Periodic Emission of Gamma-Rays in Binary Stars

Author: Juan Tornero Lucas

Facultat de Física, Universitat de Barcelona, Diagonal 645, 08028 Barcelona, Spain.

Advisor: Josep Maria Paredes

Abstract In this paper we search for an orbital modulation of the VHE γ-ray flux in LS 5039, a High Mass X-ray Binary that consists of a compact object in an 3.9 days orbit around a massive star. For this purpose we apply to archive HESS data of LS 5039 two methods: Least-Squares Spectral Analysis (LSSA) or Lomb-Scargle test, which has been proved effective in this task previously, and Period Determination Using Phase Dispersion Minimization (PDM). We'll see that both methods verify this periodic modulation and that it matches the orbital period of the binary system.

I. INTRODUCTION

LS 5039 is one of the few confirmed galactic High Mass X-ray Binaries (HMXBs) with recurrent TeV emission, along with PSR B1259-63 or LSI+61303.

X-ray binaries are a class of binary stars that happen to be luminous in X-rays, which are produced by matter falling from the donor component -a star- to the accretor, a compact object -usually a neutron star or a black hole-. In particular, HMXBs have as a donor a massive star, usually an O, B or a blue supergiant. In the case of LS 5039 the nature of the compact object is still unknown, nevertheless we have knowledge its companion is a O6.5V star in a slightly eccentric ($e \approx 0.35$) 3.90603 ± 0.00017 day orbit [1].

For decades, the existance of γ-ray emitting binaries was speculated. The new generation of observatories allowed the discovery in the mid-2000's of HMXBs emitting high-energy (HE, 0.1-100 GeV) or very high-energy(VHE, >100 GeV) γ-rays [2]. Binary stars have a specific property that distinguish them from any other object: their physical conditions vary on a regular timescale as the components revolve on their orbit.

In this work we look for a modulation in the emission of VHE γ-rays in LS 5039 that matches the orbital period of this binary system. This modulation is the result of two phenomenons: between the apastron and the inferior conjunction, pairs of very energetic positrons and electrons, produced by the compact object, induce by inverse Compton the increase of the energy in the photons that come from the star: $h\nu + e^- \rightarrow h\nu'$, with $h\nu' >> h\nu$. Therefore is in this area where we detect the maximum in the VHE γ-ray flux. As a counterpart of this effect, around the superior conjunction, where the stellar wind is more intense, we have an absorption of photons that produce pairs of positrons and electrons $\gamma + \gamma \rightarrow e^+ + e^-$. Here is where we detect the minimun in the VHE γ-rays flux. So, though it may sound non-intuitive, when the compact object is closer to the star we detect minimums and when is further maximums. In Figure 7 we can appreciate this clearly.
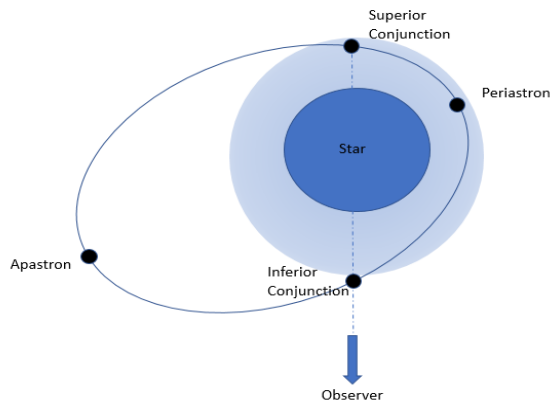

FIG. 1: *Scheme of LS-5039 system. In the periastron (phase $\phi = 0$) the compact object and the star are at minimum distance. Superior ($\phi = 0.058$) and inferior ($\phi = 0.716$) conjunctions phases correspond to the time of co-aligment along our line-of-sight of the compact object and stellar companion. The light blue cloud is a qualitative representation of the zones with intense stellar wind. The maximum of VHE γ ray flux detection lag somewhat behind the apastron epoch, aligning better with the inferior conjunction; and the minimum flux occurs at $\phi \approx 0.2$, a bit further than superior conjunction, when the compact object aligns behind the star. The orbit is actually inclined at an angle in the range $13° < i < 64°$ with respect to the view above.*

II. HESS DATA

The observations were taken with HESS (Aharonian et al. 2006b), an array of four identical Imaging Atmospheric Cherenkov Telescopes (IACT) sensitive to γ-rays above 0.1 TeV. The total dataset comprises 160 runs (or pointings) representing 69.2 h observations from both 2004 and 2005 and comprehend flux energies above 1 TeV.

III. PRECEDENT ANALYSIS

In this work we try to emulate the results obtained by Aharonian et al. (2006) [3], where they managed to find
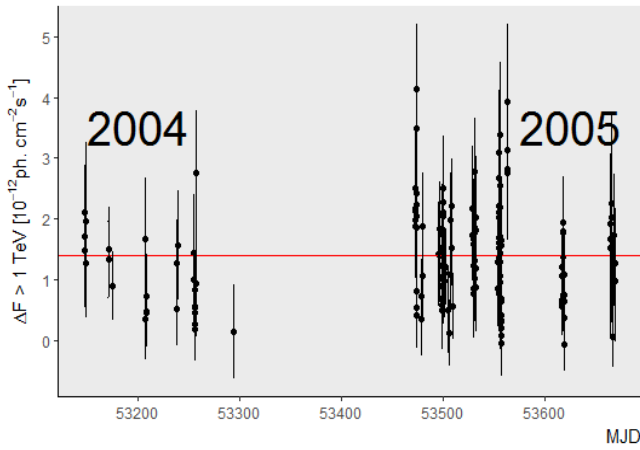
FIG. 2: *Initial disposition of data points in the runwise VHE γ-ray flux at energies > 1 TeV. Horizontal line represents the mean of the sample in order to appreciate the scattering around it of the data points.*

FIG. 3: *Lomb-Scargle (LS) periodogram of the VHE runwise flux for LS 5039 (chance probability to obtain the LS power vs. frequency). The highest peak, with a p-value of $9.4889 \cdot 10^{-13}$ is around the frequence 0.25549 days^{-1}.*

a 3.9078 ± 0.015 days orbital period that was consistent with the 3.90603 ± 0.00017 days period from Casares et al. (2005).

First, we have run a Lomb-Scargle Test, as they did in order to replicate their result. Lomb-Scargle Test, or Least-Squares Spectral Analysis (LSSA) is a method to estimate a frequency spectrum through a least squares fit. This test assumes a sinusoidal behavior, what is supposed to be true in the case of an orbital modulation.

Let it be w our estimated period, then the flux variation would fit $\phi(t) = A \sin(wt) + B \cos(wt)$ for some constants A, B. Suppose we have the pair $O_i = (x_i, t_i)$, $\forall i \in \{1, \ldots, N\}$ of the energy x_i at time t_i. Our set of data would be $O = \cup O_i$. LSSA first finds a time delay τ such that would make mutually orthogonal this pair of sinusoids in ϕ at any time t_j, and also reduce the effect of different powers in the basis of these functions. This makes LSSA equivalent to Lomb's least-squares method applied to ϕ. The time delay τ is inferred from the formula

$$\tan(2w\tau) = \frac{\sum\limits_{j=1}^{N} \sin(2wt_j)}{\sum\limits_{j=1}^{N} \cos(2wt_j)}$$

Then the periodogram at frequency w is estimated as:

$$P_O(w) = \frac{1}{2}\left(\frac{\left[\sum\limits_{j=1}^{N} x_j \cos(w(t_j - \tau))\right]^2}{\cos^2(w(t_j - \tau))} + \frac{\left[\sum\limits_{j=1}^{N} x_j \sin(w(t_j - \tau))\right]^2}{\sin^2(w(t_j - \tau))}\right)$$

If w is the period we look for, in the periodogram there will be a big peak at this frequency. We use the LSSA implementation from RStudio package *lomb* by Thomas Ruf [4], and perform it over our 160 data points set, obtaining the next periodogram.

The result of our LS Test delivered a 3.914 days orbital period, quiet consistent with the previous result of
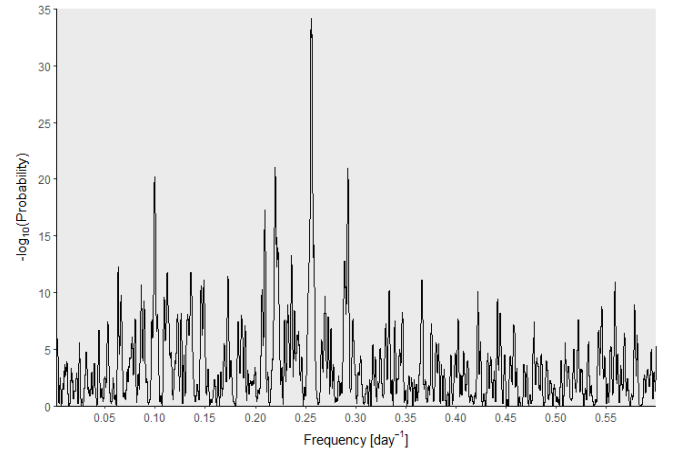
3.9078 ± 0.015, considering at Aharonian et al. (2006) they made a calculation with a p-value of less than 10^{-15}, while ours is around 10^{-12}.

IV. PERIOD DETERMINATION USING PHASE DISPERSION MINIMIZATION

Our alternative method to the LS Test, which force us to assume a sinusoidal behaviour, is the Period Determination Using Phase Dispersion Minimization (PDM) by Stellingwerf (1978)[5].

In this methodology we consider again x our energy values at times t, therefore the i-th observation is given by $(x_i, t_i), i \in \{1, \ldots, N\}$, with $N = 160$ in this case.

Let σ^2 and μ be the variance and mean of x, respectively

$$\sigma^2 = \frac{\sum\limits_{i=1}^{N} (x_i - \mu)}{N - 1} \quad (1)$$

For any sample of x we define s^2 exactly as σ^2 in (1). Suppose we have M samples, having variances $s_j, j \in \{1, \ldots, M\}$. Each of these samples would contain n_j data points and then, the overall variance for all samples is

$$s^2 = \frac{\sum\limits_{j=1}^{M} (n_j - 1)s_j^2}{\sum\limits_{k=1}^{M} n_j - 1} \quad (2)$$

Our objective is to minimize the variance of the data with respect mean light curve. Let α be our trial period, we define the phase vector ϕ as: $\phi_i = \frac{t_i}{\alpha} - [\frac{t_i}{\alpha}], i \in \{1, \ldots, N\}$. Where $[\cdot]$ means the integer part of the value.

We'll pick our samples under the condition that all data points of a sample j have similar ϕ_i. The variance of these samples will give us a measure of the scatter around the mean light curve defined by the means of the x_i in each sample, considered as a function of ϕ. Our statistic value to evaluate the accuracy of the period α will be

$$\Theta = \frac{s^2}{\sigma^2}, \quad (3)$$

where σ^2 and s^2 are given by equations (1) and (2) respectively. If α is not the right period, then $s^2 \approx \sigma^2 \rightarrow \Theta \approx 1$. When α is correct, s^2 reaches a minimum and $\Theta \approx 0$.

We'll divide our phase space in N_b intervals and take N_c covers of these N_b bins. Therefore we'll obtain $M = N_b N_c$ bins, each of length $\frac{1}{N_b}$ and whose midpoints will be uniformly spaced at a distance of $\frac{1}{N_b N_c}$. We'll denote this bin structure by (N_b, N_c).

Next, we'll study the character of Θ near a minimum. In an oscillation of amplitude A, maximum variance in x is $\sigma_0^2 = \frac{A^2}{12}$. We introduce a noise component σ_N^2, the signal-to-noise ratio is $\epsilon = \frac{\sigma_0}{\sigma_N}$ and overall variance $\sigma^2 = \sigma_0^2 + \sigma_N^2$.

Bin variance s^2 will have three components. The first one depends on the distance from line center ($\delta\phi$ or δf in period and frequency, respectively). The second one is, as in σ^2, a noise contribution. The last contribution comes from a bin width variance.

If we take $g(\phi)$ as the mean light curve with dependency on the phase ϕ, according to Stellingwerf, the average pendent in a sinusoidal function as ours is $< g' > \approx 2A = 2\sqrt{12}\sigma_0 = 4\sqrt{3}\sigma_0$.

Now let's figure out the variance due to a change in the trial period α, which would induce a scatter in the phase over a range $\Delta\phi = 0 \rightarrow \frac{\Delta\alpha}{\alpha}\frac{T}{\alpha}$, where $T = t_n - t_1$ is the time base. Therefore

$$s_\phi^2 = \frac{1}{12}\left(\frac{T}{\alpha}\frac{\Delta\alpha}{\alpha}\right)^2 < g' >^2 = 4\left(\frac{T}{\alpha}\frac{\Delta\alpha}{\alpha}\right)^2\sigma_0^2 = 4(T\Delta f)^2\sigma_0^2, \quad (4)$$

as $\frac{\Delta\alpha}{\alpha} = \frac{\Delta f}{f}$ for frequency $f = \alpha^{-1}$. Finally, we'll have also a variation of g across bins of width N_b^{-1} in the form

$$s_b^2 = \frac{1}{12}\left(\frac{g'}{N_b}\right)^2 = 4\frac{\sigma_0^2}{N_b^2}. \quad (5)$$

Then, the total bin variance is

$$s^2 = s_\phi^2 + s_b^2 + \sigma_N^2 \quad (6)$$

And we can develop the expression of Θ as

$$\Theta = \frac{s_\phi^2 + s_b^2 + \sigma_N^2}{\sigma_0^2 + \sigma_N^2} = \frac{1 + 4\epsilon^2[N_b^{-1} + (T\Delta f)^2]}{1 + \epsilon^2}. \quad (7)$$

We obtain from here that near a minimum the behaviour of the Θ statistic is parabolic. And looking at the $\Theta = 1$ level we can obtain the half-width

$$\Delta f_{1/2} = \frac{1}{T}\sqrt{\left(\frac{1}{4} - \frac{1}{N_b^2}\right)}. \quad (8)$$

As all our simulations use $N_b \geq 5$, we can approximate $\Delta f_{1/2} = \frac{1}{2T}$. Also from here we can see that the half-width in terms of the period as $\Delta\phi_{1/2} \approx \frac{\phi^2}{2T}$.

An important feature of PDM techniques is that finds all periodic components, subharmonics included -and also periodicity in the observations, what might produce false minimums-. Of course, in the subharmonics the mean light curve slope will be greater, according to Stellingwerf (1978) $< g' > = 2nA$ for the n-th harmonic. So the half-width in the frequency for subharmonics will be $(\Delta f_n)_{1/2} = \frac{1}{2nT}$, this means we can distinguish subharmonics for their narrow line widths. Detecting subharmonics is fundamental at the time of assuring the reliability in the period found by PDM; if this period induces subharmonics it will be genuinely a periodic component and not just the product of 'noise' in the data.

Additionally, the statistical significance of the line will be given by the value of Θ at $\Delta f = 0$; if we substitute at equation (7) we'll have

$$\Theta = \frac{1 + 4\epsilon^2/N_b^2}{1 + \epsilon^2} \quad (9)$$

As the noise components in our data is large ($\epsilon < 1.25$), according to equation (9) the minimum will be above 0.49 (for $N_b = 5$).

V. RESULTS

We have run various simulations with different (N_b, N_c) distributions, changing the period in 0.001 each time, and looking in a range 0.5-15 days, as we can see in Figure 4.

For $N_b > 10$ the periodogram begins to be messy, with a lot of noise due to the lack of enough data. This is because the number of data points is $N = 160$, then for $N_b \geq 10$ the average amount of points in each bin will be $n_b = 16$, and therefore the number of freedom degrees is barely enough to maintain statistically consistency. Remark that despite this fact, statistics Θ minims are still found at the right orbital period, along with its subharmonics.
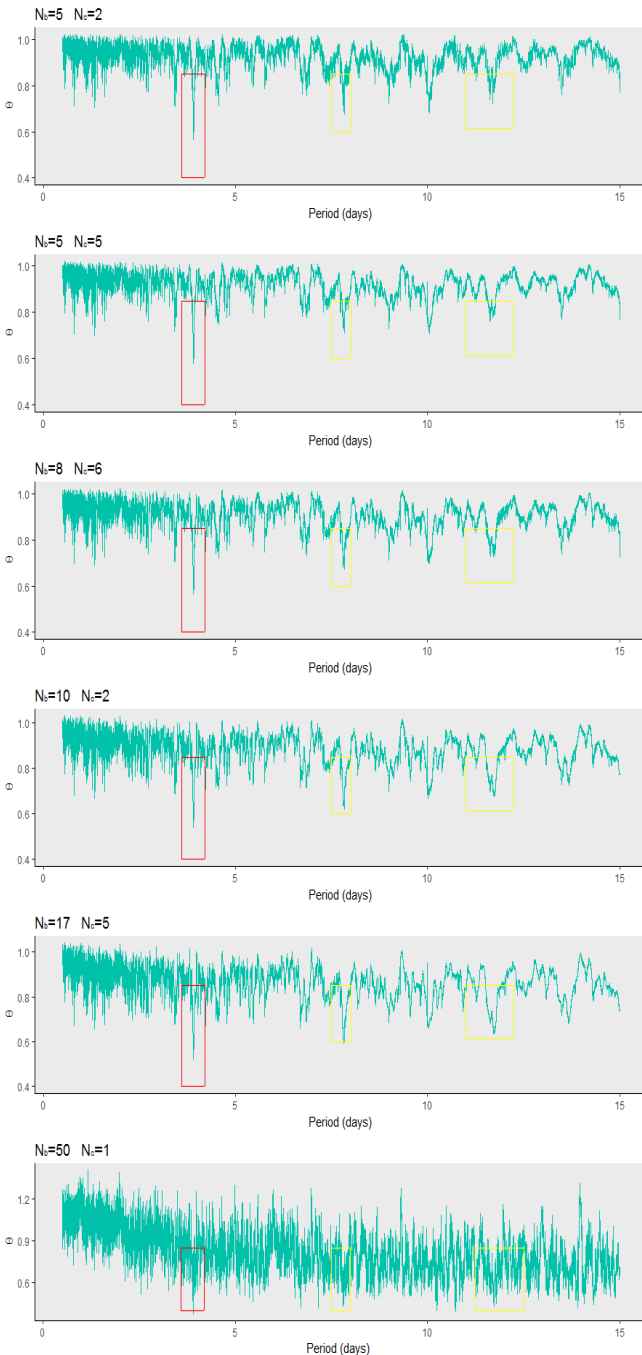
FIG. 4: *Values of statistic Θ for different bin structures (N_b, N_c). We search for the minims in a period range 0.5-15 days. In red we indicate the minimun Θ which belongs to the best fit for the period $\mathcal{P}$, and in yellow we have the first and second subharmonics, at approximately $2\mathcal{P}$ and $3\mathcal{P}$.*

In Table I we show in detail the results of our tests for $5 \leq N_b \leq 10$. We can observe that all values, except $(5,5)$, belong to the interval 3.9063 ± 0.0017 days of the estimated orbital period, being spot on in the cases $(8,6)$ and $(10,2)$ with 3.906 days in both of them. Furthermore, of all these results, $(10,2)$ happens to have the

N_b	N_c	Period Stimation	Θ
5	2	3.908	0.5653
5	4	3.908	0.5837
5	5	3.912	0.5786
6	3	3.908	0.5636
6	4	3.908	0.5666
8	4	3.908	0.5372
8	6	3.906	0.5395
10	2	3.906	0.5134
10	5	3.907	0.5209

TABLE I: *Results of our PDM tests combining different (N_b, N_c) configurations, and being $5 \leq N_b \leq 10$. More statistically relevant value (minimum Θ) found at $(10, 2)$ with a period 3.906 days; which is the exact value of the orbital period in LS 5039.*
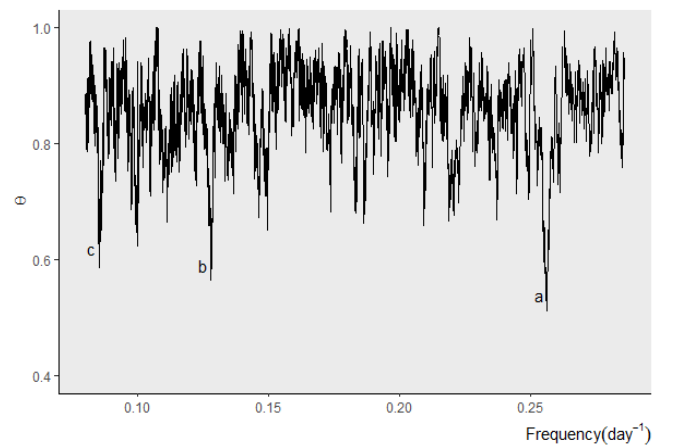

FIG. 5: *a) Frequency of the minimum at 0.256 days^{-1}, corresponding to $\mathcal{P}^{-1}$. b) First harmonic is found at 0.128 days^{-1}, whose inverse is 7.8125 days $\approx 2\mathcal{P}$. c) Second harmonic belongs to 0.0853 days^{-1} or 11.716 days $\approx 3\mathcal{P}$ days.*

smallest Θ value, what makes it the most statistically reliable value. We'll use this period $\mathcal{P} = 3.906$ days from $(10, 2)$ as our best period fit.

For the case $(10, 2)$ we also represent the Θ vs. frequency graphic in Figure 5, where we can observe clearly the minimum in Θ at the frequency 0.256 days^{-1} -inverse of period $\mathcal{P}$-. We also highlight the first two subharmonics, which have the second and third minimum value of Θ. Plus, in Figure 6 we see in detail the behavior of Θ around the minims in order to confirm the parabolic behavior predicted in Section IV.

Finally, in Fig. 7 we plot the representation in phase of the data points set for the period $\mathcal{P} = 3.906$ days inferred by PDM for the case $(10, 2)$. In the graphic we can detect, despite the big error bars, that the points follows an almost sinusoidal distribution, as we supposed before.

VI. CONCLUSION

The goal of this paper was to confirm the orbital modulation in the VHE γ-ray flux of LS 5039 first detected at Aharonian et al. (2006). We've been able to replicate their results through our own Lomb-Scargle periodogram, and we've also achieved these same results using a different method -PDM by Stellingwerf-. All this makes us affirm that LS 5039 is the first binary system with periodic emission of VHE γ-rays flux associated to its orbital period. Also, we've seen that PDM method can be a powerful tool to study emission periodicity in other systems suspicious to have the same properties as LS 5039.
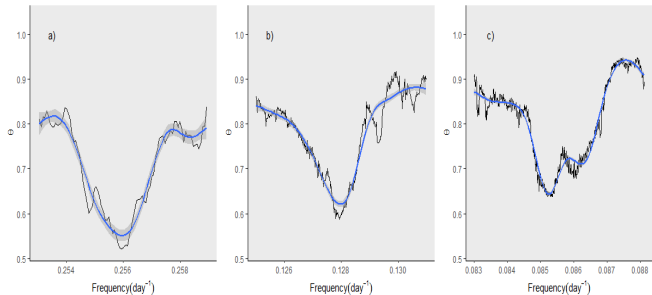
FIG. 6: *The representation of Θ behaviour around the minimums show the parabolic behavior predicted in Section IV. a) Representation of $\mathcal{P}^{-1}$ minimum, which have a width of 0.004 $days^{-1}$, the widest of them all. b) Θ minimum around the first subharmonic at $(2\mathcal{P})^{-1}$. c) Second subharmonic minimun at $(3\mathcal{P})^{-1}$, it has a width similar to the first subharmonic due to a sub-minimum. For all these graphics a quadratic regression based on a t-test with a 95% confidence band is performed. This test is performed by ggplot package of RStudio [6].*
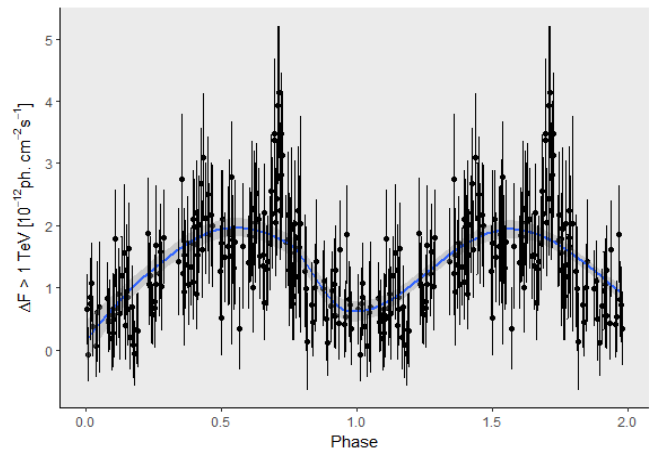
FIG. 7: *Phase distribution of data points for orbital period 3.906; we extended it for better appreciation of the periodicity in the γ ray's flux. A sinusoidal regression based on a t-test with a 95% confidence band is performed. This test is performed by ggplot package of RStudio. We see how around the superior conjunction ($\phi = 0.058$) we detect a minimum in the flux, and between the apastron and the inferior conjunction ($\phi = 0.716$) the flux reaches its maximum, as we explained in the introduction.*

[1] New Optical Results on γ-ray Binaries. J. Casares, J.M. Corral-Santana, A. Herrero et al. 2010 . High-Energy Emission from Pulsars and their Systems, Astrophysics and Space Science Proceedings, ISBN 978-3-642-17250-2. Springer-Verlag Berlin Heidelberg, 2011, p. 599-562

[2] Gamma-ray binaries and related systems. Guillaume Dubus 2013. Astronomy and Astrophysics Review, 21, 64.

[3] 3.9 day orbital modulation in the TeV γ-ray flux and spectrum from the X-ray binary LS 5039. F. Aharonian, A.G. Akhperjanian, A. R. Bazer-Bachi, et al. 2006, A&A 460 743-749.

[4] lomb: Lomb-Scargle Periodogram. R-Package implemented by Thomas Ruf, released in 2013. Repository CRAN https://CRAN.R-project.org/package=lomb.

[5] Period Determination Using Phase Dispertion Minimization. R.F. Stellingwerf, The Astrophysical Journal, 224:953-960, 1978 September 15.

[6] R-Package ggplot2. Hadley Wickham, Winston Chang v. 2.2.1 released in December 30, 2016. Repository CRAN https://cran.r-project.org/web/packages/ggplot2/ggplot2.pdf