

# On orbital elements of extrasolar planetary candidates and spectroscopic binaries

T. F. Stepinski and D. C. Black

Lunar and Planetary Institute, 3600 Bay Area Blvd., Houston, TX 77058, USA

Received 29 November 2000 / Accepted 9 March 2001

**Abstract.** We estimate probability densities of orbital elements, periods and eccentricities, for the population of extrasolar planetary candidates (EPC) and, separately, for the population of spectroscopic binaries (SB) with solar-type primaries. We construct empirical cumulative distribution functions (CDFs) in order to infer probability distribution functions (PDFs) for orbital periods and eccentricities. We also derive a joint probability density for period-eccentricity pairs in each population. Comparison of respective distributions reveals that in all cases EPC and SB populations are, in the context of orbital elements, indistinguishable from each other to a high degree of statistical significance. Probability densities of orbital periods in both populations have  $\sim P^{-1}$  functional form, whereas the PDFs of eccentricities can be best characterized as a Gaussian with a mean of about 0.35 and standard deviation of about 0.2 turning into a flat distribution at small values of eccentricity. These remarkable similarities between EPC and SB must be taken into account by theories aimed at explaining the origin of extrasolar planetary candidates, and constitute an important clue as to their ultimate nature.

**Key words.** stars: binaries: spectroscopic – stars: low-mass, brown dwarfs – stars: planetary systems

## 1. Introduction

Surveys of nearby solar-type stars using high accuracy spectroscopy have revealed so far (as of October 2000) 49 objects showing periodic residual radial motion of very low amplitude. Quantities directly provided by spectroscopy are period,  $P$ , and semi-amplitude,  $K$ . The radial velocity data form a time series that, in general, departs from a purely sinusoidal form. A measure of this departure is the third quantity supplied by spectroscopy.

The most straightforward interpretation of periodic variations in stellar motion is that the perturbation is due to the presence of an unresolved companion. Within such an interpretation, the departure of the motion from a sinusoidal form is parametrized by the orbital eccentricity,  $e$ . The mass function,  $f(M_\star, M)$ , can be related to  $K$ ,  $P$  and  $e$ ,

$$\begin{aligned} f(M_\star, M) &= \frac{(M \sin i)^3}{(M_\star + M)^2} \\ &= 1.0877 \cdot 10^{-13} K^3 P (1 - e^2)^{3/2} \end{aligned} \quad (1)$$

where  $M_\star$  is the mass of the primary,  $M$  the mass of the companion, and the inclination angle  $i$  is that between an

observer's line-of-sight to a star and the normal to the orbital plane of the companion/star system. In the above formula  $P$  is in days,  $K$  in  $\text{m s}^{-1}$ , and mass function is in units of Jupiter mass,  $M_J$ . For the 49 objects mentioned above the values of  $f(M_\star, M)$  are  $10^{-3}$ – $10^{-6} M_J$ . If we further assume that the companions are much less massive than the  $\sim 10^3 M_J$  solar-type primary, *projected* mass estimates of  $M \sin i$  of the order of  $1 M_J$  result.

The true mass of any particular companion remains unknown, subject to determination of its orbital inclination. Statistically,  $\langle M \rangle = 4/\pi \langle M \sin i \rangle$  (Chandrasekhar & Münch 1950) *assuming* that inclination angles in a sample are distributed isotropically. Classification of the aforementioned companions as extrasolar planetary candidates (EPC) is based primarily on an expectation that the distribution of  $i$  in the stellar samples chosen to monitor for the presence of low-mass companions is not biased and follows the isotropic distribution. However, Han et al. (2001), who used Hipparcos astrometric data to estimate  $i$  for 30 out of 49 EPC, have found the distribution of  $i$  to be far from isotropic and highly skewed toward small values of  $i$ . Although their estimation has to be treated as preliminary due to the noisy data, it nevertheless underscores the fact that the true masses of these companions are unknown and that values of  $M \sin i$  are not unambiguous indicators of their true nature.

Send offprint requests to: T. Stepinski,  
e-mail: tom@lpi.usra.edu

Within the unresolved companion interpretation of spectroscopic data, orbital periods and eccentricities are directly provided by the data and have no ambiguities associated with their values. The orbital elements of EPC have received much less attention than their projected masses because they cannot attest to the low-masses of these companions. Interestingly, even a cursory survey of EPC orbital elements reveals what must be considered as peculiar properties for objects presumed to be giant planets. Over half of the 49 EPC are in orbits with a period shorter than that of Mercury (88 days), and almost a quarter are in orbits with a period shorter than 10 days. Half of the EPC are in orbits with eccentricities larger than the eccentricity of Pluto ( $e > 0.25$ ). Furthermore, the EPC located far enough from their primaries to avoid tidal circularized are found in significantly eccentric orbits.

It is clear that the orbital elements of EPC carry significant information that, together with information contained in their projected masses and future astrometric data, can help to reveal the true nature of these companions. Toward that end, we have carried out calculations aimed at estimating distributions of orbital elements in a population of EPC. Our objective is twofold. First, using data from an available sample of EPC, we estimate univariate probability distribution functions (PDFs) for orbital periods and orbital eccentricities in the EPC population. We also estimate a bivariate PDF for period-eccentricity pairs in the EPC population. Second, we compare the estimated distributions to analogous distributions estimated for the population of spectroscopic binaries (SB). Such a comparison is called for and desirable inasmuch as EPC are technically single-line spectroscopic binaries, notable for their ultra-low  $f(M_*, M)$  values.

Our available EPC sample consist of 49 objects selected by applying an arbitrary limit of  $M \sin i \leq 13 M_J$ . We have collected the orbital elements data from original discovery papers, too numerous to reference here. All the data with pointers to the original source can be find on the internet at [cfa-www.harvard.edu/planets/](http://cfa-www.harvard.edu/planets/). We expect that this sample is complete enough (in sampling orbital elements but perhaps not the projected masses) to be representative, because the entire range of values of periods (up to  $P \approx 2500$  days) and eccentricities is accessible to all surveys. Our sample of SB consist of 330 objects (single-lined as well as double-lined spectroscopic binaries) found in a survey for spectroscopic binaries in a large (3347 stars) sample of G dwarfs (Udry et al. 1998). Only a subset of this sample defined by  $P \leq P_{\max}$  and  $e \leq e_{\max}$ , where  $P_{\max} = 2500$  days and  $e_{\max} = 0.72$  are maximum values of period and eccentricity in the EPCs sample, are used in our calculations. According to Udry et al., within these limits the sample is fairly complete and should be representative of the corresponding SB sub-population.

The previous work on distribution of orbital elements in EPC is limited. Heacox (1999) compared statistically low-mass companions to stellar-mass secondaries. Heacox's 28 low-mass companions consisted of all then known companions with  $M \sin i \leq 70 M_J$ . Ten of these

companions with  $17 M_J \leq M \sin i \leq 70 M_J$  have been classified as brown dwarfs companions rather than EPC. Stellar-mass companions had been chosen from Duquennoy & Mayor (1991) sample to have primary spectral types and orbital scales similar to low-mass companions objects. There are 15 such companions in the Duquennoy & Mayor sample. Heacox constructed the empirical cumulative distribution function (CDF) for various orbital properties and postulated that the actual CDF for any given property is the best-fit power-law function. He assessed the significance of the postulated distribution by means of Kolmogorov-Smirnov statistics. Stepinski & Black (2000a) performed a similar study, using the same data, but focusing on the issue of whether EPC and brown dwarfs companions came from the same population. Finally, Mazeh & Zucker (2000) compared, by means of empirical CDFs, orbital elements of EPC and stellar secondaries based on the sample of 32 EPC and 59 single-lined spectroscopic binaries.

Our current work differs from the previous efforts in a number of aspects. First, in addition to constructing empirical CDFs we actually calculate the *nonparametric* estimates of underlying PDFs for orbital periods and eccentricities. Second, we calculate bivariate PDFs for period-eccentricity pairs and devise a method to assess the significance of similarity between bivariate PDFs calculated for EPC and SB. Third, samples used for our estimations are much larger than those used in the previous studies. Preliminary results of our study, based on 32 EPC and 15 SB taken from the Duquennoy & Mayor sample, has been published in Stepinski & Black (2000b).

Section 2 contains a description of our method to estimate the underlying PDF. In Sects. 3 and 4 we present results pertaining to orbital periods and eccentricities, respectively. Section 5 is devoted to bivariate PDFs for period-eccentricity pairs. Finally, in Sect. 6 we present conclusions and discussion.

## 2. Estimating univariate PDFs

The major goal of this paper is to estimate the underlying PDFs of orbital periods and eccentricities for EPC and SB populations. Traditionally, histograms are used to estimate the PDF underlying a given sample of data. However, constructing a histogram involves binning the data and results in a loss of information. In addition, there is a considerable arbitrariness as to how the bins should be chosen.

A more reliable result is obtained by first constructing the *empirical* CDF,  $\mathcal{F}(X)$ , of quantity  $X$ , being, in our case, either an orbital period or an eccentricity. The second step is to use  $\mathcal{F}(X)$  in order to obtain a smooth approximation,  $F(X)$ , of an actual CDF. The smoothness is a desired property of a CDF inasmuch as one expects orbital elements to be distributed in a continuous fashion. The major technical issue here is to use an algorithm that would achieve an appropriate trade-off between matching

the data and achieving smoothness. The final step is to obtain an estimate of PDF for quantity  $X$ ,  $f(X) = \partial F / \partial X$ .

To find  $F(X)$  we used the method described by Gershenfeld (1999). Let  $\{x_i\}_{i=1}^N$  be a set of  $N$  observed values of quantity  $X$ , *normalized* to interval  $[0, 1]$  and *sorted* in increasing order. We construct a set of triplets  $\{(x_i, y_i, \sigma_i)\}_{i=1}^N$  where  $y_i$  denotes the value of  $\mathcal{F}(x_i)$  and  $\sigma_i$  denotes an estimate of the error of  $y_i$ .  $\mathcal{F}(x)$  is defined to be the fraction of periods below  $x$ .

$$\mathcal{F}(x) = \begin{cases} 0, & x < x_1 \\ \frac{i}{N+1}, & x_i \leq x < x_{i+1}. \end{cases} \quad (2)$$

The denominator is taken to be  $N+1$  instead of  $N$  because the normalization constraint,  $\mathcal{F}(1) = 1$ , is effectively an extra point. To estimate the error,  $\sigma_i$ , consider the zero-one discrete distribution, with  $\mathcal{F}(x_i)$  the probability to draw a quantity  $x$  below its  $x_i$  value and  $[1 - \mathcal{F}(x_i)]$  the probability to draw a quantity  $x$  at or above its  $x_i$  value. The variance of such distribution is  $\mathcal{F}(x_i)[1 - \mathcal{F}(x_i)]$ , so the standard deviation of variable  $y_i = i/(N+1)$  is,

$$\sigma_i = \frac{1}{N+1} (N\mathcal{F}(x_i)[1 - \mathcal{F}(x_i)])^{1/2}. \quad (3)$$

The smooth estimate of an underlying CDF,  $F(x)$ , is expressed by the following, convenient form,

$$F(x) = x + \sum_{i=1}^N a_i \sin(i\pi x). \quad (4)$$

Coefficients  $\mathbf{a} = \{a_i\}_{i=1}^N$  are determined by fitting (4) to the data set  $\{(x_i, y_i, \sigma_i)\}_{i=1}^N$  and  $f(x)$  is simply equal to  $\partial F / \partial x$ . Note that, from the formal point of view, we perform a parametric estimate, because the functional form of  $F(x)$  is fixed by Eq. (4), however, the estimate is *de facto* nonparametric because parametrization (4) can represent all feasible CDFs. The only constraint, of no consequence to our analysis, imposed on  $f(x)$  by representation (4) is that  $\partial f / \partial x = 0$  at  $x = 0$  and  $x = 1$ , the PDF is *locally* flat at the limits of its applicability.

To obtain a fit we minimize the following ‘‘cost’’ function,  $C(\mathbf{a})$ ,

$$C(\mathbf{a}) = \int_0^1 \left( \frac{d^2 F}{dx^2} \right)^2 dx + \lambda \frac{1}{N} \sum_{i=1}^N \left[ \frac{F(x_i, \mathbf{a}) - y_i}{\sigma_i} \right]^2. \quad (5)$$

The second term is the  $\chi^2$  expression which, by itself, is minimized (to  $\chi^2 = 0$ ) when  $F(x)$  passes through all data points, and the first term, the regularizer, reflects our expectations about smoothness of  $F(x)$ . Specifically, we postulate that  $F(x)$  should be locally linear; this is measured by the integral of the square of the curvature of  $F(x)$ . The parameter  $\lambda$  (Lagrange multiplier) controls a trade-off between matching the data and enforcing our prior belief about smoothness. Large values of  $\lambda$  overemphasize the data and would result in *overfitting* whereas small values of  $\lambda$  give extra smoothing and would result in *underfitting*. Minimizing  $C(\mathbf{a})$  by solving  $\partial C / \partial \mathbf{a}$  for a particular value of  $\lambda$  gives a set of equations to be solved to find the

corresponding  $\mathbf{a}$ . The best value for  $\lambda$  can then be found by noting that the expected value of  $\chi^2$  (the second term in Eq. (5)) is 1. An estimate of  $f(x)$  obtained using this method is continuous and smooth and offers an excellent fit to the data.

We calculate a significance of our estimated CDF (and thus PDF) using the Kolmogorov-Smirnov (thereafter referred to as K-S) statistic. The K-S statistic,  $D$ , is defined by the maximum value of the absolute difference between  $\mathcal{F}(x)$ , and  $F(x)$ ,

$$D = \max | \mathcal{F}(x) - F(x) |. \quad (6)$$

The value of  $D$  is a measure of ‘‘distance’’ between empirical and estimated CDFs. A desirable feature of  $D$  is that it is invariant under any continuous transformation to or from  $x$ , as distorting the horizontal axis cannot change the value of vertical distance. Of particular relevance to our work is that using  $F(\log x)$  instead of  $F(x)$  yields exactly the same value of  $D$ . In addition, there is no need to transform  $x$ , a normalized quantity, to  $X$ , an actual quantity, for the purpose of calculating  $D$ .

The distribution of random variable  $D$  is used to assess the significance level,  $s$ , the probability that given an observed value of  $D = D_{\text{obs}}$ , an estimated CDF is *falsely* rejected. We use the following approximate formula (Press et al. 1992) for  $s$ ,

$$s = P(D > D_{\text{obs}}) = Q_{\text{KS}} \left( \left[ \sqrt{N} + 0.12 + \frac{0.11}{\sqrt{N}} \right] D_{\text{obs}} \right) \quad (7)$$

where

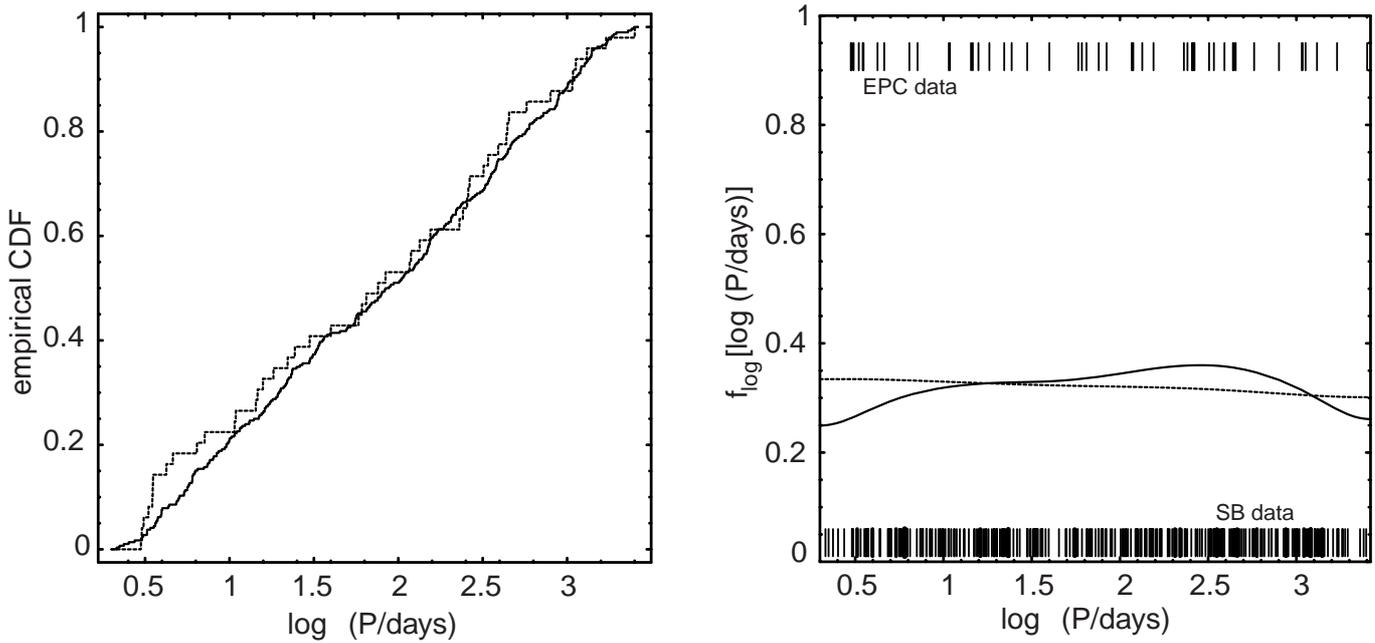
$$Q_{\text{KS}}(z) = 2 \sum_{j=1}^{\infty} (-1)^{j-1} \exp(-2j^2/z^2). \quad (8)$$

Small values of  $s$  indicate that  $F(x)$  is unlikely to be the true CDF underlying a given sample. Values of  $s$  calculated for  $F(x)$  obtained by the method described in this section are guaranteed to be high. Note that because  $F(x)$  depends on parameters estimated from the sample, the distribution of  $D$  may depart slightly from that given by the right hand side of Eq. (7), therefore calculated values of  $s$  should be treated as estimations.

We also calculate values of  $D$  and corresponding values of  $s$  for some a priori given distributions, for example, power-law or Gaussian distributions. In addition, the variant of the K-S statistic is used to assess whether two empirical CDFs reflect the same underlying distribution (see Press et al. 1992, for details).

### 3. Distributions of orbital periods

Periods in the EPC sample are in the range of 2.99 to 2518 days, whereas periods in the SB sample are in the range of 0.27 to 30348 days. It is clear that SB are much more extensively covered than EPC, therefore, in order for our comparison to be meaningful, we need to consider only binaries whose periods are commensurable with periods of extrasolar planets. Specifically, we only keep SB with



**Fig. 1.** Estimations of orbital period distribution in populations of EPC (dotted lines) and SB (solid lines). The left panel shows empirical CDFs of  $\log(P)$  and the right panel shows estimated PDFs of  $\log(P)$ . Vertical bars indicate data

2 days  $< P < 2500$  days. There are 292 such binaries. Thus, our estimated PDFs pertain only to sub-populations of SB and EPC limited to the aforementioned range of periods.

Values of periods in the samples under consideration span 3 orders of magnitude, therefore it is more convenient to estimate a PDF of a variable  $\log P$ ,  $f_{\log}(\log P)$ , instead of a PDF of a variable  $P$ ,  $f(P)$ . We use logarithms to base 10 throughout the paper. The  $f(P)$  can be easily recovered from  $f_{\log}(\log P)$  by the following transformation

$$f(P) = f_{\log}(\log P) \left| \frac{\partial \log P}{\partial P} \right| = \log e \frac{f_{\log}(\log P)}{P}. \quad (9)$$

Figure 1 (left panel) shows empirical CDFs of variable  $\log P$  constructed from EPC and SB periods, respectively. The K-S distance,  $D$ , between these CDFs is  $D = 0.1$ , and the corresponding significance is  $s = 76\%$ . Thus it is quite likely that two lists of  $\log P$  values, the one from the EPC sample, and the other from the SB sample, are drawn from the same population.

Next, we calculate the smooth estimates of  $F_{\log}(\log P)$ , as described in Sect. 2, based on EPC and SB data respectively. We choose the fits with such values of  $\lambda$  as to yield  $\chi^2$  between 1 and  $1/3$ ; slightly overemphasizing the data with respect to the formally best balance given by  $\chi^2 = 1$ . These fits have significance of 96% and 95% for EPC data and SB data, respectively. Figure 1 (right panel) shows the estimated  $f_{\log}(\log P)$  for both populations. The most interesting feature of these PDFs is that, to the first approximation, they both represent a flat (uniform) distribution.

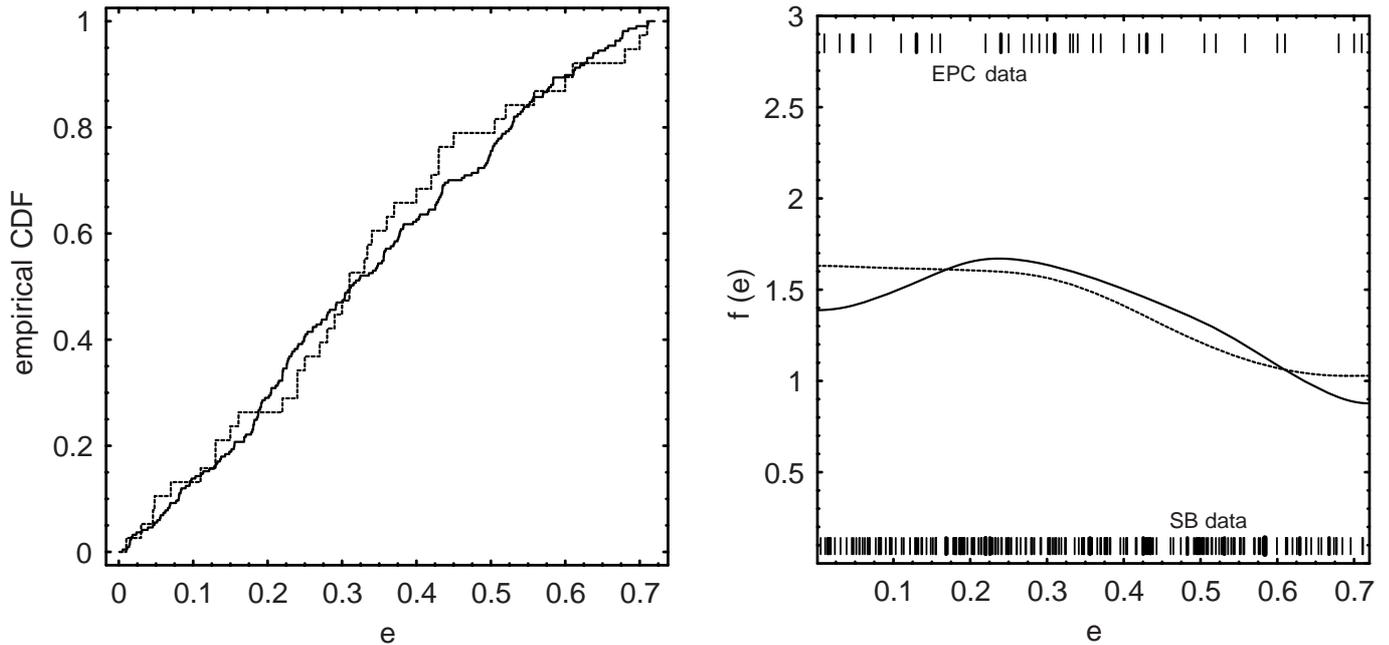
Using transformation (9) we can obtain  $f(P)$  for populations of EPC and SB. The  $f(P)$  can be viewed as

the  $P^{-1}$  distribution *modulated* by  $f_{\log}(\log P)$  in the numerator. Because, for both EPC and SB populations,  $f_{\log}(\log P)$  is a very slowly varying function of  $P$ ,  $f(P)$  for these populations is well approximated by the  $P^{-1}$  distribution. In fact, a direct parametric fit of a power-law to the period data, gives  $f(P) \sim P^{-1.02}$  with significance of 97% for EPC, and  $f(P) \sim P^{-0.99}$  with significance of 53% for SB.

#### 4. Distributions of orbital eccentricities

Orbital eccentricities in the EPC sample are in the range of 0 to 0.71, whereas eccentricities in the SB sample are in the range of 0 to 0.97. The presence of 12 SB with  $0.71 < e < 1$  and the absence of any EPC in that range may indicate a genuine difference between the two populations, but it also may be due to a particular realization of the much smaller EPC sample. Because of this latter possibility and due to the lack of completeness of the SB sample at high values of eccentricity, we base our main calculation on SB whose eccentricities are within the same range as those of EPC. Additional calculations, based on the entire range of eccentricities, are reported at the end of this section. Further, we are only interested in companions in uncircularized orbits, those located not too close to a primary. Thus, we only use EPC and SB with  $e < 0.72$  and  $P_{\text{circ}} < P < 2500$  days. The value of  $P_{\text{circ}} = 10$  days is assumed based on theoretical estimations (Zahn 1977). There are 38 EPC and 217 SB satisfying these conditions. The estimated PDFs pertain only to sub-populations of SB and EPC bounded by the above conditions.

Unlike the case of orbital periods, where we find it convenient to estimate a PDF of  $\log P$  and obtain a PDF of  $P$  by means of transformation (9), a PDF of eccentricities



**Fig. 2.** Estimations of orbital eccentricity distribution in populations of EPC (dotted lines) and SB (solid lines). The left panel shows empirical CDFs of  $e$  and the right panel shows estimated PDFs of  $e$ . Vertical bars indicate data

is estimated straightforwardly. Figure 2 (left panel) shows empirical CDFs of  $e$  constructed from eccentricities of eligible EPC and SB, respectively. The K-S distance,  $D$ , between these CDFs is  $D = 0.1$ , and the corresponding significance is  $s = 89\%$ . Note that even though the value of  $D$  happens to be the same as for empirical CDFs constructed for periods (see previous section), the significance is larger due to smaller number of data points. Thus, it is very likely that two lists of  $e$  values, the one from the EPC sample, and the other from the SB sample, are drawn from the same population.

The process of obtaining  $F(e)$  and  $f(e)$  is analogous to that described in the previous section; again values of  $\lambda$  are chosen to yield  $\chi^2$  between 1 and  $1/3$ . The obtained smooth estimations of CDFs have significance of 85% for both EPC and SB data. Figure 2 (right panel) shows the estimated  $f(e)$  for both populations. These PDFs have a similar form, that can be characterized as a combination of a Gaussian and a flat distribution. Approximately, 20% of the population, those companions with eccentricities smaller than about 0.15, have a flat distribution, whereas the rest are distributed according to a Gaussian with a mean,  $\mu \approx 0.2$ – $0.3$  and, standard deviation,  $h \approx 0.25$ – $0.3$ . Note that our method of finding PDFs tends to slightly overestimate variance because of the presence of a regularizer (see Eq. (5)) which reduces the curvature of  $f(x)$ .

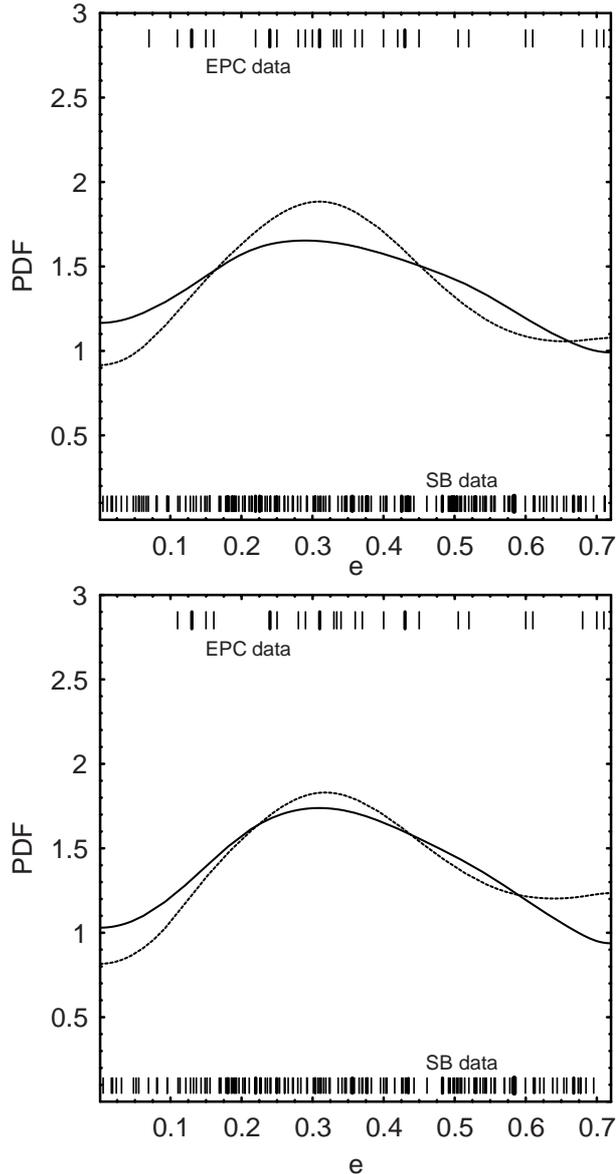
We hypothesize that the fundamental distribution of orbital eccentricities in populations of EPC and SB is Gaussian, but that companions that are relatively close to their primaries have had their orbits evolved to have a flat distribution of  $e$ . We tested this hypothesis by estimating PDFs of  $e$  for sub-populations of companions located

progressively farther from their primaries. Figure 3 (top panel) shows estimated  $f(e)$  for the sub-populations of EPC and SB limited by 20 days  $< P < 2500$  days. There are 32 EPC and 190 SB satisfying this limit. The K-S distance,  $D$ , between empirical CDFs constructed from eligible EPC and SB companions (not shown) is  $D = 0.12$ , and the corresponding significance is  $s = 77\%$ . Both PDFs appear now to be Gaussian-like. In fact, a direct parametric fit of a Gaussian to the EPC eccentricity data gives  $\mu = 0.34$  and  $h = 0.18$  with significance  $s = 98\%$ . A direct parametric fit of a Gaussian to the SB eccentricity data gives  $\mu = 0.35$  and  $h = 0.20$  with significance  $s = 25\%$ .

Figure 3 (bottom panel) shows estimated  $f(e)$  for the sub-populations of EPC and SB limited by 50 days  $< P < 2500$  days. There are 28 EPC and 156 SB satisfying this limit. The K-S distance,  $D$ , between empirical CDFs constructed from eligible EPC and SB companions (not shown) is  $D = 0.14$ , and the corresponding significance is  $s = 74\%$ . Again, both PDFs appear to be Gaussian-like. A direct parametric fit of a Gaussian to the EPC and SB eccentricity data gives  $\mu = 0.36$  and  $h = 0.19$  with significance  $s = 96\%$  and  $\mu = 0.35$  and  $h = 0.19$  with significance  $s = 51\%$ .

It is apparent that  $f(e)$  approaches a Gaussian for a sub-population of companions located far enough from a primary. Given the samples used in our calculations, this convergence occurs faster for the EPC than for the SB.

We have also estimated PDFs of  $e$  for EPC and SB on the entire possible domain of eccentricities,  $0 < e < 1$ . The K-S distance between respective empirical CDFs is  $D = 0.13$ , and the corresponding significance is  $s = 57\%$ . Thus, consideration of the entire range of eccentricities decreases the similarity between two estimated PDFs,



**Fig. 3.** Estimations of probability densities of orbital eccentricity in populations of EPC (dotted lines) and SB (solid lines). The upper panel pertain to sub-populations limited by 20 days  $< P < 2500$  days, and the lower panel pertain to sub-populations limited by 50 days  $< P < 2500$  days. Vertical bars indicate data

but there is still a 57% chance that both samples are drawn from the same population and that the absence of  $0.71 < e < 1$  orbits in the EPC sample does not indicate that there are no such orbits in the EPC population. To make this point clear we performed the following numerical experiment. The EPC eccentricity data was enriched by 2 additional entries randomly drawn from the range  $0.72 < e < 1$ . The enriched sample has about the same fraction of objects in this high eccentricity range as the SB sample. The K-S distance between empirical CDF based on the enriched EPC sample and the empirical CDF based on the SB sample is  $D = 0.08$ , and the corresponding significance is  $s = 97\%$ .

### 5. Joint distributions of orbital periods and eccentricities

So far we have considered univariate distributions of companions' orbital elements, periods and eccentricities. However, we can also study joint (bivariate) distributions  $f(\log P, e)$ . A joint distribution contains more information than two univariate distributions unless the two variables constituting a dual variable are *independent*. Black (1997) and Stepinski & Black (2000a) studied scatter plots of the  $\log P$  and  $e$  data for low-mass companions and found positive correlation between periods and eccentricities. The existence of such a correlation can be noted by visually inspecting the scatter plot for the EPC data (the left panel of Fig. 4). Thus, at least for the population of EPC, there is a reason to believe that values of periods and eccentricities are not independent, and estimating joint distribution is called for.

Visual inspection of the SB data (the right panel of Fig. 4) reveals no obvious relations between  $\log P$  and  $e$  data, except for the high concentration of companions in short periods, low eccentricity orbits. Thus, despite great similarity between distributions of periods and, separately, eccentricities in populations of EPC and SB (see Sects. 3 and 4), the similarity between respective joint distributions is not assured and cannot be anticipated from casual inspection of the respective scatter plots. However, it is often the case that a distribution estimate will detect or highlight features that are impossible to spot on a scatter plot.

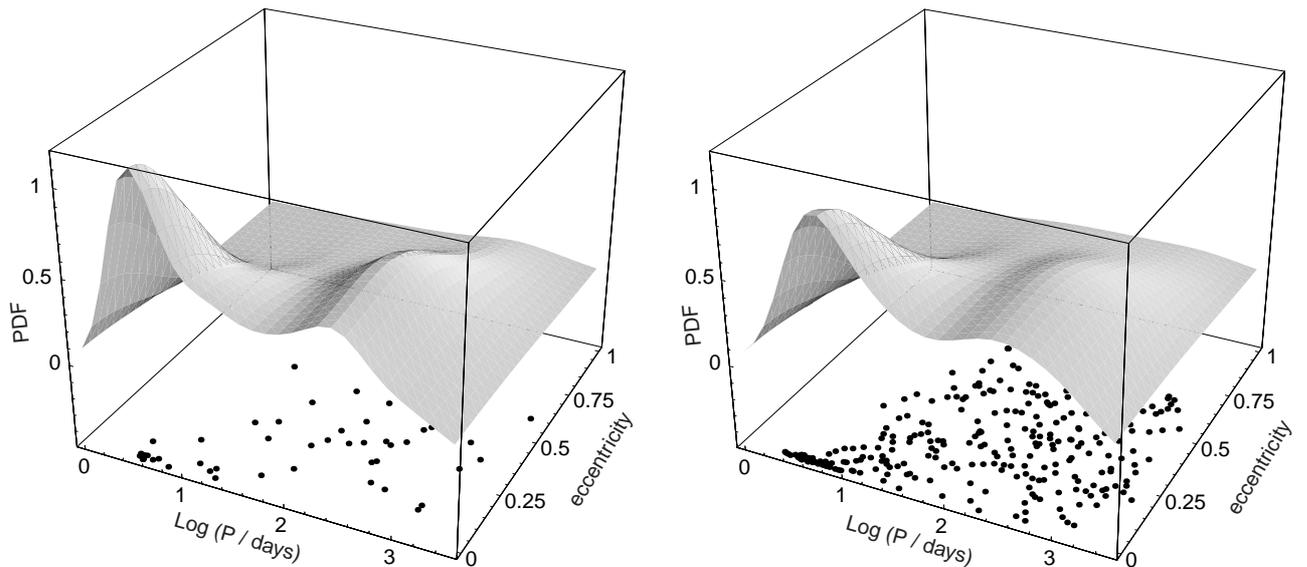
Based on the available data, we estimate PDFs of  $(\log P, e)$  for populations of EPC and SB. These estimations are based on all 49 EPC objects and 279 SB objects with  $e < 0.72$  and 2 days  $< P < 2500$  days selected for commensurability with the EPC data and thus pertain only to sub-populations of EPC and SB bounded by the above limits.

The regularization method for estimating an univariate PDF (see Sect. 2) does not generalize to the bivariate case. We use an adoptive kernel estimation method with Gaussian kernel (Silverman 1986) to infer  $f(\log P, e)$ . In this method  $f(r)$ , where  $r$  denotes a vector  $(\log P, e)$ , is approximated by an estimate,  $f_k(r)$  given by a sum of Gaussians kernels centered on the data points,

$$f_k(r) = \frac{1}{N} \sum_{i=1}^N \frac{1}{2\pi w_i^2} \exp\left(-\frac{|r_i - r|^2}{2w_i^2}\right) \quad (10)$$

where  $\{r_i\}_{i=1}^N$  are data points and  $\{w_i\}_{i=1}^N$  are the kernel widths. Finding  $f_k(r)$  is tantamount to finding optimal values of  $\{w_i\}_{i=1}^N$ , for details see Silverman (1986) or Pisani (1993, 1996).

Figure 4 shows results of our calculations. The left panel shows a surface plot depicting  $f_k^{\text{EPC}}(\log P, e)$ , an estimate of the joint distribution for the population of EPC. The right panel shows a surface plot depicting  $f_k^{\text{SB}}(\log P, e)$ , an estimate of the joint distribution for the SB population. There is a remarkable similarity, not anticipated from a visual inspection of respective scatter plots,



**Fig. 4.** Joint,  $(\log P, e)$  PDFs estimated from the EPC data (left panel) and the SB data (right panel). The PDFs are represented by the surface plots, the scale and the view point of both plots are the same. The data is displayed beneath the PDFs in the form of scatter plots

in the form of  $f_k^{\text{EPC}}(r)$  and  $f_k^{\text{SB}}(r)$ . One possible characterization of these joint distributions is that they are a superposition of two Gaussians restricted to  $e > 0$  domain.

To quantify the visually evident similarity between  $f_k^{\text{EPC}}(r)$  and  $f_k^{\text{SB}}(r)$  we calculate the value  $D_1$  defined as,

$$D_1^2 = \int [f_k^{\text{SB}}(r) - f_k(r)]^2 dr \quad (11)$$

where the integral extends over the entire domain of  $\log P$  and  $e$ . The quantity  $D_1$  (a subscript 1 is to set it apart from a quantity  $D$  given by Eq. (6)) is a measure of “distance” between  $f_k^{\text{SB}}(r)$  and any given joint distribution  $f_k(r)$ . Calculating distance between  $f_k^{\text{SB}}(r)$  and  $f_k^{\text{EPC}}(r)$  yields  $D_1^{\text{EPC}} = 0.11$ .

If the populations of EPC and SB have indeed the same joint distribution of  $(\log P, e)$ , an estimate of such a distribution based on a particular sample consisting of 49 objects is expected nevertheless to differ from an estimate based on another sample consisting of 279 objects. Thus,  $D_1$  is a random variable. We estimate the distribution of  $D_1$  from the following Monte Carlo simulation. We randomly draw 49 objects from among the 279 SBs and estimate the  $f_k(r)$  on the basis of this smaller data set. Next we calculate the distance,  $D_1$ , between such an estimation and  $f_k^{\text{SB}}(r)$ . We repeat this experiment 300 times to gather a sufficient number of  $D_1$  realizations and use methods described in Sect. 2 to infer the PDF of  $D_1$ .

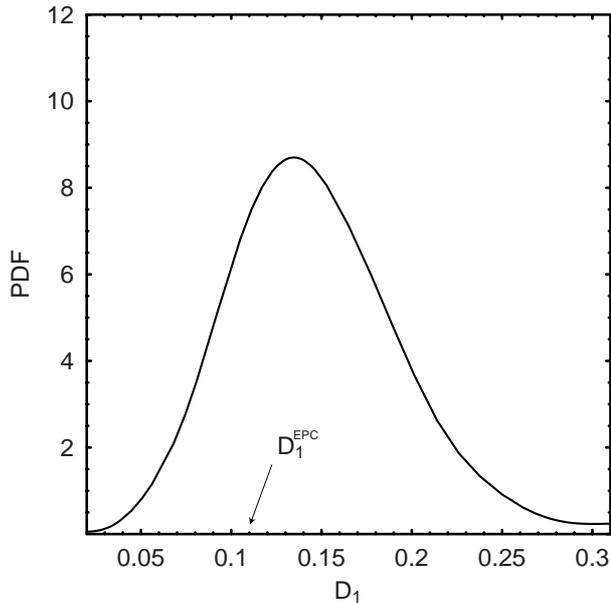
Figure 5 shows the distribution of  $D_1$  obtained from our Monte Carlo simulations. There is a 76% chance that the distance between  $f_k^{\text{SB}}(r)$  and a joint distribution estimated on the basis of only 49 objects is larger than  $D_1^{\text{EPC}} = 0.11$ . Thus the significance that  $f_k^{\text{SB}}(r)$  and  $f_k^{\text{EPC}}(r)$  are estimates of the same distribution is  $s = 76\%$ .

In order to show that the similarity between  $f_k^{\text{SB}}(r)$  and  $f_k^{\text{EPC}}(r)$  is not caused just by the fact that short period orbits in both samples are circularized, we estimated  $f_{1,k}^{\text{SB}}(r)$  and  $f_{1,k}^{\text{EPC}}(r)$ , joint distributions estimated on the basis of uncircularized objects alone, those with  $20 \text{ days} < P < 2500 \text{ days}$  and  $e < 0.72$ . Those PDFs (not shown here) have forms basically identical to  $f_k^{\text{SB}}(r)$  and  $f_k^{\text{EPC}}(r)$  restricted to the smaller  $(\log P, e)$  domain. The significance associated with these new PDFs is  $s = 75\%$ . Thus, the significance of similarity between estimated PDFs for populations of EPC and SB does not diminish when circularized orbits are removed from consideration.

Neither  $f_{1,k}^{\text{EPC}}(r)$  nor  $f_{1,k}^{\text{SB}}(r)$  are similar to hypothetical joint distributions obtained by simply multiplying respective univariate distributions of  $\log P$  (given in Sect. 3) and  $e$  (given in Sect. 4). That means that indeed periods and eccentricities are not independent in either the EPC or the SB populations. However, note that *marginalizing* estimated joint distributions with respect to either  $\log P$  or  $e$  yields respective univariate distributions (not shown here) that are consistent with those estimated in Sects. 3 and 4.

## 6. Discussion and conclusions

We have conducted a statistical analysis of orbital elements, periods and eccentricities, based on the most extensive available samples of extrasolar planetary candidates and spectroscopic binaries. The purpose of this work was to estimate, on the basis of this data, probability densities for periods, eccentricities, and the joint probability density of period-eccentricity pairs in populations of extrasolar planetary candidates and spectroscopic binaries.



**Fig. 5.** Distribution of random variable  $D_1$ , a distance between  $f_k^{\text{SB}}(r)$  and the bivariate PDF estimated on the basis of 49 SBs draw randomly from the entire SB sample. The arrow indicates  $D_1^{\text{EPC}} = 0.11$ , the distance between  $f_k^{\text{SB}}(r)$  and  $f_k^{\text{EPC}}(r)$

We also compare respective PDFs between the two populations. Our findings can be summarize as follows:

PDFs of orbital periods in both populations are very well approximated by the power-law,  $f(P) \sim P^{-1}$ . Note that the estimates are valid only for sub-populations restricted by the condition  $2 \text{ days} < P < 2500 \text{ days}$ .

It may appear that our estimation of  $f(P)$  for stellar companions is at odds with that of Duquennoy & Mayor (1991) who concluded that orbital periods of stellar companions to G-dwarfs stars in the solar neighbourhood are distributed according to the log-normal distribution (normal distribution of variable  $\log P$ ). However the Duquennoy & Mayor analysis pertains to the complete sample of binaries and is not restricted to relatively short period spectroscopic binaries. Estimation of  $f(P)$  (using methods of Sect. 2) on the basis of the Duquennoy & Mayor data set truncated to conform to condition  $P < 2500 \text{ days}$  yields  $f(P) \sim P^{-0.9}$  with significance  $s = 95\%$ , in excellent agreement with our estimation based on the much larger data set. Interestingly, the log-normal distribution estimated by Duquennoy & Mayor on the basis of the the *complete* sample is itself not markedly different from  $P^{-1}$  distribution as was first noted by Heacox (1996). Our finding that orbital periods of EPC are also distributed according to  $f(P) \sim P^{-1}$  is consistent with earlier results of Heacox (1999) and Stepinski & Black (2000a) obtained on the basis of much smaller data sets.

PDFs of orbital eccentricities in both populations can be, in general, approximated by the Gaussian with  $\mu \approx 0.35$  and  $h \approx 0.2$ , providing that only companions on

orbits with periods longer than the circularization period,  $P_{\text{circ}} \approx 10 \text{ days}$ , are used in the estimation process.

This is consistent with the eccentricity distribution reported by Duquennoy & Mayor for stellar binaries with periods  $P_{\text{circ}} < P < 1000 \text{ days}$ . However, in contrast to Duquennoy & Mayor findings, our calculations show that companions, both EPC and SB, in orbits with  $P_{\text{circ}} < P < 20 \text{ days}$  tend to have predominantly small but nonzero eccentricities. Thus, although the PDFs estimated on the basis of samples that exclude  $P_{\text{circ}} < P < 20 \text{ days}$  orbits have a purely Gaussian form, the PDFs estimated on the basis of samples that include  $P_{\text{circ}} < P < 20 \text{ days}$  orbits have a Gaussian form turning into a flatter distribution at small values of eccentricity. This suggests that the orbits of companions located just beyond the circularization orbit are nevertheless somehow altered by the proximity of the primary.

Earlier efforts (Heacox 1999; Stepinski & Black 2000a) to estimate PDFs of orbital eccentricity in populations of low-mass companions and control population of stellar binaries were based on the samples *not* truncated to exclude circularized orbits. This was dictated by the small size of samples then available their further reduction would leave too few objects to reliably estimate a PDF. In those studies, the authors reported  $f(e) \sim e^{-0.5}$  (Heacox) or  $f(e) \sim e^{-0.63}$  (Stepinski & Black) for both populations. Our current point of view is that circularized orbits should be excluded from consideration as they are not pertinent to the primordial distribution of orbital eccentricities. Nevertheless, for the sake of comparison with these earlier calculations, we have estimated PDFs of orbital eccentricity on the basis of EPC and SB samples that include all circularized orbits. Estimated PDFs are very similar to each other, evidence that EPC and SB samples came from the same population is  $s = 97\%$ . The global character of estimated PDFs is roughly consistent with a power-law form, but both PDFs have local maxima at around  $e = 0.3$ . Direct parametric fit of a power-law gives  $f(e) \sim e^{-0.48}$  with  $s = 67\%$  for EPC and  $f(e) \sim e^{-0.44}$  with  $s = 15\%$  for SB.

Joint distributions of period-eccentricity pairs in a population of companions have previously not been estimated. Duquennoy & Mayor (1991) constructed a scatter plot of  $(\log P, e)$  data for their complete sample of binaries. Similar scatter plots have been published by Black (1997), Heacox (1999) and Stepinski & Black (2000a) in the context of low-mass companions. Our estimations of  $f(\log P, e)$  for EPC and SB populations yield very similar outcomes. The form of the estimated PDF can be perhaps characterized as a superposition of two bivariate Gaussians. The narrower Gaussian, cut by the  $e > 0$  requirement, accounts for a high number of circularized companions and can be easily anticipated from the scatter plot. The broader Gaussian roughly centered at  $\log P = 2.5$  and  $e = 0.3$  cannot be anticipated from the scatter plot. It is only through a density estimation technique that this feature becomes apparent. The joint

distributions differ from what would be obtained by a simple product of independently estimated univariate distributions of  $\log P$  and  $e$ . This shows that periods and eccentricities are not independent variables, but are indeed somewhat related. This relation is most likely not causal, but rather a statistical trend, perhaps a reflection of what appears to be a stochastic mechanism responsible for formation of those companions.

Our calculations demonstrate a remarkable similarity between orbital parameters of extrasolar planetary candidates and stellar companions in spectroscopic binaries. All our tests designed to compare the two populations repeatedly yield the same result – the available samples of EPC and SB orbits come from the same distribution with a high degree of significance. Based on this result we claim that, in the context of orbital elements, the two populations are indistinguishable from each other.

One possible corollary to the above claim is that majority of EPC are indeed stellar companions viewed at very small inclination angles (Kubát et al. 1998; Imbert & Prévot 1998). Although such a corollary readily explains our findings, it also seems at odds with expected isotropy of orbits' inclination angles. However, the isotropic distribution of  $i$  in the *available sample* of EPC is not yet a foregone conclusion (see comments in Sect. 1), and the aforementioned corollary cannot be flatly rejected.

If values of  $i$  in the sample of EPC are distributed isotropically, the majority of EPC are indeed objects with masses of the order of  $1 M_J$ . We know with certainty the existence of one such object, HD 209458, for which a photometric transit has been observed (Charbonneau et al. 2000) yielding  $i = 87^\circ$  and  $M = 0.63 M_J$ . We caution however that the current masses of EPC in short-period orbits are likely to be substantially less than their masses at formation due to Roche lobe overflow (Trilling et al. 1998).

Companions with masses  $\sim M_J$  are commonly thought to form via a process analogous to that postulated for the formation of Jupiter in the Solar System – buildup by aggregation from a protoplanetary disk. This formation mechanism builds planets in circular orbits with  $P$  of about 10 years, thus subsequent “post-processing” must be invoked to modify orbits to shapes indicated by spectroscopic observations. Two types of mechanisms have been proposed, gravitational interactions between a disk and a planet that leads to inward planet migration (for example see Trilling et al. 1998) and possibly eccentricity change (Artymowicz et al. 1998), and multiplanet interactions (Weidenschilling & Marzari 1996; Lin & Ida 1997). Our present findings show a potential means for validating these various hypothesis. Inasmuch as the statistical properties of EPC orbits are now well determined, any proposed mechanism of orbit alteration, when applied to an ensemble of stars with just-formed planets, should reproduce probability densities of orbital elements as given in this paper.

If the EPC are indeed planets (in the sense described in the previous paragraph) then the most interesting

questions stemming from our findings is why their orbital elements are distributed like those of the stellar companions. The orbital elements of stellar binaries are acquired during their initial birth as small fragments of collapsing molecular cloud core, but are subject to the post-formation evolution due to accretion from an infalling gaseous envelope (Bate 2000), as well as further evolution due to possible presence of circumstellar and/or circumbinary disks (for review see Artymowicz & Lubow 2000). Thus, post-formation evolution plays a role in determining the statistics of orbital elements in both stellar binaries and planets. Future research may show that this is indeed a sufficient “common factor” to account for statistically identical orbits in otherwise very different populations. However, in our opinion, it is more likely that similar distribution of orbits arose from essentially the same formation/evolution context. Following this reasoning, some consideration should be given to a possibility that EPC formed in a way somewhat analogous to that responsible for the formation of stellar companions. Thus, EPC could be “planets” in the sense of having masses in the same range as the masses of giant planets in the Solar System, but not in the sense of originating by aggregation from a protoplanetary disk, and therefore their relevance to planets in the Solar System is unclear.

In summary, we have established distributions of orbital elements in the population of EPC and demonstrated that they are identical to corresponding distributions in the population of stellar companions. This offers a new insight into the issue of the ultimate nature of EPC. Our findings generate some doubt about the prevalent interpretation of EPC as simple analogues of giant planets in the Solar System. In this context it is crucial to get a better hold on the values of the true masses of EPC.

*Acknowledgements.* We are grateful to the anonymous referee for valuable comments on an earlier draft of this paper. This research was conducted at the Lunar and Planetary Institute, which is operated by the Universities Space Research Association under contract No. NASW-4574 with the National Aeronautics and Space Administration. This is Lunar and Planetary Institute Contribution No. 1086.

## References

- Artymowicz, P., Lubow, S. H., & Kley, W. 1998, in *Planetary Systems - the long view*, ed. L. Celnikier et al. (Éditions Frontières)
- Artymowicz, P., & Lubow, S. H. 2000, to appear in *The Formation of Binary Stars*, ed. H. Zinnecker, & R. D. Mathieu, ASP Conf. Ser., 200
- Bate, M. R. 2000, *MNRAS*, 314, 33
- Black, D. C. 1997, *ApJ*, 490, L171
- Chandrasekhar, S., & Münch, G. 1950, *ApJ*, 111, 142
- Charbonneau, D., Brown, T. M., Latham, D. W., & Mayor, M. 2000, *ApJ*, 529, L45
- Duquenois, A., & Mayor, M. 1991, *A&A*, 248, 485

- Gershenfeld, N. 1999, *The Nature of Mathematical Modeling* (Cambridge: Cambridge University Press)
- Han, I., Black, D. C., & Gatewood, G. 2001, *ApJ*, 548, L57
- Heacox, W. D. 1996, *PASP*, 108, 591
- Heacox, W. D. 1999, *ApJ*, 526, 928
- Imbert, M., & Prévot, L. 1998, *A&A*, L37
- Kubát, J., Holmgren, D., & Rentzsch-Holm, I. 1998, *A&A*, 332, 842
- Lin, D. N. C., & Ida, S. 1997, *ApJ*, 477, 781
- Mazeh, T., & Zucker, S. 2000, to appear in *The Formation of Binary Stars*, ed. H. Zinnecker, & R. D. Mathieu, *ASP Conf. Ser.*, 200
- Pisani, A. 1993, *MNRAS*, 265, 706
- Pisani, A. 1996, *MNRAS*, 278, 697
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. 1992, *Numerical Recipes: The Art of Scientific Computing* (2nd ed.; New York: Cambridge Univ. Press)
- Silverman, B. W. 1986, *Density Estimation for Statistics and Data Analysis* (Chapman & Hall, London)
- Stepinski, T. F., & Black, D. C. 2000a, *A&A*, 356, 903
- Stepinski, T. F., & Black, D. C. 2000b, in poster proceedings of *IAU Symp. 200, Birth and Evolution of Binary Stars*, 167
- Trilling, D. E., Benz, W., Guillot, T., et al. 1998, *ApJ*, 500, 428
- Udry, S., Mayor, M., Latham, D. W., et al. 1998, in *Cool Stars, Stellar Systems and the Sun*, Tenth Cambridge Workshop, *ASP Conf. Ser.*, 154, 2148
- Weidenschilling, S. J., & Marzari, F. 1996, *Nature*, 384, 619
- Zahn, J.-P. 1977, *A&A*, 57, 383