

Astronomia observacional: Pràctica periodicitats

Alumnes

- Ruiz Dern, Laura
- Peña Pérez, Dídac
- Subirà Vallés, Montserrat
- Blanco Cuaresma, Sergi

Índex

1	Objectiu	3
2	Transformada discreta de Fourier	3
2.1	Descripció	3
2.2	Funció finestra	5
3	Resultats	6
3.1	Dades inicials	6
3.2	Mètodes de càlcul	6
3.2.1	Introducció	6
3.2.2	Fast Fourier Transform (FFT)	6
3.2.3	Lomb-Scargle periodogram	7
3.2.4	Transformada discreta de Fourier	8
3.2.5	Algorisme Clean	9
3.3	Punts en fase	11
3.4	Mínims quadrats	12
3.5	Eficiència computacional	13

1 Objectiu

L'objectiu de la pràctica proposada és determinar si és periòdic un senyal de raigs X obtinguda pel satèl·lit de la NASA Rossi X-Ray Timing Explorer. En aquest document es presenten els resultats de l'anàlisi efectuat.

Els càlculs per portar a terme la transformada de Fourier, algorisme Clean, mínim quadrats i l'anàlisi en general s'han portat a terme mitjançant un programa desenvolupat en C++, pel qual s'ha redactat un manual independent. Addicionalment als requeriments de la pràctica, s'ha avaluat l'aplicació de l'algorisme Fast Fourier Transform (FFT) i el càlcul de l'espectre net amb el periodogram Lomb-Scargle per tal d'analitzar la seva validesa i comparar el consum de recursos computacionals.

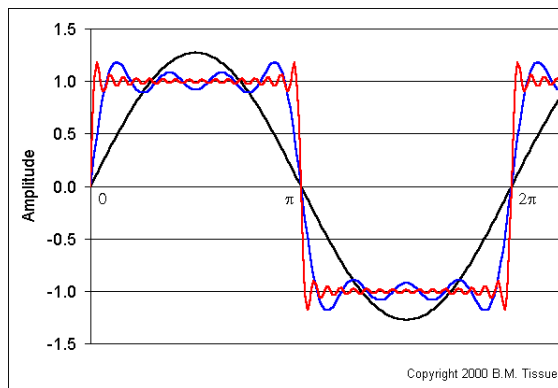
2 Transformada discreta de Fourier

2.1 Descripció

Mitjançant Fourier qualsevol funció en temps (t) pot ser representada per la suma de funcions sinusoidals en freqüència (k).

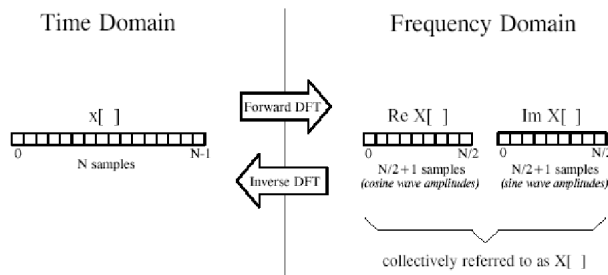
$$x(t) = \sum_{k=1}^{\infty} a_k \cos(2\pi f_k t) + \sum_{k=1}^{\infty} b_k \sin(2\pi f_k t)$$

A mode d'exemple, una funció amb dos únics valors (1 i -1) pot ser representada per la suma de diferents funcions sinusoidals com es pot apreciar a la següent imatge:



Les funcions sinusoidals es troben associades a freqüències i, en cas de que la funció temporal sigui periòdica, hi haurà una en particular que disposarà d'una contribució especialment significativa.

A partir d' N mostres ($x(n=0) \dots x(n=N-1)$) mesurades en intervals de temps variables i utilitzant la transformada de Fourier, es possible determinar la contribució de cada freqüència a la funció temporal original.



La transformada de Fourier es pot aplicar mitjançant la fórmula:

$$X(f_k) = \frac{1}{N} \sum_{n=0}^{N-1} w(n) \cdot x(n) \cos(2\pi f_k t_n) + \frac{1}{N} \sum_{n=0}^{N-1} w(n) \cdot x(n) \sin(2\pi f_k t_n) \cdot i$$

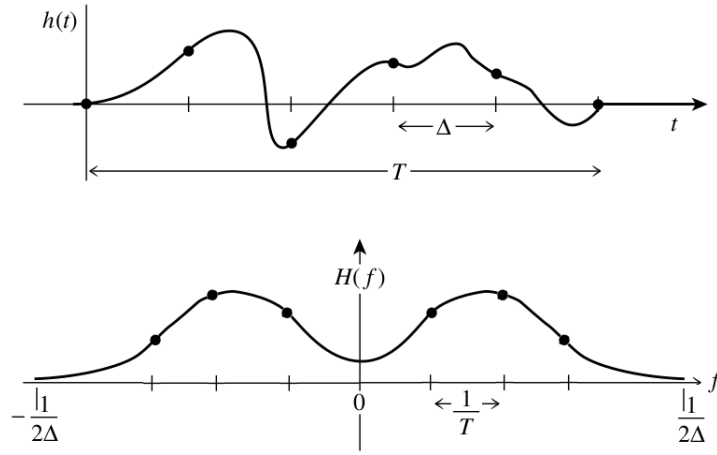
A on $X(f_k)$ és un número complex, $w(n)$ és la funció finestra, $x(n)$ el valor mesurat del senyal original, t el moment temporal al qual s'ha fet la mesura i f_k una freqüència determinada. La fórmula també pot ser expressada aplicant la identitat $e^{ix} = \cos(x) + i \sin(x)$:

$$X(f_k) = \frac{1}{N} \sum_{n=0}^{N-1} w(n) \cdot x(n) e^{-i2\pi f_k t_n}$$

Les freqüències f_k es trobaran limitades per la freqüència crítica Nyquist (amb increments de com a mínim $f_m = \frac{1}{T}$), que depèn de l'interval de mostreig:

$$f_c \equiv \frac{1}{2\Delta} \Rightarrow |f_k| \leq f_c$$

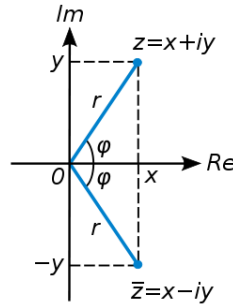
A on T és la diferència de temps entre la última mostra i la primera i Δ és la separació mínima de temps entre 2 mesures.



Cal destacar que habitualment les mesures en l'espai temporal corresponen a valors reals (no complexos) i això fa que al domini de les freqüències la transformada sigui simètrica (els valors de la zona negativa corresponen al conjugat dels valors de freqüència positiva, $X(k) = X^*(N - k)$ o \bar{z} a la imatge següent).

Donat que els valors $X(f_k)$ són números complexos, per representar la funció en una gràfica de dos dimensions es possible calcular la magnitud (mòdul) i/o la fase:

$$z = x + i \cdot y = X(k) = \text{Re}X(k) + i \cdot \text{Im}X(k) = r e^{i\psi}$$



$$Magnitud(f_k) = \sqrt{Real[X(f_k)]^2 + Imag[X(f_k)]^2}$$

$$Fase X[f_k] = \arctan\left(\frac{Imag[X(f_k)]}{Real[X(f_k)]}\right)$$

No obstant, per tal de representar la densitat espectral es recomana l'estimador denominat periodogram clàssic:

$$Densitat(f_k) = Real[X(f_k)]^2 + Imag[X(f_k)]^2$$

2.2 Funció finestra

Per realitzar el càlcul de la transformada discreta de Fourier, s'aplica la fórmula presentada a la secció anterior:

$$X(f_k) = \frac{1}{N} \sum_{n=0}^{N-1} w(n) \cdot x(n) \cos(2\pi f_k t_n) + \frac{1}{N} \sum_{n=0}^{N-1} w(n) \cdot x(n) \sin(2\pi f_k t_n) \cdot i$$

La finestra es troba representada per $w(n)$ i habitualment s'utilitza la funció quadrada, es a dir, presenta valor constant 1 per tot el rang al qual tenim valors mesurats. No obstant, existeixen altres funcions alternatives que intenten evitar el canvi brusc de 0 a 1 de la funció quadrada:

- Quadrat

$$w_{square}(n) = \begin{cases} 1 & \text{per } 0 \leq n < N \\ 0 & \text{resta} \end{cases}$$

- Bartlett

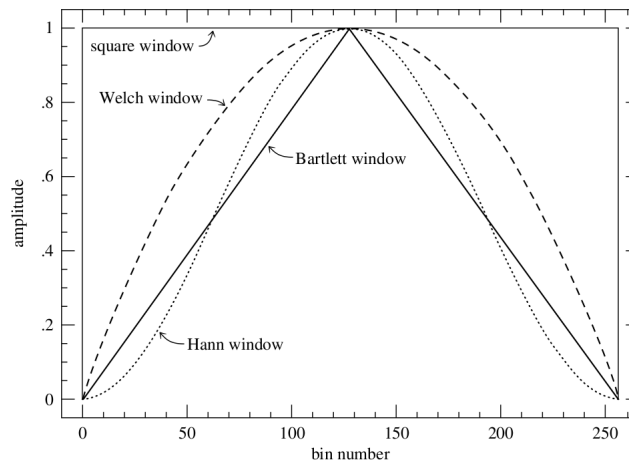
$$w_{bartlett}(n) = 1 - \left| \frac{2 \cdot n}{n_{freq} - 1} - 1 \right|$$

- Welch

$$w_{welch}(n) = 1 - \sqrt{\frac{2 \cdot n}{n_{freq} - 1} - 1}$$

- Hann

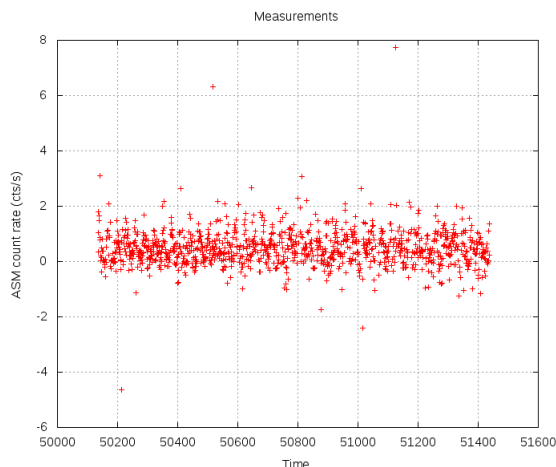
$$w_{hann}(n) = 0.5 \cdot \left(1 - \cos\left(\frac{n \cdot 2\pi}{n_{freq} - 1}\right) \right)$$



3 Resultats

3.1 Dades inicials

Les dades d'entrada es troben composades per tres columnes: temps en dies julians, comptes per segon detectats i error. No és possible identificar un senyal periòdic simplement amb la representació visual de les dades (comptes per dia Julià):



3.2 Mètodes de càlcul

3.2.1 Introducció

Per tal d'obtenir l'espectre de freqüències es poden portar a terme diverses estratègies de càlcul amb diferents graus de complexitat i requeriments computacionals.

En aquesta secció es presenten els resultats dels mètodes de càlcul utilitzats pel programa desenvolupat en C++:

- Fast Fourier Transform (FFT): Algorisme altament eficient per al càlcul de la transformada de Fourier i generació de l'espectre brut.
- Lomb-Scargle periodogram: Mètode alternatiu per a la generació d'un espectre net.
- Transformada discreta de Fourier: Generació de l'espectre brut i la finestra segons és requerida a l'enunciat de la pràctica.
- Algorisme Clean: Generació de l'espectre net, tasca opcional segons l'enunciat de la pràctica.

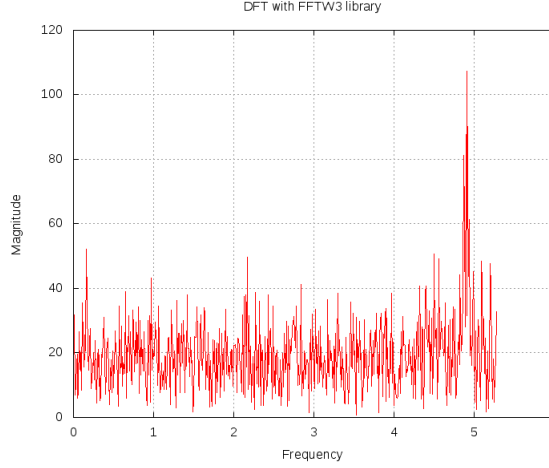
Cal destacar que els dos primers mètodes no són un requisit de l'enunciat de la pràctica i han sigut inclosos per complementar el treball realitzat.

3.2.2 Fast Fourier Transform (FFT)

Per computar la transformada de Fourier existeix un algorisme denominat Fast Fourier Transform (FFT). La versió més popular la van definir Danielson i Lanczos al 1942, quan van demostrar que una transformada de Fourier discreta de longitud N pot ser reescrita com la suma de dos transformades discretes, cadascuna de longitud $N/2$.

Aquest mètode pot ser utilitzat recursivament, reduint el temps de computació de $O(N^2)$ a $O(N \log_2 N)$. No obstant, es troba limitat estrictament a l'aplicació sobre mesures que es troben equidistants en el temps i que són potencia de dos. Aquesta darrera condició pot ser obviada si s'utilitzen algunes variacions de l'algorisme que permeten treballar amb un nombre de mesures que no siguin potencia de dos i que mantindrà la bona eficiència a no ser que es tracti d'un nombre primer.

La llibreria FFTW3 selecciona automàticament la variació de l'algorisme FFT més òptim, facilitant el tractament de dades que no són potència de dos. El programa desenvolupat en C++ fa ús d'aquesta llibreria per tal de comparar els resultats i l'eficiència computacional amb la resta de mètodes. A continuació es presenta l'espectre resultant:



Les dades d'entrada utilitzades han sigut enregistrades en intervals de temps fortament variables, per tant no es compleix un dels requeriments de l'algorisme FFT. Aquest fet provoca que els resultats no siguin d'utilitat en l'anàlisi del senyal i s'hauran d'aplicar altres mètodes.

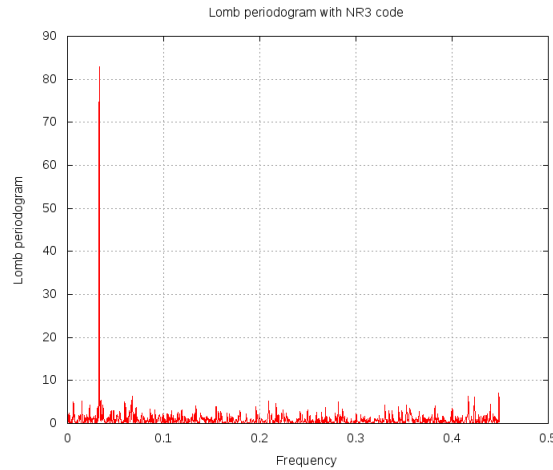
3.2.3 Lomb-Scargle periodogram

Un mètode popular per treballar amb dades que no es troben equidistants en el temps és el periodogram de Lomb-Scargle:

$$d_{LS}(f_k) = \frac{1}{2} \left(\frac{\left[\sum_{n=0}^{N-1} x(n) \cdot \cos(f_k \cdot (x(n) - \tau)) \right]^2}{\sum_{n=0}^{N-1} x(n) \cdot \cos^2(f_k \cdot (x(n) - \tau))} + \frac{\left[\sum_{n=0}^{N-1} x(n) \cdot \sin(f_k \cdot (x(n) - \tau)) \right]^2}{\sum_{n=0}^{N-1} x(n) \cdot \sin^2(f_k \cdot (x(n) - \tau))} \right)$$

A on f_k és la freqüència, $x(n)$ les mesures i $\tan(2f_k\tau) = \frac{\sum_{n=0}^{N-1} \sin(2 \cdot f_k \cdot x(n))}{\sum_{n=0}^{N-1} \cos(2 \cdot f_k \cdot x(n))}$

Computacionalment és eficient i permet construir espectres de densitat sense els efectes que provoquen les funcions finestres al càlcul de la transformada discreta de Fourier. Al cas particular en estudi, s'ha utilitzat la implementació del llibre Numerical Recipes per tal d'obtenir la següent representació:

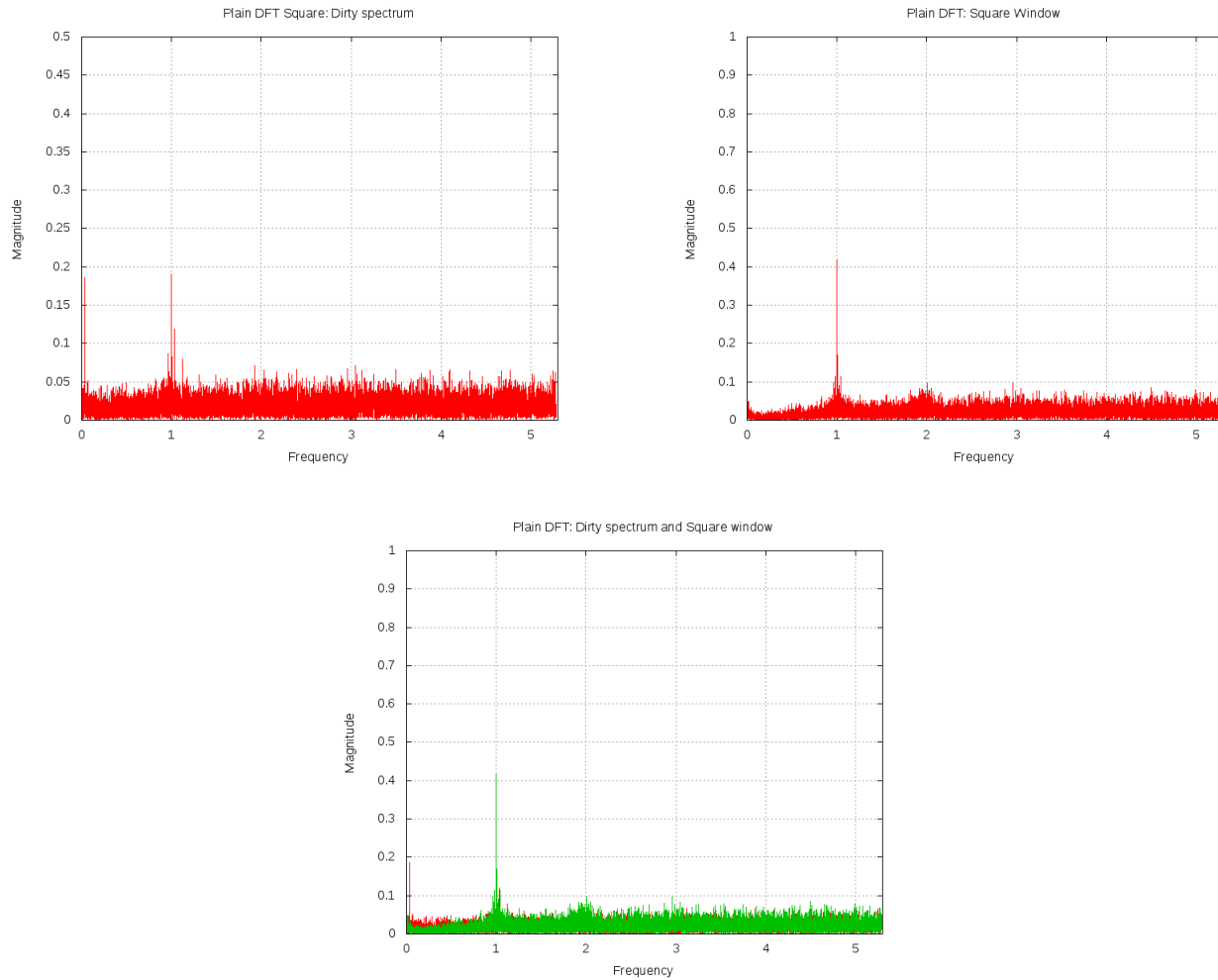


Segons aquesta tècnica, la freqüència amb una major contribució correspon a $0,033Hz$.

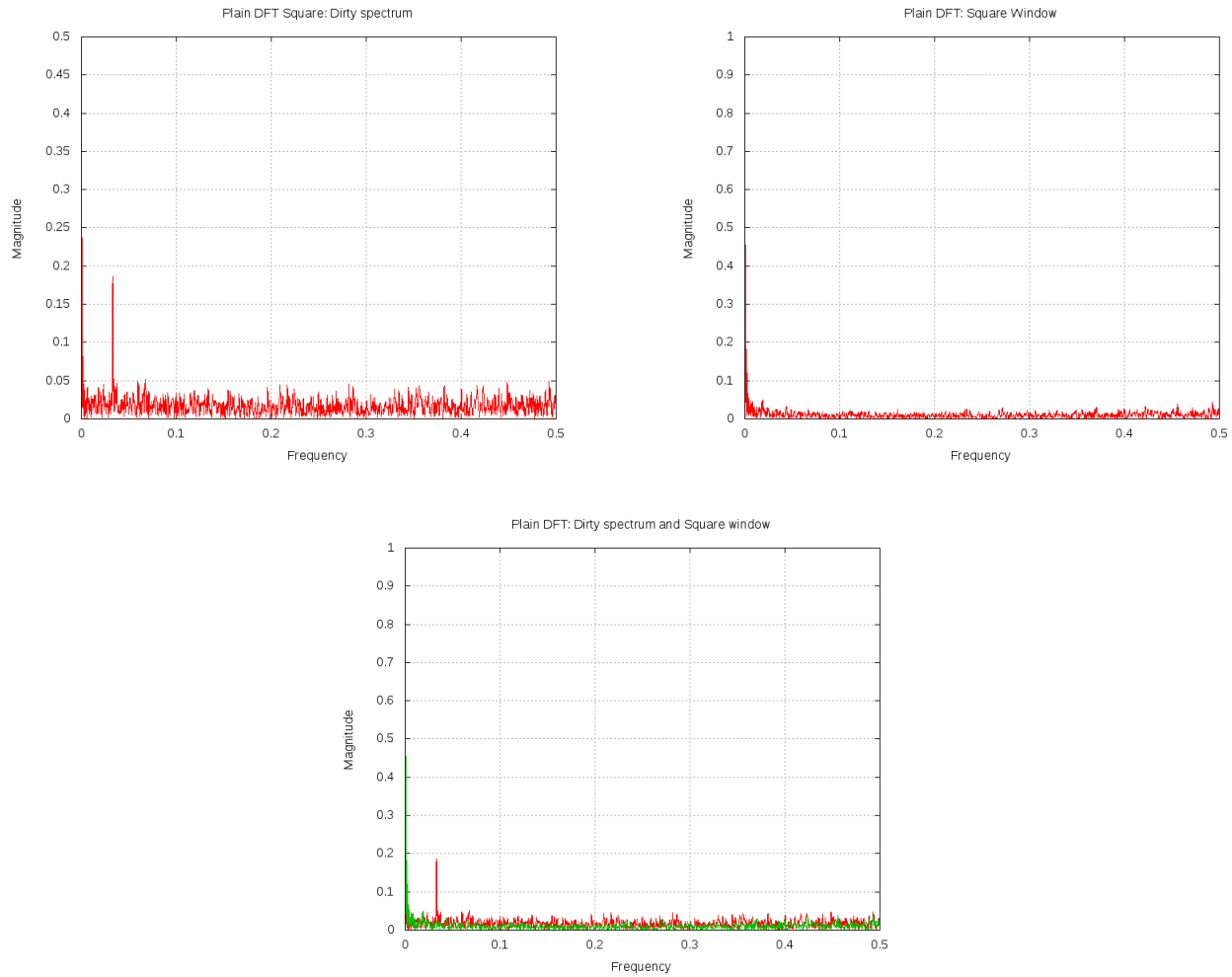
3.2.4 Transformada discreta de Fourier

L'aplicació de la transformada discreta de Fourier, sense optimitzacions computacionals especials, es pot realitzar amb diverses funcions finestres. S'ha calculat l'espectre brut i la finestra espectral utilitzant diferents funcions (presentades a la secció 2.2) i s'ha validat que no existeixen diferències significatives entre elles, sent la funció quadrada la més fàcil i ràpid d'implementar.

A continuació es presenten les gràfiques resultants de la transformada amb finestra quadrada: espectre brut, finestra espectral i la combinació d'ambdós.



En segon terme, per facilitar la lectura, s'inclouen les mateixes gràfiques però amb un rang acotat proper a la freqüència que disposa d'una major contribució:



Tal i com s'observa a la gràfica anterior a on es combina l'espectre brut i la finestra, existeix un pic a la zona dels $0,03Hz$ aproximadament, coincidint amb la freqüència amb una major contribució obtinguda pel mètode del Lomb-Scargle periodogram.

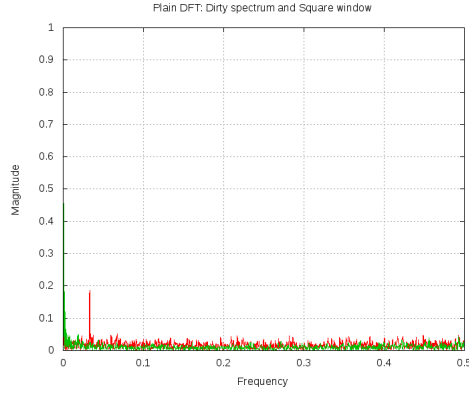
3.2.5 Algorisme Clean

S'ha desenvolupat la deconvolució de la finestra a partir de l'espectre brut (ambdós calculats i presentats a la secció anterior) utilitzant un procés iteratiu, unidimensional i complex de l'algorisme Hoegbom Clean.

Per a la programació ha sigut d'utilitat l'estudi previ del codi en Fortran proporcionat per David H. Roberts, un dels autors de l'estudi "*Time Series Analysis with Clean. I. Derivation of a spectrum*", així com una implementació en IDL per Simon Jeffery de l'Armagh Observatory.

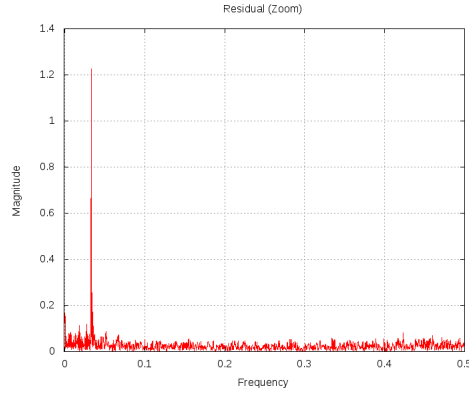
L'aplicació d'aquest algorisme consisteix en els passos següents:

1. Paràmetres inicials establerts mitjançant diverses proves d'execució:
 - (a) Guany: 0.5
 - (b) Nombre d'iteracions: 8
 - (c) Espectre brut i finestra quadrada

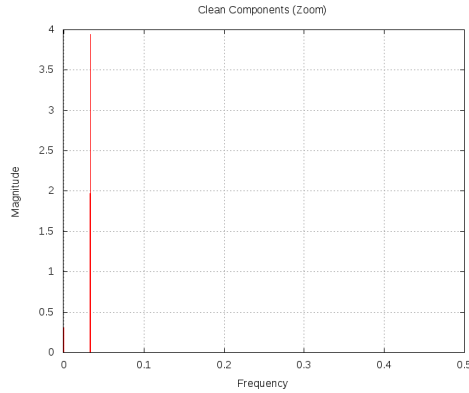


2. Procés iteratiu:

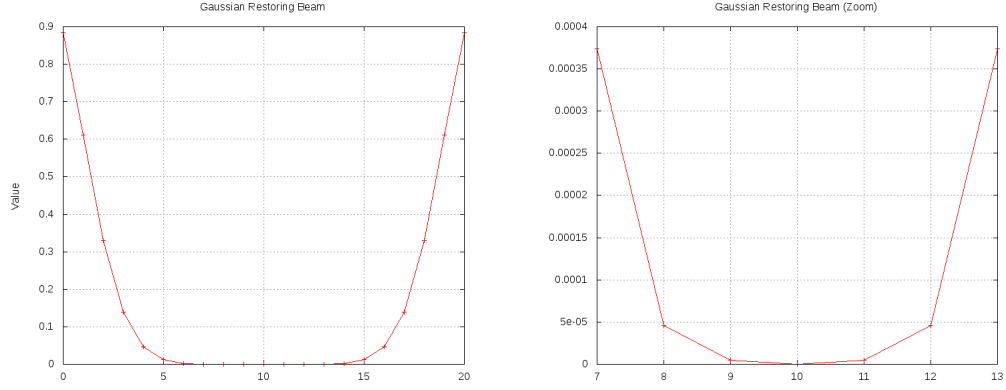
- (a) Identificació del punt de màxima contribució a l'espectre brut: L
- (b) Estimació del component net (Clean Component) d'aquest punt i multiplicació pel guany (equació 25 de l'article "*Time Series Analysis with Clean*"):
 - i. $CC = Guany \cdot \left(\frac{EspectreBrut[L] - Finestra[2 \cdot L] \cdot Conjugat(EspectreBrut[L])}{1 - Magnitud^2(Finestra[2 \cdot L])} \right)$
- (c) Extracció de la influència del component a l'espectre brut. Per cada element del vector espectre brut (equació 23 de l'article "*Time Series Analysis with Clean*"):
 - i. $EspectreBrutResidual[i] = EspectreBrutResidual[i] - CC \cdot Finestra[L + i] - Conjugat(CC) \cdot Finestra[-L + i]$



- (d) Emmagatzematge del component net en un vector



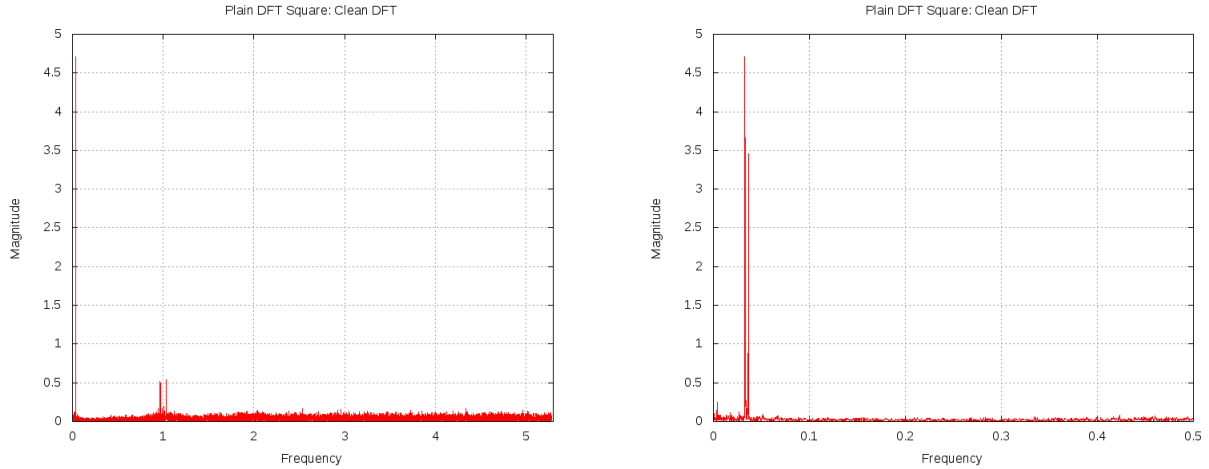
3. Construcció d'un feix gaussià de restauració



4. Convolució del feix de restauració amb el vector de components nets

5. Suma del resultat anterior amb l'espectre brut residual al que hem extret prèviament les influències dels components (punt 2.c.)

El resultat del procés proporciona un espectre net de les peculiaritats de la finestra. A continuació es presenta la gràfica general i, per facilitar la lectura, una altra visualització amb un rang acotat proper a la freqüència que disposa d'una major contribució.



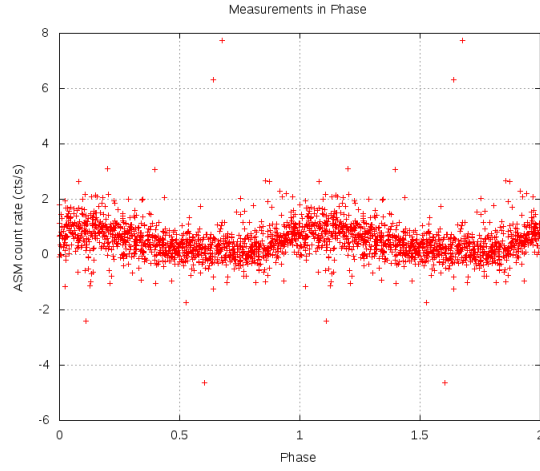
Mitjançant aquesta tècnica s'ha determinat que la freqüència amb una major contribució correspon a $0,033Hz$, coincidint amb el mètode del Lomb-Scargle periodogram.

3.3 Punts en fase

A partir del càlculs presentats als apartats anteriors, ha sigut possible determinar l'existència d'una freqüència que contribueix significativament al senyal ($0,033Hz$). Aquest fet denota que el senyal és periòdic i és possible calcular el període $P = \frac{1}{f_{pic}} = \frac{1}{0,033Hz} = 30,30s$ per representar els punts en fase:

$$\phi = \frac{t_i - t_0}{P} - Part\ Entera \left[\frac{t_i - t_0}{P} \right]$$

t_0 és el temps de la primera de les mesures registrades



3.4 Mínims quadrats

Mitjançant l'ús de GNU Scientific Libraries (GSL), s'ha realitzat un ajustament per mínims quadrats de la següent fórmula model a on volem determinar H , A , P i φ :

$$V(t) = H + A \cdot \cos\left(\frac{2\pi}{P}t + \varphi\right)$$

Inicialitzant H com el valor mitjà de les mesures i P com el període obtingut en apartats anteriors, s'ha obtingut com a resultat els següents valors estimats amb els respectius marges d'error:

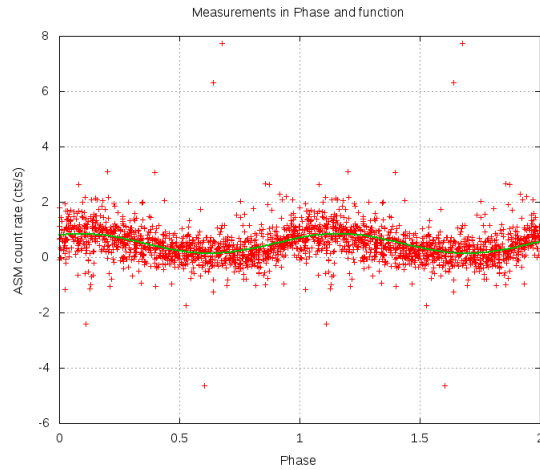
$$H = 0.49633 \pm 0.04090$$

$$A = 0.35254 \pm 0.04192$$

$$P = 1.07157 \pm 0.10738$$

$$\varphi = -0.54003 \pm 0.36976$$

La visualització de la funció anterior conjuntament amb les mesures en fase:



3.5 Eficiència computacional

Els càlculs computacionals portats a terme per cadascun dels mètodes presentats al document, han sigut mesurats per tal d'avaluar la seva eficiència i permetre una comparació relativa:

Tasca	Temps de CPU (segons)
Fast Fourier Transform (FFTW3)	0,0000917
Lomb-Scargle periodogram	0,0195532
Transformada discreta de Fourier	12,022
Clean Algorithm	1,02268
Mínims quadrats (Levenberg-Marquardt, GSL)	0,000976026

L'execució ha tingut lloc per un conjunt de dades de $N = 1.171$ elements, utilitzant un únic core de la CPU Intel(R) Core(TM)2 Duo CPU E8400 @ 3.00GHz.

Els resultats constaten l'eficiència de l'algorisme Fast Fourier Transform, no obstant els resultats d'aquest no poden ser utilitzats per no disposar de dades d'entrada equidistants en el temps. Per altra banda, amb els mateixos resultats, el Lomb-Scargle periodogram presenta un cost d'execució i desenvolupament significativament més baixos que l'ús de la transformada discreta de Fourier i l'algorisme Clean.

Referències

- [1] Roberts, D.H. Lehar, J. Dreher, J.W. (1986). *Time series analysis with Clean. I. Derivation of a spectrum*. Brandeis University.
- [2] Tarrés, F. (1995). *Introducció al tractament digital del senyal*. Editorial Bruño/EUETT.
- [3] Press, W.H. Teukolsky, S.A. Vetterling, W.T., Flannery, B.P. (2007). *Numerical Recipes 3rd Edition*. Cambridge.
- [4] Time Series Analysis. Retrieved December 26, 2010, from the World Wide Web: http://www.astrostatistics.psu.edu/su09/lecturenotes/Eric_TSA.pdf
- [5] Simon Jeffery de l'Armagh Observatory. IDL Clean implementation. Retrieved December 26, 2010, from the World Wide Web: <http://www.arm.ac.uk/~csj/idl/CLEAN/>
- [6] FFT Tutorial. University of Rhode. Retrieved December 26, 2010, from the World Wide Web: <http://www.ele.uri.edu/~hansenj/projects/ele436/fft.pdf>
- [7] Presentació Transformada de Fourier. Retrieved December 26, 2010, from the World Wide Web: http://www.slideshare.net/catita_potter/transformada-de-fourier-presentation
- [8] Wikipedia. *Transformada discreta de Fourier*. Retrieved December 26, 2010, from the World Wide Web: http://en.wikipedia.org/wiki/Discrete_Fourier_transform, http://en.wikipedia.org/wiki/Discrete-time_Fourier_transform
- [9] Wikipedia. *Mínims quadrats*. Retrieved December 26, 2010, from the World Wide Web: http://en.wikipedia.org/wiki/Least-squares_spectral_analysis
- [10] GNU Scientific Libraries (GSL). Retrieved December 26, 2010, from the World Wide Web: <http://www.gnu.org/software/gsl/>
- [11] FFTW3 libraries. Retrieved December 26, 2010, from the World Wide Web: <http://www.fftw.org/>