

Capítulo 4

Método para el análisis espectral de los datos

La confiable identificación de los parámetros de ondas gravitatorias en series de datos temporales es problemática y depende en gran parte de las características del método de análisis espectral usado. Hay que considerar que los datos no son equidistantes, son ruidosos y posiblemente no son estacionarios.

4.1 - Dificultades con diversos métodos

Se discutirá brevemente distintas alternativas para el análisis espectral.

El método clásico de transformada discreta de Fourier (DFT) requiere datos equidistantes y tiene baja resolución espectral. El uso de funciones de ventanas y/o suavizados para reducir la aparición de picos espectrales secundarios introduce un grado de libertad que afecta los resultados deteriorando aún más la resolución espectral.

El método de máxima entropía [ver, por ej., Ulrych y Bishop, 1975], aunque tiene mayor resolución en frecuencia, también requiere datos equidistantes, el resultado depende del largo de filtro de predicción elegido y el método no da información de amplitud y fase. En este trabajo se utiliza este método como una forma alternativa para el estudio de las pendientes espectrales.

La transformada de Gabor [1946] si bien puede resolver el problema de la no estacionaridad padece de los mismos inconvenientes que el DFT.

El método de wavelets [por ej. Morlet *et al.*, 1982] también promete resolver la no estacionaridad pero utiliza bases no convencionales difíciles de comparar con los modelos existentes, y especialmente, sus resultados son fuertemente dependientes del desfasaje entre la wavelets y los datos.

El método denominado "Singular Spectral Analysis" (SSA) [ver Vautard y Ghil, 1989] requiere también de datos equidistantes. SSA calcula filtros adaptivos que descomponen los datos en series temporales ortogonales entre sí. Sin embargo, estos filtros son difíciles de interpretar debido a la cantidad y al mutuo solapamiento de sus bandas de transmisión.

El problema de los datos no equidistantes puede resolverse mediante un método de interpolación (lineal, spline cúbico, etc.) aunque esto puede afectar el espectro, especialmente en las altas frecuencias.

En lo que sigue, se describirá un método que resuelve automáticamente la mayoría de las mencionadas dificultades. El método incluye la determinación de fase y amplitud, funciona con datos no equidistantes y tiene alta resolución en frecuencia. No requiere ni ventaneo para eliminar picos secundarios, ni elección de largo de filtros, ni estimación previa de cantidad de picos espectrales. Dado que el método trabaja bien con pocos datos, el problema de la posible no estacionaridad puede resolverse considerando intervalos temporales cortos.

4.2. - El periodograma Lomb-Scargle (LS):

Lomb [1976] y Scargle [1982] publicaron un periodograma generalizado que resuelve los dos problemas principales del método clásico: localiza mucho más precisamente los máximos ("picos") espectrales y funciona con datos no equidistantes. Este método se emplea mucho desde que se publicaron programas que permiten su directa utilización [Press y Teukolsky, 1988; Press y Rybicki, 1989]. Por ser la base del algoritmo que finalmente he implementado conviene describir este método más detalladamente.

Lomb y Scargle llegan, si bien por caminos distintos, a la misma fórmula para su periodograma generalizado

$$P_x(\omega) = \frac{[\sum x_j \cos \omega (t_j - \tau)]^2}{\sum \cos^2 \omega (t_j - \tau)} + \frac{[\sum x_j \sin \omega (t_j - \tau)]^2}{\sum \sin^2 \omega (t_j - \tau)} \quad (4.1)$$

donde ω es la frecuencia angular (una variable continua), x_j son los datos (con su promedio restado) correspondientes a los tiempos t_j (que no tienen la restricción de ser equidistantes), las sumas se hacen para $j = 0 \dots N-1$ (la cantidad de datos), y τ está definido por

$$\tan(2\omega\tau) = \frac{\sum \sin(2\omega t_j)}{\sum \cos(2\omega t_j)} \quad (4.2)$$

El parámetro τ es una sustitución para evitar la aparición de términos cruzados de senos y cosenos [Lomb, 1976].

En el caso equidistante y para las frecuencias armónicas de Fourier el periodograma Lomb-Scargle (LS) coincide con el periodograma clásico.

Lomb llega a su resultado haciendo un ajuste por el método de cuadrados mínimos, donde los datos son expresados como

$$x_j = a \cos \omega (t_j - \tau) + b \sin \omega (t_j - \tau) + \varepsilon_j \quad (4.3)$$

donde ε_j son los residuos después de haber restado el senoide de los datos originales.

La diferencia entre la suma de los cuadrados de los x_j y las de los ε_j es igual a $P_x(\omega)$. Por lo tanto, al restar el senoide con la frecuencia correspondiente al máximo de $P_x(\omega)$ se logra la máxima reducción en la suma de los cuadrados. El correspondiente senoide representa el mejor ajuste a los datos, respecto a este criterio. La posición de este máximo puede ser determinada con una precisión no limitada por el espaciado entre las frecuencias armónicas de los periodogramas clásicos.

Como resultado adicional, Scargle [1982] deduce la siguiente expresión para la probabilidad de falsa alarma

$$P_r = 1 - \{1 - \exp(-P_x(\omega)/2\sigma_0^2)\}^M \quad (4.4)$$

donde σ_0^2 es la varianza total de los datos y M es el número de frecuencias independientes (igual a N , en el caso equidistante). $M = N$ es una buena aproximación para varios casos de datos no equidistantes [Horne y Baliunas, 1986].

P_r representa la probabilidad que datos puramente ruidosos generen un pico con potencia no menor a P_x . Aceptando la existencia de al menos una oscilación, esta fórmula deja de tener validez [Horne y Baliunas, 1986].

4.3 - Modificaciones al periodograma LS:

El periodograma LS, a pesar de su sencillez, permite resolver en la práctica situaciones que otros métodos no pueden resolver, ya que, como se ha mencionado en la sección anterior, tiene buena resolución en frecuencia y no requiere de datos equidistantes. Sin embargo, los trabajos citados de Lomb y Scargle se limitan a detectar sólo la frecuencia de una única componente espectral. Esta limitación excluye la mayor parte de los casos en la práctica, donde los fenómenos tienen múltiples oscilaciones, como es también el caso de las ondas atmosféricas. En lo que sigue, se describirá como se extienden las ideas originales de Lomb y Scargle para construir un algoritmo capaz de superar esta limitación.

El objetivo del método consiste en calcular un conjunto mínimo de picos espectrales (dados por sus frecuencias, amplitudes y fases) que ajusten bien los datos. Este es un concepto distinto al del espectro Fourier, donde los datos son representados *exactamente* por el conjunto de $N/2$ componentes espectrales en las frecuencias armónicas. Si se pretende representar con DFT las principales oscilaciones, descartando las componentes armónicas de Fourier de menor amplitud, la aproximación obtenida, en general, no es óptima.

La extensión al periodograma LS está caracterizada por tres rasgos adicionales:

- i) Determinación de amplitud y fase.
- ii) Periodograma modificado para la máxima reducción de varianza, y con ajuste del término constante.
- iii) Método deconvolutivo por iteraciones para el cálculo de los picos espectrales.

Nótese que lo que ii) implica una generalización al periodograma DFT, alternativa al de Lomb-Scargle. En la práctica, esto tiene sólo un efecto importante en las

bajas frecuencias y, generalmente, ambos periodogramas dan casi idénticos resultados para frecuencias medias y altas.

i) Determinación de amplitud y fase:

Lomb y Scargle no dan las expresiones para calcular la amplitud y la fase de estos picos, aunque del paper de Lomb [1976] es fácil derivar las expresiones para los coeficientes a y b de (4.3),

$$a = \sum x_j \cos \omega (t_j - \tau) / \sum \cos^2 \omega (t_j - \tau) \quad (4.5)$$

$$b = \sum x_j \sin \omega (t_j - \tau) / \sum \sin^2 \omega (t_j - \tau)$$

En el periodograma modificado que se presenta a continuación, se modifican las expresiones de a y b (ec. 4.11), mientras que las expresiones siguientes (4.6), (4.7) y (4.8) siguen siendo válidas.

Se puede expresar la ec. (4.3) en términos de amplitud A y fase ϕ como

$$x_j = A \cos(\omega t_j + \phi) \quad (4.6)$$

con

$$A = \sqrt{a^2 + b^2} \quad (4.7)$$

$$\phi = -\omega \tau - \arctan(b/a) \quad (4.8)$$

y $\arctan(b/a)$ debe extenderse a los cuatro cuadrantes en la forma habitual.

ii) Periodograma modificado para la máxima reducción de varianza, y con ajuste del término constante:

Como se ha mencionado, el periodograma LS representa la máxima reducción en la suma de los cuadrados que se logra, al restar de los datos un senoide de frecuencia ω . Dado que el senoide restado, en general, no tiene promedio cero, el promedio de los datos residuales tampoco es cero, de lo que se deduce que la suma de los cuadrados de los residuos ϵ_j no es igual a su varianza. Es conveniente modificar la definición del periodograma para que represente la máxima reducción en la varianza (y no

en la suma de los cuadrados). El pico máximo en este periodograma corresponde a la frecuencia que deja menor varianza residual.

El periodograma clásico y el de Lomb implican que el término constante coincide con el promedio de los datos. Si bien esto es automáticamente garantizado para el caso de datos equidistantes y frecuencias armónicas, es, en general, una restricción innecesaria. Por eso, se puede mejorar el periodograma incluyendo el término constante en el ajuste. Esto es particularmente importante para las bajas frecuencias donde los datos cubren sólo una sección del senoide. Es en este rango espectral donde el periodograma modificado estima mejor que el de LS las componentes espectrales. En el contexto de este trabajo, la mejora es significativa, pues las componentes espectrales de gran período, dadas por la marea, suelen ser las principales en los datos nocturnos.

La modificación consiste entonces en reemplazar (4.3) por

$$x_j = a \cos \omega (t_j - \tau) + b \sin \omega (t_j - \tau) + K + \epsilon_j \quad (4.9)$$

donde la única diferencia es la introducción del término constante K .

Las expresiones del periodograma modificado son (ver Apéndice A)

$$(N-1)P_x(\omega) = a^2 [\sum (\cos \omega (t_j - \tau))^2 - (\sum \cos \omega (t_j - \tau))^2 / N] + b^2 [\sum (\sin \omega (t_j - \tau))^2 - (\sum \sin \omega (t_j - \tau))^2 / N] - 2ab \sum \cos \omega (t_j - \tau) \sum \sin \omega (t_j - \tau) / N \quad (4.10)$$

Los coeficientes a y b están dados por

$$a = A/D \quad \text{y} \quad b = B/D \quad (4.11)$$

donde

$$A = \sum x_j \cos \omega (t_j - \tau) [\sum (\sin \omega (t_j - \tau))^2 - (\sum \sin \omega (t_j - \tau))^2 / N] + \sum x_j \sin \omega (t_j - \tau) \sum \cos \omega (t_j - \tau) \sum \sin \omega (t_j - \tau) / N$$

$$B = \sum x_j \sin \omega (t_j - \tau) [\sum (\cos \omega (t_j - \tau))^2 - (\sum \cos \omega (t_j - \tau))^2 / N] + \sum x_j \cos \omega (t_j - \tau) \sum \cos \omega (t_j - \tau) \sum \sin \omega (t_j - \tau) / N$$

$$D = \sum (\cos \omega (t_j - \tau))^2 \sum (\sin \omega (t_j - \tau))^2 - \sum (\sin \omega (t_j - \tau))^2 (\sum \cos \omega (t_j - \tau))^2 / N - \sum (\cos \omega (t_j - \tau))^2 (\sum \sin \omega (t_j - \tau))^2 / N$$

El término K está dado por

$$K = -a \sum \cos \omega (t_j - \tau) / N - b \sum \sin \omega (t_j - \tau) / N \quad (4.12)$$

(El factor $N-1$ que aparece en la expresión (4.10) es debido a la definición de $P_x'(\omega)$ como la reducción de varianza. Con este factor, este periodograma es llevado a la misma escala que el de LS)

Se demuestra en el Apéndice A que para el caso de datos equidistantes, el valor de $(N-1)P_x'(\omega)$ coincide, en las frecuencias armónicas de Fourier, con el periodograma clásico.

iii) Método deconvolutivo por iteraciones para el cálculo de los picos espectrales.

Al igual que en los otros métodos, el periodograma LS (y el modificado) no están exentos de los conocidos problemas causados por la convolución del espectro verdadero con la función de ventana. Es decir, cada línea espectral se convierte en un sistema complejo de picos, análogo a la figura de difracción, causada por la distribución temporal de los datos. Para recuperar el espectro verdadero, lo ideal sería, teóricamente, calcular la deconvolución exacta, pero los algoritmos existentes requieren del conocimiento de la función de ventana y, en presencia de ruido, son muy inestables. La forma que frecuentemente se usa para reducir este problema, es el preprocesamiento de los datos mediante ventanas, filtros digitales, etc.

Una forma aproximada de deconvolución es calcular los picos espectrales iterativamente. En cada iteración se calcula los parámetros del pico máximo del periodograma. Luego, se resta a los datos x_j el senoide correspondiente al pico calculado y estos residuos constituyen los datos de la siguiente iteración, con los cuales se calcula un nuevo periodograma. Se continua las iteraciones mientras los picos tengan amplitud significativa.

La idea de este cálculo espectral iterativo es evitar el efecto de la interferencia de los picos más importantes sobre los siguientes. Al restar el senoide, no solo desaparece el pico en la siguiente iteración, sino la totalidad de la figura de difracción correspondiente, es decir, también los picos secundarios.

En un ejemplo se ilustra este método. Se simula un conjunto de 48 datos con una suma de ocho cosinoides evaluados en tiempos con un patrón similar al de nuestras mediciones. Los datos cubren un intervalo de 6053s, con un espaciado mínimo de 109s y dejando algunos "huecos" para simular una situación con pérdida de datos. Los períodos (P), amplitudes (A) fases (φ) de los cosinoides utilizados en la simulación y los calculados (P' , A' , φ') son mostrados en la tabla 4.1. También se indican los errores relativos de las amplitudes, frecuencias y los errores absolutos en las fases de los picos encontrados. El error de fase ($\Delta\varphi$) se evalúa en el centro del intervalo temporal (t_c) para minimizar la contribución trivial producida por el error de la frecuencia, o sea

$$\Delta\varphi = \varphi' - \varphi + (\omega' - \omega)t_c \quad (4.13)$$

(esta expresión se utiliza también para evaluar la diferencia de fase entre picos coincidentes, ver capítulos 6 y 7).

Primero se observa que todas las componentes espectrales son detectadas en las primeras 8 iteraciones. Los errores en frecuencia, amplitud y fase son pequeños. El máximo error relativo en frecuencia, 5%, se produce en el pico de 4000s pero en los restantes picos hay un error del 1% ó menos. En amplitud los errores máximos son del 17% y 14%, pero en el resto oscilan alrededor del 5%. Las fases tienen errores menores a 9°.

TABLA 4.1: Periodos, amplitudes y fases (P, A, ϕ) utilizados para construir un conjunto de 48 datos no equidistantes, y los resultados del método espectral (P', A', ϕ'). También se dan los errores relativos de frecuencia ($\Delta f/f$) y amplitud ($\Delta A/A$) y los errores de fase ($\Delta \phi$) se calculan en el centro del intervalo temporal según ec. (4.13).

iter.	Utilizados			Calculados			Errores		
	$P[s]$	A	$\phi [^\circ]$	$P'[s]$	A'	$\phi' [^\circ]$	$\Delta f/f$	$\Delta A/A$	$\Delta \phi [^\circ]$
1	280	25	72	279	24.48	62.6	0.003	-0.0207	2.24
2	1200	25	0	1211	24.48	9.5	-0.009	-0.0208	-0.38
3	550	18	36	552	21.10	49.7	-0.004	0.1723	6.29
4	4000	10	-90	4216	9.93	-68.4	-0.051	-0.0070	5.63
5	250	9	90	250	8.39	85.6	0.001	-0.0680	-1.86
6	300	8	-45	298	7.59	-73.5	0.007	-0.0513	-2.32
7	630	5	90	632	5.29	90.9	-0.003	0.0589	-8.93
8	320	5	0	324	4.28	48.4	-0.012	-0.1442	8.62

En la figura 4.1 se muestra el periodograma normalizado por la varianza de los datos, calculado en las distintas iteraciones. La curva de arriba representa el periodograma inicial, y hacia abajo se muestran los periodogramas de las siguientes iteraciones, cada una corrida en escala para su mejor visualización e identificada con su número consecutivo. Se muestran sombreados los picos que se restan en la siguiente iteración y con líneas rayadas verticales se indican los correspondientes períodos calculados. Se indican con líneas rayadas horizontales los niveles de probabilidad de falsa alarma del 10% y 50% para la primera iteración.

El principal aspecto a resaltar en esta figura es la comparación del periodograma de la primera iteración, con los picos encontrados por el método iterativo. En el periodograma inicial hay varios picos secundarios que son efecto de la función de ventana y que mayoritariamente son eliminados en las tres primeras iteraciones. Como ejemplos se ven los inmediatos vecinos del pico de 1200s, el pico a la derecha del de 550s y varios picos a la derecha del de 280s. Algunos de estos picos secundarios tienen amplitudes mayores a los picos reales calculados a partir de la cuarta iteración. Dos de los picos (300s y 4000s) son apenas detectables en el periodograma ini-

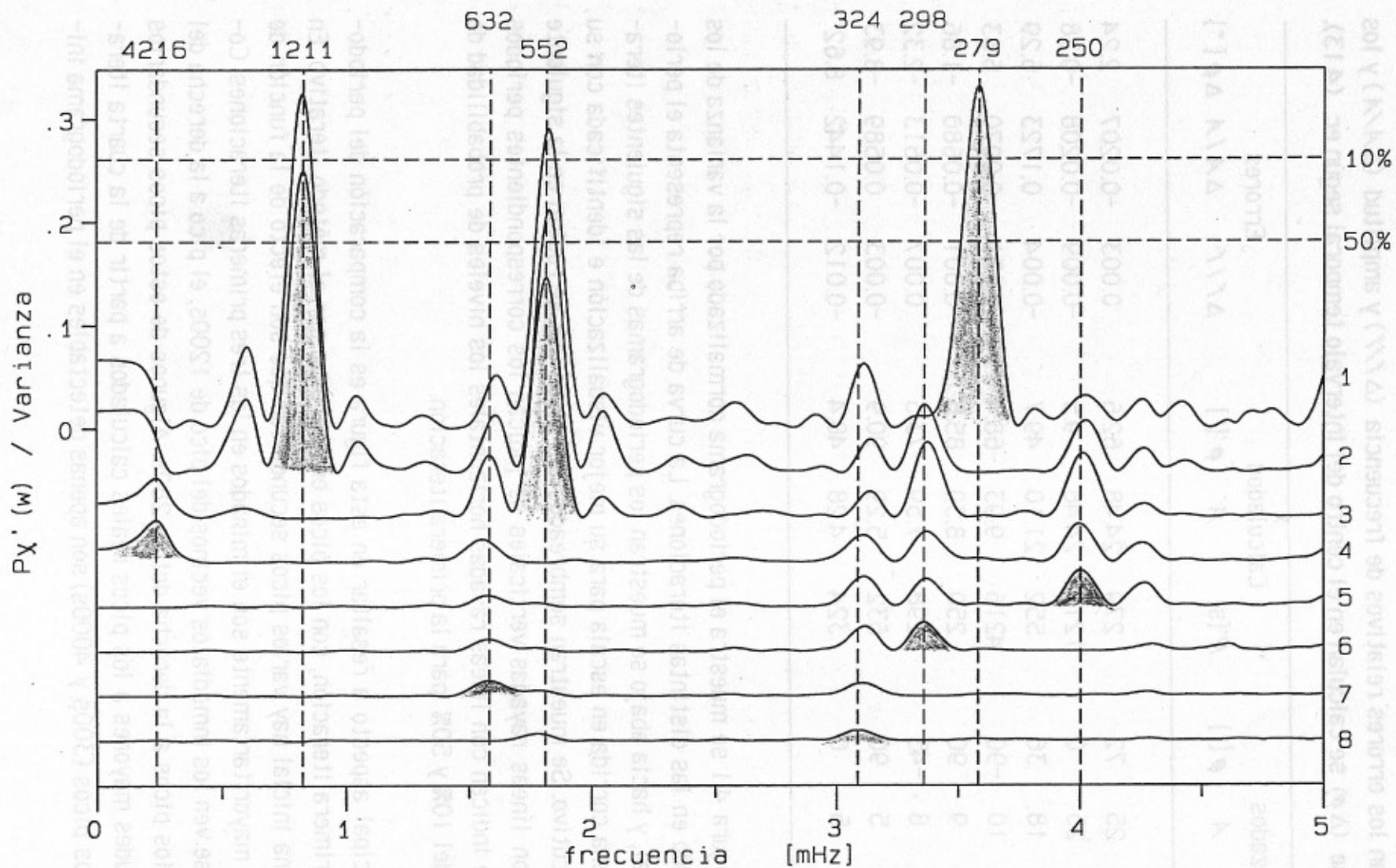


Figura 4.1: Ejemplo del desarrollo iterativo para el cálculo espectral

cial y contrariamente, otros dos picos (630s y 320s), aparecen aproximadamente en su verdadera frecuencia pero con muy exagerada amplitud.

En este ejemplo, tres picos superan la línea de probabilidad de falsa alarma del 10% y ninguno de los restantes alcanza la línea del 50%. Esto demuestra las limitaciones de este tipo de test. Además, como fue mencionado, la fórmula de Scargle no es aplicable para el caso de más de una componente espectral.

Este método iterativo no es ideal dado que no toma en cuenta la interferencia de los picos menores sobre los mayores. Sin embargo, como se puede ver con datos simulados, los resultados, en general, reproducen bien los parámetros de las componentes espectrales.

4.4 - Análisis de coincidencias:

El método descrito es utilizado para estimar las frecuencias, amplitudes y fases de las oscilaciones a partir de los datos de temperaturas e intensidades. Por supuesto, ésto sólo no alcanza para determinar si los parámetros calculados corresponden al paso de ondas. Son necesarios criterios adicionales para poder confirmar o excluir la existencia de una onda.

Cada onda en temperatura genera una onda en intensidad de igual período, al menos en el límite lineal. Entonces, las coincidencias en frecuencia entre picos de los espectros de intensidades y temperaturas de una banda (OH u O₂) pueden usarse como un criterio para la detección de las ondas.

Es necesario establecer el máximo intervalo de frecuencia para definir una coincidencia.

En cada iteración se calcula la incertidumbre de frecuencia, con la expresión de Kovacs [1981] adaptada al método iterativo

$$\delta f = \frac{3\sigma_R}{4A\sqrt{N}\Delta T} \quad (4.14)$$

donde σ_R^2 es la varianza de ruido (que aquí es reemplazada convenientemente por la varianza residual), ΔT es el intervalo temporal entre el primer y último dato, y A es la amplitud del pico. Si bien esta expresión es válida para datos equidistantes, lo es aproximadamente para datos no equidistantes.

Para cuantificar el grado de una coincidencia entre un pico en la frecuencia f_1 , de un espectro, con otro en la frecuencia f_2 del otro espectro, se define el siguiente factor de tolerancia

$$Q = \frac{|f_2 - f_1|}{\sqrt{\delta f_1^2 + \delta f_2^2}} \quad (4.15)$$

Con $Q=1$ la distancia entre picos es un poco menor al error de Kovacs. Se aceptan como coincidencias aquellas, cuyo Q no supera un valor máximo Q_{\max} . (Finalmente se usó una variante modificada de (4.15) que incluye como término adicional en el denominador el paso en frecuencia utilizado en el análisis espectral. Esto es necesario en el caso de que el error de Kovacs sea pequeño como ocurre con la marea semidiurna, pero en general su efecto es despreciable).

Para ver si la cantidad de coincidencias encontradas es estadísticamente significativa, se calcula la esperanza de la cantidad de coincidencias cuando son debidas a datos puramente ruidosos. En este caso, la probabilidad de encontrar un pico es uniforme en todo el rango espectral entre 0 y f_{\max} donde f_{\max} es la máxima frecuencia calculada. Dado los N_1 picos del primer espectro con sus errores de frecuencia δf_{1k} con $k = 1 \dots N_1$, y los N_2 picos del espectro a comparar, con sus errores de frecuencia δf_{2i} con $i = 1 \dots N_2$. Entonces, para cada pico k del primer espectro, la probabilidad de coincidir con el pico i del otro espectro es

$$p_{ki} = 2Q_{\max} \frac{\sqrt{\delta f_{1k}^2 + \delta f_{2i}^2}}{f_{\max}}$$

La esperanza del número de coincidencias aleatorias entre los dos espectros es simplemente

$$E = \sum_{i,k} p_{ki} = (2Q_{\max} / f_{\max}) \sum_{k=1}^{N_1} \sum_{i=1}^{N_2} \sqrt{\delta f_{1k}^2 + \delta f_{2i}^2} \quad (4.16)$$

Se ha verificado empíricamente la validez de (4.16) calculando un gran número de espectros de datos simulados con una distribución estadística uniforme.

Por último, cabe mencionar que una forma alternativa al uso de las coincidencias espectrales son los espectros cruzados. Sin embargo, los algoritmos clásicos de espectros cruzados tienen los mismos problemas que el FFT. Sería necesario diseñar un algoritmo para datos no equidistantes que resuelva el problema de los picos secundarios (por ejemplo, en forma iterativa).