

**PONTIFICIA UNIVERSIDAD CATÓLICA DE VALPARAISO**  
**FACULTAD DE CIENCIAS**  
**INSTITUTO DE ESTADÍSTICA**



**”ESTUDIO DE PERIODOGRAMAS  
Y EVALUACIÓN DE SU  
DESEMPEÑO EN CURVAS DE LUZ.”**

Memoria para optar al grado Magister en Estadística  
**GUNTHER ÁVILA**

PROFESOR GUÍA  
**ALEJANDRA CHRISTEN**

VALPARAÍSO 2019

La detección de exoplanetas es una de las líneas de investigación más importantes de la Astronomía actual y el método de tránsito es el que más exoplanetas ha detectado. Esta tesis se enfoca en estudiar las curvas de luz obtenidas por el observatorio satelital Kepler, estudiando la efectividad de cuatro diferentes métodos llamados periodogramas para encontrar candidatos a exoplanetas.



# Índice

<b>I</b>	<b>Introducción</b>	<b>7</b>
<b>1.</b>	<b>Introducción</b>	<b>7</b>
1.1.	Introducción al tema . . . . .	7
1.2.	Motivación: . . . . .	8
1.3.	Objetivos . . . . .	8
1.4.	Método del tránsito . . . . .	9
1.4.1.	Fundamentos del tránsito . . . . .	10
1.4.2.	Probabilidad de tránsito . . . . .	11
1.5.	Estado del arte: . . . . .	11
1.5.1.	Otros métodos de detección de exoplanetas: . . . . .	11
1.6.	organizacion de la tesis . . . . .	14
<b>II</b>	<b>Teoria de los periodogramas</b>	<b>15</b>
<b>2.</b>	<b>Introducción a Fourier y Periodogramas basados en DFT.</b>	<b>15</b>
2.1.	Familia ortogonal . . . . .	15
2.2.	Trasformación discreta de Fourier (DFT) y periodograma clásico. . . . .	17
2.3.	Period04 . . . . .	20
2.4.	Periodograma de Lomb-Scargle . . . . .	21
<b>3.</b>	<b>Periodogramas</b>	<b>22</b>
3.1.	Box-fitting Least Squares . . . . .	22
<b>4.</b>	<b>Consideraciones finales:</b>	<b>25</b>
4.0.1.	Phase: . . . . .	25
4.0.2.	Fenómeno de aliasing: . . . . .	26
4.0.3.	Efecto Eco o resonancia: . . . . .	26
<b>III</b>	<b>Analisis de desempeño de los periodogramas en curvas de luz.</b>	<b>28</b>
<b>5.</b>	<b>Introducción:</b>	<b>28</b>
5.1.	Obtención e onformación de las curvas de luz seleccionadas. . . . .	28
5.1.1.	Obtencion y eleccion: . . . . .	28
5.1.2.	Información de la curva de luz: . . . . .	30
5.2.	Aplicacion de los periodogramas . . . . .	31
5.2.1.	Clásico: . . . . .	31
5.2.2.	Period04: . . . . .	32
5.2.3.	Lomb-Scargle: . . . . .	32
5.2.4.	BLS: . . . . .	32

5.3. Metodología de obtención de desempeño. . . . .	33
<b>6. Grupo I</b>	<b>35</b>
6.0.1. Periodograma clásico . . . . .	35
6.0.2. Period04 . . . . .	36
6.0.3. Lomb-Scargle . . . . .	37
6.0.4. BLS . . . . .	38
6.0.5. Conclusiones: . . . . .	39
<b>7. Grupo II</b>	<b>40</b>
7.0.1. Periodograma clásico . . . . .	41
7.0.2. Period04 . . . . .	41
7.0.3. Lomb-Scargle . . . . .	42
7.0.4. BLS . . . . .	42
7.0.5. Conclusiones: . . . . .	44
<b>8. Grupo III</b>	<b>45</b>
8.0.1. Periodograma Clásico . . . . .	46
8.0.2. Period04 . . . . .	46
8.0.3. Lomb-Scargle . . . . .	47
8.0.4. BLS . . . . .	47
8.0.5. Conclusiones: . . . . .	49
<b>9. Grupo IV</b>	<b>50</b>
9.0.1. Periodograma clásico . . . . .	51
9.0.2. Period04 . . . . .	51
9.0.3. Lomb-Scargle . . . . .	52
9.0.4. BLS . . . . .	52
9.0.5. Conclusión: . . . . .	54
<b>10. Grupo V</b>	<b>55</b>
10.0.1. Period04 . . . . .	56
10.0.2. Lomb-Scargle . . . . .	57
10.0.3. BLS . . . . .	58
10.0.4. Conclusión: . . . . .	59
<b>11. Conclusiones generales:</b>	<b>61</b>
<b>IV Conclusiones finales:</b>	<b>63</b>
11.0.1. Trabajo futuro: . . . . .	63
<b>V Bibliografía</b>	<b>64</b>

<b>VI Anexo A: Gráficos</b>	<b>64</b>
<b>VII Anexo B: Codigos R:</b>	<b>67</b>

# Parte I

# Introducción

## 1. Introducción

Es de necesidad para los astrónomos conocer periodicidades de distintos fenómenos estelares para su análisis. Uno de ellos es la búsqueda de exoplanetas, existiendo diferentes métodos astronómicos para su descubrimiento. Uno de ellos y el más fructífero hasta la fecha, es el tránsito planetario.

El tránsito planetario consiste en ver perturbaciones de la luminosidad de la estrella, que ocurren al pasar un exoplaneta, desde nuestra perspectiva, frente a ella. Estas perturbaciones son muy características y diferentes a las que puede tener la propia estrella, como lo son estrellas del tipo variable, pulsantes, estrellas eclipsantes, etc. . . .

Sus características son: marcadamente periódicas, ocurren en un intervalo de tiempo pequeño (menor del 10 %) de todo el período. Para el análisis y búsqueda de estos tránsitos están las curvas de luz. Las curvas de luz son series de tiempo de la luminosidad de la estrella, datos que le llegan a los astrónomos desde los observatorios. y para descifrar por ejemplo si tal astro es candidato a tener exoplanetas o una estrella con algunas propiedades, se busca en ese flujo de luz periodicidades. . . para lo cual se tienen como herramienta el periodograma.

### 1.1. Introducción al tema

#### **Que es un peridograma?**

Es una representación gráfica, generalmente dada por una función, que explique y ponga en evidencia todos los períodos que pueda tener una serie de tiempo. Esta función por lo general es de la forma  $F(x) : V \subseteq \mathbb{R}^+ \rightarrow D \subseteq \mathbb{R}^+$ , donde  $V$  son los períodos, y  $D$  es la energía o poder. Esta energía o poder refleja la evidencia de existencia para el período dado. Por ejemplo, en los periodogramas basados en la descomposición discreta de Fourier DFT, esta potencia está dada por qué tan alto es el módulo de los elementos de la base que tienen un período dado. Este término fue acuñado por Schuster en 1989 para referirse a una estimación de la densidad espectral de una señal o serie temporal. Pero ahora es un concepto más amplio.

#### **¿Qué es una curva de luz?**

Una curva de luz es una serie de tiempo, del flujo de luz de una estrella. Es decir, luminosidad vs el tiempo, estas por lo general no son equiespaciadas (los intervalos de tiempo del muestreo no son constantes) y suelen tener intervalos de datos perdidos. Son series de tiempo estacionarias, por lo general cíclicas y como también tener gran cantidad de ruido. Son en este tipo de series de tiempo en las que basaremos nuestro trabajo.

## 1.2. Motivación:

Siempre me a gustado la astronomía, y en esta vez se me presenta la oportunidad de realizar la tesis con respecto a este tema, mi interés es que esta tesis sirva de introducción a personas que tengan el interés de presentando en forma lo mas didáctica posible los periodogramas, en la medida de lo posible como se usan y como se mueran los exoplanetas.

## 1.3. Objetivos

Los objetivos de este proyecto de tesis será el estudio de 4 periodogramas:

- Periodograma clásico (Schuster). Primer periodograma, se basa en reescribir la serie de tiempo en una base de funciones sinusoidales, obteniendo así que elementos de esta base son las preponderantes y con ello su período. Este periodograma tiene problemas al ser usado en curvas de luz, pues usa datos equidistantes y no esta adaptado para datos faltantes en la curva.
- El entregado por el software Period04. El cual usa el periodograma clásico, pero ajusta la data (en nuestro caso curva de luz) para solucionar los problemas antes descrito de no equiespaciedad y datos faltantes.
- Lomb-Scarle. Una sofisticación Resultado del periodograma clásico. Esta vez ajusta el algoritmo del periodograma clásico para poder ser usados en series de tiempo no equidistantes y con datos faltantes.
- Box Least Squares. Un algoritmo específicamente diseñado para la búsqueda de exoplanetas. Busca para cada período

Y se probara su rendimiento en 12 curvas de luz, 11 de ellas con exoplanetas confirmados, y una de ella no clasificada. El rendimiento será determinado por qué tan evidente se muestra en el periodograma, el período del tránsito confirmado. Si es el punto más alto del periodograma, si se podría con tal periodograma descubrir este tránsito o si se logra detectar.

## 1.4. Método del tránsito

Este método se basa en la observación de pequeños cambios de brillo en la estrella, generados por un posible exoplaneta que transita frente a esta, vistos desde nuestra perspectiva. al cruzar un planeta frente a la estrella (lo que se conoce como el tránsito), desciende el brillo que llega de esta, es recurrente en forma periódica (correspondería a su órbita) y tiene una firma particular. En la siguiente fig.1 se observa:

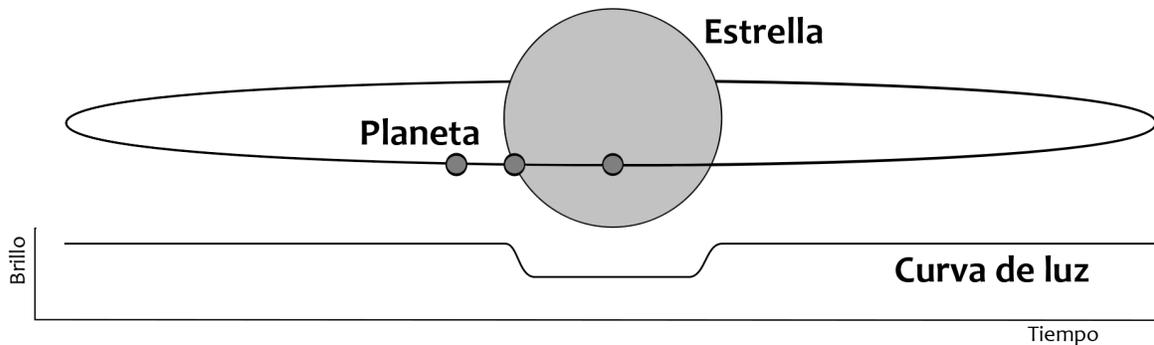


Figura 1: Esquema del tránsito de un planeta, y la luminosidad estrellas.

El método fue propuesto por primera vez en en 1952 por Otto Struve . Ya en el año 2000 el primer exoplaneta que es detectado por este método fue el HD 209458b poco tiempo después de que fuera descubierto. Y el primero exoplaneta descubierto por este método fue OGLE-TR-56 b en 2002.

Este método a sido uno de los mas exitosos en la búsqueda de exoplanetas, misiones como Kepler han multiplicado el descubrimiento de exoplanetas por esta forma, y en la actualidad es el método que más cantidad de exoplanetas encontrados.

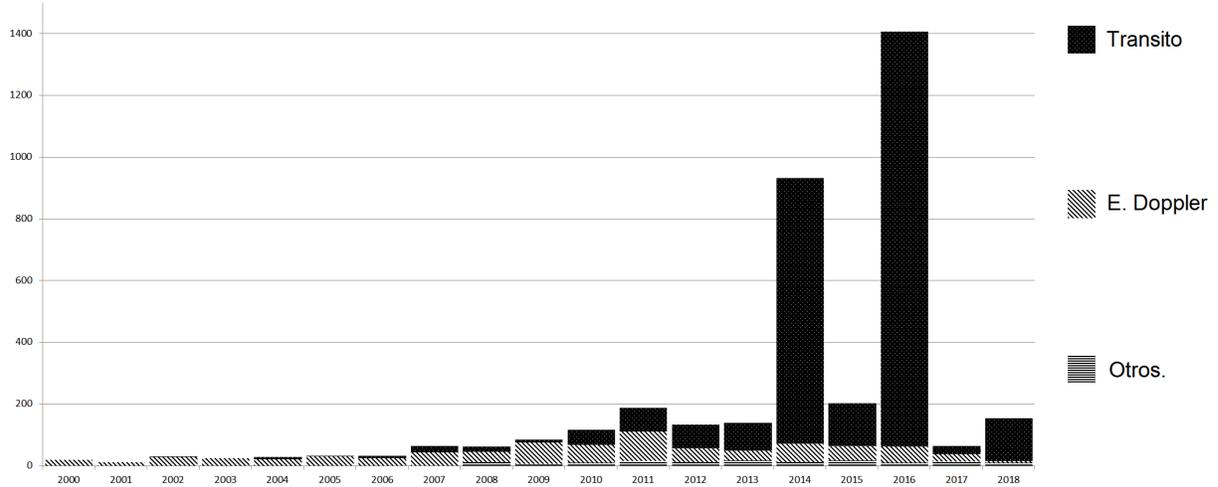


Figura 2: Exoplanetas descubiertos confirmados, diferentes métodos <http://www.openexoplanetcatalogue.com/>

### 1.4.1. Fundamentos del tránsito

Cuando el planeta pasa frente a la estrella, ocurre una disminución de su luminosidad  $F$ , este cambio de luminosidad está dado por  $\Delta F$  el cual es:

$$\Delta F \approx \left( \frac{R_p}{R_e} \right)^2 = k^2$$

Siendo  $R_p$  el radio del planeta,  $R_e$  el radio de la estrella. El cambio de luminosidad cuando el disco del planeta está totalmente dentro del disco de la estrella, es el valor  $k^2$ .  $\Delta F$  Se conoce como la profundidad del tránsito.

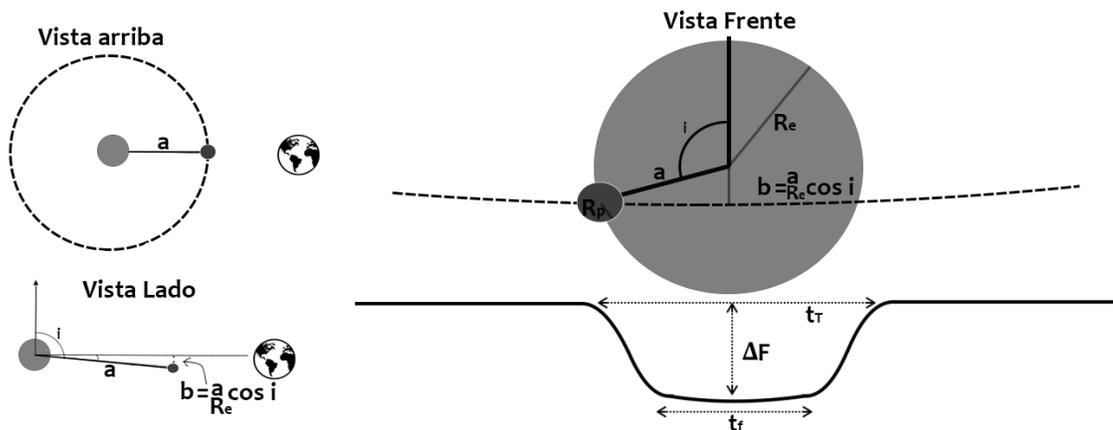


Figura 3: Esquema del tránsito de un planeta, y la luminosidad estrellas.

Luego  $t_T$  es el tiempo total en que el planeta esta en tránsito lo que se llama amplitud,  $t_g$  el tiempo en que todo el disco del planeta esta dentro de la estrella,  $a$  es el semi-eje mayor de la órbita del planeta,  $i$  es el angulo de inclinación de la órbita con respecto al observador. Y  $b$  que es la distancia mínima entre el borde del exoplaneta y el centro del disco estelar, la cual ocurre en la mitad del tránsito. La distancia  $b$  tiene la siguiente expresión: :

$$b \equiv \frac{a}{R_e} \cos i$$

### 1.4.2. Probabilidad de tránsito

Para que el exoplaneta sea observado desde la tierra (con un tránsito total o parcial), el valor  $b$  (desde nuestra perspectiva) debe ser menor a  $R_p + R_e$ . Así depende de que tan cerca este de la estrella, el radio de esta, el radio del propio planeta y obviamente la inclinación desde nuestra perspectiva (dada por el angulo  $i$ ).

Como no se conoce la cantidad de sistemas con exoplanetas (relación de sistemas con vs sin exoplanetas), que tipo de planetas son los más o menos probables (distancia y radio). No es posible calcular la probabilidad de encontrar exoplanetas por medio de tránsito, pero si es posible saber la probabilidad de que si el sistema tiene un exoplaneta de las características dadas (radio de ambos, distancia a la estrella), este pueda ser detectado por medio del tránsito. Esto está dado por la siguiente relación expuesta en el libro *Methods of Detecting Exoplanets* (2016).

$$P(\cos i < \frac{R_p + R_e}{a}) = \frac{R_p + R_e}{a}$$

Por ejemplo, para un planeta como la tierra en un sistema como el nuestro, la probabilidad de que sea detectado vía tránsito es :

$$P(\cos i < \frac{R_p + R_e}{a}) = \frac{R_p + R_e}{a} = \frac{6,371 + 695,508}{150,000,000} \approx 0,00467$$

Es decir un 0,46

## 1.5. Estado del arte:

### 1.5.1. Otros métodos de detección de exoplanetas:

**Espectroscopia Doppler (radial velocity)** Este método consiste básicamente en ver la variabilidad del espectro de luz generada por el efecto doppler. Un posible exoplaneta genera una perturbación a la estrella que órbita (oscilación de esta), pues tanto la estrella como el planeta giran alrededor del centro de masas común (suponiendo 1 solo exoplaneta), y este no se encuentra específicamente en el centro de la estrella. En la imagen fig.4 se ve que tanto el planeta como la estrella giran alrededor del centro de masas.

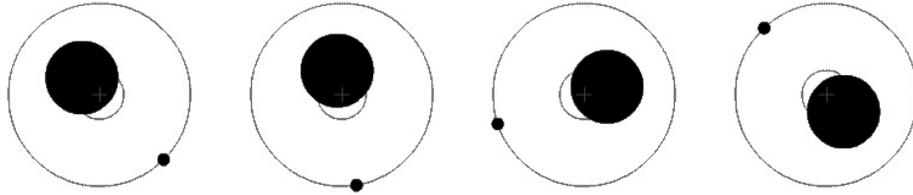


Figura 4: Ejemplo de como el efecto gravitatorio del planeta perturba la posición relativa de la estrella.

Si el sistema planetario tiene un ángulo respecto a nuestra observación, es decir no es totalmente perpendicular, al alejarse y acercarse la estrella con respecto al observador, generara una variación hacia el rojo (cuando se aleja), hacia el azul (cuando se acerca) de su espectro de luz, pudiéndose deducir la existencia y algunas características del exoplaneta.

Es el segundo método más efectivo en la búsqueda de exoplanetas, sus dificultades se basan en que debe tener un ángulo de canto con respecto al observador para obtener valores certeros de su tamaño, no es efectiva con estrellas que sean variables, además de la intrínseca dificultad de medir este efecto cuando la masa del planeta es pequeña con respecto a la estrella.

**Lente gravitacional** Consiste en usar el fenómeno de lente gravitacional para detectar un exoplaneta, lente gravitacional ocurre cuando desde la perspectiva nuestra se alinean 2 astros una estrella muy luminosa de fondo (u otro astro que emita mucha luz), y otro astro masivo (el que genera el lente) pudiendo ser en nuestro caso otra estrella. Esta última en un determinado momento de la alineación, curvara la luz del astro del fondo, aumentando el total de luz residida. Si el sistema que genera el lente tuviese un planeta, y este se cruza generara perturbaciones en ese aumento de luz. En la imagen se observa cómo funciona este método, la estrella del frente deforma la luz del fondo, creando un foco en el observador y así aumentando su luminosidad. Al cruzar un exoplaneta por donde deforma la luz la estrella, la luminosidad cae. Como se ve en la fig.5.

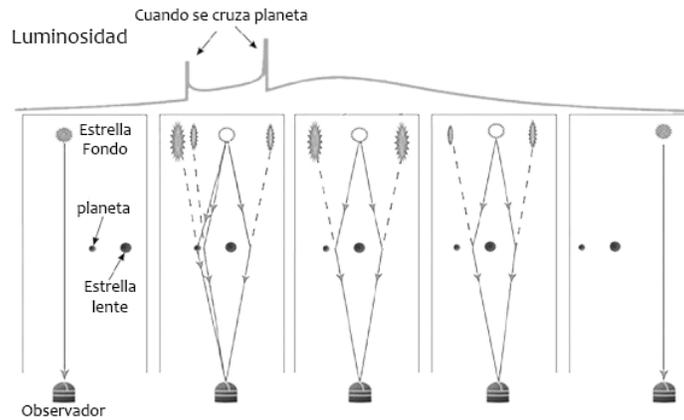


Figura 5: Esquema de cómo funciona el lente gravitacional, al cruzar un exoplaneta.

en este caso las dificultades estarían dadas por la baja probabilidad de que ocurra este fenómeno y sea documentado, la distancia y posición correcta del planeta.

**Imágenes directas** Se basa en obtener imágenes directas de la estrella, usando un Coronógrafo se eclipsa la luz de la estrella madre obteniendo así la luz reflejada de sus exoplanetas. a pesar de que muy pocos exoplanetas se han descubierto por este método, este entrega mas información al tener datos de la espectroscopia del planeta. En la Figura se muestra una imagen obtenida de este sistema.

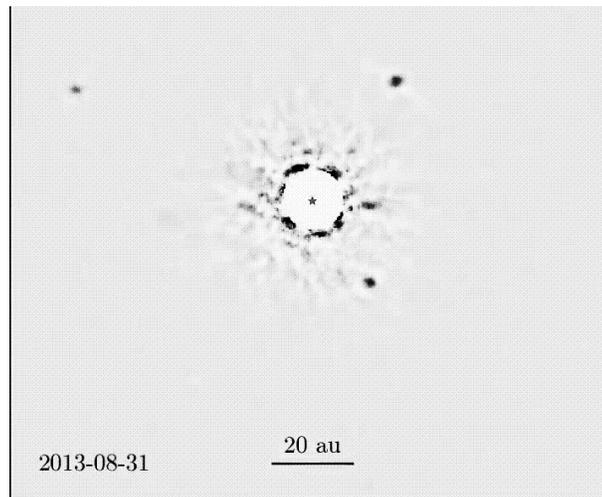


Figura 6: Imagen estrella HR 8799. W. M. Keck Observatory

## **1.6. organizacion de la tesis**

## Parte II

# Teoria de los periodogramas

## 2. Introducción a Fourier y Periodogramas basados en DFT.

### 2.1. Familia ortogonal

Dado un espacio vectorial  $V$  con producto interno  $\langle, \rangle$ , y dado dos elementos  $x, y$  de este espacio, definimos que 2 elementos son ortogonales cuando el producto interno de  $x$  e  $y$ , es igual a 0.

$$\forall x, y \in V \quad \langle x, y \rangle = 0 \quad \text{ssi } x \perp y$$

En nuestro caso usaremos el espacio vectorial de las sucesiones  $l_2$ , con el producto interno dado por  $\sum_{t \in \mathbb{N}} \phi_j(t) \phi_k(t)$ .

En este caso, y tomando un  $n \in \mathbb{N}$ , definimos la siguiente familia de funciones.

$$\left\{ \sin\left(\frac{2\pi tk}{n}\right), \cos\left(\frac{2\pi tk}{n}\right) : k = 0, 1, 2, \dots, \left\lfloor \frac{n}{2} \right\rfloor \right\}; t = 0, 1, \dots \quad (1)$$

$$\text{Ejem : } n = 5; k = \{0, 1, 2\}; \left\{ \sin\left(\frac{2\pi t0}{5}\right) = 0, \cos\left(\frac{2\pi t0}{5}\right) = 1, \sin\left(\frac{2\pi t1}{5}\right), \cos\left(\frac{2\pi t1}{5}\right), \sin\left(\frac{2\pi t2}{5}\right), \cos\left(\frac{2\pi t2}{5}\right) \right\}$$

Verificaremos que esta familia para un  $n$  arbitrario, es una familia ortogonal, es decir, que toda pareja de la familia es ortogonal entre si.

Sea la relación de euler:

$$e^{iw} = \cos(w) + i \sin(w) \quad (2)$$

$$\sin(w) = \frac{e^{iw} - e^{-iw}}{2i} \quad ; \quad \cos(w) = \frac{e^{iw} + e^{-iw}}{2}$$

$$\sum_{t=1}^n e^{iwt} = e^{iw} \left( \frac{e^{iwn} - 1}{e^{iw} - 1} \right) = e^{iw} \left[ \frac{e^{iwn/2} (e^{iwn/2} - e^{-iwn/2}) / 2i}{e^{iw/2} (e^{iw/2} - e^{-iw/2}) / 2i} \right] 1$$

$$\sum_{t=1}^n e^{iwt} = e^{iw} \left( \frac{e^{iwn} - 1}{e^{iw} - 1} \right) = e^{iw} \left[ \frac{e^{iwn/2} (e^{iwn/2} - e^{-iwn/2}) / 2i}{e^{iw/2} (e^{iw/2} - e^{-iw/2}) / 2i} \right]$$

---

<sup>1</sup>Es valido cuando  $w \neq 2\pi m$  para  $m$  entero

se factoriza por y se multiplica arriba y abajo por quedando el ultimo menos el primer termino.

$$= e^{iw(n+1)/2} \frac{\sin(wn/2)}{\sin(w/2)} = \cos\left(\frac{w(n+1)}{2}\right) \frac{\sin(wn/2)}{\sin(w/2)} + i \sin\left(\frac{w(n+1)}{2}\right) \frac{\sin(wn/2)}{\sin(w/2)}$$

pero tenemos la identidad del inicio (2) quedando finalmente.

$$\sum_{t=1}^n \cos(wt) = \cos\left(\frac{w(n+1)}{2}\right) \frac{\sin(wn/2)}{\sin(w/2)} \quad (3)$$

$$\sum_{t=1}^n \sin(wt) = \sin\left(\frac{w(n+1)}{2}\right) \frac{\sin(wn/2)}{\sin(w/2)} \quad (4)$$

para  $w = 2\pi k/n$  se tiene que

$$\frac{\sin(wn/2)}{\sin(w/2)} = \frac{\sin(\pi k)}{\sin(\pi k/n)} = \begin{cases} n & \text{si } k = 0 \\ 0 & \text{si } k \neq 0 \end{cases}$$

$\cos(w)$  es 1, y  $\sin(w)$  es 0 para  $w = 0$ , usando todo lo anterior en (3) y (4) queda:

$$\sum_{t=1}^n \cos(wt) = \begin{cases} n & \text{si } k = 0 \\ 0 & \text{si } k \text{ entero} \end{cases} \quad (5)$$

$$\sum_{t=1}^n \sin(wt) = 0, k = 0, 1, 2, \dots, \left[\frac{n}{2}\right] \quad (6)$$

usando prostaferesis y las ecuaciones (5),(6) el producto interno de cualquier par de funciones de 1 queda de la forma:

$$\sum_{t=1}^n \cos\left(\frac{2\pi tk}{n}\right) \cos\left(\frac{2\pi tj}{n}\right) = \begin{cases} n, & k = j = 0 \text{ o } n/2 \text{ (si } n \text{ par)} \\ n/2 & k = j \neq 0 \text{ o } n/2 \text{ (si } n \text{ par)} \\ 0 & k \neq j \end{cases} \quad (7)$$

$$\sum_{t=1}^n \sin\left(\frac{2\pi tk}{n}\right) \sin\left(\frac{2\pi tj}{n}\right) = \begin{cases} 0, & k = j = 0 \text{ o } n/2 \text{ (si } n \text{ par)} \\ n/2 & k = j \neq 0 \text{ o } n/2 \text{ (si } n \text{ par)} \\ 0 & k \neq j \end{cases} \quad (8)$$

$$\sum_{t=1}^n \sin\left(\frac{2\pi tk}{n}\right) \cos\left(\frac{2\pi tj}{n}\right) = 0 \text{ para cualquier } k \text{ y } j.$$

así, el producto interno de 2 funciones diferentes en (1) siempre dará 0, por lo que es una familia ortogonal.

## 2.2. Transformación discreta de Fourier (DFT) y periodograma clásico.

Teniendo esta familia ortogonal de funciones (1), y usando la teoría de espacios vectoriales, podemos usarla como base vectorial para escribir una serie finita de largo  $n$ ,  $Z = \{Z_1, Z_2, \dots, Z_n\}$  de la forma siguiente:

$$Z_t = \sum_{k=0}^{\lfloor n/2 \rfloor} \left[ a_k \cos\left(\frac{2\pi tk}{n}\right) + b_k \sin\left(\frac{2\pi tk}{n}\right) \right], t = 1, 2, \dots, n \quad (9)$$

Llamándose la serie de Fourier de la serie finita  $Z$ . Basta encontrar los valores de  $a_k$  y  $b_k$ ,  $k = 0 \dots \lfloor n/2 \rfloor$  para tener la representación de Fourier de esta serie, como sabemos que las funciones de senos y cosenos son ortogonales entre si, al escribir en forma matricial (9) tenemos una matriz de vectores ortogonal fácilmente invertible. dejando los términos  $a_k$  y  $b_k$  de la siguiente forma:

$$a_k = \begin{cases} \frac{1}{n} \sum_{t=1}^n Z_t \cos\left(\frac{2\pi tk}{n}\right) & \text{si } k = 0, \text{ y } k = n/2 \text{ si } n \text{ es par} \\ \frac{2}{n} \sum_{t=1}^n Z_t \cos\left(\frac{2\pi tk}{n}\right) & \text{si } k = 1, 2, \dots, \left[\frac{n-1}{2}\right] \end{cases} \quad (10)$$

$$b_k = \frac{2}{n} \sum_{t=1}^n Z_t \sin\left(\frac{2\pi tk}{n}\right) \text{ si } k = 1, 2, \dots, \left[\frac{n-1}{2}\right]$$

Ejemplo:

$$Z = \{1, 2, 3\}$$

$$a_0 = \frac{1}{3} \left( \cos\left(\frac{2\pi 0}{3}\right) + 2\cos\left(\frac{2\pi 2 0}{3}\right) + 3\cos\left(\frac{2\pi 3 0}{3}\right) \right) = \frac{6}{3}$$

$$a_1 = \frac{2}{3} \left( \cos\left(\frac{2\pi}{3}\right) + 2\cos\left(\frac{2\pi 2}{3}\right) + 3\cos\left(\frac{2\pi 3}{3}\right) \right) = 1$$

$$b_0 = 0$$

$$b_1 = \frac{2}{3} \left( \sin\left(\frac{2\pi}{3}\right) + 2\sin\left(\frac{2\pi 2}{3}\right) + 3\sin\left(\frac{2\pi 3}{3}\right) \right) = -\frac{2\sqrt{3}}{6}$$

así dando la representación de Fourier del ejemplo:

$$Z_t = 2 + \cos\left(\frac{2\pi t}{3}\right) - \frac{2\sqrt{3}}{6} \sin\left(\frac{2\pi t}{3}\right)$$

entregando los valores de la serie para cada  $t = 1, 2, 3 \dots$  correspondiente.

Nuestra intención es encontrar la magnitud de estos  $a_k$  y  $b_k$ , es decir la energía o potencia que aporta cada elemento de la base con una misma periodicidad. En nuestro caso sería solo 2 para cada período, pues tanto el seno como el coseno poseen la misma periodicidad para un mismo  $k$ . Lo que sigue a continuación :

Tenemos que los coeficientes de la representación de Fourier coinciden con la estimación de los mínimos cuadrados del siguiente modelo de regresión:

$$Z_t = \sum_{k=0}^{[n/2]} [a_k \cos(\omega_k t) + b_k \sin(\omega_k t)] + e_t$$

siendo  $\omega_k = \frac{2\pi k}{n}$ , así podemos usar la relación de Parseval:

Si  $X$  es un espacio vectorial normado, con producto escalar, y  $\{e_n\}$  es un conjunto ortogonal en  $X$ , con  $e_n \neq 0$  para todo  $n$ , si  $x$  es un elemento de  $X$  tal que  $x = \sum_{k=1}^{\infty} c_n e_n$  (es decir es un elemento generado por este conjunto ortogonal). Se tiene que:

$$\|x\|^2 = \sum_{k=1}^{\infty} |c_n|^2 \|e_n\|^2$$

Dado nuestra DFT9, cumple con las premisas de esta relación, tenemos lo siguiente:

$$\|Z_t\|^2 = \sum_{k=1}^{[n/2]} |a_n|^2 \|\cos(\omega_k t)\|^2 + \sum_{k=1}^{[n/2]} |b_n|^2 \|\sin(\omega_k t)\|^2$$

por 7 y 8 :

$$\|Z_t\|^2 = \sum_{k=1}^{[n/2]} |a_n|^2 \frac{n}{2} + \sum_{k=1}^{[n/2]} |b_n|^2 \frac{n}{2}$$

Resultando:

$$\sum_{k=1}^n Z_t^2 = \begin{cases} na_0^2 + \frac{n}{2} \sum_{k=1}^{[(n-1)/2]} (a_k^2 + b_k^2) & \text{si } n \text{ es impar} \\ na_0^2 + \frac{n}{2} \sum_{k=1}^{[(n-1)/2]} (a_k^2 + b_k^2) + na_{n/2}^2 & \text{si } n \text{ es par} \end{cases}$$

finalmente obteniendo lo que necesitamos, la potencia de cada período entrega de fondo(dado por los senos y cosenos). Por lo que el periodograma (clásico), es el siguiente:

$$I(\omega_k) = \begin{cases} na_0^2 & k = 0 \\ n/2(a_k^2 + b_k^2) & k = 1 \dots \left[ \frac{n-1}{2} \right] \\ na_{n/2}^2 & k = n/2 \end{cases} \quad (11)$$

Fue introducido por Schuster en 1898 para buscar los componentes periódicos de una serie.

Este sera el primer periodograma que usaremos, este algoritmo se encuentra directamente programado en R con el comando *spectrum* (modo no logarítmico), que es una de las más comunes formas de estimar la densidad espectral (a través de la DFT), y en R esta dada por este comando.

El problema de este periodograma para las curvas de luz a tratar reviste 2 problemas:

-La serie de tiempo debe ser equidistante: en la práctica, esto no suele ser un problema, las curvas de luz entregadas por Kepler son muy consistentes en el lapso de tiempo entre cada observación ( 0,020434 día, aproximadamente cada 30min) , la diferencia entre lapsos contiguos tienen una varianza de aproximadamente  $1,512191e - 13$ , por lo que se puede tomar como si fuese una serie de intervalos equidistantes, sin afectar la estimación de períodos.

-La serie de tiempo no debe tener pérdida de datos: Este es un problema más grave, pues no existen modelos para estimar estas curvas de luz de manera directa, no siendo una alternativa para completar intervalos donde no hay datos. por lo que quedan métodos como completar con la media, completar con un ruido gaussiano con media y varianzas muestrales de la curva o montecarlo.

Finalmente algo no menos importante, el periodograma de schuster entrega la potencia vs frecuencia (por eso es llamado espectro), la frecuencia mide la cantidad de repeticiones por unidad de tiempo. Nuestro interés es el período el cual es el reciproco de la frecuencia, es decir, dado una repetición cuantas unidades de tiempo han transcurrido. Por lo que el resultado entregado por este método debe ser transformado por:

$$P = \frac{1}{f} \tag{12}$$

## 2.3. Period04

Period04 es un software sucesor del Period98 especialmente dedicado al análisis estadístico de grandes series de tiempo astronómicas que contienen missing. El programa ofrece herramientas para extraer las frecuencias individuales del contenido multi-periódico de las series temporales y proporciona una interfaz simple para realizar ajustes de frecuencia múltiple.

No está diseñado para la búsqueda de exoplanetas, sino para otros fenómenos astronómicos que puedan ser explicados con la búsqueda de su periodicidad, este software usa DFT de la misma forma que el periodograma anterior, pero es una evolución directa del clásico de Schuster, pues posee herramientas de ajuste de mínimos cuadrados para lograr la equidistancia de la frecuencia(tiempo) y método de montecarlo para completar los puntos donde se encuentre datos perdidos.

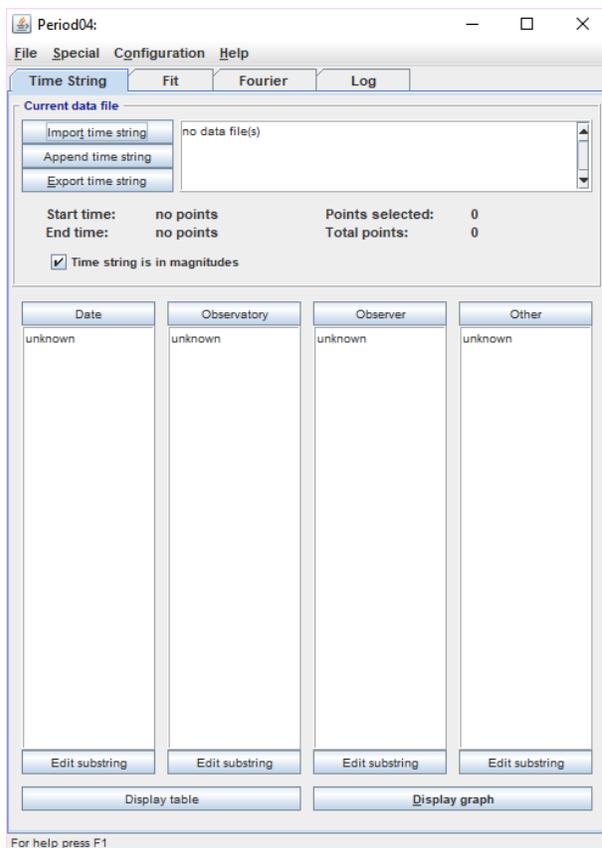


Figura 7: Interfas software Period04

Este periodograma por lo ya comentado, ofrece solución a los 2 problemas del periodograma clásico de Schuster directo, expuestos en la sección anterior. También posee herramientas para graficar tanto la serie como su espectro, el cual es entregado en términos de frecuencia, así pues para obtener el periodo se debe usar la relación 12.

## 2.4. Periodograma de Lomb-Scargle

Este periodograma fue creado como una respuesta a los problemas que posee el periodograma clásico, como los intervalos de data missing, problemas de aliasing (explicado aca), ruido gaussiano y uniformidad del muestreo (tiempo). Manejando todos estos en el mismo algoritmo del periodograma, creando una solución mucho mas robusta que la de simplemente intentar ajustar los datos para ocupar el periodograma de schuster.

Parte de la base del periodograma clásico, ajustandolo por medio de funciones auxiliares, fue desarrollado y expuesto en los Scargle 1982 y Lomb 1976 .

Este generaliza el periodograma clásico de la siguiente forma:

$$P(\omega_k) = \frac{A^2}{2} \left( \sum_m Z_{t_m} \cos(\omega_k [t_m - \tau]) \right)^2 + \frac{B^2}{2} \left( \sum_m Z_{t_m} \sin(\omega_k [t_m - \tau]) \right)^2$$

Siendo  $A, B$  y  $\tau$  funciones arbitrarias, en términos de  $\omega_k$  y  $t_m$ , pero no de  $Z_{t_m}$  que son los mismos calculados via peridograma clásico, es decir, estas 3 funciones no dependen de la serie de tiempo usada.

Vemos que volvemos al periodograma de schuster cuando asumimos  $A = B = \left(\frac{2}{n}\right)^2$ ,  $Z_{t_m} = Z_t$ ,  $\tau = 0$ .

finalmente usando:

$$A^2 = \left( \sum_m \cos(\omega_k [t_m - \tau]) \right)^{-1}$$

$$B^2 = \left( \sum_m \sin(\omega_k [t_m - \tau]) \right)^{-1}$$

obtenemos el periodograma de lomb-scargle como:

$$P(f) = \frac{1}{2} \frac{(\sum_m Z_{t_m} \cos(\omega_k [t_m - \tau]))^2}{(\sum_m \cos(\omega_k [t_m - \tau]))} + \frac{1}{2} \frac{(\sum_m Z_{t_m} \sin(\omega_k [t_m - \tau]))^2}{(\sum_m \sin(\omega_k [t_m - \tau]))}$$

con:

$$\tau = \frac{1}{2\omega_k} \tan^{-1} \left( \frac{(\sum_m \sin(2\omega_k t_m))}{(\sum_m \cos(2\omega_k t_m))} \right)$$

Estos cambios no son un capricho, estas funciones  $A, B$  reducen efectivamnete el problema de la equidistancia o uniformidad del muestreo de la serie. pudiendo usar saltos de tiempo completamente arbitrarios, esto también soluciona el problema de los datos perdidos pues estos pueden ser tomados como saltos en el muestreo. el uso de  $\tau$  ajusta para que el periodograma sea invariante en la posición inicial de toma de datos, propiedad que ya poseia el periodograma clásico.

También este periodograma hace un mejor trabajo del ruido gaussiano, todo esto esta explicitamente detallado en el archivo antes expuesto scargle 1982.

Este periodograma sera aplicado por medio de R, usando la libreria Lomb, contiene el algoritmo directamente.

### 3. Periodogramas

#### 3.1. Box-fitting Least Squares

Este método de búsqueda de períodos no se basa en búsqueda de funciones sinusoidales de fondo (como los basados en DFT), es un método mas directo. se basa en buscar o ajustar a través de mínimos cuadrados, una función escalonada de 2 valores con 1 solo intervalo diferente (o escala), a la serie de tiempo. tomando un valor cuando en la serie ocurre tránsito (cuando baja la intensidad de la luz) el momento del intervalo diferente y el resto de la serie. fue descrito por G. Kovács (2002).

El algoritmo es el siguiente:

Sea  $P$  el período, este tomara 2 valores discretos,  $H$  y  $L$ ,  $L$  cuando ocurre el tránsito, y  $H$  el resto.  $q$  sera el tiempo en el que ocurre el tránsito (en términos fraccionario, es decir,  $qP$  es tiempo de tránsito y  $P = qP + (1 - q)P$ ) este no suele ser mayor al 10 % total del período.

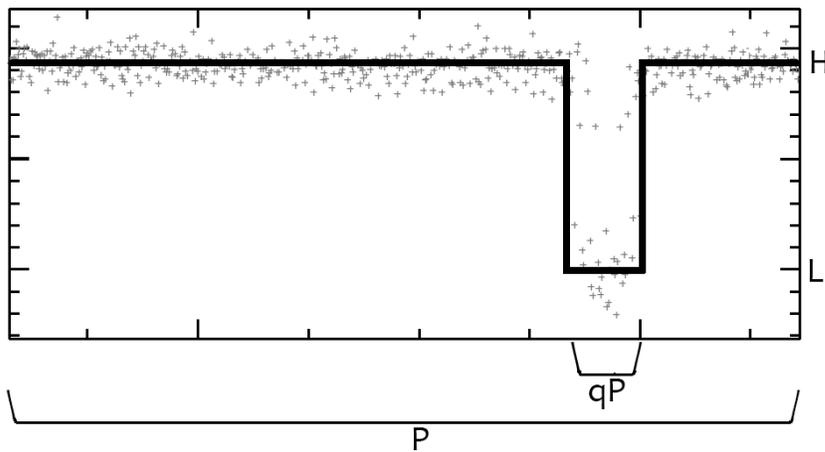


Figura 8: La forma de la función que se intenta ajustar por medio de mínimos cuadrados.

Sea la serie de tiempo, se centra por el promedio aritmético  $\mu$ , es decir, a cada valor de la serie  $x_i = x_i - \mu$ , quedando  $\sum_{t=1}^n x_t = 0$ .

por lo que queda lo siguiente en nuestra aproximación anterior:

$$(1 - q)PH + qPL = 0$$

$$H = -Lq/(1 - q)$$

Para cada  $x_i$  se asume un ruido gaussiano de media 0 y desviación estándar  $\sigma_i$ , no correlacionado, el ruido para cada punto de la serie es presentado por un peso  $\omega_i$  definido como  $\omega_i = \sigma_i^{-2} \left[ \sum_{j=1}^n \sigma_j^{-2} \right]^{-1}$ , notamos que  $\sum_{i=1}^n \omega_i = 1$  y que  $\sum_{t=1}^n \omega_i x_t = 0$ .

Finalmente reescribimos la serie de tiempo plegando la por el período  $P$  de prueba, es decir,  $\tilde{t} = \text{int}(t \text{ mod } P)$ . por lo que cada elemento de la serie quedara definido como  $\{\tilde{x}_i\}$  y con un peso correspondiente  $\{\tilde{\omega}_i\}$ .

Ajustamos la función escalonada a la serie de tiempo plegada, con los siguientes parámetros:

$$\widehat{L} \text{ el nivel en } [i_1, i_2]$$

$$\widehat{H} \text{ el nivel en } [1, i_1), (i_2, n]$$

Para cada  $(i_1, i_2)$ , debemos minimizar el error de ajustar la función escalonada, dada por la expresión :

$$D = \sum_{i=1}^{i_1-1} \tilde{\omega}_i (\tilde{x}_i - \widehat{H})^2 + \sum_{i=i_1}^{i_2} \tilde{\omega}_i (\tilde{x}_i - \widehat{L})^2 + \sum_{i=i_2+1}^n \tilde{\omega}_i (\tilde{x}_i - \widehat{H})^2$$

Pues queremos que el  $\widehat{H}$  y  $\widehat{L}$  por medio de mínimos cuadrados, se ajusten de mejor manera a la serie de tiempo, es decir minimicen la distancia a ellos. para obtener el mínimo derivamos parcialmente con respecto a  $\widehat{H}$  y  $\widehat{L}$

$$\frac{\partial D}{\partial \widehat{H}} = 2 \sum_{i=1}^{i_1-1} \tilde{\omega}_i (\widehat{H} - \tilde{x}_i) + 2 \sum_{i=i_2+1}^n \tilde{\omega}_i (\widehat{H} - \tilde{x}_i)$$

$$\frac{\partial D}{\partial \widehat{L}} = 2 \sum_{i=i_1}^{i_2} \tilde{\omega}_i (\widehat{L} - \tilde{x}_i)$$

notamos que la segunda derivada dan en ambos casos positiva  $\frac{\partial^2 D}{\partial \widehat{H}^2} = 4$ ,  $\frac{\partial^2 D}{\partial \widehat{L}^2} = 2$ , las cruzadas 0, así el punto donde la derivada son 0 es efectivamente un mínimo.

$$\frac{\partial D}{\partial \widehat{H}} = 0 \text{ Cuando } \widehat{H} \left( \sum_{i=1}^{i_1-1} \tilde{\omega}_i + \sum_{i=i_2+1}^n \tilde{\omega}_i \right) = \sum_{i=1}^{i_1-1} \tilde{\omega}_i \tilde{x}_i + \sum_{i=i_2+1}^n \tilde{\omega}_i \tilde{x}_i$$

$$\frac{\partial D}{\partial \widehat{L}} = 0 \text{ Cuando } \widehat{L} \left( \sum_{i=i_1}^{i_2} \tilde{\omega}_i \right) = \sum_{i=i_1}^{i_2} \tilde{\omega}_i \tilde{x}_i$$

nombramos  $r = \sum_{i=i_1}^{i_2} \tilde{\omega}_i$  y  $s = \sum_{i=i_1}^{i_2} \tilde{\omega}_i \tilde{x}_i$  quedando:

$$\widehat{L} = \frac{s}{r}$$

$$\widehat{H} = -\frac{s}{1-r}$$

así reemplazando el punto mínimo en  $D$ , obtenemos:

$$D = \sum_{i=1}^{i_1-1} \tilde{\omega}_i \left( \tilde{x}_i + \frac{s}{1-r} \right)^2 + \sum_{i=i_1}^{i_2} \tilde{\omega}_i \left( \tilde{x}_i - \frac{s}{r} \right)^2 + \sum_{i=i_2+1}^n \tilde{\omega}_i \left( \tilde{x}_i + \frac{s}{1-r} \right)^2$$

$$D = \sum_{i=1}^n \tilde{\omega}_i \tilde{x}_i^2 - \frac{s^2}{r(1-r)}$$

donde alcanza el mínimo, vemos que la parte izquierda no depende del período  $P$  ni del intervalo, por ello solo queda la parte derecha, que minimiza  $D$  es máxima ( $r$  es positiva menor que 1).

por lo que ahora definimos el periodograma (espectro de frecuencia) BLS, para cada período  $P$ , como:

$$SR_P = \text{MAX} \left\{ \left[ \frac{s_{(i_1, i_2)}^2}{r_{(i_1, i_2)}(1 - r_{(i_1, i_2)})} \right]^{\frac{1}{2}} \right\}$$

Para el  $P$  dado debemos recorrer todo el período con el intervalo  $(i_1, i_2)$  de puntos, luego cambiar el tamaño de este (completando todos los tamaños), de todos ellos obtener el máximo. Así para ese período  $P$  la función discreta que mejor se ajusta.

Para este método existen formas para reducir la carga computacional, uno de ellos, es buscar  $q$  en un intervalo mas reducido (por lo general se asume que el tránsito es inferior al 10 % del período), por ejemplo en KIC 11359879, el planeta tiene un corto período de 4,9 días y el tránsito es aproximadamente 1/29 del período total.

otra forma de reducir la carga, es dividir la serie de tiempo ya plegada en  $n$  bind, y calcular el SR en esos bind agrupados.

Finalmente se puede asumir que el ruido gaussiano es el mismo en todos los puntos (si desviacion estandar), quedando un  $SR_P$  aproximado de la forma:

$$\left[ \frac{\left( \sum_{i=i_1}^{i_2} \tilde{\omega}_i \tilde{x}_i \right)^2}{\left( \sum_{i=i_1}^{i_2} \tilde{\omega}_i \right) \left( 1 - \left( \sum_{i=i_1}^{i_2} \tilde{\omega}_i \right) \right)} \right]^{\frac{1}{2}}$$

Reemplazando por  $\tilde{\omega}_i = \sigma$  para todo  $i$ , y usando que  $\sum_{i=1}^n \omega_i = 1$  queda:

$$\left[ \frac{\left( \sigma \sum_{i=i_1}^{i_2} \tilde{x}_i \right)^2}{\left( \sigma \sum_{i=i_1}^{i_2} 1 \right) \left( \sigma \sum_{i=1}^n 1 - \sigma \left( \sum_{i=i_1}^{i_2} 1 \right) \right)} \right]^{\frac{1}{2}}$$

$$\left[ \frac{\sigma^2 \left( \sum_{i=i_1}^{i_2} \tilde{x}_i \right)^2}{\sigma^2 (m)(n - m)} \right]^{\frac{1}{2}}$$

Donde  $m$  es el numero total de puntos en el intervalo y  $n$  el numero total de puntos de la serie.

$$SR_P = \text{MAX} \left\{ \left[ \frac{\left( \sum_{i=i_1}^{i_2} x_i \right)^2}{m(n-m)} \right]^{\frac{1}{2}} \right\}$$

que sera el periodograma BLS.

**Ejemplo** Un ejemplo para este algoritmo seria el siguiente. sea la siguiente serie:

$$E_j = \{99, 101, 101, 70, 72, 97, 98, 102, 71, 70, 98, 97\}$$

para un período  $P = 5$ ,  $bind = 6$ ,  $|[i_1, i_2]| = 2$

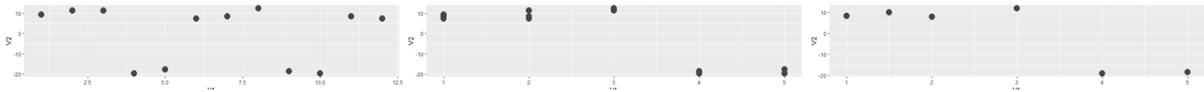


Figura 9: Paso 1, se normaliza la serie. Paso 2, se pliega a período  $P=5$ . Paso 3, se agrupa a 6 binds

Como  $|[i_1, i_2]| = 2$ ,  $\left[ \frac{\left( \sum_{i=i_1}^{i_2} x_i \right)^2}{n(N-n)} \right]^{\frac{1}{2}}$  tendrá los siguientes valores:

(1, 1,5)	(1,5, 2)	(2, 3)	(3, 4)	(4, 5)
7,4	7,2	8,0	3,0	15,4

Siendo el máximo en  $(4, 5) = 15,4$ , para calcular  $SR_5$  se debe hacer esto para  $|[i_1, i_2]| = \{1, 3, 4, 5\}$  y de todos obtener el máximo.

Finalmente este algoritmo esta incluido en el set de periodogramas de la pagina oficial de la NASA para la búsqueda de exoplanetas. Este sera el 4 tipo de periodogramas de nuestro trabajo, el algoritmo se desarrollara en R.

## 4. Consideraciones finales:

### 4.0.1. Phase:

Es el plegado de la serie de tiempo a un período  $P$ . En las imágenes se observa la curva de luz KIC 8692861 y Phase correspondiente al período  $P = 13,722$  (exoplaneta confirmado). Ejemplo en la figura 10.

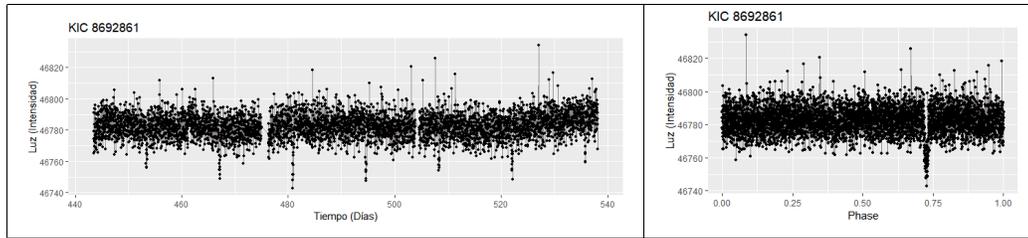


Figura 10: Ejemplo de plegar la curva de luz en el período confirmado de tránsito.

#### 4.0.2. Fenómeno de aliasing:

Esto ocurre en el periodograma clásico, cuando el período es mayor a la mitad del largo de la serie (Frecuencia de Nyquist ) o el período de la serie está en un valor intermedio del período de las bases de la descomposición.

En la siguiente imagen se observa el periodograma clásico de la serie  $\{t_i\} = \sin(\frac{23,52941\pi t}{100})$ , para  $t = [1..,1900]$

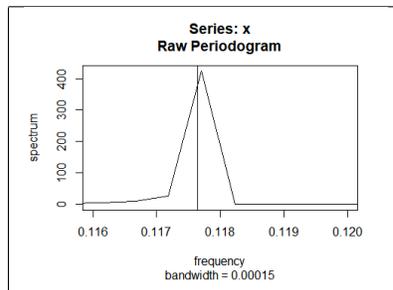


Figura 11: Ejemplo del punto intermedio.

Se observa que el máximo no coincide con el período de la serie.

El periodograma clasico, posee ambos problemas, limite de frecuencia Nyquist y la resolución del período. period04 y Lomb-scargle solo poseen el problema delimita de frecuencia Nyquist.

#### 4.0.3. Efecto Eco o resonancia:

Es un fenómeno de ruido en el periodograma, un aumento de poder en este, generado en los períodos resonantes con el período de más poder. Suele ocurrir en series de tiempo en las que la evidencia de un período está dada en un intervalo pequeño, como suelen ser las curvas de luz.

##### Caso BLS:

Ocurre en períodos resonantes, por ejemplo 1/2 o x2 del período máximo, regido por el intervalo pequeño. Donde aún existen muchos puntos diferentes al resto de la serie.

En este ejemplo se ve una simulación de una serie temporal, de período 440, con el plegado para 220, 440, 880. En estos puntos resonantes, se logran alinear muchos puntos que generan un aumento fantasma del poder, en ese período.

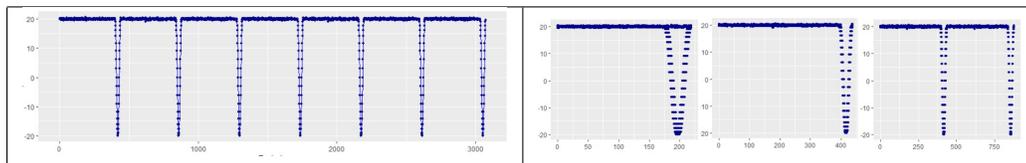


Figura 12: Al lado izquierdo la curva de luz ejemplo, lado derecho plegándola a períodos resonantes del real, donde aún puede marcar alto poder en el periodograma.

### Casos DFT:

Se genera por la resolución (un excesivo sobreajuste), como en las curvas de luz el tránsito ocurre en un pequeño intervalo, el algoritmo (elementos de la base) deben ajustar los puntos fuera de ese intervalo, al variar la cantidad de puntos o elementos de la base, se reduce (pero también lo hace la resolución). (sera bueno poner un ejemplo con imagenes de esto?)

## Parte III

# Analisis de desempeño de los periodogramas en curvas de luz.

## 5. Introducción:

En esta sección usaremos los periodogramas antes presentados para el análisis en curvas de luz de sistemas, las cuales se encuentra evidencia confirmada de existencia de exoplanetas (tránsito de exoplanetas). Primero se expondrá de donde se obtuvieron, como y porque se eligieron, para luego como serán agrupadas para su análisis. Luego se usaran los 4 periodogramas presentados para buscar en estas curvas de luz evidencia de tránsitos.

### 5.1. Obtención e onformación de las curvas de luz seleccionadas.

#### 5.1.1. Obtencion y eleccion:

Las curvas de luz utilizadas son obtenidas de la base de datos de la NASA y la elección de estas será de la siguiente manera:

- La página Open Exoplanet Catalogue contiene la base de todos los exoplanetas descubiertos y confirmados hasta la fecha, de todos los observatorios y todos los métodos de detección.
- De esta base de datos se extraerán todos los exoplanetas confirmados via tránsito.
- De ellos, se filtrarán las curvas de luz entregadas por el observatorio satelital Kepler y K2. Estas además de ser la mayoría, aproximadamente el 80 % de los exoplanetas descubiertos confirmados via tránsito han sido utilizando curvas de luz obtenidas por este observatorio, estan disponibles directamente para su descarga en la página de la NASA.
- Finalmente obtendremos una lista de 1755 sistemas de exoplanetas descubiertos confirmados. <sup>2</sup>

Esta lista se dividirá en 5 grupos, los cuales son:

- Grupo 1: Sistemas con 1 exoplaneta conformado y 1 estrella, que estén entre los primeros 100 sistemas con exoplanetas confirmados del observatorio

---

<sup>2</sup>Debemos referirnos a sistemas en general y no estrellas o exoplaneta, pues puede tener mas de 1 estrella o mas de 1 exoplaneta.

Kepler, es decir, entre Kepler-1 y Kepler-100. Se encuentran 26 sistemas que cumplen con estas características.<sup>3</sup>

- Grupo 2: Sistemas con 1 exoplaneta confirmado y 1 estrella, que estén entre los 100 y 1000 sistemas con exoplanetas confirmados por Kepler, es decir, entre Kepler-101 y Kepler-1000. Se encuentran 513 sistemas que cumplen con estas características.
- Grupo 3: Sistemas con 1 exoplaneta confirmado y 1 estrella, el resto de sistemas que contienen 1 exoplaneta confirmado y 1 estrella. Los Kepler mayor que 1000 y K2. Se encuentran 740 sistemas que cumplen con estas características.
- Grupo 4: Sistemas con más de 1 exoplaneta confirmado y 1 estrella, de toda la lista. Se encuentran 469 sistemas que cumplen con estas características.
- Grupo 5: Sistemas con más de 1 exoplaneta confirmado y más de 1 estrella, de toda la lista. Se encuentran 15 sistemas que cumplen con estas características.

De esta forma se dividirá la lista antes obtenida, ésta se encuentra en Apéndice C. Además de esta lista, habrá un 6 grupo.

- Grupo 6, obtenida de la lista de todos los sistemas observados por Kepler, es una lista de casi un millón de sistemas. A modo de tener un contraste de curvas de luz donde no hay o no se ha descubierto evidencia de tránsito de exoplaneta.

La selección al azar se realizará utilizando el comando `sample` de R (expuesto en el apéndice ).

Se crearán 5 grupos de 3 curvas de luz con las siguientes características:

GRUPO I : de los 4 elegidos en el grupo 1, se seleccionarán 3 que tengan características diferentes entre sí en la medida de lo posible.

GRUPO II : de los 8 elegidos del grupo 2 y grupo 3, se seleccionarán 3 que tengan características diferentes entre sí.

GRUPO III: de los 8 elegidos del grupos 4 y 5, se seleccionarán 3 que tengan características diferentes entre sí.

GRUPO IV : se elegirán 2 periodogramas que no tengan características de los ya elegidos más uno extraído del grupo 6, es decir del no catalogado.

Cada curva de luz será analizada con los 4 periodogramas.

GRUPO V: Este grupo se seleccionarán 3 sistemas en que los periodogramas de los anteriores grupos tuvieron un mal desempeño, y se usara todos los tiempos de observaciones (que se explican un como más adelante) para graficar los periodogramas.

Finalmente, en la siguiente tabla 1, los sistemas obtenidos aleatoriamente y en **negrita** los que serán analizados:

---

<sup>3</sup>La base de datos kepler nombra las estrellas con un código KIC, cuando en un sistema se confirma exoplaneta es renombrada Kepler-X, siendo X el orden en que fue descubierto. ej: el primer exoplaneta descubierto confirmado por data obtenida del observatorio kepler fue Kepler-1, el segundo Kepler-2 etc...

Tabla 1: Tabla correspondiente a los sistemas obtenidos y elegidos

KIC	Grupo	Periodo	Planetas	Estrellas	Nombre sistema
<b>KIC 8191672</b>	<b>Grupo I</b>	<b>3.5485±0.0000</b>	<b>1</b>	<b>1</b>	<b>Kepler-5</b>
KIC 6922244	Grupo I	3.5225+0.0003-0.0001	1	1	Kepler-8
<b>KIC 10619192</b>	<b>Grupo I</b>	<b>1.486</b>	<b>1</b>	<b>1</b>	<b>Kepler-17</b>
<b>KIC 9305831</b>	<b>Grupo I</b>	<b>3.2467±0.0000</b>	<b>1</b>	<b>1</b>	<b>Kepler-44</b>
<b>KIC 7375348</b>	<b>Grupo II</b>	<b>25.3089±0.0001</b>	<b>1</b>	<b>1</b>	<b>Kepler-508</b>
KIC 3749365	Grupo II	1.9738±0.0000	1	1	Kepler-785
KIC 3962357	Grupo II	14.5540±0.0001	1	1	Kepler-791
KIC 3749134	Grupo II	11.3012±0.0001	1	1	Kepler-792
KIC 10583761	Grupo II	2.6851±0.0000	1	1	Kepler-1179
KIC 8494617	Grupo II	22.9221±0.0002	1	1	Kepler-1218
<b>KIC 12401863</b>	<b>Grupo II</b>	<b>2.8327±0.0000</b>	<b>1</b>	<b>1</b>	<b>Kepler-1193</b>
<b>KIC 9717943</b>	<b>Grupo II</b>	<b>6.1101±0.0000</b>	<b>1</b>	<b>1</b>	<b>Kepler-1169</b>
KIC 7440748	Grupo III	4.9214±0.0000	1	2	Kepler-161
EPIC 201754305	Grupo III	19.079±0.003	1	2	EPIC 201754305
<b>KIC 8277797</b>	<b>Grupo III</b>	<b>4.3372±0.0000</b>	<b>1</b>	<b>2</b>	<b>Kepler-322</b>
<b>KIC 6381846</b>	<b>Grupo III</b>	<b>4.1670±0.0000</b>	<b>1</b>	<b>3</b>	<b>Kepler-171</b>
<b>KIC 11295426</b>	<b>Grupo III</b>	<b>5.3988±0.0000</b>	<b>2</b>	<b>3</b>	<b>Kepler-68</b>
EPIC 201384232	Grupo III	30.942±0.005	2	1	EPIC 201384232
KIC 9941662	Grupo III	1.7636±0.0000	3	1	Kepler-13
KIC 3632418	Grupo III	2.7858±0.0000	2	1	Kepler-21
<b>KIC 10723750</b>	<b>Grupo IV</b>	<b>18.7959±0.0001</b>	<b>1</b>	<b>2</b>	<b>Kepler-117</b>
<b>KIC 10577994</b>	<b>Grupo IV</b>	<b>8.1808±0.0000</b>	<b>1</b>	<b>2</b>	<b>Kepler-165</b>

### 5.1.2. Información de la curva de luz:

Las curvas de luz entregadas por Kepler tienen las siguientes características:

- Miden la intensidad de luz vs el tiempo.
- Cada sistema es observado de manera ininterrumpida (sin contar pérdida de datos) por 89 días, con tiempo espaciado de 0.020434 de día (es decir cada 30 min aprox.). Eso se llamaría un tiempo o intervalo de observación, también los hay más cortos.
- Para cada sistema hay por lo general más de 1 tiempo de observación, se elegirá el primero que tenga un largo de 89 días.

- Cada tiempo de observación tiene un largo aproximado de 4000~ contando los datos faltantes.
- En el GRUPO V se ocuparán todos los tiempos de observación para el sistema dado, por lo general, estos tienen diferentes medias entre sí, pues el observatorio realiza ajustes entre tiempos de observación, por lo que para analizarlos todos en un mismo periodograma se normalizan por medio de división (es decir cada tiempo de observación será dividido por su media).
- Se trabajará con esta curva de luz mencionada.

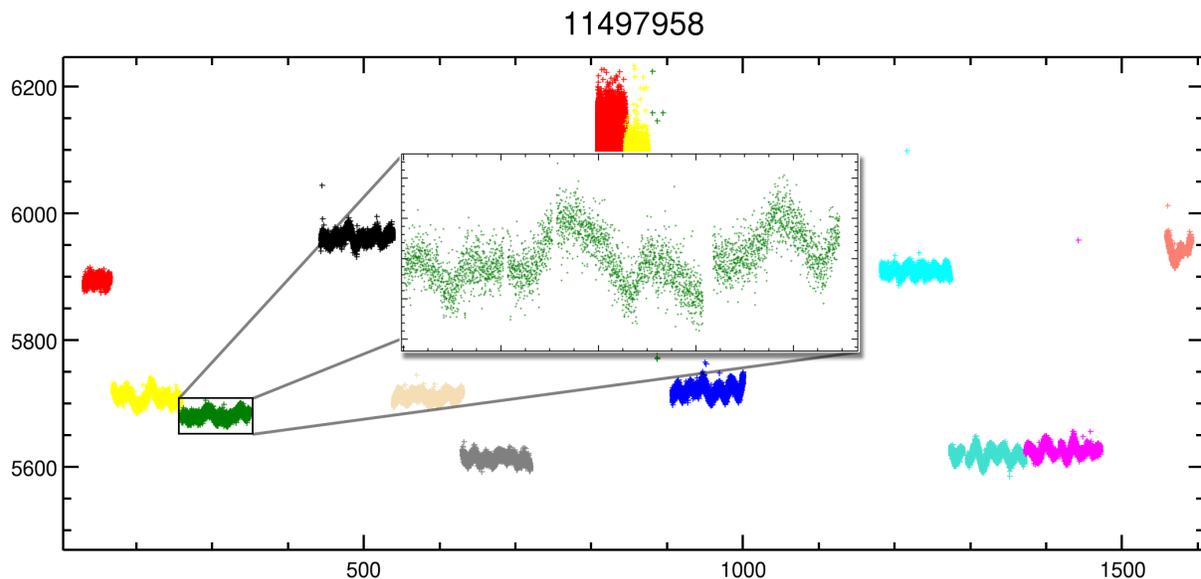


Figura 13: Se expone un tiempo de observación, de fondo los diferentes tiempos de observación para este sistema.

- La Fig.13 se expone los datos totales para un sistema (KIC 11497958) entregados por Kepler, se ven los diferentes tiempos de observación.
- No todos los sistemas tienen la misma cantidad de tiempos de observación, hay desde 1 hasta 18 o más, y no todos los tiempos de observación tienen la misma longitud.

## 5.2. Aplicación de los periodogramas

### 5.2.1. Clásico:

-Se usará el comando R, `spectrum()`, en modo no logarítmico, éste entrega el periodograma clásico antes formulado. Apéndice se adjunta un ejemplo del comando usado en R.

-Las curvas de luz suelen tener intervalos de datos faltantes, se usará la curva de luz de manera directa, pues es una debilidad propia del periodograma que tanto period04 como Lomb-Scargle la corrigen.

-Este al igual que el periodograma period04 son entregados en frecuencias, por lo que, para la obtencion del período debemos usar  $P = \frac{1}{f}$ .

-Además en este caso, como se asume equidistancia, ocupa solamente la columna de intensidad de luz, el muestreo del tiempo esta dado por 1 unidad, esta unidad es 0.020434 día en la curva de luz original, por lo que debe ser transformado a días, es decir, multiplicar por 0.020434, que es el tiempo en días de cada paso de la serie. -

### **5.2.2. Period04:**

-Se usa el software llamado period 04, <https://www.univie.ac.at/tops/Period04/>

-El software maneja internamente los missing data y el muestreo del tiempo original de las curvas de luz, por lo que el resultado (tambien entregado en frecuencias) es el del período previa transformacion,  $P = \frac{1}{f}$ .

### **5.2.3. Lomb-Scargle:**

-Se usará la librería lomb de R, esta entrega directamente los resultados necesarios para su analisis, tanto su grafico como sus datos. Apéndice se adjunta un ejemplo del comando usado en R. .

### **5.2.4. BLS:**

-Se construyó el algoritmo en R, se encuentra en la sección Apéndice, con un ejemplo.

-Al ser un método completamente no parametrico, entrega directamente el período.

-En todos los periodogramas, se ocupan 200 binds y tamaño máximo de intervalo 20 (10 %)

### 5.3. Metodología de obtención de desempeño.

**GRUPO I:** En este grupo, dado que son curva de luz donde el tránsito del exoplaneta es evidente a simple vista todos los periodogramas deberían tener un desempeño tal que detecten el tránsito correspondiente, por lo que su desempeño se basara en que tan preciso es el periodograma en encontrar este período. Se expondrán los resultados en tablas con el período real confirmado y el entregado por el periodograma.

**GRUPOS II-III-IV:** En estos grupos, los períodos por lo general no son visibles a simple vista, en análisis del desempeño se basara en lo siguiente:

**-Descubrir el período del exoplaneta:** Es decir, con el periodograma se pudiese descubrir el tránsito del exoplaneta dado, el período es un peak absoluto o dentro de los más altos del periodograma y también, características como el efecto resonancia descritos en el capítulo anterior.

Se ocuparan la siguiente clasificaron: si el peak es completamente absoluto o dentro de los más altos en forma clara, se clasificara como “SI”, si están dentro una cantidad relativamente alta de peak y no se muestra tan evidente (por ejemplo hay muchos peak altos), se clasificara como “posible”. Si no se detecta se clasificará como “No”

**-Detectar el período del exoplaneta:** Es si el periodograma logra detectar un peak local el tránsito del exoplaneta. Es decir, si se ve reflejado el tránsito confirmado en el periodograma.

Se ocuparán la siguiente clasificaron: si el peak es claro en un intervalo de tiempo aceptable, se dirá que “Sí”, si el periodograma muestra un valor alejado del cero o un peak en un intervalo muy pequeño, se describirá como “posible”. Si no detecta nada “no”.

Luego los resultados se expondrán en una tabla, a modo de ejemplo en la Tabla 2 :

Tabla 2: Tabla ejemplo

Descubrir/Detectar	Ejem 1	Ejem 2
Periodograma 1	No/Si	Posible/Si
Periodograma 2	No/No	Si/Si

En el ejemplo se ve que en el sistema Ejem 1, el periodograma 1 detecta el tránsito, pero no se logra descubrir; Ejem 2, periodograma 1 detecta el tránsito y también es posible descubrir este tránsito; Ejem 1, el periodograma 2 no logra ni detectar ni descubrir el tránsito; Ejem 2, periodograma 2 detecta y se puede descubrir el tránsito.

Finalmente, si se descubre o es posible descubrir el tránsito correspondiente, detecta este mismo tránsito.

**GRUPO V:** En este grupo se elegirán 3 curvas en las que el desempeño de los periodogramas no fuese bueno (no detectan ningún exoplaneta) o algunos de los pe-

riodogramas no tienen un buen desempeño con respecto a otro. Se usará todos los datos disponibles (todos los tiempos de observación) y con ellos se generará un nuevo periodograma.

En este grupo se comparará que tanto aumenta el desempeño al aumentar esta gran cantidad de puntos, desplegando ambos periodogramas a modo de comparación, y viendo características que se puedan deducir del aumento de puntos.

## 6. Grupo I

En este grupo se encuentran 3 curvas de luz seleccionadas de los primeros 100 sistemas catalogados por el observatorio Kepler, que cumplen la condición de ser 1 planeta 1 estrella. Estas curvas de luz se caracterizan por su evidente, a simple vista, tránsito del exoplaneta. Que como sabemos es un descenso en un intervalo pequeño de la luminosidad.

En la fig.14 se encuentran las 3 curvas seleccionadas:

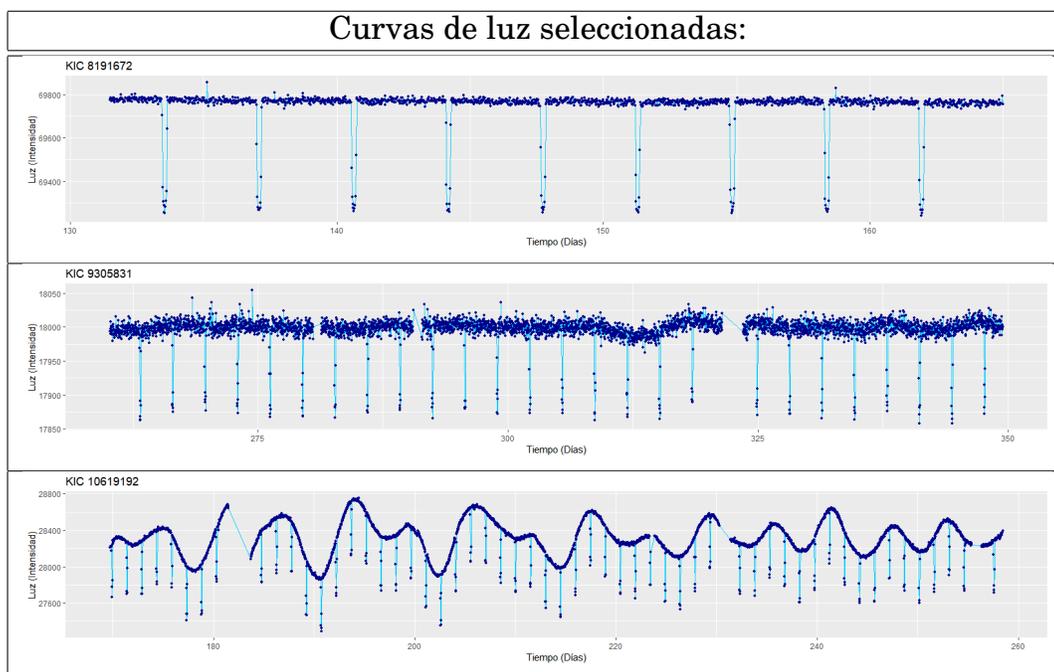


Figura 14: Curvas de luz seleccionadas Grupo I

Algunas características de estas curvas de luz seleccionadas:

**KIC 9191672:** Presenta poco ruido y un muy evidente tránsito (tanto en profundidad como amplitud), el período confirmado del exoplaneta es de 3.548 días.

**KIC 9305831:** También es muy evidente el tránsito, pero tiene un mayor ruido, el período confirmado del exoplaneta es de 3.246 días.

**KIC 10619192:** Es evidente el tránsito, pero la estrella tiene una luminosidad variable, el período confirmado del exoplaneta es de 1.486 días.

### 6.0.1. Periodograma clásico

En esta subsección se detallan los resultados obtenidos al aplicar el periodograma clásico a las 3 curvas seleccionada.

En la fig.15 están los Periodogramas correspondientes:

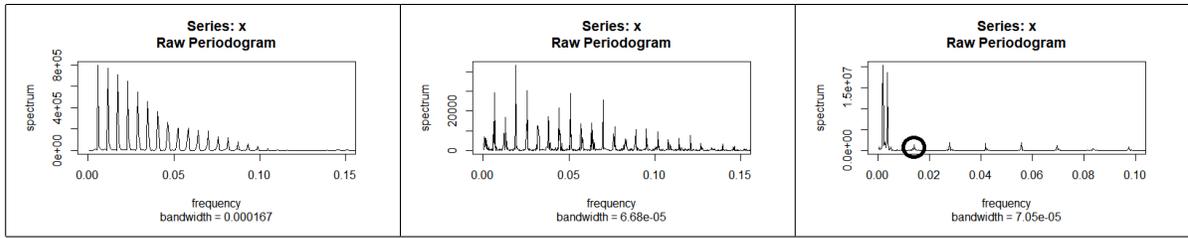


Figura 15: Periodograma clásico de curvas de luz Grupo I

**KIC 9191672:** Primero se ve claramente el período del exoplaneta, ocurriendo el fenómeno de resonancia (4.0.3), como candidato elegimos el de más a la izquierda.

**KIC 9305831:** También se observa el tránsito, Pero la curva de luz al tener mayor ruido hace que su periodograma también presente mas ruido, como candidato elegimos el de más a la izquierda.

**KIC 10619192:** Vemos que la variabilidad de la luminosidad de la propia estrella se lleva toda la fuerza del periodograma, aun así, se observa un pequeño efecto de resonancia que evidencia la presencia del tránsito. Eligiendo como candidato el indicado en la fig. 15 con un círculo negro.

Como explicamos anteriormente, el Grupo I se contrastará es el período entregado por los periodogramas con el período real. Los resultados del periodograma clásico en la siguiente Tabla. 3:

Tabla 3: Período entregado por periodograma clásico y período real.

Sistema	KIC 9191672	KIC 9305831	KIC 10619192
Clásico	3.53	3.043	1.468
Real	3.548	3.246	1.486

A pesar de ser el periodograma mas simple, y las propias limitaciones de éste, como los problemas con su desempeño en presencia de datos faltantes, son pronósticos bastante certeros.

### 6.0.2. Period04

En esta subsección se detallan los resultados obtenidos por el software Period04, a continuación en la fig.16 los periodogramas de las curvas de luz del grupo I:

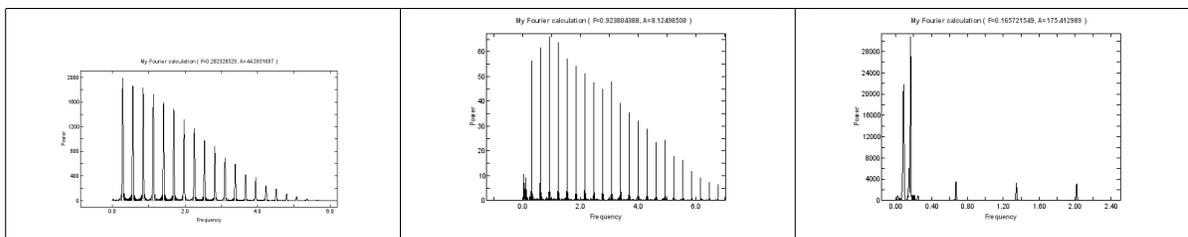


Figura 16: Periodograma entregado por Period04 de curvas de luz Grupo I

**KIC 9191672:** Se observa un periodograma muy similar al anterior, como candidato elegimos el de mas a la izquierda.

**KIC 9305831:** Ocurre lo mismo que el periodograma clasico, pero esta vez es mas definido, como candidato elegimos el de mas a la izquierda.

**KIC 10619192:** Se Observa el período de la variabilidad de la estrella, pero nuevamente el del tránsito se observa mas definido, se elige el mismo.

En la siguiente Tabla 4 se contrasta el período entregado por period04 y el período real:

Tabla 4: Período entregado por Period04 y período real.

Sistema	KIC 9191672	KIC 9305831	KIC 10619192
Period04	3.542	3.236	1.483
Real	3.548	3.246	1.486

Las mejoras al tratar los datos faltantes y la equiespaciedad se hacen evidentes, obteniendo los resultados casi idénticos a los reales.

### 6.0.3. Lomb-Scargle

A continuación, se verá los resultados entregados por el periodograma de Lomb-Scargle, en la fig.17 sus graficos:

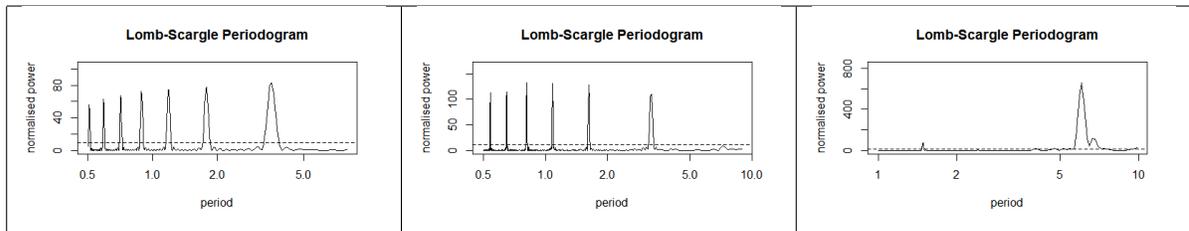


Figura 17: Periodograma Lomb-Scargle de curvas de luz Grupo I

**KIC 9191672:** Se observa un periodograma muy similar al anterior, como candidato elegimos el de mas a la derecha (pues este algoritmo entrega el resultado directamente en período).

**KIC 9305831:** Ocurre lo mismo que en los otros 2 periodogramas, como candidato elegimos el de mas a la derecha.

**KIC 10619192:** Se Observa muy fuerte el período de la variabilidad de la estrella, aun así, es distinguible el valor del período del exoplaneta, elegimos el pequeño de la izquierda.

En la Tabla 5 se entrega los resultados calculados por el periodograma de Lomb-Scargle contrastados con el período real.

Tabla 5: Período entregado por Lomb-Scargle y período real.

Sistema	KIC 9191672	KIC 9305831	KIC 10619192
Lomb-Scargle	3.56	3.265	1.486
Real	3.548	3.246	1.486

También los resultados son casi idénticos a los reales (y dentro del margen de error del tránsito).

#### 6.0.4. BLS

Finalmente, el periodograma BLS, en la fig.18 se expondrá el periodograma en las curvas de luz del Grupo I.

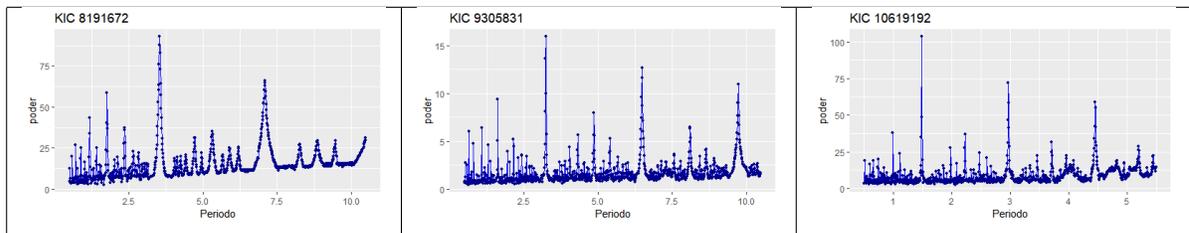


Figura 18: Periodograma BLS de curvas de luz Grupo I

**KIC 9191672:** El punto de mayor potencia es nuestro candidato, notamos también el efecto resonancia4.0.3.

**KIC 9305831:** Nuevamente el punto de mayor potencia es nuestro candidato,

**KIC 10619192:** Acá encontramos una gran diferencia, al ver el periodograma completo VI, notamos que el período«real» es bastante mas potente en relación a los períodos de la propia estrella, que en los periodogramas anteriores, es decir, este periodograma se ve menos afectado por períodos de cambio de brillo de la propia estrella.

En la Tabla 6 los períodos seleccionados por BLS, contrastados con el período real:

Tabla 6: Período entregado por Lomb-Scargle y período real.

Sistema	KIC 9191672	KIC 9305831	KIC 10619192
BLS	3.55	3.26	1.485
Real	3.548	3.246	1.486

Los resultados son los mas certeros, tomando en cuenta que el 1 y el 2 se tomaron pasos de 0.01 día, se podría encontrar un valor mas exacto, como el 3 que se uso pasos de 0.005, en el algortimo.

### **6.0.5. Conclusiones:**

Vemos que los 4 periodogramas no tienen problemas en encontrar el período real del exoplaneta, y entregan excelentes resultados con curvas del Grupo I. Notamos que el comportamiento de los 3 periodogramas basados en DFT fue muy similar, salvo el clásico que, debido a su debilidad con los intervalos de datos perdidos, pierde precisión. A pesar de todo, con cualquiera de los 4 periodogramas sería posible encontrar el exoplaneta dado, con una precisión muy buena comparada con la real. Notamos que el efecto ruido, donde es más fuerte en el sistema KIC 9305831 comparado con KIC 9191672, los periodogramas basados en DFT se muestran más “irregulares”, sobretodo el periodograma clásico, perdiendo la regularidad del efecto “resonancia” y el peak (punto más alto del periodograma) no siendo el del período real sino uno de esos puntos resonantes. El periodograma BLS no muestra diferencia dramática entre estos 2 sistemas.

También notamos que el periodograma BLS en el sistema KIC 10619192 hace un mejor manejo de la variabilidad del sistema (luminosidad de su estrella no constante) pues toda la potencia de los periodogramas DFT es dada por la variabilidad de la propia estrella, y para encontrar los períodos correspondientes de los tránsitos hay que tener en cuenta estos valores en el periodograma (los grandes peak son debidos a los cambios de luminosidad de la estrella). El periodograma BLS casi no se ve reflejado esto, y el peak más alto es efectivamente el del tránsito.

## 7. Grupo II

En este grupo se encuentran las curvas de luz seleccionadas de entre el resto de sistemas con 1 planeta y 1 estrella del observatorio satelital Kepler. Se caracterizan por no tener un tránsito evidente a simple vista (a diferencia del Grupo I). También se cambia la calificación de su desempeño que a diferencia del Grupo I, se basa en ver si es posible descubrir el exoplaneta, si es posible detectar o simplemente no muestra evidencia del período real en cada periodograma. Como se explico anteriormente se expondrá en una tabla el resultado que se basa en que tan evidente en el periodograma se muestra el período real (tránsito confirmado) del exoplaneta. En la fig. 19 se expone las curvas de luz de los sistemas a analizar:

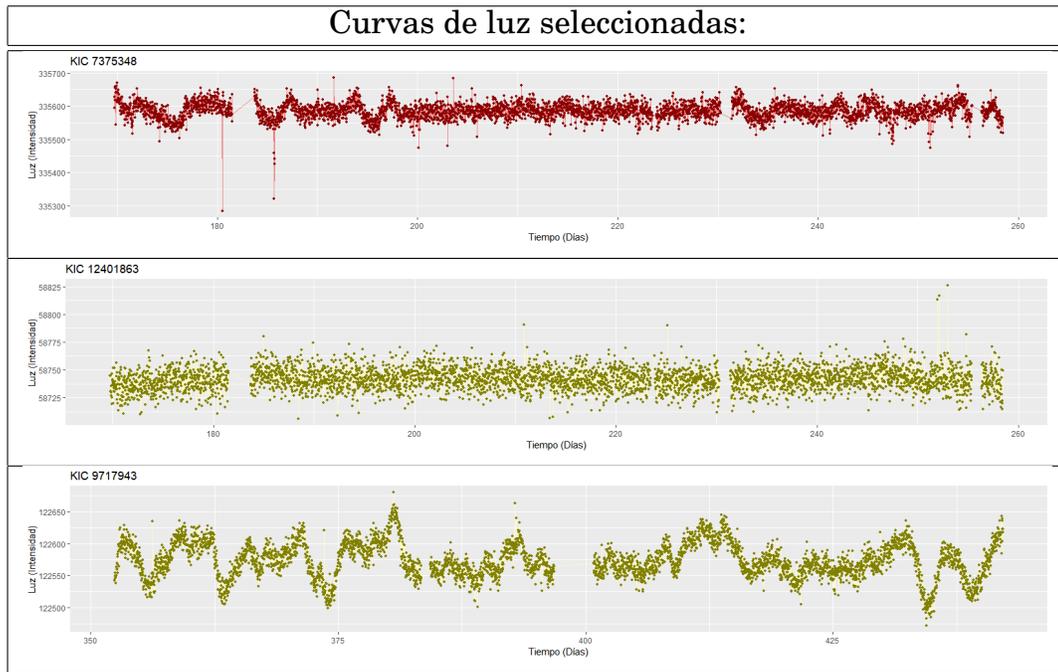


Figura 19: Curvas de luz seleccionadas Grupo II

Se eligieron las curvas con tránsito no evidente a simple vista. Y que tengan diferentes formas y período del exoplaneta. Como se detalla a continuación:

**KIC 7375348:** Es una curva de luz con período confirmado de 25.3 días, es decir, fue elegida por su relativo largo período, además de contener perturbaciones de la propia luminosidad de la estrella.

**KIC 12401863:** Es una curva de luz con período confirmado de 2.832 días, fue elegida por su muy constante luminosidad. Y similitud visual a solo un ruido gaussiano.

**KIC 9717943:** Es una curva de luz con período confirmado de 6.11 días, fue elegida por su gran cantidad de perturbaciones de la propia estrella.

### 7.0.1. Periodograma clásico

En esta subsección se detallan los resultados obtenidos al aplicar el periodograma clásico a las 3 curvas seleccionadas.

En la fig.20 están los periodogramas correspondientes, cada uno contiene una recta con el período real:

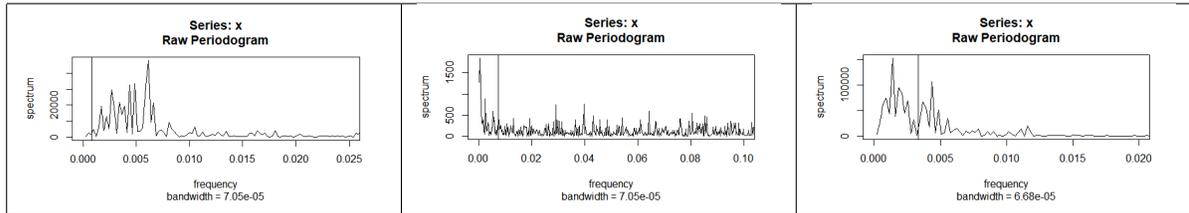


Figura 20: Periodograma clásico de curvas de luz Grupo II

**KIC 7375348:** No muestra evidencia de tránsito, el período confirmado no se ve reflejado en el periodograma.

**KIC 12401863:** No muestra evidencia de tránsito, el período confirmado no se ve reflejado en el periodograma.

**KIC 9717943:** No muestra evidencia de tránsito, el período confirmado no se ve reflejado en el periodograma.

En ninguno de los 3 sistemas logra detectar o mostrar evidencia de tránsito en el período correspondiente. Probablemente sea por el manejo de los intervalos de datos perdidos de este periodograma, donde acá denoten mayor importancia esta debilidad, al ser curvas de luz con un tránsito más difícil de detectar.

### 7.0.2. Period04

En esta subsección se detallan los resultados obtenidos por el software Period04, a continuación en la fig.21 los periodogramas de las curvas de luz del grupo II, cada uno contiene una recta con el período real:

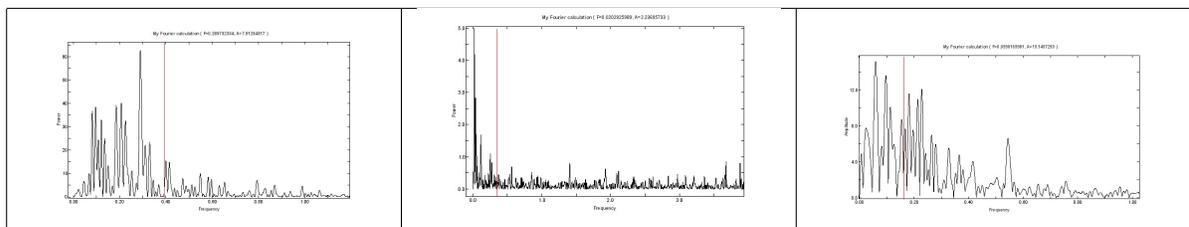


Figura 21: Periodograma entregado por Period04 de curvas de luz Grupo II

**KIC 7375348:** No muestra evidencia de tránsito, el período confirmado no se ve reflejado en el periodograma.

**KIC 12401863:** No muestra evidencia de tránsito, el período confirmado no se ve reflejado en el periodograma.

**KIC 9717943:** No muestra evidencia de tránsito, el período confirmado no se ve reflejado en el periodograma.

Nuevamente en ninguno de los 3 periodogramas se logra detectar el período real, este periodograma si hace manejo del espacio de datos perdidos, pero completando estos intervalos.

### 7.0.3. Lomb-Scargle

A continuación, se verá los resultados entregados por el periodograma de Lomb-Scargle, en la fig.22 sus graficos, con una recta en el punto del período real:

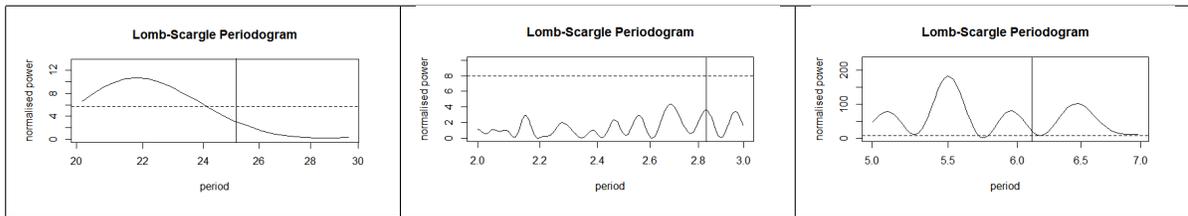


Figura 22: Periodograma Lomb-Scargle de curvas de luz Grupo II

**KIC 7375348:** No muestra evidencia de tránsito, el período confirmado no se ve reflejado en el periodograma.

**KIC 12401863:** Acá si hay cierta evidencia de detección por parte del periodograma, podríamos decir que es una posible detectar el Periodo confirmado. Se adjunta periodograma completo VI. Por otra parte, notamos que no seria posible descubrir el tránsito (que sea evidente en el periodograma).

**KIC 9717943:** No muestra evidencia de tránsito, el período confirmado no se ve reflejado en el periodograma.

A diferencia de los otros 2 periodogramas basados en DFT, este si detecta el período real del sistema KIC 12401863, aunque de forma muy tenue. Podría ser evidencia de que Lomb-Scargle, hace un mejor manejo de datos perdidos.

### 7.0.4. BLS

Finalmente, el periodograma BLS, en la fig.23 se expondrá el periodograma en las curvas de luz del Grupo II, con una recta en el punto del período real.

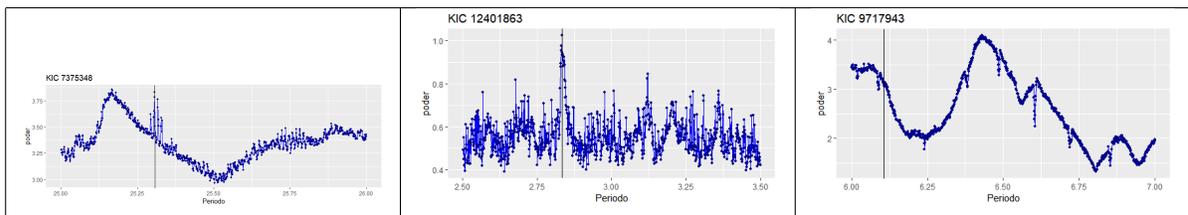


Figura 23: Periodograma BLS de curvas de luz Grupo II

**KIC 7375348:** El periodograma, con baja resolución (fig. anexo), se ve un punto extraño donde está el período confirmado, un pequeño salto abrupto. Al observar a

mayor resolución, que corresponde a la imagen de la fig.23, se ve el punto donde se encuentra el período real, un máximo en el periodograma ahí, por lo que se detecta este período.

**KIC 12401863:** Al ver con baja resolución en la fig. del anexo VI, se ve un máximo en el intervalo [0,10]. al ver mas de cerca, se confirma este punto corresponde al período real, por lo que en este caso, seria posible descubrir el exoplaneta. Es clara su detección al ser el punto mas alto.

**KIC 9717943:** No queda evidenciado el tránsito, La curva del periodograma es suave, no se observa ningún pick sobresaliente.

Acá se aprecia una gran diferencia con respecto a los periodogramas basados en DFT, se logra detectar el tránsito en 2 sistemas, y en el KIC 12401863 el período real es muy evidente en el periodograma.

### 7.0.5. Conclusiones:

Como se explicó anteriormente, descubrir implica que el período real es el de más poder o uno de los más poder en el periodograma, detectar si es un mínimo local (peak) o se logra ver un aumento en el punto donde está el período real. En la Tabla 7 se resume el resultado de los periodogramas en las curvas de luz dadas:

Tabla 7: Tabla correspondiente a desempeño de los periodogramas en el Grupo II

Descubrir/Detectar	KIC 7375348	KIC 12401863	KIC 9717943
Clásico	No/No	No/No	No/No
Period04	No/No	No/No	No/No
Lomb-Scargle	No/No	No/ <b>Posible</b>	No/No
BLS	No/ <b>Posible</b>	<b>Posible/Si</b>	No/No

Lo primero que notamos es que en este grupo donde los tránsitos en las curvas de luz están mucho más ocultos y no notorios a simple vista, el desempeño de los periodogramas basados en DFT es pésimo, pues en los 3 sistemas no detectaron ningún tránsito, salvo el periodograma Lomb-Scargle donde apenas se ve reflejado el tránsito en el sistema KIC 12401863.

Por otra parte, el periodograma BLS logra detectar el tránsito de 2 sistemas, destacando en el KIC 12401863 donde es muy notorio este máximo. Finalmente, el sistema KIC 9717943 ninguno detecta el tránsito y es un buen candidato para el grupo V donde se hace el análisis con una mayor cantidad de tiempos de observación.

En este grupo el desempeño del periodograma BLS fue mejor en forma absoluta a de los periodogramas basados en DFT.

## 8. Grupo III

En este grupo se encuentran las curvas de luz seleccionadas de sistemas con más de un exoplaneta confirmado y con una o más estrellas en el sistema. Su característica principal con respecto a las curvas de luz de los grupos anteriores es que se encuentran más de un período real confirmado en cada curva de luz. En la fig.24 se exponen las curvas de luz de estos sistemas seleccionados:

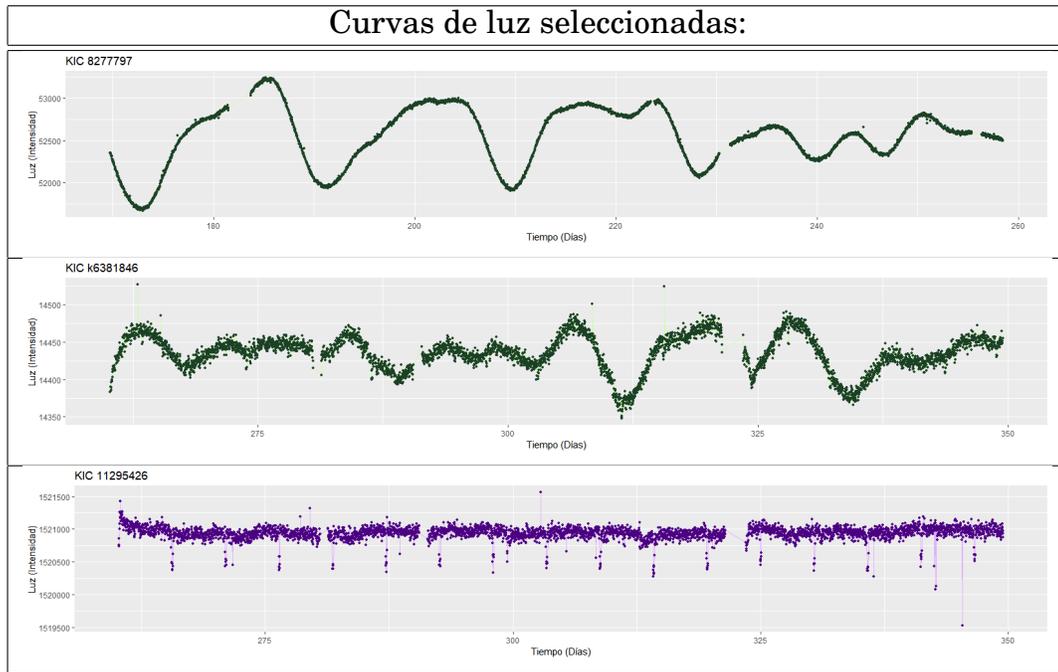


Figura 24: Curvas de luz seleccionadas Grupo III

Se eligieron por las siguientes características singulares:

**KIC 8277797:** Es una curva de luz 2 exoplanetas confirmados, de períodos 4.33 y 1.65 días, además de 1 estrella. Fue elegida por tener muy poco ruido y a la vez, ser muy evidente los cambios de luminosidad de la propia estrella, Además de que los tránsitos no son visibles a simple vista.

**KIC 6381846:** Es una curva de luz con 3 exoplanetas confirmados, de períodos 4.16, 11.46 y 39.59 días, que tampoco se detectan a simple vista, además de 1 estrella. Fue elegida por ser un sistema con 3 planetas confirmados y de muy diferente período.

**KIC 11295426:** Es una curva de luz con 3 exoplanetas confirmados, de períodos 5.39, 9.60 y 580 días. Pero solo los 2 primeros serán analizados, pues el tercero de 580 días fue descubierto por método de velocidad radial 1.5.1 además, este largo período hace imposible su detección con la curva de luz seleccionadas en este trabajo. Fue elegido por ser un sistema con 2 estrellas y un tránsito evidente a la vista.

### 8.0.1. Periodograma Clásico

En esta subsección se detallan los resultados obtenidos al aplicar el periodograma clásico a las 3 curvas seleccionadas.

En la fig.25 están los periodogramas correspondientes, cada uno contiene las rectas correspondientes a los períodos reales de cada sistema:

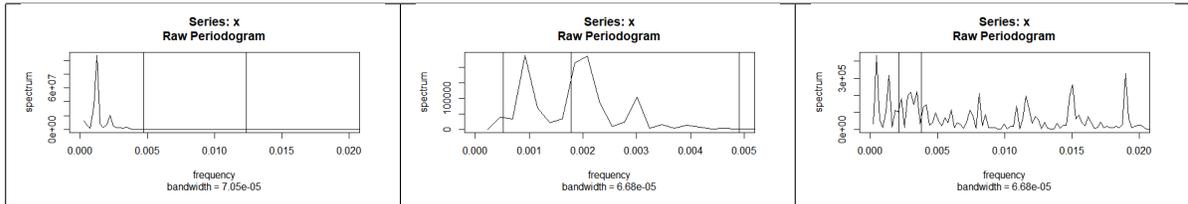


Figura 25: Periodograma clásico de curvas de luz Grupo III

**KIC 8277797:** No hay evidencia alguna de tránsito, pues no hay ningún peak en los valores correspondientes a los períodos de los tránsitos.

**KIC 6381846:** Podría haber evidencia, pues el segundo período está cerca de un máximo local, pero caer en el fenómeno de aliasing (descrito en 4.0.2), también por la interferencia de los intervalos de datos perdidos. Aunque probablemente se deba a que el período del segundo exoplaneta sea muy similar al período de la variabilidad de la luminosidad de la propia estrella, y

**KIC 11295426:** No hay evidencia alguna de tránsito, pues no hay ningún peak en los valores correspondientes a los períodos de los tránsitos.

Vemos que el periodograma clásico no muestra ningún máximo (peak) coincidente con los períodos reales, para ninguna curva de luz.

### 8.0.2. Period04

En esta subsección se detallan los resultados obtenidos por el software Period04, a continuación en la fig.26 los periodogramas de las curvas de luz del grupo III, cada uno contiene las rectas con los períodos reales:

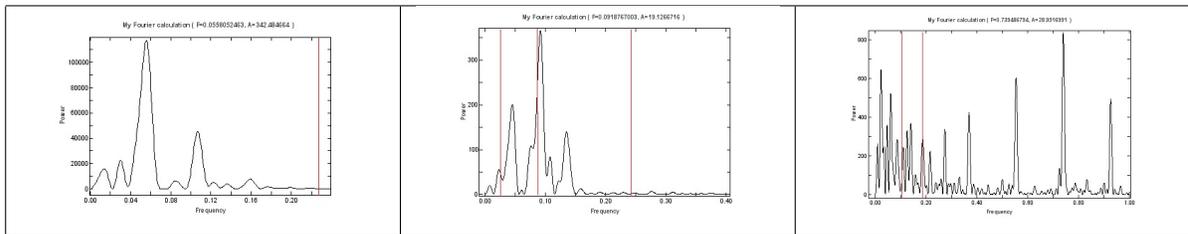


Figura 26: Periodograma entregado por Period04 de curvas de luz Grupo III

**KIC 8277797:** No hay evidencia alguna de tránsito, pues no hay ningún pick en los valores correspondientes a los períodos de los tránsitos.

**KIC 6381846:** El segundo exoplaneta podría ser pero al hacer el paso de frecuencia a días, es mas de 1 día de diferencia. Este periodogrmaa no posee el problema

del periodograma clasico, correspondiente a lo expuesto en aliasing4.0.3. Por lo que podríamos deducir que efectivamente este período corresponde al rastro de la variabilidad lumínica de la propia estrella. No se logran detectar ningún período real. **KIC 11295426:** Por la forma que va tomando el periodograma de efecto resonancia (4.0.3), llegando hasta la frecuencia 0.1852 que corresponde al período 5.39 se ve evidencia del tránsito de un exoplaneta y ademas por esta firma característica se podría descubrir por esta vía.El exoplaneta de período 9.6 no muestra evidencia alguna.

A diferencia del periodograma clásico se podría descubrir un exoplaneta del tercer sistema ha analizado, el resto de los períodos no son detectados.

### 8.0.3. Lomb-Scargle

A continuación, se verá los resultados entregados por el periodograma de Lomb-Scargle, en la fig.27 sus graficos, con rectas en los puntos de períodos reales para cada periodograma:

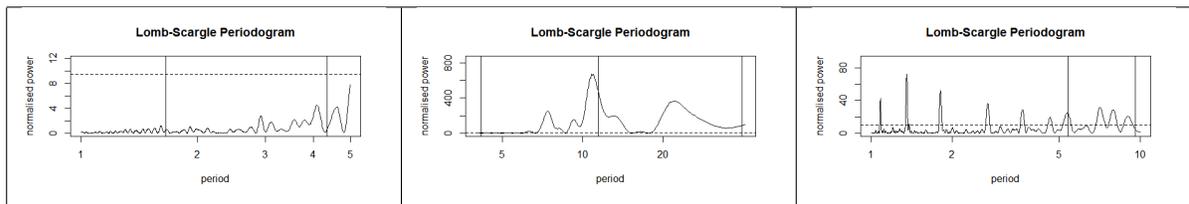


Figura 27: Periodograma Lomb-Scargle de curvas de luz Grupo III

**KIC 8277797:** No hay evidencia alguna de tránsito, pues no hay ningún pick en los valores correspondientes a los períodos de los tránsitos.

**KIC 6381846:** No hay evidencia alguna de tránsito, pues no hay ningún pick en los valores correspondientes a los períodos de los tránsitos, se confirma finalmente que el peak cercano corresponde a la variabilidad de la estrella.

**KIC 11295426:** Ocurre el mismo fenomeno que en el periodograma de Period04, por lo que es posible descubrir el exoplaneta de período 5.39. El exoplaneta de período 9.6 no muestra evidencia alguna.

Resultados muy similares a los obtenidos por el software period04, pudiendo descubrir un exoplaneta del 3 sistema de este grupo.

### 8.0.4. BLS

A continuación, se exponen los periodogramas calculador vía BLS, exponiendo también una línea negra que corresponde a los períodos confirmados. Cabe notar que, en este caso, los periodogramas para los sistemas 1 y 2 expuestos en fig.28 no se encuentran todos los períodos reales, solo los que son detectados por el periodograma. el resto de ellos se encuentran en la fig.36 en el anexo, para el sistema 3 si logran estar los 2 períodos reales expuestos en el periodograma:

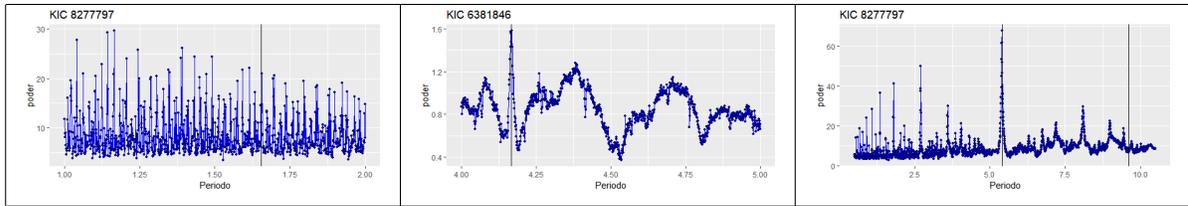


Figura 28: Periodograma BLS de curvas de luz Grupo III

**KIC 8277797:** Se detecta solo evidencia de tránsito del primer exoplaneta de período 1.65, el de período 4.33 no muestra evidencia (ver fig.36 anexo).

**KIC 6381846:** Se detecta solo evidencia de tránsito del primer exoplaneta de período 4.16, en forma mas o menos clara. los otros 2 no muestran evidencia (fig. 36 anexo)

**KIC 11295426:** Se descubre el tránsito del exoplaneta de período 5.39, no se detecta el otro exoplaneta.

A diferencia de los periodogramas basados en DFT, el periodograma de BLS detecto en cada uno de los sistemas al menos un período real. Y en el 3 sistema pudiendo descubrir este período.

### 8.0.5. Conclusiones:

Tabla que expone si es posible descubrir y/o detectar exoplaneta con el periodograma dado. Descubrir implica que el período real es el de más poder o uno de los más poder en el periodograma. Y detectar si es un mínimo local o se logra ver un aumento en el punto donde está el período real. En la Tabla 8 se resume el resultado de los periodogramas en las curvas de luz dadas:

Tabla 8: Tabla correspondiente a desempeño de los periodogramas en el Grupo III

Descubre/Detecta	8277797b	8277797c	6381846b	6381846c	6381846d	11295426b	11295426c
Clásico	No / No	No / No	No / No	No / No	No / No	No / No	No / No
Period04	No / No	No / No	No / No	No / No	No / No	<b>posible/ Si</b>	No / No
L.Scargle	No / No	No / No	No / No	No / No	No / No	<b>posible/ Si</b>	No / No
BLS	<b>No / Posible</b>	No / No	No/ <b>Si</b>	No / No	No / No	<b>Si / Si</b>	No / No

Notamos nuevamente que el periodograma BLS muestra el mejor desempeño, detectando al menos un tránsito en cada uno de los sistemas analizados, no pudiendo ningún periodograma detectar 2 o más tránsitos.

En cuando a los periodogramas basados en DFT, el clásico esta vez quedo rezagado con respecto a period04 y L-S, pues no logro detectar tránsito en KIC 11295426, probablemente por los grandes intervalos de datos perdidos en la curva de luz. El desempeño tanto de period04 como L-S fue en este grupo muy similar.

Con respecto a los tipos de curva de luz, nuevamente notamos que los DFT son muy propensos a detectar fuertemente los cambios lumínicos de la propia estrella (que tienen forma sinusoidal), quedando patente en los 2 primeros sistemas. El tercer sistema se asemeja mucho a los sistemas del grupo 1, por lo que el desempeño de los periodogramas fue muy similar a estos (salvo la acotación del periodograma clásico), el segundo tránsito de este sistema no fue detectado por ningún periodograma.

Finalmente elegimos el sistema KIC 8277797 para ser analizado en el grupo V, pues presenta poco ruido y el desempeño de los periodogramas fue malo, podía haber sido también el sistema KIC 6381846, pero este se asemeja más al ya elegido del grupo II.

## 9. Grupo IV

En este grupo se encuentran las curvas de luz seleccionadas de sistemas de elección personal, que mostraran características diferentes a los sistemas de los grupos anteriores. Como por ejemplo que apenas se pueda distinguir el tránsito, que el período sea de muchos días y de una curva de luz no catalogada. En la fig.29 se exponen las curvas de luz a analizar:

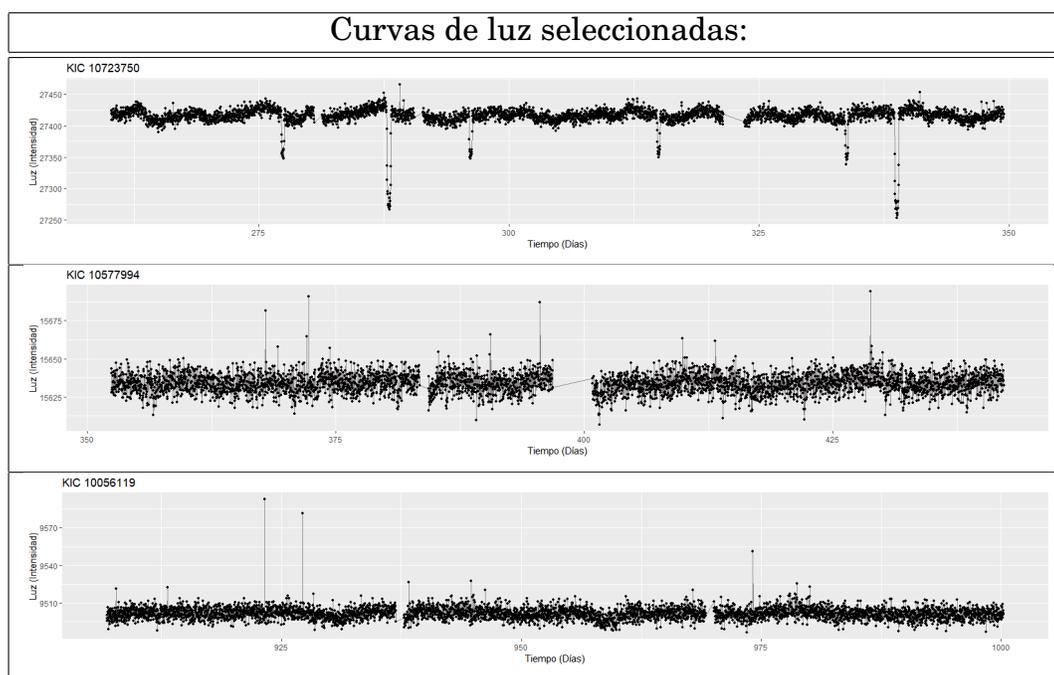


Figura 29: Curvas de luz seleccionadas Grupo IV

Las propiedades de las curvas de luz son las siguientes:

**KIC 10723750:** Es una curva de luz con 2 exoplanetas confirmados de períodos 50.79 y 18.79 con 1 estrella, la característica interesante es que son evidentes los 2 tránsitos teniendo profundidad distinta, y además, el período de estos es de muchos días. Finalmente el brillo de la propia estrella varía en períodos más cortos que los del tránsito de los exoplanetas.

**KIC 10577994:** Es una curva de luz con 2 exoplanetas confirmados de períodos 8.1808 y 15.313 con 1 estrella. La característica es que a simple vista, son apenas distinguible los tránsitos.

**KIC 10056119:** Esta curva de luz es de una estrella no catalogada. es decir, está en la base de datos de Kepler, pero no se a buscado o no se ha encontrado exoplanetas en ella. Se realizara un análisis diferente, pues no poseemos las herramientas como para poder verificar si posiblemente pudiese tener un exoplaneta. Este análisis se basara en ver cómo se comportan sus periodogramas, y si son muy diferentes para los caso en que hay tránsito extraplanetario.

### 9.0.1. Periodograma clásico

A continuación se entregan los resultados entregados por el periodograma clásico, para dado sistema, en la Fig. 30. Además, se adjuntan en una línea negra el tránsito confirmado correspondiente.

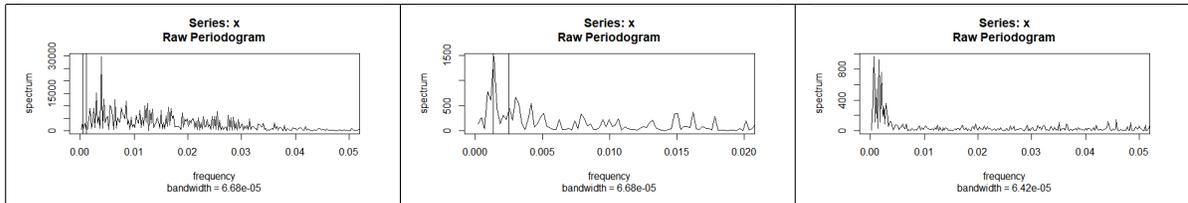


Figura 30: Periodograma clásico de curvas de luz Grupo IV

**KIC 10723750:** No se ve evidencia de tránsito alguno, no hay ningún peak en los valores correspondientes a los períodos de los tránsitos.

**KIC 10577994:** Si se detecta el tránsito de período 15.313, pero resulta ser un falso positivo, calculando su periodograma en un intervalo que no tiene pérdida de datos evidente (desde el tiempo 401 hacia adelante), el periodograma no arroja esos máximos.

**KIC 10056119:** El periodograma resulta ser similar al de KIC 12401863. Vemos que el periodograma no detecta el tránsito en ninguno de los sistemas, salvo el falso positivo. A pesar que el primer sistema KIC 10723750 es similar a los del grupo I salvo mayor distancia entre los tránsitos (visibles a simple vista).

### 9.0.2. Period04

En esta subsección se detallan los resultados obtenidos por el software Period04, a continuación en la Fig.31 los periodogramas de las curvas de luz del grupo IV, cada uno contiene las rectas con los períodos reales:

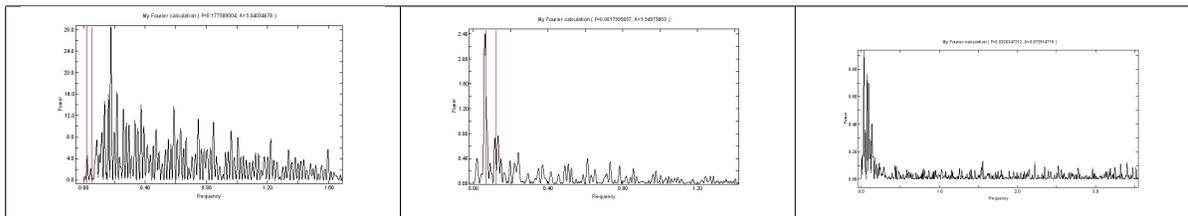


Figura 31: Periodogramas Period04 curvas de luz Grupo IV

**KIC 10723750:** No se ve evidencia de tránsito alguno, no hay ningún pick en los valores correspondientes a los períodos de los tránsitos.

**KIC 10577994:** Se aprecia casi evidencia de tránsito en el período de 15.313, Pero ese máximo no es el del exoplaneta pues el período confirmado está un poco corrido a la derecha (el peak evidenciado en el periodograma corresponde al período 0.061, es decir 16.4 días).

**KIC 10056119:** Ocurre lo mismo que el periodograma clasico, resulta ser similar al de KIC 12401863.

Nuevamente vemos que no logra detectar ningún tránsito, aun en el primero similar a los del grupo I, también notamos que el peak del sistema 2 está a 1 día de un tránsito, nos hace suponer que el propio brillo de la estrella tiene un período de ese largo.

### 9.0.3. Lomb-Scargle

A continuación, se verá los resultados entregados por el periodograma de Lomb-Scargle, en la fig.32 sus graficos, con una recta en el punto del período real:

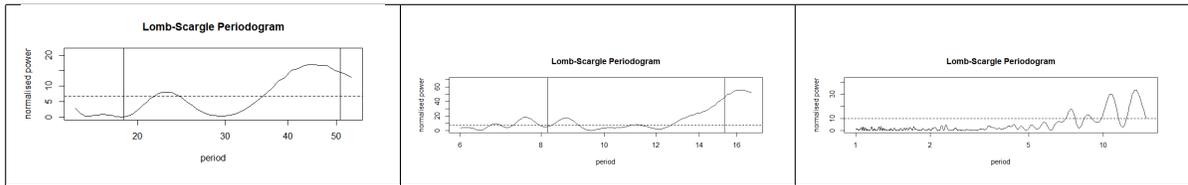


Figura 32: Periodogramas Lomb-Scargle curvas de luz Grupo IV

**KIC 10723750:** No se ve evidencia de tránsito alguno, no hay ningún peak en los valores correspondientes a los períodos de los tránsitos.

**KIC 10577994:** Acá queda más evidenciado lo visto en el periodograma anterior, muestra un peak en el punto 16.3 aprox. Confirmando aún más que la estrella tiene una variabilidad lumínica de esa forma.

**KIC 10056119:** Ocurre lo mismo que el periodograma clasico y period04, resulta ser similar al de KIC 12401863.

Desempeño muy similar a period04, tampoco detectando tránsito en el sistema similar a los del grupo I, y en el segundo sistema encontrando un período cercano al tránsito del exoplaneta.

### 9.0.4. BLS

Finalmente, los periodogramas BLS para este grupo, en la fig.33 se exponen el periodograma para cada sistema, con una recta en los puntos con un período real.

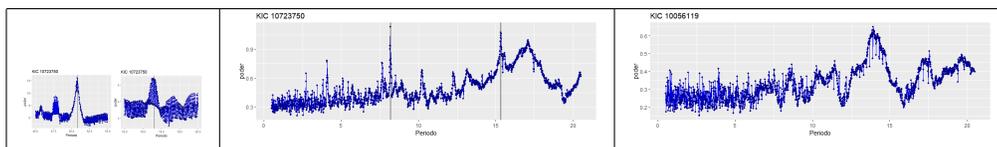


Figura 33: Periodogramas BLS curvas de luz Grupo IV

**KIC 10723750:** Se detecta el exoplaneta de tránsito 18.79, y se descubre el exoplaneta de tránsito 50.79, este ultimo es el de mas poder del periodograma (aca periodograma completo)VI.

**KIC 10577994:** Se detecta y descubre exoplaneta en ambos períodos, siendo los mas altos en cada caso.

**KIC 10056119:** Aquí ocurre algo singular, el periodograma se asemeja también al del KIC 12401863 (acá gráfico con la misma proporción[bla1-1-3-2]), pero este última evidencia el tránsito del exoplaneta confirmado. lo que nos daría a entrever que 100556119 no muestra evidencia de exoplaneta, por lo menos con esta única herramienta que tenemos.

Vemos que al contrario de los otros periodogramas, el desempeño del periodograma BLS fue muy bueno, no teniendo problemas en detectar ningún tránsito y pudiendo descubrir estos con este periodograma.

### 9.0.5. Conclusión:

Finalmente se expone la tabla 9 que contiene el resumen de los resultados para los tránsitos del grupo IV, resumiendo si detecta y descubre el tránsito correspondiente.

Tabla 9: Tabla correspondiente a desempeño de los periodogramas en el Grupo IV

Descubre/Detecta	10723750b	10723750c	10577994b	10577994c	
Clásico	No / No	No / No	No / No	No / No	
Period04	No / No	No / No	No / No	No / No	
L.Scargle	No / No	No / No	No / No	No / No	
BLS	<b>Posible / Si</b>	<b>Si / Si</b>	<b>Si / Si</b>	<b>Si / Si</b>	

Notamos en los 2 sistemas con exoplanetas confirmados de este grupo el muy buen desempeño dado por el periodograma BLS, detectan e incluso pudiendo descubrir todos los exoplanetas confirmados de este sistema. Por el contrario, para los periodogramas basados en DFT, el desempeño fue muy malo pues no detecto ningún tránsito oficial.

Más en detalle en el sistema KIC 10723750 en que los tránsitos son visibles a simple vistas, pero presentas más dificultades que los del grupo I, como por ejemplo, tener 2 tránsitos visibles, período propio de la estrella más corto que el de los tránsitos y por sobretodo una mayor distancia (período de más días) entre tránsitos, el desempeño de los basados en DFT fue pésimo, no logrando detectar ninguno. En cambio, el periodograma BLS no tuvo mayores problemas en detectar e incluso descubrir todos los tránsitos.

En el sistema KIC 10577994 sistema en que apenas son visible los tránsitos (a diferencia del sistema KIC 12401863 del grupo II donde no son visibles), nuevamente los sistemas basados en DFT no detectaron ningún exoplaneta. En cambio, BLS no tuvo problemas nuevamente en detectar y descubrir ambos tránsitos.

Como conclusión vemos que el periodograma de BLS no tiene problemas en detectar sistemas con 2 tránsitos presentando rendimientos similares (buenos) comparados con los sistemas anteriores con 1 solo tránsito.

De este grupo elegiremos el sistema KIC 10577994 para analizar en más detalle en el grupo V, pues el desempeño de los DFT fue pésimo, y por el contrario el BLS fue muy bueno. Para tener una referencia si aumentando la cantidad de tiempos de observación mejora y logra detectar los tránsitos (refiriéndose a los periodogramas basados en DFT).

De los 4 primeros grupos, este fue el grupo donde hubo mayor diferencia a favor del periodograma BLS.

## 10. Grupo V

En este último grupo se elegirán 3 sistemas de los anteriores, pero la diferencia estará en que se ocuparán todos los tiempos(intervalo) de observaciones para cada sistema, almacenados en el catálogo Kepler. Correspondiente a aproximadamente 1400 días de observación (teniendo un largo aproximado de 64792 puntos). Normalizando todos los tiempos de observación y trabajando con esta amplia curva de luz resultante. Entre cada tiempo de observación existen espacios de tiempo con datos faltantes, por lo que en este análisis obviaremos el periodograma clásico.

A continuación, en la Fig. 4, se exponen las curvas de luz elegidas:

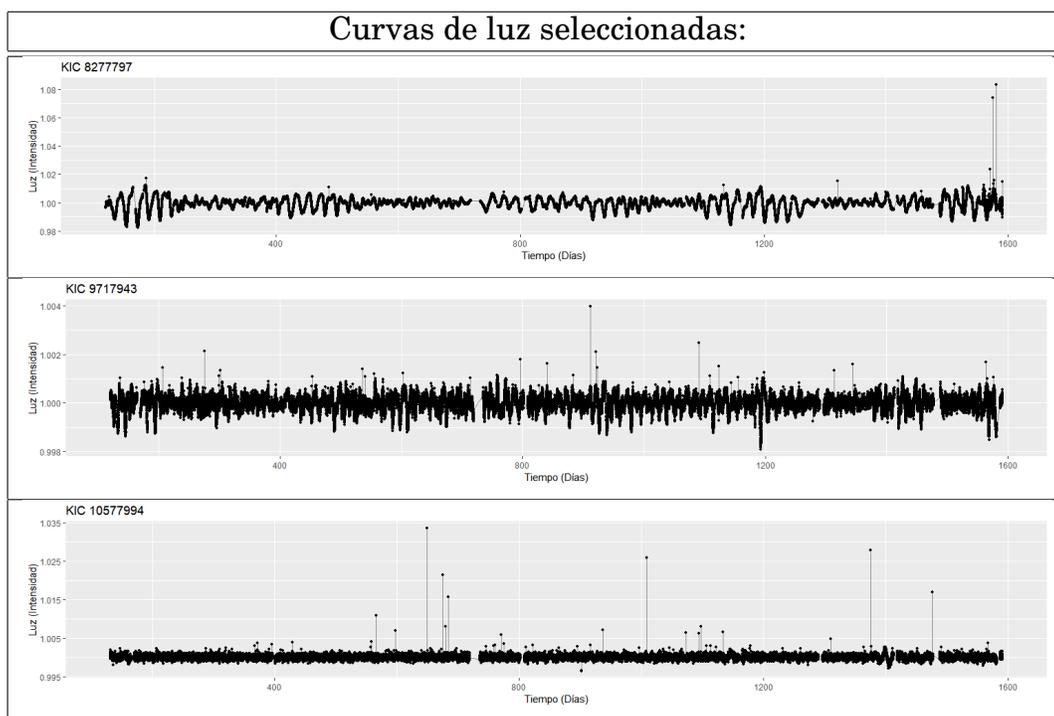


Figura 34: Curvas de luz seleccionadas Grupo V

Las propiedades de las curvas de luz son las siguientes:

**KIC 8277797:** Este sistema fue elegido porque, a pesar de tener poco ruido, no eran evidentes los tránsitos, los cambios de luz de la propia estrella son muy intensos y por sobre todo, el mal desempeño de los periodogramas (incluido el bls). Esta curva de luz tiene 2 exoplanetas confirmados, de períodos 4.33 y 1.65 días.

**KIC 9717943:** Este sistema, de manera similar al anterior, los periodogramas no obtuvieron un buen desempeño, la curva posee mucho ruido y la variabilidad de la estrella es muy pequeña (a diferencia de la curva anterior). Es una curva de luz con período confirmado de 6.11 días.

**KIC 10577994:** Finalmente este sistema fue seleccionado porque el periodograma BLS obtuvo un muy buen rendimiento comparado con los basados en DFT. Este sistema posee 2 exoplanetas confirmados de períodos 8.1808 y 15.313.

Por lo que la elección se basó en ver que tanto mejora (o no) el gran aumento de puntos en sistemas donde el desempeño de los periodogramas fue malo, como también, eligiendo el sistema en donde mayor diferencia de desempeños hay entre el periodograma BLS y los DFT.

### 10.0.1. Period04

A continuación, en la fi. se exponen los periodogramas entregados por period04, como también los obtenidos anteriormente para el mismo sistema, pero con menos tiempos de observación.

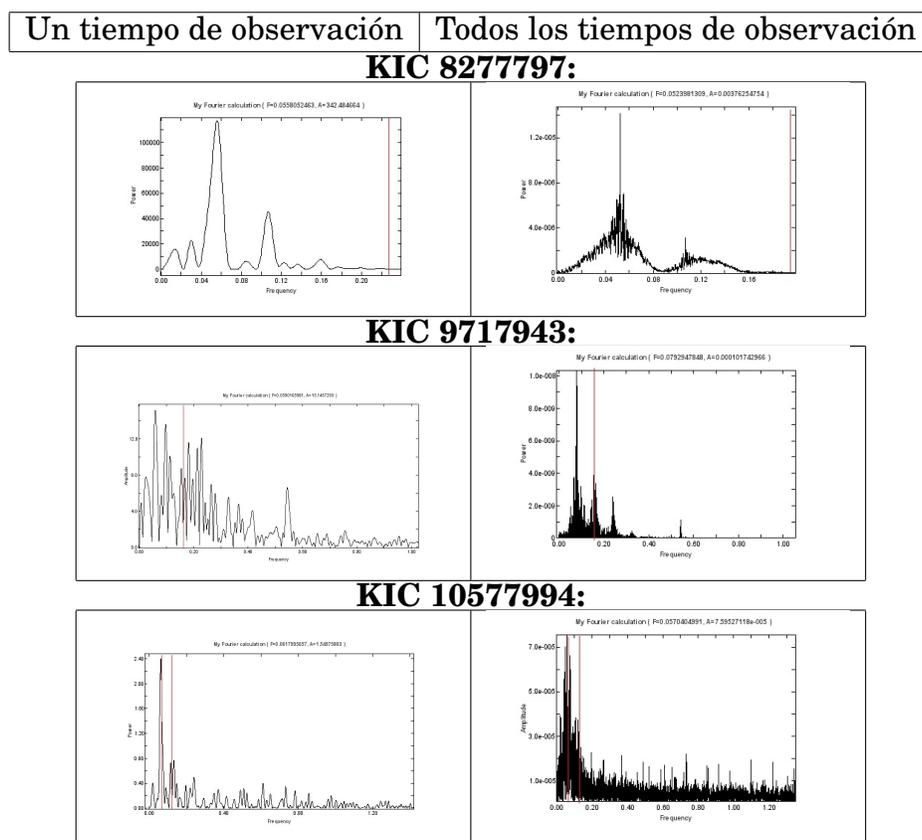


Figura 35: Periodogramas entregados por period04, a la izquierda los obtenidos anteriormente con un tiempo de observación, a la derecha con todos los tiempos de observación.

**KIC 8277797:** El periodograma se muestra mucho más detallado, pero aun así ninguno de los 2 tránsitos se ven reflejados, obteniendo el mismo desempeño que con un tiempo de observación.

**KIC 9717943:** En este sistema si denota una mejora, el exoplaneta es detectado, pero hay demasiada variabilidad y el peak del tránsito no es claro.

**KIC 10577994:** En este periodograma si son detectados ambos tránsitos, aunque más difíciles de poder descubrir, por la variabilidad del periodograma nuevamente, son solo detectados.

Notamos en 2 sistemas un mejor desempeño al usar más tiempos de observación, aunque por el contrario el periodograma entregado muestra muchas fluctuaciones, al tener mayor cantidad de puntos, se traduce directamente en mayor cantidad de elementos de la base del DFT y así generando un periodograma mucho más variable, que hacen más difícil de analizar.

### 10.0.2. Lomb-Scargle

A continuación, fig. , se exponen los periodogramas obtenidos por el método de lomb-scargle, como el periodograma anterior, se adjuntan los periodogramas obtenidos con un tiempo de observación.

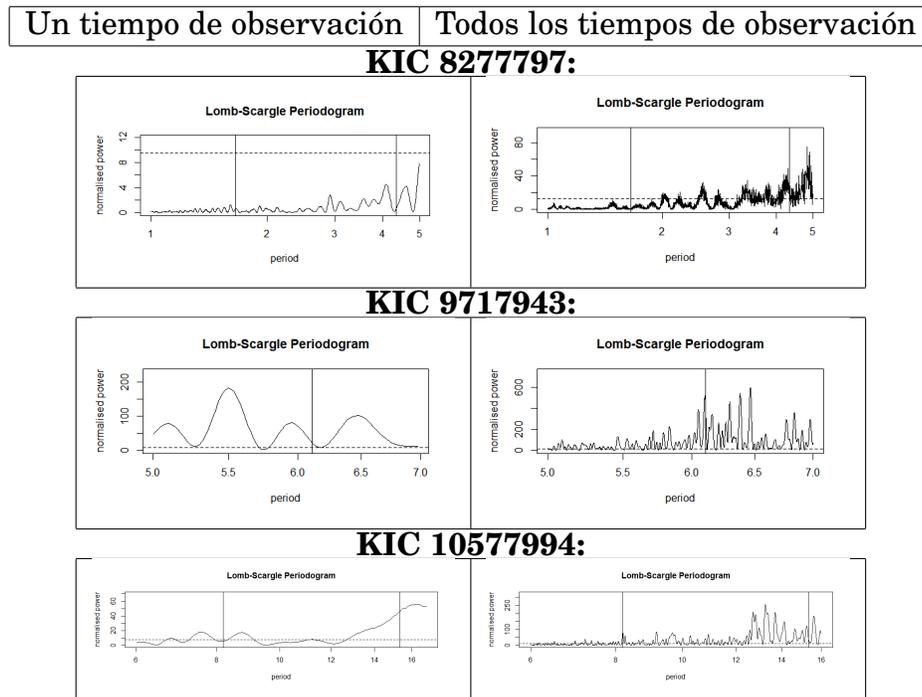


Figura 36: Periodogramas entregados por Lomb-Scargle, a la izquierda los obtenidos anteriormente con un tiempo de observación, a la derecha con todos los tiempos de observación.

**KIC 8277797:** No es posible detectar el tránsito de ninguno de los 2 exoplanetas nuevamente.

**KIC 9717943:** Este sistema, de manera similar a period04, el periodograma detecta el tránsito correspondiente, pero este se ve reflejado mucho más claro en el periodograma.

**KIC 10577994:** En este sistema si denota un mejor desempeño de este periodograma, detectando ambos tránsitos y a diferencia del period04, estos son mucho más evidentes al tener una menor variabilidad, incluso haciendo posible descubrir los tránsitos.

Nuevamente notamos que el desempeño mejoro en 2 de los sistemas analizados,

pero a diferencia del period04, el aumento de puntos no genera un periodograma con tantas fluctuaciones (variabilidad), que puedan dificultar su análisis.

### 10.0.3. BLS

Finalmente, en la fig. se exponen los periodogramas obtenidos por medio de BLS, se adjuntan los periodogramas obtenidos anteriormente con un tiempo de observación.

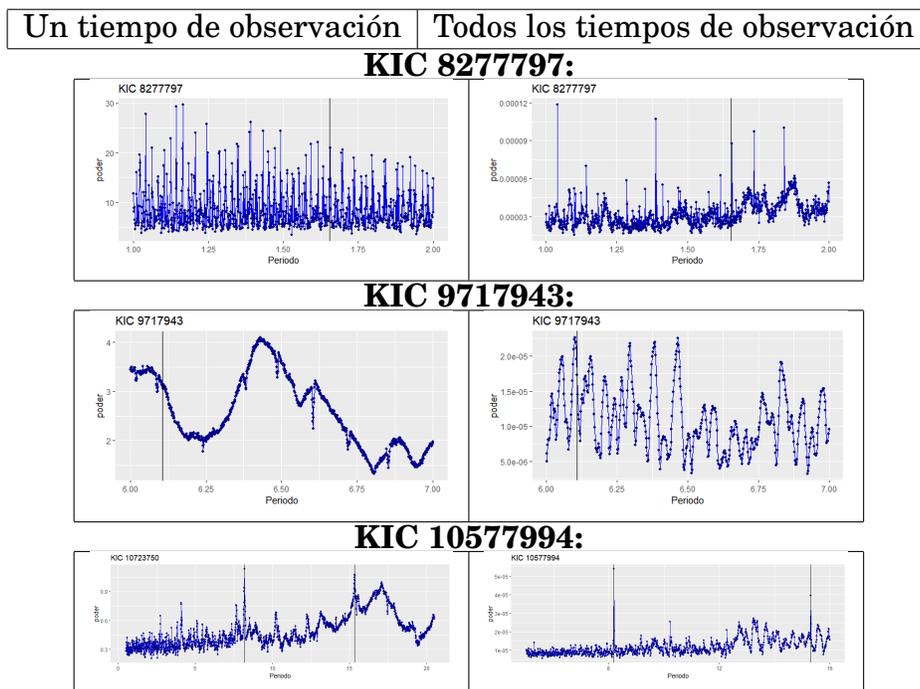


Figura 37: Periodogramas BLS, a la izquierda los obtenidos anteriormente con un tiempo de observación, a la derecha con todos los tiempos de observación.

**KIC 8277797:** En este sistema el tránsito detectado se muestra mucho más claro, en el periodograma con un tiempo de observación es detectado, pero con muchos otros tiempos donde también ocurre, en este caso el periodograma se muestra mucho más limpio, el otro tránsito no es detectado (fig, en anexo).

**KIC 9717943:** En este sistema al igual que los otros 2 periodogramas se logra detectar el tránsito del exoplaneta, a diferencia de con solo un tiempo de observación.

**KIC 10577994:** Este sistema con un tiempo de observación era posible detectar y descubrir los exoplanetas, al usar todos los tiempos, el periodograma se muestra mucho más limpio y los períodos donde existe el tránsito los peak son mucho más altos con respecto a los demás puntos.

Con BLS los 3 periodogramas muestran un mejor desempeño, detectando en el segundo el tránsito y obteniendo los períodos de tránsito mucho más explícitos en los diferentes periodogramas.

#### 10.0.4. Conclusión:

Podemos concluir que en todos los casos el aumento de puntos mejora el desempeño de los los periodogramas, teniendo en consideración la gran diferencia entre la cantidad de puntos no es un aumento de desempeño considerable, más de 15 veces mas puntos.

Viendo más en particular las curvas de luz seleccionadas, vemos que la primera elegida no se logra detectar el segundo exoplaneta con ningún periodograma y no es detectado ninguno con los basados en DFT. El periodograma de BLS mejora, exponiendo el período del tránsito en forma mucho más clara.

La respuesta a la poca mejora de desempeño en este sistema, podría estar en ver los diferentes tiempos de observación, son muy diferentes entre sí a pesar de la normalización, probablemente se deba hacer un análisis más detallado de los diferentes tiempos de observación, para luego poder usar los periodogramas en estos.

Con el segundo sistema y a diferencia del primero, los 3 periodogramas lograron detectar el período correspondiente al tránsito, por lo que, haciendo una relación con el primero, parece ser que los periodogramas mejoran su desempeño al aumentar sus puntos, si el problema es el ruido (a diferencia del primer periodograma).

En el tercer sistema vemos que el aumento de puntos logro hacer visible los períodos en los basados en DFT, esta curva de luz los tránsitos no eran distinguibles a simple vista. Pero aumentando la cantidad de puntos si los hisos visibles, en el caso del periodograma BLS la evidencia de los tránsitos es absoluta, mejorando aún más la evidencia de los períodos correspondientes.

Comparado los diferentes periodogramas, vemos que el aumento de puntos genera una mayor variabilidad y fluctuación en los periodogramas basados en DFT, pudiendo dificultar la observación y detección de tránsitos, esto es mucho más notorio en el entregado por period04, en el caso de Lomb-scargle esto es mucho menos acentuado no presentando casi problemas en los sistemas estudiados. Por otra parte, el periodograma de BLS, mejora en todos los casos y se nota superior a los DFT igualmente, pues incluso en el caso del tercer sistema (KIC 10577994) los períodos son más evidentes en BLS con un tiempo de observación, que los DFT con todos los tiempos de observación.

Finalmente, se hizo una prueba de tiempo de ejecución de los algoritmos, tanto de Lomb-Scargle como de BLS, a modo de referencia, en un sistema con un tiempo de observación vs con todos (KIC 9717943), obteniendo los siguientes resultados en la tabla. Media de 4 ejecuciones cada uno.

Puntos	L-S	BLS
4108	0.82 seg	54.85 seg
65262	227.34 seg	83.47 seg

El algoritmo BLS, al tener métodos como encajonar (bind) y plegar la curva de luz, la gran cantidad de puntos extra aumento el tiempo de ejecución en forma marginal, menos del doble. El aumento de tiempo de ejecución aumento menos del doble (considerando la gran cantidad de puntos extra) pues el mayor tiempo empleado en el algoritmo es,

Por el contrario, el algoritmo de lomb-scargle el tiempo de ejecución aumento considerablemente, en casi en 3 órdenes de magnitud.

Esto se debe tomar como referencia y da evidencias de que a gran cantidad de puntos el periodograma de BLS es más eficiente que el de lomb-scargle. Pero no da una respuesta a que el periodograma BLS es más rápido de ejecutar que el lomb-scargle, pues se requeriría un análisis mucho más profundo, tomando en consideración períodos a analizar, resolución diferentes curvas de luz y también que tan bien sea implementado el algoritmo en el software. Como única conclusión clara, tenemos que el algoritmo del periodograma de BLS es muchísimo más eficiente al aumentar la cantidad de puntos de la curva de luz, que el periodograma de Lomb-scargle.

## 11. Conclusiones generales:

En resumen, tomaremos las principales conclusiones obtenidas de los grupos analizados anteriores:

-Periodogramas Clásico, Period04 y Lomb-Scargle se comportaron de manera similar, evidenciando el periodograma clásico su debilidad en cuando a curvas con pérdida de datos como por ejemplo en KIC 11295426.

-Lomb-Scargle se mostró sensiblemente superior a period04 y Clásico, viéndose más sencillo a la hora de analizar comparado con period04, notándolo aún más cuando se analizó las curvas con muchos tiempos de observación.

-Periodograma Clásico no es alternativa por sus limitaciones en este tipo de series de tiempo, a pesar de que las curvas son muy estables en cuando a equidistancia, todas poseen en mayor o menor medida intervalos de datos perdidos.

-BLS mostro un desempeño superior al resto, en absolutamente todos los casos, muchas veces detectando en forma notoria tránsitos y DFT no detectando nada.

-BLS requiere más tiempo computacional, pero esto es solo en curvas de luz de relativamente cortas, si se usan curvas de luz cada vez más largas, el cálculo del periodograma aumenta en mucha menor medida que los basados en DFT (específicamente en Lomb-Scargle), como vimos en el ejemplo del sistema KIC 9717943, donde al aumentar 16 veces la cantidad de puntos el algoritmo BLS demoro 1.5 veces más mientras que el Lomb-Scargle demoro 277 veces más.

-Al aumentar la cantidad de puntos (aproximadamente 16 veces), todos los periodogramas mejoraron su desempeño, en el caso de los DFT logrando detectar tránsitos que antes no detectaban al igual que el BLS, y este último mejorando y evidenciando un periodograma mucho más claro de los tránsitos. A pesar de todo no es muy diferente, para la gran cantidad de puntos extra.

-El algoritmo BLS es mucho más flexible, pudiendo elegir manualmente resolución y en que períodos se realiza. En cambio, los DFT, siempre se debe realizar el proceso completo.

-En curvas de luz simples donde es evidente el tránsito planetario, y además, la distancia entre tránsitos es corta como también la profundidad de estos es alta (como el caso de los periodogramas del grupo I), los 4 periodogramas mostraron un buen desempeño.

-Cuando las curvas de luz presentan variabilidad propia de la estrella, afectan negativamente mucho más a los basados en DFT, pues todo el poder del periodograma se lo llevan estos períodos, haciendo más complicado el análisis de detectar los supuestos tránsitos.

-Cuando las curvas de luz no tienen un evidente tránsito y la luminosidad de la estrella es muy constante (KIC 12401863, KIC 10577994), los basados en DFT muestran un pésimo desempeño, aun cuando se use una mayor cantidad de puntos, logran apenas poder detectar los tránsitos. El periodograma BLS muestra un excelente desempeño en estos casos.

-Cuando el tiempo de tránsito es corto (del período total) o también el tránsito es de muchos días (KIC 10723750), los periodogramas DFT son afectados negativamente. No así BLS.

- Cuando la variabilidad de la estrella tiene un período inferior al período del tránsito, afecta muy negativamente a los basados en DFT.
- Todos los periodogramas bajan su desempeño en curvas de luz que tengan variabilidad de la estrella y además el tránsito sea no visible a simple vista.

## Parte IV

# Conclusiones finales:

Ya visto el resultado de los periodogramas, como conclusión final y mas importante diría que, el periodograma BLS es el que mejor rendimiento tubo. en ninguna detección se mostro inferior a alguno de la familia de DFT. En curvas de luz, donde hay mucha distancia entre tránsitos (tránsitos de pocos puntos de amplitud) o donde es muy poca profundidad vs el ruido gaussiano de la estrella, se nota la mayor diferencia a favor del BLS.

cuando el tránsito es evidente, los 4 periodogramas mostraron un buen rendimiento, y entre los DFT, el que mejor se comporto sensiblemente mejor fue el de lomb scargle.

### 11.0.1. Trabajo futuro:

De este trabajo considero que existen 3 caminos para complementar:

#### **Técnicas para la confirmación de que una detección es un tránsito:**

En este trabajo prácticamente no se usaron técnicas de confirmación de tránsito, simplemente se contrasto ya sabiendo el período del exoplaneta confirmado y que tan evidente se veía el período de este en el periodograma. Si se usaran técnicas de confirmación se podría desarrollar más esta área.

#### **Inclusión de más tipos de periodograma:**

En este trabajo se usaron 4 periodogramas, se podría expandir a 2 periodogramas más que me parecen interesantes:

Plavchan: este periodograma se basa en el “Phase dispersion minimization” que consiste en, de manera similar al método BLS plegar la serie de tiempo para un período P, pero en vez de buscar esa función escalonada, ve la desviación estándar de cada intervalo (caja en el caso de BLS). Este periodograma se descartó en esta instancia por ofrecer resultados muy similares al periodograma BLS, pero podría ahondarse en encontrar las diferencias y similitudes con este en diferentes instancias.

#### **Búsqueda de curvas de luz con probables exoplanetas:**

Finalmente, y dado la gran cantidad de sistemas en los que se ha hecho observaciones, por ejemplo, el catalogo del observatorio Kepler (KIC) contiene alrededor de 13.2 millones de observaciones de objetos y sistemas. Se podría desarrollar una aplicación, ya sea regresión lineal, logística o redes neuronales. En la que se clasifique cada sistema como más o menos probable de la existencia de exoplanetas, para dejar los mejores candidatos para una inspección más detallada de la existencia o no de tránsitos.

Por ejemplo, una red neuronal usando las curvas de luz de exoplanetas confirmados, extraer de estas curvas información y usarlas de variables; como seria desviación estándar, cantidad de puntos que están sobre el promedio vs los que están debajo del promedio, relación entre mediana y media, y variables más específicas, como por ejemplo cantidad de puntos que están en el 10 por ciento superior vs 10 por

ciento inferior, si el sistema es del tipo variable usando el periodograma de Lomb-Scargle , y muchas otras. Pues una característica del tránsito es que desciende la luminosidad de la estrella en un intervalo muy pequeño, y la idea es particularizar esta característica de la forma más simple posible.

Entrenando la red neuronal con todos los datos de las curvas de luz con tránsito confirmado y una buena cantidad de sistemas donde este confirmado la no detección de tránsitos, se crearía esta aplicación y usarlas en sistemas no clasificados, para luego hacer un ranking con los más probables de tránsito para una inspección mas detallada.

## Parte V

# Bibliografía

- William W.S. Wei (2006) Time Series Analysis: Univariate and Multivariate Methods. Pearson.
- Jacob T. VanderPlas (2017) Understanding the Lomb-Scargle Periodogram. arXiv: 1703.09824 [astro-ph.IM].
- Instituto Nacional de Estadística, España (2012) Análisis espectral y ajuste estacional. <https://www.ine.es/clasifi/analisisyajuste.pdf>
- Robert H. Shumway y David S. Stoffer (2016) Time Series Analysis and Applications. Springer.
- Eric D. Feigelson, G. Jogesh Babu (2012) Modern Statistical Methods for Astronomy With R Applications. Cambridge.
- G. Kovács (2002) A box-fitting algorithm in the search for periodic transits. arXiv:astro-ph/0206099 [astro-ph.IM].
- Weisstein, Eric W. "Nyquist Frequency." From MathWorld--A Wolfram Web Resource. <http://mathworld.wolfram.com/NyquistFrequency.html>
- [https://exoplanetarchive.ipac.caltech.edu/cgi-bin/nstEDAPI/nph-nstEDAPI?table=q1\\_q17\\_dr](https://exoplanetarchive.ipac.caltech.edu/cgi-bin/nstEDAPI/nph-nstEDAPI?table=q1_q17_dr)

## Parte VI

# Anexo A: Gráficos

BLS1:

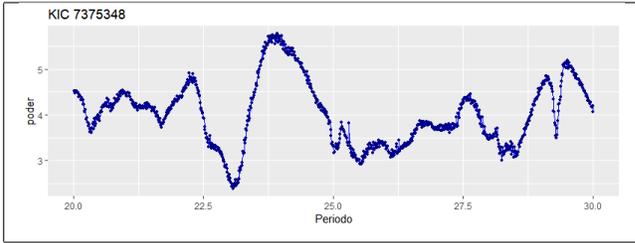


Figura 38: Menor resolucion

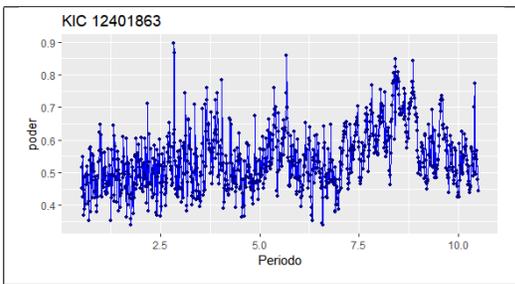


Figura 39: Menor resolucion

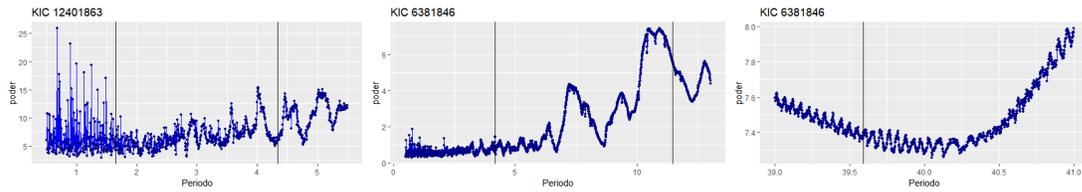


Figura 40: Resto de períodos

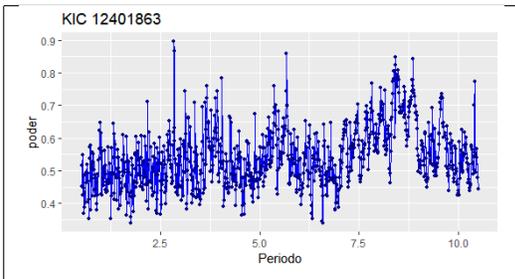
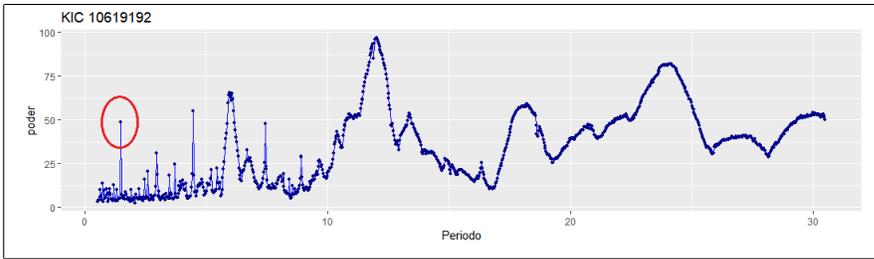
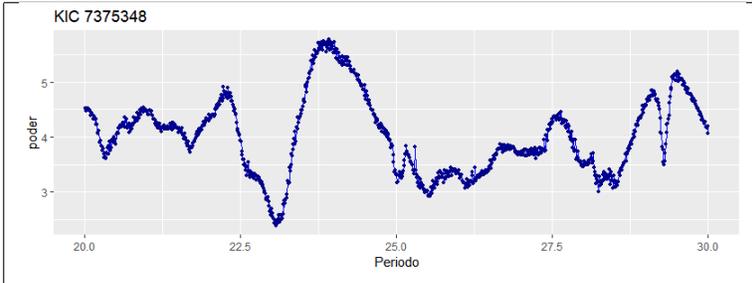


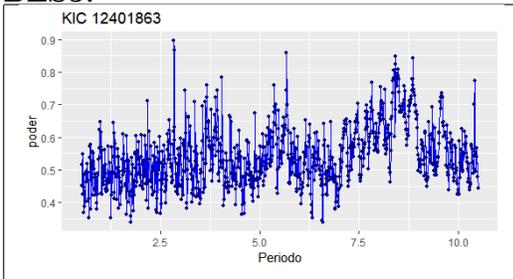
Figura 41: asdasa



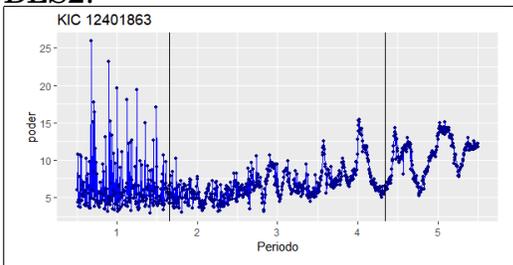
**BLS2:**



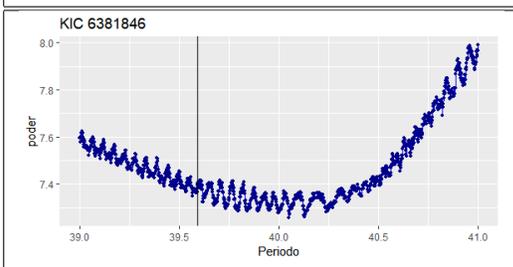
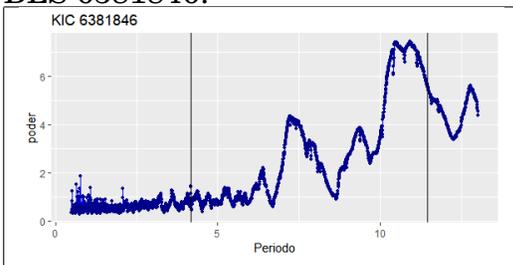
**BLS3:**



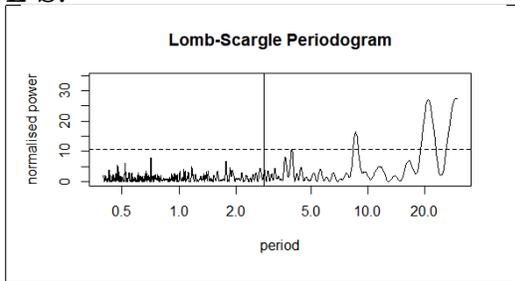
**BLS2:**



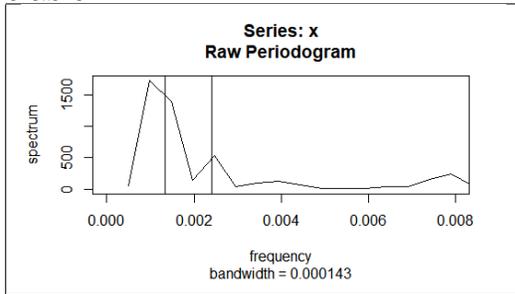
**BLS 6381846:**



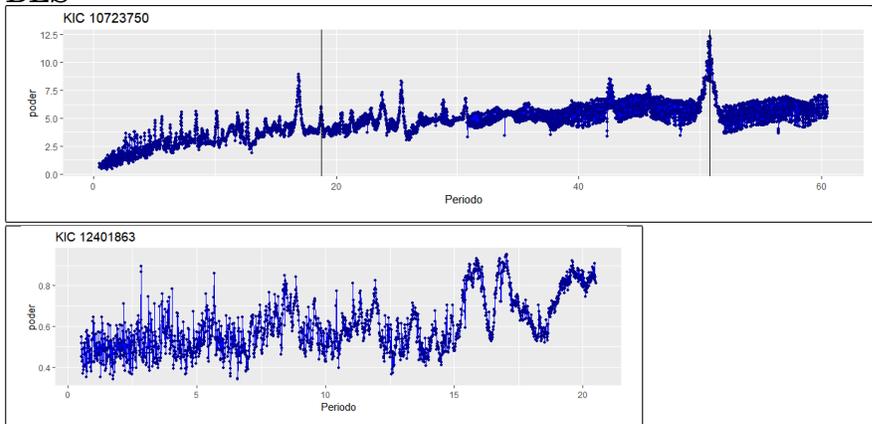
L-S:



classic



BLS



## Parte VII

# Anexo B: Codigos R:

Centra la serie en 0 (la luminosidad)

```
CentraSerie=function(x,y){  
  x[y]=x[y]-mean(data.matrix(x[y]), na.rm=TRUE)  
  return(x)  
}
```

Pliega la serie con respecto al tiempo en período P, conservando tiempo inicial

```
PeriodoTiempo=function(x,p){  
  z=min(x[2])  
  x[2]=x[2]-z  
  x[2]=x[2]%%p  
  x[2]=x[2]+z  
  return(x)  
}
```

### Pliega la serie con respecto al tiempo en período P

```
PeriodoTiempo2=function(x,p){
  x[2]=x[2]%% p
  return(x)
}
```

### Ordena la serie de menor a mayor, con respecto al tiempo ya plegado

```
OrdenaS=function(x){
  x=x[order(x$V2),]
  rownames(x)=NULL
  return(x)
}
```

### Empaqueta en m cantidad de bind, de tamaño largo/m la serie:

```
EmpaquetaBind=function(x,m){
  n=floor(nrow(x)/m)+1 l=seq(1:m)
  z=rep(1, each = n) x[1]=z[1:nrow(x)]
  x = aggregate(x,by = list(x$V1),FUN = mean)
  x = subset(x, select = -1 )
  return(x)
}
```

### Algoritmo basico del BLS para (i1,i2)

```
BloqueBLS=function(x,i1,i2){
  y=i2-i1+1
  z=sum(x[i1:i2,3])
  z=z^2 z=z/(y*(nrow(x)-y))
  z=z^(1/2)
  return(z)
}
```

### Primera iteracion,largo n, devuelve el maximo para ese largo del intervalo

```
IterMLargo=function(x,n){
  z=BloqueBLS(x,1,1+n-1)
  b=nrow(x)-n
  for (i in 2:b){
    z[i]=BloqueBLS(x,i,i+n-1)
  }
  t=max(z)
  return(t)
}
```

### Segunda iteracion, busca cada largo hasta el largo n, devuelve el maximo

```
IterMLargo2=function(x,n){
  XAXA=1
  for (i in 1:n){
    XAXA[i]=IterMLargo(x,i)
  }
  t=max(XAXA)
  return(t)
}
```

### Calcula periodograma para el valor P

```
BLSperiodP=function(x,p,n=300,m){
  x=PeriodoTiempo2(x,p)
  x=OrdenaS(x)
  x=EmpaquetaBind(x,n)
  b=IterMLargo2(x,m)
  return(b)
}
```

Calcula periodograma primero colocar data, segundo **período** p donde parte buscando, tercero tamaño de los pasos de búsqueda, 4 total de pasos a dar, 5 numero de binds, n ancho intervalo

```
BLSperiodograma=function(x,pa=0.5,pasi=0.01,tot=1000,n=300,m){
  poder=1
  período=1
  x=CentraSerie(x,3)
  for (i in 0:tot){
    poder[i]=BLSperiodP(x,pa+i*pasi,n,m)
    período[i]=pa+i*pasi
  }
  poder=cbind(período,poder)
  return(poder)
}
```