Astronomy
&
Astrophysics

The distribution of exoplanet masses

A. Jorissen[1,★], M. Mayor[2], and S. Udry[2]

[1] Institut d'Astronomie et d'Astrophysique, Université Libre de Bruxelles, CP 226, Boulevard du Triomphe,
1050 Bruxelles, Belgium

[2] Observatoire de Genève, 1290 Sauverny, Switzerland
e-mail: `Michel.Mayor, Stephane.Udry@obs.unige.ch`

Abstract. The present study derives the distribution of secondary masses M_2 for the 67 exoplanets and very low-mass brown-dwarf companions of solar-type stars, known as of April 4, 2001. This distribution is related to the distribution of $M_2 \sin i$ through an integral equation of Abel's type. Although a formal solution exists for this equation, it is known to be ill-conditioned, and is thus very sensitive to the statistical noise present in the input $M_2 \sin i$ distribution. To overcome this difficulty, we present two robust, independent approaches: (i) the formal solution of the integral equation is numerically computed after performing an optimal smoothing of the input distribution and (ii) the Lucy-Richardson algorithm is used to invert the integral equation. Both approaches give consistent results. The resulting statistical distribution of exoplanet true masses reveals that there is no reason to ascribe the transition between giant planets and brown dwarfs to the threshold mass for deuterium ignition (about 13.6 M_J). The M_2 distribution shows instead that most of the objects have $M_2 \leq 10 \ M_J$, but there is a small tail with a few heavier candidates around 15 M_J.

Key words. methods: numerical – stars: planetary systems

1. Introduction

Han et al. (2001) suggested that most of the exoplanet candidates discovered so far have masses well above the lower limit defined by $\sin i = 1$ (where i is the inclination of the orbital plane on the sky) and should therefore be considered as brown dwarfs or even stars rather than planets. The present paper shows that this claim is not consistent with the distribution of masses extracted from the observed $M_2 \sin i$ distribution (where M_2 is the mass of the companion) under the reasonable assumption that the orbits are oriented at random in space. Although the distributions of M_2 and $M_2 \sin i$ are related through an integral equation of Abel's kind (Chandrasekhar & Münch 1950; Lucy 1974), its numerical solution is ill-conditioned. Two different methods are used here to overcome this difficulty. In the first method (Sect. 2), the formal solution of Abel's equation is implemented numerically on an input $M_2 \sin i$ distribution that has been optimally smoothed with an adaptive kernel procedure (Silverman 1986) to remove the high-frequency fluctuations caused by small-number statistics. The other method (Sect. 3) is based on the Lucy-Richardson inversion algorithm (Richardson 1972; Lucy 1974).

Send offprint requests to: A. Jorissen,
e-mail: `ajorisse@astro.ulb.ac.be`
★ Research Associate, FNRS, Belgium.

The basic reason why the M_2 distribution obtained in Sect. 4 differs from that of Han et al. (2001) is because these authors concluded that most of the systems containing exoplanet candidates are seen nearly pole-on. This conclusion, based on the analysis of the Hipparcos *Intermediate Astrometric Data* (IAD), has however been shown to be incorrect (e.g. Halbwachs et al. 2000; Pourbaix 2001; Pourbaix & Arenou 2001), as summarized in Sect. 5.

While this paper was being reviewed, Zucker & Mazeh (2001) and Tabachnik & Tremaine (2001) proposed other interesting approaches to derive the exoplanet mass distribution. Zucker & Mazeh derive the binned true mass distribution by using a maximum likelihood approach to retrieve the average values of the mass distribution over the selected bins. Their results are in very good agreement with ours. Tabachnik & Tremaine suppose that the period and mass distributions follow power laws, and derive the corresponding power-law indices from a maximum likelihood method. On the contrary, the methods used in the present paper (and in Zucker & Mazeh's) are *non-parametric in nature*, since they do not require any a priori assumptions on the functional form of the mass distribution. This is especially important since the comparison of the shapes of the mass distributions for exoplanets and low-mass stellar companions may provide clues to identify the process by which they formed. By imposing a power-law function like that observed for low-mass stellar

companions, Tabachnik & Tremaine (2001) implicitly assume that these processes must be similar.

2. The integral equation of Abel's kind relating the distributions of $M_2 \sin i$ and M_2

The $M_2 \sin i$ values for low-mass companions of main sequence stars may be extracted from the spectroscopic mass function and from the primary mass as derived through e.g., isochrone fitting. Let $\Phi(Y)$ be the observed distribution of $Y \equiv M_2 \sin i$, that is easily derived from the observed spectroscopic mass functions provided that $M_2 \ll M_1$ as is expected to be the case for the systems under consideration. Then, the distribution $\Psi(M_2)$ obeys the relation

$$\Phi(Y) = \int_0^\infty \Psi(M_2) \, \Pi(Y|M_2) \, \mathrm{d}M_2. \tag{1}$$

The kernel $\Pi(Y|M_2)$ corresponds to the conditional probability of observing the value Y given M_2. Under the assumption that the orbits are oriented at random in space, the inclination angle i is distributed as $\sin i$, and the following expression is obtained for the kernel:

$$\Pi(Y|M_2) = \frac{\sin i_0}{M_2 \cos i_0}, \tag{2}$$

where i_0 satisfies the relations $M_2 \sin i_0 - Y = 0$ and $0 \le i_0 \le 90$. Eliminating the inclination i_0 in the above relation yields

$$\Pi(Y|M_2) = \frac{Y}{M_2} \frac{1}{(M_2^2 - Y^2)^{1/2}} \qquad \text{with} \quad Y \le M_2. \tag{3}$$

Equation (1) thus rewrites

$$\Phi(Y) = Y \int_Y^\infty \Psi(M_2) \frac{1}{M_2(M_2^2 - Y^2)^{1/2}} \, \mathrm{d}M_2. \tag{4}$$

Equation (4) is the integral equation to be solved for $\Psi(M_2)$. It can be reduced to Abel's integral equation by the substitutions (Chandrasekhar & Münch 1950)

$$Y^2 = 1/\eta \quad \text{and} \quad M_2^2 = 1/t. \tag{5}$$

With these substitutions, Eq. (4) becomes

$$\phi(\eta) = \int_0^\eta \frac{\psi(t)}{(\eta - t)^{1/2}} \, \mathrm{d}t, \tag{6}$$

where

$$\phi(\eta) \equiv \Phi(\frac{1}{\sqrt{\eta}}) \quad \text{and} \quad \psi(t) \equiv \frac{1}{2\sqrt{t}} \Psi(\frac{1}{\sqrt{t}}). \tag{7}$$

It is well known that the solution of Abel's equation (Eq. (6)) is given by

$$\psi(t) = \frac{1}{\pi} \int_0^t \frac{\partial \phi}{\partial \eta} \frac{1}{(t - \eta)^{1/2}} \, \mathrm{d}\eta + \frac{1}{\pi} \frac{\phi(0)}{\sqrt{t}}, \tag{8}$$

where $\phi(0) = \lim_{Y \to \infty} \Phi(Y) = 0$.

While Eq. (8) represents the formal solution of the problem, it is difficult to implement numerically, since it requires the differentiation of the observed frequency distribution $\Phi(Y)$. Unless the observations are of high precision, it is well known that this process can lead to misleading results. To overcome this difficulty, the observed frequency distribution has been smoothed in an optimal way (see Appendix) before being used in Eq. (8). The solution $\Psi(t)$ is then computed numerically using standard differentiation and integration schemes.

3. The Lucy-Richardson inversion algorithm applied to Abel's integral equation

The Lucy-Richardson algorithm provides another robust way to invert Eq. (4) (see also Cerf & Boffin 1994). The method starts from the Bayes theorem on conditional probability in the form

$$\Psi(M_2) \, \Pi(Y|M_2) = \Phi(Y) \, R(M_2|Y), \tag{9}$$

where $R(M_2|Y)$ is the reciprocal kernel corresponding to the integral equation inverse to the one that needs to be solved (Eq. (1)):

$$\Psi(M_2) = \int_0^{M_2} \Phi(Y) \, R(M_2|Y) \, \mathrm{d}Y. \tag{10}$$

The reciprocal kernel represents the conditional probability that the binary system has a companion mass M_2 when the observed $M_2 \sin i$ value amounts to Y. Thus, one has:

$$R(M_2|Y) = \frac{\Psi(M_2) \, \Pi(Y|M_2)}{\Phi(Y)} \tag{11}$$

$$= \frac{\Psi(M_2) \, \Pi(Y|M_2)}{\int_0^\infty \Psi(M_2) \, \Pi(Y|M_2) \, \mathrm{d}M_2}, \tag{12}$$

which obviously satisfies the normalization condition $\int_0^\infty R(M_2|Y) \, \mathrm{d}M_2 = 1$. The problem in solving Eq. (10) arises because $R(M_2|Y)$ also depends on $\Psi(M_2)$, so that an iterative procedure must be used. If $\Psi_r(M_2)$ is the rth estimate of $\Psi(M_2)$, it can be used to obtain the $(r+1)$th estimate in the following way:

$$\Psi_{r+1}(M_2) = \int_0^{M_2} \Phi(Y) \, R_r(M_2|Y) \, \mathrm{d}Y \tag{13}$$

with

$$R_r(M_2|Y) = \frac{\Psi_r(M_2) \, \Pi(Y|M_2)}{\Phi_r(Y)} \tag{14}$$

and

$$\Phi_r(Y) = \int_0^\infty \Psi_r(M_2) \, \Pi(Y|M_2) \, \mathrm{d}M_2. \tag{15}$$

Thus, $\Phi_r(Y)$ represents the corresponding rth estimate of the observed distribution $\Phi(Y)$. Equations (13) and (14) together yield the recurrence relation for the Ψ_r's,

$$\Psi_{r+1}(M_2) = \Psi_r(M_2) \int_0^{M_2} \frac{\Phi(Y)}{\Phi_r(Y)} \, \Pi(Y|M_2) \, \mathrm{d}Y, \tag{16}$$

with $\Pi(Y|M_2)$ given by Eq. (3) for the problem under consideration. The conditions for convergence of this recurrence relation are discussed by Lucy (1974) and Cerf & Boffin (1994). It needs only be remarked here that (i) the iterative scheme converges if $\Phi_r(Y)$ tends to $\Phi(Y)$, given the normalization of the probability $\Pi(Y|M_2)\mathrm{d}Y$, and (ii) the full convergence of the method is not necessarily desirable, as the successive estimates $\Phi_r(Y)$ will tend to match $\Phi(Y)$ on increasingly smaller scales, but the small-scale structure in $\Phi(Y)$ is likely to be dominated by the noise in the input data. This is well illustrated in Fig. 2 below.

When the number of data points is small (typically $N < 100$; Cerf & Boffin 1994), it is advantageous to express $\Phi(Y)$ as

$$\Phi(Y) = \frac{1}{N} \sum_{n=1}^{N} \delta(Y - y_n) \qquad (17)$$

where the y_n ($n = 1, \dots N$) are the N individual measured $M_2 \sin i$ values and $\delta(x)$ is the Dirac "function" such that $x_0 = \int \delta(x - x_0)\,\mathrm{d}x$. Substitution in Eq. (13) then yields

$$\Psi_{r+1}(M_2) = \frac{1}{N} \sum_{n=1}^{N} R_r(M_2|y_n) \qquad (18)$$

where $R_r(M_2|y_n)$ is defined as in Eq. (14). The sample size should nevertheless be large enough for the functions $R_r(M_2|y_n)$ to have sufficient overlap so as to produce a smooth $\Psi_{r+1}(M_2)$ function.

In the application of the method described in Sect. 4, the initial mass distribution $\Psi_0(M_2)$ was taken as a uniform distribution, but it has been verified that this choice has no influence on the final solution.

4. The frequency distribution $\Psi(M_2)$

The cumulative frequency distribution of the $M_2 \sin i$ values smaller than $17\,M_\mathrm{J}$, where M_J is the mass of Jupiter ($=M_\odot/1047.35$), available in the literature (as of April 4, 2001) is presented in Fig. 1. It appears to be sufficiently well sampled to attempt the inversion procedure. The corresponding frequency distributions smoothed with two different smoothing lengths, locally self-adapting around $h_\mathrm{opt} = 1\,M_\mathrm{J}$ and $2h_\mathrm{opt}$ (see Appendix) are presented as well for comparison.

The sample includes 60 main-sequence stars hosting 67 companions with $M_2 \sin i < 17\,M_\mathrm{J}$. Among those, 6 stars are orbited by more than one companion, namely HD 74156 and HD 82943 ($M_2 \sin i = 1.50$, $7.40\,M_\mathrm{J}$, and $M_2 \sin i = 0.84$, $1.57\,M_\mathrm{J}$, respectively; Udry & Mayor 2001)[1], HD 83433 (0.16,0.38 M_J; Mayor et al. 2000), HD 168443 (7.22, 16.2 M_J; Udry et al. 2000b), υ And (0.71, 2.20 and 4.45 M_J; Butler et al. 1999) and

---

[1] See also `http://www.eso.org/outreach/press-rel/pr-2001/pr-07-01.html`
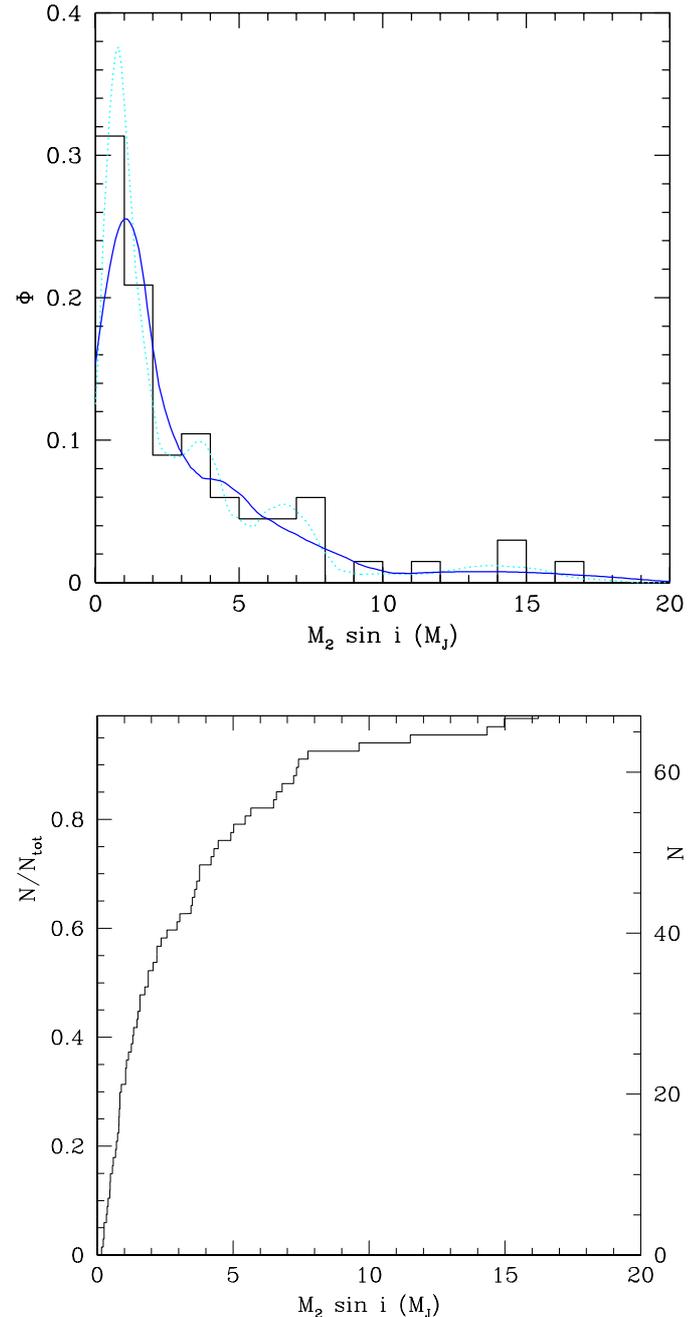

Fig. 1. Bottom panel: the cumulative frequency distribution of the observed $M_2 \sin i$ values for the 67 exoplanets and very low-mass brown dwarfs (around 60 stars) known as of April 4, 2001. Top panel: the corresponding frequency distribution, represented either as an histogram, or smoothed using an Epanechnikov kernel (see Appendix) with an optimal smoothing length of $h_\mathrm{opt} = 1\,M_\mathrm{J}$ (dashed curve). The frequency distribution smoothed with $2h_\mathrm{opt}$ is also shown for comparison (solid curve).

Gliese 876 (0.56, 1.88 M_J; Marcy et al. 2001). The inversion process is only able to treat these systems under the hypothesis that the orbits of the different planets in a given system are *not* coplanar, since Eq. (4) to hold requires random orbital inclinations. The case of coplanar

and non-coplanar orbits are discussed separately in the remainder of this section.

4.1. Non-coplanar orbits

Figure 2 compares the solutions $\Psi(M_2)$ obtained from the Lucy-Richardson algorithm (after 2 and 20 iterations, denoted Ψ_2 and Ψ_{20}, respectively) and from the formal solution of Abel's integral equation with smoothing lengths h_{opt} and $2\,h_{\text{opt}}$ on $\Phi(Y)$ (the corresponding solutions are denoted $\Psi_{h_{\text{opt}}}$ and $\Psi_{2h_{\text{opt}}}$). The solutions from the two methods basically agree with each other, although solutions with different degrees of smoothness may be obtained with each method. On the one hand, Ψ_{20} and $\Psi_{h_{\text{opt}}}$ exhibit high-frequency fluctuations that may be traced back to the statistical fluctuations in the input data. This can be seen by noting that the peaks present in $\Psi_{h_{\text{opt}}}$ correspond in fact to the high-frequency fluctuations already present in $\Phi_{h_{\text{opt}}}$ (Fig. 1). These fluctuations should thus not be given much credit. The same explanation holds true for Ψ_{20}, since it was argued in Sect. 3 that the solutions Ψ_r resulting from a large number of iterations tend to match Φ at increasingly small scales (i.e., higher frequencies) where statistical fluctuations become dominant. On the other hand, Ψ_2 and $\Psi_{2h_{\text{opt}}}$ are much smoother, and are probably better matches to the actual distribution. The local maximum around $M_2 \sim 1\ M_{\text{J}}$ is very likely, however, an artifact of the strong detection bias against low-mass companions.

The most striking feature of the $\Psi(M_2)$ distribution displayed in Fig. 2 is its decreasing character, reaching zero for the first time around $M_2 = 10\ M_{\text{J}}$, and in any case well before 13.6 M_{J}. The latter value, corresponding to the minimum stellar mass for igniting deuterium, does not in any way mark the transition between giant planets and brown dwarfs, as sometimes proposed. That transition, which is thus likely to occur at smaller masses, must rely instead on the different mechanisms governing the formation of planets and brown dwarfs. Another argument favouring a giant-planet/brown-dwarf transition mass smaller than 13.6 M_{J} is provided e.g., by the observation of *free-floating* (and thus most likely *stellar*) objects with masses probably smaller than 10 M_{J} in the σ Orionis star cluster (Zapatero Osorio et al. 2000). The $\Psi(M_2)$ distribution nevertheless clearly exhibits a tail of objects clustering around $M_2 \sim 15\ M_{\text{J}}$, due to HD 114762 ($M_2 \sin i = 11.5\ M_{\text{J}}$), HD 162020 (14.3 M_{J}), HD 202206 (15.0 M_{J}) and HD 168443 c (16.2 M_{J}). It would be interesting to investigate whether these systems differ from those with smaller masses in some identifiable way (periods, eccentricities, metallicities, ...), so as to assess whether or not they form a distinct class (Udry et al., in preparation).

The jackknife method (e.g., Lupton 1993) has been used to estimate the uncertainty on the $\Psi_{2h_{\text{opt}}}$ solution. In a first step, 67 input $\Phi_{2h_{\text{opt}}}$ distributions are computed, corresponding to all 67 possible sets with one data point
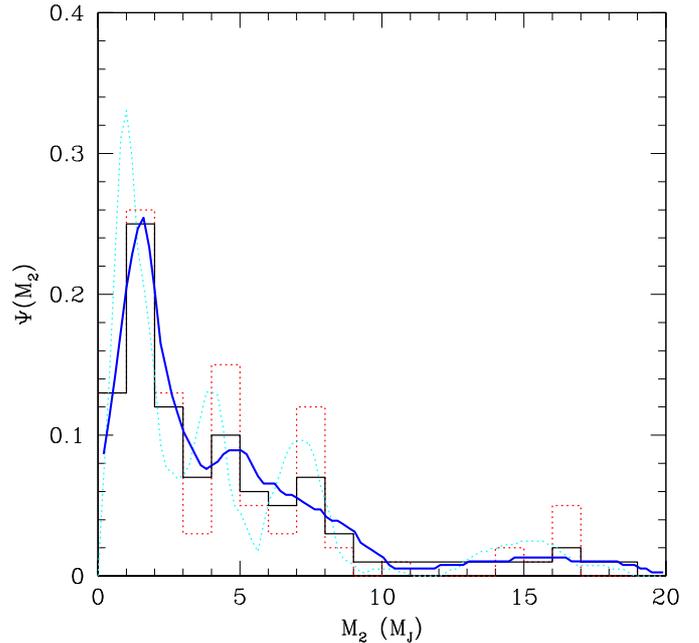

Fig. 2. Comparison of $\Psi(M_2)$ distributions obtained from the Lucy-Richardson algorithm (after 2 and 20 iterations, represented by the solid and dashed histograms, respectively), and from the formal solution of Abel's integral equation with smoothing lengths $2h_{\text{opt}}$ and h_{opt} applied on $\Phi(Y)$ (represented by solid and dashed curves, respectively). The supposedly best representation of the actual $\Psi(M_2)$ distribution is represented by the thick solid line. Planetary systems have been included in the inversion process with the assumption of non-coplanarity of the planetary orbits in a given system.

removed from the original set. Equation (8) is then applied to these 67 different input distributions. The resulting distributions are displayed in Fig. 3, which shows that the threshold observed at 10 M_{J} is a robust result not affected by the uncertainty on the solution.

4.2. Coplanar orbits in multi-planets systems

All the results discussed so far are obtained under the assumption that orbits of planets belonging to a planetary system are *not* coplanar. To evaluate the impact of this hypothesis, the following procedure has been applied. In a first step, the Lucy-Richardson algorithm is applied on the data set excluding the 13 planets belonging to planetary systems. That mass distribution obtained after 2 iterations is then completed by mass estimates for the remaining 13 planets. For each of the 6 different systems, an inclination i is drawn from a $\sin i$ distribution. This is done through the expression $i = \arccos x$, where x is a random number with uniform deviation. The same value of i is then applied to all planets in a given system to extract M_2 from the observed $M_2 \sin i$ value. The distributions of exoplanet masses obtained with and without the hypothesis of coplanarity are compared in Fig. 4, and it is seen that planetary systems are not yet numerous enough for the
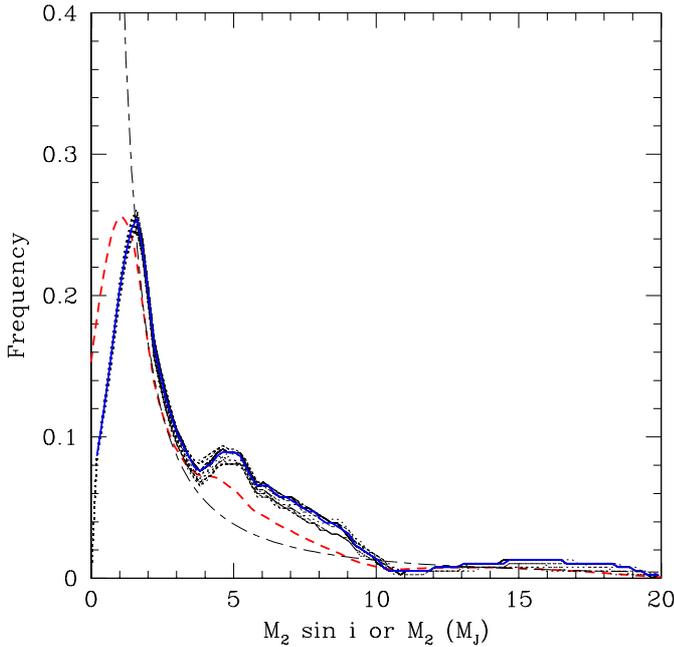
Fig. 3. Comparison of the input $\Phi_{2h_{opt}}$ distribution (thick dashed line) and the 67 $\Psi_{2h_{opt}}$ distributions (thin dotted lines) resulting from the application of the jackknife method (see text), which illustrates the uncertainty on the solution $\Psi_{2h_{opt}}$ (thick solid line). To guide the eye, a power-law of index -1.6 has been plotted as well (dot-dashed line).
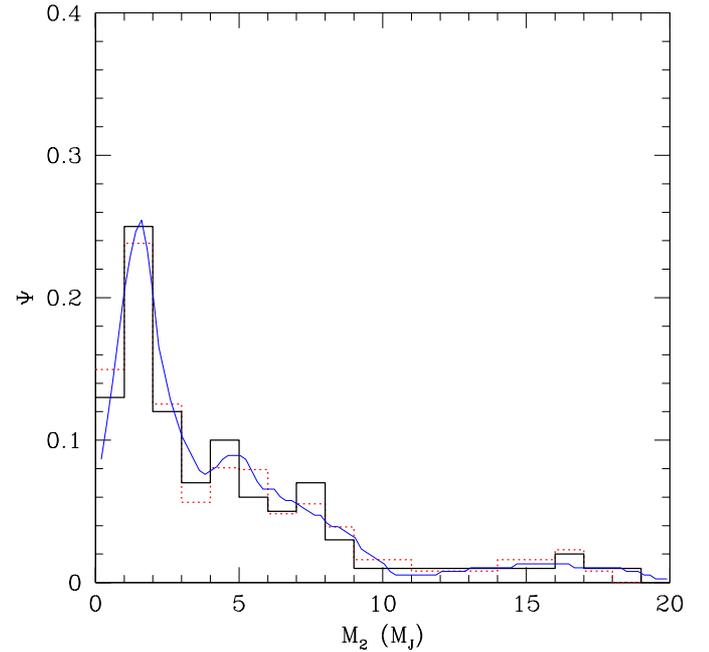
Fig. 4. Evaluation of the impact of the coplanarity hypothesis on the resulting $\Psi(M_2)$ distribution. The dashed histogram corresponds to the mass distribution obtained assuming coplanar orbits in planetary systems (see text), as compared to the Ψ_2 (solid histogram) and $\Psi_{2h_{opt}}$ solutions (see Fig. 2).

coplanarity hypothesis to alter significantly the resulting $\Psi(M_2)$ distribution.

In any case, the main result of the present paper is that *the statistical properties of the observed $M_2 \sin i$ distribution coupled with the hypothesis of randomly oriented orbital planes confine the vast majority of planetary companion masses below about 10 M_J.* Zucker & Mazeh (2001) reach the same conclusion.

It should be remarked that the above conclusion cannot be due to detection biases, since the high-mass tail of the M_2 distribution is not affected by the difficulty of finding low-amplitude, long-period orbits.

5. Discussion

Although the assumption of random orbital inclinations seems reasonable, it is at variance with the conclusion of Han et al. (2001) that most of the systems containing exoplanet candidates are seen nearly pole-on. These authors reached this conclusion by trying to extract the astrometric orbit, hence the orbital inclination, from the Hipparcos IAD. Halbwachs et al. (2000) had already cautioned that this approach is doomed to fail for systems with apparent separations on the sky that are below the Hipparcos sensitivity (i.e. $\lesssim 1$ mas). In those cases, the solution retrieved from the fit of the IAD residuals is spurious, since the true angular semi-major axis a is simply too small to be seen by Hipparcos. Since Halbwachs et al. (2000) have shown that a actually follows a Rayleigh probability distribution, the fit of the IAD residuals will yield a solution larger than

the true value, in fact of the order of the residuals. But since $a \sin i$ is constrained by the spectroscopic orbital elements, the too-large astrometric a value will force i to be close to 0 to match the spectroscopic value of the product $a \sin i$, as convincingly shown by Pourbaix (2001). Hence, this approach gives the impression that all orbits are seen nearly face-on. As an illustrative example, Pourbaix & Arenou (2001) have shown that such an approach leads to a stellar mass for the companion of HD 209458 that, on another hand, has been proven to be a 0.69 M_J planet by the photometric observation of the planet transit in front of the star (Charbonneau et al. 2000).

The Han et al. (2001) result is moreover statistically very unlikely if the orbital planes are oriented at random in space (Pourbaix & Arenou 2001). Han et al. (2001) have tried to justify this unlikely statistical occurrence by invoking biases against high-amplitude orbits in the selection process of the radial-velocity-monitoring samples. To the contrary, the planet-search surveys were specifically devised to avoid such biases, as they aim at finding not only giant planets but also brown dwarfs so as to constrain the substellar secondary mass function of solar-type stars. Furthermore, the Han et al. (2001) argument is totally invalid in the case of volume-limited, statistically well-defined samples like that of the CORALIE planet-search programme in the southern hemisphere (Udry et al. 2000a). This sample has been specifically designed to detect companions of solar-type stars all the way from $q = M_2/M_1 = 1$ down to $q \leq 0.001$.

Appendix: Non-parametric treatment of the data

To decrease the noise and allow a tractable use of the information present in small data samples, heavy smoothing techniques are often required. A common practice consists converting a set of discrete positions into binned "counts". Binning is a crude sort of smoothing and many studies in statistical analysis have shown that, unless the smoothing is done in an optimum way, some, or even most, of the information content of the data could be lost. This is especially true when a large amount of smoothing is necessary, which then changes the "shape" of the resulting function. In statistical terms, the smoothing process not only decreases the noise (i.e., the function's variance), but at the same time introduces a bias in the estimate of the function.

The variance-bias trade off is unavoidable but, for a given level of variance, the bias increase can be minimized. The correct manner of achieving that task is provided by the so-called *non-parametric density estimate* methods for the determination of the "frequency" function of a given parameter or by the *non-parametric regression* methods for the determination of a smooth function g inferred from direct measurements of g itself. Moreover, *adaptive* non-parametric methods are designed to filter the data in some local, objective way minimizing the bias, in order to get the smooth desired function without constraining its form in any way. The theory and algorithms related to those methods, originally built to handle ill-conditioned problems (either under-determined or extremely sensitive to errors or incompleteness in the data), are widely discussed in the specialized literature and summarized in easy-to-read textbooks (e.g., Silverman 1986; Härdle 1990; Scott 1992).

The simplest of the available algorithms is provided by the kernel estimator leading to the following form of the normalized "frequency" function

$$\hat{f}_K(x) = \frac{1}{Nh} \sum_{n=1}^{N} K\left(\frac{x - X_n}{h}\right) \tag{19}$$

with the normalization

$$\int_{-\infty}^{\infty} K(u)\, \mathrm{d}u = 1, \tag{20}$$

where X_n ($n = 1, \dots N$) are the n available data points. Each data point is thus simply replaced by a "bump". The *kernel function K* determines the shape of the bumps while the *bandwidth h* (also called *smoothing parameter*) determines their width[2].

---

[2] In the regression problem, the Nadaraya-Watson estimator (Nadaraya 1964; Watson 1964) is commonly used:

$$g(x) = \frac{\displaystyle\sum_{n=1}^{N} g(X_n) K(\frac{x - X_n}{h})}{\displaystyle\sum_{n=1}^{N} K(\frac{x - X_n}{h})}.$$

In the *adaptive kernel* version, a local bandwidth $h_n = h(X_n, f)$ is defined and used in Eq. (19). In order to follow the "true" underlying function in the best possible way, the amount of smoothing should be small when f is large whereas more smoothing is needed where f takes lower values. A convenient method to do so consists in deriving first a pilot estimate $\tilde{f}$ of f, e.g. by using an histogram or a kernel with fixed bandwidth h_{opt}, and then by defining the local bandwidths

$$h_n = h(X_n) = h_{\mathrm{opt}}[\tilde{f}(X_n)/s]^{-\alpha}, \tag{21}$$

where

$$\log s = \frac{1}{N} \sum_{n=1}^{N} \log \tilde{f}(X_n). \tag{22}$$

It may be shown (Abramson 1982) that $\alpha = 1/2$ leads to a bias of smaller order than for the fixed h estimate.

The optimum kernel K may be taken as the one minimizing the integrated mean square error beween f and $\hat{f}$ (MISE), where

$$\mathrm{MISE}(\hat{f}) = E \int \left[\hat{f}(x) - f(x)\right]^2 \mathrm{d}x \tag{23}$$

is usually taken as an indicator of efficiency. In the above expression, E denotes the statistical expectation. This optimum kernel is the so-called Epanechnikov (K_{e}) or quadratic kernel:

$$K_{\mathrm{e}}(u) = \frac{3}{4}(1 - u^2), \quad |u| < 1. \tag{24}$$

Other choices of K differ only slightly in their asymptotic efficiencies and could be more adapted to particular purposes as e.g., for bi-dimensional data.

The pilot smoothing length (h_{opt}) is the only subjective parameter required by the method. It relates to the quality of the sampling of the variable under consideration. There are several ways for automatically estimating an optimum value of h_{opt} (see e.g. Silverman 1986 for an extensive review). A simple approach based on the data variance gives in our case $h_{\mathrm{opt}} \simeq 1.0\, M_{\mathrm{J}}$. As the derivative of the frequency function rather than the function itself is actually used in Eq. (6), a larger pilot smoothing length ($2\, h_{\mathrm{opt}}$) was also considered in order to remove spurious small-statistics fluctuations of the density estimate.

References

Abramson, I. S. 1982, Ann. Statist., 10, 1217
Butler, R. P., Marcy, G. W., Fischer, D. A., et al. 1999, ApJ, 526, 916
Cerf, N., & Boffin, H. M. J. 1994, Inverse Problems, 10, 533
Chandrasekhar, S., & Münch, G. 1950, ApJ, 111, 142
Charbonneau, D., Brown, T. M., Latham, D. W., & Mayor, M. 2000, ApJ, 529, L45
Halbwachs, J. L., Arenou, F., Mayor, M., Udry, S., & Queloz, D. 2000, A&A, 355, 581
Han, I., Black, D., & Gatewood, G. 2001, ApJ, 548, L57
Härdle, W. 1990, Applied nonparametric regression (Cambridge University Press, Cambridge)

Lucy, L. B. 1974, AJ, 79, 745

Lupton, R. 1993, Statistics in theory and practice (Princeton University Press, Princeton)

Marcy, G. W., Butler, R. P., Fischer, D. A., Vogt, S., & Lissauer, J. 2001, ApJ, 556, 296

Mayor, M., Naef, D., Pepe, F., et al. 2000, in Planetary Systems in the Universe: Observations, Formation and Evolution, ed. A. Penny, P. Artymowicz, A.-M. Lagrange, & S. Russell, ASP Conf. Ser., in press

Nadaraya, E. A. 1964, Theor. Probab. Appl., 10, 186

Pourbaix, D. 2001, A&A, 369, L22

Pourbaix, D., & Arenou, F. 2001, A&A, 372, 935

Richardson, W. H. 1972, J. Opt. Soc. Am., 62, 55

Scott, D. W. 1992, Multivariate Density Estimation (Wiley, New York)

Silverman, B. W. 1986, Density estimation for statistics and data analysis, Monographs on Statistics and Applied Probability (London: Chapman and Hall)

Tabachnik, S., & Tremaine, S. 2001, AJ, in press

Udry, S., Mayor, M. 2001, in Astrobiology, Lect. Notes Phys. (Springer Verlag)

Udry, S., Mayor, M., Naef, D., et al. 2000a, A&A, 356, 590

Udry, S., Mayor, M., & Queloz, D. 2000b, in Planetary Systems in the Universe: Observations, Formation and Evolution, ed. A. Penny, P. Artymowicz, A.-M. Lagrange, & S. Russell, ASP Conf. Ser., in press

Watson, G. S. 1964, Sankhya, Ser. A, 26, 359

Zapatero Osorio, M. R., Bejar, V. J. S., Martin, E. L., et al. 2000, Science, 290, 103

Zucker, S., & Mazeh, T. 2001, ApJ, in press