

UNIVERSIDAD AUTÓNOMA DE MADRID

Escuela Politécnica Superior



Doble Grado en Ingeniería Informática y Matemáticas

TRABAJO DE FIN DE GRADO

ANÁLISIS Y PROCESADO DE SEÑALES MULTIDIMENSIONALES
PARA DETECCIÓN DE EVENTOS EN CONDICIONES DESFAVORABLES

Jorge Martín Pérez

Tutor: Daniel Ramos Castro

Mayo 2016

ANÁLISIS Y PROCESADO DE SEÑALES
MULTIDIMENSIONALES PARA DETECCIÓN DE EVENTOS
EN CONDICIONES DESFAVORABLES



Autor: Jorge Martín Pérez
Tutor: Daniel Ramos Castro

ATVS
Escuela Politécnica Superior
Universidad Autónoma de Madrid

Mayo 2016

RESUMEN

Resumen. Este trabajo trata la detección de irregularidades en un proceso de control de calidad de componentes de un reactor nuclear. Dichos procesos tienen como objetivo determinar si una componente está o no fuera de especificación. En el proceso de control de calidad se toman múltiples mediciones que llevan asociadas distintas señales, y de todas ellas existe una que refleja mejor que el resto las irregularidades, dicha señal es la que mide la densidad de las componentes.

Sin embargo la señal de densidad tiene un problema, y es que lleva asociado un ruido que interfiere en el algoritmo existente para la detección de picos asociados con las irregularidades de las componentes del reactor. El modelo de ruido en las señales se corresponde con un modelo aditivo, donde la señal asociada al ruido aparece sumada a la señal original en todo momento, y debido a que no se conoce la caracterización de dicho ruido, las labores de eliminación de ruido no se pueden atener a ningún tipo de conocimiento sobre este.

Para reducir el ruido en las señales de densidad, se ha hecho uso de filtros de Chebyshev y de la sustracción espectral (método ideado originalmente para la limpieza de ruido en señales de voz). Entre estos dos métodos, el último es con el que se han obtenido mejores resultados, pues consigue mejorar sustancialmente la presencia de señal frente al ruido, así como la detección de los picos existentes en las señales.

Todo esto requiere de cierto conocimiento del mundo del tratamiento de señales, por ello se comienza exponiendo la serie de Fourier y su convergencia, y cómo de ella se deriva la transformada de Fourier. Así mismo también se explica cómo interpretar el dominio frecuencial que se obtiene con la transformada de Fourier, y cómo obtener su versión discreta.

Palabras clave. serie de Fourier, transformada de Fourier, DFT, FFT, filtro de Chebyshev, sustracción espectral, enventanado, frecuencia de corte

ABSTRACT

Abstract. This paper deals with the detection of irregularities in the process of quality control of components of a nuclear reactor. These processes aim to determine whether a component is or is not out of specification. In the quality control process multiple measurements are taken, each measure has an associated signals, and among them there is one that reflects the irregularities better than the rest, that signal is the one measuring the density of the components.

However the density signal has a problem, and it's that it has an associated noise that interferes with the existing algorithm for detecting peaks associated with irregularities of the reactor components. The noise model in the signals corresponds to an additive model, where the signal associated with noise appears added to the original signal at all times, and because the noise characterization is not known, the work removal noise can not stick to any knowledge about this.

To reduce noise in density signals, Chebyshev filters and spectral subtraction have been used (the last one was originally designed for cleaning noise in voice signals). Between these two methods, the latter is the one that have given better results, as it improves substantially the presence of signal above noise, as well as the detection of existing peaks in the signals.

All this requires some knowledge of the world of signal processing, therefore the paper begins by exposing the Fourier series and their convergence, and how the Fourier transform is derived. Likewise, it also explains how to interpret the frequency domain obtained with the Fourier transform, and how to get it's discrete version.

Keywords. Fourier series, Fourier transform, DFT, FFT, Chebyshev filter, spectral subtraction, windowing, cutoff frequency

AGRADECIMIENTOS

Muchas gracias a mi tutor Daniel por el constante ofrecimiento para ayudarme, por su paciencia conmigo, por responder mis correos casi al instante, por enseñarme a cómo interpretar los resultados a lo largo de estos meses, por ponerse siempre en mi lugar y explicarme con tanta claridad todo lo que me ha ido enseñando desde que comencé el trabajo. Y sobre todo, muchas gracias por el ambiente distendido y divertido con el que siempre me ha dispensado.

Gracias a mi familia por darme esos empujones que siempre son tan necesarios cuando las cosas se tuercen, y por el incesable apoyo que siempre me han brindado.

Gracias a mis amigos de Alorcón por sacarme de casa esos fines de semana en los que me veo tan agobiado.

Gracias a los compañeros de universidad que me han echado una mano, y a los que me han sacado una sonrisa.

Gracias a mis amigos de Jarafuel por hacerme ver en más de una ocasión que los estudios no son ni mucho menos todo en la vida.

Y finalmente, gracias a la universidad pública y a todos sus profesores, porque gracias a ellos, a los buenos y a los no tan buenos, he aprendido más que nunca.

ÍNDICE GENERAL

Índice general	VI
Índice de tablas	IX
Índice de figuras	XI
1 Introducción	1
1.1 Motivación	1
1.2 Objetivos	1
1.3 Estructura de la memoria	2
2 Estado del arte	3
2.1 Conceptos fundamentales en el tratamiento de señales	3
2.1.1 Series de Fourier	3
2.1.2 Autofunciones ortogonales y coeficientes de Fourier	3
2.1.3 Convergencia uniforme de la serie de Fourier	5
2.1.4 Transformada de Fourier de una señal aperiódica	6
2.1.5 Interpretación de la transformada de Fourier	7
2.1.6 Transformada de Fourier de una señal discreta - DTFT	8
2.1.7 Transformada Discreta de Fourier - DFT	8
2.2 Filtro de Chebyshev de tipo 2	9
2.3 Sustracción espectral	11
2.3.1 Modelo de ruido aditivo	11
2.3.2 Error espectral	12
2.3.3 Promediado de la magnitud	12
2.3.4 Rectificación "a mitad de camino"	12
2.3.5 Reducción de ruido residual	13
2.4 Representaciones en el espectro de frecuencia	13
3 Desarrollo e implementación	15
3.1 Datos disponibles	15
3.2 Implementación de la sustracción espectral	18
3.2.1 Enventanado de las señales	18
3.2.2 Estimación del ruido	18
3.2.3 Modificación de la sustracción y pasos implementados	19

3.2.4	Reconstrucción de la señal	20
3.3	Elección de frecuencia de corte en filtro de Chebyshev	20
3.4	Localización de picos	21
4	Resultados	23
4.1	Sustracción espectral	23
4.1.1	Efecto del tamaño de ventana en la sustracción	24
4.2	Comparación de medida de SNR	27
4.3	Separación del ruido	27
4.4	Detección de picos tras el filtrado de la señal	28
5	Conclusiones	31
5.1	Trabajo futuro	32
	Glosario	33
	Bibliografía	35
	Anexos	37
A	Serie de Fourier de señales periódicas discretas	37
B	Muestreo de la DTFT de una señal aperiódica	38
C	Cálculo de la FFT	39
D	Otros pasos utilizados en la versión original de la sustracción espectral . .	41
D.1	Atenuación de la señal cuando no hay voz	41
D.2	Actualización de la estimación del ruido	41

ÍNDICE DE TABLAS

4.1	<i>SNR</i> promedio en sustracción espectral.	27
4.2	<i>SNR</i> promedio en el filtrado de Chebyshev	27
4.3	Coste promedio y umbrales óptimos, detección de picos sustracción espectral.	30
4.4	Coste promedio en detección de picos tras filtro de Chebyshev.	30

ÍNDICE DE FIGURAS

2.1	Señal cuadrada	6
2.2	Envolvente de los coeficientes de Fourier.	7
2.3	Transformada de Fourier de un seno.	8
2.4	Polinomios de Chebyshev de distinto orden.	10
2.5	Polos de la función de ganancia 2.26 de un filtro de Chebyshev de tipo 2 y sexto orden,	10
2.6	Ganancia de un filtro de Chebyshev de tipo 2, con frecuencia de corte $\omega_0 = \pi/2$, y caída de 40dB.	11
2.7	Representación de la "rectificación" de \hat{S} extraída de [2]	13
2.9	Espectrograma de una señal de densidad del control de calidad de una componente.	14
3.2	Visualización de intervalos de picos de densidad en las 7 señales	16
3.3	Zona activa de una de las 50 señales sin irregularidades.	17
3.4	Picos en las zonas activas de una de las 22 señales con irregularidades.	17
3.5	Proceso de enventanado de una señal con dos ventanas de Hanning solapadas con factor 0.5	18
3.6	Efectos en el offset al restar el módulo del ruido estimado en la frecuencia $\omega = 0$	19
3.7	Almacenamiento y reconstrucción de la DFT de $N = 8$ muestras de una señal real enventanada.	20
3.8	Suma de las dos componentes de una señal descompuesta en dos ventanas.	21
3.9	Transformada de Fourier de un triángulo (reescalada).	21
3.10	Ejemplo de detección de picos sobre el umbral calculado.	22
4.1	Sustracción espectral en la señal de densidad con $W = 256$	23
4.2	Espectrograma de la sustracción espectral en la densidad con $W = 256$	24
4.3	Sustracción espectral sobre la señal de densidad con $W = 32, 128$	25
4.4	Sustracción espectral en frecuencia	26
4.5	Histograma de ruido y señal antes de la sustracción espectral	28
4.6	Curva DET y tasa EER tras limpieza de la señal de densidad	29
.1	8 raíces de la unidad.	38
.2	Tren de impulsos con periodo $N = 3$	39
.3	Señal antes y después del muestreo de su DTFT.	39
.4	Diagrama del funcionamiento de la FFT para una señal con $N = 8$ muestras.	40

INTRODUCCIÓN

1.1 Motivación

Al alejarse de las situaciones ideales de estudio, las señales reales van acompañadas de ruido. La eliminación de dicho ruido es una tarea complicada, y existen varias situaciones que se pueden presentar a la hora de llevar a cabo un estudio de las señales afectadas. Una de ellas es en la que se conoce con exactitud la señal de ruido, y otra es en la que se desconoce dicha modelización (este caso es el abordado en el trabajo). Esta última situación es la más difícil de abordar, ya que la eliminación del ruido no puede basarse en ningún tipo de conocimiento sobre este.

Para llevar a cabo la limpieza de ruido es necesario el uso de matemáticas, pues sobre estas se posan las distintas técnicas y los fundamentos necesarios para afrontar este tipo de problemas. En particular este ámbito del tratamiento de señales hace un uso intenso del análisis de Fourier para la interpretación de las señales en el dominio frecuencial. Este ámbito tiene como punto de partida las series de Fourier, las cuales se estudian con asiduidad en la carrera de matemáticas al resolver ecuaciones en derivadas parciales. De dichas series se deriva lo que se conoce como la transformada de Fourier, que resulta ser una pieza imprescindible para llevar a cabo el estudio de señales, ya que permite obtener una interpretación de la señal prestando atención al dominio frecuencial.

De estos fundamentos se estudian diversas formas de realizar la limpieza de una señal, bien mediante el uso de filtros paso-bajo, o mediante la sustracción espectral. Y en este trabajo se aplican todas estas técnicas al caso de inspección de irregularidades en las componentes de un reactor nuclear. Dichas señales llevan un ruido asociado que dificulta la detección de las irregularidades, y su eliminación facilita la labor de detección.

1.2 Objetivos

Con este trabajo se pretende aprender los fundamentos del campo del tratamiento de señales. Dichos fundamentos conllevan el estudio de ciertos aspectos matemáticos, y cómo implementar las operaciones en un ordenador.

Estos fundamentos vienen motivados por la necesidad de asentar ciertos conocimientos, que son necesarios a la hora de comprender distintos acercamientos a la reducción del ruido presente en una señal, algo fundamental en el desarrollo del trabajo.

Una vez comprendidas las técnicas de reducción de ruido, el objetivo es mejorar el proceso de detección de irregularidades presentes en las señales que se obtienen en un control de calidad de componentes de un reactor nuclear.

1.3 *Estructura de la memoria*

La memoria comienza exponiendo en el estado del arte los fundamentos teóricos y matemáticos del procesamiento de señales. Se comienza introduciendo las series y transformada de Fourier, hasta llegar al cálculo de la DFT, pasando por la interpretación del espectro de frecuencia. Acto seguido se presenta el diseño de filtros paso-bajo usando los polinomios de Chebyshev, y a continuación se expone el método de reducción de ruido entorno al que gira este proyecto, la sustracción espectral. En la sección que este método tiene dedicada, se expone el modelo de ruido aditivo y cómo se lleva a cabo la limpieza de la señal mediante la eliminación del módulo del ruido estimado.

Una vez presentado el estado del arte, se continua describiendo la implementación llevada a cabo a lo largo del trabajo.

Por último se cierra el cuerpo del proyecto con los resultados obtenidos tras aplicar la limpieza a las señales del proceso de control de calidad, a la vez que se expone cómo se ha realizado la medición que permite sopesar la mejora, o el empeoramiento de la detección de dichas irregularidades.

ESTADO DEL ARTE

2.1 Conceptos fundamentales en el tratamiento de señales

En esta sección se explican y presentan conocimientos esenciales dentro del mundo del tratamiento de señales. Todo ello desde un enfoque con cierto sesgo matemático, pero sin perder de vista la interpretación de las herramientas.

2.1.1 Series de Fourier

Las series de Fourier nos permiten representar ciertas funciones como una suma infinita de senos y/o cosenos. Una función $\Phi(x)$ se puede expresar dentro del intervalo $(0, l)$, con el desarrollo en senos de la serie de Fourier:

$$\Phi(x) = \sum_{n=1}^{\infty} A_n \sin \frac{n\pi x}{l} \quad (2.1)$$

Pero también existe la conocida como expansión completa de la serie de Fourier en el intervalo $(-l, l)$, la cual tiene la siguiente expresión:

$$\Phi(x) = \frac{1}{2}A_0 + \sum_{n=1}^{\infty} \left(A_n \cos \frac{n\pi x}{l} + B_n \sin \frac{n\pi x}{l} \right) \quad (2.2)$$

Esta última versión de la serie de Fourier da pie a la versión compleja que tanto se utiliza en el mundo del tratamiento de señales. Recordando las fórmulas de DeMoivre, las cuales permiten expresar senos y cosenos de la siguiente forma:

$$\sin x = \frac{e^{ix} - e^{-ix}}{2i} \quad \cos x = \frac{e^{ix} + e^{-ix}}{2}$$

se puede reescribir 2.2 como:

$$\Phi(x) = \sum_{n=-\infty}^{\infty} c_n e^{in\pi x/l} \quad (2.3)$$

2.1.2 Autofunciones ortogonales y coeficientes de Fourier

Considérense dos funciones f, g bien definidas en el intervalo $a \leq x \leq b$, el producto interior de estas dos funciones viene definido por la integral:

$$(f, g) = \int_a^b f(x) \overline{g(x)} dx \quad (2.4)$$

En este contexto en el que se tratan las funciones “como si fuesen vectores”, dos funciones son ortogonales en caso de que cumplan $(f, g) = 0$. Del mismo modo se pueden ver los operadores “como si fuesen matrices”, de hecho en este apartado se tomará como operador: $A = -d/dx^2$, es decir, la segunda derivada. La diferencia entre este tratamiento de las funciones y el estudio de espacios vectoriales llevado a cabo en álgebra lineal, es que aquí el espacio vectorial tiene dimensión infinita.

Siguiendo esta semejanza con el álgebra lineal, se dice que $X(x)$ es una autofunción (con autovalor λ) del operador A si:

$$-X'' = -\frac{d^2 X}{dx^2} = \lambda X$$

Sabiendo esto, si se toman dos autofunciones X_1, X_2 con sus respectivos autovalores λ_1, λ_2 , donde $\lambda_1 \neq \lambda_2$; se tiene la identidad:

$$-X_1'' \overline{X_2} + X_1 \overline{X_2}'' = (-X_1 \overline{X_2} + X_1 \overline{X_2}')'$$

que al integrarla da la *segunda identidad de Green*:

$$\int_a^b (-X_1'' \overline{X_2} + X_1 \overline{X_2}'') dx = (-X_1 \overline{X_2} + X_1 \overline{X_2}') \Big|_a^b \quad (2.5)$$

La parte derecha de esta identidad se anula cuando a la ecuación $X'' + \lambda X = 0$ se le imponen condiciones de contorno (en los límites a, b sobre los que se trabaja) como las de Dirichlet, Neumann o Robin (que son condiciones simétricas). Y por tanto 2.5 se convierte en:

$$(\lambda_1 - \lambda_2) \int_a^b X_1 \overline{X_2} dx = 0 \quad (2.6)$$

Tomando como autofunciones aquellas que resuelven $X'' + \lambda X = 0$, se tiene el siguiente resultado:

Teorema 1. *Dos autofunciones cualesquiera con autovalores distintos, son ortogonales si se tienen condiciones de contorno simétricas. Además si una función cualquiera admite una expansión en término de estas autofunciones, los coeficientes quedan determinados.*

Demostración. El hecho de que dichas autofunciones son ortogonales se deduce de lo visto antes de enunciar el teorema. Supongamos que hay una función Φ que permite su desarrollo en función de ciertas autofunciones:

$$\Phi(x) = \sum_n A_n X_n(x)$$

si esta serie fuese convergente y se toma su producto interno con cierta autofunción X_m , se llega a:

$$(\Phi(x), X_m) = \left(\sum_n A_n X_n, X_m \right) = A_m (X_m, X_m)$$

gracias a la ortogonalidad enunciada por el teorema. De modo que los coeficientes quedan determinados por:

$$A_m = \frac{(\Phi(x), X_m)}{(X_m, X_m)}$$

□

Si se presta atención, se ve que $X(x) = \sin \frac{n\pi x}{l}$ satisface la ecuación $X'' + \lambda X = 0$, con $\lambda = -(n\pi/l)^2$. Por tanto la generalización desarrollada en este apartado se puede particularizar al caso del seno, y los coeficientes de la serie de Fourier vienen dados por la expresión que se acaba de presentar en la demostración. Es decir, si en vez de el intervalo (a, b) , se toma el intervalo $(0, l)$, se obtiene:

$$A_n = \frac{(\Phi(x), X_n)}{(X_n, X_n)} = \frac{1}{\int_0^l \sin^2 \frac{n\pi x}{l} dx} \int_0^l \Phi(x) \sin \frac{n\pi x}{l} dx = \frac{2}{l} \int_0^l \Phi(x) \sin \frac{n\pi x}{l} dx$$

Puesto que la función seno es un caso particular de las autofunciones, gracias al teorema 1 se puede garantizar que también cumple con la ortogonalidad en el sentido mencionado en esta sección:

$$\int_0^l \sin \frac{n\pi x}{l} \sin \frac{m\pi x}{l} dx = 0 \quad \text{si } n \neq m \quad (2.7)$$

Del mismo modo se puede aplicar este procedimiento al desarrollo en serie de Fourier en forma compleja para obtener sus coeficientes:

$$c_n = \frac{1}{2l} \int_{-l}^l \Phi(x) e^{-in\pi x/l} dx \quad (2.8)$$

2.1.3 Convergencia uniforme de la serie de Fourier

Teorema 2. La serie de Fourier $\sum A_n X_n(x)$ converge uniformemente a $f(x)$ en $[a, b]$ si:

- (i) $f(x)$, $f'(x)$ y $f''(x)$ existen y son continuas en $[a, b]$
- (ii) $f(x)$ satisface condiciones de contorno dadas

Demostración. Considerando que f, f' son continuas con periodo 2π , y tomando A'_n, B'_n como coeficientes de Fourier de la derivada $f'(x)$, si se integra por partes los coeficientes de la expansión completa de la serie Fourier de f :

$$A_n = \int_{-\pi}^{\pi} f(x) \cos nx \frac{dx}{\pi} = \frac{1}{n\pi} f(x) \sin nx \Big|_{-\pi}^{\pi} - \int_{-\pi}^{\pi} f'(x) \sin nx \frac{dx}{n\pi}$$

se tiene que:

$$A_n = -\frac{1}{n} B'_n, \quad \text{para } n \neq 0$$

y del mismo modo se obtiene que:

$$B_n = \frac{1}{n} A'_n$$

Si ahora se hace uso de la desigualdad de Bessel aplicada a $f'(x)$ (véase [7]), se tiene que la expansión completa de la serie de Fourier cumple:

$$\sum_{n=1}^{\infty} (|A'_n|^2 + |B'_n|^2) < \infty$$

de modo que:

$$\sum_{n=1}^{\infty} (|A_n \cos nx| + |B_n \sin nx|) \leq \sum_{n=1}^{\infty} (|A_n| + |B_n|) = \quad (2.9)$$

$$= \sum_{n=1}^{\infty} \frac{1}{n} (|B'_n| + |A'_n|) \leq \left(\sum_{n=1}^{\infty} \frac{1}{n^2} \right)^{1/2} \left[\sum_{n=1}^{\infty} 2 (|A'_n|^2 + |B'_n|^2) \right]^{1/2} < \infty \quad (2.10)$$

La última desigualdad no es más que la desigualdad de Schwarz, la cual afirma lo siguiente:

$$\left| \sum_i a_i b_i \right| \leq \left(\sum_i |a_i|^2 \right)^{1/2} \left(\sum_i |b_i|^2 \right)^{1/2}$$

Denotando S_N como la suma parcial de la serie de Fourier y sabida la convergencia puntual de las series de Fourier (véase de nuevo [7]), se tiene que:

$$\text{máx} |f(x) - S_N(x)| \leq \text{máx} \sum_{n=N+1}^{\infty} |A_n \cos nx + B_n \sin nx| \leq \sum_{n=N+1}^{\infty} (|A_n| + |B_n|) < \infty$$

esta última suma es la cola de una serie convergente, y por tanto se hace cero cuando $N \rightarrow \infty$. De este modo se obtiene convergencia uniforme. \square

2.1.4 Transformada de Fourier de una señal aperiódica

La transformada de Fourier se puede derivar a partir de las series de Fourier expuestas anteriormente, y para enseñarlo se comienza tomando como ejemplo de partida la señal cuadrada periódica de la figura 2.1:

$$x(t) = \begin{cases} 1, & |t| < T_1 \\ 0, & T_1 < |t| < T/2 \end{cases} \quad (2.11)$$

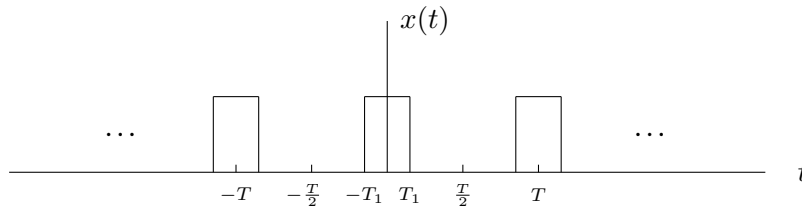


Figura 2.1: Señal cuadrada

Calculando los coeficientes de Fourier mediante 2.8, se obtiene que estos tienen la siguiente expresión:

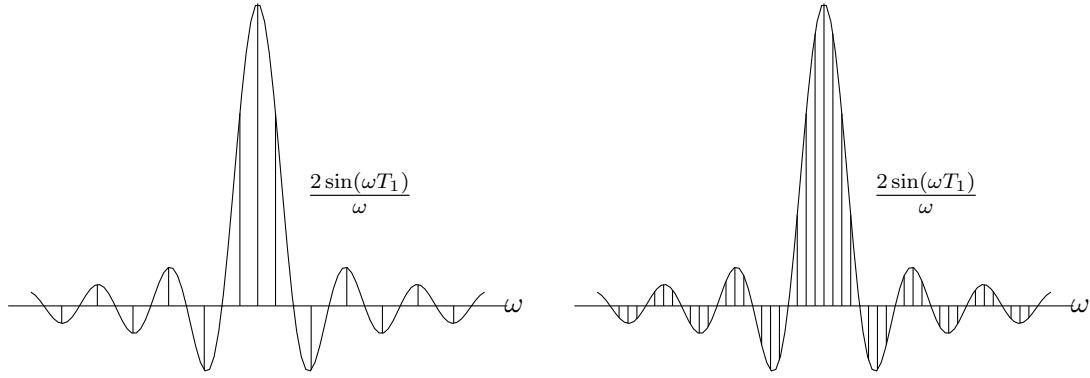
$$c_n = \frac{2 \sin(n\omega_0 T_1)}{n\omega_0 T} \quad (2.12)$$

donde $\omega_0 = 2\pi/T$ se conoce como la frecuencia fundamental de la señal $x(t)$. Viendo los coeficientes de Fourier se puede interpretar que estos son muestras equiespaciadas de la función envolvente $2 \sin(\omega T_1)/\omega$:

$$T c_n = \frac{2 \sin(\omega T_1)}{\omega} \Big|_{\omega = n\omega_0} \quad (2.13)$$

donde las muestras estarán más juntas cuanto mayor sea T , ya que de ese modo ω_0 se hace más pequeña. Esto se puede apreciar perfectamente en la figura 2.2. Viendo el efecto de coger T muy grande, se intuye que haciendo tender $T \rightarrow \infty$ se obtiene la envolvente sobre la que se hace el muestreo. Por otra parte a medida que se hace crecer la T , el periodo de la señal periódica tomada en 2.11 aumenta; de modo que si se hace tender $T \rightarrow \infty$ dicha señal quedará como un pulso cuadrado centrado en el origen. Es decir se pasa de tener una señal periódica, a atener una señal aperiódica.

Esta idea es la que permite desarrollar la transformada de Fourier para una señal aperiódica. Si se toma una señal aperiódica $x(t)$, esta puede extenderse a una señal periódica



(a) Coeficientes Ta_k de 2.13 tomando $T = 4T_1$. (b) Coeficientes Ta_k de 2.13 tomando $T = 8T_1$.

Figura 2.2: Envolvente de los coeficientes de Fourier.

$\tilde{x}(t)$ (véase [3]) con periodo T que cumpla $x(t) = \tilde{x}(t)$ cuando $|t| < T/2$, y $x(t) = 0$ cuando $|t| > T/2$.

Tomando la representación en serie de Fourier de la $\tilde{x}(t)$ en el intervalo $[-T/2, T/2]$ se tiene:

$$\tilde{x}(t) = \sum_{n=-\infty}^{\infty} c_n e^{jn\omega_0 t} \quad c_n = \frac{1}{T} \int_{-\infty}^{\infty} x(t) e^{-jn\omega_0 t} dt \quad (2.14)$$

Nótese que en el cálculo de los coeficientes c_n , se puede tomar la integral en toda la recta real de la función $x(t)$. Ya que esto es idéntico a integrar $\tilde{x}(t)$ en $|t| < T/2$. Razonando como antes se toma la función envolvente:

$$X(e^{j\omega}) = \int_{-\infty}^{\infty} x(t) e^{-j\omega t} dt \quad (2.15)$$

para poder expresar \tilde{x} como (véase [3]):

$$\tilde{x}(t) = \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} X(jn\omega_0) e^{jn\omega_0 t} \omega_0 \quad (2.16)$$

Llegado este punto, basta con hacer tender $T \rightarrow \infty$ para que \tilde{x} sea x (en el límite). Tomar este límite hace que el término ω_0 que figura en 2.16 tienda a 0, de modo que se puede interpretar como el incremento infinitesimal que se lleva a cabo en una integral. De hecho cuando $T \rightarrow \infty$, 2.16 pasa a ser:

$$x(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} X(e^{j\omega}) e^{j\omega t} d\omega \quad (2.17)$$

esta expresión se conoce como ecuación de síntesis en la transformada de Fourier. Y a la función a la que previamente se ha nombrado como “envolvente”, es decir 2.15, se la conoce como transformada de Fourier de la función aperiódica $x(t)$.

2.1.5 Interpretación de la transformada de Fourier

En la jerga del tratamiento de señales, se dice que la señal original (la función $x(t)$) está en el dominio del tiempo, mientras que su transformada de Fourier estará en el espectro de frecuencia. Este espectro proporciona información acerca de las componentes periódicas que tiene una señal. Es por ello por lo que la función $x(t) = \sin(\omega_0 t)$ tiene como transformada de Fourier dos impulsos situados en la frecuencia fundamental a la que oscila, figura 2.3.

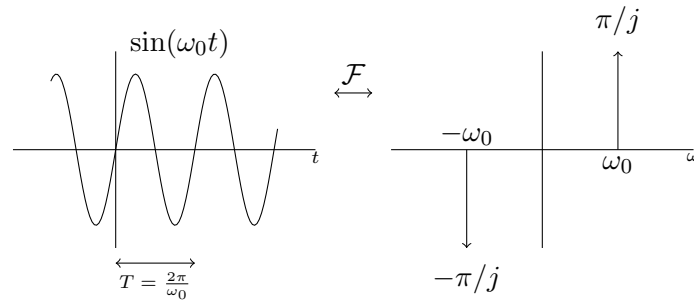


Figura 2.3: Transformada de Fourier de un seno.

2.1.6 Transformada de Fourier de una señal discreta - DTFT

Hasta ahora todo lo explicado versaba sobre señales continuas, pero para poder trabajar con una señal en un ordenador es necesario discretizarla. La señal $x(t)$ que antes era continua en el tiempo, pasa a ser una señal discreta $x[n]$ que se obtiene mediante un muestreo de la señal original con un tren de impulsos cuyo periodo debe cumplir el teorema de Nyquist (véase [3] para más información). Escogiendo un periodo adecuado para el muestreo, al final se acaba teniendo que $x[n] = x(nT)$ donde T es el periodo del tren de impulsos con el que se muestrea la señal.

Tomando las versiones discretizadas de las señales, las ecuaciones de la transformada de Fourier (véase [3] para conocer cómo se llegan a ellas) pasa a ser:

$$x[n] = \frac{1}{2\pi} \int_{2\pi} X(e^{j\omega}) e^{j\omega n} d\omega \quad (2.18)$$

$$X(e^{j\omega}) = \sum_{n=-\infty}^{\infty} x[n] e^{-j\omega n} \quad (2.19)$$

Pero aún se tiene la frecuencia ω como variable continua. Así que si se quiere trabajar con la señal en un ordenador, aún queda por ver cómo trabajar con tiempo y frecuencia como variables discretas.

2.1.7 Transformada Discreta de Fourier - DFT

Atendiendo a lo expuesto en los anexos A y B, y a lo mostrado en 3, los coeficientes de la serie de Fourier para una señal discreta tienen periodo N . Si en el caso de señales continuas la transformada de Fourier no era más que los coeficientes de Fourier de una señal periódica cuyo periodo se hace tender a infinito, en el caso de señales discretas se sigue el mismo procedimiento.

Si se tiene una señal finita $x[n]$ que se anula para $n \notin [0, N - 1]$, lo primero que se hace es extenderla periódicamente y obtener su representación en serie de Fourier. La extensión periódica de $x[n]$ con periodo N es:

$$\tilde{x}[n] = \sum_{r=-\infty}^{\infty} x[n - rN]$$

Acto seguido se calculan los coeficientes de Fourier de su representación en serie:

$$\tilde{X}[k] = \sum_{n=0}^{N-1} \tilde{x}[n] W_N^{kn}$$

Lo siguiente es hacer tender el periodo N de la señal $\tilde{x}[n]$ a infinito, de modo que pase a ser idénticamente la señal original $x[n]$. Esto hace que la señal $\tilde{X}[k]$ deje de ser periódica,

pues fuera del intervalo $[0, N - 1]$ la señal $x[n]$ es nula. De modo que ahora $\tilde{X}[k] = X[k]$ se anula también fuera de dicho intervalo.

Así se llega a que la transformada discreta de Fourier de una señal finita es:

$$X[k] = \sum_{n=0}^{N-1} x[n] W_N^{kn} \quad (2.20)$$

donde $X[k] = 0$ si $k \notin [0, N - 1]$. Del mismo modo se podrá reconstruir la señal original $x[n]$ con la fórmula:

$$x[n] = \frac{1}{N} \sum_{k=0}^{N-1} X[k] W_N^{-kn} \quad (2.21)$$

Una propiedad muy importante de la DFT es que cuando la señal $x[n]$ es real, su transformada cumple que:

$$X[N - k] = X^*[k] \quad (2.22)$$

donde $X^*[k]$ es el conjugado de $X[k]$. Esta propiedad se comprende fácilmente atendiendo a lo siguiente:

$$X[N - k] = \sum_{n=0}^{N-1} x[n] e^{-j2\pi n} e^{j\frac{2\pi}{N}kn} = \left(\sum_{n=0}^{N-1} x[n] e^{-j\frac{2\pi}{N}kn} \right)^* = X^*[k]$$

2.2 Filtro de Chebyshev de tipo 2

Los filtros de Chebyshev se caracterizan por tener crestas en las zonas de paso, o bien en las zonas de atenuación de la señal sobre la que actúan. Estos filtros funcionan basándose en los polinomios de Chebyshev, los cuales vienen dados por la siguiente recurrencia:

$$T_0(x) = 1 \quad (2.23)$$

$$T_1(x) = x \quad (2.24)$$

$$T_{n+1}(x) = 2x \cdot T_n(x) - T_{n-1}(x) \quad (2.25)$$

Entre las múltiples propiedades de los polinomios de Chebyshev, las que más interesan para el diseño de los filtros de Chebyshev son las siguientes:

1. $T_n(x) \in [-1, 1], \quad |x| \leq 1$
2. $T'_n(x) > 0, \quad |x| > 1$
3. $T_n(1) = 1, \quad \forall n$
4. $T_{2n+1}(0) = 0, \quad T_{2n}(0) = \pm 1$
5. $T_n(x) = 0 \iff |x| < 1$
6. $T_n(x) = \cos(n \arccos x)$

La representación de polinomios de Chebyshev de distinto orden (véase la figura 2.4), facilita la comprensión de alguna de estas propiedades.

Haciendo uso de estos polinomios, la función de ganancia de un filtro de Chebyshev de tipo 2 con n -ésimo orden viene dada por:

$$G(\omega) = \frac{1}{\sqrt{1 + \frac{1}{\epsilon^2 T_n^2(\omega_0/\omega)}}} \quad (2.26)$$

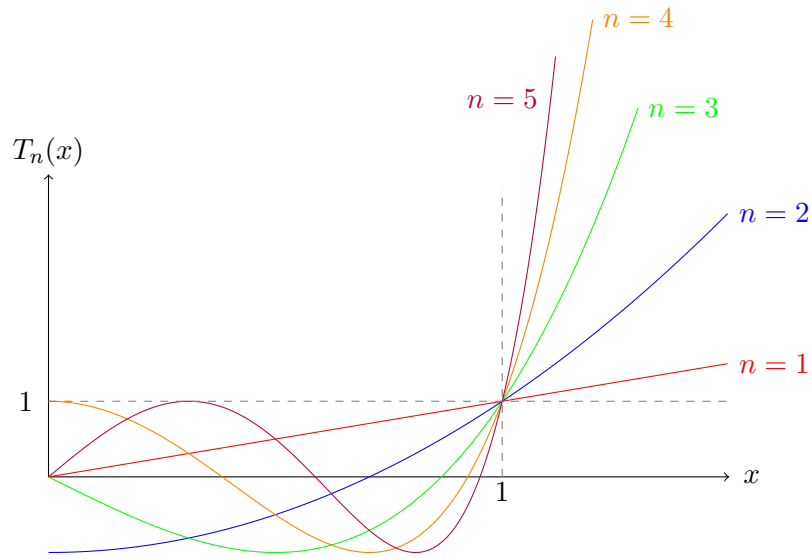


Figura 2.4: Polinomios de Chebyshev de distinto orden.

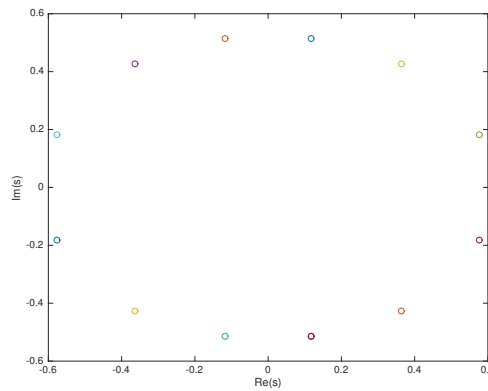


Figura 2.5: Polos de la función de ganancia 2.26 de un filtro de Chebyshev de tipo 2 y sexto orden,

donde ω_0 es la frecuencia de corte del filtro, y ϵ el factor que determina la amplitud de las crestas. Este es un filtro paso bajo con crestas en la zona de atenuación de la señal.

Para obtener los polos del filtro bastará con ver dónde se anula el denominador, y si se considera como frecuencia de corte $\omega_0 = 1$ y se toma $-1/i \cdot s = \cos \theta$, se tiene que los polos son:

$$\begin{aligned}
 1 + \epsilon^2 T_n^2 \left(-\frac{1}{i \cdot s} \right) &= 0 \\
 1 + \epsilon^2 \cos^2(n\theta) &= 0 \\
 \theta &= \frac{1}{n} \arccos \left(\pm \frac{i}{\epsilon} \right) + \frac{m\pi}{n} \\
 \text{dado que } s &= \frac{i}{\cos \theta} \\
 s &= \frac{i}{\cos \left(\frac{1}{n} \arccos \left(\pm \frac{i}{\epsilon} \right) + \frac{m\pi}{n} \right)}, \quad m = 1, \dots, n
 \end{aligned}$$

En [6] se puede ver que los polos se pueden expresar como:

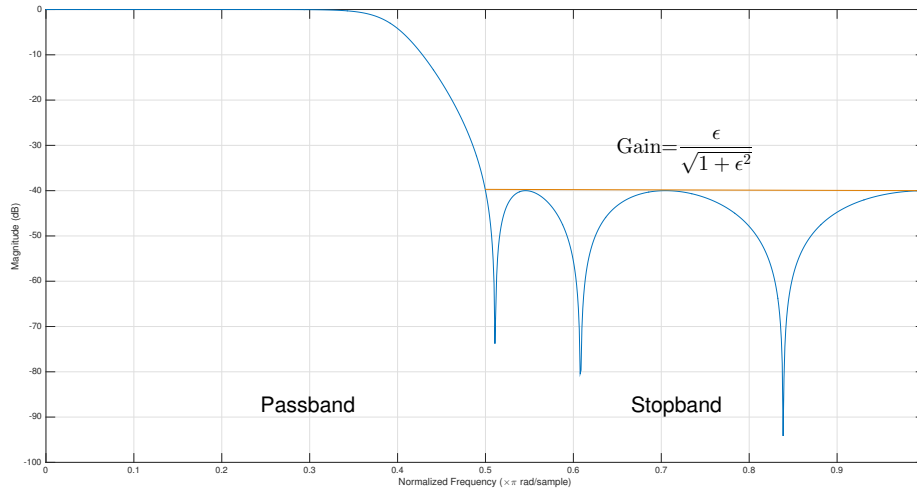


Figura 2.6: Ganancia de un filtro de Chebyshev de tipo 2, con frecuencia de corte $\omega_0 = \pi/2$, y caída de $40dB$.

$$s = \sinh(\varphi_1) \sin\left(\frac{(2m-1)\pi}{2n}\right) + i \cosh(\varphi_1) \cos\left(\frac{(2m-1)\pi}{2n}\right), \quad m = 1, \dots, 2n \quad (2.27)$$

donde $\varphi_1 = 1/n \sinh^{-1}(1/\epsilon)$. Como los polos no toman valores reales (véase un ejemplo en la figura 2.5), la ganancia expresada en 2.26 no presenta problemas. Dicha expresión tomará valores altos cuando ω es menor que la frecuencia de corte ω_0 , ya que en dichos casos el polinomio de Chebyshev tomará valores por encima de uno, tal y como se puede apreciar en 2.4. Sin embargo cuando se tiene $\omega > \omega_0$, el polinomio de Chebyshev oscila entre 1 y -1 , por lo que los valores de la ganancia irán oscilando por valores menores a los que se obtenían cuando $\omega < \omega_0$ (véase la figura 2.6).

2.3 Sustracción espectral

En [2] se describe cómo llevar a cabo la sustracción del ruido en una señal de voz, en la que hay presencia de un ruido que se asume estacionario. Del mismo modo se asume que eliminando la componente de la magnitud del ruido, se "limpia" la señal.

2.3.1 Modelo de ruido aditivo

Definiendo $n(k)$ como la señal discretizada del ruido presente, y $s(k)$ como la señal discretizada de voz, su suma será:

$$x(k) = s(k) + n(k)$$

con transformada de Fourier:

$$X(e^{j\omega}) = S(e^{j\omega}) + N(e^{j\omega}) \quad (2.28)$$

La sustracción espectral lleva a cabo una resta del valor medio de la magnitud del ruido, es decir, tomando $\mu(e^{j\omega}) = E\{|N(e^{j\omega})|\}$ se lleva a cabo la siguiente estimación de la señal de voz $\hat{S}(e^{j\omega})$:

$$\hat{S}(e^{j\omega}) = [|X(e^{j\omega})| - \mu(e^{j\omega})] e^{\theta_X(e^{j\omega})} \quad (2.29)$$

que equivale a pasar $X(e^{j\omega})$ por el filtrado:

$$\hat{S}(e^{j\omega}) = H(e^{j\omega})X(e^{j\omega})$$

con:

$$H(e^{j\omega}) = 1 - \frac{\mu(e^{j\omega})}{|X(e^{j\omega})|}$$

2.3.2 Error espectral

Esta estimación de la señal de voz induce el siguiente error (denominado como error espectral en [2]):

$$\epsilon(e^{j\omega}) = \hat{S}(e^{j\omega}) - S(e^{j\omega}) = N(e^{j\omega}) - \mu(e^{j\omega})e^{j\theta_X(e^{j\omega})}$$

Para reducir este error el artículo propone 4 pasos a seguir, que son los presentados en las siguientes cuatro subsecciones.

2.3.3 Promediado de la magnitud

Puesto que en la estimación \hat{S} de la señal de voz se resta el valor medio del ruido, se puede tomar también el valor medio de X para intentar reducir el error espectral; puesto que de este modo la señal original promediada será la suma de los promedios del ruido y de la voz.

Tomando:

$$\overline{|X(e^{j\omega})|} = \frac{1}{M} \sum_{i=0}^{M-1} |X_i(e^{j\omega})|$$

con X_i la transformada de Fourier de la i -ésima ventana tomada para $x[k]$, se tiene que:

$$S_A(e^{j\omega}) = \left[\overline{|X(e^{j\omega})|} - \mu(e^{j\omega}) \right] e^{j\theta_X(e^{j\omega})}$$

De este modo el error espectral pasa a ser aproximadamente:

$$\epsilon(e^{j\omega}) = S_A(e^{j\omega}) - S(e^{j\omega}) \approx \overline{|N|} - \mu$$

donde:

$$\overline{|N(e^{j\omega})|} = \frac{1}{M} \sum_{i=0}^{M-1} |N_i(e^{j\omega})|$$

y por tanto al tomar más y más enventanados del ruido, se consigue que $|N| \approx \mu$, y consecuentemente se suprime el error espectral.

Pero esta idea tiene un problema, y es que el ruido no es estacionario; de modo que el promediado debe ser tomado en franjas de tiempo limitadas.

2.3.4 Rectificación "a mitad de camino"

Este método consiste en que sabiendo que \hat{S} encierra la diferencia entre $|X|$ y μ , se puede trincar \hat{S} a 0 en aquellas frecuencias ω en las que $|X| \leq \mu$; pues se da por supuesto que en dichas zonas el habla no tiene presencia. Esto se implementa mediante el siguiente filtrado de la señal $X(e^{j\omega})$:

$$\hat{S}(e^{j\omega}) = H_R(e^{j\omega})X(e^{j\omega})$$

donde:

$$H_R = \frac{H(e^{j\omega}) + |H(e^{j\omega})|}{2}$$

Todo esto equivale a que:

$$\hat{S}(e^{j\omega}) = \begin{cases} [|X(e^{j\omega})| - \mu(e^{j\omega})] e^{j\theta_X(e^{j\omega})} & , |X| > \mu \\ 0, & , |X| \leq \mu \end{cases}$$

Sin embargo también es cierto que este proceso puede conducir al error de eliminar habla en aquellos momentos en los que X toma valores de magnitud pequeños, y estas situaciones pueden ser momentos en los que no haya presencia de ruido, y se tenga una presencia de habla con poca magnitud.

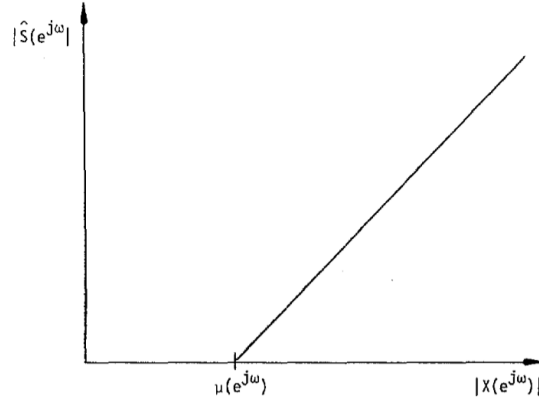


Figura 2.7: Representación de la "rectificación" de \hat{S} extraída de [2]

2.3.5 Reducción de ruido residual

Todavía queda por eliminar aquel ruido por encima de su valor medio $\mu(e^{j\omega})$, y para ello el siguiente paso es reducir en lo posible lo que se conoce como ruido residual $N_R = N - \mu e^{j\theta_N}$. Se sabe que el ruido varía más rápido que el habla, por tanto la idea es tomar los valores más pequeños de $\hat{S}(e^{j\omega})$ de varias ventanas adyacentes en el espectro de frecuencia, pues sabemos que esto no debería modificar el contenido del habla. Basándose en esta idea, [2] termina llegando a que N_R se reduce aplicando el siguiente proceso:

$$|\hat{S}_i(e^{j\omega})| = \begin{cases} |\hat{S}_i(e^{j\omega})| & , |\hat{S}_i(e^{j\omega})| \geq \text{máx } |N_R(e^{j\omega})| \\ \text{mín } |\hat{S}_j(e^{j\omega})|, j = i - 1, i, i + 1 & , |\hat{S}_i(e^{j\omega})| < \text{máx } |N_R(e^{j\omega})| \end{cases}$$

donde $S_i(e^{j\omega}) = H_R(e^{j\omega})X_i(e^{j\omega})$. Detrás de esto está que en las tres ventanas adyacentes sobre las que se toma el mínimo, el habla (S) de la señal original no varía; sin embargo el ruido sí, y por ello la ventana que tome menor $|\hat{S}_j|$ será aquella con menos ruido. Esta afirmación se debe a que $X = S + N$, y por tanto el valor más pequeño entre ventanas en las que sabemos que S no varía, se debe a que N es también el valor más pequeño de las tres ventanas.

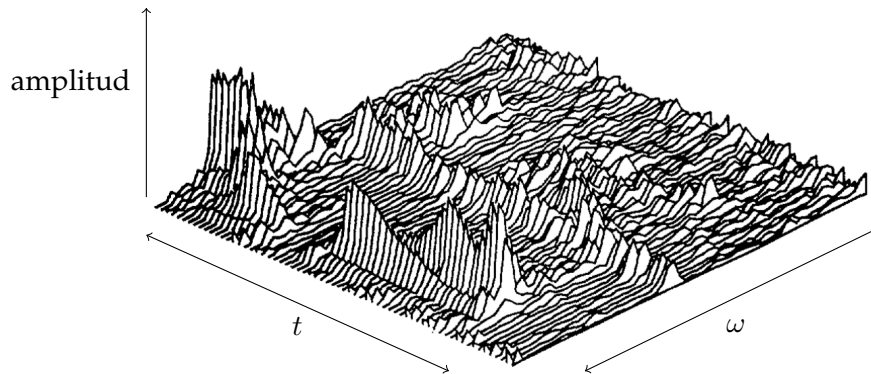
Por otra parte tiene sentido no modificar $|\hat{S}_i|$ cuando toma valores mayores al máximo ruido residual, pues en ese caso la presencia de voz es notoria para esas frecuencias, tanto, que queda por encima del máximo N_R .

En los diagramas de waterfall (explicados en la siguiente sección) obtenidos tras realizar la sustracción espectral en un audio, en el que un reportero graba su voz con el ruido de fondo de un helicóptero (véanse las figuras 2.8a y 2.8b), se aprecia cómo las frecuencias altas asociadas al ruido ven muy mermada su amplitud. Del mismo modo, picos abruptos de ruido en frecuencia son notoriamente atenuados al mismo tiempo que se respeta la señal original (véase [2]).

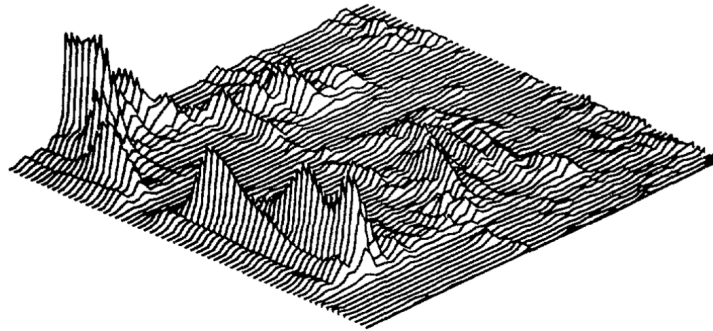
2.4 Representaciones en el espectro de frecuencia

En el tratamiento de señales existen dos formas muy comunes de representar las características espectrales de la señal a lo largo del tiempo. Una de ellas es el uso de los conocidos

como diagramas “waterfall”, los cuales representan la transformada de Fourier de la señal a lo largo de distintos intervalos de tiempo. Este diagrama es en 3 dimensiones tal y como se puede apreciar en las figuras 2.8a y 2.8b.



(a) Diagrama waterfall de voz con ruido de un helicóptero (tomado de [2]).



(b) Diagrama waterfall de voz con ruido de un helicóptero, tras la sustracción espectral con rectificación a mitad de camino y promediado de tres ventanas (tomado de [2]).

La otra forma de representar las características espectrales de la señal a lo largo del tiempo es el espectrograma, esta representación es la versión en 2 dimensiones del diagrama waterfall. En uno de sus ejes se representa el tiempo, y en el otro las distintas frecuencias; y la amplitud de la transformada de Fourier de la señal va asociada con colores más intensos a mayor amplitud, y con colores menos intensos a menor amplitud (véase la figura 2.9).

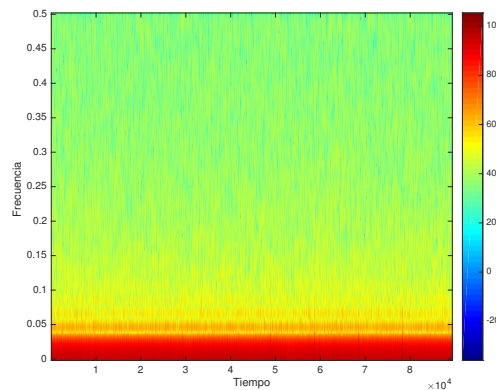


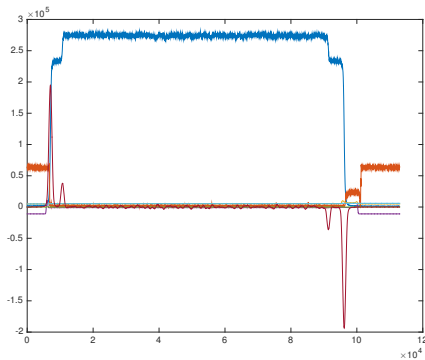
Figura 2.9: Espectrograma de una señal de densidad del control de calidad de una componente.

DESARROLLO E IMPLEMENTACIÓN

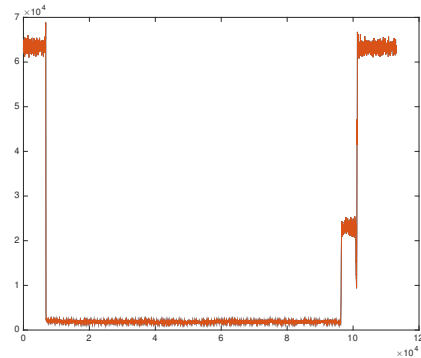
3.1 Datos disponibles

En este trabajo se ha tenido acceso a mediciones procedentes de un proceso de inspección de componentes de un reactor nuclear. Cada inspección de una componente da lugar a 7 señales (véase la figura 3.1a) que tienen como objetivo encontrar materiales fuera de especificación. En estas 7 señales interesa que haya el menor número de irregularidades.

Se dispone de 50 mediciones asociadas a componentes sin irregularidades, y de 22 mediciones asociadas a componentes fuera de especificación.

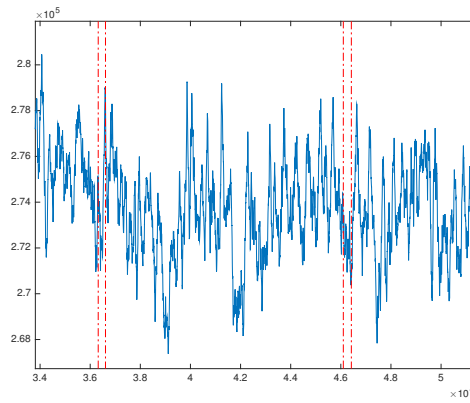


(a) Representación de las 7 señales asociadas a las distintas mediciones de un objeto pasando el control de calidad.

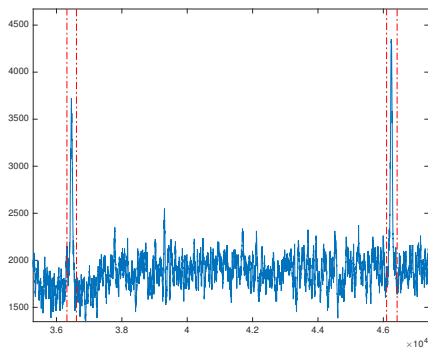


(b) Señal de densidad en el control de calidad.

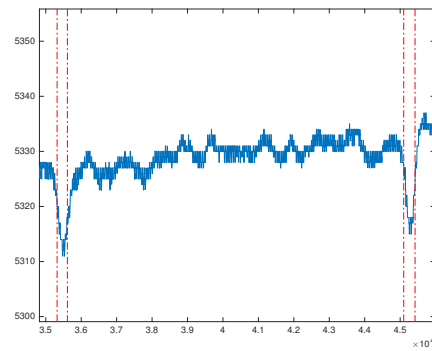
Entre los distintas señales mostradas en la figura 3.1a, la representada con color naranja mide la densidad del objeto (en la figura 3.1b aparece exclusivamente dicha señal). En la figura 3.2 se muestra en rojo los intervalos de dos irregularidades encontradas en la señal de densidad de uno de los 22 objetos (también llamados 22 componentes a inspeccionar) fuera de especificación, y se puede observar como el resto de señales (a excepción de la tercera) las irregularidades no presentan picos tan exageradamente abruptos como en el caso de la señal de densidad.



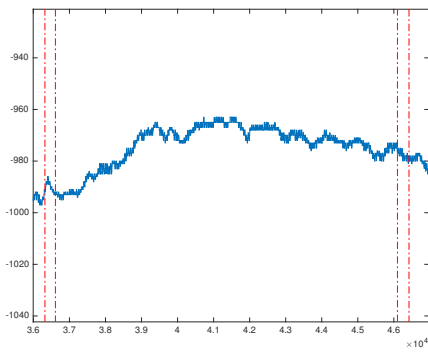
(a) Señal 1



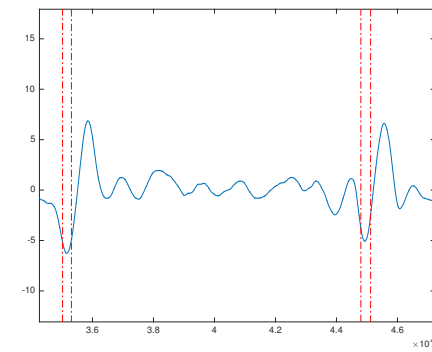
(b) Señal 2, menor densidad, mayor amplitud



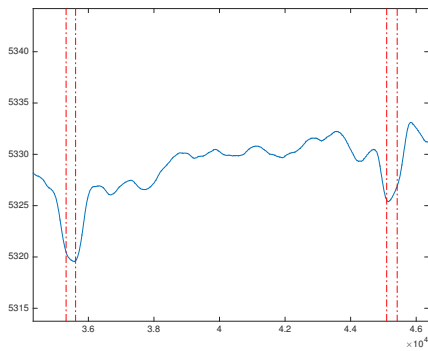
(c) Señal 3, mayor valor, mayor contenido de materialX



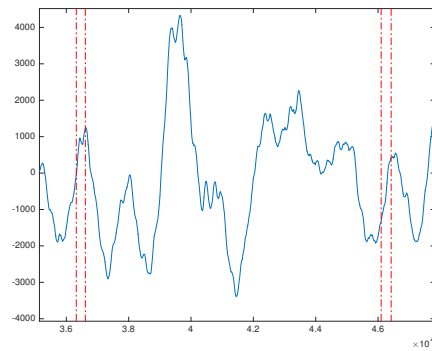
(d) Señal 4



(e) Señal 5



(f) Señal 6



(g) Señal 7

Figura 3.2: Visualización de intervalos de picos de densidad en las 7 señales. Nótese que por diferencias en los sensores, el alineamiento temporal no coincide.

El propósito inicial del trabajo era realizar un análisis detallado de las 7 señales y la relación existente entre ellas, sin embargo el análisis de la señal de densidad (la que mejor refleja las irregularidades) ha sido tan extenso, y se ha obtenido tanta información, que las tareas de filtrado se han centrado en esta señal, la cual tomará valores altos cuando haya poca densidad en la zona de medición, y valores bajos cuando más densidad haya.

Para que una componente del reactor pase el control de calidad en las mediciones de densidad, basta con que esta no presente picos en una zona que recibe el nombre de "zona activa". En la figura 3.3 se presenta una de las 50 señales que pasa el control de calidad por tener su zona activa con una densidad regular.

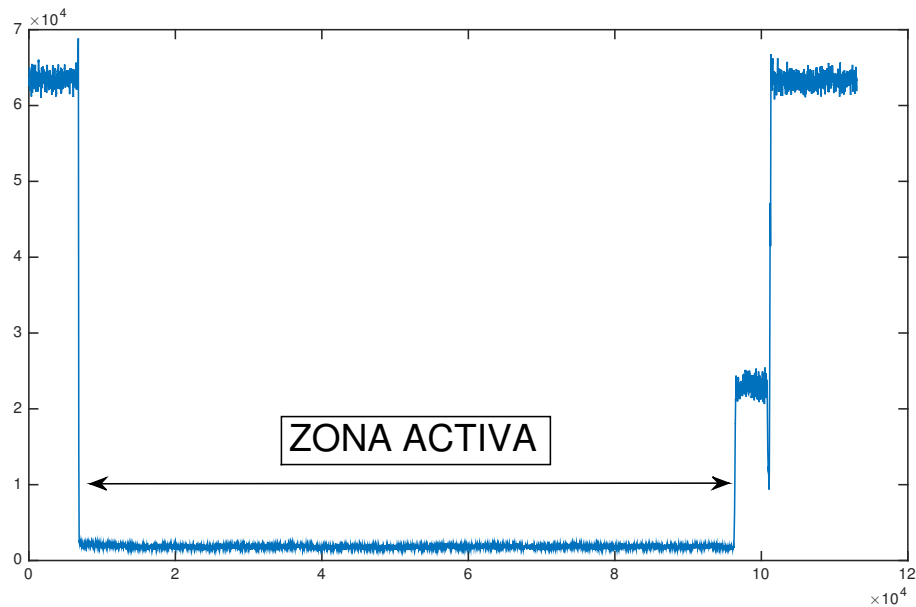


Figura 3.3: Zona activa de una de las 50 señales sin irregularidades.

En las 22 señales con irregularidades existen picos que van asociados a zonas en las que el objeto presenta menos densidad de la necesaria (véase la figura 3.4).

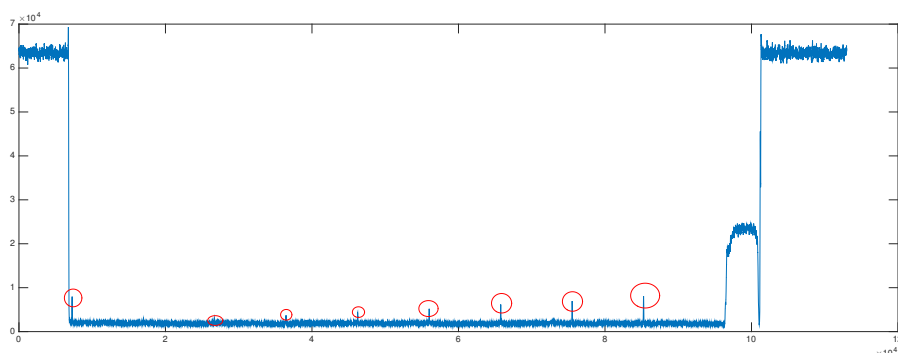


Figura 3.4: Picos en las zonas activas de una de las 22 señales con irregularidades.

Cada pico corresponde a imperfecciones en el material analizado, que pueden dar lugar a problemas muy serios si no se detectan convenientemente, y todos ellos han sido etiquetados indicando su principio y final para usos posteriores durante las pruebas. El principio y final de cada pico de una señal ha sido almacenado como un listado de arrays de dos números.

3.2 Implementación de la sustracción espectral

3.2.1 Enventanado de las señales

En todo momento en el que ha sido necesario llevar a cabo un enventanado de alguna señal, se han utilizado ventanas de Hanning cuya longitud ha estado controlada por el parámetro W (típicamente toma como valores potencias de dos que van desde las 64 muestras hasta las 1024 muestras), y el solapamiento entre las ventanas ha sido de factor 0.5 en todos los casos. Este tipo de ventanas presenta un buen compromiso entre los diversos tipos de distorsión que se introducen en frecuencia. En la figura 3.5 se puede observar cómo se ha realizado el enventanado.

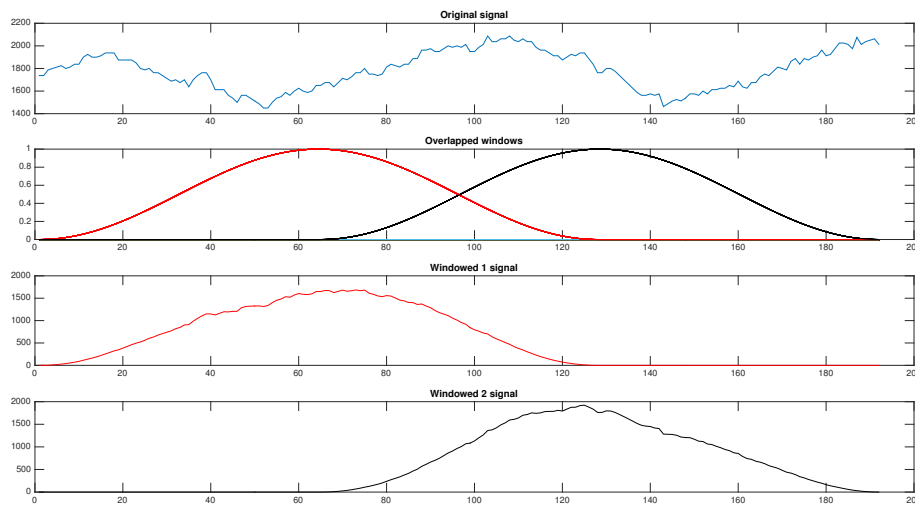


Figura 3.5: Proceso de enventanado de una señal con dos ventanas de Hanning solapadas con factor 0.5

La señal enventanada se almacena en una matriz en la que la n -ésima columna corresponde con las muestras de multiplicar la señal por la n -ésima ventana.

Al aplicar la FFT sobre una señal que ha sido enventanada y guardada como una matriz, basta con almacenar la primera mitad de las filas. Esto es así gracias a que la DFT de una señal real tiene la propiedad 2.22, la cual establece que la k -ésima muestra de la DFT equivale a la $(N - k)$ -ésima muestra conjugada, es decir:

$$X[N - k] = \overline{|X[k]|e^{i\theta(X[k])}} = |X[k]|e^{-i\theta(X[k])} = \overline{X[k]} \quad (3.1)$$

donde \bar{a} denota la conjugación del número complejo a , y $\theta(X[k])$ representa la fase de $X[k]$. Por tanto es suficiente almacenar los $\lfloor N/2 \rfloor + 1$ primeros valores de la señal enventanada, y posteriormente reconstruir la señal enventanada tomando el conjugado de las primeras muestras.

3.2.2 Estimación del ruido

Partiendo de la hipótesis de que el ruido presente en la señal se ajusta al modelo aditivo expresado en la ecuación 2.28 de la sustracción espectral, se ha hecho uso de las zonas activas de las 50 señales sin irregularidades para obtener una estimación del ruido, considerando que en dichas zonas tan solo hay presencia de ruido.

Para cada una de estas 50 señales se ha tomado un enventanado de la zona activa, para posteriormente aplicar la FFT sobre cada una de las ventanas. Dicha FFT tiene el

mismo número de muestras que cada una de las ventanas por las que se multiplica la señal.

Acto seguido se calcula la media del módulo del ruido en cada una de las ventanas de las señales:

$$\tilde{N}_l[k] = \frac{\sum_i |N_l[k, i]|}{\# \text{ windows}} \quad (3.2)$$

donde N_l es la DFT de la zona activa de la l -ésima señal sin irregularidades enventanada, y $N_l[k, i]$ es el valor que toma la DFT de la señal N_l en la k -ésima muestra de la i -ésima ventana.

Por último se realiza el promedio de las medias de ruido obtenidas para cada señal:

$$\mu[k] = \frac{\sum_l \tilde{N}_l[k]}{\# \text{ nice signals}} \quad (3.3)$$

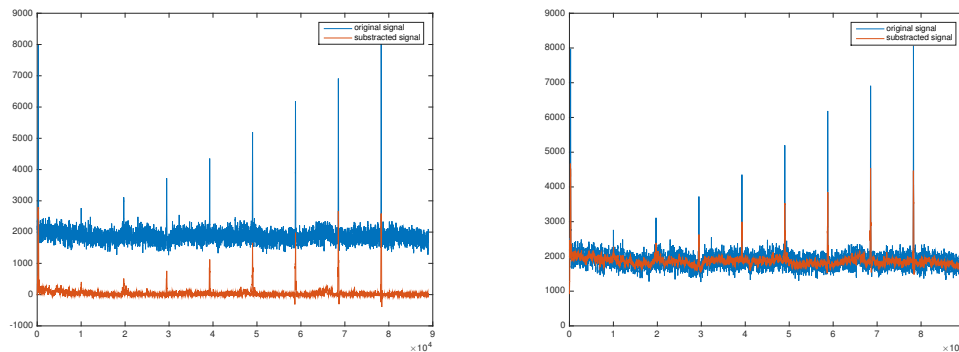
3.2.3 Modificación de la sustracción y pasos implementados

La implementación de la sustracción espectral hace uso de todos los pasos citados en [2]. De hecho no se ha implementado ni la atenuación de la señal cuando no hay presencia de voz, ni la actualización del ruido (véase el anexo D). Esto es así porque en las señales con las que se ha trabajado, la señal que mide la densidad (equivalente a la voz en el artículo [2]) no cesa en ningún momento de la medición; y porque se considera que el ruido es estacionario.

Otra de las diferencias con respecto a la explicación de la sección 2.3, es que en el momento de la resta de módulos de las DFT mostrada en la ecuación 2.29, se restan todas las frecuencias menos la asociada a $\omega = 0$. El motivo es que esta frecuencia es la que contiene la información del offset de la señal original, y si se ve modificada, la reconstrucción tras la sustracción dará lugar a una señal con menor amplitud que la original. La explicación de este fenómeno es que la transformada de Fourier de una constante resulta ser un impulso centrado en el origen, esto se puede comprobar tomando la transformada inversa de Fourier:

$$\int_{-\infty}^{\infty} \delta(\omega) e^{k\omega t} d\omega = 1 \quad (3.4)$$

Estos efectos se observan perfectamente en la figura 3.6.



(a) Sustracción espectral en zona activa de la primera señal con irregularidades. Restando en $\omega = 0$.

(b) Sustracción espectral en zona activa de la primera señal con irregularidades. Sin restar en $\omega = 0$.

Figura 3.6: Efectos en el offset al restar el módulo del ruido estimado en la frecuencia $\omega = 0$. En azul la señal original, y en rojo la reconstrucción de la señal sustraída.

El orden de los pasos llevados a cabo en la implementación de la sustracción espectral (sin contar la estimación del ruido, el enventanado y la reconstrucción de la señal) han sido:

1. FFT de la señal con ruido.
2. Promediado de magnitudes descrito en 2.3.3.
3. La reducción de ruido residual descrita en 2.3.5.
4. La rectificación a mitad de camino expuesta en 2.3.4.

3.2.4 Reconstrucción de la señal

Tras la sustracción espectral llevada a cabo sobre el módulo de la señal con ruido, se tiene en memoria dos matrices que almacenan el módulo y fase de la señal resultante. Multiplicando dichas matrices elemento a elemento (multiplicando cada módulo por la exponencial del número imaginario i por su fase), se obtiene la reconstrucción de la señal resultante. Además se hace uso de la propiedad citada en 2.22 para ahorrar en almacenamiento, tal y como se muestra en la figura 3.7.

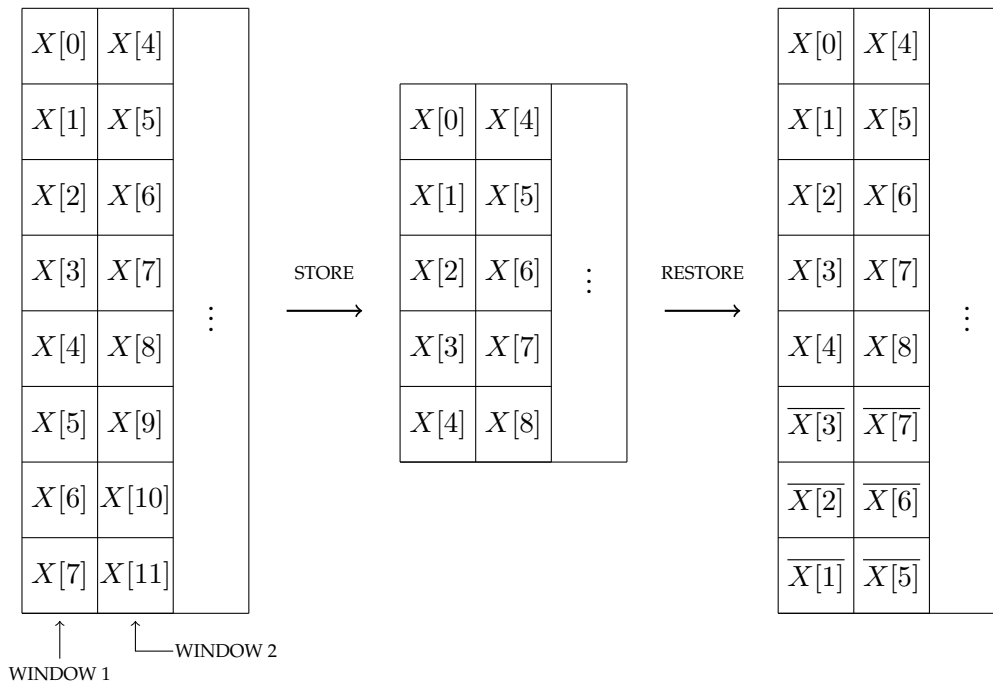


Figura 3.7: Almacenamiento y reconstrucción de la DFT de $N = 8$ muestras de una señal real enventanada.

Teniendo la matriz de la señal sustraída enventanada, basta con tomar la suma de cada una de las ventanas presentes en la matriz para obtener la señal resultante, tal y como se observa en el ejemplo de la figura 3.8.

3.3 Elección de frecuencia de corte en filtro de Chebyshev

Puesto que la señal del densitómetro presenta picos en las zonas con irregularidades, se considera que la señal que se desea mantener tras el filtrado es la señal de un triángulo. Siguiendo este planteamiento se ha tomado la transformada de Fourier de un triángulo (la cual se puede ver en la figura 3.9) cuya anchura fuera la anchura promedio de los picos presentes en las 22 señales del densitómetro con irregularidades.

Una vez conseguida la transformada del triángulo, se han obtenido filtros de Chebyshev de tipo 2 y sexto grado, cuyas frecuencias de corte variaban a lo largo de los distintos ceros de la transformada del triángulo. Para ello se ha hecho uso de la función de Matlab

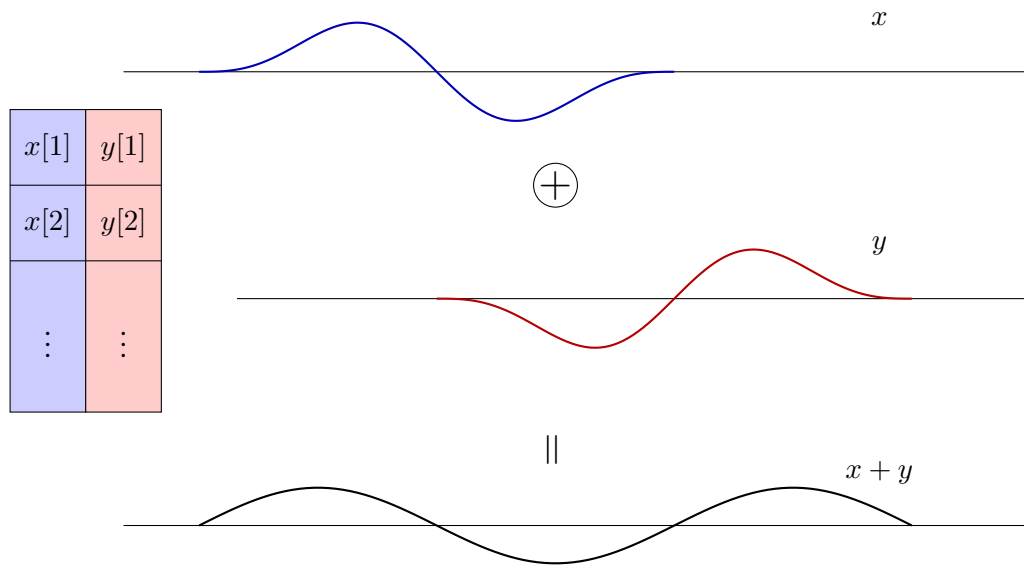


Figura 3.8: Suma de las dos componentes de una señal descompuesta en dos ventanas.

$cheby2(6, Rs, Ws)$, donde Rs indica la caída en decibelios, y Ws la frecuencia de corte del filtro.

De este modo se consigue mantener más o menos componente espectral de los picos del densitómetro, en función de qué cero se escoja como frecuencia de corte.

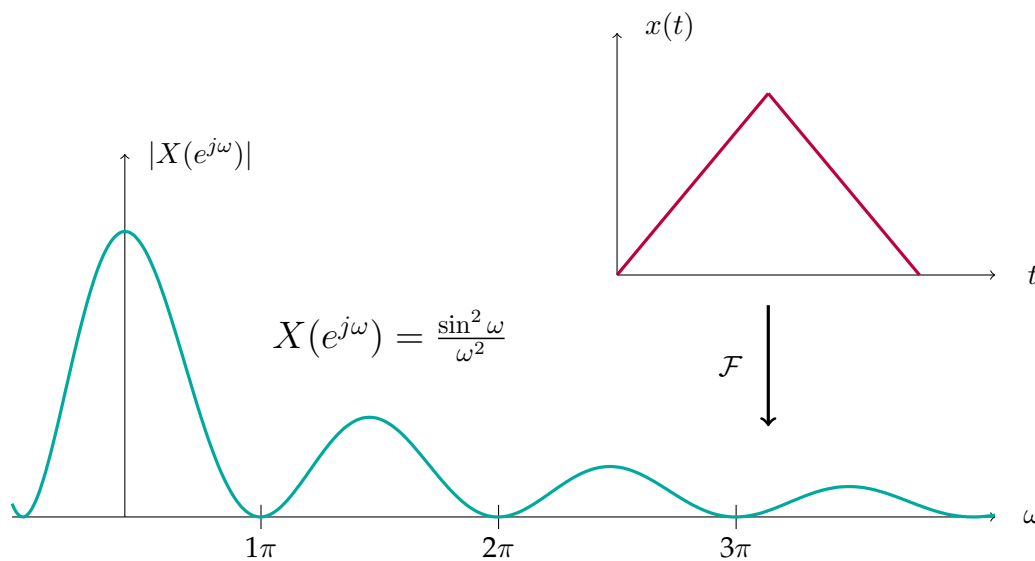


Figura 3.9: Transformada de Fourier de un triángulo (reescalada).

3.4 Localización de picos

Las señales con irregularidades presentan picos que son tenidos en cuenta si sobrepasan cierto umbral establecido por el control de calidad. Para ser más específico, un pico será tenido en cuenta únicamente si contiene más de una muestra con amplitud superando un umbral que viene dado por:

$$th(x) = f(\text{mean}(x[k])), \quad k \in \text{zona_activa}(x) \tag{3.5}$$

es decir, el proceso de control de calidad obtiene el umbral a partir de la amplitud media de la zona activa de la señal $x[k]$.

Una vez obtenido el umbral $th(x)$ para la señal, basta con recorrerla por todas las zonas activas para encontrar qué valores tienen una amplitud superior. De este modo se obtiene un listado de arrays que indican el comienzo y fin de cada uno de los picos localizados en la zona activa.

Acto seguido, para cada uno de los picos localizados, se asignará como amplitud de dicho pico el valor máximo presente entre el principio y fin de este:

$$peak_value_{p_i} = \text{máx}\{x[k]\}, \quad k \in p_i$$

donde p_i denota el intervalo del i -ésimo pico detectado en la zona activa de la señal $x[k]$. En la figura 3.10 se puede apreciar cómo se detectan los picos siguiendo el procedimiento descrito.

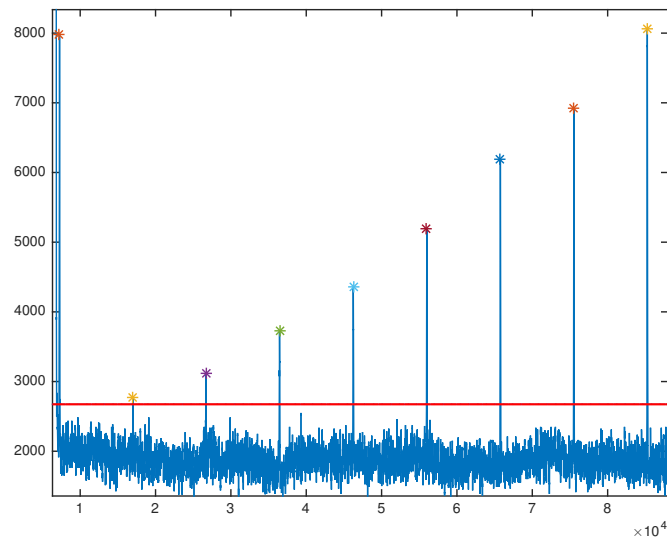


Figura 3.10: Zona activa de la señal 12 con picos. En rojo el umbral $th(x)$, y en asteriscos de colores los picos detectados en la zona activa.

RESULTADOS

4.1 *Sustracción espectral*

La eficacia de la sustracción espectral (descrita en 3.2.3) sobre la señal de densidad está muy sujeta al tamaño de ventana usado en el proceso de enventanado de la señal. Sin embargo la limpieza de la señal queda patente desde el primer momento en que se observa la señal reconstruida en el tiempo, y tras la obtención del espectrograma para cualquier tamaño de ventana. En las figuras 4.1 y 4.2, se ve qué aspecto tiene la señal de densidad antes y después de la sustracción espectral.

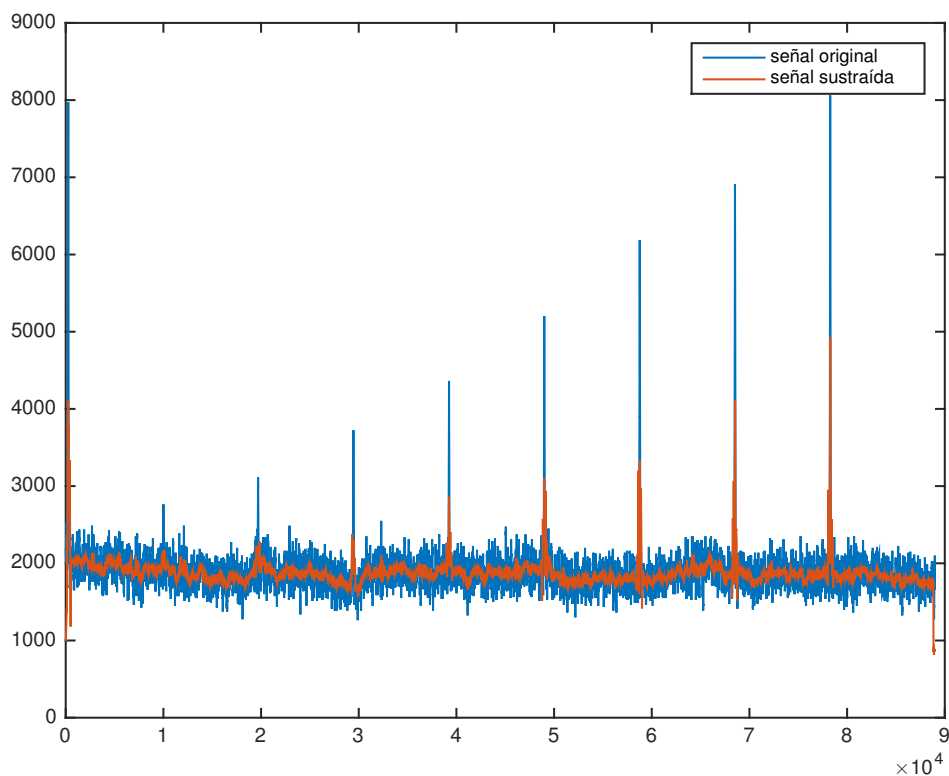


Figura 4.1: Sustracción espectral en la señal de densidad con $W = 256$, en rojo la señal sustraída, y en azul la original.

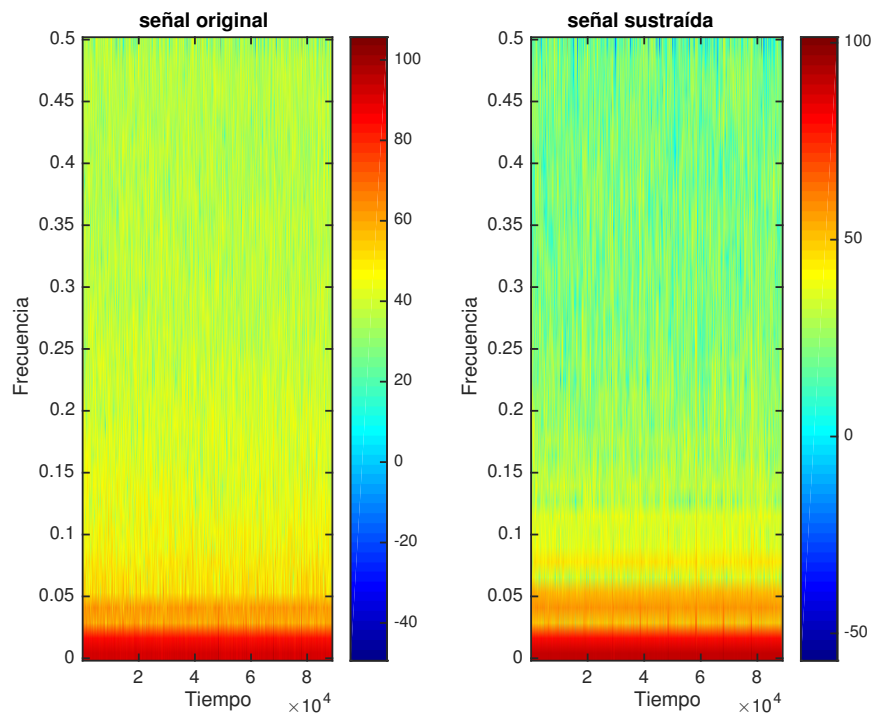


Figura 4.2: Espectrograma de la sustracción espectral en la señal de densidad tomando $W = 256$.

En la figura 4.2 se aprecia cómo se atenúa la señal en la zona de alta frecuencia (colores menos intensos en altas frecuencias del espectrograma), lo cual significa la reducción del ruido de alta frecuencia, que es la parte de la señal que varía más bruscamente. Por tanto la sustracción suaviza la señal a limpiar, pues esta pasará a tener menores variaciones repentinas.

En este capítulo se trata exclusivamente la sustracción espectral sobre la señal de densidad, dado que es en esta señal donde la detección de picos puede trabajar, ya que en el resto de señales los picos no son tan pronunciados y pueden ser valles de amplitud, en lugar de crestas de amplitud (que es lo que contempla el proceso descrito en 3.4).

4.1.1 Efecto del tamaño de ventana en la sustracción

Los tamaños de ventana escogidos son potencias de dos, pues el cálculo de la FFT está enfocado a estos casos. Así mismo la resolución escogida para la FFT coincide con el tamaño de ventana, el motivo es que de este modo se extrae la máxima cantidad de información del dominio temporal, puesto que al tomar como resolución de la FFT un número mayor al tamaño de ventana, lo único se hace es añadir ceros a la señal original en el tiempo.

Fijando la atención en el dominio del tiempo, la señal se ajusta más a la señal original cuando el tamaño de ventana es pequeño ($W = 32, 64, 128$), mientras que a medida que la ventana se va haciendo más y más grande, la señal empieza a quedar menos ajustada a la original y la limpieza es más drástica; tanto que la amplitud de los picos se reduce más de lo deseado.

De hecho al escoger tamaños de ventana grandes, en los entornos de los picos aparecen transitorios que distorsionan la señal, ya que la ventana está abarcando gran parte de dicha irregularidad, lo cual da lugar a perturbaciones de amplitudes significativas. La disminución de la amplitud de los picos más pronunciados complica la localización de estos dentro de la señal, tal y como se verá en la sección 4.4.

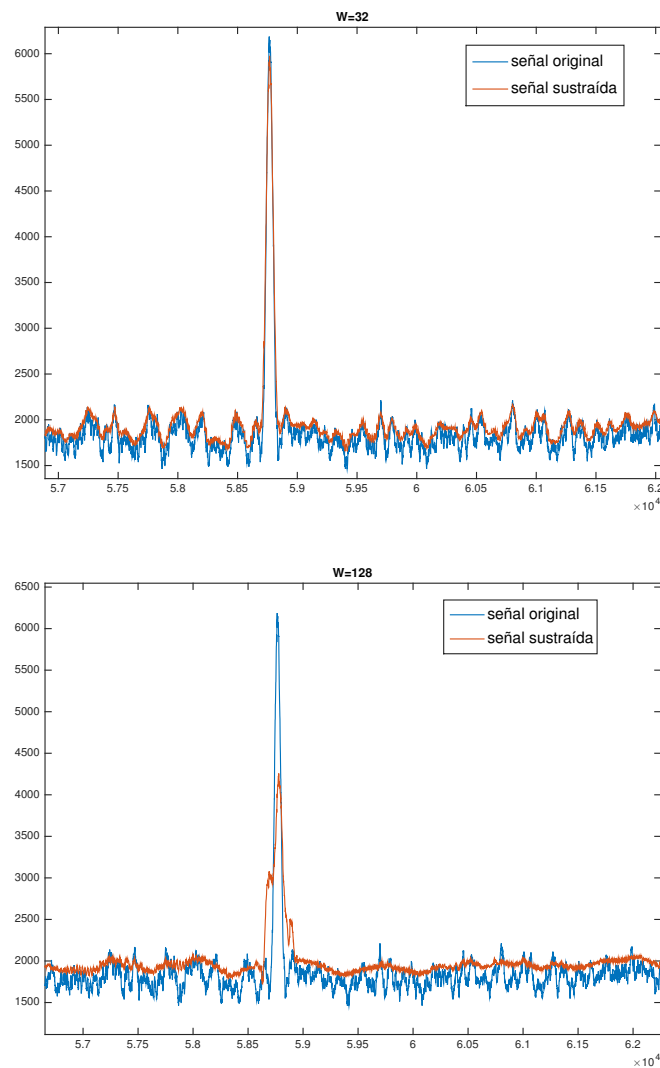


Figura 4.3: Sustracción espectral sobre la señal de densidad con $W = 32, 128$.

Todo esto se puede ver en la figura 4.3.

Observando el espectro de frecuencia durante la sustracción espectral en las ventanas del pico mostrado en la figura 4.3, se puede apreciar que un tamaño de ventana pequeño, lo cual también supone que la resolución de la FFT será menor en cada ventana, da lugar a que la sustracción espectral trabaje con una señal con menor número de muestras en el espectro de frecuencia. Atendiendo a la implementación se sabe que la frecuencia asociada a $\omega = 0$ nunca se sustrae, es decir, la frecuencia que determina el offset de la señal se conserva. Por tanto la señal sustraída siempre tendrá una amplitud muy similar a la señal original.

Tomando ventanas grandes la resolución es mayor y la limpieza es más contundente, ya que existen más muestras en frecuencia a ser sustraídas/limpiadas, pero la señal sustraída no está tan ajustada a la original. Véase la figura 4.4 para visualizar las diferencias de resolución en el espectro de frecuencia al variar el tamaño de ventana.

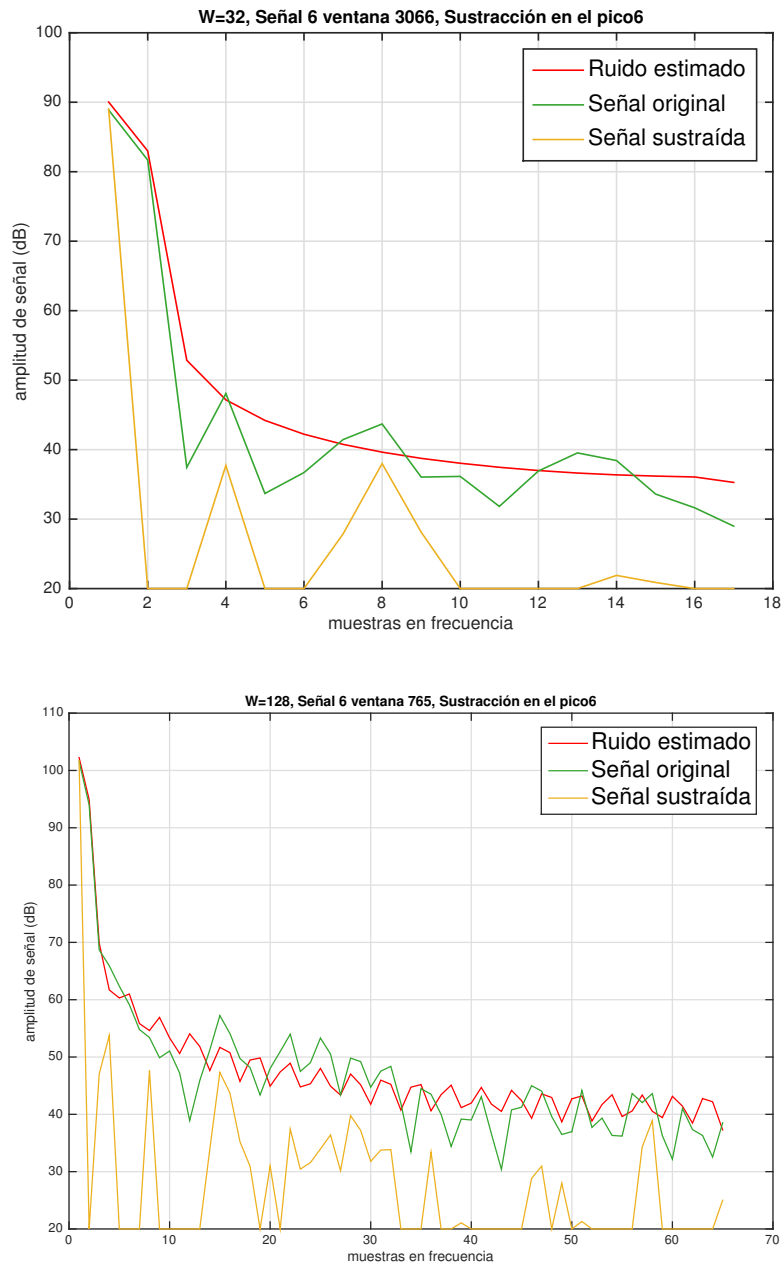


Figura 4.4: Visualización de la estimación de ruido, y de la señal antes y después de la sustracción espectral durante la primera ventana del 6 pico mostrado en la figura 4.3. En la figura de arriba el tamaño de ventana escogido es $W = 32$, y en la de abajo es $W = 128$.

4.2 Comparación de medida de SNR

Para poder cuantificar la disminución de ruido frente a la señal original sin ruido, se hace uso de lo que se conoce como coeficiente SNR (signal-to-noise-ratio), el cual no es más que una fracción entre la potencia de una señal y la potencia del ruido presente:

$$SNR = \frac{S}{N} \quad (4.1)$$

donde S es la potencia de la señal original sin ruido, y N es la potencia del ruido. En el caso de las señales estudiadas en este trabajo no se dispone de la señal original ni de la modelización del ruido, sin embargo dado que conocemos las etiquetas donde la señal presenta las irregularidades (picos), se puede tomar las zonas sin picos como exclusivamente ruido (N), y las zonas con picos como señal más ruido ($X = S + N$). De este modo tras descomponer la señal original como $S = X - N$, se puede llevar a cabo el cálculo del coeficiente SNR como:

$$SNR = \frac{S}{N} = \frac{X}{N} - 1 \quad (4.2)$$

Este proceso se realiza tras eliminar el offset de la zona activa, el cual se estima mediante el promediado del ruido en toda la zona activa de las 50 señales de densidad exentas de irregularidades. En las 22 señales con irregularidades, el SNR promedio es de 168.2361. En la tabla 4.1 se muestra el SNR promedio obtenido tras pasar dichas señales por la sustracción espectral.

	original	$W = 16$	$W = 32$	$W = 64$	$W = 128$	$W = 256$	$W = 512$	$W = 1024$
SNR	168.2	199.7	230.1	230.4	119	50	58.4	24.3

Tabla 4.1: SNR promedio de la señal original sin sustraer, y de las 22 señales de prueba tras la sustracción espectral con distintos tamaños de ventana.

El aumento del SNR indica que la presencia de la señal a detectar, es decir los picos que denotan irregularidades, aumenta frente al ruido. Por tanto los resultados obtenidos en la sustracción espectral indican que los tamaños de ventana entre 16 y 64 muestras, suponen una limpieza de la señal considerable; mientras que la elección de ventanas mayores baja tanto la amplitud de los picos, que la presencia de la señal empeora drásticamente frente al ruido.

En el caso del filtrado usando Chebyshev (véanse resultados en la tabla 4.2), los mejores resultados obtenidos son para frecuencias de corte situadas en los primeros ceros de la transformada del triángulo. Esto se debe a que al tomar ceros más alejados como frecuencia de corte, se filtran menos las frecuencias de la señal original. De hecho al tomar como frecuencia de corte, los n -ésimos ceros con $n > 25$, el SNR promedio de las 22 señales con irregularidades, se acerca al SNR de la señal original.

	original	$n = 1$	$n = 2$	$n = 4$	$n = 6$	$n = 8$	$n = 12$	$n = 14$
SNR	168.2	173.2	180.2	175.7	173.6	172.3	170.9	170.4

Tabla 4.2: SNR promedio de la señal original sin filtrar, y de las 22 señales de prueba tras su limpieza con un filtro de Chebyshev de tipo dos, sexto grado, caída de $40dB$, y frecuencia de corte situada en el n -ésimo cero de la transformada de un triángulo.

4.3 Separación del ruido

Para medir la disminución del solapamiento existente entre señal y ruido, se ha hecho uso de curvas DET e histogramas (véase la figura 4.5). En estos últimos se ha representado

en el eje x la amplitud de la señal, y para sacar el histograma de las zonas de ruido se ha hecho uso de todas las secciones que no están etiquetadas como picos en las 22 regiones de zona activa de las señales de densidad fuera de especificación. Así mismo, para sacar el histograma de las zonas de señal, se han tomado las regiones etiquetadas como picos.

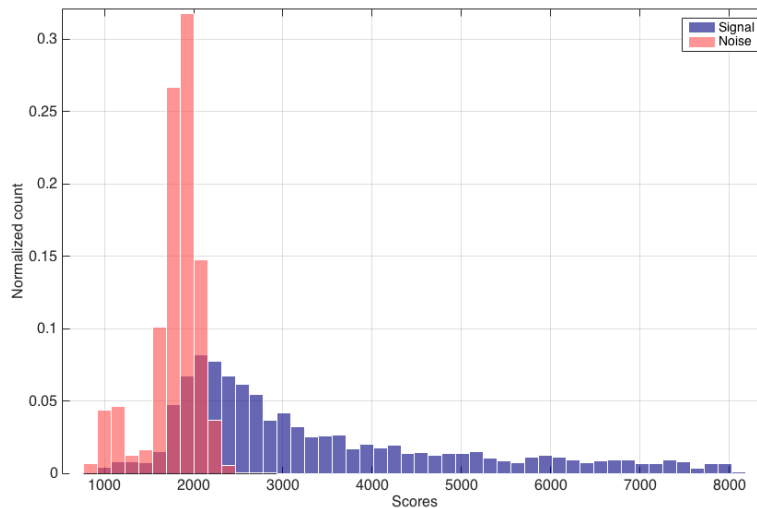


Figura 4.5: Histograma de amplitudes de ruido y señal presente en las 22 señales de densitómetro fuera de especificación. En rojo el ruido, y en azul la señal.

Una vez obtenidos estos histogramas se ha realizado un barrido de umbrales con el objetivo de tomar como ruido las amplitudes a la izquierda de este umbral, y como señal las amplitudes a la derecha de este umbral. Cada umbral dará una tasa de falsos positivos (amplitudes etiquetadas como señal siendo ruido), y de falsos negativos (amplitudes etiquetadas como ruido siendo señal). Estos porcentajes quedan recogidos por la curva DET, la cual está formada por puntos asociados a cada umbral donde la coordenada x establece el número de falsos positivos para dicho umbral, y la coordenada y el número de falsos negativos.

Con estas curvas se pretende encontrar el umbral que da lugar al mismo número de falsos positivos y falsos negativos. Este umbral se conoce como EER (Equal Error Rate). Que el EER baje supone una mejora en la detección, puesto que esto significa que el umbral óptimo devuelve un menor porcentaje de falsos positivos y negativos. Del mismo modo, el hecho de que la curva DET se acerque al origen significa que los múltiples umbrales tomados para su elaboración dan lugar a una menor tasa de falsos positivos y falsos negativos.

En la figura 4.6 se pueden ver los resultados obtenidos sobre las 22 señales de densidad de componentes fuera de especificación. En dicha figura se puede observar cómo la sustracción espectral obtiene mejores EER que la señal original cuando se toman tamaños de ventana menores o iguales a $W = 256$. De hecho en los casos de $W = 32$ y $W = 64$ las curvas DET quedan siempre por debajo de la de la señal original, lo cual significa que para todos los umbrales se obtiene una menor tasa de falsos positivos y negativos. Del mismo modo en la figura 4.6 se ve cómo la curva DET del filtro de Chebyshev queda por encima de la de la señal original (al igual que su EER).

4.4 Detección de picos tras el filtrado de la señal

Para poder evaluar si mejora la detección de picos tras la sustracción espectral, es necesario tener una forma de medir la eficacia de la detección de picos. Para ello se ha tenido

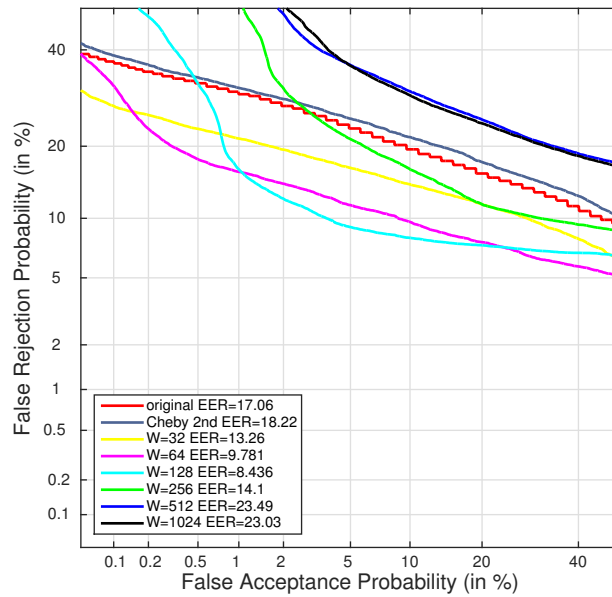


Figura 4.6: Curva DET y tasa EER para las 22 señales de densidad fuera de especificación, antes y después de la limpieza con diferentes tamaños de ventana en la sustracción espectral, y tras el uso de un filtro de Chebyshev de tipo 2, sexto grado, y caída de 40dB.

en consideración los picos etiquetados en la señal de densidad, es decir, para cada una de las irregularidades que aparecen en las 22 señales de densidad de los objetos que pasan por el control de calidad, se han etiquetado el principio y fin de los picos asociados (en la figura 3.2 estos intervalos aparecen en rojo).

En la detección de picos existen varios tipos de fallos que dependen del etiquetado de los picos asociados a las irregularidades de la señal de densidad:

- Falso positivo grave.
- Falso positivo leve.
- Pico no detectado grave.
- Pico no detectado leve.
- Pico demasiado próximo a otro detectado.

Los falsos positivos son los picos que son detectados según el procedimiento implementado en la 3.4, pero que no están etiquetados como picos. Otro tipo de error es aquel en que un pico etiquetado no es detectado. Y el último tipo de error sucede cuando un pico dentro de los límites de otro pico (estando muy próximo) es considerado como pico diferente, cuando ambos deberían detectarse como una sola irregularidad.

Para poder cuantificar cada aparición de este tipo de errores, se ha utilizado la siguiente función de coste:

$$f(\text{pico}) = \begin{cases} 5, & \text{si falso negativo grande} \\ 4, & \text{si falso negativo pequeño} \\ 3, & \text{si falso positivo grande} \\ 2, & \text{si falso positivo pequeño} \\ 1, & \text{si vecino extra} \end{cases} \quad (4.3)$$

4. RESULTADOS

Una vez las señales son sustraídas, estas bajan en amplitud, por ello es necesario encontrar el umbral de detección para cada tamaño de ventana. En cuanto se obtiene dicho umbral se evalúa la detección de picos en las 22 posibles sustracciones de las señales de densidad de componentes fuera de especificación. Tras ello se utiliza la función de coste 4.3 sobre cada señal, y finalmente se obtiene el promedio de todos los costes.

Escogiendo el tamaño de ventana adecuado (véase la tabla 4.3) se consigue mejorar el rendimiento del algoritmo de detección de picos cuando este actúa sobre las 22 señales sin filtrar (en este caso se obtiene un coste promedio de 1.1364).

	original	$W = 32$	$W = 64$	$W = 128$	$W = 256$	$W = 512$	$W = 1024$
coste promedio	1.1	0.7	1.1	1.6	8.7	14	30
umbral óptimo	2574	2419	2294	2178	2147	2143	2247

Tabla 4.3: Coste promedio y umbrales óptimos en la detección de picos de la señal original, y de las 22 señales del densitómetro tras las sustracción espectral con distintos tamaños de ventana.

Con el filtro de Chebyshev de tipo 2 no sucede lo mismo que en la sustracción espectral, puesto que el filtrado asociado a la mayor reducción de ruido mostrado en la tabla 4.2 (correspondiente al filtrado con frecuencia de corte en el segundo cero), no está asociado con el mejor coste promedio en detección de picos (véase en la tabla 4.4). De hecho el coste disminuye a medida que se toman frecuencias de corte mayores, hasta el punto en que se obtiene el mismo coste que en las señales originales, ya que cuanto mayor es la frecuencia de corte, más similar será la señal filtrada a la original.

	original	$n = 1$	$n = 2$	$n = 4$	$n = 6$	$n = 8$	$n = 12$	$n = 14$
coste promedio	1.1	3.4	2.4	1.8	1.7	1.7	1.7	1.6
umbral óptimo	2574							

Tabla 4.4: Coste promedio en la detección de picos de la señal original, y de las 22 señales del densitómetro tras la limpieza con un filtro de Chebyshev de tipo 2, sexto grado, caída de 40dB, y frecuencia de corte en el n -ésimo cero de la transformada de un triángulo.

CONCLUSIONES

Queda claro tras terminar la exposición de resultados, que la sustracción espectral consigue mejores resultados en los SNR obtenidos, y también consigue mejores resultados en la detección de picos tras la limpieza de la señal. Esto se debe a que los mejores resultados obtenidos para el filtrado de Chebyshev, se obtienen cuando la frecuencia de corte va aumentando, y haciendo esto lo único que se consigue es dejar la señal original prácticamente intacta, menos en altísimas frecuencias que han resultado ser insignificantes en la limpieza de las señales.

A la vista de los resultados obtenidos, en la figura 4.6 se puede observar que la separación entre ruido y señal original mejora notoriamente al tomar ventanas de tamaños menores o iguales a $W = 128$. De hecho con tamaños de ventana de $W = 32, 64$ muestras, cualquier umbral escogido como criterio para discernir entre señal o ruido, produce menos falsos positivos y negativos que en las señales originales. Además la obtención de una tasa $EER = 8.436$ en el caso $W = 128$ supone una reducción del 49.45 % frente a al $EER = 17.06$ obtenido en las 22 señales de densidad originales fuera de especificación.

Los resultados obtenidos para la función de coste que evalúa la detección de picos 4.3, podrían dar a pensar que son contradictorios a la vista de los resultados obtenidos para la separación de ruido. Ya que para la separación de ruido, es el tamaño de $W = 128$ muestras el que obtiene mejor tasa EER , sin embargo con tamaños de ventana $W = 32$ se consigue el mejor coste promedio en la detección de picos.

Esto tiene una explicación, y es que hay que tener en consideración que la función de coste presentada en 4.3 penaliza más los falsos negativos (peak misses), que los falsos positivos; de tal forma que los costes óptimos corresponden con los umbrales con menor número de falsos negativos, es decir, con los puntos que se encuentran a la derecha del todo de la curva DET de la figura 4.3, puesto que estos puntos corresponden con los valores de menor porcentaje de falsos negativos. En dicha zona se puede apreciar como la curva de $W = 128$ está mínimamente encima de la de $W = 32$, lo cual significa que la sustracción espectral es ligeramente peor tomando dicho tamaño de ventana.

Además también hay que tener en cuenta que las penalizaciones de la función de coste vienen dadas en función de si el falso positivo/negativo tiene mayor o menor amplitud. Esto explicaría el por qué $W = 64$ tiene su curva DET por debajo de la de $W = 32$, aunque para este último tamaño de ventana se haya encontrado menos coste promedio que para el de $W = 64$, tal y como se ve en la tabla 4.3.

5.1 *Trabajo futuro*

- Estudiar los motivos de la aparición de picos periódicos al tomar resoluciones de la FFT mayores al tamaño de ventana.
- Estudiar cómo influye el tamaño de falsos positivos/negativos que hacen que curvas DET de tamaños de ventana con mejor coste promedio, queden por encima de tamaños de ventana con pero coste promedio.
- Hacer un estudio más exhaustivo de los parámetros involucrados.
- Publicar el trabajo, pues los resultados obtenidos tienen cierta relevancia.

GLOSARIO

Autofunciones funciones que permanecen invariantes al aplicarles un operador. [VI, 3](#)

Convergencia uniforme tipo de convergencia más fuerte que la puntual. [VI, 5](#)

DET Detection Error Tradeoff. [27](#)

DFT Discrete Fourier Transform. [VI, 8](#)

DTFT Discrete Time Fourier Transform. [VI, 7](#)

EER Equal Error Rate. [28](#)

espectro de frecuencia dominio en el que se maneja la transformada de Fourier. [1](#)

FFT Fast Fourier Transform. [VII, 39](#)

frecuencia fundamental frecuencia más baja a la que oscila una onda. [6](#)

ganancia módulo de la respuesta al impulso de un filtro. [9](#)

polos del filtro frecuencias en las que la ganancia del filtro se dispara a infinito. [9](#)

Series de Fourier descomposición de una función como suma de sinusoides. [VI, 3](#)

SNR Signal to Noise Reduction. [VII, 27](#)

Transformada de Fourier descomposición de una función en las frecuencias que contiene. [VI, 6](#)

BIBLIOGRAFÍA

- [1] T. K. M. O. M. P. A. Martin, G. Doddington. The det curve in assessment of detection task performance. 1997.
- [2] S. F. Boll. Supression of acoustic noise in speech using spectral subtraction. 1979.
- [3] A. V. Oppenheim. *Signals And Systems*. Prentice Hall, 1983.
- [4] A. V. Oppenheim. *Discrete-Time Signal Processing*. Prentice Hall, 1989.
- [5] J. Peyton Z. Peebles. *Probability, Random Variables and Random Signal Principles*. Mc Graw Hill, 2001.
- [6] B. A. Shenoi. *Introduction to Digital Signal Processing and Filter Design*. Wiley-Interscience, 2005.
- [7] W. A. Strauss. *Partial Differential Equations, An Introduction*. Wiley, 2008.

ANEXOS

A Serie de Fourier de señales periódicas discretas

Considerando una secuencia de valores periódica $\tilde{x}[n]$ con periodo N , tal que $\tilde{x}[n] = \tilde{x}[n+rN]$ con r entero. Si se quiere representar dicha secuencia como una combinación lineal de funciones exponenciales (lo que precisamente hace la serie de Fourier), no se pueden tomar exponentes puramente continuos del tipo $e^{j\omega}$ como se hacía en la transformada para funciones continuas.

Para poder tener las frecuencias discretizadas se toman exponenciales que sean múltiplos de $2\pi/N$. De este modo se obtienen exponenciales del tipo:

$$e_k[n] = e^{j\frac{2\pi}{N}kn} = e_k[n+rN] \quad (1)$$

estas exponenciales presentan la propiedad de tener periodo N . Si ahora se toma la serie de Fourier de $\tilde{x}[n]$, la expresión debe tener el siguiente aspecto:

$$\tilde{x}[n] = \sum_{k=-\infty}^{\infty} \tilde{X}[k] e^{j\frac{2\pi}{N}kn}$$

La periodicidad de las exponenciales tomadas hace posible la representación mediante los N primeros términos:

$$\tilde{x}[n] = \sum_{k=0}^{N-1} \tilde{X}[k] e^{j\frac{2\pi}{N}kn} \quad (2)$$

Para calcular los coeficientes $\tilde{X}[n]$ se hace uso de la ortogonalidad de las exponenciales, al igual que el planteamiento expuesto en 2.1.2. Lo primero que se hace es multiplicar ambos lados de 2 por $e^{-j(2\pi/N)rn}$ y sumar sobre n :

$$\sum_{n=0}^{N-1} \tilde{x}[n] e^{-j\frac{2\pi}{N}rn} = \sum_{n=0}^{N-1} \frac{1}{N} \sum_{k=0}^{N-1} \tilde{X}[k] e^{j(2\pi/N)(k-r)n} = \sum_{k=0}^{N-1} \tilde{X}[k] \left[\frac{1}{N} \sum_{n=0}^{N-1} e^{j\frac{2\pi}{N}(k-r)n} \right]$$

La suma que aparece en el término entre corchetes será 1 cuando $k = r$, y en otro caso la suma será nula debido a que se toman raíces sobre el círculo unidad. De modo que el único término que cuenta en la suma más externa es en el que k vale exactamente r , y por tanto queda que:

$$\tilde{X}[r] = \sum_{n=0}^{N-1} \tilde{x}[n] e^{-j\frac{2\pi}{N}rn} \quad (3)$$

Frecuentemente se suele abreviar $e^{-j(2\pi/N)n}$ como W_N .

B Muestreo de la DTFT de una señal aperiódica

Este anexo es el eslabón que conecta el desarrollo en serie de Fourier de una señal discreta y periódica del anexo A, con la obtención de la transformada discreta de Fourier expuesta en 2.1.7.

Tomando $x[n]$ como señal aperiódica, con 2.18 se obtiene su transformada DTFT $X(e^{j\omega})$. El siguiente paso es muestrear la DTFT tomando frecuencias separadas a una distancia $2\pi/N$:

$$\tilde{X}[k] = X(e^{j\frac{2\pi}{N}k}) \quad (4)$$

Como la DTFT de una señal tiene periodo 2π , consecuentemente $\tilde{X}[k]$ tendrá periodo N . Los puntos sobre los que se muestrea $X(e^{j\omega})$ se encuentran en la circunferencia unidad, tal y como se puede observar en la figura .1. Puesto que la secuencia $\tilde{X}[k]$ obtenida tras

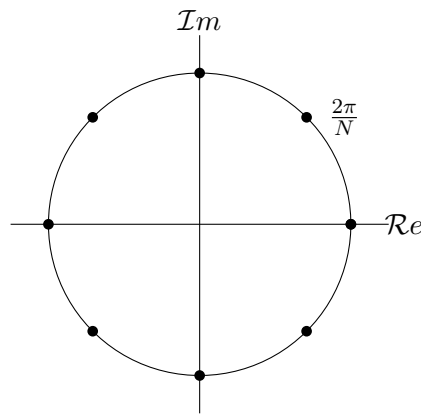


Figura .1: 8 raíces de la unidad.

el muestreo es periódica, puede ser precisamente la secuencia de coeficientes de Fourier de una secuencia periódica a la que se denotará como $\tilde{x}[n]$.

Ahora mismo está por ver qué es exactamente $\tilde{x}[n]$. Habiendo supuesto que a secuencia de 4 son coeficientes de Fourier de $\tilde{x}[n]$, es lógico escribir la señal periódica como en 3:

$$\tilde{x}[n] = \frac{1}{N} \sum_{k=0}^{N-1} \tilde{X}[k] W_N^{-kn}$$

Suponiendo que la transformada de Fourier DTFT de $x[n]$ exista, su expresión será:

$$X(e^{j\omega}) = \sum_{m=-\infty}^{\infty} x[m] e^{-j\omega m}$$

Llevando a cabo el muestreo sobre la DTFT, y sustituyendo los coeficientes en el supuesto desarrollo en serie de Fourier, se obtiene:

$$\tilde{x}[n] = \frac{1}{N} \sum_{k=0}^{N-1} \left[\sum_{m=-\infty}^{\infty} x[m] e^{-j\frac{2\pi}{N}km} \right] W_N^{-kn} = \quad (5)$$

$$= \sum_{m=-\infty}^{\infty} x[m] \left[\frac{1}{N} \sum_{k=0}^{N-1} W_N^{-k(n-m)} \right] = \quad (6)$$

$$= \sum_{m=-\infty}^{\infty} x[m] \tilde{p}[n-m] \quad (7)$$

La expresión que figura entre corchetes en 7 es la representación en serie de Fourier de un tren de impulsos (véase [4]), y $\tilde{p}[n]$ denota un tren de impulsos centrado en 0 con periodo N , tal y como se puede apreciar en la figura .2.

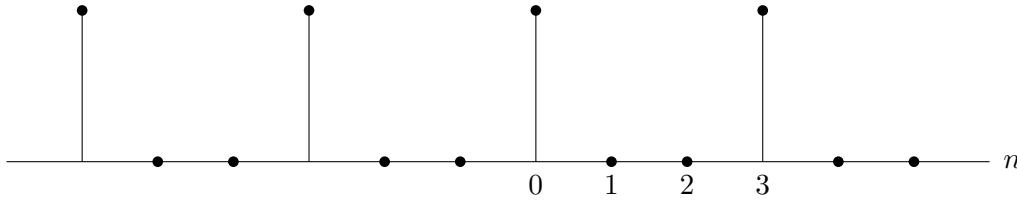


Figura .2: Tren de impulsos con periodo $N = 3$.

Por otra parte la operación que aparece en la ecuación 7 resulta ser la convolución de la señal $x[n]$ con el tren de impulsos $\tilde{p}[n]$. Es bien conocido (véase [3]) que la convolución con un tren de impulsos resulta replicar una señal, centrando cada una de las réplicas cada periodo del tren de impulsos, es decir:

$$\tilde{x}[n] = x[n] * \tilde{p}[n] = x[n] * \sum_{r=-\infty}^{\infty} \delta[n - rN] = \sum_{r=-\infty}^{\infty} x[n - rN] \quad (8)$$

De modo que el muestreo de la DTFT da lugar a la repetición de la señal original cada cierto periodo N marcado por la precisión del muestreo, tal y como se puede apreciar en la figura .3.

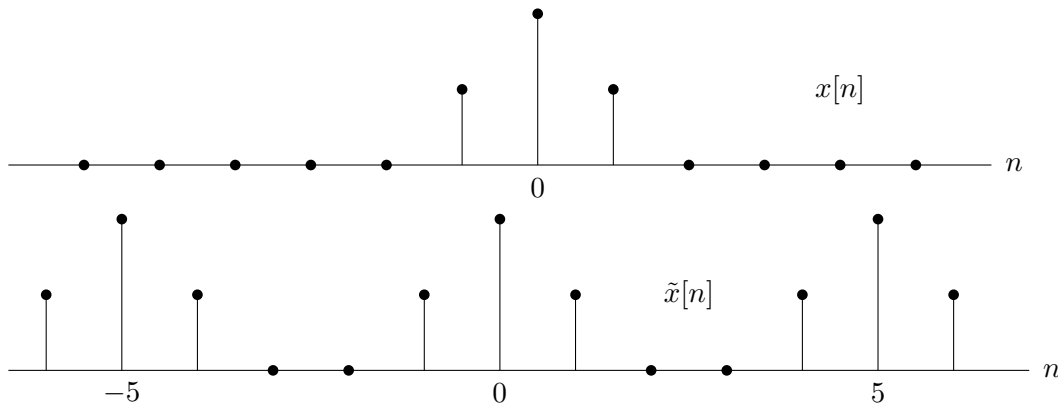


Figura .3: Señal antes y después del muestreo de su DTFT.

C Cálculo de la FFT

Para calcular la DFT de una señal discreta se recurre a algoritmos que reducen el número de operaciones mediante diversos procedimientos, como por ejemplo los métodos de diezmado en el tiempo. En estos métodos se expresa la periodicidad de las exponenciales complejas W_N^{kn} , y se asume que el número de muestras de la señal a transformar es una potencia de 2, es decir $N = 2^m$.

Lo primero que se hace es separar el cálculo de la DFT en los índices pares e impares del sumatorio:

$$X[k] = \sum_{r=0}^{(N/2)-1} x[2r]W_N^{2rk} + \sum_{r=0}^{(N/2)-1} x[2r + 1]W_N^{(2r+1)k} \quad (9)$$

Si en el último sumatorio se saca W_N^k como factor común, se obtiene:

$$X[k] = \sum_{r=0}^{(N/2)-1} x[2r]W_N^{2rk} + W_N^k \sum_{r=0}^{(N/2)-1} x[2r+1]W_N^{2rk}$$

Por último ya solo queda darse cuenta de que

$$W_N^2 = e^{-2j\frac{2\pi}{N}} = e^{-j\frac{2\pi}{N/2}} = W_{N/2} \quad (10)$$

para terminar reescribiendo la DFT como:

$$X[k] = \sum_{r=0}^{(N/2)-1} x[2r]W_{N/2}^{rk} + W_N^k \sum_{r=0}^{(N/2)-1} x[2r+1]W_{N/2}^{rk} = E[k] + W_N^k O[k] \quad (11)$$

En esta última ecuación $E[k]$, $O[k]$ hacen referencia, respectivamente, a las DFT de los coeficientes pares e impares de la señal $x[n]$. En el diagrama presentado en la figura .4 se muestra el funcionamiento de la FFT sobre una señal $x[n]$ de $N = 8$ muestras. En la imagen los puntos significan operaciones de multiplicación, y las cajas hacen alusión a las operaciones necesarias para calcular las DFT de los términos pares e impares.

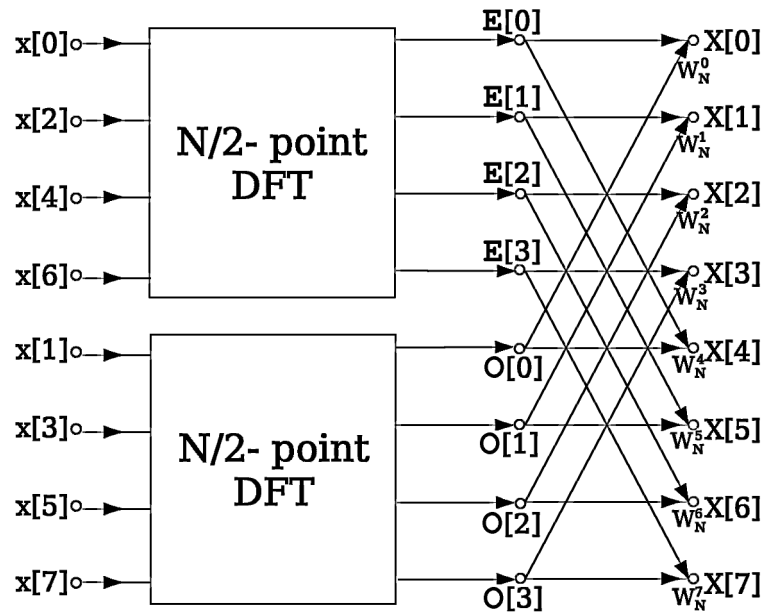


Figura .4: Diagrama del funcionamiento de la FFT para una señal con $N = 8$ muestras.

Si la DFT se computase sin seguir ningún procedimiento para ahorrar el número de operaciones, habría que realizar N^2 productos y $N(N - 1)$ sumas. A la hora de decir el orden de operaciones se suelen contar exclusivamente las multiplicaciones, por ello se dice que calcular la DFT de una señal con N muestras tiene coste $\mathcal{O}(N^2)$.

En el ejemplo de la FFT mostrado en la figura .4 se realizan:

$$2 \left(\frac{N}{2} \right)^2 + N$$

multiplicaciones. $(N/2)^2$ multiplicaciones para el cálculo de cada DFT de $N/2$ muestras (los coeficientes pares e impares), y otras N multiplicaciones para combinar los coeficientes pares e impares. Si se volviera a dividir el cálculo de las transformadas de $N/2$ puntos,

habría que hacer

$$2 \left(2 \left(\frac{N}{4} \right)^2 + \frac{N}{2} \right) + N = \frac{N^2}{4} + 2N$$

multiplicaciones. A lo sumo se pueden hacer $p = \log_2 N$ divisiones de la operación, y llegado a dicho número de divisiones se tienen que hacer

$$\frac{N^2}{2^p} + pN = N + N \log_2 N$$

multiplicaciones, y por tanto el cálculo de la FFT es de orden $\mathcal{O}(N \log N)$ frente al orden $\mathcal{O}(N^2)$ necesario para calcular la DFT.

D Otros pasos utilizados en la versión original de la sustracción espectral

En esta sección se presentan dos de los pasos que se sugieren en [2] para conseguir la atenuación de ruido cuando no hay presencia de voz, y para la actualización de ruido cuando este varía con el tiempo.

D.1 Atenuación de la señal cuando no hay voz

Tras llevar a cabo el paso descrito en 2.3.5, [2] determina que nos encontramos en una sección que tiene exclusivamente ruido, cuando:

$$T = 20 \log_{10} \left[\int_{-\pi}^{\pi} \left| \frac{S(e^{j\omega})}{\mu(e^{j\omega})} \right| d\omega \right]$$

es menor que -12 dB. Por tanto en esas frecuencias se puede decidir llevar a cabo una atenuación multiplicando mediante un factor de -30 dB, tal y como establece [2]:

$$\hat{S}(e^{j\omega}) = \begin{cases} \hat{S}(e^{j\omega}) & , T \geq -12 \text{ dB} \\ c\hat{S}(e^{j\omega}) & , T < -12 \text{ dB} \end{cases}$$

donde $20 \log_{10} c = -30$ dB.

D.2 Actualización de la estimación del ruido

Al comienzo de [2] se recalca la asunción de que el ruido es estacionario, Sin embargo si el ruido varía, es necesario volver a realizar estimaciones según va avanzando el tiempo.

Para llevar a cabo la actualización de la estimación de ruido, una posible opción (la que se utiliza en este trabajo) es la siguiente:

1. Detectar si en la ventana actual tenemos presencia de voz. Esto se consigue realizando el cálculo de la **distancia espectral**, que no es más que el valor T calculado en la subsección anterior D.1. Una vez calculado se establece que si T cae debajo de un umbral (que en [2] se establece como -12 dB), es porque en dicha ventana hay exclusivamente ruido.
2. En caso de estar ante una ventana sin voz, **actualizamos la estimación de ruido** de la siguiente manera:

$$\mu'(e^{j\omega}) = \frac{|\hat{S}_i(e^{j\omega})| + W \cdot \mu(e^{j\omega})}{W + 1}$$

Donde W es la longitud de la ventana, y el subíndice i indica la transformada de Fourier de la i -ésima ventana (con ventana nos referimos al producto de un intervalo de la señal original, por una ventana de Hanning o Hamming). Obsérvese que con esta operación se lleva a cabo un suavizado del ruido, y el primer factor introduce la información asociada al ruido presente en la ventana actual.

3. Tras actualizar nuestra estimación del ruido, también se **actualiza el ruido residual** N_R :

$$|N'_R(e^{j\omega})| = \text{máx} \left\{ |N_R(e^{j\omega})|, |\hat{S}_i(e^{j\omega})| - \mu'(e^{j\omega}) \right\}$$

Es decir, tomamos el máximo entre el anterior valor del ruido residual, y el valor estimado de ruido residual en la ventana actual. Recordemos que $N_R = N - \mu e^{j\theta N}$, y puesto que en la ventana actual no hay presencia de voz, tenemos que $\hat{S}_i(e^{j\omega}) = N(e^{j\omega})$.