

Estudiando el libro “Discrete-Time Signal Processing”
de Alan V. Oppenheim y Ronald W. Schafer

Ernesto López

31 de mayo de 2024

Índice general

1. Señales y sistemas en tiempo discreto	2
1.1. Señales en tiempo discreto	2
1.2. Sistemas en tiempo discreto	4
1.2.1. Sistemas sin memoria	4
1.2.2. Sistemas lineales	5
1.2.3. Sistemas invariantes en el tiempo	5
1.2.4. Sistemas causales	5
1.2.5. Sistemas estables	6
1.3. Sistemas lineales e invariantes en el tiempo	6
1.4. Propiedades de los sistemas lineales e invariantes en el tiempo	6
1.4.1. Estabilidad de los sistemas LTI	7
1.4.2. Causalidad de los sistemas LTI	8
1.5. Ecuaciones en diferencias lineales con coeficientes constantes	8
1.6. Representación de señales y sistemas en el dominio de la frecuencia	15
1.6.1. Funciones propias de los sistemas LTI	15
1.7. Representación de secuencias mediante transformadas de Fourier	21
1.8. Propiedades de simetría de la transformada de Fourier	23
1.9. Teoremas de la transformada de Fourier	24
1.9.1. Linealidad de la transformada de Fourier	24
1.9.2. Desplazamiento temporal y desplazamiento en frecuencia	24
1.9.3. Inversión temporal	25
1.9.4. Diferenciación en frecuencia	25
1.9.5. Teorema de la convolución	25
1.9.6. Teorema de enventanado o modulación	26
1.9.7. Teorema de Parseval	26
1.10. Señales aleatorias en tiempo discreto	28
2. La transformada z	31
2.1. La transformada z	31
2.2. Propiedades de la ROC de la transformada z	33
2.3. La transformada z inversa	34
2.3.1. Evaluación directa de la integral de contorno	35
2.3.2. Método de inspección	37
2.3.3. Método de fracciones simples	37
2.3.4. Expansión en series de potencias	40
2.4. Propiedades de la transformada z	42
2.4.1. Linealidad	42
2.4.2. Desplazamiento en el tiempo	42
2.4.3. Multiplicación por una secuencia exponencial	43
2.4.4. Diferenciación de $X(z)$	43
2.4.5. Conjugación de una secuencia compleja	43
2.4.6. Inversión temporal	43
2.4.7. Convolución de secuencias	43
2.4.8. Teorema del valor inicial	44

3. Análisis en el dominio transformado de sistemas lineales e invariantes en el tiempo	45
3.1. Introducción	45
3.2. La respuesta en frecuencia de sistemas LTI	45
3.2.1. Respuesta en fase y retardo de grupo	45
3.3. Sistemas caracterizados por ecuaciones en diferencias lineales con coeficientes constantes .	57
3.3.1. Estabilidad y causalidad	57
3.3.2. Sistemas inversos	58
3.4. Respuesta en frecuencia de funciones de sistema racionales	59
3.4.1. Respuesta en frecuencia de sistemas de primer orden	61
3.4.2. Ejemplos con múltiples ceros y polos	63
3.5. Relación entre la magnitud y la fase	65
3.6. Sistemas pasa todo	67
3.7. Sistemas de fase mínima	70
3.7.1. Descomposición en fase mínima y pasa todo	70
3.7.2. Compensación de la respuesta en frecuencia de sistemas de fase no mínima	72
3.7.3. Propiedades de los sistemas de fase mínima	73
3.8. Sistemas lineales con fase lineal generalizada	79
3.8.1. Sistemas con fase lineal	79
3.8.2. Fase lineal generalizada	81
3.8.3. Sistemas causales con fase lineal generalizada	84
3.8.4. Relación entre sistemas FIR de fase lineal y sistemas de fase mínima	87
4. Estructuras de sistemas en tiempo discreto	91
4.1. Representación en diagramas de bloques de ecuaciones en diferencias con coeficientes lineales	91
4.2. Representación en grafos de flujo de ecuaciones en diferencias con coeficientes lineales . .	94
4.2.1. Representación en grafos de flujo de sistemas en tiempo discreto	94
4.2.2. Representación matricial de redes digitales	98
4.2.3. Teorema de Tellegen para filtros digitales	100
4.3. Estructuras básicas para sistemas IIR	104
4.3.1. Formas directas	104
4.3.2. Forma en cascada	105
4.3.3. Forma en paralelo	106
4.3.4. Estructuras de muestreo en frecuencia	108
4.4. Formas transpuestas	116
4.5. Estructuras básicas para sistemas FIR	118
4.5.1. Forma directa	118
4.5.2. Forma en cascada	119
4.5.3. Estructuras para sistemas FIR de fase lineal	119
4.6. Filtros lattice	122
4.6.1. Filtros lattice FIR	123
4.6.2. Estructura lattice todo polos	129
4.6.3. Generalización de los sistemas lattice	131

Prefacio

El presente documento consiste en apuntes sobre procesamiento de señales discretas. Se seguirá fielmente el libro [1]. Solo se incluye un resumen de los conceptos básicos y se profundizará en los tópicos mas avanzados de procesamiento de señales incluidos en el libro.

Capítulo 1

Señales y sistemas en tiempo discreto

1.1. Señales en tiempo discreto

En la teoría de señales y sistemas en tiempo discreto, hay secuencias son de particular importancia. Algunas son las siguientes:

- *Impulso unidad:*

$$\delta[n] = \begin{cases} 0, & n \neq 0, \\ 1, & n = 0. \end{cases} \quad (1.1)$$

Un aspecto importante del impulso unidad es que cualquier secuencia puede expresarse como la suma de impulsos desplazados y escalados como

$$x[n] = \sum_{k=-\infty}^{\infty} x[k]\delta[n-k]. \quad (1.2)$$

- *Escalón unidad:*

$$u[n] = \begin{cases} 1, & n \geq 0, \\ 0, & n < 0. \end{cases} \quad (1.3)$$

El escalón unidad está relacionado con el impulso unidad por

$$u[n] = \sum_{k=-\infty}^n \delta[k],$$

es decir, el escalón unidad en el índice (tiempo) n es la suma acumulada de todas las muestras del impulso unidad hasta el tiempo n . Una representación alternativa del escalón unidad en términos de impulsos es

$$u[n] = \sum_{k=0}^{\infty} \delta[n-k],$$

como en la ecuación 1.2. Por otro lado, el impulso puede expresarse como la *primera diferencia regresiva* del escalón unidad, es decir,

$$\delta[n] = u[n] - u[n-1].$$

- *Exponencial:*

$$u[n] = A\alpha^n, \quad (1.4)$$

donde en el caso mas general, α y A son números complejos. Específicamente, si $\alpha = |\alpha|e^{j\omega_0}$ y $A = |A|e^{j\phi}$,

$$\begin{aligned} x[n] &= |A||\alpha|^n e^{j(\omega_0 n + \phi)} \\ &= |A||\alpha|^n \cos(\omega_0 n + \phi) + j|A||\alpha|^n \sin(\omega_0 n + \phi). \end{aligned}$$

Si $|\alpha| = 1$, la secuencia toma la forma

$$\begin{aligned} x[n] &= |A|e^{j(\omega_0 n + \phi)} \\ &= |A|\cos(\omega_0 n + \phi) + j|A|\sin(\omega_0 n + \phi). \end{aligned}$$

En esta secuencia, la parte real e imaginaria oscilan sinusoidalmente con n . Por analogía con el caso en tiempo continuo, ω_0 es la *frecuencia* y ϕ es la *fase*. Sin embargo, como n es adimensionado, las unidades de ω_0 son radianes. Para mantener un analogía mas cercana al caso continuo, puede especificarse que las unidades de ω_0 son radianes por muestra y las unidades de n son muestras.

El hecho de que n sea un número entero conduce a diferencias importantes de las propiedades de las señales exponenciales complejas y sinusoidales en tiempo continuo y en tiempo discreto. A diferencia del caso en tiempo continuo, las secuencias exponenciales complejas y sinusoidales con frecuencias $\omega_0 + 2\pi r$, con r entero, son indistinguibles entre si, ya que

$$e^{j[(\omega_0 + 2\pi r)n + \phi]} = e^{j(\omega_0 n + \phi)}.$$

Por lo tanto, al estudiar secuencias exponenciales complejas de la forma $x[n] = Ae^{j(\omega_0 n + \phi)}$ o señales sinusoidales reales de la forma $x[n] = A\cos(\omega_0 n + \phi)$ solo se necesita considerar frecuencias en un intervalo de 2π . Típicamente se eligen los intervalos $-\pi < \omega_0 \leq \pi$ o $0 \leq \omega_0 < 2\pi$.

A diferencia del caso en tiempo continuo, una secuencia exponencial compleja o sinusoidal no necesariamente es periódica. Las secuencias periódicas son aquellas que cumplen que

$$x[n] = x[n + N], \quad \text{para todo } n,$$

donde el período N es entero. Si se prueba esta condición para una secuencia sinusoidal,

$$A\cos(\omega_0 n + \phi) = A\cos(\omega_0 n + \omega_0 N + \phi),$$

se requiere que

$$\omega_0 N = 2\pi k, \tag{1.5}$$

con k entero. Por lo tanto, las secuencias exponenciales complejas y sinusoidales no necesariamente son periódicas con período $2\pi/\omega_0$, y dependiendo del valor de ω_0 , pueden no ser periódicas en absoluto.

Si se combina la condición de la ecuación 1.5 con la observación previa de que las frecuencias ω_0 y $\omega_0 + 2\pi r$ son indistinguibles, se deduce que una secuencia de período N puede ser de N frecuencias distinguibles. El conjunto de frecuencias distintas es

$$\omega_k = \frac{2\pi k}{N}, \quad k = 0, 1, \dots, N-1.$$

Vinculado a la discusión precedente es el hecho de que la interpretación de bajas y altas frecuencias es distinto para secuencias exponenciales complejas y sinusoidales. Para la secuencia sinusoidal $x[n] = A\cos(\omega_0 n + \phi)$, cuando la frecuencia ω_0 crece desde $\omega_0 = 0$ a $\omega_0 = \pi$, $x[n]$ oscila progresivamente mas rápidamente, pero cuando la frecuencia ω_0 crece desde $\omega_0 = \pi$ a $\omega_0 = 2\pi$, las oscilaciones se hacen mas lentas. Esto es porque las frecuencias ω_0 y $2\pi - \omega_0$ tienen la misma velocidad de oscilación, ya que, por ejemplo

$$\sin[(2\pi - \omega_0)n] = \sin(2\pi n - \omega_0 n) = \sin(-\omega_0 n) = -\sin(\omega_0 n).$$

Como consecuencia, para secuencias sinusoidales y exponenciales complejas, las frecuencias cercanas a $\omega_0 = 2\pi k$, con k entero, son referidas como bajas frecuencias, y las frecuencias en la vecindad de $\pi + 2\pi k$, con k entero, son referidas como altas frecuencias.

1.2. Sistemas en tiempo discreto

Un sistema en tiempo discreto se define matemáticamente como una transformación que se aplica sobre una secuencia de entrada $x[n]$ para obtener una secuencia de salida $y[n]$. Se denota como

$$y[n] = T\{x[n]\}.$$

A continuación se dan ejemplos de algunos sistemas básicos que serán útiles para ilustrar las propiedades de los sistemas en tiempo discreto mas adelante.

- *Retardo ideal:*

$$y[n] = x[n - n_d], \quad -\infty < n < \infty, \quad (1.6)$$

donde n_d es un entero positivo. El retardo ideal desplaza la entrada hacia la derecha n_d muestras. En el caso en que n_d es negativo, desplaza la señal de entrada a la izquierda n_d muestras, correspondiendo a un adelanto temporal.

- *Media móvil:*

$$y[n] = \frac{1}{M_1 + M_2 + 1} \sum_{k=-M_1}^{M_2} x[n - k]. \quad (1.7)$$

En este sistema, la n -ésima muestra de la salida es el promedio de $M_1 + M_2 + 1$ muestras de la entrada en torno a la n -ésima muestra.

- *Acumulador:*

$$y[n] = \sum_{k=-\infty}^n x[k]. \quad (1.8)$$

La salida en el tiempo n es la suma de la muestra presente y todas las muestras pasadas de la entrada.

- *Compresor:*

$$y[n] = x[Mn], \quad -\infty < n < \infty, \quad (1.9)$$

donde M es un entero positivo. Este sistema descarta $M - 1$ por cada M muestras de la entrada.

- *Diferencia progresiva y regresiva:* el sistema diferencia progresiva esta dado por la ecuación

$$y[n] = x[n + 1] - x[n] \quad (1.10)$$

y el sistema diferencia regresiva esta dado por la ecuación

$$y[n] = x[n] - x[n - 1]. \quad (1.11)$$

Imponiendo restricciones sobre las propiedades de la transformación $T\{\cdot\}$ se definen clases de sistemas. En las siguientes secciones se brinda una clasificación de sistemas con restricciones y propiedades especialmente importantes.

1.2.1. Sistemas sin memoria

Un sistema se dice sin memoria si la salida $y[n]$ en cada instante n depende solo de la entrada $x[n]$ en el mismo instante n .

Los sistemas retardo ideal y media móvil no son sin memoria a menos que $n_d = 0$ y $M_1 = M_2 = 0$ respectivamente. El resto de los sistemas definidos al comienzo de la sección también son sin memoria. Un ejemplo de un sistema sin memoria es

$$y[n] = (x[n])^2. \quad (1.12)$$

1.2.2. Sistemas lineales

La clase de sistemas lineales está definida por el *principio de superposición*. Si $y_1[n]$ e $y_2[n]$ son las respuestas de un sistema cuando las entradas son $x_1[n]$ y $x_2[n]$ respectivamente, el sistema es lineal si y solo si

$$T\{ax_1[n] + bx_2[n]\} = aT\{x_1[n]\} + bT\{x_2[n]\}, \quad (1.13)$$

donde a y b son constantes arbitrarias. La ecuación 1.13 es el principio de superposición.

La prueba de que un sistema es lineal requiere una demostración general sin realizar hipótesis específicas sobre las señales de entrada. Por el contrario, la prueba de que un sistema no es lineal solo requiere un contraejemplo.

A partir del principio de superposición puede mostrarse que todos los sistemas presentados al comienzo de esta sección son lineales. El sistema definido por la ecuación 1.12 no es lineal.

1.2.3. Sistemas invariantes en el tiempo

Un sistema invariante en el tiempo es un sistema para el cual un desplazamiento temporal de la secuencia de entrada produce el mismo desplazamiento temporal en la secuencia de salida. Específicamente, supóngase que un sistema transforma la secuencia de entrada $x[n]$ en la secuencia de salida $y[n]$. El sistema es invariante en el tiempo si, para todo n_0 , la secuencia de entrada $x_1[n] = x[n - n_0]$ produce la secuencia de salida $y_1[n] = y[n - n_0]$.

Al igual que con los sistemas lineales, la prueba de la invarianza temporal requiere una demostración general sin imposición de hipótesis sobre la secuencia de entrada, y alcanza un contraejemplo para probar lo contrario.

El sistema compresor, dado por la ecuación 1.9, no es invariante en el tiempo. Si lo fuera, dado que la entrada $x[n]$ produce la salida $y[n] = x[Mn]$, la entrada $x_1[n] = x[n - n_0]$ debería producir la salida

$$y[n - n_0] = x[M(n - n_0)] = x[Mn - Mn_0].$$

Sin embargo, la entrada $x_1[n]$ produce la salida

$$y_1[n] = x_1[Mn] = x[Mn - n_0],$$

que es distinta a $y[n - n_0]$. La no invarianza también se podría mostrar con un contraejemplo: sea $M = 2$, $x[n] = \delta[n]$ y $n_0 = 1$, es decir, $x_1[n] = x[n - 1] = \delta[n - 1]$. La salida es $y[n] = \delta[n]$, pero $y_1[n] = 0$, y por lo tanto, como $y[n - 1] \neq y_1[n]$, el sistema no es invariante en el tiempo. Puede probarse que el resto de los sistemas al comienzo de la sección son invariantes en el tiempo.

1.2.4. Sistemas causales

Un sistema es causal si para cada n_0 , la secuencia de salida en el instante $n = n_0$ depende solo de valores de la secuencia de entrada en $n \leq n_0$. Esto implica que si $x_1[n] = x_2[n]$ para $n \leq n_0$ se cumple que $y_1[n] = y_2[n]$ para $n \leq n_0$. Conceptualmente significa que el sistema es *no anticipativo*.

Para verificar si el sistema retardo ideal es causal, sean las entradas $x_1[n]$ y $x_2[n]$ que cumplen que $x_1[n] = x_2[n]$ para $n \leq n_0$, y sean $y_1[n] = x_1[n - n_d]$ y $y_2[n] = x_2[n - n_d]$ las salidas correspondientes. De esta forma, $y_1[n] = y_2[n]$ si $x_1[n - n_d] = x_2[n - n_d]$, y por hipótesis esto es cierto si

$$n - n_d \leq n_0 \quad \Leftrightarrow \quad n \leq n_0 + n_d. \quad (1.14)$$

Pero para que el sistema sea causal se requiere que $y_1[n] = y_2[n]$ para $n \leq n_0$, lo que en la condición 1.14 ocurre si $n_d \geq 0$. Se concluye que el sistema es causal si $n_d \geq 0$ y no es causal si $n_d < 0$. En general, en la práctica es directo determinar si un sistema es causal inspeccionando si en la ecuación del sistema la salida depende de muestras futuras de la entrada. Así, el sistema media móvil de la ecuación 1.7 es causal si $M_1 \leq 0$ y $M_2 \geq 0$, y en otro caso es no causal. El sistema acumulador y el sistema diferencia regresiva de las ecuaciones 1.8 y 1.11 son causales. El sistema compresor de la ecuación 1.9 no es causal si $M > 1$ y el sistema diferencia progresiva dado por la ecuación 1.10 no es causal.

1.2.5. Sistemas estables

Un sistema es estable en el sentido *entrada acotada, salida acotada* (*BIBO*, Bounded Input, Bounded Output) si y solo si cada entrada acotada produce una salida acotada. La entrada es acotada si existe un número positivo fijo B_x tal que

$$|x[n]| \leq B_x < \infty, \quad \text{para todo } n.$$

La propiedad de estabilidad requiere que para toda entrada acotada exista un número positivo fijo B_y tal que

$$|y[n]| \leq B_y < \infty, \quad \text{para todo } n.$$

Puede probarse que los sistemas retardo ideal, media móvil y diferencia progresiva y regresiva son estables y el sistema acumulador no es estable.

1.3. Sistemas lineales e invariantes en el tiempo

Una clase importante de sistemas son aquellos que son tanto lineales como invariantes en el tiempo (*LTI*, Linear Time Invariant). La importancia radica en que conociendo la respuesta al impulso del sistema, es posible calcular la respuesta a cualquier entrada, y por lo tanto, tienen una gran cantidad de aplicaciones en procesamiento de señales. Específicamente, si $h[n]$ es la respuesta al impulso de un sistema LTI, la salida es

$$y[n] = \sum_{k=-\infty}^{\infty} x[k]h[n-k], \quad \text{para todo } n. \quad (1.15)$$

Esta ecuación se denomina *suma de convolución* y se representa mediante la notación

$$y[n] = x[n] * h[n].$$

1.4. Propiedades de los sistemas lineales e invariantes en el tiempo

Como el comportamiento de todos los sistemas LTI está determinado por la suma de convolución, las propiedades de esta clase de sistemas están establecidas por las propiedades de la convolución.

Algunas propiedades generales de la convolución son las siguientes:

- *Conmutativa:*

$$x[n] * h[n] = h[n] * x[n].$$

Esto implica que un sistema LTI con entrada $x[n]$ y respuesta al impulso $h[n]$ tiene la misma salida que un sistema LTI con entrada $h[n]$ y respuesta al impulso $x[n]$, como se ilustra en la figura 1.1.

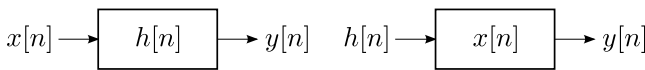


Figura 1.1: Propiedad conmutativa de la convolución.

- *Distributiva sobre la adición:*

$$x[n] * (h_1[n] + h_2[n]) = x[n] * h_1[n] + x[n] * h_2[n].$$

Esto significa que el sistema equivalente a dos sistemas LTI en paralelo con respuestas al impulso $h_1[n]$ y $h_2[n]$ es un sistema LTI con respuesta al impulso $h_1[n] + h_2[n]$, como se muestra en la figura 1.2.

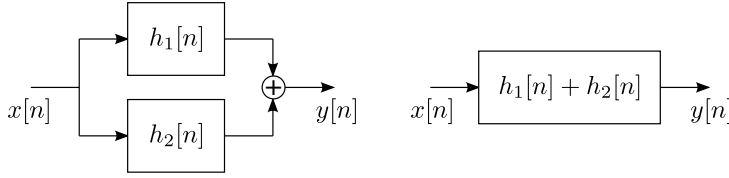


Figura 1.2: Propiedad conmutativa de la convolución y equivalencia de sistemas en paralelo.

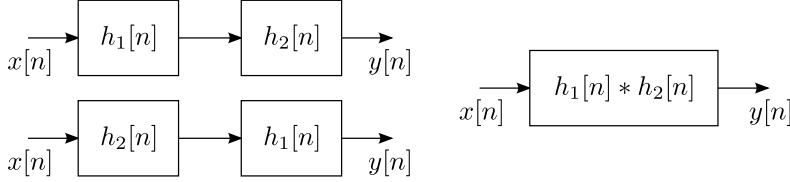


Figura 1.3: Propiedad asociativa de la convolución y equivalencia de sistemas en serie.

■ *Asociativa:*

$$y[n] = (x[n] * h_1[n]) * h_2[n] = x[n] * (h_1[n] * h_2[n]).$$

Esto implica que la respuesta al impulso de dos sistemas LTI en serie es la convolución de sus respuestas al impulso, como se muestra en la figura 1.3.

Dos propiedades de los sistemas adicionales a la linealidad y la invarianza temporal son la estabilidad y la causalidad, y es importante poder determinar cuando un sistema LTI es estable y cuando es causal. Como los sistemas LTI están caracterizados por la respuesta al impulso, es posible determinar las propiedades del sistema a partir de las características de la respuesta al impulso.

1.4.1. Estabilidad de los sistemas LTI

Recordar de la sección 1.2.5 que un sistema es estable si toda entrada acotada produce una salida acotada. Puede mostrarse que un sistema LTI es estable si y solo si su respuesta al impulso es absolutamente sumable, es decir, si

$$B_h = \sum_{k=-\infty}^{\infty} |h[k]| \leq \infty. \quad (1.16)$$

Para estudiar la estabilidad de un sistema, alcanza con calcular la respuesta al impulso a partir de la ecuación del sistema imponiendo como entrada $\delta[n]$ y luego verificar si se cumple la ecuación 1.16.

Por ejemplo, la respuesta al impulso de los sistemas media móvil y acumulador, dados respectivamente por las 1.7 y 1.8 son

■ *Media móvil:*

$$\begin{aligned} h[n] &= \frac{1}{M_1 + M_2 + 1} \sum_{k=-M_1}^{M_2} \delta[n - k] \\ &= \frac{1}{M_1 + M_2 + 1} (\delta[n + M_1] + \delta[n + M_1 - 1] + \cdots + \delta[n] + \cdots + \delta[n - M_2 + 1] + \delta[n - M_2]) \\ &= \begin{cases} \frac{1}{M_1 + M_2 + 1}, & -M_1 \leq n \leq M_2 \\ 0, & \text{en otro caso.} \end{cases} \\ &= \frac{1}{M_1 + M_2 + 1} (u[n + M_1] - u[n - M_2 - 1]) \end{aligned} \quad (1.17)$$

Por lo tanto,

$$B_h = \sum_{k=-M_1}^{M_2} \frac{1}{M_1 + M_2 + 1} = 1,$$

concluyendo que el sistema media móvil es estable.

■ *Acumulador:*

$$\begin{aligned} h[n] &= \sum_{k=-\infty}^n \delta[k] \\ &= \begin{cases} 1, & n \geq 0 \\ 0, & n < 0 \end{cases} \\ &= u[n]. \end{aligned}$$

Por lo tanto,

$$B_h = \sum_{k=-\infty}^{\infty} u[k] = \sum_{k=0}^{\infty} 1 = \infty,$$

concluyendo que el sistema acumulador no es estable.

Además, los sistemas retardo ideal y diferencia progresiva y regresiva, dados por las ecuaciones 1.6, 1.10 y 1.11 respectivamente, son estables. En general, un sistema con respuesta al impulso de duración finita (*FIR*, finite-duration impulse response) es siempre estable si las muestras de la respuesta al impulso tienen magnitud finita. Observar que la respuesta al impulso del sistema acumulador es de duración infinita (*IIR*, infinite-duration impulse response).

1.4.2. Causalidad de los sistemas LTI

Puede mostrarse a partir de la definición 1.15 de la convolución que la condición necesaria y suficiente para que un sistema sea causal es que la respuesta al impulso cumpla que

$$h[n] = 0, \quad n < 0. \quad (1.18)$$

Por este motivo, cualquier secuencia nula en $n < 0$ es referida como *secuencia causal*.

Para estudiar la causalidad de un sistema LTI, hay que calcular la respuesta al impulso y verificar si se cumple la ecuación 1.18. Al hacerlo, se comprobarán los resultados discutidos en la sección 1.2.4 sobre la causalidad de los sistemas empleados como ejemplo.

1.5. Ecuaciones en diferencias lineales con coeficientes constantes

Una clase importante de sistemas LTI consiste en los sistemas en los que la entrada $x[n]$ y la salida $y[n]$ satisfacen una ecuación en diferencias lineales con coeficientes constantes de orden N de la forma

$$\sum_{k=0}^N a_k y[n-k] = \sum_{m=0}^M b_m x[n-m]. \quad (1.19)$$

Ejemplo: representación como ecuación en diferencias del acumulador. El sistema acumulador se define como (ver la ecuación 1.8)

$$y[n] = \sum_{k=-\infty}^n x[k].$$

Para mostrar que la entrada y la salida satisfacen una ecuación en diferencias de la forma de la ecuación 1.19, escríbase la ecuación del acumulador como

$$y[n] = x[n] + \sum_{k=-\infty}^{n-1} x[k],$$

y además como

$$y[n-1] = \sum_{k=-\infty}^{n-1} x[k]$$

se obtiene que

$$y[n] = x[n] + y[n-1] \quad (1.20)$$

o equivalentemente

$$y[n] - y[n-1] = x[n].$$

La ecuación en diferencias 1.20 sugiere una implementación simple del acumulador: para cada valor de n , se suma el valor actual $x[n]$ de la entrada a la suma acumulada $y[n-1]$. Esta interpretación del acumulador se muestra en el diagrama de bloques de la figura 1.4.

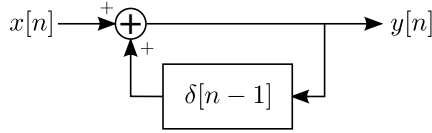


Figura 1.4: Diagrama de bloques de la ecuación en diferencias del acumulador.

Ejemplo: representación como ecuación en diferencias del sistema de media móvil. Considérese el sistema de media móvil, dado por la ecuación 1.7, con $M_1 = 0$ de forma de que sea causal. En ese caso, la ecuación del sistema es

$$y[n] = \frac{1}{M_2 + 1} \sum_{k=0}^{M_2} x[n-k], \quad (1.21)$$

que es una ecuación en diferencias de la forma de la ecuación 1.19.

Por otro lado, la respuesta al impulso, dada por la ecuación 1.17, puede escribirse como

$$h[n] = \frac{1}{M_2 + 1} (\delta[n] - \delta[n - M_2 - 1]) * u[n],$$

que sugiere la representación del sistema media móvil causal puede representarse mediante los sistemas en cascada de la figura 1.5. Para obtener la ecuación en diferencias de este diagrama de bloques, se

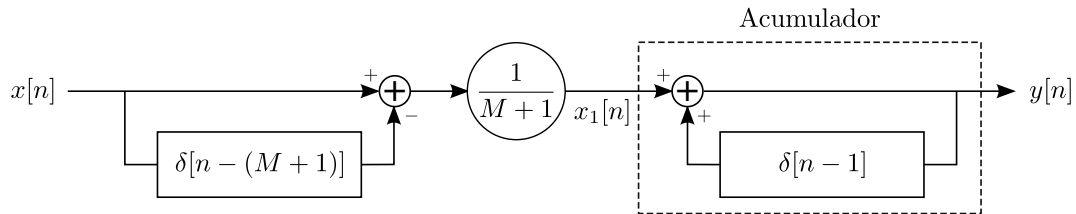


Figura 1.5: Diagrama de bloques del sistema media móvil.

observa primero que

$$x_1[n] = \frac{1}{M + 1} (x[n] - x[n - M_2 - 1]).$$

Además, de la ecuación 1.20, la salida del sistema acumulador cumple la ecuación en diferencias

$$y[n] - y[n-1] = x_1[n],$$

resultando en que

$$y[n] - y[n-1] = \frac{1}{M + 1} (x[n] - x[n - M_2 - 1]),$$

que es una ecuación en diferencias de la forma 1.19. En este ejemplo se ve que el sistema media móvil admite mas de una representación en ecuaciones en diferencia, una recursiva y otra no recursiva. Este es un caso muy particular de un sistema recursivo que tiene respuesta al impulso finita.

Al igual que en el caso de ecuaciones diferenciales en tiempo continuo, sin restricciones adicionales, una ecuación en diferencias lineal con coeficientes constantes en tiempo discreto no provee una única

especificación de la salida dada una entrada. Concretamente, supóngase que para una entrada dada $x_p[n]$, se determinó de alguna forma que la salida es $y_p[n]$, tal que se satisface la ecuación 1.19. Entonces, la misma ecuación con la misma entrada es satisfecha para cualquier salida de la forma

$$y[n] = y_p[n] + y_h[n],$$

donde $y_h[n]$ es cualquier solución de la ecuación 1.19 con $x[n] = 0$, es decir, cualquier solución de la ecuación

$$\sum_{k=0}^N a_k y_h[n-k] = 0. \quad (1.22)$$

La ecuación 1.22 se llama *ecuación en diferencias homogénea* y la solución $y_h[n]$ se llama *solución homogénea*. La secuencia $y_h[n]$ es un miembro de la familia de soluciones de la forma

$$y_h[n] = \sum_{m=1}^N A_m z_m^n, \quad (1.23)$$

donde los coeficientes A_m pueden elegirse para satisfacer un conjunto de condiciones auxiliares sobre $y[n]$. Efectivamente, sustituyendo la ecuación 1.23 en la ecuación 1.22 se observa que

$$\sum_{k=0}^N a_k \left(\sum_{m=1}^N A_m z_m^{n-k} \right) = 0,$$

y por lo tanto

$$\sum_{m=1}^N A_m z_m^n \left(\sum_{k=0}^N a_k z_m^{-k} \right) = 0.$$

Esta ecuación se cumple para cualquier conjunto de valores A_m cuando los números complejos z_m con $m = 1, 2, \dots, N$ son las raíces del polinomio

$$A(z) = \sum_{k=0}^N a_k z^{-k}. \quad (1.24)$$

Ejemplo (problema 2.50). Sea la ecuación en diferencias

$$y[n] - \frac{3}{4}y[n-1] + \frac{1}{8}y[n-2] = 2x[n-1].$$

Se determinará la forma general de la solución homogénea. Según la ecuación 1.23, en este caso la solución homogénea es de la forma

$$y_h[n] = A_1 z_1^n + A_2 z_2^n,$$

donde, como indica la ecuación 1.24, z_1 y z_2 son las raíces del polinomio

$$A(z) = 1 - \frac{3}{4}z^{-1} + \frac{1}{8}z^{-2},$$

que son

$$z_{1,2} = \frac{\frac{3}{4} \pm \sqrt{\frac{9}{16} - \frac{4}{8}}}{2} = \frac{\frac{3}{4} \pm \sqrt{\frac{1}{16}}}{2} = \frac{\frac{3}{4} \pm \frac{1}{4}}{2} \Rightarrow z_1 = \frac{1}{2} \quad \text{y} \quad z_2 = \frac{1}{4},$$

resultando en que

$$y_h[n] = A_1 \left(\frac{1}{2} \right)^n + A_2 \left(\frac{1}{4} \right)^n.$$

Puede comprobarse que el resultado obtenido es la solución homogénea verificando que cumple la ecuación homogénea para cualquier valor de A_1 y A_2 . Efectivamente,

$$\begin{aligned} y_h[n] - \frac{3}{4}y_h[n-1] + \frac{1}{8}y_h[n-2] &= A_1 \left(\frac{1}{2}\right)^n + A_2 \left(\frac{1}{4}\right)^n - \frac{3}{4} \left[A_1 \left(\frac{1}{2}\right)^{n-1} + A_2 \left(\frac{1}{4}\right)^{n-1} \right] \\ &\quad + \frac{1}{8} \left[A_1 \left(\frac{1}{2}\right)^{n-2} + A_2 \left(\frac{1}{4}\right)^{n-2} \right] \\ &= A_1 \left(\frac{1}{2}\right)^n \left(1 - \frac{3}{2} + \frac{1}{2}\right) + A_2 \left(\frac{1}{4}\right)^n (1 - 3 + 2) \\ &= 0. \end{aligned}$$

Finalmente se determinarán los coeficientes A_1 y A_2 de la solución homogénea si $y[-1] = 1$ y $y[0] = 0$. Con estas condiciones,

$$\begin{cases} y[-1] &= 2A_1 + 4A_2 &= 1 \\ y[0] &= A_1 + A_2 &= 0 \end{cases} \Rightarrow A_1 = -\frac{1}{2} \quad \text{y} \quad A_2 = \frac{1}{2},$$

resultando en que

$$y_h[n] = -\frac{1}{2} \left(\frac{1}{2}\right)^n + \frac{1}{2} \left(\frac{1}{4}\right)^n.$$

La ecuación 1.23 asume que todas las raíces del polinomio 1.24 son distintas. Si el polinomio tiene raíces múltiples, la ecuación 1.23 tiene una forma ligeramente diferente, como se ilustra en el siguiente ejemplo, pero siempre hay N coeficientes indeterminados.

Ejemplo (problema 2.50, continuación). Sea la ecuación en diferencias

$$y[n] - y[n-1] + \frac{1}{4}y[n-2] = 2x[n-1].$$

Se intentará determinar la forma general de la solución homogénea empleando la ecuación 1.23. Según esa ecuación, la solución homogénea es de la forma

$$y_h[n] = A_1 z_1^n + A_2 z_2^n,$$

donde, como indica la ecuación 1.24, z_1 y z_2 son las raíces del polinomio

$$A(z) = 1 - z^{-1} + \frac{1}{4}z^{-2},$$

que son

$$z_{1,2} = \frac{1 \pm \sqrt{1-1}}{2} = \frac{1}{2}.$$

En este caso, el polinomio $A(z)$ tiene una raíz doble, resultando en que

$$y_h[n] = A_1 \left(\frac{1}{2}\right)^n + A_2 \left(\frac{1}{2}\right)^n = (A_1 + A_2) \left(\frac{1}{2}\right)^n = C \left(\frac{1}{2}\right)^n.$$

Puede mostrarse que el resultado efectivamente satisface la ecuación homogénea para todo valor de C . Sin embargo, como hay solo un coeficiente no pueden imponerse el par de condiciones iniciales, como por ejemplo, $y[-1] = 1$ y $y[0] = 0$. En el caso en que el polinomio 1.24 tiene una raíz doble, en lugar de la ecuación 1.23, la forma general de la solución homogénea es

$$y_h[n] = \sum_{m=1}^{N-1} A_m z_m^n + n B_1 z_1^n,$$

donde la raíz duplicada es z_1 . De esta ecuación, se obtiene que la solución homogénea general es

$$y_h[n] = A_1 \left(\frac{1}{2}\right)^n + n B_1 \left(\frac{1}{2}\right)^n = (A_1 + n B_1) \left(\frac{1}{2}\right)^n.$$

Puede verificarse que efectivamente es la solución homogénea general, ya que

$$\begin{aligned}
y_h[n] - y_h[n-1] + \frac{1}{4}y_h[n-2] &= (A_1 + nB_1) \left(\frac{1}{2}\right)^n - [A_1 + (n-1)B_1] \left(\frac{1}{2}\right)^{n-1} \\
&\quad + \frac{1}{4}[A_1 + (n-2)B_1] \left(\frac{1}{2}\right)^{n-2} \\
&= [(A_1 + nB_1) - 2(A_1 + nB_1 - B_1) + (A_1 + nB_1 - 2B_1)] \left(\frac{1}{2}\right)^n \\
&= 0.
\end{aligned}$$

Si se imponen las condiciones $y[-1] = 1$ y $y[0] = 0$ pueden determinarse los coeficientes A_1 y B_1 :

$$\begin{cases} y[-1] &= 2A_1 - 2B_1 &= 1 \\ y[0] &= A_1 &= 0 \end{cases} \Rightarrow A_1 = 0 \quad \text{y} \quad B_1 = -\frac{1}{2},$$

resultando en que

$$y_h[n] = -\frac{n}{2^{n+1}}.$$

Como $y_h[n]$ tiene N coeficientes indeterminados, se necesita un conjunto de N condiciones auxiliares para la especificación unívoca de $y[n]$ dado $x[n]$. Esas condiciones auxiliares pueden consistir en la especificación de los valores de $y[n]$ para ciertos valores de n , como por ejemplo $y[-1]$, $y[-2]$, \dots , $y[-N]$. Los coeficientes se determinan resolviendo el sistema de N ecuaciones lineales, como se hizo en los ejemplos precedentes.

De particular interés son los sistemas lineales e invariantes en el tiempo, y en ese caso, las condiciones auxiliares deben ser consistentes con ese requerimiento. En el capítulo 2, donde se discute la solución de ecuaciones en diferencias empleando la transformada z , se incorporan implícitamente condiciones de linealidad e invarianza temporal. Como se verá, incluso con las condiciones de linealidad e invarianza temporal, la solución de la ecuación en diferencias no queda unívocamente determinada, y por lo tanto, tampoco el sistema. En particular, en general hay tanto un sistema LTI causal como uno no causal consistente con una ecuación en diferencias dada.

Ejemplo (problema 2.51). Se considera un sistema con entrada $x[n]$ y salida $y[n]$. La relación entre la entrada y la salida está definida por las siguientes dos propiedades:

1. $y[n] - ay[n-1] = x[n]$,
2. $y[0] = 1$,

y se quiere determinar si el sistema es lineal e invariante en el tiempo.

Como $y[0] = 1$ para cualquier entrada, el sistema no es lineal. Por ejemplo, en un sistema lineal, si se duplica la amplitud de la entrada, se duplica la amplitud de la salida, y en este sistema $y[0]$ tiene un valor fijo.

Para estudiar la invarianza temporal, sean las entradas $x_1[n] = \delta[n]$ y $x_2[n] = \delta[n-1]$. Como $x_2[n] = x_1[n-1]$, si el sistema fuera invariante en el tiempo se debería cumplir que $y_2[n] = y_1[n-1]$. La salida cuando la entrada es $x_1[n] = \delta[n]$ cuando $n \leq 0$ es

$$\begin{aligned}
y_1[0] &= 1 \\
y_1[1] &= ay_1[0] + \delta[1] = a \\
y_1[2] &= ay_1[1] + \delta[2] = a^2 \\
&\vdots \\
y_1[n] &= a^n,
\end{aligned}$$

y la salida cuando la entrada es $x_2[n] = \delta[n-1]$ cuando $n \leq 0$ es

$$\begin{aligned}
y_2[0] &= 1 \\
y_2[1] &= ay_2[0] + \delta[0] = a + 1 \\
y_2[2] &= ay_2[1] + \delta[1] = a^2 + a \\
&\vdots \\
y_2[n] &= a^n + a^{n-1}.
\end{aligned}$$

Como $y_2[n] \neq y_1[n-1]$, el sistema no es invariante en el tiempo.

Se considera ahora el caso en que la condición inicial es $y[0] = 0$. Para estudiar si el sistema es lineal, se considera primero una entrada arbitraria $x[n]$. La salida del sistema cuando $n \geq 0$ es

$$\begin{aligned} y[0] &= 0 \\ y[1] &= ay[0] + x[1] = x[1] \\ y[2] &= ay[1] + x[2] = ax[1] + x[2] \\ y[3] &= ay[2] + x[3] = a^2x[1] + ax[2] + x[3] \\ &\vdots \\ y[n] &= a^{n-1}x[1] + a^{n-2}x[2] + \cdots + ax[n-1] + x[n], \end{aligned}$$

es decir,

$$y[n] = \sum_{k=0}^{n-1} a^k x[n-k], \quad \text{para } n \geq 0.$$

Para calcular la salida recursivamente para $n < 0$, se escribe la ecuación en diferencias como

$$y[n-1] = a^{-1}(y[n] - x[n]) \quad \text{con} \quad y[0] = 0.$$

Luego,

$$\begin{aligned} y[-1] &= a^{-1}(y[0] - x[0]) = -a^{-1}x[0] \\ y[-2] &= a^{-1}(y[-1] - x[-1]) = -a^{-2}x[0] - a^{-1}x[-1] \\ y[-3] &= a^{-1}(y[-2] - x[-2]) = -a^{-3}x[0] - a^{-2}x[-1] - a^{-1}x[-2] \\ &\vdots \\ y[n] &= -a^n x[0] - a^{n+1}x[-1] - \cdots - a^{-2}x[n+2] - a^{-1}x[n+1], \end{aligned}$$

es decir,

$$y[n] = - \sum_{k=-1}^n a^k x[n-k], \quad \text{para } n < 0.$$

Se obtuvo que dada una entrada $x[n]$, la salida es

$$y[n] = \begin{cases} \sum_{k=0}^{n-1} a^k x[n-k], & n \geq 0 \\ - \sum_{k=-1}^n a^k x[n-k], & n < 0. \end{cases} \quad (1.25)$$

Sean $y_1[n]$ y $y_2[n]$ las correspondientes salidas cuando las entradas son $x_1[n]$ y $x_2[n]$, y se considera la entrada

$$x_3[n] = \alpha x_1[n] + \beta x_2[n].$$

De la ecuación 1.25, la salida cuando $n \geq 0$ es

$$\begin{aligned} y_3[n] &= \sum_{k=0}^{n-1} a^k x_3[n-k] \\ &= \sum_{k=0}^{n-1} a^k (\alpha x_1[n-k] + \beta x_2[n-k]) \\ &= \alpha \sum_{k=0}^{n-1} a^k x_1[n-k] + \beta \sum_{k=0}^{n-1} a^k x_2[n-k] \\ &= \alpha y_1[n] + \beta y_2[n]. \end{aligned}$$

De forma análoga, cuando $n < 0$, la salida es

$$\begin{aligned}
y_3[n] &= - \sum_{k=-1}^n a^k x_3[n-k] \\
&= - \sum_{k=-1}^n a^k (\alpha x_1[n-k] + \beta x_2[n-k]) \\
&= \alpha \left(- \sum_{k=-1}^n a^k x_1[n-k] \right) + \beta \left(- \sum_{k=-1}^n a^k x_2[n-k] \right) \\
&= \alpha y_1[n] + \beta y_2[n].
\end{aligned}$$

Se concluye que el sistema es lineal.

Si un sistema es caracterizado por una ecuación en diferencias lineal con coeficientes constantes y además es restringido a ser lineal, invariante en el tiempo y causal, entonces la solución es única. En este caso, se dice que las condiciones auxiliares son *condiciones iniciales de reposo*. Concretamente, la información auxiliar es que si la entrada $x[n]$ es nula para n menor a algún instante n_0 , la salida $y[n]$ se restringe a ser nula para n menor que n_0 . Esto provee condiciones iniciales suficientes para obtener $y[n]$ unívocamente para $n \geq n_0$ recursivamente.

Ejemplo. Se considera nuevamente el sistema definido por la ecuación en diferencias

$$y[n] - ay[n-1] = x[n]$$

del ejemplo anterior, y se impone la condición adicional de que si la entrada es nula para $n \leq n_0$, la salida es nula en $n \leq n_0$. Se mostrará que con esta imposición, el sistema es invariante en el tiempo. Sea la entrada

$$x_1[n] = x[n]u[n],$$

donde $x[n]$ es una secuencia arbitraria. Definida de esta forma se cumple que $x_1[n] = 0$ en $n < 0$, y se impone que la salida $y_1[n] = 0$ en $n < 0$. Esto equivale a imponer la condición auxiliar $y_1[-1] = 0$. En ese caso, la salida para $n \geq 0$ es

$$\begin{aligned}
y_1[0] &= ay_1[-1] + x_1[0] = x[0] \\
y_1[1] &= ay_1[0] + x_1[1] = ax[0] + x[1] \\
y_1[2] &= ay_1[1] + x_1[2] = a^2x[0] + ax[1] + x[2] \\
&\vdots \\
y_1[n] &= a^n x[0] + a^{n-1}x[1] + \cdots + ax[n-1] + x[n],
\end{aligned}$$

es decir,

$$y_1[n] = \sum_{k=0}^n a^k x[n-k], \quad \text{para } n \geq 0. \quad (1.26)$$

Para calcular la salida recursivamente para $n < 0$, se escribe la ecuación en diferencias como

$$y[n-1] = a^{-1}(y[n] - x[n]) \quad \text{con} \quad y[-1] = 0.$$

Luego,

$$\begin{aligned}
y_1[-2] &= a^{-1}(y_1[-1] - x_1[-1]) = 0 \\
y_1[-3] &= a^{-1}(y_1[-2] - x_1[-2]) = 0 \\
&\vdots \\
y_1[n] &= 0, \quad \text{para } n < 0.
\end{aligned}$$

Puede verse imponiendo como entrada $x[n] = \delta[n]$, que la respuesta al impulso del sistema es

$$h[n] = a^n u[n],$$

y por lo tanto, el sistema es causal. Además, la ecuación 1.26 se puede escribir como

$$y_1[n] = \sum_{k=0}^n a^k x[n-k] = \sum_{k=-\infty}^{\infty} h[k] x_1[n-k] = h[n] * x_1[n].$$

Se considera ahora la entrada $x_2[n] = x_1[n - n_0] = x[n - n_0]u[n - n_0]$, donde n_0 es un entero arbitrario. Ahora, $x_2[n] = 0$ en $n < n_0$, y la restricción de que $y_2[n] = 0$ en $n < n_0$ corresponde a la condición inicial $y_2[n_0 - 1] = 0$ de la ecuación en diferencias. Con esta condición inicial, la salida para $n \geq n_0$ es

$$\begin{aligned} y_2[n_0] &= a y_2[n_0 - 1] + x_2[n_0] = x[0] \\ y_2[n_0 + 1] &= a y_2[n_0] + x_2[n_0 + 1] = a x[0] + x[1] \\ &\vdots \\ y_2[n_0 + m] &= a^m x[0] + a^{m-1} x[1] + \cdots + a x[m-1] + x[m], \end{aligned}$$

es decir,

$$y_2[n_0 + m] = \sum_{k=0}^m a^k x[m-k], \quad \text{para } m \geq 0.$$

Esto también se puede escribir como

$$y_2[n] = \sum_{k=0}^{n-n_0} a^k x[n-n_0-k], \quad \text{para } n \geq n_0.$$

resolviendo la recursión puede verificarse además que $y_2[n] = 0$ si $n < n_0$. Por lo tanto, reemplazando n por $n - n_0$ en la ecuación 1.26, se deduce que

$$y_2[n] = y_1[n - n_0]$$

para todo n_0 . Considerando que $x_2[n] = x_1[n - n_0]$, se concluye que el sistema es invariante en el tiempo.

1.6. Representación de señales y sistemas en el dominio de la frecuencia

Las secuencias sinusoidales y exponenciales complejas juegan un rol particularmente importante en la representación de señales en tiempo discreto. Esto se debe a que las secuencias exponenciales complejas son funciones propias de los sistemas LTI, y la respuesta de un sistema LTI a una entrada sinusoidal es una secuencia sinusoidal con la misma frecuencia de la entrada pero con amplitud y fase determinadas por el sistema.

1.6.1. Funciones propias de los sistemas LTI

La propiedad de función propia de las exponenciales complejas para los sistemas LTI se obtiene directamente de la definición 1.15 de la suma de convolución. Específicamente, si la entrada a un sistema LTI con respuesta al impulso $h[n]$ es $x[n] = e^{j\omega n}$ para $-\infty < n < \infty$, de la ecuación 1.15, la salida es

$$\begin{aligned} y[n] &= \sum_{k=-\infty}^{\infty} h[k] e^{j\omega(n-k)} \\ &= e^{j\omega n} \sum_{k=-\infty}^{\infty} h[k] e^{-j\omega k}. \end{aligned}$$

Definiendo a la constante compleja

$$H(e^{j\omega}) = \sum_{k=-\infty}^{\infty} h[k]e^{-j\omega k}, \quad (1.27)$$

la salida es

$$y[n] = H(e^{j\omega})e^{j\omega n}. \quad (1.28)$$

En consecuencia, $e^{j\omega n}$ es una función propia del sistema, y valor propio asociado es $H(e^{j\omega})$. El valor propio $H(e^{j\omega})$ se llama *respuesta en frecuencia* del sistema y describe el cambio de la magnitud y la fase de la exponencial compleja de entrada en función de la frecuencia.

En la sección 1.7 se verá que un amplio conjunto de señales se puede representar como la combinación lineal de exponenciales complejas como

$$x[n] = \sum_k \alpha_k e^{j\omega_k n}. \quad (1.29)$$

Del principio de superposición, dado por la ecuación 1.13, y la ecuación 1.28, la correspondiente salida de un sistema LTI es

$$y[n] = \sum_k \alpha_k H(e^{j\omega_k}) e^{j\omega_k n}.$$

Por lo tanto, si se puede encontrar una representación de la entrada como superposición de secuencias exponenciales complejas, es posible obtener la salida si se conoce la respuesta en frecuencia del sistema para todas las frecuencias ω .

Ejemplo: respuesta sinusoidal de los sistemas LTI. Sea la entrada sinusoidal

$$x[n] = A \cos(\omega_0 n + \phi) = \underbrace{\frac{A}{2} e^{j\phi} e^{j\omega_0 n}}_{x_1[n]} + \underbrace{\frac{A}{2} e^{-j\phi} e^{-j\omega_0 n}}_{x_2[n]}.$$

De la ecuación 1.28, la respuesta a $x_1[n]$ es

$$y_1[n] = H(e^{j\omega_0}) \frac{A}{2} e^{j\phi} e^{j\omega_0 n}$$

y la respuesta a $x_2[n]$ es

$$y_2[n] = H(e^{-j\omega_0}) \frac{A}{2} e^{-j\phi} e^{-j\omega_0 n}.$$

Por lo tanto, por el principio de superposición, la respuesta total es

$$\begin{aligned} y[n] &= \frac{A}{2} [H(e^{j\omega_0}) e^{j\phi} e^{j\omega_0 n} + H(e^{-j\omega_0}) e^{-j\phi} e^{-j\omega_0 n}] \\ &\stackrel{(a)}{=} \frac{A}{2} [H(e^{j\omega_0}) e^{j\phi} e^{j\omega_0 n} + H^*(e^{j\omega_0}) e^{-j\phi} e^{-j\omega_0 n}] \\ &= A \operatorname{Re} [H(e^{j\omega_0}) e^{j\phi} e^{j\omega_0 n}], \end{aligned}$$

donde en (a) se consideró que si la respuesta al impulso $h[n]$ del sistema es real, se cumple que $H(e^{-j\omega_0}) = H^*(e^{j\omega_0})$. Expresando la respuesta en frecuencia en ω_0 como

$$H(e^{j\omega_0}) = G_0 e^{j\theta_0}, \quad \text{donde} \quad G_0 = |H(e^{j\omega_0})| \quad \text{y} \quad \theta_0 = \angle H(e^{j\omega_0}),$$

resulta en que

$$y[n] = A G_0 \cos(\omega_0 n + \phi + \theta_0). \quad (1.30)$$

Si la entrada a un sistema LTI es una secuencia sinusoidal de frecuencia ω_0 y la respuesta en magnitud y fase del sistema en esa frecuencia es respectivamente G_0 y θ_0 , como se muestra en la figura 1.6, la ecuación 1.30 indica que la salida es una secuencia sinusoidal con la misma frecuencia ω_0 pero con

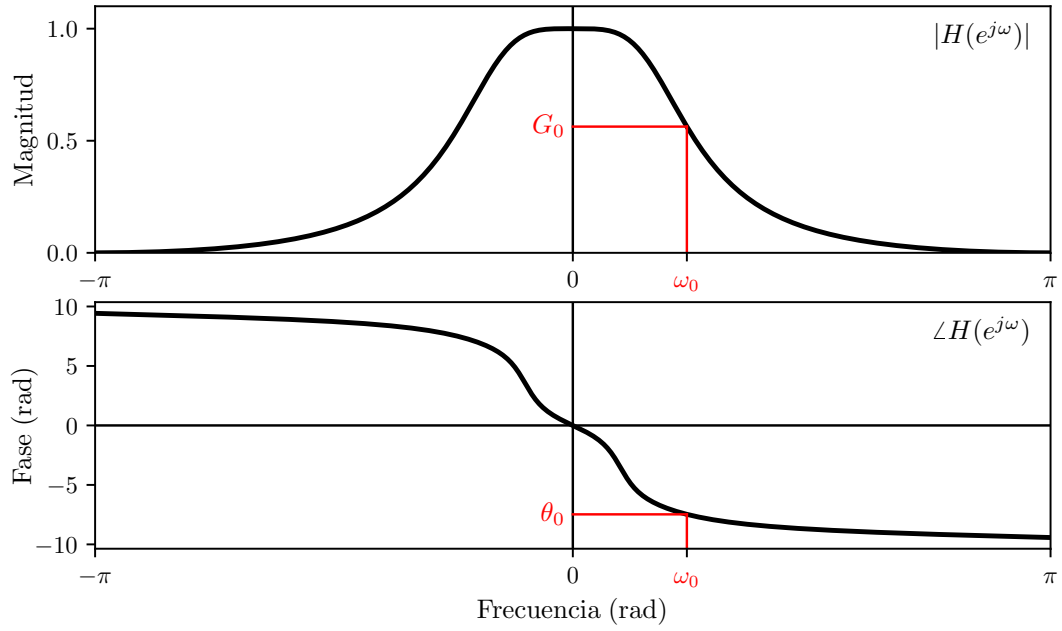


Figura 1.6: Respuesta en magnitud $|H(e^{j\omega})|$ y fase $\angle H(e^{j\omega})$ de un sistema LTI. Se trata de un filtro pasabajos.

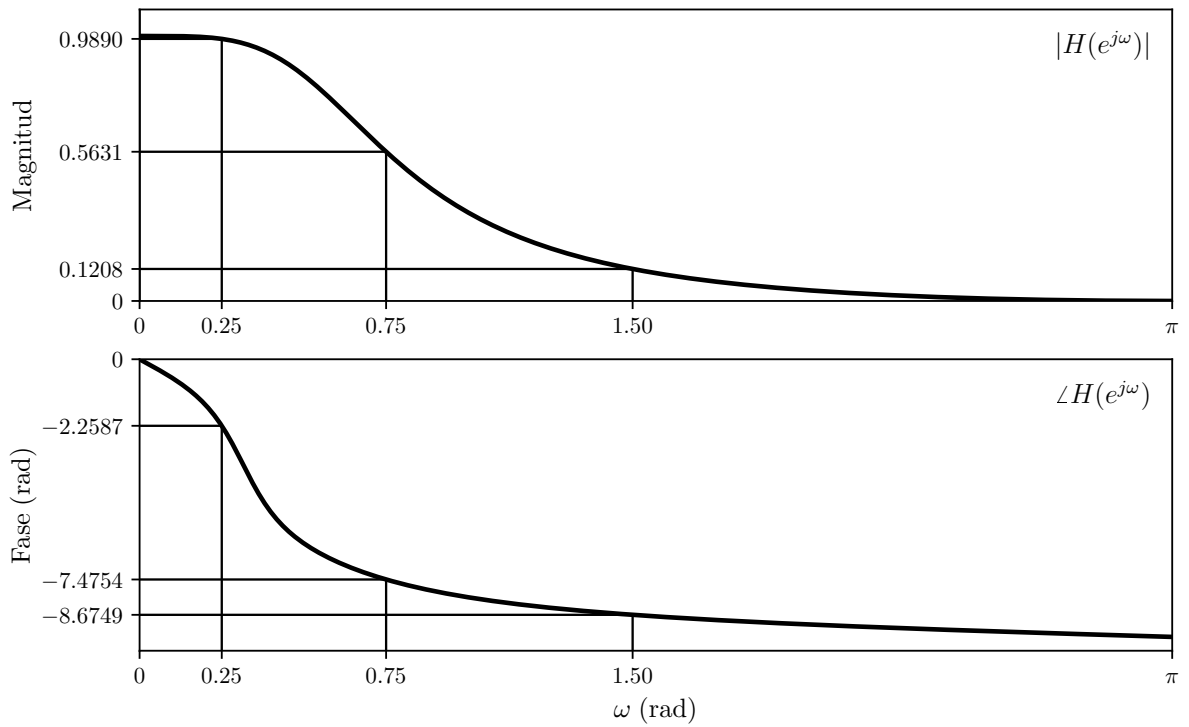


Figura 1.7: Sistema pasabajos LTI. Se especifica la respuesta en magnitud y en fase en las frecuencias 0.25, 0.75 y 1.5 radianes. Observar que la respuesta en fase está *desdoblada* (*unwrapped*).

amplitud alterada un factor G_0 y fase desplazada una cantidad θ_0 radianes. Este sistema se llama *filtro pasabajos*, ya que deja pasar las bajas frecuencias inalteradas e impide pasar las frecuencias altas.

A continuación se ilustra el resultado de la ecuación 1.30 con un ejemplo numérico. Sea el filtro pasabajos LTI cuya respuesta en frecuencia se muestra en la figura 1.7, y se consideran las entradas

$$x_i[n] = \text{sen } \omega_i n, \quad \text{para} \quad i = 0, 1, 2,$$

donde $\omega_0 = 0.25$, $\omega_1 = 0.75$ y $\omega_2 = 1.5$ radianes. Según la ecuación 1.30, las correspondientes salidas son

$$y_i[n] = G_i \text{sen}(\omega_i n + \theta_i), \quad (1.31)$$

donde G_i y θ_i son respectivamente la respuesta en magnitud y en fase del sistema en la frecuencia ω_i , es decir,

$$G_i = |H(e^{j\omega_i})| \quad \text{y} \quad \theta_i = \angle H(e^{j\omega_i}).$$

En la figura 1.8 se muestran las entradas $x_i[n]$ y las correspondientes salidas $y_i[n]$. Se observa que, acorde a la ecuación 1.31, la salida tiene la misma frecuencia que la entrada, pero con amplitud modificada según la repuesta en magnitud del sistema en la frecuencia ω_i .

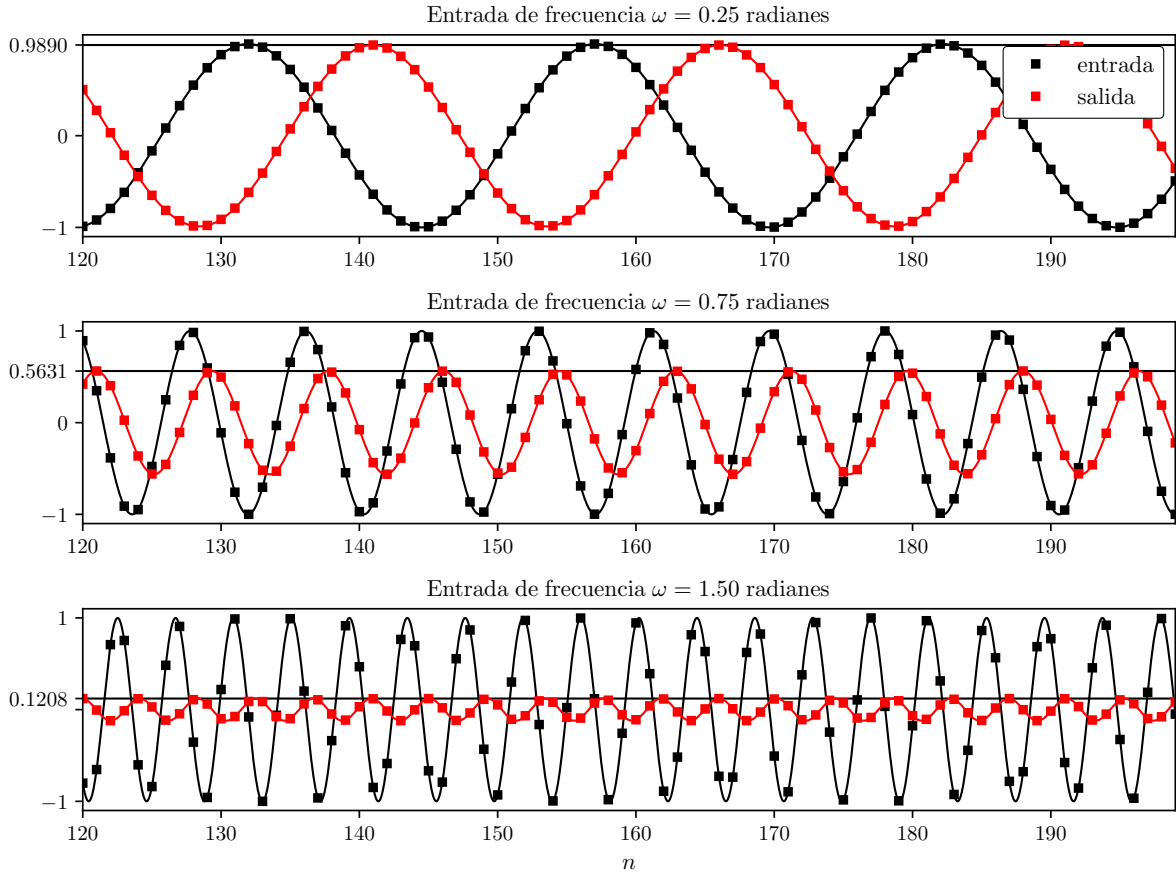


Figura 1.8: Ilustración de la respuesta en magnitud de un sistema LTI para una señal sinusoidal. La amplitud de la salida es acorde a la repuesta en magnitud del sistema. La línea continua que interpola las muestras es solo para mejor visualización.

Respecto a la fase de la senoide de salida, la ecuación 1.30 indica que a la fase de la entrada se le adiciona la respuesta en fase del sistema en la frecuencia de la senoide. La alteración de la fase puede interpretarse como un retardo de la señal. Efectivamente, considerando que la ecuación 1.31 puede escribirse como

$$y_i[n] = G_i \text{sen} \left\{ \omega_i \left[n - \left(-\frac{\theta_i}{\omega_i} \right) \right] \right\},$$

la senoide de salida está retardada una cantidad de $-\theta_i/\omega_i$ muestras respecto a la senoide de entrada. Esto motiva a la definición del *retardo de fase* de un sistema como

$$\tau_\phi(\omega) = -\frac{\angle H(e^{j\omega})}{\omega}. \quad (1.32)$$

En la figura 1.9 se muestra la respuesta en fase y el retardo de fase para el sistema del ejemplo, y se

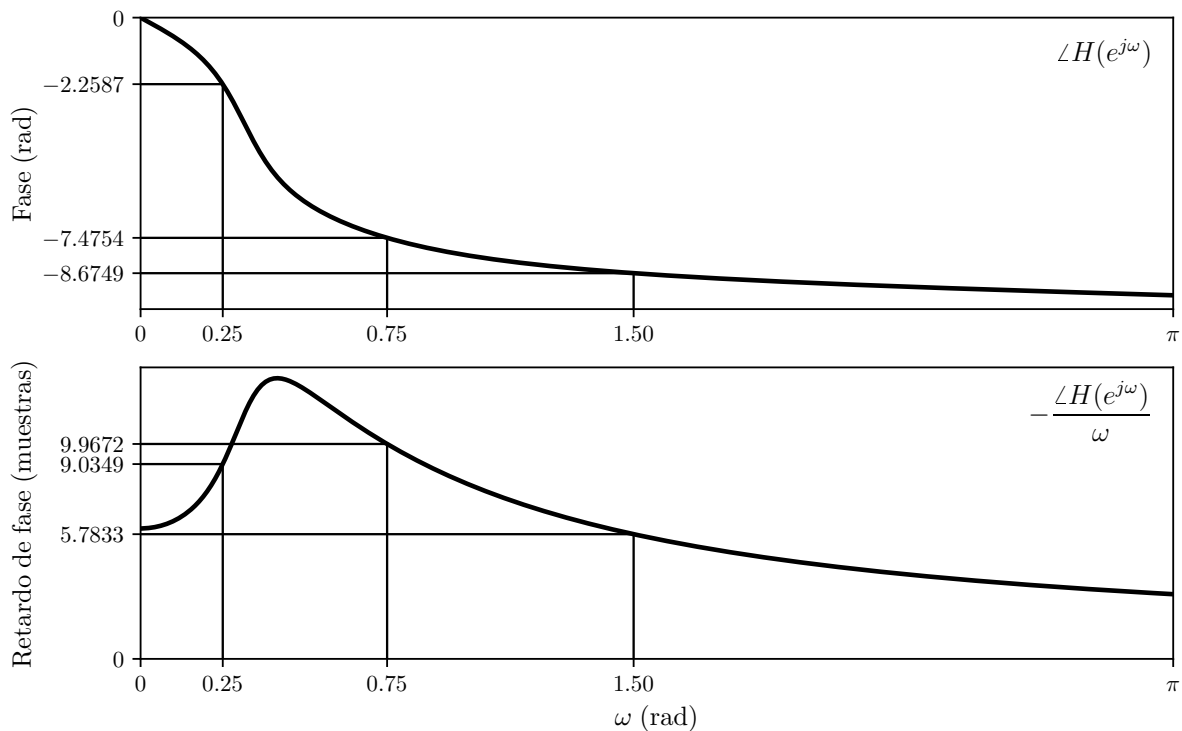


Figura 1.9: Respuesta en fase y retardo de fase del sistema. El retardo de fase se interpreta como la cantidad de muestras que el sistema retarda a una senoide de entrada.

especifica la fase y el retardo de fase en las frecuencias de las sinusoides de entrada. En la figura 1.10 se muestran las entradas y las salidas y se ilustra la interpretación del retardo de fase. Observar que el retardo de fase, que no necesariamente es un número entero, es el retardo temporal de la senoide de tiempo continuo subyacente a la senoide en tiempo discreto. En el caso de las señales de entrada de este ejemplo, las señales en tiempo continuo subyacentes son

$$x_i(t) = \text{sen } \omega_i t,$$

donde las unidades de t son segundos y en los instantes de muestreo $t = n$ segundos, y el valor ω_i es el mismo que el de la señal en tiempo discreto pero las unidades son radianes/segundo en lugar de radianes o radianes/muestra. En la figura 1.10 la señal continua subyacente en cada caso está representada con trazo continuo. La señal equivalente en tiempo continuo de las salidas es

$$y_i(t) = G_i \text{sen } \{\omega_i [t - \tau_\phi(\omega_i)]\},$$

donde el retardo de fase $\tau_\phi(\omega_i)$ tiene el mismo valor que en el caso discreto pero las unidades son segundos en lugar de adimensionado o muestras. En la sección 4.5 de [1] se da una interpretación de un retardo no entero.

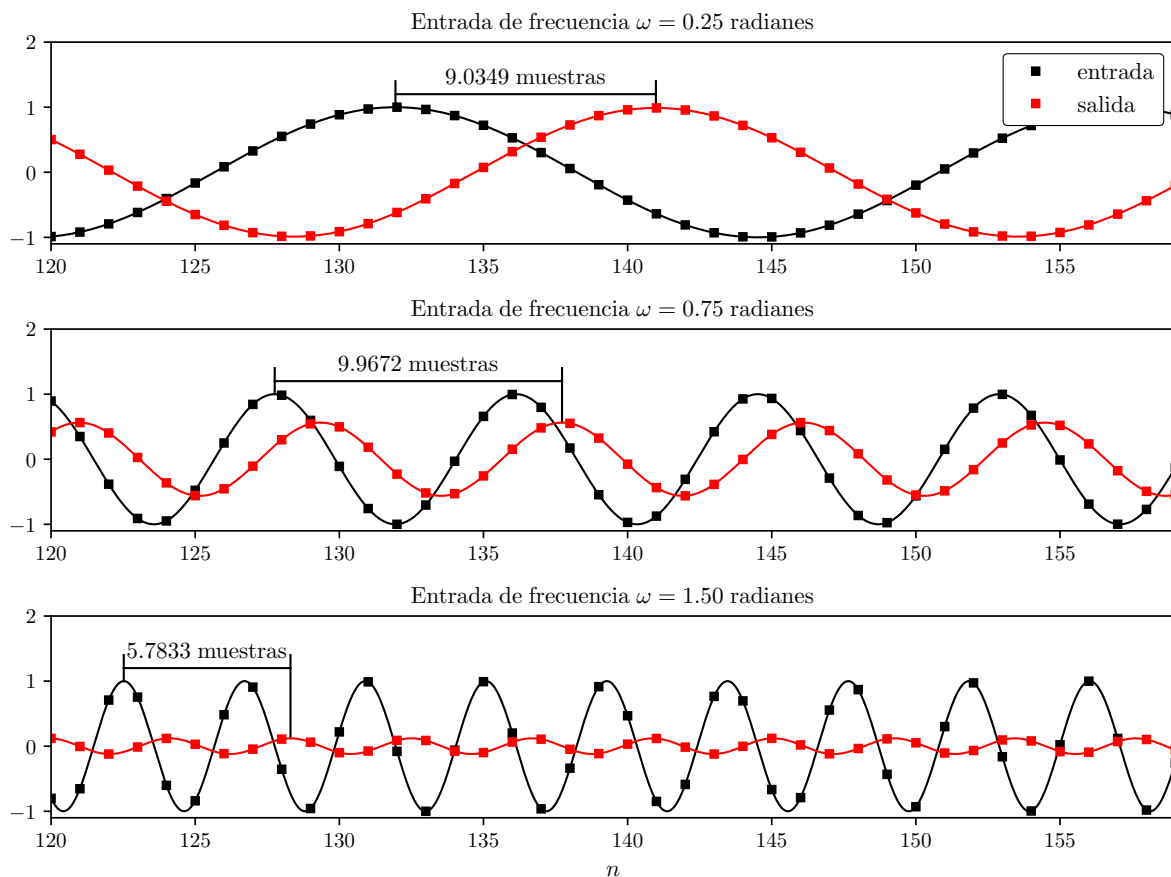


Figura 1.10: Ilustración del retardo de fase de un sistema. Se interpreta como el retardo en muestras de la señal sinusoidal de tiempo continuo subyacente a la señal sinusoidal de tiempo discreto. Por supuesto que los instantes de muestreo no cambian.

Ejemplo: respuesta en frecuencia del sistema de media móvil La respuesta al impulso del sistema de media móvil está dado por la ecuación 1.7 es (ver la ecuación 1.17)

$$h[n] = \begin{cases} \frac{1}{M_1 + M_2 + 1}, & -M_1 \leq n \leq M_2, \\ 0, & \text{en otro caso.} \end{cases}$$

Por lo tanto, la respuesta en frecuencia es

$$H(e^{j\omega}) = \frac{1}{M_1 + M_2 + 1} \sum_{n=-M_1}^{M_2} e^{-j\omega n}.$$

Para el sistema causal, $M_1 = 0$ y la respuesta en frecuencia queda

$$H(e^{j\omega}) = \frac{1}{M_2 + 1} \sum_{n=0}^{M_2} e^{-j\omega n},$$

y empleando el resultado de la suma de los primeros $M_2 + 1$ términos de una serie geométrica, se puede escribir como

$$\begin{aligned} H(e^{j\omega}) &= \frac{1}{M_2 + 1} \left(\frac{1 - e^{-j\omega M_2 + 1}}{1 - e^{-j\omega}} \right) \\ &= \frac{1}{M_2 + 1} \frac{[e^{j\omega(M_2+1)/2} - e^{-j\omega(M_2+1)/2}] e^{-j\omega(M_2+1)/2}}{(e^{j\omega/2} - e^{-j\omega/2}) e^{j\omega/2}} \end{aligned}$$

resultando en

$$H(e^{j\omega}) = \frac{1}{M_2 + 1} \frac{\text{sen} [\omega(M_2 + 1)/2]}{\text{sen } \omega/2} e^{-j\omega M_2/2}. \quad (1.33)$$

La magnitud y la fase de $H(e^{j\omega})$ para este caso con $M_2 = 4$ se muestra en la figura 1.11.

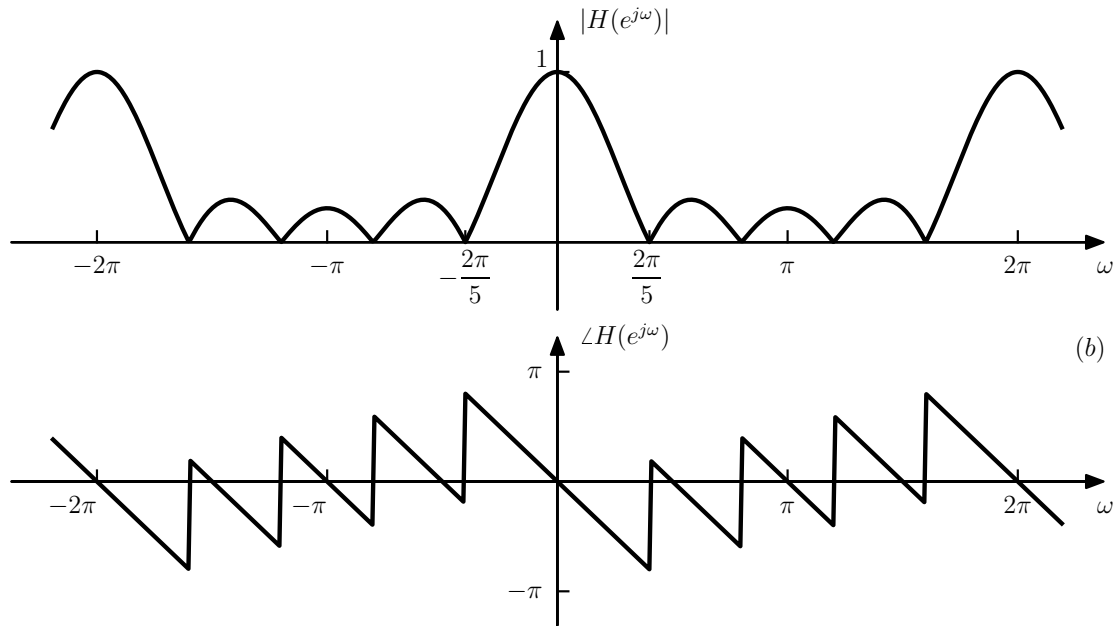


Figura 1.11: Magnitud y fase de la respuesta en frecuencia del sistema de media móvil causal con $M_2 = 4$.

1.7. Representación de secuencias mediante transformadas de Fourier

Una ventaja de la representación en frecuencias de los sistemas LTI es que a menudo la interpretación del comportamiento del sistema se observa inmediatamente en el dominio de las frecuencias, como es el caso de los filtros selectores de frecuencias. En esta sección se estudia como encontrar representaciones de señales de la forma de la ecuación 1.29.

Muchas secuencias pueden representarse mediante una integral de Fourier como

$$x[n] = \frac{1}{2\pi} \int_{-\pi}^{\pi} X(e^{j\omega}) e^{j\omega n} d\omega, \quad (1.34)$$

donde

$$X(e^{j\omega}) = \sum_{n=-\infty}^{\infty} x[n] e^{-j\omega n}. \quad (1.35)$$

Las ecuaciones 1.34 y 1.35 son la representación de Fourier de la secuencia. La ecuación 1.34, la *transformada de Fourier inversa*, es una ecuación de síntesis. Representa a $x[n]$ como una superposición de sinusoidales complejas infinitesimalmente pequeñas de la forma

$$\frac{1}{2\pi} X(e^{j\omega}) e^{j\omega n} d\omega,$$

con ω variando en un intervalo de largo 2π y donde $X(e^{j\omega})$ determina el peso relativo de cada componente sinusoidal complejo. La ecuación 1.35, la *transformada de Fourier*, es una expresión para calcular $X(e^{j\omega})$, es decir, para analizar la secuencia $x[n]$ para determinar cuanto de cada componente de frecuencia se necesita para sintetizar $x[n]$ mediante la ecuación 1.34. La ecuación 1.35 es referida explícitamente como la *transformada de Fourier en tiempo discreto* (*DTFT*, Discrete-Time Fourier Transform).

En general la transformada de Fourier es una función compleja de ω , y se puede representar en forma rectangular como

$$X(e^{j\omega}) = X_R(e^{j\omega}) + jX_I(e^{j\omega})$$

o en forma polar como

$$X(e^{j\omega}) = |X(e^{j\omega})| e^{j\angle X(e^{j\omega})},$$

donde $|X(e^{j\omega})|$ representa la magnitud y $\angle X(e^{j\omega})$ la fase. Observar que la fase $\angle X(e^{j\omega})$ no está unívocamente especificada, ya que puede sumarse cualquier múltiplo de 2π a cualquier valor de frecuencia ω sin alterar el resultado de la exponencial compleja. Cuando se hace referencia al valor principal, es decir, cuando $\angle X(e^{j\omega})$ se restringe al rango de valores entre $-\pi$ y π , se denota como $\text{ARG}[X(e^{j\omega})]$. Cuando se hace referencia a una función de fase que es continua en ω para $0 < \omega < \pi$ se emplea la notación $\arg[X(e^{j\omega})]$.

Comparando las ecuaciones 1.27 y 1.35, es claro que la respuesta en frecuencia de un sistema LTI es la transformada de Fourier de la respuesta al impulso. La respuesta al impulso puede obtenerse a partir de la respuesta en frecuencia mediante la transformada de Fourier inversa como

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

La transformada de Fourier es una función periódica de período 2π . Para representar funciones periódicas puede emplearse la serie de Fourier, y efectivamente la ecuación 1.35 es de la forma de la representación en series de Fourier de la función periódica $X(e^{j\omega})$, y la ecuación 1.34 es de la forma de la integral empleada para obtener los coeficientes de la serie de Fourier.

La clase de señales que pueden representarse mediante la ecuación 1.34 son aquellas para las que la suma infinita de la ecuación 1.35 converge. Una condición suficiente para la convergencia de dicha suma es que la secuencia $x[n]$ sea absolutamente sumable. En ese caso, puede mostrarse que la serie converge uniformemente a una función continua de ω . Como una secuencia estable es absolutamente sumable, todas las secuencias estables tienen transformada de Fourier. Equivalentemente, todo sistema estable tiene respuesta en frecuencia finita y continua. Las secuencias de largo finito y los sistemas FIR también tienen transformada de Fourier.

Algunas secuencias no son absolutamente sumables pero son cuadráticamente sumables, es decir, tienen energía finita

$$E_x = \sum_{n=-\infty}^{\infty} |x[n]|^2 < \infty,$$

que es una condición mas débil que la de sumabilidad absoluta. Estas secuencias pueden representarse mediante una transformada de Fourier pero relajando la condición de convergencia uniforme de la suma que define a $X(e^{j\omega})$. En este caso, la convergencia es en media cuadrática, es decir,

$$\lim_{M \rightarrow \infty} \int_{-\pi}^{\pi} |X(e^{j\omega}) - X_M(e^{j\omega})|^2 d\omega = 0,$$

donde

$$X_M(e^{j\omega}) = \sum_{n=-M}^M x[n] e^{-j\omega n}.$$

El error $|X(e^{j\omega}) - X_M(e^{j\omega})|$ puede no tender a cero para cada valor de ω cuando $M \rightarrow \infty$ pero si lo hace la energía total del error. Un ejemplo de este caso es el pasabajos ideal.

Ejemplo: transformada de Fourier de secuencias exponenciales complejas Se considera la secuencia $x[n]$ cuya transformada de Fourier es el tren de impulsos periódico

$$X(e^{j\omega}) = \sum_{r=-\infty}^{\infty} 2\pi\delta(\omega - \omega_0 + 2\pi r), \quad (1.36)$$

donde $-\pi < \omega_0 < \pi$. Para determinar $x[n]$ puede aplicarse la ecuación 1.34 de transformada inversa de Fourier a $X(e^{j\omega})$. Como el intervalo de integración en la ecuación 1.34 se extiende sobre $-\pi < \omega < \pi$ solo el impulso con $r = 0$ pertenece a dicho intervalo, y por lo tanto,

$$x[n] = \frac{1}{2\pi} \int_{-\pi}^{\pi} 2\pi\delta(\omega - \omega_0) e^{j\omega n} d\omega = e^{j\omega_0 n}.$$

Se obtuvo que

$$x[n] = e^{j\omega_0 n} \xleftrightarrow{\mathcal{F}} X(e^{j\omega}) = \sum_{r=-\infty}^{\infty} 2\pi\delta(\omega - \omega_0 + 2\pi r). \quad (1.37)$$

En este caso, $x[n]$ no es absolutamente ni cuadráticamente sumable, y $|X(e^{j\omega})|$ no es finita para todo ω . La igualdad matemática

$$\sum_{n=-\infty}^{\infty} e^{j\omega n} e^{-j\omega_0 n} = \sum_{r=-\infty}^{\infty} 2\pi\delta(\omega - \omega_0 + 2\pi r),$$

que proviene de sustituir $x[n] = e^{j\omega_0 n}$ en la ecuación 1.35 e igualar el resultado con la ecuación 1.36, debe interpretarse en el sentido de las funciones generalizadas. Con $\omega_0 = 0$ se obtiene el par de transformadas

$$x[n] = 1 \xleftrightarrow{\mathcal{F}} X(e^{j\omega}) = \sum_{r=-\infty}^{\infty} 2\pi\delta(\omega + 2\pi r). \quad (1.38)$$

Además, considerando que

$$\cos \omega_0 n = \frac{1}{2} e^{j\omega_0 n} + \frac{1}{2} e^{-j\omega_0 n} \quad \text{y} \quad \sin \omega_0 n = \frac{1}{2j} e^{j\omega_0 n} - \frac{1}{2j} e^{-j\omega_0 n},$$

y empleando la propiedad de linealidad de la transformada de Fourier, de 1.37 se deduce inmediatamente que

$$\begin{aligned} \cos \omega_0 n &\xleftrightarrow{\mathcal{F}} \pi\delta(\omega - \omega_0) + \pi\delta(\omega + \omega_0) \\ \sin \omega_0 n &\xleftrightarrow{\mathcal{F}} \frac{\pi}{j}\delta(\omega - \omega_0) - \frac{\pi}{j}\delta(\omega + \omega_0) \end{aligned} \quad \text{en} \quad -\pi < \omega < \pi. \quad (1.39)$$

1.8. Propiedades de simetría de la transformada de Fourier

Las propiedades de simetría de la transformada de Fourier son útiles para simplificar el cálculo de transformadas. Previo a la enumeración de algunas propiedades, se brindan algunas definiciones.

Una *secuencia simétrica conjugada* $x_e[n]$ es una secuencia que cumple que $x_e[n] = x_e^*[-n]$. También suele llamarse secuencia hermitica. Una *secuencia antisimétrica conjugada* es una secuencia que cumple que $x_o[n] = -x_o^*[-n]$. Puede mostrarse que cualquier secuencia $x[n]$ puede expresarse como la suma de una secuencia simétrica conjugada y una secuencia antisimétrica conjugada,

$$x[n] = x_e[n] + x_o[n],$$

donde

$$x_e[n] = \frac{1}{2}(x[n] + x^*[-n]) = x_e^*[-n] \quad \text{y} \quad x_o[n] = \frac{1}{2}(x[n] - x^*[-n]) = -x_o^*[-n].$$

Una secuencia real que cumple que $x_e[n] = x_e[-n]$ se llama *secuencia par* y una secuencia real que cumple que $x_o[n] = -x_o[-n]$ se llama *secuencia impar*. De forma analoga, la transformada de Fourier puede descomponerse en la suma de una función simétrica conjugada y una función antisimétrica como

$$X(e^{j\omega}) = X_e(e^{j\omega}) + X_o(e^{j\omega}),$$

donde

$$X_e(e^{j\omega}) = \frac{1}{2}[X_e(e^{j\omega}) + X_e^*(e^{-j\omega})] = X_e^*(e^{-j\omega}) \quad \text{y} \quad X_o(e^{j\omega}) = \frac{1}{2}[X_o(e^{j\omega}) - X_o^*(e^{-j\omega})] = -X_o^*(e^{-j\omega}).$$

A continuación se incluyen algunas propiedades de simetría de la transformada de Fourier. Si

$$x[n] \xleftrightarrow{\mathcal{F}} X(e^{j\omega}).$$

se cumple que:

$$\begin{array}{lll} x^*[n] \xleftrightarrow{\mathcal{F}} X^*(e^{-j\omega}) & \text{Re}(x[n]) \xleftrightarrow{\mathcal{F}} X_e(e^{j\omega}) & x_e[n] \xleftrightarrow{\mathcal{F}} X_R(e^{j\omega}) = \text{Re}[X(e^{j\omega})] \\ x^*[-n] \xleftrightarrow{\mathcal{F}} X^*(e^{j\omega}) & j \text{Im}(x[n]) \xleftrightarrow{\mathcal{F}} X_o(e^{j\omega}) & x_o[n] \xleftrightarrow{\mathcal{F}} jX_I(e^{j\omega}) = j \text{Im}[X(e^{j\omega})] \end{array}$$

Además, si $x[n]$ es real se cumple que:

- La transformada de Fourier es simétrica conjugada:

$$X(e^{j\omega}) = X^*(e^{-j\omega}).$$

- La magnitud es una función par:

$$|X(e^{j\omega})| = |X(e^{-j\omega})|.$$

- La fase es una función impar:

$$\angle X(e^{j\omega}) = -\angle X(e^{-j\omega}).$$

- La parte real es par:

$$X_R(e^{j\omega}) = X_R(e^{-j\omega}).$$

- La parte imaginaria es impar:

$$X_I(e^{j\omega}) = -X_I(e^{-j\omega}).$$

1.9. Teoremas de la transformada de Fourier

1.9.1. Linealidad de la transformada de Fourier

Si

$$x_1[n] \xleftrightarrow{\mathcal{F}} X_1(e^{j\omega}) \quad \text{y} \quad x_2[n] \xleftrightarrow{\mathcal{F}} X_2(e^{j\omega})$$

se cumple que

$$ax_1[n] + bx_2[n] \xleftrightarrow{\mathcal{F}} aX_1(e^{j\omega}) + bX_2(e^{j\omega}).$$

1.9.2. Desplazamiento temporal y desplazamiento en frecuencia

Si

$$x[n] \xleftrightarrow{\mathcal{F}} X(e^{j\omega})$$

se cumple que

$$x[n - n_d] \xleftrightarrow{\mathcal{F}} e^{-j\omega n_d} X(e^{j\omega})$$

y

$$e^{j\omega_0 n} x[n] \xleftrightarrow{\mathcal{F}} X(e^{j(\omega - \omega_0)}).$$

1.9.3. Inversión temporal

Si

$$x[n] \xleftrightarrow{\mathcal{F}} X(e^{j\omega}),$$

la transformada de la secuencia invertida temporalmente es

$$x[-n] \xleftrightarrow{\mathcal{F}} X(e^{-j\omega}).$$

Si $x[n]$ es real, el teorema se convierte en

$$x[-n] \xleftrightarrow{\mathcal{F}} X^*(e^{j\omega}).$$

1.9.4. Diferenciación en frecuencia

Si

$$x[n] \xleftrightarrow{\mathcal{F}} X(e^{j\omega}),$$

diferenciando la transformada de Fourier, se deduce que

$$nx[n] \xleftrightarrow{\mathcal{F}} j \frac{dX(e^{j\omega})}{d\omega}.$$

1.9.5. Teorema de la convolución

Si

$$x[n] \xleftrightarrow{\mathcal{F}} X(e^{j\omega}) \quad \text{y} \quad h[n] \xleftrightarrow{\mathcal{F}} H(e^{j\omega})$$

y además

$$y[n] = \sum_{k=-\infty}^{\infty} x[k]h[n-k] = x[n] * h[n]$$

se cumple que

$$Y(e^{j\omega}) = X(e^{j\omega})H(e^{j\omega}).$$

Este teorema indica que la transformada de Fourier transforma la operación de convolución en el tiempo en la operación de producto en la frecuencia.

Para la prueba, se parte aplicando la transformada de Fourier a $y[n]$, obteniendo que

$$Y(e^{j\omega}) = \sum_{n=-\infty}^{\infty} \left(\sum_{k=-\infty}^{\infty} x[k]h[n-k] \right) e^{-j\omega n}.$$

Intercambiando el orden de las sumas y observando que $x[k]$ no depende de n , se obtiene que

$$\begin{aligned} Y(e^{j\omega}) &= \sum_{k=-\infty}^{\infty} x[k] \left(\sum_{n=-\infty}^{\infty} h[n-k] e^{-j\omega n} \right) \\ &= \sum_{k=-\infty}^{\infty} x[k] \mathcal{F}\{h[n-k]\} \\ &= \sum_{k=-\infty}^{\infty} x[k] H(e^{j\omega}) e^{-j\omega k} \\ &= H(e^{j\omega}) \sum_{k=-\infty}^{\infty} x[k] e^{-j\omega k} \\ &= H(e^{j\omega}) X(e^{j\omega}), \end{aligned}$$

que es lo que se quería probar.

1.9.6. Teorema de enventanado o modulación

Si

$$x[n] \xleftrightarrow{\mathcal{F}} X(e^{j\omega}) \quad \text{y} \quad w[n] \xleftrightarrow{\mathcal{F}} W(e^{j\omega})$$

y además

$$y[n] = x[n]w[n]$$

se cumple que

$$Y(e^{j\omega}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} X(e^{j\theta}) W(e^{j(\omega-\theta)}) d\theta$$

$Y(e^{j\omega})$ es la convolución periódica entre $X(e^{j\omega})$ y $W(e^{j\omega})$. Este teorema indica que la transformada de Fourier convierte la operación de multiplicación en el tiempo en la operación de convolución periódica en la frecuencia.

Para la prueba, puede aplicarse la transformada de Fourier a $y[n]$,

$$\begin{aligned} Y(e^{j\omega}) &= \sum_{n=-\infty}^{\infty} x[n]w[n]e^{-j\omega n} \\ &= \sum_{n=-\infty}^{\infty} \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} X(e^{j\theta}) e^{j\theta n} d\theta \right) w[n]e^{-j\omega n} \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} X(e^{j\theta}) \left(\sum_{n=-\infty}^{\infty} w[n]e^{-j(\omega-\theta)n} \right) d\theta \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} X(e^{j\theta}) W(e^{j(\omega-\theta)}) d\theta, \end{aligned}$$

que es lo que se quería probar. Esta prueba es la incluida en la sección 4.4.2 de [2].

1.9.7. Teorema de Parseval

Si

$$x[n] \xleftrightarrow{\mathcal{F}} X(e^{j\omega}),$$

se cumple que

$$E_x = \sum_{n=-\infty}^{\infty} |x[n]|^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} |X(e^{j\omega})|^2 d\omega.$$

La función $|X(e^{j\omega})|^2$ se llama *densidad espectral de potencia*, ya que determina como se distribuye la energía en el dominio de la frecuencia. Una forma mas general del teorema de Parseval es

$$\sum_{n=-\infty}^{\infty} x[n]y^*[n] = \frac{1}{2\pi} \int_{-\pi}^{\pi} X(e^{j\omega})Y^*(e^{j\omega}) d\omega. \quad (1.40)$$

La siguiente prueba del teorema de Parseval es la dada en la sección 4.4.2 de [2]. Se parte reemplazando $X(e^{j\omega})$ por su definición en el lado derecho de la ecuación 1.40,

$$\begin{aligned} \frac{1}{2\pi} \int_{-\pi}^{\pi} X(e^{j\omega})Y^*(e^{j\omega}) d\omega &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \left(\sum_{n=-\infty}^{\infty} x[n]e^{-j\omega n} \right) Y^*(e^{j\omega}) d\omega \\ &= \sum_{n=-\infty}^{\infty} x[n] \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} Y^*(e^{j\omega}) e^{-j\omega n} d\omega \right) \\ &= \sum_{n=-\infty}^{\infty} x[n] \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} Y(e^{j\omega}) e^{j\omega n} d\omega \right)^* \\ &= \sum_{n=-\infty}^{\infty} x[n]y^*[n]. \end{aligned}$$

Otra forma de probar el teorema de Parseval es empleando el teorema de la convolución presentado en la sección 1.9.5. Teniendo en cuenta que $y^*[-n] \xleftrightarrow{\mathcal{F}} Y^*(e^{j\omega})$ (ver la sección 1.8), por el teorema de la convolución,

$$x[n] * y^*[-n] = \sum_{k=-\infty}^{\infty} x[k]y^*[-(n-k)] \xleftrightarrow{\mathcal{F}} X(e^{j\omega})Y^*(e^{j\omega}).$$

Aplicando la transformada inversa de Fourier resulta en que

$$\sum_{k=-\infty}^{\infty} x[k]y^*[-(n-k)] = \frac{1}{2\pi} \int_{-\pi}^{\pi} X(e^{j\omega})Y^*(e^{j\omega})e^{j\omega n} d\omega.$$

Finalmente, evaluando en $n = 0$ se concluye que

$$\sum_{k=-\infty}^{\infty} x[k]y^*[k] = \frac{1}{2\pi} \int_{-\pi}^{\pi} X(e^{j\omega})Y^*(e^{j\omega}) d\omega.$$

Ejemplo: transformada de Fourier de la secuencia escalón Se calculará la transformada de Fourier de la secuencia escalón unidad $u[n]$, dada por la ecuación 1.3, para ilustrar el uso de algunas propiedades. Observar que la secuencia escalón no es absolutamente sumable y por lo tanto, la suma 1.35 de la transformada de Fourier puede no converger para todo ω , como se explicó en la sección 1.6. Para el cálculo de la transformada se expresará la secuencia escalón como la suma de una secuencia par y una secuencia impar (ver la sección 1.8),

$$u[n] = u_e[n] + u_o[n],$$

donde

$$u_e[n] = \frac{1}{2}(u[n] + u[-n]) = \frac{1}{2} + \frac{1}{2}\delta[n] \quad \text{y} \quad u_o[n] = \frac{1}{2}(u[n] - u[-n]) = u[n] - \frac{1}{2} - \frac{1}{2}\delta[n],$$

como se muestra en la figura 1.12. Por la propiedad de linealidad de la transformada de Fourier, se

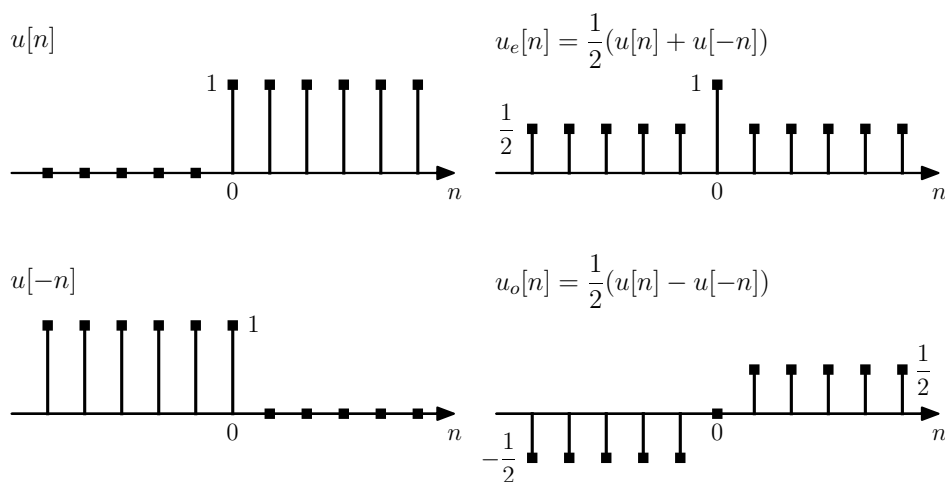


Figura 1.12: Representación de la función escalón $u[n]$ como la suma de la función par $u_e[n]$ y la función impar $u_o[n]$.

cumple que

$$U(e^{j\omega}) = U_e(e^{j\omega}) + U_o(e^{j\omega}), \quad (1.41)$$

donde

$$u[n] \xleftrightarrow{\mathcal{F}} U(e^{j\omega}), \quad u_e[n] \xleftrightarrow{\mathcal{F}} U_e(e^{j\omega}), \quad u_o[n] \xleftrightarrow{\mathcal{F}} U_o(e^{j\omega}).$$

Por un lado, del par de transformadas de una constante dada por la ecuación 1.38, la transformada de Fourier de $u_e[n]$ es

$$U_e(e^{j\omega}) = \pi\delta(\omega) + \frac{1}{2} \quad \text{para} \quad -\pi < \omega < \pi. \quad (1.42)$$

Para calcular la transformada de Fourier de $u_o[n]$ se observa primero que

$$u_o[n-1] = u[n-1] - \frac{1}{2} - \frac{1}{2}\delta[n-1],$$

y por lo tanto,

$$u_o[n] - u_o[n-1] = \underbrace{u[n] - u[n-1]}_{\delta[n]} - \frac{1}{2}(\delta[n] - \delta[n-1]) = \frac{1}{2}(\delta[n] + \delta[n-1]).$$

Aplicando la transformada de Fourier en ambos lados de esta igualdad, considerando la propiedad de desplazamiento temporal de la DTFT en la sección 1.9.2, se deduce que

$$U_o(e^{j\omega}) - U_o(e^{j\omega})e^{-j\omega} = \frac{1 + e^{-j\omega}}{2},$$

es decir,

$$U_o(e^{j\omega}) = \frac{1 + e^{-j\omega}}{2(1 - e^{-j\omega})}. \quad (1.43)$$

Combinando los resultados de las ecuaciones 1.41, 1.42 y 1.43 resulta en que

$$U(e^{j\omega}) = \pi\delta(\omega) + \frac{1}{2} \left(1 + \frac{1 + e^{-j\omega}}{1 - e^{-j\omega}} \right) = \pi\delta(\omega) + \frac{1}{1 - e^{-j\omega}}.$$

Se concluye que

$$u[n] \xleftrightarrow{\mathcal{F}} \pi\delta(\omega) + \frac{1}{1 - e^{-j\omega}} \quad \text{para} \quad -\pi < \omega < \pi. \quad (1.44)$$

1.10. Señales aleatorias en tiempo discreto

Esta sección consiste en un resumen de conceptos básicos sobre señales aleatorias en tiempo discreto.

Una señal aleatoria se considera como un miembro de un conjunto de señales en tiempo discreto caracterizado por un conjunto de funciones de densidad de probabilidad. Cada muestra individual $x[n]$ de una señal particular se asume ser la realización de una variable aleatoria \mathbf{x}_n subyacente. La señal completa es representada por una colección de dichas variables aleatorias, una por cada muestra $-\infty < n < \infty$. Esta colección de variables aleatorias es referida como *proceso aleatorio*, y se asume que una secuencia de muestras particular $x[n]$ para $-\infty < n < \infty$ ha sido generada por el proceso aleatorio subyacente a la señal. Para describir completamente al proceso aleatorio se necesita especificar la distribuciones de probabilidad individuales y conjuntas de todas las variables aleatorias.

Considérese un sistema LTI estable con respuesta al impulso $h[n]$ real, y sea $x[n]$ la realización de un proceso aleatorio en tiempo discreto estacionario en sentido amplio. La salida del sistema lineal también es la realización de un proceso aleatorio vinculado a la salida mediante la transformación

$$y[n] = \sum_{k=-\infty}^{\infty} h[n-k]x[k] = \sum_{k=-\infty}^{\infty} h[k]x[n-k].$$

La media de los procesos de entrada y salida son respectivamente

$$m_x[n] = E\{x[n]\} \quad \text{y} \quad m_y[n] = E\{y[n]\}.$$

Como $x[n]$ es estacionario,¹ $m_x[n]$ es independiente de n y será denotado como m_x . La media de la salida es

$$m_y[n] = E\{y[n]\} = \sum_{k=-\infty}^{\infty} h[k]E\{x[n-k]\} = m_x \sum_{k=-\infty}^{\infty} h[k],$$

también constante. Observar que la media del proceso de salida puede expresarse en función de la respuesta en frecuencia del sistema como

$$m_y = H(e^{j0})m_x.$$

La función de autocorrelación de la salida es

$$\begin{aligned} \phi_{yy}[n, n+m] &= E\{y[n]y[n+m]\} \\ &= E\left\{\sum_{k=-\infty}^{\infty} h[k]x[n-k] \sum_{r=-\infty}^{\infty} h[r]x[n+m-r]\right\} \\ &= \sum_{k=-\infty}^{\infty} h[k] \sum_{r=-\infty}^{\infty} h[r]E\{x[n-k]x[n+m-r]\}. \end{aligned}$$

Considerando que como $x[n]$ es estacionario, se cumple que $E\{x[n-k]x[n+m-r]\} = \phi_{xx}[m+k-r]$ depende solo de la diferencia temporal $m+k-r$, y por o tanto

$$\phi_{yy}[n, n+m] = \sum_{k=-\infty}^{\infty} h[k] \sum_{r=-\infty}^{\infty} h[r]\phi_{xx}[m+k-r] = \phi_{yy}[m], \quad (1.45)$$

es decir, la autocorrelación de la salida también depende solo de la diferencia temporal m . Se concluye que la salida de un sistema LTI cuando la entrada es un proceso estacionario en sentido amplio es también un proceso estacionario en sentido amplio.

Realizando la sustitución $l = r - k$, la ecuación 1.45 puede escribirse como

$$\begin{aligned} \phi_{yy}[m] &= \sum_{k=-\infty}^{\infty} h[k] \sum_{l=-\infty}^{\infty} h[l+k]\phi_{xx}[m-l] \\ &= \sum_{l=-\infty}^{\infty} \phi_{xx}[m-l] \sum_{k=-\infty}^{\infty} h[k]h[l+k] \\ &= \sum_{l=-\infty}^{\infty} \phi_{xx}[m-l]c_{hh}[l], \end{aligned} \quad (1.46)$$

donde se definió

$$c_{hh}[l] = \sum_{k=-\infty}^{\infty} h[k]h[l+k] \stackrel{(a)}{=} \sum_{m=-\infty}^{\infty} h[m-l]h[m] = h[l] * h[-l], \quad (1.47)$$

donde en (a) se realizó el cambio de variable $m = l + k$. La secuencia $c_{hh}[l]$ es referida como *secuencia de autocorrelación determinística* de $h[n]$. La ecuación 1.46 indica que la autocorrelación de la salida de un sistema lineal es la convolución entre la autocorrelación de la entrada y la autocorrelación determinística de la respuesta al impulso del sistema.

Si $\Phi_{xx}(e^{j\omega})$, $\Phi_{yy}(e^{j\omega})$ y $C_{hh}(e^{j\omega})$ son las transformadas de Fourier de $\phi_{xx}[m]$, $\phi_{yy}[m]$ y $c_{hh}[l]$ respectivamente, al tomar la transformada de Fourier de la ecuación 1.46, se obtiene que

$$\Phi_{yy}(e^{j\omega}) = C_{hh}(e^{j\omega})\Phi_{xx}(e^{j\omega}),$$

y al tomar la transformada de Fourier en la ecuación 1.47, resulta en que

$$C_{hh}(e^{j\omega}) = H(e^{j\omega})H^*(e^{j\omega}) = |H(e^{j\omega})|^2,$$

¹De aquí en mas un proceso estacionario en sentido amplio será referido simplemente como estacionario.

y por lo tanto,

$$\Phi_{yy}(e^{j\omega}) = |H(e^{j\omega})|^2 \Phi_{xx}(e^{j\omega}),$$

Esta ecuación brinda la motivación del término *densidad espectral de potencia*. Específicamente,

$$E\{y^2[n]\} = \phi_{yy}[0] = \frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_{yy}(e^{j\omega}) d\omega,$$

que es la potencia promedio de la salida. Puede probarse que la función densidad espectral de potencia de una señal real es una función par, real y no negativa.

Otro resultado importante es sobre la correlación cruzada entre la entrada y la salida de un sistema LTI,

$$\begin{aligned} \phi_{yx}[m] &= E\{x[n]y[n+m]\} \\ &= E\left\{x[n] \sum_{k=-\infty}^{\infty} h[k]x[n+m-k]\right\} \\ &= \sum_{k=-\infty}^{\infty} h[k]E\{x[n]x[n+m-k]\} \\ &= \sum_{k=-\infty}^{\infty} h[k]\phi_{xx}[m-k]. \end{aligned}$$

Se observa que la correlación cruzada entre la entrada y la salida es la convolución entre la respuesta al impulso y la autocorrelación de la entrada. Aplicando la transformada de Fourier se obtiene que

$$\Phi_{yx}(e^{j\omega}) = H(e^{j\omega})\Phi_{xx}(e^{j\omega}).$$

Capítulo 2

La transformada z

2.1. La transformada z

La transformada z de una secuencia $x[n]$ se define como

$$X(z) = \sum_{n=-\infty}^{\infty} x[n]z^{-n}. \quad (2.1)$$

Comparando esta definición con la definición 1.35 de la transformada de Fourier es evidente que hay una relación cercana entre ambas transformadas. En particular, si se reemplaza la variable compleja z en la ecuación 2.1 por el valor complejo $e^{j\omega}$, la transformada z se reduce a la transformada de Fourier. De hecho, esa es la motivación para la notación $X(e^{j\omega})$ de la transformada de Fourier. Mas en general, expresando z en coordenadas polares como

$$z = re^{j\omega},$$

la ecuación 2.1 queda

$$X(re^{j\omega}) = \sum_{n=-\infty}^{\infty} x[n](re^{-j\omega})^{-n},$$

o

$$X(re^{j\omega}) = \sum_{n=-\infty}^{\infty} (x[n]r^{-n})e^{-j\omega n}, \quad (2.2)$$

que es la transformada de Fourier de la secuencia $x[n]r^{-n}$. Con $r = 1$, la ecuación 2.2 se reduce a la transformada de Fourier de $x[n]$. En el plano complejo z , el contorno $z = e^{j\omega}$ o $|z| = 1$ es una circunferencia de radio unidad y es referida como *circunferencia unidad*. La transformada z evaluada en la circunferencia unidad corresponde a la transformada de Fourier.

Como se discutió en el capítulo 1, la representación en series de potencias de la transformada de Fourier no converge para todas las secuencias. De forma similar, la transformada z no converge para todas las secuencias o para todos los valores de z . Para una secuencia dada, el conjunto de valores de z para los cuales la serie de potencias de la transformada z converge se llama *región de convergencia* (*ROC*, region of convergence) de la transformada z . Como se estableció en la sección 1.7, si la secuencia es absolutamente sumable, la transformada de Fourier converge a una función continua de ω . Este criterio aplicado en la ecuación 2.2 conduce a la condición

$$|X(re^{j\omega})| \leq \sum_{n=-\infty}^{\infty} |x[n]r^{-n}| < \infty$$

para la convergencia de la transformada z . Debido a la multiplicación de la secuencia por la exponencial real r^{-n} , es posible que la transformada z converja incluso aunque la transformada de Fourier ($r = 1$) no lo haga. Por ejemplo, la secuencia $u[n]$ no es absolutamente sumable y por lo tanto, la transformada de Fourier no converge absolutamente (ver el ejemplo al final de la sección 1.9). Sin embargo, la secuencia

$u[n]r^{-n}$ es absolutamente sumable si $r > 1$, lo que significa que la transformada z existe con ROC $r = |z| > 1$.

La convergencia de la serie de potencias 2.1 depende únicamente de $|z|$, ya que $|X(z)| < \infty$ si

$$\sum_{n=-\infty}^{\infty} |x[n]| |z|^{-n} < \infty,$$

es decir, la ROC consiste en todos los valores de z en donde esta ecuación se cumple. Por lo tanto, si cierto valor z_1 pertenece a la ROC, todos los valores de z definidos en el círculo $|z| = |z_1|$ pertenecen a la ROC. Una consecuencia de esto es que la ROC consiste en un anillo en el plano z centrado en el origen. Su frontera exterior será un círculo o puede extenderse por el exterior hasta el infinito, y su frontera interior será un círculo o puede extenderse hacia el interior hasta contener el origen. Si la ROC incluye la circunferencia unidad, es decir, si la transformada z converge para $|z| = 1$, la transformada de Fourier de la secuencia converge. Por el contrario, si la ROC no incluye la circunferencia unidad, la transformada de Fourier de la secuencia no converge absolutamente.

Una serie de potencias de la forma de la ecuación 2.1 es una serie de Laurent. Por lo tanto, existen una cantidad de teoremas elegantes y poderosos de la teoría de funciones de variable compleja que pueden emplearse en el estudio de la transformada z (ver [3], por ejemplo). Por ejemplo, una serie de Laurent, y por lo tanto una transformada z , representa una función analítica en cada punto de la ROC. En consecuencia, la transformada z y todas sus derivadas deben ser funciones continuas en z dentro de la ROC. Esto implica que si la ROC incluye la circunferencia unidad, la transformada de Fourier y todas sus derivadas respecto a ω deben ser funciones continuas de ω . Además, de la discusión de la sección 1.7, la secuencia debe ser absolutamente sumable o estable.

La transformada z es especialmente útil cuando la suma infinita puede expresarse de forma cerrada. Entre las transformadas z más importantes y útiles están las que $X(z)$ es una función racional dentro de la ROC, es decir,

$$X(z) = \frac{P(z)}{Q(z)},$$

donde $P(z)$ y $Q(z)$ son polinomios en z . En general, los valores de z para los cuales $X(z) = 0$ son los *ceros* de $X(z)$ y los valores de z para los cuales $X(z)$ es infinito son los *polos* de z . En el caso de las funciones racionales, los ceros son las raíces del polinomio del numerador y los polos son las raíces del polinomio del denominador. Para transformadas z racionales existe una relación importante entre la ubicación de los polos de $X(z)$ y la ROC de la transformada z .

Ejemplo: secuencia exponencial de largo finito Se considera la señal

$$x[n] = \begin{cases} a^n, & 0 \leq n \leq N-1 \\ 0, & \text{en otro caso.} \end{cases}$$

La transformada z es

$$X(z) = \sum_{n=0}^{N-1} a^n z^{-n} = \sum_{n=0}^{N-1} (az^{-1})^n = \frac{1 - (az^{-1})^N}{1 - az^{-1}} = \frac{1}{z^{N-1}} \frac{z^N - a^N}{z - a}.$$

La ROC está determinada por los valores de z para los cuales

$$\sum_{n=0}^{N-1} |az^{-1}|^n < \infty$$

Como la cantidad de sumandos es finita, la suma es finita siempre que az^{-1} sea finito. Asumiendo que a es finito, az^{-1} es finito salvo en $z = 0$. Por lo tanto, la ROC consiste en el plano complejo completo excepto el origen $z = 0$. Para encontrar los ceros y polos de la transformada z se observa que las raíces del numerador son

$$z^N - a^N = 0 \quad \Rightarrow \quad z^N = a^N e^{j2\pi k} \quad \Rightarrow \quad z_k = a e^{j2\pi k/N}, \quad k = 0, 1, \dots, N-1.$$

El cero en $z = a$ correspondiente a $k = 0$ se cancela con el polo en $z = a$. En consecuencia, los únicos polos son los $N - 1$ en el origen y los $N - 1$ ceros están en

$$z_k = ae^{j2\pi k/N}, \quad k = 1, \dots, N - 1.$$

En la figura 2.1 se muestra el diagrama de polos y ceros de este ejemplo para el caso en que $N = 16$ y $0 < a < 1$ real.

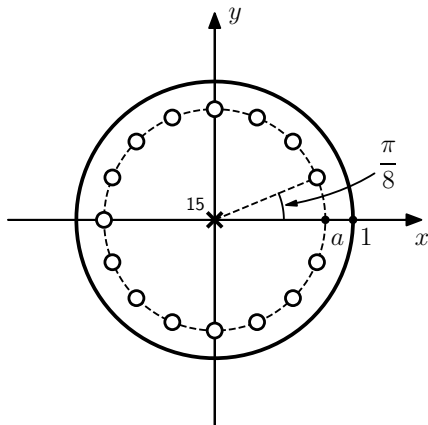


Figura 2.1: Diagrama de polos y ceros en el plano complejo $z = x + jy$ para $N = 16$ y $0 < a < 1$ real. La ROC consiste en todo el plano complejo excepto $z = 0$.

2.2. Propiedades de la ROC de la transformada z

En esta sección se resumen las propiedades de la transformada z .

1. La ROC es un anillo o un disco centrado en el origen, es decir,

$$0 \leq r_R \leq |z| \leq r_L \leq \infty.$$

2. La transformada de Fourier converge absolutamente si y solo si la ROC incluye la circunferencia unidad.
3. La ROC no puede contener ningún polo.
4. Si $x[n]$ es una *secuencia de duración finita* la ROC es todo el plano complejo excepto tal vez $z = 0$ o $z = \infty$.
5. Si $x[n]$ es una *secuencia hacia adelante*, la ROC se extiende desde el polo de mayor magnitud hasta infinito.
6. Si $x[n]$ es una *secuencia hacia atrás*, la ROC se extiende desde el polo de menor magnitud hasta el origen del plano complejo.
7. Si $x[n]$ es una *secuencia hacia ambos lados*, la ROC consiste en un anillo, cuyas fronteras interior y exterior están acotadas por polos.
8. La ROC debe ser una región conexa.

Para especificar completamente a la transformada z de una secuencia no alcanza únicamente con la expresión algebraica o el diagrama de polos y ceros, se necesita además especificar la ROC. Las propiedades consideradas en esta sección limitan las posibles ROC asociadas a determinado patrón de polos y ceros. Para ilustrar esto, considérese el diagrama de polos y ceros mostrado en la figura 2.2. De las propiedades 1, 3 y 8, se deduce que hay solo cuatro posibilidades de ROC, las cuales se muestran en la figura. Si se asume que la circunferencia unidad está entre los polos en $z = b$ y $z = c$, es decir, $b < 1 < c$, la secuencia correspondiente a la ROC que consiste en el anillo $b < |z| < c$ es la única cuya transformada de Fourier converge.

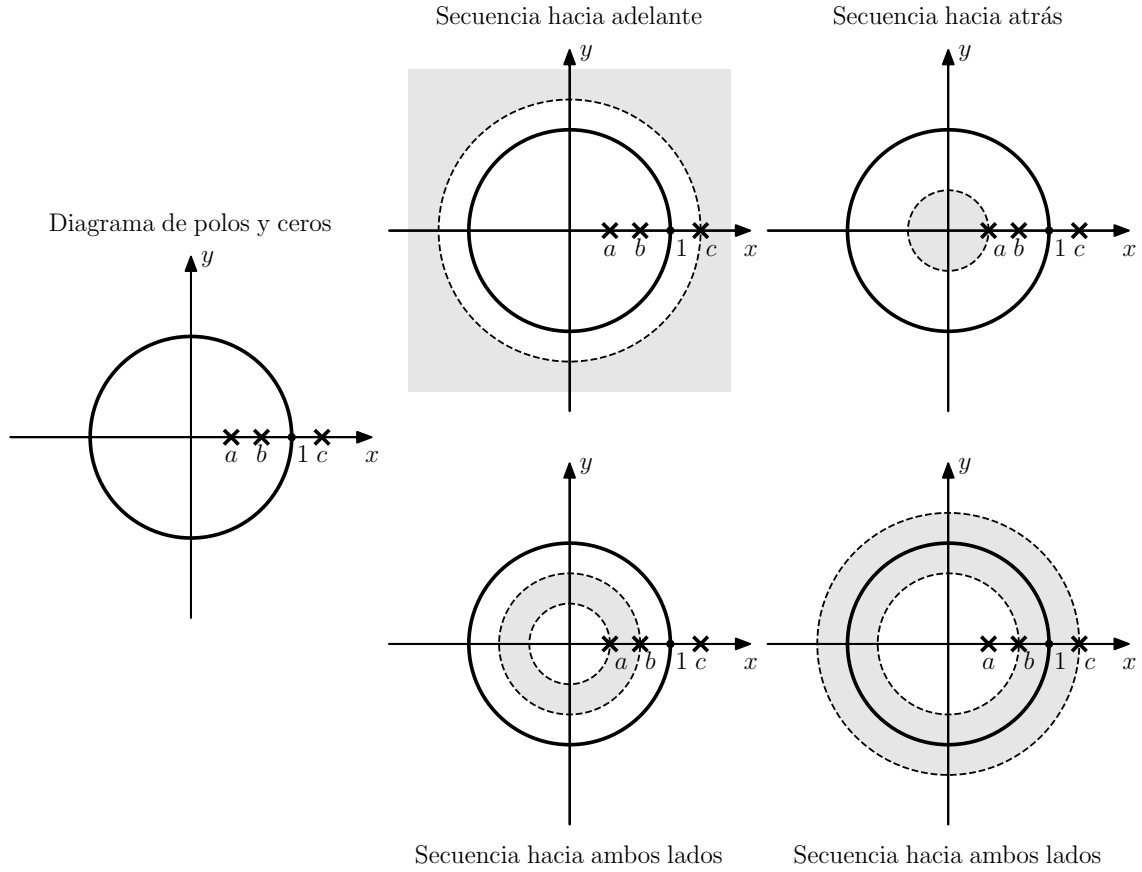


Figura 2.2: Diagrama de polos y ceros de una transformada z y las posibles ROC. Una corresponde a una secuencia hacia adelante, otra a una secuencia hacia atrás y dos a secuencias hacia ambos lados distintas.

2.3. La transformada z inversa

En el empleo de la transformada z para el análisis de secuencias en tiempo discreto debe ser posible el pasaje entre las representaciones en el dominio del tiempo y en el dominio z . A menudo, este análisis implica calcular la transformada z y luego de algo de manipulación de las expresiones algebraicas, calcular la transformada z inversa. La transformada z inversa es la integral de contorno

$$x[n] = \frac{1}{2\pi j} \oint_C X(z) z^{n-1} dz, \quad (2.3)$$

donde C es un contorno cerrado orientado en sentido antihorario dentro de la ROC que encierra al origen. Esta expresión se deriva del teorema integral de Cauchy de la teoría de funciones de variable compleja. Específicamente, $x[n]$ son los coeficientes de la representación en series de Laurent de la función $X(z)$ (ver la sección 66 de [3] sobre la representación en series de Laurent de una función de variable compleja, y también el ejercicio 8 de la sección 68). Sin embargo, para el tipo de secuencias y transformadas z encontradas en el análisis de sistemas LTI, procedimientos menos formales son suficientes y preferibles para la evaluación de la ecuación 2.3. En las secciones siguientes, además de la evaluación directa de la integral de contorno 2.3, se consideran dichos procedimientos, en particular, el método de inspección, la expansión en fracciones simples y la expansión en series de potencia.

2.3.1. Evaluación directa de la integral de contorno

En ciertas condiciones, para evaluar la integral 2.3 puede emplearse el *teorema de los residuos de Cauchy* (ver la sección 76 de [3]), que indica que si C un contorno cerrado simple orientado en sentido antihorario y f es una función analítica dentro de C excepto por un número finito de puntos singulares aislados z_k con $k = 1, 2, \dots, n$ dentro de C , entonces

$$\oint_C f(z) dz = 2\pi j \sum_{k=1}^n \text{Res}_{z=z_k} f(z). \quad (2.4)$$

El teorema indica que la integral de contorno es $2\pi j$ veces la suma de los residuos de f en los puntos singulares contenidos dentro del contorno C .

Se considerará el caso en que los todos los puntos singulares de f interiores a C son polos. Una forma conveniente de calcular los residuos en los polos es considerando que si z_0 es un polo de orden m de f , con $m = 1, 2, \dots$, $f(z)$ se puede escribir de la forma

$$f(z) = \frac{\phi(z)}{(z - z_0)^m} \quad \text{con} \quad m = 1, 2, \dots,$$

donde $\phi(z)$ es analítica y no nula en z_0 , y el residuo de f en z_0 es

$$\text{Res}_{z=z_0} f(z) = \frac{\phi^{(m-1)}(z_0)}{(m-1)!} \quad \text{cuando } m = 1, 2, \dots, \quad (2.5)$$

con las convenciones $\phi^{(0)}(z_0) = \phi(z_0)$ y $0! = 1$ (ver la sección 80 de [3]).

Ejemplo: antitransformada z correspondiente a la secuencia exponencial Se evaluará la transformada inversa de

$$X(z) = \frac{1}{1 - az^{-1}},$$

que tiene un polo en $z = a$, y en consecuencia, las posibles ROC son $|z| > |a|$ y $|z| < |a|$. En este caso, el integrando en la ecuación 2.3 es

$$f(z) = \frac{z^{n-1}}{1 - az^{-1}} = \frac{z^n}{z - a}.$$

Se considera primero el caso con ROC $|z| > |a|$. En consecuencia, el contorno de integración en la ecuación 2.3 es un círculo centrado en el origen con radio mayor que $|a|$. Se distinguen los siguientes dos casos:

1. $n \geq 0$: en este caso, el único polo de f interior a C es $z = a$. Expresando

$$f(z) = \frac{\phi(z)}{z - a} \quad \text{donde} \quad \phi(z) = z^n.$$

Como $\phi(z)$ es analítica en $z = a$ y $\phi(a) = a^n \neq 0$, de la ecuación 2.5 se cumple que

$$\text{Res}_{z=a} f(z) = \phi(a) = a^n,$$

concluyendo que para $n > 0$,

$$x[n] = a^n$$

2. $n < 0$: definiendo $m = -n > 0$, el integrando es

$$f(z) = \frac{1}{z^m(z - a)},$$

y ahora, además del polo en $z = a$, también hay un polo de orden m en $z = 0$, y por lo tanto, la integral de contorno es la suma de los residuos en ambos polos. Al igual que antes, para calcular el residuo en $z = a$, se expresa

$$f(z) = \frac{\phi(z)}{z - a} \quad \text{donde} \quad \phi(z) = \frac{1}{z^m}.$$

Como $\phi(z)$ es analítica en $z = a$ y $\phi(a) = 1/a^m \neq 0$, de la ecuación 2.5 se cumple que

$$\operatorname{Res}_{z=a} f(z) = \phi(a) = \frac{1}{a^m}.$$

Para calcular el residuo en el polo de orden m en $z = 0$, se expresa a f como

$$f(z) = \frac{\phi(z)}{z^m} \quad \text{donde} \quad \phi(z) = \frac{1}{z-a}, \quad \text{para } m = 1, 2, \dots$$

Para calcular el residuo mediante la ecuación 2.5, hay que calcular las derivadas $\phi^{(m-1)}(z)$ para $m = 1, 2, \dots$. Observando que

$$\phi'(z) = -\frac{1}{(z-a)^2}, \quad \phi''(z) = \frac{2}{(z-a)^3}, \quad \phi'''(z) = -\frac{3 \times 2}{(z-a)^4},$$

y así sucesivamente, el resultado se puede generalizar como

$$\phi^{(m-1)}(z) = \frac{(-1)^{m-1}(m-1)!}{(z-a)^m}.$$

Luego,

$$\phi^{(m-1)}(0) = \frac{(-1)^{m-1}(m-1)!}{(-1)^m a^m} = -\frac{(m-1)!}{a^m},$$

resultando en que

$$\operatorname{Res}_{z=0} f(z) = -\frac{1}{a^m}, \quad \text{para } m = 1, 2, \dots$$

Finalmente,

$$x[-m] = \operatorname{Res}_{z=a} f(z) + \operatorname{Res}_{z=0} f(z) = \frac{1}{a^m} - \frac{1}{a^m} = 0, \quad \text{para } m = 1, 2, \dots$$

Se concluye que

$$x[n] = a^n u[n]$$

cuando $|z| > |a|$. En el caso en que la ROC es $|z| < |a|$, el contorno de integración C es un circunferencia centrada en el origen de radio menor a $|a|$ y se consideran nuevamente los dos siguientes casos:

1. $n \geq 0$: en este caso, el integrando f no tiene polos dentro C y por lo tanto, $x[n] = 0$.
2. $n < 0$: nuevamente definiendo $m = -n > 0$, el integrando es

$$f(z) = \frac{1}{z^m(z-a)},$$

y solo el polo de orden m en $z = 0$ es interior a C . El residuo en dicho polo ya fue calculado antes y es

$$\operatorname{Res}_{z=0} f(z) = -\frac{1}{a^m}, \quad \text{para } m = 1, 2, \dots$$

Por lo tanto,

$$x[-m] = \operatorname{Res}_{z=0} f(z) = -\frac{1}{a^m}, \quad \text{para } m = 1, 2, \dots$$

o

$$x[n] = -a^n, \quad \text{para } n = -1, -2, \dots$$

Se concluye que

$$x[n] = -a^n u[-n-1]$$

cuando $|z| < |a|$.

Observar que el factor z^{n-1} en el integrando de la transformada inversa 2.3 introduce un polo de multiplicidad ($m = -n + 1$) en $z = 0$ para $n = -1, -2, \dots$. Por lo tanto, como indica la ecuación 2.5, el cálculo de residuo en $z = 0$ requiere el cálculo de la derivada m -ésima de $\phi(z)$, o equivalentemente, de $z^m f(z)$ para todo valor de m . Dicho cálculo fue posible en el ejemplo precedente debido a que la transformada z a invertir es una función sencilla, pero en otros casos puede no ser viable. Sin embargo, si la ROC es $|z| > |a|$ correspondiente a una secuencia hacia adelante, ya se sabe de antemano que los valores de la secuencia para n negativos son nulos, por lo que el cálculo de la derivada no es necesario.

2.3.2. Método de inspección

El método de inspección consiste reconocer mediante inspección pares de transformadas mediante el uso de una tabla de pares de transformadas básicos, como la que se muestra en el cuadro 2.1. Por

Secuencia, $x[n]$	Transformada Z , $X(z)$	ROC
$\delta[n]$	1	$\forall z$
$\delta[n - n_0]$	z^{-n_0}	$z \neq 0$ o $z \neq \infty$
$a^n u[n]$	$\frac{1}{1 - az^{-1}}$	$ z > a $
$-a^n u[-n - 1]$	$\frac{1}{1 - az^{-1}}$	$ z < a $
$nu[n]$	$\frac{z^{-1}}{(1 - z^{-1})^2}$	$ z > 1$
$-nu[-n - 1]$	$\frac{z^{-1}}{(1 - z^{-1})^2}$	$ z < 1$
$n^2 u[n]$	$\frac{z^{-1}(1 + z^{-1})}{(1 - z^{-1})^3}$	$ z > 1$
$-n^2 u[-n - 1]$	$\frac{z^{-1}(1 + z^{-1})}{(1 - z^{-1})^3}$	$ z < 1$
$\cos(\omega_0 n)u[n]$	$\frac{1 - z^{-1} \cos(\omega_0)}{1 - 2z^{-1} \cos(\omega_0) + z^{-2}}$	$ z > 1$
$\sin(\omega_0 n)u[n]$	$\frac{z^{-1} \sin(\omega_0)}{1 - 2z^{-1} \cos(\omega_0) + z^{-2}}$	$ z > 1$
$a^n \cos(\omega_0 n)u[n]$	$\frac{1 - az^{-1} \cos(\omega_0)}{1 - 2az^{-1} \cos(\omega_0) + a^2 z^{-2}}$	$ z > a $
$a^n \sin(\omega_0 n)u[n]$	$\frac{az^{-1} \sin(\omega_0)}{1 - 2az^{-1} \cos(\omega_0) + a^2 z^{-2}}$	$ z > a $

Cuadro 2.1: Pares de transformadas z básicos.

ejemplo, si se necesita encontrar la transformada z inversa de

$$X(z) = \frac{1}{1 - \frac{1}{2}z^{-1}}, \quad |z| > \frac{1}{2},$$

del cuadro se observa que

$$a^n u[n] \xleftrightarrow{\mathcal{Z}} \frac{1}{1 - az^{-1}}, \quad |z| > a,$$

y por lo tanto,

$$x[n] = \left(\frac{1}{2}\right)^n u[n].$$

Si una transformada z dada puede expresarse como suma de términos que se encuentran en una tabla de transformadas, su transformada inversa puede encontrarse solo empleando la tabla. Esto está vinculado al método de fracciones simples descrito en la siguiente sección.

2.3.3. Método de fracciones simples

Como se describió, las transformadas z pueden encontrarse mediante inspección si la expresión de la transformada z está tabulada. A veces, $X(z)$ no se encuentra explícitamente en una tabla, pero puede

descomponerse como la suma de términos mas simples, cada uno de los cuales están tabulados. Este es el caso de las funciones racionales, que pueden expandirse en fracciones simples en términos que están tabulados.

Para ver como se obtiene la expansión en fracciones simples se asume que $X(z)$ se puede expresar como un cociente de polinomios en z^{-1} de la siguiente forma:

$$X(z) = \frac{\sum_{k=0}^M b_k z^{-k}}{\sum_{k=0}^N a_k z^{-k}}. \quad (2.6)$$

Una expresión equivalente como cociente de polinomios en z es

$$X(z) = \frac{z^N \sum_{k=0}^M b_k z^{M-k}}{z^M \sum_{k=0}^N a_k z^{N-k}}.$$

Esta expresión muestra que hay explícitamente M ceros y N polos distintos de cero asumiendo que a_0 , b_0 , a_N y b_M son no nulos. Además hay $M - N$ polos en $z = 0$ si $M > N$ o $N - M$ ceros en $z = 0$ si $N > M$. Esto indica que en cualquier caso, las transformadas z de la forma de la ecuación 2.6 tienen la misma cantidad de polos y ceros en el plano complejo finito, y no hay polos o ceros en $z = \infty$. Para obtener la expansión en fracciones simples de $X(z)$ de la ecuación 2.6, es mas conveniente factorizar el numerador y el denominador en monomios,

$$X(z) = \frac{b_0 \prod_{k=1}^M (1 - c_k z^{-1})}{a_0 \prod_{k=1}^N (1 - d_k z^{-1})}, \quad (2.7)$$

donde c_k y d_k son las raíces de numerador y denominador respectivamente. Se distinguen tres casos:

1. $M < N$ y todos los polos son de primer orden. Cuando $M < N$, la función racional se dice que es *propia*. En este caso, $X(z)$ puede expresarse como

$$X(z) = \sum_{k=1}^N \frac{A_k}{1 - d_k z^{-1}}. \quad (2.8)$$

Observar que el numerador que resulta al sumar los términos es a lo sumo de grado $N - 1$ en z^{-1} . Los parámetros A_k pueden calcularse sacando denominador común en los sumandos del lado derecho de la igualdad y luego igualando los coeficientes de los polinomios del numerados en ambos lados de la igualdad. Esto conduce a un sistema de N ecuaciones donde los parámetros A_k son las N incógnitas. Una forma mas directa de obtener los parámetros A_k es multiplicando ambos lados de la ecuación 2.8 por $(1 - d_k z^{-1})$ y luego evaluando en $z = d_k$, es decir,

$$A_k = (1 - d_k z^{-1})X(z) \Big|_{z=d_k}. \quad (2.9)$$

2. $M \geq N$ y todos los polos son de primer orden. Cuando $M \geq N$ la función racional se dice que es *impropia*. En este caso debe sumarse un polinomio de grado $M - N$ en el lado derecho de la ecuación 2.8,

$$X(z) = \sum_{r=0}^{M-N} B_r z^{-r} + \sum_{k=1}^N \frac{A_k}{1 - d_k z^{-1}}. \quad (2.10)$$

Los coeficientes B_r pueden obtenerse mediante división larga del numerador entre el denominador, donde el proceso de división finaliza cuando el resto es de menor grado que el denominador. La división larga se realiza en orden reverso, es decir, se escriben ambos polinomios comenzando por la potencia mas negativa de z a la izquierda (ver la sección 3.4.3 de [2]), como se hace en el ejemplo siguiente. Los coeficientes A_k pueden obtenerse igual que antes mediante la ecuación 2.9.

3. $M \geq N$ y existen polos de orden múltiple. Si $X(z)$ tiene un polo de orden s en $z = d_i$ y el resto de los polos son de primer orden, la ecuación 2.10 se convierte en

$$X(z) = \sum_{r=0}^{M-N} B_r z^{-r} + \sum_{k=1, k \neq i}^N \frac{A_k}{1 - d_k z^{-1}} + \sum_{m=1}^s \frac{C_m}{(1 - d_i z^{-1})^m}. \quad (2.11)$$

Los coeficientes A_k y B_r se obtienen como antes. Los coeficientes C_m se obtienen de la ecuación

$$C_m = \frac{1}{(s-m)!(-d_i)^{s-m}} \left\{ \frac{d^{s-m}}{dw^{s-m}} [(1 - d_i w)^s X(w^{-1})] \right\}_{w=d_i^{-1}}.$$

La ecuación 2.11 es la forma mas general de la expansión en funciones simples de una transformada z racional expresada como función de z^{-1} para el caso en que $M \geq N$ y d_i es un polo múltiple de orden s . Si hay mas de un polo múltiple, la ecuación 2.11 tendrá un término como el tercer sumando por cada polo múltiple. Si el grado M del numerador es menor al grado N del denominador, el término polinómico en la ecuación 2.11 desaparece.

Ejemplo: inversión mediante fracciones simples Considérese la secuencia $x[n]$ con transformada z

$$X(z) = \frac{1 + 2z^{-1} + z^{-2}}{1 - \frac{3}{2}z^{-1} + \frac{1}{2}z^{-2}} = \frac{(1 + z^{-1})^2}{\left(1 - \frac{1}{2}z^{-1}\right)(1 - z^{-1})}, \quad |z| > 1.$$

$X(z)$ tiene polos en $z = 1/2$ y en $z = 1$, y por la definición de la ROC, $x[n]$ es una secuencia hacia adelante. Como $M = N = 2$ y los polos son de primer orden, $X(z)$ puede descomponerse en la forma de la ecuación 2.10, en este caso como

$$X(z) = B_0 + \frac{A_1}{1 - \frac{1}{2}z^{-1}} + \frac{A_2}{1 - z^{-1}}.$$

La constante B_0 puede obtenerse mediante división larga como

$$\begin{array}{r} 2 \\ \hline \frac{1}{2}z^{-2} - \frac{3}{2}z^{-1} + 1 \quad \begin{array}{r} z^{-2} \quad +2z^{-1} \quad +1 \\ z^{-2} \quad -3z^{-1} \quad +2 \\ \hline 5z^{-1} \quad -1 \end{array} \end{array}$$

Como el grado del resto en la variable z^{-1} es de menor grado que el denominador luego de un paso de la división larga, no es necesario continuar con la división. Se obtuvo que

$$X(z) = 2 + \frac{5z^{-1} - 1}{\left(1 - \frac{1}{2}z^{-1}\right)(1 - z^{-1})}.$$

Los coeficientes A_1 y A_2 pueden calcularse mediante la ecuación 2.9 como

$$A_1 = \left[\left(1 - \frac{1}{2}z^{-1}\right) \frac{(1 + z^{-1})^2}{\left(1 - \frac{1}{2}z^{-1}\right)(1 - z^{-1})} \right]_{z=1/2} = \frac{(1+2)^2}{1-2} = -9$$

$$A_2 = \left[(1 - z^{-1}) \frac{(1 + z^{-1})^2}{\left(1 - \frac{1}{2}z^{-1}\right)(1 - z^{-1})} \right]_{z=1} = \frac{(1+1)^2}{1 - \frac{1}{2}} = 8.$$

Se concluye que

$$X(z) = 2 - \frac{9}{1 - \frac{1}{2}z^{-1}} + \frac{8}{1 - z^{-1}},$$

y considerando que la ROC es $|z| > 1$, del cuadro 2.1 resulta en que

$$x[n] = 2\delta[n] - 9\left(\frac{1}{2}\right)^n u[n] + 8u[n].$$

2.3.4. Expansión en series de potencias

La expresión de la definición de la transformada z es una serie de Laurent donde los valores $x[n]$ de la secuencia son los coeficientes de z^{-n} . Por lo tanto, si la transformada z está dada como una serie de potencias de la forma

$$\begin{aligned} X(z) &= \sum_{n=-\infty}^{\infty} x[n]z^{-n} \\ &= \cdots + x[-2]z^2 + x[-1]z + x[0] + x[1]z^{-1} + x[2]z^{-2} + \cdots, \end{aligned} \quad (2.12)$$

se puede determinar cualquier valor de la secuencia encontrando el coeficiente de la serie de Laurent de la potencia apropiada de z^{-1} (ver el capítulo 5 de [3]). Este enfoque ya se empleó en encontrar la parte polinómica de la expansión en fracciones simples cuando $M \geq N$.

Ejemplo: transformada inversa mediante expansión en series de potencias Se considera la transformada z

$$X(z) = \log(1 + az^{-1}) \quad \text{cuando} \quad |z| > |a|.$$

Empleando la expansión en series de Taylor conocida

$$\log z = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} (z-1)^n \quad \text{cuando} \quad |z-1| < 1,$$

(ver el ejercicio 6 de la sección 72 de [3]), al reemplazar z por $1 + az^{-1}$ en esta expresión se obtiene la representación en series de Laurent

$$\log(1 + az^{-1}) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} a^n z^{-n} \quad \text{cuando} \quad |az^{-1}| < 1 \quad \text{o} \quad |z| > |a|.$$

Comparando esta expresión con la definición 2.12 de la transformada z , se deduce que la antitransformada z es

$$x[n] = \begin{cases} (-1)^{n+1} \frac{a^n}{n}, & n \geq 1 \\ 0, & n \leq 0. \end{cases}$$

Ejemplo: antitransformada z correspondiente a la secuencia exponencial Sea la transformada z

$$X(z) = \frac{1}{1 - az^{-1}},$$

que tiene un polo en $z = a$. Se quiere encontrar las secuencias correspondientes a las regiones de convergencia $|z| > |a|$ y $|z| < |a|$. Para hacerlo se calculará la expansión en series de Laurent de $X(z)$. Un resultado conocido de la teoría de análisis complejo es que la siguiente expansión en series de Laurent:

$$\frac{1}{1 - z} = \sum_{n=0}^{\infty} z^n = 1 + z + z^2 + \cdots \quad \text{cuando} \quad |z| < 1, \quad (2.13)$$

identidad mostrada en el ejemplo de la sección 61 y en el ejemplo 1 de la sección 64 de [3]. Reemplazando z por az^{-1} en esta identidad conduce a la expansión en series de Laurent

$$\frac{1}{1 - az^{-1}} = \sum_{n=0}^{\infty} a^n z^{-n} \quad \text{cuando} \quad |az^{-1}| < 1 \quad \text{o} \quad |z| > |a|.$$

Comparando este resultado con la ecuación 2.12, se deduce que la antitransformada z cuando $|z| > |a|$ es

$$x[n] = a^n u[n].$$

Por otro lado, se tiene que

$$X(z) = \frac{1}{1 - az^{-1}} = \frac{a^{-1}z}{a^{-1}z(1 - az^{-1})} = -a^{-1}z \left(\frac{1}{1 - a^{-1}z} \right). \quad (2.14)$$

Además, reemplazando z por $a^{-1}z$ en la series de potencias 2.13 se obtiene la expansión en series de potencias

$$\frac{1}{1 - a^{-1}z} = \sum_{n=0}^{\infty} a^{-n} z^n \quad \text{cuando} \quad |a^{-1}z| < 1 \quad \text{o} \quad |z| < |a|.$$

Combinando ambos resultados, se obtiene la expansión en series de Laurent

$$\frac{1}{1 - az^{-1}} = -a^{-1}z \sum_{n=0}^{\infty} a^{-n} z^n = - \sum_{n=0}^{\infty} a^{-(n+1)} z^{n+1} \stackrel{(a)}{=} - \sum_{m=-\infty}^{-1} a^m z^{-m} \quad \text{cuando} \quad |z| < |a|,$$

donde en (a) se realizó el cambio de variable $m = -(n+1)$. Comparando este resultado con la ecuación 2.12, se obtiene que la antitransformada z cuando $|z| < |a|$ es

$$x[n] = -a^n u[-n-1].$$

Observar que la expansión en series de potencias puede obtenerse mediante división larga (ver la sección 73 de [3]). Efectivamente, con

$$X(z) = \frac{1}{1 - az^{-1}} \quad \text{cuando} \quad |z| > |a|,$$

la división larga conduce a

$$\begin{array}{r} \begin{array}{r} 1 \quad +az^{-1} \quad +a^2z^{-2} \quad +\dots \\ \hline 1 - az^{-1} \end{array} \begin{array}{r} 1 \\ \hline 1 - az^{-1} \\ \hline az^{-1} \\ \hline az^{-1} - a^2z^{-2} \\ \hline a^2z^{-2} \\ \hline a^2z^{-2} - a^3z^{-3} \\ \hline a^3z^{-3} \\ \hline \vdots \end{array} \end{array}$$

es decir,

$$\frac{1}{1 - az^{-1}} = \sum_{n=0}^{\infty} a^n z^{-n} \quad \text{cuando} \quad |z| > |a|.$$

Observar que al realizar la división de polinomios en z^{-1} como se acaba de hacer, se obtuvo un desarrollo en series de potencias en z^{-1} . Para el caso en que la ROC es $|z| < |a|$ hay que expresar los polinomios del numerador y el denominador en potencias de z , como en la ecuación 2.14, y luego dividir. De esta forma se obtiene una expresión en series de potencias en z correspondiente a una secuencia hacia la izquierda. Equivalentemente, cuando la ROC corresponde a una secuencia hacia la izquierda, puede realizarse la división larga escribiendo ambos polinomios en orden reverso, es decir, comenzando con la potencia mas negativa de z en la izquierda (ver la sección 3.4.2 de [2]). En el caso del ejemplo, la división larga sería de la siguiente forma:

$$\begin{array}{r} \begin{array}{cccc} -a^{-1}z & -a^{-2}z^2 & -a^{-3}z^3 & -\dots \end{array} \\ \hline -az^{-1} + 1 \bigg) \begin{array}{cc} 1 & \\ 1 & -a^{-1}z \end{array} \\ \hline \begin{array}{cc} a^{-1}z & \\ a^{-1}z & -a^{-2}z^2 \end{array} \\ \hline \begin{array}{cc} a^{-2}z^2 & \\ a^{-2}z^2 & -a^{-3}z^3 \end{array} \\ \hline \begin{array}{cc} a^{-3}z^3 & \\ & \vdots \end{array} \end{array}$$

En general, el método de división larga no es viable debido a que hay que continuar con la división para todo n , y la obtención de una expresión analítica no es posible, a menos que el patrón sea sencillo y pueda generalizarse, como en el caso de la secuencia exponencial en el ejemplo precedente. Sin embargo, el método es útil si se desea obtener los valores de unas pocas muestras de $x[n]$.

2.4. Propiedades de la transformada z

En la siguiente discusión, $X(z)$ denota la transformada z de $x[n]$ y R_x denota la ROC de $X(z)$.

2.4.1. Linealidad

La propiedad de linealidad establece que

$$ax_1[n] + bx_2[n] \xrightarrow{\mathcal{Z}} aX_1(z) + bX_2(z), \quad \text{ROC contiene } R_{x_1} \cap R_{x_2}.$$

Para el caso de secuencias con transformada z racional, si $aX_1(z) + bX_2(z)$ tiene todos los polos de $X_1(z)$ y $X_2(z)$, es decir, no hay cancelación de polos y ceros, la ROC es la intersección de las ROCs individuales. Si en la combinación lineal se introducen ceros que se cancelan con polos, la ROC puede ser mayor.

2.4.2. Desplazamiento en el tiempo

La propiedad de desplazamiento en el tiempo indica que

$$x[n - n_0] \xrightarrow{\mathcal{Z}} z^{-n_0} X(z), \quad \text{ROC} = R_x \text{ excepto por la adición o el borrado de } z = 0 \text{ o } z = \infty.$$

2.4.3. Multiplicación por una secuencia exponencial

Esta propiedad indica que

$$z_0^n x[n] \xleftrightarrow{\mathcal{Z}} X(z/z_0), \quad \text{ROC} = |z_0| R_x.$$

La notación $\text{ROC} = |z_0| R_x$ significa que la ROC está escalada por el valor $|z_0|$, es decir, si R_x es el conjunto de valores $r_R < |z| < r_L$, $|z_0| R_x$ es el conjunto de valores $|z_0| r_R < |z| < |z_0| r_L$. Como si $X(z)$ tiene un cero o un polo en $z = z_1$, $X(z/z_0)$ tiene un cero o un polo en $z = z_0 z_1$, todas las posiciones de los polos y ceros están escaladas un factor z_0 . Si z_0 es real, el escalamiento puede interpretarse como un encogimiento o expansión del plano complejo. Si z_0 es un número complejo $z_0 = e^{j\omega_0}$ de magnitud unidad, el escalamiento es a una rotación del plano complejo un ángulo ω_0 y corresponde a la propiedad de desplazamiento en frecuencia de la transformada de Fourier indicada en la sección 1.9.2.

2.4.4. Diferenciación de $X(z)$

La propiedad establece que

$$nx[n] \xleftrightarrow{\mathcal{Z}} -z \frac{dX(z)}{dz}, \quad \text{ROC} = R_x.$$

2.4.5. Conjugación de una secuencia compleja

Se cumple que

$$x^*[n] \xleftrightarrow{\mathcal{Z}} X^*(z^*), \quad \text{ROC} = R_x.$$

Una consecuencia de esta propiedad es que si $x[n]$ es real, se cumple que $x[n] = x^*[n]$ y por lo tanto, $X(z) = X^*(z^*)$. En palabras, la transformada z en el semiplano inferior es la conjugada del semiplano superior. Se concluye que la transformada z de una secuencia real tiene polos y ceros reales y pares de polos y ceros complejos conjugados.

2.4.6. Inversión temporal

La propiedad de inversión temporal indica que

$$x^*[-n] \xleftrightarrow{\mathcal{Z}} X^*(1/z^*), \quad \text{ROC} = \frac{1}{R_x}.$$

La notación $\text{ROC} = 1/R_x$ indica que si R_x es el conjunto de valores $r_R < |z| < r_L$, $1/R_x$ es el conjunto de valores $1/r_L < |z| < 1/r_R$.

2.4.7. Convolución de secuencias

La propiedad de convolución indica que

$$x_1[n] * x_2[n] \xleftrightarrow{\mathcal{Z}} X_1(z)X_2(z), \quad \text{ROC contiene } R_{x_1} \cap R_{x_2}$$

Para la prueba, sea

$$y[n] = x_1[n] * x_2[n] = \sum_{k=-\infty}^{\infty} x_1[k]x_2[n-k].$$

Por lo tanto,

$$\begin{aligned}
Y(z) &= \sum_{n=-\infty}^{\infty} y[n]z^{-n} \\
&= \sum_{n=-\infty}^{\infty} \left\{ \sum_{k=-\infty}^{\infty} x_1[k]x_2[n-k] \right\} z^{-n} \\
&\stackrel{(a)}{=} \sum_{k=-\infty}^{\infty} x_1[k] \sum_{n=-\infty}^{\infty} x_2[n-k]z^{-n} \\
&\stackrel{(b)}{=} \sum_{k=-\infty}^{\infty} x_1[k] \sum_{m=-\infty}^{\infty} x_2[m]z^{-(m+k)} \\
&= \sum_{k=-\infty}^{\infty} x_1[k]z^{-k} \sum_{m=-\infty}^{\infty} x_2[m]z^{-m} \\
&= X_1(z)X_2(z),
\end{aligned}$$

donde en (a) se intercambi6 el orden de las sumatorias, que es posible para los valores z en la ROC, y en (b) se realiz6 el cambio de variable $m = n - k$. La ROC incluye la intersecci6n de las ROCs de $X_1(z)$ y $X_2(z)$, pero si un polo en la frontera de la ROC de una de las transformadas z es cancelado por un cero de la otra, la ROC puede ser mayor.

2.4.8. Teorema del valor inicial

Sea $x[n]$ una secuencia causal, es decir, $x[n] = 0$ para $n < 0$. Entonces se cumple que

$$\lim_{z \rightarrow \infty} X(z) = x[0].$$

Para la prueba (ver el problema 3.57 de [1] o la secci6n 3.2 de [2]), se observa que como la secuencia es causal,

$$X(z) = \sum_{n=0}^{\infty} x[n]z^{-n}.$$

Por lo tanto,

$$\lim_{z \rightarrow \infty} X(z) = \lim_{z \rightarrow \infty} \sum_{n=0}^{\infty} x[n]z^{-n} = x[0] + \lim_{z \rightarrow \infty} \sum_{n=1}^{\infty} x[n]z^{-n} = x[0].$$

Este resultado se llama *teorema del valor inicial*.

En el caso en que la secuencia es anticausal, es decir, $x[n] = 0$ para $n > 0$, el resultado equivalente es

$$\lim_{z \rightarrow 0} X(z) = x[0].$$

Efectivamente,

$$\lim_{z \rightarrow 0} X(z) = \lim_{z \rightarrow 0} \sum_{n=-\infty}^0 x[n]z^{-n} = x[0] + \lim_{z \rightarrow 0} \sum_{n=1}^{\infty} x[-n]z^n = x[0].$$

Capítulo 3

Análisis en el dominio transformado de sistemas lineales e invariantes en el tiempo

3.1. Introducción

En este capítulo se desarrolla con mayor detalle la representación y el análisis de sistemas LTI empleando la transformada de Fourier y la transformada z .

Como se discutió en el capítulo 1, un sistema LTI queda completamente caracterizado en el dominio del tiempo por la respuesta al impulso $h[n]$, quedando la salida $y[n]$ correspondiente a una entrada $x[n]$ especificada por la suma de convolución

$$y[n] = \sum_{k=-\infty}^{\infty} x[k]h[n-k].$$

Alternativamente, como la respuesta en frecuencia está directamente vinculada a la respuesta al impulso mediante la transformada de Fourier, la respuesta en frecuencia, si existe, provee igualmente una caracterización completa de los sistemas LTI. Además, la transformada z de la salida de un sistema LTI se relaciona con la transformada z de la entrada y la respuesta al impulso como

$$Y(z) = H(z)X(z), \quad (3.1)$$

donde $Y(z)$, $X(z)$ y $H(z)$ son las transformadas z de $y[n]$, $x[n]$ y $h[n]$ respectivamente, con las regiones de convergencia apropiadas. $H(z)$ es referida usualmente como la *función del sistema* o *función de transferencia*. Como una secuencia y su transformada z consisten en un par único, cualquier sistema LTI queda completamente caracterizado por la función del sistema, nuevamente asumiendo convergencia.

3.2. La respuesta en frecuencia de sistemas LTI

3.2.1. Respuesta en fase y retardo de grupo

La respuesta en frecuencia es en general un número complejo en cada frecuencia. Con la respuesta en frecuencia expresada en forma polar, la magnitud y la fase de las transformadas de Fourier de la entrada y la salida se relacionan como

$$|Y(e^{j\omega})| = |H(e^{j\omega})||X(e^{j\omega})|, \quad (3.2)$$

$$\angle Y(e^{j\omega}) = \angle H(e^{j\omega}) + \angle X(e^{j\omega}), \quad (3.3)$$

donde $|H(e^{j\omega})|$ es la *respuesta en magnitud* o la *ganancia* del sistema y $\angle H(e^{j\omega})$ es la *respuesta en fase* o *desplazamiento de fase* del sistema.

La transformación en magnitud y fase representada por las ecuaciones 3.2 y 3.3 puede ser tanto deseable, si la señal de entrada es modificada de forma útil, o indeseable, si la señal de entrada es modificada de forma perjudicial. En este último caso, el efecto del sistema LTI correspondiente a las ecuaciones 3.2 y 3.3 sobre la señal de entrada se llama *distorsión de magnitud* y *distorsión de fase* respectivamente.

El ángulo de la fase de un número complejo no está unívocamente definido, ya que la suma de cualquier múltiplo entero de 2π no cambia al número complejo. Al calcular la fase de forma numérica, típicamente se obtiene el valor principal. El valor principal de la fase de $H(e^{j\omega})$ se denotará como $\text{ARG}[H(e^{j\omega})]$, con

$$-\pi < \text{ARG}[H(e^{j\omega})] \leq \pi.$$

Cualquier otro ángulo que de el valor complejo correcto de $H(e^{j\omega})$ puede ser representado en términos del valor principal como

$$\angle H(e^{j\omega}) = \text{ARG}[H(e^{j\omega})] + 2\pi r(\omega),$$

donde $r(\omega)$ es un entero positivo o negativo que puede ser diferente para cada valor de ω .

En general, la respuesta en frecuencia de un sistema LTI es una función continua de la frecuencia. Un ejemplo de una respuesta en frecuencia $H(e^{j\omega})$ en el plano complejo se muestra en figura 3.1. El

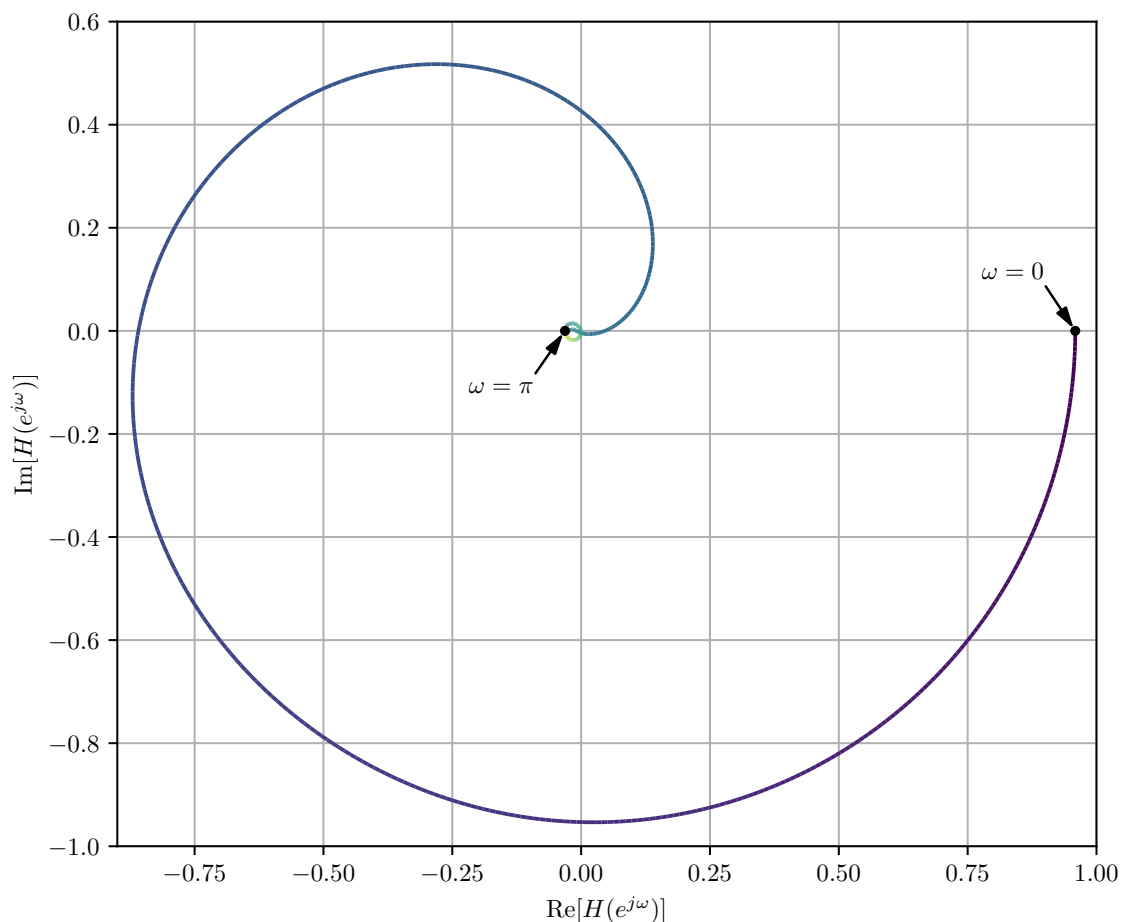


Figura 3.1: Respuesta en frecuencia de un filtro FIR pasabajos en el plano complejo.

sistema de la figura corresponde a un filtro pasabajos, ya que tiene ganancia cercana a la unidad en $\omega = 0$ y ganancia cercana a cero en $\omega = \pi$. Sin embargo, el valor principal de la fase en función de la frecuencia exhibe discontinuidades de 2π radianes. Esto se ilustra en la figura 3.2, donde se muestra el valor principal de la fase $\text{ARG}[H(e^{j\omega})]$ del filtro pasabajos en el rango de frecuencias $0 \leq \omega \leq \pi$. Es común referirse a $\text{ARG}[H(e^{j\omega})]$ como la fase *enrollada*. En la representación en magnitud y fase, suele

ser útil *desenrollar* la fase para obtener una función de fase continua con ω . Esto se logra sumando a la fase enrollada el múltiplo entero de 2π apropiado $r(\omega)$, como se muestra en la figura 3.2. La fase continua desenrollada se denotará como $\arg[H(e^{j\omega})]$. Observar que es común que la fase, tanto enrollada como desenrollada, tenga saltos de π radianes. Esto ocurre en las frecuencias en donde la magnitud se anula, ya que en un entorno de esas frecuencias el número complejo $H(e^{j\omega})$ puede pasar de un valor a su opuesto, lo que corresponde a un cambio de fase de π radianes.

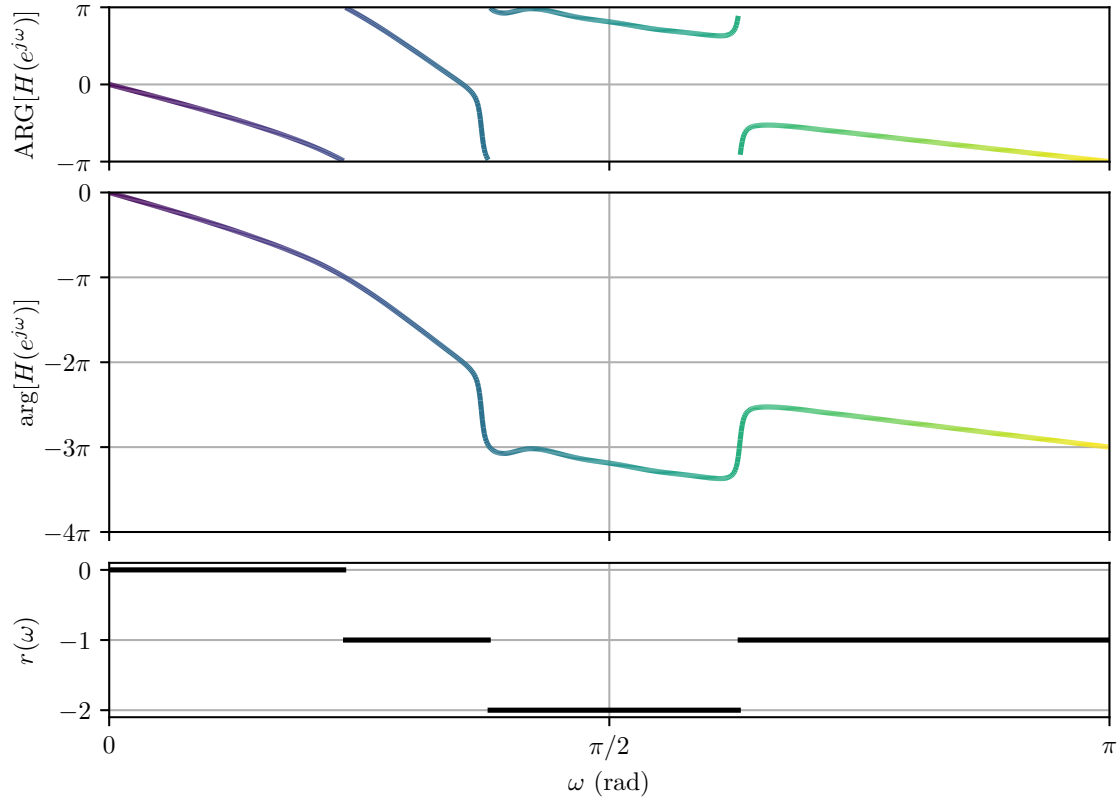


Figura 3.2: Desenrollamiento del valor principal de la fase mediante la suma del múltiplo entero $r(\omega)$ de 2π apropiado. El color de las líneas de la respuesta en fase para cada frecuencia corresponde al mismo color en la misma frecuencia que la línea de $H(e^{j\omega})$ en la figura 3.1.

Otra representación particularmente útil de la fase es el retardo de grupo $\tau(\omega)$, definido como

$$\tau(\omega) = \text{grd}[H(e^{j\omega})] = -\frac{d}{d\omega}\{\arg[H(e^{j\omega})]\}. \quad (3.4)$$

Para entender el efecto de la fase y específicamente del retardo de grupo, se considera primero el sistema retardo ideal. La respuesta al impulso es

$$h_{\text{id}}[n] = \delta[n - n_d],$$

y la respuesta en frecuencia es

$$H_{\text{id}}(e^{j\omega}) = e^{-j\omega n_d},$$

o

$$\begin{aligned} |H_{\text{id}}(e^{j\omega})| &= 1, \\ \angle H_{\text{id}}(e^{j\omega}) &= -\omega n_d, \quad |\omega| < \pi, \end{aligned}$$

donde se asume periodicidad de 2π en ω . Observar que el retardo temporal (o adelanto temporal si $n_d < 0$) está asociado a una fase que varía linealmente con la frecuencia. En muchas aplicaciones, la

distorsión de retardo se considera una forma moderada de distorsión de fase, ya que su efecto es solo desplazar a la señal en el tiempo.

El retardo de grupo definido por la ecuación 3.4, provee una medida apropiada de la linealidad de la fase. Específicamente, considérese la salida de un sistema con respuesta en frecuencia $H(e^{j\omega})$ correspondiente a una entrada de banda angosta de la forma $x[n] = s[n] \cos(\omega_0 n)$. Como se asume que $X(e^{j\omega})$ es únicamente no nula en torno a $\omega = \omega_0$, el efecto de la fase del sistema en la banda de la señal puede aproximarse de forma lineal en torno a $\omega = \omega_0$ como

$$\arg[H(e^{j\omega})] \approx -\phi_0 - \omega n_d,$$

donde n_d es el retardo de grupo. En el siguiente ejemplo se muestra que con esta aproximación, la respuesta $y[n]$ a $x[n] = s[n] \cos(\omega_0 n)$ es aproximadamente $y[n] = |H(e^{j\omega})|s[n - n_0] \cos(\omega_0 n - \phi_0 - \omega_0 n_d)$. En consecuencia, el retardo temporal de la envolvente $s[n]$ de la señal de banda angosta $x[n]$ con transformada de Fourier centrada en ω_0 está dado por el retardo de grupo del sistema en ω_0 . En general, una señal de banda ancha se puede considerar como la superposición de señales de banda angosta centradas en distinta frecuencia. Si el retardo de grupo es constante con la frecuencia, cada componente de banda angosta sufrirá un retardo idéntico. Si el retardo de grupo no es constante, el retardo de cada componente de banda angosta será distinto, resultando en la dispersión temporal de la energía de la señal. Por lo tanto, una respuesta en fase no lineal, o equivalentemente, un retardo de grupo no constante produce dispersión temporal.

Ejemplo: estudio del efecto del retardo de grupo en una señal de banda angosta Este ejemplo es el problema 5.63 de [1]. Sea un sistema con respuesta en frecuencia $H(e^{j\omega})$ con entrada de la forma $x[n] = s[n] \cos(\omega_0 n)$, donde $s[n]$ es una señal banda base de banda relativamente angosta, es decir, $S(e^{j\omega}) = 0$ en $|\omega| > \Delta$, con $\Delta \ll \omega_0$, de forma de que $X(e^{j\omega})$ es de banda angosta en torno a $\omega = \pm\omega_0$. Del teorema del desplazamiento en frecuencia (ver la sección 1.9) de la transformada de Fourier, la transformada de $x[n]$ en términos de la transformada de $s[n]$ es

$$X(e^{j\omega}) = \frac{1}{2} [S(e^{j(\omega-\omega_0)}) + S(e^{j(\omega+\omega_0)})].$$

Se asume que la respuesta en magnitud del sistema es $|H(e^{j\omega})| = 1$, y para la respuesta en fase se

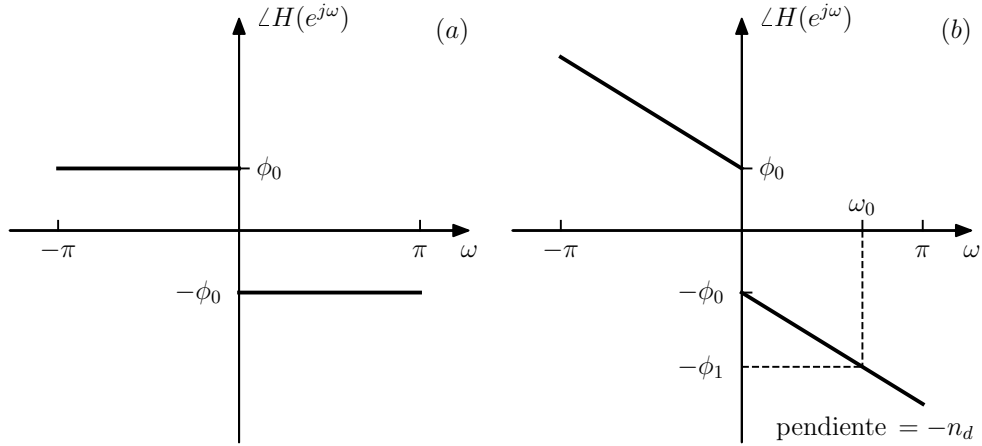


Figura 3.3: Respuesta en fase del sistema de los casos (a) y (b) del ejemplo.

consideran los siguientes casos:

(a) $\angle H(e^{j\omega})$ es como se muestra en la gráfica (a) de la figura 3.3. La transformada de Fourier de la

salida del sistema es

$$\begin{aligned} Y(e^{j\omega}) &= H(e^{j\omega})X(e^{j\omega}) \\ &= \frac{1}{2} \left[S(e^{j(\omega-\omega_0)})e^{-j\phi_0} + S(e^{j(\omega+\omega_0)})e^{j\phi_0} \right], \end{aligned}$$

donde se consideró que como $S(e^{j\omega})$ es de banda angosta, los componentes $S(e^{j(\omega-\omega_0)})$ y $S(e^{j(\omega+\omega_0)})$ no se interfieren, y el primero solo es no nulo en las frecuencias positivas en torno a ω_0 y el segundo solo es no nulo en las frecuencias negativas en torno a $-\omega_0$. Luego, considerando nuevamente el teorema del desplazamiento en frecuencia y la linealidad de la transformada de Fourier, la salida en el dominio del tiempo es

$$\begin{aligned} y[n] &= \frac{1}{2} (s[n]e^{j\omega_0 n}e^{-j\phi_0} + s[n]e^{-j\omega_0 n}e^{j\phi_0}) \\ &= s[n] \frac{e^{j(\omega_0 n - \phi_0)} + e^{-j(\omega_0 n - \phi_0)}}{2}, \end{aligned}$$

resultando en que

$$y[n] = s[n] \cos(\omega_0 n - \phi_0).$$

(b) Ahora, $\angle H(e^{j\omega})$ es como se muestra en la gráfica (b) de la figura 3.3, es decir,

$$\angle H(e^{j\omega}) = \begin{cases} \phi_0 - \omega n_d & -\pi < \omega < 0 \\ -\phi_0 - \omega n_d & 0 < \omega < \pi \end{cases}$$

con periodicidad 2π . La transformada de Fourier de la salida es

$$\begin{aligned} Y(e^{j\omega}) &= H(e^{j\omega})X(e^{j\omega}) \\ &= \frac{1}{2} \left[S(e^{j(\omega-\omega_0)})e^{j(-\phi_0-\omega n_d)} + S(e^{j(\omega+\omega_0)})e^{j(\phi_0-\omega n_d)} \right] \\ &= \frac{1}{2} \left[S(e^{j(\omega-\omega_0)})e^{-j\omega n_d}e^{-j\phi_0} + S(e^{j(\omega+\omega_0)})e^{-j\omega n_d}e^{j\phi_0} \right] \end{aligned} \quad (3.5)$$

donde se consideró el mismo argumento que antes. Luego, teniendo en cuenta que

$$s[n]e^{j\omega_0 n} \xleftrightarrow{\mathcal{F}} S(e^{j(\omega-\omega_0)}) \quad \text{y} \quad \delta[n-n_d] \xleftrightarrow{\mathcal{F}} e^{-j\omega n_d},$$

donde se asumió que n_d es un número entero, se cumple que

$$s[n]e^{j\omega_0 n} * \delta[n-n_d] \xleftrightarrow{\mathcal{F}} S(e^{j(\omega-\omega_0)})e^{-j\omega n_d},$$

es decir,

$$s[n-n_d]e^{j\omega_0(n-n_d)} \xleftrightarrow{\mathcal{F}} S(e^{j(\omega-\omega_0)})e^{-j\omega n_d}.$$

Realizando el mismo razonamiento para el segundo sumando de la ecuación 3.5, se cumple que

$$s[n-n_d]e^{-j\omega_0(n-n_d)} \xleftrightarrow{\mathcal{F}} S(e^{j(\omega+\omega_0)})e^{-j\omega n_d},$$

por lo que que la salida en el dominio del tiempo es

$$\begin{aligned} y[n] &= \frac{1}{2} \left\{ s[n-n_d]e^{j\omega_0(n-n_d)}e^{-j\phi_0} + s[n-n_d]e^{-j\omega_0(n-n_d)}e^{j\phi_0} \right\} \\ &= s[n-n_d] \frac{e^{j[\omega_0(n-n_d)-\phi_0]} + e^{-j[\omega_0(n-n_d)-\phi_0]}}{2} \\ &= s[n-n_d] \cos[\omega_0(n-n_d) - \phi_0], \end{aligned}$$

o

$$y[n] = s[n-n_d] \cos[\omega_0 n - \phi_0 - \omega_0 n_d]. \quad (3.6)$$

Si la fase en ω_0 es ϕ_1 , como se muestra en la figura 3.3, es decir,

$$-\phi_0 - \omega_0 n_d = -\phi_1,$$

resulta en que

$$y[n] = s[n-n_d] \cos[\omega_0 n - \phi_1].$$

- (c) Ahora la fase es una función arbitraria con retardo de grupo y retardo de fase en ω_0 los números enteros $\tau_{\text{gr}}(\omega_0)$ y $\tau_{\text{ph}}(\omega_0)$ respectivamente (ver las ecuaciones 3.4 y 1.32). Como la transformada de Fourier de la entrada es

$$X(e^{j\omega}) = \frac{1}{2} \left[S(e^{j(\omega-\omega_0)}) + S(e^{j(\omega+\omega_0)}) \right],$$

y $S(e^{j(\omega-\omega_0)})$ es de banda angosta en torno a ω_0 , la respuesta en fase del sistema se puede aproximar por una recta en el ancho de banda de la señal, como se muestra en la figura 3.4. Realizando un desarrollo de Taylor de primer orden de la función de fase en torno al punto $\omega = \omega_0$, se cumple que

$$\arg H(e^{j\omega}) \approx \arg H(e^{j\omega_0}) + \left. \frac{d}{d\omega} [\arg H(e^{j\omega})] \right|_{\omega=\omega_0} (\omega - \omega_0)$$

en $\omega \approx \omega_0$. Como por definición,

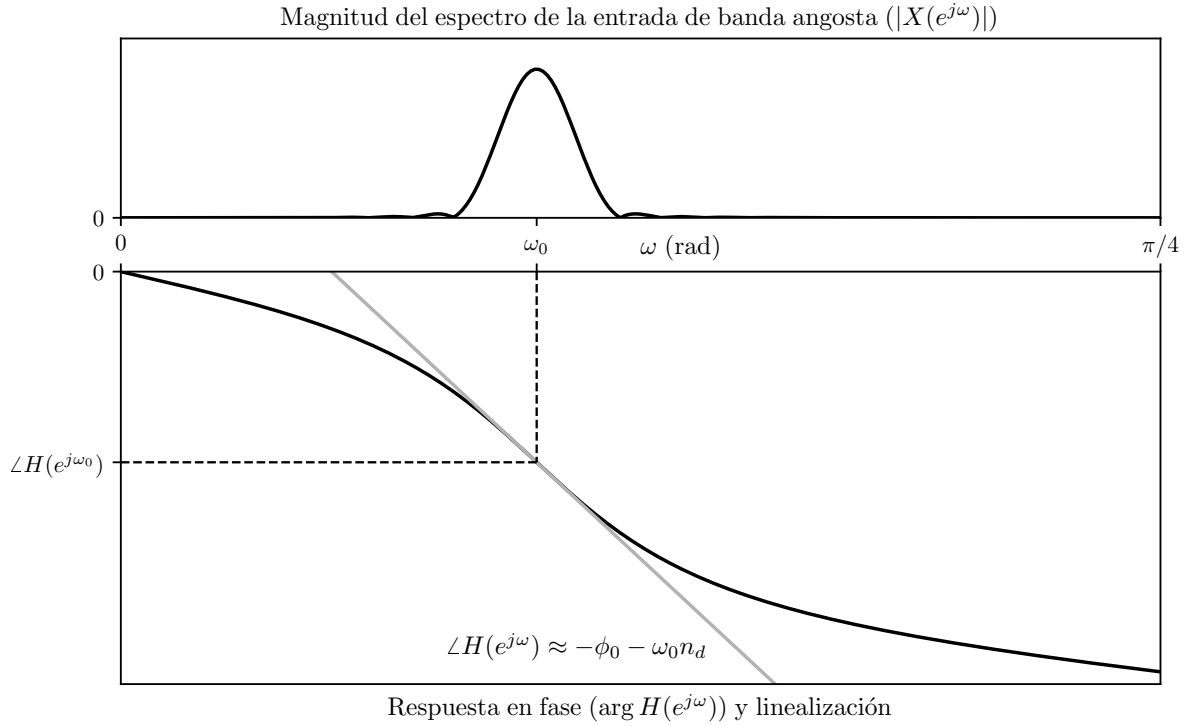


Figura 3.4: Magnitud de la entrada $|X(e^{j\omega})|$ y respuesta en fase del sistema $\arg H(e^{j\omega})$. Como la entrada es de banda angosta, la fase se puede aproximar como una función lineal en la banda de la entrada.

$$\tau_{\text{gr}}(\omega_0) = -\left. \frac{d}{d\omega} [\arg H(e^{j\omega})] \right|_{\omega=\omega_0} \quad \text{y} \quad \tau_{\text{gr}}(\omega_0) = -\frac{\arg H(e^{j\omega_0})}{\omega_0},$$

la aproximación lineal de la fase puede expresarse en términos del retardo de grupo y el retardo de fase como

$$\begin{aligned} \arg H(e^{j\omega}) &\approx -\tau_{\text{ph}}(\omega_0)\omega_0 - \tau_{\text{gr}}(\omega_0)(\omega - \omega_0) \\ &\approx -\omega_0[\tau_{\text{ph}}(\omega_0) - \tau_{\text{gr}}(\omega_0)] - \omega\tau_{\text{gr}}(\omega_0). \end{aligned}$$

Por lo tanto, los parámetros ϕ_0 y n_d correspondientes al corte por cero y a la pendiente de la linealización de la fase en la parte (b) son respectivamente

$$\phi_0 = \omega_0[\tau_{\text{ph}}(\omega_0) - \tau_{\text{gr}}(\omega_0)] \quad \text{y} \quad n_d = \tau_{\text{gr}}(\omega_0).$$

Sustituyendo estos parámetros en la ecuación 3.6, la salida queda

$$y[n] \approx s[n - \tau_{\text{gr}}(\omega_0)] \cos \{ \omega_0 n - \omega_0 [\tau_{\text{ph}}(\omega_0) - \tau_{\text{gr}}(\omega_0)] - \omega_0 \tau_{\text{gr}}(\omega_0) \}$$

resultando en

$$y[n] \approx s[n - \tau_{\text{gr}}(\omega_0)] \cos \{ \omega_0 [n - \tau_{\text{ph}}(\omega_0)] \}. \quad (3.7)$$

Esta ecuación indica que para una señal de banda angosta $x[n]$, el efecto de la fase $\angle H(e^{j\omega})$ del filtro es aplicar un retardo $\tau_{\text{gr}}(\omega_0)$ a la envolvente $s[n]$ y un retardo $\tau_{\text{ph}}(\omega_0)$ a la portadora $\cos \omega_0 n$.

Observar que la ecuación 3.7 es válida solo si $n_d = \tau_{\text{gr}}(\omega_0)$ es un número entero, ya que la transformada inversa de $e^{-j\omega n_d}$ es $\delta[n - n_d]$ si n_d es un número entero, hecho que se empleó para obtener la ecuación 3.6. Si $n_d = \Delta$ no es entero, aplicando la transformada Fourier inversa, se puede mostrar que

$$\frac{\sin \pi(n - \Delta)}{\pi(n - \Delta)} \xleftrightarrow{\mathcal{F}} e^{-j\omega \Delta}.$$

De todas formas, si $\tau_{\text{gr}}(\omega_0)$ o $\tau_{\text{ph}}(\omega_0)$ o ambos no son enteros, el efecto sobre la envolvente y la portadora es el mismo, con la interpretación de un retardo no entero explicada en la sección 4.5 de [1], que consiste en convertir la señal de tiempo discreto a tiempo continuo (ver la sección 4.3 de [1]), aplicar un retardo arbitrario en tiempo continuo y muestrear la señal resultante para obtener una señal en tiempo discreto.

Como ilustración del efecto del retardo de fase y el retardo de grupo, considérese el sistema con repuesta en fase que se muestra en la figura 3.5. En la figura se muestra además el correspondiente

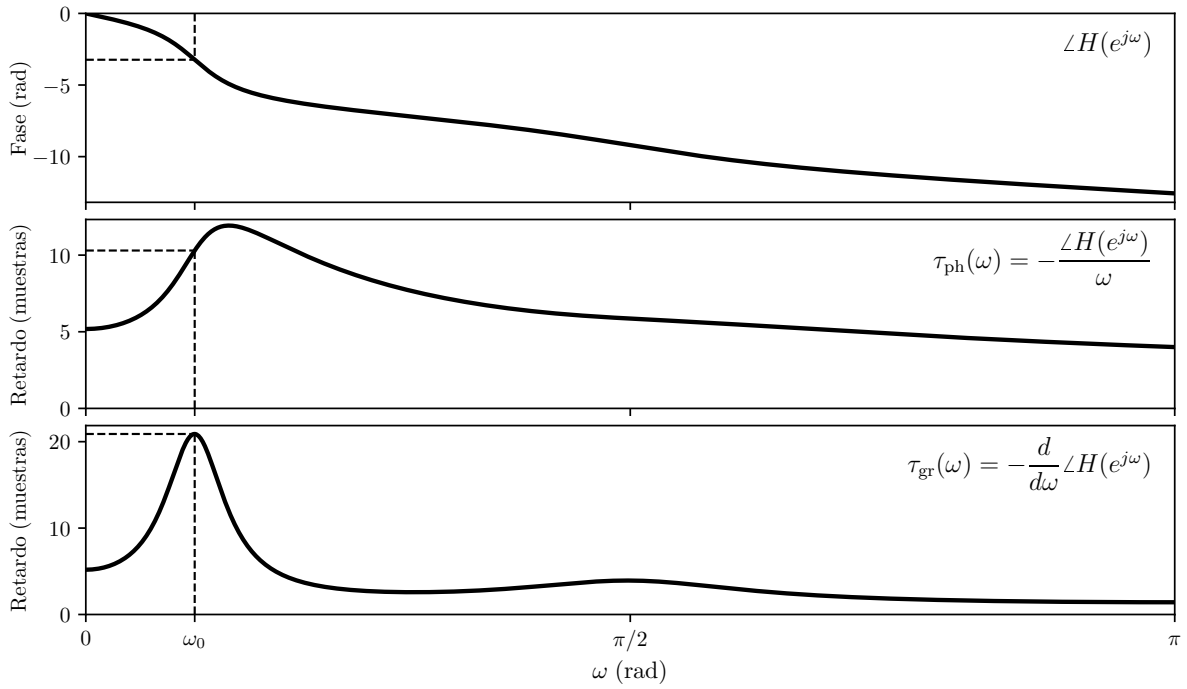


Figura 3.5: Respuesta en fase, retardo de fase y retardo de grupo. En la frecuencia ω_0 el retardo de fase es aproximadamente 10.3 muestras y el retardo de grupo es aproximadamente 20.9 muestras.

retardo de grupo $\tau_{\text{gr}}(\omega)$ y retardo de fase $\tau_{\text{ph}}(\omega)$. Sea la entrada de banda angosta $x[n] = s[n] \cos \omega_0 n$ con espectro como el que se muestra en la figura 3.4. La respuesta en magnitud del sistema es la unidad en la frecuencia ω_0 de la portadora. Como se observa en la figura 3.5, en la frecuencia ω_0 , el retardo de fase es $\tau_{\text{ph}}(\omega_0) \approx 10.3$ muestras y el retardo de grupo es $\tau_{\text{gr}}(\omega_0) \approx 20.9$ muestras. En la figura 3.6 se muestra la entrada y la correspondiente salida. Se observa que el retardo de la portadora, por ejemplo midiendo la distancia entre un máximo de la senoide de entrada y

el correspondiente máximo de la senoide de salida, es el retardo de fase $\tau_{\text{ph}}(\omega_0)$, mientras que el retardo de la envolvente, que se puede obtener midiendo la distancia entre los máximos de la envolvente de la entrada y de la salida, es el retardo de grupo $\tau_{\text{gr}}(\omega_0)$, acorde con el resultado de la ecuación 3.7.

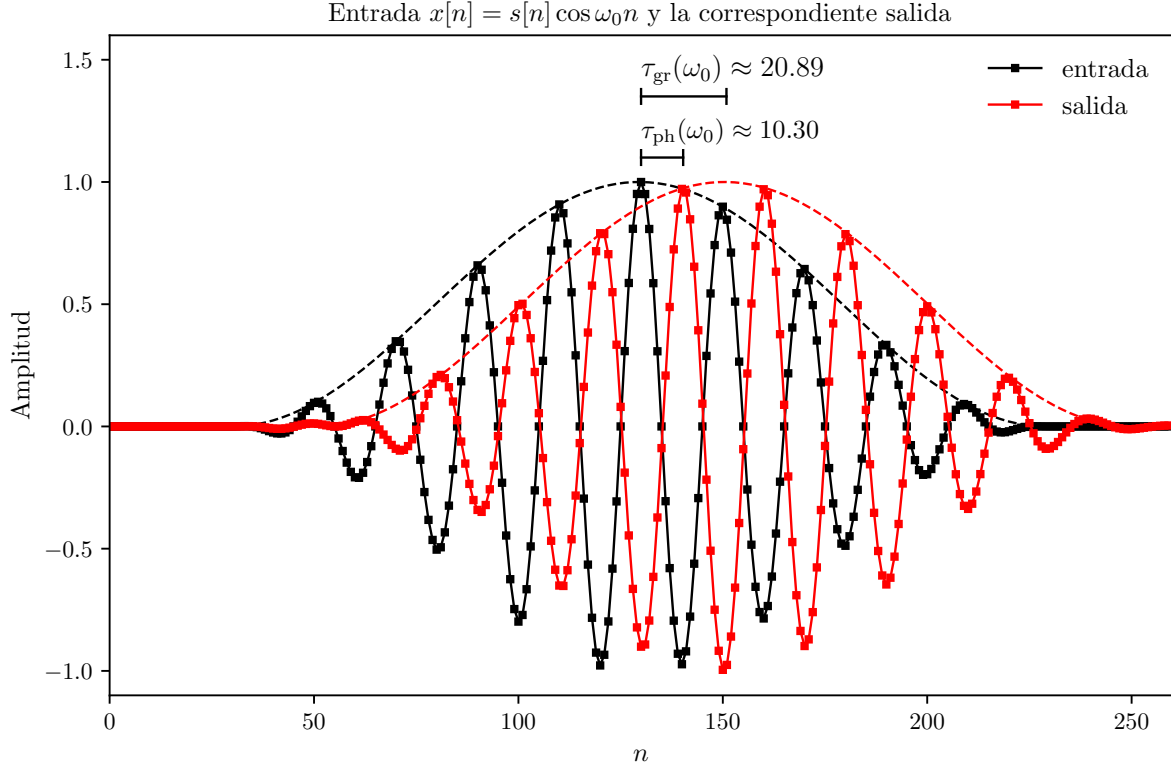


Figura 3.6: Señal de entrada y señal de salida. El retardo de la portadora es $\tau_{\text{ph}}(\omega_0) \approx 10.3$ muestras y el retardo de la envolvente es $\tau_{\text{gr}}(\omega_0) \approx 20.9$ muestras.

Ejemplo: otra ilustración del efecto del retardo de grupo en una señal de banda angosta
Este ejemplo es una variación del ejemplo de la sección 5.1.2 de [1]. Se considera ahora el sistema con función de transferencia

$$H(z) = \prod_{k=1}^4 \left[\frac{(c_k^* - z^{-1})(c_k - z^{-1})}{(1 - c_k z^{-1})(1 - c_k^* z^{-1})} \right]^2, \quad (3.8)$$

donde $c_k = 0.95e^{j(0.15\pi + 0.02\pi k)}$ para $k = 1, 2, 3, 4$. Este sistema tiene polos dobles en $z = c_k = 0.95e^{j(0.15\pi + 0.02\pi k)}$ y en $z = c_k^* = 0.95e^{-j(0.15\pi + 0.02\pi k)}$ y ceros dobles en $z = 1/c_k = 0.95e^{-j(0.15\pi + 0.02\pi k)}$ y en $z = 1/c_k^* = 0.95e^{j(0.15\pi + 0.02\pi k)}$. Como se verá mas adelante en la sección 3.6, $H(z)$ es un sistema pasa todo, es decir, $|H(e^{j\omega})| = 1$ para todo ω . Como se ilustrará, este sistema introduce un retardo de grupo grande sobre una banda angosta de frecuencias.

En la figura 3.8 se muestra la respuesta en frecuencia del sistema. Como ya se mencionó, se observa que la respuesta en magnitud es unitaria en todas las frecuencias, como corresponde a un sistema pasa todo. Además, el valor principal de la fase tiene exhibe múltiples discontinuidades de tamaño 2π en torno a la frecuencia $\omega = 0.2\pi$ radianes. Esta frecuencia corresponde la fase central del grupo de polos y ceros del sistema, cuyas fases son $\pm 0.17\pi$, $\pm 0.19\pi$, $\pm 0.21\pi$ y $\pm 0.23\pi$ radianes, tanto para los ceros como para los polos.

En la figura 3.9 se muestra la respuesta en fase desenrollada y el retardo de grupo del sistema. Se observa que, como la fase es monótonamente decreciente, el retardo de grupo es positivo. Además se ve que el retardo de grupo tiene un pico positivo grande en una banda de frecuencia en torno a 0.2π radianes, que

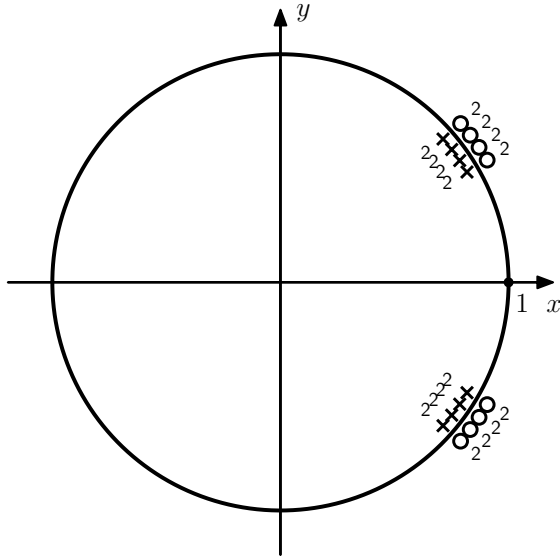


Figura 3.7: Diagrama de polos y ceros del sistema $H(z)$.

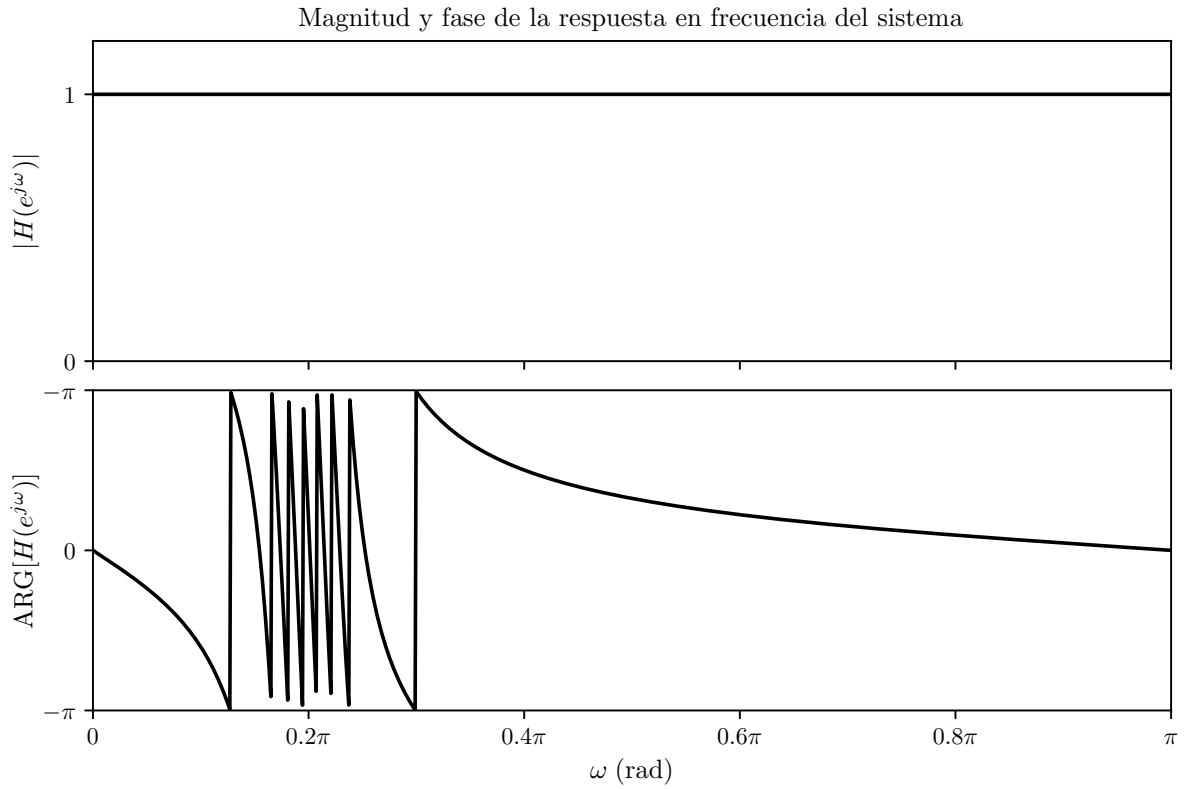


Figura 3.8: Respuesta en magnitud y respuesta en fase del sistema $H(z)$.

corresponde a la ubicación angular del grupo de polos y ceros mostrados en la figura 3.7. Como entrada al sistema se considera una señal que consiste en dos pulsos de banda angosta separados en el tiempo. Los pulsos son

$$x_1[n] = w[n] \cos(0.2\pi n) \quad \text{y} \quad x_2[n] = w[n] \cos(0.4\pi n - \pi/2),$$

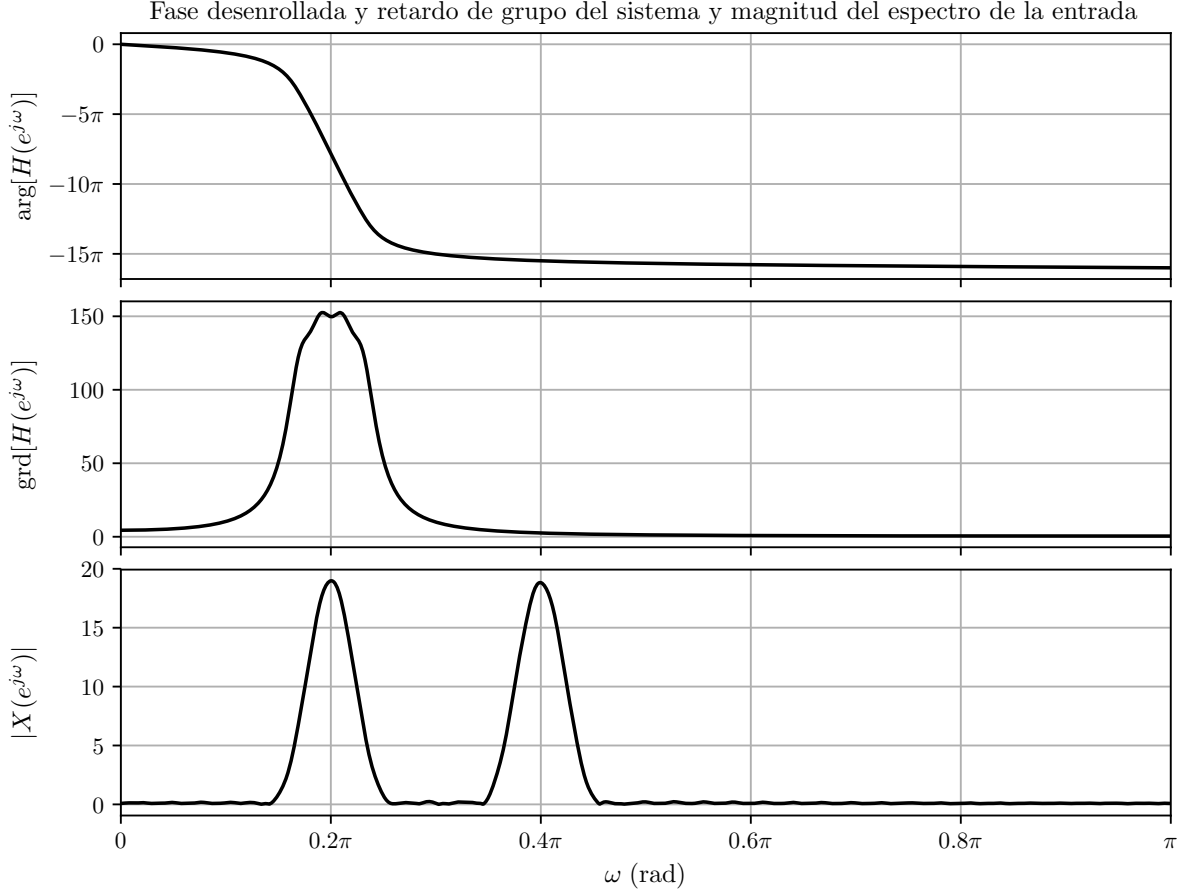


Figura 3.9: Fase desenrollada y retardo de grupo del sistema. También se muestra la magnitud del espectro de la señal de entrada. Se observa que la sinusoide de entrada de frecuencia $\omega_1 = 0.2\pi$ se ve afectada por un retardo de grupo de alrededor de 150 muestras mientras que la sinusoide de frecuencia $\omega_1 = 0.4\pi$ se ve afectada por un retardo de grupo pequeño (2.5 muestras aproximadamente).

donde la envolvente $w[n]$ de cada sinusoide es un ventana de Hamming de 71 muestras,

$$w[n] = \begin{cases} 0.54 - 0.46 \cos(2\pi n/M) & 0 \leq n \leq M \\ 0 & \text{en otro caso,} \end{cases}$$

donde $M = 70$. La secuencia completa de entrada es

$$x[n] = x_1[n] + x_2[n - M - 1],$$

es decir, la sinusoide de menor frecuencia viene primero seguida de la sinusoide de mayor frecuencia. En la figura 3.9 se muestra la magnitud del espectro de la señal de entrada. Como es de esperar por el teorema de modulación (ver la sección 1.9), la señal tiene alta energía en torno a las frecuencias de las sinusoides $\omega_1 = 0.2\pi$ y $\omega_2 = 0.4\pi$. La forma de los pulsos de alta energía son la transformada de Fourier de la ventana temporal.

Cuando se emplea $x[n]$ como entrada al sistema $H(z)$, cada banda de frecuencia asociado a cada una de las sinusoides será afectada por la repuesta en magnitud y el retardo de grupo del sistema en cada banda de frecuencia. Como el sistema es un filtro pasa todo, no afecta la magnitud de las sinusoides. Examinando el retardo de grupo en la figura 3.9 se ve que el retardo de grupo en torno a la frecuencia $\omega = \omega_1 = 0.2\pi$ es significativamente mayor que el correspondiente a la frecuencia $\omega = \omega_2 = 0.4\pi$, y en consecuencia, la sinusoide de menor frecuencia experimentará un mayor retardo.

La entrada $x[n]$ y la correspondiente salida $y[n]$ son mostradas en la figura 3.10. El pulso sinusoidal de

frecuencia $\omega_1 = 0.2\pi$ es retardado unas 150 muestras, mientras que el pulso sinusoidal de frecuencia $\omega_2 = 0.4\pi$ es retardado unas 2 muestras. De hecho, como la senoide de baja frecuencia es retardada unas 147 muestras mas que la senoide de alta frecuencia y el largo de los pulsos sinusoidales es de 71 muestras, los pulsos son intercambiados de orden temporal en la salida, como se ve en la figura.

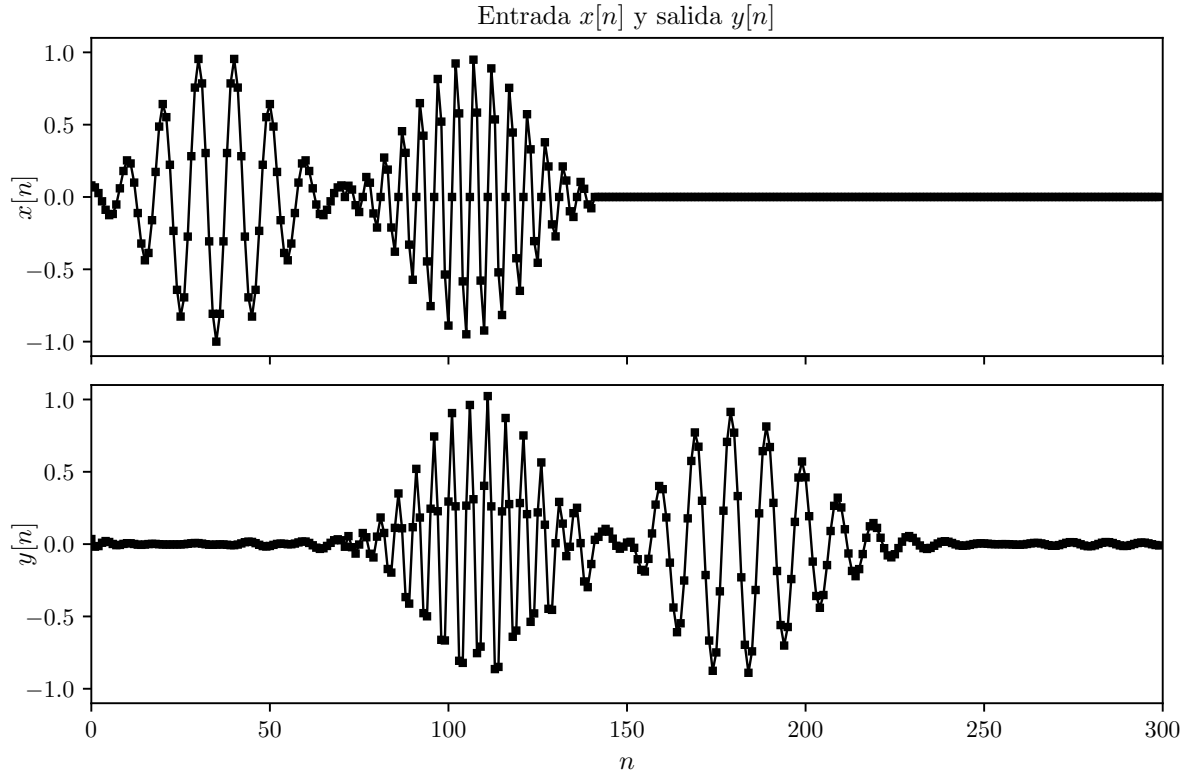


Figura 3.10: Entrada y salida del sistema. Debido al alto retardo de grupo que afecta al pulso sinusoidal de menor frecuencia, el orden de ambos pulsos es intercambiado en la salida.

Finalmente, se realizan algunas consideraciones sobre la implementación del filtro en el lenguaje de programación Python. Cuando la función de transferencia es racional, como es el caso de $H(z)$ dado por la ecuación 3.8 del presente ejemplo, el sistema en el dominio del tiempo está dado por una ecuación en diferencias de la forma 1.19, y hay una relación directa entre los coeficientes de la ecuación en diferencias y los coeficientes de los polinomios del numerador y denominador de la función de transferencia. Concretamente, la función de transferencia del sistema dado por 1.19 es

$$H(z) = \frac{Y(z)}{X(z)} = \frac{\sum_{k=0}^M b_k z^{-k}}{\sum_{k=0}^N a_k z^{-k}}. \quad (3.9)$$

En el caso de este ejemplo, la función del sistema está dada por los polos y ceros del sistema, es decir, por las raíces de los polinomios del numerador y denominador. Una forma de obtener los coeficientes de los polinomios a partir de las raíces de los polinomios es mediante la función `zpk2tf` de la biblioteca

```

import numpy as np
from scipy import signal

# polos
c_k = 0.95 * np.exp(1j * (0.15 * np.pi + 0.02 * np.pi * np.arange(1, 5)))
# construcción del vector de ceros y polos y cálculo de la ganancia: (z, p, k)
# ceros dobles en 1/c_k, 1/c_k^*. polos dobles en c_k, c_k^*.
p = np.concatenate((c_k, c_k, np.conjugate(c_k), np.conjugate(c_k)))
z = np.concatenate((1 / c_k, 1 / c_k, 1 / np.conjugate(c_k), 1 / np.conjugate(c_k)))
k = np.abs(np.prod(c_k)) ** 4
# cálculo de los coeficientes (b, a) a partir de (z, p, k)
b, a = signal.zpk2tf(z, p, k)

```

Figura 3.11: Implementación en el lenguaje Python del cálculo de los coeficientes **b** y **a** a partir del conjunto de ceros **z**, polos **p** y la ganancia **k** del filtro empleado en este ejemplo.

SciPy [4]. Dicha función asume que la función de transferencia racional está factorizada como

$$H(z) = k \frac{\prod_{k=1}^M (z - c_k)}{\prod_{k=1}^N (z - d_k)} \quad (3.10)$$

y recibe como parámetros los vectores **z** y **p** conteniendo respectivamente los ceros c_k y los polos d_k , además del valor k de la ganancia de la función del sistema 3.10 y calcula y devuelve los vectores **b** y **a** conteniendo respectivamente los coeficientes b_k y a_k de la representación 3.9, que como se mencionó, son los mismos que los de la ecuación en diferencias 1.19. Para expresar la función del sistema 3.8 de este ejemplo de la forma 3.10 para emplear la función `zpk2tf`, se observa que cada término pasa todo en 3.8, puede expresarse como

$$\frac{c_k^* - z^{-1}}{1 - c_k z^{-1}} = \frac{z c_k^* - 1}{z - c_k} = c_k^* \frac{z - 1/c_k^*}{z - c_k},$$

y por lo tanto, dicha ecuación puede expresarse como

$$\begin{aligned}
H(z) &= \prod_{k=1}^4 \left[c_k c_k^* \frac{(z - 1/c_k^*)(z - 1/c_k)}{(z - c_k)(z - c_k^*)} \right]^2 \\
&= \prod_{k=1}^4 \left[|c_k|^2 \frac{(z - 1/c_k^*)(z - 1/c_k)}{(z - c_k)(z - c_k^*)} \right]^2 \\
&= \prod_{k=1}^4 |c_k|^4 \prod_{k=1}^4 \left[\frac{(z - 1/c_k^*)(z - 1/c_k)}{(z - c_k)(z - c_k^*)} \right]^2.
\end{aligned}$$

De esta forma,

$$k = \prod_{k=1}^4 |c_k|^4,$$

el vector **z** de ceros contiene los valores $1/c_k$ y $1/c_k^*$ duplicados, ya que son ceros dobles debido a la elevación al cuadrado, y el vector **p** de polos contiene los valores c_k y c_k^* , también duplicados por la misma razón. En la figura 3.11 se incluye el código en el lenguaje Python para el cálculo de los coeficientes **b** y **a** a partir del conjunto de ceros **z**, polos **p** y la ganancia **k** del filtro dado por la ecuación 3.8.

3.3. Sistemas caracterizados por ecuaciones en diferencias lineales con coeficientes constantes

Los filtros en el dominio del tiempo discreto son típicamente implementados mediante una ecuación en diferencias lineal con coeficientes constantes de la forma

$$\sum_{k=0}^N a_k y[n-k] = \sum_{k=0}^M b_k x[n-k]. \quad (3.11)$$

En el capítulo 4 se discuten varias estructuras para la realización de estos sistemas y en el capítulo ?? se discuten varios procedimientos para la obtención de los parámetros de la ecuación en diferencias para aproximar una respuesta en frecuencia deseada. En esta sección se examinan las propiedades y características de los sistemas LTI representados por la ecuación 3.11 con la ayuda de la transformada z .

Aplicando la transformada z a ambos lados de la ecuación 3.11 se obtiene que la función del sistema tiene la forma algebraica

$$H(z) = \frac{Y(z)}{X(z)} = \frac{\sum_{k=0}^M b_k z^{-k}}{\sum_{k=0}^N a_k z^{-k}}. \quad (3.12)$$

En ocasiones, es conveniente expresar la ecuación 3.12 en forma factorizada como

$$H(z) = \left(\frac{b_0}{a_0} \right) \frac{\prod_{k=1}^M (1 - c_k z^{-1})}{\prod_{k=1}^N (1 - d_k z^{-1})}. \quad (3.13)$$

Cada uno de los factores $(1 - c_k z^{-1})$ en el numerador contribuye con un cero en $z = c_k$ y con un polo en $z = 0$ y cada uno de los factores $(1 - d_k z^{-1})$ en el denominador contribuye con un cero en $z = 0$ y con un polo en $z = d_k$.

3.3.1. Estabilidad y causalidad

Para obtener la ecuación 3.12 a partir de la ecuación 3.11 se asume que el sistema es lineal e invariante en el tiempo de forma de que se cumple la ecuación 3.1, pero no se realizó ninguna hipótesis sobre estabilidad o causalidad. De forma correspondiente, si bien de la ecuación en diferencias puede obtenerse la función del sistema, no puede especificarse la región de convergencia. Específicamente, la ROC de $H(z)$ no queda determinada en la deducción que conduce a la ecuación 3.12, ya que lo único que se requiere para que se cumple dicha ecuación es que $X(z)$ y $Y(z)$ tengan ROCs que se solapen (para la existencia de $H(z)$). Esto es consistente con el hecho de que la ecuación en diferencias no especifica unívocamente la respuesta al impulso del sistema LTI, como se discutió en la sección 1.5. Para el sistema de la ecuación 3.12 o de la ecuación 3.13, hay cierta cantidad de elecciones de ROC distintas. Dado un cociente de polinomios dado, cada posible ROC conduce a una respuesta al impulso distinta, pero a la misma ecuación en diferencias. Sin embargo, si se asume que el sistema es causal, $h[n]$ debe ser una secuencia hacia adelante, y en consecuencia, la ROC de $H(z)$ es la región exterior al polo de mayor magnitud (ver la figura 2.2). Alternativamente, si se asume que el sistema es estable, como se indicó en la sección 1.4, la respuesta al impulso debe ser absolutamente sumable, es decir,

$$\sum_{n=-\infty}^{\infty} |h[n]| < \infty.$$

Como esta condición es equivalente a la condición

$$\sum_{n=-\infty}^{\infty} |h[n] z^{-n}| < \infty$$

para $|z| = 1$, es decir, a la condición de que $H(z)$ converja en $|z| = 1$, la condición de estabilidad es que la ROC contenga a la circunferencia unidad.

De esta discusión se deduce que para que un sistema LTI sea causal y estable, la ROC de la función del sistema debe ser la región exterior al polo de mayor magnitud y además debe contener a la circunferencia unidad. Esto requiere que todos los polos de la función del sistema estén contenidos dentro de la circunferencia unidad.

3.3.2. Sistemas inversos

Dado un sistema LTI con función de transferencia $H(z)$, el correspondiente sistema inverso se define como el sistema con función de transferencia $H_i(z)$ que al ser colocado en serie con $H(z)$ produce un sistema con función de transferencia global unidad, es decir,

$$G(z) = H(z)H_i(z) = 1. \quad (3.14)$$

Esto implica que

$$H_i(z) = \frac{1}{H(z)}.$$

La condición equivalente a la ecuación 3.14 en el dominio del tiempo es

$$g[n] = h[n] * h_i[n] = \delta[n]. \quad (3.15)$$

De la ecuación 3.14, la respuesta en frecuencia del sistema inverso, si existe, es

$$H_i(e^{j\omega}) = \frac{1}{H(e^{j\omega})}.$$

Equivalentemente, la magnitud logarítmica, la fase y el retardo de grupo del sistema inverso son los opuestos de las funciones correspondientes del sistema original. No todos los sistemas tienen inverso. Por ejemplo, el pasabajos ideal no lo tiene. No hay forma de recuperar los componentes espectrales de la señal sobre la frecuencia de corte que son eliminados por el filtro.

Muchos sistemas tienen sistema inverso, y la clase de sistemas con función de transferencia racional brindan un ejemplo interesante y útil. Específicamente, considérese el sistema

$$H(z) = \left(\frac{b_0}{a_0} \right) \frac{\prod_{k=1}^M (1 - c_k z^{-1})}{\prod_{k=1}^N (1 - d_k z^{-1})}$$

con ceros en $z = c_k$ y polos en $z = d_k$, además de posibles ceros o polos en $z = 0$. El sistema inverso es

$$H_i(z) = \left(\frac{a_0}{b_0} \right) \frac{\prod_{k=1}^N (1 - d_k z^{-1})}{\prod_{k=1}^M (1 - c_k z^{-1})},$$

es decir, los polos de $H_i(z)$ son los ceros de $H(z)$ y viceversa. Surge la pregunta de que ROC asociar a $H_i(z)$. La respuesta es provista por el teorema de la convolución, que en este caso está dado por la ecuación 3.15. Para que se cumpla la ecuación 3.15 la ROC de $H(z)$ y $H_i(z)$ deben solaparse.

Si $H(z)$ es causal, su ROC es

$$|z| > \max_k |d_k|,$$

y por lo tanto, cualquier ROC apropiada de $H_i(z)$ que se solape con esta ROC, es una ROC válida de $H_i(z)$. El sistema inverso será también causal si y solo si se asocia la ROC

$$|z| > \max_k |c_k|$$

a $H_i(z)$. Si se requiere además que el sistema inverso sea estable, la ROC de $H_i(z)$ debe incluir la circunferencia unidad, para lo cual se debe cumplir que

$$\max_k |c_k| < 1,$$

es decir, todos los ceros de $H(z)$ deben estar contenidos en el círculo unidad. Por lo tanto, un sistema LTI es causal y estable y además tiene un sistema inverso causal y estable si y solo si todos los polos y ceros de $H(z)$ están contenidos dentro de la circunferencia unidad. Los sistemas con estas características son referidos como sistemas de *fase mínima*, y será discutidos con mayor detalle en la sección 3.7.

3.4. Respuesta en frecuencia de funciones de sistema racionales

Si un sistema LTI estable tiene una función racional, es decir, si su entrada y salida satisfacen una ecuación en diferencias de la forma 3.11, su respuesta en frecuencia expresada en términos de los ceros y polos del sistema está dada por la ecuación 3.13 evaluada en la circunferencia unidad $z = e^{j\omega}$,

$$H(e^{j\omega}) = \left(\frac{b_0}{a_0} \right) \frac{\prod_{k=1}^M (1 - c_k e^{-j\omega})}{\prod_{k=1}^N (1 - d_k e^{-j\omega})}.$$

La respuesta en magnitud es

$$|H(e^{j\omega})| = \left| \frac{b_0}{a_0} \right| \frac{\prod_{k=1}^M |1 - c_k e^{-j\omega}|}{\prod_{k=1}^N |1 - d_k e^{-j\omega}|},$$

y la respuesta en magnitud al cuadrado es

$$|H(e^{j\omega})|^2 = H(e^{j\omega})H^*(e^{j\omega}) = \left(\frac{b_0}{a_0} \right)^2 \frac{\prod_{k=1}^M (1 - c_k e^{-j\omega})(1 - c_k^* e^{j\omega})}{\prod_{k=1}^N (1 - d_k e^{-j\omega})(1 - d_k^* e^{j\omega})}.$$

Expresada en decibels, la ganancia se define como

$$\text{Ganancia en dB} = 20 \log_{10} |H(e^{j\omega})|$$

y en este caso es

$$\text{Ganancia en dB} = 20 \log_{10} \left| \frac{b_0}{a_0} \right| + \sum_{k=1}^M 20 \log_{10} |1 - c_k e^{-j\omega}| - \sum_{k=1}^N 20 \log_{10} |1 - d_k e^{-j\omega}|. \quad (3.16)$$

La respuesta en fase es

$$\arg [H(e^{j\omega})] = \arg \left(\frac{b_0}{a_0} \right) + \sum_{k=1}^M \arg(1 - c_k e^{-j\omega}) - \sum_{k=1}^N \arg(1 - d_k e^{-j\omega}), \quad (3.17)$$

donde $\arg()$ representa la fase desenrollada. El retardo de grupo es

$$\text{grd} [H(e^{j\omega})] = \sum_{k=1}^N \frac{d}{d\omega} \arg(1 - d_k e^{-j\omega}) - \sum_{k=1}^M \frac{d}{d\omega} \arg(1 - c_k e^{-j\omega}).$$

Una expresión equivalente del retardo de grupo es

$$\text{grd} [H(e^{j\omega})] = \sum_{k=1}^M \frac{|c_k|^2 - \text{Re}(c_k e^{-j\omega})}{1 + |c_k|^2 - 2 \text{Re}(c_k e^{-j\omega})} - \sum_{k=1}^N \frac{|d_k|^2 - \text{Re}(d_k e^{-j\omega})}{1 + |d_k|^2 - 2 \text{Re}(d_k e^{-j\omega})}. \quad (3.18)$$

Para deducir esta identidad, considérese un factor de la forma $1 - ce^{-j\omega}$ correspondiente a un cero. Se calculará el retardo de grupo aportado por dicho factor. Expresando este número complejo como

$$1 - ce^{-j\omega} = 1 - \operatorname{Re}(ce^{-j\omega}) - j \operatorname{Im}(ce^{-j\omega})$$

su repuesta en fase puede escribirse como

$$\arg(1 - ce^{-j\omega}) = \arctan \left[\frac{-\operatorname{Im}(ce^{-j\omega})}{1 - \operatorname{Re}(ce^{-j\omega})} \right] = -\arctan \left[\frac{\operatorname{Im}(ce^{-j\omega})}{1 - \operatorname{Re}(ce^{-j\omega})} \right],$$

donde en la última igualdad se consideró que la función arcotangente es impar. El retardo de grupo es

$$\begin{aligned} \operatorname{grd}(1 - ce^{-j\omega}) &= -\frac{d}{d\omega} \arg(1 - ce^{-j\omega}) \\ &= \frac{d}{d\omega} \arctan \left[\frac{\operatorname{Im}(ce^{-j\omega})}{1 - \operatorname{Re}(ce^{-j\omega})} \right] \\ &= \frac{1}{1 + \left[\frac{\operatorname{Im}(ce^{-j\omega})}{1 - \operatorname{Re}(ce^{-j\omega})} \right]^2} \frac{d}{d\omega} \left[\frac{\operatorname{Im}(ce^{-j\omega})}{1 - \operatorname{Re}(ce^{-j\omega})} \right]. \end{aligned} \quad (3.19)$$

Para calcular la derivada en la última igualdad, se escribe el número complejo $ce^{-j\omega}$ como

$$\begin{aligned} ce^{-j\omega} &= [\operatorname{Re}(c) + j \operatorname{Im}(c)](\cos \omega - j \operatorname{sen} \omega) \\ &= [\operatorname{Re}(c) \cos \omega + \operatorname{Im}(c) \operatorname{sen} \omega] + j [\operatorname{Im}(c) \cos \omega - \operatorname{Re}(c) \operatorname{sen} \omega], \end{aligned}$$

y se observa que

$$\frac{d}{d\omega} \operatorname{Re}(ce^{-j\omega}) = -\operatorname{Re}(c) \operatorname{sen} \omega + \operatorname{Im}(c) \cos \omega = \operatorname{Im}(ce^{-j\omega}) \quad (3.20)$$

y

$$\frac{d}{d\omega} \operatorname{Im}(ce^{-j\omega}) = -\operatorname{Im}(c) \operatorname{sen} \omega - \operatorname{Re}(c) \cos \omega = -\operatorname{Re}(ce^{-j\omega}). \quad (3.21)$$

Por lo tanto, la derivada en la ecuación 3.19 es

$$\begin{aligned} \frac{d}{d\omega} \left[\frac{\operatorname{Im}(ce^{-j\omega})}{1 - \operatorname{Re}(ce^{-j\omega})} \right] &= \frac{\left[\frac{d}{d\omega} \operatorname{Im}(ce^{-j\omega}) \right] [1 - \operatorname{Re}(ce^{-j\omega})] - \operatorname{Im}(ce^{-j\omega}) \left[-\frac{d}{d\omega} \operatorname{Re}(ce^{-j\omega}) \right]}{[1 - \operatorname{Re}(ce^{-j\omega})]^2} \\ &\stackrel{(a)}{=} \frac{-\operatorname{Re}(ce^{-j\omega})[1 - \operatorname{Re}(ce^{-j\omega})] - \operatorname{Im}(ce^{-j\omega}) [-\operatorname{Im}(ce^{-j\omega})]}{[1 - \operatorname{Re}(ce^{-j\omega})]^2} \\ &= \frac{[\operatorname{Re}(ce^{-j\omega})]^2 + [\operatorname{Im}(ce^{-j\omega})]^2 - \operatorname{Re}(ce^{-j\omega})}{[1 - \operatorname{Re}(ce^{-j\omega})]^2}, \end{aligned}$$

donde en (a) se emplearon los resultados de las ecuaciones 3.20 y 3.21. Sustituyendo este resultado en la ecuación 3.19, se obtiene que

$$\begin{aligned} \operatorname{grd}(1 - ce^{-j\omega}) &= \frac{1}{1 + \left[\frac{\operatorname{Im}(ce^{-j\omega})}{1 - \operatorname{Re}(ce^{-j\omega})} \right]^2} \times \frac{[\operatorname{Re}(ce^{-j\omega})]^2 + [\operatorname{Im}(ce^{-j\omega})]^2 - \operatorname{Re}(ce^{-j\omega})}{[1 - \operatorname{Re}(ce^{-j\omega})]^2} \\ &= \frac{[\operatorname{Re}(ce^{-j\omega})]^2 + [\operatorname{Im}(ce^{-j\omega})]^2 - \operatorname{Re}(ce^{-j\omega})}{[1 - \operatorname{Re}(ce^{-j\omega})]^2 + [\operatorname{Im}(ce^{-j\omega})]^2} \\ &= \frac{[\operatorname{Re}(ce^{-j\omega})]^2 + [\operatorname{Im}(ce^{-j\omega})]^2 - \operatorname{Re}(ce^{-j\omega})}{1 - 2 \operatorname{Re}(ce^{-j\omega}) + [\operatorname{Re}(ce^{-j\omega})]^2 + [\operatorname{Im}(ce^{-j\omega})]^2}. \end{aligned}$$

Finalmente, considerando que

$$[\operatorname{Re}(ce^{-j\omega})]^2 + [\operatorname{Im}(ce^{-j\omega})]^2 = |ce^{-j\omega}|^2 = |c|^2,$$

resulta en que

$$\text{grd}(1 - ce^{-j\omega}) = \frac{|c|^2 - \text{Re}(ce^{-j\omega})}{1 + |c|^2 - 2\text{Re}(ce^{-j\omega})}, \quad (3.22)$$

que es la forma de los términos correspondientes a los ceros en la ecuación 3.18.

Las ecuaciones 3.16, 3.17 y 3.18 representan respectivamente la magnitud en dB, la fase y el retardo de grupo como la suma de las contribuciones de cada polo y cero de la función del sistema. Para ganar entendimiento en como la ubicación de los polos y los ceros de sistemas estables influyen en la respuesta en frecuencia, es útil considerar en detalle la respuesta en frecuencia de los sistemas de primer y segundo orden en relación a la ubicación de sus polos y ceros.

3.4.1. Respuesta en frecuencia de sistemas de primer orden

En esta sección se examinan las propiedades de un único factor de la forma $(1 - re^{j\theta}e^{-j\omega})$, donde r es la magnitud y θ es el ángulo de un cero o un polo en el plano z . Esto es equivalente a lo que se hizo para el retardo de grupo en la sección anterior que condujo a la ecuación 3.22, pero allí se expresó el polos en la parte real e imaginaria.

La magnitud al cuadrado de dicho factor es

$$\begin{aligned} |1 - re^{j\theta}e^{-j\omega}|^2 &= (1 - re^{j\theta}e^{-j\omega})(1 - re^{-j\theta}e^{j\omega}) \\ &= 1 - re^{-j\theta}e^{j\omega} - re^{j\theta}e^{-j\omega} + r^2 \\ &= 1 + r^2 - 2r \cos(\omega - \theta). \end{aligned} \quad (3.23)$$

La ganancia en dB asociada con ese factor es

$$(+/-)20 \log_{10} |1 - re^{j\theta}e^{-j\omega}| = (+/-)10 \log_{10} [1 + r^2 - 2r \cos(\omega - \theta)], \quad (3.24)$$

con signo positivo si el factor representa a un cero y signo negativo si el factor representa a un polo.

Expresando al factor como

$$1 - re^{j\theta}e^{-j\omega} = 1 - re^{-j(\omega - \theta)} = 1 - r \cos(\omega - \theta) + jr \sin(\omega - \theta),$$

se deduce que el aporte a la respuesta en fase del factor es

$$(+/-) \text{ARG}(1 - re^{j\theta}e^{-j\omega}) = (+/-) \arctan \left[\frac{r \sin(\omega - \theta)}{1 - r \cos(\omega - \theta)} \right]. \quad (3.25)$$

El retardo de grupo para el factor de signo positivo correspondiente a un cero es

$$\begin{aligned} \text{grd}(1 - re^{j\theta}e^{-j\omega}) &= -\frac{d}{d\omega} \arctan \left[\frac{r \sin(\omega - \theta)}{1 - r \cos(\omega - \theta)} \right] \\ &= -\frac{1}{1 + \left[\frac{r \sin(\omega - \theta)}{1 - r \cos(\omega - \theta)} \right]^2} \\ &\quad \times \frac{r \cos(\omega - \theta)[1 - r \cos(\omega - \theta)] - r \sin(\omega - \theta)r \sin(\omega - \theta)}{[1 - r \cos(\omega - \theta)]^2} \\ &= -\frac{r \cos(\omega - \theta) - r^2 \cos^2(\omega - \theta) - r^2 \sin^2(\omega - \theta)}{[1 - r \cos(\omega - \theta)]^2 + r^2 \sin^2(\omega - \theta)}, \end{aligned}$$

resultando en que

$$\text{grd}(1 - re^{j\theta}e^{-j\omega}) = -\frac{d}{d\omega} \arctan \left[\frac{r \sin(\omega - \theta)}{1 - r \cos(\omega - \theta)} \right] = \frac{r^2 - r \cos(\omega - \theta)}{1 + r^2 - 2r \cos(\omega - \theta)}. \quad (3.26)$$

Observar que este resultado es idéntico al de la ecuación 3.22, teniendo en cuenta que con $c = re^{j\theta}$, $|c|^2 = r^2$ y $\text{Re}(ce^{-j\omega}) = r \cos(\omega - \theta)$. Finalmente, considerando también el término correspondiente a un polo,

$$(+/-) \text{grd}(1 - re^{j\theta}e^{-j\omega}) = (+/-) \frac{r^2 - r \cos(\omega - \theta)}{1 + r^2 - 2r \cos(\omega - \theta)} = (+/-) \frac{r^2 - r \cos(\omega - \theta)}{|1 - re^{j\theta}e^{-j\omega}|^2}, \quad (3.27)$$

donde nuevamente el signo positivo corresponde a un cero y el signo negativo corresponde a un polo.

Las funciones de las ecuaciones 3.24, 3.25 y 3.27 son periódicas con período 2π . En la figura 3.12 se muestran dichas funciones sobre un período ($0 \leq \omega < 2\pi$) para distintos valores de θ y $r = 0.9$. Se

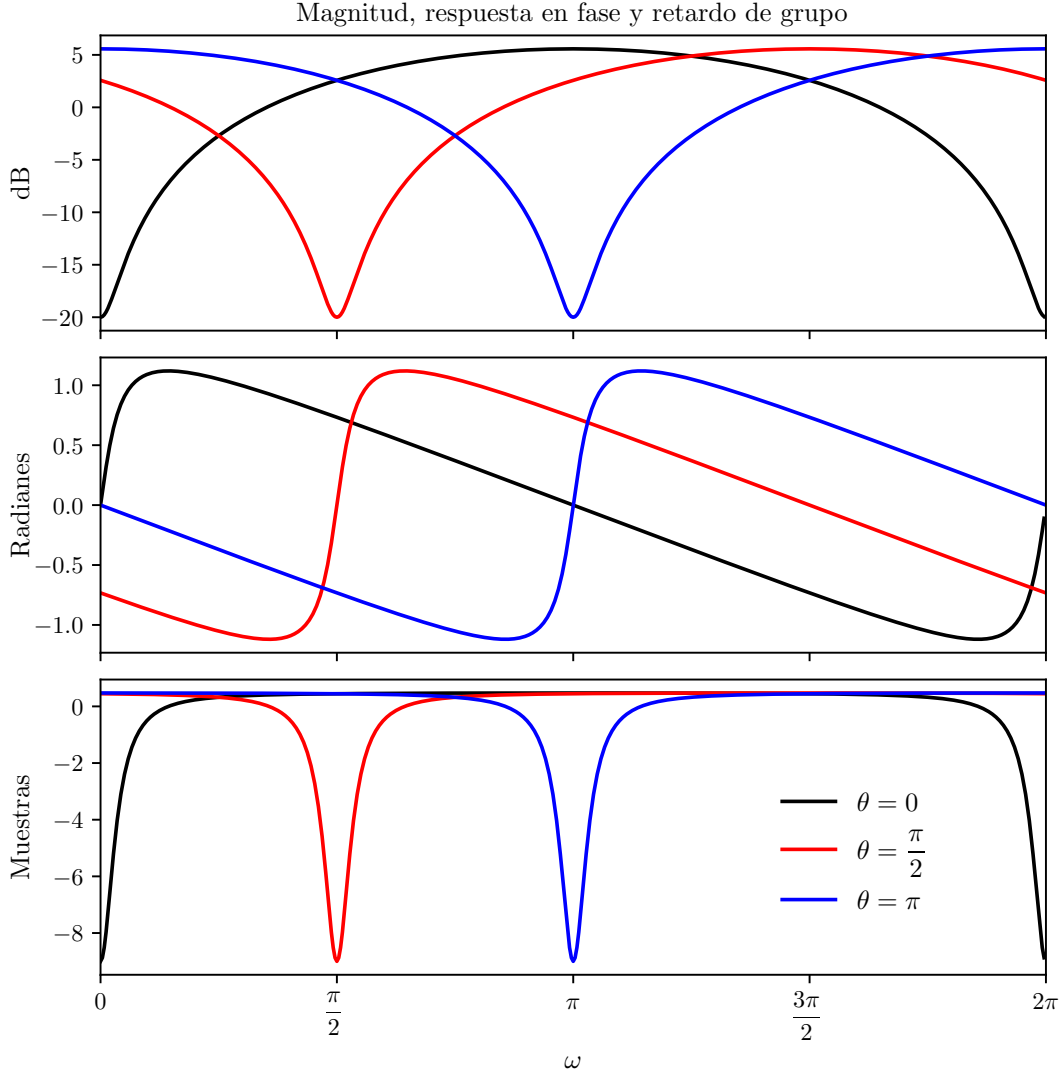


Figura 3.12: Respuesta en frecuencia para un único cero, con $r = 0.9$ y los valores de θ indicados en la figura.

observa que la magnitud tiene valles en $\omega = \theta$. También se aprecia en la respuesta en fase que la fase es nula en $\omega = \theta$ y tiene pendiente positiva máxima en dicho valor. De forma correspondiente, en las funciones del retardo de grupo se observa que hay un pico negativo en $\omega = \theta$.

Las características de la respuesta en frecuencia pueden inferirse a partir del diagrama de ceros y polos del sistema. Cada factor correspondiente a un cero o a un polo puede ser representado en plano z por un vector desde el cero o polo hasta un punto en la circunferencia unidad. En el caso de un sistema de primer orden de la forma

$$H(z) = 1 - re^{j\theta}z^{-1} = \frac{z - re^{j\theta}}{z}, \quad r < 1$$

el diagrama de ceros y polos se muestra en la figura 3.13. En la figura se indican los vectores v_1 , v_2 y $v_3 = v_1 - v_2$ representando los números complejos $e^{j\omega}$, $re^{j\theta}$ y $e^{j\omega} - re^{j\theta}$. La magnitud de la respuesta

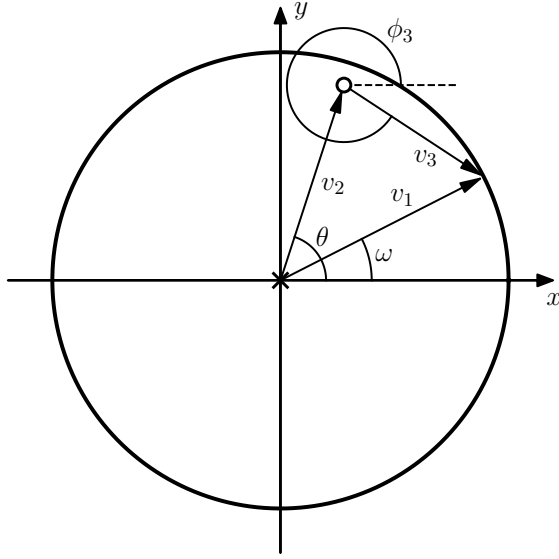


Figura 3.13: Diagrama de polos y ceros de un sistema de primer orden con cero en $z = re^{j\theta}$ con $r < 1$.

en frecuencia del sistema de primer orden en términos de estos vectores es

$$|1 - re^{j\theta}e^{-j\omega}| = \left| \frac{e^{j\omega} - re^{j\theta}}{e^{j\omega}} \right| = \frac{|v_3|}{|v_1|} = |v_3|.$$

Esto indica que la contribución de un único factor $1 - re^{j\theta}e^{-j\omega}$ a la magnitud de la respuesta en frecuencia es la longitud del vector v_3 desde el cero al punto $z = e^{j\omega}$ en la circunferencia unidad. Este vector tiene longitud mínima cuando $\omega = \theta$, acorde con los valles en la magnitud de la respuesta en frecuencia en la figura 3.12. La fase es

$$\angle(1 - re^{j\theta}e^{-j\omega}) = \angle(e^{j\omega} - re^{j\theta}) - \angle(e^{j\omega}) = \angle v_3 - \angle v_1 = \phi_3 - \omega.$$

Del diagrama de polos y ceros se deduce que cuando $\omega = \theta$, los vectores v_1 , v_2 y v_3 son colineales, y entonces también $\phi_3 = \theta$, resultando en que la fase es nula.

La dependencia del radio r de la contribución del factor de primer orden a la respuesta en frecuencia se muestra en la figura 3.14 para $\theta = \pi$ y varios valores de r . Notar que el valle en la magnitud en decibels es mas pronunciado a medida que r se acerca a 1. De hecho, la magnitud en dB tiende a $-\infty$ en $\omega = \theta$ cuando r tiende a 1. La fase tiene pendiente positiva en torno a $\omega = \theta$, y se hace infinita cuando r tiende a 1. Es decir, para $r = 1$ la fase es discontinua, con un salto de π radianes en $\omega = \theta$. El retardo de grupo, que es el opuesto de la pendiente de la fase, tiene un valle cada vez mas pronunciado en $\omega = \theta$ cuando r se acerca a 1, y se convierte en un impulso en $r = 1$, no mostrado en la figura. En la figura se aprecia también que el retardo de grupo es relativamente plano en los valores de ω lejanos a θ .

3.4.2. Ejemplos con múltiples ceros y polos

En esta sección se emplea y se expande la discusión de la sección 3.4.1 para determinar la respuesta en frecuencia de sistemas con función de transferencia racional.

Ejemplo: sistema IIR de segundo orden Se considera el sistema de segundo orden

$$H(z) = \frac{1}{(1 - re^{j\theta}z^{-1})(1 - re^{-j\theta}z^{-1})} = \frac{1}{1 - 2r \cos \theta z^{-1} + r^2 z^{-2}}.$$

Se calculará la respuesta al impulso del sistema causal, lo cual puede hacerse mediante el método de descomposición en fracciones simples. La función del sistema puede expresarse en la forma de la ecuación 2.8 como

$$H(z) = \frac{A_1}{1 - re^{j\theta}z^{-1}} + \frac{A_2}{1 - re^{-j\theta}z^{-1}},$$

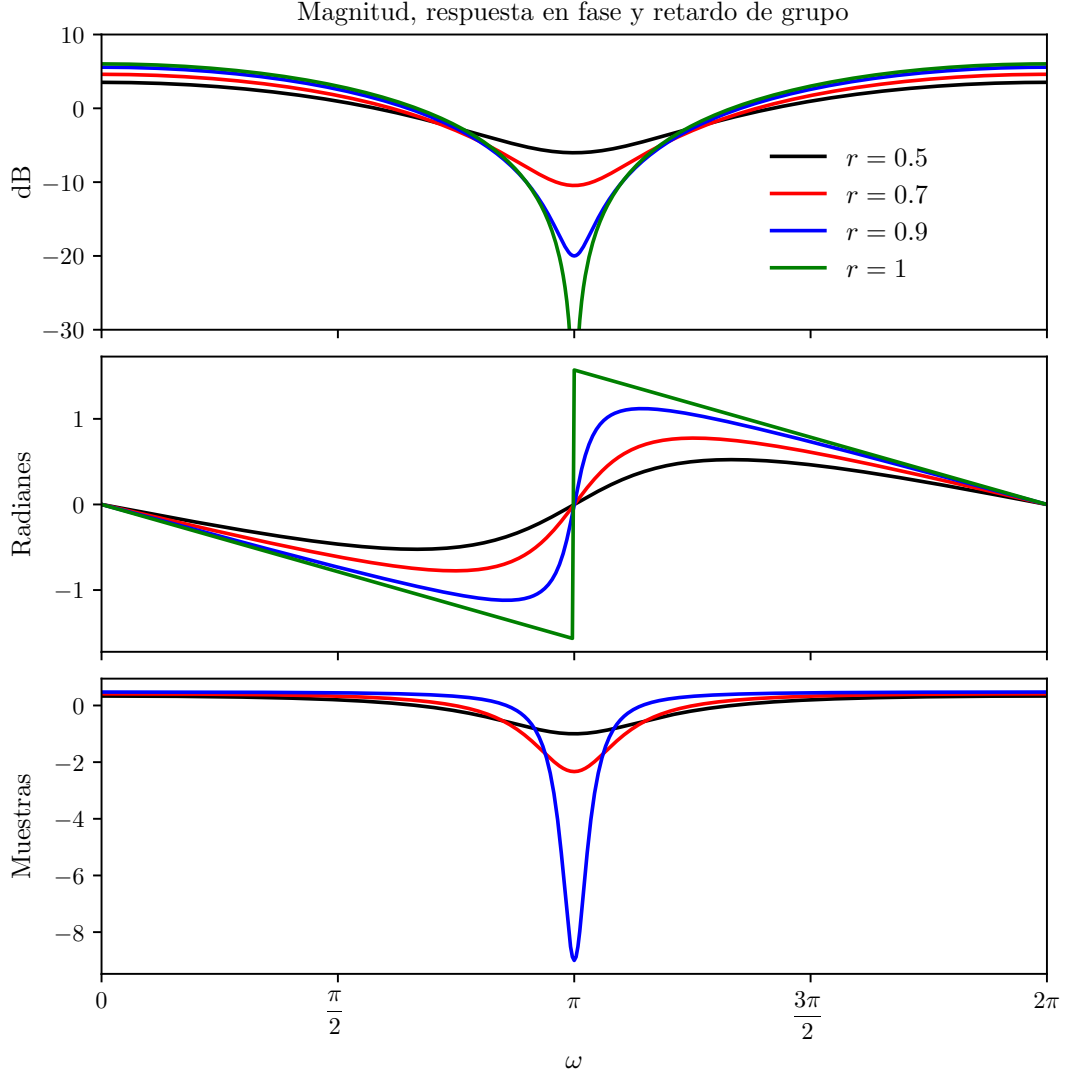


Figura 3.14: Respuesta en frecuencia para un único cero, con $\theta = \pi$ y los valores de r indicados en la figura.

donde los coeficientes A_1 y A_2 pueden obtenerse empleando la ecuación 2.9. De esta forma,

$$A_1 = (1 - re^{j\theta}z^{-1})H(z)|_{z=re^{j\theta}} = \frac{1}{1 - re^{-j\theta}(re^{j\theta})^{-1}} = \frac{1}{1 - e^{-2j\theta}} = \frac{e^{j\theta}}{e^{j\theta} - e^{-j\theta}} = \frac{e^{j\theta}}{2j \sin \theta}.$$

y

$$A_2 = (1 - re^{-j\theta}z^{-1})H(z)|_{z=re^{-j\theta}} = \frac{1}{1 - re^{j\theta}(re^{-j\theta})^{-1}} = \frac{1}{1 - e^{2j\theta}} = \frac{e^{-j\theta}}{e^{-j\theta} - e^{j\theta}} = -\frac{e^{-j\theta}}{2j \sin \theta}.$$

Luego, aplicando la transformada inversa considerando la ROC $|z| > r$ para que el sistema sea causal (ver la tabla 2.1), se obtiene que

$$\begin{aligned} h[n] &= \frac{e^{j\theta}}{2j \sin \theta} r^n e^{jn\theta} u[n] - \frac{e^{-j\theta}}{2j \sin \theta} r^n e^{-jn\theta} u[n] \\ &= \frac{r^n}{\sin \theta} \left[\frac{e^{j(n+1)\theta} - e^{-j(n+1)\theta}}{2j} \right] u[n] \end{aligned}$$

resultando en

$$h[n] = \frac{r^n \sin[(n+1)\theta]}{\sin \theta} u[n]$$

La función del sistema tiene un par de polos complejos conjugados en $z = re^{\pm j\theta}$ y dos ceros en $z = 0$. El diagrama de polos y ceros se muestra en la figura 3.15. De los resultados de la sección 3.4.1 se deduce

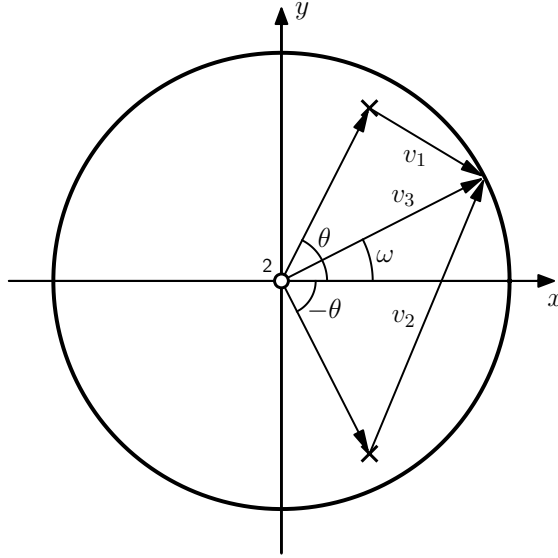


Figura 3.15: Diagrama de polos y ceros del sistema de segundo orden con polos en $z = re^{\pm j\theta}$ con $r < 1$.

que la respuesta en magnitud en dB es

$$20 \log_{10} |H(e^{j\omega})| = -10 \log_{10}[1 + r^2 - 2r \cos(\omega - \theta)] - 10 \log_{10}[1 + r^2 - 2r \cos(\omega + \theta)],$$

la respuesta en fase es

$$\angle H(e^{j\omega}) = -\arctan \left[\frac{r \sin(\omega - \theta)}{1 - r \cos(\omega - \theta)} \right] - \arctan \left[\frac{r \sin(\omega + \theta)}{1 - r \cos(\omega + \theta)} \right]$$

y el retardo de grupo es

$$\text{grd } H(e^{j\omega}) = -\frac{r^2 - r \cos(\omega - \theta)}{1 + r^2 - 2r \cos(\omega - \theta)} - \frac{r^2 - r \cos(\omega + \theta)}{1 + r^2 - 2r \cos(\omega + \theta)}.$$

Estas funciones se muestran en la figura 3.16 para el valor de $r = 0.9$ y varios valores de θ . La figura 3.15 muestra los vectores de ceros y polos v_1 , v_2 y v_3 . La respuesta en magnitud puede expresarse en función de estos vectores como

$$|H(e^{j\omega})| = \frac{|e^{j\omega}|^2}{|e^{j\omega} - re^{j\theta}||e^{j\omega} - re^{-j\theta}|} = \frac{|v_3|^2}{|v_1||v_2|} = \frac{1}{|v_1||v_2|}.$$

Esto indica que la respuesta en magnitud es el producto de las longitudes de los vectores de los ceros, que en este caso son la unidad, dividido el producto de los vectores de los polos. Cuando $\omega \approx \theta$, la longitud del vector $v_1 = e^{j\omega} - re^{j\theta}$ se hace pequeña y cambia significativamente cuando ω varía en torno a θ , mientras que la longitud del vector $v_2 = e^{j\omega} - re^{-j\theta}$ no cambia significativamente cuando ω varía en torno a θ . Por lo tanto, el polo con ángulo θ domina en la respuesta en frecuencia en torno a $\omega = \theta$, como se observa en la respuesta en magnitud en la figura 3.16. Por simetría, el polo con ángulo $-\theta$ domina en la respuesta en frecuencia en torno a $\omega = -\theta$.

3.5. Relación entre la magnitud y la fase

En general el conocimiento de la magnitud de la respuesta en frecuencia de un sistema LTI no brinda conocimiento de la fase y viceversa. Sin embargo, para sistemas descritos por ecuaciones en diferencias,

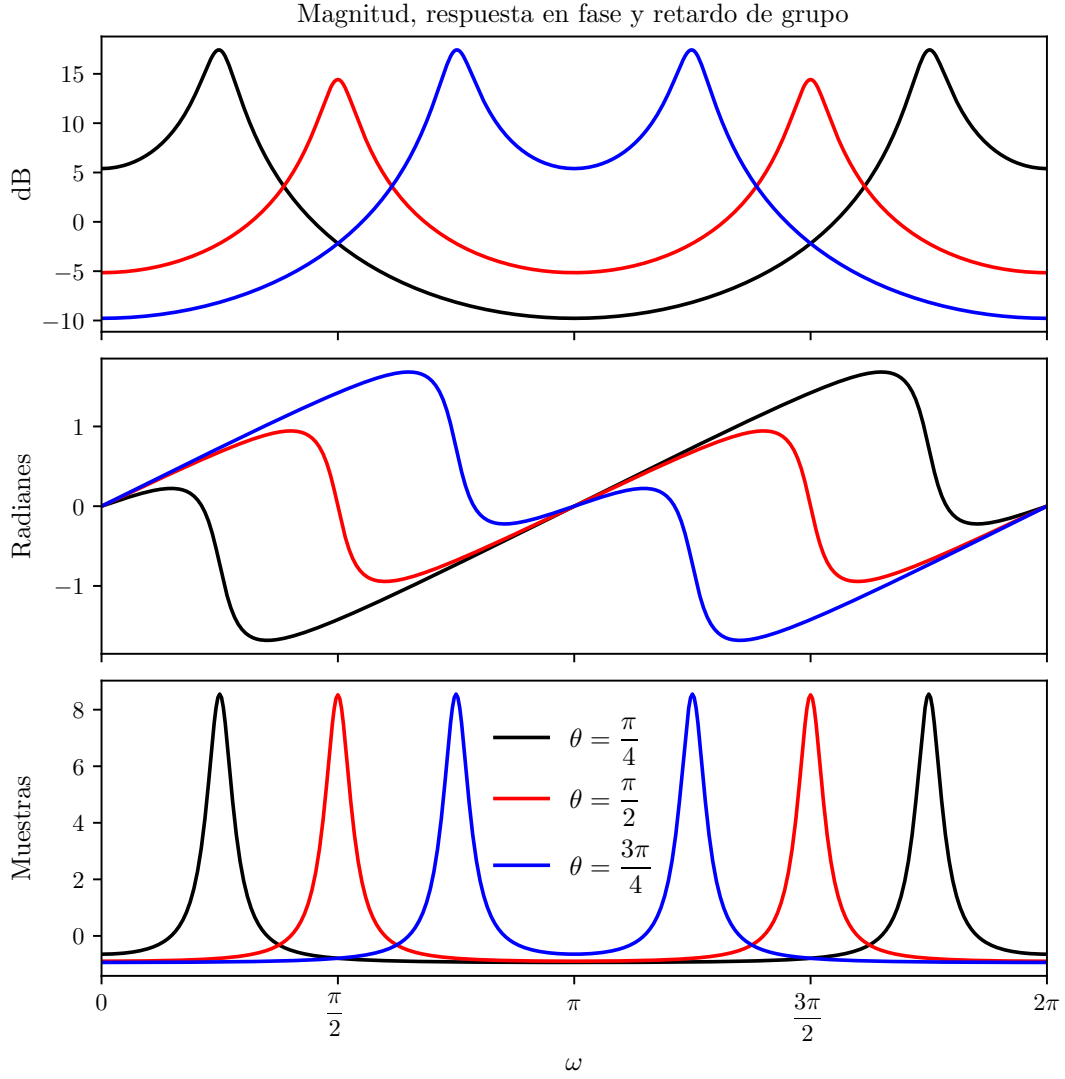


Figura 3.16: Respuesta en frecuencia del sistema de segundo orden, con $r = 0.9$ y los valores de θ indicados en la figura.

es decir, para sistemas con función del sistema racional, hay algunas restricciones entre la magnitud y la fase. En particular, como se verá en esta sección, si se conoce la magnitud de la respuesta en frecuencia y el número de polos y ceros, hay solo una cantidad finita de posibilidades de la fase asociada. De forma similar, si se conoce la cantidad de polos y ceros y la fase, hay solo un número finito de elecciones de la magnitud, salvo un factor de escala. Es mas, si se impone la restricción referida como fase mínima, la magnitud de la respuesta en frecuencia especifica a la fase unívocamente y la respuesta en fase especifica a la respuesta en magnitud excepto por un factor de escala.

Para explorar las elecciones posibles de la función del sistema dado el cuadrado de la respuesta en magnitud, se parte considerando que

$$\begin{aligned} |H(e^{j\omega})|^2 &= H(e^{j\omega})H^*(e^{j\omega}) \\ &= H(z)H^*(1/z^*) \Big|_{z=e^{j\omega}}, \end{aligned} \quad (3.28)$$

donde la última igualdad proviene de que

$$H^*(1/z^*) = \left[\sum_{n=-\infty}^{\infty} h[n] \left(\frac{1}{z^*} \right)^{-n} \right]^* = \sum_{n=-\infty}^{\infty} h^*[n] z^n$$

y por lo tanto,

$$H^*(1/z^*)|_{z=e^{j\omega}} = \sum_{n=-\infty}^{\infty} h^*[n] e^{j\omega n} = \left(\sum_{n=-\infty}^{\infty} h[n] e^{-j\omega n} \right)^* = H^*(e^{j\omega}).$$

Restringiendo a la función del sistema a una función racional, puede expresarse de la forma de la ecuación 3.13,

$$H(z) = \left(\frac{b_0}{a_0} \right) \frac{\prod_{k=1}^M (1 - c_k z^{-1})}{\prod_{k=1}^N (1 - d_k z^{-1})}$$

y por lo tanto,

$$H^* \left(\frac{1}{z^*} \right) = \left(\frac{b_0}{a_0} \right) \frac{\prod_{k=1}^M (1 - c_k^* z)}{\prod_{k=1}^N (1 - d_k^* z)}$$

asumiendo que a_0 y b_0 son reales. Considerando estos resultados, la ecuación 3.28 indica que la magnitud al cuadrado es la evaluación en la circunferencia unidad de la transformada z

$$\begin{aligned} C(z) &= H(z) H^*(1/z^*) \\ &= \left(\frac{b_0}{a_0} \right)^2 \frac{\prod_{k=1}^M (1 - c_k z^{-1})(1 - c_k^* z)}{\prod_{k=1}^N (1 - d_k z^{-1})(1 - d_k^* z)}. \end{aligned} \quad (3.29)$$

Si se conoce la expresión de $|H(e^{j\omega})|^2$ como función de $e^{j\omega}$, reemplazando $e^{j\omega}$ por z se obtiene la expresión de $C(z)$, y a partir de $C(z)$ se quiere extraer la mayor cantidad de información posible sobre $H(z)$. Para hacerlo, se observa que por cada polo d_k de $H(z)$ hay un polo de $C(z)$ en d_k y en $(d_k^*)^{-1}$, y también por cada cero c_k de $H(z)$ hay un cero de $C(z)$ en c_k y en $(c_k^*)^{-1}$. En consecuencia, los polos y ceros de $C(z)$ ocurren en pares conjugados inversos, donde un elemento del par lo contribuye $H(z)$ y el otro lo contribuye $H^*(1/z^*)$. Observar que si un elemento del par se encuentra dentro de círculo unidad, entonces el otro se encuentra fuera del círculo unidad. La otra alternativa es que ambos se ubiquen sobre la circunferencia unidad en la misma posición.

Si se asume que $H(z)$ es la función de un sistema causal y estable, todos sus polos deben estar contenidos dentro del círculo unidad. Por lo tanto, con esta restricción, los polos de $H(z)$ pueden identificarse de los polos de $C(z)$. Sin embargo, los ceros de $H(z)$ no pueden identificarse unívocamente de los polos de $C(z)$.

3.6. Sistemas pasa todo

Una función estable de la forma

$$H_{\text{ap}}(z) = \frac{z^{-1} - a^*}{1 - a z^{-1}} \quad (3.30)$$

tiene una respuesta en frecuencia con magnitud independiente de ω . Esto puede verse escribiendo la respuesta en frecuencia como

$$H_{\text{ap}}(e^{j\omega}) = \frac{e^{-j\omega} - a^*}{1 - a e^{-j\omega}} = e^{-j\omega} \frac{1 - a^* e^{j\omega}}{1 - a e^{-j\omega}} \quad (3.31)$$

y observando que en la expresión de la derecha, el término $e^{-j\omega}$ tiene magnitud unidad, y la fracción también tiene magnitud unidad debido a que el numerador y el denominador son complejos conjugados. En consecuencia, $|H_{ap}(e^{j\omega})| = 1$. Este sistema tiene un polo en a y un cero en $1/a^*$, es decir, el polo y el cero son un par de números complejos conjugados inversos. Un sistema que tiene la magnitud de la frecuencia constante es referido como sistema pasa todo, y permite pasar a todos los componentes espectrales de la entrada con igual ganancia.

La forma mas general de un sistema pasa todo con respuesta al impulso real es el producto de factores como los de la ecuación 3.30, con los polos complejos en pares conjugados, es decir,

$$H_{ap}(z) = A \prod_{k=1}^{M_r} \frac{z^{-1} - d_k}{1 - d_k z^{-1}} \prod_{k=1}^{M_c} \frac{(z^{-1} - e_k^*)(z^{-1} - e_k)}{(1 - e_k z^{-1})(1 - e_k^* z^{-1})}, \quad (3.32)$$

donde A es una constante positiva, d_k son los polos reales y e_k son los polos complejos. Para que el sistema sea causal y estable se requiere que $|d_k| < 1$ y $|e_k| < 1$.

La respuesta en frecuencia de un sistema pasa todo general puede expresarse en términos de las respuestas en frecuencia de los sistemas de primer orden que lo componen, dados por la ecuación 3.30. Para sistema causales, cada término consiste en un único polo dentro de la circunferencia unidad y un cero en la posición conjugada inversa. La respuesta en magnitud de cada término es la unidad, como ya fue mostrado. Con a expresado en forma polar como $a = re^{j\theta}$, la función de fase de 3.30 es

$$\angle \left[\frac{e^{-j\omega} - re^{-j\theta}}{1 - re^{j\theta} e^{-j\omega}} \right] = -\omega - 2 \arctan \left[\frac{r \sin(\omega - \theta)}{1 - r \cos(\omega - \theta)} \right], \quad (3.33)$$

resultado que surge de expresar la respuesta en frecuencia como en la ecuación 3.31 y considerando el resultado de la ecuación 3.25 teniendo en cuenta que la función arcotangente es impar. De forma análoga se deduce que la fase de un sistema de pasa todo de segundo orden con polos en $z = re^{j\theta}$ y $z = re^{-j\theta}$ es

$$\angle \left[\frac{(e^{-j\omega} - re^{-j\theta})(e^{-j\omega} - re^{j\theta})}{(1 - re^{j\theta} e^{-j\omega})(1 - re^{-j\theta} e^{-j\omega})} \right] = -\omega - 2 \arctan \left[\frac{r \sin(\omega - \theta)}{1 - r \cos(\omega - \theta)} \right] - 2 \arctan \left[\frac{r \sin(\omega + \theta)}{1 - r \cos(\omega + \theta)} \right],$$

Ejemplo: sistemas pasa todo de primer y segundo orden La figura 3.17 muestra las características de la respuesta en frecuencia de dos sistemas pasa todo de primer orden, uno con el polo en $z = 0.9$ y otro con el polo en $z = -0.9$. También se muestra la repuesta en frecuencia de un sistema pasa todo de segundo orden con polos en $z = 0.9e^{j\pi/4}$ y $z = 0.9e^{-j\pi/4}$.

El ejemplo ilustra una propiedad general de los sistemas pasa todo. En la figura 3.17 se observa que si se eliminan las discontinuidades de 2π radianes que se producen en la frecuencia de los polos en la respuesta en fase, la respuesta en fase es no positiva en el intervalo $0 < \omega < \pi$, o en otras palabras, la fase desenrollada es no negativa en dicho intervalo. Como el pasa todo general dado por la ecuación 3.32 es el producto de sistemas pasa todo de primer y segundo orden, es posible concluir que la respuesta en fase desenrollada $\arg[H_{ap}(e^{j\omega})]$ de un sistema pasa todo causal es siempre no positiva en $0 < \omega < \pi$. Este resultado puede establecerse analíticamente considerando el retardo de grupo. El retardo de grupo del sistema pasa todo de primer orden 3.30 es la función opuesta a la derivada de la fase, dada por la ecuación 3.33, es decir,

$$\begin{aligned} \text{grd} \left[\frac{e^{-j\omega} - re^{-j\theta}}{1 - re^{j\theta} e^{-j\omega}} \right] &= -\frac{d}{d\omega} \angle \left[\frac{e^{-j\omega} - re^{-j\theta}}{1 - re^{j\theta} e^{-j\omega}} \right] \\ &\stackrel{(a)}{=} 1 - 2 \left[\frac{r^2 - r \cos(\omega - \theta)}{1 + r^2 - 2r \cos(\omega - \theta)} \right] \\ &= \frac{1 + r^2 - 2r \cos(\omega - \theta) - 2r^2 + 2r \cos(\omega - \theta)}{1 + r^2 - 2r \cos(\omega - \theta)} \\ &= \frac{1 - r^2}{1 + r^2 - 2r \cos(\omega - \theta)}, \end{aligned}$$

donde en (a) se empleó el resultado de la ecuación 3.26. Finalmente, considerando el resultado de la ecuación 3.23, resulta en que

$$\text{grd} \left[\frac{e^{-j\omega} - re^{-j\theta}}{1 - re^{j\theta} e^{-j\omega}} \right] = \frac{1 - r^2}{1 + r^2 - 2r \cos(\omega - \theta)} = \frac{1 - r^2}{|1 - re^{j\theta} e^{-j\omega}|^2}. \quad (3.34)$$

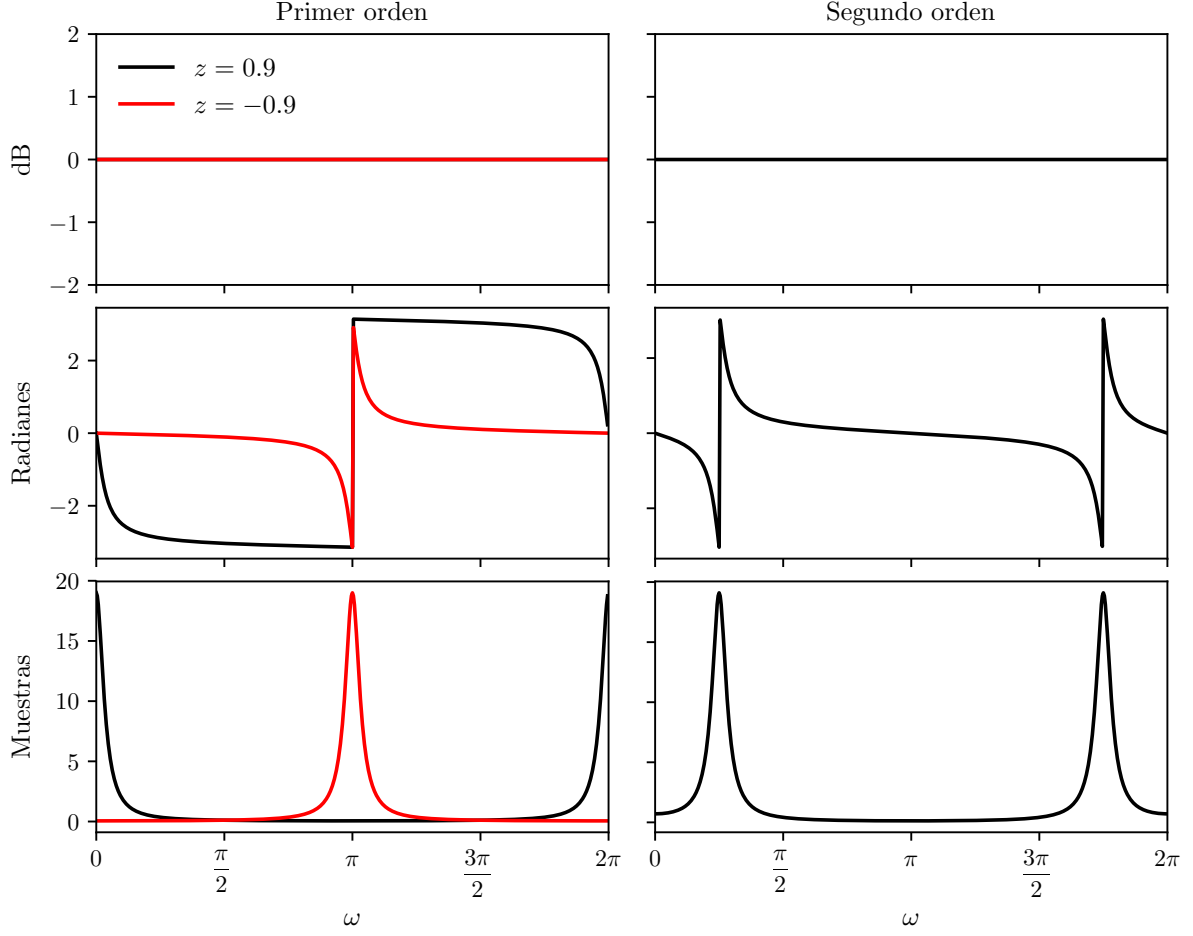


Figura 3.17: Sistemas pasa todo. Respuesta en magnitud, respuesta en fase y retardo de grupo de sistemas pasa todo de primer y segundo orden.

Como $r < 1$ para un sistema pasa todo causal y estable, el retardo de grupo contribuido por un único factor pasa todo es siempre positivo. Como el retardo de grupo de sistemas pasa todo de mayor orden es la suma de los retardos de grupo de los factores de primer orden, será siempre positivo en general.

La positividad del retardo de grupo de los sistemas pasa todo causales es la base de una prueba simple de la no negatividad de la fase de dichos sistemas. Efectivamente, de la definición de retardo de grupo y el teorema fundamental del cálculo, se cumple que

$$\arg[H_{\text{ap}}(e^{j\omega})] = - \int_0^\omega \text{grd}[H_{\text{ap}}(e^{j\phi})] d\phi + \arg[H_{\text{ap}}(e^{j0})] \quad (3.35)$$

en el intervalo $0 < \omega < \pi$. De la ecuación 3.32,

$$\arg[H_{\text{ap}}(e^{j0})] = A \prod_{k=1}^{M_r} \frac{1-d_k}{1-d_k} \prod_{k=1}^{M_c} \frac{(1-e_k^*)(1-e_k)}{(1-e_k)(1-e_k^*)} = A \prod_{k=1}^{M_r} \frac{1-d_k}{1-d_k} \prod_{k=1}^{M_c} \frac{|1-e_k|^2}{|1-e_k|^2} = A.$$

Por lo tanto, $\arg[H_{\text{ap}}(e^{j0})] = 0$, y como

$$\text{grd}[H_{\text{ap}}(e^{j\omega})] \geq 0$$

de la ecuación 3.35 se concluye que

$$\arg[H_{\text{ap}}(e^{j\omega})] \leq 0 \quad \text{para} \quad 0 \leq \omega < \pi.$$

La positividad del retardo de grupo y la no positividad de la fase desenrollada son propiedades importantes de los sistemas pasa todo causales.

Los sistemas pasa todo tienen importancia en diversos contextos. Pueden emplearse para contrarrestar el efecto de la distorsión de fase. También son útiles para transformar filtros pasabajos en los otros tipos de filtros de selección de frecuencias. Estos aspectos se verán en el capítulo ?? . Además, también son útiles en la teoría de los sistemas de fase mínima, como se verá en la sección 3.7

3.7. Sistemas de fase mínima

En la sección 3.5 se mostró que la magnitud de la respuesta en frecuencia de un sistema LTI con función del sistema racional no caracteriza unívocamente al sistema. Si el sistema es causal y estable, los polos deben estar dentro de la circunferencia unidad, pero dichas condiciones no imponen restricciones sobre los ceros. En ciertas clases de problemas es útil imponer la condición adicional de que el sistema inverso también sea causal y estable. Como se discutió en la sección 3.3.2, esta condición restringe a que tanto los polos como los ceros estén contenidos en el círculo unidad. Esta clase de sistemas es referida como *sistemas de fase mínima*. El nombre fase mínima proviene de una propiedad de la respuesta en fase que no es obvia a partir de la definición precedente. Esta y otras propiedades, que se verán en esta sección, son únicas de esta clase de sistemas y por lo tanto, cada una de ellas puede considerarse como la definición de la clase.

Si se conoce el cuadrado de la respuesta en magnitud en la forma de la ecuación 3.29 y se sabe que tanto el sistema como el sistema inverso son causales y estables, $H(z)$ queda unívocamente determinado y consistirá en los polos y ceros de $C(z) = H(z)H^*(1/z^*)$ contenidos dentro del círculo unidad. Este enfoque es empleado en diseño de filtros cuando únicamente es especificada la respuesta en magnitud por el método de diseño empleado (ver el capítulo ??).

3.7.1. Descomposición en fase mínima y pasa todo

En la sección 3.5 se vio que a partir de la respuesta en magnitud al cuadrado no es posible determinar unívocamente la función del sistema $H(z)$, ya que cualquier elección de respuesta en frecuencia con la magnitud dada puede ponerse en cascada con un número arbitrario de factores pasa todo sin afectar la magnitud. Una observación relacionada es que cualquier función del sistema racional puede expresarse como

$$H(z) = H_{\min}(z)H_{\text{ap}}(z), \quad (3.36)$$

donde $H_{\min}(z)$ es un sistema de fase mínima y $H_{\text{ap}}(z)$ es un sistema pasa todo.

Para mostrar esto, supóngase que $H(z)$ tiene un cero fuera del círculo unidad en $z = 1/c^*$, donde $|c| < 1$, y el resto de los polos y ceros se encuentran dentro del círculo unidad. De esta forma, $H(z)$ puede expresarse como

$$H(z) = H_1(z)(z^{-1} - c^*), \quad (3.37)$$

donde por definición, $H_1(z)$ es de fase mínima. Una expresión equivalente para $H(z)$ es

$$H(z) = H_1(z)(1 - cz^{-1}) \frac{z^{-1} - c^*}{1 - cz^{-1}}. \quad (3.38)$$

Como $|c| < 1$, el factor $H_1(z)(1 - cz^{-1})$ es de fase mínima y difiere de $H(z)$ únicamente en que el cero de $H(z)$ que estaba fuera del círculo unidad en $z = 1/c^*$ fue reflejado dentro del círculo unidad a la posición conjugada inversa $z = c$. El factor $(z^{-1} - c^*)/(1 - cz^{-1})$ es pasa todo. El ejemplo puede generalizarse para incluir mas ceros fuera del círculo unidad. Por lo tanto, en la ecuación 3.36, $H_{\min}(z)$ contiene todos los polos y ceros de $H(z)$ contenidos dentro del círculo unidad, y además, ceros en la posición conjugada inversa de los ceros de $H(z)$ fuera del círculo unidad. El sistema $H_{\text{ap}}(z)$ se compone de todos los ceros de $H(z)$ fuera del círculo unidad y polos en la posición conjugada inversa que se cancelan con los ceros que fueron reflejados hacia adentro del círculo unidad en $H_{\min}(z)$.

Mediante la ecuación 3.36 se puede construir un sistema de fase no mínima a partir de un sistema de fase mínima reflejando uno o mas ceros dentro del círculo unidad a la posición conjugada inversa fuera del círculo unidad, o viceversa, se puede construir un sistema de fase mínima a partir de un sistema de fase no mínima reflejando los ceros fuera del círculo unidad a la posición conjugada inversa dentro

del círculo unidad. En ambos casos, el sistema de fase mínima y el sistema de fase no mínima tienen la misma respuesta en magnitud.

Ejemplo: descomposición en fase mínima y pasa todo Considérese el sistema

$$H(z) = \frac{\left(1 + \frac{3}{2}e^{j\pi/4}z^{-1}\right)\left(1 + \frac{3}{2}e^{-j\pi/4}z^{-1}\right)}{1 - \frac{1}{3}z^{-1}}.$$

Este sistema tiene dos ceros complejos en

$$z = \frac{3}{2}e^{-j3\pi/4} \quad \text{y} \quad z = \frac{3}{2}e^{j3\pi/4}$$

fuera del círculo unidad y un polo real en $z = 1/3$ dentro del círculo unidad. $H(z)$ puede expresarse en la forma de la ecuación 3.37 como

$$\begin{aligned} H(z) &= \frac{\frac{3}{2}e^{j\pi/4}\left(\frac{2}{3}e^{-j\pi/4} + z^{-1}\right)\frac{3}{2}e^{-j\pi/4}\left(\frac{2}{3}e^{j\pi/4} + z^{-1}\right)}{1 - \frac{1}{3}z^{-1}} \\ &= \frac{9\left(z^{-1} + \frac{2}{3}e^{-j\pi/4}\right)\left(z^{-1} + \frac{2}{3}e^{j\pi/4}\right)}{1 - \frac{1}{3}z^{-1}}, \end{aligned}$$

y factorizando como en la ecuación 3.38 se obtiene que

$$H(z) = \underbrace{\left[\frac{9\left(1 + \frac{2}{3}e^{j\pi/4}z^{-1}\right)\left(1 + \frac{2}{3}e^{-j\pi/4}z^{-1}\right)}{1 - \frac{1}{3}z^{-1}}\right]}_{H_{\min}(z)} \underbrace{\left[\frac{\left(z^{-1} + \frac{2}{3}e^{-j\pi/4}\right)\left(z^{-1} + \frac{2}{3}e^{j\pi/4}\right)}{\left(1 + \frac{2}{3}e^{j\pi/4}z^{-1}\right)\left(1 + \frac{2}{3}e^{-j\pi/4}z^{-1}\right)}\right]}_{H_{\text{ap}}(z)}.$$

El primer sistema entre paréntesis rectos tiene dos ceros en

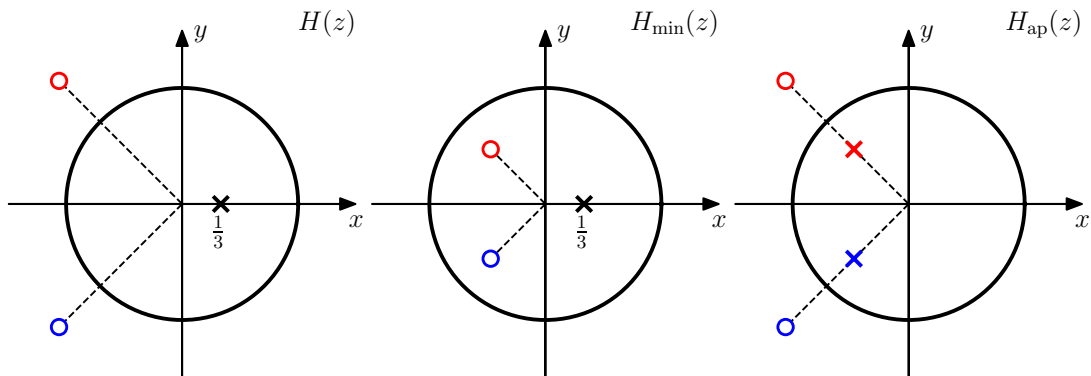


Figura 3.18: Diagrama de polos y ceros del sistema $H(z)$ y del sistema de fase mínima $H_{\min}(z)$ y pasa todo $H_{\text{ap}}(z)$ en que se descompone.

$$z = \frac{2}{3}e^{-j3\pi/4} \quad \text{y} \quad z = \frac{2}{3}e^{j3\pi/4},$$

que son los reflejos de los ceros de $H(z)$ fuera del círculo unidad a la posición conjugada inversa dentro del círculo unidad, y un polo en $z = 1/3$ dentro del círculo unidad, y por lo tanto, es de fase mínima. El segundo sistema es un pasa todo, cuyos polos se cancelan con los ceros del sistema de fase mínima. En la figura 3.18 se muestra el diagrama de polos de los sistemas $H(z)$, $H_{\min}(z)$ y $H_{\text{ap}}(z)$.

3.7.2. Compensación de la respuesta en frecuencia de sistemas de fase no mínima

En muchas aplicaciones de procesamiento de señales ocurre que una señal puede ser distorsionada por un sistema LTI con respuesta en frecuencia no deseada, y es de interés procesar la señal distorsionada con un sistema compensador, como se muestra en la figura 3.19. Si se logra compensación perfecta, $s_c[n] = s[n]$, o equivalentemente, $H_c(z)$ es el sistema inverso de $H_d(z)$. Sin embargo, si se asume que el sistema distorsionador es causal y estable y se requiere que el sistema compensador sea causal y estable, la compensación perfecta solo es posible si $H_d(z)$ es de fase mínima, de forma tal que tenga un sistema inverso causal y estable.

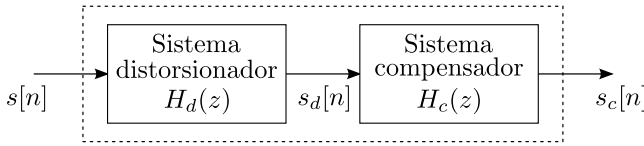


Figura 3.19: Compensación de la distorsión mediante filtrado.

En base a lo discutido previamente, asumiendo que $H_d(z)$ es conocido o aproximado por una función racional, se puede construir un sistema de fase mínima $H_{d\min}(z)$ reflejando todos los ceros de $H_d(z)$ que están fuera del círculo unidad a las posiciones conjugadas inversas dentro del círculo unidad. De esta forma, $H_d(z)$ y $H_{d\min}(z)$ tienen la misma magnitud de respuesta en frecuencia y están vinculados mediante un sistema pasa todo $H_{\text{ap}}(z)$,

$$H_d(z) = H_{d\min}(z)H_{\text{ap}}(z)$$

Eligiendo el sistema compensador como

$$H_c(z) = \frac{1}{H_{d\min}(z)},$$

la función del sistema completo entre la entrada $s[n]$ y la salida $s_c[n]$ es

$$G(z) = H_d(z)H_c(z) = H_{\text{ap}}(z).$$

es decir, $G(z)$ corresponde a un sistema pasa todo. En consecuencia, la distorsión en magnitud es perfectamente compensada mientras que la respuesta en fase es modificada como $\angle H_{\text{ap}}(z)$.

El siguiente ejemplo ilustra la compensación de la magnitud de la respuesta en frecuencia para un sistema FIR.

Ejemplo: compensación para un sistema FIR Se considera el sistema distorsionador con función del sistema

$$H_d(z) = (1 - 0.9e^{j0.6\pi}z^{-1})(1 - 0.9e^{-j0.6\pi}z^{-1})(1 - 1.25e^{j0.8\pi}z^{-1})(1 - 1.25e^{-j0.8\pi}z^{-1}). \quad (3.39)$$

El diagrama de polos y ceros se muestra en la figura 3.20. Como $H_d(z)$ tiene solo ceros (todos los polos están en $z = 0$), el sistema tiene respuesta al impulso de duración finita. Por lo tanto, el sistema es estable, y además, como $H_d(z)$ es un polinomio que tiene solo potencias negativas de z , el sistema es causal. Sin embargo, como el sistema tiene dos ceros fuera del círculo unidad, el sistema no es fase mínima.

El sistema correspondiente de fase mínima se obtiene reflejando los ceros en $z = 1.25e^{\pm j0.8\pi}$ a sus posiciones inversas conjugadas dentro del círculo unidad. Expresando $H_d(z)$ como

$$\begin{aligned} H_d(z) &= (1 - 0.9e^{j0.6\pi}z^{-1})(1 - 0.9e^{-j0.6\pi}z^{-1})1.25e^{j0.8\pi}(0.8e^{-j0.8\pi} - z^{-1})1.25e^{-j0.8\pi}(0.8e^{j0.8\pi} - z^{-1}) \\ &= (1.25)^2(1 - 0.9e^{j0.6\pi}z^{-1})(1 - 0.9e^{-j0.6\pi}z^{-1})(z^{-1} - 0.8e^{-j0.8\pi})(z^{-1} - 0.8e^{j0.8\pi}), \end{aligned}$$

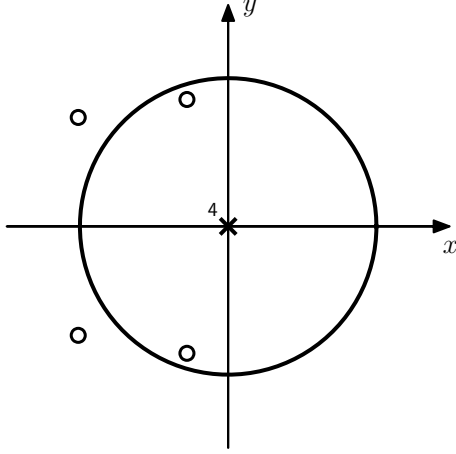


Figura 3.20: Diagrama de polos y ceros del sistema FIR del ejemplo.

se deduce que en la descomposición el sistema de fase mínima es

$$H_{d\min}(z) = (1.25)^2(1 - 0.9e^{j0.6\pi}z^{-1})(1 - 0.9e^{-j0.6\pi}z^{-1})(1 - 0.8e^{-j0.8\pi}z^{-1})(1 - 0.8e^{j0.8\pi}z^{-1}), \quad (3.40)$$

y el sistema pasa todo es

$$H_{\text{ap}}(z) = \frac{(z^{-1} - 0.8e^{-j0.8\pi})(z^{-1} - 0.8e^{j0.8\pi})}{(1 - 0.8e^{j0.8\pi}z^{-1})(1 - 0.8e^{-j0.8\pi}z^{-1})}. \quad (3.41)$$

La respuesta en magnitud logarítmica, la fase desenrollada y el retardo de grupo de $H_d(e^{j\omega})$, $H_{d\min}(e^{j\omega})$ y $H_{\text{ap}}(e^{j\omega})$ se muestran en la figura 3.21. Como corresponde, la respuesta en magnitud de $H_d(e^{j\omega})$ y $H_{d\min}(e^{j\omega})$ son idénticas. Observar que el sistema inverso de $H_d(z)$ tiene polos en $z = 1.25e^{\pm j0.8\pi}$ y por lo tanto, el sistema causal inverso sería inestable. El sistema inverso de fase mínima sería el inverso de $H_{d\min}(z)$, dado por la ecuación 3.40, y si este sistema inverso fuera empleado como sistema compensador en la figura 3.19, la respuesta global del sistema sería $H_{\text{ap}}(z)$, dado por la ecuación 3.41.

3.7.3. Propiedades de los sistemas de fase mínima

Se ha empleado el término *fase mínima* para referirse a un sistema causal y estable que tiene un sistema inverso causal y estable. El nombre está motivado por una propiedad de la función de fase de esta clase de sistemas. En esta sección se discuten las propiedades de los sistemas de fase mínima en comparación con otros sistemas con la misma magnitud de respuesta en frecuencia.

Propiedad de disminución de fase mínima

El término sistema de fase mínima como nombre descriptivo de los sistemas que tienen todos los polos y ceros dentro del círculo unidad está sugerido por el ejemplo de la sección 3.7.2. Como consecuencia de la ecuación 3.36, la fase desenrollada de un sistema de fase no mínima puede expresarse como

$$\arg[H(e^{j\omega})] = \arg[H_{\min}(e^{j\omega})] + \arg[H_{\text{ap}}(e^{j\omega})].$$

Por ejemplo, en la figura 3.21, la función de fase de $H_d(e^{j\omega})$ es la suma de la función de fase de $H_{d\min}(e^{j\omega})$ y la función de fase de $H_{\text{ap}}(e^{j\omega})$. Como se mostró en la sección 3.6 y se observa en las gráficas de la respuesta en fase en las figuras 3.17 y 3.21, la fase desenrollada de los sistemas pasa todo es negativa en $0 \leq \omega \leq \pi$. Por lo tanto, la reflexión de los ceros de $H_{\min}(z)$ dentro del círculo unidad a su posición conjugada inversa fuera del círculo unidad producida por el sistema pasa todo siempre disminuye la respuesta en fase desenrollada o incrementa el negativo de fase, que se llama función de *disminución de fase*. De esta forma, el sistema causal y estable que tiene a $|H_{\min}(e^{j\omega})|$ como magnitud de la respuesta en frecuencia tiene la mínima función de disminución de fase de todos los sistemas con la misma respuesta en magnitud, como se observa en la función de fase de $H_{d\min}(e^{j\omega})$ en la figura 3.21. Entonces, una

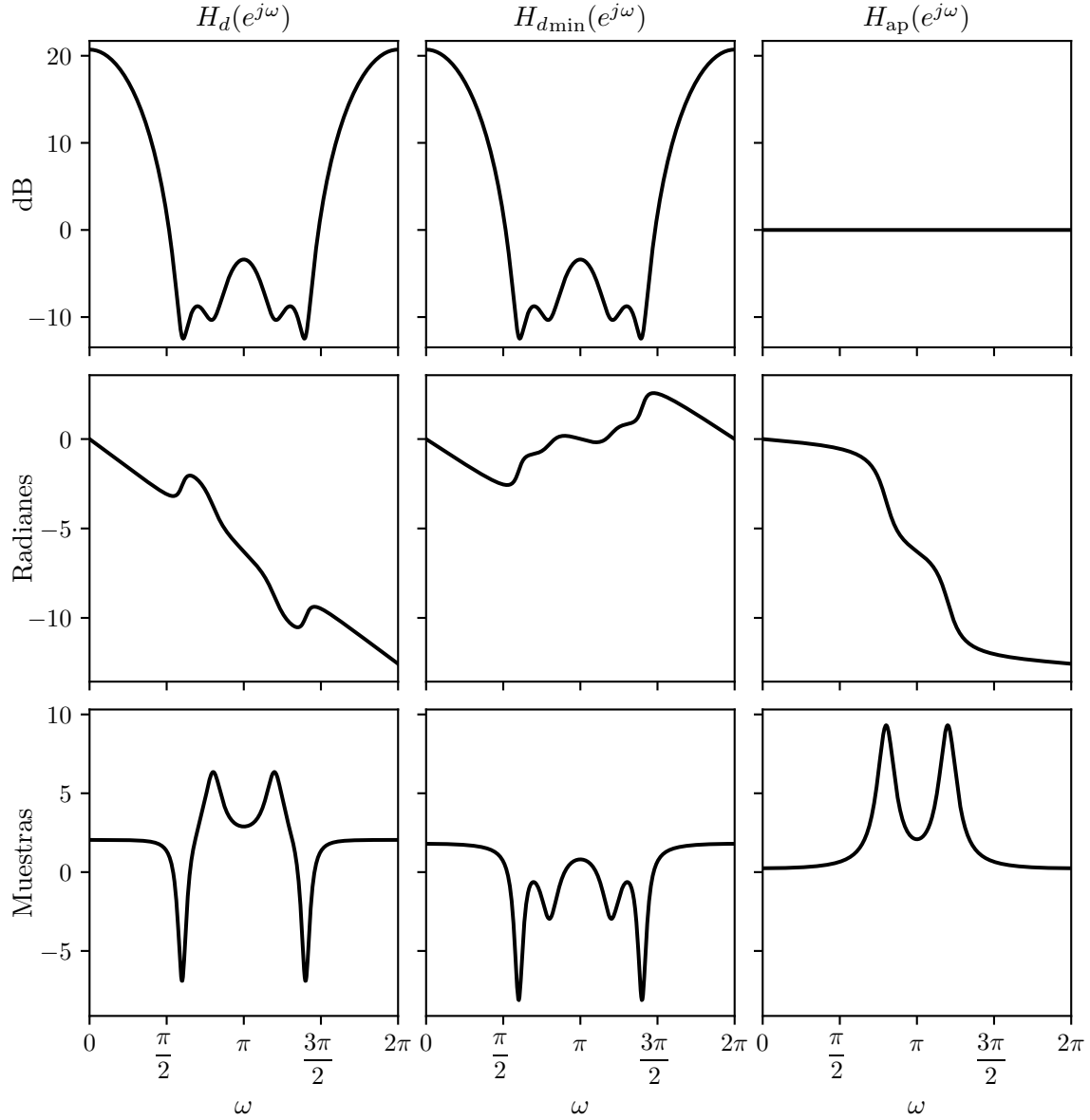


Figura 3.21: Respuesta en magnitud logarítmica, la fase desenrollada y retardo de grupo de los sistemas $H_d(z)$, $H_{dmin}(z)$ y $H_{ap}(z)$.

terminología mas precisa sería sistema de *disminución de fase mínima*, pero sistema de *fase mínima* es el término establecido históricamente.

Para hacer mas precisa la interpretación de sistema de disminución de fase mínima, es necesario imponer la restricción adicional de que $H(e^{j\omega})$ sea positivo en $\omega = 0$, es decir,

$$H(e^{j0}) = \sum_{n=-\infty}^{\infty} h[n] > 0.$$

Notar que $H(e^{j0})$ es real si $h[n]$ es real. Esta condición es necesaria debido a que un sistema con respuesta al impulso $-h[n]$ tiene los mismos polos y ceros que un sistema con respuesta al impulso $h[n]$. Sin embargo, la multiplicación por -1 altera la fase en π radianes. Así, para eliminar esta ambigüedad y asegurar que el sistema con los polos y ceros dentro del círculo unidad sea el que tiene disminución de fase mínima se

impone la condición $H(e^{j0}) > 0$. Esta condición es de poca importancia, y la definición al comienzo de esta sección, que es la definición aceptada de sistemas de fase mínima, no la incluye.

Propiedad de retardo de grupo mínimo

El ejemplo de la sección 3.7.2 ilustra otra propiedad de los sistemas con los polos y ceros dentro del círculo unidad. Nuevamente, como consecuencia de la ecuación 3.36, para dos sistemas con la misma magnitud de respuesta en frecuencia se cumple que

$$\text{grd}[H(e^{j\omega})] = \text{grd}[H_{\min}(e^{j\omega})] + \text{grd}[H_{\text{ap}}(e^{j\omega})].$$

El retardo de grupo del sistema de fase mínima es siempre menor que el retardo de grupo del sistema con fase no mínima, como se observa en la figura 3.21. Esto es porque el sistema pasa todo que convierte al sistema de fase mínima en el sistema de fase no mínima tiene retardo de grupo positivo, como se aprecia en la figura 3.21 para el ejemplo. Como se mostró en la sección 3.6, el retardo de grupo positivo es una propiedad de todos los sistemas pasa todo (ver también la figura 3.17). Se concluye que entre todos los sistemas con una magnitud de respuesta en frecuencia $|H_{\min}(e^{j\omega})|$ dada, el que tiene todos los polos y ceros dentro del círculo unidad es el de retardo de grupo mínimo. Un nombre apropiado para estos sistemas sería también sistemas de *retardo de grupo mínimo*, pero este término no es usado en general.

Propiedad de retardo de energía mínimo

En el ejemplo de la sección 3.7.2 hay un total de cuatro sistemas FIR causales con respuesta al impulso real que tienen la misma magnitud de respuesta en frecuencia que el sistema dado por la ecuación 3.39. Los diagramas de polos y ceros asociados se muestran en la figura 3.22, donde $H_a(z)$ corresponde al sistema de la ecuación 3.39 y el sistema $H_a(z)$ corresponde al sistema de fase mínima de la ecuación 3.40. Las respuestas al impulso de los cuatro sistemas se muestran en la figura 3.23. Si se comparan las cuatro secuencias de la figura, se observa que la secuencia de fase mínima parecería tener muestras mas grandes en el lado izquierdo que el resto de las secuencias. Esto es cierto para este ejemplo, y también en general. En general, siempre se cumple que

$$|h[0]| \leq |h_{\min}[0]|$$

para cualquier secuencia $h[n]$ causal y estable tal que

$$|H(e^{j\omega})| = |H_{\min}(e^{j\omega})|.$$

Para la prueba (ver el problema 5.71 de [1]), se parte considerando que cualquier sistema de fase no mínima puede expresarse como la cascada del sistema de fase mínima y un sistema pasa todo. Por lo tanto, el sistema $H(z)$ puede escribirse como

$$H(z) = H_{\min}(z)H_{\text{ap}}(z),$$

donde $H_{\min}(z)$ y $H_{\text{ap}}(z)$ se obtienen mediante el reflejo de los ceros fuera del círculo unidad, como se discutió en la sección 3.7.1. Tomado el límite para z tendiendo a infinito en ambos lados de la igualdad, se obtiene que

$$\lim_{z \rightarrow \infty} H(z) = \lim_{z \rightarrow \infty} [H_{\min}(z)H_{\text{ap}}(z)].$$

Como los sistemas involucrados son causales, puede aplicarse el teorema del valor inicial de la transformada z (ver la sección 2.4), resultando por un lado que

$$\lim_{z \rightarrow \infty} H(z) = h[0]$$

y además, que

$$\lim_{z \rightarrow \infty} [H_{\min}(z)H_{\text{ap}}(z)] = \left[\lim_{z \rightarrow \infty} H_{\min}(z) \right] \left[\lim_{z \rightarrow \infty} H_{\text{ap}}(z) \right] = h_{\min}[0]h_{\text{ap}}[0],$$

y combinando los resultados, se concluye que

$$h[0] = h_{\min}[0]h_{\text{ap}}[0]$$

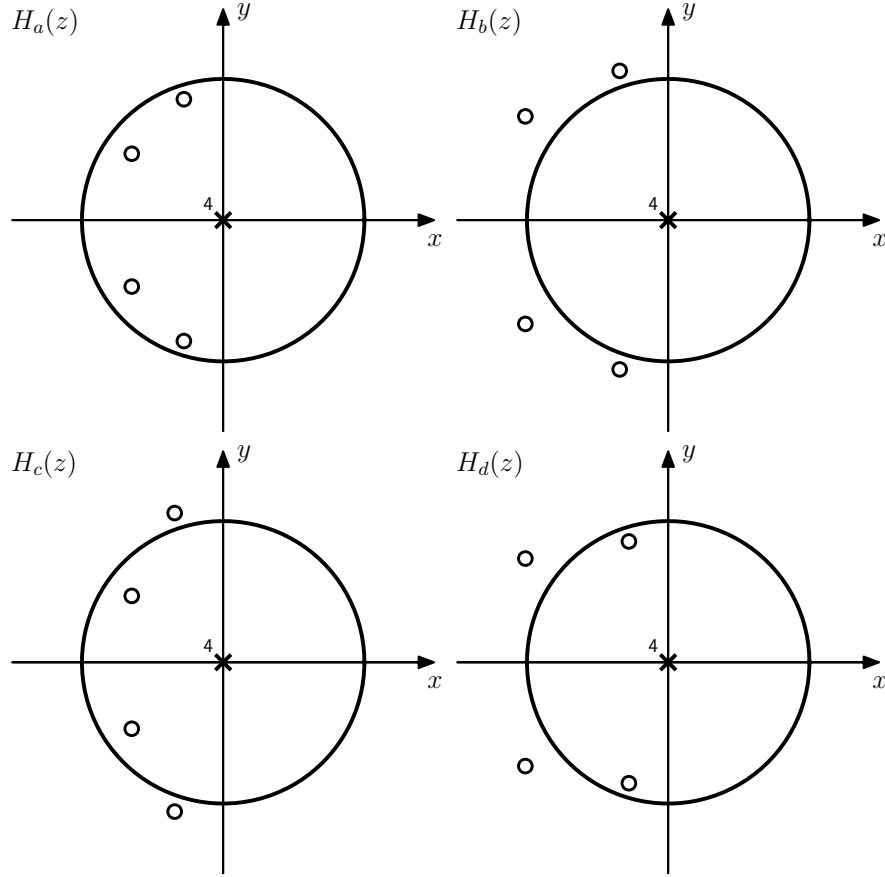


Figura 3.22: Diagrama de polos y ceros de cuatro sistemas que tienen la misma magnitud de respuesta en frecuencia. Los ceros son todas las combinaciones de los pares de ceros conjugados $0.9e^{\pm j0.6\pi}$ y $0.8e^{\pm j0.8\pi}$ y los inversos.

o

$$|h[0]| = |h_{\min}[0]| |h_{\text{ap}}[0]|. \quad (3.42)$$

Considerando la forma genérica 3.32 de un sistema pasa todo de magnitud unidad ($A = 1$), se observa que

$$\begin{aligned} |h_{\text{ap}}[0]| &= \left| \lim_{z \rightarrow \infty} H_{\text{ap}}(z) \right| \\ &= \left| \lim_{z \rightarrow \infty} \prod_{k=1}^{M_r} \frac{z^{-1} - d_k}{1 - d_k z^{-1}} \prod_{k=1}^{M_c} \frac{(z^{-1} - e_k^*)(z^{-1} - e_k)}{(1 - e_k z^{-1})(1 - e_k^* z^{-1})} \right| \\ &= \left| \prod_{k=1}^{M_r} (-d_k) \prod_{k=1}^{M_c} |e_k|^2 \right| \\ &= \prod_{k=1}^{M_r} |d_k| \prod_{k=1}^{M_c} |e_k|^2, \end{aligned}$$

y como $|d_k| < 1$ y $|e_k| < 1$, resulta en que

$$|h_{\text{ap}}[0]| < 1.$$

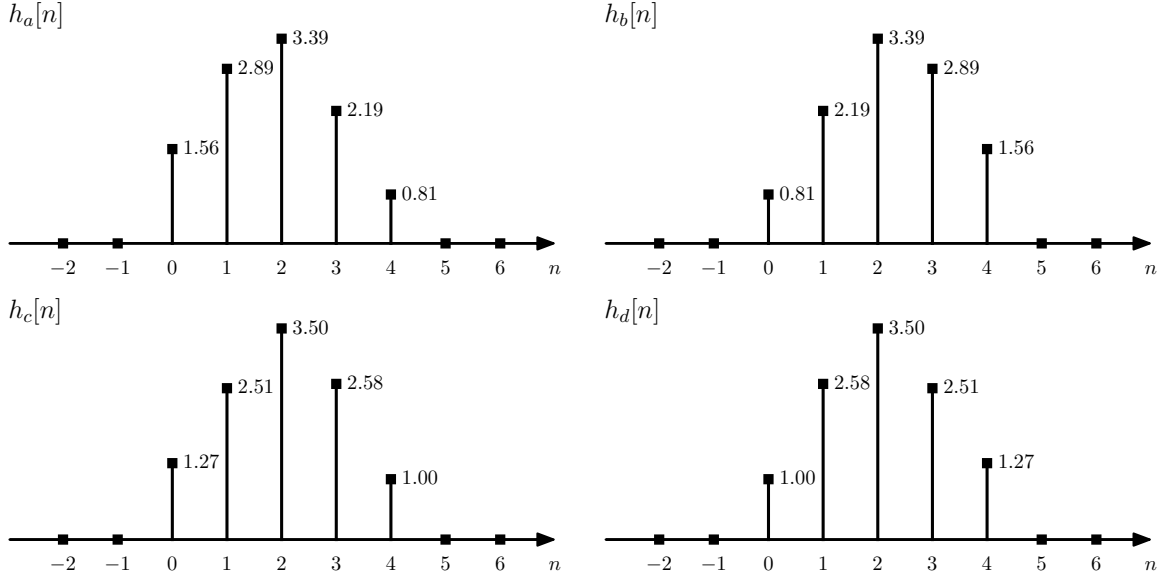


Figura 3.23: Secuencias correspondientes a los sistemas de con diagrama de polos y ceros de la figura 3.22.

Aplicando este resultado en la ecuación 3.42 se obtiene que

$$|h[0]| < |h_{\min}[0]|,$$

que es lo que se quería probar.

Todas las respuestas al impulso correspondientes a sistemas con la misma magnitud de la respuesta en frecuencia tienen siempre la misma energía total que $h_{\min}[n]$, ya que del teorema de Parseval, se cumple que

$$\sum_{n=0}^{\infty} |h[n]|^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} |H(e^{j\omega})|^2 d\omega = \frac{1}{2\pi} \int_{-\pi}^{\pi} |H_{\min}(e^{j\omega})|^2 d\omega = \sum_{n=0}^{\infty} |h_{\min}[n]|^2.$$

Si se define la *energía parcial* de la respuesta al impulso como

$$E[n] = \sum_{m=0}^n |h[m]|^2,$$

se cumple que

$$\sum_{m=0}^n |h[m]|^2 \leq \sum_{m=0}^n |h_{\min}[m]|^2. \quad (3.43)$$

Para probar este resultado (ver el problema 5.72 de [1]), sea z_k un cero de $H_{\min}(z)$, por lo que

$$H_{\min}(z) = Q(z)(1 - z_k z^{-1}), \quad |z_k| < 1, \quad (3.44)$$

donde $Q(z)$ continúa siendo de fase mínima. Sea $h[n]$ una secuencia con la misma magnitud de la respuesta en frecuencia, tal que su función del sistema $H(z)$ tiene un cero en $z = 1/z_k^*$ en vez del cero en $z = z_k$. De esta forma, $H(z)$ puede descomponerse en el sistema de fase mínima y un pasa todo como

$$H(z) = H_{\min}(z) \frac{z^{-1} - z_k^*}{1 - z_k z^{-1}},$$

y por lo tanto,

$$H(z) = Q(z)(z^{-1} - z_k^*). \quad (3.45)$$

Transformando las ecuaciones 3.44 y 3.45 al dominio del tiempo, las secuencias $h_{\min}[n]$ y $h[n]$ pueden expresarse en función de la secuencia $q[n]$ como

$$h_{\min}[n] = q[n] - z_k q[n-1] \quad (3.46)$$

y

$$h[n] = q[n-1] - z_k^* q[n], \quad (3.47)$$

donde $q[n]$ es una secuencia compleja causal y estable. Para ver que es causal, puede emplearse el teorema del valor inicial de la transformada z en la ecuación 3.44 o 3.45 obteniendo que $h_{\min}[0] = q[0]$ o $h[0] = -z_k^* q[0]$ respectivamente, y sustituyendo este resultado en la ecuación 3.46 o 3.47 respectivamente, se obtiene que $q[-1] = 0$. Luego, considerando que $h_{\min}[n]$ y $h[n]$ son causales, se concluye que $q[n]$ es causal. Continuando, se calculará el valor

$$\varepsilon = \sum_{m=0}^n |h_{\min}[m]|^2 - \sum_{m=0}^n |h[m]|^2$$

y se mostrará que $\varepsilon \geq 0$, lo que concluye la prueba. De la ecuación 3.46, se observa que

$$\begin{aligned} |h_{\min}[n]|^2 &= (q[n] - z_k q[n-1])(q^*[n] - z_k^* q^*[n-1]) \\ &= |q[n]|^2 - z_k^* q[n] q^*[n-1] - z_k q[n-1] q^*[n] + |z_k|^2 |q[n-1]|^2, \end{aligned}$$

y de forma similar, de la ecuación 3.47 se obtiene que

$$\begin{aligned} |h[n]|^2 &= (q[n-1] - z_k^* q[n])(q^*[n-1] - z_k q^*[n]) \\ &= |q[n-1]|^2 - z_k q[n-1] q^*[n] - z_k^* q[n] q^*[n-1] + |z_k|^2 |q[n]|^2. \end{aligned}$$

Considerando estos resultados,

$$\begin{aligned} \varepsilon &= \sum_{m=0}^n (|h_{\min}[m]|^2 - |h[m]|^2) \\ &= \sum_{m=0}^n (|q[m]|^2 + |z_k|^2 |q[m-1]|^2 - |q[m-1]|^2 - |z_k|^2 |q[m]|^2) \\ &= \sum_{m=0}^n [(1 - |z_k|^2) |q[m]|^2 - (1 - |z_k|^2) |q[m-1]|^2] \\ &= (1 - |z_k|^2) \left(\sum_{m=0}^n |q[m]|^2 - \sum_{m=0}^n |q[m-1]|^2 \right) \\ &\stackrel{(a)}{=} (1 - |z_k|^2) \left(\sum_{m=0}^n |q[m]|^2 - \sum_{l=-1}^{n-1} |q[l]|^2 \right) \\ &\stackrel{(b)}{=} (1 - |z_k|^2) (|q[n]|^2 - |q[-1]|^2), \end{aligned}$$

donde en (a) se realizó el cambio de variable $l = m - 1$ y en (b) se consideró que los términos de las sumatorias se cancelan excepto los correspondientes a $m = n$ y $l = -1$. Teniendo en cuenta que $q[-1] = 0$ se concluye que

$$\varepsilon = (1 - |z_k|^2) |q[n]|^2 \geq 0,$$

ya que $|z_k| < 1$ por tratarse de un cero de un sistema de fase mínima. Este resultado conduce a la ecuación 3.43 concluyendo la prueba. De acuerdo a la ecuación 3.43 la energía parcial de un sistema de fase mínima está mas concentrada en torno a $n = 0$, es decir, un sistema de fase mínima produce el menor retardo de energía entre todos los sistemas con la misma magnitud de respuesta en frecuencia. Por esta razón, los sistemas de fase mínima son también llamados *sistemas de retardo de energía mínimo* o *sistemas de retardo mínimo*. Esta propiedad del retardo es ilustrada en la figura 3.24, donde se muestra la energía parcial de las cuatro secuencias de la figura 3.23. Se observa que el retardo de energía mínimo ocurre para el sistema que tiene todos los ceros dentro del círculo unidad, es decir, para el sistema de fase mínima, y el mayor retardo de energía máximo ocurre para el sistema que tiene todos los ceros fuera del círculo unidad. Los sistemas de retardo de energía máximo son llamados *sistemas de fase máxima*.

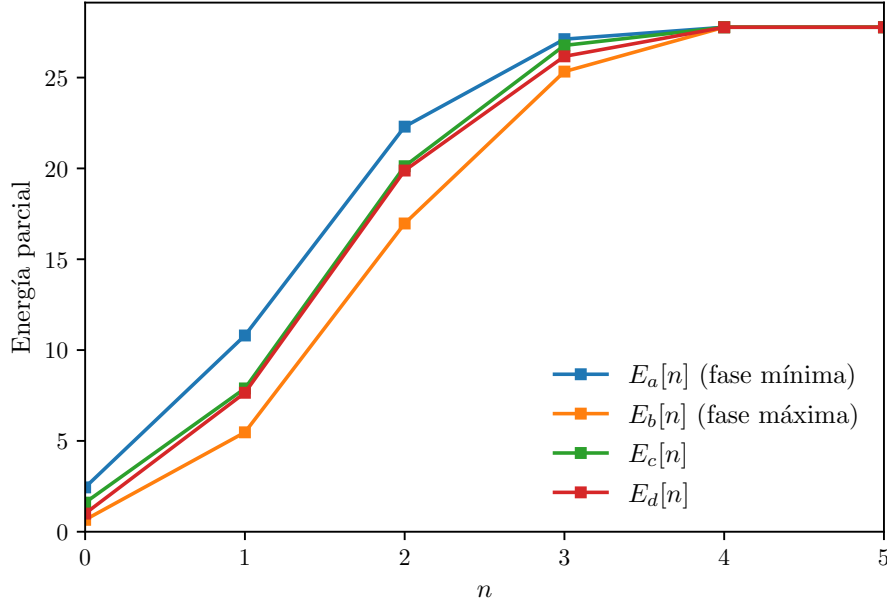


Figura 3.24: Energía parcial de las cuatro secuencias de la figura 3.23. El sistema de fase mínima produce el menor retardo de energía y el sistema de fase máxima produce el mayor retardo de energía.

3.8. Sistemas lineales con fase lineal generalizada

En el diseño de sistemas de selección de frecuencias, es deseable lograr una magnitud de respuesta en frecuencia aproximadamente constante y fase nula en la banda pasante. Para sistemas causales, fase nula es inalcanzable, y por lo tanto, debe permitirse cierta distorsión de fase. Como se vio en la sección 3.2 el efecto de fase lineal con pendiente entera es un simple retardo temporal. Por otro lado, una fase no lineal puede alterar considerablemente la forma de la señal de entrada incluso cuando la magnitud de la respuesta en frecuencia es constante. En consecuencia, en muchas situaciones es importante el diseño de sistemas con fase exactamente o aproximadamente lineal. En esta sección se formalizan y generalizan las nociones de fase lineal y retardo ideal mediante la consideración de la clase de sistemas con retardo de grupo constante.

3.8.1. Sistemas con fase lineal

Sea el sistema LTI cuya respuesta en frecuencia en un período es

$$H_{\text{id}}(e^{j\omega}) = e^{-j\omega\alpha}, \quad |\omega| < \pi,$$

donde α es un número real no necesariamente entero. Este sistema es un retardo ideal, donde α es el retardo introducido por el sistema. Este sistema tiene repuesta en magnitud constante, fase lineal y retardo de grupo constante,

$$|H_{\text{id}}(e^{j\omega})| = 1, \quad \angle H_{\text{id}}(e^{j\omega}) = -\omega\alpha, \quad \text{grd}[H_{\text{id}}(e^{j\omega})] = \alpha.$$

La transformada inversa de Fourier de $H_{\text{id}}(e^{j\omega})$ es la respuesta al impulso

$$h_{\text{id}}[n] = \frac{\text{sen } \pi(n - \alpha)}{\pi(n - \alpha)}, \quad -\infty < n < \infty. \quad (3.48)$$

La salida de este sistema cuando la entrada es $x[n]$ es

$$y[n] = x[n] * \frac{\text{sen } \pi(n - \alpha)}{\pi(n - \alpha)} = \sum_{k=-\infty}^{\infty} x[k] * \frac{\text{sen } \pi(n - k - \alpha)}{\pi(n - k - \alpha)}.$$

Si $\alpha = n_d$, donde n_d es un numero entero,

$$h_{id}[n] = \delta[n - n_d]$$

y la salida es

$$y[n] = x[n] * \delta[n - n_d] = x[n - n_d].$$

Es decir, si $\alpha = n_d$ entero el sistema con fase lineal y ganancia unidad simplemente desplaza n_d muestras a la señal. Si α no es entero también es un retardo que se interpreta como se indica en la sección 4.5 de [1] (ver en particular el ejemplo 4.7). Si α es positivo se trata de un retardo temporal y si α es negativo es un avance temporal.

Esta discusión también brinda una interpretación útil de la fase lineal cuando está asociada a una respuesta en magnitud no constante. Por ejemplo, considérese la respuesta en frecuencia con fase lineal mas general dada por

$$H(e^{j\omega}) = |H(e^{j\omega})|e^{-j\omega\alpha}, \quad |\omega| < \pi,$$

Esta ecuación sugiere la interpretación de la figura 3.25. La señal $x[n]$ es filtrada por la respuesta en

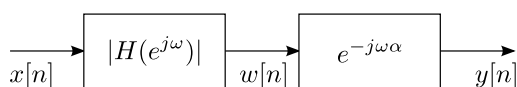


Figura 3.25: representación de un sistema LTI de fase lineal como la casacada de un sistema de fase nula y un retardo temporal.

frecuencia de fase nula $|H(e^{j\omega})|$ y la salida filtrada es desplazada temporalmente una cantidad α entera o no entera. Supóngase por ejemplo que $H(e^{j\omega})$ es el pasabajos ideal de fase lineal

$$H_{lp}(e^{j\omega}) = \begin{cases} e^{-j\omega\alpha}, & |\omega| < \omega_c, \\ 0, & \omega_c < |\omega| < \pi. \end{cases}$$

La correspondiente respuesta al impulso es

$$h_{lp}[n] = \frac{\sin \omega_c(n - \alpha)}{\pi(n - \alpha)}. \quad (3.49)$$

Observar que cuando $\omega_c = \pi$ se obtiene la ecuación 3.48.

Ejemplo: pasabajos ideal con fase lineal La respuesta al impulso del pasabajos ideal ilustra algunas propiedades de los sistemas de fase lineal. Cuando $\alpha = n_d$ es un entero, la respuesta en frecuencia es simétrica en torno a $n = n_d$, es decir, se cumple que $h_{lp}[n_d - n] = h_{lp}[n_d + n]$, o equivalentemente, $h_{lp}[n] = h_{lp}[2n_d - n]$. Efectivamente,

$$h_{lp}[2n_d - n] = \frac{\sin \omega_c(2n_d - n - n_d)}{\pi(2n_d - n - n_d)} = \frac{\sin \omega_c(n_d - n)}{\pi(n_d - n)} = \frac{\sin \omega_c(n - n_d)}{\pi(n - n_d)} = h_{lp}[n].$$

En este caso se puede definir un *sistema de fase cero*

$$\hat{H}_{lp}(e^{j\omega}) = H(e^{j\omega})e^{j\omega n_d} = |H(e^{j\omega})|,$$

donde la respuesta es desplazada a la izquierda n_d muestras (si $n_d > 0$) resultando en una secuencia par

$$\hat{h}_{lp}[n] = \frac{\sin \omega_c n}{\pi n} = \hat{h}_{lp}[-n].$$

Otro caso a distinguir es cuando la pendiente α de la fase lineal es un número entero mas un medio, o equivalentemente, 2α es entero. Puede mostrarse que también se cumple que

$$h_{lp}[2\alpha - n] = h_{lp}[n].$$

Observar que por ser 2α entero, el argumento $2\alpha - n$ es entero por lo que $h_{lp}[2\alpha - n]$ está bien definido. En este caso, el punto de simetría es α , que no es entero sino que es el punto medio entre dos muestras sucesivas. Como el punto de simetría no es un punto de la secuencia no es posible desplazar la secuencia

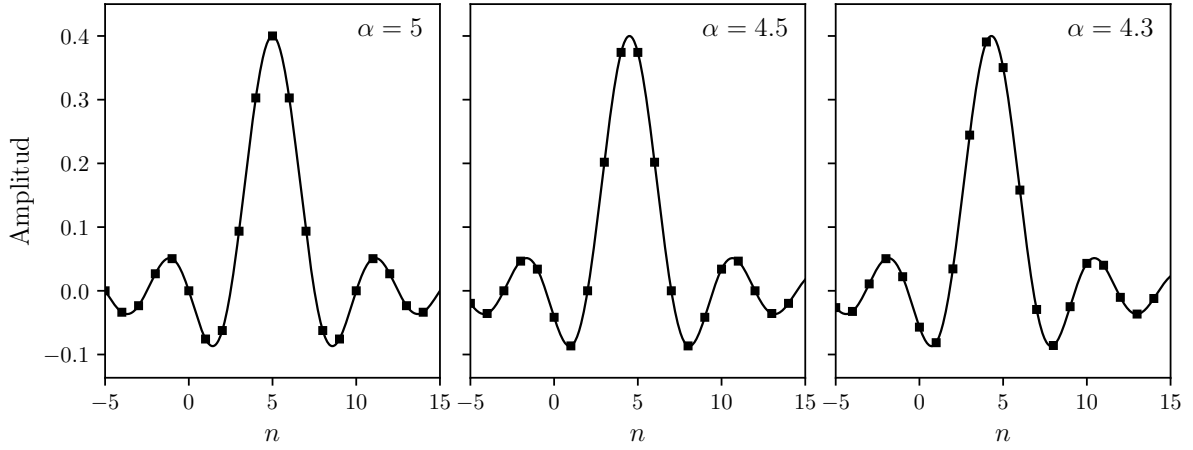


Figura 3.26: Respuesta al impulso de pasabajos ideal con $\omega_c = 0.4$. Cuando $\alpha = n_d$ es entero o 2α es entero, la respuesta al impulso es simétrica respecto a α , pero solo en el primer caso puede definirse un sistema de fase cero mediante un desplazamiento temporal. Cuando α es un número real arbitrario no hay simetría.

para obtener una secuencia par. Un último tercer caso es cuando α es un número real arbitrario, en donde la respuesta al impulso no tiene simetría. En ejemplo de cada caso se muestra en la figura 3.26.

En general, un sistema de fase lineal tiene respuesta en frecuencia

$$H(e^{j\omega}) = |H(e^{j\omega})|e^{-j\omega\alpha}. \quad (3.50)$$

Como se ilustró en el ejemplo, si 2α es un número entero, la respuesta al impulso tiene simetría par en torno a α , es decir,

$$h[2\alpha - n] = h[n].$$

Si 2α no es entero, la respuesta al impulso no tiene simetría, pero igualmente el sistema tiene fase lineal o equivalentemente, retardo de grupo constante.

3.8.2. Fase lineal generalizada

En el caso de los sistemas con respuesta en frecuencia de la forma de la ecuación 3.50, la fase está completamente establecida por el factor de fase lineal $e^{-j\omega\alpha}$ y es $-\omega\alpha$. En consecuencia, los sistemas con respuesta en frecuencia de esa forma se llaman sistemas de fase lineal. Pero considérese por ejemplo el sistema de media móvil, cuya respuesta en frecuencia está dada por la ecuación 1.33. Esta respuesta en frecuencia consiste en una función real de ω multiplicada por un factor de fase lineal pero no es estrictamente de fase lineal debido a los cambios de signo de la función real de ω , que contribuyen con saltos de π en la función de fase, como se muestra en la figura 1.11. Notar que la fase en dicha figura no tiene saltos de 2π por lo que es tanto el argumento principal como la fase desenrollada, es decir, el argumento principal de la fase y la fase desenrollada coinciden.

Muchas de las ventajas de los sistemas de fase lineal se aplican a los que tienen respuesta en frecuencia de la forma de la ecuación 1.33 del sistema de media móvil, y en consecuencia es útil generalizar la definición del concepto de fase lineal. Específicamente, un sistema es referido como *sistema de fase lineal generalizada* si su respuesta en frecuencia se puede expresarse como

$$H(e^{j\omega}) = A(e^{j\omega})e^{-j\omega\alpha+j\beta}, \quad (3.51)$$

donde α y β son constantes y $A(e^{j\omega})$ es una función real de ω que puede tomar tanto valores positivos como negativos. La fase de estos sistemas consiste en términos constantes sumados a la fase lineal $-\omega\alpha$. Por lo tanto, la fase es una recta con discontinuidades producidas por la adición de los términos

constantes, como en la figura 1.11, y si se ignoran dichas discontinuidades, el sistema se caracteriza por tener un retardo de grupo constante. Es decir, la clase de sistemas tales que

$$\tau(\omega) = \text{grd}[H(e^{j\omega})] = -\frac{d}{d\omega}\{\arg[H(e^{j\omega})]\} = \alpha$$

tienen fase lineal de la forma mas general

$$\arg[H(e^{j\omega})] = \beta - \omega\alpha, \quad 0 < \omega < \pi,$$

donde β y α son constantes.

En la sección 3.8.1 se mostró que la respuesta al impulso de un sistemas con fase lineal tiene simetría en torno a α si 2α es entero. Para ver la implicancia de esto para los sistemas de fase lineal generalizada se derivará una condición que debe ser satisfecha por $h[n]$, α y β . Se parte notando que para dichos sistemas, la respuesta en frecuencia puede expresarse como

$$\begin{aligned} H(e^{j\omega}) &= A(e^{j\omega})e^{j(\beta-\omega\alpha)} \\ &= A(e^{j\omega})\cos(\beta-\omega\alpha) + jA(e^{j\omega})\sin(\beta-\omega\alpha), \end{aligned}$$

o equivalentemente, como

$$\begin{aligned} H(e^{j\omega}) &= \sum_{n=-\infty}^{\infty} h[n]e^{-j\omega n} \\ &= \sum_{n=-\infty}^{\infty} h[n]\cos\omega n - j \sum_{n=-\infty}^{\infty} h[n]\sin\omega n, \end{aligned}$$

donde se asumió que $h[n]$ es real. La tangente de la fase de $H(e^{j\omega})$ puede expresarse como

$$\tan(\beta - \omega\alpha) = \frac{\sin(\beta - \omega\alpha)}{\cos(\beta - \omega\alpha)} = \frac{-\sum_{n=-\infty}^{\infty} h[n]\sin\omega n}{\sum_{n=-\infty}^{\infty} h[n]\cos\omega n},$$

por lo que la condición buscada es

$$\sin(\beta - \omega\alpha) \sum_{n=-\infty}^{\infty} h[n]\cos\omega n = -\cos(\beta - \omega\alpha) \sum_{n=-\infty}^{\infty} h[n]\sin\omega n,$$

o equivalentemente,

$$\sum_{n=-\infty}^{\infty} h[n] \{\sin(\beta - \omega\alpha)\cos\omega n + \cos(\beta - \omega\alpha)\sin\omega n\} = 0,$$

y empleando la identidad trigonométrica $\sin(x \pm y) = \sin x \cos y \pm \cos x \sin y$ resulta en que

$$\sum_{n=-\infty}^{\infty} h[n] \sin[\omega(n - \alpha) + \beta] = 0 \quad \text{para todo } \omega. \quad (3.52)$$

Esta ecuación establece una condición necesaria sobre $h[n]$, α y β para que el sistema tenga retardo de grupo constante. Sin embargo, no es una condición suficiente, y debido a su naturaleza implícita no indica como encontrar un sistema de fase lineal generalizada.

Una clase de sistemas de fase lineal generalizada son aquellos que cumplen que

$$\begin{aligned} \beta &= 0 \quad \text{o} \quad \beta = \pi \\ 2\alpha &= M \text{ entero} \\ h[2\alpha - n] &= h[n] \end{aligned} \quad (3.53)$$

Con $\beta = 0$ o $\beta = \pi$, la condición de la ecuación 3.52 es

$$\sum_{n=-\infty}^{\infty} h[n] \sin[\omega(n - \alpha)] = 0, \quad (3.54)$$

ya que en el caso en que $\beta = \pi$ se cumple que $\sin[\omega(n - \alpha) + \pi] = -\sin[\omega(n - \alpha)]$. Si 2α es un entero, los términos de la sumatoria 3.54 pueden agruparse de a pares que se anulan para todo ω . Por ejemplo, si α es entero, pueden agruparse las muestras equidistantes de la muestra de simetría α como

$$\begin{aligned} \sum_{n=-\infty}^{\infty} h[n] \sin[\omega(n - \alpha)] &= \sum_{n=-\infty}^{\alpha-1} \{h[n] \sin[\omega(n - \alpha)] + h[2\alpha - n] \sin[\omega(2\alpha - n - \alpha)]\} + h[\alpha] \sin 0 \\ &= \sum_{n=-\infty}^{\alpha-1} \{h[n] \sin[\omega(n - \alpha)] - h[2\alpha - n] \sin[\omega(n - \alpha)]\} \\ &= \sum_{n=-\infty}^{\alpha-1} (h[n] - h[2\alpha - n]) \sin[\omega(n - \alpha)] \\ &= 0. \end{aligned}$$

En el caso en que α es un número entero mas un medio, puede hacerse el razonamiento análogo, con el límite superior de la sumatoria siendo $\alpha - 1/2$. Las condiciones 3.53 además implican que la respuesta en frecuencia 3.51 es

$$H(e^{j\omega}) = A(e^{j\omega})e^{-j\omega\alpha}e^{j\beta} \stackrel{(a)}{=} \pm A(e^{j\omega})(\cos \omega\alpha - j \sin \omega\alpha) = \pm A(e^{j\omega}) \cos \omega\alpha \mp j A(e^{j\omega}) \sin \omega\alpha,$$

donde en (a) se consideró que $e^{j\beta} = \pm 1$ con $\beta = 0$ o $\beta = \pi$ respectivamente. Como si $h[n]$ es real, $H(e^{j\omega})$ tiene como parte real una función par y como parte imaginaria una función impar (ver la sección 1.8), las condiciones 3.53 implican que $A(e^{j\omega})$ es una función par.

Otra clase de sistemas de fase lineal generalizada son aquellos que cumplen que

$$\begin{aligned} \beta &= \pi/2 \quad \text{o} \quad \beta = 3\pi/2 \\ 2\alpha &= M \text{ entero} \\ h[2\alpha - n] &= -h[n] \end{aligned} \quad (3.55)$$

Esta vez la condición de la ecuación 3.52 es

$$\sum_{n=-\infty}^{\infty} h[n] \cos[\omega(n - \alpha)] = 0, \quad (3.56)$$

ya que $\sin[\omega(n - \alpha) + \pi/2] = \cos[\omega(n - \alpha)]$ y $\sin[\omega(n - \alpha) + 3\pi/2] = -\cos[\omega(n - \alpha)]$, y se cumple para todo ω . Efectivamente, si por ejemplo α es entero,

$$\begin{aligned} \sum_{n=-\infty}^{\infty} h[n] \cos[\omega(n - \alpha)] &= \sum_{n=-\infty}^{\alpha-1} \{h[n] \cos[\omega(n - \alpha)] + h[2\alpha - n] \cos[\omega(2\alpha - n - \alpha)]\} + h[\alpha] \cos 0 \\ &= \sum_{n=-\infty}^{\alpha-1} (h[n] + h[2\alpha - n]) \cos[\omega(n - \alpha)] + h[\alpha] \\ &= 0. \end{aligned}$$

Observar que si α es entero, la condición $h[2\alpha - n] = -h[n]$ implica que $h[\alpha] = 0$. Además, la respuesta en frecuencia 3.51 es

$$H(e^{j\omega}) = A(e^{j\omega})e^{-j\omega\alpha}e^{j\beta} \stackrel{(a)}{=} \pm j A(e^{j\omega})(\cos \omega\alpha - j \sin \omega\alpha) = \pm A(e^{j\omega}) \sin \omega\alpha \pm j A(e^{j\omega}) \cos \omega\alpha,$$

donde en (a) se consideró que $e^{j\beta} = \pm j$ con $\beta = \pi/2$ y $\beta = 3\pi/2$ respectivamente. Esto implica que $A(e^{j\omega})$ es una función impar ya que $h[n]$ es real.

Se concluye que las ecuaciones 3.53 y 3.55 brindan dos conjuntos de condiciones suficientes que garantizan fase lineal generalizada o retardo de grupo constante, pero también hay otros sistemas que cumplen la ecuación 3.51 sin las condiciones de simetría, como el de la figura 3.26 cuando 2α no es entero.

3.8.3. Sistemas causales con fase lineal generalizada

Causalidad y las condiciones de las ecuaciones 3.53 y 3.55 implica que

$$h[n] = 0, \quad n < 0 \quad \text{y} \quad n > M,$$

es decir, sistemas FIR causales tienen fase lineal generalizada si tienen respuesta al impulso de largo $M + 1$ y satisfacen las condiciones de simetría en 3.53 o 3.55. Específicamente, si

$$h[n] = \begin{cases} h[M - n], & 0 \leq n \leq M, \\ 0, & \text{en otro caso,} \end{cases} \quad (3.57)$$

la respuesta en frecuencia es

$$H(e^{j\omega}) = A_e(e^{j\omega})e^{-j\omega M/2}, \quad (3.58)$$

donde $A_e(e^{j\omega})$ es una función real, par y periódica de ω . Análogamente, si

$$h[n] = \begin{cases} -h[M - n], & 0 \leq n \leq M, \\ 0, & \text{en otro caso,} \end{cases} \quad (3.59)$$

la respuesta en frecuencia es

$$H(e^{j\omega}) = jA_o(e^{j\omega})e^{-j\omega M/2} = A_o(e^{j\omega})e^{-j\omega M/2 + j\pi/2}, \quad (3.60)$$

donde $A_o(e^{j\omega})$ es una función real, impar y periódica de ω .

Las condiciones 3.57 y 3.59 son suficientes para garantizar un sistema causal con fase lineal generalizada, pero no son condiciones necesarias. Es útil definir los siguientes cuatro tipos de sistemas FIR con fase lineal generalizada que se diferencian en el tipo de simetría y en si M es un número par o impar.

Sistema FIR de fase lineal tipo I Un sistema tipo I tiene respuesta al impulso simétrica como en la ecuación 3.57 con M un número entero par. El retardo $M/2$, que es el punto de simetría, es un entero. La respuesta en frecuencia es

$$\begin{aligned} H(e^{j\omega}) &= \sum_{n=0}^M h[n]e^{-j\omega n} \\ &= \sum_{n=0}^{M/2-1} \left[h[n]e^{-j\omega n} + h[M-n]e^{-j\omega(M-n)} \right] + h[M/2]e^{-j\omega M/2} \\ &\stackrel{(a)}{=} \sum_{n=0}^{M/2-1} h[n]e^{-j\omega M/2} \left[e^{-j\omega(n-M/2)} + e^{-j\omega(M/2-n)} \right] + h[M/2]e^{-j\omega M/2} \\ &= e^{-j\omega M/2} \sum_{n=0}^{M/2-1} 2h[n] \cos[\omega(M/2 - n)] + h[M/2]e^{-j\omega M/2} \\ &\stackrel{(b)}{=} e^{-j\omega M/2} \sum_{k=M/2}^1 2h[M/2 - k] \cos \omega k + h[M/2]e^{-j\omega M/2}, \end{aligned}$$

donde (a) se empleó la condición de simetría y en (b) se realizó el cambio de variable $k = M/2 - n$. La expresión resultante se puede escribir como

$$H(e^{j\omega}) = e^{-j\omega M/2} \left(\sum_{k=0}^{M/2} a[k] \cos \omega k \right),$$

donde

$$\begin{aligned} a[0] &= h[M/2] \\ a[k] &= 2h[M/2 - k], \quad k = 1, 2, \dots, M/2. \end{aligned}$$

Se observa que $H(e^{j\omega})$ tiene la forma de la ecuación 3.58, y en particular, β en la ecuación 3.51 es 0 o π .

Sistema FIR de fase lineal tipo II Un sistema tipo II tiene respuesta al impulso simétrica como en la ecuación 3.57 con M un número entero impar. El punto de simetría $M/2$ no es un entero si no que es un entero mas $1/2$. La respuesta en frecuencia es

$$\begin{aligned}
H(e^{j\omega}) &= \sum_{n=0}^M h[n]e^{-j\omega n} \\
&= \sum_{n=0}^{(M-1)/2} \left[h[n]e^{-j\omega n} + h[M-n]e^{-j\omega(M-n)} \right] \\
&\stackrel{(a)}{=} \sum_{n=0}^{(M-1)/2} h[n]e^{-j\omega M/2} \left[e^{-j\omega(n-M/2)} + e^{-j\omega(M/2-n)} \right] \\
&= e^{-j\omega M/2} \sum_{n=0}^{(M-1)/2} 2h[n] \cos[\omega(M/2 - n)] \\
&\stackrel{(b)}{=} e^{-j\omega M/2} \sum_{k=(M+1)/2}^1 2h[(M+1)/2 - k] \cos[\omega(k - 1/2)],
\end{aligned}$$

donde (a) se empleó la condición de simetría 3.57 y en (b) se realizó el cambio de variable $k = (M + 1)/2 - n$. La expresión se puede escribir como

$$H(e^{j\omega}) = e^{-j\omega M/2} \left\{ \sum_{k=1}^{(M+1)/2} b[k] \cos \left[\omega \left(k - \frac{1}{2} \right) \right] \right\},$$

donde

$$b[k] = 2h[(M+1)/2 - k], \quad k = 1, 2, \dots, (M+1)/2.$$

Nuevamente se observa que $H(e^{j\omega})$ tiene la forma de la ecuación 3.58.

Sistema FIR de fase lineal tipo III Si el sistema tiene respuesta al impulso antisimétrica como en la ecuación 3.59 con M un número entero par, puede mostrarse de forma análoga a como se hizo para el sistema de tipo I que la respuesta en frecuencia tiene la forma

$$H(e^{j\omega}) = je^{-j\omega M/2} \left(\sum_{k=0}^{M/2} c[k] \sen \omega k \right),$$

donde

$$c[k] = 2h[M/2 - k], \quad k = 1, 2, \dots, M/2.$$

En este caso $H(e^{j\omega})$ tiene la forma de la ecuación 3.60, y en particular, β en la ecuación 3.51 es $\pi/2$ o $3\pi/2$.

Sistema FIR de fase lineal tipo IV Si el sistema tiene respuesta al impulso antisimétrica como en la ecuación 3.59 y M es un número entero impar, puede mostrarse igual que antes que la respuesta en frecuencia es

$$H(e^{j\omega}) = je^{-j\omega M/2} \left\{ \sum_{k=1}^{(M+1)/2} d[k] \sen \left[\omega \left(k - \frac{1}{2} \right) \right] \right\},$$

donde

$$d[k] = 2h[(M+1)/2 - k], \quad k = 1, 2, \dots, (M+1)/2.$$

Nuevamente se observa que $H(e^{j\omega})$ tiene la forma de la ecuación 3.60 y β en la ecuación 3.51 es $\pi/2$ o $3\pi/2$. En la figura 3.27 se muestra un ejemplo de respuesta al impulso de cada tipo de filtro FIR de fase lineal.

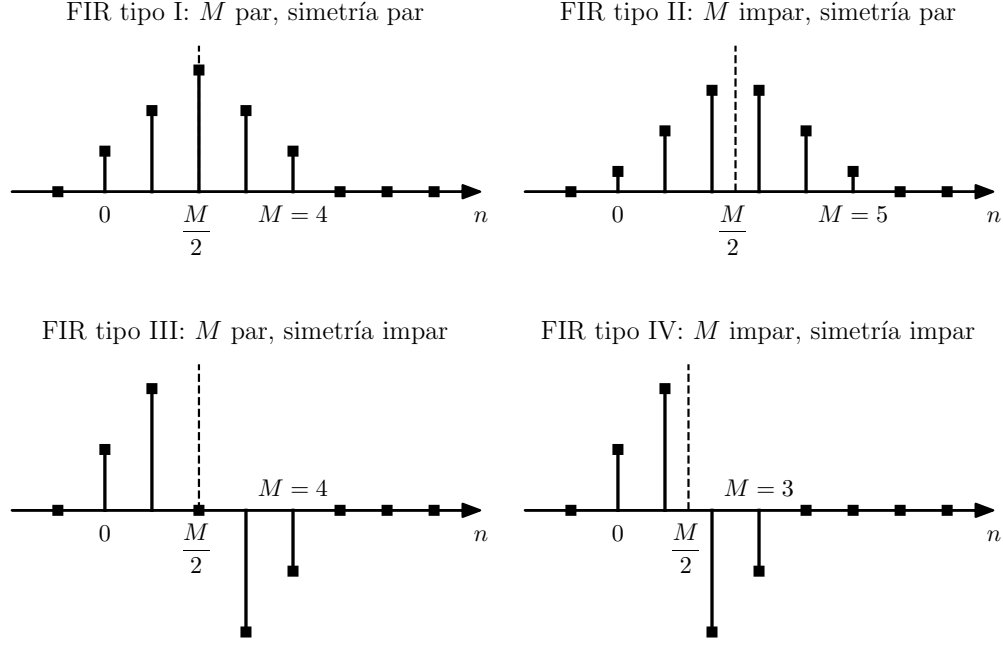


Figura 3.27: Ejemplos de la respuesta al impulso de los cuatro tipos de filtros FIR de fase lineal.

Ubicación de los ceros de los sistemas FIR de fase lineal

Resulta instructivo considerar la ubicación de los ceros de los sistemas FIR de fase lineal. La función del sistema es

$$H(z) = \sum_{n=0}^M h[n]z^{-n}.$$

En los casos en que la respuesta al impulso es simétrica (tipos I y II), se puede emplear la ecuación 3.57 para expresar $H(z)$ como

$$H(z) = \sum_{n=0}^M h[M-n]z^{-n} = \sum_{k=M}^0 h[k]z^k z^{-M} = z^{-M} H(z^{-1}). \quad (3.61)$$

Esto indica que si z_0 es un cero de $H(z)$, se cumple que

$$H(z_0) = z_0^{-M} H(z_0^{-1}) = 0,$$

es decir, si $z_0 = re^{j\theta}$ es un cero de $H(z)$, $z_0^{-1} = r^{-1}e^{-j\theta}$ también es un cero de $H(z)$. Además, si $h[n]$ es real, los ceros deben estar en pares complejos conjugados, y por lo tanto $z_0^* = re^{-j\theta}$ y $(z_0^*)^{-1} = r^{-1}e^{j\theta}$ también son ceros de $H(z)$. Por lo tanto, si $h[n]$ es real, cada cero complejo que no está en el círculo unidad será parte de un conjunto de cuatro ceros conjugados inversos de la forma

$$(1 - re^{j\theta}z^{-1})(1 - re^{-j\theta}z^{-1})(1 - r^{-1}e^{j\theta}z^{-1})(1 - r^{-1}e^{-j\theta}z^{-1}).$$

Si un cero de $H(z)$ está en el círculo unidad, es decir, $z_0 = e^{j\theta}$, entonces $z_0^{-1} = e^{-j\theta} = z_0^*$, lo que indica que los ceros sobre el círculo unidad están de a pares complejo conjugados,

$$(1 - e^{j\theta}z^{-1})(1 - e^{-j\theta}z^{-1}).$$

Si un cero de $H(z)$ es real y no está en el círculo unidad, el inverso también es un cero de $H(z)$ y $H(z)$ tendrá factores de la forma

$$(1 \pm rz^{-1})(1 \pm r^{-1}z^{-1}).$$

Finalmente, un cero de $H(z)$ en $z \pm 1$ puede aparecer solo, ya que ± 1 es su propio inverso y su propio conjugado. Por lo tanto, $H(z)$ puede tener factores de la forma

$$(1 \pm z^{-1}).$$

El caso de un cero en $z = -1$ es particularmente importante. De la ecuación 3.61 se cumple que

$$H(-1) = (-1)^M H(-1).$$

Si M es par, esta ecuación es solo una simple identidad, pero si M es impar, $H(-1) = -H(-1)$, lo que indica que $H(-1)$ debe ser cero. Por lo tanto, en el caso de respuesta al impulso simétrica con M impar, la función del sistema debe tener un cero en $z = -1$.

Si la respuesta al impulso es antisimétrica (tipos III y IV), el razonamiento análogo del que resultó la ecuación 3.61 conduce en este caso a que

$$H(z) = -z^{-M} H(z^{-1}). \quad (3.62)$$

Esta ecuación puede emplearse para mostrar que los ceros de $H(z)$ en el caso antisimétrico están restringidos de la misma forma que los ceros en el caso simétrico. Sin embargo, en el caso antisimétrico son de especial interés tanto $z = 1$ como $z = -1$. Con $z = 1$, la ecuación 3.62 es

$$H(1) = -H(1).$$

Por lo tanto, $H(z)$ debe tener un cero en $z = 1$ para M par o impar. Con $z = -1$, la ecuación 3.62 es

$$H(-1) = (-1)^{-M+1} H(-1).$$

En este caso, si $M - 1$ es impar, $H(-1) = -H(-1)$. Por lo tanto, $z = -1$ debe ser un cero de $H(z)$ cuando M es par.

En la figura 3.28 se muestra la magnitud de la respuesta en frecuencia de los cuatro tipos de filtros FIR de la figura 3.27. Se observa que cuando la repuesta al impulso es simétrica y M es impar (tipo II), la respuesta en frecuencia se anula en $\omega = \pi$, acorde al hecho de que la función del sistema tiene un cero en $z = -1$. En el caso en que la respuesta al impulso es antisimétrica se observa que la respuesta en frecuencia se anula en $\omega = 0$ y $\omega = \pi$ si M es par (tipo III), ya que la función del sistema tiene ceros en $z = \pm 1$, y si M es impar (tipo IV), la respuesta en frecuencia se anula en $\omega = \pi$ acorde al hecho de que la función del sistema tiene un cero en $z = -1$. Estas restricciones en los ceros son importantes en el diseño de filtros FIR de fase lineal, ya que imponen limitaciones sobre los tipos de respuesta en frecuencia que pueden obtenerse. Por ejemplo, si se quiere aproximar un filtro pasa alto empleando una respuesta al impulso simétrica, M no puede ser impar, ya que la respuesta en frecuencia está restringida a ser nula en $\omega = \pi$.

3.8.4. Relación entre sistemas FIR de fase lineal y sistemas de fase mínima

En la discusión previa se mostró que los sistemas FIR de fase lineal con respuesta al impulso real tienen ceros sobre el círculo unidad o en posiciones conjugadas inversas. Por lo tanto, la función del sistema de un sistema FIR de fase lineal puede factorizarse en la cascada de un sistema de fase mínima $H_{\min}(z)$, un sistema de fase máxima $H_{\max}(z)$ y un sistema $H_{\text{uc}}(z)$ que contiene ceros únicamente sobre el círculo unidad, es decir,

$$H(z) = H_{\min}(z) H_{\text{uc}}(z) H_{\max}(z) \quad (3.63)$$

donde

$$H_{\max}(z) = H_{\min}(z^{-1}) z^{-M_i}, \quad (3.64)$$

y M_i es la cantidad de ceros de $H_{\min}(z)$. En la ecuación 3.63, $H_{\min}(z)$ tiene todos sus M_i ceros dentro del círculo unidad, $H_{\text{uc}}(z)$ tiene todos sus M_o ceros sobre el círculo unidad. $H_{\max}(z)$ tiene todos sus M_i ceros fuera del círculo unidad, y de la ecuación 3.64, sus ceros son los inversos de los M_i ceros de $H_{\min}(z)$. El orden de la función del sistema $H(z)$ es entonces $2M_i + M_o$.

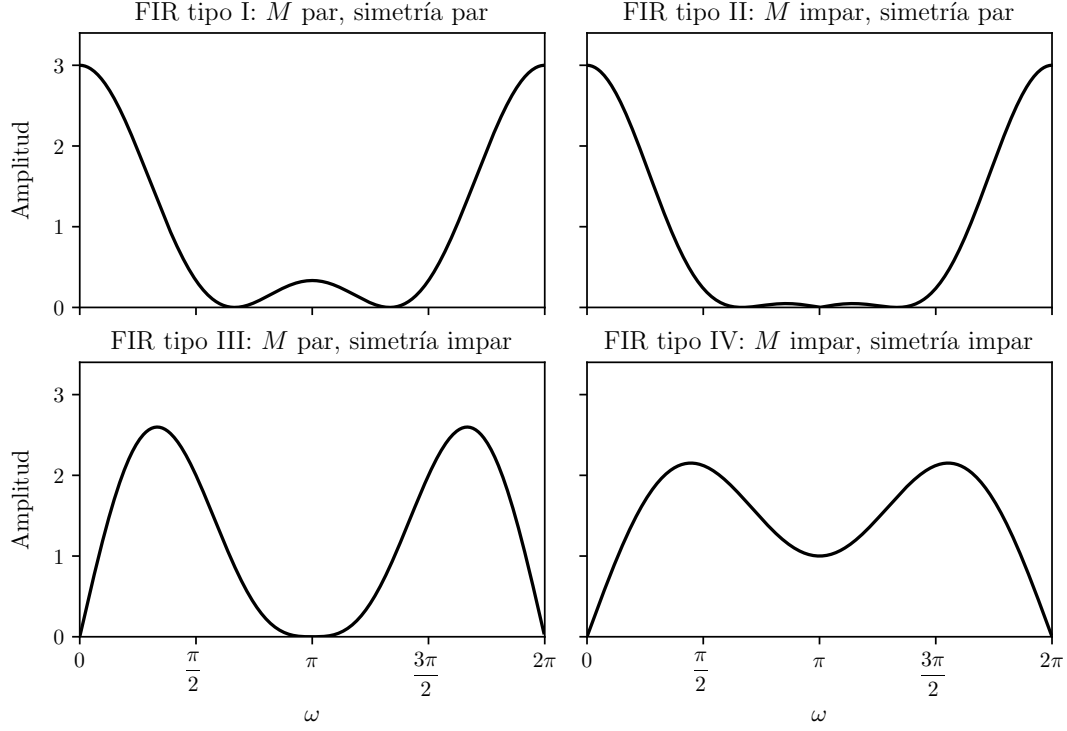


Figura 3.28: Ejemplos de la magnitud de la respuesta en frecuencia de los cuatro tipos de filtros FIR de fase lineal. En todos los casos la fase es lineal con a lo sumo saltos de π radianes y el retardo de grupo es $\alpha = M/2$ muestras constante.

Ejemplo: descomposición de un sistema de fase lineal Como ejemplo del uso de las ecuaciones 3.63 y 3.64, considérese el sistema de fase mínima dado por la ecuación 3.40,

$$H_{\min}(z) = (1.25)^2(1 - 0.9e^{j0.6\pi}z^{-1})(1 - 0.9e^{-j0.6\pi}z^{-1})(1 - 0.8e^{j0.8\pi}z^{-1})(1 - 0.8e^{-j0.8\pi}z^{-1}).$$

El sistema de fase máxima obtenido al aplicar la ecuación 3.64 es

$$\begin{aligned} H_{\max}(z) &= (1.25)^2(1 - 0.9e^{j0.6\pi}z)(1 - 0.9e^{-j0.6\pi}z)(1 - 0.8e^{j0.8\pi}z)(1 - 0.8e^{-j0.8\pi}z)z^{-4} \\ &= (1.25)^2(z^{-1} - 0.9e^{j0.6\pi})(z^{-1} - 0.9e^{-j0.6\pi})(z^{-1} - 0.8e^{j0.8\pi})(z^{-1} - 0.8e^{-j0.8\pi}) \\ &= (1.25)^2(0.9)^2(0.8)^2(1 - 1.1111e^{-j0.6\pi}z^{-1})(1 - 1.1111e^{j0.6\pi}z^{-1}) \\ &\quad \times (1 - 1.25e^{-j0.8\pi}z^{-1})(1 - 1.25e^{j0.8\pi}z^{-1}), \end{aligned}$$

resultando en

$$H_{\max}(z) = (0.9)^2(1 - 1.1111e^{j0.6\pi}z^{-1})(1 - 1.1111e^{-j0.6\pi}z^{-1})(1 - 1.25e^{j0.8\pi}z^{-1})(1 - 1.25e^{-j0.8\pi}z^{-1}),$$

Si estos dos sistemas se colocan en cascada, por la ecuación 3.64, el sistema global

$$H(z) = H_{\min}(z)H_{\max}(z),$$

tiene fase lineal. El diagrama de polos y ceros de los tres sistemas se muestra en la figura 3.29. Observar que por la ecuación 3.64 y considerando que la respuesta al impulso es real,

$$H_{\max}(e^{j\omega}) = H_{\min}(e^{-j\omega})e^{-j\omega M_i} = H_{\min}^*(e^{j\omega})e^{-j\omega M_i}$$

y por lo tanto,

$$|H_{\max}(e^{j\omega})| = |H_{\min}(e^{j\omega})| \quad \text{y} \quad \angle H_{\max}(e^{j\omega}) = -\angle H_{\min}(e^{j\omega}) - \omega M_i.$$

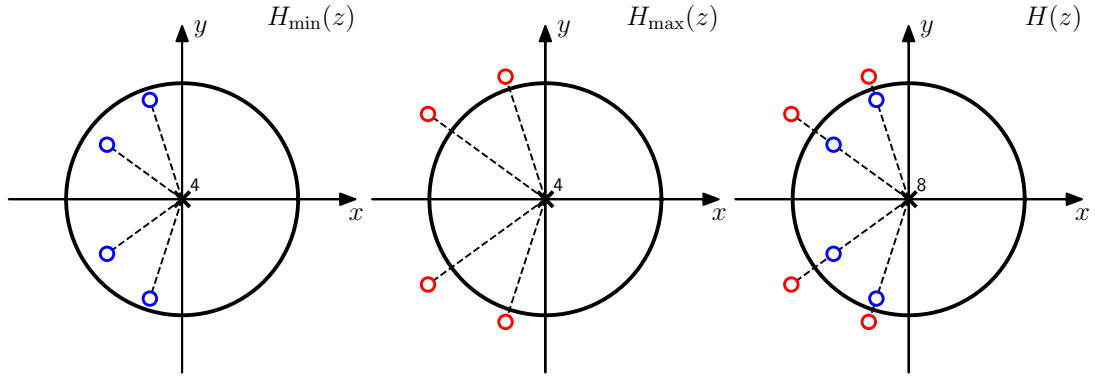


Figura 3.29: Diagrama de polos y ceros de los sistemas $H_{\min}(z)$, $H_{\max}(z)$ y $H(z)$.

La respuesta en frecuencia del sistema global se obtiene sumando la magnitud logarítmica, la fase y el retardo de grupo de los sistemas individuales. De esta forma,

$$\begin{aligned} 20 \log_{10} |H(e^{j\omega})| &= 20 \log_{10} |H_{\min}(e^{j\omega})| + 20 \log_{10} |H_{\max}(e^{j\omega})| \\ &= 40 \log_{10} |H_{\min}(e^{j\omega})|. \end{aligned}$$

Además,

$$\begin{aligned} \angle H(e^{j\omega}) &= \angle H_{\min}(e^{j\omega}) + \angle H_{\max}(e^{j\omega}) \\ &= \angle H_{\min}(e^{j\omega}) - \angle H_{\min}(e^{j\omega}) - \omega M_i \\ &= -\omega M_i, \end{aligned}$$

donde $M_i = 4$ es la cantidad de ceros de $H_{\min}(z)$. Por último, el retardo de grupo de los sistemas combinados es

$$\text{grd}[H(e^{j\omega})] = M_i = 4.$$

En la figura 3.30 se muestra la respuesta en frecuencia de los tres sistemas.

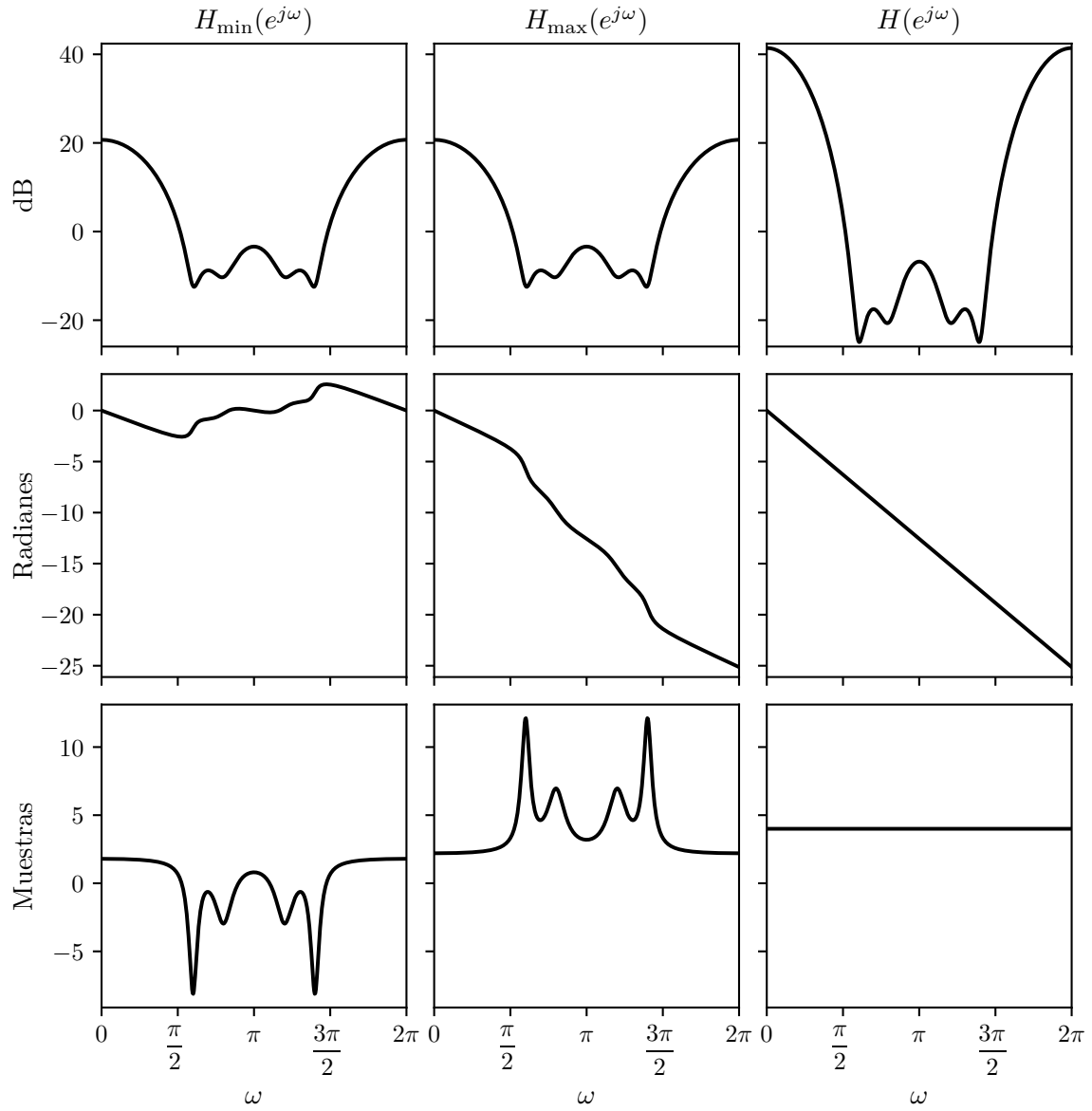


Figura 3.30: Respuesta en magnitud logarítmica, la fase desenrollada y retardo de grupo de los sistemas $H_{\min}(z)$, $H_{\max}(z)$ y $H(z)$.

Capítulo 4

Estructuras de sistemas en tiempo discreto

4.1. Representación en diagramas de bloques de ecuaciones en diferencias con coeficientes lineales

Considérese el sistema definido por la ecuación en diferencias genérica¹

$$y[n] - \sum_{k=1}^N a_k y[n-k] = \sum_{k=0}^M b_k x[n-k], \quad (4.1)$$

con la correspondiente función del sistema

$$H(z) = \frac{\sum_{k=0}^M b_k z^{-k}}{1 - \sum_{k=1}^N a_k z^{-k}} \quad (4.2)$$

Reescribiendo a la ecuación 4.1 como una ecuación en recurrencia de $y[n]$ en términos de una combinación lineal de las muestras pasadas de la salida y la muestra actual y pasadas de la entrada se obtiene la relación

$$y[n] = \sum_{k=1}^N a_k y[n-k] + \sum_{k=0}^M b_k x[n-k]. \quad (4.3)$$

El diagrama de bloques en la figura 4.1 es una representación explícita de la ecuación 4.3. Mas precisamente, representa el par de ecuaciones

$$\begin{aligned} v[n] &= \sum_{k=0}^M b_k x[n-k] \\ y[n] &= \sum_{k=1}^N a_k y[n-k] + v[n] \end{aligned}$$

Un diagrama de bloques puede ser reordenado o modificado de diversas formas sin cambiar la función del sistema global. Cada reorganización apropiada representa una algoritmo computacional diferente para implementar el mismo sistema. Por ejemplo, el diagrama de bloques de la figura 4.1 puede verse como

¹Observar que la forma general de una ecuación en diferencias de orden N empleada en los capítulos previos, dada por la ecuación 1.19, es ligeramente diferente a la empleada de aquí en mas, donde el coeficiente $a_0 = 1$ y los coeficientes correspondientes a la salida retardada son positivos luego de ser despejados al lado derecho de la igualdad.

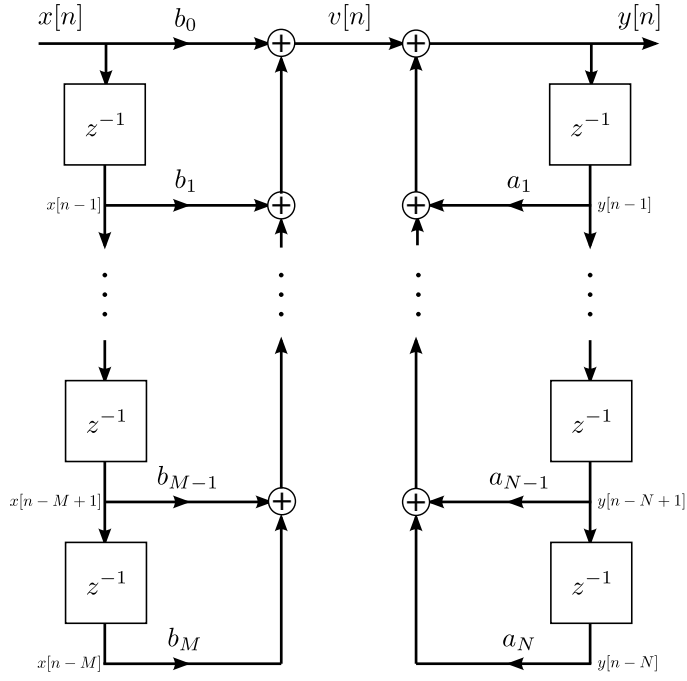


Figura 4.1: Representación en diagrama de bloques de una ecuación en diferencias de orden N .

dos sistemas en cascada, donde el primero calcula $v[n]$ a partir de $x[n]$ y el segundo calcula $y[n]$ a partir de $v[n]$. Como ambos sistemas son LTI, asumiendo condiciones iniciales nulas, puede intercambiarse el orden de los sistemas en la cascada, como se muestra en la figura 4.2, sin afectar la función del sistema global. En la figura 4.2 se asumió que $M = N$. Esto no implica pérdida de generalidad, ya que si $M \neq N$, algunos de los coeficientes a_k o b_k en la figura son nulos.

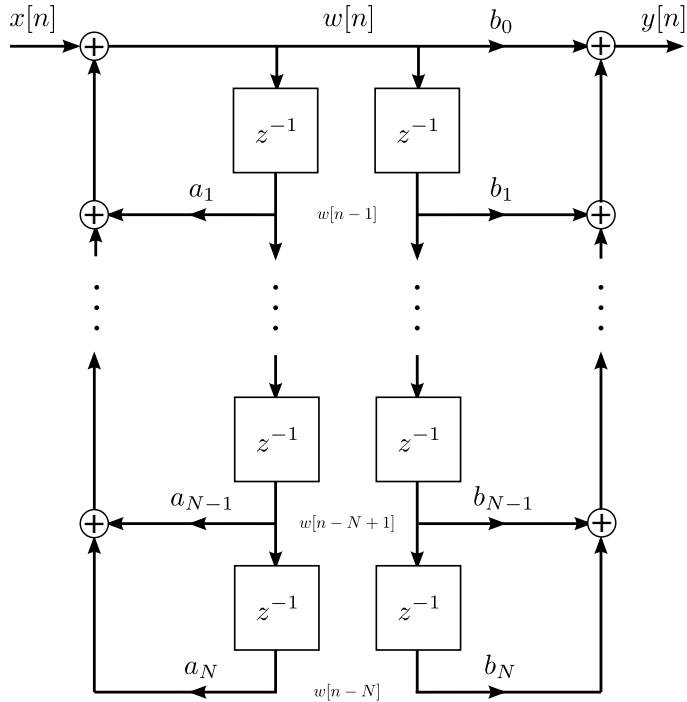


Figura 4.2: Reorganización del diagrama de bloques de la figura 4.1. Se asume por conveniencia que $M = N$ sin perder generalidad.

En términos de la función del sistema $H(z)$ en la ecuación 4.2, la figura 4.1 puede verse como la

implementación de $H(z)$ mediante la descomposición

$$H(z) = H_2(z)H_1(z) = \left(\frac{1}{1 - \sum_{k=1}^N a_k z^{-k}} \right) \left(\sum_{k=0}^M b_k z^{-k} \right)$$

o equivalentemente, mediante el par de ecuaciones

$$V(z) = H_1(z)X(z) = \left(\sum_{k=0}^M b_k z^{-k} \right) X(z)$$

y

$$Y(z) = H_2(z)V(z) = \left(\frac{1}{1 - \sum_{k=1}^N a_k z^{-k}} \right) V(z).$$

Por otro lado, la figura 4.2 implementa a $H(z)$ como

$$H(z) = H_1(z)H_2(z) = \left(\sum_{k=0}^M b_k z^{-k} \right) \left(\frac{1}{1 - \sum_{k=1}^N a_k z^{-k}} \right),$$

o equivalentemente, mediante las ecuaciones

$$W(z) = H_2(z)X(z) = \left(\frac{1}{1 - \sum_{k=1}^N a_k z^{-k}} \right) X(z)$$

y

$$Y(z) = H_1(z)W(z) = \left(\sum_{k=0}^M b_k z^{-k} \right) W(z)$$

En el dominio del tiempo, la figura 4.2 puede representarse mediante el par de ecuaciones en diferencias

$$\begin{aligned} w[n] &= \sum_{k=1}^N a_k w[n-k] + x[n] \\ y[n] &= \sum_{k=0}^M b_k w[n-k]. \end{aligned}$$

Los diagramas de bloque de las figuras 4.1 y 4.2 tiene varias diferencias importantes. En la figura 4.1, los ceros de $H(z)$, dados por $H_1(z)$ se implementan primero, seguidos de los polos, representados por $H_2(z)$. En la figura 4.2, los polos se implementan primero, seguido de los ceros. Teóricamente, el orden de implementación no afecta la función del sistema global. Sin embargo, cuando una ecuación en diferencias se implementa con aritmética de precisión finita, como ocurre en la práctica, puede haber una diferencia significativa entre los dos sistemas. Otra diferencia importante concierne a la cantidad de elementos de retardo de los dos sistemas. Ambos sistemas en las figuras 4.1 y 4.2 tienen $N + M$ elementos de retardo. Sin embargo, el diagrama de bloques de la figura 4.2 puede modificarse observando que la misma señal

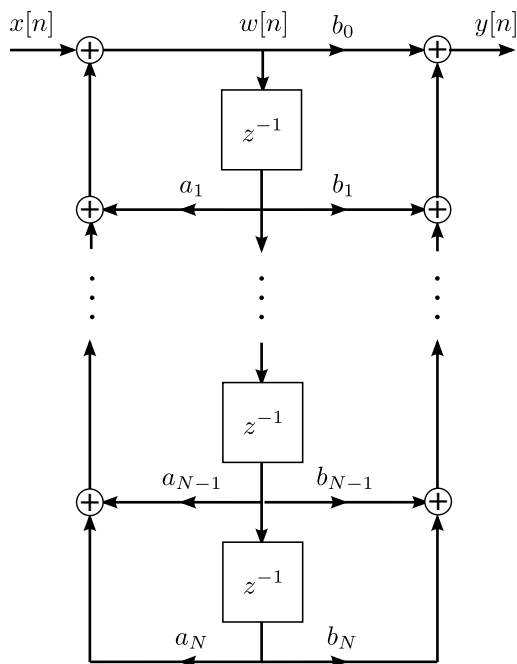


Figura 4.3: Ecuación en diferencias en forma canónica obtenida mediante la combinación de los retardos en la figura 4.2.

$w[n]$ está almacenada en las dos cadenas de elementos de retardo. En consecuencia, las dos cadenas de retardos pueden colapsarse en un única cadena de retardos, como se muestra en la figura 4.3.

La cantidad total de retardos en la figura 4.3 es menor o igual que la cantidad de retardos en las figuras 4.1 y 4.2, y de hecho, es la mínima cantidad de retardos para implementar el sistema con función del sistema dado por la ecuación 4.2. Concretamente, la cantidad mínima de elementos de retardo requerida es $\max(N, M)$. Una implementación con el número mínimo de elementos de retardos es referida como implementación en *forma canónica*. El diagrama de bloques de la figura 4.1 es referido como implementación en *forma directa I* del sistema de orden N debido a que es una realización directa de la ecuación en diferencias satisfecha por la entrada $x[n]$ y la salida $y[n]$. El diagrama de la figura 4.3 es referido como implementación en *forma directa II* o *forma canónica*.

4.2. Representación en grafos de flujo de ecuaciones en diferencias con coeficientes lineales

En esta sección se introduce a la representación en grafo de flujo de señales de una ecuación en diferencias, que es esencialmente lo mismo que la representación en diagramas de bloques excepto por algunas diferencias de notación. Lo que sigue está basado en la sección 4.1 de [5].

4.2.1. Representación en grafos de flujo de sistemas en tiempo discreto

Un grafo de flujo de señales es una red de *ramas* con dirección que conectan *nodos*. Asociado a cada nodo hay una variable o valor del nodo. El valor asociado al nodo k se denotará w_k , o, como las variables de nodos de sistemas digitales son secuencias, a menudo se indicará explícitamente con la notación $w_k[n]$. La rama (jk) denota a la rama con origen en el nodo j y destino en el nodo k , con la dirección de j a k indicada por una flecha en la rama. Esto se muestra en la figura 4.4. Cada rama tiene una señal de entrada y una señal de salida. La señal de entrada desde el nodo j a la rama (jk) es el valor del nodo w_j y la señal de salida de la rama (jk) en el nodo k se denotará v_{jk} . La dependencia de la salida de una rama con la entrada a la rama se denotará

$$v_{jk} = f_{jk}[w_j],$$

es decir, $f_{jk}[\]$ indica la transformación de la entrada a la rama en la salida de la rama.

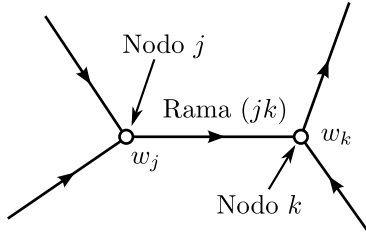


Figura 4.4: Nodos y ramas en un grafo de flujo.

Para representar la inyección de entradas externas o fuentes al grafo se emplean *nodos fuente*. Un nodo fuente no tiene ramas entrantes. El valor del nodo del nodo fuente j se denotará como x_j o $x_j[n]$ y la salida de una rama que conecta el nodo fuente j con el nodo k se denotará como s_{jk} o $s_{jk}[n]$.

Las salidas de un grafo de flujo es a través de *nodos sumideros*, que son nodos que solo tienen entradas. El valor del nodo de un nodo sumidero k será denotado como y_k o $y_k[n]$ y la salida de una rama que conecta al nodo j con el nodo sumidero k se denotará como r_{jk} o $r_{jk}[n]$.

Por definición el valor del nodo en cada nodo del grafo es la suma de la salida de todas las ramas que entran al nodo. En ocasiones es conveniente notacionalmente asumir que hay una rama en cada dirección entre cada par de nodos en la red y también que cada nodo fuente está conectado a cada nodo del grafo, aunque la salida de algunas ramas será nula. Con esta notación y asumiendo que hay N nodos en el grafo numerados de 1 a N , M nodos fuente numerados de 1 a M y P nodos sumideros numerados de 1 a P , el conjunto de ecuaciones representadas por el grafo es

$$\begin{aligned} w_k &= \sum_{j=1}^N v_{jk} + \sum_{j=1}^M s_{jk}, & k = 1, 2, \dots, N \\ y_k &= \sum_{j=1}^N r_{jk}, & k = 1, 2, \dots, P. \end{aligned} \quad (4.4)$$

Para ilustrar estos conceptos considérese el grafo de flujo de la figura 4.5. En este caso, las variables de rama son secuencias. Hay un nodo fuente conectado al nodo 1 y un nodo sumidero conectado al nodo 3. Las ecuaciones 4.4 para este grafo son

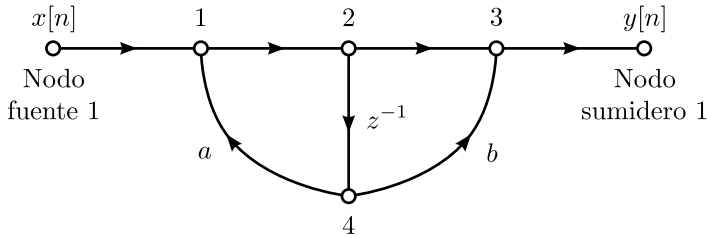


Figura 4.5: Grafo de flujo de un sistema de primer orden en forma directa II.

$$\begin{aligned} w_1[n] &= s_{11}[n] + v_{41}[n] \\ w_2[n] &= v_{12}[n] \\ w_3[n] &= v_{23}[n] + v_{43}[n] \\ w_4[n] &= v_{24}[n] \\ y[n] &= r_{31}[n], \end{aligned}$$

y las salidas de rama son

$$\begin{aligned}
s_{11}[n] &= x[n] \\
v_{12}[n] &= f_{12}(w_1[n]) = w_1[n] \\
v_{23}[n] &= f_{23}(w_2[n]) = w_2[n] \\
v_{43}[n] &= f_{43}(w_4[n]) = bw_4[n] \\
v_{41}[n] &= f_{41}(w_4[n]) = aw_4[n] \\
v_{24}[n] &= f_{24}(w_2[n]) = w_2[n-1] \\
r_{31}[n] &= w_3[n].
\end{aligned}$$

Este conjunto de ecuaciones puede resolverse para $y[n]$ en términos de $x[n]$ con la ayuda de la transformada z para obtener la ecuación en diferencias

$$y[n] = ay[n-1] + x[n] + bx[n-1].$$

En el caso de sistemas en tiempo discreto invariantes en el tiempo caracterizados por ecuaciones en diferencias, el diagrama de flujo puede representar relaciones entre transformadas z . En este caso, cada rama se representa por su función de transferencia. De esta forma,

$$V_{jk}(z) = F_{jk}(z)W_j(z).$$

En estos casos, la transmitancia de cada rama se coloca al lado de la flecha que indica la dirección de la rama. Se asume que una rama que no indica explícitamente la transmitancia tiene transmitancia unidad.

Cuando todas las ramas en un grafo de flujo pueden ser representadas mediante transmitancias, el conjunto de ecuaciones que representan al grafo es un conjunto de ecuaciones lineales, y la manipulación del grafo es equivalente a la manipulación de este conjunto de ecuaciones. Por lo tanto, una descripción alternativa de un diagrama de flujo lineal es mediante dicho conjunto de ecuaciones, que pueden representarse en forma matricial, como se considera en la siguiente sección.

A continuación se ilustra el empleo de grafos de flujo mediante un ejemplo.

Ejemplo: determinación de la función del sistema a partir del grafo de flujo En este ejemplo se ilustra el uso de la transformada z para determinar la función del sistema de un grafo de flujo. Considérese el grafo de flujo de la figura 4.6, el cual no está en forma directa. En consecuencia, la función del sistema no puede obtenerse mediante inspección del grafo. El conjunto de ecuaciones en diferencias que representan al grafo consiste en el valor de la variable de cada nodo en función de la variable de los otros nodos. En este caso, las cinco ecuaciones son

$$w_1[n] = w_4[n] - x[n] \quad (4.5)$$

$$w_2[n] = \alpha w_1[n] \quad (4.6)$$

$$w_3[n] = w_2[n] + x[n] \quad (4.7)$$

$$w_4[n] = w_3[n-1] \quad (4.8)$$

$$y[n] = w_2[n] + w_4[n]. \quad (4.9)$$

En este punto, para obtener la función del sistema es conveniente transformar el conjunto de ecuaciones al dominio z , pero a modo ilustrativo se continuará en el dominio del tiempo. Sustituyendo la ecuación 4.5 en la ecuación 4.6 y la ecuación 4.7 en la ecuación 4.8, el conjunto de ecuaciones que representa al grafo queda

$$w_2[n] = \alpha(w_4[n] - x[n]) \quad (4.10)$$

$$w_4[n] = w_2[n-1] + x[n-1] \quad (4.11)$$

$$y[n] = w_2[n] + w_4[n]. \quad (4.12)$$

Continuando en el dominio del tiempo, si se sustituye la ecuación 4.10 en la ecuación 4.11 se obtiene una ecuación en diferencias para $w_4[n]$ en términos de la entrada $x[n]$,

$$\begin{aligned}
w_4[n] &= \alpha(w_4[n-1] - x[n-1]) + x[n-1] \\
&= \alpha w_4[n-1] + (1 - \alpha)x[n-1].
\end{aligned}$$

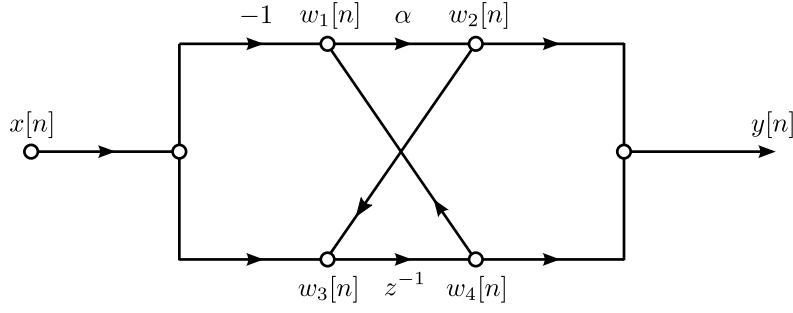


Figura 4.6: Grafo de flujo que no está en forma directa.

Esta ecuación en diferencias puede resolverse recursivamente como se hizo en algunos ejemplos de la sección 1.5 para obtener que

$$w_4[n] = (1 - \alpha) \sum_{k=0}^{n-1} \alpha^{n-1-k} x[k] = (1 - \alpha)(\alpha^{n-1} u[n-1] * x[n]).$$

El mismo resultado puede obtenerse aplicando la transformada z a la ecuación en diferencias. Realizando el mismo razonamiento para obtener $w_2[n]$ en términos de la entrada $x[n]$, pueden sustituirse las expresiones de $w_2[n]$ y $w_4[n]$ en la ecuación 4.12 resultando en la ecuación de la salida $y[n]$ en términos de la entrada $x[n]$, pero dicha expresión no es recursiva. De forma similar, también recursivamente o mediante la transformada z , puede encontrarse la respuesta al impulso, que en el caso de la ecuación en diferencias para $w_4[n]$ es

$$h_4[n] = (1 - \alpha)\alpha^{n-1}u[n-1],$$

y por lo tanto, obtener la respuesta el impulso de la salida $y[n]$. Sin embargo, no es directo el cálculo en el dominio del tiempo de la ecuación en diferencias involucrando únicamente muestras de la entrada y la salida.

La función del sistema se obtiene de forma mas directa mediante la transformada z . Aplicando la transformada z a las ecuaciones 4.10-4.12, se obtiene que

$$W_2(z) = \alpha[W_4(z) - X(z)] \quad (4.13)$$

$$W_4(z) = z^{-1}[W_2(z) + X(z)] \quad (4.14)$$

$$Y(z) = W_2(z) + W_4(z). \quad (4.15)$$

Las ecuaciones 4.13 y 4.14 pueden resolverse para $W_2(z)$ y $W_4(z)$ resultando en

$$W_2(z) = \frac{\alpha(z^{-1} - 1)}{1 - \alpha z^{-1}} X(z)$$

$$W_4(z) = \frac{z^{-1}(1 - \alpha)}{1 - \alpha z^{-1}} X(z)$$

y la sustitución de estas ecuaciones en la ecuación 4.15 conduce a

$$Y(z) = \frac{\alpha(z^{-1} - 1) + z^{-1}(1 - \alpha)}{1 - \alpha z^{-1}} X(z) = \frac{z^{-1} - \alpha}{1 - \alpha z^{-1}} X(z). \quad (4.16)$$

Se concluye que la ecuación del sistema del grafo de flujo de la figura 4.6 es

$$H(z) = \frac{z^{-1} - \alpha}{1 - \alpha z^{-1}},$$

de donde se obtiene que la respuesta al impulso es

$$h[n] = \alpha^{n-1}u[n-1] + \alpha^n u[n].$$

También de la ecuación 4.16 se obtiene que la ecuación en diferencias del sistema es

$$y[n] - \alpha y[n-1] = -\alpha x[n] + x[n-1],$$

que si se expresa como el par de ecuaciones

$$\begin{aligned} v[n] &= -\alpha x[n] + x[n-1] \\ y[n] &= \alpha y[n-1] + v[n], \end{aligned}$$

se deduce inmediatamente que el grafo de flujo en forma directa I es el mostrado en la figura 4.7. Este

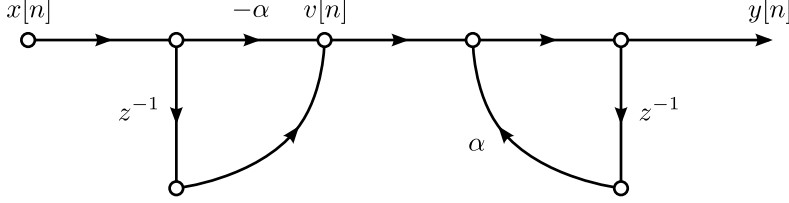


Figura 4.7: Grafo de flujo en forma directa I equivalente al grafo de la figura 4.6.

ejemplo ilustra como la transformada z convierte las expresiones en el tiempo, que involucran retroalimentación y por lo tanto son difíciles de resolver, en ecuaciones lineales que pueden resolverse mediante técnicas algebraicas. El ejemplo también muestra que representaciones en grafos de flujo equivalentes definen algoritmos computacionales que requieren una cantidad distinta de recursos computacionales. Comparando las figuras 4.6 y 4.7, se observa que la implementación original requiere únicamente una multiplicación y un elemento de retardo, mientras que la implementación en forma directa I requiere dos multiplicaciones y dos elementos de retardo. La implementación en forma directa II requiere un retardo menos, pero aún requiere dos multiplicaciones.

4.2.2. Representación matricial de redes digitales

Las ecuaciones 4.4 representadas por un diagrama de flujo expresadas en términos de transformadas z son

$$\begin{aligned} W_k(z) &= \sum_{j=1}^N V_{jk}(z) + \sum_{j=1}^M S_{jk}(z), \quad k = 1, 2, \dots, N \\ Y_k(z) &= \sum_{j=1}^N R_{jk}(z), \quad k = 1, 2, \dots, P. \end{aligned} \tag{4.17}$$

Si el grafo representa a un sistema lineal e invariante en el tiempo, cada rama puede representarse mediante una transmitancia. Es conveniente asumir que las ramas desde los nodos fuente a los nodos de red y desde los nodos de red a los nodos sumideros tienen transmitancia constante independiente de z . No hay pérdida de generalidad en esto, ya que si es necesario, puede insertarse un nodo de red conectado a un nodo fuente y este nodo puede tener ramas con transmitancia no constante a los otros nodos de red. El mismo procedimiento puede emplearse para los nodos sumidero si es necesario. Por lo tanto

$$\begin{aligned} V_{jk}(z) &= F_{jk}(z)W_j(z) \\ S_{jk}(z) &= b_{jk}X_j(z) \\ R_{jk}(z) &= c_{jk}W_j(z), \end{aligned}$$

y sustituyendo estas ecuaciones en la ecuación 4.17 se obtiene el conjunto de ecuaciones algebraicas lineales

$$\begin{aligned} W_k(z) &= \sum_{j=1}^N F_{jk}(z)W_j(z) + \sum_{j=1}^M b_{jk}X_j(z), \quad k = 1, 2, \dots, N \\ Y_k(z) &= \sum_{j=1}^N c_{jk}W_j(z), \quad k = 1, 2, \dots, P. \end{aligned} \tag{4.18}$$

Estas ecuaciones pueden escribirse en forma matricial como

$$\begin{aligned} \mathbf{W}(z) &= \mathbf{F}^t(z)\mathbf{W}(z) + \mathbf{B}^t\mathbf{X}(z) \\ \mathbf{Y}(z) &= \mathbf{C}^t\mathbf{W}(z), \end{aligned} \tag{4.19}$$

donde $\mathbf{W}(z)$ es un vector columna de valores $W_k(z)$ con $k = 1, 2, \dots, N$, $\mathbf{X}(z)$ es un vector columna de valores $X_k(z)$ con $k = 1, 2, \dots, M$ y $\mathbf{Y}(z)$ es un vector columna de valores $Y_k(z)$ con $k = 1, 2, \dots, P$. La matriz $\mathbf{F}(z)$ tiene dimensiones $N \times N$ con valores $\mathbf{F}(z) = \{F_{jk}(z)\}$, la matriz $\mathbf{B}(z)$ tiene dimensiones $M \times N$ con valores $\mathbf{B} = \{b_{jk}\}$ y la matriz \mathbf{C} tiene dimensiones $N \times P$ con valores $\mathbf{C} = \{c_{jk}\}$. Para ramas que no existen en el grafo de flujo, es decir, ramas que tienen transmitancia nula, los elementos correspondiente de las matrices es nulo.

La primera ecuación en 4.19 puede resolverse para \mathbf{W} mediante inversión resultando en

$$\begin{aligned}\mathbf{W}(z) &= [\mathbf{I} - \mathbf{F}^t(z)]^{-1} \mathbf{B}^t \mathbf{X}(z) \\ &= \mathbf{T}^t(z) \mathbf{X}(z),\end{aligned}\tag{4.20}$$

donde

$$\mathbf{T}^t(z) = [\mathbf{I} - \mathbf{F}^t(z)]^{-1} \mathbf{B}^t \quad \text{o} \quad \mathbf{T}(z) = \mathbf{B}[\mathbf{I} - \mathbf{F}(z)]^{-1} = \{T_{jk}(z)\}.$$

La matriz $\mathbf{T}(z)$ se llama *matriz función de transferencia* del sistema. Como consecuencia de la ecuación 4.20, la señal en el nodo k -ésimo $\mathbf{W}_k(z)$ puede expresarse como

$$W_k(z) = \sum_{j=1}^M T_{jk}(z) X_j(z),$$

es decir, cada variable de nodo puede expresarse como una combinación lineal de las fuentes. Si solo hay un nodo fuente no nulo, el nodo a , con valor $X_a(z)$ y hay solo un nodo sumidero con valor $Y(z)$, la salida es

$$Y(z) = \mathbf{C}^t \mathbf{W}(z) = \mathbf{C}^t \mathbf{T}^t(z) X_a(z),$$

por lo que el sistema se caracteriza por la función del sistema

$$H(z) = \mathbf{C}^t \mathbf{T}^t(z).$$

En el caso en que la función del sistema de cada rama es a lo sumo de primer orden, es decir, es una constante multiplicativa o una constante multiplicativa junto con un retardo unidad, los elementos de la matriz $\mathbf{F}^t(z)$ en la ecuación 4.19 son constantes o constantes multiplicadas por z^{-1} . Por lo tanto, pueden separarse los elementos de la matriz que no involucran un retardo con aquellos que si para expresar a $\mathbf{F}^t(z)$ como

$$\mathbf{F}^t(z) = \mathbf{F}_c^t + z^{-1} \mathbf{F}_d^t,$$

donde \mathbf{F}_c^t y \mathbf{F}_d^t son matrices $N \times N$ constantes. La primera ecuación en 4.19 puede escribirse como

$$\mathbf{W}(z) = \mathbf{F}_c^t \mathbf{W}(z) + z^{-1} \mathbf{F}_d^t \mathbf{W}(z) + \mathbf{B}^t \mathbf{X}(z).$$

Como \mathbf{F}_c^t y \mathbf{F}_d^t son constantes, puede aplicarse la transformada z inversa obteniendo que

$$\begin{aligned}\mathbf{w}[n] &= \mathbf{F}_c^t \mathbf{w}[n] + \mathbf{F}_d^t \mathbf{w}[n-1] + \mathbf{B}^t \mathbf{x}[n] \\ \mathbf{y}[n] &= \mathbf{C}^t \mathbf{w}[n],\end{aligned}\tag{4.21}$$

donde la ecuación de la salida se obtiene aplicando la transformada z inversa a la segunda ecuación en 4.19.

Ejemplo Considérese el grafo de flujo de primer orden de la figura 4.5. El conjunto de ecuaciones que representan al grafo son

$$\begin{aligned}\begin{pmatrix} w_1[n] \\ w_2[n] \\ w_3[n] \\ w_4[n] \end{pmatrix} &= \begin{pmatrix} 0 & 0 & 0 & a \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & b \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} w_1[n] \\ w_2[n] \\ w_3[n] \\ w_4[n] \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} w_1[n-1] \\ w_2[n-1] \\ w_3[n-1] \\ w_4[n-1] \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} x[n] \\ y[n] &= (0 \quad 0 \quad 1 \quad 0) \begin{pmatrix} w_1[n] \\ w_2[n] \\ w_3[n] \\ w_4[n] \end{pmatrix} = w_3[n].\end{aligned}$$

Observar que la forma de las matrices \mathbf{F}_c^t y \mathbf{F}_d^t depende del orden de las ecuaciones o, equivalentemente, de la numeración de los nodos. En el caso de este grafo, se observa que las variables de nodo no pueden calcularse en secuencia, es decir, primero $w_1[n]$, luego $w_2[n]$, y así sucesivamente, ya que por ejemplo, se necesita $w_4[n]$ para calcular $w_1[n]$. Un grafo de flujo con esta característica se llama *no computable*. Debe enfatizarse que esto no implica que las ecuaciones del grafo no puedan resolverse, si no que no pueden resolverse sucesivamente para cada variable de nodo. En algunos casos, la renumeración de los nodos puede transformar al grafo en computable. Puede mostrarse que una condición necesaria y suficiente para la computabilidad de un grafo es que la matriz \mathbf{F}_c^t tenga valores no nulos solo bajo la diagonal. Una condición necesaria y suficiente equivalente es que el grafo no tenga ramas de realimentación sin retardos.

En las ecuaciones 4.21, el valor presente de las variables de nodo están expresadas en términos del valor presente y el valor pasado, es decir, $w[n]$ está expresado en términos de $w[n]$ y $w[n-1]$. En ocasiones es conveniente que tener una representación de la red en donde los valores presentes de las variables de nodo dependan solo de sus valores pasados y el valor presente de la entrada. Esto es similar a lo que comúnmente se llama representación en variables de estado, aunque en un grafo de flujo, la cantidad de estados (nodos) es en general mayor a la cantidad de estados esenciales necesarios para representar el grafo. Una representación en variables de estado puede obtenerse a partir de las ecuaciones 4.21. Específicamente, escribiendo la primera ecuación en 4.21 como

$$(\mathbf{I} - \mathbf{F}_c^t)\mathbf{w}[n] = \mathbf{F}_d^t\mathbf{w}[n-1] + \mathbf{B}^t\mathbf{x}[n],$$

y asumiendo que la matriz $\mathbf{I} - \mathbf{F}_c^t$ es no singular, se puede resolver para $\mathbf{w}[n]$ obteniendo que

$$\mathbf{w}[n] = (\mathbf{I} - \mathbf{F}_c^t)^{-1}\mathbf{F}_d^t\mathbf{w}[n-1] + (\mathbf{I} - \mathbf{F}_c^t)^{-1}\mathbf{B}^t\mathbf{x}[n].$$

Definiendo

$$\mathbf{A} = (\mathbf{I} - \mathbf{F}_c^t)^{-1}\mathbf{F}_d^t \quad \text{y} \quad \mathbf{D} = (\mathbf{I} - \mathbf{F}_c^t)^{-1}\mathbf{B}^t$$

se obtiene la representación matricial

$$\begin{aligned} \mathbf{w}[n] &= \mathbf{A}\mathbf{w}[n-1] + \mathbf{D}\mathbf{x}[n] \\ \mathbf{y}[n] &= \mathbf{C}^t\mathbf{w}[n]. \end{aligned} \tag{4.22}$$

Puede mostrarse que si el sistema es computable, la matriz $\mathbf{I} - \mathbf{F}_c^t$ es no singular. En consecuencia puede encontrarse una representación matricial de la forma de las ecuaciones 4.22 para todo sistema computable.

4.2.3. Teorema de Tellegen para filtros digitales

En la sección 4.4 se verán las formas traspuestas de redes digitales como una representación equivalente que tiene la misma función del sistema que la red original. Una forma de mostrar esta equivalencia es mediante el teorema de Tellegen, que se formula y se demuestra en esta sección. Esta sección está basada en la sección 4.7 de[5].

En el teorema Tellegen para diagramas de flujo, se consideran dos diagramas de flujo con la misma topología pero no relacionados entre si. El hecho de tener igual topología significa que hay una correspondencia uno a uno entre los nodos de las dos redes y entre las ramas en cuanto a ubicación y dirección únicamente. Notar que este requerimiento es de poca importancia. En particular, si se considera que en un diagrama de flujo hay una rama en cada dirección entre todos los pares de nodos, con algunas ramas con transmisión nula, cualquier par de diagramas de flujo con la misma cantidad de nodos son topológicamente equivalentes. En la discusión se adoptará la convención de que cada nodo de red tiene asociado un nodo fuente conectado mediante una rama con transmitancia unidad. Adicionalmente se asume que dicho nodo fuente no está conectado a ningún otro nodo de la red. Los nodos fuente se numeran coincidiendo con el nodo de red asociado, de forma tal que la fuente x_k es la entrada al nodo de red k . Si el nodo fuente y la rama correspondiente no aparecen en el diagrama significa que el valor del nodo fuente es nulo.

Teorema de Tellegen Se consideran dos grafos de flujo con la misma topología. Sea N el número de nodos de red de cada grafo. Las variables de nodo, salidas de rama y valores de nodos fuente de la primera red se denotan w_k , v_{jk} y x_k respectivamente, y w'_k , v'_{jk} y x'_k denotan los valores correspondientes de la segunda red. Entonces se cumple que

$$\sum_{k=1}^N \sum_{j=1}^N (w_k v'_{jk} - w'_k v_{jk}) + \sum_{k=1}^N (w_k x'_k - w'_k x_k) = 0. \quad (4.23)$$

La prueba de la ecuación 4.23 se obtiene casi directamente de la definición de grafo de flujo. Recordando que las variables de nodo son (ver la ecuación 4.4)

$$w_k = \sum_{j=1}^N v_{jk} + \sum_{j=1}^M s_{jk},$$

con la convención respecto a los nodos fuente se cumple que

$$w_k = \sum_{j=1}^N v_{jk} + x_k. \quad (4.24)$$

Reemplazando este resultado en la identidad trivial

$$\sum_{k=1}^N (w_k w'_k - w'_k w_k) = 0$$

se obtiene que

$$\sum_{k=1}^N \left[w_k \left(\sum_{j=1}^N v'_{jk} + x'_k \right) - w'_k \left(\sum_{j=1}^N v_{jk} + x_k \right) \right] = 0,$$

que es la ecuación 4.23.

La ecuación 4.23 será referida como el teorema de Tellegen para grafos de flujo o redes digitales. La derivación solo depende de la ecuación 4.24 y por lo tanto no requiere que el grafo de flujo sea lineal. Se puede mostrar que si las variables W_k , W'_k , V_{jk} , V'_{jk} , X_k y X'_k se obtienen mediante una operación lineal de las variables w_k , w'_k , v_{jk} , v'_{jk} , x_k y x'_k respectivamente, se cumple que

$$\sum_{k=1}^N \sum_{j=1}^N (W_k V'_{jk} - W'_k V_{jk}) + \sum_{k=1}^N (W_k X'_k - W'_k X_k) = 0. \quad (4.25)$$

En consecuencia, el teorema de Tellegen se cumple para las variables del grafo en el dominio del tiempo (ecuación 4.23) y en el dominio $w_1[n]$ z (ecuación 4.25).

Los diagramas de flujo en la figura 4.8 ilustran el concepto de topologías equivalentes. En ambos diagramas se agregaron ramas con transmitancia nula para que ambos diagramas tengan ramas con igual ubicación y dirección, como requiere el teorema de Tellegen. Se verificará que estas redes cumplen con el teorema Tellegen. Los valores de nodo y las salidas de rama de la primera red son

$$\begin{aligned} w_1[n] &= x[n] & v_{12}[n] &= w_1[n-1] \\ w_2[n] &= w_1[n-1] & v_{13}[n] &= b_0 w_1[n] \\ w_3[n] &= b_0 w_1[n] + b_1 w_2[n] & v_{23}[n] &= b_1 w_2[n], \end{aligned}$$

y los de la segunda red son

$$\begin{aligned} w'_1[n] &= a w'_3[n] + x'[n] & v'_{31}[n] &= a w'_3[n] \\ w'_2[n] &= w'_1[n] & v'_{12}[n] &= w'_1[n] \\ w'_3[n] &= w'_2[n-1] & v'_{23}[n] &= w'_2[n-1]. \end{aligned}$$

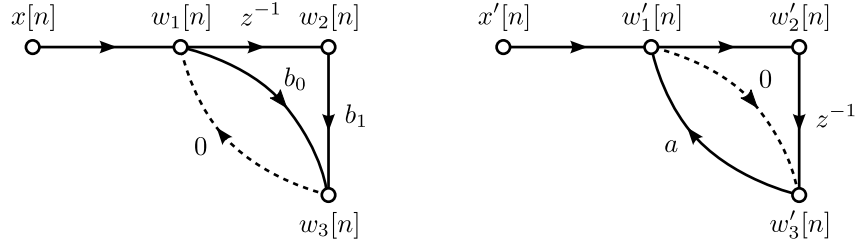


Figura 4.8: Ejemplo de dos diagramas de flujo que son topológicamente equivalentes y por lo tanto satisfacen las hipótesis de teorema de Tellegen.

Escribiendo la ecuación 4.24 para las variables que no son nulas, se tiene que

$$\begin{aligned}
& \{w_1[n]v'_{31}[n] + w_2[n]v'_{12}[n] + w_3[n]v'_{23}[n]\} \\
& - \{w'_2[n]v_{12}[n] + w'_3[n](v_{13}[n] + v_{23}[n])\} \\
& + w_1[n]x'_1[n] - w'_1[n]x[n] \\
& = \{x[n]aw'_3[n] + w_1[n-1]w'_1[n] + (b_0w_1[n] + b_1w_2[n])w'_2[n-1]\} \\
& - \{w'_1[n]w_1[n-1] + w'_2[n-1](b_0w_1[n] + b_1w_2[n])\} \\
& + x[n]x'[n] - (aw'_3[n] + x'[n])x[n] \\
& = 0,
\end{aligned}$$

que es lo que se quería comprobar.

Redes digitales recíprocas e interrecíprocas

Para la noción de reciprocidad de redes digitales, se considera una red digital excitada por dos conjuntos distintos de fuentes. La transformada z de los valores de los nodos fuente son X_k y X'_k y los valores correspondientes de los nodos de red son W_k y W'_k . La red se dice *recíproca* si para todo par de distribuciones de señales se cumple que

$$\sum_{k=1}^N [W_k X'_k - W'_k X_k] = 0. \quad (4.26)$$

Como una consecuencia de la reciprocidad considérese dos nodos a y b de la red. Adicionalmente, sean todos los valores de los nodos fuente X_k nulos excepto X_a y todos los valores de los nodos fuente X'_k nulos excepto X'_b . Por lo tanto, de la ecuación 4.26 se cumple que

$$W_b X'_b = W'_a X_a,$$

y si $X_a = X'_b$ resulta en que

$$W_b = W'_a.$$

Esto significa que si se excita el nodo a y se observa la salida del nodo b , para un grafo recíproco, la misma excitación en el nodo b producirá la misma salida en el nodo a .

Las redes digitales pueden no ser recíprocas. Un concepto vinculado mas útil para redes digitales es el de *interreciprocidad*. En este caso, se consideran dos grafos de flujo distintos. Sean X_k los valores de los nodos fuente y W_k los valores de los nodos de red de una de las redes y X'_k y W'_k los valores de los nodos fuente y los nodos de red de la otra red. Las dos redes se dicen interrecíprocas si

$$\sum_{k=1}^N [W_k X'_k - W'_k X_k] = 0. \quad (4.27)$$

Si bien esta ecuación es similar a la ecuación 4.26 es importante tener en cuenta que para reciprocidad ambas redes solo difieren en las fuentes mientras que para interreciprocidad tanto las fuentes como las transmitancias de las ramas pueden diferir en ambas redes. Una red que es recíproca es interrecíproca consigo misma. Por otro lado, dos redes pueden ser interrecíprocas sin que ninguna de ellas sea recíproca.

Prueba del teorema de transposición

Una propiedad útil de las redes digitales es que son interrecíprocas con su transpuesta. El transpuesto de un diagrama de flujo se obtiene invirtiendo la dirección de todas las ramas pero manteniendo la transmitancia.

Se considera una red digital donde W_k denota la variable del nodo de red k -ésimo. La transmisión del nodo j al nodo k se denota F_{jk} , y por lo tanto,

$$V_{jk} = F_{jk}W_j. \quad (4.28)$$

En la red transpuesta el nodo de red k -ésimo se denota W'_k y la transmitancia de la rama entre los nodos j y k se denota F'_{jk} , de forma tal que

$$V'_{jk} = F'_{jk}W'_j. \quad (4.29)$$

Por definición de red transpuesta se cumple que

$$F'_{jk} = F_{kj}.$$

Para probar que una red y su transpuesta son interrecíprocas, es decir, para probar que se cumple la ecuación 4.27 en las condiciones mencionadas, se emplea el hecho de que una red y su transpuesta tienen la misma topología y por lo tanto, se cumple el teorema de Tellegen dado por la ecuación 4.25,

$$\sum_{k=1}^N \sum_{j=1}^N (W_k V'_{jk} - W'_k V_{jk}) + \sum_{k=1}^N (W_k X'_k - W'_k X_k) = 0.$$

Reemplazando las ecuaciones 4.28 y 4.29 en esta ecuación se obtiene que

$$\sum_{k=1}^N \sum_{j=1}^N (W_k F'_{jk} W'_j - W'_k F_{jk} W_j) + \sum_{k=1}^N (W_k X'_k - W'_k X_k) = 0.$$

o equivalentemente

$$\sum_{k=1}^N \sum_{j=1}^N W_k W'_j F'_{jk} - \sum_{k=1}^N \sum_{j=1}^N W'_k W_j F_{jk} + \sum_{k=1}^N (W_k X'_k - W'_k X_k) = 0.$$

Intercambiando los índices de la primer doble sumatoria se obtiene que

$$\sum_{k=1}^N \sum_{j=1}^N (W'_k W_j F'_{kj} - W'_k W_j F_{jk}) + \sum_{k=1}^N (W_k X'_k - W'_k X_k) = 0,$$

pero como las redes son transpuestas y por lo tanto $F'_{kj} = F_{jk}$, la doble sumatoria se anula resultando en que

$$\sum_{k=1}^N (W_k X'_k - W'_k X_k) = 0, \quad (4.30)$$

lo que prueba que una red y su transpuesta son interrecíprocas.

Una consecuencia importante de este hecho es que para redes con entrada única y salida única, una red y su transpuesta tienen la misma función de transferencia. Para mostrar esto a partir de la ecuación 4.30 se consideran dos nodos cualquiera a y b . Adicionalmente, sean nulos todos los nodos fuentes excepto X_a en la red original y excepto X'_b en la red transpuesta. De esta forma, de la ecuación 4.30,

$$W_b X'_b = W'_a X_a.$$

Por lo tanto, si la misma excitación es aplicada en el nodo a en la red original y en el nodo b en la red transpuesta, es decir, $X'_b = X_a$, la misma respuesta se observa en el nodo a en la red transpuesta que en el nodo b en la red original, es decir, $W'_a = W_b$. Esto se muestra esquemáticamente en la figura 4.9

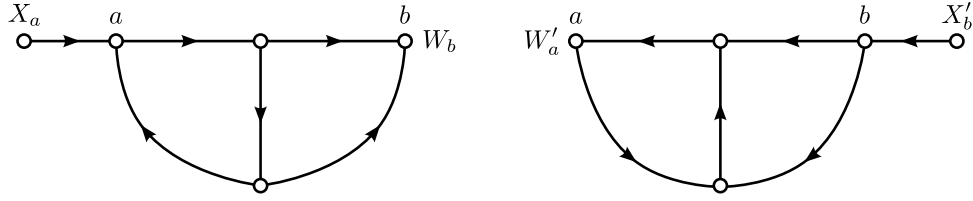


Figura 4.9: Ilustración de la interreciprocidad de redes transpuestas. Para sistemas de una entrada y una salida, la función de transferencia de redes transpuestas es la misma.

4.3. Estructuras básicas para sistemas IIR

En la sección 4.1 se introdujo dos estructuras equivalentes para la representación de sistemas LTI con función del sistema como en la ecuación 4.2. En esta sección se presentan los grafos de flujo de esos sistemas y se desarrollan otras estructuras comunes en redes de grafos equivalentes. Una consideración para la elección entre diversas estructuras equivalentes es la complejidad computacional. Por ejemplo, puede preferirse el menor número posible de multiplicadores por una constante para incrementar la velocidad y el menor número posible de retardos para reducir la memoria requerida. En la elección también hay otros factores, como la capacidad de particionar el algoritmo para la implementación en multiprocesadores, o la sensibilidad de la estructura de grafo a la representación numérica con precisión finita.

4.3.1. Formas directas

En la sección 4.1 se obtuvo la representación en diagrama de bloques de la forma directa I (figura 4.1) y la forma directa II o forma canónica directa (figura 4.3) de sistemas LTI con entrada y salida que satisfacen una ecuación en diferencias de la forma dada por la ecuación 4.1 con función del sistema correspondiente dada por la ecuación 4.2. En las figuras 4.10 y 4.11 se muestra el grafo de flujo de

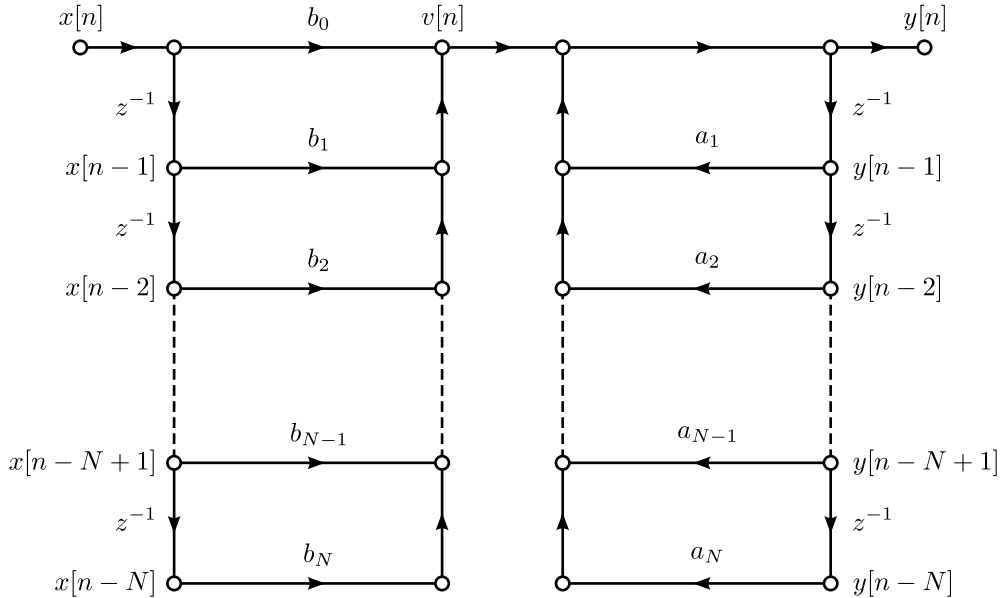


Figura 4.10: Diagrama de flujo de una estructura en forma directa I de un sistema de orden N .

la estructura en forma directa I y en forma directa II respectivamente. Nuevamente se asumió por conveniencia que $N = M$.

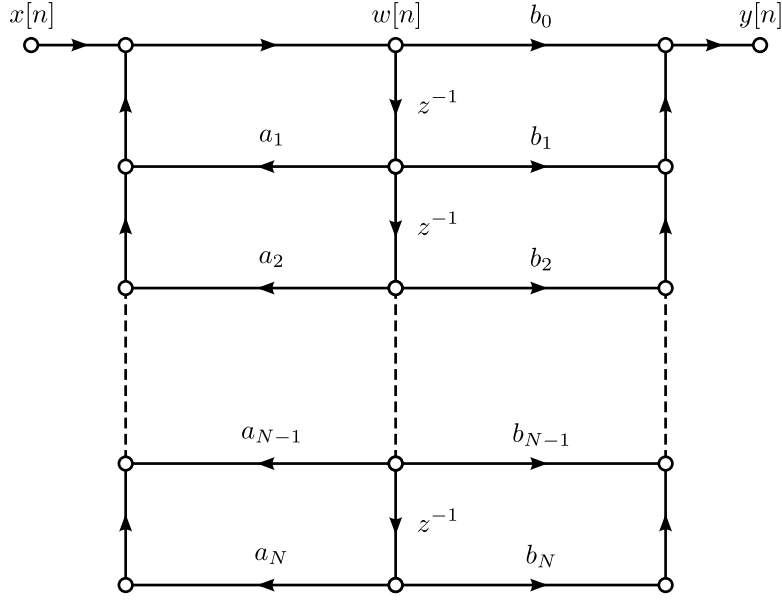


Figura 4.11: Diagrama de flujo de una estructura en forma directa II de un sistema de orden N .

4.3.2. Forma en cascada

Las estructuras en forma directa surgen directamente de la función del sistema $H(z)$ expresada como cociente de polinomios en la variable z^{-1} . Factorizando los polinomios del numerador y del denominador, $H(z)$ se puede expresar como

$$H(z) = A \frac{\prod_{k=1}^{M_1} (1 - f_k z^{-1}) \prod_{k=1}^{M_2} (1 - g_k z^{-1})(1 - g_k^* z^{-1})}{\prod_{k=1}^{N_1} (1 - c_k z^{-1}) \prod_{k=1}^{N_2} (1 - d_k z^{-1})(1 - d_k^* z^{-1})}, \quad (4.31)$$

donde $M = M_1 + 2M_2$ y $N = N_1 + 2N_2$, los factores de primer orden representan ceros reales en f_k y polos reales en c_k , y los factores de segundo orden representan pares de ceros complejos conjugados en g_k y g_k^* y pares de polos complejos conjugados en d_k y d_k^* . Esto es la distribución mas general de polos y ceros cuando todos los coeficientes en la ecuación 4.2 son reales. La ecuación 4.31 sugiere una clase de estructuras que consiste en la cascada de sistemas de primer y segundo orden. Hay una considerable libertad en la elección de la composición de los subsistemas y en el orden de los subsistemas en la cascada. En la practica, puede ser deseable implementar la cascada con requerimientos de memoria y cálculo mínimo. En este sentido, una estructura modular ventajosa en muchas implementaciones se obtiene combinando pares de factores reales y pares de complejos conjugados en factores de segundo orden de forma que la ecuación 4.31 puede expresarse como

$$H(z) = \prod_{k=1}^{N_s} \frac{b_{0k} + b_{1k} z^{-1} + b_{2k} z^{-2}}{1 + a_{1k} z^{-1} + a_{2k} z^{-2}}, \quad (4.32)$$

donde $N_s = \lfloor (N+1)/2 \rfloor$. Al expresar $H(z)$ de esta forma, se asume que $M \leq N$ y los polos y ceros reales se combinaron de a pares. Si hay un número impar de ceros, uno de los coeficientes b_{2k} es cero, y lo análogo con los polos. Las estructuras de segundo orden individuales pueden implementarse empleando cualquiera de las formas directas, pero la forma directa II emplea menor cantidad de retardos y multiplicaciones. Una estructura en cascada de un sistema de orden seis empleando tres subsistemas de segundo orden en forma directa II se muestra en la figura 4.12. Las ecuaciones en diferencias correspondientes a una

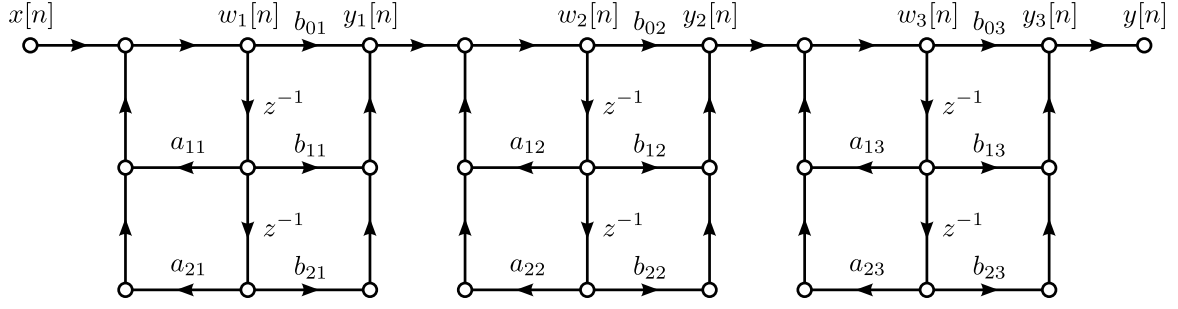


Figura 4.12: Estructura en cascada de un sistema de orden seis empleando tres subsistemas de segundo orden en forma directa II.

estructura en cascada general de subsistemas de segundo orden son

$$\begin{aligned} y_0[n] &= x[n], \\ w_k[n] &= a_{1k}w_k[n-1] + a_{2k}w_k[n-2] + y_{k-1}[n], \quad k = 1, 2, \dots, N_s, \\ y_k[n] &= b_{0k}w_k[n] + b_{1k}w_k[n-1] + b_{2k}w_k[n-2], \quad k = 1, 2, \dots, N_s, \\ y[n] &= y_{N_s}[n]. \end{aligned}$$

Puede verse que se pueden formar una cantidad variada de sistemas teóricamente equivalentes emparejando los polos y los ceros de distintas formas y ordenando los subsistemas de segundo orden de formas distintas. Efectivamente, si hay N_s secciones de segundo orden, hay $N_s!$ formas distintas de emparejar ceros con polos y $N_s!$ formas distintas de ordenar los subsistemas en la cascada, resultando en un total de $(N_s!)^2$ sistemas distintos equivalentes. Si bien todos los sistemas tienen la misma función del sistema y la misma relación entrada-salida, el comportamiento al emplear aritmética de precisión finita puede ser muy diferente.

Finalmente, se realizan algunos comentarios respecto a la cantidad de operaciones de las estructuras. Cada subsistema de segundo orden de la estructura en cascada tiene cinco constantes multiplicativas cuando se define mediante la ecuación 4.32. Como comparación, asúmase que $M = N$ en $H(z)$ dado por la ecuación 4.2 y además, que N es un número par, y por lo tanto $N_s = N/2$. De esta forma, las estructuras de forma directa I y II tienen $2N + 1$ constantes multiplicativas, mientras que la estructura en cascada definida por la ecuación 4.32 tiene $5N/2$ constantes multiplicativas. Para el sistema de orden seis de la figura 4.12 se requieren 15 multiplicadores, mientras que las formas directas requieren 13 multiplicadores. Otra definición de la forma en cascada es

$$H(z) = b_0 \prod_{k=1}^{N_s} \frac{1 + \tilde{b}_{1k}z^{-1} + \tilde{b}_{2k}z^{-2}}{1 + a_{1k}z^{-1} + a_{2k}z^{-2}},$$

donde b_0 es el primer coeficiente en el numerador en la ecuación 4.2 y $\tilde{b}_{ik} = b_{ik}/b_{0k}$ para $i = 1, 2$ y $k = 1, 2, \dots, N_s$. Esta forma de $H(z)$ sugiere la cascada de secciones de segundo orden con cuatro multiplicadores con una única constante de ganancia global b_0 , resultando en un total de $2N + 1$ multiplicadores, igual que las formas directas.

4.3.3. Forma en paralelo

Como alternativa de factorización de los polinomios del numerador y denominador de $H(z)$, una función del sistema racional como la dada por la ecuación 4.2 o 4.31 se puede expandir en fracciones simples como (ver la sección 2.3.3)

$$H(z) = \sum_{k=0}^{N_p} C_k z^{-k} + \sum_{k=1}^{N_1} \frac{A_k}{1 - c_k z^{-k}} + \sum_{k=1}^{N_2} \frac{B_k(1 - e_k z^{-1})}{(1 - d_k z^{-1})(1 - d_k^* z^{-1})}, \quad (4.33)$$

donde $N = N_1 + 2N_2$. Si $M > N$, es decir, si la función racional es impropia, $N_p = M - N$; en caso contrario, la primera sumatoria en la ecuación 4.33 no se incluye. Si los coeficientes a_k y b_k son reales en

la ecuación 4.2, las constantes A_k , B_k , C_k , c_k y e_k son todas reales. Expresado de esta forma, el sistema puede interpretarse como el paralelo de sistemas IIR de primer y segundo orden, con posiblemente N_p caminos de escalado y retardo. Alternativamente, se puede agrupar a los polos reales de a pares para expresar a $H(z)$ como

$$H(z) = \sum_{k=0}^{N_p} C_k z^{-k} + \sum_{k=1}^{N_s} \frac{e_{0k} - e_{1k} z^{-1}}{1 + a_{1k} z^{-1} + a_{2k} z^{-2}},$$

donde, al igual que en la forma en cascada, $N_s = \lfloor (N+1)/2 \rfloor$, y si $N_p = M - N$ es negativo, la primera sumatoria no esta presente. Un ejemplo con $N = M = 6$ se muestra en la figura 4.13. Las ecuaciones en diferencias generales correspondientes a una estructura en paralelo de subsistemas de segundo orden son

$$\begin{aligned} w_k[n] &= a_{1k} w_k[n-1] + a_{2k} w_k[n-2] + x[n], & k = 1, 2, \dots, N_s, \\ y_k[n] &= e_{0k} w_k[n] + e_{1k} w_k[n-1], & k = 1, 2, \dots, N_s, \\ y[n] &= \sum_{k=0}^{N_p} C_k x[n-k] + \sum_{k=1}^{N_s} y_k[n]. \end{aligned}$$

Si $M < N$, la primera sumatoria en la ecuación de $y[n]$ no se incluye.

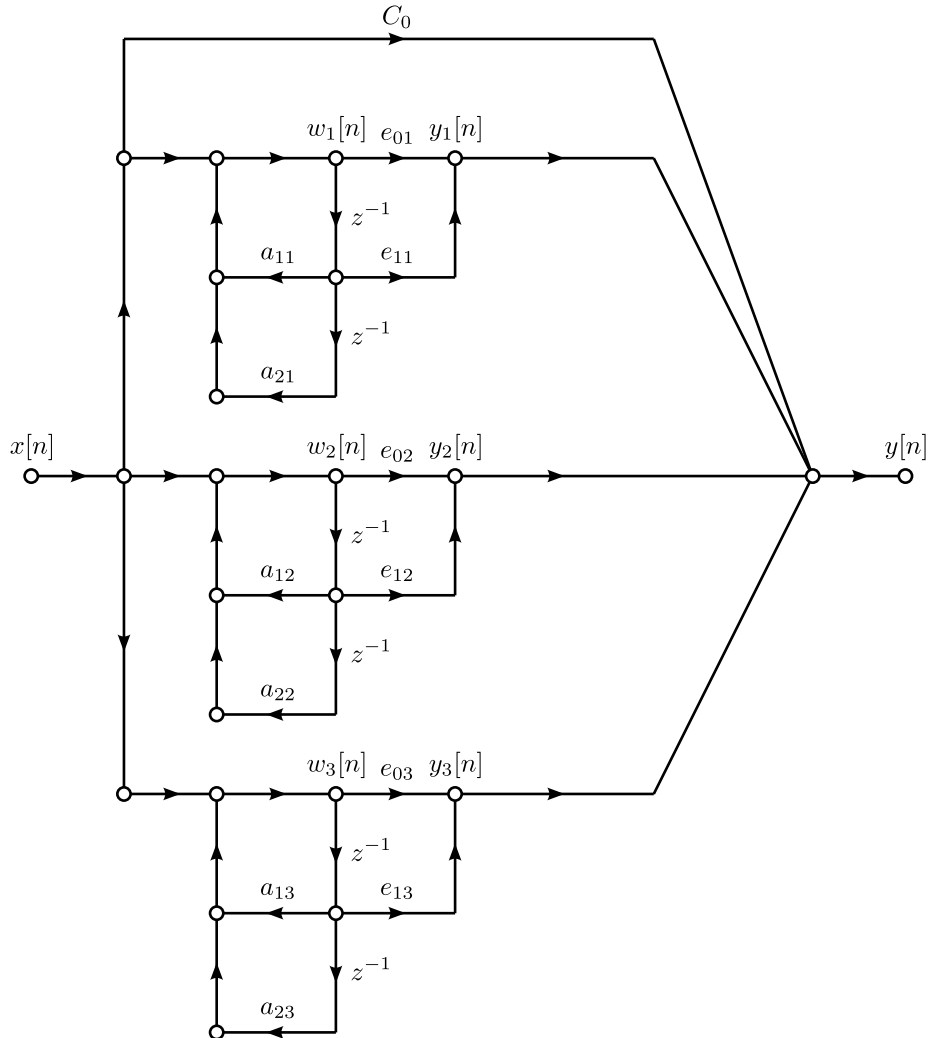


Figura 4.13: Estructura de forma en paralelo de un sistema de sexto orden $N = M = 6$ con los polos reales y complejos agrupados de a pares.

Ejemplo: ilustración de estructuras en paralelo Sea el sistema con función del sistema

$$H(z) = \frac{1 + 2z^{-1} + z^{-2}}{1 - 0.75z^{-1} + 0.125z^{-2}}.$$

Para expresar la función del sistema en forma en paralelo, hay que descomponerlo en fracciones simples como en la ecuación 4.33. En este caso, la función del sistema es racional impropia con $N_p = 2 - 2 = 0$. El término C_0 de la ecuación 4.33 se obtiene mediante división larga con los polinomios del numerador y denominador ordenados en orden reverso, como se explicó en la sección 2.3.3, de la siguiente forma:

$$\begin{array}{r} 8 \\ 0.125z^{-2} - 0.75z^{-1} + 1 \overline{) \begin{array}{r} z^{-2} + 2z^{-1} + 1 \\ z^{-2} - 6z^{-1} + 8 \\ \hline -8z^{-1} - 7 \end{array}} \end{array}$$

Esto indica que

$$H(z) = \frac{1 + 2z^{-1} + z^{-2}}{1 - 0.75z^{-1} + 0.125z^{-2}} = 8 + \frac{-7 + 8z^{-1}}{1 - 0.75z^{-1} + 0.125z^{-2}}.$$

Esta es la expresión del sistema en forma en paralelo con una estructura de segundo orden y se muestra en la figura 4.14. Considerando que el polinomio del denominador tiene raíces reales 0.5 y 0.25, el sistema también se puede expresar como el paralelo de estructuras de primer orden continuando con la expansión en fracciones simples,

$$H(z) = 8 + \frac{A_1}{1 - 0.5z^{-1}} + \frac{A_2}{1 - 0.25z^{-1}},$$

donde A_1 y A_2 se obtienen mediante la ecuación 2.9 como

$$\begin{aligned} A_1 &= \left. \frac{-7 + 8z^{-1}}{1 - 0.25z^{-1}} \right|_{z=0.5} = \frac{9}{0.5} = 18 \\ A_2 &= \left. \frac{-7 + 8z^{-1}}{1 - 0.5z^{-1}} \right|_{z=0.25} = \frac{25}{-1} = -25, \end{aligned}$$

es decir,

$$H(z) = 8 + \frac{9}{1 - 0.5z^{-1}} - \frac{25}{1 - 0.25z^{-1}},$$

cuyo diagrama de flujo se muestra en la figura 4.14.

4.3.4. Estructuras de muestreo en frecuencia

Todos los grafos de flujo en esta sección tienen bucles de realimentación. Estos bucles son necesarios, pero no suficientes, para generar sistemas con respuesta al impulso infinita. Si una función del sistema tiene polos, el correspondiente diagrama de bloques o grafo de flujo tendrá bucles de realimentación. Sin embargo, ni los polos de la función del sistema ni los bucles de realimentación son suficientes para que la respuesta al impulso sea infinitamente larga. Esto es porque los polos pueden cancelarse con ceros. Una clase de sistemas con bucles de realimentación pero respuesta al impulso finita son los *sistemas de muestreo en frecuencia*, y se introducen a continuación. Lo siguiente está basado en la sección 9.2.3 de [2]. Ver también la sección 4.5.4 de [5].

Los sistemas de muestreo en frecuencia son estructuras alternativas de sistemas FIR en los cuales los parámetros que caracterizan al filtro son los valores de la respuesta en frecuencia en lugar de la respuesta al impulso. Para su diseño se parte especificando la respuesta en frecuencia deseada en el conjunto equiespaciado de frecuencias

$$\omega_k = \frac{2\pi}{N}(k + \alpha), \quad k = 0, 1, \dots, N - 1,$$

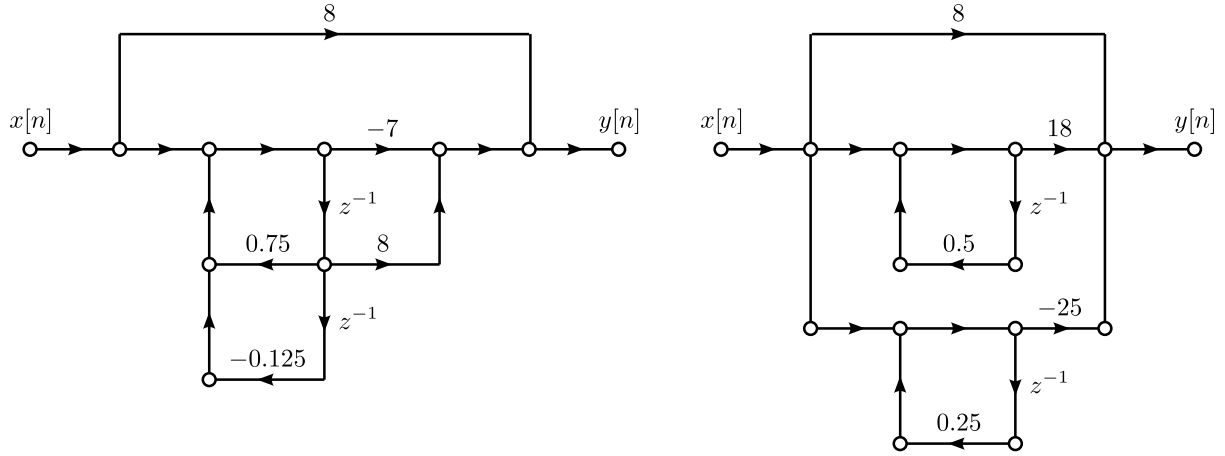


Figura 4.14: Estructuras de forma en paralelo con subsistemas de segundo orden y con subsistemas de primer orden correspondientes al ejemplo.

donde $\alpha = 0$ o $\alpha = 1/2$. Estas frecuencias se muestran en la figura 4.15 cuando N es par e impar. Notar que cuando $\alpha = 1/2$, las frecuencias se desplazan π/N radianes respecto a cuando $\alpha = 0$ manteniendo la estructura de puntos complejos conjugados en el plano complejo. En el caso de un sistema con respuesta

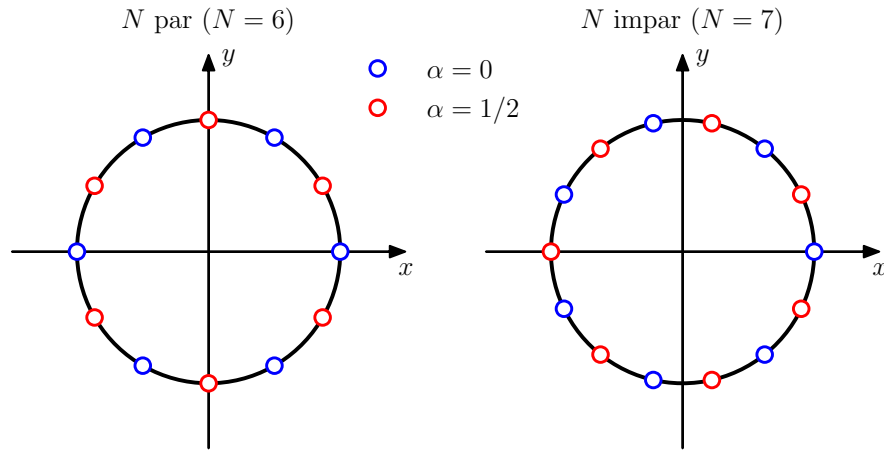


Figura 4.15: Frecuencias en donde se especifica la respuesta en frecuencia deseada de un sistema de muestreo en frecuencia.

al impulso real, hay que especificar únicamente la respuesta en frecuencia en el semiplano superior, es decir, para

$$k = 0, 1, \dots, \frac{N}{2} \pm 1 \quad \text{cuando } N \text{ es par,}$$

donde el signo positivo corresponde al caso en que $\alpha = 0$ y el signo negativo al caso en que $\alpha = 1/2$ (ver la figura 4.15) y

$$k = 0, 1, \dots, \frac{N-1}{2} \quad \text{cuando } N \text{ es impar.}$$

Especificando la respuesta en frecuencia del sistema en dichas frecuencias, se quiere calcular la función del sistema. Para hacerlo, se parte considerando que la respuesta en frecuencia es

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

y los valores especificados de $H(e^{j\omega})$ en las frecuencias ω_k son

$$\tilde{H}[k] = H(e^{j2\pi(k+\alpha)/N}) = \sum_{n=0}^{N-1} h[n] e^{-j2\pi(k+\alpha)n/N}, \quad k = 0, 1, \dots, N-1. \quad (4.34)$$

El conjunto de valores $\tilde{H}[k]$ se llaman muestras en frecuencia de $H(e^{j\omega})$. En el caso en que $\alpha = 0$, $\tilde{H}[k]$ es la *transformada discreta de Fourier* (*DFT*, Discrete Fourier Transform) de N puntos de $h[n]$, como se verá mas adelante en el capítulo ???. También en dicho capítulo se verá que la ecuación 4.34 puede invertirse para obtener la respuesta al impulso $h[n]$ en función de las muestras en frecuencia $\tilde{H}[k]$ como

$$h[n] = \frac{1}{N} \sum_{k=0}^{N-1} \tilde{H}[k] e^{j2\pi(k+\alpha)n/N}, \quad n = 0, 1, \dots, N-1. \quad (4.35)$$

Este resultado puede comprobarse sustituyendo esta ecuación en la ecuación 4.34 y observando que se obtiene una identidad. Cuando $\alpha = 0$, la ecuación 4.35 es la *transformada discreta de Fourier inversa* (*IDFT*, Inverse Discrete Fourier Transform) de $\tilde{H}[k]$ (ver el capítulo ???). Continuando, la función del sistema es

$$\begin{aligned} H(z) &= \sum_{n=0}^{N-1} h[n] z^{-n} \\ &\stackrel{(a)}{=} \sum_{n=0}^{N-1} \left\{ \frac{1}{N} \sum_{k=0}^{N-1} \tilde{H}[k] e^{j2\pi(k+\alpha)n/N} \right\} z^{-n} \\ &= \frac{1}{N} \sum_{k=0}^{N-1} \tilde{H}[k] \sum_{n=0}^{N-1} \left[e^{j2\pi(k+\alpha)/N} z^{-1} \right]^n \\ &= \frac{1}{N} \sum_{k=0}^{N-1} \tilde{H}[k] \frac{1 - e^{j2\pi(k+\alpha)} z^{-N}}{1 - e^{j2\pi(k+\alpha)/N} z^{-1}}, \end{aligned}$$

donde en (a) se reemplazó $h[n]$ por el resultado de la ecuación 4.35. Considerando que $e^{j2\pi(k+\alpha)} = e^{j2\pi\alpha}$ resulta en que la función del sistema es

$$H(z) = \frac{1 - e^{j2\pi\alpha} z^{-N}}{N} \sum_{k=0}^{N-1} \frac{\tilde{H}[k]}{1 - e^{j2\pi(k+\alpha)/N} z^{-1}}. \quad (4.36)$$

Se observa que la función del sistema tiene ceros en

$$z_k = e^{j2\pi(k+\alpha)/N} \quad k = 0, 1, \dots, N-1,$$

que se muestran en la figura 4.15 para valores específicos de N , y polos en las mismas posiciones. Además, se caracteriza por el conjunto de muestras en frecuencia $\tilde{H}[k]$ en lugar de la respuesta al impulso $h[n]$. Se confirmará que $H(z_m) = \tilde{H}[m]$ cuando $z_m = e^{j2\pi(m+\alpha)/N}$ para $m = 0, 1, \dots, N-1$. Se parte evaluando a la función del sistema 4.36 en z_m ,

$$\begin{aligned} H(z_m) &\stackrel{(a)}{=} \frac{1}{N} \sum_{k=0}^{N-1} \frac{\tilde{H}[k] (1 - e^{j2\pi\alpha} z^{-N})}{1 - z_k z^{-1}} \Bigg|_{z=z_m} \\ &\stackrel{(b)}{=} \frac{1}{N} \frac{\tilde{H}[m] (1 - e^{j2\pi\alpha} z^{-N})}{1 - z_m z^{-1}} \Bigg|_{z=z_m} \\ &\stackrel{(c)}{=} \frac{\tilde{H}[m]}{N} \frac{(-e^{j2\pi\alpha})(-N) z^{-(N+1)}}{(-z_m)(-1) z^{-2}} \Bigg|_{z=z_m} \\ &= \tilde{H}[m] \frac{e^{j2\pi\alpha} z_m^{-N}}{z_m z_m^{-1}} \\ &= \tilde{H}[m] e^{j2\pi\alpha} e^{-j2\pi(m+\alpha)} \\ &\stackrel{(d)}{=} \tilde{H}[m], \end{aligned}$$

donde en (a) se consideró que la expresión no se puede evaluar directamente porque cuando $k = m$ queda una indeterminación $0/0$, en (b) se tuvo en cuenta que todos los sumandos de la sumatoria se anulan excepto cuando $k = m$, ya que si $k \neq m$ se anula el numerador pero no el denominador, y cuando $k = m$ se anula tanto el numerador como el denominador y en (c) se aplicó la regla de L'Hôpital para evaluar a la indeterminación. Esto concluye la confirmación.

De la ecuación del sistema 4.36, se observa que este sistema FIR consiste en la cascada de un sistema FIR, dado por el factor que multiplica a la sumatoria, con el paralelo de sistemas IIR de primer orden. La ubicación de los ceros del sistema FIR coincide con la ubicación de los polos de los sistemas IIR, y dichos polos y ceros tienen la frecuencia en donde la respuesta en frecuencia deseada es especificada. En la figura 4.16 se muestra el diagrama de flujo del sistema. Los factores de ganancia $\tilde{H}[k]$ de los sistemas IIR en paralelo son números complejos, así como las ganancias $z_k = e^{j2\pi(k+\alpha)/N}$ de los bucles de realimentación.

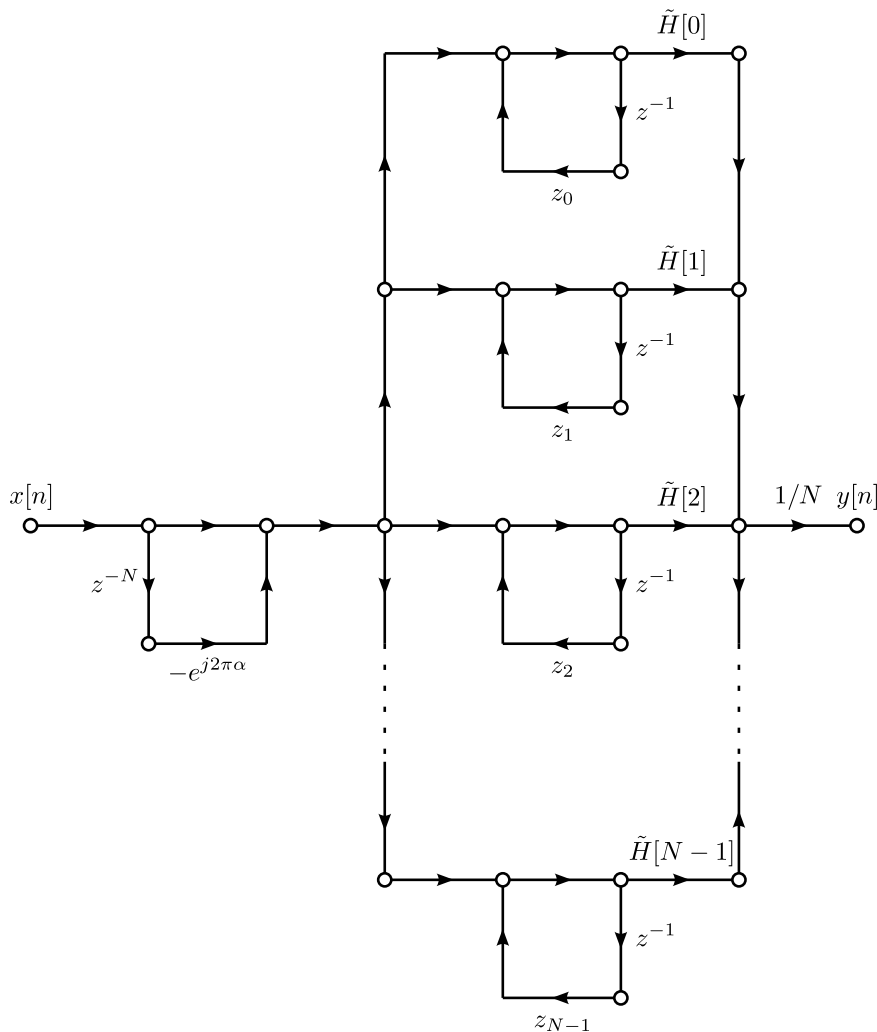


Figura 4.16: Diagrama de flujo de la estructura de muestreo en frecuencia, cuya función del sistema está dada por la ecuación 4.36.

En el caso en que $h[n]$ es real, la implementación del sistema de muestreo en frecuencia involucra únicamente parámetros reales. Para ver esto, se considerará únicamente el caso en que $\alpha = 0$, pero lo mismo se cumple para $\alpha = 1/2$. De la ecuación 4.34 puede comprobarse que si $h[n]$ es real, se cumple que (ver la figura 4.15)

N par:

$\tilde{H}[0], \tilde{H}[N/2]$ reales

$$\tilde{H}[N-k] = \tilde{H}^*[k], \quad k = 1, \dots, \frac{N}{2} - 1$$

N impar:

$\tilde{H}[0]$ real

$$\tilde{H}[N-k] = \tilde{H}^*[k], \quad k = 1, \dots, \frac{N-1}{2}.$$

Por lo tanto, si N es par, la ecuación del sistema 4.36 puede escribirse como (ver el problema 6.51 de [1])

$$\begin{aligned}
H(z) &= \frac{1 - z^{-N}}{N} \left\{ \frac{\tilde{H}[0]}{1 - z^{-1}} + \frac{\tilde{H}[N/2]}{1 + z^{-1}} + \sum_{k=1}^{N/2-1} \left(\frac{\tilde{H}[k]}{1 - e^{j2\pi k/N} z^{-1}} + \frac{\tilde{H}[N-k]}{1 - e^{j2\pi(N-k)/N} z^{-1}} \right) \right\} \\
&= \frac{1 - z^{-N}}{N} \left\{ \frac{\tilde{H}[0]}{1 - z^{-1}} + \frac{\tilde{H}[N/2]}{1 + z^{-1}} + \sum_{k=1}^{N/2-1} \left(\frac{\tilde{H}[k]}{1 - e^{j2\pi k/N} z^{-1}} + \frac{\tilde{H}^*[k]}{1 - e^{-j2\pi k/N} z^{-1}} \right) \right\} \\
&= \frac{1 - z^{-N}}{N} \left\{ \frac{\tilde{H}[0]}{1 - z^{-1}} + \frac{\tilde{H}[N/2]}{1 + z^{-1}} + \sum_{k=1}^{N/2-1} \frac{\tilde{H}[k](1 - e^{-j2\pi k/N} z^{-1}) + \tilde{H}^*[k](1 - e^{j2\pi k/N} z^{-1})}{(1 - e^{j2\pi k/N} z^{-1})(1 - e^{-j2\pi k/N} z^{-1})} \right\} \\
&= \frac{1 - z^{-N}}{N} \left\{ \frac{\tilde{H}[0]}{1 - z^{-1}} + \frac{\tilde{H}[N/2]}{1 + z^{-1}} + \sum_{k=1}^{N/2-1} \frac{\tilde{H}[k] + \tilde{H}^*[k] - (\tilde{H}[k]e^{-j2\pi k/N} + \tilde{H}^*[k]e^{j2\pi k/N})z^{-1}}{1 - 2\cos(2\pi k/N)z^{-1} + z^{-2}} \right\}.
\end{aligned}$$

Expresando $\tilde{H}[k] = |\tilde{H}[k]|e^{j\theta[k]}$, resulta en

$$H(z) = \frac{1 - z^{-N}}{N} \left\{ \frac{\tilde{H}[0]}{1 - z^{-1}} + \frac{\tilde{H}[N/2]}{1 + z^{-1}} + \sum_{k=1}^{N/2-1} 2|\tilde{H}[k]| \frac{\cos(\theta[k]) - \cos(\theta[k] - 2\pi k/N)}{1 - 2\cos(2\pi k/N)z^{-1} + z^{-2}} \right\}.$$

De esta forma, el paralelo de filtros IIR en la figura 4.16 consiste en filtros de primer orden y de segundo orden todos con coeficientes reales. Expresiones similares pueden obtenerse para N impar y $\alpha = 1/2$.

Ejemplo: sistema de muestreo en frecuencia Este ejemplo es el problema 6.39 de [1]. Se considera el sistema FIR con respuesta al impulso

$$h[n] = \begin{cases} \frac{1}{15} \left(1 + \cos \left[\frac{2\pi}{15}(n - n_0) \right] \right), & 0 \leq n \leq 14 \\ 0, & \text{en otro caso.} \end{cases} \quad (4.37)$$

En la figura se muestra la respuesta al impulso para $n_0 = 0$ y $n_0 = 15/2$. Observar que para $n_0 = 15/2$ la respuesta al impulso cumple que $h[n] = h[15 - n]$ para todo n , es decir, es simétrica respecto al punto $15/2$. Para $n_0 = 0$, la condición $h[n] = h[15 - n]$ no se cumple cuando $n = 0$, por lo que no hay simetría en ese caso.

Se calculará la función del sistema. Para eso, aplicando la transformada z a la respuesta al impulso 4.37 se observa que

$$\begin{aligned}
H(z) &= \sum_{n=0}^{14} \frac{1}{15} \left(1 + \cos \left[\frac{2\pi}{15}(n - n_0) \right] \right) z^{-n} \\
&= \frac{1}{15} \sum_{n=0}^{14} \left[1 + \frac{e^{j2\pi(n-n_0)/15} + e^{-j2\pi(n-n_0)/15}}{2} \right] z^{-n} \\
&= \frac{1}{15} \left(\sum_{n=0}^{14} z^{-n} + \frac{e^{-j2\pi n_0/15}}{2} \sum_{n=0}^{14} e^{j2\pi n/15} z^{-n} + \frac{e^{j2\pi n_0/15}}{2} \sum_{n=0}^{14} e^{-j2\pi n/15} z^{-n} \right) \\
&= \frac{1}{15} \left(\frac{1 - z^{-15}}{1 - z^{-1}} + \frac{e^{-j2\pi n_0/15}}{2} \cdot \frac{1 - e^{j2\pi} z^{-15}}{1 - e^{j2\pi/15} z^{-1}} + \frac{e^{j2\pi n_0/15}}{2} \cdot \frac{1 - e^{-j2\pi} z^{-15}}{1 - e^{-j2\pi/15} z^{-1}} \right),
\end{aligned}$$

resultando en que

$$H(z) = \frac{1 - z^{-15}}{15} \left(\frac{1}{1 - z^{-1}} + \frac{e^{-j2\pi n_0/15}}{2} \cdot \frac{1}{1 - e^{j2\pi/15} z^{-1}} + \frac{e^{j2\pi n_0/15}}{2} \cdot \frac{1}{1 - e^{-j2\pi/15} z^{-1}} \right). \quad (4.38)$$

El factor $(1 - z^{-15})$ en la función del sistema introduce ceros en $z = e^{j2\pi k/15}$ para $k = 0, 1, \dots, 14$, y en este caso, solo hay tres polos en $z = 1$, $z = e^{j2\pi/15}$ y $z = e^{-j2\pi/15}$, que se cancelan con los ceros

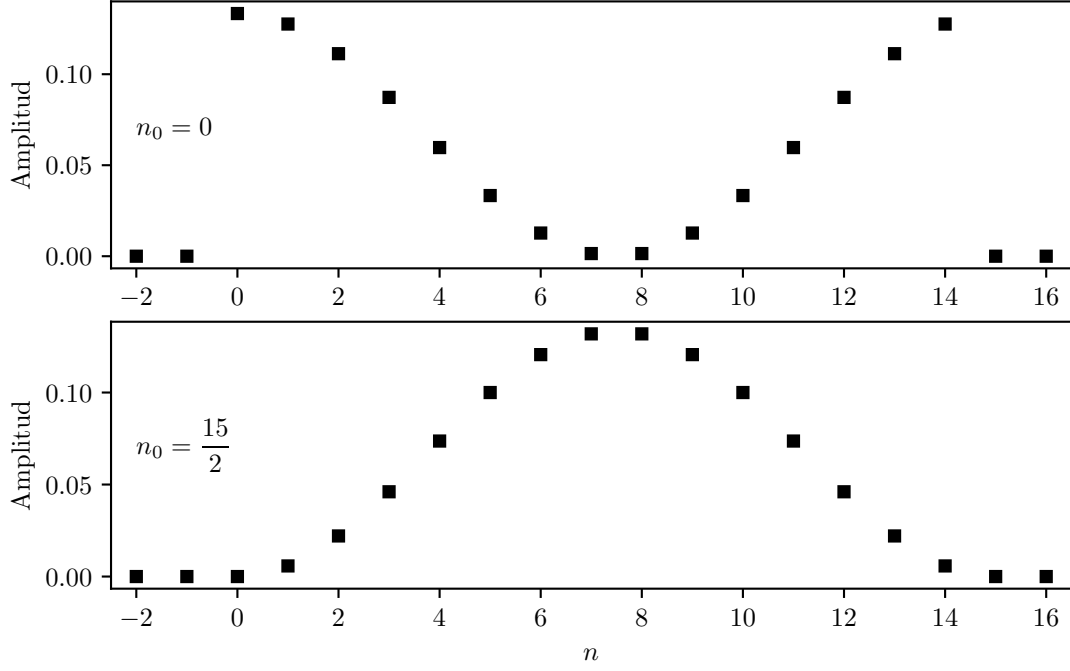


Figura 4.17: Respuestas al impulso correspondientes a la ecuación 4.37 para $n_0 = 0$ y $n_0 = 15/2$.

cuando $k = 0$, $k = 1$ y $k = 14$ respectivamente. Esta cancelación de los polos hace que el sistema tenga respuesta al impulso finita. Notar que la función del sistema 4.38 tiene la forma de la ecuación 4.36 de un sistema de muestreo en frecuencia. Comparando ambas ecuaciones se deduce que la respuesta en frecuencia cumple que

$$H(e^{j2\pi k/15}) = \begin{cases} 1, & k = 0, \\ \frac{e^{-j2\pi n_0/15}}{2}, & k = 1, \\ 0, & k = 2, 3, \dots, 13, \\ \frac{e^{j2\pi n_0/15}}{2}, & k = 14. \end{cases}$$

Para $n_0 = 0$ los factores $e^{\pm j2\pi n_0/15} = 1$ en la ecuación 4.38 y para $n_0 = 15/2$ los factores $e^{\pm j2\pi n_0/15} = e^{\pm j\pi} = -1$ en dicha ecuación. Por lo tanto, para $n_0 = 0$ y $n_0 = 15/2$ la función del sistema es respectivamente

$$H(z) = \frac{1 - z^{-15}}{15} \left(\frac{1}{1 - z^{-1}} \pm \frac{1/2}{1 - e^{j2\pi/15} z^{-1}} \pm \frac{1/2}{1 - e^{-j2\pi/15} z^{-1}} \right).$$

y la respuesta en frecuencia cumple que

$$H(e^{j2\pi k/15}) = \begin{cases} 1, & k = 0, \\ \pm \frac{1}{2}, & k = 1, \\ 0, & k = 2, 3, \dots, 13, \\ \pm \frac{1}{2}, & k = 14. \end{cases} \quad (4.39)$$

Para continuar, se calculará la respuesta en frecuencia para $n_0 = 0$ y $n_0 = 15/2$. Para hacerlo, se evalúa $H(z)$ en $z = e^{j\omega}$,

$$\begin{aligned}
H(e^{j\omega}) &= \frac{1 - e^{-j15\omega}}{15} \left(\frac{1}{1 - e^{-j\omega}} \pm \frac{1/2}{1 - e^{j2\pi/15}e^{-j\omega}} \pm \frac{1/2}{1 - e^{-j2\pi/15}e^{-j\omega}} \right) \\
&= \frac{e^{-j15\omega/2}(e^{j15\omega/2} - e^{-j15\omega/2})}{15} \left\{ \frac{1}{e^{-j\omega/2}(e^{j\omega/2} - e^{-j\omega/2})} \right. \\
&\quad \pm \frac{1/2}{e^{-j(\omega-2\pi/15)/2} [e^{j(\omega-2\pi/15)/2} - e^{-j(\omega-2\pi/15)/2}]} \\
&\quad \left. \pm \frac{1/2}{e^{-j(\omega+2\pi/15)/2} [e^{j(\omega+2\pi/15)/2} - e^{-j(\omega+2\pi/15)/2}]} \right\} \\
&= \frac{e^{-j15\omega/2} 2j \operatorname{sen}(15\omega/2)}{15} \left\{ \frac{e^{j\omega/2}}{2j \operatorname{sen}(\omega/2)} \pm \frac{e^{j(\omega-2\pi/15)/2}}{2j \operatorname{sen}[(\omega-2\pi/15)/2]} \pm \frac{e^{j(\omega+2\pi/15)/2}}{2j \operatorname{sen}[(\omega+2\pi/15)/2]} \right\},
\end{aligned}$$

y sacando de factor común $e^{j\omega/2}$ en la expresión entre corchetes y eliminando el factor $2j$ en el numerador y el denominador, resulta en

$$H(e^{j\omega}) = \frac{e^{-j7\omega} \operatorname{sen}(15\omega/2)}{15} \left\{ \frac{1}{\operatorname{sen}(\omega/2)} \pm \frac{e^{-j\pi/15}}{2 \operatorname{sen}[(\omega-2\pi/15)/2]} \pm \frac{e^{j\pi/15}}{2 \operatorname{sen}[(\omega+2\pi/15)/2]} \right\}, \quad (4.40)$$

donde el signo positivo corresponde a $n_0 = 0$ y el signo negativo a $n_0 = 15/2$. Se continuará desarrollando la expresión para $n_0 = 15/2$, correspondiente a los signos negativos en la ecuación 4.40. Sacando denominador común, el factor entre corchetes es

$$\begin{aligned}
&\frac{1}{4 \operatorname{sen}(\omega/2) \operatorname{sen}[(\omega-2\pi/15)/2] \operatorname{sen}[(\omega+2\pi/15)/2]} \left\{ 4 \operatorname{sen}[(\omega-2\pi/15)/2] \operatorname{sen}[(\omega+2\pi/15)/2] \right. \\
&\quad - 2e^{-j\pi/15} \operatorname{sen}(\omega/2) \operatorname{sen}[(\omega+2\pi/15)/2] \\
&\quad \left. - 2e^{j\pi/15} \operatorname{sen}(\omega/2) \operatorname{sen}[(\omega-2\pi/15)/2] \right\}. \quad (4.41)
\end{aligned}$$

Aplicando la identidad trigonométrica

$$\sin \theta \sin \varphi = \frac{\cos(\theta - \varphi) - \cos(\theta + \varphi)}{2},$$

el numerador queda

$$\begin{aligned}
&2[\cos(2\pi/15) - \cos \omega] - e^{-j\pi/15}[\cos(\pi/15) - \cos(\omega + \pi/15)] - e^{j\pi/15}[\cos(\pi/15) - \cos(\omega - \pi/15)] \\
&= 2 \cos(2\pi/15) - 2 \cos \omega - (e^{j\pi/15} + e^{-j\pi/15}) \cos(\pi/15) + e^{-j\pi/15} \cos(\omega + \pi/15) + e^{j\pi/15} \cos(\omega - \pi/15) \\
&\stackrel{(a)}{=} 2 \cos(2\pi/15) - 2 \cos \omega - 2 \cos(\pi/15) \cos(\pi/15) + e^{-j\pi/15}[\cos \omega \cos(\pi/15) - \operatorname{sen} \omega \operatorname{sen}(\pi/15)] \\
&\quad + e^{j\pi/15}[\cos \omega \cos(\pi/15) + \operatorname{sen} \omega \operatorname{sen}(\pi/15)] \\
&= 2 \cos(2\pi/15) - 2 \cos \omega - 2 \cos^2(\pi/15) + (e^{j\pi/15} + e^{-j\pi/15}) \cos \omega \cos(\pi/15) \\
&\quad + (e^{j\pi/15} - e^{-j\pi/15}) \operatorname{sen} \omega \operatorname{sen}(\pi/15) \\
&= 2 \cos(2\pi/15) - 2 \cos \omega - 2 \cos^2(\pi/15) + 2 \cos \omega \cos^2(\pi/15) + 2j \operatorname{sen} \omega \operatorname{sen}^2(\pi/15),
\end{aligned}$$

donde en (a) se empleó la identidad trigonométrica $\cos(\alpha \pm \beta) = \cos \alpha \cos \beta \mp \sin \alpha \sin \beta$. Aplicando la identidad trigonométrica $\cos(2\theta) = \cos^2 \theta - \sin^2 \theta$ al primer sumando de la última igualdad y eliminando

el factor 2, que se cancela con el factor 4 en el denominador de la ecuación 4.40, se obtiene que

$$\begin{aligned}
& \cos^2(\pi/15) - \sin^2(\pi/15) - \cos \omega - \cos^2(\pi/15) + \cos \omega \cos^2(\pi/15) + j \sin \omega \sin^2(\pi/15) \\
&= -\sin^2(\pi/15) - \cos \omega [1 - \cos^2(\pi/15)] + j \sin \omega \sin^2(\pi/15) \\
&= -\sin^2(\pi/15) - \cos \omega \sin^2(\pi/15) + j \sin \omega \sin^2(\pi/15) \\
&= -\sin^2(\pi/15)(1 + \cos \omega - j \sin \omega) \\
&= -\sin^2(\pi/15)(1 + e^{-j\omega}) \\
&= -\sin^2(\pi/15)e^{-j\omega/2}(e^{j\omega/2} + e^{-j\omega/2}) \\
&= -2e^{-j\omega/2} \sin^2(\pi/15) \cos(\omega/2)
\end{aligned}$$

Finalmente, combinando este resultado con las ecuaciones 4.40 y 4.41 resulta en que la respuesta en frecuencia cuando $n_0 = 15/2$ es

$$H(e^{j\omega}) = -e^{-j15\omega/2} \frac{\sin^2(\pi/15) \cos(\omega/2) \sin(15\omega/2)}{15 \sin(\omega/2) \sin[(\omega - 2\pi/15)/2] \sin[(\omega + 2\pi/15)/2]}. \quad (4.42)$$

Observar que la respuesta en frecuencia tiene la forma

$$H(e^{j\omega}) = A_e(e^{j\omega})e^{-j15\omega/2},$$

donde $A_e(e^{j\omega})$ es una función real, par y periódica en ω . Esto implica que el sistema tiene fase lineal generalizada (ver la ecuación 3.58) siempre que $n_0 = 15k/2$ donde k es un número entero impar, acorde al hecho de que la respuesta al impulso es simétrica en ese caso. En la figura 4.18 se muestra la respuesta en frecuencia para $n_0 = 0$ y $n_0 = 15/2$. Como es de esperar, se observa que la respuesta en frecuencia cumple con los valores de las muestras en frecuencia dados por la ecuación 4.39. La función del sistema

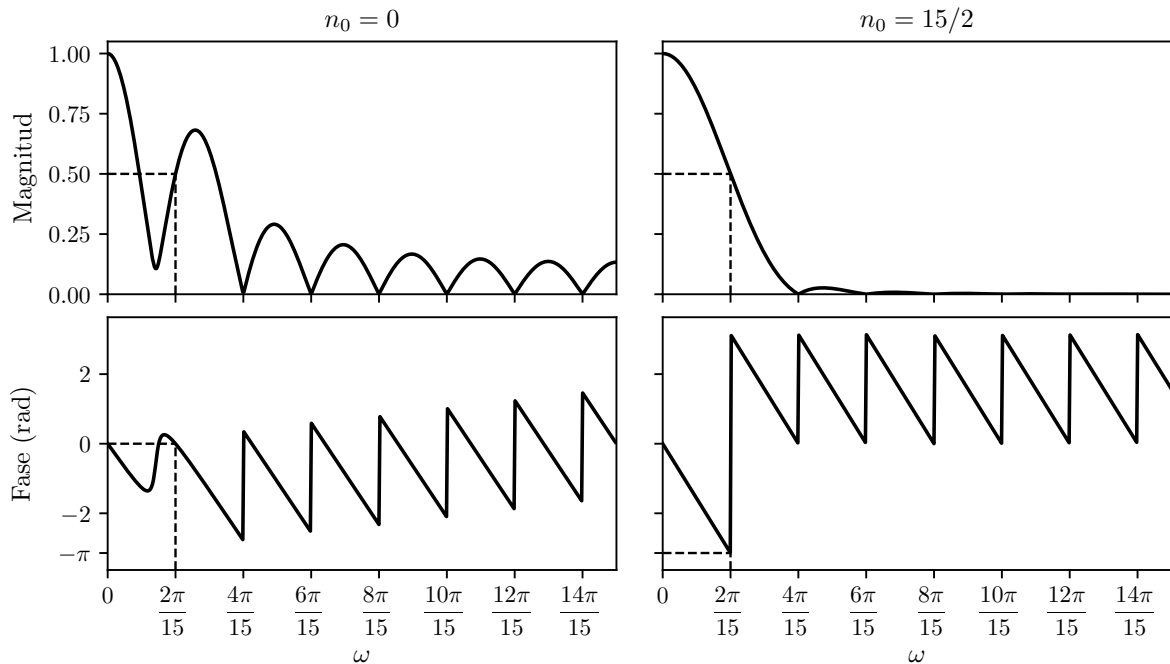


Figura 4.18: Respuesta en magnitud y respuesta en fase para $n_0 = 0$ y $n_0 = 15/2$ dadas por la ecuación 4.40. Cuando $n_0 = 15/2$ la respuesta al impulso es simétrica y el sistema es de fase lineal generalizada.

4.38 se puede reescribir como

$$\begin{aligned}
H(z) &= \frac{1 - z^{-15}}{15} \left[\frac{1}{1 - z^{-1}} + \frac{1}{2} \cdot \frac{e^{-j2\pi n_0/15}(1 - e^{-j2\pi/15}z^{-1}) + e^{j2\pi n_0/15}(1 - e^{j2\pi/15}z^{-1})}{(1 - e^{j2\pi/15}z^{-1})(1 - e^{-j2\pi/15}z^{-1})} \right] \\
&= \frac{1 - z^{-15}}{15} \left\{ \frac{1}{1 - z^{-1}} + \frac{1}{2} \cdot \frac{e^{j2\pi n_0/15} + e^{-j2\pi n_0/15} - [e^{j2\pi(n_0+1)/15} + e^{-j2\pi(n_0+1)/15}]z^{-1}}{1 - (e^{j2\pi/15} - e^{-j2\pi/15})z^{-1} + z^{-2}} \right\} \\
&= \frac{1 - z^{-15}}{15} \left\{ \frac{1}{1 - z^{-1}} + \frac{\cos(2\pi n_0/15) - \cos[2\pi(n_0+1)/15]z^{-1}}{1 - 2\cos(2\pi/15)z^{-1} + z^{-2}} \right\}.
\end{aligned}$$

La función del sistema indica que el sistema se puede implementar como la cascada de un sistema FIR, dado por el factor que multiplica al término entre corchetes, con el paralelo de un sistemas IIR de primer orden con uno de segundo orden, dados por el término entre corchetes. En la figura 4.19 se muestra el diagrama de flujos del sistema, donde $\alpha = \cos(2\pi n_0/15)$, $\beta = -\cos[2\pi(n_0+1)/15]$ y $\gamma = 2\cos(2\pi/15)$.

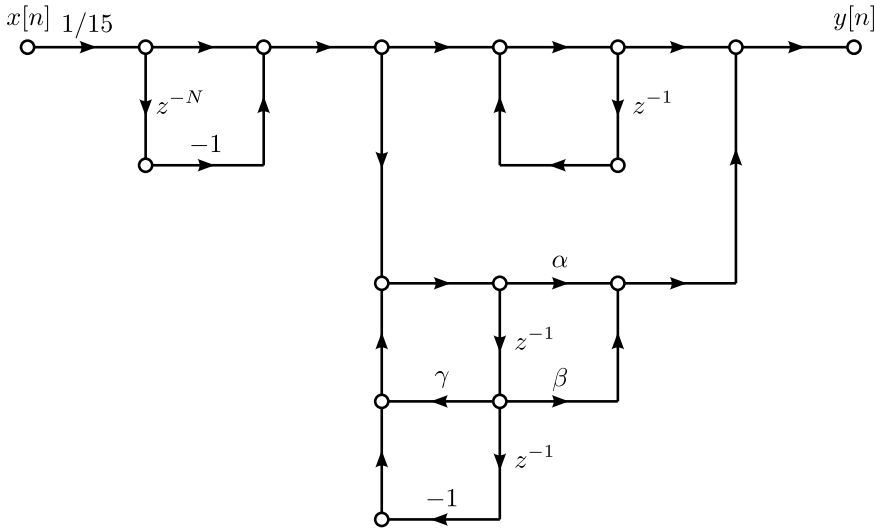


Figura 4.19: Diagrama de flujo de una estructura en forma directa II de un sistema de orden N .

4.4. Formas transpuestas

La teoría de grafos de flujo lineales de señales provee una variedad de procedimientos para transformar grafos de diferentes formas manteniendo incambiada la respuesta global entre la entrada y la salida. Uno de esos procedimientos es el de transposición, como se mostró en la sección 4.2.3.

La transposición de un grafo de flujo consiste en invertir la dirección de todas las ramas del grafo manteniendo la transmitancia e invirtiendo los roles de la entrada y la salida de forma que los nodos fuente pasan a ser nodos sumideros y viceversa. Como se mostró en la sección 4.2.3, para sistemas de un entrada y una salida el grafo de flujo resultante tiene la misma función del sistema.

Ejemplo: forma transpuesta de una sección de segundo orden Se considera el sistema en forma directa II de la figura 4.20. Las ecuaciones en diferencias correspondientes son

$$\begin{aligned}
w[n] &= a_1 w[n-1] + a_2 w[n-2] + x[n] \\
y[n] &= b_0 w[n] + b_1 w[n-1] + b_2 w[n-2].
\end{aligned}$$

En la misma figura se muestra el diagrama de flujo transpuesto. Las ecuaciones en diferencias son

$$\begin{aligned}
v_0[n] &= b_0 x[n] + v_1[n-1] \\
y[n] &= v_0[n] \\
v_1[n] &= a_1 y[n] + b_1 x[n] + v_2[n-1] \\
v_2[n] &= a_2 y[n] + b_2 x[n].
\end{aligned} \tag{4.43}$$

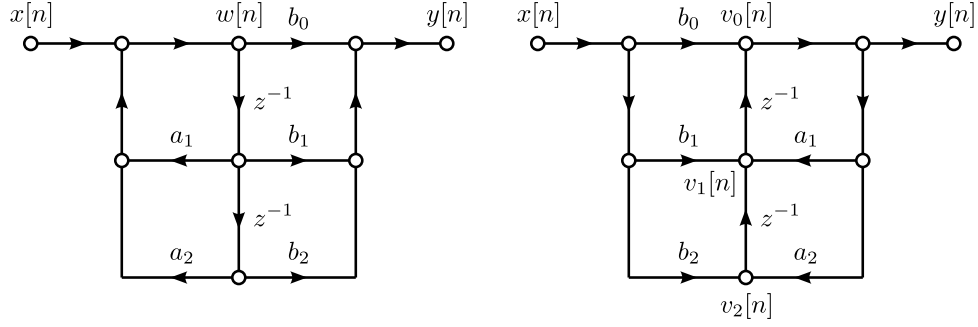


Figura 4.20: Sistema de segundo orden en forma directa II y su forma transpuesta.

Ambos grupos de ecuaciones son formas distintas de organizar el cálculo de las muestras $y[n]$ de la salida a partir de las muestras $x[n]$ de la entrada, y no es obvio que sean equivalentes. Una forma de mostrarlo es mediante la transformada z , calculando en ambos casos el cociente $H(z) = Y(z)/X(z)$ y comparando los resultados. Para mostrarlo en el dominio del tiempo puede sustituirse la segunda ecuación en 4.43 en la primera, y la cuarta ecuación en la tercera obteniendo en que

$$\begin{aligned} y[n] &= b_0 x[n] + v_1[n-1] \\ v_1[n] &= a_1 y[n] + b_1 x[n] + a_2 y[n-1] + b_2 x[n-1], \end{aligned}$$

y al sustituir la segunda ecuación en la primera de este resultado se concluye que

$$y[n] = a_1 y[n-1] + a_2 y[n-2] + b_0 x[n] + b_1 x[n-1] + b_2 x[n-2].$$

Además, como el sistema original está en forma directa II, de la sección 4.1 se sabe que cumple la misma ecuación en diferencias. Por lo tanto, en condiciones iniciales nulas, ambos sistemas son equivalentes.

El teorema de la transposición puede aplicarse a cualquiera de las estructuras discutidas previamente. Por ejemplo, el resultado de aplicar el teorema a la estructura en forma directa I de la figura 4.10 se muestra en la figura 4.21, y la estructura obtenida mediante la transposición de la estructura en forma directa II de la figura 4.11 se muestra en la figura 4.22. Si un grafo de flujo es transpuesto, el número de

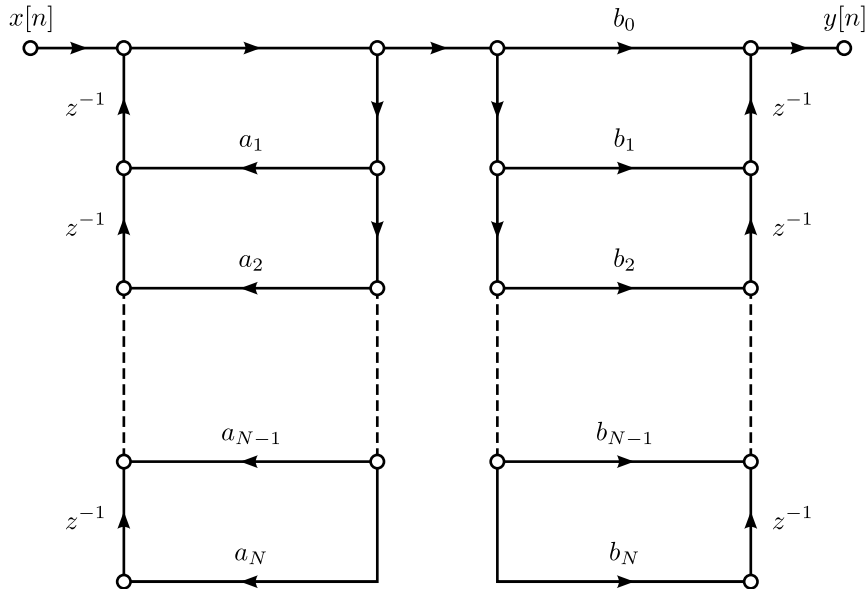


Figura 4.21: Diagrama de flujo obtenido al aplicar el teorema de la transposición a la estructura en forma directa I de la figura 4.10.

ramas de retardo y ramas de coeficientes multiplicativos permanece incambiado. Por lo tanto, estructura en forma directa II transpuesta también es canónica. Las estructuras transpuestas derivadas de formas directas son también directas, en el sentido de que pueden construirse por inspección del numerador y el denominador de la función del sistema.

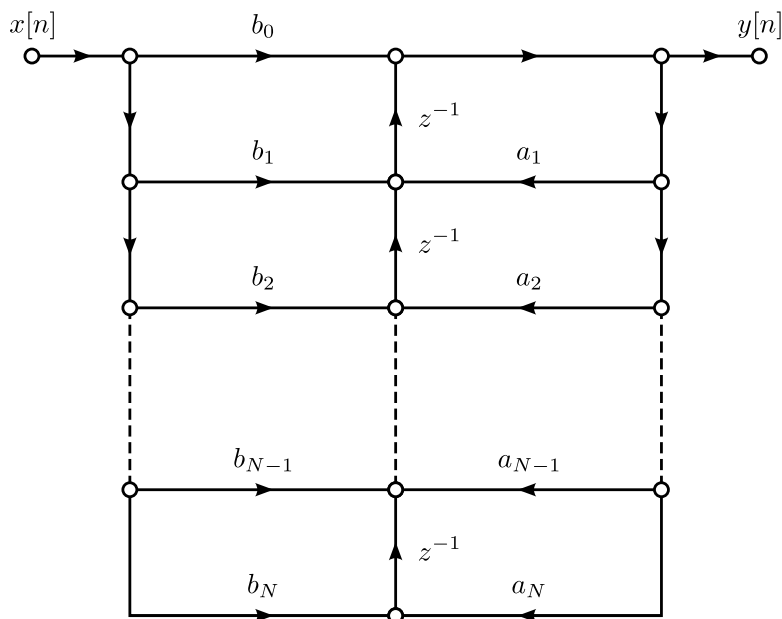


Figura 4.22: Diagrama de flujo obtenido al aplicar el teorema de la transposición a la estructura en forma directa II de la figura 4.11.

Comparando las figuras 4.11 y 4.22 se observa que mientras la estructura en forma directa II implementa primero los polos y luego los ceros, la forma directa II transpuesta implementa primero los ceros y luego los polos. Esta diferencia puede ser importante en presencia de cuantización cuando el sistema se implementa mediante precisión finita.

4.5. Estructuras básicas para sistemas FIR

Las estructuras en forma directa, en cascada y en paralelo discutidas en la sección 4.1 y 4.3 son las estructuras básicas más comunes para sistemas IIR. Aunque las estructuras en forma directa y en cascada para sistemas IIR incluyen a los sistemas FIR como caso especial, existen formas adicionales específicas para sistemas FIR.

4.5.1. Forma directa

En el caso de sistemas FIR, la función del sistema tiene únicamente ceros (excepto por polos en $z = 0$), y como los coeficientes a_k son todos nulos, la ecuación en diferencias 4.1 se reduce a

$$y[n] = \sum_{k=0}^M b_k x[n - k].$$

Esta expresión se reconoce como la convolución discreta entre $x[n]$ y la respuesta al impulso

$$h[n] = \begin{cases} b_n, & n = 0, 1, \dots, M, \\ 0, & \text{en otro caso.} \end{cases}$$

En este caso, las estructuras en forma directa I y en forma directa II en las figuras 4.10 y 4.11 se reducen ambas a la estructura directa FIR de la figura 4.23. Debido a la cadena de elementos de retardo a lo largo de la parte superior del diagrama, esta estructura es también referida como estructura de *línea de retardo* o estructura de *filtro transversal*.

La forma directa transpuesta de un sistema FIR se obtiene aplicando el teorema de la transposición al grafo de la figura 4.23, que es equivalente a establecer los coeficientes a_k en cero en las figuras 4.21 y 4.22.

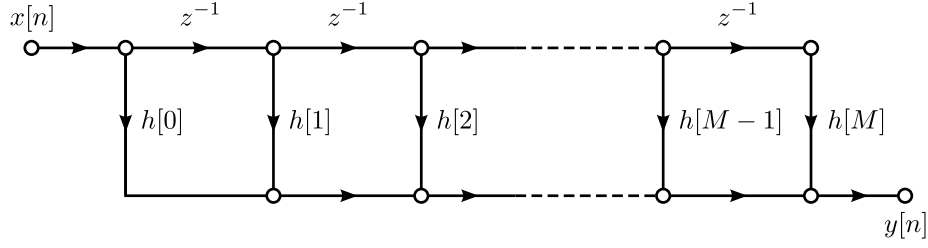


Figura 4.23: Realización en forma directa de un sistema FIR.

4.5.2. Forma en cascada

La forma en cascada para filtros FIR se obtiene factorizando la función del sistema polinómica. Concretamente, se expresa $H(z)$ como

$$H(z) = \sum_{n=0}^M h[n]z^{-n} = \prod_{k=1}^{M_s} (b_{0k} + b_{1k}z^{-1} + b_{2k}z^{-2}),$$

donde $M_s = \lfloor (M+1)/2 \rfloor$. Si M es impar, uno de los coeficientes b_{2k} es nulo, ya que en ese caso $H(z)$ tiene un número impar de ceros reales. El grafo de flujo se muestra en la figura 4.24, y es idéntico al de la figura 4.12 con los coeficientes a_{1k} y a_{2k} nulos. Cada una de las secciones de segundo orden emplea la forma directa de la figura 4.23. Una alternativa es emplear formas directas transpuestas de segundo orden, o equivalentemente, aplicar el teorema de la transposición al grafo de flujo de la figura 4.24.

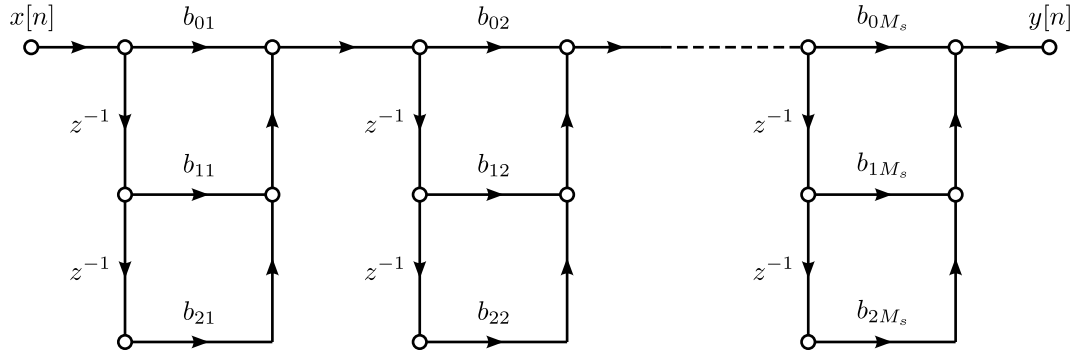


Figura 4.24: Realización en cascada de un sistema FIR.

4.5.3. Estructuras para sistemas FIR de fase lineal

En la sección 3.8.3 se mostró que un filtro FIR causal tiene fase lineal generalizada si la respuesta al impulso satisface las condiciones de simetría 3.57 o 3.59, es decir,

$$h[M-n] = h[n] \quad \text{o} \quad h[M-n] = -h[n], \quad n = 0, 1, \dots, M. \quad (4.44)$$

Con estas condiciones, el número de coeficientes puede ser esencialmente dividido a la mitad. Para ver esto, considérese la siguiente manipulación de la convolución discreta, asumiendo que M es un entero

par, lo que corresponde a sistemas FIR tipo I o tipo III (ver la figura 3.27),

$$\begin{aligned}
y[n] &= \sum_{k=0}^M h[k]x[n-k] \\
&= \sum_{k=0}^{M/2-1} h[k]x[n-k] + h[M/2]x[n-M/2] + \sum_{k=M/2+1}^M h[k]x[n-k] \\
&\stackrel{(a)}{=} \sum_{k=0}^{M/2-1} h[k]x[n-k] + h[M/2]x[n-M/2] + \sum_{l=M/2-1}^0 h[M-l]x[n-M+l] \\
&\stackrel{(b)}{=} \sum_{k=0}^{M/2-1} h[k]x[n-k] + h[M/2]x[n-M/2] + \sum_{k=0}^{M/2-1} h[M-k]x[n-M+k],
\end{aligned}$$

donde en (a) se realizó el cambio de variable $l = M - k$ en la segunda sumatoria y en (b) se invirtieron los límites y se renombró a l como k en la segunda sumatoria. Para sistemas de tipo I se emplea la condición de simetría en 4.44 para obtener que

$$y[n] = \sum_{k=0}^{M/2-1} h[k](x[n-k] + x[n-M+k]) + h[M/2]x[n-M/2]. \quad (4.45)$$

Para sistemas de tipo III se emplea la condición de antisimetría en 4.44 para obtener

$$y[n] = \sum_{k=0}^{M/2-1} h[k](x[n-k] - x[n-M+k]). \quad (4.46)$$

Observar que el producto fuera de la sumatoria se anula ya que $h[M/2] = 0$ para sistemas de tipo III (ver la figura 3.27). Realizando el mismo razonamiento cuando M es un entero impar, se obtiene que la ecuación correspondiente para sistemas de tipo II es

$$y[n] = \sum_{k=0}^{(M-1)/2} h[k](x[n-k] + x[n-M+k]). \quad (4.47)$$

y para sistemas de tipo IV es

$$y[n] = \sum_{k=0}^{(M-1)/2} h[k](x[n-k] - x[n-M+k]). \quad (4.48)$$

Las ecuaciones 4.45-4.48 implican $M/2 + 1$ (tipo I), $M/2$ (tipo III) y $(M + 1)/2$ (tipos II y IV) coeficientes multiplicativos en la estructura de forma directa general de la figura 4.23. En la figura 4.25 se muestra el grafo de flujo correspondiente a la ecuación 4.45.

En la discusión sobre los sistemas de fase lineal de la sección 3.8.3 se mostró que las condiciones de simetría 4.44 causan que los ceros ocurran en pares inversos, es decir, si z_0 es un cero de $H(z)$, $1/z_0$ también es un cero de $H(z)$. Además, si $h[n]$ es real, los ceros ocurren en pares de complejos conjugados. Como consecuencia, los ceros reales que no pertenecen a la circunferencia unidad ocurren en pares de números inversos. Los ceros complejos que no están en la circunferencia unidad ocurren en grupos de cuatro, correspondiendo a complejos conjugados e inversos. Si un cero está sobre el círculo unidad, si inverso es también su conjugado. Por lo tanto, ceros complejos sobre el círculo unidad ocurren en pares. Finalmente, como los ceros en $z = \pm 1$ son su propio conjugado e inverso, pueden ocurrir solos. Los cuatro casos se resumen en la figura 4.26, donde los ceros complejos z_1 , z_1^* , $1/z_1$ y $1/z_1^*$ que no están sobre la circunferencia unidad forman un grupo de cuatro. Los ceros reales z_2 y $1/z_2$ que no están sobre la circunferencia unidad forman un grupo de dos, así como los ceros complejos z_3 y z_3^* sobre la circunferencia unidad. El cero z_4 real sobre la circunferencia unidad puede aparecer solo. Si $H(z)$ tiene los ceros mostrados en la figura 4.26 puede factorizarse como el producto de factores de primer, segundo

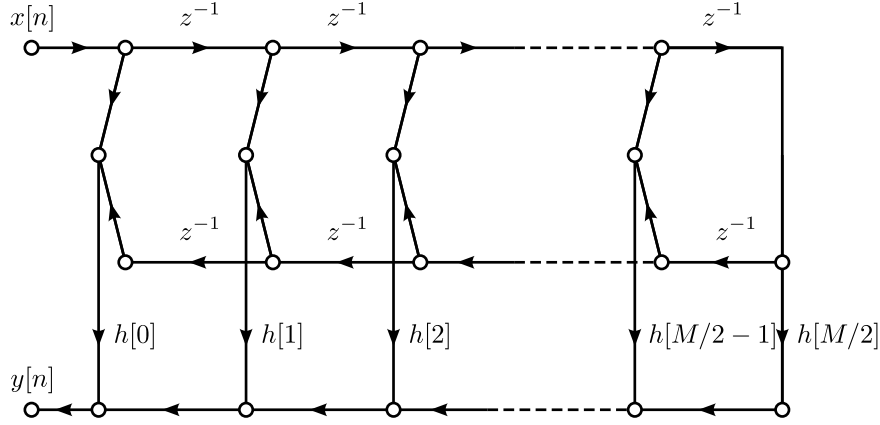


Figura 4.25: Estructura en forma directa de un sistema FIR de fase lineal tipo I.

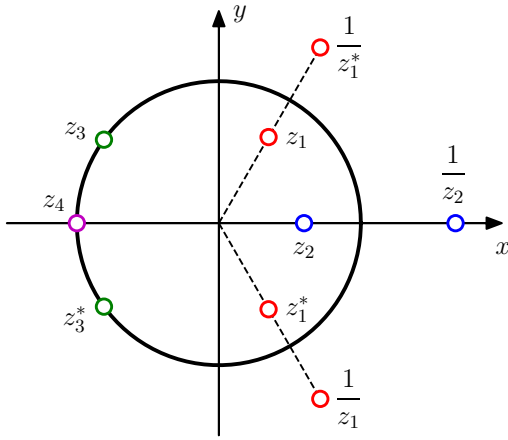


Figura 4.26: Simetría de los ceros de un sistema FIR de fase lineal.

y cuarto orden. Cada uno de estos factores tienen la misma simetría que los coeficientes de $H(z)$. En consecuencia, el sistema puede implementarse como la cascada de sistemas FIR de primer, segundo y cuarto orden. Por ejemplo, la función del sistema correspondiente al diagrama de ceros de la figura 4.26 puede expresarse como

$$H(z) = h[0](1 + z^{-1})(1 + az^{-1} + z^{-2})(1 + bz^{-1} + z^{-2})(1 + cz^{-1} + dz^{-2} + cz^{-3} + z^{-4}) \quad (4.49)$$

donde

$$a = -\left(z_2 + \frac{1}{z_2}\right), \quad b = -2 \operatorname{Re}\{z_3\}, \quad c = -2 \operatorname{Re}\left\{z_1 + \frac{1}{z_1}\right\}, \quad d = 2 + \left|z_1 + \frac{1}{z_1}\right|^2.$$

Esto es porque el segundo factor polinómico en la ecuación 4.49 corresponde a los ceros reales, y por lo tanto,

$$(1 - z_2 z^{-1})\left(1 - \frac{z^{-1}}{z_2}\right) = 1 - \left(z_2 + \frac{1}{z_2}\right)z^{-1} + z^{-2}, \quad (4.50)$$

el tercer factor polinómico en la ecuación 4.49 corresponde a los ceros complejos sobre el círculo unidad y entonces

$$(1 - z_3 z^{-1})(1 - z_3^* z^{-1}) = 1 - (z_3 + z_3^*)z^{-1} + |z_3|^2 z^{-2} = 1 - 2 \operatorname{Re}\{z_3\}z^{-1} + z^{-2},$$

donde en la última igualdad se tuvo en cuenta que $|z_3| = 1$. El cuarto factor polinómico en la ecuación 4.49 corresponde al grupo de cuatro ceros complejos, y si se agrupan de a pares inversos en factores de

segundo orden como en la ecuación 4.50, se puede escribir como

$$\begin{aligned}
& \left[1 - \left(z_1 + \frac{1}{z_1} \right) z^{-1} + z^{-2} \right] \left[1 - \left(z_1^* + \frac{1}{z_1^*} \right) z^{-1} + z^{-2} \right] \\
&= 1 - \left(z_1 + \frac{1}{z_1} + z_1^* + \frac{1}{z_1^*} \right) z^{-1} + \left[2 + \left(z_1 + \frac{1}{z_1} \right) \left(z_1^* + \frac{1}{z_1^*} \right) \right] z^{-2} \\
&\quad - \left(z_1 + \frac{1}{z_1} + z_1^* + \frac{1}{z_1^*} \right) z^{-3} + z^{-4} \\
&= 1 - 2 \operatorname{Re} \left\{ z_1 + \frac{1}{z_1} \right\} z^{-1} + \left(2 + \left| z_1 + \frac{1}{z_1} \right|^2 \right) z^{-2} - 2 \operatorname{Re} \left\{ z_1 + \frac{1}{z_1} \right\} z^{-3} + z^{-4}.
\end{aligned}$$

4.6. Filtros lattice

Una estructura en cascada útil e interesante es basada en la conexión en cascada de la estructura básica mostrada en la figura 4.27. En la figura, el bloque básico que compone la cascada tiene dos entradas y dos salidas, y su implementación para el caso de una estructura lattice FIR se muestra en la misma figura. En la figura 4.28 se muestra la conexión en cascada de M de esos bloques básicos con la entrada

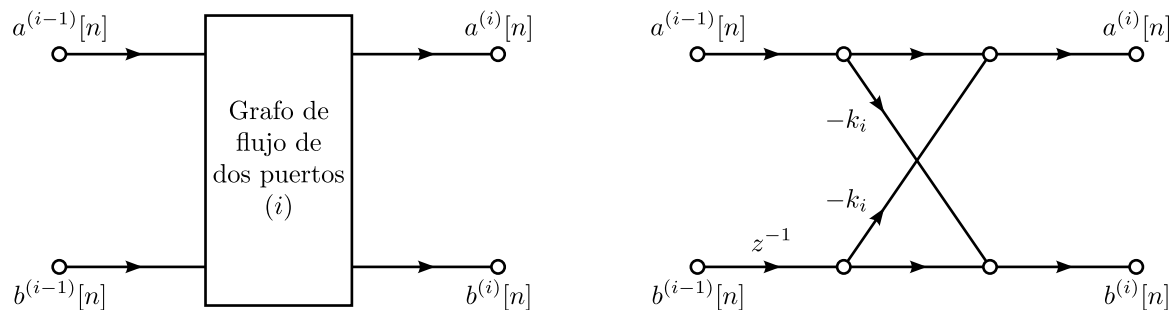


Figura 4.27: Bloque constitutivo de una estructura lattice FIR. Se muestra una representación en diagrama de bloques genérico de dos entradas y dos salidas y su implementación como grafo de flujo.

$x[n]$ alimentando ambas entradas del bloque 1 y con salida $y[n]$ la primer salida del bloque M , de forma que el sistema global tiene una entrada y una salida. La segunda salida del bloque M se ignora. Si bien la estructura en cascada puede tomar muchas formas distintas dependiendo de la definición del bloque constitutivo, se limitará la atención a la implementación en la figura 4.27, la cual conduce a una clase de filtros FIR e IIR ampliamente usada.

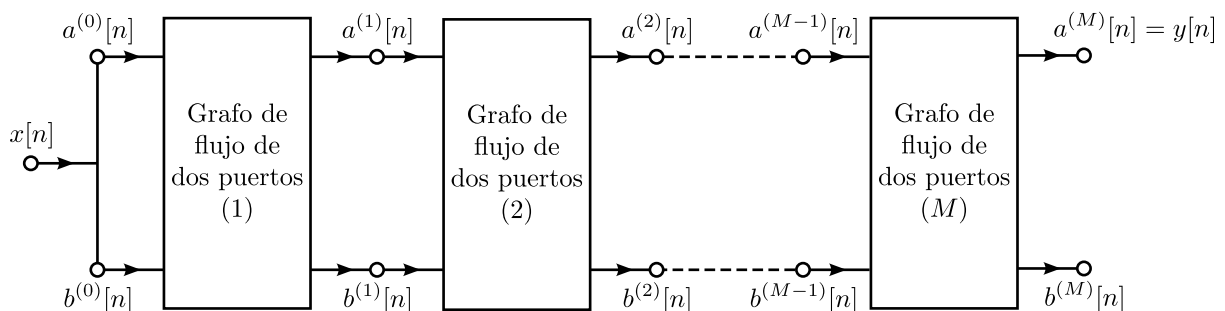


Figura 4.28: Conexión en cascada de M bloques constitutivos.

4.6.1. Filtros lattice FIR

Si la estructura en forma de mariposa del grafo de flujo la figura 4.27 es empleada como bloque básico en la cascada de la figura 4.28 se obtiene el grafo de flujo mostrado en la figura 4.29, cuya forma motiva el nombre de *filtro lattice* (*filtro en celosía*). Los coeficientes k_1, k_2, \dots, k_M son referidos como parámetros k de la estructura. En el capítulo ?? se verá que los parámetros k tienen un significado especial en el contexto de modelado todo polos de señales, y que el filtro lattice de la figura 4.29 es una estructura de implementación para predicción lineal de señales. En este capítulo solo se ilustra el empleo de filtros lattice para implementar funciones de transferencia FIR e IIR todo polos.

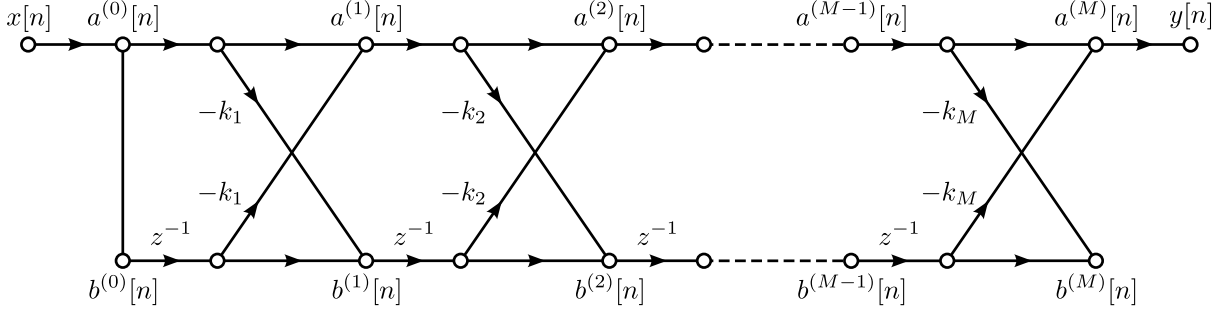


Figura 4.29: Grafo de flujo de un filtro lattice FIR basado en la conexión en cascada de M bloques básicos de dos puertos como el de la figura 4.27.

Las variables de nodo $a^{(i)}[n]$ y $b^{(i)}[n]$ en la figura 4.29 son secuencias intermedias que dependen de la entrada mediante el conjunto de ecuaciones en diferencias

$$\begin{aligned} a^{(0)}[n] &= b^{(0)}[n] = x[n] \\ a^{(i)}[n] &= a^{(i-1)}[n] - k_i b^{(i-1)}[n-1] & i = 1, 2, \dots, M \\ b^{(i)}[n] &= b^{(i-1)}[n-1] - k_i a^{(i-1)}[n] & i = 1, 2, \dots, M \\ y[n] &= a^{(M)}[n]. \end{aligned} \quad (4.51)$$

Observar que las ecuaciones deben calcularse en el orden $i = 0, 1, \dots, M$ ya que la salida de la etapa $(i-1)$ se necesita como entrada de la etapa (i) .

La estructura lattice de la figura 4.29 es claramente un sistema LTI por tratarse de un grafo de flujo lineal únicamente con ramas con retardos y constantes. Además, no hay bucles de realimentación, lo que implica que el sistema tiene respuesta al impulso de duración finita. Concretamente, considérese la respuesta al impulso desde $x[n]$ a la variable de nodo $a^{(i)}[n]$. Con $x[n] = \delta[n]$ es claro que $a^{(i)}[0] = 1$ para todo i , ya que el impulso se propaga sin retardo por la rama superior de todas las etapas. Todos los otros caminos hacia cualquier variable de nodo $a^{(i)}[n]$ o $b^{(i)}[n]$ pasan por al menos una rama de retardo, siendo el camino de mayor retardo el camino por la rama inferior de todas las etapas y luego subiendo a la variable de nodo $a^{(i)}[n]$ a través de la rama con coeficiente $-k_i$. Este camino produce un retardo total de i muestras, por lo que la respuesta al impulso tiene un largo de $i+1$ muestras.

En lo que sigue, se asume que $a^{(i)}[n]$ y $b^{(i)}[n]$ son las respuestas al impulso a la salida de la etapa i en lugar de las variables de nodo correspondientes a cualquier entrada $x[n]$, y las respectivas transformadas $z A^{(i)}(z)$ y $z B^{(i)}(z)$ son la función de transferencia entre la entrada y los nodos i -ésimos. En consecuencia, la función de transferencia entre la entrada y el nodo i -ésimo superior es

$$A^{(i)}(z) = \sum_{n=0}^i a^{(i)}[n]z^{-n} = 1 - \sum_{m=1}^i \alpha_m^{(i)} z^{-m}, \quad (4.52)$$

donde en el lado derecho de la igualdad, los coeficientes $\alpha_m^{(i)}$ para $m \leq i$ se componen de sumas de productos de los coeficientes k_j con $j \leq m$. En particular, como ya se mencionó, el coeficiente para el mayor retardo entre la entrada y el nodo i -ésimo superior es $\alpha_i^{(i)} = k_i$. Con esta notación, la respuesta

al impulso desde $x[n]$ a la variable de nodo $a^{(i)}[n]$ es

$$a^{(i)}[n] = \begin{cases} 1 & n = 0 \\ -\alpha_n^{(i)} & 1 \leq n \leq i \\ 0 & \text{en otro caso.} \end{cases}$$

Del diagrama de flujo de la figura 4.27 o aplicando la transformada z a las ecuaciones en diferencias 4.51, se obtiene que

$$A^{(i)}(z) = A^{(i-1)}(z) - k_i z^{-1} B^{(i-1)}(z) \quad (4.53)$$

$$B^{(i)}(z) = -k_i A^{(i-1)}(z) + z^{-1} B^{(i-1)}(z), \quad (4.54)$$

donde $B^{(i)}(z)$ es la función de transferencia entre la entrada y el nodo i -ésimo inferior. Además, en la entrada

$$A^{(0)}(z) = B^{(0)}(z) = 1. \quad (4.55)$$

Empleando las ecuaciones 4.53 y 4.54 y comenzando con la ecuación 4.55, se puede calcular $A^{(i)}(z)$ y $B^{(i)}(z)$ recursivamente hasta cualquier valor de i . Se mostrará que al hacerlo, el patrón de la relación entre $A^{(i)}(z)$ y $B^{(i)}(z)$ que surge es

$$B^{(i)}(z) = z^{-i} A^{(i)}(1/z), \quad (4.56)$$

o equivalentemente, reemplazando z por $1/z$,

$$A^{(i)}(z) = z^{-i} B^{(i)}(1/z). \quad (4.57)$$

La prueba puede hacerse mediante inducción, es decir, verificando que si la expresión es cierta para algún valor $i - 1$, se cumple para el valor i . Se parte observando que para $i = 0$, la validez de las identidades 4.56 y 4.57 se obtiene inmediatamente de la ecuación 4.55. Además, para $i = 1$

$$\begin{aligned} A^{(1)}(z) &= A^{(0)}(z) - k_1 z^{-1} B^{(0)}(z) \\ &= 1 - k_1 z^{-1} \\ B^{(1)}(z) &= -k_1 A^{(0)}(z) + z^{-1} B^{(0)}(z) \\ &= -k_1 + z^{-1} \\ &= z^{-1}(1 - k_1 z) \\ &= z^{-1} A^{(1)}(1/z), \end{aligned}$$

y para $i = 2$,

$$\begin{aligned} A^{(2)}(z) &= A^{(1)}(z) - k_2 z^{-1} B^{(1)}(z) \\ &= 1 - k_1 z^{-1} - k_2 z^{-1}(-k_1 + z^{-1}) \\ &= 1 - k_1(1 - k_2)z^{-1} - k_2 z^{-2} \\ B^{(2)}(z) &= -k_2 A^{(1)}(z) + z^{-1} B^{(1)}(z) \\ &= -k_2(1 - k_1 z^{-1}) + z^{-1}(-k_1 + z^{-1}) \\ &= -k_2 - k_1(1 - k_2)z^{-1} + z^{-2} \\ &= z^{-2}[1 - k_1(1 - k_2)z - k_2 z^2] \\ &= z^{-2} A^{(2)}(1/z). \end{aligned}$$

Para probar el resultado general, se asume que las ecuaciones 4.56 y 4.57 se cumplen para $i - 1$ y se sustituyendo esos resultados en la ecuación 4.54, se obtiene que

$$\begin{aligned} B^{(i)}(z) &= -k_i z^{-(i-1)} B^{(i-1)}(1/z) + z^{-1} z^{-(i-1)} A^{(i-1)}(1/z) \\ &= z^{-i} A^{(i-1)}(1/z) - k_i z^{-(i-1)} B^{(i-1)}(1/z) \\ &= z^{-i} [A^{(i-1)}(1/z) - k_i z^{-1} B^{(i-1)}(1/z)] \\ &= z^{-i} A^{(i)}(1/z), \end{aligned}$$

que es lo que se quería probar.

Las funciones de transferencia $A^{(i)}(z)$ y $B^{(i)}(z)$ son polinomios de orden i , y es particularmente útil obtener una relación directa entre los coeficientes de los polinomios. Con el fin de obtener una expresión de los coeficientes $\alpha_m^{(i)}$ en términos de $\alpha_m^{(i-1)}$ y k_i se parte de la ecuación 4.52 de la función de transferencia para el nodo $i-1$,

$$A^{(i-1)}(z) = 1 - \sum_{m=1}^{i-1} \alpha_m^{(i-1)} z^{-m} \quad (4.58)$$

y se combina con la ecuación 4.56 para el mismo nodo, resultando en

$$B^{(i-1)}(z) = z^{-(i-1)} A^{(i-1)}(1/z) = z^{-(i-1)} \left[1 - \sum_{m=1}^{i-1} \alpha_m^{(i-1)} z^m \right].$$

Luego, sustituyendo este resultado y la ecuación 4.58 en la ecuación 4.53, se observa que

$$\begin{aligned} A^{(i)}(z) &= \left[1 - \sum_{m=1}^{i-1} \alpha_m^{(i-1)} z^{-m} \right] - k_i z^{-1} \left(z^{-(i-1)} \left[1 - \sum_{m=1}^{i-1} \alpha_m^{(i-1)} z^m \right] \right) \\ &\stackrel{(a)}{=} \left[1 - \sum_{m=1}^{i-1} \alpha_m^{(i-1)} z^{-m} \right] - k_i z^{-i} \left[1 - \sum_{l=i-1}^1 \alpha_{i-l}^{(i-1)} z^{i-l} \right] \\ &= 1 - \sum_{m=1}^{i-1} \alpha_m^{(i-1)} z^{-m} + \sum_{l=1}^{i-1} k_i \alpha_{i-l}^{(i-1)} z^{-l} - k_i z^{-i} \end{aligned}$$

donde en (a) se realizó el cambio de variable $l = i - m$ en la segunda sumatoria. Finalmente, combinando ambas sumatorias resulta en

$$A^{(i)}(z) = 1 - \sum_{m=1}^{i-1} \left[\alpha_m^{(i-1)} - k_i \alpha_{i-m}^{(i-1)} \right] z^{-m} - k_i z^{-i}.$$

Comparando este resultado con la ecuación 4.52 se concluye que

$$\begin{aligned} \alpha_m^{(i)} &= \alpha_m^{(i-1)} - k_i \alpha_{i-m}^{(i-1)}, & m = 1, 2, \dots, i-1, \\ \alpha_i^{(i)} &= k_i. \end{aligned} \quad (4.59)$$

Estas ecuaciones son la recursión buscada entre los coeficientes de $A^{(i)}(z)$ y $A^{(i-1)}(z)$. Observar que como se mencionó previamente, el coeficiente correspondiente a z^{-i} , que representa al mayor retardo, es $-k_i$. Estas ecuaciones junto con la ecuación 4.56 también determinan la función de transferencia $B^{(i)}(z)$, ya que sustituyendo la ecuación 4.52 en la ecuación 4.56 se observa que

$$\begin{aligned} B^{(i)}(z) &= z^{-i} \left[1 - \sum_{m=1}^i \alpha_m^{(i)} z^m \right] \\ &= z^{-i} - \sum_{m=1}^i \alpha_m^{(i)} z^{m-i} \\ &\stackrel{(a)}{=} z^{-i} - \sum_{l=i-1}^0 \alpha_{i-l}^{(i)} z^{-l} \\ &= - \sum_{m=0}^{i-1} \alpha_{i-m}^{(i)} z^{-m} + z^{-i}. \end{aligned}$$

Por lo tanto, si

$$B^{(i)}(z) = \sum_{n=0}^i b^{(i)}[n] z^{-n} = \sum_{m=0}^i \beta_m^{(i)} z^{-m},$$

los coeficientes $\beta_m^{(i)}$ del polinomio de la función de transferencia $B^{(i)}(z)$ son

$$\begin{aligned}\beta_m^{(i)} &= \alpha_{i-m}^{(i)}, & m = 0, 1, \dots, i-1, \\ \beta_i^{(i)} &= 1.\end{aligned}$$

Observar en la figura 4.29 que el resultado es acorde al hecho de que el único camino sin retardo entre la entrada y la variable de nodo $b^{(i)}[n]$ es el camino por las ramas superiores y luego descendiendo al nodo por la rama de constante $-k_i$, que corresponde en la ecuación al coeficiente $\beta_0^{(i)} = \alpha_i^{(i)} = -k_i$, y también al hecho de que el único camino con retardo de i muestras, correspondiente al mayor retardo, es el que va por las ramas inferiores y no pasa por ninguna rama con constante multiplicativa, por lo que el coeficiente es $\beta_i^{(i)} = 1$.

La recursión de la ecuación 4.59 puede expresarse de forma matricial. Denotando como α_{i-1} al vector de los coeficientes de la función de transferencia $A^{(i-1)}(z)$ y como $\check{\alpha}_{i-1}$ al mismo vector con los elementos en orden invertido, es decir,

$$\alpha_{i-1} = \begin{bmatrix} \alpha_1^{(i-1)} & \alpha_2^{(i-1)} & \dots & \alpha_{i-1}^{(i-1)} \end{bmatrix}^T \quad \text{y} \quad \check{\alpha}_{i-1} = \begin{bmatrix} \alpha_{i-1}^{(i-1)} & \alpha_{i-2}^{(i-1)} & \dots & \alpha_1^{(i-1)} \end{bmatrix}^T,$$

la ecuación 4.59 se puede escribir en forma matricial como

$$\alpha_i = \begin{bmatrix} \alpha_{i-1} \\ \dots \\ 0 \end{bmatrix}^T - k_i \begin{bmatrix} \check{\alpha}_i \\ \dots \\ -1 \end{bmatrix}^T. \quad (4.60)$$

La recursión de las ecuaciones 4.59 o 4.60 son la base de un algoritmo para el análisis de una estructura lattice FIR para obtener la función de transferencia. Se parte del grafo de flujo de la figura 4.29 donde se especifican los parámetros $k \{k_1, k_2, \dots, k_M\}$. Luego se emplean las ecuaciones 4.59 recursivamente para calcular la función de transferencia del filtro FIR de orden sucesivamente mas alto hasta llegar al final de la cascada, obteniendo

$$A(z) = 1 - \sum_{m=1}^M \alpha_m z^{-m} = \frac{Y(z)}{X(z)}, \quad (4.61)$$

donde

$$\alpha_m = \alpha_m^{(M)}, \quad m = 1, 2, \dots, M.$$

En la figura 4.30 se muestra la implementación en el lenguaje Python de una función para obtener los coeficientes de la función de transferencia a partir de los parámetros k .

También es de interés obtener los parámetros k de una estructura lattice FIR que implementa una función de transferencia dada, es decir, se quiere obtener el conjunto de parámetros k de la estructura lattice en la figura 4.29 a partir del polinomio $A(z)$ de la ecuación 4.61. Esto puede hacerse invirtiendo la recursión de las ecuaciones 4.59 o 4.60 para obtener sucesivamente la función de transferencia $A^{(i-1)}(z)$ en términos de $A^{(i)}(z)$ para $i = M, M-1, \dots, 2$. Los parámetros k surgen como consecuencia del proceso, ya que $k_i = \alpha_i^{(i)}$.

Concretamente, se asume que se especifican los coeficientes $\alpha_m^{(M)} = \alpha_m$ para $m = 1, 2, \dots, M$ y se quiere obtener los parámetros $k \{k_1, k_2, \dots, k_M\}$ para implementar la función de transferencia como una estructura lattice. Para hacerlo, se parte de la última etapa de la estructura lattice FIR, es decir, con $i = M$. De la ecuación 4.59,

$$k_M = \alpha_M^{(M)} = \alpha_M.$$

Se continúa invirtiendo la ecuación 4.59 o la ecuación 4.60 para obtener α_{M-1} , el vector de coeficientes de la función de transferencia desde la entrada a la etapa $i = M-1$, y así sucesivamente hasta llegar a $i = 1$.

Para obtener la fórmula general de recursión de $\alpha_m^{(i-1)}$ en términos de $\alpha_m^{(i)}$ de la ecuación 4.59, se observa que el término $\alpha_{i-m}^{(i-1)}$ debe ser eliminado. Para hacerlo, se reemplaza m por $i-m$ en la ecuación 4.59 y se multiplican ambos lados de la igualdad por k_i para obtener

$$k_i \alpha_{i-m}^{(i)} = k_i \alpha_{i-m}^{(i-1)} - k_i^2 \alpha_m^{(i-1)},$$

```

import numpy as np

def k_to_alpha(k):
    """Cálculo de los coeficientes de A(z) a partir de los parámetros k.

    Parámetros
    -----
    k : vector numpy de dimension (M, )
        vector con los parámetros k

    Retorna
    -----
    vector numpy de dimensión (M, )
        vector con los coeficientes
    """
    M = len(k)
    for i in np.arange(M):
        # inicialización del vector \alpha^{(i)}
        alpha_i = np.zeros((i + 1, ))
        for m in np.arange(i):
            # cálculo de los elementos del vector \alpha^{(i)}[m]
            # para m = 0, 1, ..., i - 1
            alpha_i[m] = alpha_iprev[m] - k[i] * alpha_iprev[i - m - 1]
        # cálculo del último elemento del vector: \alpha^{(i)}[i]
        alpha_i[i] = k[i]
        # se almacena \alpha^{(i)}, necesario para calcular \alpha^{(i+1)}
        # en la siguiente iteración.
        alpha_iprev = alpha_i
    return alpha_i

```

Figura 4.30: Implementación en el lenguaje Python de la conversión de los parámetros k a los coeficientes de la función de transferencia. Tener en cuenta que en el lenguaje Python, la indexación de los vectores comienza en 0, y por lo tanto, es levemente diferente a la indexación empleada en la explicación de la conversión.

y sumando este resultado con la ecuación 4.59 resulta en

$$\alpha_m^{(i)} + k_i \alpha_{i-m}^{(i)} = \alpha_m^{(i-1)} - k_i^2 \alpha_m^{(i-1)},$$

de donde surge que

$$\alpha_m^{(i-1)} = \frac{\alpha_m^{(i)} + k_i \alpha_{i-m}^{(i)}}{1 - k_i^2}, \quad m = 1, 2, \dots, i-1, \quad (4.62)$$

que es la expresión que se necesita para hacer la recursión. Con $\alpha_m^{(i-1)}$ calculado para $m = 1, 2, \dots, i-1$, la ecuación 4.59 indica que

$$k_{i-1} = \alpha_{i-1}^{(i-1)}. \quad (4.63)$$

Por lo tanto, comenzando con los coeficientes conocidos $\alpha_m^M = \alpha_m$ para $m = 1, 2, \dots, M$, se emplean las ecuaciones 4.62 para calcular α_m^{M-1} para $m = 1, 2, \dots, M-1$ y de la ecuación 4.63, $k_{M-1} = \alpha_{M-1}^{(M-1)}$

Parámetros k para un sistema FIR de tercer orden Se considera el sistema FIR mostrado en la figura 4.32, cuya función de transferencia es

$$A(z) = 1 - 0.9z^{-1} + 0.64z^{-2} - 0.576z^{-3}.$$

En consecuencia, $M = 3$ y los coeficientes $\alpha_k^{(3)}$ de la ecuación 4.61 son

$$\alpha_1^{(3)} = 0.9, \quad \alpha_2^{(3)} = -0.64, \quad \alpha_3^{(3)} = 0.576.$$

Se quieren calcular los parámetros k de la estructura lattice equivalente. Se comienza observando que $k_3 = \alpha_3^{(3)} = 0.576$. Luego hay que calcular los coeficientes de la función de transferencia $A^{(2)}(z)$, para lo que se emplea la ecuación 4.62, resultando en

$$\alpha_1^{(2)} = \frac{\alpha_1^{(3)} + k_3 \alpha_2^{(3)}}{1 - k_3^2} = 0.795, \quad \alpha_2^{(2)} = \frac{\alpha_2^{(3)} + k_3 \alpha_1^{(3)}}{1 - k_3^2} = -0.182,$$

```

import numpy as np

def alpha_to_k(alpha):
    """Cálculo de los parámetros k a partir de los coeficientes de A(z).

    Parámetros
    alpha : vector numpy de dimension (M, )
            vector con los coeficientes de A(z)

    Retorna
    numpy vector de dimensión (M, )
            vector con los parámetros k
    """
    M = len(alpha)
    # inicialización del vector con los parámetros k
    k = np.zeros((M, ))
    k[M - 1] = alpha[M - 1]
    alpha_iprev = alpha
    for i in np.arange(M - 1, 0, -1):
        # se almacena \alpha^{(i)}, necesario para calcular \alpha^{(i-1)}
        # en la siguiente iteración.
        alpha_i = alpha_iprev
        # inicialización del vector \alpha^{(i-1)}
        alpha_iprev = np.zeros((i, ))
        for m in np.arange(i):
            # cálculo de los elementos del vector \alpha^{(i-1)}[m]
            # para m = 0, 1, ..., i - 1
            alpha_iprev[m] = (alpha_i[m] + k[i] * alpha_i[i - m - 1]) / (1 - k[i] ** 2)
        # se asigna el valor de k, que es el último valor del vector \alpha^{(i-1)}
        k[i - 1] = alpha_iprev[i - 1]
    return k

```

Figura 4.31: Implementación en el lenguaje Python del cálculo de los parámetros k a partir de los coeficientes de la función de transferencia.

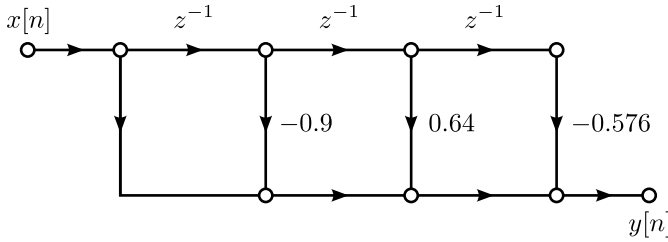


Figura 4.32: Estructura FIR en forma directa.

y de la ecuación 4.63, $k_2 = \alpha_2^{(2)} = -0.182$. Finalmente, se aplica nuevamente la ecuación 4.62 para obtener $A^{(1)}(z)$,

$$\alpha_1^{(1)} = \frac{\alpha_1^{(2)} + k_2 \alpha_1^{(2)}}{1 - k_2^2} = 0.673,$$

donde se identifica que $k_1 = \alpha_1^{(1)} = 0.673$. La estructura lattice equivalente se muestra en la figura 4.33. También puede realizarse el proceso inverso empleando las ecuaciones 4.59 para obtener la función de transferencia $A(z)$ a partir de los parámetros k de la estructura lattice. De la segunda ecuación en 4.59, $\alpha_1^{(1)} = k_1 = 0.673$, quedando determinada $A^{(1)}(z)$. Luego, de las ecuaciones 4.59,

$$\alpha_1^{(2)} = \alpha_1^{(1)} - k_2 \alpha_1^{(1)} = 0.795, \quad \alpha_2^{(2)} = k_2 = -0.182,$$

determinado $A^{(2)}(z)$. Finalmente, se aplican las ecuaciones 4.59 nuevamente para obtener

$$\alpha_1^{(3)} = \alpha_1^{(2)} - k_3 \alpha_2^{(2)} = 0.9, \quad \alpha_2^{(3)} = \alpha_2^{(2)} - k_3 \alpha_1^{(2)} = -0.64, \quad \alpha_3^{(3)} = k_3 = 0.576,$$

que son los coeficientes de la función de transferencia en la ecuación 4.61.

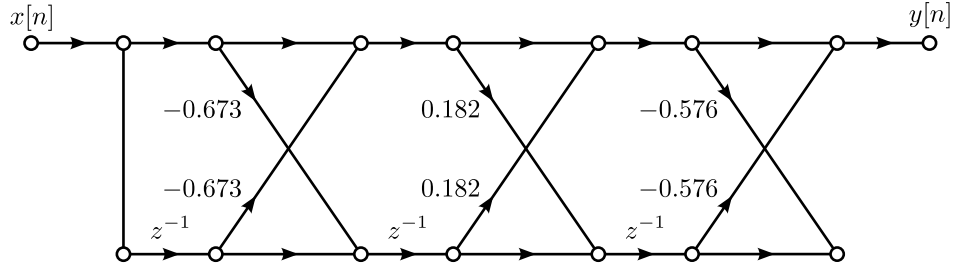


Figura 4.33: Estructura lattice con la misma función de transferencia que la estructura FIR de la figura 4.32.

4.6.2. Estructura lattice todo polos

Una estructura lattice para implementar la función de transferencia todo polos $H(z) = 1/A(z)$ puede obtenerse de la estructura lattice FIR de la sección previa observando que $H(z)$ es el filtro inverso de la función de transferencia $A(z)$ del sistema FIR. Para hacerlo, se asume que la salida $y[n] = a^M[n]$ es dada y se quiere calcular la entrada $x[n] = a^{(0)}[n]$ que produce dicha salida. Esto se logra realizando los cálculos desde la derecha hacia la izquierda en la estructura de la figura 4.29. Específicamente, expresando $A^{(i-1)}(z)$ en términos $A^{(i)}(z)$ en la ecuación 4.53 y dejando a la ecuación 4.54 incambiada, se obtiene el par de ecuaciones

$$A^{(i-1)}(z) = A^{(i)}(z) + k_i z^{-1} B^{(i-1)}(z) \quad (4.64)$$

$$B^{(i)}(z) = -k_i A^{(i-1)}(z) + z^{-1} B^{(i-1)}(z), \quad (4.65)$$

que tienen la representación en grafo de flujo de la figura 4.34. Notar que ahora el flujo de la señal es

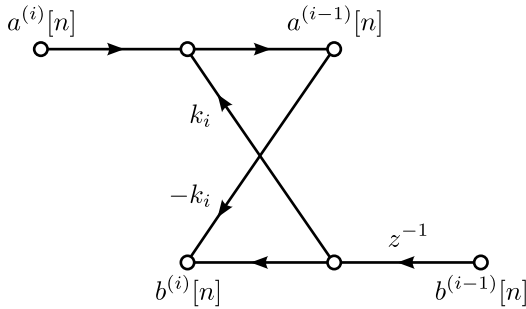


Figura 4.34: Diagrama de flujo de una etapa de un sistema lattice todo polos.

desde i a $i-1$ por las ramas de la parte superior del diagrama y desde $i-1$ a i por las ramas de la parte inferior. La conexión en cascada de M etapas como la de la figura 4.34 con los valores de k_i apropiados en cada etapa toma una entrada $a^{(M)}[n]$ y produce la salida $a^{(0)}[n]$, como se muestra en la figura 4.35. Finalmente, la condición $x[n] = a^{(0)}[n] = b^{(0)}[n]$ en las terminales de la última etapa produce la realimentación que propaga a $b^{(i)}[n]$ en dirección inversa. Esta realimentación es necesaria para un sistema IIR.

El conjunto de ecuaciones en diferencias representado por la figura 4.35 es²

$$\begin{aligned} a^{(M)}[n] &= y[n] \\ a^{(i-1)}[n] &= a^{(i)}[n] + k_i b^{(i-1)}[n-1] \\ b^{(i)}[n] &= b^{(i-1)}[n-1] + k_i a^{(i-1)}[n] \\ x[n] &= a^{(0)}[n] = b^{(0)}[n]. \end{aligned} \quad (4.66)$$

²Observar que debido a la deducción basada en el sistema lattice FIR, se nombró $y[n]$ a la entrada y $x[n]$ a la salida, contrariamente a la convención usual. Esta nomenclatura es arbitraria y puede restituirse una vez concluida la deducción.

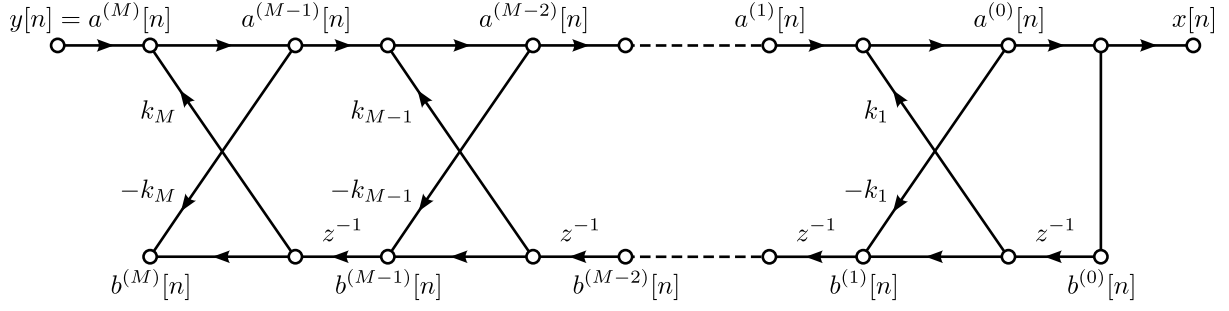


Figura 4.35: Grafo de flujo de un sistema lattice todo polos.

Debido a la realimentación en la figura 4.35 y las ecuaciones que representan al sistema, deben especificarse condiciones iniciales para todas las variables de nodo asociadas a retardos. Típicamente se establece $b^{(i)}[-1] = 0$ para condiciones iniciales de reposo.

La deducción en la presente sección permite afirmar que todo el análisis realizado en la sección 4.6.1 para el sistema lattice FIR es válido para el sistema lattice todo polos de la figura 4.35. Si se desea determinar la implementación lattice de un sistema todo polos con función de transferencia $H(z) = 1/A(z)$ se puede emplear los algoritmos de las figuras 4.31 y 4.30 para obtener los parámetros k a partir de los coeficientes del polinomio del denominador y viceversa.

Ejemplo: implementación lattice de un sistema IIR Como ejemplo, considérese el sistema IIR dado por la función de transferencia

$$H(z) = \frac{1}{1 - 0.9z^{-1} + 0.64z^{-2} - 0.576z^{-3}}$$

$$= \frac{1}{(1 - 0.8z^{-1})(1 + 0.8z^{-1})(1 - 0.9z^{-1})},$$

que es el sistema inverso del sistema del ejemplo de la sección 4.6.1. La figura 4.36 muestra la implementación en forma directa del sistema, mientras que la figura 4.37 muestra el sistema lattice IIR equivalente con los parámetros k calculados como en el ejemplo de la sección 4.6.1. Observar que la estructura lattice tiene la misma cantidad de retardos que la forma directa pero el doble de multiplicadores. Esto se cumple para cualquier orden M , como es evidente de la observación de las estructuras.

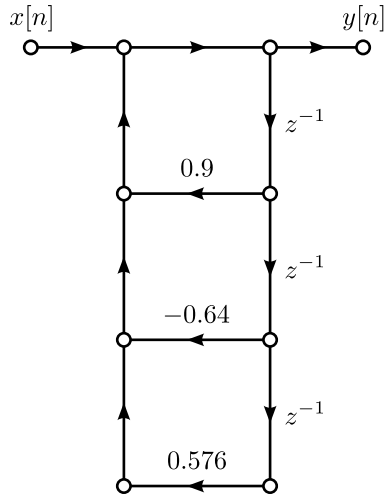


Figura 4.36: Diagrama de flujo del sistema IIR en forma directa.

Dado que la estructura lattice de la figura 4.35 es un sistema IIR, es necesario considerar su estabilidad. En el capítulo ?? se verá que una condición necesaria y suficiente para que todos los ceros del polinomio $A(z)$ estén dentro de círculo unidad es que $|k_i| < 1$ para $i = 1, 2, \dots, M$. Esto se ve en el ejemplo

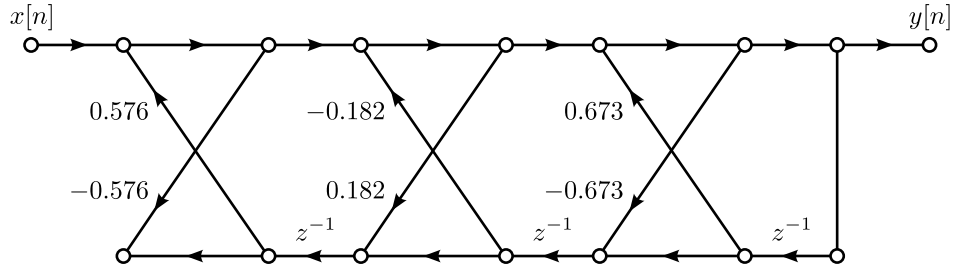


Figura 4.37: Estructura lattice con la misma función de transferencia que la estructura de la figura 4.36.

anterior, en donde los polos de la función de transferencia $H(z)$ están dentro de círculo unidad y todos los parámetros k tienen valor absoluto menor que 1. Para sistemas IIR, la garantía de estabilidad inherente a la condición $|k_i| < 1$ es particularmente importante. Si bien la estructura lattice requiere el doble de multiplicaciones por muestra de salida que la forma directa, es insensible a la cuantización de los parámetros k . Esta propiedad es la razón de la popularidad de los filtros lattice en diversas aplicaciones, como por ejemplo, en síntesis de voz.

4.6.3. Generalización de los sistemas lattice

En las secciones previas se vió que los sistemas FIR y los sistemas IIR tienen una representación como estructura lattice. Cuando la función del sistema tiene tanto polos como ceros, es posible obtener una estructura lattice en base a una modificación de la estructura lattice todo polos de la figura 4.35. La deducción de esto no se hará, pero se resume en la sección ??.

Bibliografía

- [1] A. V. Oppenheim and R. W. Schaffer, *Discrete-Time Signal Processing*. Prentice Hall, 3rd ed., 2009.
- [2] J. G. Proakis and D. K. Manolakis, *Digital Signal Processing: Principles, Algorithms and Applications*. Prentice Hall, 4th ed., 2006.
- [3] J. W. Brown and R. V. Churchill, *Complex variables and applications*. McGraw-Hill, 9th ed., 2013.
- [4] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, Í. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, and SciPy 1.0 Contributors, “SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python,” *Nature Methods*, vol. 17, pp. 261–272, 2020.
- [5] A. V. Oppenheim and R. W. Schaffer, *Digital Signal Processing*. Englewood Cliffs, N. J., Prentice-Hall, Inc., 1975.