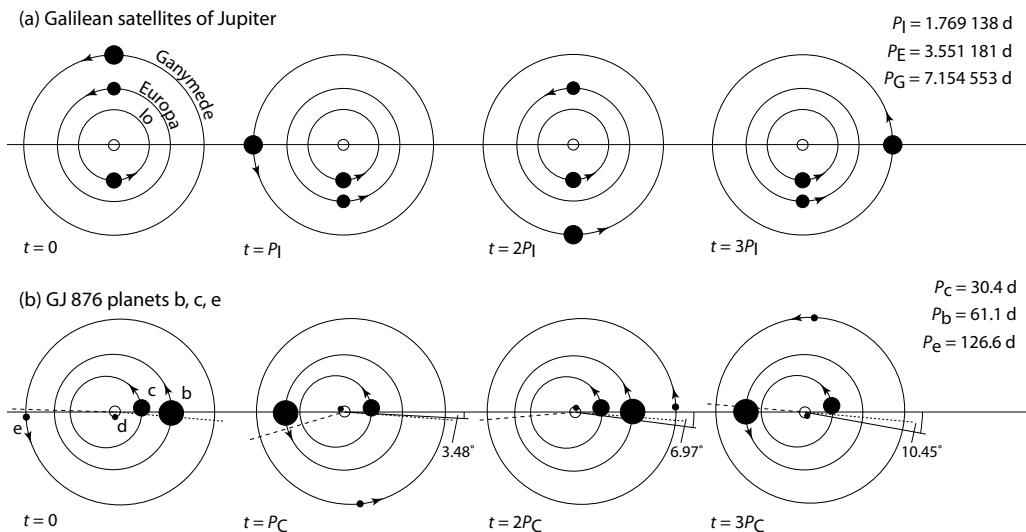


Figure 2.42: The first two known examples of Laplace resonance, in which three (or more) orbiting bodies have a simple integer ratio between their orbital periods: (a) Jupiter's inner satellites Ganymede, Europa, and Io are in a 1:2:4 Laplace resonance, with Ganymede completing 1 orbit in the time that Europa makes 2, and Io makes 4. They never experience triple conjunctions. Their mean motions (Equation 2.8) are related by $n_I - 3n_E + 2n_G = 0$, satisfied to nine significant figures (Peale, 1976); (b) planets c, b, e of the M dwarf host star GJ 876 are an exoplanet triplet in a Laplace resonance. The reference frame rotates at the mean orbital precession of planet b, $-10^\circ.45$ over 90 d (solid lines). The apsidal line of b coincides with the x-axis, while those for planets c and e are shown with short-dash and long-dash lines respectively. In each system, the orbits and object radii (assuming $R \propto M^{1/3}$) are shown to scale. The lower figure is adapted from Rivera et al. (2010b, Figure 7).

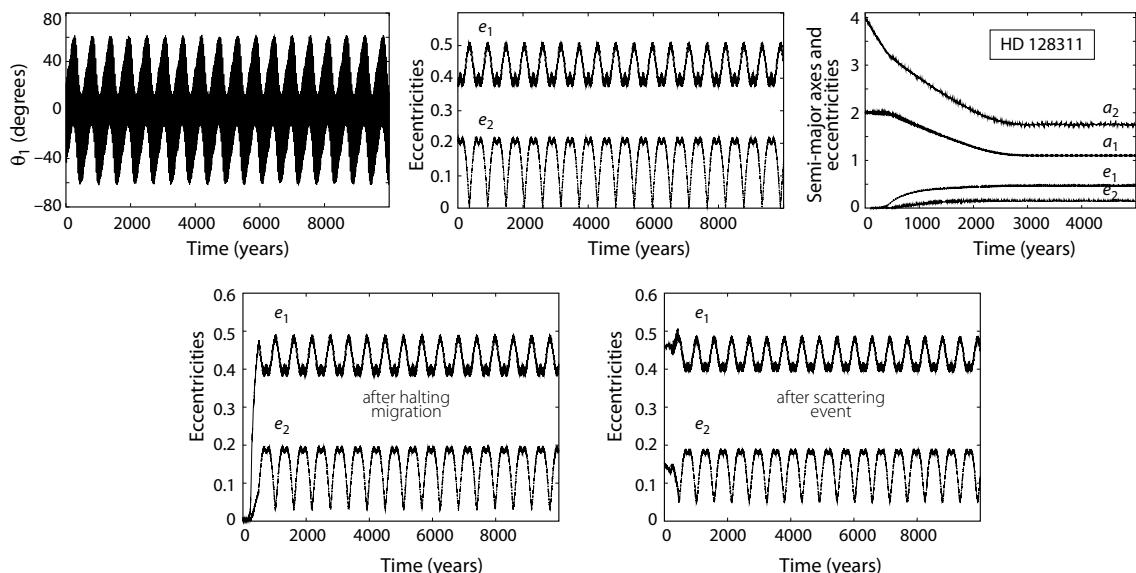


Figure 2.43: Resonant evolution of HD 128311. Evolution of the resonant angle ϕ_1 (top left) and eccentricities e_1 and e_2 (top middle) obtained by numerical integration of the orbital parameters of Vogt et al. (2005). Top right: predicted behaviour of the semi-major axes and eccentricities during adiabatic migration with an e-folding time of 2×10^3 yr; for the simulations, the ongoing migration is progressively stopped between $2 - 3 \times 10^3$ yr. Bottom left: evolution of the planetary eccentricities, both originally in circular orbits, following a sudden halting in the migration of the outer planet. Bottom right: evolution of the eccentricities following a scattering event with a low-mass planet migrating outwards. From Sándor & Kley (2006, Figures 1, 2, 3, 5), reproduced with permission © ESO.

Solar system 1:1 resonances and the Trojans: In the circular restricted three-body problem, three bodies move in circular co-planar orbits, with the mass of the third being negligible. In 1772, Lagrange proved the existence of five equilibrium points where the third particle has zero velocity and zero acceleration in the rotating frame. These are the three colinear Lagrangian equilibrium points L_1, L_2, L_3 , and two (leading and trailing) triangular equilibrium points L_4, L_5 . Although L_1-L_3 are unstable, special starting conditions of the third particle, in position and velocity, can nevertheless result in semi-stable periodic orbits in their vicinity. This is exploited by artificial satellites such as SOHO which observes the Sun, interior to Earth's orbit, from L_1 (Domingo et al., 1995), and deep-space missions which observe 'outward' from L_2 (e.g., WMAP, Herschel, Planck, and Gaia).

Stable orbits in the vicinity of L_4 or L_5 were first demonstrated by Routh (1875) for the special case of circular orbits, and by Danby (1964a,b) for elliptical orbits. They comprise a short-term component of motion with a period close to the orbital period, and a longer term periodic motion about the equilibrium point referred to as libration. Resulting orbits can be described as a short-period epicyclic motion around a long-period motion of the epicentre, the relative contributions of the two components determined by the starting conditions. The resulting elongated orbits around L_4 or L_5 are referred to as *tadpole orbits*. For an increased radial separation from L_4 or L_5 , the resulting *horseshoe orbits* can encompass both L_4 and L_5 .

Examples of solar system bodies in such stable 1:1 resonances are referred to as *Trojans*, and they include both asteroids and satellites. Librating mostly in tadpole orbits around the Sun–Jupiter triangular equilibrium points is the leading group around L_4 (the 'Greeks'), including the first known (588) Achilles discovered by Max Wolf in 1906, and the trailing group around L_5 (the 'Trojans'). Typical libration amplitudes are 15–20° (Shoemaker et al., 1989).

Other Trojan asteroids orbit the Sun–Mars system, with the first (5261) Eureka, discovered in 1990, librating about L_5 (Mikkola et al., 1994). Asteroid (3753) Cruithne is in a horseshoe orbit around the Sun–Earth system (Wiegert et al., 1997).

The dynamics of planet–satellite systems around their triangular equilibrium points are identical, and result in the *co-orbital* or *Trojan satellites*. Telesto and Calypso librate around the Saturn–Tethys system, and Helene and Polydeuces around Saturn–Dione. In the Janus–Epimetheus Saturnian system, Epimetheus moves in a horseshoe orbit, exchanging altitude with Janus as they approach in a 4 yr repeating cycle (Dermott & Murray, 1981a,b), and resulting in forced rotational libration (Tiscareno et al., 2009). Solitary density waves are launched when they exchange radial position (Rehnberg et al., 2016).

Dynamical insights Sándor & Kley (2006) performed a 3-body integration of the (Newtonian) orbital solution of Vogt et al. (2005), and examined the evolution of the difference in longitudes of pericentre $\Delta\tilde{\omega} = \tilde{\omega}_2 - \tilde{\omega}_1$, and the two appropriate resonant angles ϕ_1, ϕ_2 (Equation 10.33 corresponding to $\tilde{\omega}_1, \tilde{\omega}_2$ respectively). Of these, only ϕ_1 librates around 0° with an amplitude of ~60°, while ϕ_2 and $\Delta\tilde{\omega}$ rotate. At the same time, the eccentricities show large oscillations (Figure 2.43).

This behaviour contrasts with the apsidal corotation predicted by adiabatic migration (Figure 2.43, top right). Instead, Sándor & Kley (2006) considered that the system's state arises from a strong scattering event in the past. Assuming that the two planets were once locked in a 2:1 resonance with apsidal corotation as a result of inward migration, they then studied two distinct perturbation scenarios which might have disrupted the corotation (Figure 2.43, bottom). The first scenario invokes a sudden halting of migration. This might arise if the outer planet reaches some disk discontinuity with an empty region inside, possibly due to stellar photoevaporation (§10.3.6).

In the second scenario, an encounter with a low-mass $10M_{\oplus}$ planet approaching from inside or outside (similar to that proposed by Ford et al. 2005 for v And) could also break the apsidal libration. The smaller planet would be ejected from the system, or perturbed into a larger orbit, in the process. Particularly in the case of a small planet approaching from the inside, the apsidal corotation of the giant planets is indeed broken. Consistent with the behaviour predicted from long-term orbit integration, ϕ_2 and $\Delta\tilde{\omega}$ then circulate, while the giant planets remain in the 2:1 resonance, with ϕ_1 still librating around 0°.

The 3:1 antisymmetric resonance system HD 60532

Amongst suggested systems in a 3:1 mean motion resonance are the two giants orbiting HD 60532 (Desort et al., 2008; Laskar & Correia, 2009; Sándor & Kley, 2010). The two orbital solutions found by (Laskar & Correia, 2009) have very different inclinations, $i = 20^\circ$ or $i = 90^\circ$, with planet masses differing by a factor $1/\sin i \sim 3$.

Dynamical insights Sándor & Kley (2010) carried out detailed 2d hydrodynamical simulations of appropriate thin disks with an embedded pair of massive planets, studying the effects of migration and resonant capture using N-body simulations.

For resonant capture, planet pairs must undergo convergent migration, such that if the inner planet migrates inwards at the same speed, or faster than the outer one, then no resonant capture occurs. Resonant capture through migration therefore depends on the inner planet opening a suitably large gap in the inner disk to slow down its inward migration rate. In their simulations, Sándor & Kley (2010) found that capture into the observed 3:1 resonance takes place only for higher planetary masses, thus favouring orbital solutions having the smaller inclination of $i = 20^\circ$ (Figure 2.44).

They also found that the inner disk, between the inner planet and the star, plays a key role in determining its final configuration. Specifically, fast inward migration of the outer planet may result in it crossing the 3:1 mean motion resonance without capture, ending the migration in the more robust 2:1 mean motion resonance. In the case of HD 60532, the damping effect of the disk on the inner planet's eccentricity is responsible for an antisymmetric (antialigned) resonance configuration in which $\tilde{\omega}_2 - \tilde{\omega}_1$ oscillates around 180°, rather than the 0° in the case of aligned pericentres.

Other mean motion resonances Other systems believed to be in a 2:1 mean motion resonance include HD 82943 (Goździewski & Maciejewski, 2001; Ji et al., 2003a; Mayor et al., 2004; Ferraz-Mello et al., 2005; Lee et al., 2006); HD 73526 (Tinney et al., 2003b, 2006; Sándor et al., 2007a); and μ Ara d–b (Pepe et al., 2007).

Other possible 3:1 resonances include 14 Her (Goździewski et al., 2006, 2008b), and the more contested case of 55 Cnc b–c (Marcy et al., 2002; Ji et al., 2003b; Fischer et al., 2008).

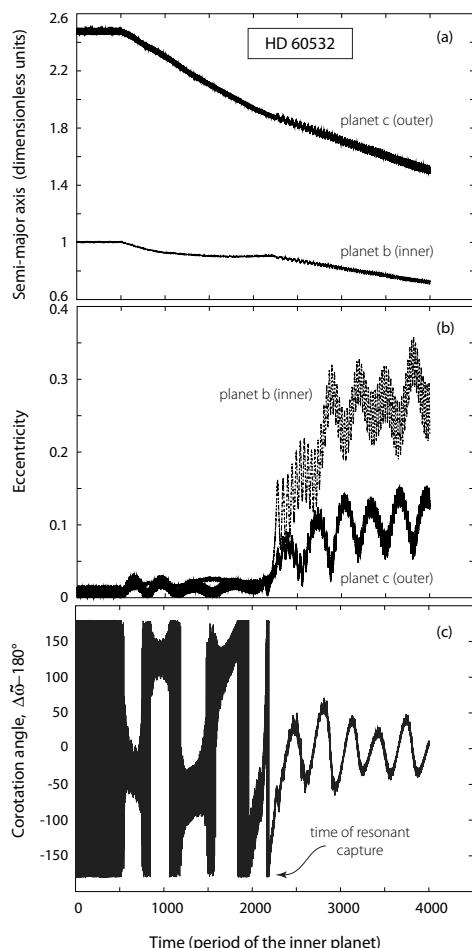


Figure 2.44: Dynamical simulation of the two giant planets orbiting HD 60532, embedded in a protoplanetary disk, showing (a) semi-major axes, (b) eccentricities, and (c) the evolution of the resonant angle $\Delta\tilde{\omega} = \tilde{\omega}_2 - \tilde{\omega}_1$. During the first 500 orbits of the inner planet, the planets are ‘fixed’ to obtain a steady state in the disk. Capture into the 3:1 mean motion resonance occurs after ~ 2300 periods of the inner planet, subsequent to which the orbits exhibit (antialigned) apsidal corotation. The planetary masses ($3.15M_J$ and $7.46M_J$) correspond to a (co-planar) system inclination of $i = 20^\circ$. From Sándor & Kley (2010, Figures 4 and 5b), reproduced with permission © ESO.

A number of other possible resonances have been identified: the 4:1 resonance of HD 108874 (Butler et al., 2003; Goździewski et al., 2006) and perhaps 14 Her (Wittenmyer et al., 2007); the 5:1 resonance of HD 202206 (Udry et al., 2002; Correia et al., 2005; Goździewski et al., 2006); a possible 3:2 resonance of HD 45364 (Correia et al., 2009; Rein et al., 2010); a possible 5:2 resonance of HD 37124 (Goździewski et al., 2006); a possible 6:1 or 11:2 resonance of HD 12661 (Goździewski, 2003b; Lee & Peale, 2003; Rodríguez & Gallardo, 2005; Zhang & Zhou, 2006; Veras & Ford, 2009); and a possible 5:2 or 7:3 resonance in 47 UMa (Laughlin et al., 2002).

Proximity to resonance Proximity to a resonance may still affect the secular motion, as in the case of the near 5:2 resonance between Jupiter and Saturn, Laplace’s ‘great inequality’ (e.g. Varadi et al., 1999; Michtchenko & Ferraz-Mello, 2001a). Such proximity has been suggested for HD 12661 and ν And (Libert & Henrard, 2007).

Inclination resonance In the absence of observational data on the mutual inclination of exoplanet orbits, most studies of mean motion resonances have focused on co-planar configurations. The extension to non-coplanar orbits is relevant for an understanding of eccentricity excitation and migration.

Thommes & Lissauer (2003) found that, subsequent to a 2:1 eccentric resonance capture, a subsequent capture into a 4:2 *inclination resonance* is possible. This is the lowest-order inclination resonance at the 2:1 commensurability (since the mutual inclination appears first as $(\Delta i)^2$ in the series expansion of the disturbing function), and is characterised by the mean motion resonance variables (cf. Equation 10.36)

$$\begin{aligned}\phi_{11} &= 2\lambda_1 - 4\lambda_2 + 2\Omega_1 \quad \text{and} \\ \phi_{22} &= 2\lambda_1 - 4\lambda_2 + 2\Omega_2,\end{aligned}\quad (2.56)$$

where Ω_j are the longitudes of the ascending nodes. The properties and subsequent evolution of this resonance are detailed by Lee & Thommes (2009).

Capture into higher-order resonances (such as the 3:1, 4:1, and 5:1) can also result in the excitation of inclinations, with mutual inclinations reaching $\Delta i \sim 20 - 70^\circ$ in the simulations of Libert & Tsiganis (2009b), at least for low-mass inner planets and for one or both planets developing eccentricities $e \gtrsim 0.4$. Amongst systems believed to be in higher-order resonance, HD 60532 (3:1), HD 108874 (4:1) and HD 102272 (4:1) do not satisfy these requirements, and are likely to reside in co-planar resonances, as verified by numerical simulations.

Lidov–Kozai oscillations Large mutual inclinations can be generated by Lidov–Kozai oscillations (§10.10.6), as well as through planet–planet scattering (§10.10.4) and planetesimal-driven migration (§10.10.3). Whether any of the known systems are in a stable Lidov–Kozai resonant state, with mutual orbital inclinations of order $\Delta i \sim 40 - 60^\circ$, was investigated parametrically by Libert & Tsiganis (2009a). They found that four of the systems studied (ν And, HD 12661, HD 74156, and HD 169830) could in principle be in Kozai resonance, provided that their mutual inclination is at least 45° . Direct astrometric determination of the mutual inclinations in ν And by HST–FGS (§3.7, McArthur et al., 2010) gave $\Delta i_{cd} = 29.9 \pm 1^\circ$.

The 1:1 resonance A number of objects in the solar system move in 1:1 resonance orbits. They are examples of Lagrange’s celebrated solution to the restricted three-body problem, and generally satisfy the condition that

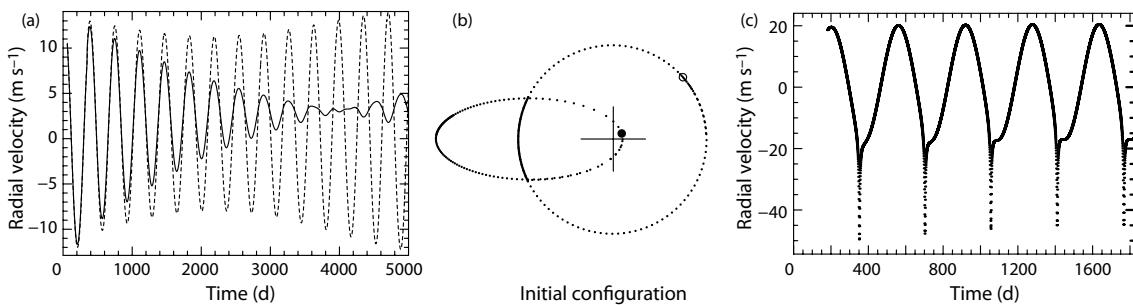


Figure 2.45: Theoretical 1:1 resonance orbits for equal mass planets: (a) synthetic radial velocity variations predicted for the tadpole-type (dotted) and horseshoe-type (solid) orbits; (b) initial configuration for the 1:1 eccentric resonance; solid lines indicate parts of the respective orbits swept out over 15% of the orbital period; (c) associated radial velocity variations for the case of eccentric resonance. From Laughlin & Chambers (2002, Figures 3, 4, 6), by permission of IOP Publishing/AAS.

the mass of the smallest is negligible compared with that of the other two (see box, page 74).

For the more general three-body problem, and in particular for two equal mass planets which share a time-averaged orbital period, Laughlin & Chambers (2002) identified a wide parameter space in which stable coorbital configurations exist (Figure 2.45).

Possible orbits These stable coorbital configurations fall into rather distinct regimes. In the first, the star and both planets participate in tadpole-like librations around the vertices of an equilateral triangle. The dynamics resembles that of Jupiter's Trojan asteroids, and the configuration is stable for mass ratios $2M_p/(2M_p + M_*) < 0.03812$. In the second regime, for larger perturbations, the planets execute symmetrical horseshoe-type orbits similar to those of Janus and Epimetheus. Stability analysis indicates that a pair of Saturn-mass planets could survive in this resonance over long periods (Figure 2.45a).

In practice, recognising such systems from the radial velocity data might be problematic, since the periodogram signal can be indistinguishable from that of a single planet in an eccentric orbit (Laughlin & Chambers, 2002), or from that of a system in 2:1 mean motion resonance (Goździewski & Konacki, 2006).

Their third configuration, described as *eccentric resonance*, is qualitatively distinct from Lagrange's solution (Figure 2.45b). One orbit starts as highly eccentric, the other is more nearly circular, the pericentres are aligned, and conjunctions occur near pericentre; the pair of planets then exchange angular momentum and eccentricity (in the case studied, with a periodicity of ~ 800 yr). With appropriate initial conditions, they can avoid close encounters indefinitely. In this regime, the radial velocity variations are distinctive, perhaps recognisable by a single deviating measurement superposed on an otherwise sinusoidal pattern (Figure 2.45c). A similar configuration was found independently by Nauenberg (2002b). These eccentric resonances might arise from the same type of planet–planet interactions responsible for non-resonant high eccentricity systems.

The theoretical existence and properties of the various 1:1 resonance configurations (Hadjidemetriou et al., 2009) has led some investigations to consider them as alternative models for certain specific systems (e.g. Érdi et al., 2007), in particular in cases where long-term orbit integration has identified dominant instabilities.

Goździewski & Konacki (2006) examined the long-term stability of HD 128311 (§2.12.4), searching for alternative 2:1 res-

onances using a genetic algorithm, and examining their long-term stability using MEGNO. They also searched for 1:1 resonances consistent with the observed radial velocity variations, and identified extended zones of stability in which such a system would survive (Figure 2.46). They used the term 'Trojan planets' for all 1:1 mean motion resonances, including the eccentric resonance configurations identified by Laughlin & Chambers (2002), i.e. not only for the Lagrange-type solutions, but also for those with similar semi-major axes, but possibly with large relative inclinations and time-varying eccentricities.

In the multiple planet system μ Ara studied by Pepe et al. (2007) one of their solutions, located using a genetic algorithm search, gave an inner planet at 9.64 d, an outer planet at 2741 d, and with an intermediate pair of Trojan planets in a 1:1 mean motion resonance. The solution was discarded as a result of its high dynamical instability; orbit integration indicating that it would be disrupted in less than 100 years.

Schwarz et al. (2007a,b,c) demonstrated the stability of terrestrial Trojan planets in resonance with a giant planet in the habitable zone for a number of systems.

The formation of a hypothetical terrestrial-type body at the Lagrange points of a giant exoplanet has been studied using an N-body code by Beaugé et al. (2007). Although such planets may form *in situ*, the accretion process is inefficient, and the mass of the final planet never exceeds $\sim 0.6M_\oplus$.

Retrograde resonances Systems with multiple planets are generally considered to revolve in the same direction around their respective system barycentres, and most radial velocity orbit elements are derived assuming *prograde* or *regular* orbits, i.e. all orbiting in the same direction. Such behaviour is expected, at least for unperturbed states, according to planet formation theories.

The theoretical existence and dynamical stability of *retrograde* mean motion resonances, i.e. with one planet counter-revolving with respect to the other in a two-planet system, or to all other planets in a multiple planet system, has been investigated and demonstrated (Gayon & Bois, 2008a,b; Gayon et al., 2009).

Formation of retrograde resonances A planet might conceivably be found in a retrograde mean motion resonance either as a result of capture of a pre-existing free-floating planet directly into a retrograde orbit, or through (violent) dynamical

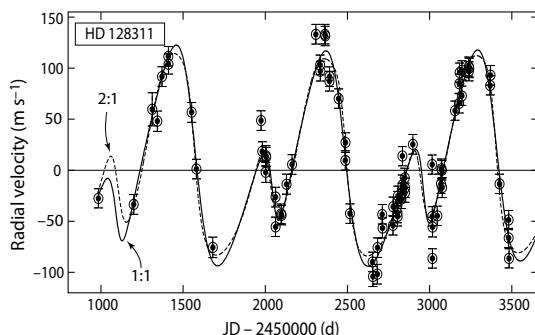


Figure 2.46: The two-planet system HD 128311, showing radial velocity curves for the best-fit solutions corresponding to the 2:1 (dashed) and 1:1 (solid) mean motion resonances. Both curves give an rms of 15 m s^{-1} . Error bars include the stellar jitter of 9 m s^{-1} . From Goździewski & Konacki (2006, Figure 4), by permission of IOP Publishing/AAS.

evolution as a result of planet–planet interactions. The sling-shot effects seen in the simulations of Nagasawa et al. (2008) show that close-in planets, some in retrograde motion as a result of Lidov–Kozai resonances, may indeed result.

In the solar system, a small number of asteroids, a number of asteroid-size planetary moons, and Triton (the largest of Neptune’s moons) have retrograde orbits.

Gayon & Bois (2008a) used the genetic algorithm PIKAIA and the fast-chaos indicator MEGNO to search for stable 2:1 retrograde resonant configurations for HD 73526, consistent with the published initial conditions. They found prominent stability islands, which occur when the apsidal longitudes of the two orbits precess at the same average rate, albeit with complex structure, and even though they do not precess in the same direction. The resulting radial velocity curve is very similar to that given by the prograde solutions of Tinney et al. (2006) and Sándor et al. (2007a).

Similar counter-revolving solutions have been investigated by Gayon-Markt & Bois (2009) for eight compact multi-planet systems. Their results suggest that six of them (HD 69830, HD 73526, HD 108874, HD 128311, HD 155358 and HD 202206) could be dynamically regulated by such a configuration.

Retrograde packing Systems with retrograde planets can be packed substantially more closely than prograde systems with an equal number of planets (Smith & Lissauer, 2009).

Resonant planets disguised as single Various studies have shown how different orbital architectures, such as planet pairs near a 2:1 orbital resonance, or even a pair of (Trojan) planets in a hypothetical co-orbital motion, can be difficult to distinguish from a single (eccentric) planet (Giuppone et al., 2012a; Dobrovolskis, 2013a, 2015b; Leleu et al., 2015).

A specific example is the case of HD 27894, originally discovered as a single-planet system (Moutou et al., 2005a), with evidence that the eccentric orbit might be a 2:1 resonant pair (Kürster et al., 2015), followed by evidence for an inner resonant pair and an outer eccentric planet (Trifonov et al., 2017).

2.12.5 Interacting two-planet systems

A number of systems contain two giant planets in which planet–planet interactions are significant. These include: HD 12661 (Goździewski & Maciejewski, 2003); HD 14810 (Wright et al., 2009b); HD 155358 (Cochran et al., 2007); HD 169830 (Goździewski & Konacki, 2004); and HD 183263 (Wright et al., 2009b).

2.12.6 Non-interacting two-planet systems

A number of systems contain planet pairs with very large ratios of their orbital periods (sometimes referred to as ‘hierarchical’, but see the footnote on page 511), which are therefore unlikely to be interacting. These include: HD 11964 with $P = 38 \text{ d}$ and 5.3 yr (Wright et al., 2009b); HD 68988 with $P = 6.3 \text{ d}$ and $11\text{--}60 \text{ yr}$ (Wright et al., 2007); HD 168443 with $P = 58 \text{ d}$ and 4.8 yr (Marcy et al., 2001b); HD 187123 with $P = 3 \text{ d}$ and 10.5 yr (Wright et al., 2009b); HD 190360 with $P = 17.1 \text{ d}$ and 8.0 yr (Vogt et al., 2005); HD 217107 with $P = 7.1 \text{ d}$ and 11.7 yr (Wright et al., 2009b); and HD 38529 with $P = 14.3 \text{ d}$ and 5.9 yr (Fischer et al., 2003b).

2.12.7 Super-Earth systems

Systems comprising only low-mass planets are being discovered in greater numbers as radial velocity accuracies improve. The first three discovered were:

HD 69830: the triple Neptune system with $P \sim 9, 32$, and 200 d (Lovis et al., 2006). The outer planet is located near the inner edge of the habitable zone. Beichman et al. (2005a) reported Spitzer photometry and spectroscopy, which shows an infrared excess characteristic of a large cloud of fine silicate dust within a few au of the star, suggestive of a large asteroid belt.

GJ 581: a five- or six-planet system (Bonfils et al., 2005b; Udry et al., 2007; Vogt et al., 2010a), with some lying within, or close to the inner and outer edges, of the star’s habitable zone (§11.7.4). Further details are given below.

HD 40307: the third triple super-Earth system, with masses $4.2, 6.9$, and $9.2 M_{\oplus}$ (Mayor et al., 2009b).

GJ 581 GJ 581 is a nearby (6.27 pc) M3V dwarf. Radial velocity data, including 119 measurements over 4.3 yr from HARPS, and 122 over 11 yr from HIRES (Bonfils et al., 2005b; Udry et al., 2007; Mayor et al., 2009a; Vogt et al., 2010a) suggest the presence of up to six planets (not all of which have been independently confirmed to date), listed in Table 2.9, and illustrated schematically in Figure 2.47. Dynamical stability (Beust et al., 2008), issues of aliasing (Dawson & Fabrycky, 2010), and the possibility of hidden and stable orbits (Zollinger & Armstrong, 2009; Anglada-Escudé et al., 2010a) have been studied.

Udry et al. (2007) originally suggested that planet c ($M_p \sin i = 5.1 M_{\oplus}$, $P = 12.9 \text{ d}$) lies close to the inner edge of the habitable zone, while planet d ($M_p \sin i = 8.3 M_{\oplus}$, $P = 83.4 \text{ d}$) lies close to its outer edge. The original estimate of the equilibrium temperature of planet c ($T_{\text{eq}} \sim 320 \text{ K}$, assuming a Bond albedo $A_B \sim 0.5$) was revised by later studies, but

Table 2.9: The planets of GJ 581, ordered by increasing semi-major axis. Data are from Vogt et al. (2010a).

Planet	$M_p \sin i$ (M_\oplus)	a (au)	P (d)
e	1.7	0.028	3.15
b	15.6	0.041	5.37
c	5.6	0.073	12.92
g	3.1	0.146	36.56
d	5.6	0.218	66.87
f	7.0	0.758	433

with the emerging consensus that planets c and d could nevertheless support habitability depending on their atmospheres, and whether they are tidally locked (Selsis et al., 2007; von Bloh et al., 2007a; Chylek & Perez, 2007).

Further data and analysis by Vogt et al. (2010a) suggests that a proposed sixth planet, GJ 581 g ($M_p \sin i = 3.1 M_\oplus$, $P = 36.6$ d, $a = 0.146$ au, $e = 0$) could place it in the middle of the star's habitable zone. The radius of GJ 581 g is estimated as $1.3 - 1.5 R_\oplus$ if primarily composed of magnesium silicates, $1.7 - 2 R_\oplus$ if predominantly H₂O-ice, and with all radii smaller by 20% if it is significantly differentiated. The mass and radius together imply a surface gravity $\sim 1.1 - 1.7 g_\oplus$.

Its estimated $T_{\text{eq}} \sim 209 - 228$ K, depending on A_B , satisfies the conditions necessary for habitability of $T_{\text{eq}} \lesssim 270$ K derived by Selsis et al. (2007). The actual surface temperature is expected to be higher than T_{eq} . For Earth ($T_{\text{eq}} \sim 255$ K), atmospheric greenhouse heating results in a mean surface temperature $T_s \sim 288$ K. For a comparable or more massive atmosphere, greenhouse heating, combined with its probable synchronous-rotation (tidal locking), will result in higher surface temperature, plausibly implying the presence of liquid water.

Stability of habitable zone systems Numerous investigations have been carried out to examine the stability of planets which might exist in the habitable zones of known systems (§11.7.1). Using fast indicators of stability or chaos, stability maps can help to establish, for example, whether and where Earth-like planets could exist in systems having one giant planet (Figure 2.48).

Orbits of terrestrial-type planets within the habitable zone have been classified as fully stable if no mean motion resonance exists within that zone, partially stable if only a few resonances exist, marginally stable if many resonances exist, and very unstable if the region is found to be strongly chaotic (Sándor et al., 2007b).

2.13 Binary and multiple stars

Early radial velocity surveys tended to exclude spectroscopic and visual binary stars with separations below 2–6 arcsec because of problems of light contamination from the other component at the spectrograph entrance (Udry et al., 2000; Perrier et al., 2003; Marcy et al., 2005b; Jones et al., 2006b). This resulted in limited or biased surveys for binary star separations $\lesssim 100 - 200$ au.

At larger separations, viz. for visual binaries with separations $\gtrsim 100$ au, circumstellar planet searches can treat the targets as distinct single stars, and such surveys

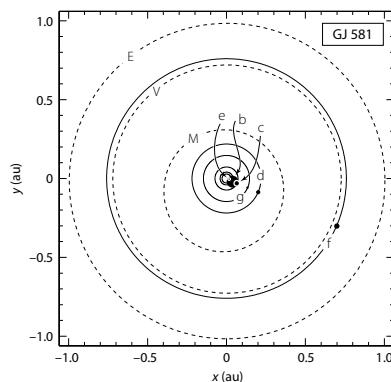


Figure 2.47: View of the GJ 581 system from above, with coordinates in au. Planets, shown with a circle size proportional to mass, are from inner to outer: e, b, c, g, d, f, of orbital period 3.1, 5.4, 12.9, 36.6, 66.9, 433 d respectively, and mass ($M_p \sin i$) 1.7, 15.6, 5.6, 3.1, 5.6, 7.0 M_\oplus respectively. The orbits of M(ercury), V(enus), and E(arth) are shown to the same scale. From Vogt et al. (2010a, Figure 6), by permission of IOP Publishing/AAS.

have faced limited technical difficulty (Gratton et al., 2003; Desidera & Barbieri, 2007; Toyota et al., 2009).

Observational confirmation that giant planets existed orbiting one component of a spectroscopic binary star came with the discoveries of GJ 86 Ab (Queloz et al., 2000b) with its white dwarf companion, and γ Cep Ab (Hatzes et al., 2003a), a binary with $P = 57$ yr, with the planet orbiting the primary with $P = 2.48$ yr.

The growing interest in exploring broader planetary domains as a way of constraining formation models has led to some more recent surveys preferentially including, rather than excluding, binary systems (Table 2.6).

Amongst specific radial velocity surveys focusing on binaries are a sample of 101 single-lined spectroscopic binaries (SB1) with $P > 1.5$ yr, selected on the basis of the earlier CORAVEL surveys (Duquennoy & Mayor, 1991; Halbwachs et al., 2003). A programme to supplement the original CORAVEL data with 10–15 additional high-precision measurements of each system, using ELODIE in the north and CORALIE in the south, was initiated in 2001 (Eggenberger, 2010). Analysis employs the spectral deconvolution programme TODCOR. No promising candidates have been reported, and constraints indicate that $\lesssim 20\%$ of SB1 systems host a short-period, $P < 40$ d circumbinary giant ($M_p \sin i \gtrsim 0.5 M_J$) planet.

No extensive surveys have yet been undertaken for double-lined spectroscopic binaries (SB2) (Eggenberger, 2010). The principal reason for this is related to the problem of wavelength calibration: the standard technique using iodine cells (§2.3) is only applicable to single stars, and the time-varying spectra of SB2s cannot provide the template spectrum required by Equation 2.45. A modified procedure applicable to SB2s, using interleaved observations with and without the iodine cell, has been described by Konacki (2005b).

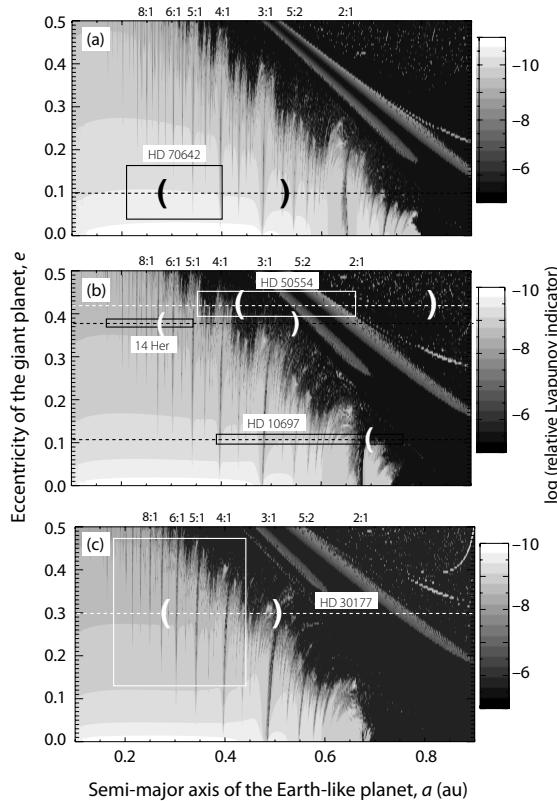


Figure 2.48: Stability maps, computed by the method of relative Lyapunov indicators, for orbits of possible habitable-zone planets interior to the orbit of a known giant planet, for various values of mass ratios $\mu = M_p/(M_\star + M_p)$, where M_p is the mass of the giant planet, and the mass of the hypothetical terrestrial-type planet is negligible: (a) $\mu = 0.002$, (b) $\mu = 0.005$, (c) $\mu = 0.009$. The habitable zones corresponding to the star's zero-age main sequence luminosity are indicated by elongated rectangles, whose heights correspond to the uncertainty on the giant planet's eccentricity, while the current habitable zones are indicated by parentheses. Positions of mean motion resonances are indicated. In (c), for example, the eccentricity of the giant planet in the HD 30177 system is highly uncertain. The present habitable zone is strongly chaotic for large values of e , while its inner part is marginally stable for small e . From Sánchez et al. (2007b, Figures 7, 9 and 10), © Oxford University Press.

Other surveys for binary identification In parallel, the fraction of planets known to reside in binary systems has risen as a result of searches for common proper motion companions to known planet-hosting stars.

Other searches for planets around binary systems include transit studies of eclipsing binaries (§6.3) which may allow the detection of transiting circumstellar or circumbinary planets (Deeg et al., 1998; Doyle et al., 2000; Ofir, 2008), or Trojan planets (Caton et al., 2000), and of non-transiting giant planets in circumbinary orbits through eclipse timing. The latter has resulted in the discovery of a few circumbinary systems (§4.4).

2.13.1 Present inventory

By late 2010, more than 50 planets had been found associated with binary or multiple stars (Eggenberger et al., 2004b; Eggenberger, 2010). Although most were giants orbiting the primary, with projected separations in the range 20–12 000 au, one possibly circumbinary planet, HD 202206, had been reported (Correia et al., 2005).

More recent statistics of planetary systems discovered orbiting binary stars are given by Crepp et al. (2016), and their distribution as a function of binary separation are given in their Figure 2. Few small-separation systems have been searched for with the radial velocity technique because existing spectrometers are seeing-limited. A diffraction-limited Doppler spectrometer for the Large Binocular Telescope (LBT), iLocater, to measure close-in system, is described in Section 2.6.6.

At small binary separations, $\lesssim 100$ au, only a small number of planets have been found, and giant planets may be rarer in these systems than around wider separation binaries or single stars (e.g. Bonavita & Desidera, 2007). This may be consistent with studies which suggest that giant planet formation is inhibited under the influence of a nearby orbiting stellar companion (Nelson, 2000; Mayer et al., 2005; Thébault et al., 2006).

At the smallest binary separations, $\lesssim 20$ au, the absence of circumstellar planets is consistent with the idea that giant planet formation in these systems is suppressed (Nelson, 2000; Thébault et al., 2004; Mayer et al., 2005; Thébault et al., 2006; Boss, 2006a), although Doppler surveys to confirm this remain incomplete.

Properties Roell et al. (2012) estimated that, as of 2012, 57 exoplanet (radial velocity or transit) host stars have a *stellar* companion. The resulting multiplicity rate of (at least) 12% for exoplanet host stars is about four times smaller than that of solar-like stars in general. The mass and the number of planets in these multiple stellar systems depends on the separation between the two stellar components, e.g. the planetary mass decreases with an increasing stellar separation.

Zucker & Mazeh (2002) suggested that planets in binaries follow a different mass–period relation than those orbiting single stars, with the most massive short-period planets being found in binary or multiple systems. Later studies showed a lower significance (Eggenberger et al., 2004b; Desidera & Barbieri, 2007; Mugrauer et al., 2007; Eggenberger, 2010), and the most recent distributions (Figure 2.49a) reveal mixed evidence for such an effect. Meanwhile, short period ($P < 50$ d) planets around binary stars do appear to have smaller eccentricities, $e \lesssim 0.05$, than around single stars (Figure 2.49b).

Still based on somewhat limited statistics, the few planets with the highest eccentricities, $e > 0.8$, tend to be accompanied by a stellar or brown dwarf companion (Tamuz et al., 2008), a correlation that may extend to lower masses (Ribas & Miralda-Escudé, 2007). This has been taken to support the idea of eccentricities being excited by the Kozai mechanism, in which hierarchical triple systems with high relative inclinations cause large-amplitude periodic oscillations of the eccentricity of the inner pair (Holman et al., 1997; Innanen et al., 1997; Mazeh et al., 1997; Ford et al., 2000b; Wu & Murray, 2003; Takeda & Rasio, 2005; Moutou et al., 2009a).

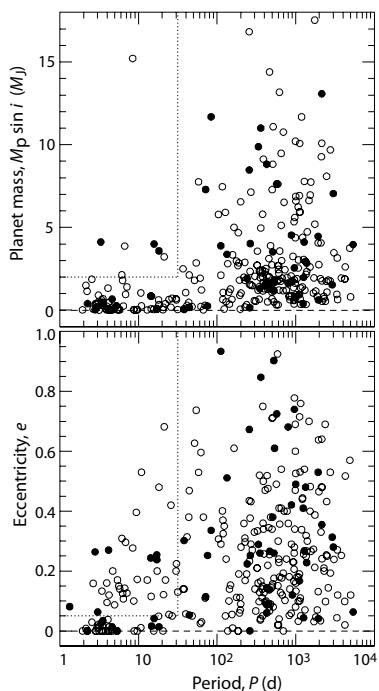


Figure 2.49: Planets detected from radial velocities (317 planets from exoplanets.org, 2010–11–01), with those orbiting single stars shown as open circles, and those in binary or multiple systems as filled circles. Top: mass ($M_p \sin i$) versus orbital period. Bottom: eccentricity versus period. Dashed regions correspond to those delineated by Eggenberger (2010, Figure 1).

The coupling of Lidov–Kozai oscillation with tidal friction, also referred to as Lidov–Kozai migration, might also lead to short-period planets preferentially following nearly circular orbits (Figure 2.49b). In multiple systems this mechanism may be more effective than type II migration in bringing massive planets close to their host stars. This may also explain why the most massive short-period planets are found in binary or multiple systems (Takeda & Rasio, 2006; Fabrycky & Tremaine, 2007).

2.13.2 Specific examples

γ Cep The binary star system γ Cep harbours a stable giant planet orbiting the primary star at a distance some one tenth of the stellar separation (Hatzes et al., 2003a). The primary and secondary stars, of masses 1.6 and $0.4M_\odot$, are separated by 18.5 au ($P \sim 57$ yr, $e = 0.36$), with the planet of mass $1.7M_J$ orbiting the primary with $a = 2.13$ au, $P = 2.5$ yr, $e = 0.12$.

Numerical integrations as a function of a_b , e_b , and i_b (the inclination of the planet with respect to the stellar binary) by Haghighipour (2006), show that the planet orbit is stable for values of the binary eccentricity $0.2 \leq e_b \leq 0.45$, with the system becoming unstable within a few thousand years for $e_b > 0.5$. Within the stable range of orbital eccentricity, the planet remains stable for $i_b < 40^\circ$, becoming unstable within a few thousand years for $i_b > 40^\circ$. For large values of the inclination, the system may be locked in a Kozai resonance (Haghighipour, 2004). The stability map for test particles interior to the planet orbit show complex stability patterns influenced by mean motion resonances (Figure 2.50).

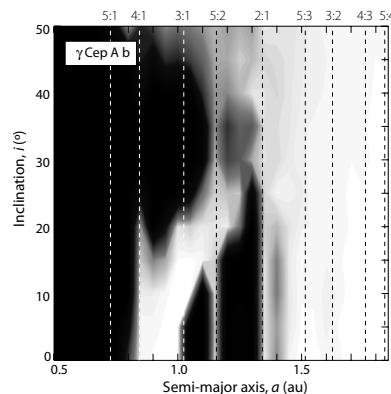


Figure 2.50: Stability of test particles interior to the planet γ Cep Ab, for the co-planar case $i_b = 0$. Stability is indicated by the test particle's survival time averaged over longitudes and normalised to 1 Myr. Lighter regions are unstable, darker regions indicate higher stability. Dashed vertical lines are nominal locations of mean motion resonances with γ Cep Ab. From Verrier & Evans (2006, Figure 2), © Oxford University Press.

Various studies of planet formation in this system have been made (Kley & Nelson, 2008; Jang-Condell et al., 2008; Paardekooper & Leinhardt, 2010; Beaugé et al., 2010; Giuppone et al., 2011; Müller & Kley, 2012; Martí & Beaugé, 2012; Rafikov, 2013b; Rafikov & Silsbee, 2015b; Funk et al., 2015).

16 Cyg B The planet around 16 Cyg B has a highly eccentric orbit, $e \sim 0.7$. 16 Cyg B is one component of a wide-separation binary whose eccentricity is itself large, $e > 0.54$. The orbital period of the binary star is long and difficult to measure accurately, but from astrometry over 170 years is estimated to be $> 18\,000$ yr (Hauser & Marcy, 1999). 16 Cyg A may also have a close stellar companion (Patience et al., 2002).

The system is an interesting case for planetary formation theories, which try to explain the large eccentricities in binary systems as a result of gravitational perturbations by the other stellar component, here 16 Cyg A. If the planet was originally formed in a circular orbit, with an orbital plane inclined to that of the stellar binary by $> 45^\circ$, then the planet orbit is predicted to oscillate chaotically between high- and low-eccentricity states, on time scales of $10^7 - 10^{10}$ yr, (Holman et al., 1997; Mazeh et al., 1997; Hauser & Marcy, 1999; Plávalová & Solovaya, 2013). Further planets with $M_p \sim M_J$ within 30 au would require this conclusion to be revisited.

HD 188753 Some controversy surrounds this triple system (Mazeh et al., 2009b), comprising a 155 d spectroscopy binary (HD 188753B) in a 25.7-yr orbit ($a = 12.3$ au) about the primary (HD 188753A). Konacki (2005a) reported a 3.35 d $1.1M_J$ planet around the primary, based on 10 m s^{-1} accuracy Keck–HIRES data, unconfirmed by Eggenberger et al. (2007b).

According to theoretical studies, the pericentre distance of the AB pair is too small to have allowed giant planet formation. Modeling by Jang-Condell (2007) has shown that the circumstellar disk around the primary would have been truncated at around 1.5–2.7 au, leaving insufficient material, and in any case too hot, for a Jovian planet to have formed *in situ*. Alternative formation mechanisms were forthcoming, in particular proposing that the primary could have acquired the planet through interactions within an open star cluster that has now dissolved (Pfahl, 2005; Portegies Zwart & McMillan, 2005).