

C. R. Acad. Sci. Paris, t. 0, Série IV, p. 1-7, 2001
Techniques astronomiques/Astronomical techniques

Exo-planet detection with the COROT space mission. I. A multi-transit detection criterion

Pascal BORDÉ^a, Daniel ROUAN^b, Alain LÉGER^b

^a DESPA – CNRS, observatoire de Paris, 92195 Meudon, France

^b Institut d'astrophysique spatiale, 91405 Orsay, France

E-mail: pascal.borde@obspm.fr; daniel.rouan@obspm.fr; alain.leger@ias.fr

(Reçu le 24 janvier 2001, accepté le 25 juin 2001)

Abstract.

In this Note, we present a detection criterion for exo-planets to be used with the space mission COROT. This criterion is based on the transit method, which suggests the observation of star dimming caused by partial occultations by planetary companions. When at least three transits are observed, we show that a cross-correlation technique can yield a detection threshold, thus enabling the evaluation of the number of possible detections assuming a model for the stellar population in the Galaxy. © 2001 Académie des sciences/Éditions scientifiques et médicales Elsevier SAS

COROT / photometry / data analysis / statistical methods / exo-planets

Détection d'exo-planètes avec la mission spatiale COROT. I. Un critère de détection pour les multi-transits

Résumé.

Dans cette Note, nous présentons un critère de détection d'exo-planètes destiné à la mission spatiale COROT. Ce critère est fondé sur la méthode des transits qui suggère de rechercher des affaiblissements de flux stellaires dus à des occultations partielles par des compagnons planétaires. Lorsqu'au moins trois transits sont observés, nous montrons qu'une technique d'inter-corrélation conduit à un seuil de détection utilisable pour évaluer le nombre de détections potentielles (dans le cadre d'un modèle de population stellaire de la Galaxie). © 2001 Académie des sciences/Éditions scientifiques et médicales Elsevier SAS

COROT / photometrie / traitement des données / méthodes statistiques / exo-planètes

Version française abrégée

Une des deux tâches de COROT, une petite mission du CNES dont le lancement est prévu en 2004, sera de détecter des exo-planètes par la méthode des transits. Cette méthode consiste à chercher des baisses de flux stellaires dues à des occultations très partielles par des compagnons planétaires. Ces petites baisses de flux sont caractérisées par leur amplitude (équation (1)), leur durée (équation (2)) et leur probabilité géométrique qui dépend de l'inclinaison de l'orbite (équation (3)). Grâce à sa très haute sensibilité photométrique, COROT sera capable de détecter des planètes dont la taille avoisine celle de la Terre. Pour ce programme, COROT utilise deux CCD 2048 × 2048 lui permettant de surveiller de 5000 à 12000 étoiles pendant 150 jours, pour chacun des 5 champs prévus.

En amont de la procédure de traitement, on moyenne les données sur la durée supposée du transit pour augmenter le rapport signal à bruit S/B . Soit k le nombre de transits cherchés dans l'échantillon. Le signal de multi-transit est modélisé par la somme d'un peigne de Dirac et d'un bruit blanc gaussien (équations (4))

Note présentée par Pierre ENCRENAZ.

P. Bordé et al.

et (5)). On forme alors un produit d'inter-corrélation avec tous les peignes de Dirac possibles (*figure 1*). Une grande valeur de ce produit C_k vis-à-vis des fluctuations statistiques est interprétée comme la signature d'un transit (*figure 2* et équation (7)). L'équation (8) donne la probabilité que la valeur observée de C_k résulte d'une fluctuation statistique pour un sous-ensemble associé à k transits, et l'équation (9) donne le degré de confiance associé à toutes les valeurs de k (voir aussi la *figure 3*). Pour un degré de confiance donné, on peut calculer le S/B nécessaire à la détection (équation (12)), ainsi que le rayon planétaire minimal associé (équation (15)).

Nous avons utilisé le modèle de population stellaire développé à l'Observatoire de Besançon pour avoir une distribution réaliste d'étoiles en terme de magnitudes et de types spectraux, dans une direction donnée de la Galaxie. Nous avons supposé que toutes les étoiles observées ont 100% de chance d'avoir un compagnon planétaire à la distance considérée, et nous avons choisi un taux de fausse alarme de 10^{-5} . Sous ces hypothèses et pour toute la durée de la mission, nous avons tracé un réseau de courbes qui montre le nombre de détections attendues pour différentes valeur du rayon minimal détectable (*figure 4*). Il est ensuite possible d'appliquer à ces résultats le facteur d'échelle de son choix. Par exemple, on en déduit que COROT pourrait détecter plusieurs dizaines de Jupiters chauds si on leur affecte une probabilité de 2%. On remarque par ailleurs que COROT est théoriquement capable de détecter des planètes de taille terrestre, si celles-ci orbitent très près de leur étoile.

1. Overview of COROT

1.1. Introduction

Exo-planet detection has now become a very active field in astrophysics. In august 2000, about fifty Jupiter-like planets have been discovered around nearby stars. The next challenging step is to find much less massive planets with the hope to detect life on some of them afterwards. COROT, a CNES space mission to be launched in 2004 [1], is currently a funded project capable of detecting planets with a radius close to $1R_{\oplus}$ thanks to the 'transit method' [2]. In this section, we briefly describe the transit method and the COROT satellite. In Section 2, we present a detection criterion for multi-transits, and in Section 3, we discuss the number of possible dectections assuming a model for the stellar population in the Galaxy. The specific case of mono-transits will be considered in a forthcoming article.

1.2. The transit method

The transit method for searching extrasolar planets is based on the idea that if a planet crosses the disk of its parent star, it would result in a dimming of the star's observed light [3,4]. The expected amplitude of the relative stellar flux dimming is:

$$\varepsilon \equiv \frac{\Delta F}{F} = \left(\frac{R_P}{R_*} \right)^2 \quad (1)$$

where F is the flux, R_P the radius of the planet and R_* the radius of the parent star. We define the impact parameter, expressed in stellar radii, as the apparent height of the planet trajectory above the star's equator. For an impact parameter equal to 0.5, the transit duration is:

$$tr = \frac{\sqrt{3} R_*}{\sqrt{GM_*}} \sqrt{a} \quad (2)$$

where a is the orbital radius. This duration is respectively 11.2 and 25.7 hours for the Earth and Jupiter in the solar system, and 3 hours for 51 Peg-b [5]. The geometrical probability that the orbital inclination on

Exo-planet detection with COROT

the sky is close enough to 90° to make a transit visible is:

$$p_g = \frac{R_\star}{a} \quad (3)$$

p_g is equal to 0.5% for the Earth, 0.1% for Jupiter and 16% for 51 Peg-b. This method has already been successfully used from ground and space on the planetary companion of HD209458, previously detected by radial velocity techniques, e.g. [6].

1.3. Focusing on multi-transits

COROT features a 27 cm telescope and four 2048×2048 CCDs, two of which are devoted to the exoplanets program. During its 30 month mission, the satellite will monitor 5 fields of 5000 to 12000 stars ($11 \leq m_V \leq 16.5$, galactic latitude = $15-20^\circ$), each of them for 150 days, thus leading up to 60000 lightcurves. The task will be then to look in the data for the signatures of planetary transits. In order to estimate the number of potential detections by COROT, we here propose consideration of a rather basic method based on a cross-correlation technique. More accurate techniques are under study at the laboratoire d'astronomie spatiale in Marseille [7], or have been already suggested [8]. Here we will consider only multi-transits, which means that a given planet will have to transit at least 3 times in front of its parent star to be detected by this method. The sought after signal looks like a repeated dimming in the parent star lightcurve, at even time intervals corresponding to the planetary orbital period P .

2. Principle of data processing

2.1. Raw data averaging

To increase the signal to noise ratio (SNR) and also to reduce the amount of data, we begin by averaging the raw data, i.e. samples of the total flux F every 16 minutes over 150 days, on the duration of a presumed transit tr (up to 15 hrs). This operation leads to a $N = 150 \times 24/tr$ point vector in which a potential event is reduced to one point. It means that in the whole process of the transit search, all relevant values of tr should be tried.

2.2. Transit signal modelling

The transit signal s we are looking for can be modeled by the sum of a Dirac comb $\alpha\Pi_k$ (k teeth, amplitude α) and a Gaussian white noise b of standard deviation σ_b :

$$\forall i \in \{1, \dots, N\}, \quad s[i] = -\alpha\Pi_k[i] + b[i] \quad (4)$$

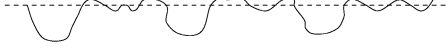
$$\text{where } \Pi_k[i] = \sum_{m=0}^{k-1} \delta\left[i - m\frac{N}{k}\right] \quad (5)$$

(δ is Kroenecker's symbol). The term b is the sum of several types of noise:

- the photon noise: $\sigma_{\text{ph}} = \sqrt{N_{\text{ph}}}$ (N_{ph} refers actually to the number of stellar photo-electrons detected by the detector during tr);
- the electrical read-out noise: $\sigma_{\text{ro}} = 12e^- \cdot \text{pixel}^{-1}$;
- the background noise (essentially zodiacal light): $\sigma_{\text{bg}} = 16e^- \cdot \text{pixel}^{-1}$ (exposure time is 32 s);
- the stellar irradiance variability noise depending on the averaging interval whose length is tr : $\sigma_{\text{st}}(tr) = N_{\text{ph}}(tr) \sigma_{\text{st}}(tr)_{\text{ppm}}$. For G and K stars, an estimated value of this quantity can be obtained through filtering of the SOHO-VIRGO data acquired on the Sun [9]: $\sigma_{\text{st}}(tr)_{\text{ppm}} \sim 30-75$ ppm for $tr = 2-15$ hrs;
- the noise introduced by the random pointing error of the satellite is supposed to be perfectly corrected by onboard processing.

P. Bordé et al.

Raw data



Dirac comb



Averaged data

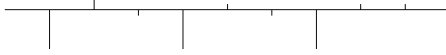


Figure 1. Raw data are averaged on the duration of the presumed transit, then cross-correlated with a Dirac comb.

Figure 1. *Les données brutes sont moyennées sur la durée supposée du transit, puis inter-corrélées avec un peigne de Dirac.*

For simplicity, all those noises are supposed independant, white and Gaussian noises. Consequently, we get:

$$\sigma_b = \sqrt{N_{\text{ph}} + n(\sigma_{\text{ro}}^2 + \sigma_{\text{bg}}^2) + \sigma_{\text{st}}^2} \quad (6)$$

where n denotes the total number of read-out pixels during the transit.

2.3. Detection with cross-correlation

The idea is to compute cross-correlation products between the data and a Dirac comb (amplitude unity) which has the shape of a noise free multi-transit signal (see *figure 1*). For a given value of k , we get N/k products of this kind, denoted C_k :

$$C_k = \frac{1}{N} \sum_{i=0}^{N-1} s[i] \Pi_k[i] = \frac{1}{N} \left[\alpha k + \sum_{m=0}^{k-1} b \left[m \frac{N}{k} \right] \right] \quad (7)$$

Trying all values of k between 3 and 50 ($3 \leq P \leq 50$ days) means computing $\sum_{k=3}^{50} N/k = \gamma N$ products ($\gamma \approx 3.0$). The presence of a transit must be somehow related to a high value of C_k , but the question is: where to draw the line?

2.4. Statistical analysis

The answer to this question lies in a statistical analysis of the problem. To begin with, let us consider the set of N/k products C_k for a given k . C_k can be seen as a random variable. Because we assumed that b is a Gaussian noise with a null mean and a standard deviation σ_b , the probability law of C_k should also be Gaussian with a standard deviation $\sigma_C = \sqrt{k} \sigma_b / N$. Its mean value should be equal to $\alpha k / N$ in case of a star actually showing transits or zero otherwise. Let p_k be the probability of having $C_k \leq \beta_k \sigma_C$ in case of noise only (no transit):

$$p_k = \Pr\{C_k \leq \beta_k \sigma_C\} = \frac{1}{2} \left[1 + \operatorname{erf} \left(\frac{\beta_k}{\sqrt{2}} \right) \right] \quad (8)$$

The probability that all the N/k values of C_k will remain inferior to $\beta_k \sigma_C$ is $a_k = p_k^{N/k}$. This gives the level of confidence that the statistical noise would not generate a high value of C_k that would be mistaken for a transit (see *figure 2*). If the statistics bear on all values of k (γN cross-correlation products), the level of confidence becomes:

$$a = \prod_{k=3}^{50} a_k = \prod_{k=3}^{50} \left\{ \frac{1}{2} \left[1 + \operatorname{erf} \left(\frac{\beta_k}{\sqrt{2}} \right) \right] \right\}^{N/k} \quad (9)$$

Exo-planet detection with COROT

Figure 2. The probability distribution of the cross-correlation products C_k is centered around zero, because the absence of a transit is the general rule. A large value of a C_k with respect to σ_C is indicative of a transit with a level of confidence a_k . This confidence is also measured by β , which is supposed common to all k .

Figure 2. La distribution de probabilité des produits de corrélation C_k est centrée sur zéro parce que l'absence de transit est la règle générale. Une grande valeur de C_k par rapport à σ_C indique la présence d'un transit avec un degré de confiance a_k . Cette confiance est aussi mesurée par β , supposé commun à toutes les valeurs de k .

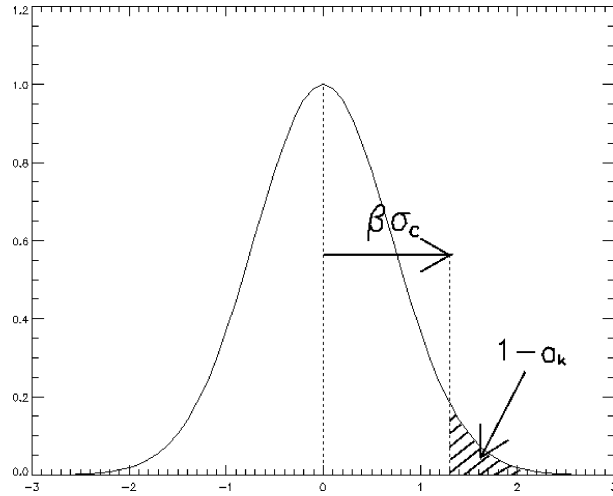
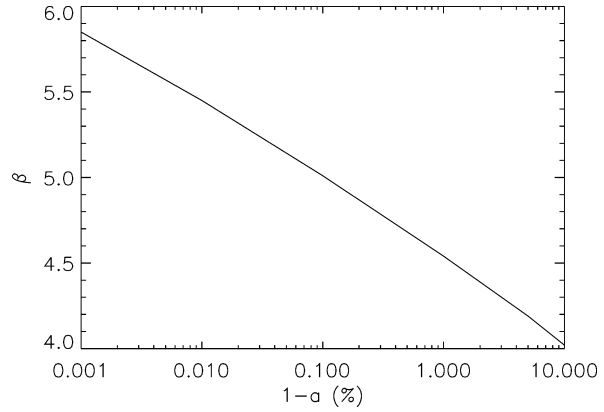


Figure 3. β as a function of the false alarm percentage $1 - a$. This quasi logarithmic relation allows the obtaining of high confidence levels with reasonable SNR.

Figure 3. Paramètre β en fonction du taux de fausse alarme. La dépendance quasi-logarithmique autorise à imposer un haut niveau de confiance $1 - a$ avec un S/B raisonnable.



One can numerically show that β depends weakly on k provided a is close enough to one (no more than 10% variation). That is why we will assume that β is common to all k , in order to solve for β , given the global level of confidence a :

$$\operatorname{erf}\left(\frac{\beta}{\sqrt{2}}\right) + 1 - 2a^{1/(\gamma N)} = 0 \quad (10)$$

For example, the transit of a planet orbiting a sun-like star at 0.05 AU (such as 51 Peg-b) would last approximately 3 hrs if the line of sight belongs to the orbital plane. In this case $N = 1200$, and solving (10) for $a = 99.9\%$ leads to $\beta = 5.0$ (see figure 3).

2.5. SNR detection criterion

Assume that a global level of confidence a has been chosen and has yielded up a value of β : a detection could be claimed with that confidence, if among the γN cross-correlation products, one is greater than $\beta\sigma_C$. As it is a signature of a transit, its value can be written as $\alpha k/N$, so the detection criterion translates into the inequation:

$$\frac{\alpha k}{N} \geq \frac{\sqrt{k} \beta \sigma_b}{N} \quad (11)$$

P. Bordé et al.

Introducing the SNR on a single event (a single dimming of the lightcurve) defined by $S/N = \alpha/\sigma_b$, we get:

$$\frac{S}{N} \geq \frac{\beta}{\sqrt{k}} \quad (12)$$

As expected, the required SNR increases with the level of confidence and decreases with the number of observed transits. To go on with the previous example, $P = 4.1$ days so $k = 36$ and $S/N \geq 0.7$! As we see here, this cross-correlation technique is powerful enough to detect repetitive transits *with lower than unity SNR on single events*.

2.6. Link with the planetary radius

The amplitude of the relative dimming of the parent star lightcurve is connected to the radius of the transiting planet R_P by:

$$\frac{\Delta F}{F} = \left(\frac{R_P}{R_\star} \right)^2 \quad (13)$$

$$\text{besides } \frac{\Delta F}{F} = -\frac{\alpha}{N_{\text{ph}}} = -\frac{S}{N} \frac{\sigma_b}{N_{\text{ph}}} \quad (14)$$

Thus (12) translates into:

$$R_P \geq R_\star \left(\frac{\beta}{\sqrt{k}} \frac{\sigma_b}{N_{\text{ph}}} \right)^{1/2} \quad (15)$$

which gives the minimum planetary radius that can be detected with a level of confidence a . In our example, if we assume for the star $m_V = 14$ and given the characteristics of COROT, we compute: $R_P \geq 1.5R_\oplus$.

3. Expected number of detections

3.1. Assumptions

To have a realistic distribution of stars per spectral type and magnitude interval in the observed fields, we have used a model of the stellar population in the Galaxy developed at Besançon Observatory [10]. Only stars on the Main Sequence have been considered (type V stars). We will assume that every observed star has 100% probability of having a planet orbiting at the distance tested, and that orbits are circular. For every distance, we compute the minimum radius the planet should have for our criterion to be able to detect it with a false alarm level of 10^{-5} .

3.2. Algorithm

Let a be the orbital of the planetary orbit. We define the reduced orbital radius by $a_r = a(L_\star/L_\odot)^{-0.5}$ so that $a_r = 1$ AU would always correspond to the distance where the planet receives as much flux from its parent star as the Earth from the Sun. Computations have been done in the range $a_r = 0.03-1$ AU.

Given the star characteristics: mass M_\star , radius R_\star , effective temperature T_\star , luminosity L_\star , spectral type Sp and magnitude m_V , we compute:

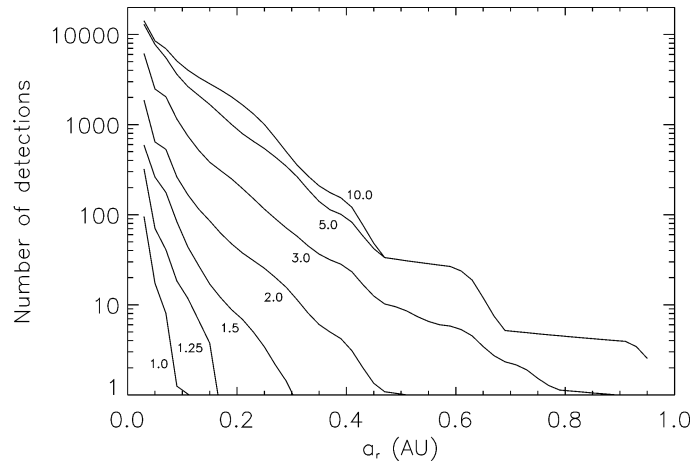
- the orbital radius a in AU;
- the revolution period P in days;
- the probability of seeing 3 transits in 150 days: $p_g \times \frac{150}{P}$ ($P \leq 50$ d);
- the transit duration tr (assuming a mean impact parameter of 0.5);
- the number of photo-electrons N_{ph} received during tr .

Then, the SNR detection criterion imposes a minimum value for the signal $\varepsilon = (S/N)/\sqrt{N_{\text{ph}}}$, and consequently for the detectable planetary radius $R_P = R_\star \sqrt{\varepsilon}$.

Exo-planet detection with COROT

Figure 4. This plot shows the number of expected detections during the full life-time of the mission, provided 100% of the stars have a planetary companion of radius R_P at the distance a_r , and assuming a false alarm rate of 10^{-5} . The number attached to each curve is the minimum value of R_P (expressed in Earth unit) enabling the detection.

Figure 4. *Nombre de détections attendues pour toute la durée de la mission, à condition que 100% des étoiles aient un compagnon planétaire à la distance a_r , et en considérant un taux de fausse alarme de 10^{-5} . Le nombre associé à chaque courbe est la valeur minimale détectable du rayon de la planète en rayons terrestres.*



3.3. Results

We have plotted the number of detections as a function of the reduced orbital distance for various planetary radii (*figure 4*) and for a false alarm rate of 10^{-5} . For every curve, it is the minimum detectable radius (expressed in Earth unit) that is considered. Results are given for the whole mission. This way of presenting our results was chosen because of the unknown frequency of the different planetary types. Anyone can apply to our curves the scaling factor of his choice. For instance, assuming a 2% probability of existence, COROT could detect several tens of ‘hot Jupiters’ enhancing significantly the statistics on this class of planets. What is more, COROT has the potential to spot ‘hot Earths’ if any, since for example around 4 events would be expected if 20% of the stars exhibit earth planets at 0.05 AU.

References

- [1] Baglin A. et al., Asteroseismology from space – The COROT experiment, *New Eyes to See Inside the Sun and Stars*, IAU 185 (1998) 301.
- [2] Rouan et al., Searching for exosolar planets with the COROT space mission, *Phys. Chem. of the Earth Part C* 24 (5) (1999) 567–571.
- [3] Rosenblatt F., A two-color photometric method for detection of extra-solar planetary systems, *Icarus* 14 (1971) 71–93.
- [4] Schneider J., Extra-solar planets transits: detection and follow up, in: *VLT Opening Symposium Antofagasta*, Springer, 1999.
- [5] Mayor M., Queloz D., A Jupiter-mass companion to a solar-type star, *Nature* 378 (1995) 355–359.
- [6] Charbonneau et al., Detection of planetary transits across a sun-like star, *Astrophys. J.* 529 (2000) L45–L48.
- [7] Defaÿ C. et al., A Bayesian method for the detection of planetary transits, *Astron. Astrophys.*, submitted.
- [8] Jenkins et al., A matched filter method for ground-based sub-noise detection of terrestrial extrasolar planets in eclipsing binaries: Application to CM Draconis, *Icarus* 119 (1996) 244–260.
- [9] Fröhlich et al., First results from VIRGO, the experiment for helioseismology and irradiance monitoring on SOHO, *Solar Phys.* 170 (1997) 1–25.
- [10] Robin A., Crézé M., Stellar population in the Milky Way – A synthetic model, *Astron. Astrophys.* 157 (1986) 71–90.