
Tratamiento digital de señales analógicas

Introducción al diseño de filtros digitales y a las operaciones de cambio de frecuencia de muestreo

PID_00262133

José Antonio López Salcedo

Tiempo mínimo de dedicación recomendado: 3 horas



José Antonio López Salcedo

El encargo y la creación de este recurso de aprendizaje UOC han sido coordinados por el profesor: José Antonio Morán Moreno (2019)

Primera edición: febrero 2019
© José Antonio López Salcedo
Todos los derechos reservados
© de esta edición, FUOC, 2019
Av. Tibidabo, 39-43, 08035 Barcelona
Diseño: Manel Andreu
Realización editorial: Oberta UOC Publishing, SL

Ninguna parte de esta publicación, incluido el diseño general y la cubierta, puede ser copiada, reproducida, almacenada o transmitida de ninguna forma, ni por ningún medio, sea este eléctrico, químico, mecánico, óptico, grabación, fotocopia, o cualquier otro, sin la previa autorización escrita de los titulares de los derechos.

Índice

Introducción.....	5
Objetivos.....	7
1. Motivación del tratamiento digital de señales analógicas.....	9
2. Operaciones de diezmado e interpolación.....	12
2.1. Diezmado	13
2.2. Interpolación	15
2.3. Cambio no entero de frecuencia de muestreo	16
3. Introducción al filtrado digital.....	18
3.1. Especificaciones de un filtro digital	18
3.2. Diseño de filtros FIR	20
3.2.1. Método por muestreo en frecuencia	21
3.2.2. Método por enventanado	22
3.2.3. Método óptimo de rizado constante	25
3.3. Diseño de filtros IIR	26
3.3.1. Método de la transformación bilineal	26
3.4. Comparativa entre filtros FIR e IIR	29
4. Implementación del filtrado en bloques.....	31
4.1. Método <i>overlap-save</i>	31
4.2. Método <i>overlap-add</i>	33
Ejercicios de autoevaluación.....	35
Solucionario.....	37
Bibliografía.....	38

Introducción

El tratamiento digital de señales analógicas tiene sus raíces a principios de la década de 1980, coincidiendo con el auge de la integración a gran escala de circuitos integrados (VLSI) y el desarrollo de los primeros dispositivos digitales de cálculo. Estos dispositivos constituyeron una revolución al permitir programar en *software* sofisticadas operaciones que hasta aquel momento se llevaban a cabo de forma manual o mecanizada. La programación de dispositivos digitales aportó una gran flexibilidad al permitir modificar de forma sencilla las operaciones que se llevan a cabo, facilitar su repetitividad y, por tanto, su automatización de forma sistemática. Además, el uso de sistemas digitales aportó en general, una mayor fiabilidad y precisión en comparación con los sistemas analógicos. Finalmente, su bajo coste, debido a las economías de gran escala adoptadas en la industria electrónica, han hecho que hoy en día los dispositivos digitales estén presentes en la práctica totalidad de los sistemas que nos rodean, tanto a nivel profesional como de consumo.

Dentro de los sistemas digitales es importante destacar el papel que han jugado los procesadores digitales de señal (DSP). Estos dispositivos están específicamente diseñados para implementar de forma eficiente las operaciones básicas del tratamiento de señales y sistemas lineales, como por ejemplo la convolución, haciendo que operaciones que antes se llevaban a cabo mediante circuitos analógicos puedan ser implementadas de forma eficiente mediante DSP. Para ello es necesario llevar a cabo tres pasos:

- 1) Convertir las señales analógicas al dominio discreto en el que trabajan los dispositivos digitales.
- 2) Llevar a cabo el tratamiento digital de la señal mediante un sistema en tiempo discreto.
- 3) Convertir el resultado en una señal analógica, de forma que, extremo a extremo, el resultado sea equivalente a haber utilizado un sistema analógico.

Dejando de lado ciertas consideraciones de los sistemas digitales, como las posibles pérdidas de cuantificación o el hecho de que la señal de entrada tenga que estar limitada en banda, el comportamiento extremo a extremo del sistema depende principalmente de la elección de la frecuencia de muestreo y del sistema en tiempo discreto que se implemente. El análisis de estos dos aspectos ya ha sido objeto de estudio en asignaturas anteriores, con lo cual el objetivo del presente módulo es dar un paso más allá y centrarse en aspectos prácticos que tienen que ver con:

- Cómo cambiar la frecuencia de muestreo de una señal en tiempo discreto que ya ha sido muestreada (apartado 2).
- Cómo implementar un sistema lineal e invariante en tiempo discreto a partir de una serie de especificaciones de su respuesta frecuencial, dando lugar al diseño de filtros digitales (apartado 3), y cómo implementarlos de forma eficiente (apartado 4).

Este último aspecto comprende un campo muy amplio de estudio para el cual, lejos de dar un tratamiento exhaustivo, el presente módulo pretende dar una visión introductoria que permita al estudiante adquirir los conocimientos básicos.

Objetivos

Los objetivos de este módulo son los siguientes:

1. Comprender el interés del tratamiento digital de señales analógicas.
2. Entender la relación entre la respuesta de un sistema en tiempo continuo y su equivalencia en el dominio del tiempo discreto.
3. Conocer las operaciones de diezmado e interpolación para llevar a cabo cambios en la frecuencia de muestreo y comprender el impacto que tienen en el dominio frecuencial.
4. Conocer las alternativas para implementar sistemas en tiempo discreto, lineales e invariantes para una respuesta frecuencial dada.
5. Distinguir entre los diferentes tipos de filtros FIR en función de la longitud y simetría de su respuesta impulsional.
6. Conocer los parámetros clave a la hora de especificar la respuesta frecuencial de un filtro.
7. Aplicar los métodos básicos de diseño de filtros FIR y filtro IIR conociendo sus ventajas y sus limitaciones.
8. Conocer los métodos básicos de filtrado en bloque para implementar de forma eficiente la operación de filtrado de señales de duración muy larga con filtros FIR.

1. Motivación del tratamiento digital de señales analógicas

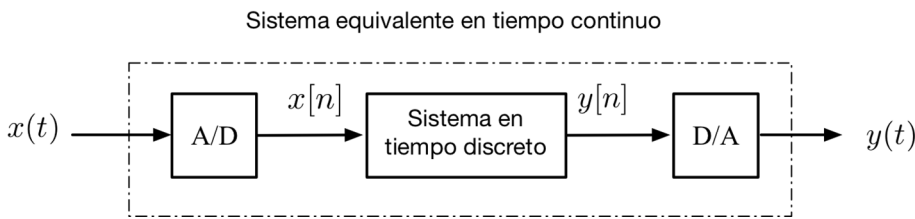
Existen multitud de aplicaciones donde es necesario procesar una señal y acondicionarla para que sea almacenada, transmitida o analizada con el fin de extraer cierta información de ella. Este es el caso del procesado de señales de audio, la recepción de señales de comunicaciones o el análisis de señales geofísicas, biomédicas o de sistemas de radar. En todos estos casos es habitual procesar las señales en cuestión mediante un sistema lineal e invariante que típicamente está destinado a reducir el nivel de ruido, eliminar contribuciones indeseadas (por ejemplo, interferencias) o seleccionar una cierta componente espectral de la señal recibida (por ejemplo, como ocurre en sistemas con acceso múltiple por división de frecuencia). Como se ha descrito en la introducción a este módulo, el tratamiento digital de estas señales ofrece grandes ventajas en términos de flexibilidad, sencillez de implementación, mayor fiabilidad y reducción de costes en comparación con el uso de técnicas y circuitos analógicos.

A nivel conceptual, el tratamiento digital de señales analógicas consta de tres pasos que implican:

- 1) convertir las señales analógicas al dominio en tiempo discreto;
- 2) llevar a cabo el tratamiento digital de la señal utilizando un sistema en tiempo discreto;
- 3) convertir el resultado del tratamiento digital de nuevo en una señal analógica, de forma que el resultado extremo a extremo sea equivalente a haber utilizado un sistema analógico.

Estos tres pasos se corresponden con los tres bloques que aparecen en la figura 1. Como puede observarse, el resultado equivalente a todo el proceso es una señal analógica de entrada $x(t)$ que es procesada y convertida en una señal analógica de salida $y(t)$, entendiendo las operaciones intermedias como equivalentes (dentro de una serie de condicionantes) a las que llevaría a cabo un sistema analógico.

Figura 1. Representación en forma de diagrama de bloques del tratamiento digital de señales analógicas y su equivalencia con un sistema en tiempo continuo



La limitación más importante de este esquema es el hecho de que la señal $x(t)$ tenga que estar limitada en banda; esto es, $X(\Omega) = 0$ para $|\Omega| \geq \pi/T$ con $X(\Omega)$ la transformada de Fourier de $x(t)$ y T el tiempo de muestreo. Como es conocido, este requisito es necesario para garantizar que su conversión al dominio en tiempo discreto no incurra en *aliasing*. De la misma forma, también se darán por supuesto las posibles pérdidas que introduce la cuantificación del conversor analógico-digital (A/D) y que hacen que la correspondencia entre un sistema analógico y su implementación digital no sea perfecta. En este caso, sin embargo, ajustando el número de bits del cuantificador o teniendo en cuenta las operaciones que se llevan a cabo en el dominio del tiempo discreto, es posible acotar el impacto de la cuantificación y hacer que pueda llegar a ser despreciable.

La importancia de la figura 1 reside en que, a fin de cuentas, una señal analógica de entrada $x(t)$ es convertida en otra señal analógica de salida $y(t)$. Por lo tanto, la respuesta global de entrada/salida puede entenderse como un sistema analógico, independientemente de que alguno de sus bloques intermedios estén implementados en tiempo discreto. De hecho, puede demostrarse que la señal de salida $y(t)$ viene dada por la convolución de la señal de entrada $x(t)$ mediante un sistema analógico equivalente $h_{eq}(t)$,

$$Y(\Omega) = H_{eq}(\Omega)X(\Omega) \quad (1)$$

con $\{X(\Omega), Y(\Omega), H_{eq}(\Omega)\}$ las transformadas de Fourier de las señales $\{x(t), y(t)\}$ y del sistema equivalente con respuesta impulsional $h_{eq}(t)$, respectivamente. Teniendo en cuenta la limitación en banda de $x(t)$, la respuesta frecuencial de este sistema equivalente ha de venir dada por:

$$H_{eq}(\Omega) = \begin{cases} H(e^{j\omega})_{\omega=\Omega T} & , \quad |\Omega| < \pi/T \\ 0 & , \quad |\Omega| \geq \pi/T \end{cases} \quad (2)$$

donde $H(e^{j\omega})$ es la respuesta frecuencial del sistema en tiempo discreto que se muestra en la figura 1. En definitiva, la señal de salida $y(t)$ puede verse como el resultado de procesar la señal de entrada $x(t)$ mediante un sistema en tiempo

Nota

Los pasos correspondientes a la conversión A/D y D/A han sido ya objeto de estudio en otras asignaturas, así como la caracterización y las propiedades de los sistemas en tiempo discreto. En este sentido, estos conceptos se darán por supuestos en el presente módulo.

continuo cuya respuesta frecuencial viene dada por (2). Este resultado conecta la respuesta del sistema que se desea en tiempo continuo con la respuesta del sistema que ha de implementarse en tiempo discreto.

La motivación de este módulo es profundizar en esa dirección y, por tanto, tratar acerca del diseño de sistemas en tiempo discreto que presentan el mismo comportamiento que un cierto sistema en tiempo continuo. En ocasiones este problema implica realizar operaciones previas que conllevan un cambio de la frecuencia de muestreo, problema que es tratado en el apartado 2. Una vez llevado a cabo, el diseño de sistemas lineales invariantes en tiempo discreto (filtros digitales) es tratado en el apartado 3 y su implementación práctica en el apartado 4.

2. Operaciones de diezmado e interpolación

En ocasiones resulta conveniente cambiar la frecuencia de muestreo con la que fue adquirida una señal, bien sea para reducirla o para incrementarla.

Considere por ejemplo una señal en tiempo continuo $x(t)$ con ancho de banda B_x que es muestreada a la frecuencia mínima de muestreo $F_m = 2B_x$ de acuerdo con el criterio de Nyquist. El resultado es una señal o secuencia en tiempo discreto $x[n]$. A continuación, dicha secuencia sufre un proceso de filtrado digital que da lugar a una nueva secuencia $x_f[n]$, cuyo ancho de banda B_f es muy inferior al de la señal original $x[n]$, por ejemplo, $B_f = B_x/10$. Se da la circunstancia de que la frecuencia de muestreo de la secuencia filtrada $x_f[n]$ es ahora muy superior a la que podría haber tenido si hubiera sido muestreada desde un inicio en base a su ancho de banda B_f . En ese caso, la frecuencia mínima de muestreo podría haber sido $F_m' = 2B_f = B_x/5 = F_m/10$, diez veces inferior a la frecuencia de muestreo inicial. La importancia de esta observación radica en que muchas implementaciones *hardware* son sensibles a la frecuencia de muestreo, puesto que esta tiene un impacto importante en términos de consumo de potencia, complejidad de implementación y, por tanto, coste. Es por ello que, ante un ejemplo como el que se acaba de ilustrar, se hace necesario poder hacer un cambio de frecuencia de muestreo para pasar de F_m a una F_m' mucho más pequeña a la salida del filtrado digital. Esta operación de reducción de frecuencia de muestreo, en general por un factor M , se conoce como diezmado.

Una situación como la que acaba de describirse ocurre por ejemplo en sistemas de comunicaciones basados en acceso por división de frecuencia (FDMA), donde la estación base necesita demultiplexar cada uno de los usuarios individuales (cuyo ancho de banda es pequeño) a partir de la señal que recibe con todos los usuarios multiplexados en frecuencia (cuyo ancho de banda es grande, puesto que depende del número de usuarios).

Por el contrario, existen situaciones donde más que reducir la frecuencia de muestreo, lo que interesa es aumentarla. Siguiendo con el ejemplo de un sistema FDMA, sería el caso de un transmisor que, a partir de las señales en tiempo discreto de cada usuario individual, ha de generar una señal en tiempo discreto de ancho de banda mucho mayor y que contenga a todos estos usuarios multiplexados en frecuencia. La operación de aumento de frecuencia de muestreo, en general por un factor L , se conoce como interpolación.

Ambas operaciones, diezmado e interpolación, serán objeto de estudio en este apartado.

2.1. Diezmado

El diezmado es una operación que reduce la frecuencia a la que ha sido muestreada una secuencia $x[n]$ pasando de $F_m = 1/T_m$ (o $\Omega_m = 2\pi/T_m$ en rad/s) a $F'_m = 1/MT_m$ (o $\Omega'_m = 2\pi/MT_m$ en rad/s). Partiendo de la secuencia $x[n]$,

$$x[n] = x(nT_m) \quad (3)$$

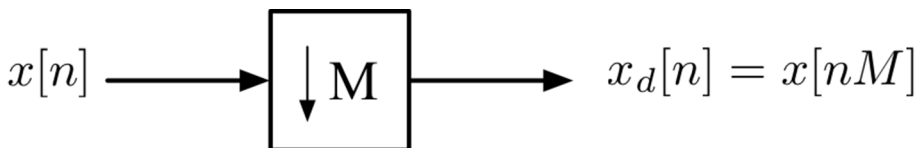
el resultado del proceso de diezmado es una nueva secuencia $x_d[n]$ según:

$$x_d[n] = x[Mn] = x(nMT_m) = x(nT_m). \quad (4)$$

donde puede observarse que el nuevo tiempo entre muestras de la señal original $x(t)$ pasa a ser $T'_m = MT_m$, lo cual corresponde efectivamente a una reducción de la frecuencia de muestreo por un factor M .

La operación de diezmado se suele representar en forma de diagrama de bloques según se indica en la figura a continuación:

Figura 2. Representación en forma de diagrama de bloques de la operación de diezmado



donde el bloque que realiza el cambio de frecuencia de muestreo se conoce como *compresor*.

A nivel temporal, el diezmado puede entenderse como una operación que genera una nueva secuencia $x_d[n]$ conservando una de cada M muestras de la señal original $x[n]$ y, por tanto, despreciando las $M - 1$ restantes. Por lo tanto, en el tiempo equivalente a M muestras de la secuencia original, la secuencia diezmada contiene solo una muestra, que es precisamente lo que ocurre al utilizar una frecuencia de muestreo M veces más pequeña.

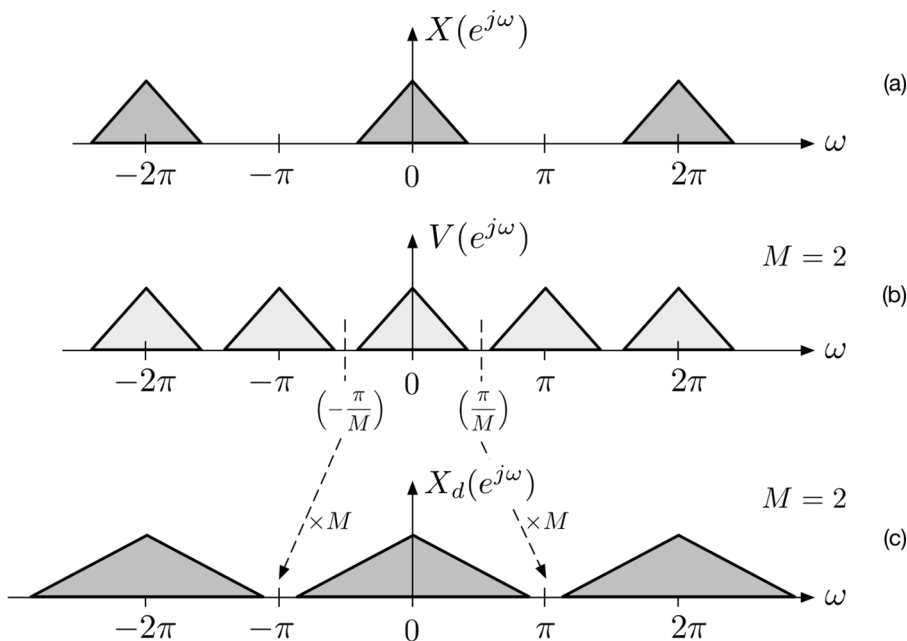
A nivel frecuencial, la operación de diezmado implica una serie de transformaciones que es interesante analizar con detalle, pues tiene importantes consecuencias en la secuencia de salida. La operación de diezmado se lleva a cabo en dos pasos, que implican:

1) Multiplicar la secuencia original $x[n]$ con un tren de deltas que selecciona una de cada M muestras. Ello da lugar a una nueva secuencia intermedia $v[n]$ cuya representación frecuencial se muestra en la figura 3b, la cual presenta en réplicas de la transformada de Fourier de la secuencia $x[n]$ a frecuencias múltiplos de $2\pi/M$.

2) Comprimir el eje temporal de la secuencia $v[n]$ para eliminar las muestras nulas que existen entre las muestras seleccionadas, haciendo, por tanto, que el tiempo entre muestras no nulas sea $T_m = MT_m$. En el dominio frecuencial, ello implica una expansión del eje de frecuencias que da lugar a la representación mostrada en la figura 3c, que se corresponde con la transformada de Fourier de la secuencia diezmada.

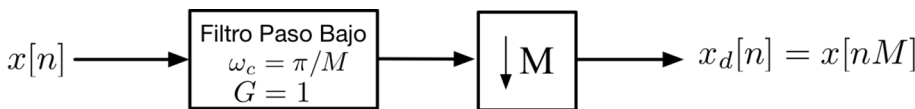
Es interesante hacer notar que, como resultado del proceso de diezrado, la secuencia resultante $x_d[n]$ presenta un ancho de banda que, en tiempo discreto, es M veces mayor que el de la secuencia inicial $x[n]$, tal y como se puede observar comparando la figura 5a y 5c.

Figura 3. Representación frecuencial de la operación de diezrado



Por otro lado, hay que tener en cuenta que, según se observa en la transformada de Fourier de la secuencia $v[n]$ (ver figura 3b), la aparición de réplicas a frecuencias múltiplos de $2\pi/M$ implica que el ancho de banda de la secuencia original $x[n]$ ha de estar limitado a π/M . De lo contrario, aparecerá *aliasing* en la secuencia intermedia $v[n]$ y, por lo tanto, también en la secuencia diezmada $x_d[n]$. En ciertas aplicaciones puede que la aparición de *aliasing* no sea algo muy crítico. Si no es el caso, entonces habrá que asegurarse de que la secuencia original $x[n]$ se filtre con un filtro paso bajo para asegurar que no hay componentes frecuenciales más allá de la frecuencia discreta $\omega = \pi/M$. En términos de la señal analógica $x(t)$, esto implica asegurar que no existen componentes frecuenciales más allá de la frecuencia $F = F_m/2M$ o, equivalentemente, $\Omega = \pi/MT_m$.

Figura 4. Representación en forma de diagrama de bloques de la operación de diezmado, incluyendo el filtro paso bajo para limitar en banda la secuencia a diezmar y evitar así el posible *aliasing*



Finalmente, es importante mencionar que, si bien la operación de diezmado es lineal y estable, no es sin embargo causal ni invariante. Ello ha de ser tenido en cuenta a la hora de analizar sistemas en tiempo discreto que hagan uso de esta operación.

2.2. Interpolación

La interpolación es una operación que implica el aumento de la frecuencia a la que ha sido muestreada una secuencia $x[n]$ pasando de $F_m = 1/T_m$ a $F'_m = L/T_m = 1/T'_m$ con T_m el tiempo de muestreo inicial y $T'_m = T_m/L$ el tiempo de muestreo final. Partiendo de la secuencia:

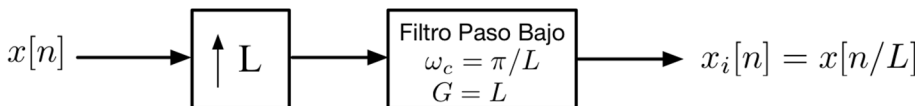
$$x[n] = x(nT_m) \quad (5)$$

el objetivo es, por tanto, llegar a una nueva secuencia:

$$x_i[n] = x[n/L] = x(nT_m/L) = x(nT'_m). \quad (6)$$

La operación de interpolación se representa en forma de diagrama de bloques según se indica en la figura a continuación:

Figura 5. Representación en forma de diagrama de bloques de la operación de interpolación



A nivel temporal, la nueva secuencia $x_i[n]$ se genera introduciendo $L - 1$ nuevas muestras entre cada una de las muestras de la señal inicial $x[n]$. El proceso se resume en los siguientes pasos:

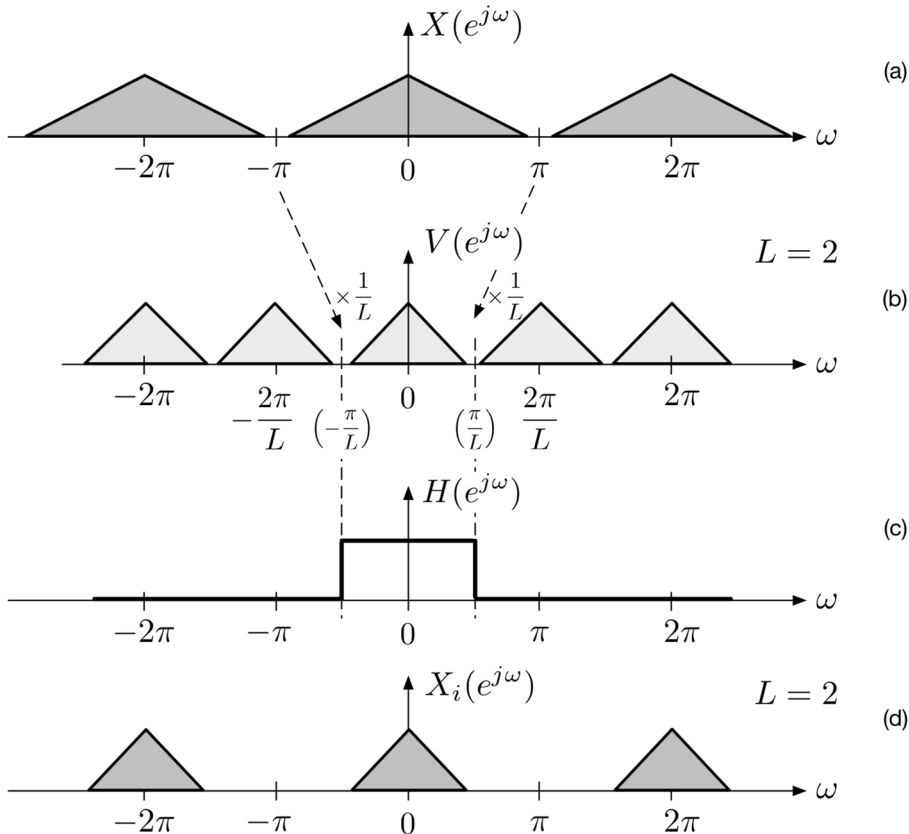
1) Se expande el eje temporal introduciendo $L - 1$ nuevas muestras de valor cero entre cada muestra de la secuencia $x[n]$, lo que hace que el nuevo tiempo de muestreo sea $T'_m = T_m/L$. Este paso es equivalente a generar una secuencia

intermedia $v[n] = \sum_{k=-\infty}^{\infty} x[k]\delta(n - kL)$ basada en un tren de deltas. En el dominio

frecuencial, la transformada de Fourier de esta secuencia intermedia consiste en réplicas de la transformada de Fourier de la secuencia inicial centradas en frecuencias múltiples de $2\pi/L$, tal y como se muestra en la figura 6b.

2) Se filtra la secuencia $v[n]$ con un filtro paso bajo que actúa como filtro interpolador y cuya representación frecuencial se muestra en la figura 5c. Este filtrado paso bajo elimina las réplicas para dar lugar a la representación frecuencial de la secuencia interpolada, tal y como se muestra en la figura 6d.

Figura 6. Representación frecuencial de la operación de interpolación. Señal original (ver figura 6a), señal interpolada con $L = 2$ (ver figura 6d)



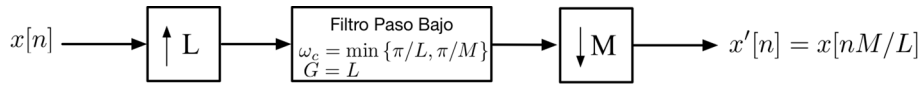
Es interesante hacer notar que como resultado del proceso de interpolación, la secuencia resultante $x_i[n]$ presenta un ancho de banda que, en tiempo discreto, es aparentemente L veces menor que el de la secuencia inicial $x[n]$, tal y como se puede observar comparando la figura 6a y 6d. Por otro lado, puede demostrarse que la operación de interpolación es lineal y estable, pero su causalidad y estabilidad vendrá dada por el filtro interpolador que se utilice.

2.3. Cambio no entero de frecuencia de muestreo

Las operaciones de diezmado e interpolación que se han descrito anteriormente suponen que el factor de cambio de frecuencia de muestreo (M y L) es siempre un número entero. Es posible, sin embargo, llevar a cabo un cambio de frecuencia de muestreo por un factor que no sea entero a partir de combinar en serie ambas operaciones. Ello da lugar a un esquema, como el mostrado en la figura 7, donde la operación de interpolación viene seguida de la operación de diezmado, lo cual da lugar a un cambio de frecuencia de muestreo por un

factor L/M . Cabe destacar que el filtro paso bajo que se considera proviene de combinar los dos filtros que se utilizan en las dos operaciones y, por tanto, puede implementarse con un único bloque.

Figura 7. Representación en forma de diagrama de bloques de la operación de cambio de frecuencia de muestreo por un factor no entero



3. Introducción al filtrado digital

En este apartado se entienden como «filtros» los sistemas lineales e invariantes que procesan una señal de entrada con el objetivo de seleccionar una cierta banda de frecuencias a costa de atenuar o eliminar el resto de frecuencias. Estos sistemas se caracterizan de forma única por su respuesta impulsional $h(t)$ o, equivalentemente, su respuesta en frecuencia $H(\Omega)$. La implementación de estos filtros en el dominio del tiempo discreto, que dan lugar a los llamados «filtros digitales» con respuesta impulsional $h[n]$ y respuesta frecuencial $H(e^{j\omega})$, resulta en dos familias bien diferenciadas de filtros que se conocen como: filtros de respuesta impulsional finita (FIR) y filtros de respuesta impulsional infinita (IIR). En el caso de los filtros FIR, las muestras de la respuesta impulsional $h[n]$ se conocen como los coeficientes del filtro. Sin embargo, en el caso de los filtros IIR, lo que se conoce como coeficientes del filtro suelen ser las constantes que determinan las ecuaciones recursivas, en forma de diferencias finitas, que caracterizan el comportamiento del filtro.

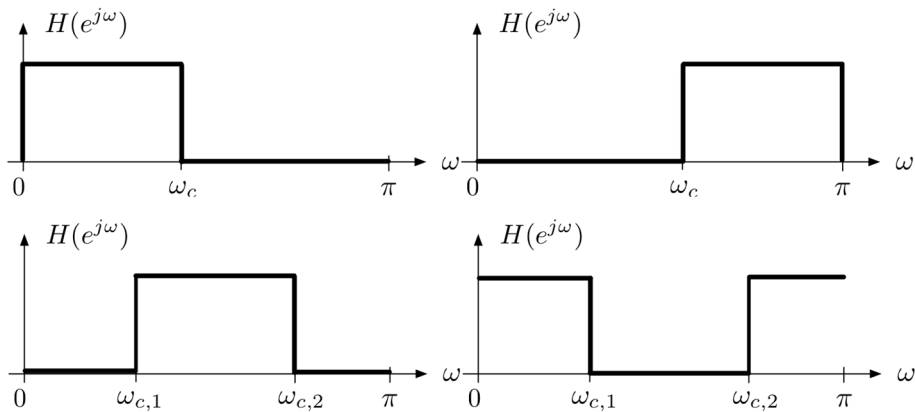
El diseño de filtros digitales es una disciplina bien conocida y en la cual los principales avances fueron llevados a cabo en las últimas dos décadas del siglo XX, a partir de la aparición de los primeros procesadores digitales de señal (DSP) hacia 1980 (Parks, 1987). El procedimiento para el diseño de filtros suele ser sistemático y se basa en:

- 1) Determinar las especificaciones que se necesitan cumplir.
- 2) Encontrar un filtro cuya respuesta en frecuencia se aproxime lo máximo posible a las especificaciones.

3.1. Especificaciones de un filtro digital

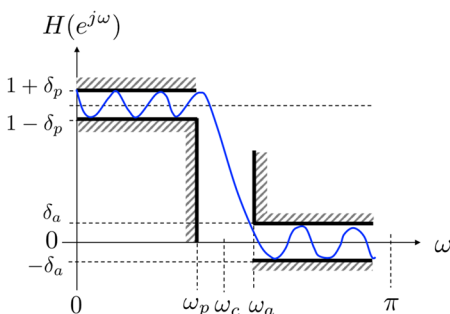
El primer paso en el diseño de un filtro digital consiste en determinar las especificaciones que este debe cumplir. Ello implica habitualmente determinar la respuesta en frecuencia que debería tener el filtro, bien en términos de magnitud y fase, bien en forma de parámetros como frecuencias de corte, ancho de banda, nivel en banda de paso, nivel de atenuación fuera de banda, etc. Existen cuatro tipos de respuestas frecuenciales según se trate de un filtro que deja pasar solo las bajas frecuencias (filtro paso bajo) o el caso inverso, en el que solo deje pasar las altas frecuencias (filtro paso alto). De la misma forma, se puede considerar un filtro que deje pasar solo una banda concreta de frecuencias (filtro paso banda) o el caso inverso, en el que atenúe dicha banda (filtro banda eliminada). Las respuestas frecuenciales ideales de estos cuatro tipos de filtros se muestran a continuación en la figura 8.

Figura 8. Respuesta frecuencial ideal de los cuatro tipos de filtros digitales: paso bajo (a), paso alto (b), paso banda (c) y banda eliminada (d)



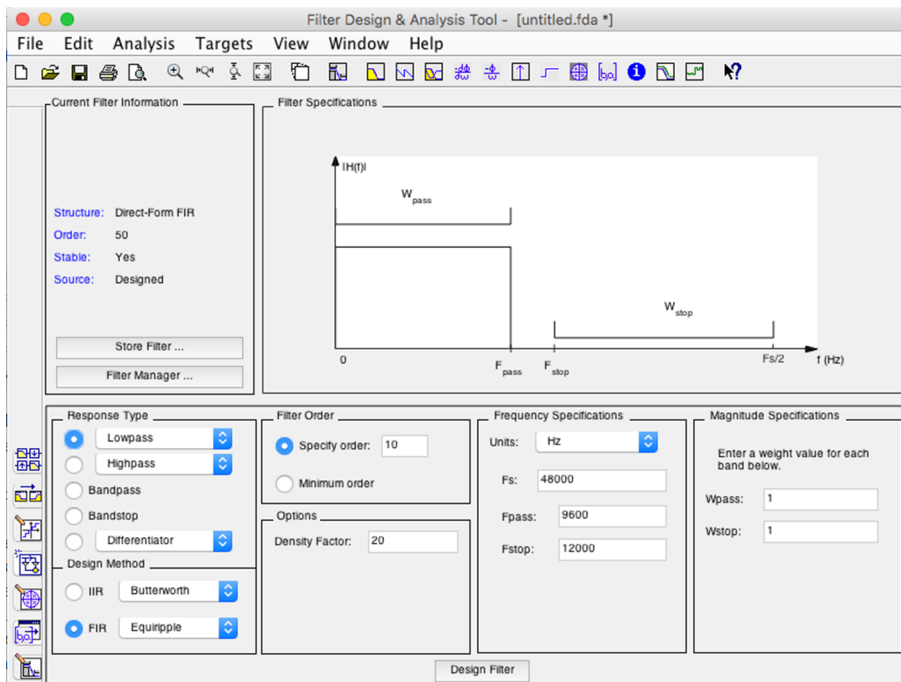
Las respuestas frecuenciales de la figura 8 se caracterizan todas por tener una banda de paso a través de la cual se dejan pasar ciertas componentes frecuenciales de la señal que se filtra y una banda atenuada en donde se atenúan o eliminan ciertas frecuencias. El punto que delimita la banda de paso de la banda atenuada es la frecuencia de corte, indicada según ω_c en la figura 8. En el caso de los filtros paso banda y banda atenuada se definen dos frecuencias de corte, una inferior y otra superior indicadas según $\omega_{c,1}$ y $\omega_{c,2}$. Cabe notar que las respuestas frecuenciales de la figura 8 son ideales en el sentido de que la transición de la banda de paso a la banda atenuada (o al revés) es instantánea. En la práctica estos cambios abruptos en frecuencia ocasionan respuestas muy lentas en el dominio temporal, ya que hacen que la longitud del filtro digital sea muy larga (teóricamente infinita) y, por tanto, dificulten o hagan imposible su implementación. Es por ello que para facilitar la implementación del filtro, se define la transición de la banda de paso a la banda atenuada (o al revés) a través de una región de transición que va desde la denominada frecuencia de paso ω_p hasta la frecuencia atenuada ω_a , siendo la frecuencia de corte el punto intermedio entre ambas, $\omega_c = (\omega_p + \omega_a)/2$. De la misma forma, en lugar de especificar el nivel que presenta la banda de paso y la banda atenuada, se suele especificar una cierta tolerancia de $\pm\delta_p$ para la banda de paso, respecto a un nivel nominal de 1, y $\pm\delta_a$ para la banda atenuada, tal y como se muestra en la figura 9. Esta tolerancia cubre el posible rizado que pueda presentar la respuesta frecuencial que se obtenga en la práctica para cumplir con las especificaciones fijadas.

Figura 9. Plantilla de especificaciones de un filtro digital, para el caso de un filtro paso bajo



Una vez fijadas las especificaciones que hay que cumplir, el diseño del filtro suele implicar una serie de operaciones matemáticas para aproximar el diseño especificado a unas ciertas plantillas, optimizar el número de coeficientes y las prestaciones (en el caso de filtros FIR), o bien el mapeo de las especificaciones a un dominio transformado sobre el cual se determinan los parámetros del filtro (en el caso de filtros IIR). Todas estas operaciones se llevan a cabo hoy en día mediante paquetes de *software* específico, como por ejemplo la *Filter design and analysis tool* de MATLAB, de la cual se muestra una captura en la figura 10.

Figura 10. Captura de pantalla de la *Filter design and analysis tool* de MATLAB



A pesar de la existencia de *software* especializado, es importante conocer los fundamentos en los que se sustentan para comprender las líneas que hay que seguir en el proceso de diseño de un filtro digital. A continuación, se describen de forma resumida los métodos más habituales para el diseño de filtros FIR e IIR que cumplen con unas especificaciones predeterminadas. Puede encontrarse información detallada en diversas referencias especializadas, como por ejemplo Oppenheim (1999) y Karam (2010).

3.2. Diseño de filtros FIR

Los filtros FIR se caracterizan por tener una respuesta frecuencial de tipo polinómica según:

$$H(e^{j\omega}) = \sum_{n=0}^{L-1} h[n]e^{-j\omega n} \quad (7)$$

y presentan la propiedad de ser siempre filtros estables y sin distorsión de fase en el caso de que se cumplan ciertas condiciones de simetría. Los métodos existentes para el diseño de filtros FIR se centran en conseguir filtros de fase lineal y respuesta impulsional $h[n]$ con coeficientes reales. En relación con la simetría de estos coeficientes y su longitud, se distingue entre cuatro tipos de filtros FIR diferentes según se indica en la tabla 1. En dicha tabla se indica además para cada tipo de filtro el número de ceros K , la longitud del filtro $L = K + 1$ en número de coeficientes, el retardo $\alpha = K/2$ que introduce en muestras y el *offset* β que presenta la fase lineal de la respuesta del filtro en frecuencia.

Tabla 1. Resumen de los cuatro tipos de filtros FIR

		Simetría																				
		Par $h[n] = h[2\alpha - n]$	Impar $h[n] = -h[2\alpha - n]$																			
Longitud	Impar	Tipo I																				
		<table border="1"> <tr> <td>Ceros forzados</td> <td>-</td> </tr> <tr> <td>K</td> <td>Par</td> </tr> <tr> <td>$L = K+1$</td> <td>Impar</td> </tr> <tr> <td>$\alpha = K/2$</td> <td>Entero</td> </tr> <tr> <td>β</td> <td>0</td> </tr> </table>	Ceros forzados	-	K	Par	$L = K+1$	Impar	$\alpha = K/2$	Entero	β	0	<table border="1"> <tr> <td>Ceros forzados</td> <td>$z = \{1, -1\}$ $\omega = \{0, \pi\}$</td> </tr> <tr> <td>K</td> <td>Par</td> </tr> <tr> <td>$L = K+1$</td> <td>Impar</td> </tr> <tr> <td>$\alpha = K/2$</td> <td>Entero</td> </tr> <tr> <td>β</td> <td>$\pi/2$</td> </tr> </table>	Ceros forzados	$z = \{1, -1\}$ $\omega = \{0, \pi\}$	K	Par	$L = K+1$	Impar	$\alpha = K/2$	Entero	β
	Ceros forzados	-																				
	K	Par																				
$L = K+1$	Impar																					
$\alpha = K/2$	Entero																					
β	0																					
Ceros forzados	$z = \{1, -1\}$ $\omega = \{0, \pi\}$																					
K	Par																					
$L = K+1$	Impar																					
$\alpha = K/2$	Entero																					
β	$\pi/2$																					
Par	Tipo II																					
	<table border="1"> <tr> <td>Ceros forzados</td> <td>$z = -1$ $\omega = \pi$</td> </tr> <tr> <td>K</td> <td>Impar</td> </tr> <tr> <td>$L = K+1$</td> <td>Par</td> </tr> <tr> <td>$\alpha = K/2$</td> <td>Entero + $\frac{1}{2}$</td> </tr> <tr> <td>β</td> <td>0</td> </tr> </table>	Ceros forzados	$z = -1$ $\omega = \pi$	K	Impar	$L = K+1$	Par	$\alpha = K/2$	Entero + $\frac{1}{2}$	β	0	<table border="1"> <tr> <td>Ceros forzados</td> <td>$z = 1$ $\omega = 0$</td> </tr> <tr> <td>K</td> <td>Impar</td> </tr> <tr> <td>$L = K+1$</td> <td>Par</td> </tr> <tr> <td>$\alpha = K/2$</td> <td>Entero + $\frac{1}{2}$</td> </tr> <tr> <td>β</td> <td>$\pi/2$</td> </tr> </table>	Ceros forzados	$z = 1$ $\omega = 0$	K	Impar	$L = K+1$	Par	$\alpha = K/2$	Entero + $\frac{1}{2}$	β	$\pi/2$
Ceros forzados	$z = -1$ $\omega = \pi$																					
K	Impar																					
$L = K+1$	Par																					
$\alpha = K/2$	Entero + $\frac{1}{2}$																					
β	0																					
Ceros forzados	$z = 1$ $\omega = 0$																					
K	Impar																					
$L = K+1$	Par																					
$\alpha = K/2$	Entero + $\frac{1}{2}$																					
β	$\pi/2$																					

En cuanto al diseño de filtros FIR, a continuación se describen brevemente los métodos más representativos.

3.2.1. Método por muestreo en frecuencia

Suponiendo que tenemos especificada una cierta respuesta en frecuencia $H(e^{j\omega})$, una primera opción para obtener un filtro digital que implemente de forma aproximada dicha respuesta es muestrear $H(e^{j\omega})$ para obtener lo que vendría a ser la DFT de la respuesta impulsional del filtro. Esto es,

$$H[k] = H(e^{j\omega})_{\omega=2\pi k/L} \quad \text{con } k = 0, 1, \dots, L - 1. \quad (8)$$

A continuación, pueden obtenerse los coeficientes del filtro aplicando la IDFT a la respuesta frecuencial que acaba de muestrearse, dando lugar a:

$$h_m[n] = \text{IDFT}_L\{H[k]\}. \quad (9)$$

Los coeficientes en (9) son una manera fácil y sencilla de obtener un filtro FIR que se corresponde exactamente con la respuesta frecuencial que se ha especificado, aunque únicamente en aquellos puntos del eje de frecuencia que han sido muestreados. La desventaja de este método es que, para el resto de frecuencias intermedias, la respuesta en frecuencia del filtro FIR queda fuera de nuestro control y puede que no se corresponda exactamente con la respuesta especificada en $H(e^{j\omega})$.

3.2.2. Método por enventanado

Otra opción relacionada de alguna forma con la anterior sería, a partir de la respuesta en frecuencia especificada, obtener su correspondencia en tiempo discreto a través de la TFSD inversa,

$$h[n] = \frac{1}{2\pi} \int_{-\pi}^{\pi} H(e^{j\omega}) e^{j\omega n} d\omega. \quad (10)$$

La respuesta impulsional que se obtiene presenta en general una duración infinita, con lo cual, es necesario limitarla a un conjunto finito de L muestras para obtener un filtro FIR. Este truncado se suele llevar a cabo multiplicando $h[n]$ por $w[n]$, una ventana de duración L muestras. El resultado es la siguiente respuesta impulsional truncada:

$$h_w[n] = h[n]w[n] \quad (11)$$

que obviamente incurrirá en un cierto error en comparación con la respuesta impulsional ideal. Mediante el teorema de Parseval, dicho error puede calcularse equivalentemente tanto en el dominio frecuencial como en el temporal mediante:

$$\epsilon^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} |H(e^{j\omega}) - H_w(e^{j\omega})|^2 d\omega = \sum_{n=-\infty}^{\infty} |h[n] - h_w[n]|^2 \quad (12)$$

expresión que puede utilizarse como función de coste para encontrar la ventana más adecuada. Esto es, aquella ventana que minimice (12). Algunos ejemplos de las ventanas más habituales se muestran a continuación en la tabla 2.

Por otro lado, es interesante mencionar que, debido al enventanado en (11), el filtro FIR que se obtiene presenta una respuesta frecuencial que viene dada por la convolución (periódica 2π) entre la respuesta frecuencial ideal $H(e^{j\omega})$ y la transformada de Fourier de la ventana utilizada, $W(e^{j\omega})$; esto es:

$$H_w(e^{j\omega}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} H(e^{j\lambda}) W(e^{j(\omega-\lambda)}) d\lambda \quad (13)$$

Este aspecto ha de ser tenido en cuenta a la hora de evaluar la conveniencia de utilizar una ventana u otra, en función de la respuesta en frecuencia que se obtenga. Para ello es necesario comparar la respuesta frecuencial obtenida con la especificada en origen y evaluar el error que se está incurriendo a través de (12).

Tabla 2. Ejemplos (con $L = 16$) de ventanas habituales en el diseño de filtros FIR por enventanado

Ventana	Respuesta temporal / frecuencial
<p>Blackman $w[n] = 0.42 - 0.5\cos\left(\frac{2\pi}{L-1}n\right) + 0.08\cos\left(\frac{4\pi}{L-1}n\right)$, $n = 0, 1, \dots, M-1$ siendo $M = L/2$ para L par y $M = (L+1)/2$ para L impar; los coeficientes correspondientes a $M \leq n \leq L-1$ se obtienen girando los coeficientes para $n = 0, 1, \dots, M-1$ alrededor del punto medio.</p>	<p>The left plot shows the window coefficients $w[n]$ for $n = 0, \dots, 15$. The coefficients are symmetric around $n=7.5$ and have a main peak of 1.0 at $n=7.5$. The right plot shows the magnitude response $H(e^{j\omega})$ in dB versus normalized frequency $\omega/2\pi$. The magnitude starts at 0 dB at $\omega=0$ and has a very deep null at $\omega/2\pi = 0.25$.</p>
<p>Hamming $w[n] = 0.54 - 0.46\cos\left(\frac{2\pi}{L-1}n\right)$, $n = 0, 1, \dots, L-1$</p>	<p>The left plot shows the window coefficients $w[n]$ for $n = 0, \dots, 15$. The coefficients are symmetric around $n=7.5$ and have a main peak of 1.0 at $n=7.5$. The right plot shows the magnitude response $H(e^{j\omega})$ in dB versus normalized frequency $\omega/2\pi$. The magnitude starts at 0 dB at $\omega=0$ and has a deep null at $\omega/2\pi = 0.25$.</p>
<p>Hanning $w[n] = 0.5 \left[1 - \cos\left(\frac{2\pi}{L-1}n\right) \right]$, $n = 0, 1, \dots, L-1$</p>	<p>The left plot shows the window coefficients $w[n]$ for $n = 0, \dots, 15$. The coefficients are symmetric around $n=7.5$ and have a main peak of 1.0 at $n=7.5$. The right plot shows the magnitude response $H(e^{j\omega})$ in dB versus normalized frequency $\omega/2\pi$. The magnitude starts at 0 dB at $\omega=0$ and has a deep null at $\omega/2\pi = 0.25$.</p>

3.2.3. Método óptimo de rizado constante

Existen métodos que utilizan diferentes criterios de optimalidad para el diseño de filtros FIR, como puede ser por ejemplo la minimización del error cuadrático o del error máximo entre la respuesta frecuencial del filtro FIR y la respuesta especificada (Karam, 2010). De forma similar, también es posible diseñar el filtro FIR de forma que no minimice una función de coste concreta, sino que minimice el nivel de rizado que presenta su respuesta frecuencial, esto es, de forma que sea lo más plana posible. Finalmente se añaden criterios prácticos como minimizar el número de coeficientes del filtro que son necesarios para una anchura de la banda de transición también mínima.

Estos son criterios cualitativos que, a menudo, resultan más útiles a nivel práctico que no fijar un cierto valor para el error cuadrático. Para ello, el problema se suele formular a partir de la teoría de la aproximación utilizando una función de coste basada en el error ponderado de Chebyshev. La solución a este problema no es sencilla ni presenta una expresión cerrada, pero puede obtenerse con relativa facilidad utilizando el algoritmo iterativo ideado por Parks y McClellan, cuyos detalles pueden consultarse por ejemplo en Karam (2010). El algoritmo tan solo necesita que se le proporcionen cinco parámetros, a partir de los cuales encontrará la solución que mejor se ajuste a ellos. Estos son: la longitud deseada del filtro FIR, L , las frecuencias que delimitan la banda de paso y la banda atenuada, ω_p y ω_a , así como el error máximo en la banda de paso y en la banda atenuada, δ_p y δ_a , en definitiva, los parámetros que se muestran en la figura 9.

Estos cinco parámetros están relacionados entre ellos de tal forma que cualquiera de ellos depende de los otros cuatro. En la práctica es habitual especificar los requerimientos de ω_p , ω_a , δ_p y δ_a y dar libertad al algoritmo para que encuentre el filtro FIR que mejor cumple con estos requisitos, siendo la salida del algoritmo los coeficientes del filtro FIR y la longitud necesaria. De forma aproximada, la longitud necesaria en el criterio óptimo de rizado constante viene dada por:

$$L \approx \frac{-10 \log_{10}(\delta_p \delta_a) - 13}{14.6 \Delta f} + 1 \quad (14)$$

donde $\Delta f = (\omega_a - \omega_p) / (2\pi)$ es la anchura de la banda de transición. En general, puede concluirse que los filtros FIR diseñados bajo el criterio de rizado constante ofrecen una atenuación (definida como $-10 \log_{10}(\delta_p \delta_a)$), que es del orden de 5 dB superior a la que ofrecen los métodos de diseño de filtros FIR basados en el enventanado para un mismo número de coeficientes.

3.3. Diseño de filtros IIR

Los filtros IIR presentan una respuesta impulsional infinita que proviene de un sistema cuya caracterización se lleva a cabo mediante ecuaciones en diferencias finitas, las cuales relacionan entradas y salidas actuales con instantes pasados. Ello hace que la respuesta frecuencial de los filtros IIR sea de tipo racional basada en el cociente de dos polinomios según:

$$H(e^{j\omega}) = \frac{\sum_{k=0}^P b_k e^{-j\omega k}}{\sum_{k=0}^Q a_k e^{-j\omega k}} \quad (15)$$

donde el orden del filtro IIR viene dado por el orden del denominador, Q , ya que en la habitualmente $P < Q$. El orden del filtro coincide con el número de muestras de salida en instantes anteriores que ha de almacenarse para calcular una nueva muestra a la salida del filtro, siendo los coeficientes $\{a_k\}_{k=0}^Q$ y $\{b_k\}_{k=0}^P$ las incógnitas que hay que determinar en el diseño del filtro. Como puede verse, no existe una dependencia lineal entre los coeficientes del filtro y la respuesta frecuencial, lo cual hace más difícil el diseño de filtros IIR en comparación con los FIR.

A pesar de las dificultades, existen métodos analíticos que permiten obtener directamente la respuesta de un filtro IIR deseado a partir de filtros tipo de respuesta analógica. Uno de estos métodos es la llamada transformación bilineal que se describe a continuación.

3.3.1. Método de la transformación bilineal

La transformación bilineal es un método que permite convertir el problema de diseñar un filtro IIR en tiempo discreto, cuya función de transferencia es $H(z)$ en el dominio transformado z , en el problema de diseñar un filtro en tiempo continuo (i. e. analógico), cuya función de transferencia es $H_a(s)$ en el dominio de Laplace. De esta forma, se aprovecha que para el caso del diseño de filtros analógicos ya existen soluciones bien conocidas. Una vez diseñado el filtro analógico, se realiza la transformación para volver al dominio en tiempo discreto, en donde se asegura que las especificaciones se siguen cumpliendo.

La relación entre ambas funciones de transferencia viene dada por:

$$s = K \left(\frac{1-z^{-1}}{1+z^{-1}} \right) \quad (16)$$

en donde K es un parámetro de diseño. La transformación en (16) es además invertible, puesto que:

$$z^{-1} = \frac{1 - s/K}{1 + s/K} \quad (17)$$

y garantiza la estabilidad del filtro al pasar de uno a otro dominio, puesto que el semiplano izquierdo en el plano-S de la transformada de Laplace queda dentro del círculo unitario al convertir a la representación de la transformada z .

Por otro lado, el eje imaginario en el plano-S se mapea directamente sobre el círculo unidad del plano-Z, con lo cual, existe una relación directa entre las frecuencias de uno y otro dominio. De hecho, sustituyendo $s = e^{j\Omega}$ y $z = e^{j\omega}$ en (16), obtenemos:

$$\Omega = K \tan\left(\frac{\omega}{2}\right). \quad (18)$$

Con el resultado en (18) puede verse que la constante K permite mapear una frecuencia concreta en el dominio analógico a una frecuencia concreta del dominio digital, lo cual es útil para controlar la posición de la frecuencia de corte del filtro.

Finalmente, una vez convertido el problema al dominio analógico, el siguiente paso consiste en seleccionar uno de los filtros analógicos conocidos, cuya respuesta frecuencial puede obtenerse de forma analítica. Los cuatro tipos de filtros analógicos más conocidos son: Butterworth, Chebyshev, inverso de Chebyshev y elíptico. Sus respuestas en frecuencia se muestran a continuación en la tabla 3, en donde la atenuación $\alpha(\Omega)$ se define como:

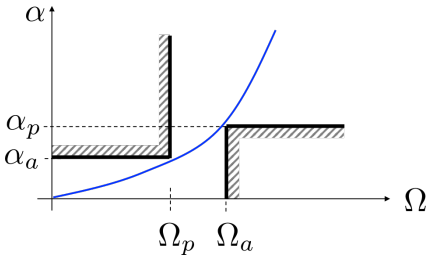
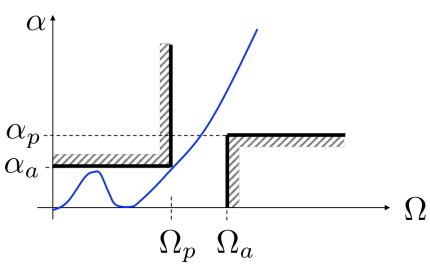
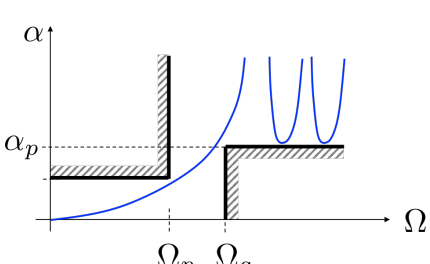
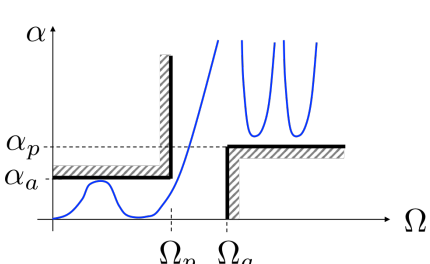
$$\alpha(\Omega) = -20 \log_{10} H_a(\Omega) \quad (19)$$

y, por lo tanto,

$$\alpha_a = -20 \log_{10} \delta_a \quad (20)$$

$$\alpha_p = -20 \log_{10} \frac{1 + \delta_p}{1 - \delta_p}. \quad (21)$$

Tabla 3. Tipos básicos de filtros analógicos y su respuesta en términos de atenuación en frecuencia

Tipo de filtro	Atenuación del filtro analógico en frecuencia
<p>Butterworth, orden N:</p> $ H_a(\Omega) ^2 = \frac{1}{\left 1 + \left(\frac{\Omega}{\Omega_c}\right)^{2N}\right ^2}$ <p>Presenta un comportamiento maximalmente plano en la banda de paso a cambio de una modesta selectividad en frecuencia.</p>	
<p>Chebyshev, de polinomio $T_N(x)$:</p> $ H_a(\Omega) ^2 = \frac{1}{1 + \epsilon^2 T_N^2\left(\frac{\Omega}{\Omega_c}\right)}$ <p>Presenta rizado constante en la banda de paso y decrece monótonamente en la banda atenuada. El rizado se controla a través de ϵ, resolviendo:</p> $(1 - \delta_p)^2 = \frac{1}{1 + \epsilon^2}$	
<p>Inverso de Chebyshev, de polinomio $T_N(x)$:</p> $ H_a(\Omega) ^2 = \frac{1}{1 + \left[\epsilon^2 T_N^2\left(\frac{\Omega}{\Omega_c}\right)\right]^{-1}}$ <p>Decrece monótonamente en la banda de paso y presenta rizado constante en la banda atenuada y. El rizado se controla a través de ϵ, resolviendo:</p> $(1 - \delta_p)^2 = \frac{\epsilon^2}{1 + \epsilon^2}$	
<p>Elíptico, de orden N:</p> $ H_a(\Omega) ^2 = \frac{1}{1 + \epsilon^2 E_N^2(\Omega)}$ <p>Presenta rizado constante tanto en la banda atenuada como en la banda de paso y la banda de transición más estrecha de entre todos los filtros con igual banda de paso y banda atenuada. El denominador depende de la función elíptica Jacobiana $E_N(\Omega)$.</p>	

A partir de los cuatro tipos de filtros paso bajo analógicos descritos anteriormente, el procedimiento para diseñar un filtro paso bajo IIR utilizando el método de la transformación bilineal puede resumirse en los siguientes pasos:

1) Indicar las especificaciones que ha de cumplir la respuesta frecuencial $H(e^{j\omega})$ del filtro IIR en los términos de la figura 9 y convertirlas al dominio analógico utilizando la relación (18).

2) Seleccionar el filtro analógico que mejor se ajuste a las especificaciones y asegurar que proporcione una respuesta causal y estable. Para ello, comprobar que los polos de $|H_a(j\Omega)|^2$ yacen todos en el semiplano izquierdo del plano-S.

3) Obtener la función de transferencia $H(z)$ del filtro digital aplicando la transformación bilineal a $H_a(\Omega)$ utilizando (16).

4) Obtener la respuesta en frecuencia $H(e^{j\omega})$ del filtro digital sustituyendo $z = e^{j\omega}$.

3.4. Comparativa entre filtros FIR e IIR

La decisión entre el uso de un filtro FIR o IIR depende de un compromiso entre estabilidad, facilidad de implementación y selectividad en términos de relación entre el nivel de banda de paso y el nivel de banda atenuada.

Por un lado, los filtros FIR permiten diseñar filtros causales y de fase lineal, los cuales son muy importantes en diversas aplicaciones de procesamiento de señal, donde interesa filtrar una banda de frecuencias de la señal de entrada introduciendo la mínima distorsión de fase. Los filtros FIR que cumplen simetría par o impar en respuesta impulsional son los únicos que pueden presentar fase lineal. Además, los filtros FIR resultan en una respuesta impulsional finita que permite implementar la operación de filtrado mediante la convolución y, por tanto, beneficiarse de las implementaciones eficientes de filtrado mediante el uso de la FFT y el filtrado en bloques. La sencillez de implementación, sin retroalimentaciones, evita posibles problemas de estabilidad que pueden surgir al cuantificar los coeficientes del filtro a la hora de implementarlo en *hardware*. Como contrapartida, los filtros FIR suelen necesitar muchos coeficientes para conseguir una respuesta frecuencial que sea suficientemente selectiva para ciertas aplicaciones.

En cuanto a los filtros IIR, su diseño es *a priori* más complejo porque implica una serie de transformaciones para mapear los requisitos del dominio digital al dominio analógico y posteriormente de vuelta al dominio digital. Además, al tratarse de una respuesta frecuencial racional, las raíces del denominador han de escogerse con cuidado para evitar que el filtro resultante sea inestable. No solo eso, sino que los coeficientes resultantes han de ser analizados también con cuidado para evitar que al redondearlos o cuantificarlos en implementaciones *hardware*, los polos resultantes queden desviados respecto de su posición inicial y el filtro se vuelva inestable. Esto hace que el análisis de estabilidad sea un aspecto muy importante en el diseño de filtros IIR. Otra característica de los filtros IIR es que su respuesta impulsional infinita hace imposible el uso del operador convolución. Ello hace que la salida del filtro tenga que calcularse mediante ecuaciones en diferencias finitas y, por tanto, realimentando la salida del filtro hacia la entrada. La contrapartida de los fil-

Los filtros IIR necesitan muchos menos coeficientes que los filtros FIR para obtener respuestas selectivas en frecuencia, lo cual los hace muy atractivos en implementaciones de baja complejidad.

4. Implementación del filtrado en bloques

Una de las propiedades de la transformada de Fourier es que permite implementar de forma muy sencilla la convolución entre dos secuencias, lo que hace que dicha operación se convierta en un mero producto en el dominio transformado. Ello hace que la DFT sea muy utilizada en el análisis de sistemas discretos y sobre todo en la implementación de la operación de filtrado con filtros FIR, donde la existencia de algoritmos que implementan la DFT de forma eficiente hace que el proceso de convolución en el dominio frecuencial sea aún más ventajoso (Blahut, 2010).

El filtrado de secuencias mediante la DFT implica utilizar bloques de muestras de tamaño limitado, puesto que la complejidad de cálculo de la DFT depende del número de muestras que se procesan por bloque. Esto hace que el filtrado de secuencias muy largas tenga que llevarse a cabo segmentando la secuencia de entrada en bloques de tamaño fijo para que puedan ser procesados secuencialmente por la DFT. Además, si lo que se pretende es implementar la convolución lineal entre una secuencia de entrada y la respuesta impulsional de un cierto filtro FIR, interesa evitar el efecto de convolución circular que introduce por defecto la DFT. Para ello existen diversos métodos, como el *overlap-save* y el *overlap-add*, que se describen a continuación, y que precisamente tienen como objetivo implementar de forma eficiente la convolución lineal entre bloques de la secuencia de entrada y la respuesta impulsional de un filtro FIR de L coeficientes.

4.1. Método *overlap-save*

Este método consiste en tomar bloques de longitud $N = K + L - 1$ muestras de la secuencia de entrada, con típicamente $K \gg L$, y donde N pasa a ser el tamaño de la DFT e IDFT que llevan a cabo la convolución en el dominio de la frecuencia. Interesa, por tanto, que N sea potencia de 2 para poder aplicar el algoritmo FFT en el cálculo de la DFT. Dada la respuesta impulsional del filtro FIR de L coeficientes, $h(n)$, se añaden $K - 1$ ceros a partir del último coeficiente, de forma que queda una respuesta impulsional con *zero padding* hasta un total de N coeficientes. De esta respuesta impulsional, se calcula la DFT de N puntos,

$$H_N[k] = DFT_N \left\{ h(0), h(1), \dots, h(L-1), \underbrace{0, \dots, 0}_{K-1} \right\} \quad (22)$$

A partir de este punto se va procesando la secuencia de entrada en bloques de N muestras, para los cuales se calcula su DFT:

$$X_{N,m}[k] = \text{DFT}_N\{x_m(0), x_m(1), \dots, x_m(N-1)\} \quad (23)$$

con:

$$x_m(n) = x(mN + n). \quad (24)$$

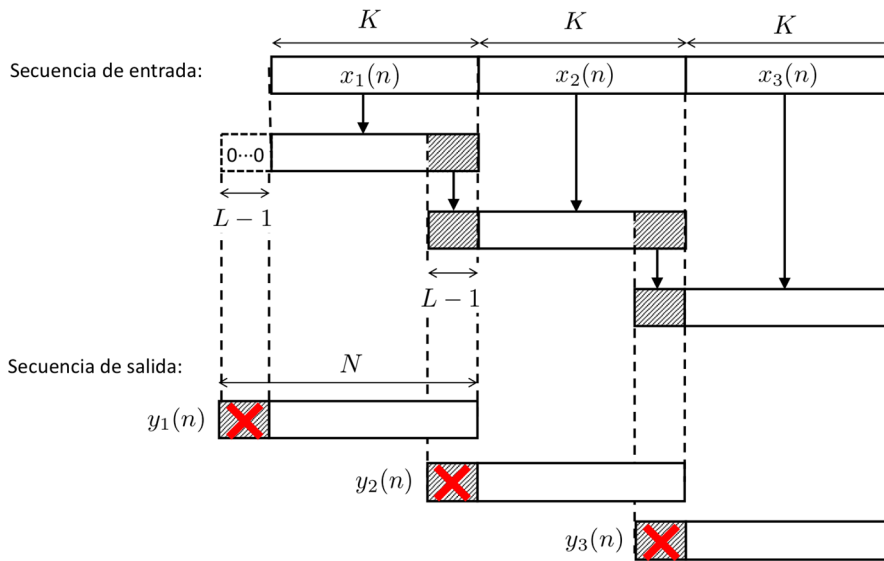
A continuación, se multiplican ambas DFT y se lleva a cabo la IDFT del resultado, dando lugar a la secuencia de muestras:

$$y_m(n) = \text{IDFT}_N\{X_{N,m}[k]H_N[k]\}. \quad (25)$$

Puesto que estamos convolucionando un segmento de N muestras de entrada con un filtro de longitud L , si se tratara de una convolución lineal tendríamos una secuencia de salida de longitud $N + L - 1$. Sin embargo, la longitud de la secuencia de salida $y_m(n)$ viene limitada a N , el tamaño de la IDFT, con lo cual, las primeras $L - 1$ muestras se habrán visto afectadas por el *aliasing* ocasionado por la convolución circular. Estas $L - 1$ primeras muestras de $y_m(n)$ hay que descartarlas y quedarse únicamente con las K muestras a continuación, puesto que son las que coinciden exactamente con el resultado de la convolución lineal. Por lo tanto, por cada bloque de N muestras de la secuencia de entrada obtenemos tan solo K muestras de la secuencia filtrada. Es por ello que para mantener la continuidad de la secuencia de salida y evitar que se pierdan muestras, los bloques de muestras de la secuencia de entrada han de tomarse con un solape de $L - 1$ muestras entre ellos. Es decir, las $L - 1$ últimas muestras del bloque $x_m(n)$ pasan a ser las primeras $L - 1$ del bloque siguiente $x_{m+1}(n)$.

El método *overlap-save* se describe gráficamente en la figura 11, donde se indican claramente las muestras que se solapan entre bloques consecutivos de muestras de entrada y cómo el procedimiento comienza inicializando la secuencia de entrada con $L - 1$ ceros para el primer bloque.

Figura 11. Ilustración gráfica de la operación de convolución por bloques utilizando el método *overlap-save*



4.2. Método *overlap-add*

A diferencia del método anterior, el método *overlap-add* evita el efecto de *aliasing*, que aparece al filtrar bloques consecutivos de la secuencia de entrada, ya que procesa bloques de tamaño K , mientras que el tamaño de la DFT e IDFT ya tiene en cuenta la longitud del filtro de L coeficientes, y es precisamente de $N = K + L - 1$ puntos; esto es, la longitud que tendría la secuencia de salida al convolucionar linealmente un bloque de K muestras con un filtro de L coeficientes. Para ello es necesario añadir $L - 1$ ceros a cada bloque de K muestras que se procesa, y del que se calcula su DFT de N puntos según:

$$X'_{N,m}[k] = DFT_N \left\{ x_m(0), x_m(1), \dots, x_m(K-1), \underbrace{0, \dots, 0}_{L-1} \right\} \quad (26)$$

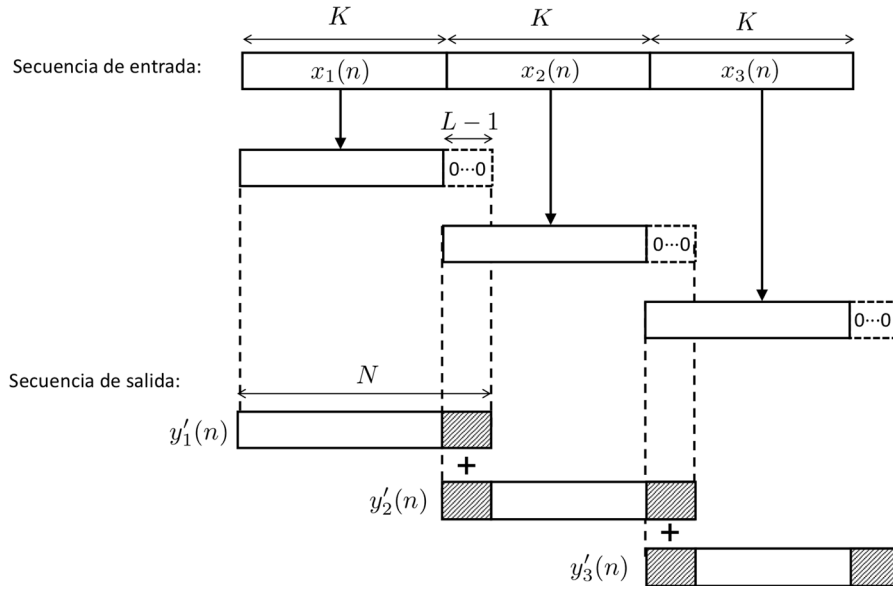
A continuación se multiplica la DFT del bloque de entrada con la DFT del filtro y se lleva a cabo la IDFT, dando lugar a la secuencia de muestras:

$$y'_m(n) = IDFT_N \{ X'_{N,m}[k] H_N[k] \}. \quad (27)$$

Puesto que se ha evitado el efecto de *aliasing*, todas las N muestras de este bloque de salida son útiles y no hay que descartar ninguna. Estas muestras de salida se corresponden con las muestras $x_m(n)$ de un cierto bloque de entrada de K muestras, con lo cual, para procesar el siguiente bloque $x_{m+1}(n)$ hay que desplazarse K muestras hacia delante. Sobre estas nuevas K muestras el procedimiento es similar al descrito anteriormente, es decir, añadir $L - 1$ ceros al nuevo bloque de muestras según (26), calcular la DFT del bloque, multiplicar por la DFT de los coeficientes del filtro y hacer la IDFT para volver al dominio temporal. Como resultado se obtiene el bloque de muestras de salida $y'_{m+1}(n)$,

cuyas $L - 1$ primeras muestras han de sumarse a las $L - 1$ últimas muestras del bloque $y'_m(n)$ anterior para así garantizar la continuidad de la señal de salida, tal y como se indica gráficamente en la figura 12.

Figura 12. Ilustración gráfica de la operación de convolución por bloques utilizando el método *overlap-add*



Ejercicios de autoevaluación

1. Sea la señal $x[n] = \sin(\omega_0 n)$ con $\omega_0 = 0.8\pi/3$. Si aplicamos un diezmado con filtro *anti-aliasing* a esta señal por un factor $M = 3$, la transformada de Fourier de la señal resultante presenta:

- a) Una delta a la frecuencia $\omega'_0 = 0.8\pi$.
- b) Una delta a la frecuencia $\omega'_0 = 0.8\pi/9$.
- c) Una delta a la frecuencia $\omega'_0 = 0.8\pi$ y otra delta a la frecuencia $\omega'_0 = 1.2\pi$.
- d) Una delta a la frecuencia $\omega'_0 = 0.8\pi/9$ y otra delta a la frecuencia $\omega'_0 = 2\pi - 0.8\pi/9$.

2. Sea la señal $x[n] = \sin(\omega_0 n)$ con $\omega_0 = 1.2\pi/3$. Si aplicamos un diezmado sin filtro *anti-aliasing* a esta señal por un factor $M = 3$, la transformada de Fourier de la señal resultante presenta:

- a) Una delta a la frecuencia $\omega'_0 = 1.2\pi$.
- b) Una delta a la frecuencia $\omega'_0 = 1.2\pi/9$.
- c) Una delta a la frecuencia $\omega'_0 = 0.8\pi$.
- d) Una delta a la frecuencia $\omega'_0 = 0.8\pi$ y otra delta a la frecuencia $\omega'_0 = 1.2\pi$.

3. Sea la señal $x[n] = \sin(\omega_0 n)$ con $\omega_0 = 0.8\pi/3$. Si interpolamos esta señal por un factor $L = 3$, la transformada de Fourier de la señal resultante presenta:

- a) Una delta a la frecuencia $\omega'_0 = 0.8\pi$.
- b) Una delta a la frecuencia $\omega'_0 = 0.8\pi/9$.
- c) Una delta a la frecuencia $\omega'_0 = 0.8\pi$ y otra delta a la frecuencia $\omega'_0 = 1.2\pi$.
- d) Una delta a la frecuencia $\omega'_0 = 0.8\pi/9$ y otra delta a la frecuencia $\omega'_0 = 2\pi - 0.8\pi/9$.

4. Sea la señal $x[n] = \sin(\omega_0 n)$ con $\omega_0 = 0.8\pi/3$. Si aplicamos un cambio de frecuencia por un factor $3/2$, ¿qué afirmaciones son ciertas?

- a) Es necesario implementar un filtro paso bajo con frecuencia de corte $\omega_c = \pi/2$.
- b) Es necesario implementar un filtro paso bajo con frecuencia de corte $\omega_c = \pi/3$.
- c) La transformada de Fourier de la señal resultante presenta una delta a la frecuencia $\omega'_0 = 1.6\pi/9$ y otra delta a la frecuencia $\omega'_0 = 16.4\pi/9$.
- d) La transformada de Fourier de la señal resultante presenta una delta a la frecuencia $\omega'_0 = 0.4\pi$ y otra delta a la frecuencia $\omega'_0 = 1.6\pi$.

5. Se necesita implementar un filtro FIR y se quiere que la señal de salida presente un retardo entero de muestras y un *offset* de fase nulo. ¿Qué tipo de filtro FIR hay que implementar?

- a) Tipo I.
- b) Tipo II.
- c) Tipo III.
- d) Tipo IV.

6. Con relación al diseño de filtros FIR e IIR, indique qué afirmaciones son ciertas:

- a) Los filtros FIR suelen requerir menos coeficientes que los filtros IIR para implementar respuestas selectivas en frecuencia.
- b) Los filtros FIR son siempre causales y de fase lineal.
- c) Los filtros IIR dan lugar a implementaciones de filtrado en bloque más eficientes, gracias al uso de la FFT.
- d) Los filtros FIR obtenidos mediante el método rizado constante tienen como desventaja que su atenuación es del orden de 5 dB peor que los filtros FIR obtenidos por enventanado.

7. Con relación al diseño de filtros IIR, indique qué afirmaciones son ciertas:

- a) Los filtros de Butterworth presentan un comportamiento plano en la banda de paso y una gran selectividad en frecuencia.
- b) Los filtros de Chebyshev decrecen monótonamente en la banda de paso y presentan rizado constante en la banda atenuada.
- c) Los filtros inversos de Chebyshev presentan rizado constante en la banda de paso y decrecen monótonamente en la banda atenuada.
- d) Los filtros elípticos presentan rizado constante tanto en la banda de paso como en la banda atenuada.

8. Indique la expresión correcta de la transformación bilineal para el diseño de filtros IIR:

- a) $s = K \left(\frac{1-z^{-1}}{1+z^{-1}} \right)$.
- b) $s = K \left(\frac{1+z^{-1}}{1-z^{-1}} \right)$.
- c) $z^{-1} = K \left(\frac{1-s^{-1}}{1+s^{-1}} \right)$.
- d) $z^{-1} = \frac{1+s/K}{1-s/K}$.

Solucionario

Ejercicios de autoevaluación

1. c

2. d

3. d

4. b c

5. a

6. b

7. d

8. a

Bibliografía

Blahut, R. E. (2010). *Fast Algorithms for Signal Processing*. Cambridge University Press.

Karam, L. J.; McClellan, J. H.; Selesnick, I. W.; Sidney, C. (2010). *Digital Signal Processing Fundamentals* (Capítulo 11, «Digital filtering»), *The Digital Signal Processing Handbook* (2.^a ed.). CRC Press.

Oppenheim, A. V.; Schafer, R. W.; Buck, J. R. (1999). *Discrete-Time Signal Processing* (2.^a ed.). Prentice Hall, Signal Processing Series.

Parks, T. W.; Burrus, C. S. (1987). *Digital Filter Design*. John Wiley & Sons.