

# THE EFFECTS OF BINARITY ON PLANET OCCURRENCE RATES MEASURED BY TRANSIT SURVEYS

L. G. BOUMA,<sup>1</sup> J. N. WINN,<sup>1</sup> AND K. MASUDA<sup>1</sup>

<sup>1</sup>*Department of Astrophysical Sciences, Princeton University, 4 Ivy Lane, Princeton, NJ 08540, USA*

Submitted to AAS journals.

## ABSTRACT

This note formally derives a general equation for the apparent occurrence rate used in the paper. For some analytic cases, this lets us vet our Monte Carlo code.

*Keywords:* methods: data analysis — planets and satellites: detection  
— surveys

## 1. PRELIMINARIES

### 1.1. Searchable Distance

We can detect a signal if

$$\frac{\text{signal}}{\text{noise}} \sim \frac{\delta_{\text{obs}}}{(L/d^2)^{-1/2}}, \quad \delta_{\text{obs}} : \text{observed depth} \quad (1)$$

is above some threshold (it would probably make more sense to include the duration information, but that would anyway be a trivial extension and so we omit it here for brevity). Thus, the maximum searchable distance scales as

$$d(\delta_{\text{obs}}, L_{\text{sys}}) \propto \delta_{\text{obs}} \cdot L_{\text{sys}}^{1/2}. \quad (2)$$

We assume that the signal is detected if and only if a given star is searchable.

Assuming that stars are uniformly distributed in space, the number of searchable stars  $N_s$  is then proportional to

$$N_s(\delta_{\text{obs}}, L_{\text{sys}}) \propto n \delta_{\text{obs}}^3 L_{\text{sys}}^{3/2}, \quad (3)$$

where  $n$  is the volume density of *e.g.*, single star, or binary systems. We neglect the dependence of  $n$  on stellar type.

### 1.2. Relation between Apparent and Actual Stellar Properties

We assume that the apparent properties of an unresolved binary are the same as those of the primary:

$$M_a = M_1, \quad R_a = R_1, \dots \quad (4)$$

We also assume the stellar radius and luminosity is uniquely related to the stellar mass.

Given these assumptions, the total luminosity of a system is

$$L_{\text{sys}} = L_1 + L_2 = L(M_a) + L(qM_a), \quad (5)$$

where  $q = M_2/M_1$ . Note that  $L_{\text{sys}}$  is the true value (because  $M_a$  is the true primary mass), while an observer would estimate a system luminosity of  $L(M_a)$ , based on the apparent stellar parameters. We presume the observer has no a priori knowledge of the distance to a given system – they only measure a flux, assume a stellar mass  $M_a$ , and use relations for  $L(M_a)$  and  $R(M_a)$  to estimate system properties.

### 1.3. Apparent Number of Searchable Stars

Given the apparent signal  $\delta_{\text{obs}}$  and stellar mass  $M_a$ , the maximum searchable distance for singles and binaries are proportional to  $\delta_{\text{obs}} \cdot L(M_a)^{1/2}$  and  $\delta_{\text{obs}} \cdot [L(M_a) + L(qM_a)]^{1/2}$ . Thus, the apparent number of searchable stars (i.e. points in the sky), which will be selected by an ignorant observer, is proportional to

$$N_{s,a}(\delta_{\text{obs}}, M_a) \propto n_s \delta_{\text{obs}}^3 L(M_a)^{3/2} [1 + \mu(\text{BF}, M_a)],$$

and can be *defined* as

$$N_{s,a}(\delta_{\text{obs}}, M_a) \equiv N_s^0(\delta_{\text{obs}}, L(M_a)) [1 + \mu(\text{BF}, M_a)], \quad (6)$$

where  $N_s^0$  is the number of searchable singles (this agrees with the actual value),  $\text{BF} = n_b/(n_s + n_b)$  is the binary fraction in a volume-limited sample, and  $n_s$  is the number density of singles in a volume-limited sample. Proportionality constants subsumed into the definition of  $N_s^0$  include *e.g.*, the telescope area and the survey duration.

### 1.3.1. What is $\mu$ ?

In Eq. 6,  $\mu$  is the ratio of the number of single to double star systems (that are searchable for an observed signal  $\delta_{\text{obs}}$ ). If we make various simplifying assumptions, it turns out that this number depends only on the binary fraction and the binary mass ratio distribution in a volume-limited sample. The proof is as follows.

In a shell of thickness  $dr$ , there are  $n_d(r)4\pi r^2 dr$  searchable double star systems. The number density of searchable double star systems  $n_d(r)$  is constant up to the maximum searchable distance for singles,  $d_0$ . At greater distances,  $n_d(r)$  decreases, because only some binaries have sufficient flux to make the SNR cut. Specifically,

$$n_d(r) = \begin{cases} n_b, & r \leq d_0 \\ n_b \int_{q_{\min}(r)}^1 dq f_v(q), & d_0 < r < \sqrt{2}d_0 \end{cases} \quad (7)$$

where  $f_v(q)$  is the mass ratio distribution in a volume-limited sample,  $q_{\min}(r)$  is the minimum mass ratio required to provide sufficient flux for searchability. Assuming  $L \propto M^\alpha$ ,

$$q_{\min}(r) = \left( \left( \frac{r}{d_0} \right)^2 - 1 \right)^{\frac{1}{\alpha}}. \quad (8)$$

Using the common parametrization  $f_v(q) \sim q^\beta$ , and setting  $\beta = 0$ ,

$$n_d(r) = \begin{cases} n_b, & r \leq d_0 \\ n_b(1 - q_{\min}(r)), & d_0 < r < \sqrt{2}d_0. \end{cases} \quad (9)$$

Integrating,

$$N_d = \int dN_d = \int n_d(r)4\pi r^2 dr \quad (10)$$

$$= 4\pi n_b \left( \frac{d_0^3}{3} + \left[ \frac{r^3}{3} \right]_{d_0}^{\sqrt{2}d_0} - \int_{d_0}^{\sqrt{2}d_0} \left( \left( \frac{r}{d_0} \right)^2 - 1 \right)^{\frac{1}{\alpha}} r^2 dr \right) \quad (11)$$

$$= \frac{4\pi n_b}{3} d_0^3 \left( 2^{3/2} - 3 \int_1^{\sqrt{2}} du u^2 (u^2 - 1)^{1/\alpha} \right), \quad (12)$$

$$\equiv \frac{4\pi n_b}{3} d_0^3 (2^{3/2} - 3I), \quad (13)$$

where the penultimate line substituted  $u = r/d_0$ . For the common case  $\alpha = 3.5$ , the dimensionless integral  $I$  evaluates to  $\approx 0.484174$ .

Thus

$$\mu \equiv \frac{N_d}{N_s^0} = \frac{n_b}{n_s} (2^{3/2} - 3I) = \frac{\text{BF}}{1 - \text{BF}} (2^{3/2} - 3I). \quad (14)$$

For the twin binary case,  $I = 0$ , and for the power law volume-limited distribution,  $I$  takes the form given in Eq. 12.

## 2. APPARENT OCCURRENCE RATE — GENERAL FORMULA

A group of astronomers wants to measure the mean number of planets of a certain type per star of a certain type. They observe a set points on the sky and detect a set of planets that appear to be of the desired class. They then choose the stars (among those initially selected) around which the planets of interest appeared to be searchable. Given these quantities, they compute the apparent rate density,

$$\Gamma_a(\mathcal{P}_a, \mathcal{S}_a) = \frac{N_{\text{det}}(\mathcal{P}_a, \mathcal{S}_a)}{N_{s,a}(\mathcal{P}_a, \mathcal{S}_a)} \times \frac{1}{p_{\text{tra}}(\mathcal{P}_a, \mathcal{S}_a)}. \quad (15)$$

where  $\mathcal{P}_a, \mathcal{S}_a$  are the apparent planetary/stellar parameters. Here  $N_{\text{det}}(\mathcal{P}_a, \mathcal{S}_a)$  is the number of detected planets with parameters  $\mathcal{P}_a, \mathcal{S}_a$ , *per unit*  $\mathcal{P}_a$  and  $\mathcal{S}_a$ . The quantity  $N_{s,a}(\mathcal{P}_a, \mathcal{S}_a)$  is the apparent number of searchable stars. The apparent rate for a given class of planets is found by integrating the rate density. For instance, if  $\mathcal{P}_a = r_a$  and  $\mathcal{S}_a = M_a$ , the apparent rate is

$$\int d^2\Lambda_a(r_a, M_a) = \int dr_a dM_a \Gamma_a(\mathcal{P}_a, \mathcal{S}_a) = \int dr_a dM_a \frac{N_{\text{det}}(r_a, M_a)}{N_{s,a}(r_a, M_a)} \times \frac{1}{p_{\text{tra}}(r_a, M_a)}. \quad (16)$$

In the presence of dilution, planets with  $(\mathcal{P}_a, \mathcal{S}_a)$  are associated with systems of many different planetary and stellar properties, so  $N_{\text{det}}(\mathcal{P}_a, \mathcal{S}_a)$  is given by the convolution of the true rate density,  $\Gamma(\mathcal{P}, \mathcal{S})$ , and  $\mathcal{N}(\mathcal{P}_a, \mathcal{S}_a; \mathcal{P}, \mathcal{S})$ , the number (per unit  $(\mathcal{P}_a, \mathcal{S}_a)$ ) of searchable stars that give  $(\mathcal{P}_a, \mathcal{S}_a)$  when the true system actually has  $(\mathcal{P}, \mathcal{S})$ :

$$N_{\text{det}}(\mathcal{P}_a, \mathcal{S}_a) = \sum_i N_{\text{det}}^i(\mathcal{P}_a, \mathcal{S}_a) = \sum_i \int d\mathcal{P} d\mathcal{S} \mathcal{N}_s^i(\mathcal{P}_a, \mathcal{S}_a; \mathcal{P}, \mathcal{S}) \cdot \Gamma^i(\mathcal{P}, \mathcal{S}) \cdot p_{\text{tra}}(\mathcal{P}, \mathcal{S}), \quad (17)$$

where  $i$  specifies the type of true host stars (0: single, 1: primary, 2: secondary).

So the problem reduces to the evaluation of

$$\mathcal{N}_s^i(\mathcal{P}_a, \mathcal{S}_a; \mathcal{P}, \mathcal{S}) \quad (18)$$

for planets around single stars, primaries in binaries, and secondaries in binaries.

### 3. EVALUATION OF $\mathcal{N}_s^i$

Let us explicitly write  $\mathcal{P} = r$  and  $\mathcal{S} = M$ ;  $R$  and  $L$  are uniquely determined from the assumed mass–radius–luminosity relation. We neglect the dependence on planetary orbital period. We proceed by evaluating

$$N_{\text{det}}(r_a, M_a) = \sum_i N_{\text{det}}^i(r_a, M_a) \quad (19)$$

$$N_{\text{det}}(r_a, M_a) = \sum_i \int \text{d}r \text{d}M \mathcal{N}_s^i(r_a, M_a; r, M) \cdot \Gamma^i(r, M) \cdot p_{\text{tra}}(r, M), \quad (20)$$

term by term.

#### 3.1. *Single Stars*

For  $i = 0$ ,

$$\mathcal{N}_s^0(r_a, M_a; r, M) = \delta(r_a - r) \delta(M_a - M) N_s^0(r, M), \quad (21)$$

so

$$N_{\text{det}}^0(r_a, M_a) = N_s^0(r_a, M_a) \cdot \Gamma^0(r_a, M_a) \cdot p_{\text{tra}}(r_a, M_a). \quad (22)$$

If all the stars are singles, this yields

$$\Gamma_a(r_a, M_a) = \Gamma^0(r_a, M_a), \quad (23)$$

as expected (now  $\mu = 0$  and  $N_{s,a} = N_s^0$ ) — the true occurrence is recovered.

#### 3.2. *Primaries in Binaries*

Since we assume  $\mathcal{S}_a = \mathcal{S}_1$ ,

$$\mathcal{N}_s^1(r_a, M_a; r, M) \propto \delta(M_a - M). \quad (24)$$

In this case,  $\mathcal{N}_s^1$  is non-zero only at  $r_a = R_a \sqrt{\delta_{\text{obs}}}$ , and the observed depth is

$$\delta_{\text{obs}} = \left[ \frac{r}{R(M_a)} \right]^2 \times \frac{L(M_a)}{L_{\text{sys}}(M_a, q)}. \quad (25)$$

The normalization of  $\mathcal{N}_s^1$  is given by the number of binaries that are searchable for a signal  $\delta_{\text{obs}}$ :

$$N_s^0(\delta_{\text{obs}}, L(M_a)) \cdot \mu(\text{BF}, M_a). \quad (26)$$

Thus, the number of primaries with apparent parameters  $(r_a, M_a)$  given the true parameters  $(r, M)$  is

$$\mathcal{N}_s^1(r_a, M_a; r, M) = \int \text{d}q f(q) \mathcal{N}_{s,q}^1(r_a, M_a; r, M; q), \quad (27)$$

where  $f(q)$  is the binary mass ratio distribution in the SNR-limited sample and

$$\begin{aligned} \mathcal{N}_{s,q}^1(r_a, M_a; r, M; q) &= N_s^0(\delta_{\text{obs}}, L(M_a)) \cdot \mu(\text{BF}, M_a) \\ &\times \delta \left( r_a - r \sqrt{\frac{L(M_a)}{L_{\text{sys}}(M_a, q)}} \right) \delta(M_a - M). \end{aligned} \quad (28)$$

Note that while the observed depth is a function of the true parameters (Eq. 25), since we are counting detected planets of a given *apparent* size, we will not need to worry about this dependence. As discussed in the main text, in Eq. 27 we will have  $f(q) \sim q^\beta(1 + q^\alpha)^{3/2}$ .

### 3.3. Secondaries in Binaries

In this case,  $M = qM_1 = qM_a$ , so

$$\mathcal{N}_s^2(r_a, M_a; r, M) \propto \delta\left(M_a - \frac{M}{q}\right). \quad (29)$$

Again  $\mathcal{N}_s^2$  is non-zero only at  $r_a = R_a\sqrt{\delta_{\text{obs}}}$ , but this time

$$\delta_{\text{obs}} = \left[\frac{r}{R(qM_a)}\right]^2 \times \frac{L(qM_a)}{L_{\text{sys}}(M_a, q)}. \quad (30)$$

The normalization remains the same as the previous case (we are counting the searchable stars at a given observed depth  $\delta_{\text{obs}}$ , total luminosity of the binary is the same). Thus,

$$\mathcal{N}_s^2(r_a, M_a; r, M) = \int dq f(q) \mathcal{N}_{s,q}^2(r_a, M_a; r, M; q), \quad (31)$$

where

$$\begin{aligned} \mathcal{N}_s^2(r_a, M_a; r, M; q) &= N_s^0(\delta_{\text{obs}}, L(M_a)) \cdot \mu(\text{BF}, M_a) \\ &\times \delta\left(r_a - r\sqrt{\left[\frac{R(M_a)}{R(qM_a)}\right]^2 \frac{L(qM_a)}{L_{\text{sys}}(M_a, q)}}\right) \delta\left(M_a - \frac{M}{q}\right). \end{aligned} \quad (32)$$

One might worry in Eq. 29 that we opt to write  $\mathcal{N}_s^2 \propto \delta(M_a - M/q)$ , rather than  $\propto \delta(M_a q - M)$  or some other seemingly function with the same functional dependence, but a different normalization once integrated. We do this because the delta function in Eq. 29 is defined with respect to the measure  $dM_a$ , not  $dM$ . This means that it should be unity when integrated over  $dM_a$ . Ultimately, this is based on how we defined  $\mathcal{N}_s^2$  as a number per  $r_a$ , per  $M_a$ .

## 4. RESULT

### 4.1. Marginalization over the True Properties

Let's integrate out  $\mathcal{P}$  and  $\mathcal{S}$ .

$$N_{\text{det}}^0(r_a, M_a) = \int dr dM \mathcal{N}_s^0(r_a, M_a; r, M) \cdot \Gamma^0(r, M) \cdot p_{\text{tra}}(M) \quad (33)$$

$$= \int dr dM N_s^0(\delta_{\text{obs}}, L(M_a)) \delta(r_a - r) \delta(M_a - M) \cdot \Gamma^0(r, M) \cdot p_{\text{tra}}(M) \quad (34)$$

$$= N_s^0(\delta_{\text{obs}}, L(M_a)) \cdot \Gamma^0(r_a, M_a) \cdot p_{\text{tra}}(M_a). \quad (35)$$

$$N_{\text{det}}^1(r_a, M_a) = \int dr dM \mathcal{N}_s^1(r_a, M_a; r, M) \cdot \Gamma^1(r, M) \cdot p_{\text{tra}}(M) \quad (36)$$

$$= \int dq f(q) \int dr dM \mathcal{N}_{s,q}^1(r_a, M_a; r, M; q) \cdot \Gamma^1(r, M) \cdot p_{\text{tra}}(M) \quad (37)$$

$$= N_s^0(\delta_{\text{obs}}, L(M_a)) \cdot p_{\text{tra}}(M_a) \cdot \mu(\text{BF}, M_a) \int \frac{dq}{\mathcal{A}} f(q) \Gamma^1(r_a/\mathcal{A}, M_a), \quad (38)$$

where

$$\mathcal{A}(q, M_a) = \sqrt{\frac{L(M_a)}{L_{\text{sys}}(M_a, q)}}. \quad (39)$$

Finally,

$$N_{\text{det}}^2(r_a, M_a) = \int dr dM \mathcal{N}_s^2(r_a, M_a; r, M) \cdot \Gamma^2(r, M) \cdot p_{\text{tra}}(M) \quad (40)$$

$$= \int dq f(q) \int dr dM \mathcal{N}_{s,q}^2(r_a, M_a; r, M; q) \cdot \Gamma^2(r, M) \cdot p_{\text{tra}}(M) \quad (41)$$

$$= N_s^0(\delta_{\text{obs}}, L(M_a)) \mu(\text{BF}, M_a) \int dr dq q f(q) \delta(r_a - r\mathcal{B}) \Gamma^2(r, qM_a) p_{\text{tra}}(qM_a) \quad (42)$$

$$= N_s^0(\delta_{\text{obs}}, L(M_a)) \cdot \mu(\text{BF}, M_a) \int \frac{qdq}{\mathcal{B}} f(q) \Gamma^2(r_a/\mathcal{B}, qM_a) p_{\text{tra}}(qM_a), \quad (43)$$

where

$$\mathcal{B}(q, M_a) = \frac{R(M_a)}{R(qM_a)} \sqrt{\frac{L(qM_a)}{L_{\text{sys}}(M_a, q)}}. \quad (44)$$

#### 4.2. Final Formula

Note again that the denominators of  $\Gamma_a$  are the same for singles, primaries, and secondaries:  $N_{s,a}(\delta_{\text{obs}}, M_a)$  and  $p_{\text{tra}}(M_a)$ . This is because we are calculating the occurrence at the same apparent planet/star properties, and the observer can never distinguish binaries from singles (and adopt the primary properties for binaries). Using the results above, the apparent rate density,

$$\Gamma_a(r_a, M_a) = \frac{1}{N_{s,a}(r_a, M_a) p_{\text{tra}}(r_a, M_a)} \times \sum_i N_{\text{det}}^i(r_a, M_a), \quad (45)$$

evaluates to

$$\begin{aligned} \Gamma_a(r_a, M_a) = \frac{1}{1 + \mu(\text{BF}, M_a)} \times & \left\{ \Gamma^0(r_a, M_a) + \mu(\text{BF}, M_a) \left[ \int \frac{dq}{\mathcal{A}} f(q) \Gamma^1\left(\frac{r_a}{\mathcal{A}}, M_a\right) \right. \right. \\ & \left. \left. + \int \frac{qdq}{\mathcal{B}} f(q) \Gamma^2\left(\frac{r_a}{\mathcal{B}}, qM_a\right) \frac{R(qM_a)}{R(M_a)} q^{-1/3} \right] \right\}. \end{aligned} \quad (46)$$

#### 4.2.1. Simplifying above general result

To simplify, drop the explicit  $M_a$  and BF dependences, and assume  $R \propto M$ , so  $R(qM_a)/R(M_a) = q$ . Also, move all superscript numbers to subscripts. Then the apparent rate density is

$$\Gamma_a(r_a) = \frac{1}{1 + \mu} \left( \Gamma_0(r_a) + \mu \left[ \int \frac{dq}{\mathcal{A}(q)} f(q) \Gamma_1 \left( \frac{r_a}{\mathcal{A}(q)} \right) + \int \frac{q dq}{\mathcal{B}(q)} f(q) \Gamma_2 \left( \frac{r_a}{\mathcal{B}(q)} \right) q^{2/3} \right] \right). \quad (47)$$

For the power law case when  $L \sim M^\alpha$ ,

$$\mathcal{A}(q) = (1 + q^\alpha)^{-1/2}, \quad \mathcal{B}(q) = q^{-1}(1 + q^{-\alpha})^{-1/2} \quad (48)$$

### 5. EXAMPLES

#### 5.1. Twin Binary

We have  $f(q) = \delta(q - 1)$ . The apparent rate density is

$$\Gamma_a(r_a) = \frac{1}{1 + \mu(\text{BF})} \left\{ \Gamma_0(r_a) + \mu(\text{BF}) \cdot \sqrt{2} \cdot \left[ \Gamma_1(\sqrt{2} r_a) + \Gamma_2(\sqrt{2} r_a) \right] \right\}, \quad (49)$$

where

$$\mu(\text{BF}) = \int dq f(q) \frac{\text{BF}}{1 - \text{BF}} \left[ 1 + \frac{L(qM_a)}{L(M_a)} \right]^{3/2} = 2^{3/2} \cdot \frac{\text{BF}}{1 - \text{BF}}. \quad (50)$$

##### 5.1.1. Same Planets (Model # 1)

If  $\Gamma_i(r, M) = Z_i \cdot \delta(r - r_0)$  (all the planets have the same radius), by using the identity  $\delta(ax + b) = \delta(x + b/a)/|a|$ ,

$$\Gamma_a(r_a) = \frac{1}{1 + \mu(\text{BF})} \left[ Z_0 \cdot \delta(r_a - r_0) + (Z_1 + Z_2) \cdot \mu(\text{BF}) \cdot \delta \left( r_a - \frac{r_0}{\sqrt{2}} \right) \right]. \quad (51)$$

This reproduces Eq.(9) of the draft (after removing the  $p_{\text{det}}$  terms)<sup>1</sup>. In the limit of  $\text{BF} \rightarrow 1$  ( $\mu \rightarrow \infty$ ), the above formula yields  $\Gamma_a = (Z^1 + Z^2) \delta(r_a - r_0/\sqrt{2})$ .

##### 5.1.2. Power Law Distribution of Planets

Imagine instead that we took  $\Gamma_i(r, M) = Z_i f(r) = Z_i r^\delta / \mathcal{N}_r$ , for  $f(r)$  the radius shape function,  $\mathcal{N}_r$  the shape function's normalization, and  $Z_i$  the number of planets per star. From Eq. 49, we get an apparent rate density

$$\Gamma_a(r_a) = \frac{r_a^\delta}{\mathcal{N}_r} \left[ \frac{Z_0}{1 + \mu(\text{BF})} + 2^{\frac{\delta+1}{2}} \frac{\mu(\text{BF})}{1 + \mu(\text{BF})} (Z_1 + Z_2) \right]. \quad (52)$$

Consider the case when the  $Z_i$ 's are identical. If we compare  $\Gamma_a(r_a)$  with  $\Gamma(r)$ , we must impose that  $r_a$  and  $r$  are indistinguishable (i.e.: the observer cannot tell them apart). Note that the true rate density for stars selected in the sample is

$$\Gamma_{\text{sel}}(r) = \frac{r^\delta}{\mathcal{N}_r} \left[ \frac{Z_0}{1 + 2\mu} + \frac{\mu}{1 + 2\mu} (Z_1 + Z_2) \right], \quad (53)$$

<sup>1</sup> The above equation also follows by just thinking about the problem and explicitly writing out the answer (cf. LB's handwritten notes 2017/11/26.9), though the formal derivation is a good check.



while the true rate density for a volume-limited sample of stars is

$$\Gamma_{\text{vl}}(r) = \frac{r^\delta}{\mathcal{N}_r} \left[ \frac{Z_0}{1 + 2(n_b/n_s)} + \frac{(n_b/n_s)}{1 + 2(n_b/n_s)}(Z_1 + Z_2) \right], \quad (54)$$

for  $n_b$  and  $n_s$  the number density of binaries and singles in a volume-limited sample.

The stand-alone term “true rate density” is ill defined, and should probably be avoided. In different contexts, it might refer to either a)  $\Gamma_{\text{sel}}(r)$ , b)  $\Gamma_{\text{vl}}(r)$ , or even c)  $\Gamma_0(r)$ , the rate around singles. The meaning that most people probably *want* true rate density to have is  $\Gamma_0(r)$  — simply “I have a star that is single; how many planets are there?”. This is also a sensible definition because it is *independent* of any complications of planet formation that would make  $Z_1$  or  $Z_2$  different from  $Z_0$ . One way to deal with this, rather than making the arbitrary choice, would be to just plot out all the possible options.

We could define a “correction factor” relative to  $\Gamma_0$ , or to  $\Gamma_{\text{sel}}$ , or to  $\Gamma_{\text{vl}}$ . If the  $Z_i$ ’s are equal, these are all identical:

$$\begin{aligned} X_\Gamma &\equiv \frac{\Gamma_a(r_a)}{\Gamma_0(r)} \Big|_{r_a \rightarrow r, Z_i\text{'s equal}} \\ &= \frac{\Gamma_a(r_a)}{\Gamma_{\text{sel}}(r)} \Big|_{r_a \rightarrow r, Z_i\text{'s equal}} \\ &= \frac{\Gamma_a(r_a)}{\Gamma_{\text{vl}}(r)} \Big|_{r_a \rightarrow r, Z_i\text{'s equal}} \\ &= \frac{1 + 2^{\frac{\delta+3}{2}} \mu}{1 + \mu}. \end{aligned} \quad (55)$$

For the case of  $\text{BF} = 0.1$ ,  $\mu \approx 0.153$ . Taking  $\delta = -2.92$  from Howard et al. (2012), we get a correction factor  $X_\Gamma = \Gamma_a/\Gamma = 1.004$ . In other words, the apparent rate density is an *overestimate* of the rate density of selected stars, with a relative error  $\delta\Gamma = |\Gamma - \Gamma_a|/\Gamma$  of 0.4%.

As a side-note, this particular model (twin binaries, power law planet distribution) is a good unit test for the numerics. More realistically, we’ll take a power law above some cutoff (to avoid the divergence for  $\delta < 0$ ), but Eq. 52 will tell us the apparent rate density within the region of parameter space where only the power law matters.

## 5.2. Power Law World

If we assume

$$L(M) \sim M^\alpha \sim R^\alpha, \quad (56)$$

we find

$$\mathcal{A}(q) = (1 + q^\alpha)^{-1/2}, \quad \mathcal{B}(q) = q^{-1}(1 + q^{-\alpha})^{-1/2}. \quad (57)$$

This gives a value of  $\mu(\text{BF})$  stated in Eq. 14. Now recall Eq. 47 for the apparent rate density. We restate it here:

$$\Gamma_a(r_a) = \frac{1}{1+\mu} \left( \Gamma_0(r_a) + \mu \left[ \int \frac{dq}{\mathcal{A}(q)} f(q) \Gamma_1 \left( \frac{r_a}{\mathcal{A}(q)} \right) + \int \frac{q dq}{\mathcal{B}(q)} f(q) \Gamma_2 \left( \frac{r_a}{\mathcal{B}(q)} \right) q^{2/3} \right] \right). \quad (58)$$

### 5.2.1. Same Planets (Model # 2)

For our “Model #2”, we again want  $\Gamma_i(r) = Z_i \cdot \delta(r - r_0)$ . To achieve this, write

$$f(q) = \frac{q^\beta (1 + q^\alpha)^{3/2}}{\mathcal{N}_q}, \quad (59)$$

for  $\mathcal{N}_q$  the normalization. For  $i = 1$ , call the integral on the left the  $[\dots]$  “ $I_1$ ”. Then

$$I_1(r_a) \equiv \int \frac{dq}{\mathcal{A}(q)} f(q) \Gamma_1 \left( \frac{r_a}{\mathcal{A}(q)} \right) \quad (60)$$

$$= \frac{1}{\mathcal{N}_q} \int dq q^\beta (1 + q^\alpha)^2 \Gamma_1(r_a (1 + q^\alpha)^{1/2}) \quad (61)$$

$$= \frac{Z_1}{\mathcal{N}_q} \int dq q^\beta (1 + q^\alpha)^2 \delta(r_a (1 + q^\alpha)^{1/2} - r_0) \quad (62)$$

$$= \frac{Z_1}{\mathcal{N}_q} \frac{2 r_0}{\alpha r_a^2} \left( \left( \frac{r_0}{r_a} \right)^2 - 1 \right)^{\frac{1-\alpha}{\alpha}} \int dq q^\beta (1 + q^\alpha)^2 \delta \left( q - \left( \left( \frac{r_0}{r_a} \right)^2 - 1 \right)^{\frac{1}{\alpha}} \right) \quad (63)$$

$$= \frac{Z_1}{\mathcal{N}_q} \frac{2 r_0}{\alpha r_a^2} \left( \left( \frac{r_0}{r_a} \right)^2 - 1 \right)^{\frac{\beta-\alpha+1}{\alpha}} \left( \frac{r_0}{r_a} \right)^4. \quad (64)$$

The penultimate line used the fact that  $r_a (1 + q^\alpha)^{1/2} = r_0$  has only one root:

$$q(r_a) = \left( \left( \frac{r_0}{r_a} \right)^2 - 1 \right)^{1/\alpha}, \quad (65)$$

and for any function  $f(q)$  with a single root  $q_i$ ,

$$\delta(f(q)) = \frac{\delta(q - q_i)}{|f'(q_i)|}. \quad (66)$$

Eq. 64 is a good unit test for Monte Carlo numerics.

Attempting the same thing for  $i = 2$ ,

$$I_2(r_a) \equiv \int \frac{q dq}{\mathcal{B}(q)} f(q) \Gamma_2 \left( \frac{r_a}{\mathcal{B}(q)} \right) q^{2/3} \quad (67)$$

$$= \frac{Z_2}{\mathcal{N}_q} \int dq q^{\beta+\frac{8}{3}} (1 + q^\alpha)^{3/2} (1 + q^{-\alpha})^{1/2} \delta(r_a q (1 + q^{-\alpha})^{1/2} - r_0). \quad (68)$$

Unfortunately, analytic roots of  $r_a q(1 + q^{-\alpha})^{1/2} = r_0$  do not exist. If the roots are found numerically then

$$I_2(r_a) = \frac{Z_2}{\mathcal{N}_q} \sum_{q_i \in \{\text{roots}\}} q_i^{\beta + \frac{8}{3}} (1 + q_i^\alpha)^{3/2} (1 + q_i^{-\alpha})^{1/2}, \quad (69)$$

but this doesn't really give much insight.

### 5.2.2. Power law planets (Model # 3)

Power law planets, a realistic mass ratio distribution. In other words, now assume

$$\Gamma_i(r, M) = Z_i f(r) = Z_i r^\delta / \mathcal{N}_r. \quad (70)$$

Surprisingly, this might give more insight than Model #2.

From Eq. 47, we can write

$$\Gamma_a(r_a) = \frac{1}{1 + \mu} (\Gamma_0(r_a) + \mu [I_1(r_a) + I_2(r_a)]), \quad (71)$$

where  $I_1(r_a)$  and  $I_2(r_a)$  are defined in the same way as in Eqs. 60 and 67. Doing the calculation, we find

$$I_1(r_a) = \frac{Z_1}{\mathcal{N}_q \mathcal{N}_r} r_a^\delta \int dq q^\beta (1 + q^\alpha)^{\frac{\delta}{2} + 2} \quad (72)$$

and

$$I_2(r_a) = \frac{Z_2}{\mathcal{N}_q \mathcal{N}_r} r_a^\delta \int dq q^{\beta + \delta + \frac{8}{3}} (1 + q^\alpha)^{\frac{3}{2}} (1 + q^{-\alpha})^{\frac{\delta + 1}{2}}. \quad (73)$$

The net effect of binarity is to just *scale* the overall distribution!

Note that the true rate density for the selected stars is

$$\Gamma_{\text{sel}}(r) = \frac{r^\delta}{\mathcal{N}_r} \left[ \frac{Z_0}{1 + 2\mu} + \frac{\mu}{1 + 2\mu} (Z_1 + Z_2) \right]. \quad (74)$$

The correction factor with respect to selected stars,  $X_\Gamma = \Gamma_a / \Gamma_{\text{sel}}$ , will be

$$X_\Gamma \equiv \frac{\Gamma_a(r_a)}{\Gamma_{\text{sel}}(r)} \Big|_{r_a \rightarrow r, Z_i \text{'s equal}} \quad (75)$$

$$= \frac{1 + 2\mu}{3(1 + \mu)} \left[ 1 + \mu \left( \int dq q^\beta (1 + q^\alpha)^{\frac{\delta}{2} + 2} + \int dq q^{\beta + \delta + \frac{8}{3}} (1 + q^\alpha)^{\frac{3}{2}} (1 + q^{-\alpha})^{\frac{\delta + 1}{2}} \right) \right]. \quad (76)$$

For  $\alpha = 3.5$ ,  $\beta = 0$ ,  $\delta = -2.92$ , the term in (...)  $\approx 1.41425 \approx \sqrt{2}$ , where the latter surprising numerical fluke holds to the first four decimal places. For  $\text{BF} = 0.44$ , this yields  $\Gamma_a / \Gamma_{\text{sel}} = 1.28$ , an apparent rate which is *higher* than the true rate by 28%! Note that this result is the same as  $\Gamma_a / \Gamma_0$  and  $\Gamma_a / \Gamma_{\text{vl}}$  in the limit where all the  $Z_i$ 's are equal.

More realistically, we'll take a power law above some cutoff (to avoid the divergence for  $\delta < 0$ ), but nonetheless the above indicates we're going to need to write a different argument about whether binarity might explain the HJ rate discrepancy (apparently, it goes the wrong way!)