

Análisis de series

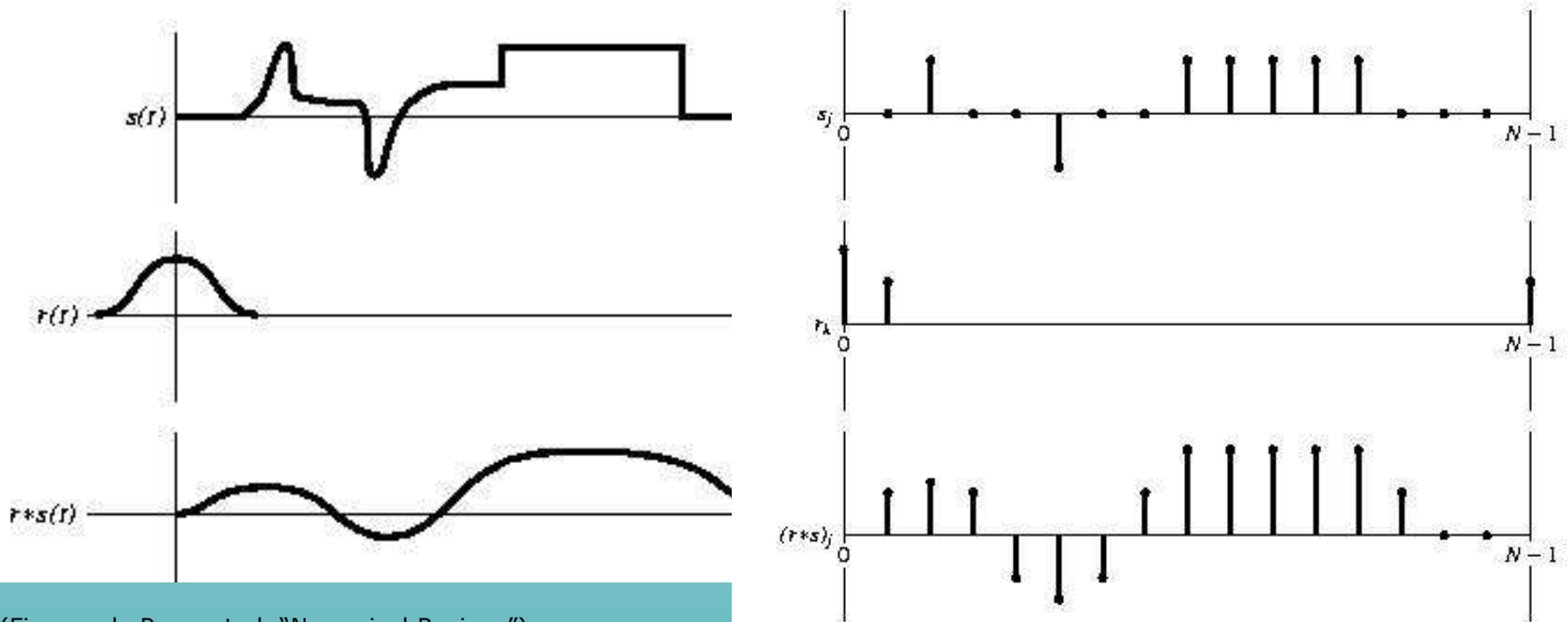
Métodos espectrales: convolución

Sea $s(t)$ una señal función de la variable t y $r(t)$ una función de respuesta. Se define la operación convolución de r con s

$$r \otimes s \equiv \int_{-\infty}^{\infty} s(t - \tau) r(\tau) d\tau$$

$$(r \otimes s)_j \equiv \sum_{k=-M/2+1}^{M/2} s_{j-k} r_k$$

El efecto de la convolución es la suavización de la función señal y su contaminación para un t dado con señal proveniente de t adyacentes.



(Figuras de Press et al. "Numerical Recipes")

Métodos espectrales: correlación

♦ La **correlación** entre dos señales $g(t)$, $h(t)$ viene dada por

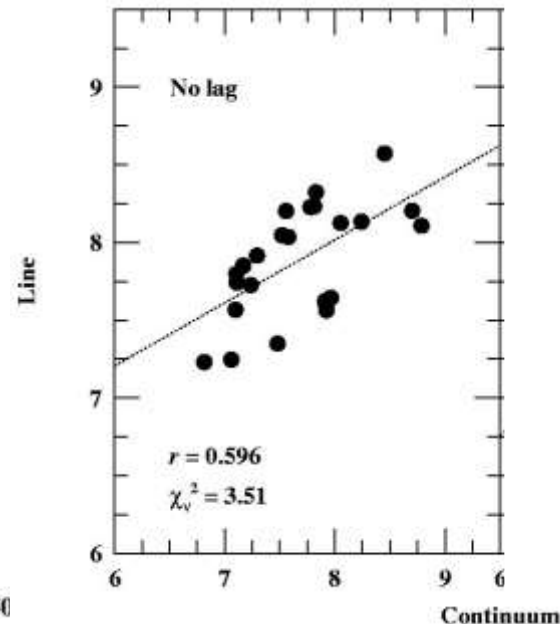
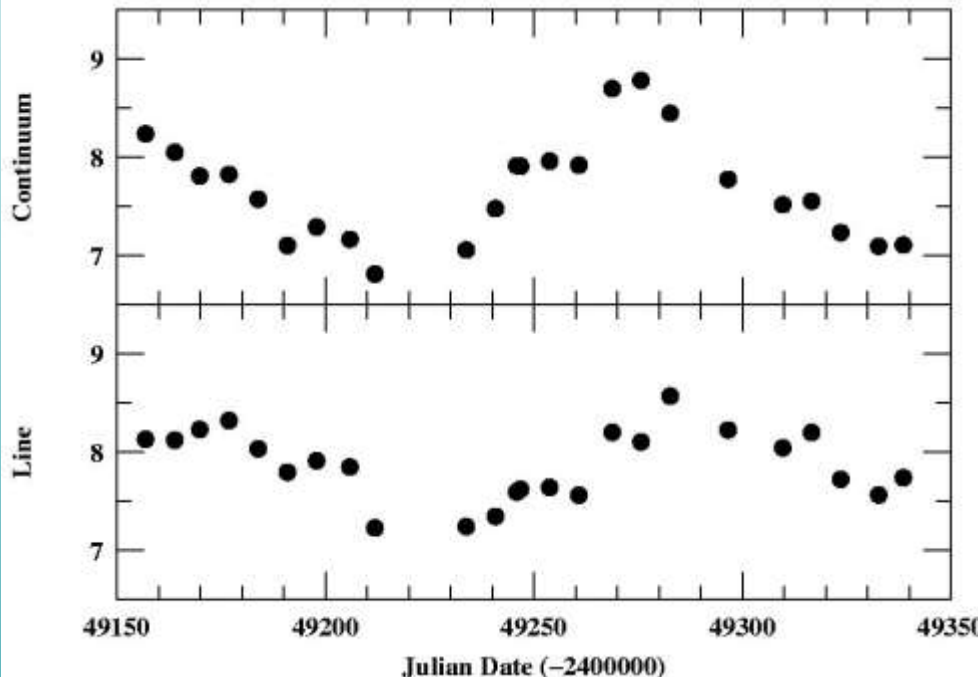
$$CF(g, h)(t) \equiv \int_{-\infty}^{\infty} g(t + \tau) h(\tau) d\tau$$

$$DCF(g, h)_j \equiv \sum_{k=0}^{N-1} g_{j+k} h_k$$

donde **CF** denota *correlation function* y **DCF** *discrete correlation function*. Cuando las señales son funciones del tiempo, t se denomina retraso (*lag* en inglés)

♦ La **autocorrelación** se define, de forma análoga, como $ACF(g) \equiv CF(g, g)$ donde **ACF** denota *autocorrelation function*.

La ACF se suele emplear para encontrar períodos.



Métodos espectrales: correlación

♦ La **correlación** entre dos señales $g(t)$, $h(t)$ viene dada por

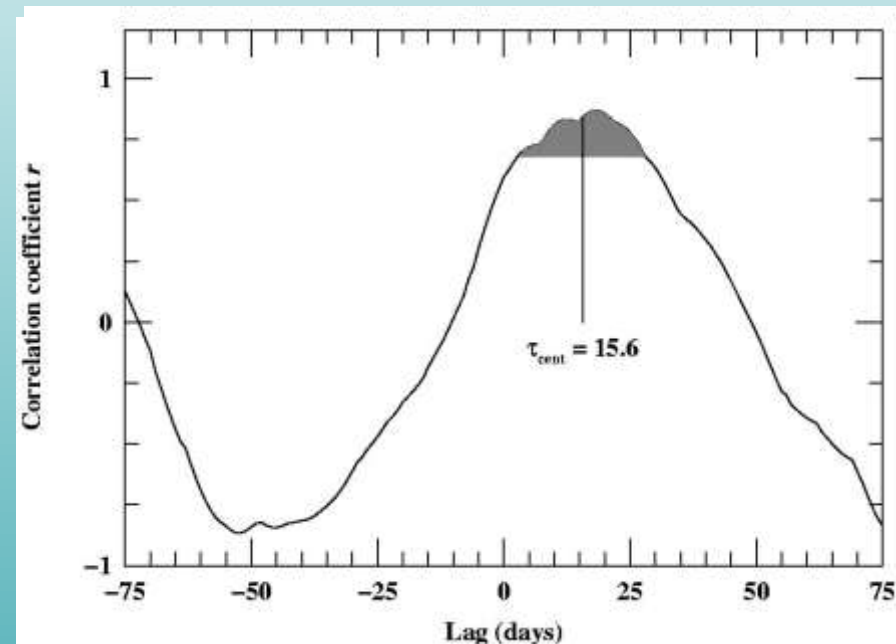
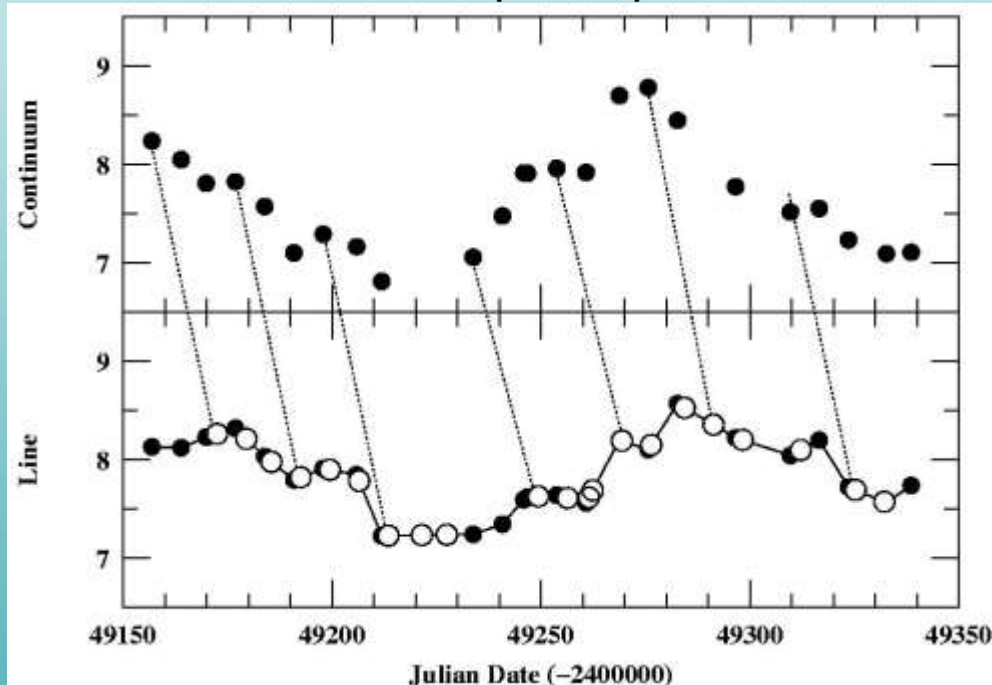
$$CF(g, h)(t) \equiv \int_{-\infty}^{\infty} g(t + \tau) h(\tau) d\tau$$

$$DCF(g, h)_j \equiv \sum_{k=0}^{N-1} g_{j+k} h_k$$

donde **CF** denota *correlation function* y **DCF** *discrete correlation function*. Cuando las señales son funciones del tiempo, t se denomina retraso (**lag** en inglés)

♦ La **autocorrelación** se define, de forma análoga, como $ACF(g) \equiv CF(g, g)$ donde **ACF** denota *autocorrelation function*.

La ACF se suele emplear para encontrar períodos.



Métodos espectrales: correlación

♦ La **correlación** entre dos señales $g(t)$, $h(t)$ viene dada por

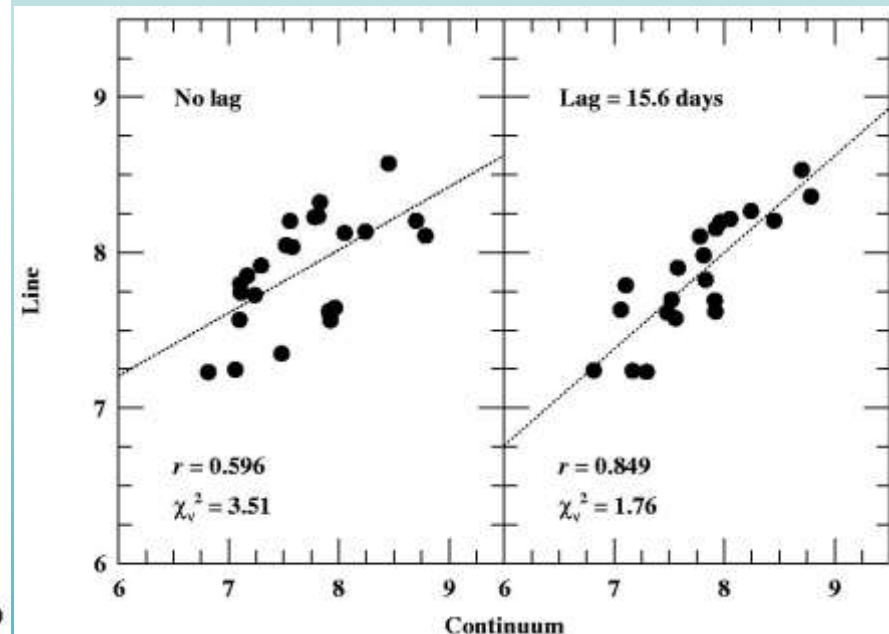
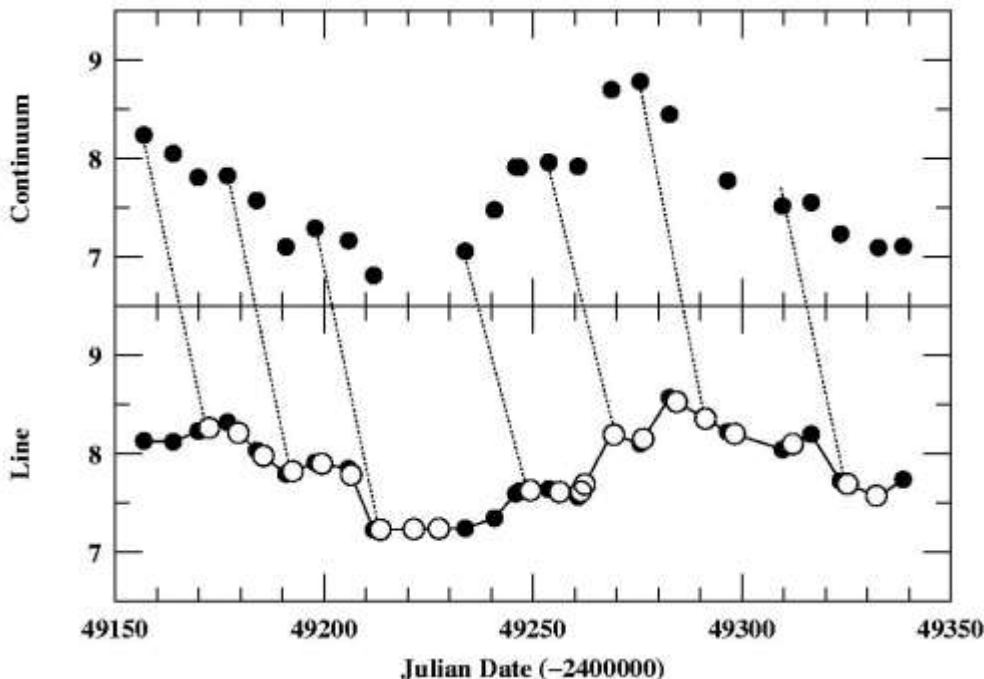
$$CF(g, h)(t) \equiv \int_{-\infty}^{\infty} g(t + \tau) h(\tau) d\tau$$

$$DCF(g, h)_j \equiv \sum_{k=0}^{N-1} g_{j+k} h_k$$

donde **CF** denota *correlation function* y **DCF** *discrete correlation function*. Cuando las señales son funciones del tiempo, t se denomina retraso (**lag** en inglés)

♦ La **autocorrelación** se define, de forma análoga, como $ACF(g) \equiv CF(g, g)$ donde **ACF** denota *autocorrelation function*.

La ACF se suele emplear para encontrar períodos.



Métodos espectrales: correlación



8.4

Métodos espectrales: función de estructura

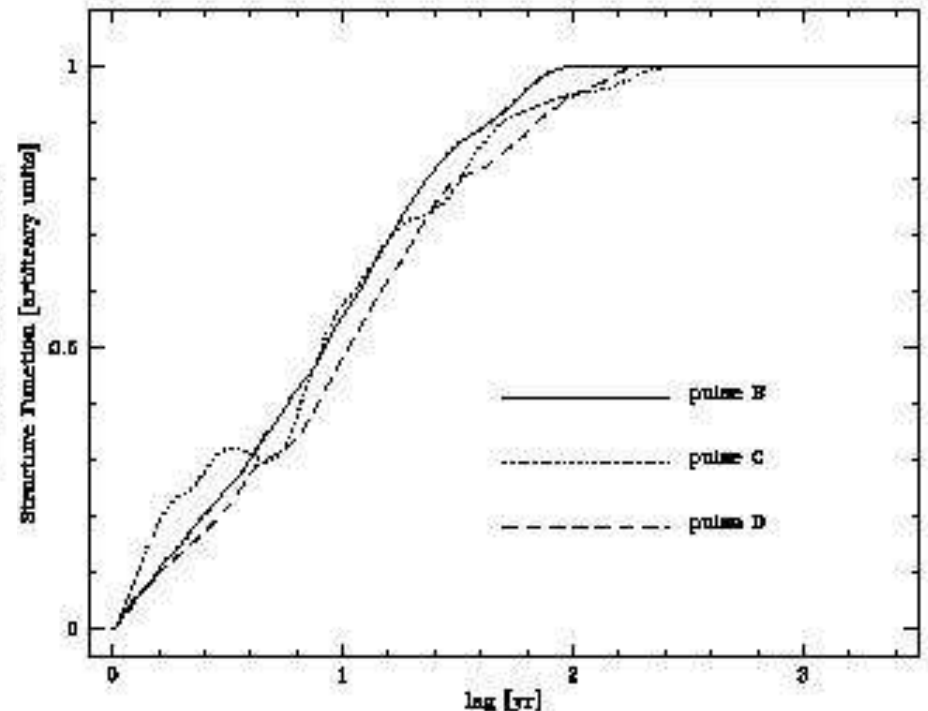
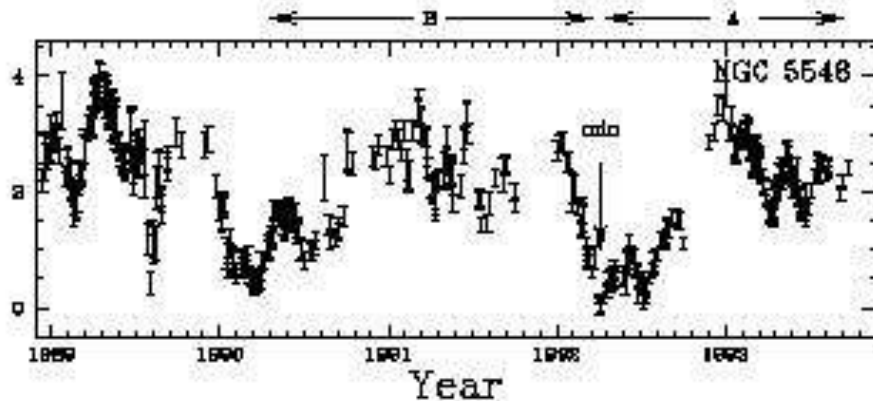
Se define la **función de estructura** de una señal $s(t)$, como

$$SF(t) \equiv \frac{1}{t} \int_0^t [s(t+\tau) - s(\tau)]^2 d\tau$$

$$SF(\Delta t) \equiv \left\langle [s(t + \Delta t) - s(t)]^2 \right\rangle$$

donde **SF** denote *structure function*.

SF da cuenta del incremento de la variación como función de t , es decir, mide la escala de t en la que se produce la variancia máxima (Simoneti et al....???)



Métodos espectrales: función de estructura

Ejemplo: SF como promedio de una población de fuentes (Aretxaga et al. 1997, MNRAS, 286, 271)

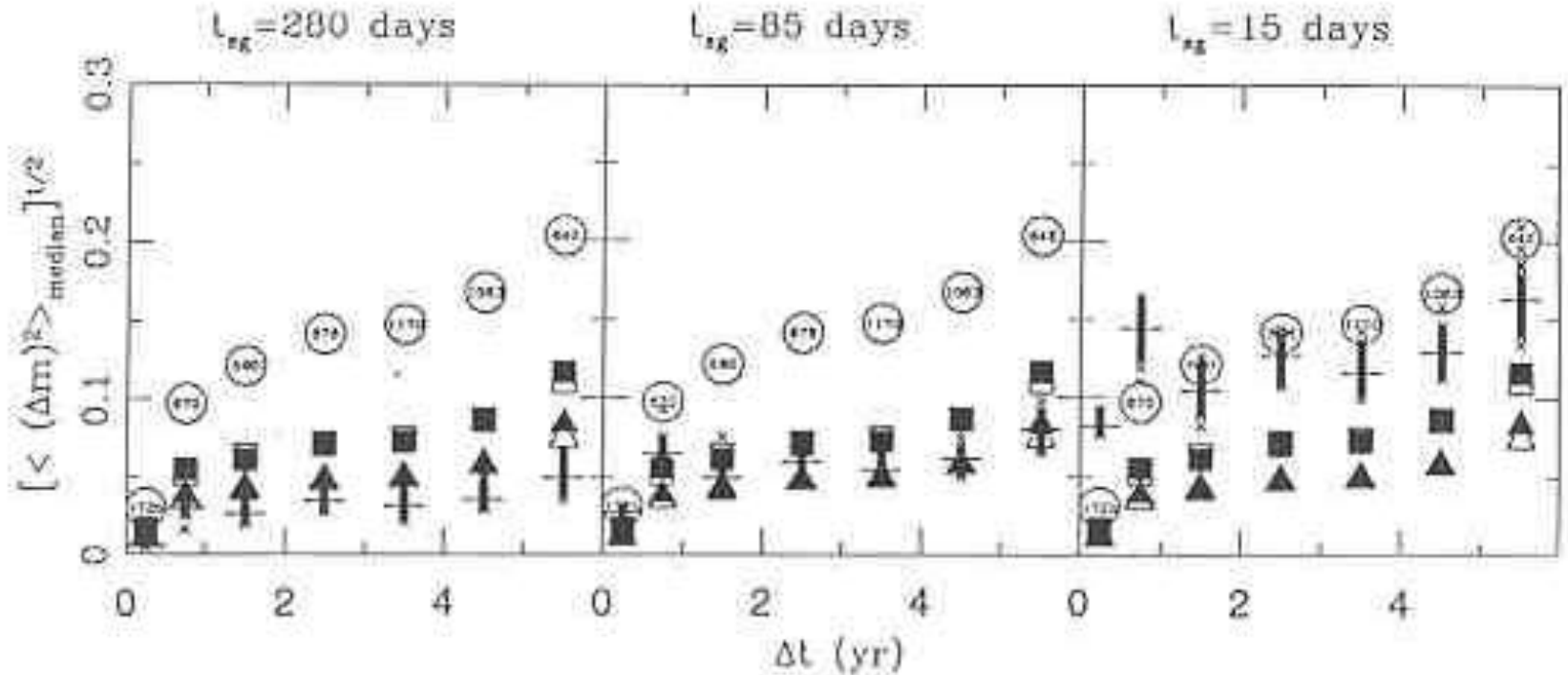


Figure 7: Square root of the structure function measured in the Monte Carlo simulations of the SGP sample, compared with the measured function of the observed sample. All the QSO light curves are treated as an ensemble light curve. Crosses represent the individual SFs obtained for 100 simulated samples, and thick dashes represent the median values of all the simulated samples. Open circles represent the SF of the observed sample. Open squares and triangles represent the same SF after applying a linear variability-wavelength correction, such that a factor 2 in the $\sigma(\lambda 1350 \text{ \AA})/x_g$ ratio is described by squares and a factor 3 by triangles. Filled squares and triangles represent the data after correcting the wavelength dependence of variability with a power law of indices -0.6 and -1.0 respectively (factors 2 and 3 for the same ratio). The numbers next to the open circles give the total number of differences involved in the derivation of the structure function for the corresponding time-bin: 1725, 630, 590, 838, 1170, 1063, 642 from $\Delta t = 0.5$ to 0.55 yr.

Métodos espectrales: espectro de potencias

Se define la potencia de una señal $s(t)$ a una frecuencia ν , como

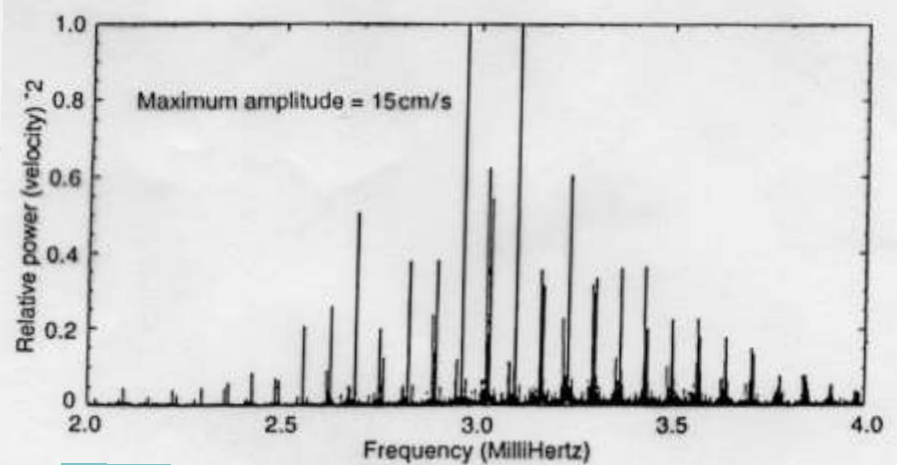
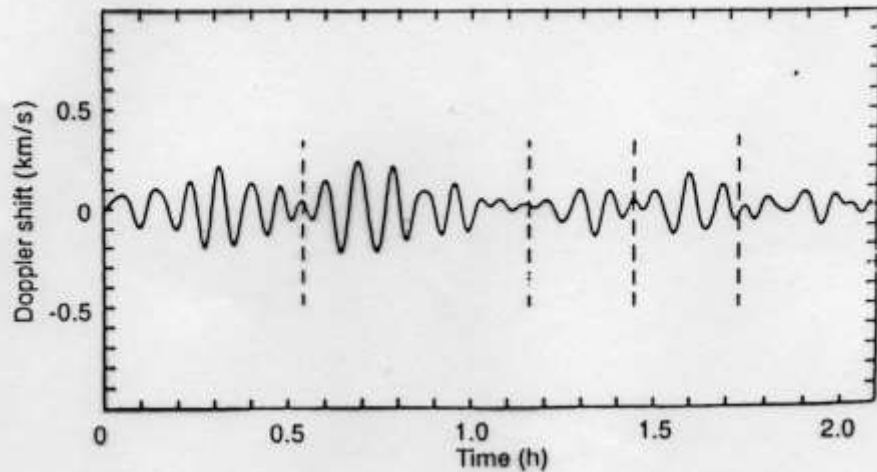
$P(\nu) \equiv |S(\nu)|^2 \quad -\infty \leq \nu \leq \infty$ donde

$$S(\nu) \equiv \text{FT}(s(t)) \equiv \int_{-\infty}^{\infty} dt s(t) e^{+2\pi i \nu t}$$

$$S(\nu) \equiv \text{DFT}(s(t)) \equiv \frac{1}{N} \sum_{k=0}^{N-1} s_k e^{+2\pi i \nu t_k}$$

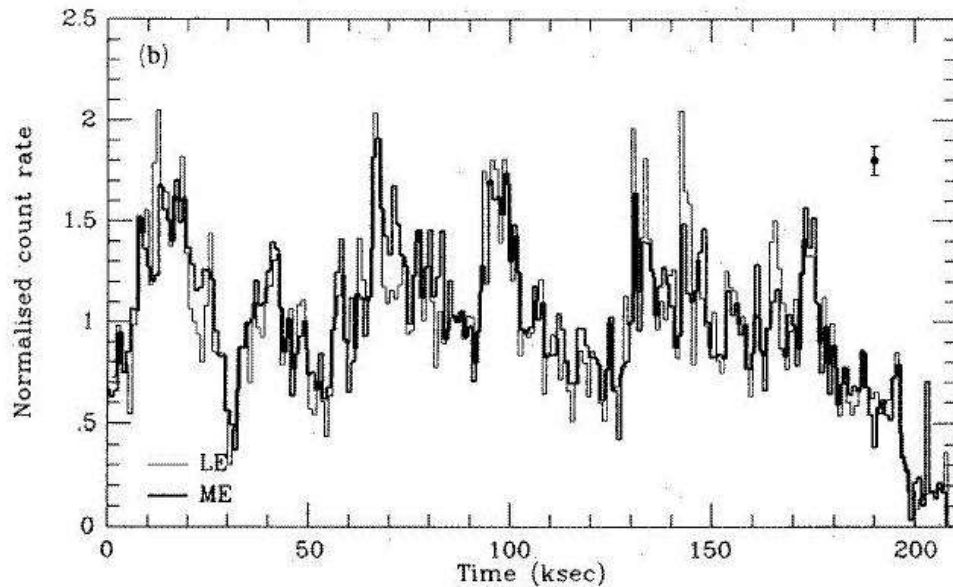
donde **FT** denota *Fourier transform*. Alternativamente, la potencia también se define como $P(\nu) \equiv |S(\nu)|^2 + |S(-\nu)|^2 = 2|S(\nu)|^2 \quad 0 \leq \nu \leq \infty$

El espectro de potencias mide la contribución a la variancia de una cierta frecuencia ν .

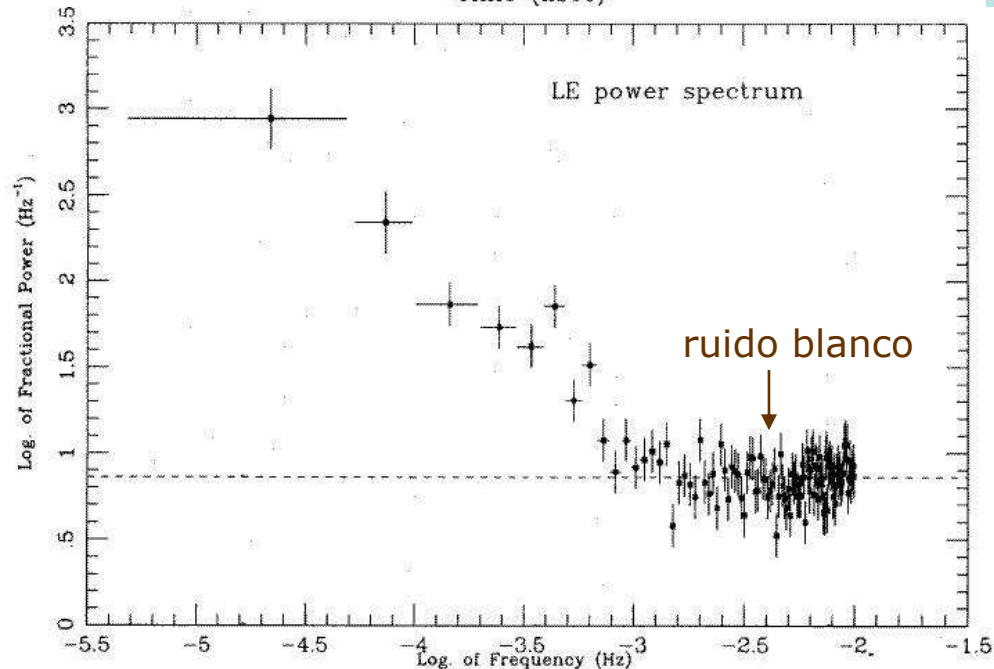


Ejemplo: espectro de potencias de corrimientos Doppler en el sol. Medida de periodos (Wentzel 1989)

Métodos espectrales: espectro de potencias

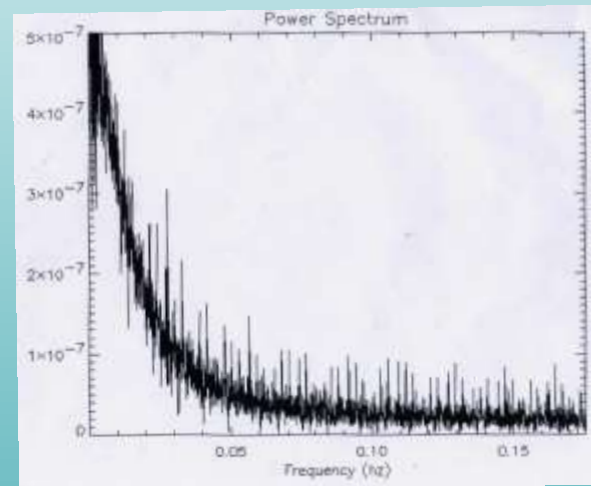
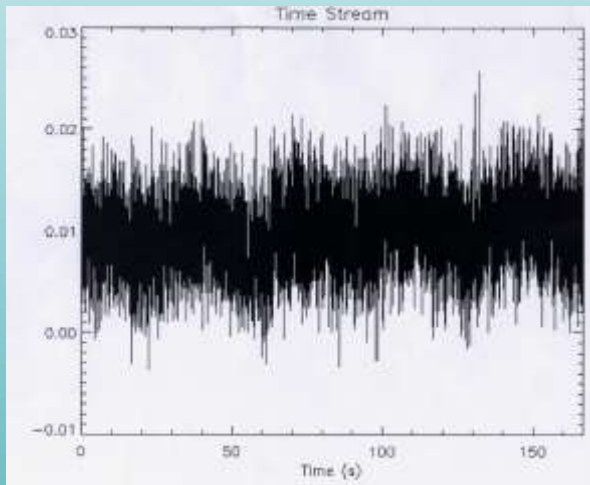


Ejemplo: espectro de potencias de la curva de luz en LE (0.05-2keV) y ME (2-10keV) de NGC4051 medida por ROSAT (Papadakis & Lawrence 1995, MNRAS, 272, 161)



Métodos espectrales: espectro de potencias

- ♦ Ruido blanco (*white noise*), es un espectro de potencias plano, con igual potencia por cada intervalo de frecuencia (Hz). Ejemplos de procesos que crean ruidos blancos son los proceso poissonianos no correlacionados (*shot noise*), o el ruido térmico en circuitos electrónicos (ruido de Johnson).
- ♦ Ruido rosa o ruido $1/f$ (*flicker noise*, ruido de exceso), está caracterizado por una potencia constante por cada década de frecuencia. Es inherente a muchos sistemas de eléctricos, asociado a variaciones en la resistencia. Por lo tanto, es algo a tener en cuenta al diseñar observaciones, por ejemplo con bolómetros. También aparece frecuentemente en la naturaleza: la velocidad de las corrientes marinas, el flujo de los granos en un reloj de arena, la potencia de una pieza musical, ...



- ♦ Ruido rojo o ruido $1/f^2$

(Horowitz & Hill, 1980, "The Art of Electronics", Cambridge University Press)

Transformadas de Fourier

La transformación de Fourier (FT) convierte funciones del espacio temporal o espacial al espacio de frecuencias temporales o espaciales. Esta transformación simplifica el cálculo de los métodos espectrales.

transformada

$$H(\nu) \equiv \text{FT}(h) \equiv \int_{-\infty}^{\infty} dt h(t) e^{+2\pi i \nu t}$$

antitransformada

$$h(t) \equiv \text{FT}^{-1}(H) \equiv \int_{-\infty}^{\infty} d\nu H(\nu) e^{-2\pi i \nu t}$$

En Física se utiliza la frecuencia expresada en rad/s en vez de en ciclos/s,

$$\omega = 2\pi\nu \quad H(\omega) \equiv \int_{-\infty}^{\infty} dt H(t) e^{+i\omega t} \quad h(t) \equiv \int_{-\infty}^{\infty} d\omega H(\omega) e^{-i\omega t}$$

◆ Propiedades

si $h(t)$ es real

$$\rightarrow H(-\nu) = H^*(\nu)$$

imaginario

$$\rightarrow H(-\nu) = -H^*(\nu)$$

par

$$\rightarrow H(-\nu) = H(\nu)$$

impar

$$\rightarrow H(-\nu) = -H(\nu)$$

real y par

$$\rightarrow H(\nu) \text{ es real y par}$$

real e impar

$$\rightarrow H(\nu) \text{ es imaginario e impar}$$

imaginario y par

$$\rightarrow H(\nu) \text{ es real y par}$$

imaginario e impar $\rightarrow H(\nu)$ es real e impar

Transformadas de Fourier

◆ Propiedades:

escalaje

$$\text{FT}(h(at)) = \frac{1}{|a|} H(\nu / a) \quad , \quad \text{FT}\left(\frac{1}{|b|} h(t/|b|)\right) = H(b\nu)$$

corrimiento

$$\text{FT}(h(t-t_0)) = H(\nu) e^{+2\pi i \nu t_0} \quad , \quad \text{FT}(h(t) e^{-2\pi i \nu_0 t}) = H(\nu - \nu_0)$$

teorema de **convolución**: $\text{FT}(g \otimes h) = G(\nu) H(\nu)$

teorema de **correlación**: $\text{FT}(\text{CF}(g, h)) = G(\nu) H^*(\nu)$

teorema de **Wiener-Khinchin**: $\text{FT}(\text{ACF}(g)) = |G(\nu)|^2$

teorema de **Parseval**: $\int_{-\infty}^{\infty} dt |h(t)|^2 = \int_{-\infty}^{\infty} d\nu |H(\nu)|^2$

Transformadas de Fourier discretas (DFT)

Sea una señal $s(t)$ medida a intervalos equidistantes de frecuencia $1/\Delta$:

$$s_n = s(n\Delta) \text{ donde } n = \dots, -3, -2, -1, 0, 1, 2, 3, \dots$$

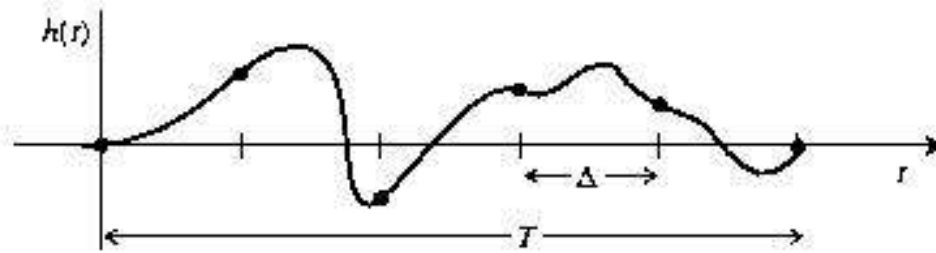
Se define la **frecuencia crítica de Nyquist** como $\nu_c \equiv 1/2\Delta$ donde $S(\nu) = 0$ para $|\nu| \geq \nu_c$

♦ **Teorema de muestreo:** si una función continua $s(t)$, evaluada a intervalos equidistantes Δ , está acotada en su ancho de banda ($S(\nu) = 0$ para $|\nu| \geq \nu_c$), entonces la función está completamente determinada por su muestra s_n .

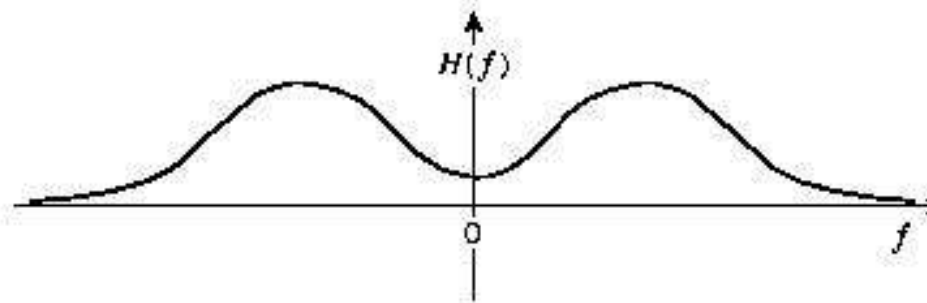
$$s(t) = \Delta \sum_{n=-\infty}^{\infty} s_n \frac{\sin[2\pi\nu_c(t - n\Delta)]}{\pi(t - n\Delta)}$$

Si la función no se evalúa con una frecuencia mayor que la frecuencia de Nyquist, entonces ocurre el **engorroso problema de la creación de sobrenombres** (*aliasing*)

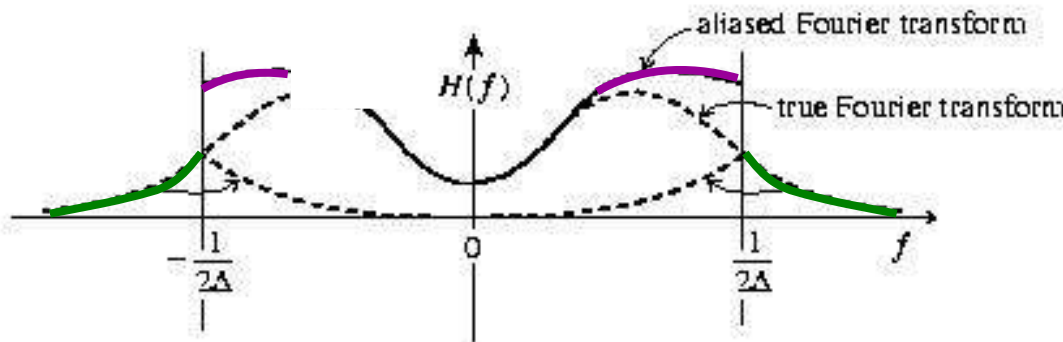
Transformadas de Fourier discretas (DFT)



(a)



(b)



Recomendación: examínese la FT cerca de los valores correspondientes a la frecuencia crítica de Nyquist. Si la FT no se aproxima a 0, probablemente existan contribuciones de frecuencias que se han doblado dentro del ancho de banda. (Press et al. "Numerical Recipes")

Transformadas de Fourier discretas (DFT)

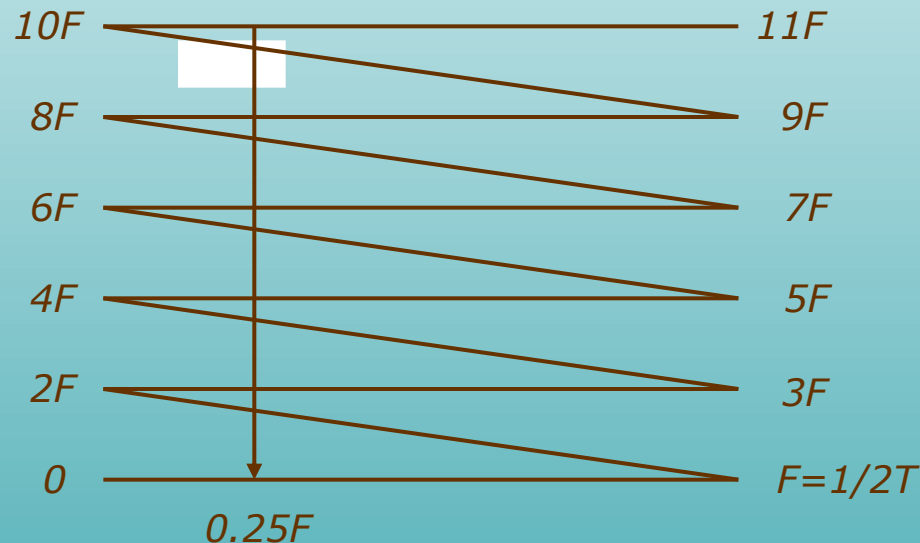
Ejemplo de creación de sobrenombres: senoide de frecuencia mayor que $1/2T$, con muestras tomadas a intervalos T .

$$x(t) = \sin 2\pi\nu_0 t \quad \text{donde} \quad \nu_0 = \frac{p+q}{2T} \quad \text{y } p \text{ es entero y } q \text{ es fracción.}$$

$$x(t) = \sin 2\pi \frac{p+q}{2T} kT = \sin \cancel{\pi p k} \cos \pi q k + \cos \pi p k \sin \pi q k = \cos \pi p k \sin \pi q k$$

- si p es par $\cos \pi p k = 1 \quad \forall k$ y la señal queda $x(k) = \sin \pi q k = \sin 2\pi \frac{q}{2T} kT = \sin 2\pi \nu'_0 t$
- si p es impar $\cos \pi p k = (-1)^k$ entonces $x(k) = (-1)^k \sin \pi q k = \cos \pi k \sin \pi q k = -\sin(\pi k - \pi q k) = -\sin \pi k(1 - q) = -\sin 2\pi \left(\frac{1}{2T} - \frac{q}{2T} \right) kT = -\sin 2\pi (F - \nu'_0) t$

Para $\nu_0 = 1025$ Hz y $T = 0.005$ s : $F = 1/2T = 100$ Hz, $p = 10$, $q = 0.25$ \longrightarrow $\nu'_0 = 25$ Hz



Transformadas de Fourier discretas (DFT)

Método de Danielson-Lanczos para el cálculo de transformadas rápidas de Fourier (FFT)

Se basa en que una DFT de tamaño N puede reescribirse como dos DFT de tamaño $N/2$, la primera formada por los datos pares, y la segunda por los impares.

$$\begin{aligned} H_n &= \sum_{k=0}^{N-1} h_k e^{+2\pi i k n / N} = \sum_{k=0}^{N/2-1} h_{2k} e^{+2\pi i n (2k) / N} + \sum_{k=0}^{N/2-1} h_{2k+1} e^{+2\pi i n (2k+1) / N} = \\ &= \sum_{k=0}^{N/2-1} h_{2k} e^{+2\pi i n k / (N/2)} + \underbrace{\left(e^{2\pi i / N} \right)^n}_{\equiv W^n} \sum_{k=0}^{N/2-1} h_{2k+1} e^{+2\pi i n k / (N/2)} = H_n^e + W^n H_n^o \end{aligned}$$

Así se puede ir descomponiendo a DFT de menor complejidad, con $N/4$, $N/8$, $N/16$, ... Hasta llegar a 1 $H_n^{eoeo...oe} = h_k$ para algún k .

El algoritmo para encontrar qué k es:

1. reviertase el diseño de e y o
2. hágase $e=0$, $o=1$
3. k = número resultante en base binaria

Existen otros algoritmos para calcular FFT

Convolución y deconvolución

$$r \otimes s \equiv \int_{-\infty}^{\infty} s(t - \tau) r(\tau) d\tau$$

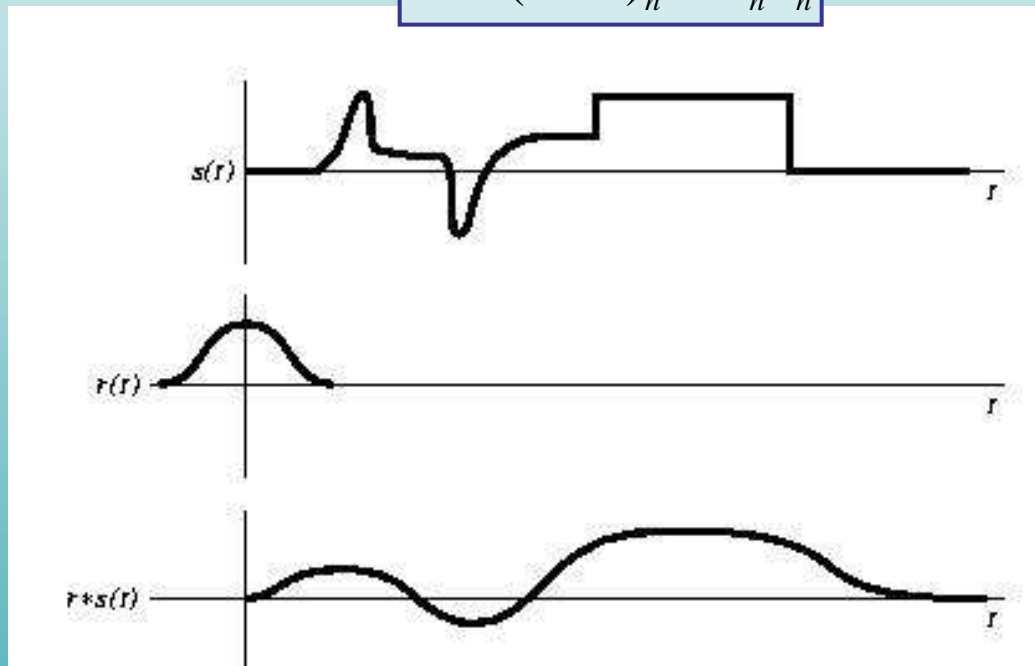
$$(r \otimes s)_j \equiv \sum_{k=-M/2+1}^{M/2} s_{j-k} r_k$$

- Teorema de convolución $\boxed{\text{FT}(r \otimes s) = R(\nu)S(\nu)}$

- Teorema de convolución discreta:

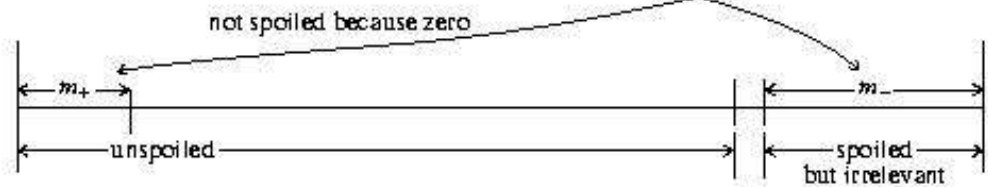
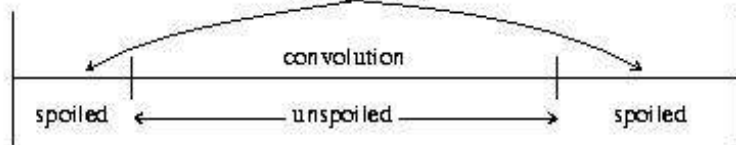
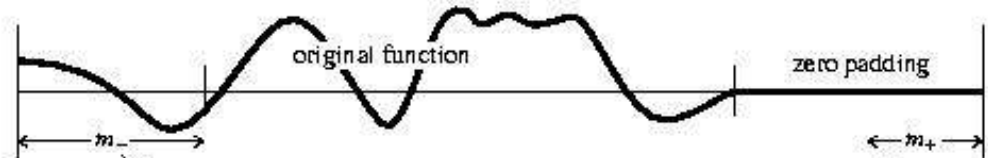
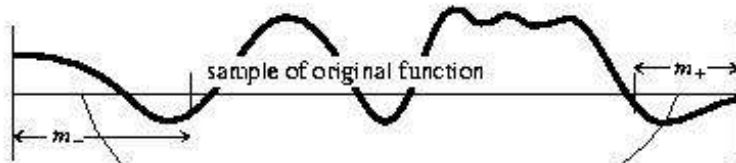
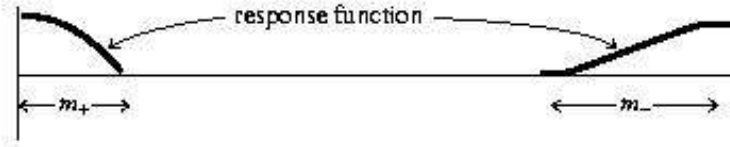
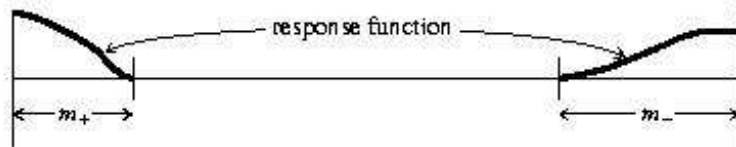
Si s_j es una señal periódica, de periodo N , completamente determinada s_0, \dots, s_{N-1} y r es una función de respuesta de impulso infinito (no 0 sólo en $-M/2 \leq k \leq M/2$, intervalo menor que N), entonces:

$$\boxed{\text{DFT}(r \otimes s)_n = R_n S_n}$$



Convolución y deconvolución

- **Cálculos con funciones no periódicas:** se introducen artificialmente 0s al final de la función de señal s , tantos como número de índices positivos o negativos en r . Esto impide la contaminación de frecuencias con tiempos que toma como periódicos.



Correlación y autocorrelación

$$CF(g, h)(t) \equiv \int_{-\infty}^{\infty} g(t + \tau) h(\tau) d\tau$$

$$DCF(g, h)_j \equiv \sum_{k=0}^{N-1} g_{j+k} h_k$$

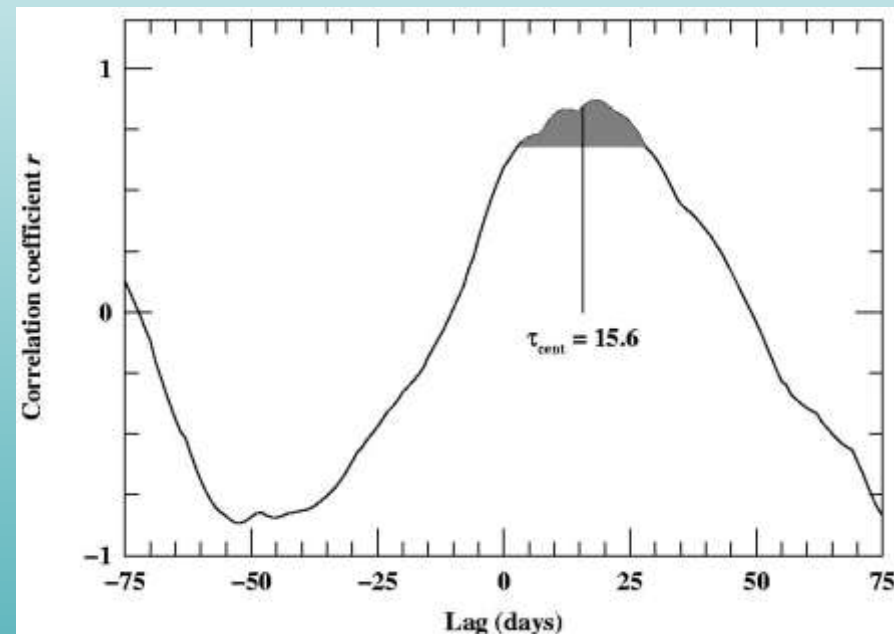
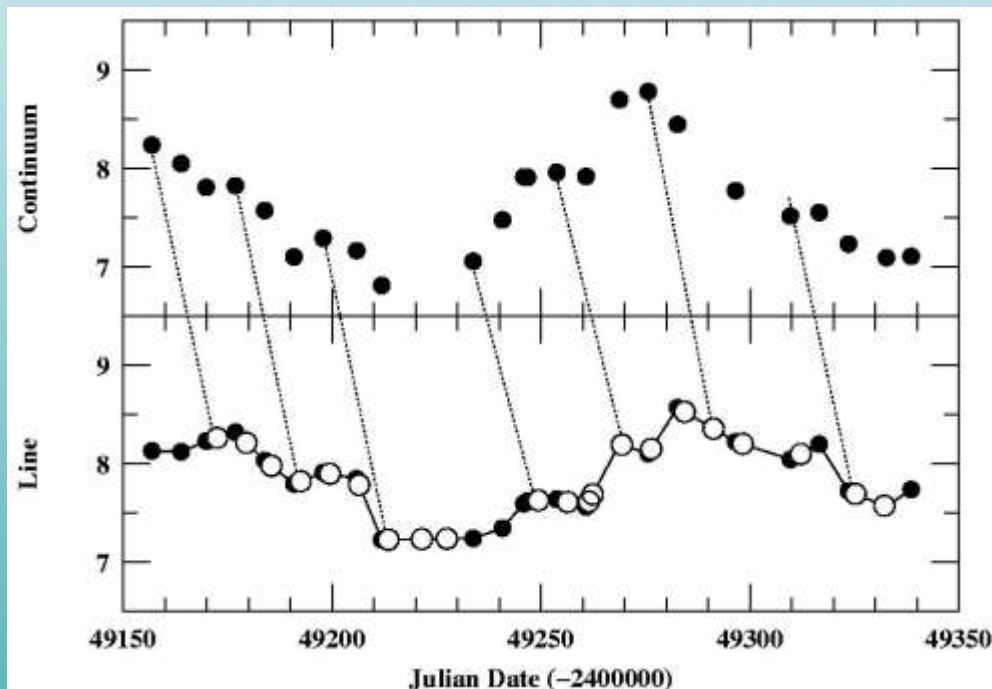
- Teorema de correlación

$$CF(g, h) = G(\nu) H^*(\nu)$$

- Teorema de correlación discreta:

$$DCF(g, h)_j = G_n H_n^*$$

Si s_j es una señal periódica, de periodo N , completamente determinada s_0, \dots, s_{N-1} y r es una función de respuesta de impulso infinito (no 0 sólo en $-M/2 \leq k \leq M/2$, intervalo menor que N), entonces:



Espectro de potencias: periodograma

Sea una señal $s(t)$ evaluada en intervalos equidistantes, su transformada de Fourier discreta viene dada por

$$S_n = \sum_{k=0}^{N-1} s_k e^{+2\pi i n k / N} \quad , \quad n = 0, \dots, N-1$$

y su **periodograma** en $N/2+1$ frecuencias viene definido por

$$P(\nu_0) = \frac{1}{N^2} |S_0|^2$$

$$P(\nu_j) = \frac{1}{N^2} \left[|S_j|^2 + |S_{N-j}|^2 \right] \quad , \quad j = 1, 2, \dots, (N/2 - 1)$$

$$P(\nu_c) = P(\nu_{N/2}) = \frac{1}{N^2} |S_{N/2}|^2$$

donde ν_j se define sólo para frecuencias positivas o cero:

$$\nu_j = \frac{j}{N\Delta} = 2\nu_c \frac{j}{N} \quad , \quad j = 0, \dots, N/2$$

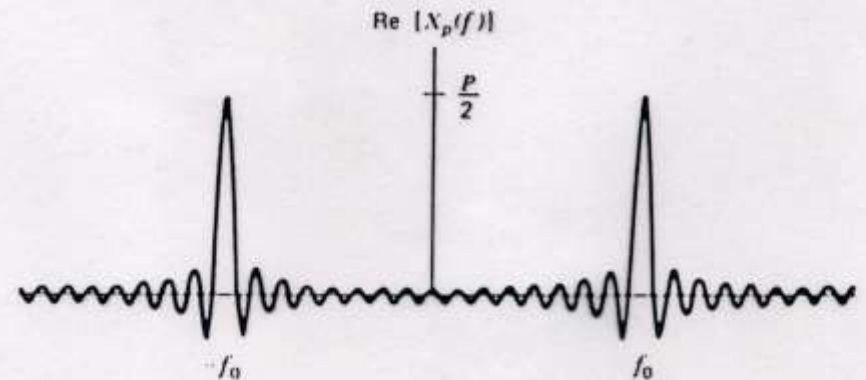
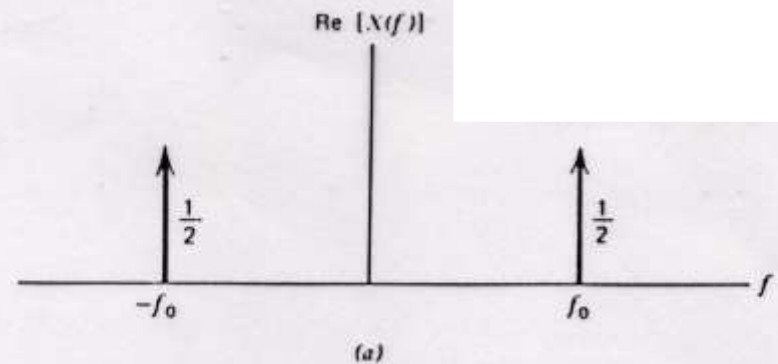
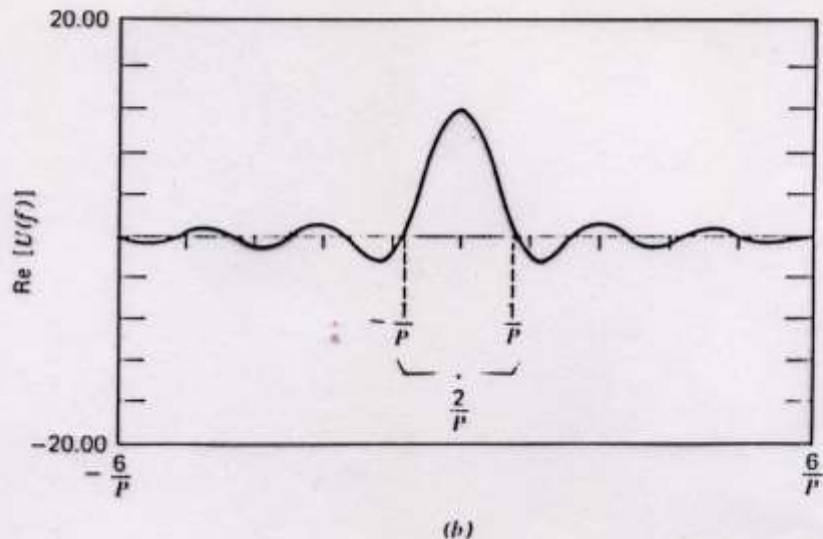
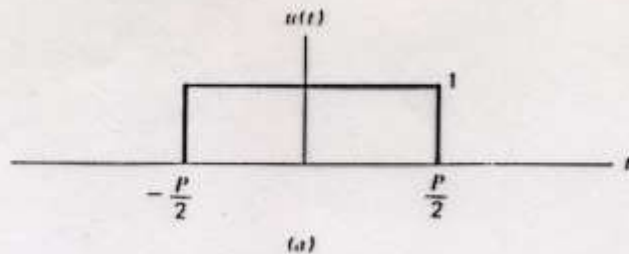
Por el teorema de Parseval, vemos que la suma de los $N/2+1$ valores del periodograma dan la amplitud media cuadrada de s_k .

Espectro de potencias: periodograma

$P(v_j)$ es un promedio de $P(v)$ sobre una ventana estrecha (W) centrada en v_j . Si definimos s como la diferencia de frecuencias en casillas

$$W(s) = \frac{j}{N^2} \left[\frac{\sin \pi s}{\sin(\pi s / N)} \right]^2$$

lo que implica un derrame (*leakage*) de una frecuencia a otra



Espectro de potencias: periodograma

El periodograma no se vuelve más preciso si N aumenta: la desviación estándar es siempre 100% del valor del periodograma, independiente de N .

- Métodos para aumentar la precisión:

1. producir un periodograma con N grande y sumar los valores de frecuencias adyacentes. No promediar.

2. dividir los datos en k segmentos de $2M$ puntos consecutivos. Calcular FFT de cada segmento, y promediarlo para obtener k puntos del periodograma. La desviación estándar será así \sqrt{k} .

- Ventanas para disminuir el derrame:

$$P(\nu_0) = \frac{1}{W_{ss}} |S_0|^2$$

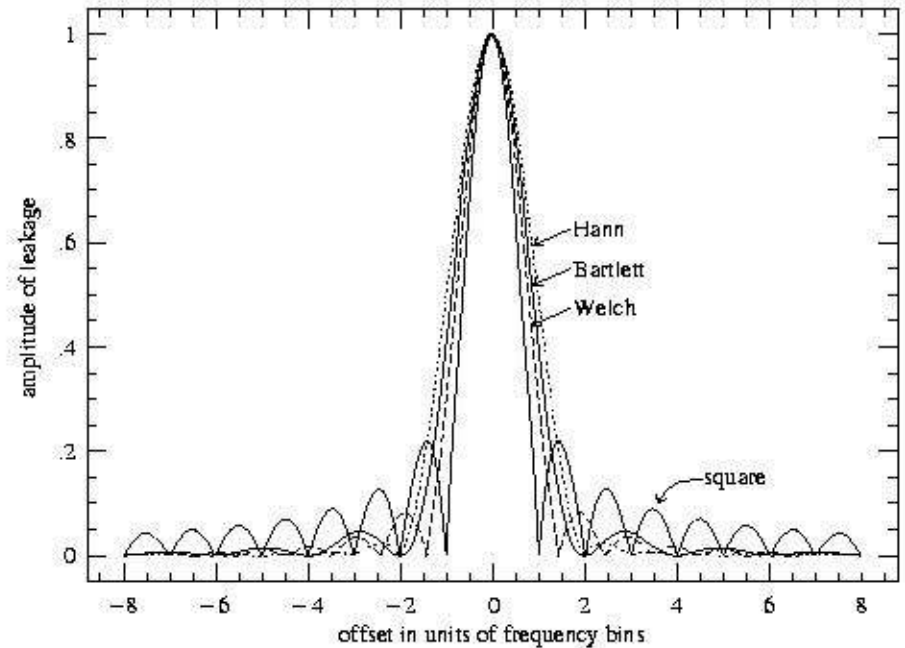
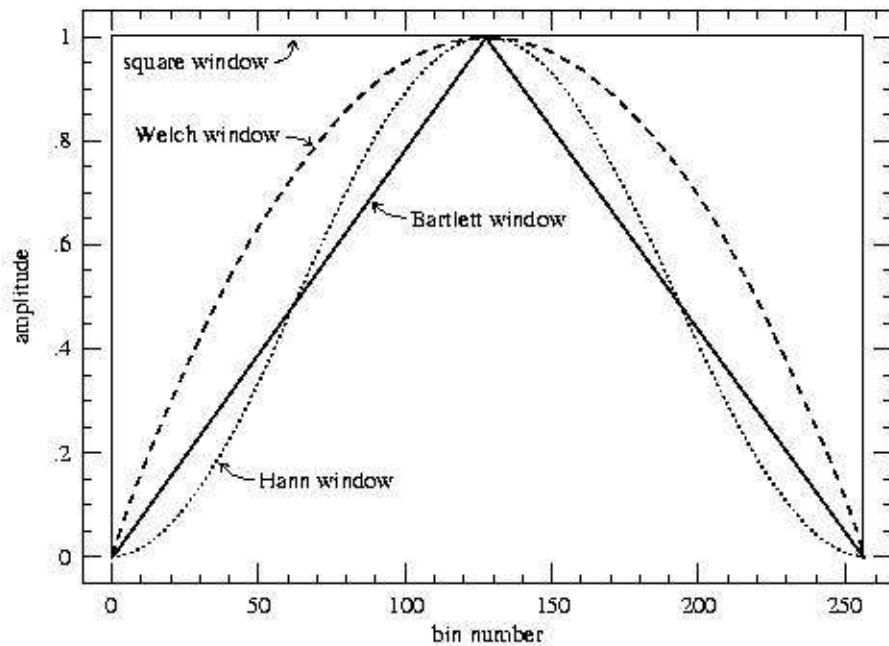
$$P(\nu_j) = \frac{1}{W_{ss}} \left[|S_j|^2 + |S_{N-j}|^2 \right], \quad j = 1, 2, \dots, (N/2 - 1)$$

$$P(\nu_c) = P(\nu_{N/2}) = \frac{1}{W_{ss}} |S_{N/2}|^2$$

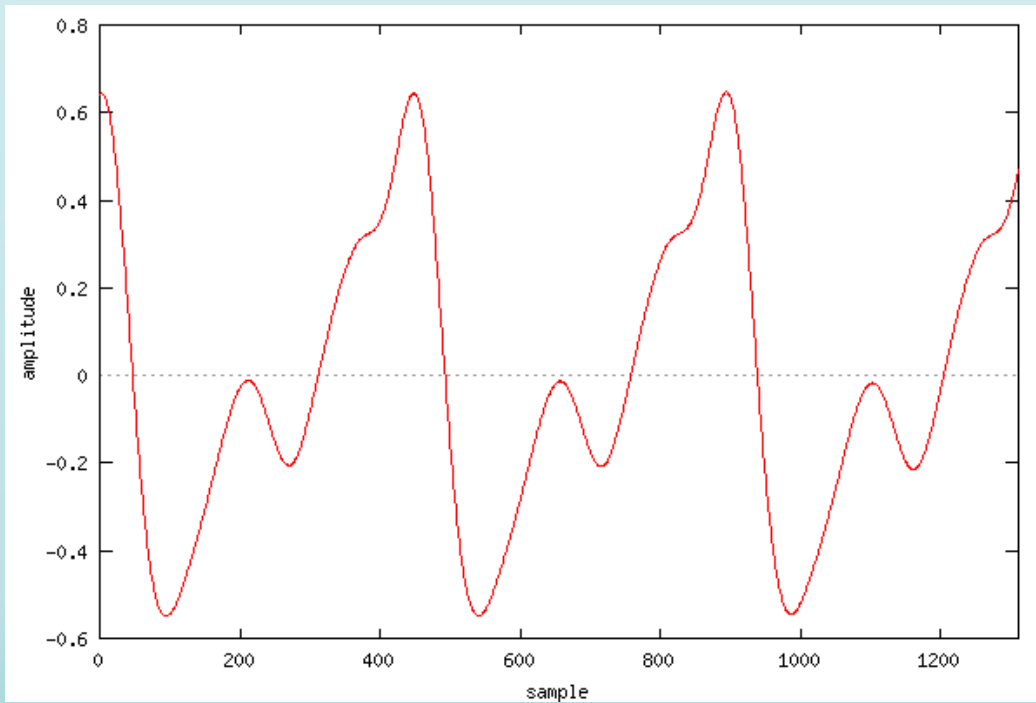
$$W_{ss} = N \sum_{j=0}^{N-1} W_j^2$$

$$\text{donde } W_j = \begin{cases} 1 - \frac{j - N/2}{N/2} & \text{Bartlett} \\ 1/2 [1 - \cos(2\pi j / N)] & \text{Hann} \\ 1 - \left(\frac{j - N/2}{N/2} \right)^2 & \text{Welch} \end{cases}$$

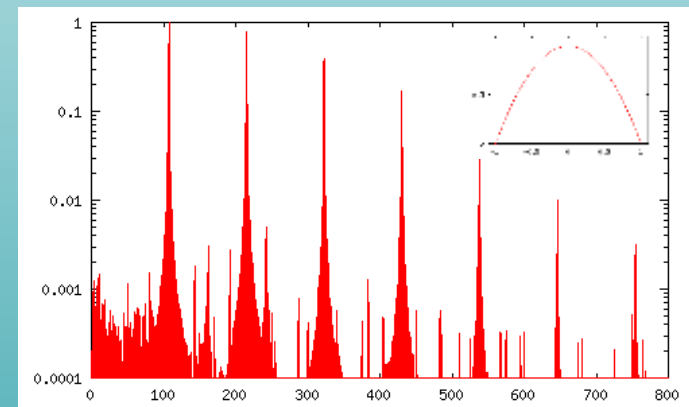
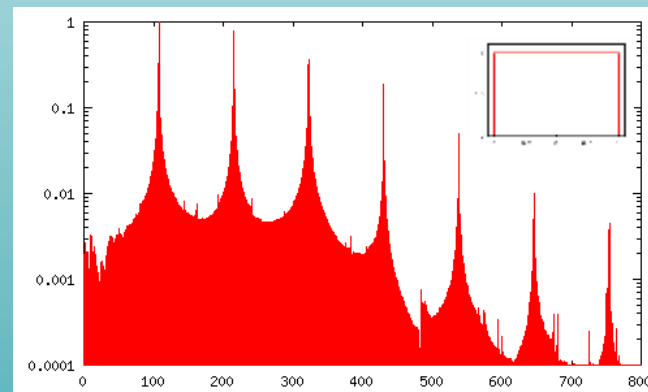
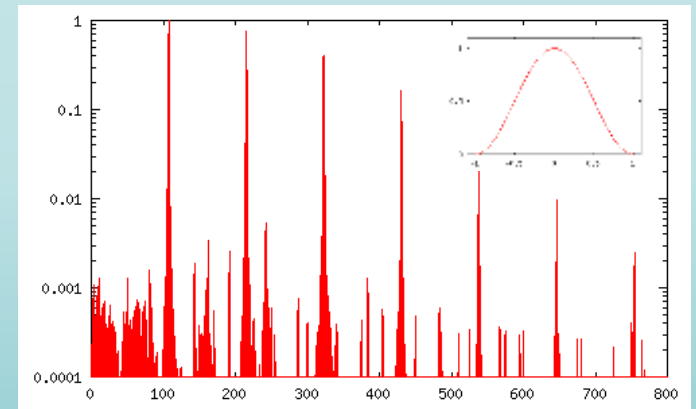
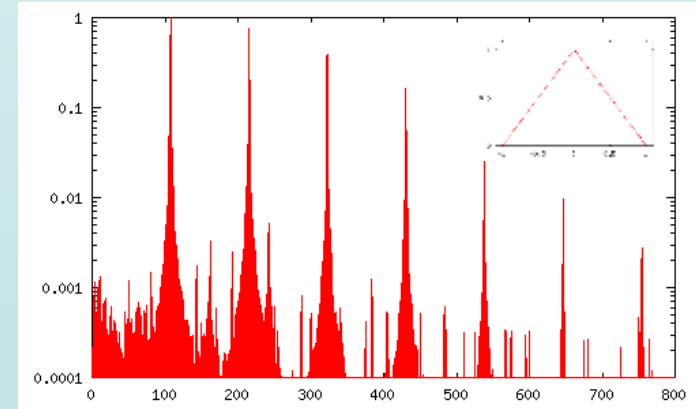
Espectro de potencias: periodograma



Espectro de potencias: periodograma



1



Espectro de potencias: periodograma

♦ Periodograma normalizado de Lomb-Scargle para series evaluadas de forma no equidistante $h_j = h(t_j)$, $j=1, \dots, N$. Se define la media y la variancia de forma estándar (\bar{h}, σ^2)

En función de la frecuencia angular $\omega \equiv 2\pi\nu > 0$:

$$P_N(\omega) \equiv \frac{1}{2\sigma^2} \left\{ \frac{\left[\sum_j (h_j - \bar{h}) \cos \omega(t_j - \tau) \right]^2}{\sum_j \cos^2 \omega(t_j - \tau)} + \frac{\left[\sum_j (h_j - \bar{h}) \sin \omega(t_j - \tau) \right]^2}{\sum_j \sin^2 \omega(t_j - \tau)} \right\}$$

donde τ viene definido por $\tan 2\omega\tau = \frac{\sum_j \sin 2\omega t_j}{\sum_j \cos 2\omega t_j}$

Habitualmente h_j contiene señal y ruido. Para calcular la significancia de un pico, se toman M frecuencias independientes. La probabilidad de que ninguna dé un valor más alto que z es $P(> z) = 1 - (a - e^{-z})^M$, de forma que un valor pequeño tiene una alta significancia. M depende del número de frecuencias definidas por la muestra:

$M \approx N$ si están aproximadamente espaciadas

$M \approx N$ si son aleatorias

$M \neq N$ si los puntos están acumulados en alguna zona de t . Si es así se debe realizar un Monte Carlo para calcular M (ejem. Horne & Baliunas 1986, ApJ, 302, 757).

Espectro de potencias: periodograma

◆ Periodograma normalizado de Lomb-Scargle

