

Índice general

1. Filtros digitales	3
1.1. Introducción	3
1.2. Filtros digitales: ceros y polos	7
1.2.1. Filtros digitales elementales	7
1.2.2. Característica de amplitud y fase	8
1.2.3. Filtros digitales de realimentación negativa	13
1.2.4. Filtros generales con ceros y polos	16
1.2.5. Filtro con un solo cero	16
1.2.6. Filtro con un solo polo	18
1.2.7. Filtro inverso	20
1.2.8. Filtro con un solo cero y un solo polo: filtro pasa-todo	20
1.2.9. Filtros pasa-todo inversos	24
1.2.10. Recursión	25
1.2.11. MA, AR, ARMA, FIR, IIR	27
1.2.12. Síntesis de filtros vía ceros y polos	27
1.2.13. Ceros, polos y fase mínima: revisión de conceptos	32
1.2.14. Teorema de retardo de energía de Robinson	34
1.3. Diseño de filtros digitales	38
1.3.1. Filtro ideal	39
1.3.2. Filtros mediante el uso ventanas	40
1.3.3. Filtros mediante digitalización de funciones racionales de ω	42

Capítulo 1

Filtros digitales

1.1. Introducción

Un filtro digital es un sistema que discrimina cierta parte de la información que ingresa. Usualmente la discriminación se realiza en base a consideraciones de contenido en frecuencias, pero puede ser en base a otros criterios (longitud de onda, moveout, velocidad, etc.). En este capítulo nos concentraremos en aquellos filtros que discriminan de acuerdo al contenido en frecuencias.

Recordando el teorema de la convolución en el Análisis de Fourier, el siguiente concepto debe tenerse bien presente: *convolución en el dominio del tiempo equivale a multiplicación en el dominio de las frecuencias, y viceversa*. O sea, la convolución de dos funciones significa la multiplicación de sus transformadas. Esto nos permitirá aplicar el filtro en uno u otro de los dominios produciendo resultados equivalentes.

Supongamos que tenemos una traza sísmica y queremos atenuar el “ruido” presente en las altas frecuencias. Podemos diseñar un filtro *pasa-bajos* en base a la observación del contenido en frecuencias de la traza (ver Figura 1.1). Claramente, las frecuencias altas resultarán atenuadas si multiplicamos el espectro de la traza por el espectro del filtro, obteniendo como resultado una traza filtrada con menor contenido de ruido. Es importante destacar que en la práctica el ruido siempre resulta **atenuado** y no **eliminado**, ya que no es posible construir un filtro “ideal” por las limitaciones impuestas por el manejo digital de la información, como veremos más adelante.

Se suele caracterizar un filtro por su

- a) **respuesta impulsiva** (dominio del tiempo): son los “coeficientes” o “muestras” del filtro.
- b) **respuesta en frecuencia** (dominio de las frecuencias): representa la transformada de Fourier de la respuesta impulsiva. En este caso se tienen en cuenta los espectros de amplitud y de fase, y

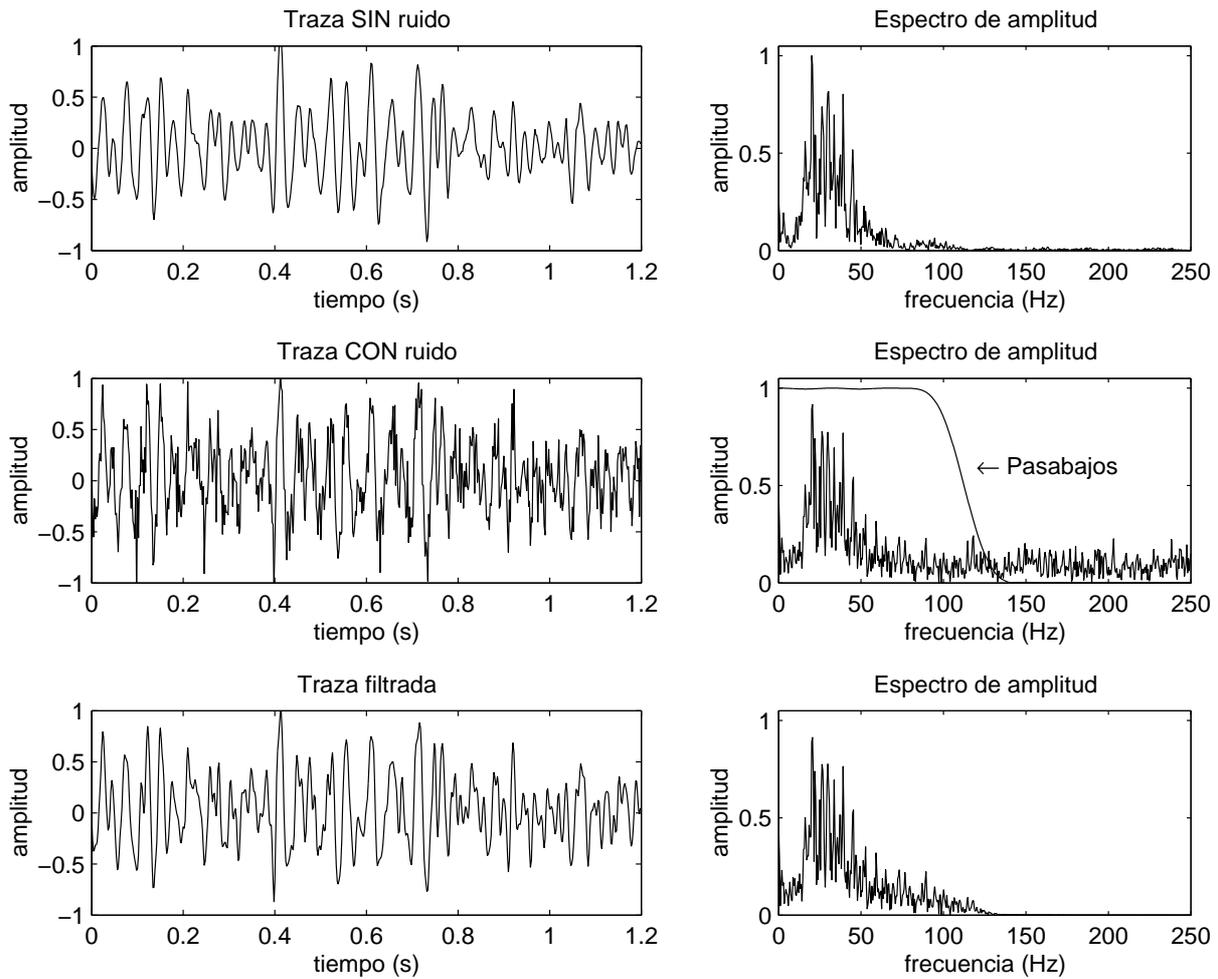


Figura 1.1: Ejemplo del efecto de un filtro pasa-bajos en la traza sísmica.

entran en juego conceptos tales como frecuencia de corte, bandas de paso y de rechazo, etc., como veremos más adelante. Usualmente un filtro se diseña a partir de la respuesta en frecuencia (también llamada función de transferencia) deseada.

Filtros típicos

Existen varios tipos de filtros, entre los más utilizados en sísmica podemos mencionar:

- filtro *pasa-bajos*: se especifica la frecuencia de corte.
- filtro *anti-alias*: pasa-bajos con corte abrupto en la frecuencia de Nyquist.
- filtro *pasa-banda*: se especifican las frecuencias de corte alta y baja.
- filtro *ranura*: elimina una parte muy estrecha del espectro (por ejemplo, 50 Hz). Es un filtro “rechaza-banda” muy angosto.

Filtro de fase cero

Idealmente, todos los filtros anteriores son filtros de *fase cero*, o sea que solamente alteran el espectro de amplitud de la entrada (no actúan sobre la fase). Estos filtros son fácilmente interpretables y su uso es más generalizado. Sin embargo, también existen filtros que actúan exclusivamente sobre la fase de la entrada (filtros dispersivos), o bien sobre ambos componentes del espectro. Estos últimos son utilizados, por ejemplo, para la deconvolución de la traza sísmica, como veremos más adelante.

Los filtros de fase cero suelen ser los más deseables, pues son simétricos, tienen mayor resolución y son más fáciles de interpretar ya que no producen cambios de fase (recordar que el producto de dos transformadas de Fourier es complejo e implica la multiplicación de los espectros de amplitud y la suma de los espectros de fase). La Figura 1.4 muestra la respuesta impulsiva de un filtro de fase cero típico. Por sus características, un filtro de fase cero es un filtro **no causal**, ya que tiene componentes que actúan sobre valores futuros (no disponibles en un sistema analógico), también denominados *componentes de anticipación*. Veremos esto con algún detalle más adelante.

Aplicación

Los filtros digitales se aplican usualmente en el dominio del tiempo convolucionando el dato (traza) con los **coeficientes del filtro**. Permiten diseñar filtros que no se pueden realizar analógicamente, o es muy difícil hacerlo (por ejemplo un filtro que actúa sobre las muestras futuras de la señal). De todos modos, siempre es más práctico aplicar el filtro en el **dominio de las frecuencias**, que es donde suele diseñárselos.

Esencialmente, los filtros digitales se aplican siguiendo los lineamiento de la Figura 1.2. Una vez diseñado el filtro, a través de la especificación de sus espectros de amplitud y fase deseados, se aplica el filtro o bien mediante una convolución (dominio del tiempo), o mediante una multiplicación compleja (dominio de las frecuencias).

Ventajas y desventajas

Dominio del tiempo: un solo paso (convolución), pero más operaciones si la traza y/o el filtro son relativamente “largos”. Por esta razón existe una mayor propagación de los errores. Si L_h y L_x son las longitudes del filtro y la traza respectivamente, se necesitan

$$N_{op} = L_h \times L_x \quad (1.1)$$

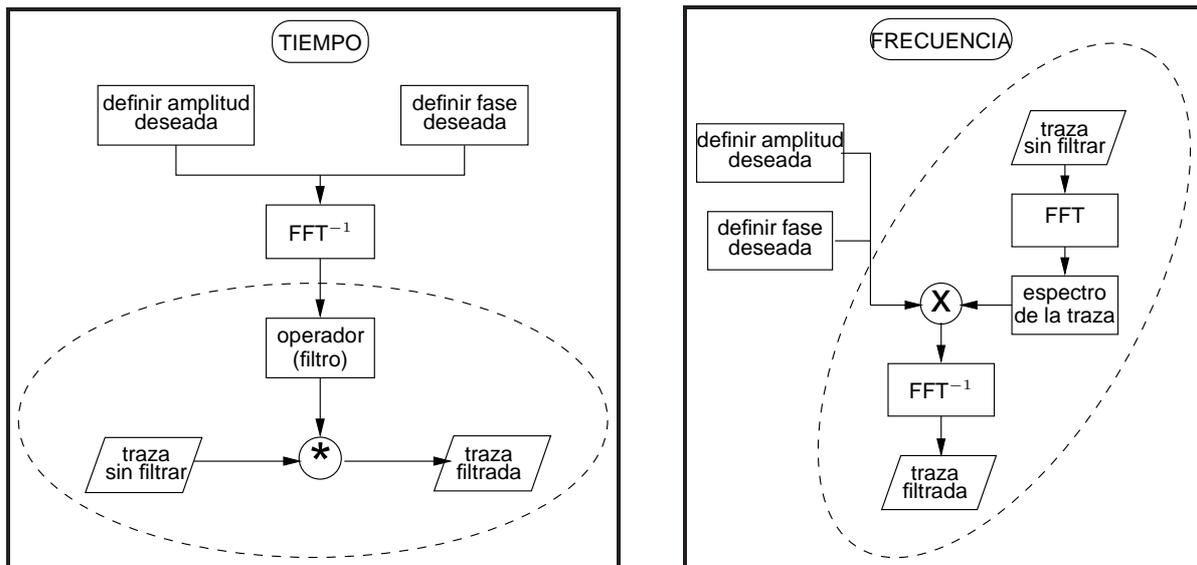


Figura 1.2: Aplicación típica de los filtros digitales en los dominios del tiempo y de las frecuencias.

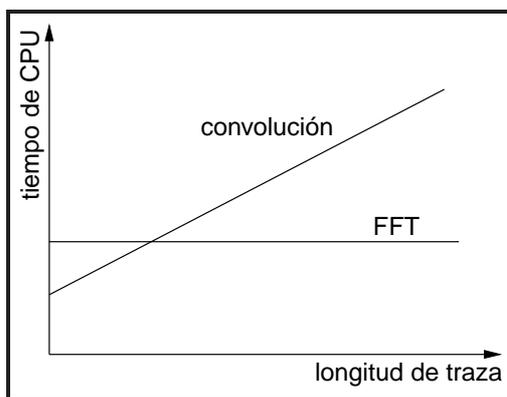


Figura 1.3: Tiempo de CPU vs longitud de la traza.

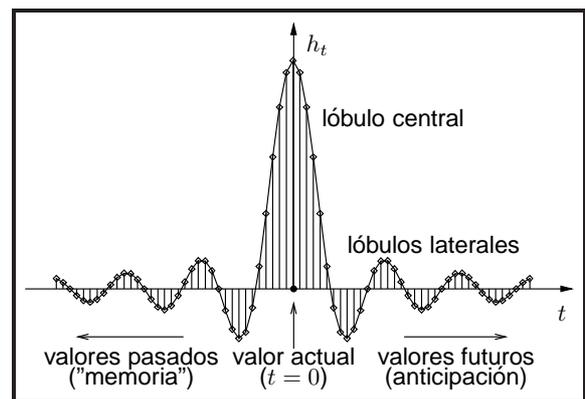


Figura 1.4: Típico filtro de fase cero.

operaciones.

Dominio de las frecuencias: más pasos (FFT, FFT^{-1} , etc.), pero menos operaciones, menor propagación del error y más fácil de interpretar. El número de operaciones es aproximadamente

$$N_{op} = 2M \log_2 M + M, \quad (1.2)$$

donde M es la longitud utilizada para la FFT. Claramente, para un filtro de una longitud determinada, cuanto más larga sea la traza, mayor será la proporción del número de operaciones en el dominio del tiempo con relación al de operaciones en el dominio de las frecuencias, como se ilustra en la Figura 1.3. Por ejemplo, para un filtro de 100 muestras y una traza de 1000, se tienen 100000 operaciones (tiempo) contra sólo 20000 (frecuencia, $M = 1024$).

1.2. Filtros digitales: ceros y polos

En esta sección veremos algunas características de los filtros digitales que están definidos a través de “ceros” y “polos”. El concepto de ceros y polos está íntimamente relacionado con el concepto de Transformada Z, y veremos cómo la posición de los ceros y polos en relación con el círculo unitario (conjunto de puntos del plano complejo tales que $|z| = 1$) juega un papel fundamental para la “realizabilidad” de los filtros digitales.

1.2.1. Filtros digitales elementales

El filtro más simple tiene un solo coeficiente, h_0 , y se puede representar mediante el siguiente diagrama de bloques, donde x_t es la entrada y y_t es la salida¹

$$x_t \rightarrow \boxed{h_0} \rightarrow y_t = h_0 x_t$$

La transformada z es $Y(z) = h_0 X(z)$, donde $z = \exp(-i\omega)$, y $H(z) = h_0$. Otro filtro elemental es aquel que produce un retardo unitario:

$$x_t \rightarrow \boxed{h_1 z} \rightarrow y_t = h_1 x_{t-1}$$

cuya transformada z está dada por $Y(z) = h_1 z X(z)$, con $H(z) = h_1 z$. En forma análoga podríamos definir un filtro elemental que produce un adelantamiento unitario, simplemente cambiando z por z^{-1} en el filtro anterior, o bien filtros que producen 2 o más retardos (o adelantamientos). La combinación de estos diagramas de bloque elementales nos permite diseñar filtros más complejos, como el filtro general que se muestra en la Figura 1.5. En este caso, se tiene

$$\begin{aligned} Y(z) &= h_0 X(z) + h_1 z X(z) + h_2 z^2 X(z) + \cdots + h_{m-1} z^{m-1} X(z) \\ &= (h_0 + h_1 z + h_2 z^2 + \cdots + h_{m-1} z^{m-1}) X(z) \\ &= H(z) X(z). \end{aligned} \tag{1.3}$$

En el dominio del tiempo, este producto implica la convolución entre la respuesta impulsiva del filtro (operador) y la entrada. En el dominio de las frecuencias, donde $z = \exp(-i\omega)$, tenemos un producto complejo.

¹Notar la nomenclatura cuando trabajamos con señales muestreadas: $x_t = x(t\Delta T)$, donde ΔT es el intervalo de muestreo, y $t = 1, 2, \dots$ es el índice de la muestra. Para simplificar la notación, y sin temor a perder la generalidad, de ahora en adelante asumiremos que $\Delta T = 1$.

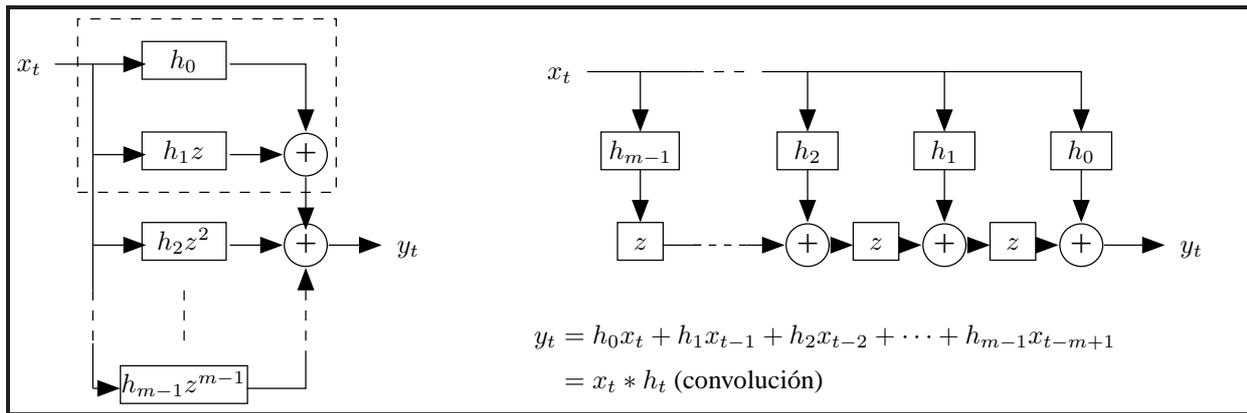


Figura 1.5: Diagrama en bloques de un filtro digital causal de m coeficientes diseñado a partir de bloques elementales. Los dos diagramas son equivalentes.

Filtros en serie y en paralelo

Claramente, cada filtro de retardo no unitario que se muestra en la figura anterior, es el resultado de aplicar dos o más filtros de retardo unitario **en serie**. Por ejemplo, el filtro que produce dos retardos, h_2z^2 , se obtiene tras aplicar dos veces un mismo filtro de retardo unitario, además del filtro de retardo nulo:

$$x_t \rightarrow \boxed{h_2} \rightarrow \boxed{z} \rightarrow \boxed{z} \rightarrow y_t = h_2x_{t-2}$$

La figura anterior también muestra filtros conectados **en paralelo**. Los dos primeros términos de la salida del filtro general de la figura, $h_0x_t + h_1x_{t-1}$, son el resultado de aplicar dos filtros en paralelo, y luego sumar las salidas respectivas, como se destaca en la figura (recuadro punteado).

Causalidad

Los filtros que acabamos de describir se denominan *causales*, pues la salida para el tiempo t (presente) depende solamente de valores presentes y pasados (*memoria* del filtro). El filtro de la Figura 1.5 es el filtro causal más general posible, dado un número finito de coeficientes. Este tipo de filtros también suelen denominarse *moving average (MA) filters* (media en movimiento), ya que la salida es un promedio pesado en “movimiento” de los datos de la entrada.

1.2.2. Característica de amplitud y fase

Veamos algunos casos particulares:

a) Supongamos que $x_t = e^{i\omega t}$ y que $h_t = h_0$. Luego la salida es el vector

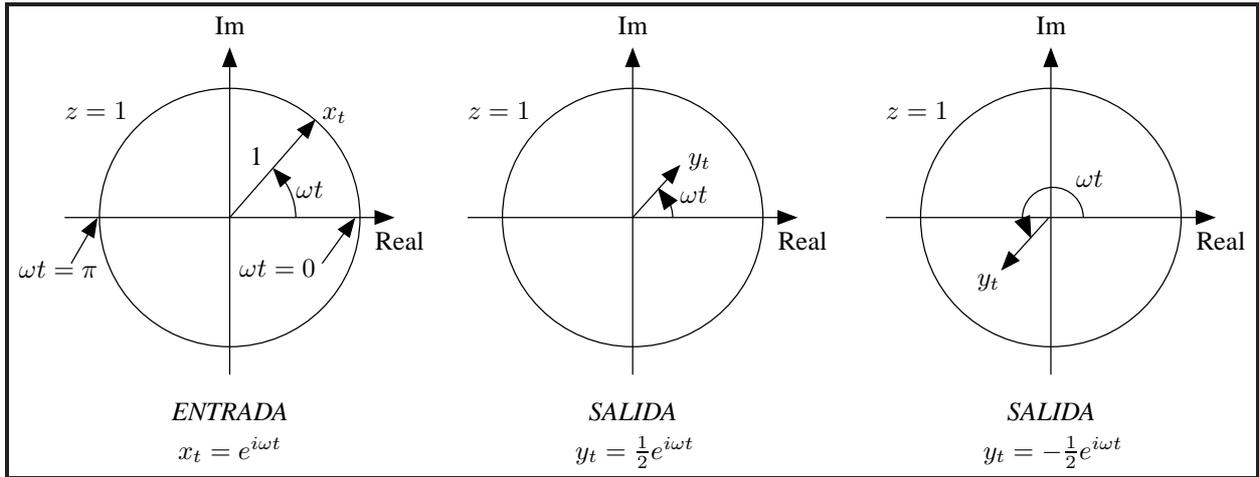


Figura 1.6: Resultado de la aplicación de un filtro elemental. Notar el cambio de fase cuando el coeficiente del filtro es negativo.

$$y_t = h_0 e^{i\omega t}. \tag{1.4}$$

Claramente, si $h_0 > 0$, y_t conserva la misma dirección y sentido que x_t (están *en fase*), pero su módulo (amplitud) aparece multiplicado por h_0 . En cambio, si $h_0 < 0$, se conserva la dirección y cambia el sentido en 180 grados (están *en contra-fase*), porque $(-1) = e^{i\pi}$. En efecto:

$$y_t = -|h_0| e^{i\omega t} = |h_0| e^{i\pi} e^{i\omega t} = |h_0| e^{i(\omega t + \pi)}. \tag{1.5}$$

La **transferencia** del sistema, que se define como la salida sobre la entrada, sería

$$H(\omega) = \frac{\text{salida}}{\text{entrada}} = \frac{h_0 e^{i\omega t}}{e^{i\omega t}} = \begin{cases} h_0 & h_0 > 0 \\ -|h_0| & h_0 < 0 \end{cases} \tag{1.6}$$

En general, $H(\omega)$ es compleja (tiene amplitud y fase). La amplitud (también llamada *espectro de amplitud*) es $|h_0|$ y la fase (también llamada *espectro de fase*) es 0 para $h_0 > 0$ y π para $h_0 < 0$. La Figura 1.6 muestra el comportamiento de estos filtros elementales para $h_0 = \pm 1/2$. El vector x_t gira a una velocidad angular de ω radianes por unidad de tiempo, en el sentido contrario a las agujas del reloj.

b) Consideremos ahora el filtro de retardo unitario y la misma entrada. La salida será

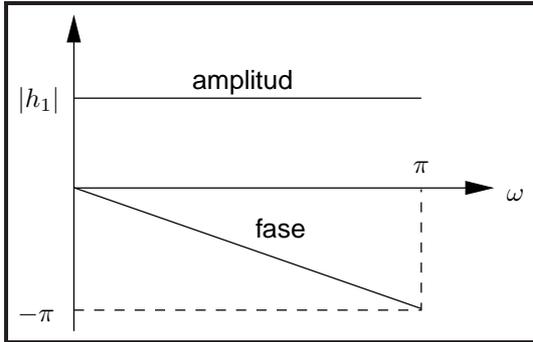


Figura 1.7: Característica de amplitud y fase de un filtro elemental de retardo unitario.

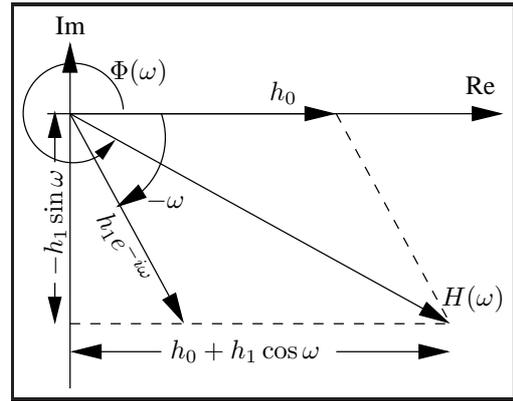


Figura 1.8: Filtro de dos coeficientes y transformada zeta $H(z) = h_0 + h_1z$.

$$y_t = h_1 x_{t-1} = h_1 e^{i\omega(t-1)} \Rightarrow H(\omega) = \frac{h_1 e^{i\omega t} e^{-i\omega}}{e^{i\omega t}} = h_1 e^{-i\omega}. \quad (1.7)$$

El espectro de amplitud es entonces igual a $|h_1|$ y el de fase es $-\omega$, como se ve en la Figura 1.7.

c) Sea ahora un filtro de dos coeficientes ($m = 1$ en la Figura 1.5). Tendremos entonces:

$$y_t = h_0 e^{i\omega t} + h_1 e^{i\omega(t-1)} \Rightarrow H(\omega) = h_0 + h_1 e^{-i\omega}. \quad (1.8)$$

d) Para el caso general, donde el filtro tiene $m + 1$ coeficientes, se tiene

$$y_t = h_0 e^{i\omega t} + \dots + h_m e^{i\omega(t-m)} \Rightarrow H(\omega) = h_0 + h_1 e^{-i\omega} + \dots + e^{-i\omega m}. \quad (1.9)$$

Esta última expresión no es otra cosa que la transformada z de h_t evaluada en el círculo unitario ($z = e^{-i\omega}$).

Es más común escribir la transferencia $H(\omega)$ en forma polar, donde quedan expresados en forma explícita los espectros de amplitud y fase:

$$\boxed{H(\omega) = |H(\omega)| e^{i\Phi(\omega)}} \quad (1.10)$$

donde

$$|H(\omega)|^2 = H(\omega)H^*(\omega) \quad \text{y} \quad \Phi(\omega) = \arctan \left[\frac{\Im H(\omega)}{\Re H(\omega)} \right]. \quad (1.11)$$

Para $m = 1$ tenemos el caso del filtro en (c), donde

$$H(\omega) = h_0 + h_1 e^{-i\omega} = h_0 + h_1 \cos \omega - i h_1 \sin \omega. \quad (1.12)$$

Los espectros de amplitud y fase estarán dados por la magnitud del vector $H(\omega)$ y el ángulo de fase, como se indica en la Figura 1.8:

$$|H(\omega)| = \sqrt{(h_0 + h_1 \cos \omega)^2 + (h_1 \sin \omega)^2} = \sqrt{h_0^2 + h_1^2 + 2h_0 h_1 \cos \omega}, \quad (1.13)$$

$$\Phi(\omega) = - \arctan \frac{h_1 \sin \omega}{h_0 + h_1 \cos \omega}. \quad (1.14)$$

Filtros con fase lineal

Claramente, el filtro de retardo unitario de la sección anterior es un filtro de **fase lineal** (el espectro de fase es una recta de pendiente negativa). Una fase lineal produce un desplazamiento (retardo) de la señal. El desplazamiento o retardo está dado por la pendiente de la fase. Si tenemos dos retardos unitarios (equivalente a la aplicación sucesiva de dos filtros de retardo unitario), la salida resultante será:

$$y_t = h_1 x_{t-2} = h_1 e^{i\omega(t-2)} \quad \Rightarrow \quad H(\omega) = \frac{h_1 e^{i\omega t} e^{-i2\omega}}{e^{i\omega t}} = h_1 e^{-i2\omega}. \quad (1.15)$$

Como vemos, la fase es una recta de pendiente -2 . Esto quiere decir que la salida será igual a la entrada pero retardada en dos unidades de tiempo. En general se pueden tener tantos retardos (o adelantamientos si la pendiente es positiva) como uno quiera.

Como corolario podemos decir que si tomamos una señal cualquiera (una ondícula, una traza sísmica o un filtro, por ejemplo) y la desplazamos en el tiempo, el efecto en el dominio de las frecuencias es tal que el espectro de amplitud de la señal desplazada permanece igual, en tanto que el de fase se ve modificado por la suma de una recta cuya pendiente está dada por el número de muestras o unidades de tiempo que fue desplazada la señal original. Y viceversa: si tomamos una señal cualquiera y le agregamos a su fase una recta con una pendiente determinada, el efecto en el dominio del tiempo

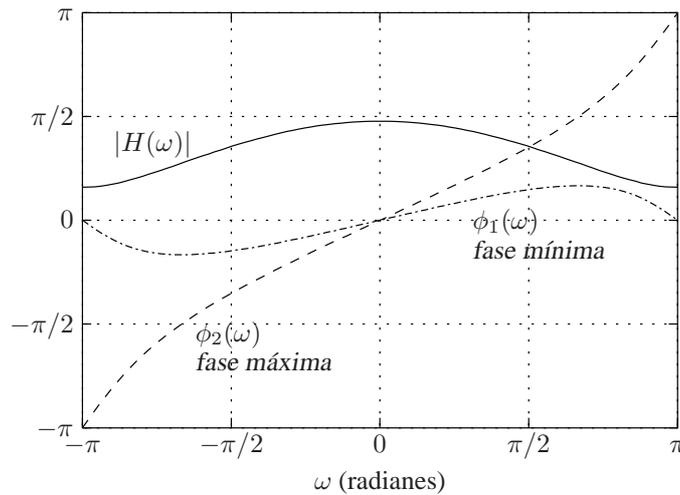


Figura 1.9: Espectros de amplitud y retardo de fase de los filtros $(1, 0.5)$ y $(0.5, 1)$. Ambos tienen el mismo espectro de amplitud, pero sus espectros de retardo de fase son diferentes. El filtro de fase mínima tiene $\phi(\pi) = \phi(-\pi)$, en tanto que el de fase máxima, $\phi(\pi) = \phi(-\pi) + 2\pi$.

es tal que la forma de la señal no cambia pero la misma aparece retrasada o adelantada de acuerdo al signo de la pendiente de la recta sumada a su fase.

Filtros con fase de retardo mínimo

Examinemos los espectros de amplitud y fase de los filtros $h^{(1)} = (1, 0.5)$ (fase mínima) y $h^{(2)} = (0.5, 1)$ (fase máxima). Claramente, $|H(\omega)|$ es invariante si intercambiamos los coeficientes h_0 y h_1 de los dos filtros. Por lo tanto, independientemente de sus fases, los filtros tienen el mismo espectro de amplitud, que es igual a

$$|H(\omega)| = \sqrt{1 + 0.25 + \cos \omega} = \sqrt{1.25 + \cos \omega}. \quad (1.16)$$

Definiendo el **espectro de retardo de fase** como el espectro de fase cambiado de signo, o sea $\phi(\omega) = -\Phi(\omega)$, los mismos son diferentes dependiendo del orden de los coeficientes. En efecto,

$$\phi_1 = \arctan \frac{\sin \omega}{2 + \cos \omega} \quad (1.17)$$

y

$$\phi_2 = \arctan \frac{\sin \omega}{0.5 + \cos \omega}. \quad (1.18)$$

La Figura 1.9 muestra los espectros de amplitud y retardo de fase para estos dos filtros. Notar las diferencias entre los espectros de retardo de fase. El cambio en la fase (al variar ω en todo el rango posible) es menor para $h^{(1)}$ que para $h^{(2)}$.

De todos los posibles filtros de dos elementos con igual espectro de amplitud (en este caso son solo dos filtros posibles), aquel que muestra la menor variación de la fase se denomina *filtro de retardo mínimo* (o filtro de fase mínima).

Familia de filtros con igual espectro de amplitud

El concepto anterior es bien general y se extiende a otros conjuntos de filtros causales con igual número de coeficientes y cuyos espectros de amplitud son iguales. Para un filtro x_t de m coeficientes tenemos

$$X(z) = x_0 + x_1z + x_2z^2 + \cdots + x_{m-1}z^{m-1} = \alpha(z - z_0)(z - z_1) \cdots (z - z_{m-1}), \quad (1.19)$$

donde α es una constante y z_i son los ceros del polinomio en z . El espectro de amplitud del filtro será el producto de los espectros de amplitud de cada uno de los dipolos anteriores. Cada uno de ellos tiene un espectro de amplitud dado por la ecuación (1.13), con $h_0 = -z_i$ y $h_1 = 1$. Si intercambiamos h_0 y h_1 vimos que el espectro de amplitud no varía. Esto es, el espectro de amplitud del dipolo $(z - z_i)$ es idéntico al espectro de amplitud del dipolo $(1 - z_i z)$. Si realizamos estos intercambios para todas las combinaciones posibles de los dipolos de la expresión anterior, tendremos entonces un total 2^m filtros u operadores con igual espectro de amplitud. Naturalmente, los espectros de fase de cada uno de estos 2^m operadores serán diferentes. De todo ellos, el que muestra menor variación total de fase es de fase mínima. Dicho de otro modo, si cada uno de los dipolos es de fase mínima, el operador de la expresión anterior es de fase mínima ².

1.2.3. Filtros digitales de realimentación negativa

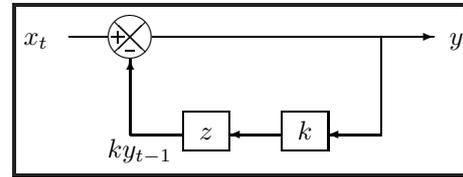
Los filtros de realimentación negativa son aquellos en los cuales la salida (o una parte de ella) ingresa nuevamente a la entrada mediante un *lazo* de retroalimentación. La Figura 1.10 muestra un filtro sencillo de este tipo. La salida tiene dos contribuciones: por un lado tenemos la entrada x_t , y por el otro (el lazo) la salida y_t que pasa por los filtros constante y de retardo unitario, cambiada de signo. Es decir:

²Nota: el filtro será de fase máxima si todos los dipolos son de fase máxima. Y será de fase mixta en todos los otros casos: o sea con al menos un dipolo de fase mínima y un dipolo de fase máxima.

$$y_t = x_t - ky_{t-1}. \quad (1.20)$$

En términos de la transformada z , tenemos

$$Y(z) = X(z) - kzY(z), \quad (1.21)$$



luego

Figura 1.10: Filtro de realimentación negativa

$$X(z) = Y(z) + kzY(z) = Y(z)(1 + kz). \quad (1.22)$$

La función de transferencia de este filtro será pues

$$H(z) = \frac{Y(z)}{X(z)} = \frac{1}{1 + kz} \quad (1.23)$$

Esta ecuación es clave, ya que tiene mucho que ver con la *estabilidad* del filtro (un filtro inestable no puede ser aplicado correctamente). Veamos:

$$H(z) = \frac{1}{1 + kz} = 1 - kz + k^2z^2 - k^3z^3 + \dots \quad (1.24)$$

Claramente, la estabilidad del filtro depende de las características de fase del dipolo $(1 + kz)$. Si éste es de fase mínima (es decir $|k| < 1$), el filtro (1.24) será estable, pues la serie es convergente. En tanto que si el doblete es de fase máxima (es decir $|k| > 1$), el filtro es inestable, pues la serie es no convergente. Este resultado se puede ver mejor si escribimos la salida

$$Y(z) = H(z)X(x) \Rightarrow y_t = x_t - kx_{t-1} + k^2x_{t-2} - k^3x_{t-3} + \dots \quad (1.25)$$

Si $|k| < 1$, el filtro constante de la figura anterior produce una *atenuación*, y el efecto del lazo de retroalimentación decrece con el tiempo (filtro estable). En cambio, si $|k| > 1$, el lazo produce una *amplificación* sin límites. Cabe entonces la pregunta: ¿Cómo se puede resolver este problema si se necesita un filtro de realimentación negativa que produzca una amplificación? La respuesta está vinculada al **retardo**. Introduciendo un retardo en la salida, veremos cómo se puede lograr la estabilidad del filtro aún cuando $|k| > 1$.

Podemos reescribir la ecuación (1.20) de la siguiente manera:

$$y_{t-1} = \frac{1}{k}x_t - \frac{1}{k}y_t. \quad (1.26)$$

Ahora la salida depende de valores *futuros* exclusivamente. Por ejemplo, para $t = 3$, tenemos $y_2 = \frac{1}{k}x_3 - \frac{1}{k}y_3$. Recurriendo a la transformada z , obtenemos:

$$Y(z) = \frac{1}{kz}X(z) - \frac{1}{kz}Y(z) \quad \Rightarrow \quad X(z) = kz\left(1 + \frac{1}{kz}\right)Y(z). \quad (1.27)$$

La transferencia será luego

$$H(z) = \frac{Y(z)}{X(z)} = \frac{1}{kz\left(1 + \frac{1}{kz}\right)} \quad (1.28)$$

Ahora tenemos una serie convergente, ya que $\frac{1}{|k|} < 1$ para $|k| > 1$:

$$\begin{aligned} H(z) &= \frac{1}{kz\left(1 + \frac{1}{kz}\right)} = \frac{1}{kz} \left[1 - \frac{1}{kz} + \frac{1}{(kz)^2} - \frac{1}{(kz)^3} + \dots \right] \\ &= \frac{1}{kz} - \frac{1}{(kz)^2} + \frac{1}{(kz)^3} - \frac{1}{(kz)^4} + \dots, \end{aligned} \quad (1.29)$$

y por lo tanto la salida será

$$y_t = \frac{1}{k}x_{t+1} - \frac{1}{k^2}x_{t+2} + \frac{1}{k^3}x_{t+3} - \frac{1}{k^4}x_{t+4} + \dots \quad (1.30)$$

En este caso hemos logrado un filtro estable con $|k| > 1$. Como contrapartida, ya no tenemos un filtro *causal*, pues la salida depende de valores futuros. Este tipo de filtros se denomina *no causal*, y es un filtro **no realizable**. Por “no realizable” queremos decir que no es posible aplicar el filtro en “tiempo real”, pues necesitamos “esperar” cierto tiempo para acumular las muestras de la entrada (por ejemplo almacenándolas en memoria) para luego aplicar la ecuación anterior. Es decir, este filtro no puede ser aplicado en forma analógica, pero no hay inconvenientes en hacerlo en forma digital, en tanto la serie anterior se pueda trunca sin cometer mayores errores a partir de cierto término (numéricamente despreciable), y el filtro no deba ser aplicado online.

Contrariamente a los filtros causales (filtros MA), cuyos componentes son de “memoria” pues actúan sobre valores pasados (y presente) de la entrada, los filtros *no-causales* tienen además componentes de anticipación, y por ello son filtros *no-realizables*. En el caso que estos filtros sean no-causales puros (es decir que actúen solamente sobre valores futuros de la entrada como el filtro que acabamos de ver), se los conoce con el nombre de *autoregressive (AR) filters*, ya que suponen una autoregresión, como veremos más adelante. Los filtros que contienen tanto elementos de memoria como de anticipación, se conocen como filtros ARMA (autoregressive-moving average filters). Más adelante veremos con más detalles estos sistemas.

1.2.4. Filtros generales con ceros y polos

El filtro digital más general posible se puede expresar como el cociente de dos polinomios en z . Cada uno de estos polinomios se puede factorizar en dobletes o dipolos (hallando las raíces correspondientes), de manera que la función de transferencia se puede expresar, excepto por un factor de escala, mediante la siguiente función racional:

$$H(z) = \frac{a_0 + a_1z + a_2z^2 + \cdots + a_{m-1}z^{m-1}}{b_0 + b_1z + b_2z^2 + \cdots + b_{n-1}z^{n-1}} \propto \frac{(z - \tilde{a}_0)(z - \tilde{a}_1)(z - \tilde{a}_2) \cdots (z - \tilde{a}_{m-1})}{(z - \tilde{b}_0)(z - \tilde{b}_1)(z - \tilde{b}_2) \cdots (z - \tilde{b}_{n-1})}, \quad (1.31)$$

donde los \tilde{a}_i y \tilde{b}_i se denominan *ceros* y *polos* respectivamente. Los filtros vistos en las secciones anteriores constaban solamente de ceros. El agregado de polos nos permitirá elaborar filtros más versátiles para obtener los resultados deseados según la aplicación. Cada uno de los dobletes de la ecuación anterior constituye una unidad básica cuyo comportamiento podemos analizar por separado del resto. Si estudiamos en detalle el espectro de amplitud y fase de cada una de estas unidades (filtro con un solo cero y filtro con un solo polo), luego podremos conocer el espectro del filtro $H(z)$, porque no olvidemos que el espectro de amplitud total será el producto de los espectros de amplitud de los dobletes del numerador dividido por el producto de los espectros de amplitud de los dobletes del denominador. En tanto que el espectro de fase será la suma de los espectros de fase de los dobletes del numerador menos la suma de los espectros de fase de los dobletes del denominador.

1.2.5. Filtro con un solo cero

Ya hemos analizado el comportamiento del espectro de un filtro con un solo cero (fase mínima y fase máxima). ¿Qué podemos decir sobre la posición del cero con respecto al círculo unitario?.

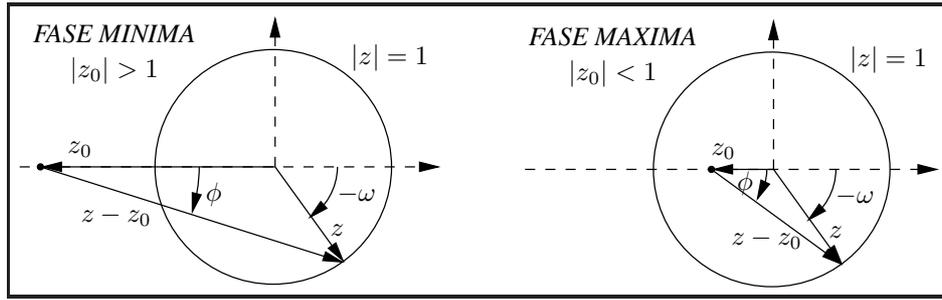


Figura 1.11: Comportamiento del filtro elemental $z - z_0$ de acuerdo a la posición del cero con respecto al círculo unitario. El módulo del vector $z - z_0$ es el espectro de amplitud del filtro, $|H(\omega)|$. El espectro de retardo de fase, $\phi(\omega)$, es el ángulo que forma $z - z_0$ con el eje horizontal, medido en sentido contrario al que aumenta ω (pues $\phi = -\Phi$).

Veremos que conocer la posición de los ceros (y de los polos) es de vital importancia para conocer el comportamiento del filtro $H(z)$. Sea $H(z)$ un filtro elemental con cero z_0 (o sea $H(z_0) = 0$):

$$H(z) = z - z_0. \tag{1.32}$$

Dependiendo del valor de z_0 , el filtro será de retardo de fase mínima o máxima, como se analizó anteriormente para el filtro $H(z) = h_0 + h_1z$ en la Figura 1.9³. Como veremos, lo que determina el comportamiento del filtro es la posición del cero respecto del círculo unitario. Si $|z_0| > 1$ (cero fuera del círculo unitario), el filtro será de fase mínima, pues $|h_0| > |h_1|$. Si $|z_0| < 1$ (cero dentro del círculo unitario), el filtro será de fase máxima, pues $|h_0| < |h_1|$.

Como hemos visto, la posición del cero con relación al círculo unitario es fundamental, ya que determina el comportamiento del filtro. La Figura 1.11 muestra el vector (1.32) al variar ω para $z_0 = -2$ y $z_0 = -1/2$. El espectro de amplitud del filtro es simplemente el módulo de este vector, en tanto que el de retardo de fase es el ángulo indicado por ϕ (sentido antihorario). Visualmente se puede apreciar cómo varían $|H(\omega)| = |z - z_0|$ y $\phi(\omega) = -\Phi(\omega)$ al rotar el vector z desde $\omega = 0$ hasta $\omega = 2\pi$. Cuando el cero está dentro del círculo unitario (fase máxima), el vector $z - z_0$ rodea por completo al origen, generando un cambio de fase total de 2π (comparar con la Figura 1.9). O sea, $\phi(0) - \phi(2\pi) = 2\pi$. Cuando el cero está fuera del círculo unitario (fase mínima), la fase varía mucho menos, oscilando entre dos valores que en ningún caso llegan a π , comenzando y terminando en el mismo valor. El cambio de fase total en este caso es nulo, ya que $\phi(0) - \phi(2\pi) = 0$.

Este análisis es muy importante. Si tuviéramos un filtro con varios ceros, y recordando que el

³Notar que estos filtros son equivalentes excepto por un factor de escala, pues $h_0 + h_1z = h_1(z + h_0/h_1) = h_1(z - z_0)$, con $z_0 = h_0/h_1$. Aquí el factor de escala h_1 es irrelevante.

espectro de fase total será la suma de las fases de cada uno de los dipolos, la fase total dará tantas vueltas al origen como ceros dentro del círculo unitario tenga el filtro. Es decir, el cambio de fase total será $2\kappa\pi$, donde κ es el número de ceros dentro del círculo unitario.

¿Qué ocurre si el cero está sobre el círculo unitario?. A medida que $|z_0|$ se aproxima a $|z| = 1$, $|H(\omega)|$ se aproximará a cero en las proximidades de la frecuencia dada por el vector z_0 . Si el cero está exactamente sobre el círculo unitario, $|H(\omega)| = 0$ para la frecuencia correspondiente. Esto nos permitirá diseñar un filtro que rechace una frecuencia determinada (filtro ranura), como veremos más adelante, utilizando solamente un cero convenientemente ubicado en el plano complejo.

Otra posición particular del cero sería el origen. En este caso $|H(\omega)|$ sería constante y la fase lineal, coincidiendo con el filtro de retardo unitario elemental $H(z) = z$, tal como se vio en la Figura 1.7, pero con $|h_1| = 1$.

1.2.6. Filtro con un solo polo

En el caso de un filtro con un solo polo, el análisis es similar. El filtro con un polo en $z = z_0$ sería

$$H(z) = \frac{1}{z - z_0}. \quad (1.33)$$

En la sección anterior (filtro de realimentación negativa) ya analizamos la estabilidad de un filtro con un solo polo. En ese caso habíamos analizado el filtro $H(z) = 1/(1 + kz)$, que tiene un polo en $z_0 = -1/k$. El análisis de estabilidad realizado para este filtro es entonces totalmente válido para el filtro de la expresión anterior, ya que, salvo factor de escala, se trata del mismo filtro. En efecto,

$$\frac{1}{1 + kz} = \frac{1}{k(z + 1/k)} = \frac{1}{k(z - z_0)}, \quad (1.34)$$

con $z_0 = -1/k$. Es decir, el filtro (1.33) es el mismo filtro que teníamos antes excepto por el factor de escala $1/k$. Luego, dependiendo del valor de $|k|$, tendremos un polo afuera ($|z_0| > 1 \Rightarrow |k| < 1$) o dentro ($|z_0| < 1 \Rightarrow |k| > 1$) del círculo unitario, lo que nos lleva a un filtro estable o inestable, respectivamente.

Resulta interesante analizar qué ocurre cuando el polo está próximo al círculo unitario. Bajo estas circunstancias, $|H(\omega)|$ crecerá mucho para la frecuencia correspondiente. El espectro de potencia (cuadrado del espectro de amplitud) es:

$$|H(z)|^2 = \frac{1}{(z - z_0)(z^* - z_0^*)} = \frac{1}{z_0^2 + z^2 - 2\Re(z_0 z^*)}. \quad (1.35)$$

En coordenadas polares, escribiendo $z = e^{-i\omega}$ y $z_0 = \rho e^{-i\omega_0}$, donde ρ es una constante, se obtiene:

$$|H(\omega)|^2 = \frac{1}{1 + \rho^2 - 2\rho \cos(\omega - \omega_0)}. \quad (1.36)$$

Por simplicidad consideremos ahora sólo los valores de ω próximos a ω_0 , lo que nos permite reemplazar el coseno por los dos primeros términos de su expansión en serie de potencias: $\cos(\omega - \omega_0) = 1 - (\omega - \omega_0)^2/2 + \dots$. Finalmente,

$$|H(\omega)|^2 \simeq \frac{1}{(1 - \rho)^2 + \rho(\omega - \omega_0)^2}. \quad (1.37)$$

Como estamos interesados en un polo muy cerca del círculo unitario, elegimos un filtro causal y estable ubicando el polo a una pequeña distancia ϵ **fuera** del círculo unitario, o sea:

$$\rho = 1 + \epsilon. \quad (1.38)$$

Por lo tanto,

$$|H(\omega)|^2 \simeq \frac{1}{\epsilon^2 + (\omega - \omega_0)^2}, \quad \omega \simeq \omega_0. \quad (1.39)$$

La Figura 1.12 muestra esta función para $\epsilon = 0.1$ y $\omega_0 = \pi/4$. Este tipo de filtros, basados en una conveniente ubicación de polos en el plano complejo, tiene aplicación en el diseño, por ejemplo, de filtros pasa-banda muy angostos, o filtros pasa-bajos y pasa-altos con frecuencias de corte muy abruptas. Un polo producirá un pico muy agudo en el espectro de amplitud.

En la práctica, los filtros con un solo polo (y un solo cero) no son de mucha utilidad ya que tienen una función de transferencia compleja y producen, por ende, salidas complejas. Para remediar este inconveniente, se suele agregar un polo (o un cero) en $z = z_0^*$ (o sea el complejo conjugado). Se obtiene de esta manera un filtro de dos polos (o dos ceros) cuyo espectro de amplitud mantiene las características esenciales del filtro con un solo polo o cero (con un pico adicional en $-\pi/4$ para este ejemplo en particular), pero que produce una salida real.

1.2.7. Filtro inverso

Notar que el filtro anterior puede verse como la inversa de un filtro de un solo cero. Entonces se habla de un filtro inverso estable o inestable, causal o no causal, realizable o no realizable, dependiendo del valor de $|z_0|$ (o de $|k|$). En cuanto al espectro de este filtro inverso, pensemos en la inversa del espectro de amplitud y la fase cambiada de signo correspondiente al filtro de un solo cero $(z - z_0)$, y obtendremos el espectro del filtro “inverso”.

1.2.8. Filtro con un solo cero y un solo polo: filtro pasa-todo

Además del filtro con un cero o un polo más su complejo conjugado, como se mencionó en la sección anterior, el filtro que le sigue en complejidad es un filtro que tiene un cero y un polo. El más interesante de estos filtros es aquel que presenta un espectro de amplitud plano, generando sólo cambios de fase. Este filtro se denomina **filtro pasa-todo**.

Esencialmente, hay tres tipos de filtros pasa-todo:

1. filtro constante (trivial): $H(z) = \text{constante}$, no afecta ni la fase ni la amplitud (*tipo 0*);
2. filtro de un solo cero y un solo polo, que veremos a continuación, también denominado **filtro dispersivo** o *tipo 1*.
3. filtro de retardo puro: $H(z) = z^m$, produce cambios lineales de fase (*tipo 2*)⁴;

Los filtros pasa-todo tipo 0 y tipo 2 ya los hemos descripto en secciones anteriores. El más interesante es el filtro pasa-todo tipo 1 o filtro *dispersivo*. La clave para lograr un filtro pasa-todo (espectro de amplitud plano) con un cero y un polo está en ubicarlos de manera tal que sus efectos sobre el espectro de amplitud se cancelen. Esto se logra haciendo que sean el inverso uno del otro, y complejos conjugados:

$$H(z) = \frac{H_2(z)}{H_1(z)} = \frac{z + k^*}{1 + kz}. \quad (1.40)$$

Como veremos, el cero y el polo de la función anterior son inversos y complejos conjugados. Además, el espectro de amplitud es plano e igual a la unidad. En efecto, hallamos el polo z_1 y el cero z_2 igualando a cero denominador y numerador, respectivamente:

⁴Existe un cuarto tipo de filtro pasatodo denominado filtro pasatodo de retardo impuro, una curiosidad matemática que no detallaremos (*tipo 3*).

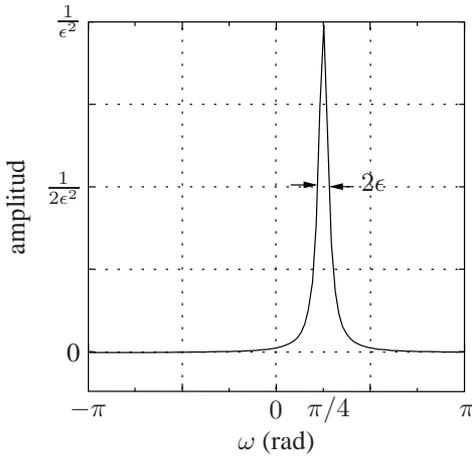


Figura 1.12: Espectro de potencia de un filtro con un polo en $z_0 = (1 + \epsilon)e^{i\pi/4}$, con $\epsilon = 0.1$.

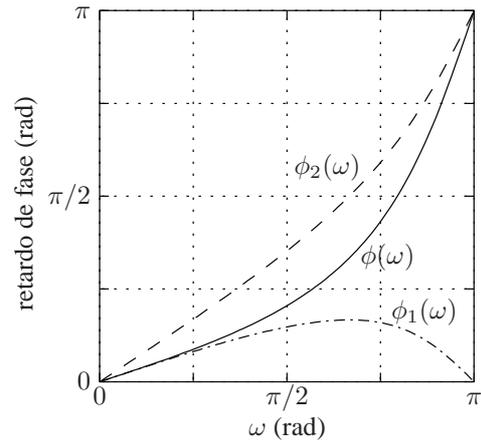


Figura 1.13: Espectro de retardo de fase del filtro pasa-todo tipo 1 (dispersivo) con $k = 0.5$.

$$H_1(z) = 1 + kz = 0 \quad \Rightarrow \quad z = z_1 = -1/k \quad \leftarrow \text{polo} \quad (1.41)$$

$$H_2(z) = z + k^* = 0 \quad \Rightarrow \quad z = z_2 = -k^* \quad \leftarrow \text{cero} \quad (1.42)$$

Claramente $z_1 = 1/z_2^*$. Para hallar el espectro de amplitud hallamos primero el espectro de potencia en el dominio z , que se escribe

$$P(z) = |H(z)|^2 = H(z)H^*(1/z), \quad (1.43)$$

donde $H^*(1/z)$ implica reemplazar z por $1/z$ y tomar el complejo conjugado de todos los coeficientes.

Luego

$$P(z) = \left(\frac{z + k^*}{1 + kz} \right) \left(\frac{1/z + k}{1 + k^*/z} \right) = 1. \quad (1.44)$$

Luego $|H(\omega)| = 1$. Cualquiera sea la entrada a este filtro, el espectro de amplitud de la salida será igual al de la entrada.

Ahora bien, para que el filtro pasa-todo sea **causal** y **estable**, $H_1(z)$ debe ser de fase mínima, por lo que el polo debe estar **fuera** del círculo unitario⁵. O sea $|k| < 1$. Esto implica que el cero estará **dentro** del círculo unitario, resultando $H_2(z)$ de fase máxima. Por lo tanto $H(z)$ será un filtro causal y estable de **fase mixta** (o máxima según diferentes autores).

⁵Notar que este filtro es idéntico al filtro de realimentación negativa visto anteriormente.

Este mismo análisis nos muestra que el filtro dispersivo $H(z)$ tiene el efecto de transformar una ondícula de fase mínima en una de fase máxima:

$$1 + kz \rightarrow \boxed{H(z)} \rightarrow z + k^*,$$

ya que $|k| < 1$. El filtro dispersivo anula el efecto del cero en $-1/k$ y lo reemplaza por un cero en k^* .

En la práctica un filtro dispersivo de tipo 1 puede generalizarse para cualquier número de ceros y polos, siempre y cuando exista un polo por cada cero que guarden la relación inversa y compleja conjugada, y todos los polos se encuentren fuera del círculo unitario. Un filtro dispersivo de tipo 1 se define entonces como aquél cuya transformada z es una función racional formada por un polinomio de mínima fase como denominador, y el correspondiente polinomio de máxima fase como numerador.

Volviendo al filtro dispersivo formado por polinomios de grado 1, sabemos que el mismo es de fase mixta y, debido a que $H_2(z)$ es de fase máxima (y $H_1(z)$ de fase mínima), la misma aumentará en 2π al variar ω de $-\pi$ a π (o de 0 a 2π). Pero ¿cómo es el espectro de fase del filtro dispersivo en sí?. Veamos el caso en que k es real. Los espectros de amplitud de $H_1(z)$ y $H_2(z)$ son iguales, pero no así los de fase, que se restarán para darnos el espectro de fase de $H(z)$:

$$\Phi(z) = \Phi_2(z) - \Phi_1(z). \quad (1.45)$$

Si k es real, entonces $k^* = k$, y las fases surgen a partir de las siguientes expresiones:

$$H_1(z) = 1 + kz = 1 + ke^{-i\omega} \Rightarrow H_1(\omega) = 1 + k \cos \omega - ik \sin \omega \Rightarrow \tan \Phi_1(\omega) = \frac{-k \sin \omega}{1 + k \cos \omega} \quad (1.46)$$

$$H_2(z) = k + z = k + e^{-i\omega} \Rightarrow H_2(\omega) = k + \cos \omega - i \sin \omega \Rightarrow \tan \Phi_2(\omega) = \frac{-\sin \omega}{k + \cos \omega}. \quad (1.47)$$

Luego

$$\begin{aligned} \tan \Phi &= \tan(\Phi_2 - \Phi_1) = \frac{\tan \Phi_2 - \tan \Phi_1}{1 + \tan \Phi_2 \tan \Phi_1} \\ &= \frac{(k^2 - 1) \sin \omega}{2k + (k^2 + 1) \cos \omega}. \end{aligned} \quad (1.48)$$

Aquí hemos utilizado una igualdad trigonométrica para las tangentes y la última igualdad requiere unas pocas operaciones algebraicas. La Figura 1.13 muestra el espectro de retardo de fase $\phi(\omega) = -\Phi(\omega)$ para el caso en que $k = 0.5$.

El filtro dispersivo visto se puede representar mediante el diagrama en bloques de la Figura 1.14a. Para construir este diagrama simplemente hemos tenido en cuenta que el filtro dispersivo se puede reemplazar por dos filtros en serie de la siguiente manera:

$$x_t \rightarrow \boxed{z + k^*} \rightarrow \boxed{1/(1 + kz)} \rightarrow y_t.$$

El primer filtro es un filtro causal que es la suma de un filtro de retardo unitario y uno constante (se aplican en paralelo). El segundo es un filtro de realimentación negativa, estable y causal. Estos filtros se encuentran representados en las Figuras 1.5 y 1.10, respectivamente. Combinando ambos diagramas, se obtiene la Figura 1.14a. La salida estará dada por

$$y_t = k^* x_t + x_{t-1} - k y_{t-1}. \quad (1.49)$$

Si quisiéramos representar el filtro dispersivo *sin* el lazo de retroalimentación, debemos primero calcular los coeficientes del filtro, para lo cual realizamos la división de polinomios:

$$\begin{aligned} H(z) &= \frac{z + k^*}{1 + kz} = (z + k^*)(1 - kz + k^2 z^2 - k^3 z^3 + \dots) \\ &= k^* + (1 - k^* k)z - k(1 - k^* k)z^2 + k^2(1 - k^* k)z^3 - \dots \\ &= h_0 + h_1 z + h_2 z^2 + h_3 z^3 + \dots, \end{aligned} \quad (1.50)$$

donde

$$h_i = \begin{cases} k^* & i = 0 \\ (-k)^{i-1}(1 - k^* k) & i = 1, 2, \dots \end{cases} \quad (1.51)$$

La representación en diagrama de bloques se muestra en la Figura 1.14b. En la práctica, los coeficientes del filtro h_i se hacen cada vez más pequeños, lo que permite realizar un truncamiento para un determinado i .

La salida será entonces una ondícula de longitud infinita, cuyo espectro de amplitud es igual al de la entrada (es un filtro pasa-todo cuyo espectro de amplitud es igual a la unidad). Esta ondícula constituye un operador adicional que se agrega al conjunto de operadores con idéntico espectro de amplitud que teníamos en secciones anteriores:

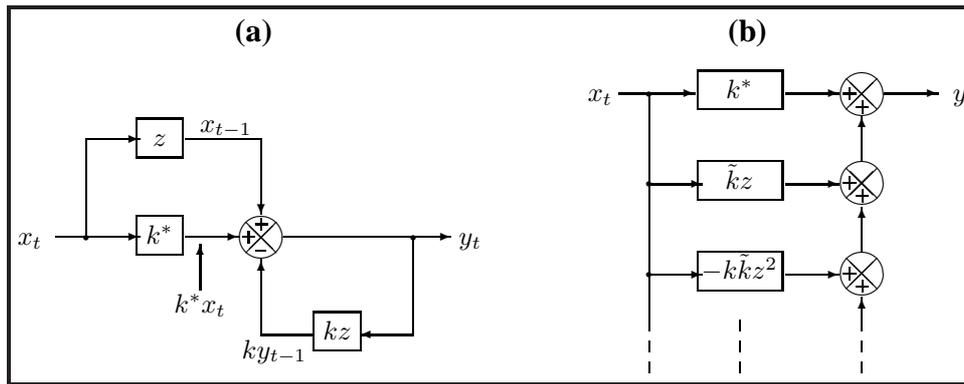


Figura 1.14: (a) Diagrama en bloques de un filtro dispersivo con un polo en $z = -1/k$ y un cero en $z = -k^*$, $|k| < 1$. (b) Representación alternativa sin lazo de realimentación negativa. Aquí $\tilde{k} = (1 - k^*k)$.

1. la entrada x_t y todos los operadores formados por diferentes combinaciones que se obtienen tras factorizar $X(z)$ en sus dipolos e intercambiar los coeficientes de los mismos (2^m combinaciones posibles, para un operador de m coeficientes); y
2. la salida $y_t = x_t * h_t$, donde h_t es un filtro dispersivo pasa-todo.

Naturalmente, de todos ellos sólo uno será de fase mínima y sólo uno de fase máxima. El resto tendrá fase mixta.

1.2.9. Filtros pasa-todo inversos

Son simplemente las inversas de los filtros anteriores. Como aquéllos, estos filtros tienen un espectro de amplitud plano, de manera que $|H(z)| = 1$, y producen solamente corrimientos de fase. Pero, a diferencia de los anteriores, poseen solamente componentes de anticipación, pues son filtros no-causales que trabajan sobre valores futuros.

El filtro pasa-todo inverso de tipo 0 es trivial. El inverso de tipo 2 produce un adelantamiento, y se denomina filtro de *avance puro*. Su transformada z es $1/H(z) = 1/z^m = z^{-m}$. El filtro inverso pasa-todo de tipo 1 (dispersivo inverso) tiene ciertas características interesantes. Recordando la ecuación (1.40), la transformada z del filtro dispersivo inverso será entonces:

$$\frac{1}{H(z)} = \frac{H_1(z)}{H_2(z)} = \frac{1 + kz}{z + k^*}. \quad (1.52)$$

El cero y el polo invierten su posición relativa con respecto al círculo unitario. Desarrollando la

ecuación anterior, y teniendo en cuenta que, por razones de causalidad y estabilidad, k debía ser menor que la unidad, luego

$$\begin{aligned} \frac{1}{H(z)} &= \frac{1 + kz}{z + k^*} = (1 + kz)(z^{-1} - k^*z^{-2} + k^{*2}z^{-3} - \dots) \\ &= k + (1 - kk^*)z^{-1} - k^*(1 - kk^*)z^{-2} + k^{*2}(1 - kk^*)z^{-3} \\ &= h_0^* + h_1^*z^{-1} + h_2^*z^{-2} + h_3^*z^{-3} + \dots, \end{aligned} \quad (1.53)$$

donde los coeficientes h_i^* no son otra cosa que el conjugado de los coeficientes h_i del filtro dispersivo (comparar con ecuación (1.50)). En general, esta relación es válida para todos los filtros pasa-todo.

1.2.10. Recursión

La ecuación (1.31), y que repetimos a continuación, corresponde a un filtro digital general que está expresado como el cociente de dos polinomios de orden n y m , que ha sido factorizada por conveniencia:

$$H(z) = \frac{a_0 + a_1z + a_2z^2 + \dots + a_{m-1}z^{m-1}}{b_0 + b_1z + b_2z^2 + \dots + b_{n-1}z^{n-1}}, \quad (1.54)$$

Si quisiéramos aplicar un filtro de este tipo, por ejemplo a través de la convolución de su respuesta impulsiva con la entrada, necesitamos primero realizar el cociente de polinomios de donde extraer los coeficientes correspondientes. Si estos coeficientes convergen rápidamente (o sea, tienden a cero relativamente rápido), la aplicación del filtro mediante la convolución no presenta mayores inconvenientes, ya que siempre es posible truncar la secuencia para valores pequeños de los coeficientes. Si la convergencia es lenta, el filtro será más largo y por lo tanto su aplicación comienza a traer dificultades (mayor tiempo de CPU, etc.). La recursión nos permite aplicar el filtro sin necesidad de recurrir a la división de polinomios, lo que implica una mayor eficiencia computacional porque se requiere un número mucho menor de operaciones.

Sea el filtro

$$H(z) = \frac{a_0 + a_1z + a_2z^2}{1 + b_1z + b_2z^2 + b_3z^3}, \quad (1.55)$$

con $b_0 = 1$ ⁶. Siendo $H(z) = Y(z)/X(z)$, donde $X(z)$ es la entrada y $Y(z)$ es la salida, entonces

⁶Notar que si $b_0 \neq 1$, basta con dividir todo por b_0 y obtenemos la unidad como primer término en el denominador. Si $b_0 = 0$, no se puede aplicar la recursión.

$$\begin{aligned}
 Y(z) + b_1 z Y(z) + b_2 z^2 Y(z) + b_3 z^3 Y(z) \\
 = a_0 X(z) + a_1 z X(z) + a_2 z^2 X(z).
 \end{aligned}
 \tag{1.56}$$

Por lo tanto

$$y_t = a_0 x_t + a_1 x_{t-1} + a_2 x_{t-2} - b_1 y_{t-1} - b_2 y_{t-2} - b_3 y_{t-3}. \tag{1.57}$$

Esta expresión recursiva nos permite obtener la salida sin necesidad de obtener los coeficientes de la división de polinomios. Por supuesto, asumimos que $y_t = 0$ para $t < 0$. En general, para el caso de la ecuación (1.54), y habiendo previamente forzado un “1” en el denominador, se tiene

$$y_t = \sum_{i=0}^{m-1} a_i x_{t-i} - \sum_{i=1}^{n-1} b_i y_{t-i} \tag{1.58}$$

La Figura 1.15 muestra el diagrama en bloques de la recursión para el ejemplo de la ecuación (1.57). En este caso, para obtener cada y_t se requieren seis operaciones (sumas y productos). Si la división de polinomios permitiera truncar la serie en seis o menos coeficiente, la recursión no representaría ningún ahorro computacional. Pero en general el truncamiento lleva a un número mucho mayor de coeficientes significativos, poniendo en evidencia la importancia de este procedimiento.

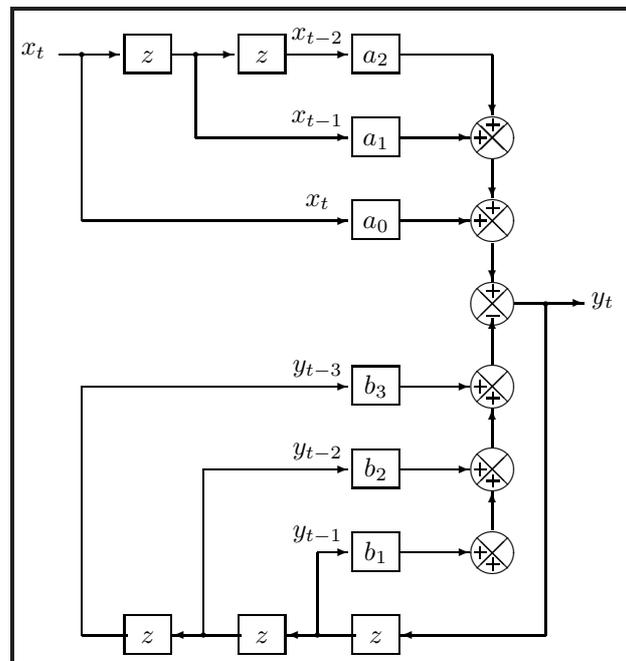


Figura 1.15: Diagrama en bloques de un filtro recursivo (ver texto).

1.2.11. MA, AR, ARMA, FIR, IIR

La expresión anterior está compuesta por dos sumatorias o partes. La primera de ellas indica que la salida depende de una suma ponderada de valores presentes y pasados de la entrada, siendo este “promedio” variable con el tiempo. Este tipo de filtro se denomina “moving average” (MA), pues la salida representa un promedio en movimiento (o promedio móvil) de valores de la entrada. La otra sumatoria involucra valores pasados de la salida misma, e implica una recursión o regresión, por lo que se lo denomina “autoregresivo” (AR). Un filtro que contiene solamente m coeficientes a_i se denomina entonces filtro MA de orden m , y uno que contiene solamente n coeficientes b_i se denomina filtro AR de orden n . Si los dos conjuntos de coeficientes están presentes, se tiene un filtro ARMA de orden (m, n) . En definitiva, MA implica ceros, AR implica polos, y ARMA implica ceros y polos simultáneamente.

Notar además que un filtro AR (o un filtro ARMA) tiene en general infinitos coeficientes MA, por lo que además se lo suele llamar filtro IIR (infinite impulse response). Un filtro MA tiene un número finitos de coeficientes y se lo conoce como filtro FIR (finite impulse response).

Los sistemas MA, AR y ARMA son además modelos matemáticos muy utilizados en el análisis de series temporales (o no) que apuntan por ejemplo a poder predecir valores futuros de las mismas, y tienen en realidad una connotación mucho más profunda que la de un filtro con ceros y/o polos, como veremos más adelante. Estos modelos matemáticos permite ajustar determinada serie de tiempo (por ejemplo la evolución de cierta variable económica, física, etc.) mediante un número finito de parámetros. El ajuste suele hacerse mediante el método de mínimos cuadrados. Más allá de las operaciones necesarias para hallar los coeficientes del modelo elegido para representar la serie de tiempo en cuestión, la determinación del orden del modelo es crucial, y muchas veces muy difícil de determinar sin alguna información adicional acerca de los datos.

1.2.12. Síntesis de filtros vía ceros y polos

Como se mencionó anteriormente, la ubicación estratégica de ceros y polos nos permite diseñar filtros digitales con ciertas características deseadas. Por ejemplo, un filtro con un solo cero nos permitirá rechazar cierta frecuencia. Consideremos un ejemplo sencillo: un filtro con un solo cero capaz de eliminar la componente continua (frecuencia cero) de una entrada dada, y analicemos su espectro.

Como se vio anteriormente, está claro que si colocamos un cero en $z_0 = 1 + i0$ rechazaremos la frecuencia $\omega_0 = 0$, ya que $z_0 = e^{-i\omega_0} = \cos \omega_0 - i \sin \omega_0 = 1$. Un filtro que elimina la componente

continua podrá ser entonces

$$H(z) = z - 1. \quad (1.59)$$

Este filtro está representado en la Figura 1.16. Para $z = 1$ (o sea $\omega = 0$), $|H(z)| = 0$. A medida que el vector z recorre el círculo unitario, $|H(z)|$ aumenta hasta un máximo para $\omega = \pi$, donde $|H(z)| = 2$. Luego decrece hasta cero nuevamente. El espectro de amplitud de este filtro está dado por

$$|H(\omega)| = |e^{-i\omega} - 1| = \sqrt{2(1 - \cos \omega)}, \quad (1.60)$$

y se encuentra representado en la Figura 1.17. Cabe aquí una observación. De la figura anterior se desprende claramente que además de rechazar la frecuencia cero, el filtro atenúa significativamente las frecuencias cercanas al cero. Este es un efecto no deseado, pues el interés está en rechazar el cero y alterar lo menos posible el resto de las frecuencias.

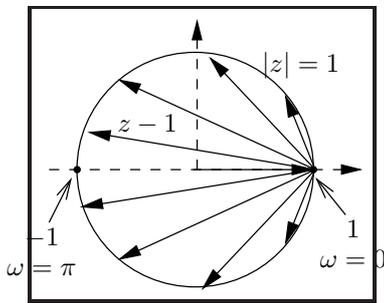


Figura 1.16: Filtro con un cero en $z_0 = 1$. Sólo para $z = 1$ ($\omega = 0$), $|H(z)| = 0$.

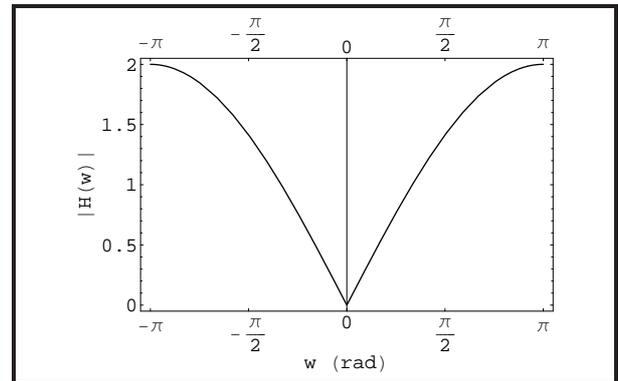


Figura 1.17: Espectro de amplitud del filtro con un sólo cero en $z_0 = 1$.

Esto se puede remediar parcialmente recurriendo a los polos. En los filtros dispersivos vimos cómo un polo colocado estratégicamente anulaba el efecto de un cero. Lo que queremos nosotros es anular el efecto del cero en las frecuencias cercanas al origen, de manera de minimizar el efecto no deseado mencionado. Para ello, colocamos un polo próximo al cero, sobre el eje real, y fuera del círculo unitario (para no tener problemas con la estabilidad). Por ejemplo elegimos $z_0 = 1 + \epsilon + 0i$, donde ϵ es una cantidad pequeña. Entonces

$$H(z) = \frac{z - 1}{z - (1 + \epsilon)}. \quad (1.61)$$

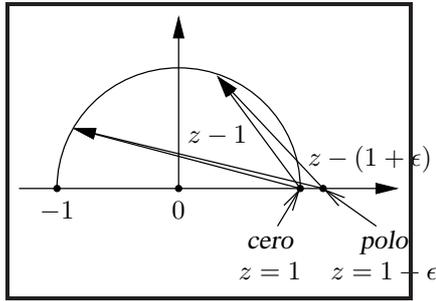


Figura 1.18: Filtro con un cero en $z_0 = 1$ y un polo en $z_0 = 1 + \epsilon$. Notar que $|H(z)| \simeq 1$ excepto para $z = 1$ ($\omega = 0$), donde $|H(z)| = 0$.

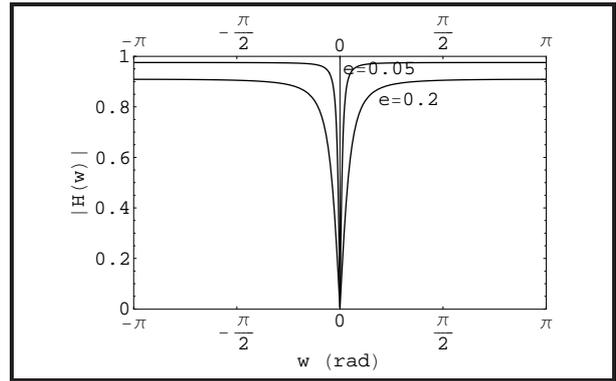


Figura 1.19: Espectros de amplitud de los filtros con un cero en $z_0 = 1$ y un polo en $z_0 = 1 + \epsilon$ ($\epsilon = 0.2$ y 0.05).

Este filtro es el cociente de dos vectores cuyos módulos son muy similares, de manera que el espectro de amplitud será próximo a la unidad excepto para $z = 1$, donde valdrá cero. La Figura 1.18 muestra el comportamiento de este filtro a medida que z recorre el círculo unitario. El efecto obtenido es el de un filtro ranura muy abrupto, como se ve en la Figura 1.19, donde hemos representado el espectro de amplitud para $\epsilon = 0.2$ y 0.05 .

Por supuesto, en la práctica no tiene mucho sentido usar este filtro para eliminar la frecuencia cero (bastaría con restar el promedio a la serie de entrada). Usualmente se requiere filtrar otras frecuencias, como por ejemplo la frecuencia $f_0 = 50\text{Hz}$ (y $f_0 = -50\text{Hz}$), que suele ser un ruido introducido en la señal sísmica por la presencia del cableado de alta tensión. Por simplicidad hemos considerado el caso de $f_0 = 0$ pues solamente se requiere un cero y un polo. En el caso de interés ($f_0 = \pm 50\text{Hz}$), necesitaremos 2 ceros y 2 polos.

Veamos ahora otro ejemplo algo más complejo y con mayor sentido práctico: un filtro ranura abrupto capaz de eliminar las frecuencias $f_0 = \pm 50\text{Hz}$, con $\Delta t = 4\text{ms}$ (frecuencia de muestreo). En este caso, ubicaremos 2 ceros en $\pm 50\text{Hz}$ y dos polos próximos a ellos, pero fuera del círculo unitario. Primero debemos obtener la frecuencia en radianes y el ángulo correspondiente (recordemos que en los análisis anteriores habíamos asumido $\Delta t = 1$, por lo cual teníamos $\omega\Delta t = \omega$):

$$\omega = 2\pi f \Rightarrow \omega_0 = \pm 2\pi 50 \text{ s}^{-1} = \pm 314.16 \text{ s}^{-1} \Rightarrow \omega_0 \Delta t = \pm 1.257 \text{ radianes} \equiv \pm 72^\circ.$$

Alternativamente, la conversión de Hz a grados se puede obtener sabiendo que la frecuencia de Nyquist, $1/(2\Delta t)$, se corresponde con π radianes. Luego, por regla de tres simple, 50 Hz se corresponde con 1.255 radianes (72 grados).

Luego elegimos $\epsilon = 0.1$ (por ejemplo) y ubicamos los ceros y polos de la siguiente manera (ver

Figura 1.20):

ceros

$$z_{01} = \cos 72^\circ + i \sin 72^\circ = 0.309 + i0.951$$

$$z_{02} = \cos 72^\circ - i \sin 72^\circ = 0.309 - i0.951$$

polos

$$z_{p1} = 1.1 \cos 72^\circ + 1.1 \sin 72^\circ i = 0.340 + 1.046i$$

$$z_{p2} = 1.1 \cos 72^\circ - 1.1 \sin 72^\circ i = 0.340 - 1.046i$$

El filtro ranura puede escribirse finalmente como la siguiente función racional:

$$H(z) = \frac{(z - z_{01})(z - z_{02})}{(z - z_{p1})(z - z_{p2})} \simeq \frac{1 - 0.618z + z^2}{1.210 - 0.680z + z^2} = \frac{0.826 - 0.511z + 0.826z^2}{1 - 0.562z + 0.826z^2}. \quad (1.62)$$

La división de estos polinomios no es exacta, y si bien es posible truncar la serie para cierto retardo a partir del cual los coeficientes son muy pequeños, es conveniente aplicar la recursión como se vio en la sección anterior (ver ecuación (1.55) con $b_3 = 0$), resultando

$$y_t = 0.826x_t - 0.511x_{t-1} + 0.826x_{t-2} + 0.562y_{t-1} - 0.826y_{t-2}. \quad (1.63)$$

El espectro de amplitud, que está representado en la Figura 1.21, y resulta de realizar el cociente entre los espectros de amplitud del numerador y del denominador de la ecuación anterior, muestra claramente el efecto que tendrá este filtro sobre el espectro de una señal de entrada (por ejemplo una traza sísmica). Sólo las frecuencias muy próximas a $f_0 = \pm 50\text{Hz}$ resultan atenuadas considerablemente, que es lo que queríamos obtener. Notar que la frecuencia de Nyquist es $f_N = 1/2\Delta t = 125\text{Hz}$.

¿Y la fase?

Hasta ahora no hemos hablado del espectro de fase de este tipo de filtros. En general la fase no es lineal, y mucho menos nula, lo que implica un inconveniente si queremos aplicar estos filtros a señales sísmicas. Para evitar cambios en la fase, estos filtros se aplican a través de los siguientes pasos:

- 1) se aplica el filtro como se vio recién;
- 2) se revierte en tiempo la salida y se aplica el filtro nuevamente;
- 3) se revierte en tiempo el resultado final.

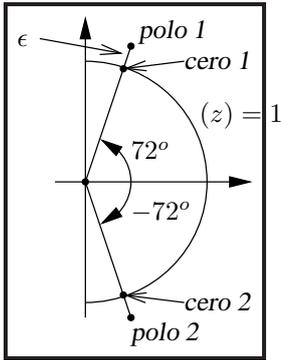


Figura 1.20: Ubicación de los ceros y los polos en un filtro ranura con $f_0 = \pm 50\text{Hz}$.

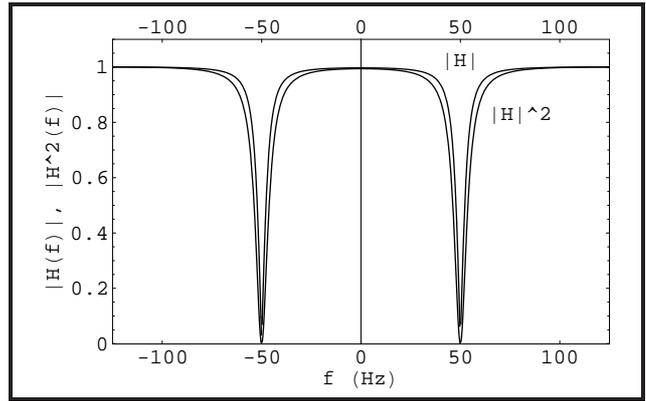


Figura 1.21: Espectros de amplitud y de potencia (normalizados) del filtro ranura de dos ceros y dos polos con $f_0 = \pm 50\text{Hz}$ ($\Delta t = 4\text{ms}$) y $\epsilon = 0.1$.

Como consecuencia de este proceso el efecto en la fase se cancela y el efecto en la amplitud se duplica. En efecto, si la traza a filtrar es x_t , la salida del paso (1) será

$$y_t^{(1)} = h_t * x_t. \tag{1.64}$$

Luego del paso (2), tendremos

$$y_t^{(2)} = h_t * y_{-t}^{(1)} = h_t * h_{-t} * x_{-t}. \tag{1.65}$$

Finalmente, tras el último paso, tendremos

$$y_t^{(3)} = y_{-t}^{(2)} = h_{-t} * h_t * x_t. \tag{1.66}$$

Recordando la propiedad de escalado (o teorema de similitud) para la transformada de Fourier, con $\alpha = \text{constante}$,

$$h_t \leftrightarrow H(\omega) \Rightarrow h(\alpha t) \leftrightarrow H(\omega/\alpha)/|\alpha|,$$

y haciendo $a = -1$ vemos que

$$h_t * h_{-t} \leftrightarrow H(\omega)H(-\omega).$$

Luego

$$\begin{aligned}
 H(\omega)H(-\omega) &= |H(\omega)|e^{i\Phi(\omega)}|H(-\omega)|e^{i\Phi(-\omega)} \\
 &= |H(\omega)||H(-\omega)|e^{i[\Phi(\omega)+\Phi(-\omega)]} \\
 &= |H(\omega)|^2,
 \end{aligned}
 \tag{1.67}$$

en virtud de las conocidas simetrías de la transformada de Fourier: $|H(-\omega)| = |H(\omega)|$ y $\Phi(-\omega) = -\Phi(\omega)$. El efecto en el espectro se puede apreciar en la Figura 1.21.

1.2.13. Ceros, polos y fase mínima: revisión de conceptos

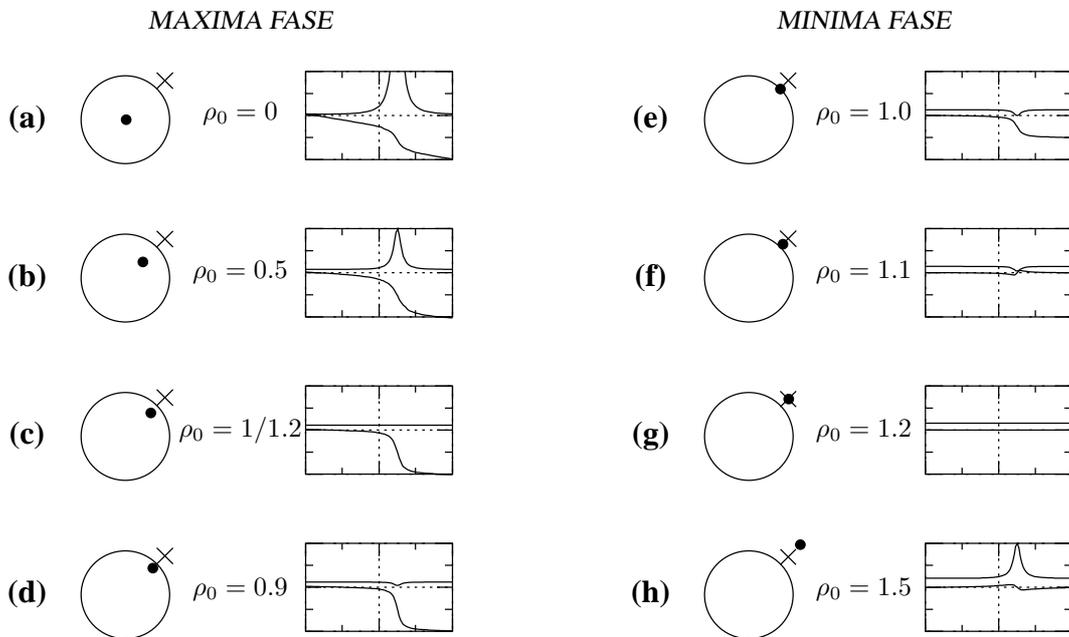


Figura 1.22: Filtro con un solo polo (cruz) y un solo cero (punto) sobre la misma línea radial. El polo está ubicado en $z_p = 1.2 \exp(-i\omega_0)$ y el cero en $z_0 = \rho_0 \exp(-i\omega_0)$, con $\omega_0 = \pi/4$ y diferentes ρ_0 . Los espectros de potencia y de fase están graficados para ω en $[-\pi, \pi]$, en tanto que la escala vertical es $[-2\pi, 2\pi]$. Cero *dentro* del círculo unitario (fase máxima): (a) Filtro con un solo polo (filtro pasa-banda angosto); (b) el cero comienza a cancelar al polo; (c) el cero cancela totalmente al polo (filtro pasa-todo); (d) el cero domina al polo (filtro ranura). Cero *fuera* del círculo unitario (fase mínima, excepto (e)): (e) cero sobre el círculo unitario (filtro ranura); (f) el cero comienza a cancelarse con el polo (filtro ranura débil); (g) el cero cancela totalmente al polo (filtro constante); y (h) el cero más allá del polo produce un filtro pasa-banda angosto “elevado”.

De la misma manera que hemos *sintetizado* este filtro ranura en base a la ubicación estratégica de ceros y polos, es posible sintetizar filtros más elaborados simplemente agregando un mayor número de ceros (y polos). Antes de avanzar sobre este tema, revisemos cualitativamente el efecto que tiene

la ubicación relativa de un polo con respecto a un cero. Para ello consideremos el caso de un filtro con un solo cero y un solo polo colocados sobre la misma línea radial:

$$H(z) = \frac{z - z_0}{z - z_p} = \frac{e^{-i\omega} - \rho_0 e^{-i\omega_0}}{e^{-i\omega} - \rho_p e^{-i\omega_0}} \quad (1.68)$$

Realizando algunas manipulaciones algebraicas⁷, se obtiene el espectro de potencia y el de fase siguiente:

$$|H(\omega)|^2 = \frac{1 + \rho_0^2 - 2\rho_0 \cos(\omega - \omega_0)}{1 + \rho_p^2 - 2\rho_p \cos(\omega - \omega_0)} \quad (1.69)$$

y

$$\Phi(\omega) = -\arctan\left(\frac{\sin \omega - \rho_0 \sin \omega_0}{\cos \omega - \rho_0 \cos \omega_0}\right) + \arctan\left(\frac{\sin \omega - \rho_p \sin \omega_0}{\cos \omega - \rho_p \cos \omega_0}\right) \quad (1.70)$$

La Figura 1.22 muestra estas funciones para $\rho_p = 1.2$, $\omega_0 = \pi/4$ y varios valores de ρ_0 . El epígrafe de la figura explica el significado de cada panel. Notar que todos estos filtros son estables debido a que el polo está ubicado *fuera* del círculo unitario. La condición de mínima o máxima fase está dada por la posición del cero, y se puede apreciar viendo el espectro de fase que produce cada uno de los filtros de la figura.

En secciones anteriores vimos que un filtro con un solo polo es estable y causal si el polo está fuera del círculo unitario, ya que la serie es convergente y depende solamente de valores presentes y pasados (ver sección sobre el filtro de realimentación negativa)⁸. Si le agregáramos un cero fuera del círculo unitario, tendríamos un filtro que, además de causal y estable, sería de fase mínima. Si en cambio el cero estuviera dentro del círculo unitario, el filtro sería de fase máxima. Un análisis similar se podría hacer para un filtro general con varios polos y ceros como el de la ecuación (1.31). Si el denominador es mínima fase (todos los polos fuera del círculo unitario), el filtro será estable y causal, y la condición de mínima fase o no estará dada por el numerador: todos los ceros fuera del círculo unitario nos darían un filtro de fase mínima, todos dentro, uno de fase máxima. Si tenemos algunos ceros fuera y otros dentro del círculo unitario, el filtro resultante será de fase mixta.

⁷Considerar numerador y denominador por separado y recordar que los espectros de amplitud se dividen y los de fase se restan.

⁸Si el polo estuviera *dentro* del círculo unitario, el filtro sería (1) inestable y causal, o (2) estable y no-causal, dependiendo de valores futuros, es decir tendría componentes de anticipación.

Cabe aclarar que algunos autores hablan de fase máxima si todos los ceros y polos están dentro del círculo unitario (y es anticausal puro por cuestiones de estabilidad). Otros autores hablan de fase máxima si todos los polos están fuera del círculo unitario (por cuestiones de estabilidad) y todos los ceros están dentro (en este caso se trata de un filtro causal). Karl, por ejemplo, menciona el segundo caso como de fase mixta, lo que implicaría que se inclina por la primera de las definiciones. En todo caso, conviene hacer la distinción entre sistemas causales y no-causales, y luego hablar de fase máxima. Si uno piensa un sistema de fase máxima como lo opuesto a uno de fase mínima, entonces la primera definición sería la apropiada. Pero si hablamos de sistemas causales exclusivamente, entonces la segunda opción sería la adecuada.

¿Por qué insistimos tanto con el concepto de fase mínima? La respuesta tiene dos aspectos. El primer aspecto se refiere a la relación que existe entre mínima fase y un sistema real físicamente realizable (estabilidad y causalidad de la inversa, etc.). El segundo aspecto, y tal vez con una implicación más profunda, reside en una importante aplicación en sismica de reflexión: la deconvolución. Como veremos más adelante, en la deconvolución es esencial el concepto de filtro inverso. Resulta que para obtener la inversa de un operador de fase mínima (o fase máxima) basta con conocer su auto-correlación. Esto es equivalente a requerir solamente el espectro de amplitud. La información de fase está contenida en el hecho de que el operador es mínima fase. En la deconvolución, este “operador” suele ser la ondícula sísmica, que no siempre es conocida con precisión. Pero si la ondícula es de fase mínima, no es difícil estimar su auto-correlación a partir de la auto-correlación de la traza sísmica, con lo cual podremos determinar la ondícula y luego aplicar el operador inverso en la deconvolución.

A modo de conclusión, la Figura 1.2.13 muestra una clasificación de todas las secuencias posibles, de acuerdo a lo visto en este capítulo.

1.2.14. Teorema de retardo de energía de Robinson

Veremos ahora otro concepto importante asociado a un filtro de mínima fase, que nos permitirá comprender más claramente qué significa “físicamente” que un filtro sea de mínima fase o no. Como veremos, este concepto está asociado al retardo de la energía. Ya hemos visto anteriormente que el cambio de fase es mínimo para un dipolo cuyo primer coeficiente es mayor que el segundo, luego extendimos este concepto al caso de muchos dipolos, concluyendo que un operador o filtro será de mínima fase si todos los dipolos que lo conforman (tras hallar los ceros de su transformada z) son de fase mínima.

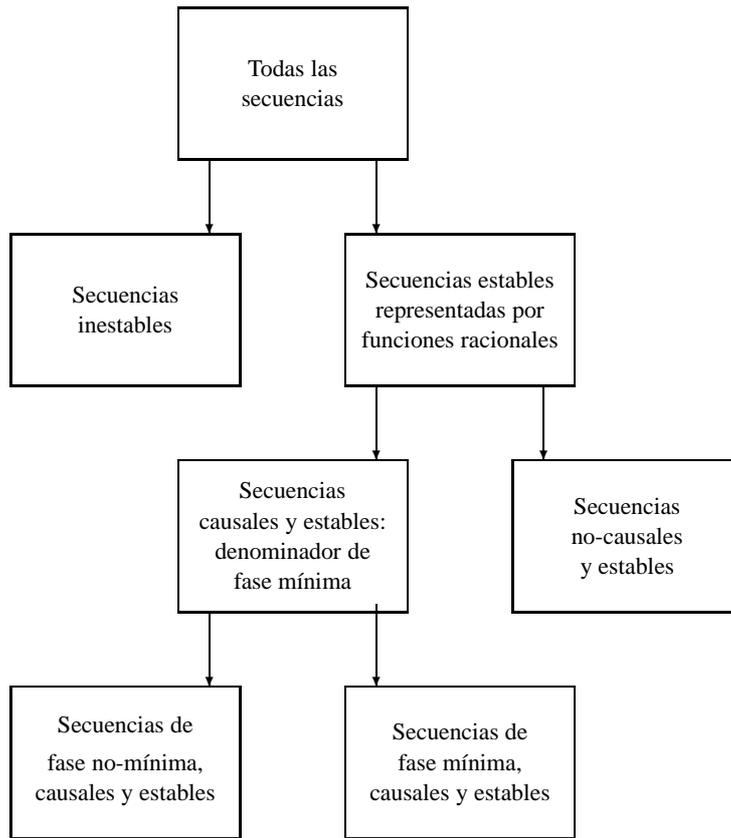


Figura 1.23: Clasificación de todas las secuencias posibles. La secuencia de fase mínima es una sub-clase de una secuencia causal y estable.

Ahora bien, la **energía parcial** de un filtro u operador se define como la energía acumulada para el tiempo t . O sea, si el operador es (h_0, h_1, \dots, h_m) , la energía parcial para el tiempo t es

$$E(t) = \sum_{k=0}^t h_k^2. \tag{1.71}$$

Notar que $E(m)$ es igual a la **energía total** del operador. Si consideramos un dipolo (h_0, h_1) , con $|h_0| > |h_1|$, la energía parcial será mayor que la correspondiente al dipolo (h_1, h_0) , cuyo espectro de amplitud es idéntico al del primer dipolo. Esto quiere decir que le energía arriba **antes** en el caso de mínima fase que en el de máxima. Naturalmente, la energía total de ambos dipolos es la misma, y vale $h_0^2 + h_1^2$.

Veamos un ejemplo. Si $h_t = (1, 0.5)$, su energía parcial es $(1, 1.25)$, en tanto que para $h_t = (0.5, 1)$, la energía parcial es $(0.25, 1.25)$. Esta propiedad se verifica para operadores de mayor longitud. Es decir, **la energía sumada desde el tiempo 0 hasta cualquier tiempo t para un operador de mínima fase, es mayor o igual que la energía correspondiente para cualquier otro operador con**

igual espectro de amplitud. En el caso del operador de máxima fase y con igual espectro de amplitud, la energía parcial será siempre mayor, y en el caso de fase mixta, tendrá valores intermedios.

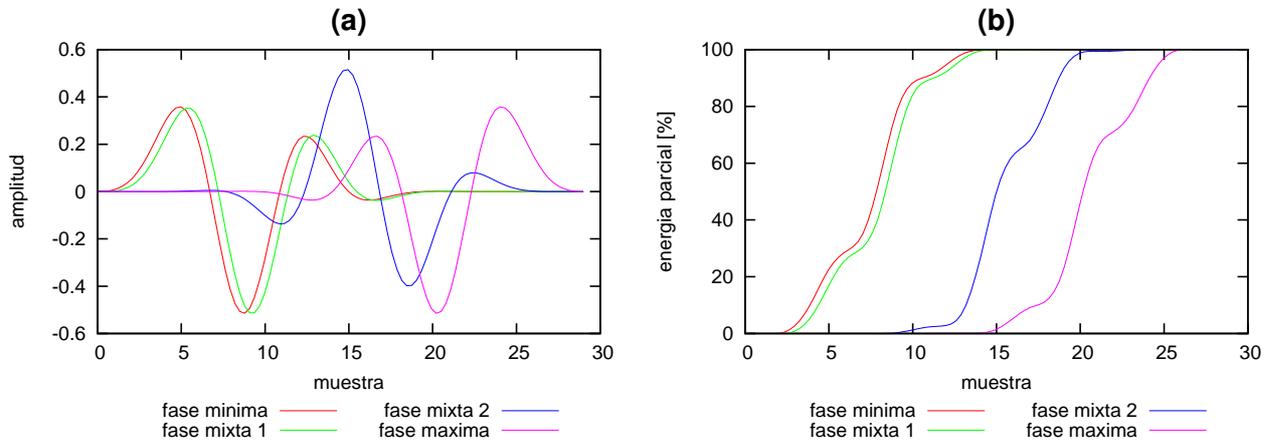


Figura 1.24: (a) Cuatro operadores con igual espectro de amplitud construido a partir de la convolución de 29 dipolos. (b) Energía parcial de los cuatro operadores de (a).

La Figura 1.2.14a muestra la respuesta impulsiva de cuatro operadores o filtros de 30 coeficientes con igual espectro de amplitud. Estos operadores fueron construidos a partir de la convolución de 29 dipolos reales, acomodando los coeficientes de cada dipolo de manera tal que:

1. todos los ceros están fuera del círculo unitario (fase mínima);
2. todos los ceros excepto uno están fuera del círculo unitario (fase mixta 1);
3. 12 ceros están fuera del círculo unitario y 17 están dentro (fase mixta 2); y
4. todos los ceros están dentro del círculo unitario (fase máxima).

Como se puede apreciar, el retardo de energía de cada operador es diferente. El de fase mínima concentra su energía más cerca del origen que el resto, en tanto que el de fase máxima lo hace hacia el otro extremo. La energía parcial de cada uno de estos cuatro operadores se muestra en la Figura 1.2.14b, donde queda claramente demostrado cómo la energía del operador de mínima fase arriba antes que ningún otro operador de la misma longitud e igual espectro de amplitud. Este resultado queda demostrado con el teorema que veremos a continuación.

El teorema

Supongamos que tenemos dos operadores $h_1(t)$ y $h_2(t)$ que son idénticos excepto por la posición de uno de sus ceros: $h_1(t)$ tiene ese cero fuera del círculo unitario, y $h_2(t)$ lo tiene dentro. En el dominio z , podemos escribir

$$H_1(z) = (a + bz)H(z) \quad \text{y} \quad H_2(z) = (b + az)H(z), \tag{1.72}$$

donde $|a| > |b|$ y $H(z) = h_0 + h_1z + \dots + h_{m-1}z^{m-1}$ es cualquier polinomio de grado $m - 1$. Por simplicidad asumamos que a y b son reales. Notar que los espectros de amplitud de $(a + bz)$ y $(b + az)$ son iguales, como se desprende de la ecuación (1.13).

Calculemos ahora los coeficientes de ambos operadores, cuyos valores están tabulados en las columnas 2 y 3 de la siguiente tabla:

t	$h_1(t)$	$h_2(t)$	$h_1(t)^2 - h_2(t)^2$	$\Delta E(t)$
0	ah_0	bh_0	$(a^2 - b^2)h_0^2$	$(a^2 - b^2)h_0^2$
1	$ah_1 + bh_0$	$bh_1 + ah_0$	$(a^2 - b^2)(h_1^2 - h_0^2)$	$(a^2 - b^2)h_1^2$
\vdots	\vdots	\vdots	\vdots	\vdots
k	$ah_k + bh_{k-1}$	$bh_k + ah_{k-1}$	$(a^2 - b^2)(h_k^2 - h_{k-1}^2)$	$(a^2 - b^2)h_k^2$
\vdots	\vdots	\vdots	\vdots	\vdots
m	bh_{m-1}	ah_{m-1}	$(a^2 - b^2)(-h_{m-1}^2)$	0

La idea es calcular la diferencia de las energías parciales de los dos operadores, que es igual a

$$\Delta E(t) = \sum_{k=0}^t h_1(k)^2 - \sum_{k=0}^t h_2(k)^2 = \sum_{k=0}^t [h_1(k)^2 - h_2(k)^2]. \tag{1.73}$$

La columna 4 de la tabla muestra cada uno de los sumandos del miembro de la derecha de la expresión anterior, y la columna 5 muestra finalmente la suma acumulada de estos sumandos, que naturalmente es igual a la diferencia de las energías parciales de los dos operadores. Claramente, como $|a| > |b|$ todos los valores de la última columna son positivos, por lo que la energía parcial de $h_1(t)$ es siempre mayor o igual que $h_2(t)$.

Finalmente supongamos que repetimos el procedimiento para todos los ceros de $H(z)$ sucesivamente. Es decir, intercambiamos sucesivamente el orden de los coeficientes de cada uno de los dipolos de $H(z)$, y cada vez que hacemos el intercambio para un dipolo determinado, podremos verificar que

la energía parcial será menor para el operador cuyo dipolo en cuestión sea de mínima fase, como lo acabamos de demostrar para uno de ellos. Eventualmente, este intercambio de coeficientes nos llevará a tener un operador con todos los ceros fuera del círculo unitario (fase mínima). Este operador, que comparte el mismo espectro de amplitud con todos los otros $2^{m+1} - 1$ operadores, exhibirá la mayor energía parcial para cada tiempo. Como consecuencia, la energía del operador de mínima fase llegará antes que la energía de cualquier otro operador de la misma longitud y con el mismo espectro de amplitud. El retardo de energía es menor para el operador de mínima fase que para uno de fase mixta o máxima.

1.3. Diseño de filtros digitales

Existen varias estrategias alternativas para el diseño de filtros digitales. La mayoría busca reproducir, de la manera más precisa posible con un número predeterminado de coeficientes, cierta respuesta en frecuencia (espectro) deseada. Ya vimos cómo diseñar ciertos filtros a partir de la ubicación estratégica de ceros y polos.

Pero esta técnica es bastante limitada pues no nos permite obtener filtros muy precisos a partir de una respuesta en frecuencia deseada más general. La mayor dificultad reside en que, cuando se tienen varios polos y varios ceros, los mismos interactúan entre sí de una forma bastante impredecible, limitando su aplicación a aquellos casos en los que se trata de diseñar filtros con unos pocos ceros y polos (por ejemplo el filtro ranura visto).

En la mayor parte de las aplicaciones la forma más usual de obtener la respuesta impulsiva de un filtro digital es a través del cálculo de la transformada de Fourier inversa de cierta respuesta en frecuencia deseada. Lamentablemente en general no es posible obtener un filtro de longitud finita que se ajuste en forma exacta al espectro deseado, siendo en este punto donde entran en juego diversas estrategias que buscarán obtener un filtro lo más aproximado posible a lo que se pretende. La longitud del filtro es, entonces, uno de los elementos más importantes a tener en cuenta. Por razones de practicidad, cuanto más corta es la respuesta impulsiva del filtro, mejor. Pero un filtro muy corto puede producir efectos indeseados debido al truncamiento prematuro de los lóbulos laterales (ver Figura 1.4), en tanto que un filtro muy largo, si bien se aproxima más a la respuesta en frecuencias deseada, presenta otras desventajas (mayor tiempo de CPU, salida más larga, que al truncarla pierde sus características deseadas, etc.). El uso de ventanas apropiadas para truncar la respuesta impulsiva convenientemente es una técnica muy usual. Otra técnica consiste en modificar iterativamente los

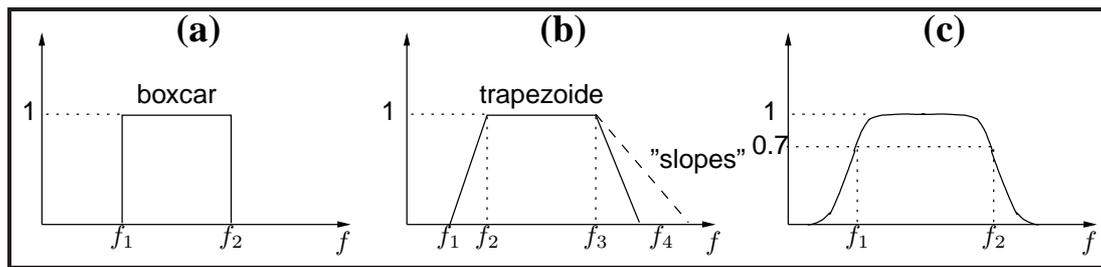


Figura 1.25: (a) Filtro pasa-banda ideal. (b) Filtro pasa-banda realista. (c) Filtro pasa-banda real.

coeficientes del filtro obtenidos luego de la transformada de Fourier inversa hasta satisfacer cierto criterio. Un ejemplo de estos filtros son los filtros *equi-ripple* o *minimax*, que se obtienen con el algoritmo de Parks-McClellan (ver por ejemplo el Karl). Estos filtros se obtienen tras minimizar la máxima amplitud del ripple en la banda de paso, de rechazo, o las dos simultáneamente. Como resultado todas las oscilaciones exhiben la misma amplitud para un largo de filtro determinado. Otro método consiste en digitalizar funciones racionales de la frecuencia (filtros de Chebyshev, Butterworth, elípticos, etc.). Antes de describir con algún detalle algunas de estas técnicas de aproximación, veamos cómo es un filtro ideal.

1.3.1. Filtro ideal

Un filtro “ideal” es aquel cuya respuesta en frecuencias se ajusta exactamente a lo que se desea filtrar. En la vida “real” no existe tal filtro, y uno se tiene que conformar con una aproximación al filtro ideal. El grado de aproximación dependerá de la técnica que se utilice para diseñar el filtro, de su longitud, de las limitaciones impuestas por su aplicación, etc. Por ejemplo, la Figura 1.25 muestra un filtro pasa-banda ideal (no-realizable) y otras opciones realizables en las cuales se resignaron las pendientes abruptas del filtro ideal. El problema con el filtro en (a) radica en que los lóbulos laterales de su respuesta impulsiva tienden a cero muy lentamente, y un corte abrupto de los mismos produce el efecto indeseado conocido como “ripple”. En cambio, el filtro en (b) tiene lóbulos laterales que decaen más rápidamente, minimizándose este efecto. Pero como tampoco se permiten “queiebres” tan marcados, las transiciones en un filtro real son siempre suaves, como el filtro en (c). El uso de pendientes más moderadas permite entonces contar con filtros más cortos. Las **rampas** o pendientes (*banda de transición*) se especifican con dos frecuencias cada una, y se usan unidades como el **decibel (dB)**⁹ para medir la caída o **atenuación** en amplitud (o energía) para dos frecuencias:

⁹El término “decibel” se utiliza en honor a Alexander Graham Bell (1847–1922), inventor estadounidense. La definición original de decibel se expresa en términos del cociente de energías (que representan el cuadrado de la amplitud). La razón por la cual se utiliza el logaritmo y no otra función para medir cambios de energía es que el oído humano reacciona

$$D = 20 \log_{10} \frac{\text{amplitud 2}}{\text{amplitud 1}} = 10 \log_{10} \frac{(\text{amplitud 2})^2}{(\text{amplitud 1})^2} \quad (1.74)$$

Un valor típico es -3dB. Para este valor, la amplitud decae a un 70 %:

$$\text{amplitud 2} = 10^{-\frac{3}{20}} \times \text{amplitud 1} \simeq 0.70 \text{ amplitud 1}, \quad (1.75)$$

lo que puede considerarse razonable para muchas aplicaciones, en tanto que la energía decae a la mitad. Para -20 dB, la amplitud decae a un 10 % y la energía a un 1 % solamente. En general, las frecuencias de corte del filtro se definen en estos puntos (ver Figura 1.25c).

1.3.2. Filtros mediante el uso ventanas

Un filtro pasa-banda puede diseñarse a partir de dos filtros pasa-bajos, como muestra la Figura 1.26. En este caso, si restamos los espectros de amplitud de los pasa-bajos con diferentes frecuencias de corte, resulta un filtro pasa-banda como se ve en la figura. La figura anterior muestra un filtro pasa-banda ideal. En la práctica, el mayor problema reside, como mencionamos, en el truncamiento y el consiguiente ripple. La solución al problema del ripple consiste en determinar dónde y cómo truncar para controlar los efectos del mismo, y, a la vez, cómo obtener un filtro con una respuesta próxima a la ideal. Para ello dos o tres décadas atrás en la literatura surgió toda una batería de métodos y ventanas especialmente diseñadas para este fin (“window carpentry”). Cada ventana tiene sus propias características en el sentido que producen filtros con determinadas propiedades en la banda de paso, de rechazo y/o de transición. Entre las ventanas más conocidas podemos mencionar: *triangular* (*Bartlett*), *Hamming*, *hanning*, *Parzen*, *Daniell*, etc., siendo la de Hamming una de las más utilizadas.

La ventana de Hamming se construye con la siguiente expresión, donde L_v , $L_v \leq L_h$, es la longitud de la ventana:

$$h_t = \begin{cases} 0.54 + 0.46 \cos(\pi t/L_v) & |t| \leq L_v, \\ 0 & |t| > L_v. \end{cases} \quad (1.76)$$

La Figura 1.27 muestra el efecto que tienen sobre el espectro de amplitud diversas ventanas. En la primera fila se ve un filtro pasa-banda ideal. La aplicación de la ventana en el dominio del tiempo es equivalente a convolucionar en el de las frecuencias. Claramente, un simple truncamiento con una ventana rectangular (filas 2 y 3) no produce los mejores resultados, ya que los lóbulos de la diferencia de esta manera.

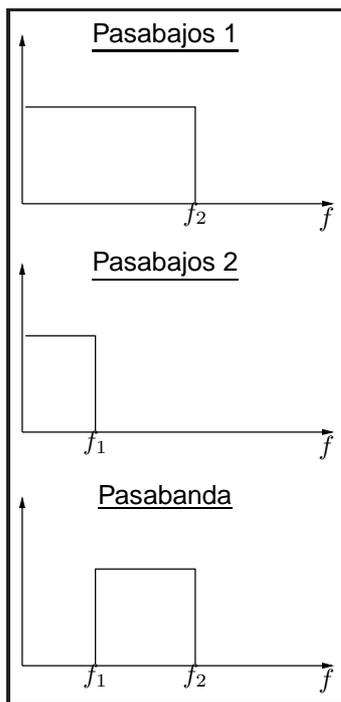


Figura 1.26: Construcción de un filtro pasa-banda a partir de dos filtros pasa-bajos.

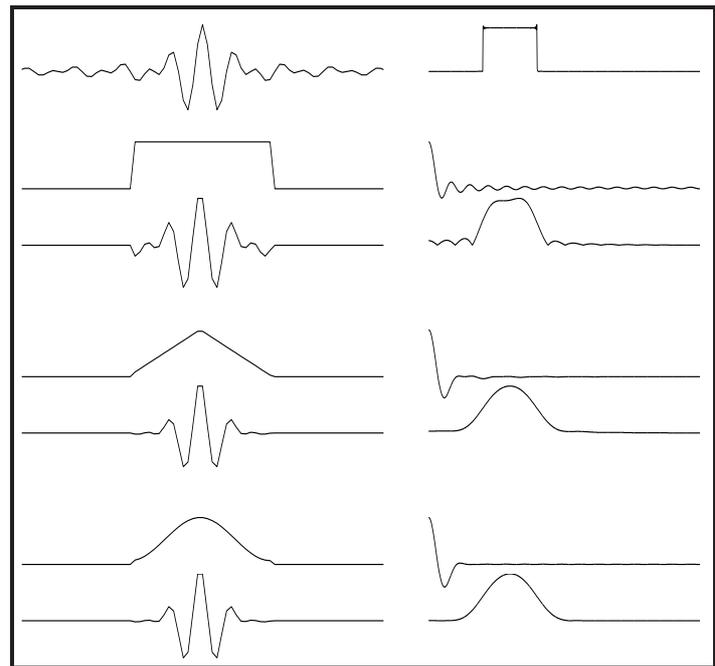


Figura 1.27: La columna de la izquierda representa el dominio temporal, con el cero centrado. La columna de la derecha es su equivalente en el dominio de las frecuencias (frecuencias positivas solamente).

de dos senos cardinales (transformada de Fourier de la “boxcar”) decaen lentamente. El caso de la ventana triangular (filas 4 y 5) funciona bastante mejor para el ripple ya que el filtro resultante exhibe lóbulos laterales menores. Como contrapartida, genera pendientes más suaves porque su lóbulo central es más ancho. Aunque en la figura no se aprecia bien, existe todavía cierto ripple que no está presente cuando se utiliza la ventana de Hamming (filas 6 y 7), que es una función muy suave, y produce resultados mucho más convenientes.

En resumen, cualquiera sea la ventana utilizada, en general el filtro original resulta “ensanchado” y con los bordes “suavizados” y/o afectados de ripple en mayor o menor medida. Si bien el diseño de filtros mediante el uso de ventanas es muy práctico, simple y efectivo para muchas aplicaciones, presenta algunas desventajas. Por ejemplo, son poco flexibles si se pretende especificar a priori y con precisión las frecuencias de corte, ya que la convolución en el dominio de las frecuencias con el espectro de la ventana correspondiente produce cambios relativamente impredecibles en estas frecuencias. Tampoco es posible predeterminedar un largo de filtro fijo para determinada respuesta en frecuencia. Por supuesto, cuanto más largo el filtro (y la ventana), mejores resultados, pero a expensas de mayor longitud de la salida ante una entrada dada.

Estos conceptos son válidos también para el caso de una ondícula de fase cero de largo fijo. La Figura 1.28 muestra una serie de ondículas u operadores de fase cero cuyo espectro de amplitud se asemeja a un filtro pasa-banda. Claramente, a mayor pendiente, mayor ripple tanto en el espectro como en el tiempo (lóbulos laterales grandes). El ripple en el espectro proviene del truncamiento abrupto de la ondícula. Si en cambio se suavizan los bordes resignando pendiente (ver panel (c)), el ripple es mucho menor ya que la ondícula resultante es más corta y no se ve afectada por truncamientos abruptos.

Otro ejemplo es el que se muestra en la Figura 1.29. Aquí se muestra un mismo filtro pasa-banda con forma trapezoidal donde se ha aplicado una ventana rectangular de ancho progresivo. La idea es ver el efecto del truncamiento sobre la forma del espectro. Si el truncamiento es severo, la forma original se pierde significativamente.

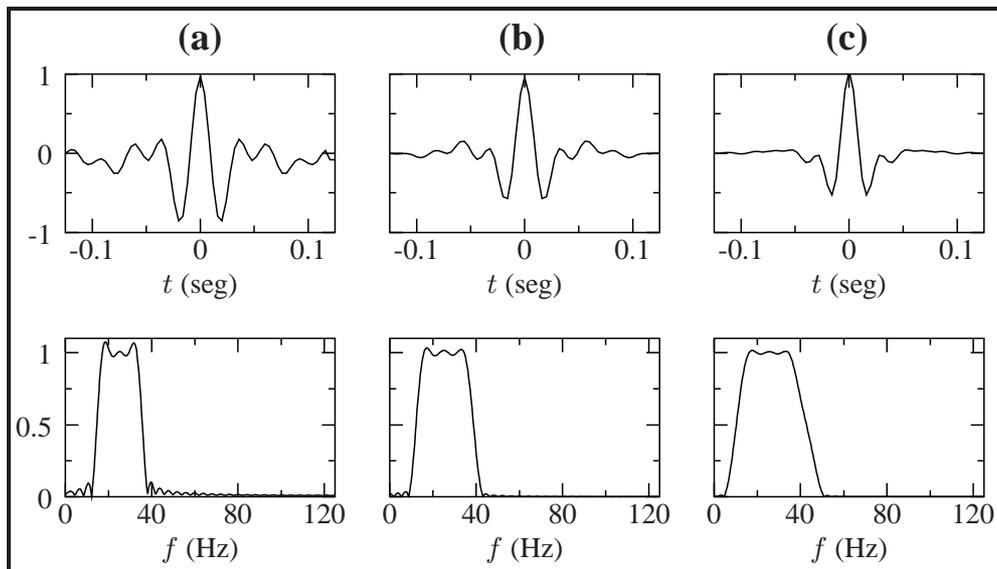


Figura 1.28: Tres operadores de fase cero de largo fijo (filtro pasa-banda) y sus respectivos espectros de amplitud. Se seleccionaron tres tipos de bandas de transición: (a) muy abruptas, (b) moderadas, y (c) muy suaves. Notar el efecto en el ripple y en la longitud del operador.

1.3.3. Filtros mediante digitalización de funciones racionales de ω

Estos filtros derivan de filtros RC analógicos cuya respuesta en frecuencia es bien conocida y ajustable de acuerdo a la selección apropiada de sus componentes. Entre los más conocidos podemos mencionar los filtros de Butterworth, de Chebyshev (tipo I y tipo II) y de Jacobi (elípticos).

Resulta que los espectros de potencia de estos filtros se pueden expresar como funciones racionales de ω , lo que permite, en principio, su factorización en términos de sus ceros y polos. En base a este

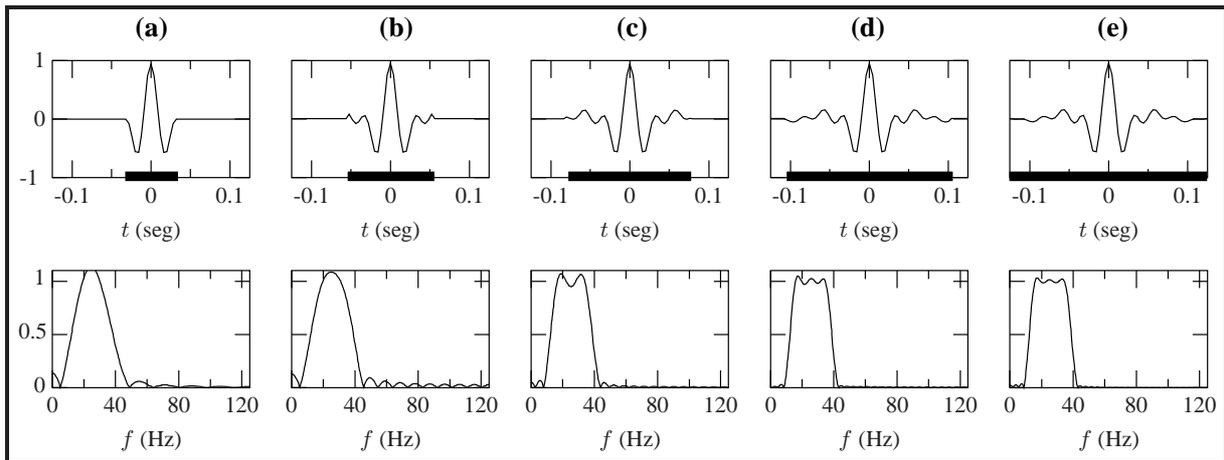


Figura 1.29: Cinco operadores de fase cero (filtro pasa-banda trapezoidal) y sus respectivos espectros de amplitud. La línea gruesa en los paneles de la primera fila indica la longitud del operador: (a) 0.05 s, (b) 0.1 s, (c) 0.15 s, (d) 0.2 s, y (e) 0.25 s. Un truncamiento muy severo deforma la respuesta deseada e introduce un ripple excesivo.

principio se “construye” el operador digital cuyo espectro de potencia se ajusta al sistema analógico¹⁰. Es importante destacar que por razones de estabilidad, es necesario descartar aquellos polos que yacen dentro del círculo unitario. Lo que se obtiene es un operador estable cuyo espectro de potencia se aproxima al espectro de potencia del sistema analógico, y cuyo espectro de fase está dictado por cuestiones de estabilidad. No brindaremos mayores detalles pues el procedimiento es bastante engorroso, excepto para filtros muy simples. Simplemente describiremos las principales características de algunos de estos filtros.

Filtros de Butterworth

El espectro de potencia (o magnitud cuadrática) de un filtro (pasa-bajos) de Butterworth se puede expresar como

$$|H(\omega)|^2 = \frac{1}{1 + (\omega/\omega_c)^{2N}}, \tag{1.77}$$

donde ω_c es la *frecuencia de corte* y N es el orden del filtro. Esta función tiene $2N$ polos, en tanto que $|H(\omega)|$ tiene N polos. Este filtro se usa como anti-alias y se caracteriza por tener una banda de paso muy plana (“maximally flat filter”). Esto es así porque las primeras $2N - 1$ derivadas del

¹⁰En la práctica hay que transformar la función racional en ω por una función racional en z . Esto se logra utilizando la poderosa transformada bilineal que vincula z con ω en forma aproximada pero precisa. Esta transformada bilineal se obtiene luego de expandir en serie las variables $z = \exp(-i\omega)$ y $i\omega = -\ln z$. Luego, las expresiones resultantes se truncan y se obtienen las relaciones bilineales $z = (1 - i\omega/2)/(1 + i\omega/2)$ y $i\omega = 2(1 - z)/(1 + z)$.

espectro de potencia en $\omega = 0$ son nulas. La Figura 1.30a muestra el espectro de amplitud de este filtro para $\omega_c = \pi/4$ y $N = 2, 4$ y 8 . Cuanto más alto es el orden del filtro, más abrupta es la banda de transición. Otra característica es que tanto la banda de paso como la de rechazo son *monotónicas* y que $|H(\omega_c)| = 1/\sqrt{2}$, independientemente del orden del filtro.

Filtros de Chebyshev (tipo I)

Como los filtros de Butterworth, los filtros de Chebyshev también son filtros que contienen solamente polos. Un filtro pasa-bajos de Chebyshev¹¹ (tipo I) de orden N y frecuencia de corte ω_c tiene por espectro de potencia a la expresión

$$|H(\omega)|^2 = \frac{1}{1 + \epsilon^2 T_N^2(\omega/\omega_c)}, \quad (1.78)$$

donde ϵ es un parámetro (tiene que ver con el ripple) y $T_N(x)$ es el polinomio de Chebyshev¹² de orden N .

Estos filtros se obtienen a partir de minimizar el máximo de la amplitud del ripple en la banda de paso (criterio *minimax*). El espectro de amplitud resultante presenta como consecuencia oscilaciones (ripple) de amplitud constante en la banda de paso, y luego decae monotónicamente en la banda de rechazo, como se muestra en la Figura 1.30b. El número de oscilaciones está controlado por el orden del filtro, en tanto que la amplitud de las mismas depende de ϵ exclusivamente. El ancho de la banda

¹¹Tchebyscheff, matemático ruso (1821–1894).

¹²El polinomio de Chebyshev de orden N se define como

$$T_N(x) = \begin{cases} \cos(N \arccos x) & |x| \leq 1 \\ \cosh(N \operatorname{arccosh} x) & |x| > 1. \end{cases}$$

Como en la banda de paso $|\omega/\omega_c| \leq 1$, entonces $|x| \leq 1$. Además $0 \leq T_N^2 \leq 1$. En este caso se tienen los siguientes polinomios para órdenes $N = 0, 1$ y 2 :

$$\begin{aligned} N = 0 : \quad T_0(x) &= \cos(0) &= 1 \\ N = 1 : \quad T_1(x) &= \cos[\arccos(x)] &= x \\ N = 2 : \quad T_2(x) &= \cos[2 \arccos(x)] &= 2x^2 - 1. \end{aligned}$$

Existe fórmulas recursivas para obtener los polinomios de órdenes superiores, como ser

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x).$$

Para $|\omega/\omega_c| > 1$, o sea $|x| > 1$, el arco-coseno es imaginario y se comporta como un coseno hiperbólico que decrece monotónicamente.

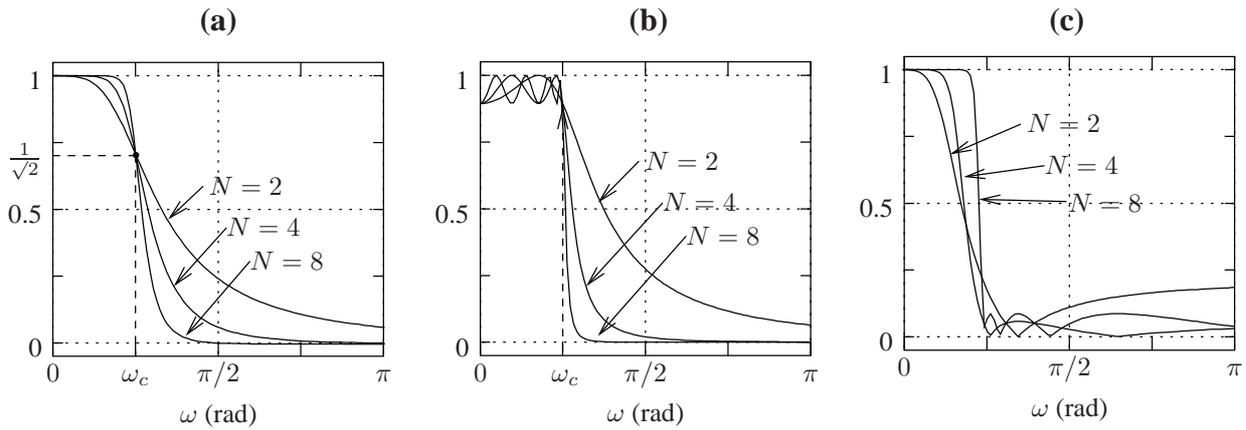


Figura 1.30: Espectros de amplitud de (a) filtro de Butterworth, (b) filtro de Chebyshev de Tipo I ($\epsilon = 0.5$), y (c) filtro de Chebyshev de Tipo II ($\epsilon = 20$). En todos los casos $\omega_c = \pi/4$, $N = 2, 4$ y 8 .

de transición depende de N y de ϵ . Para N fijo, cuanto más grande es ϵ , más abrupta es la banda de transición, pero la amplitud de las oscilaciones crece en igual medida.

Filtros de Chebyshev (Tipo II)

Los filtros de Chebyshev de Tipo II (ver Figura 1.30c), a diferencia de los del Tipo I, presentan oscilaciones de amplitud constante en la banda de rechazo, satisfaciendo entonces el criterio *minimax* en esa región. La expresión del espectro de potencia es

$$|H(\omega)|^2 = \frac{1}{1 + \epsilon^2 [T_N(\omega_c)/T_N(\omega_c/\omega)]^2}. \tag{1.79}$$

Filtros de Jacobi (elípticos)

Estos filtros también recurren al criterio minimax, pero lo hacen tanto para determinar la banda de paso como la de rechazo, obteniéndose un filtro cuyo espectro de potencia es

$$|H(\omega)|^2 = \frac{1}{1 + \epsilon^2 R_N^2(\omega, L)}, \tag{1.80}$$

donde $R_N(\omega, L)$ es la función elíptica de Jacobi con factor L que controla el ripple. Los filtros de Jacobi presentan la banda de transición más abrupta posible para un filtro de orden fijo N .

Bibliografía

- [Claerbout, 1992] Claerbout, J. F. (1992). *Earth Sounding Analysis: Processing versus Inversion*. Blackwell Scientific Publications. [Electronic version available on the Internet].
- [Karl, 1989] Karl, J. H. (1989). *An Introduction to Digital Signal Processing*. Academic Press, Inc.
- [Oppenheim and Schaffer, 1975] Oppenheim, A. and Schaffer, R. (1975). *Digital Signal Processing*. Prentice-Hall, Englewood Cliffs, New Jersey.
- [Rabiner and Gold, 1975] Rabiner, L. R. and Gold, B. (1975). *Theory and Applications of Digital Signal Processing*. Prentice-Hall, Englewood Cliffs, New Jersey.
- [Robinson and Treitel, 1980] Robinson, E. and Treitel, S. (1980). *Geophysical Signal Analysis*. Prentice-Hall, Inc., Englewood Cliffs, NJ 07632.
- [Robinson and Treitel, 2002] Robinson, E. A. and Treitel, S. (2002). *Geophysical Signal Analysis*. SEG.
- [Yilmaz, 1987] Yilmaz, O. (1987). *Seismic Data Processing*, volume 2 of *Investigations in Geophysics*. SEG.