Draft version October 19, 2017
Typeset using LaTeX **modern** style in AASTeX61

BINARITY'S EFFECTS ON OCCURRENCE RATES MEASURED BY TRANSIT SURVEYS

LGB, KM, JNW

GENERALITIES

Define the occurrence rate density, Γ, as the expected number of planets per star per natural logarithmic bin of planetary and stellar phase space:

$$\Gamma(\vec{x}) = \frac{d^n \Lambda}{\prod_{i=1}^{n} \mathrm{d} \ln x_i}. \tag{1}$$

$\vec{x}$ is an n-dimensional list of the continuous physical properties that might affect the occurrence rate density. For example, $\vec{x} = (r, P, R)$ where r is the planet radius, P is its orbital period, and R is the host star radius. The occurrence rate Λ is found by integrating the rate density over a specified volume of phase space.

The previous definition implicitly marginalizes the rate density over stellar multiplicity. For simplicity, this work only considers single and binary star systems. Then for a selected population of stars and planets the rate density can be written

$$\Gamma(\vec{x}) = \sum_{i=0}^{2} w_i \Gamma_i(\vec{x}) = \sum_i w_i \Lambda_i p_i(\vec{x}) \tag{2}$$

where $i = 0$ corresponds to single star systems, $i = 1$ primaries of binaries, and $i = 2$ secondaries of binaries. Λ_i is the occurrence rate integrated over all possible phase space for the i^{th} system type, $p_i(\vec{x})$ is the joint probability density function so that $\Gamma_i(\vec{x}) = \Lambda_i p_i(\vec{x})$, and the weights are given by

$$w_i = N_i / N_{\text{tot}}, \tag{3}$$

for $N_{\text{tot}} = \sum_i N_i$ the total number of selected stars, and N_0, N_1, N_2 the number of selected single stars, primaries, and secondaries respectively[1]. The relationship between the rate Λ over a desired volume of phase space Ω_{desired} and Λ_i is

$$\Lambda = \sum_i \left[\left(w_i \Lambda_i \int_{\Omega_{\text{desired}}} p_i(\vec{x})\, \mathrm{d}\Omega \right) \cdot \left(w_i \Lambda_i \int_{\Omega_{\text{total}}} p_i(\vec{x})\, \mathrm{d}\Omega \right)^{-1} \right], \tag{4}$$

where the inverse term is unity if $p_i(\vec{x})$ is appropriately normalized.

A transit survey will have a rate density of detected planets $\hat{\Gamma}$, which will be the (dot) product of the rate density and the detection efficiency $Q(\vec{x})$:

$$\hat{\Gamma}(\vec{x}) = \sum_i Q_i(\vec{x}) \Gamma_i(\vec{x}) \equiv \sum_i \hat{\Gamma}_i(\vec{x}), \tag{5}$$

where again the index i is over each type of system (singles, primaries, and secondaries). The detection efficiency includes the geometric transit probability, as well as any incompleteness effects.

[1] Eq. 2 follows by writing the i^{th} system type's rate density as some normalization multiplied by a probability density: $\Gamma_i(\vec{x}) = \mathcal{Z}_i p_i(\vec{x})$. For Eq. 1 to hold, we must have $\mathcal{Z}_i = \Lambda_i$.

1. MODEL #1: FIXED STARS, FIXED PLANETS, TWIN BINARIES

Consider a universe in which all planets are identical, and all stars are either single or twin stars with otherwise identical physical properties. Then $\vec{x} = (r, R, a)$, and

$$p_i(\vec{x}) = \delta(r - r_p)\delta(R - R_\star)\delta(a - a_p) \equiv \delta^3(r_p, R_\star, a_p), \tag{6}$$

for r_p and $R_\star$ some fixed planet and stellar radii, a_p a fixed semi-major axis, and δ the Dirac delta function, whose latter compact form will be used for brevity.

The occurrence rate density for this model is

$$\Gamma(r, R, a) = \sum_i w_i \Lambda_i \delta^3(r_p, R_\star, a_p), \tag{7}$$

and the occurrence rate over any interval that includes r_p, $R_\star$, and a_p is

$$\Lambda = \sum_i w_i \Lambda_i = \frac{\sum_i N_i \Lambda_i}{N_{\rm tot}}. \tag{8}$$

The rate is zero over intervals that do not.

To express the rate density of detected planets, $\hat{\Gamma} = \sum Q_i \Gamma_i$, we need the detection efficiencies for each system type, which are products of the geometric and selection probabilities:

$$Q_i(\vec{x}) = Q_{g,i}(\vec{x}) Q_{c,i}(\vec{x}), \quad \text{where } \vec{x} = (r, R, a). \tag{9}$$

Similar to Pepper et al. (2003), but in a new context, we take Q_c as the ratio of the number of stars that were searchable to the number of stars that were selected. Assuming a homogeneous distribution of stars, this gives

$$Q_{c,i}(\vec{x}) = \left(\frac{d_{{\rm det},i}(\vec{x})}{d_{\rm sel}(\vec{x})} \right)^3, \tag{10}$$

for $d_{\rm sel}$ the maximum distance to which surveyed stars are selected, and $d_{{\rm det},i}$ the maximum distance to which planets can actually be detected about the $i^{\rm th}$ system type. Note that $d_{\rm sel} \geq d_{{\rm det},i}$. In a signal-to-noise limited transit survey in which the observer does not know which stars are binaries,

$$d_{\rm sel} \propto (r/R)^2 (L_{\rm sys} T_{\rm dur} A N_{\rm tra})^{1/2}, \tag{11}$$

for $L_{\rm sys} = L_1(1 + \gamma_R)$ the system luminosity, $T_{\rm dur}$ the transit duration, A the detector area, and $N_{\rm tra}$ the number of observed transits. However,

$$d_{{\rm det},i} \propto \mathcal{D}_i (r/R)^2 (L_{\rm sys} T_{\rm dur} A N_{\rm tra})^{1/2}, \tag{12}$$

for the dilution $\mathcal{D}_i$ given by

$$\mathcal{D}_i = \begin{cases} 1 & \text{for } i = 0, \text{ single} \\ L_1/L_{\rm sys} = (1 + \gamma_R)^{-1}, & \text{for } i = 1, \text{ primary} \\ \gamma_R L_1/L_{\rm sys} = (1 + \gamma_R^{-1})^{-1}, & \text{for } i = 2, \text{ secondary,} \end{cases} \tag{13}$$

where the light ratio γ_R of a given binary is defined as the ratio of the luminosity of the secondary to the primary.

The maximum detectable distance to single stars is assumed to be known, and so $d_{\mathrm{sel},0} = d_{\mathrm{det},0}$. For binary systems there is a necessary incompleteness, and combining Eqs. 9 through 13 in the case of twin binaries yields

$$Q_i(\vec{x}) = \begin{cases} R/a, & \text{for } i = 0 \text{ single} \\ (R/a)(1+\gamma_R)^{-3} = (R/8a), & \text{for } i = 1 \text{ primary} \\ (R/a)(1+\gamma_R^{-1})^{-3}\gamma_R^{-5/\alpha} = (R/8a), & \text{for } i = 2 \text{ secondary.} \end{cases} \tag{14}$$

The factor of $\gamma_R^{-5/\alpha}$ in the latter expression of Eq. 7 comes from the mass-luminosity-radius relation of stars in the model: we assume $R \propto M \propto L^{1/\alpha}$. For $\gamma_R = 1$ this term is unimportant, but it will later become relevant.

Summarizing, since we have written the rate density for each system type (Eq. 7) and the detection efficiency for each system type (Eq. 14), we have fully specified the rate density of detected planets, $\hat{\Gamma} = \sum Q_i \Gamma_i$, in addition to the true rate density.

1.1. *What does an observer ignoring binarity infer?*

Consider an observer who ignores binarity. They assume a detection efficiency $\tilde{Q} = Q_0$, measure a detected planet rate density $\tilde{\Gamma}$, and infer an apparent rate density Γ_a. Analogous to Eq. 5,

$$\tilde{\Gamma} = \Gamma_a \tilde{Q}. \tag{15}$$

Accounting for dilution, one can show

$$\Gamma_a = w_a \Lambda_0 \delta^3(r_p, R_\star, a_p) + w_b(\Lambda_1 Q_{c,1} + \Lambda_2 Q_{c,2})\delta^3(r_p/\sqrt{2}, R_\star, a_p), \tag{16}$$

for $w_a = N_0/(N_0 + N_1)$, and $w_b = N_1/(N_0 + N_1)$. This observer miscounts the number of total searched stars, does not correct for completeness, and misclassifies the planetary radii because of dilution.

1.2. *Correction to inferred rate density and inferred rate*

Define the rate density correction factor, X_Γ, as the ratio of the apparent to true rate densities:

$$X_\Gamma \equiv \frac{\Gamma_a}{\Gamma}. \tag{17}$$

This factor can be a function of whatever parameters Γ_a and Γ depend on; in this study, the planet radius is most relevant. For the twin-binaries model,

$$X_\Gamma(r) = \frac{w_a \Lambda_0 \delta^3(r_p) + w_b(\Lambda_1 Q_{c,1} + \Lambda_2 Q_{c,2})\delta^3(r_p/\sqrt{2})}{(w_0\Lambda_0 + w_1\Lambda_1 + w_2\Lambda_2)\delta^3(r_p)} \tag{18}$$

where $\delta^3(r_p)$ is shorthand for $\delta^3(r - r_p, R - R_\star, a - a_p)$.

If we take the rates Λ_i to be equal, applying the definitions of the weights gives a rate density correction factor at $r = r_p$ of $X_\Gamma(r_p) = (1+\mu)^{-1}$, where

$$\mu \equiv \frac{N_1}{N_0} = \frac{n_b}{n_s}\left(\frac{d_{\rm sel,b}}{d_{\rm sel,s}}\right)^3 = \frac{\rm BF}{1-{\rm BF}}(1+\gamma_R)^{3/2}, \tag{19}$$

for n_b and n_s the number density of binaries and singles in a volume limited sample. Using Raghavan et al. (2010)'s $0.7{-}1.3M_\odot$ multiplicity fraction as our binary fraction, we set BF $= 0.44$. The resulting correction to the rate density is $X_\Gamma(r_p) \approx 0.31$. The correction at $r_p/\sqrt{2}$ is infinite.

If we assume that $\Lambda_0 = \Lambda_1$, but that $\Lambda_2 = 0$, we find $X_\Gamma(r_p) = (1+2\mu)/(1+\mu)^2$. Taking the same binary fraction, this evaluates to $X_\Gamma(r_p) \approx 0.53$.

In passing, note that a correction to the inferred rate, X_Λ, can be defined analogously:

$$X_\Lambda \equiv \frac{\Lambda_a}{\Lambda}. \tag{20}$$

For the twin binary model, the correction to the rate is the same as that to the rate density.

2. MODEL #2: FIXED PLANETS AND PRIMARIES, VARYING SECONDARIES

The main use of our binary-twin model was to help develop intuition. Gradually introducing complexity, we now let the light ratio $\gamma_R = L_2/L_1$ vary across the binary population. It does so because the underlying mass ratio $q = M_2/M_1$ varies. We keep the primary mass fixed as M_1, which is also the mass of all single stars.

We parametrize the distribution of binary mass ratios in a volume-limited sample as a power law: $p(q) \propto q^\beta$. For binaries with solar-type primaries[2], β is probably between 0 and 0.3. Since we assume stars are a one-parameter family, $R \propto M \propto L^{1/\alpha}$, a drawn value of q determines everything about a secondary.

The rate density in this model, $\Gamma(\vec{x})$, is the sum over system types of $w_i\Lambda_i p_i(\vec{x})$:

$$\Gamma(\vec{x}) = \delta^4(r_p, R_\star, a_p, P_p)(w_0\Lambda_0 + w_1\Lambda_1) + w_2\Lambda_2\delta^3(r_p, P_p, a_p)p_2(q), \tag{21}$$

where the semimajor axis of the planet must be such that its period is P_p, and $p_2(q)$ is expressed in terms of the mass ratio instead of the secondary star's radius for convenience (q and R_2 are interchangeable). The probability that a secondary hosts a planet, as a function of the mass ratio, is

$$p_2(q) = p(\text{has planet} \,|\, \text{secondary with } q) \times p(\text{secondary with } q) \tag{22}$$

$$p_2(q) \propto q^{\gamma+\beta}(1+q^\alpha)^{3/2}. \tag{23}$$

[2] Duchene and Kraus (2013), fitting all the multiple systems of Raghavan et al. (2010)'s Fig 16, find $\beta = 0.28 \pm 0.05$ for $0.7 < M_\star/M_\odot < 1.3$. Examining only the binary systems of Rhagavan et al 2010, Fig 16, the distribution seems roughly uniform, $\beta \approx 0$, except for a claimed excess of twin binaries with $q \approx 1$, and a deficiency of $q < 0.1$ stellar companions.

We take first term, $p(\text{has planet} \,|\, \text{secondary with } q)$, as a power law of q with exponent γ. For the second term, since the selected sample at a given (r, P, a) is magnitude-limited, $p(\text{secondary with } q)$ is the product of the volume limited probability and a Malmquist-like bias $(1 + q^\alpha)^{3/2}$.

The occurrence rate corresponding to Eq. 21's rate density for a desired volume of phase space Ω_{desired} is given by Eq. 4. Specifying the desired mass ratios of interest as $q_{\min} < q < q_{\max}$, this simplifies to

$$\Lambda = \frac{N_0\Lambda_0 + N_1\Lambda_1 + N_2\Lambda_2 f_2}{N_{\text{tot}}}, \tag{24}$$

for

$$f_2 \equiv \left(\int_{q_{\min}}^{q_{\max}} p_2(q)\, \mathrm{d}q \right) \cdot \left(\int_0^1 p_2(q)\, \mathrm{d}q \right)^{-1}. \tag{25}$$

The detected rate density, $\hat{\Gamma} = \sum_i Q_i \Gamma_i$, will be specified by the detection efficiencies for each type of system. These are nearly identical to Eq. 14:

$$Q_0 = Q_{g,0}Q_{c,0} = Q_{g,0} \tag{26}$$
$$Q_1 = Q_{g,1}Q_{c,1} = Q_{g,0}(1 + q^\alpha)^{-3} \tag{27}$$
$$Q_2 = Q_{g,2}Q_{c,2} = Q_{g,0} q^{2/3}(1 + q^{-\alpha})^{-3} q^{-5}, \tag{28}$$

for $Q_{g,0} = R/a$, the transit probability in single star systems. The detection efficiency for secondaries (Eq. 28) includes the transit probability from the smaller stellar radius, and combines dilution, the transit duration, and stellar radius for the completeness probability.

2.1. *What does an observer ignoring binarity infer?*

As a reminder, the apparent rate density is found by correcting the detected apparent rate density for the transit probability: $\Gamma_a = \tilde{\Gamma} Q_{g,0}^{-1}$. The observer's errors are as follows:

1. The true planetary radii r are interpreted as apparent radii r_a.
2. The true stellar radii R are all thought to be $R_\star$. In "reality", this only holds for single stars and primaries.
3. The selected and searchable stars are miscounted.

Writing the apparent rate density as a function of $\vec{x} = (r_a, P, a, R)$, where R is written interchangeably with the binary mass ratio q,

$$\Gamma_a(\vec{x}) = Q_0 w_a \Lambda_0 \delta^4(r_p) + Q_1 w_b \Lambda_1 \delta^3(P_p) p_1(r_a) + Q_2 w_b \Lambda_2 \delta^2(P_p, a_p) p_2(q) p_2(r_a), \tag{29}$$

where the detection efficiencies are given in Eqs. 26-28, $p_2(q)$ is given by Eq. 23, and as in the first model, $w_a = N_0/(N_0 + N_1)$, $w_b = N_1/(N_0 + N_1)$.

The apparent radii depend on the system type:

$$r_a = \begin{cases} r_p(1+q^\alpha)^{-1/2} & \text{for } i=1, \text{ primary} \\ r_p(1+q^{-\alpha})^{-1/2}q^{-1}, & \text{for } i=2, \text{ secondary.} \end{cases} \tag{30}$$

The factor of q^{-1} for the secondary case accounts for the observer assuming that the host star is the primary.

Since the apparent radii directly depend on the mass ratio, we wish to write

$$p_i(r_a) = p_i(q(r_a)) \left| \frac{\mathrm{d}q}{\mathrm{d}r_a} \right|, \quad i \in \{1,2\}, \tag{31}$$

where $q(r_a)$ must be found by inverting Eq. 30. For $i=1$, this gives

$$p_1(r_a) \propto \frac{r_p^5}{r_a^6} \left(\left(\frac{r_p}{r_a} \right)^2 - 1 \right)^{\frac{\beta-\alpha+1}{\alpha}}, \tag{32}$$

but for $i=2$, $r_a(q)$ is not invertible. Instead, we find $p_2(r_a)$ numerically[3], and show it in Fig. 1.

2.2. *Correction to inferred rate density*

Recall that the rate density correction factor, X_Γ, is the ratio of the apparent to true rate densities. Applying Eqs. 21 and 29,

$$X_\Gamma(r,q) = \frac{w_a \Lambda_0 \delta(r_p, R_\star) + Q_{c,1} w_b \Lambda_1 \delta(R_\star) p_1(r_a) + Q_{c,2} q^{2/3} w_b \Lambda_2 p_2(q) p_2(r_a)}{w_0 \Lambda_0 \delta^2(r_p, R_\star) + w_1 \Lambda_1 \delta^2(r_p, R_\star) + w_2 \Lambda_2 p_2(q) \delta(r_p)}, \tag{33}$$

where the stellar radius-dependent δ-functions are not important. To marginalize over q, we must write

$$X_\Gamma(r) = \frac{\Gamma_a(r)}{\Gamma(r)} = \frac{\sum_i \int_0^1 \Gamma_{a,i}(r,q)\,\mathrm{d}q}{\sum_i \int_0^1 \Gamma_i(r,q)\,\mathrm{d}q}. \tag{34}$$

Using the same definition of $\mu = N_1/N_0$ from Eq. 19, it is important when performing each integral to note that μ is a function of q:

$$\mu = \mathcal{B}(1+q^\alpha)^{3/2}, \quad \mathcal{B} \equiv \frac{\mathrm{BF}}{1-\mathrm{BF}}. \tag{35}$$

Thus the weights in the apparent and actual rate densities, (w_a, w_b) and (w_0, w_1, w_2), are all functions of q.

The crucial point of Eq. 34 is that since both the real and apparent rate densities are separable, the expressions will simplify.

[3] We sample from $p_2(q)$ (Eq. 23), and use Eq. 31 to compute the resulting apparent radii

Our "nominal model" is a stellar population similar to Sun-like stars in the local neighborhood: BF $= 0.44$, $\alpha = 3.5$, $\beta = 0$. We also assume that the occurrence of planets is independent of stellar mass ($\gamma = 0$). Under these nominal assumptions, the planetary rate density is

$$\Gamma(r) \approx \delta(r - r_p)\left(\Lambda_0 + \Lambda_1 + \Lambda_2\right)/3, \tag{36}$$

where the coefficients of $1/3$ are accurate to within one percent of the true coefficients. Ignoring binarity, the observer finds an apparent rate density

$$\Gamma_a(r) = c_0\Lambda_0\delta(r - r_p) + c_1\Lambda_1 p_1(r_a) + c_2\Lambda_2 p_2(r_a), \tag{37}$$

for $c_0 \approx 0.49$, $c_1 \approx 0.32$, $c_2 \approx 0.03$, and the two latter apparent-radius dependences are shown in Fig. 1.

While the detailed radial dependence of the error between Eqs. 36 and 37 is interesting, the general point is that dilution produces a spectrum of apparent planetary radii, which skews occurrence rate and occurrence rate density measurements.

Evaluating the correction term at $r = r_p$, since $\lim_{q\to 0} p_i(r_a) = 0$ for $i \in \{1, 2\}$,

$$X_\Gamma(r = r_p) \approx \frac{3c_0\Lambda_0}{\Lambda_0 + \Lambda_1 + \Lambda_2}. \tag{38}$$

If all the rates are equal, $X_\Gamma(r = r_p) \approx 0.49$. If there are no planets around the secondaries, $X_\Gamma(r = r_p) \approx 0.74$.

3. MODEL #3: FIXED PRIMARIES, VARYING PLANETS AND SECONDARIES

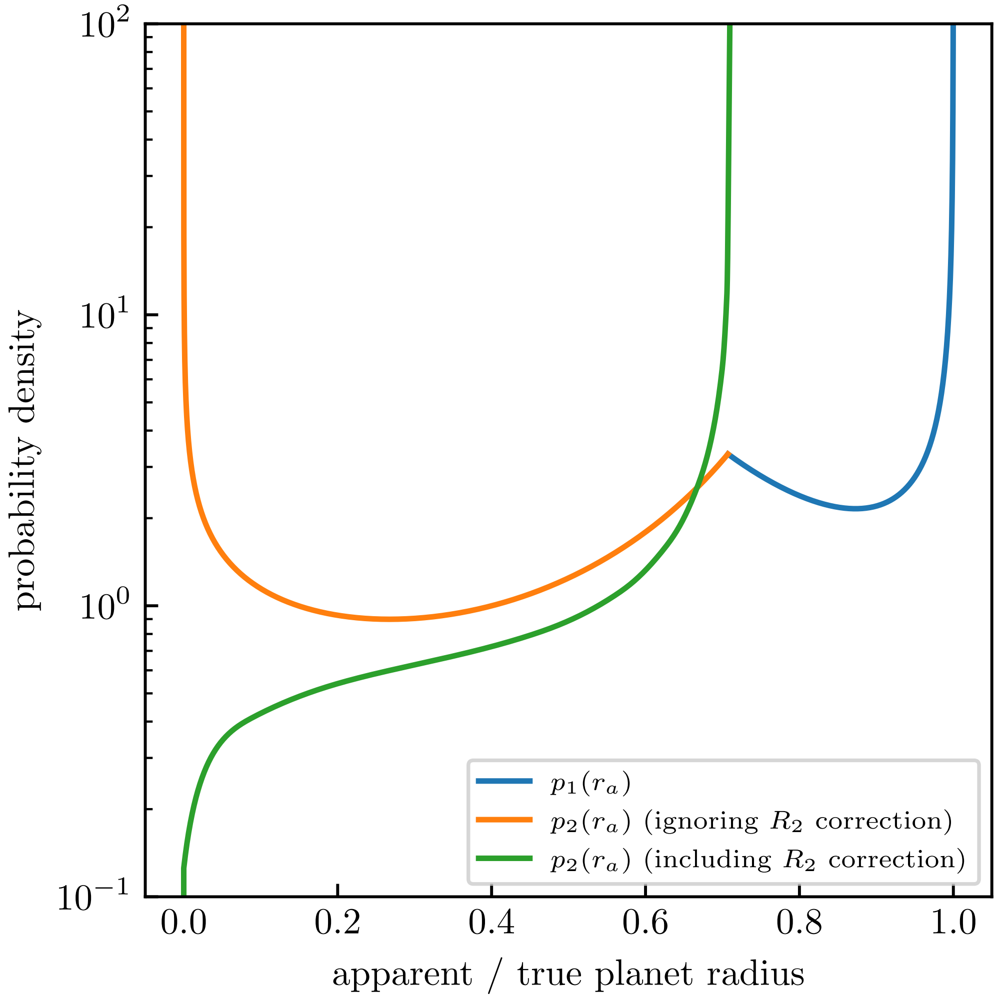

Figure 1. Model #2 (Sec. 2): probability of observing an apparent radius r_a given a planet orbiting the primary, $p_1(r_a)$, or the secondary, $p_2(r_a)$ of a binary system. The true planet radius is fixed – a delta function centered on "1". This plot takes $\alpha = 3.5, \beta = 0$, for α the exponent in the mass-luminosity relation $L \propto M^\alpha$, and β the exponent in the distribution of mass ratios in a volume-limited sample. Each distribution is separately normalized.