

Tecnológico de Costa Rica
Escuela de Ingeniería Electrónica

Notas de Clase

Procesamiento Digital de Señales

EL-5805

Dr. José Pablo Alvarado Moya

Prefacio

Estas notas de clase pretenden ofrecer al estudiante una guía para el curso “Procesamiento Digital de Señales”. No deben considerarse bajo ninguna circunstancia como fuente de información única. En cada capítulo se hace referencia a fuentes bibliográficas adicionales que el estudiante deberá revisar por su cuenta, con ejercicios adicionales y mayor detalle en su presentación.

El concepto del curso se ha orientado en las sugerencias de los autores John G. Proakis y Dimitris G. Manolakis [15] para un programa semestral de curso de pregrado, con algunas adiciones y principalmente resúmenes de la materia considerados necesarios. Más detalles de los temas tratados aquí se podrán entonces encontrar del capítulo 1 al capítulo 8 del mencionado libro de texto.

Las presentes *notas de clase* son para uso exclusivo del curso “Procesamiento Digital de Señales” de la Escuela de Ingeniería Electrónica del Instituto Tecnológico de Costa Rica.

Dr. José Pablo Alvarado Moya

Cartago, 22 de noviembre de 2010

Reservados todos los derechos. Ni la totalidad ni parte de este documento pueden reproducirse, registrarse o transmitirse en ninguna forma ni por ningún medio sin permiso previo del autor.

© 2005-2007 P. Alvarado Escuela de Electrónica Instituto Tecnológico de Costa Rica

Índice general

Índice de tablas	v
Índice de ejemplos	vii
Revisar	ix
Lista de símbolos y abreviaciones	xi
1 Introducción	1
1.1 Señales	1
1.2 Sistemas	3
1.3 Elementos de un sistema de procesamiento de señales	4
1.4 Ventajas del procesamiento digital sobre el analógico	5
1.5 Aplicaciones	6
1.5.1 Áreas de aplicación	6
1.5.2 Algoritmos	8
1.5.3 Implementación	9
1.5.4 Casos	9
1.6 Problemas	11
2 Señales y Sistemas de Variable Discreta	13
2.1 Señales de variable discreta	13
2.1.1 Manipulaciones elementales de señales de variable discreta	14
2.1.2 Clasificación de señales de variable discreta	16
2.2 Señales sinusoidales y el concepto de frecuencia	21
2.2.1 Señal sinusoidal continua	22
2.2.2 Señal sinusoidal discreta	22
2.2.3 Exponenciales complejos relacionados armónicamente	26
2.3 Sistemas en tiempo discreto	27
2.3.1 Descripción entrada-salida de sistemas	27
2.3.2 Diagramas de bloques	29
2.3.3 Clasificación de los sistemas discretos	31
2.3.4 Interconexión de sistemas discretos	34
2.4 Análisis de sistemas discretos lineales e invariantes en el tiempo	35
2.4.1 Técnicas de análisis	35

2.4.2	Descomposición de una señal en impulsos	36
2.4.3	Convolución	36
2.4.4	Propiedades de la convolución	41
2.4.5	Sistemas LTI causales	42
2.4.6	Estabilidad de sistemas lineales e invariantes en el tiempo	43
2.4.7	Sistemas de respuesta finita e infinita	45
2.5	Sistemas discretos descritos mediante ecuaciones de diferencias	45
2.5.1	Sistemas discretos recursivos y no recursivos	45
2.5.2	Sistemas LTI caracterizados por ecuaciones de diferencias con coeficientes constantes	47
2.5.3	Solución de ecuaciones de diferencias con coeficientes constantes	50
2.5.4	Respuesta impulsional de un sistema recursivo LTI	55
2.6	Correlación	57
2.6.1	Autocorrelación y correlación cruzada	57
2.6.2	Propiedades de la correlación	58
2.6.3	Correlación de secuencias periódicas	59
2.6.4	Secuencias de correlación de entrada-salida	60
2.7	Problemas	63
3	Análisis de sistemas LTI discretos con la transformada z	67
3.1	La transformada z	67
3.1.1	La transformada z directa	67
3.2	Transformadas z racionales	69
3.2.1	Polos y ceros	69
3.2.2	Localización de los polos y el comportamiento en el dominio de n para señales causales	74
3.2.3	La función de transferencia de un sistema LTI	76
3.3	Análisis de sistemas LTI en el dominio z	78
3.3.1	Respuesta de sistemas con función de transferencia racional	78
3.3.2	Condiciones iniciales no nulas	79
3.3.3	Respuesta transitoria y en régimen permanente	81
3.3.4	Causalidad y Estabilidad	82
3.3.5	Cancelación polo-cero	82
3.3.6	Polos de orden múltiple y estabilidad	82
3.3.7	Estabilidad de sistemas de segundo orden	83
3.4	Problemas	86
4	Análisis frecuencial	91
4.1	Espectro de señales continuas	91
4.1.1	Espectro de señales continuas periódicas	91
4.1.2	Espectro de señales continuas aperiódicas	92
4.2	Espectro de señales en tiempo discreto	96
4.2.1	Espectro de señales discretas periódicas	96
4.2.2	Espectro de señales discretas aperiódicas	98

4.2.3	Relación entre las transformadas de Fourier y z	99
4.2.4	El teorema del muestreo	102
4.3	Propiedades de la transformada de Fourier de señales discretas	108
4.4	Sistemas LTI en el dominio de la frecuencia	111
4.4.1	La función de respuesta en frecuencia	111
4.4.2	Respuesta transitoria y en régimen permanente a entradas sinusoidales	113
4.4.3	Respuesta en régimen permanente a señales de entrada periódicas . .	114
4.4.4	Respuesta a señales de entrada aperiódicas	115
4.4.5	Relaciones entre la función de transferencia y la respuesta en frecuencia	115
4.4.6	Cálculo de la respuesta en frecuencia	116
4.5	Sistemas LTI como filtros selectivos en frecuencia	117
4.5.1	Filtros ideales	118
4.5.2	Filtros paso alto, paso bajo y paso banda	119
4.5.3	Resonadores digitales	122
4.5.4	Filtros ranura	124
4.5.5	Filtros peine	124
4.5.6	Filtros paso todo	127
4.5.7	Osciladores digitales	129
4.6	Sistemas inversos	129
4.6.1	Invertibilidad de sistemas LTI	130
4.6.2	Sistemas de fase mínima, fase máxima y fase mixta	132
4.6.3	Identificación de sistemas	134
4.7	Transformada Discreta de Fourier	135
4.7.1	Muestreo en el dominio de la frecuencia	136
4.7.2	La transformada discreta de Fourier	139
4.7.3	Relación de la DFT con otras transformadas	141
4.7.4	Propiedades de la DFT	142
4.7.5	Filtrado lineal basado en la DFT	146
4.7.6	Filtrado de secuencias de larga duración	148
4.7.7	Análisis espectral de señales usando la DFT	150
5	Conversión analógica/digital y digital/analógica	153
5.1	Muestreo de señales analógicas	154
5.1.1	Muestreo de señales pasa-bajos	156
5.1.2	Muestreo de señales pasa-banda	156
5.2	Cuantificación de señales de amplitud continua	160
5.3	Codificación de los valores cuantificados	162
5.3.1	Números codificados con coma fija	162
5.3.2	Números codificados con coma flotante	164
5.4	Circuitos de conversión analógica digital	166
5.4.1	Circuito de muestreo y retención	166
5.4.2	Contador	167
5.4.3	Aproximaciones sucesivas	168
5.4.4	Convertidor paralelo	169

5.4.5	Convertidor en subrangos	169
5.4.6	Convertidor delta-sigma	170
5.5	Conversión Digital/Analógica	176
5.5.1	Fuentes de tensión conmutadas	177
5.5.2	Resistencias conmutadas	177
5.5.3	Condensadores conmutados	178
5.5.4	Redes resistivas $R - 2R$	179
5.5.5	Conversión DAC delta-sigma	179
5.6	Problemas	180
6	Implementación de sistemas discretos	181
6.1	Número de Condición	181
6.2	Estructuras directas	182
6.2.1	Estructuras directas para filtros FIR simétricos	185
6.3	Grafos de flujo de señal y estructuras transpuestas	186
6.4	Muestreo en frecuencia	188
6.5	Sistemas en cascada	190
6.6	Sistemas paralelos	191
7	Introducción al diseño de filtros digitales	193
7.1	Causalidad y sus implicaciones	193
7.2	Filtros de respuesta de impulso finita	196
7.2.1	Diseño de filtros FIR por el método de ventanas	198
7.2.2	Diseño de filtros óptimos	203
7.3	Diseño de filtros de respuesta impulsional infinita a partir de filtros analógicos	206
7.3.1	Diseño por aproximación de derivadas	207
7.3.2	Diseño por invarianza impulsional	208
7.3.3	La transformada z adaptada	209
7.3.4	Diseño por transformación bilineal	209
7.3.5	Filtros Analógicos	210
7.4	Transformación de Frecuencia	210
	Bibliografía	213
	A Octave y Matlab	215
	Índice alfabético	217

Índice de tablas

1.1	Características de las señales	1
2.1	Propiedades de señales de variable discreta.	16
2.2	Propiedades de sistemas discretos.	31
2.3	Ejemplo de convolución de dos secuencias finitas.	38
2.4	Forma general de la solución particular para diversos tipos de señal de entrada	54
3.1	Transformada z de algunas funciones comunes	71
3.2	Funciones de transferencia de segundo orden y equivalentes temporales . . .	85
4.1	Propiedades de la Serie de Fourier	94
4.2	Propiedades de la Transformada de Fourier	97
4.3	Propiedades de la DFT	146
5.1	Estándar de coma flotante IEEE 754-1985	166
7.1	Simetrías en filtros FIR de fase lineal	198
7.2	Funciones utilizadas como ventanas.	199
7.3	Descomposición de filtros FIR en $P(\omega)$ y $Q(\omega)$	205
7.4	Características de Filtros Paso Bajo Analógicos	210
7.5	Transformaciones de frecuencia para filtros analógicos.	211
7.6	Transformaciones de frecuencia para filtros digitales.	212

Índice de ejemplos

2.1	Escalón unitario como suma de impulsos desplazados	14
2.2	Operaciones básicas con señales	15
2.3	Señales de energía y potencia	18
2.4	Simetría de funciones	19
2.5	Descripción entrada-salida	27
2.6	Salida de sistema acumulador	28
2.7	Diagrama de bloques	30
2.8	Invarianza en el tiempo	32
2.9	Sistemas lineales	33
2.10	Descomposición en impulsos	36
2.11	Convolución	37
2.12	Longitud de la convolución de dos señales finitas	37
2.13	Reacción de sistemas con respuesta exponencial	39
2.14	Estabilidad y convergencia	44
2.15	Estabilidad y convergencia	44
2.16	Sistemas discretos recursivos y no recursivos	46
2.17	Linealidad de sistema de primer orden	49
2.18	Solución homogénea	51
2.19	Respuesta de entrada nula para sistema de segundo orden	52
2.20	Respuesta de entrada nula con raíz de multiplicidad 2	53
2.21	Solución particular	53
2.22	Solución total	54
2.23	Respuesta impulsional de un sistema recursivo LTI	56
4.1	Serie de Fourier de pulso rectangular continuo periódico	92
4.2	Serie de Fourier de pulso rectangular continuo aperiódico	96

Revisar

Por hacer:.....	68
Por hacer:.....	166
Por hacer:.....	187
Por hacer:.....	191
Por hacer:.....	191

Lista de símbolos y abreviaciones

Notación general

A	Matriz.
	$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix}$
$\mathbb{N}^+, \mathbb{N}^*$	Conjunto de los números naturales sin cero $\mathbb{N}^+ = \mathbb{N} \setminus \{0\}$.
\mathbb{N}, \mathbb{N}_0	Conjunto de los números naturales $\mathbb{N} = \{0, 1, 2, \dots\}$.
\mathbb{Z}	Conjunto de los números enteros $\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$.
\mathbb{Q}	Conjunto de los números racionales $\mathbb{Q} = \{q \mid q = \frac{n}{d}; n, d \in \mathbb{Z}\}$.
\mathbb{R}	Conjunto de los números reales.
\mathbb{C}	Conjunto de los números complejos.
Δ	Cuanto. Distancia entre dos niveles de cuantificación.
$e_q(t)$ o $e_q(n)$	Error de cuantificación.
F	Frecuencia en ciclos por unidad de tiempo para señales de variable continua.
f	Frecuencia en ciclos por muestra para señales de variable discreta.
F_s	Frecuencia de muestreo de una señal analógica. $F_s = 1/T$
$h(n)$	Respuesta al impulso
$H(\omega)$	Respuesta en frecuencia
$\angle H(\omega)$	Respuesta en fase
$ H(\omega) $	Respuesta en magnitud
$H(z)$	Función de transferencia
$\text{Im}(z)$ o z_{Im}	Parte imaginaria del número complejo z
j	$j = \sqrt{-1}$
$\circ \text{---} \bullet$	Mapeo de un dominio temporal al dominio frecuencial o z
$\bullet \text{---} \circ$	Mapeo de un dominio frecuencial al dominio temporal o z
$\text{Re}(z)$ o z_{Re}	Parte real del número complejo z
$\mathcal{T}[\cdot]$	Transformación realizada por un sistema
T	Intervalo de muestreo. $T = 1/F_s$
T_p	Periodo fundamental $T_p = 1/F$ para señales de variable continua.
$x(n)$	Señal de variable discreta.

$\underline{\mathbf{x}}$	Vector.
	$\underline{\mathbf{x}} = [x_1 \ x_2 \ \dots \ x_n]^T = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$
y	Escalar.
z^*	Complejo conjugado de z
Ω	Frecuencia angular en radianes por unidad de tiempo para señales de variable continua.
ω	Frecuencia angular en radianes por muestra para señales de variable discreta.

Abreviaciones

BIBO	Entrada acotada – Salida acotada (<i>bounded input – bounded output</i>)
DSP	<i>Digital Signal Processing</i> (o <i>Processor</i>).
FIR	Respuesta finita al impulso (<i>Finite Impulse Response</i>)
IIR	Respuesta infinita al impulso (<i>Infinite Impulse Response</i>)
LTI	Sistema lineal e invariante en el tiempo (<i>Linear and Time Invariant</i>)
PDS	Procesamiento Digital de Señales.
SQNR	Relación señal a ruido de cuantificación (<i>signal to quantization noise ratio</i>).

Capítulo 1

Introducción

El Procesamiento Digital de Señales (PDS) es un área de la ciencia y la ingeniería que se ha desarrollado rápidamente desde la segunda mitad del siglo XX. Tanto los aportes teóricos como de aplicación continúan extendiéndose desde y hacia varias áreas del saber. Los avances en el procesamiento y compresión de audio y video, así como las nuevas tecnologías en comunicaciones digitales (telefonía celular, modems ADSL, etc.) son quizá los ejemplos de aplicación más representativos del PDS.

1.1 Señales

Para definir las tareas del PDS se requiere primero precisar el concepto de *señal*, considerada aquí como aquella observación de una magnitud física en función de variables independientes de tiempo y espacio, realizada de tal modo que la señal contenga información de los procesos observados.

En general, toda señal contiene información que se desea extraer o modificar de acuerdo a los requisitos de cada aplicación particular. Sismógrafos, por ejemplo, registran señales sísmicas que contienen información sobre intensidad y características espectrales de los sismos, con ayuda de las cuales pueden determinarse entre otras cosas la ubicación de epicentros y la naturaleza de los sismos. Las señales electrocardiográficas permiten al médico determinar el estado del corazón de sus pacientes.

Tabla 1.1: Características de las señales

Característica	Valores	
Número de variables	una variable	múltiples variables
Dimensionalidad	escalar	vectorial (multicanal)
Variables independientes	discretas	continuas
Valores de la señal	discretos	continuos
Naturaleza estadística	deterministas	aleatorias

La tabla 1.1 resume las características utilizadas para clasificar las señales. Las señales son representadas por funciones matemáticas de una o más variables. Una señal de voz, por ejemplo, puede representarse como una función de una variable temporal $f(t)$, imágenes se pueden considerar como funciones de dos variables espaciales $f(x, y)$, y vídeo como una señal espacio-temporal $f(x, y, t)$.

Las funciones pueden ser además escalares o vectoriales. Si la voz se captura con un micrófono monofónico, la señal eléctrica de salida tendrá por ejemplo un solo valor de tensión eléctrica en cada instante de tiempo. Por otro lado, un electroencefalograma provee un conjunto o *vector* de señales eléctricas provenientes de los diferentes electrodos para cada instante t :

$$\mathbf{f}(t) = [f_1(t) \quad f_2(t) \quad \dots \quad f_n(t)]^T$$

Otro ejemplo de señales vectoriales utilizadas frecuentemente en ingeniería son las imágenes en color, en las que cada elemento de la imagen o *pixel* se representa como un vector en un espacio de color, donde las componentes del vector pueden, por ejemplo, representar los valores de los colores primarios rojo, verde y azul. A cada una de las componentes de las señales vectoriales se les denomina usualmente *canales* y por lo tanto a la señal se le denota como *multicanal*.

Las variables de las que depende la señal pueden ser discretas o continuas. La salida de un foto-transistor puede, por ejemplo, ser obtenida en todo instante de tiempo t (variable continua), mientras que el número de llamadas realizado por hora es una señal que el ICE puede generar para instantes discretos de tiempo nT distanciados por un intervalo de $T = 1$ h (variable discreta). Los puntos donde la variable independiente de una señal discreta está definida no deben ser necesariamente equidistantes; sin embargo, usualmente este tipo de distribución homogénea de las *muestras* se utiliza por su conveniencia computacional y manejabilidad matemática.

Los valores que puede tomar una señal pueden ser también discretos o continuos. Así, el voltaje del fototransistor puede tomar cualquier valor real en un intervalo, mientras que el número de llamadas es siempre un valor entero. También los valores de una función discreta pueden ser equidistantes o seguir otros patrones más complejos (como el logarítmico). El término *digital* se utiliza para señales de variables independientes discretas y de valores discretos, mientras que *analógica* es una señal con variables independientes continuas y valores continuos. El análisis matemático involucrado en el tratamiento de señales digitales pueden simplificarse si se realiza a través de funciones de valor continuo y variable discreta, llamadas usualmente *señales en tiempo discreto*, por representar la variable independiente generalmente instantes de tiempo definidos. Este será el enfoque utilizado en este documento.

Un último criterio de carácter matemático para clasificar las señales es su naturaleza estadística: las señales pueden ser deterministas si puede especificarse con precisión la forma de la función. Por ejemplo, para señales determinísticas definidas en el tiempo, sus valores en el pasado, presente y futuro son siempre conocidos (por ejemplo, una señal senoidal). Por otro lado, las señales aleatorias o estocásticas solo permiten una descripción aproximada de la forma de su función, por tener asociado un comportamiento impredecible (por ejemplo, un generador de ruido, una señal sísmica, una señal acústica de voz).

Asociado a la naturaleza estadística de la señal se distingue además entre señales estacionarias y no estacionarias. Las señales estacionarias son aquellas cuyos parámetros estadísticos no varían en el tiempo, lo que de ninguna manera implica que el valor de la señal se constante. Por otro lado, las señales no estacionarias tienen parámetros estadísticos que varían en el tiempo.

En el presente texto se estudiarán señales de una variable, de valor escalar, digitales y de naturaleza determinista. En cursos de procesamiento de imágenes se extienden los conceptos a señales vectoriales (usualmente tres dimensiones) de dos variables discretas espaciales (x, y) . El análisis de señales estocásticas es tema usual para cursos de posgrado; sin embargo, es en este curso introductorio donde se presentan todas las bases necesarias para comprender los conceptos avanzados.

En principio las señales pueden corresponder a cualquier tipo de magnitud físicas observada; sin embargo, por medios electrónicos solo señales eléctricas pueden ser procesadas, por lo que usualmente se requieren transductores o sensores que realicen la correspondiente conversión.

1.2 Sistemas

El término *sistema* denota a una colección o conjunto de elementos interrelacionados que conforman un todo unificado. Su raíz etimológica es el término latino *systema*, que a su vez proviene del griego *σύστημα* relacionado con los conceptos *combinar* e *instalar*.

Un sistema puede formar parte de otro sistema de mayor nivel, en cuyo caso al primero se le denomina *subsistema* del segundo. Los diferentes subsistemas intercambian por lo general información, materia o energía para lograr algún objetivo. Los términos *señales de entrada* o *de salida* se utilizan entonces para abstraer ese flujo de información, materia o energía en el concepto matemático de *funciones*.

El sistema entonces puede interpretarse como un conjunto de subsistemas que logran transformar una señal en otra. Estos dispositivos pueden ser entes físicos, como un circuito electrónico, o virtuales, como algoritmos implementados en software.

En la literatura actual se tiende a diferenciar entre dos tipos de tareas de los sistemas: *procesamiento* y *análisis*. Se dice que un sistema *procesa* una señal si la señal de salida tiene las mismas características semánticas de la entrada: por ejemplo, si la entrada representa una señal de voz, la salida de un sistema procesador será también voz aunque quizá modificada para cumplir ciertos requisitos de la aplicación. Se dice que un sistema realiza *análisis* de la señal, si la salida tiene otra naturaleza semántica a la entrada. Por ejemplo, un módulo de un reconocedor de habla puede extraer de una señal de voz información sobre la presencia de la vocal 'a' en ella. Usualmente el análisis de una señal se realiza a través de diversos pasos de procesamiento de la señal, junto con tareas de reconocimiento o codificación de patrones.

En resumen, el procesamiento digital de señales¹, abreviado PDS o DSP por sus siglas en inglés (*Digital Signal Processing*) se refiere al proceso de modificación de una señal digital

¹En la literatura en castellano se encuentran los términos *tratamiento* o *procesamiento* de señales como

en un sistema, realizado para destacar o suprimir diferentes características de la señal que tienen algún significado especial para una aplicación en particular.

1.3 Elementos de un sistema de procesamiento de señales

La mayoría de las señales en ciencia e ingeniería tienen una naturaleza *analógica*, es decir, tanto las variables independientes de las funciones que las representan como sus valores son continuos. Matemáticamente estas señales se representan como funciones $f(t)$

$$f : \mathbb{R} \rightarrow \mathbb{R}$$

es decir, relaciones matemáticas que mapean valores reales en valores reales. Este tipo de señales pueden ser tratadas directamente utilizando sistemas analógicos, como por ejemplo filtros pasivos o analizadores de frecuencia (figura 1.1).



Figura 1.1: Procesamiento analógico de una señal [15].

El procesamiento digital (figura 1.2) requiere transformar las señales de entrada a un formato digital, es decir, a funciones $f(n)$

$$f : \mathbb{Z} \rightarrow \mathbb{Z}.$$

Esto ocurre en una etapa llamada conversión analógica-digital (A/D).



Figura 1.2: Procesamiento digital de una señal analógica [15].

La señal digitalizada es tratada luego en el procesador digital de señales, que puede ser desde un computador de propósito general, pasando por sistemas empujados basados en microcontroladores, hasta circuitos digitales específicamente diseñados para realizar las tareas de procesamiento deseadas; sin embargo, las configuraciones programables tanto en software como en hardware reconfigurable son las que han brindado al procesamiento digital una flexibilidad inalcanzable con sistemas analógicos equivalentes. El vertiginoso avance en la electrónica digital ha permitido el uso cada vez más generalizado de las técnicas digitales de

sinónimos. La preferencia de autores y traductores españoles por el primero radica en la acepción principal de *proceso* en la variante dialectal ibérica como “causa civil o criminal”, que no se aplica tanto en Latino América.

procesamiento. En la actualidad hasta los más pequeños teléfonos celulares utilizan algoritmos de alta complejidad que hace tan solo 15 años hubieran requerido computadores de gran tamaño y costo.

El último paso del procesamiento digital consiste en convertir la salida del bloque procesador a una señal analógica, lo que ocurre en el llamado convertidor digital-analógico (D/A). En muchos casos de módulos de *análisis* de señal, la última etapa no es necesaria, pues la información a extraer puede obtenerse fácilmente de las representaciones digitales.

1.4 Ventajas del procesamiento digital sobre el analógico

Para poder representar una señal analógica por medio de una señal digital sin que sufra pérdidas de información considerables, la señal analógica debe ser *digitalizada* con una tasa de muestreo suficientemente alta (esto se analizará con suficiente detalle en capítulos posteriores). La tecnología digital impone límites de velocidad de procesamiento, que, aunque cada vez menos restrictivos, determinan los anchos de banda de señales que pueden ser tratadas digitalmente. Es por esto que los sistemas analógicos siguen siendo irremplazables en aplicaciones con señales de anchos de banda en el orden de los gigahertz. En casos donde las señales pueden ser digitalizadas sin perder información de forma considerable y se cuenta con suficiente tiempo para realizar los cálculos necesarios, es preferible utilizar un sistema digital sobre su equivalente analógico. Esto por varias razones:

Los sistemas digitales son usualmente más baratos y confiables para el procesamiento de señales. Como ya se mencionó, pueden utilizarse sistemas programables, por lo que a través de cambios en el software pueden modificarse o adaptarse sus características, proveyendo así un alto grado de flexibilidad en el diseño. Además, las precisiones alcanzables con sistemas digitales son usualmente mucho mayores que los circuitos analógicos, en los que el error acumulado en forma de ruido aumenta con cada etapa de procesamiento.

Un sistema digital funciona en toda su vida útil exactamente de la misma manera, y la fabricación de dispositivos asegurará en todos ellos un comportamiento idéntico. Esto contrasta con los diseños analógicos, donde las características de los componentes, pasivos y activos, varían con el tiempo y donde la tolerancia de cada componente alterará en alguna medida el funcionamiento del sistema total. Además del envejecimiento de los circuitos, el funcionamiento de los sistemas analógicos tiende a ser más sensible a cambios en la temperatura y a fuentes externas de interferencia que los sistemas digitales. En este sentido se dice que los sistemas digitales son más robustos que los sistemas analógicos.

Otra ventaja de los sistemas de procesamiento digital tiene que ver con las posibilidades de almacenamiento. Los niveles de ruido introducidos en sistemas de almacenamiento analógicos (como cintas magnéticas) son extremadamente altos comparados con el almacenamiento prácticamente sin pérdidas (excepto las introducidas por la propia digitalización) de señales digitales. Por este motivo, con señales digitales es más factible realizar los llamados procesamientos “fuera de línea” (*off-line*), donde el tratamiento de la señal se realiza en otro tiempo al de la captura de la señal. Esto es muy útil por ejemplo en astronomía, donde

las altas cantidades de información capturadas por los radio-telescopios pueden ser entonces analizadas mucho después de la adquisición de datos, sin riesgos de producir conclusiones incorrectas producto de imprecisiones del almacenaje. Otro ejemplo es el análisis de imágenes médicas, donde altos volúmenes de información son analizados en procesos automáticos o semi-automáticos a posteriori en la detección de enfermedades y el planeamiento de operaciones quirúrgicas.

Pero quizá una de las ventajas fundamentales del procesamiento digital es la complejidad alcanzable por medio de algoritmos de software, para los cuales pueden incluso no existir equivalentes analógicos. Debido a las consiguientes simplificaciones en los procesos de diseño de sistemas digitales y considerando la predictibilidad en el incremento de las capacidades de procesamiento y memoria (por ejemplo, por medio de la Ley de Moore), se acostumbra desarrollar los modernos y complejos algoritmos para el tratamiento digital de señales utilizando equipos de alto costo y tal vez de dimensiones volumétricas que exceden las limitaciones espaciales, puesto que se asume que en los próximos años se podrá integrar y mejorar el hardware utilizado hasta satisfacer las expectativas de aparatos domésticos. Como ejemplo, los nuevos estándares de codificación de video del grupo MPEG necesitan varios microprocesadores de última generación para funcionar, y aún así no alcanzan las velocidades necesarias para desplegar los videos con la naturalidad deseada. Se parte del hecho que en un corto plazo los prototipos actuales podrán ser integrados y comercializados hasta en sistemas portátiles.

1.5 Aplicaciones

1.5.1 Áreas de aplicación

Las aplicaciones del procesamiento digital de señal son hoy en día incontables. Las más conocidas, pero no las únicas, se resumen a continuación:

- Aplicaciones automotrices
Control del motor, sistemas antibloqueo (ABS), sistemas de navegación, análisis de vibración, etc.
- Electrónica de consumo
Radio y televisión digital, sistemas de video (DVD, Blue-Ray, etc.), juguetes educativos, instrumentos musicales, sistemas de impresión y despliegue, como monitores de plasma, LED, LCD, etc.
- Industria
Control numérico, monitorización de líneas de potencia, robótica, sistemas de seguridad.
- Instrumentación
Generación de funciones, emparejamiento de patrones, procesamiento sísmico, análisis espectral, análisis de transcientes.
- Medicina

Equipo de diagnóstico, monitorización de pacientes, prótesis auditivas, visuales y mecánicas, equipos de ultrasonido, tomografía, MRI, etc.

- Telecomunicaciones

Modems, ecualizadores de señal, codificadores y decodificadores, telefonía celular, multiplexación, cancelación de eco, repetidores de señal, compensación de canal, modulaciones de espectro ensanchado, video-conferencia, cifrado de datos

- Voz/Habla

Verificación de locutor, mejoramiento de señal, reconocimiento de habla, síntesis de habla

El tratamiento de señales acústicas es utilizado entre otros en el almacenamiento y transmisión eficientes de sonido digital (MP3, OggVorbis, etc.), el procesamiento profesional de sonido en industria musical y cinematográfica, el manejo de señales de ultrasonido para elaboración de imágenes médicas, o el procesamiento de voz humana, necesario para codificar, encriptar, reconocer o sintetizar el habla.

El procesamiento de imágenes bidimensionales permite analizar las señales obtenidas por medio de cámaras industriales, hoy en día frecuentemente encontradas en las líneas de producción; además, el procesamiento de imágenes tomadas por satélite permiten identificar entre otras cosas el tipo de uso del suelo, facilitan la construcción de mapas actualizados, etc. Esta área es central en la codificación y compresión de señales de video, tal y como los establecen los estándares MPEG (*Motion Picture Expert Group*).

El procesamiento de imágenes tridimensionales se utiliza por ejemplo en el análisis y generación de imágenes de resonancia magnética (MRI), utilizadas en medicina como instrumento de observación de tejidos internos de un paciente, sin tener la necesidad de utilizar procedimientos quirúrgicos.

Las técnicas modernas de análisis permiten obtener mejores resoluciones y aumentar la confiabilidad de la información producida por sonares y radares. Por otro lado, el estudio digital de señales sísmicas y volcánicas permite incorporar técnicas de simulación y reconocimiento de patrones que mejoran la predicción de zonas y periodos de riesgo.

En los procesos de automatización industrial el procesamiento digital es en la actualidad omnipresente, pues a pesar de que la mayoría de los sensores producen salidas analógicas, estas son transformadas casi inmediatamente a señales digitales para permitir una transmisión más confiable y sin mayores pérdidas a las unidades de procesamiento, para facilitar la aplicación de algoritmos de extracción de la información de interés, y para hacer posible la utilización de técnicas confiables de almacenamiento de la información, que puede ser la base luego para el mejoramiento de los procesos productivos, en el cálculo de costos, etc.

En la preparación de señales para su transmisión y en su decodificación y mejoramiento del lado de los receptores, el procesamiento digital juega un papel cada vez más importante. Un ejemplo lo representan los módems utilizados actualmente para permitir enlaces de alta velocidad a través de las líneas telefónicas de cobre, denominado ADSL (*Asymmetric Digital Subscriber Line*), donde el procesamiento digital es utilizado para codificar y decodificar las tramas y las señales de acuerdo a los estándares de modulación digitales.

1.5.2 Algoritmos

Conceptos algorítmicos clásicos del procesamiento digital, encontrados en las áreas de aplicación anteriores son: compresión, cifrado, reconocimiento, identificación, sintetización, eliminación de ruido, estimación espectral y filtrado, solo por mencionar algunos.

La *compresión* consiste en la reducción de capacidad necesaria para almacenar o transmitir una señal. En telefonía celular señal de la voz es comprimida para poder transmitirla en anchos de banda relativamente pequeños, comparados con los utilizados en telefonía fija. Los estándares MPEG contienen sofisticados algoritmos de compresión de imágenes que permiten reducir en factores de 8 a 12 veces las señales de video.

El *cifrado* es necesario cuando la confidencialidad de la información en las señales debe ser asegurada. Algoritmos complejos codifican la información de forma tal que solo el destinatario pueda decodificarla.

Tareas de *reconocimiento* intentan inferir de patrones en la señal, información contenida de forma implícita. Por ejemplo, de una señal de voz puede reconocerse tanto el mensaje hablado, como el hablante (reconocimiento de habla y de voz, respectivamente). En imágenes médicas pueden utilizarse algoritmos para reconocer tejidos malignos y benignos, o en imágenes industriales pueden ser reconocidos caracteres, formas de productos, el ensamblaje correcto de partes, etc.

La *identificación* está relacionada con el reconocimiento. Aquí no se intenta descubrir una identificación para el contenido de una señal, sino verificar que una identidad previamente dada es compatible con la señal. Métodos de identificación se utilizan, junto con la encriptación, en aplicaciones de alta seguridad.

La *sintetización* permite producir señales artificiales similares a aquellas generadas a través de fenómenos físicos. Es utilizada por ejemplo en la elaboración de efectos acústicos e imitación de instrumentos musicales en sintetizadores de sonido. Otro ejemplo es la sintetización de voz humana, utilizada en interfaces avanzadas hombre-máquina.

Las señales transmitidas por canales analógicos usualmente son perturbadas con *ruido*, es decir, con alteraciones indeseables que no contienen ninguna información relevante para la aplicación. Por medio del procesamiento digital es posible aplicar diversos algoritmos que permiten reducir el efecto de dichas distorsiones.

El *filtrado* es un concepto básico del procesamiento digital que forma parte de prácticamente cualquier otro algoritmo. El sistema que realiza esta tarea se denomina *filtro*, y permite el paso de solo ciertas “componentes” de su señal de entrada, y bloqueando el paso de otras. Algunos detalles serán presentados en los capítulos 4.5 y 7.

La estimación espectral es utilizada en varias áreas de las comunicaciones para encontrar los rangos de frecuencias en que se concentra la energía de una señal (como en el caso del arriba mencionado ADSL).

1.5.3 Implementación

El diseño de algoritmos para resolver las tareas mencionadas anteriormente se fundamenta en teorías matemáticas que aseguran su funcionamiento bajo limitaciones pre-establecidas. Para cada algoritmo se puede elegir aquella estrategia de implementación que mejor satisfaga los requisitos de una aplicación particular. Esto puede involucrar

- plataformas de propósito general, como un computador personal;
- plataformas empuotradas, incluyendo teléfonos celulares, PDA, controles en máquinas, etc. Según sea la demanda computacional requerida, éstas plataformas se pueden basar en
 - microprocesadores de propósito general, o
 - microcontroladores especializados en el tratamiento de señales digitales (PDS, Procesadores Digitales de Señales)
- hardware reconfigurable, que se emplea en aplicaciones de alto desempeño, para el cual los PDS no tienen suficientes prestaciones,
- circuitos integrados de aplicación específica (ASIC), utilizados si se espera una producción en masa (como decodificadores del formato de audio Mp3)
- implementación de circuitos de procesamiento en tiempo discreto.

En la última estrategia la señal nunca es convertida a un formato digital, pero se incluye aquí por basarse su diseño en las teorías de análisis de señales en tiempo discreto, a tratar en este documento.

El desarrollo de tecnologías de implementación de algoritmos ha tendido en los últimos años a simplificar los procesos, de modo que utilizando lenguajes de alto nivel (como por ejemplo el lenguaje C, o MatLab) sea posible obtener con los compiladores adecuados el código optimizado para el procesador particular utilizado o incluso la descripción en VHDL o Verilog del circuito que realiza la tarea descrita en C. Esto reduce costos al acelerar los procesos de implementación y evita el entrenamiento de los ingenieros en tecnologías cuya vigencia se restringe a unos pocos años. Sin embargo, en casos con restricciones críticas es inevitable recurrir ya sea a la optimización de algoritmos en ensamblador del PDS empleado o la optimización manual de los circuitos.

1.5.4 Casos

Los codificadores de voz (o codecs) utilizados en telefonía celular son un ejemplo de algoritmos complejos imposibles de realizar con circuitos analógicos, al menos en el tamaño de un teléfono celular actual. En GSM, por ejemplo, la señal analógica adquirida por el micrófono es muestreada a 8 kHz, con 13 bits por muestra, lo que equivale a 104 000 bit/s. Según la calidad deseada o disponible, la salida de los codecs reducen el ancho de banda a un rango entre 4,75 kbit/s y 13 kbit/s, es decir, permiten factores de compresión de 21 a 8 veces.

Otro caso impresionante es el de los diversos algoritmos para compresión de video. Considerando que una imagen de televisión tiene 320x200 pixels aproximadamente y cada pixel

necesita 24 bits para ser representado digitalmente con suficiente precisión, entonces una sola imagen necesita, sin compresión alguna, 187,5 kB para ser almacenada. Un segundo de video, asumiendo una frecuencia de 30 imágenes por segundo, necesitaría 5,6 MB. Una película que tarde 90 minutos requeriría entonces más de 30 GB, lo que sería imposible de almacenar en las tecnologías actuales de DVD. Solo con los sofisticados algoritmos de codificación de video y sonido pueden almacenarse no solo la señal de video, sino varias versiones acústicas en diferentes idiomas, y materiales adicionales con video y audio, en poco más de 4 GB de espacio, con factores de compresión más de 10 veces. La implementación de los codificadores y decodificadores se realiza en circuitos integrados de aplicación específica tanto en las cámaras comerciales, como en las unidades de DVD y Blue-ray.

Incluso los sistemas de transmisión de radio y televisión modernos están cambiando de ser completamente analógicos a conceptos digitales. Los nuevos formatos de televisión de alta definición (HDTV) son completamente digitales, que no serían prácticos sin la presencia de los algoritmos para el procesamiento y compresión adecuadas de las señales en cuestión.

Los discos Blue-Ray se utilizan para almacenar vídeo apto para HDTV. Soportan el formato 1080p24, lo que implica una resolución de 1920×1080 píxeles con exploración progresiva², así como una tasa de 24 cuadros por segundo. Una imagen tiene así casi 2,1 millones de píxeles que requieren un espacio de 5,93 MB de almacenamiento sin compresión. Esto implica que para la tasa de 24 imágenes por segundo se requiere almacenar 142,4 MB por cada segundo de vídeo sin comprimir. La tasa de transmisión de datos estandarizada para este formato de HDTV es de 4,5 MB/s, que es un factor aproximadamente 32 veces menor al de la señal *cruda* (sin compresión). Las técnicas de compresión del vídeo utilizan algoritmos de PDS más avanzados que sus equivalentes en un DVD, para lograr alcanzar dicha tasa y así poder almacenar todos los datos (vídeo, texto, audio) en un disco con 25 GB de capacidad.

Medicina es quizá una de las áreas que más ventajas ha tomado del procesamiento digital de señales. Las nuevas tecnologías de visualización de ultrasonido permiten por ejemplo la reconstrucción de una imagen tridimensional del feto en el vientre de la madre, ayudando así al ginecólogo a tomar las precauciones del caso cuando algo no esté en orden. El manejo digital de tomografías computarizadas y de imágenes de resonancia magnética permite la visualización de tejidos internos en formatos legibles para los médicos. Los equipos más modernos permiten incluso hacer uso de la llamada *realidad aumentada*, donde imágenes reconstruidas a partir de las mediciones se superponen a las reales para guiar a los cirujanos en operaciones delicadas.

La creación de prótesis cada vez más complejas es también posible gracias a las técnicas de procesamiento digital. El implante de cóclea, por ejemplo, permite a personas sordas volver a escuchar utilizando el análisis digital de las señales acústicas obtenidas con un micrófono. Similar a este último se trabaja en la actualidad en los implantes de retina, donde complejos algoritmos de PDS intentan transformar la señal capturada por una cámara en impulsos eléctricos que pueden ser acoplados al nervio óptico en el ojo de personas ciegas.

²la exploración progresiva implica que la imagen se genera progresivamente línea por línea en un solo cuadro, y es lo opuesto a la exploración entrelazada, donde la imagen se genera a través de dos cuadros, uno que contiene las líneas pares y otro con las impares.

1.6 Problemas

Problema 1.1. Busque un ejemplo de señal para cada una de las 32 posibles combinaciones de las características de las señales indicadas en la Tabla 1.1.

Problema 1.2. Una señal puede clasificarse según cinco criterios:

Número de variables	(U na variable o M últiples variables)
Dimensión del valor	(E scalar o V ectorial)
Variables independientes	(D iscretas o C ontinuas)
Valores de la señal	(D iscretas o C ontinuas)
Naturaleza estadística	(D eterminista o A leatoria)

Utilice las letras mayúsculas indicadas en negrita para identificar las características de las señales a continuación. Si alguna característica no se aplica o puede interpretarse con cualquier valor de la propiedad, indíquelo con *.

Señal	Característica				
	Núm. Variables (U/M/*)	Dimensión (E/V/*)	Variables (D/C/*)	Valores (D/C/*)	Estadística (D/A/*)
Imagen tomada por una cámara digital a color					
Registro mensual de la posición de una bandada de aves migratorias					
Señales de salida de un micrófono estéreo					
Señal digital					
Señal analógica					

Capítulo 2

Señales y Sistemas de Variable Discreta

2.1 Señales de variable discreta

Sea $x(n)$ una señal de variable discreta, es decir, una función definida para n entero. La figura 2.1 muestra una típica representación gráfica de una señal de este tipo. A n se le denomina número de muestra y a $x(n)$ la n -ésima muestra de la señal.

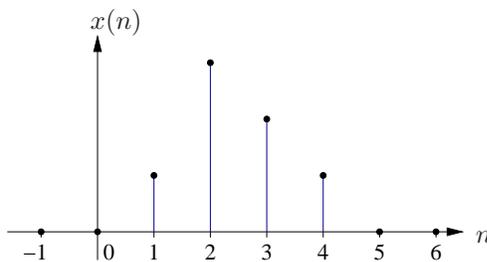


Figura 2.1: Representación gráfica de una función real de variable discreta $x(n)$.

Nótese que $x(n)$ no está definida para n no entero. Un error común es considerar que la señal es cero “entre” las muestras, cuando en realidad allí simplemente la función no está definida. La suposición de que la señal es cero entre muestras es válido solo para una señal de variable real $x(t)$, que es otro concepto diferente.

Además de la representación gráfica se utilizan aquí otras tres notaciones:

1. Funcional:

$$x(n) = \begin{cases} 1 & \text{para } n = 1 \\ 5 - n & \text{para } 2 \leq n \leq 4 \\ 0 & \text{el resto} \end{cases}$$

Esta representación es la forma más compacta de representar señales cuyo comportamiento puede ser expresado directamente por expresiones algebraicas cerradas.

2. Tabular

n	...	-1	0	1	2	3	4	5	...
$x(n)$...	0	0	1	3	2	1	0	...

En programas computacionales como MatLab[12] u Octave[5] las señales se representan con esta notación, donde las dos filas de la tabla se interpretan como dos arreglos de datos independientes.

3. Como secuencia.

Una secuencia de duración infinita con el origen en $n = 0$ (indicado con “↑”) se representa como

$$x(n) = \{\dots, 0, 0, \underset{\uparrow}{1}, 3, 2, 1, 0, \dots\}$$

Si la secuencia es 0 para $n < 0$ se suele representar como

$$x(n) = \{\underset{\uparrow}{0}, 1, 3, 2, 1, 0, \dots\}$$

y si es finita

$$x(n) = \{\underset{\uparrow}{0}, 1, 3, 2, 1\}$$

Si la secuencia inicia en cero, entonces usualmente se omite la flecha:

$$x(n) = \{0, 1, 3, 2, 1\}$$

Por su brevedad y simpleza notacional, ésta será la representación de señales de preferencia en el presente texto.

2.1.1 Manipulaciones elementales de señales de variable discreta

Se describen a continuación algunas transformaciones elementales para señales de variable independiente discreta (o tiempo discreto). Estas transformaciones son la base de operaciones más complejas en el procesamiento digital de señales y se utilizarán a menudo en éste y los siguientes capítulos.

Desplazamiento

La señal $x(n)$ se desplaza k muestras sustituyendo la variable n por $n - k$. Si $k > 0$ la señal se retarda k muestras y si $k < 0$ la señal se adelanta k muestras.

En la manipulación fuera de línea (*off-line*) ambos tipos de desplazamiento son posibles; sin embargo, en sistemas de tiempo real o en línea (*on-line*), solo el retraso de la función es realizable.

Ejemplo 2.1 Utilizando desplazamientos exprese el escalón unitario $u(n)$ en términos de una suma de impulsos $\delta(n)$ desplazados.

Solución:

$$u(n) = \delta(n) + \delta(n-1) + \delta(n-2) + \dots = \sum_{i=0}^{\infty} \delta(n-i)$$

2.1

Reflexión

Consiste en “plegar” la señal $x(n)$ en el instante $n = 0$, sustituyendo la variable n por su inverso aditivo $-n$.

Nótese que las operaciones de reflexión y desplazamiento no son conmutativas, es decir, no es lo mismo reflejar una señal y luego retardarla k unidades que retardar la señal y luego reflejarla, puesto que:

$$\underbrace{x((-n) - k) = x(-n - k)}_{\substack{\text{Reflexión,} \\ \text{Desplazamiento}}} \neq \underbrace{x(-(n - k)) = x(-n + k)}_{\substack{\text{Desplazamiento,} \\ \text{Reflexión}}}$$

Escalado de variable o submuestreo

En el escalado de variable o *submuestreo* (en inglés *downsampling*) se sustituye la variable discreta n por κn con $\kappa \in \mathbb{N}^+$.

Si la señal analógica original $x_a(t)$ fue muestreada de forma homogénea con el intervalo de muestreo T entonces $x(n) = x_a(nT)$, de donde se deriva que $x(\kappa n) = x_a(n(\kappa T))$. Esto implica que el submuestreo equivale a utilizar un intervalo de muestreo de mayor duración, igual a κT .

Suma, multiplicación y escalado de secuencias

Todas estas operaciones afectan la amplitud de las muestras de una secuencia.

- Escalado de amplitud: $y(n) = Ax(n)$
- Suma de secuencias: $y(n) = x_1(n) + x_2(n)$
- Producto: $y(n) = x_1(n)x_2(n)$

Ejemplo 2.2 Dadas las secuencias

$$\begin{aligned} x_1(n) &= u_r(n) = \{0, 1, 2, 3, 4, 5, 6, \dots\} \\ x_2(n) &= (-1)^n u_r(n) = \{0, -1, 2, -3, 4, -5, 6, \dots\} \\ x_3(n) &= \{0, 0, 1\} \end{aligned}$$

Calcule la secuencia

$$x_4(n) = 2x_3(2 - n) - x_1(2n - 1) + x_2(4 - n)u(n)$$

Solución: El primer término $2x_3(2 - n)$ corresponde a una reflexión seguida por un atraso (de la reflexión), escalado por un factor 2. Esto resulta en

$$2x_3(2 - n) = \{2\}_{\uparrow}$$

El segundo término representa un submuestreo retrasado:

$$x_1(2n - 1) = \{0, 1, 3, 5, 7 \dots\}_{\uparrow}$$

El primer factor del tercer término contiene una inversión atrasada:

$$x_2(4 - n) = \{\dots, 6, -5, 4, -3, 2, -1, 0\}_{\uparrow}$$

y al multiplicarlo por el escalón unitario $u(n)$ se eliminan todas las muestras anteriores a $n = 0$:

$$x_2(4 - n)u(n) = \{4, -3, 2, -1, 0\}_{\uparrow}$$

Finalmente, se deben combinar estos tres resultados parciales aditivamente:

$$x_4(n) = \{6, -4, -1, -6, -7, -9, -11, -13 \dots\}_{\uparrow}$$

2.2

2.1.2 Clasificación de señales de variable discreta

La tabla 2.1 presenta cinco propiedades que caracterizan a las señales de variable discreta.

Tabla 2.1: Propiedades de señales de variable discreta.

Propiedad	Clasificación	
Energía	Señal de energía	Señal de potencia
Periodicidad	Señal periódica	Señal aperiódica
Simetría	Señal simétrica	Señal asimétrica
Acotación	Señal acotada	Señal no acotada
Longitud	Señal finita	Señal infinita

Señales de energía y potencia

La figura 2.2 muestra el cálculo de potencia instantánea $p(t)$ y de energía disipada $e(t)$ en un circuito eléctrico simple. El valor de resistencia R representa una constante que únicamente

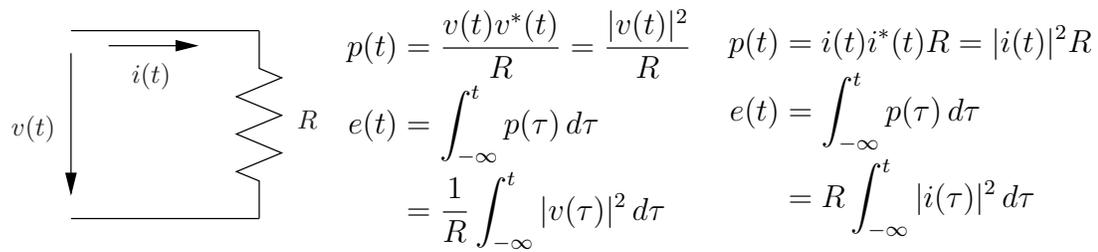


Figura 2.2: Potencia instantánea $p(t)$ y energía $e(t)$ de una señal analógica.

repercutirá en factores de escala de las funciones $p(t)$ y $e(t)$, es decir, el valor de R no altera la *forma* de $p(t)$ o $e(t)$.

Usualmente, a las funciones de forma $|f(t)|^2$ se les denomina entonces *funciones de potencia* de $f(t)$ y a su integral hasta el instante t , $\int_{-\infty}^t |f(t)|^2 dt$, *función de energía* de $f(t)$, por la similitud con las funciones de energía y potencia en el circuito mostrado (asuma por ejemplo $R = 1 \Omega$).

De forma análoga, para una señal de variable discreta $x(n)$ se define su *energía* total como

$$E = \sum_{n=-\infty}^{\infty} x(n)x^*(n) = \sum_{n=-\infty}^{\infty} |x(n)|^2$$

donde el uso de la magnitud de la señal permite aplicar la definición a señales de valor complejo.

Si E es finita entonces a $x(n)$ se le denomina *señal de energía*. Muchas señales de energía infinita poseen *potencia promedio* finita:

$$P = \lim_{N \rightarrow \infty} \frac{1}{2N+1} \sum_{n=-N}^N |x(n)|^2$$

En caso de que P sea finita, se dice que $x(n)$ es una *señal de potencia*.

Si se define la energía E_N de $x(n)$ en un intervalo finito como

$$E_N = \sum_{n=-N}^N |x(n)|^2$$

entonces

$$E = \lim_{N \rightarrow \infty} E_N$$

y la potencia promedio puede entonces también expresarse como

$$P = \lim_{N \rightarrow \infty} \frac{1}{2N+1} E_N$$

con lo que se deriva que si E es finita entonces $P = 0$ y a su vez que si $P > 0$ entonces $E \rightarrow \infty$, o en otras palabras, toda señal de potencia tiene energía infinita.

Ejemplo 2.3 Especifique si el escalón unitario $u(n)$, la rampa $u_r(n)$ y la señal exponencial $x(n) = Ae^{j\omega_0 n}$ son señales de energía o potencia.

Solución:

1. Si $x(n) = u(n)$ entonces

$$\begin{aligned} P &= \lim_{N \rightarrow \infty} \frac{1}{2N+1} \sum_{n=-N}^N |u(n)|^2 \\ &= \lim_{N \rightarrow \infty} \frac{1}{2N+1} \sum_{n=0}^N 1 \\ &= \lim_{N \rightarrow \infty} \frac{N+1}{2N+1} = \frac{1}{2} \end{aligned}$$

lo que implica que $u(n)$ es una señal de potencia.

2. Para la señal rampa unitaria $u_r(n)$ se tiene que

$$\begin{aligned} P &= \lim_{N \rightarrow \infty} \frac{1}{2N+1} \sum_{n=-N}^N |u(n)|^2 \\ &= \lim_{N \rightarrow \infty} \frac{1}{2N+1} \sum_{n=0}^N n^2 \\ &= \lim_{N \rightarrow \infty} \frac{1}{2N+1} \frac{N(N+1)(2N+1)}{6} = \infty \end{aligned}$$

por lo que no es ni señal de energía ni señal de potencia al ser tanto E como P infinitas.

3. La señal $x(n) = Ae^{j\omega_0 n}$, $A \in \mathbb{R}$, tiene una potencia media

$$\begin{aligned} P &= \lim_{N \rightarrow \infty} \frac{1}{2N+1} \sum_{n=-N}^N |u(n)|^2 \\ &= \lim_{N \rightarrow \infty} \frac{1}{2N+1} \sum_{n=0}^N |A|^2 \\ &= \lim_{N \rightarrow \infty} \frac{1}{2N+1} (N+1)|A|^2 = \frac{|A|^2}{2} \end{aligned}$$

y es por lo tanto una señal de potencia.

2.3

Señales acotadas y no acotadas

Una señal $x(n)$ se dice ser acotada (en inglés *bounded signal*) si existe un número real positivo finito M tal que $|x(n)| < M$ para todo n . Por el contrario, si la magnitud de cualquier muestra es infinita, entonces la señal no es acotada. Esta propiedad es fundamental para el concepto de estabilidad de sistemas, que se revisará con detalle posteriormente.

Señales periódicas y aperiódicas

Una señal es periódica con periodo N ($N > 0$) si y solo si

$$x(n + N) = x(n), \quad \text{para todo } n \quad (2.1)$$

El valor más pequeño de N para el que se cumple lo anterior se denomina *periodo fundamental*. Si no existe ningún N que satisfaga (2.1) la señal es entonces aperiódica.

Se demostrará en la siguiente sección como ejemplo particular que la señal $x(n) = \cos(2\pi f_0 n)$ es periódica si y solo si f_0 es un número racional $f_0 = \frac{k}{N}$, donde si k y N son primos relativos entonces el periodo es exactamente N .

Si una señal periódica no toma valores infinitos (es decir, es una señal acotada) entonces es a su vez una señal de potencia, por ser su potencia promedio finita e igual al promedio en un periodo:

$$P = \frac{1}{N} \sum_{n=0}^{N-1} |x(n)|^2$$

Señales pares e impares

Una señal es simétrica o par si

$$x(n) = x(-n) \quad (2.2)$$

Se dice ser asimétrica o impar si

$$x(-n) = -x(n) \quad (2.3)$$

en cuyo caso siempre debe cumplirse que $x(0) = 0$.

Asúmase que toda señal $x(n)$ se puede expresar como la suma de una señal par $x_p(n)$ y otra señal impar $x_i(n)$, de forma que

$$x(n) = x_p(n) + x_i(n) \quad (2.4)$$

Considerando la naturaleza par e impar de las componentes se debe cumplir

$$\begin{aligned} x(-n) &= x_p(-n) + x_i(-n) \\ &= x_p(n) - x_i(n) \end{aligned} \quad (2.5)$$

Sumando y restando (2.4) y (2.5) se despeja

$$x_p(n) = \frac{x(n) + x(-n)}{2} \quad x_i(n) = \frac{x(n) - x(-n)}{2} \quad (2.6)$$

Ejemplo 2.4 Dada una señal $x(n) = \{0, 1, 2\}$ encuentre sus componentes par e impar.

Solución:

Para calcular la componente par se realiza la suma de la señal con su reflexión, y luego se divide por dos:

$$\begin{array}{rcl}
 x(n) & = & \{ 0, \quad 0, \quad \underset{\uparrow}{0}, \quad 1, \quad 2 \} \\
 x(-n) & = & \{ 2, \quad 1, \quad \underset{\uparrow}{0}, \quad 0, \quad 0 \} \\
 \hline
 x(n) + x(-n) & = & \{ 2, \quad 1, \quad \underset{\uparrow}{0}, \quad 1, \quad 2 \} \\
 x_p(n) = \frac{x(n)+x(-n)}{2} & = & \{ 1, \quad 1/2, \quad \underset{\uparrow}{0}, \quad 1/2, \quad 1 \}
 \end{array}$$

Para calcular la componente impar se le subtrae la reflexión a la señal, y luego se divide por dos:

$$\begin{array}{rcl}
 x(n) & = & \{ 0, \quad 0, \quad \underset{\uparrow}{0}, \quad 1, \quad 2 \} \\
 x(-n) & = & \{ 2, \quad 1, \quad \underset{\uparrow}{0}, \quad 0, \quad 0 \} \\
 \hline
 x(n) - x(-n) & = & \{ -2, \quad -1, \quad \underset{\uparrow}{0}, \quad 1, \quad 2 \} \\
 x_i(n) = \frac{x(n)-x(-n)}{2} & = & \{ -1, \quad -1/2, \quad \underset{\uparrow}{0}, \quad 1/2, \quad 1 \}
 \end{array}$$

Se comprueba directamente que se cumple $x(n) = x_p(n) + x_i(n)$. □ 2.4

Señales hermíticas y anti-hermíticas

Una señal de valor complejo se denomina hermítica (simétrica conjugada o par conjugada) si cumple

$$x(n) = x^*(-n) \quad (2.7)$$

de lo que se deduce que

$$\begin{aligned}
 |x(n)| &= |x(-n)| \\
 \operatorname{Re}\{x(n)\} &= \operatorname{Re}\{x(-n)\}
 \end{aligned}$$

es decir, tanto la magnitud como la parte real de una señal hermítica tienen simetría par. Por otro lado

$$\begin{aligned}
 \arg x(n) &= -\arg x(-n) \\
 \operatorname{Im}\{x(n)\} &= -\operatorname{Im}\{x(-n)\}
 \end{aligned}$$

lo que indica que el argumento (ángulo o fase) y la parte imaginaria de una señal hermítica son señales impares.

La señal es anti-hermítica (asimétrica conjugada o impar conjugada) si cumple

$$x(n) = -x^*(-n) \quad (2.8)$$

de lo que se deduce que

$$\begin{aligned}
 |x(n)| &= |x(-n)| \\
 \operatorname{Im}\{x(n)\} &= \operatorname{Im}\{x(-n)\}
 \end{aligned}$$

es decir, tanto la magnitud como la parte imaginaria de una señal anti-hermítica tienen simetría par. Por otro lado

$$\arg x(n) = \pi - \arg x(-n)$$

que no tiene exhibe ninguna simetría, aunque si se multiplica ambos lados de la ecuación por dos se obtiene

$$2 \arg x(n) = 2\pi - 2 \arg x(-n)$$

$$2 \arg x(n) = -2 \arg x(-n)$$

es decir, el doble del ángulo tiene simetría impar.

Para la parte real de la función anti-hermítica se cumple

$$\operatorname{Re}\{x(n)\} = -\operatorname{Re}\{x(-n)\}$$

es decir, tiene simetría impar. Asíumase ahora que es posible realizar una descomposición lineal de cualquier función $x(n)$ de valor complejo en una componente hermítica $x_h(n)$ y otra anti-hermítica $x_{\bar{h}}(n)$, de forma que

$$x(n) = x_h(n) + x_{\bar{h}}(n) \quad (2.9)$$

Considerando la naturaleza hermítica y anti-hermítica de las componentes se debe cumplir

$$\begin{aligned} x^*(-n) &= x_h^*(-n) + x_{\bar{h}}^*(-n) \\ &= x_h(n) - x_{\bar{h}}(n) \end{aligned} \quad (2.10)$$

Sumando y restando (2.9) y (2.10) se despeja

$$x_h(n) = \frac{x(n) + x^*(-n)}{2} \quad x_{\bar{h}}(n) = \frac{x(n) - x^*(-n)}{2} \quad (2.11)$$

2.2 Señales sinusoidales y el concepto de frecuencia

El estudio clásico de señales se basa en su representación como combinación de una serie de señales elementales con un comportamiento conocido o predecible en los sistemas con que se trabaja. En cursos anteriores se han introducido conceptos del análisis de Fourier, transformada de Laplace, donde las señales elementales se caracterizan por su frecuencia y fase respecto a alguna referencia común. El concepto de frecuencia será revisado en las siguientes secciones, donde se apreciarán las diferencias que existen entre los dominios discreto y continuo. El punto de partida para el análisis serán las funciones sinusoidales, como representantes espectrales “puras”.

2.2.1 Señal sinusoidal continua

Una oscilación armónica simple se describe matemáticamente a través de la señal continua [15]:

$$x_a(t) = A \cos(\Omega t + \theta), \quad -\infty < t < \infty \quad (2.12)$$

donde el subíndice a indica que es una señal analógica. A representa la *amplitud* de la señal, Ω es la *frecuencia angular* en radianes por segundo (rad/s) y θ es la *fase* en radianes. Es también común utilizar la frecuencia F en ciclos por segundo o Hertz (Hz) con

$$\Omega = 2\pi F$$

con lo que (2.12) se puede expresar como

$$x_a(t) = A \cos(2\pi F t + \theta), \quad -\infty < t < \infty .$$

Si F es constante, entonces x_a es periódica

$$x_a(t + T_p) = x_a(t)$$

con período fundamental $T_p = 1/F$ (figura 2.3).

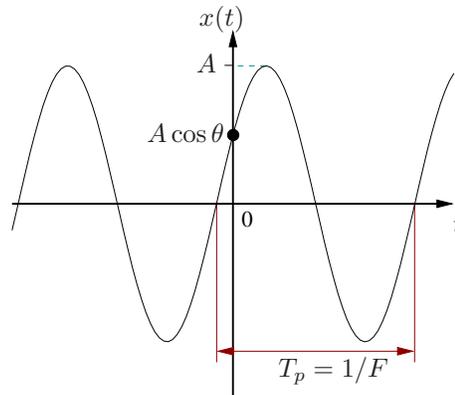


Figura 2.3: Ejemplo de una señal analógica sinusoidal.

2.2.2 Señal sinusoidal discreta

La señal sinusoidal en tiempo discreto se expresa como

$$x(n) = A \cos(\omega n + \theta), \quad -\infty < n < \infty \quad (2.13)$$

con la variable entera $n \in \mathbb{Z}$ (también denominada *número de muestra*), A la magnitud, ω es la frecuencia en radianes por muestra y θ la fase en radianes. De forma similar al caso continuo, puede utilizarse $\omega = 2\pi f$ para reescribir (2.13) como

$$x(n) = A \cos(2\pi f n + \theta), \quad -\infty < n < \infty . \quad (2.14)$$

En este caso las dimensiones de la frecuencia son ciclos por muestra, y tal como estas unidades lo indican, usualmente tiene valores menores que la unidad. La figura 2.4 muestra un ejemplo con $f = 1/12$ y $\theta = \pi/3$.

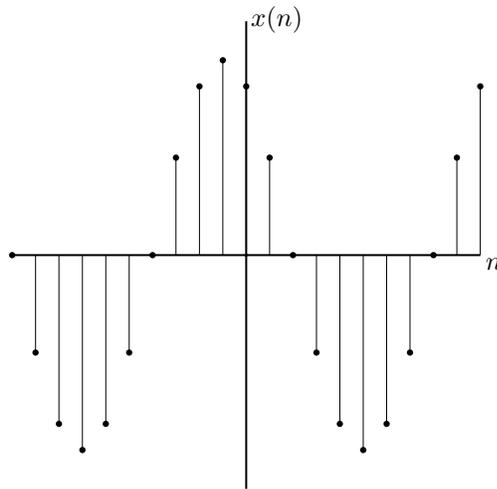


Figura 2.4: Ejemplo de una señal sinusoidal discreta ($\omega = \pi/6$ y $\theta = \pi/3$) [15].

Periodicidad de un senoide discreto

Por definición una señal de variable discreta $x(n)$ es periódica con periodo N ($N > 0$) si y solo si

$$x(n + N) = x(n) \quad \text{para todo } n \quad (2.15)$$

El menor valor de N para el que se cumple (2.15) se denomina *periodo fundamental*. Una señal sinusoidal de frecuencia f_0 es periódica si

$$\cos(2\pi f_0(N + n) + \theta) = \cos(2\pi f_0 n + \theta)$$

lo que se cumple solo si existe un entero k tal que

$$2\pi f_0 N = 2k\pi$$

o, en otros términos

$$f_0 = \frac{k}{N} \quad (2.16)$$

que es siempre un número racional. Esto quiere decir, que una señal sinusoidal discreta con un valor irracional de frecuencia no es periódica.

Para determinar el periodo fundamental de una señal sinusoidal se expresa su frecuencia como en (2.16) y se simplifican los factores comunes de tal forma que k y N formen números primos relativos. Por ejemplo si $f_1 = 5/32 = 0,15625$ el periodo fundamental es $N = 32$, pero si la frecuencia es $f_2 = 4/32 = 0,125$ el periodo fundamental es $N = 8$ (figura 2.5). Esto quiere decir que aún cuando los valores de frecuencia se encuentren relativamente cerca, sus periodos pueden cambiar radicalmente.

Equivalencia de frecuencias en sinusoides discretos

Considérese de nuevo la señal sinusoidal $\cos(\omega_0 n + \theta)$. Es fácil de obtener que:

$$\cos((\omega_0 + 2\pi)n + \theta) = \cos(\omega_0 n + 2\pi n + \theta) = \cos(\omega_0 n + \theta)$$

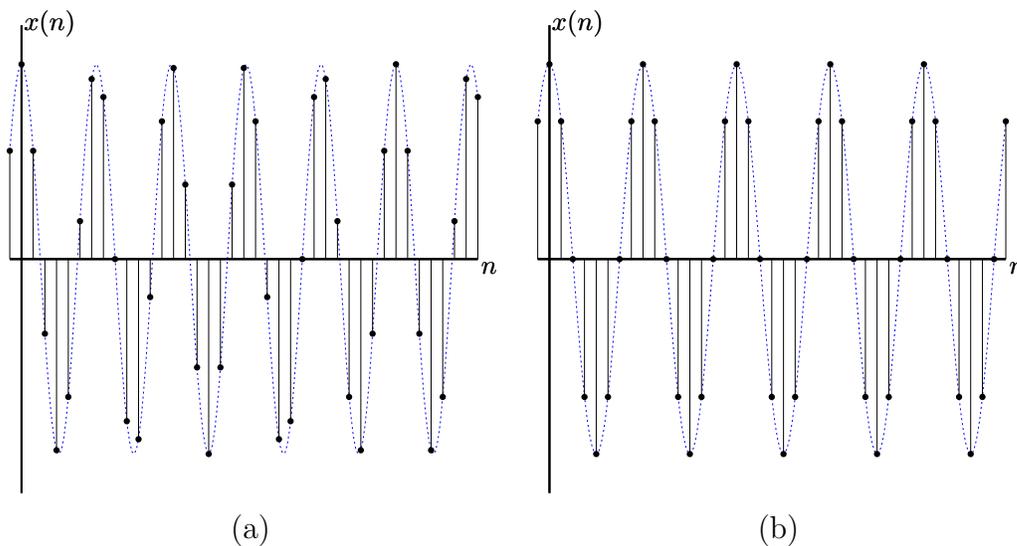


Figura 2.5: Comparación de dos frecuencias cercanas en señales de variable discreta. (a) Frecuencia $f = 5/32$, periodo $N = 32$. (b) Frecuencia $f = 4/32$, periodo $N = 8$. La línea punteada denota una señal continua con frecuencia equivalente para simplificar la comparación.

por lo que todas las secuencias sinusoidales

$$x_k(n) = A \cos(\omega_k n + \theta), \quad k = 0, 1, 2, \dots$$

con

$$\omega_k = \omega_0 + 2k\pi$$

son idénticas. Por otro lado, las secuencias de cualesquiera dos señales sinusoidales discretas con frecuencias en el rango $-\pi \leq \omega \leq \pi$ (ó $-\frac{1}{2} \leq f \leq \frac{1}{2}$) son diferentes. Combinando los resultados anteriores se obtiene que cualquier secuencia sinusoidal de frecuencia $|\omega| > \pi$ (ó $|f| > \frac{1}{2}$) tiene una señal equivalente con $|\omega| \leq \pi$ (ó $|f| \leq \frac{1}{2}$). A las frecuencias $|\omega| \leq \pi$ (ó $|f| \leq \frac{1}{2}$) se les considera *frecuencias fundamentales* y a las frecuencias $|\omega| > \pi$ (ó $|f| > \frac{1}{2}$) se les denomina *alias*.

Equivalencia de frecuencias entre sinusoides continuos y discretos

Para analizar la relación entre frecuencias continuas y discretas asúmase que una señal analógica $x_a(t) = \cos(2\pi Ft)$ se muestrea cada T_s unidades de tiempo para dar origen a la señal en tiempo discreto $x(n) = x_a(nT_s) = \cos(2\pi fn)$. La igualdad es posible solo si los argumentos de las señales cosenoidales son iguales, por lo que

$$2\pi fn \stackrel{!}{=} 2\pi F n T_s$$

y definiendo la frecuencia de muestreo $F_s = 1/T_s$ entonces

$$f = \frac{F}{F_s}$$

de donde se deriva el nombre de *frecuencia normalizada* para f .

De modo similar se deriva para las frecuencias angulares con $\Omega_s = 2\pi F_s$ que

$$\omega = 2\pi \frac{\Omega}{\Omega_s}$$

Tasa máxima de oscilación

Considérese ahora la secuencia sinusoidal $x_0(n) = \cos(\omega_0 n)$ para el rango de frecuencias $\omega_0 \in [0, \pi]$. La figura 2.6 muestra algunos casos particulares como ejemplo. Se puede apreciar que la tasa de oscilación aumenta conforme la frecuencia aumenta.

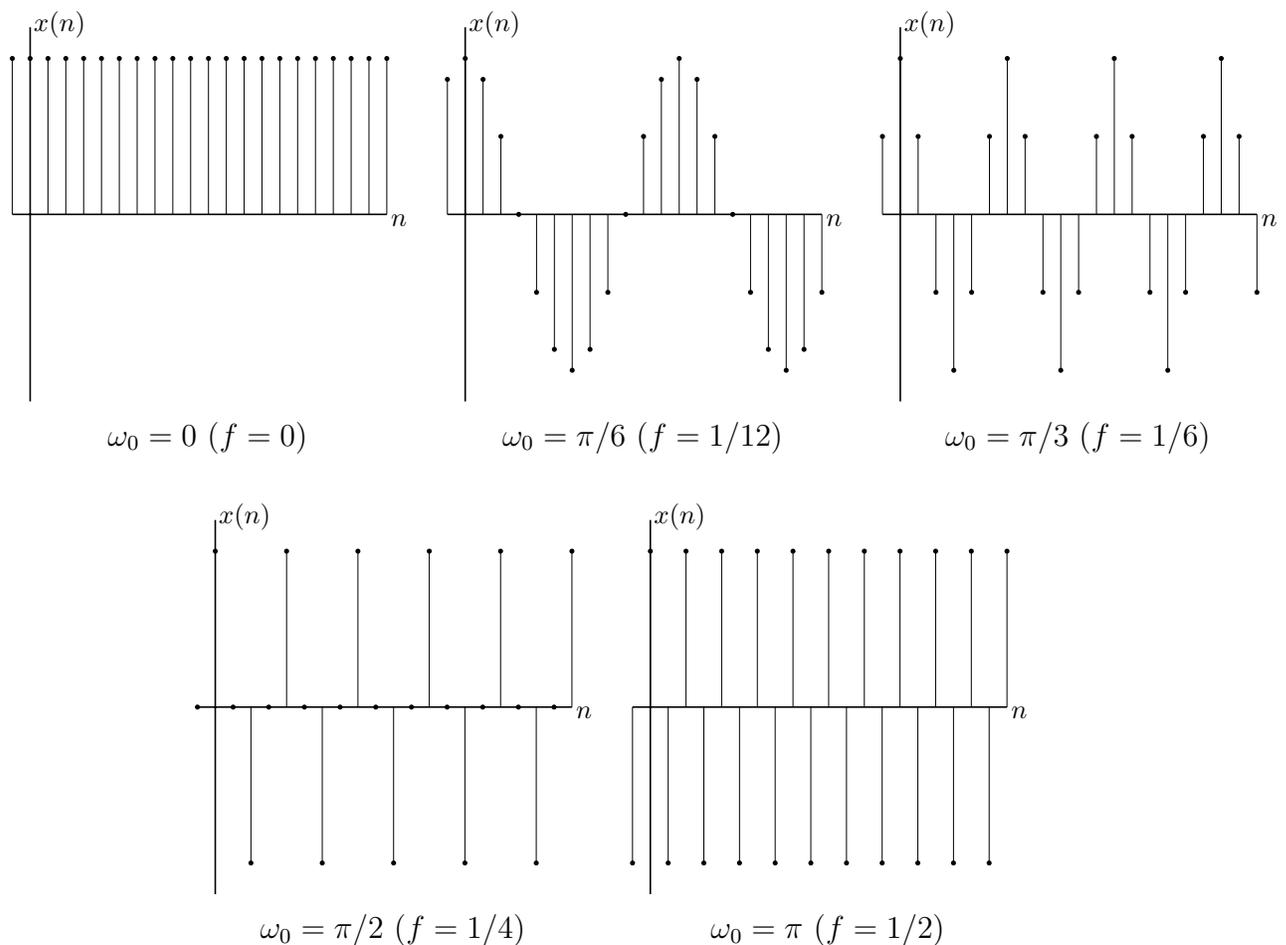


Figura 2.6: Secuencia sinusoidal con diferentes frecuencias.

Ahora, para analizar el caso de frecuencias angulares mayores que π considérese el caso especial de $\omega_1 = 2\pi - \omega_0$. Si $\omega_0 \in [0, \pi]$ entonces $\omega_1 \in [\pi, 2\pi]$ de tal forma que si ω_0 aumenta ω_1 disminuye.

Debido a que

$$x_1(n) = A \cos(\omega_1 n) = A \cos((2\pi - \omega_0)n) = A \cos(-\omega_0 n) = x_0(n)$$

la frecuencia angular ω_1 es un alias de ω_0 . De aquí se concluye que si la frecuencia angular aumenta de π hacia 2π la tasa de oscilación se reduce, al representar estos alias frecuencias que bajan de π a 0.

De forma similar al caso de senoides de variable continua puede utilizarse la identidad de Euler en el caso discreto para introducir el concepto de frecuencia negativa:

$$x(n) = A \cos(\omega n + \theta) = \frac{A}{2} e^{j(\omega n + \theta)} + \frac{A}{2} e^{-j(\omega n + \theta)}$$

Debido a que las señales sinusoidales de variable discreta con frecuencias separadas por un entero múltiplo de 2π son idénticas, entonces todas las frecuencias en un intervalo $[\omega_0, \omega_0 + 2\pi]$ representan todas las frecuencias existentes para estas señales. En otras palabras, el rango de frecuencias distinguibles para sinusoides de variable discreta es finito con un ancho de banda de 2π . Usualmente se utilizan los rangos de frecuencias angulares $\omega \in [-\pi, \pi]$ ($f \in [-\frac{1}{2}, \frac{1}{2}]$) o $\omega \in [0, 2\pi]$ ($f \in [0, 1]$) y reciben el nombre de *rango fundamental*.

2.2.3 Exponenciales complejos relacionados armónicamente

Las señales de variable continua

$$s_k(t) = e^{jk\Omega_0 t} = e^{j2\pi k F_0 t} \quad k = 0, \pm 1, \pm 2, \dots$$

se denominan señales exponenciales relacionadas armónicamente. El periodo fundamental de la señal $s_k(t)$ es $1/(kF_0) = T_p/k$, lo que equivale a una frecuencia fundamental de kF_0 . Puesto que una señal periódica con periodo T_p/k es también periódica con periodo $k(T_p/k) = T_p$ con $k \in \mathbb{Z}$ entonces todas las señales $s_k(t)$ tienen como periodo común T_p . El nombre proviene de la relación de estas funciones con las oscilaciones de una cuerda, que tienen relaciones “armónicas” (mantienen los nodos externos fijos) cuando las frecuencias son múltiplos enteros de una frecuencia fundamental.

En este caso de variable continua, para $k_1 \neq k_2$ se cumple siempre que $s_{k_1}(t) \neq s_{k_2}(t)$, o en otros términos, existe un número infinito de señales complejas relacionadas armónicamente con $e^{j2\pi F_0 t}$.

Para el caso de señales exponenciales complejas de variable discreta, puesto que son señales periódicas solo si su frecuencia es racional, se escoge $f_0 = 1/N$ y se define el conjunto de señales relacionadas armónicamente como

$$s_k(n) = e^{jk\omega_0 n} = e^{j2\pi k f_0 n} \quad k = 0, \pm 1, \pm 2, \dots \quad (2.17)$$

A diferencia del caso continuo se tiene para $k_1 = k + N$

$$s_{k+N}(n) = e^{j2\pi \frac{(k+N)}{N} n} = e^{j2\pi \frac{kn}{N}} e^{j2\pi n} = e^{j2\pi \frac{kn}{N}} = s_k(n)$$

Esto significa que en el caso discreto solo existen N señales complejas relacionadas armónicamente, donde todos los miembros del conjunto descrito en (2.17) tienen como periodo común N muestras.

2.3 Sistemas en tiempo discreto

A los dispositivos que operan sobre señales de variable discreta (o tiempo discreto) se les denomina *sistemas discretos*. En general, reciben una señal de entrada $x(n)$ para producir una señal de salida $y(n)$. Se dice que el sistema transforma $x(n)$ en $y(n)$, lo que se expresa como

$$y(n) = \mathcal{T}[x(n)]$$

donde $\mathcal{T}[\cdot]$ representa al operador de transformación o procesamiento realizado por el sistema sobre $x(n)$ para producir $y(n)$.

2.3.1 Descripción entrada-salida de sistemas

La descripción de entrada-salida define la relación entre $x(n)$ y $y(n)$. La estructura interna del sistema es desconocida o ignorada, es decir, el sistema se considera como una *caja negra* cuyo funcionamiento interno no interesa, sino el comportamiento específico ante cierta entrada (figura 2.7).

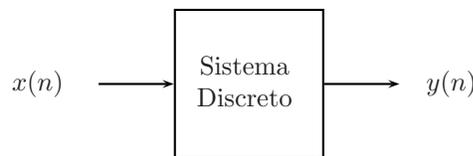


Figura 2.7: Entrada-salida de un sistema discreto

Ejemplo 2.5 Determine la salida de los siguientes sistemas para la entrada

$$x(n) = \begin{cases} 3 - |n| & \text{para } -2 \leq n \leq 2 \\ 0 & \text{en el resto} \end{cases}$$

1. $y(n) = x(n)$
2. $y(n) = x(n - 2)$
3. $y(n) = x(n + 1)$
4. $y(n) = \frac{1}{3} [x(n + 1) + x(n) + x(n - 1)]$
5. $y(n) = \max \{x(n + 1), x(n), x(n - 1)\}$
6. $y(n) = \sum_{k=-\infty}^n x(k)$

Solución:

1. Al sistema $y(n) = x(n)$ se le denomina *identidad*, pues su salida es idéntica a la entrada: $y(n) = \{1, 2, 3, 2, 1\}$
2. El sistema $y(n) = x(n - 2)$ retarda la entrada dos muestras: $y(n) = \{1, 2, 3, 2, 1\}$.
3. El sistema $y(n) = x(n + 1)$ adelanta la señal una unidad y solo puede ser realizado fuera de línea, por ser imposible en un sistema de tiempo real determinar el valor de una muestra en el futuro: $y(n) = \{1, 2, 3, 2, 1\}$.

4. El filtro de media móvil $y(n) = \frac{1}{3} [x(n+1) + x(n) + x(n-1)]$ calcula el promedio de tres muestras: $y(n) = \{1/3, 1, 2, 7/3, 2, 1, 1/3\}$.
5. El filtro de rango $y(n) = \max \{x(n+1), x(n), x(n-1)\}$ entrega el valor máximo de la muestra actual, la anterior y la futura: $y(n) = \{1, 2, 3, 3, 3, 2, 1\}$. Este filtro puede considerarse como filtro paso bajos.
6. El acumulador $y(n) = \sum_{k=-\infty}^n x(k)$ realiza la “integración” discreta de la entrada: $y(n) = \{1, 3, 6, 8, 9, 9, \dots\}$. Nóte que el acumulador puede reescribirse como

$$y(n) = \sum_{k=-\infty}^n x(k) = \underbrace{\sum_{k=-\infty}^{n-1} x(k)}_{y(n-1)} + x(n) = y(n-1) + x(n)$$

2.5

En general, la salida $y(n)$ en el instante n puede depender no solo de la muestra actual $x(n)$, sino también de la señal en instantes anteriores y posteriores a n . Además, la salida de un sistema puede depender de un estado interno. Por ejemplo, en el acumulador $y(n) = y(n-1) + x(n)$ si una secuencia de entrada se aplica en dos instantes de tiempo distintos, las dos reacciones del sistema difieren, dependiendo de la historia anterior del sistema “ $y(n-1)$ ”. Para determinar la salida del acumulador en un instante n_0 es necesario conocer $y(n_0 - 1)$. El cálculo de la secuencia de salida $y(n)$ para todo instante $n > n_0$ tiene como *condición inicial* al valor $y(n_0 - 1)$, que en cierta forma resume todo el pasado del sistema.

Si todas las condiciones iniciales son cero se dice que el sistema estaba en *reposo*. Siempre se asume que en $n = -\infty$ todo sistema estuvo en reposo. La salida de un sistema en reposo puede expresarse entonces utilizando únicamente la señal de entrada, puesto que por definición todas las salidas anteriores son cero.

En la literatura se encuentra el término *condiciones de reposo iniciales* (IRC, *Initial Rest Conditions*) para indicar que se asume $y(n) = 0$ para $n < n_0$ donde la entrada deja de ser cero justo en n_0 y $x(n) = 0$ para $n < n_0$. Además, se utiliza el concepto de condiciones de reposo *finales* (FRC, *Final Rest Conditions*) para indicar el caso contrario, es decir, se asume que $y(n) = 0$ para $n > n_0$, y la entrada se hace cero justo después de n_0 , es decir, $x(n) = 0$ para $n > n_0$.

Ejemplo 2.6 Determine la salida del sistema acumulador para la entrada $x(n) = nu(n)$ con condición inicial $y(-1) = \alpha$.

Solución:

$$y(n) = \sum_{k=-\infty}^n x(k) = \underbrace{\sum_{k=-\infty}^{-1} x(k)}_{y(-1)=\alpha} + \underbrace{\sum_{k=0}^n k}_{\frac{n(n+1)}{2}} = \alpha + \frac{n(n+1)}{2}$$

Recuerde que

$$\begin{aligned} \sum_{k=0}^n k &= 1 + 2 + 3 + \dots + n \\ \sum_{k=0}^n k &= n + n-1 + n-2 + \dots + 1 \\ \hline 2 \sum_{k=0}^n k &= n+1 + n+1 + n+1 + \dots + n+1 \\ 2 \sum_{k=0}^n k &= n(n+1) \\ \sum_{k=0}^n k &= \frac{n(n+1)}{2} \end{aligned}$$

2.6

2.3.2 Diagramas de bloques

En el capítulo anterior se indicó que un sistema está conformado por subsistemas, es decir, bloques de procesamiento de señal con tareas específicas dentro del sistema total. Este concepto se aplica recursivamente, de tal modo que un subsistema se descompone a su vez en otros subsistemas más sencillos, que a su vez pueden estar constituidos por otros subsistemas, y así sucesivamente, hasta llegar a ciertos componentes elementales descritos a continuación.

Sumador

El sumador es un bloque sin memoria, que se representa como lo indica la figura 2.8. Tal y

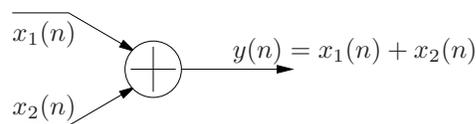


Figura 2.8: Diagrama de un sumador.

como lo indica su nombre, su tarea es sumar dos señales o más señales muestra por muestra.

Multiplicador por constante

El multiplicador por constante es un bloque sin memoria, que se representa como lo indica la figura 2.9. Se utiliza para escalar toda una secuencia por un mismo factor.



Figura 2.9: Diagrama de un multiplicador por constante.

Multiplicador de señal

El multiplicador de señal es un bloque sin memoria, que se representa como lo indica la figura 2.10, y que genera una nueva secuencia a partir de los productos entre muestras de

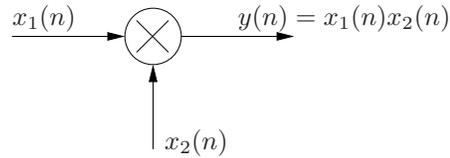


Figura 2.10: Diagrama de un multiplicador de señales.

las señales de entrada correspondientes a un mismo instante n .

Retardador de un elemento

El retardador es un bloque con memoria de longitud 1, representado como lo indica la figura 2.11. Es uno de los elementos de procesamiento digital más utilizados. Su símbolo está

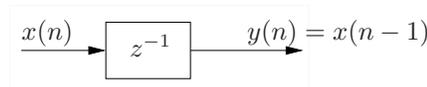


Figura 2.11: Diagrama de elemento retardador.

muy relacionado con las propiedades de la transformada z que se analizarán posteriormente.

Adelantador de un elemento

El adelantador no es realizable físicamente y solo existe en sistemas de tratamiento de señales fuera de línea. Se representa como lo indica la figura 2.12.

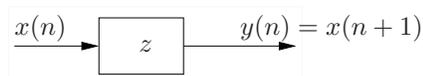


Figura 2.12: Diagrama de elemento adelantador.

Ejemplo 2.7 Realice el diagrama de bloques para

$$y(n) = \frac{1}{4}y(n-1) + \frac{1}{2}x(n) + \frac{1}{2}x(n-1)$$

Solución:

Nótese primero que esta expresión puede reescribirse de la siguiente forma:

$$y(n) = \frac{1}{4}y(n-1) + \frac{1}{2}(x(n) + x(n-1)) \quad (2.18)$$

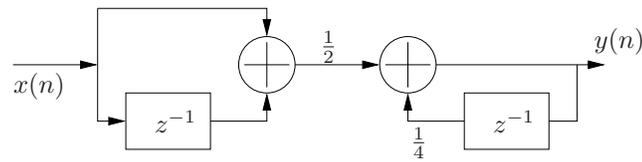


Figura 2.13: Diagrama de bloques de la ecuación 2.18.

con lo que se deriva fácilmente el diagrama mostrado en la figura 2.13.

2.7

2.3.3 Clasificación de los sistemas discretos

Un sistema se dice tener una determinada propiedad, si dicha propiedad se cumple para *todas* las señales de entrada posibles. Un contraejemplo basta para descartar que un sistema posea una determinada propiedad. La tabla 2.2 resume las propiedades de sistemas discretos, que se detallan a continuación.

Tabla 2.2: Propiedades de sistemas discretos.

Propiedad	Clasificación	
Memoria	Sistema estático	Sistema dinámico
Varianza	Sistema variante en el tiempo	Sistema invariante en el tiempo
Linealidad	Sistema lineal	Sistema no lineal
Causalidad	Sistema causal	Sistema no causal
Estabilidad	Sistema estable	Sistema inestable

Sistemas estáticos y dinámicos

En un sistema estático, o sin memoria, la salida $y(n)$ depende solo de la entrada $x(n)$ en el mismo instante n , y no de las entradas pasadas o futuras. Todos los otros casos son sistemas dinámicos.

Si la salida $y(n)$ depende de las entradas de $x(n - N)$ a $x(n)$ se dice que el sistema tiene *memoria* de duración N , donde N puede ser finita o infinita.

Por ejemplo, $y(n) = ax(n) + nx^2(n) + bx^3(n)$ representa un sistema estático, mientras que $y(n) = \sum_{k=0}^n a_k x(n - k)$ es un sistema dinámico.

Sistemas variantes e invariantes en el tiempo

Un sistema \mathcal{T} en reposo es invariante en el tiempo o invariante al desplazamiento si y solo si

$$x(n) \xrightarrow{\mathcal{T}} y(n) \quad \Rightarrow \quad x(n - k) \xrightarrow{\mathcal{T}} y(n - k)$$

En el diagrama de bloques de la figura 2.14 el sistema \mathcal{T} es invariante en el tiempo si y solo si la salida $d(n)$ es cero para todas las entradas $x(n)$ y para todos los valores de retardo k .

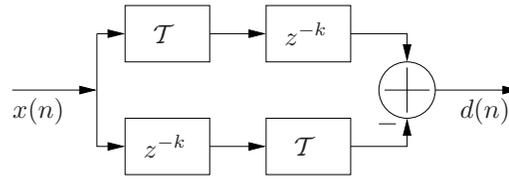


Figura 2.14: Verificación de la invarianza en el tiempo.

Ejemplo 2.8 Determine si los siguientes sistemas son o no invariantes en el tiempo:

1. $y(n) = x(n) - x(n - 1)$
2. $y(n) = x(n) \cos(\omega_0 n)$

Solución:

El sistema $y(n) = x(n) - x(n - 1)$ es invariante en el tiempo. Para demostrarlo se calcula la respuesta del sistema en reposo a la entrada desplazada $x(n - k)$, que resulta en $y_k(n) = x(n - k) - x(n - k - 1)$. La respuesta a $x(n)$, retrasada k muestras es $y(n - k) = x(n - k) - x(n - k - 1)$. Como $y(n - k) = y_k(n)$ el sistema es invariante en el tiempo.

El sistema modulador $y(n) = x(n) \cos(\omega_0 n)$ es variante en el tiempo, puesto que su respuesta $y_k(n)$ a $x(n - k)$ es $y_k(n) = x(n - k) \cos(\omega_0 n)$, y la repuesta a $x(n)$, retardada k muestras es $y(n - k) = x(n - k) \cos(\omega_0(n - k))$ que es diferente a $y_k(n)$. 2.8

Sistemas lineales y no lineales

Un sistema es lineal si satisface el teorema de superposición, es decir, para cualesquiera dos constantes a_1, a_2 y para toda señal $x_1(n)$ y $x_2(n)$ se cumple

$$\mathcal{T} [a_1 x_1(n) + a_2 x_2(n)] = a_1 \mathcal{T} [x_1(n)] + a_2 \mathcal{T} [x_2(n)]. \quad (2.19)$$

La figura 2.15 ilustra un esquema de verificación de linealidad, donde el sistema es lineal si y solo si para todo par de entradas $x_1(n)$ y $x_2(n)$ y para cualesquiera dos constantes a_1, a_2 siempre la salida $d(n)$ es cero.

Como consecuencia de (2.19) todo sistema lineal cumple la propiedad multiplicativa o de escalado

$$\mathcal{T} [a_1 x_1(n)] = a_1 \mathcal{T} [x_1(n)] \quad (2.20)$$

y la propiedad aditiva

$$\mathcal{T} [x_1(n) + x_2(n)] = \mathcal{T} [x_1(n)] + \mathcal{T} [x_2(n)].$$

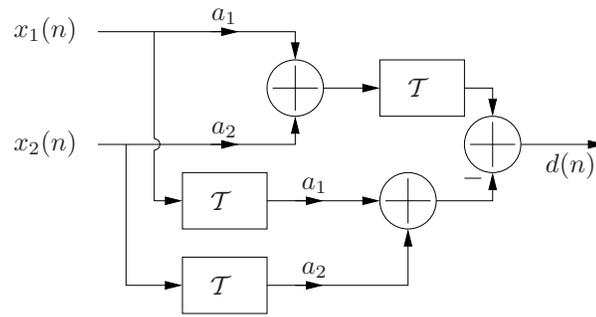


Figura 2.15: Verificación de linealidad.

El principio de superposición puede generalizarse para M señales como

$$x(n) = \sum_{k=1}^M a_k x_k(n) \xrightarrow{\mathcal{T}} y(n) = \sum_{k=1}^M a_k \mathcal{T}[x_k(n)]$$

De la propiedad de escalado se deduce además que un sistema lineal en reposo con entrada cero ($x(n) = 0$, $n \in \mathbb{Z}$) debe tener como salida una señal nula $y(n) = 0$.

Si para un sistema la propiedad de superposición no se cumple, entonces el sistema se dice ser *no lineal*.

Ejemplo 2.9 Compruebe si los siguientes sistemas son lineales.

1. $y(n) = nx(n)$
2. $y(n) = x(n^2)$
3. $y(n) = x^2(n)$
4. $y(n) = Ax(n) + B$
5. $y(n) = e^{x(n)}$

Solución:

1. Para el sistema 1 se obtiene primero la respuesta del sistema a una entrada igual a la suma ponderada de dos señales $x_1(n)$ y $x_2(n)$, es decir, para una entrada *total* $x(n) = a_1x_1(n) + a_2x_2(n)$ y se obtiene $y_T(n) = n(a_1x_1(n) + a_2x_2(n)) = a_1nx_1(n) + a_2nx_2(n)$. Ahora, la suma ponderada de las salidas del sistema para $x_1(n)$ y $x_2(n)$ por separado es $y_S(n) = a_1y_1(n) + a_2y_2(n)$, con $y_1(n) = nx_1(n)$ y $y_2(n) = nx_2(n)$, lo que es igual a $y_S(n) = a_1nx_1(n) + a_2nx_2(n)$. Como $y_T(n) = y_S(n)$ se puede afirmar que el sistema $y(n) = nx(n)$ es lineal.
2. Para $y(n) = x(n^2)$ las salidas $y_T(n) = a_1x_1(n^2) + a_2x_2(n^2)$ y $y_S = a_1x_1(n^2) + a_2x_2(n^2)$ son idénticas y por tanto el sistema es lineal.
3. Para $y(n) = x^2(n)$ la salida $y_T(n) = (a_1x_1(n) + a_2x_2(n))^2 = a_1^2x_1^2(n) + a_2^2x_2^2(n) + 2a_1a_2x_1(n)x_2(n)$ y la salida $y_S(n) = a_1^2x_1^2(n) + a_2^2x_2^2(n)$ evidentemente son diferentes y por tanto el sistema no es lineal.
4. Para $y(n) = Ax(n) + B$ la salida $y_T(n) = A(a_1x_1(n) + a_2x_2(n)) + B$ y la salida $y_S(n) = Aa_1x_1(n) + B + Ax_2(n) + B = Aa_1x_1(n) + Ax_2(n) + 2B$ difieren para todo

$B \neq 0$ y por tanto el sistema, a pesar de su apariencia, no es lineal.

5. Para $y(n) = e^{x(n)}$ la salida $y_T = e^{a_1 x_1(n) + a_2 x_2(n)} = e^{a_1 x_1(n)} e^{a_2 x_2(n)}$ y la salida $y_S = e^{a_1 x_1(n)} + e^{a_2 x_2(n)}$ son diferentes y por tanto el sistema tampoco es lineal.

2.9

Sistemas causales y no causales

Un sistema es causal si $y(n)$ depende únicamente de las entradas presentes y pasadas ($x(n), x(n-1), x(n-2), \dots$), y salidas pasadas ($y(n-1), y(n-2), \dots$), pero no de las entradas o salidas futuras ($x(n+1), x(n+2), \dots; y(n+1), y(n+2), \dots$). En caso contrario, el sistema es no causal.

Sistemas que funcionan “en línea” deben ser causales por la imposibilidad de determinar el valor de la entrada o la salida en el futuro.

A una señal $x(n)$ que es cero para $n \geq 0$ y diferente de cero para $n < 0$ se le denomina señal *anticausal*. Las implicaciones de la causalidad junto con la linealidad e invarianza en el tiempo no son triviales, y se profundizará en ellas en los próximos capítulos.

Sistemas estables e inestables

Un sistema arbitrario en reposo se denomina estable de entrada acotada - salida acotada (BIBO: *bounded input - bounded output*) si toda entrada acotada produce una salida acotada:

$$|x(n)| \leq M_x < \infty \xrightarrow{\mathcal{T}} |y(n)| \leq M_y < \infty, \quad \forall n \in \mathbb{Z}$$

Basta con que alguna entrada acotada produzca una salida no acotada (es infinita), para que el sistema sea *inestable*.

2.3.4 Interconexión de sistemas discretos

Hay dos maneras fundamentales de interconectar sistemas discretos: interconexión en cascada (o serie) e interconexión paralela (figura 2.16). La interconexión en cascada se describe con sistemas de la forma:

$$y(n) = \mathcal{T}_2[\mathcal{T}_1[x(n)]] = \mathcal{T}_c[x(n)]$$

En general, el orden de los bloques para la conexión en cascada es relevante. Sin embargo, si los sistemas son lineales e invariantes en el tiempo entonces \mathcal{T}_c es a su vez lineal e invariante en el tiempo, y $\mathcal{T}_1[\mathcal{T}_2[\cdot]] = \mathcal{T}_2[\mathcal{T}_1[\cdot]]$. La interconexión en paralelo se describe por

$$y(n) = \mathcal{T}_1[x(n)] + \mathcal{T}_2[x(n)] = \mathcal{T}_p[x(n)].$$

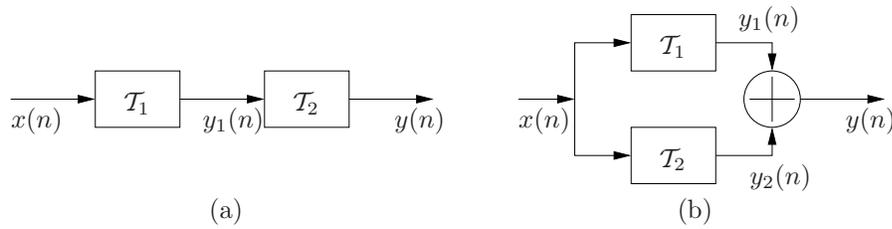


Figura 2.16: Interconexión de sistemas discretos. (a) Cascada. (b) Paralelo

2.4 Análisis de sistemas discretos lineales e invariantes en el tiempo

El análisis de sistemas se simplifica enormemente si estos son lineales e invariantes en el tiempo (LTI: *Linear and Time Invariant*).

2.4.1 Técnicas de análisis

Existen dos métodos básicos para el análisis del comportamiento de un sistema:

1. Descomposición de la señal de entrada en señales elementales, cuya respuesta es conocida.
2. Solución de la *ecuación de diferencias*.

La solución de la ecuación de diferencias se revisará en la sección 2.5.

El concepto fundamental del análisis de sistemas lineales por descomposición es el siguiente: supóngase que la entrada $x(n)$ puede expresarse como una suma ponderada de funciones elementales $\{x_k(n)\}$

$$x(n) = \sum_k c_k x_k(n)$$

donde c_k son los coeficientes de ponderación o *pesos* de la descomposición de la señal $x(n)$. Si la respuesta del sistema en reposo a $x_k(n)$ es $y_k(n)$, es decir

$$y_k(n) = \mathcal{T}[x_k(n)]$$

entonces con la propiedad de linealidad se obtiene

$$y(n) = \mathcal{T}[x(n)] = \mathcal{T}\left[\sum_k c_k x_k(n)\right] = \sum_k c_k \mathcal{T}[x_k(n)] = \sum_k c_k y_k(n)$$

En otras palabras, si el sistema es lineal, la respuesta $y(n)$ del sistema a una entrada $x(n)$ es igual a la suma ponderada de las repuestas $y_k(n)$ a cada una de las componentes $x_k(n)$ en que se puede descomponer $x(n)$, donde además los coeficientes c_k de ponderación de las salidas corresponden a los coeficientes de ponderación de la entrada.

En esta descomposición, la elección de las funciones elementales dependerá de las características de las señales a evaluar. Dos clases de funciones son usuales: impulsos ($\delta(n - k)$) y funciones exponenciales complejas $e^{j\omega_k n}$, esta última utilizándose frecuentemente en el llamado análisis frecuencial (ver capítulo 4). En los últimos años ha cobrado fuerza el uso de otras familias de funciones elementales caracterizadas a partir de la teoría de *wavelets*, tema que sin embargo excede el marco de este curso.

2.4.2 Descomposición de una señal en impulsos

Utilizando impulsos desplazados como funciones elementales puede expresarse cualquier señal $x(n)$ como:

$$x(n) = \sum_{k=-\infty}^{\infty} x(k)\delta(n - k)$$

Ejemplo 2.10 Descomponga la señal

$$x(n) = \{0, 1, 2, -1, -1/2, 1\}$$

en sus impulsos.

Solución:

Esta señal puede expresarse como

$$x(n) = 1 \cdot \delta(n - 1) + 2 \cdot \delta(n - 2) - 1 \cdot \delta(n - 3) - \frac{1}{2} \cdot \delta(n - 4) + 1 \cdot \delta(n - 5)$$

2.10

2.4.3 Convolución

Si se denota con $h'(n, k)$ la respuesta del sistema a un impulso desplazado k unidades $\delta(n - k)$

$$h'(n, k) = \mathcal{T} [\delta(n - k)]$$

entonces la salida de un sistema lineal puede calcularse con las repuestas elementales a dichos impulsos desplazados:

$$y(n) = \sum_k c_k y_k(n) = \sum_k x(k)h'(n, k)$$

donde se han utilizado $x_k(n) = \delta(n - k)$ como entradas elementales y por lo tanto los coeficientes $c_k = x(k)$.

Si el sistema es además invariante en el tiempo, entonces con $h(n) = \mathcal{T} [\delta(n)]$ se tiene que $h'(n, k) = h(n - k)$ y por lo tanto

$$y(n) = \sum_{k=-\infty}^{\infty} x(k)h(n - k) = x(n) * h(n)$$

que se denomina sumatoria de *convolución*. Se dice entonces que la respuesta del sistema LTI $y(n)$ a una entrada $x(n)$ es igual a la convolución de $x(n)$ con la respuesta al impulso $h(n)$.

Esto quiere decir que en un sistema LTI en reposo su respuesta a cualquier entrada puede determinarse con solo conocer dicha entrada y la respuesta al impulso $h(n)$.

El cálculo de la convolución involucra cuatro pasos:

1. *Reflexión* de $h(k)$ con respecto a $k = 0$ para producir $h(-k)$.
2. *Desplazamiento* del origen de $h(-k)$ hacia el punto n que se desea calcular.
3. *Multiplicación* de $x(k)$ y $h(n-k)$ para obtener la secuencia producto $v_n(k) = x(k)h(n-k)$.
4. *Suma* de todos los valores de $v_n(k)$ para obtener $y(n)$.

Los pasos del 2 al 4 deben realizarse para todo instante n que se deseé calcular.

Ejemplo 2.11 Determine la respuesta a la señal de entrada

$$x(n) = \{1, 2, 3, 1\}$$

de un sistema LTI con respuesta al impulso

$$h(n) = \{1, 2, 1, -1\}$$

Solución:

Siguiendo el procedimiento indicado, primero se calcula la reflexión de la respuesta al impulso

$$h(-k) = \{-1, 1, 2, 1\}.$$

Los pasos siguientes se resumen en la tabla 2.3.

Con lo que resulta la señal de salida en

$$y(n) = \{1, 4, 8, 8, 3, -2, -1\}$$

2.11

Ejemplo 2.12 Para el caso en que la entrada $x(n)$ y la respuesta impulsional $h(n)$ sean de longitud finita, calcule la longitud de la salida en términos de las longitudes de $x(n)$ y $h(n)$.

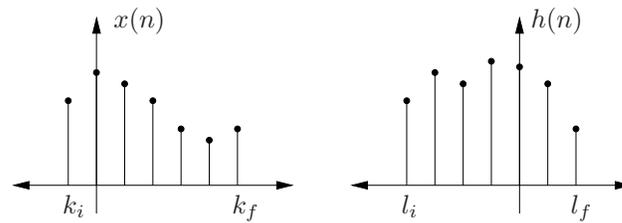
Solución:

Una señal finita tiene longitud N si el intervalo más pequeño con muestras diferentes de cero tiene N muestras.

Asuma que los índices menor y mayor de la entrada $x(n)$ son k_i y k_f respectivamente, y los de la respuesta impulsional $h(n)$ son l_i y l_f (figura 2.17).

Tabla 2.3: Ejemplo de convolución de dos secuencias finitas.

2. Desplazamiento	3. Multiplicación por $x(k) = \{1, 2, 3, 1\}$	4. Suma
$h(-1 - k) = \{-1, 1, 2, 1\}$	$v_{-1} = \{0, 0, 0, 1, 0, 0, 0\}$	$y_{-1} = 1$
$h(0 - k) = \{-1, 1, 2, 1\}$	$v_0 = \{0, 0, 2, 2, 0, 0\}$	$y_0 = 4$
$h(1 - k) = \{-1, 1, 2, 1\}$	$v_1 = \{0, 1, 4, 3, 0\}$	$y_1 = 8$
$h(2 - k) = \{-1, 1, 2, 1\}$	$v_2 = \{-1, 2, 6, 1\}$	$y_2 = 8$
$h(3 - k) = \{0, -1, 1, 2, 1\}$	$v_3 = \{0, -2, 3, 2\}$	$y_3 = 3$
$h(4 - k) = \{0, 0, -1, 1, 2, 1\}$	$v_4 = \{0, 0, -3, 1\}$	$y_4 = -2$
$h(5 - k) = \{0, 0, 0, -1, 1, 2, 1\}$	$v_5 = \{0, 0, 0, -1\}$	$y_5 = -1$

**Figura 2.17:** Índices de las señales de entrada y respuesta impulsional.

La longitud de $x(n)$ es entonces $N_x = k_f - k_i + 1$, y la longitud de la respuesta impulsional es $N_h = l_f - l_i + 1$.

El desplazamiento menor n_{min} para el que puede haber salida es (figura 2.18):

$$n_{min} = k_i + l_i$$

El desplazamiento mayor n_{max} para el que puede haber salida es (figura 2.19):

$$n_{max} = k_f + l_f$$

La longitud de la señal de salida es entonces

$$\begin{aligned}
 N_y &= n_{max} - n_{min} + 1 \\
 &= (k_f + l_f) - (k_i + l_i) + 1 + (1 - 1) \\
 &= (k_f - k_i + 1) + (l_f - l_i + 1) - 1 \\
 &= N_x + N_h - 1
 \end{aligned}$$

es decir, el resultado de la convolución tiene una longitud que es tan solo una muestra menor que la suma de las longitudes de las dos señales sobre las que ella opera. 2.12

En el ejemplo anterior se observa que para los valores n_{min} y n_{max} , es decir, los límites del intervalo de muestras donde existe salida no nula, es irrelevante qué función representa la

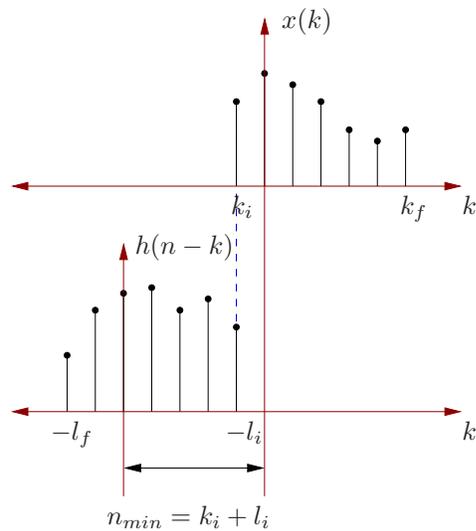


Figura 2.18: Menor desplazamiento para el que puede haber salida.

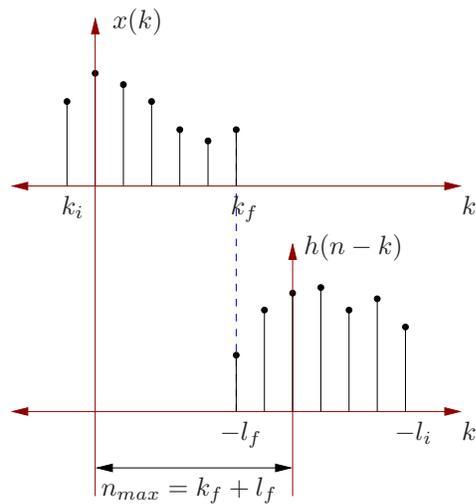


Figura 2.19: Mayor desplazamiento para el que puede haber salida.

entrada y qué función representa la respuesta al impulso del sistema. Esto indica cierta simetría en la convolución. Con un simple cambio de variables es posible demostrar que la convolución es de hecho totalmente conmutativa, y no solo el valor de sus índices:

$$\begin{aligned}
 y(n) &= x(n) * h(n) = \sum_{k=-\infty}^{\infty} x(k)h(n-k) \underset{m=n-k}{=} \sum_{m=-\infty}^{\infty} x(n-m)h(m) \\
 &= \sum_{k=-\infty}^{\infty} h(k)x(n-k) \\
 &= h(n) * x(n)
 \end{aligned}$$

Ejemplo 2.13 Encuentre a través de la convolución la respuesta de un sistema con respuesta exponencial al impulso:

$$h(n) = a^n u(n), \quad |a| < 1.$$

ante el escalón unitario $u(n)$.

Solución:

Obsérvese que $a^n = e^{n \ln a}$, con lo que se establece la relación con los sistemas analógicos de respuesta exponencial $h_a(t) = e^{-t/\tau}$ (figura 2.20). Para encontrar más detalles de esta relación asúmase que la señal $h_a(t)$ se discretiza en el tiempo con un intervalo de muestreo T :

$$h_a(nT) = e^{-nT/\tau} \stackrel{!}{=} e^{n \ln a} \Rightarrow \ln a = -\frac{T}{\tau} \Rightarrow a = e^{-T/\tau}$$

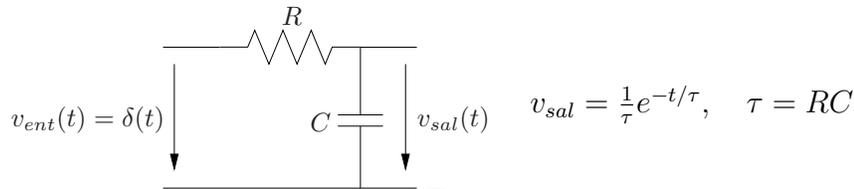


Figura 2.20: Circuito analógico con respuesta exponencial al impulso.

Puede observarse que este tipo de sistemas analógicos solo permiten $a \geq 0$, mientras que los sistemas discretos permiten además $a < 0$ sin ninguna complejidad computacional adicional.

Para determinar la salida $y(n)$ del sistema con la entrada escalón unitario $u(n)$ se utiliza la sumatoria de convolución:

$$y(n) = \sum_{k=-\infty}^{\infty} x(n-k)h(k) = \sum_{k=-\infty}^{\infty} u(n-k)h(k) \quad (2.21)$$

Puesto que las secuencias producto son 0 para $n < 0$ entonces $y(n) = 0$ para $n < 0$. Evaluando entonces (2.21) para varios n se obtiene:

$$\begin{aligned}
 y(0) &= h(0) = 1 \\
 y(1) &= h(0) + h(1) = 1 + a \\
 y(2) &= h(0) + h(1) + h(2) = 1 + a + a^2 \\
 &\vdots \\
 y(n) &= \sum_{k=0}^n h(k) = \sum_{k=0}^n a^k
 \end{aligned}$$

Recordando

$$\begin{array}{rcl}
 \sum_{k=0}^n a^k & = & 1 + a + a^2 + \dots + a^n \\
 a \sum_{k=0}^n a^k & = & a + a^2 + \dots + a^n + a^{n+1} \\
 \hline
 (1-a) \sum_{k=0}^n a^k & = & 1 - a^{n+1}
 \end{array}$$

con lo que se deriva

$$\sum_{k=0}^n a^k = \frac{1 - a^{n+1}}{1 - a}$$

Si $|a| < 1$ entonces $\lim_{n \rightarrow \infty} a^{n+1} = 0$ lo que implica que $y(\infty) = \frac{1}{1-a}$. La figura 2.21 muestra un ejemplo de la respuesta para $a = 0,9$.

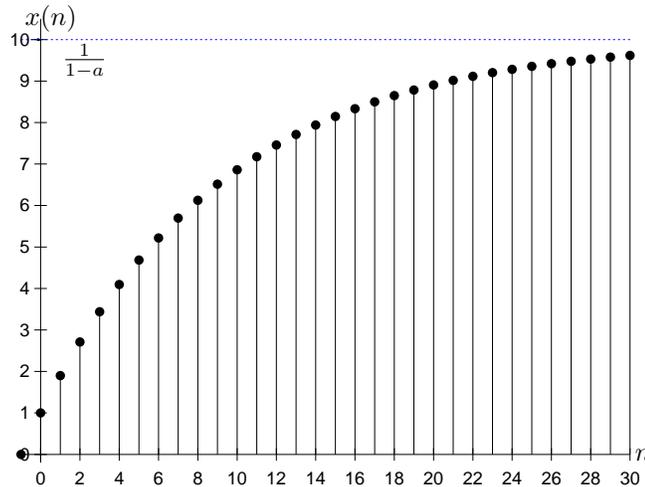


Figura 2.21: Respuesta al sistema exponencial con $a = 0,9$.

Nótese la similitud con la respuesta del sistema analógico al escalón continuo de la forma $k(1 - e^{-t/\tau})$. Más tarde en el curso se retomarán estas similitudes. 2.13

2.4.4 Propiedades de la convolución

Conmutatividad

$$y(n) = x(n) * h(n) = h(n) * x(n)$$



Figura 2.22: Representación de la propiedad de la convolución de conmutatividad.

Asociatividad

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



Figura 2.23: Representación de la propiedad de la convolución de asociatividad.

Distributividad

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

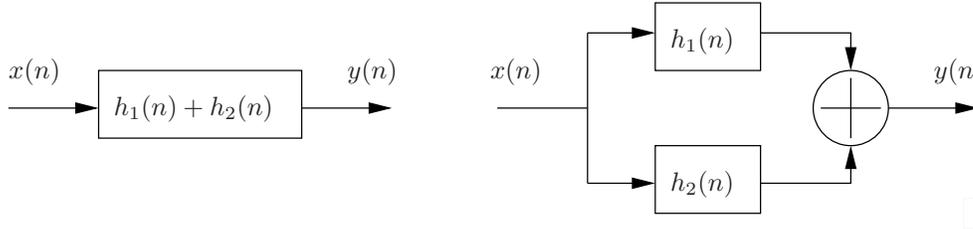


Figura 2.24: Representación de la propiedad de la convolución de distributividad.

Las tres propiedades pueden generalizarse para más de dos subsistemas.

2.4.5 Sistemas LTI causales

En un sistema causal la salida en $n = n_0$ depende solo de valores de entrada $x(n)$ para $n \leq n_0$, es decir, de valores pasados. Como

$$y(n_0) = \sum_{k=-\infty}^{\infty} h(k)x(n_0 - k) = \underbrace{\sum_{k=0}^{\infty} h(k)x(n_0 - k)}_{\substack{\text{Muestras} \\ \text{pasadas y} \\ \text{actual}}} + \underbrace{\sum_{k=-\infty}^{-1} h(k)x(n_0 - k)}_{\substack{\text{Muestras} \\ \text{futuras}}}$$

se deriva que $h(k) = 0$ para todo $k \leq -1$, pues solo así la salida podrá ser independiente de entradas futuras. Dado que $h(n)$ es la respuesta impulsional de un sistema LTI en reposo, $h(n) = 0$ para $n < 0$ es condición necesaria y suficiente para la causalidad. Un sistema es causal entonces si y solo si $h(n) = 0, \forall n < 0$.

Si un sistema es causal entonces la convolución puede simplificarse en

$$y(n) = \sum_{k=0}^{\infty} h(k)x(n - k) = \sum_{k=-\infty}^n x(k)h(n - k)$$

Generalizando, a una secuencia $x(n)$ con $x(n) \neq 0$ para algún $n < 0$ se le denomina secuencia no causal, y de lo contrario, secuencia causal. A secuencias con $x(n) = 0$ para $n \geq 0$ y con $x(n) \neq 0$ para alguna muestra con $n < 0$ se les denomina secuencias *anticausales*.

Si tanto la entrada $x(n)$ como la respuesta impulsional son causales, entonces la convolución se simplifica en:

$$y(n) = \sum_{k=0}^n h(k)x(n-k) = \sum_{k=0}^n x(k)h(n-k)$$

Nótese que esta respuesta es a su vez causal, es decir, $y(n) = 0$ para todo $n < 0$.

2.4.6 Estabilidad de sistemas lineales e invariantes en el tiempo

Un sistema se dice estable si para toda entrada acotada su salida es también acotada:

$$|x(n)| \leq M_x < \infty \quad \rightarrow \quad |y(n)| \leq M_y < \infty \quad \forall n$$

Dada la convolución

$$y(n) = \sum_{k=-\infty}^{\infty} h(k)x(n-k)$$

y su valor absoluto

$$|y(n)| = \left| \sum_{k=-\infty}^{\infty} h(k)x(n-k) \right| \leq \sum_{k=-\infty}^{\infty} |h(k)||x(n-k)| \leq \sum_{k=-\infty}^{\infty} |h(k)|M_x$$

$$|y(n)| \leq M_x \sum_{k=-\infty}^{\infty} |h(k)|$$

lo que implica que $|y(n)|$ es acotada solo si

$$S_h = \sum_{k=-\infty}^{\infty} |h(k)| \leq \infty$$

En consecuencia un sistema LTI es estable si su respuesta impulsional es absolutamente sumable. Esta condición es necesaria y suficiente.

Si $h(n)$ no es absolutamente sumable entonces, derivado de lo anterior, debería existir al menos una entrada que produce una salida no acotada y por tanto el sistema es inestable. Si para un sistema con respuesta al impulso $h(n)$ se elige como entrada la señal

$$x(n) = \begin{cases} \frac{h^*(-n)}{|h(-n)|} & \text{para } h(-n) \neq 0 \\ 0 & \text{para } h(-n) = 0 \end{cases}$$

basta entonces con demostrar que existe algún valor de n para el que $y(n)$ no está acotada para concluir inestabilidad. Por ejemplo, calcúlese $y(0)$

$$y(0) = \sum_{k=-\infty}^{\infty} x(-k)h(k) = \sum_{k=-\infty}^{\infty} \frac{|h(k)|^2}{|h(k)|} = \sum_{k=-\infty}^{\infty} |h(k)| = S_h$$

Quiere decir que si $h(n)$ no es absolutamente sumable ($S_h \rightarrow \infty$) entonces la salida no es acotada y el sistema es inestable.

La condición $\sum_{k=-\infty}^{\infty} |h(k)| < \infty$ implica que $h(n)$ tiende a cero cuando $n \rightarrow \infty$, lo que a su vez implica que si la entrada $x(n)$ es cero para todo $n > n_0$ entonces la salida tiende a cero para $n \rightarrow \infty$:

$$y(n_0 + N) = \underbrace{\sum_{k=-\infty}^{N-1} h(k)x(n_0 + N - k)}_{=0 \quad (x(n)=0, n>n_0)} + \sum_{k=N}^{\infty} h(k)x(n + N - k)$$

$$\Rightarrow |y(n_0 + N)| = \left| \sum_{k=N}^{\infty} h(k)x(n + N - k) \right| \leq \sum_{k=N}^{\infty} |h(k)| \underbrace{|x(n + N - k)|}_{\leq M_x} \leq M_x \sum_{k=N}^{\infty} |h(k)|$$

Si $N \rightarrow \infty$ entonces $\lim_{N \rightarrow \infty} \sum_{k=N}^{\infty} |h(k)| = 0$ y de ahí que $\lim_{N \rightarrow \infty} y(n_0 + N) = 0$.

Esto implica que en un sistema estable cualquier excitación de duración finita a la entrada produce una respuesta transitoria, es decir, una respuesta cuya amplitud decrece y se anula con el tiempo.

Ejemplo 2.14 Determine el rango del parámetro a para el que el sistema LTI de respuesta al impulso $h(n) = a^n u(n)$ es estable.

Solución:

El sistema es estable si

$$\sum_{k=0}^{\infty} |a^k| = \sum_{k=0}^{\infty} |a|^k = 1 + |a| + |a|^2 + \dots$$

converge. Esto ocurre si y solo si $|a| < 1$ y esta suma converge a

$$\sum_{k=0}^{\infty} |a^k| = \frac{1}{1 - |a|}$$

2.14

Ejemplo 2.15 Determine el rango de valores de a y b para los cuales el sistema LTI de respuesta

$$h(n) = \begin{cases} a^n & n \geq 0 \\ b^n & n < 0 \end{cases}$$

es estable.

Solución:

La condición de estabilidad es

$$\sum_{n=-\infty}^{\infty} |h(n)| = \sum_{n=-\infty}^{-1} |b|^n + \sum_{n=0}^{\infty} |a|^n = \underbrace{\sum_{n=1}^{\infty} \left| \frac{1}{b} \right|^n}_{\substack{\text{Converge} \\ \text{si } |b| > 1}} + \underbrace{\sum_{n=0}^{\infty} |a|^n}_{\substack{\text{Converge} \\ \text{si } |a| < 1}} = \frac{|b|}{|b| - 1} + \frac{1}{1 - |a|}$$

El sistema es estable si $|a| < 1$ y si $|b| > 1$.

2.15

2.4.7 Sistemas de respuesta finita e infinita

Es útil distinguir los sistemas entre aquellos con una respuesta finita al impulso (FIR: *Finite Impulse Response*) y aquellos con respuesta infinita al impulso (IIR: *Infinite Impulse Response*) por las consideraciones prácticas que los distintos conceptos conllevan: la implementación de los primeros se puede realizar directamente, mientras que algunos de los segundos pueden realizarse por medio de ecuaciones de diferencias.

Para los sistemas causales FIR la convolución se reduce a

$$y(n) = \sum_{k=0}^{M-1} h(k)x(n-k), \quad h(k) = 0, \quad k < 0 \wedge k \geq M \quad (2.22)$$

El sistema se comporta como una “ventana” que solo permite ver M muestras para calcular la salida. Se dice que el sistema FIR tiene memoria finita de M muestras.

2.5 Sistemas discretos descritos mediante ecuaciones de diferencias

El cálculo de la convolución:

$$y(n) = \sum_{k=-\infty}^{\infty} h(k)x(n-k)$$

sólo es posible para sistemas FIR, puesto que en el caso de sistemas IIR se requeriría de una memoria infinita para almacenar $h(n)$, y un número infinito de multiplicaciones y adiciones para calcular una sola muestra de la salida.

Las llamadas *ecuaciones de diferencias* permiten trabajar con sistemas IIR.

2.5.1 Sistemas discretos recursivos y no recursivos

Un sistema es *recursivo* si su salida en el instante n depende de valores anteriores de la salida, $y(n-1), y(n-2), \dots$:

$$y(n) = F[y(n-1), y(n-2), \dots, y(n-N), x(n), x(n-1), \dots, x(n-M)]$$

donde $F[\cdot]$ denota una función cualquiera, y sus argumentos son las entradas y salidas presentes y pasadas.

El sistema es *no recursivo* si solo depende de las entradas presentes y pasadas, mas no de salidas pasadas:

$$y(n) = F[x(n), x(n-1), \dots, x(n-M)]$$

Nótese que los sistemas LTI FIR causales, cuya salida se expresa como la suma finita de convolución (2.22) no son recursivos.

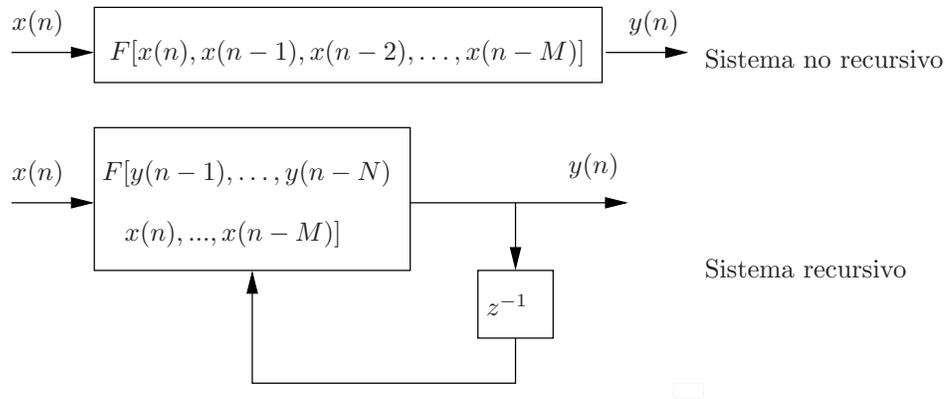


Figura 2.25: Diagrama de bloques de sistemas recursivos y no recursivos.

El lazo de realimentación da origen a una diferencia fundamental en el cálculo de la salida: La salida recursiva debe determinarse en orden, pues para calcular $y(n_0)$ se necesitan todos los N valores anteriores, mientras que en un sistema no recursivo las salidas anteriores no son necesarias y se puede calcular un valor directamente para cualquier n .

Ejemplo 2.16 Determine una expresión recursiva para el sistema de media acumulativa.

Solución:

El sistema de media acumulativa

$$y(n) = \frac{1}{n+1} \sum_{k=0}^n x(k)$$

es recursivo pues

$$\begin{aligned} (n+1)y(n) &= \sum_{k=0}^n x(k) \\ \Rightarrow ny(n-1) &= \sum_{k=0}^{n-1} x(k) \\ \Rightarrow (n+1)y(n) &= \sum_{k=0}^{n-1} x(k) + x(n) = ny(n-1) + x(n) \\ y(n) &= \frac{n}{n+1}y(n-1) + \frac{1}{n+1}x(n) \end{aligned}$$

$$= \frac{1}{n+1} (ny(n-1) + x(n)) \quad (2.23)$$

Nótese los coeficientes *no* constantes para la salida anterior $y(n-1)$ y la entrada actual $x(n)$, que son $\frac{n}{n+1}$ y $\frac{1}{n+1}$ respectivamente; estos implican que el sistema es variante en el tiempo. La figura 2.26 muestra el diagrama de bloques correspondiente.

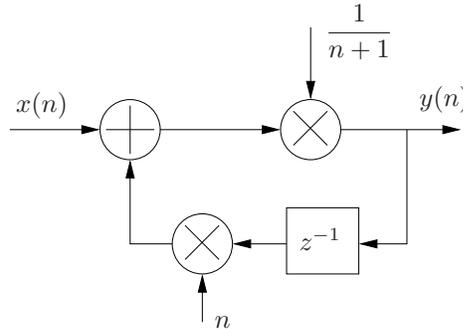


Figura 2.26: Diagrama de bloques del sistema de media acumulativa.

2.16

2.5.2 Sistemas LTI caracterizados por ecuaciones de diferencias con coeficientes constantes

Los sistemas descritos por ecuaciones de diferencias con coeficientes constantes son una subclase de los sistemas recursivos y no recursivos.

Considérese un sistema causal representado por (ver además la figura 2.27)

$$y(n] = ay(n-1) + x(n)$$

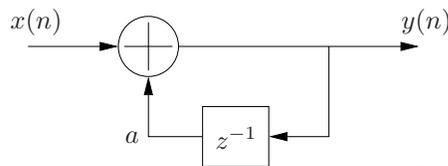


Figura 2.27: Diagrama de bloques de un sistema LTI.

que a pesar de su similitud con (2.23), difiere por la naturaleza de los coeficientes, lo que tiene implicaciones sobre la invarianza en el tiempo. En este último caso, el coeficiente a es constante y el sistema es invariante en el tiempo. Para la media acumulativa, el coeficiente es dependiente del tiempo y el sistema es entonces variante en el tiempo.

Evalúese ahora la respuesta de este sistema ante una entrada $x(n)$ con $x(n) = 0$, $n < 0$, y

con una condición inicial $y(-1)$:

$$\begin{aligned}
 y(0) &= ay(-1) + x(0) \\
 y(1) &= ay(0) + x(1) = a^2y(-1) + ax(0) + x(1) \\
 y(2) &= ay(1) + x(2) = a^3y(-1) + a^2x(0) + ax(1) + x(2) \\
 &\vdots \\
 y(n) &= a^{n+1}y(-1) + a^n x(0) + a^{n-1}x(1) + \dots + a^0x(n) \\
 &= \underbrace{a^{n+1}y(-1)}_{y_{zi}(n)} + \underbrace{\sum_{k=0}^n a^k x(n-k)}_{y_{zs}(n)}, n \geq 0
 \end{aligned}$$

El término $y_{zi}(n)$ depende de las condiciones iniciales y se obtendría si la entrada fuese cero (*zero input*), como producto del estado inicial del sistema y de sus características propias. A $y_{zi}(n)$ se le denomina respuesta natural o libre del sistema, o también, respuesta a entrada nula.

El término $y_{zs}(n)$ se obtiene cuando el estado del sistema es cero (*zero state*), es decir, con una entrada $x(n)$ cuando el sistema está en reposo, y se le denomina respuesta en estado nulo o respuesta forzada. Nótese que en el ejemplo $y_{zs}(n)$ puede interpretarse como la convolución de $x(n)$ con la respuesta impulsional:

$$h(n) = a^n u(n)$$

donde los índices son finitos por la causalidad de ambas señales $x(n)$ y $h(n)$.

Este ejemplo corresponde a una ecuación de diferencias de primer orden, y es un caso particular de la ecuación de diferencias:

$$y(n) = - \sum_{k=1}^N a_k y(n-k) + \sum_{k=0}^M b_k x(n-k)$$

o con $a_0 = 1$:

$$\sum_{k=0}^N a_k y(n-k) = \sum_{k=0}^M b_k x(n-k)$$

donde el entero N recibe el nombre de orden de la ecuación de diferencias u orden del sistema.

Las condiciones iniciales $y(-1), \dots, y(-N)$ resumen toda la historia pasada del sistema, y son necesarias para efectuar el cálculo de las salidas presentes y futuras.

Nótese que la ecuación de diferencias por sí sola no describe un *sistema*, en el sentido de que una entrada $x(n)$ no produce *una* única salida: para cada conjunto de condiciones iniciales se obtiene una salida particular. De este modo, un sistema es caracterizado tanto por la ecuación de diferencias como por las condiciones iniciales.

Anteriormente se introdujo la propiedad de escalado (2.20) de un sistema lineal, de la que se deduce que un sistema lineal con entrada cero debe tener una entrada nula, puesto que

de otro modo aparece una inconsistencia de escalado: multiplicando cualquier entrada con cero, debe producir una salida multiplicada también por cero, que debe ser nula. Con el esquema anterior compuesto de respuestas forzada y natural se nota que la respuesta natural de un sistema descrito con una ecuación de diferencias produce una salida a pesar de que la entrada es nula, y por tanto viola la propiedad de escalado lo que indica que la ecuación en el caso general representa a un sistema no lineal.

El concepto de linealidad debe entonces replantearse como la satisfacción de tres requisitos:

1. La respuesta total es igual a la suma de las respuestas de entrada nula y en estado nulo ($y(n) = y_{zi}(n) + y_{zs}(n)$).
2. El principio de superposición se aplica a la respuesta en estado nulo (lineal en estado nulo).
3. El principio de superposición se aplica a la respuesta a la entrada nula (lineal a entrada nula).

La redefinición del concepto para cada una de las componentes de la respuesta permite resolver las inconsistencias introducidas por las condiciones iniciales.

Ejemplo 2.17 Determine que un sistema descrito por la ecuación de diferencias de primer orden es lineal.

Solución:

1.

$$y(n) = ay(n-1) + x(n) = a^{n+1}y(-1) + \sum_{k=0}^n a^k x(n-k)$$

Con la entrada cero se obtiene:

$$y_{zi}(n) = a^{n+1}y(-1)$$

Con estado inicial cero se obtiene:

$$y_{zs}(n) = \sum_{k=0}^n a^k x(n-k) \Rightarrow y(n) = y_{zi}(n) + y_{zs}(n)$$

2. Para $x_1(n)$,

$$y_{zs_1}(n) = \sum_{k=0}^n a^k x_1(n-k)$$

Para $x_2(n)$,

$$y_{zs_2}(n) = \sum_{k=0}^n a^k x_2(n-k)$$

$$\begin{aligned} y_{zs_1}(n) + y_{zs_2}(n) &= \sum_{k=0}^n a^k x_1(n-k) + \sum_{k=0}^n a^k x_2(n-k) \\ &= \sum_{k=0}^n a^k (x_1(n-k) + x_2(n-k)) \end{aligned}$$

que equivale a la respuesta del sistema a $x_1(n) + x_2(n)$ y, por lo tanto, el sistema es lineal en estado nulo.

3. Con $y_{zi_1}(n) = a^{n+1}y_1(-1)$ y $y_{zi_2}(n) = a^{n+1}y_2(-1)$, se obtiene

$$\begin{aligned} y_{zi_1} + y_{zi_2} &= a^{n+1}y_1(-1) + a^{n+1}y_2(-1) \\ &= a^{n+1}(y_1(-1) + y_2(-1)) \end{aligned}$$

que equivale a la respuesta de entrada cero del sistema con condiciones iniciales $y(-1) = y_1(-1) + y_2(-1)$, por lo que el sistema es lineal a entrada nula.

De forma similar, es posible demostrar que todo sistema descrito por una ecuación de diferencias con coeficientes constantes es lineal. Estos sistemas son a su vez invariantes en el tiempo, al ser sus coeficientes constantes.

El estudio de estabilidad en sistemas de orden mayor a 1, requiere de otras técnicas que serán estudiadas posteriormente. 2.17

2.5.3 Solución de ecuaciones de diferencias con coeficientes constantes

El objetivo de solucionar la ecuación de diferencias

$$\sum_{k=0}^N a_k y(n-k) = \sum_{k=0}^M b_k x(n-k) \quad (2.24)$$

es encontrar la salida $y(n)$ de un sistema para una determinada entrada $x(n)$ dado un conjunto de condiciones iniciales (o finales). Se puede asumir que la ecuación de diferencias corresponde a un sistema causal, o anticausal, para lo cual usualmente se plantea la ecuación de $y(n)$ en términos de salidas anteriores o futuras, respectivamente. En el presente documento se tratará el caso causal, que es el más usual.

Si se tiene la ecuación de diferencias más condiciones de reposo iniciales $y(-1) = \dots = y(-N) = 0$, se asume que se tiene un sistema causal. Por otro lado, la misma ecuación de diferencias junto a condiciones de reposo finales $y(0) = y(1) = \dots = y(N-1) = 0$ caracterizan a un sistema anticausal.

Utilizando el llamado *método directo*, se asume que la solución se compone de dos partes:

$$y(n) = y_h(n) + y_p(n) \quad (2.25)$$

donde $y_h(n)$ es la *solución homogénea* o complementaria, y $y_p(n)$ es la *solución particular*.

La solución homogénea de una ecuación de diferencias

Para solucionar la ecuación (2.24) se asume primero una entrada nula $x(n) = 0$, para así resolver:

$$\sum_{k=0}^N a_k y(n-k) = 0 \quad (2.26)$$

De forma análoga a la solución de ecuaciones diferenciales, se asume que la solución es de forma exponencial:

$$y_h(n) = \lambda^n \quad (2.27)$$

Sustituyendo (2.27) como solución tentativa en (2.26), se obtiene:

$$\sum_{k=0}^N a_k \lambda^{n-k} = 0$$

que se puede reescribir como:

$$\lambda^{n-N} \underbrace{(\lambda^N + a_1 \lambda^{N-1} + a_2 \lambda^{N-2} + \dots + a_{N-1} \lambda + a_N)}_{\text{polinomio característico}} = 0$$

El término entre paréntesis es llamado *polinomio característico* y tiene en general N raíces, denotadas $\lambda_1, \lambda_2, \dots, \lambda_N$, que pueden ser reales o complejas, apareciendo en el último caso como pares complejos conjugados, siempre y cuando los coeficientes $\{a_k\}$ sean reales. Algunas de las raíces pueden ser idénticas (de orden múltiple).

Asumiendo que las raíces son distintas, la solución más general a la ecuación de diferencias homogénea es:

$$y_h(n) = C_1 \lambda_1^n + C_2 \lambda_2^n + \dots + C_N \lambda_N^n$$

con los coeficientes de ponderación C_k , que se calculan a partir de las condiciones iniciales especificadas para el sistema.

Ejemplo 2.18 Determine la solución homogénea para la ecuación de diferencias de primer orden:

$$y(n) + a_1 y(n-1) = x(n) \quad (2.28)$$

Solución:

Con $x(n) = 0$ y $y_h(n) = \lambda^n$ se obtiene:

$$\begin{aligned} \lambda^n + a_1 \lambda^{n-1} &= 0 \\ \lambda^{n-1}(\lambda + a_1) &= 0 \Rightarrow \lambda = -a_1 \end{aligned}$$

y la solución sería:

$$y_h(n) = C \lambda^n = C(-a_1)^n$$

Sustituyendo en (2.28) para $n = 0$, junto con $x(n) = 0$, se obtiene:

$$y(0) + a_1 y(-1) = 0 \Rightarrow y(0) = -a_1 y(-1)$$

y con $y_h(0) = C = -a_1 y(-1)$

$$y_h(n) = -a_1 y(-1)(-a_1)^n = y(-1)(-a_1)^{n+1} = y_{zi}(n), n \geq 0$$

que corresponde además con la respuesta de entrada nula. 2.18

Ejemplo 2.19 Determine la respuesta a entrada nula del sistema descrito por la ecuación de diferencias homogénea de segundo orden

$$y(n) - 3y(n-1) - 4y(n-2) = 0.$$

Solución:

Suponiendo $y_h(n) = \lambda^n$, se obtiene:

$$\lambda^n - 3\lambda^{n-1} - 4\lambda^{n-2} = \lambda^{n-2}(\lambda^2 - 3\lambda - 4) = \lambda^{n-2}(\lambda - 4)(\lambda + 1)$$

Por lo que la forma general de la solución homogénea es:

$$y_h(n) = C_1 \lambda_1^n + C_2 \lambda_2^n = C_1 (-1)^n + C_2 (4)^n$$

Para la respuesta a entrada nula se plantea:

$$\begin{aligned} y(0) &= 3y(-1) + 4y(-2) = C_1 + C_2 \\ y(1) &= 3y(0) + 4y(-1) \\ &= 3(3y(-1) + 4y(-2)) + 4y(-1) \\ &= 13y(-1) + 12y(-2) = -C_1 + 4C_2 \end{aligned} \tag{2.29}$$

Sumando $y(0)$ y $y(1)$ resulta en

$$C_2 = \frac{16y(-1) + 16y(-2)}{5}$$

y considerando (2.29) se deriva además

$$\begin{aligned} 3y(-1) + 4y(-2) - C_2 &= C_1 \\ 3y(-1) + 4y(-2) - \frac{16y(-1) + 16y(-2)}{5} &= C_1 \\ \Rightarrow C_1 &= \frac{4y(-2) - y(-1)}{5} \end{aligned}$$

Finalmente

$$y_{zi}(n) = \frac{4y(-2) - y(-1)}{5} (-1)^n + \frac{16y(-1) + 16y(-2)}{5} 4^n.$$

2.19

Si λ_1 es una raíz de multiplicidad m , entonces:

$$y_h(n) = C_1 \lambda_1^n + C_2 n \lambda_1^n + C_3 n^2 \lambda_1^n + \dots + C_m n^{m-1} \lambda_1^n + C_{m+1} \lambda_2^n + \dots$$

Ejemplo 2.20 Repita el ejemplo anterior para:

$$y(n) - 4y(n-1) + 4y(n-2) = 0$$

Solución:

Suponiendo $y_h(n) = \lambda^n$, resulta en el polinomio característico

$$\lambda^{n-2}(\lambda^2 - 4\lambda + 4) = \lambda^{n-2}(\lambda - 2)^2 = 0$$

por lo que la solución general es

$$y_h(n) = C_1\lambda_1^n + C_2n\lambda_1^n, \quad \lambda_1 = 2$$

Se deja como ejercicio al lector demostrar que

$$y_{zi}(n) = [y(-1) - y(-2)]2^{n+2} + [y(-1) - 2y(-2)]n2^{n+1}$$

2.20

La solución particular de la ecuación de diferencias

La solución particular $y_p(n)$ debe satisfacer:

$$\sum_{k=0}^N a_k y_p(n-k) = \sum_{k=0}^M b_k x(n-k), \quad a_0 = 1$$

lo que generalmente depende de la forma de $x(n)$.

Ejemplo 2.21 Determine la solución particular de

$$y(n) + a_1 y(n-1) = x(n), \quad |a_1| < 1 \tag{2.30}$$

para $x(n) = u(n)$

Solución:

Como $x(n)$ es constante para $n \geq 0$, se asume que la solución también es constante para $n \geq 0$. La solución particular es entonces:

$$y_p(n) = Ku(n)$$

Sustituyendo esto en (2.30) se obtiene:

$$Ku(n) + a_1 Ku(n-1) = u(n) \tag{2.31}$$

y para $n \geq 1$ esto equivale a:

$$\begin{aligned} K + a_1 K &= 1 \\ K(1 + a_1) &= 1 \\ K &= \frac{1}{1 + a_1} \end{aligned}$$

Tabla 2.4: Forma general de la solución particular para diversos tipos de señal de entrada

Señal de entrada $x(n)$	Solucion particular $y_p(n)$
$\delta(n)$	0
$A(\text{cte})$	K
AM^n	KM^n
An^M	$K_0n^M + K_1n^{M-1} + \dots + K_M$
$A^n n^M$	$A^n(K_0n^M + K_1n^{M-1} + \dots + K_M)$
$\begin{cases} A \cos \omega_0 n \\ A \sin \omega_0 n \end{cases}$	$K_1 \cos \omega_0 n + K_2 \sin \omega_0 n$

lo que implica que la solución particular es:

$$y_p(n) = \frac{1}{1 + a_1} \cdot u(n)$$

2.21

En el ejemplo anterior se asumió $y_p(n)$ como un escalón porque la entrada tenía esta forma. Si $x(n)$ fuera exponencial, $y_p(n)$ también lo sería. Si $x(n)$ fuera senoidal, entonces $y_p(n)$ también lo sería. La tabla 2.4 muestra algunas formas de solución particular para entradas particulares.

Solución total a la ecuación de diferencias

Por la linealidad de los sistemas descritos por ecuaciones de diferencias con coeficientes constantes, la solución total se calcula como la suma de las soluciones homogénea y particular (2.25).

Las constantes $\{C_k\}$, de la solución homogénea, sin embargo, deben recalcularse para poder satisfacer las condiciones iniciales.

Ejemplo 2.22 Determine la solución total $y(n)$, $n \geq 0$ de la ecuación de diferencias $y(n) + a_1 y(n-1) = x(n)$, con $x(n) = u(n)$ y condición inicial $y(-1)$.

Solución:

De los ejemplos anteriores se tiene:

$$y_h(n) = C(-a_1)^n, \quad y_p(n) = \frac{1}{1 + a_1}$$

$$y(n) = C(-a_1)^n + \frac{1}{1 + a_1}$$

Con $n = -1$ se obtiene:

$$y(-1) = -\frac{C}{a_1} + \frac{1}{1 + a_1} \Rightarrow -a_1 y(-1) + \frac{a_1}{1 + a_1} = C$$

y por lo tanto:

$$y(n) = (-a_1)^{n+1}y(-1) + \frac{1 - (-a_1)^{n+1}}{1 + a_1}$$

Nótese que el valor de C depende no sólo de las condiciones iniciales, sino también de la entrada. 2.22

2.5.4 Respuesta impulsional de un sistema recursivo LTI

En general, las condiciones iniciales de una ecuación de diferencias además de obligar a la revisión del concepto de linealidad, tienen efecto sobre la invarianza temporal del sistema. Esto obliga entonces a redefinir el concepto de respuesta impulsional.

La respuesta impulsional $h(n)$ en sistemas recursivos es igual a la respuesta *de estado cero* del sistema cuando la entrada $x(n) = \delta(n)$, es decir, para la respuesta impulsional siempre se considera que el sistema está inicialmente en reposo.

Por ejemplo, en el sistema recursivo de primer orden tratado anteriormente, la respuesta de estado cero es:

$$y_{zs}(n) = \sum_{k=0}^n a^k x(n-k)$$

y con $x(n) = \delta(n)$, se obtiene:

$$\begin{aligned} y_{zs}(n) &= \sum_{k=0}^n a^k \delta(n-k) \\ &= a^n, \quad n \geq 0 \end{aligned}$$

Así, su respuesta impulsional es $h(n) = a^n u(n)$, que coincide con lo calculado previamente.

En el caso general de sistemas recursivos LTI (considerados causales), la respuesta de estado cero en términos de convolución con una entrada causal se expresa como

$$y_{zs}(n) = \sum_{k=0}^n h(k)x(n-k) = h(n) * x(n), \quad n \geq 0$$

Si la entrada es un impulso, lo anterior se reduce a:

$$y_{zs}(n) = h(n) * \delta(n) = h(n)$$

Con la entrada $\delta(n)$, la solución particular es cero, puesto que $x(n) = 0$ para $n > 0$. Consecuentemente, la respuesta de un sistema a un impulso consiste en la solución a la ecuación homogénea con los coeficientes C_k determinados para satisfacer las condiciones iniciales.

Ejemplo 2.23 Determine la respuesta impulsional $h(n)$ del sistema

$$y(n) - 3y(n-1) - 4y(n-2) = x(n) + 2x(n-1) \quad (2.32)$$

Solución:

En el ejemplo 2.19 se obtuvo:

$$y_h(n) = C_1(-1)^n + C_2(4)^n, \quad n \geq 0 \quad (2.33)$$

que corresponde también a la respuesta impulsional $h(n)$, debido a que para $x(n) = \delta(n)$, $y_p(n) = 0$. Sustituyendo (2.33) en (2.32), y considerando que el sistema está en reposo ($y(-1) = y(-2) = 0$):

$$\begin{aligned} y(0) &= 1 \stackrel{!}{=} C_1 + C_2 \\ y(1) - 3y(0) &= 2 \Rightarrow y(1) = 5 \stackrel{!}{=} -C_1 + 4C_2 \\ &\Rightarrow 5C_2 = 6 \Rightarrow C_2 = \frac{6}{5}, C_1 = -\frac{1}{5} \end{aligned}$$

y finalmente

$$h(n) = \left[-\frac{1}{5}(-1)^n + \frac{6}{5}(4)^n \right] u(n)$$

2.23

Cualquier sistema recursivo descrito por una ecuación de diferencias lineal con coeficientes constantes es un sistema IIR, pero no todo sistema IIR LTI puede ser descrito con estas ecuaciones.

La respuesta impulsional de un sistema de orden N , descrito por una ecuación de diferencias lineal de orden N :

$$y(n) = -\sum_{k=1}^N a_k y(n-k) + \sum_{k=0}^M b_k x(n-k)$$

se obtiene a través de la solución homogénea:

$$y_h(n) = \sum_{k=1}^N C_k \lambda_k^n = h(n)$$

cuando las raíces del polinomio característico son distintas, donde los coeficientes $\{C_k\}$ se obtienen con las condiciones iniciales

$$y(-1) = y(-2) = \dots = y(-N) = 0$$

Considerando que

$$\sum_{n=0}^{\infty} |h(n)| = \sum_{n=0}^{\infty} \left| \sum_{k=1}^N C_k \lambda_k^n \right| \leq \sum_{k=1}^N |C_k| \sum_{n=0}^{\infty} |\lambda_k|^n$$

si $|\lambda_k| < 1$ para todo k , entonces $\sum_{n=0}^{\infty} |\lambda_k|^n < \infty$, y por lo tanto $\sum_{n=0}^{\infty} |h(n)| < \infty$. Si uno o más de los $|\lambda_k| > 1$, entonces $h(n)$ no es absolutamente sumable y en consecuencia el sistema es inestable.

Por lo tanto, un sistema IIR causal descrito por una ecuación de diferencias con coeficientes constantes, es estable si todas las raíces del polinomio son menores que uno en valor absoluto. Esto es así aún con raíces de multiplicidad m .

2.6 Correlación

La *correlación* es una operación similar a la convolución que se utiliza para medir la similitud entre dos secuencias. Se aplica en diversas áreas de la ingeniería como radar, sonar, comunicaciones digitales, geología, etc.

2.6.1 Autocorrelación y correlación cruzada

La *correlación cruzada* de las secuencias $x(n)$ e $y(n)$, ambas reales y de energía finita, se define como:

$$r_{xy}(n) = \sum_{k=-\infty}^{\infty} x(k)y(k-n), n = 0, \pm 1, \pm 2, \dots$$

o lo que es equivalente:

$$r_{xy}(n) = \sum_{k=-\infty}^{\infty} x(k+n)y(k), n = 0, \pm 1, \pm 2, \dots$$

El índice n es el parámetro de desplazamiento o retardo en el tiempo, y los subíndices xy indican cuáles señales han sido correlacionadas. De esta forma se tiene entonces que:

$$r_{yx}(n) = \sum_{k=-\infty}^{\infty} y(k)x(k-n) = \sum_{k=-\infty}^{\infty} y(k+n)x(k) = r_{xy}(-n)$$

Por lo tanto, $r_{yx}(n)$ es la correlación $r_{xy}(n)$ reflejada con respecto a $n = 0$.

Nótese que exceptuando el paso de reflexión de una de las señales, la correlación es idéntica con la convolución: se desplaza la segunda señal, se calcula la secuencia producto, y se determina su suma. El lector puede demostrar que:

$$r_{xy}(n) = x(n) * y(-n)$$

Si $y(n) = x(n)$, entonces se obtiene la *autocorrelación* de $x(n)$, que se define como la secuencia:

$$r_{xx}(n) = \sum_{k=-\infty}^{\infty} x(k)x(k-n) = \sum_{k=-\infty}^{\infty} x(k+n)x(k)$$

Si $x(n)$ e $y(n)$ son señales causales finitas de longitud N (esto es, $x(n) = y(n) = 0, \forall n < 0, n \geq N$) la correlación puede expresarse como:

$$r_{xy}(n) = \sum_{k=\max\{n,0\}}^{N+\min\{0,n\}-1} x(k)y(k-n)$$

2.6.2 Propiedades de la correlación

Sean $x(n)$ y $y(n)$ dos señales de energía finita. La energía de la señal:

$$w(k) = ax(k) + by(k - n)$$

es:

$$\begin{aligned} \sum_{k=-\infty}^{\infty} [ax(k) + by(k - n)]^2 &= a^2 \sum_{k=-\infty}^{\infty} x^2(k) + b^2 \sum_{k=-\infty}^{\infty} y^2(k - n) + 2ab \sum_{k=-\infty}^{\infty} x(k)y(k - n) \\ &= a^2 r_{xx}(0) + b^2 r_{yy}(0) + 2abr_{xy}(n) \end{aligned}$$

Nótese que $r_{xx}(0) = E_x$ es la energía de la señal $x(n)$ y $r_{yy}(0) = E_y$ la energía de $y(n)$. Puesto que la energía es una suma de valores positivos debe cumplirse que:

$$a^2 r_{xx}(0) + b^2 r_{yy}(0) + 2abr_{xy}(n) \geq 0$$

y suponiendo que $b \neq 0$, se obtiene dividiendo por b^2 :

$$r_{xx}(0) \left(\frac{a}{b}\right)^2 + 2r_{xy}(n) \left(\frac{a}{b}\right) + r_{yy}(0) \geq 0$$

que se puede considerar como ecuación cuadrática de variable (a/b) con coeficientes $r_{xx}(0)$, $2r_{xy}(n)$ y $r_{yy}(0)$, que para ser no negativa requiere que el discriminante sea no positivo:

$$4[r_{xy}^2 - r_{xx}(0)r_{yy}(0)] \leq 0 \Rightarrow |r_{xy}(n)| \leq \sqrt{r_{xx}(0)r_{yy}(0)} = \sqrt{E_x E_y}$$

y para $y(n) = x(n)$:

$$|r_{xx}(n)| \leq r_{xx}(0) = E_x$$

es decir, la autocorrelación alcanza su valor máximo para el retardo cero, cuando por decirlo así la señal es idéntica a sí misma.

Un escalado de la señal por un escalar, escala la correlación de la misma manera, y por tanto carece en el análisis de importancia por ser más relevante la forma de la correlación. Por ello en la práctica se *normaliza* la correlación para producir así valores entre -1 y 1. La autocorrelación normalizada se define como:

$$\rho_{xx}(n) = \frac{r_{xx}(n)}{r_{xx}(0)}$$

y la correlación cruzada como:

$$\rho_{xy}(n) = \frac{r_{xy}(n)}{\sqrt{r_{xx}(0)r_{yy}(0)}}$$

Como $r_{xy}(n) = r_{yx}(-n)$, se cumple para la autocorrelación $r_{xx}(n) = r_{xx}(-n)$, es decir, es una función par.

Ejemplo 2.24 Calcule la autocorrelación de la señal:

$$x(n) = a^n u(n), \quad 0 < a < 1$$

Solución:

Para $n \geq 0$:

$$r_{xx}(n) = \sum_{k=n}^{\infty} x(k)x(k-n) = \sum_{k=n}^{\infty} a^k a^{k-n} = a^{-n} \sum_{k=n}^{\infty} (a^2)^k$$

y puesto que $\lim_{N \rightarrow \infty} \alpha^{N+1} = 0$ si $\alpha < 1$ entonces

$$\begin{aligned} \sum_{k=n}^{\infty} a^k &= \lim_{N \rightarrow \infty} \frac{\alpha^n - \alpha^{N+1}}{1 - \alpha} = \frac{\alpha^n}{1 - \alpha} \\ r_{xx}(n) &= a^{-n} \left(\frac{a^{2n}}{1 - a^2} \right) = \frac{a^n}{1 - a^2} \end{aligned}$$

para $n < 0$:

$$\begin{aligned} r_{xx}(n) &= \sum_{k=0}^{\infty} x(k)x(k-n) = \sum_{k=0}^{\infty} a^k a^{k-n} = a^{-n} \sum_{k=0}^{\infty} (a^2)^k \\ &= \frac{a^{-n}}{1 - a^2} = \frac{a^{|n|}}{1 - a^2} \end{aligned}$$

y combinando ambas partes se obtiene:

$$r_{xx}(n) = \frac{a^{|n|}}{1 - a^2}$$

que es, como se predijo, una función par $r_{xx}(n) = r_{xx}(-n)$. Con $r_{xx}(0) = \frac{1}{1-a^2}$ se obtiene la autocorrelación normalizada:

$$\rho_{xx}(n) = \frac{r_{xx}(n)}{r_{xx}(0)} = a^{|n|}$$

2.24

2.6.3 Correlación de secuencias periódicas

La correlación anterior fue definida para señales de energía. Si las señales $x(n)$ e $y(n)$ son señales de potencia, su correlación cruzada se define como:

$$r_{xy}(n) = \lim_{M \rightarrow \infty} \frac{1}{2M+1} \sum_{k=-M}^M x(k)y(k-n)$$

y la autocorrelación de una señal de potencia como:

$$r_{xx}(n) = \lim_{M \rightarrow \infty} \frac{1}{2M+1} \sum_{k=-M}^M x(k)x(k-n)$$

Si las señales son periódicas con periodo N , entonces este promedio puede calcularse en un solo periodo:

$$r_{xy}(n) = \frac{1}{N} \sum_{k=0}^{N-1} x(k)y(k-n)$$

En la práctica se utiliza la correlación para encontrar periodicidades en señales físicas alteradas por interferencias aleatorias.

Si por ejemplo $y(n) = x(n) + \omega(n)$, donde $x(n)$ es periódica y $\omega(n)$ es una señal aleatoria, entonces la correlación es:

$$\begin{aligned} r_{yy}(n) &= \frac{1}{M} \sum_{k=0}^{M-1} y(k)y(k-n) \\ &= \frac{1}{M} \sum_{k=0}^{M-1} [x(k) + \omega(k)][x(k-n) + \omega(k-n)] \\ &= \frac{1}{M} \sum_{k=0}^{M-1} [x(k)x(k-n)] + [x(k)\omega(k-n)] + [\omega(k)x(k-n)] + [\omega(k)\omega(k-n)] \\ &= r_{xx}(n) + r_{x\omega}(n) + r_{\omega x}(n) + r_{\omega\omega}(n) \end{aligned}$$

considerando sólo M muestras y asumiendo $y(n) = 0$ para $n < 0, n \geq M$.

La periodicidad de $x(n)$ permite asumir que $r_{xx}(n)$ también es periódica y presenta picos en $n = 0, N, 2N$, etc.. Por la consideración de tan solo M muestras, conforme N tiende a M la magnitud de estos picos tiende a cero, por lo que debe escogerse $M \gg N$ y n no debe evaluarse, como regla empírica, para $n > M/2$.

Si $\omega(n)$ es aleatoria puede asumirse que $r_{x\omega}(n)$ y $r_{\omega x}(n)$ son muy pequeñas, por no estar relacionada con $x(n)$. Por la naturaleza aleatoria de $\omega(n)$ puede asumirse además que $r_{\omega\omega}(n)$ tiende rápidamente a cero, por lo que en $r_{\omega\omega}(n)$ dominará $r_{xx}(n)$, lo que permite detectar el periodo.

La figura 2.28 muestra un segmento de 701 muestras de una señal acústica representando la vocal "a". La señal original tiene una longitud de 57000 muestras. La figura 2.29 muestra el resultado de la autocorrelación, donde se aprecia claramente un periodo de aproximadamente 100 muestras, lo que se puede corroborar con la señal original en la figura 2.28.

2.6.4 Secuencias de correlación de entrada-salida

Si se aplica la señal $x(n)$ con autocorrelación $r_{xx}(n)$ a un sistema con respuesta al impulso $h(n)$, entonces:

$$y(n) = h(n) * x(n) = \sum_{k=-\infty}^{\infty} h(k)x(n-k)$$

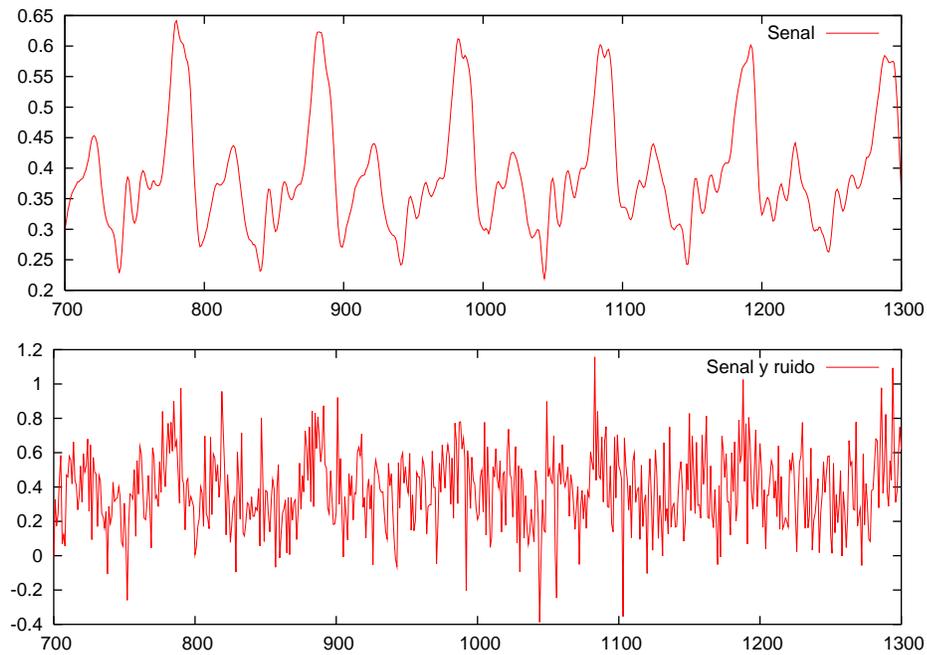


Figura 2.28: Señal acústica representando la vocal “a”, en la parte superior pura, y en la parte inferior con ruido.

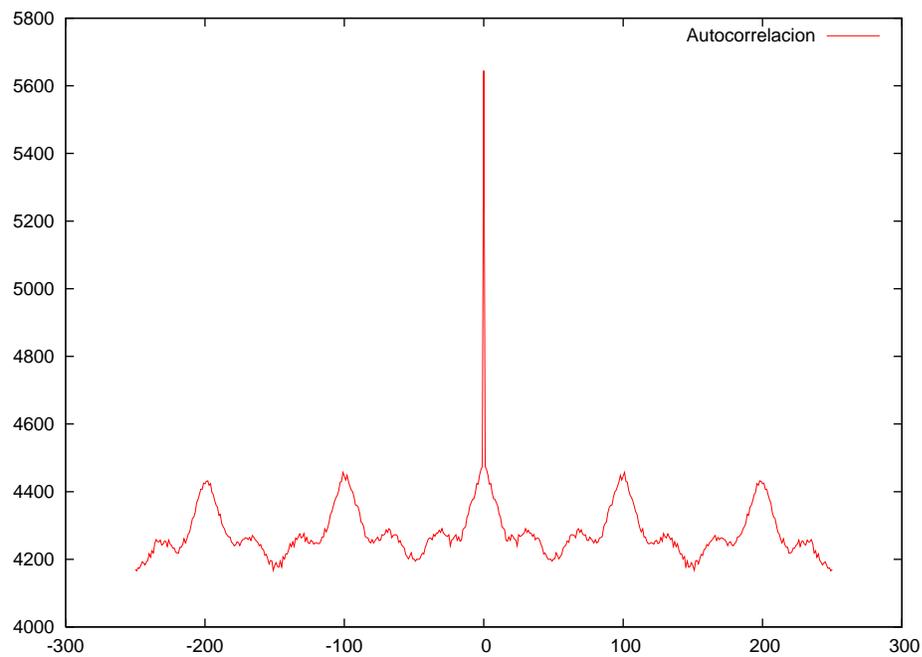


Figura 2.29: Autocorrelación de la señal con ruido en la figura 2.28.

La correlación entre la salida y la entrada es

$$\begin{aligned}
 r_{yx}(n) &= y(n) * x(-n) \\
 &= [h(n) * x(n)] * x(-n) \\
 &= h(n) * r_{xx}(n)
 \end{aligned}$$

que depende solo de la respuesta impulsional $h(n)$ y la autocorrelación de la entrada $r_{xx}(n)$.

La autocorrelación de la salida es

$$\begin{aligned}r_{yy}(n) &= y(n) * y(-n) \\ &= [h(n) * x(n)] * [h(-n) * x(-n)] \\ &= [h(n) * h(-n)] * [x(n) * x(-n)] \\ &= r_{hh}(n) * r_{xx}(n)\end{aligned}$$

que equivale a la convolución de la autocorrelación de la respuesta impulsional $r_{hh}(n)$ con la autocorrelación de la entrada $r_{xx}(n)$.

2.7 Problemas

Problema 2.1. Encuentre expresiones algebraicas que permitan generar a partir de la señal $x(n)$ las siguientes secuencias, utilizando para ello las operaciones fundamentales de reflexión y desplazamiento.

$$x(n) = \{\dots, a_{-2}, a_{-1}, \underset{\uparrow}{a_0}, a_1, a_2, a_3, \dots\}$$

1. $x_1(n) = \{\dots, a_{-4}, a_{-3}, \underset{\uparrow}{a_{-2}}, a_{-1}, a_0, a_1, a_2, \dots\}$
2. $x_2(n) = \{\dots, a_1, a_2, \underset{\uparrow}{a_3}, a_4, a_5, a_6, a_7, \dots\}$
3. $x_3(n) = \{\dots, a_4, a_3, \underset{\uparrow}{a_2}, a_1, a_0, a_{-1}, a_{-2}, \dots\}$
4. $x_4(n) = \{\dots, a_{-1}, a_{-2}, \underset{\uparrow}{a_{-3}}, a_{-4}, a_{-5}, a_{-6}, a_{-7}, \dots\}$

Problema 2.2. Si $x_1(n) = \{4, 2, \underset{\uparrow}{-1}, 5, 1, 3\}$ y $x_2(n) = \{1, -2, 3, -4, 5\}$, calcule

1. $x_3(n) = x_1(n) + x_2(n)$
2. $x_4(n) = x_1(n)x_2(n)$
3. $x_5(n) = 3x_1(n)x_2(n)$

Problema 2.3. Sea $f = k/N$ la frecuencia normalizada de una señal $x(n)$ periódica en tiempo discreto con periodo N muestras. Si se asume que k y N son números primos relativos, y se sabe que $x(n) = x_a(nT)$ donde $x_a(t)$ es una señal analógica periódica de periodo T_a que es muestreada entonces con un intervalo de muestreo T , encuentre el número de veces que cabe el periodo de la señal analógica dentro del periodo de la señal muestreada.

Problema 2.4. Determine si cada una de las señales siguientes es periódica, y de ser así determine su periodo fundamental. Asuma $t \in \mathbb{R}$ y $n \in \mathbb{Z}$.

1. $x_a(t) = \sqrt{5} \operatorname{sen}(4t + \pi/3)$
2. $x(n) = \sqrt{5} \operatorname{sen}(4n + \pi/3)$
3. $x(n) = 2e^{j(n\pi/6 - \pi)}$
4. $x(n) = \cos(n/8) \cos(n\pi/8)$
5. $x(n) = \operatorname{sen}(\pi n/4) + \operatorname{sen}(\pi n/8) + 3 \cos(\pi n/2 + \pi/6)$

Problema 2.5. En la sección 2.2.3 se demostró que una señal exponencial compleja discreta $s_k(n) = e^{jk\omega_0 n}$ solo tiene $N - 1$ señales armónicamente relacionadas, cuando se cumple $\omega_0 = 2\pi/N$. Demuestre que si la frecuencia es $\omega_0 = 2\pi M/N$ con $M < N$, M y N números primos relativos, el número de frecuencias armónicamente relacionadas sigue siendo N .

Problema 2.6.

Genere y muestre las siguientes secuencias en GNU/Octave:

1. $x_1(n) = 0,75\delta(n - 5)$ para $0 \leq n \leq 10$
2. $x_2(n) = 0,9\delta(n)$ para $-7 \leq n \leq 7$
3. $x_3(n) = 1,5\delta(n - 250)$ para $-300 \leq n \leq 350$
4. $x_4(n) = 4,5\delta(n + 7)$ para $-10 \leq n \leq 0$
5. $x_5(n) = \dots, 0, 1, 2, 3, 4, 0, 1, 2, 3, 4, 0, 1 \dots$ para $-5 \leq n \leq 9$
6. $x_6(n) = \text{sen } \frac{\pi}{17}n$ para $0 \leq n \leq 25$
7. $x_7(n) = \text{sen } \frac{\pi}{17}n$ para $-15 \leq n \leq 25$
8. $x_8(n) = \text{sen}(3\pi n + \frac{\pi}{2})$ para $-10 \leq n \leq 10$
9. $x_9(n) = \text{cos}(\frac{\pi}{\sqrt{23}}n)$ para $0 \leq n \leq 50$

Problema 2.7. Dada una señal $x(n) = \{-2, 3, 4, -1, 0, 2, 3, 2\}$.

Represente esta señal en GNU/Octave y muéstrela gráficamente.

Calcule las componentes par e impar de esta señal $x_p(n)$ e $x_i(n)$.

Verifique que la suma de ambas componentes generan la señal de entrada original.

Problema 2.8. Genere y visualice en GNU/Octave una representación para la señal $h(n) = \{2/7, 2/7, 3/14, 1/7, 1/14, 0\}$, y otra para la señal $x(n) = \{1, 0, 0, 0, 0, 1\}$.

Calcule y visualice la convolución de las señales $h(n)$ y $x(n)$ utilizando el operador `conv`.

Para una señal constante, ¿qué salida tiene un sistema con respuesta impulsional $h(n)$?

¿Qué tipo de *filtro* es $h(n)$?

Problema 2.9. Genere y visualice en GNU/Octave una representación para la señal $h(n) = \{1, -1\}$, y otra para la señal $x(n) = \{1, 0, 0, 0, 0, 1\}$.

Calcule y visualice la convolución de ambas señales

Para una señal constante, ¿qué salida tiene un sistema con respuesta impulsional $h(n)$?

¿Qué tipo de *filtro* es $h(n)$?

Problema 2.10. Programe una función de GNU/Octave que reciba dos señales finitas y retorna su convolución.

Verifique el funcionamiento de su función comparándola con el operador `conv`.

Problema 2.11. Genere una función de GNU/Octave que reciba un vector con los coeficientes a_k y otro vector con los coeficientes b_k , además de la señal de entrada, y que utilizando la expresión del sistema como ecuación de diferencias genere la señal de salida para un número dado de muestras. Las condiciones iniciales deben poder darse en forma opcional, asumiendo

un valor de cero en todas ellas si se omiten.

Capítulo 3

Análisis de sistemas LTI discretos con la transformada z

La *transformada z* es a los sistemas discretos lo que la transformada de Laplace es a los sistemas continuos. Ambas representan herramientas para el análisis de ciertas propiedades de las señales, que en el dominio del tiempo sólo pueden ser evaluadas con un mayor número de pasos.

3.1 La transformada z

3.1.1 La transformada z directa

La *transformada z directa* de una señal $x(n)$ se define como la serie de potencias:

$$X(z) \equiv \sum_{n=-\infty}^{\infty} x(n)z^{-n} \quad (3.1)$$

que mapea la señal $x(n)$ en el dominio del tiempo discreto a la función $X(z)$ en el dominio z , lo que se denota como:

$$X(z) \equiv \mathcal{Z} \{x(n)\}$$

La relación entre ambos dominios se indica como:

$$x(n) \circ \bullet X(z) \quad \text{ó} \quad x(n) \xleftrightarrow{z} X(z)$$

Como la transformada z es una serie infinita de potencias, ésta existe solo para los valores de z para los que la serie converge. La región de convergencia (ROC, *region of convergence*) de $X(z)$ es entonces el conjunto de valores de z para los que $X(z)$ es finita.

Nótese que siempre que se hable de la transformada z de una señal $x(n)$, debe incluirse la ROC.

Si se expresa z en su forma polar $z = re^{j\varphi}$, con $r = |z|$ y $\varphi = \angle z$, entonces

$$X(z) = \sum_{n=-\infty}^{\infty} x(n)r^{-n}e^{-j\varphi n}.$$

Puesto que dentro de la ROC de $X(z)$ se debe cumplir $|X(z)| < \infty$, entonces:

$$|X(z)| = \left| \sum_{n=-\infty}^{\infty} x(n)r^{-n}e^{-j\varphi n} \right| \leq \sum_{n=-\infty}^{\infty} |x(n)r^{-n}| \quad (3.2)$$

es decir, $|X(z)|$ es finita si y sólo si $x(n)r^{-n}$ es absolutamente sumable.

Para encontrar la ROC se debe entonces encontrar el rango de valores de r para los que la secuencia $x(n)r^{-n}$ es absolutamente sumable.

Ahora bien, la ecuación (3.2) puede reescribirse como:

$$|X(z)| = \sum_{n=-\infty}^{-1} |x(n)r^{-n}| + \sum_{n=0}^{\infty} |x(n)r^{-n}| = \sum_{n=1}^{\infty} |x(-n)r^n| + \sum_{n=0}^{\infty} |x(n)r^{-n}|$$

y ambas sumatorias deben converger si $|X(z)|$ ha de ser finito. Para la primera suma, que corresponde a los elementos anticausales, deben existir valores de r suficientemente pequeños para que $x(-n)r^n$ sea absolutamente sumable ($r < r_1$) (figura 3.1).

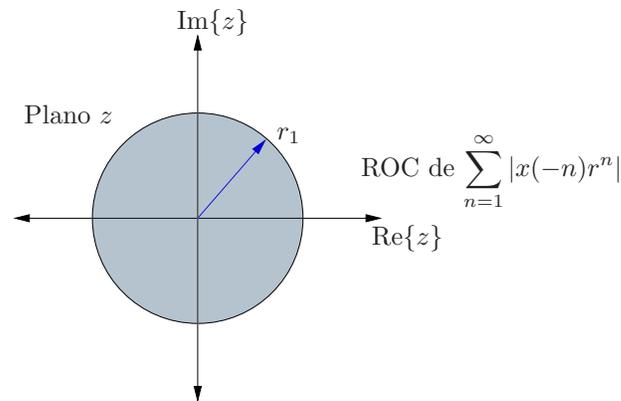


Figura 3.1: Representación gráfica de la ROC para r suficientemente pequeños.

Para la segunda suma, correspondiente a los elementos causales de $x(n)$, se necesitan valores de r suficientemente grandes para que $x(n)r^{-n}$ sea absolutamente sumable y por tanto converja. Por ello, la ROC serán los puntos fuera de una circunferencia $r > r_2$ (figura 3.2).

Como ambas sumas deben converger, la ROC de $X(z)$ es la región anular del plano z , $r_2 < r < r_1$ (figura 3.3).

La figura 3.4 muestra un resumen de lo discutido hasta el momento en cuanto a la relación de la causalidad de una señal con respecto a la ROC de su transformada z .

Por hacer:

Agregar tabla de propiedades

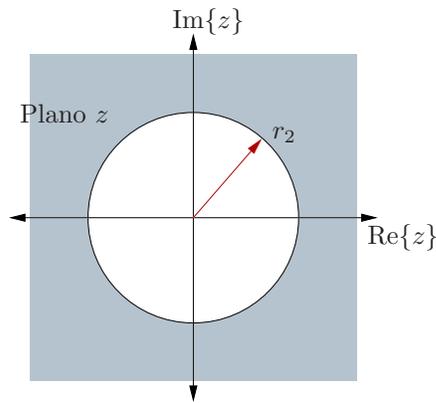


Figura 3.2: Representación gráfica de la ROC para r suficientemente grandes.

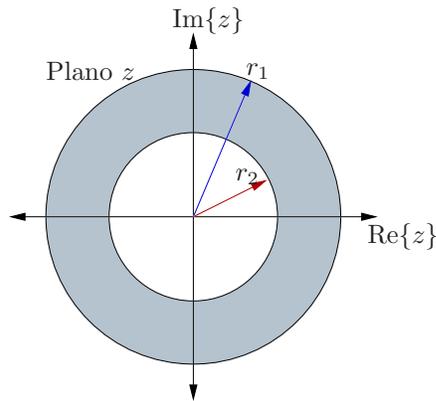


Figura 3.3: Representación gráfica completa de la ROC.

3.2 Transformadas z racionales

Una clase de transformadas z encontradas con frecuencia en sistemas digitales presentan una forma racional, es decir, un cociente de polinomios en z^{-1} ó z . La tabla 3.1 resume la transformada z de algunas funciones comunes con transformadas de este tipo.

3.2.1 Polos y ceros

Los *ceros* de la transformada z racional $X(z)$ son aquellos valores de z para los que $X(z) = 0$, y los *polos*, aquellos z para los que $X(z) = \infty$ y la serie de Laurent centrada en z contiene un número finito de términos en su parte principal. Como $X(z)$ es racional, entonces:

$$X(z) = \frac{N(z)}{D(z)} = \frac{b_0 + b_1 z^{-1} + \dots + b_M z^{-M}}{a_0 + a_1 z^{-1} + \dots + a_N z^{-N}} = \frac{\sum_{k=0}^M b_k z^{-k}}{\sum_{k=0}^N a_k z^{-k}}$$

Asumiendo $a_0 \neq 0$ y $b_0 \neq 0$ y factorizando $b_0 z^{-M}$ y $a_0 z^{-N}$ para eliminar las potencias negativas se obtiene:

$$X(z) = \frac{N(z)}{D(z)} = \frac{b_0 z^{-M}}{a_0 z^{-N}} \cdot \frac{z^M + \frac{b_1}{b_0} z^{M-1} + \dots + \frac{b_M}{b_0}}{z^N + \frac{a_1}{a_0} z^{N-1} + \dots + \frac{a_N}{a_0}}$$

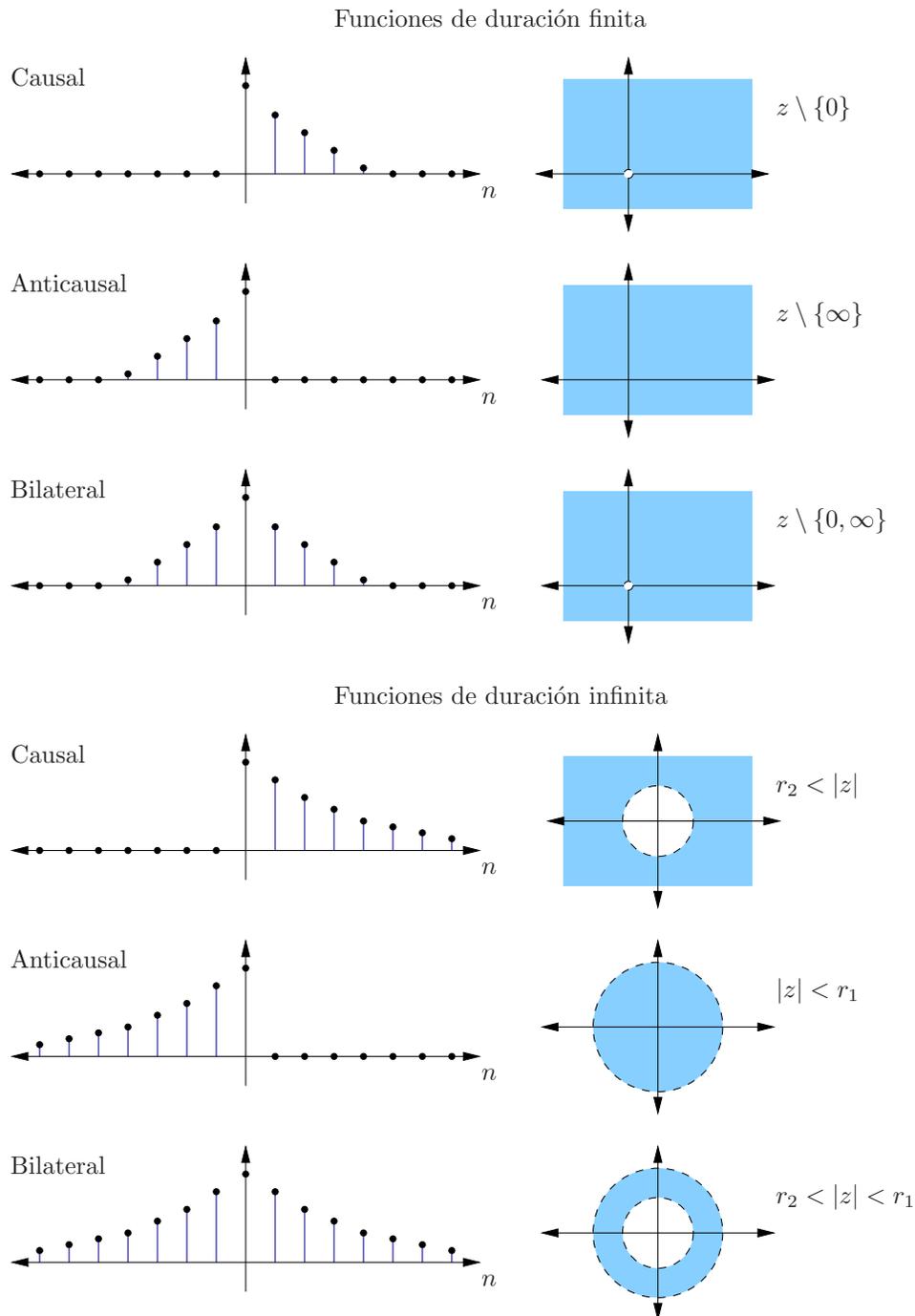


Figura 3.4: Familia de Señales y sus ROC[15].

Tabla 3.1: Transformada z de algunas funciones comunes

Señal $x(n)$	Transformada z , $X(z)$	ROC
$\delta(n)$	1	Plano z
$u(n)$	$\frac{1}{1 - z^{-1}}$	$ z > 1$
$a^n u(n)$	$\frac{1}{1 - az^{-1}}$	$ z > a $
$na^n u(n)$	$\frac{az^{-1}}{(1 - az^{-1})^2}$	$ z > a $
$-(a^n)u(-n - 1)$	$\frac{1}{1 - az^{-1}}$	$ z < a $
$-n(a^n)u(-n - 1)$	$\frac{az^{-1}}{(1 - az^{-1})^2}$	$ z < a $
$\cos(\omega_0 n)u(n)$	$\frac{1 - z^{-1} \cos \omega_0}{1 - 2z^{-1} \cos \omega_0 + z^{-2}}$	$ z > 1$
$\text{sen}(\omega_0 n)u(n)$	$\frac{z^{-1} \text{sen } \omega_0}{1 - 2z^{-1} \cos \omega_0 + z^{-2}}$	$ 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 > 1$
$a^n \text{sen}(\omega_0 n)u(n)$	$\frac{az^{-1} \text{sen } \omega_0}{1 - 2az^{-1} \cos \omega_0 + a^2 z^{-2}}$	$ z > 1$

Puesto que $N(z)$ y $D(z)$ son polinomios en z , y debido al teorema fundamental del álgebra, $X(z)$ se puede expresar como:

$$X(z) = \frac{N(z)}{D(z)} = \frac{b_0}{a_0} z^{N-M} \frac{(z - z_1)(z - z_2) \dots (z - z_M)}{(z - p_1)(z - p_2) \dots (z - p_N)} = G z^{N-M} \frac{\prod_{k=1}^M (z - z_k)}{\prod_{k=1}^N (z - p_k)}$$

donde la constante G es igual a $\frac{b_0}{a_0}$; los ceros $\{z_k\}$ y los polos $\{p_k\}$ corresponden a las raíces de los polinomios $N(z)$ y $D(z)$ respectivamente, y hay $N - M$ ceros en el origen si $N > M$, o $M - N$ polos en el origen si $M > N$. También se dice que hay un cero en infinito si $X(\infty) = 0$, o un polo si $X(\infty) = \infty$. Si se cuentan estos últimos, el número de polos y ceros es siempre idéntico.

$X(z)$ puede representarse mediante un diagrama de polos y ceros en el plano complejo z , donde los ceros se denotan con “o” y los polos con “x”. La multiplicidad se denota con un índice al lado del símbolo (por ejemplo, un polo de tercer orden se denota con \times^3). Evidentemente la ROC no puede contener ningún polo, puesto que en él la función no converge.

Si los coeficientes de los polinomios son reales, entonces sus raíces serán o reales, o aparecerán en pares complejos conjugados. De lo anterior se deriva que si un polinomio es de orden impar, deberá entonces tener al menos una raíz real. Por otro lado, para la factorización de

dicho polinomio se pueden agrupar las raíces complejas conjugadas para conformar términos de segundo orden con coeficientes reales.

Ejemplo 3.1 Determine el diagrama de polos y ceros de la señal $x(n) = a^n u(n)$, para a real positivo.

Solución:

Con $X(z) = \frac{1}{1-az^{-1}} = \frac{z}{z-a}$, ROC : $|z| > |a|$, se obtiene un cero en $z = 0$ y un polo en $z = a$ (figura 3.5).

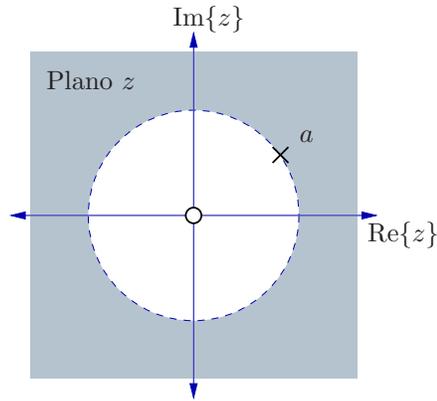


Figura 3.5: Diagrama de polos y ceros del ejemplo 3.1.

3.1

Ejemplo 3.2 Determine la transformada z de

$$x(n) = \begin{cases} a^n, & 0 \leq n \leq M-1 \\ 0 & \text{en el resto} \end{cases}$$

con a real positivo.

Solución:

$$X(z) = \sum_{n=-\infty}^{\infty} x(n)z^{-n} = \sum_{n=0}^{M-1} a^n z^{-n} = \sum_{n=0}^{M-1} (az^{-1})^n$$

Puesto que $\sum_{n=0}^{M-1} a^n = \frac{1-a^M}{1-a}$, entonces:

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

con $a = a \cdot 1 = a \cdot e^{j2k\pi}$, $z^M - a^M = 0 \Rightarrow z^M = a^M = a^M e^{j2k\pi}$ de donde se deduce que $z = ae^{j2\pi k/M}$. Es decir, $z^M - a^M$ tiene M raíces $ae^{j2\pi k/M}$, $k = 0, \dots, M-1$. En resumen,

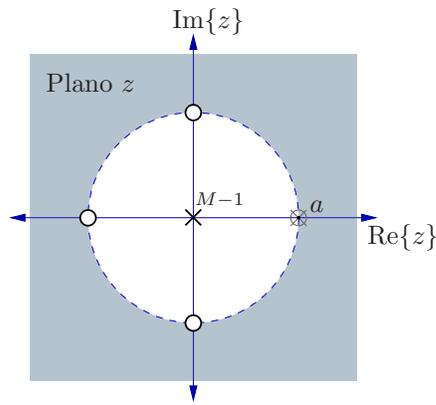


Figura 3.6: Diagrama de polos y ceros del ejemplo 3.2 con $M = 4$.

$X(z)$ tiene un polo en $z = 0$ con multiplicidad $M - 1$ y otro polo en $z = a$, que se cancela con el cero en a , y $M - 1$ ceros distribuidos en un círculo de radio $|a|$ (figura 3.6).

3.2

Ejemplo 3.3 Determine la transformada z y la señal correspondiente al diagrama de polos y ceros de la figura 3.7.

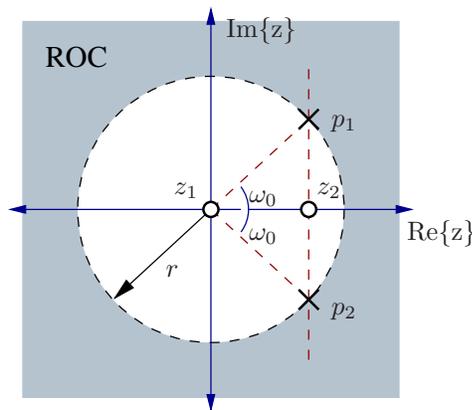


Figura 3.7: Diagrama de polos y ceros para el ejemplo (3.3).

Solución:

La transformada tiene dos ceros y dos polos: $z_1 = 0$, $z_2 = r \cos \omega_0$, $p_1 = r(\cos \omega_0 + j \sen \omega_0) = r e^{j\omega_0}$, $p_2 = r(\cos \omega_0 - j \sen \omega_0) = r e^{-j\omega_0}$.

$$\begin{aligned} X(z) &= G \frac{(z - z_1)(z - z_2)}{(z - p_1)(z - p_2)} = G \frac{z(z - r \cos \omega_0)}{(z - r e^{j\omega_0})(z - r e^{-j\omega_0})} \\ &= G \frac{1 - r z^{-1} \cos \omega_0}{1 - 2r z^{-1} \cos \omega_0 + r^2 z^{-2}}, \quad \text{ROC: } |z| > r \end{aligned}$$

De la tabla 3.1 se obtiene $x(n) = G(r^n \cos \omega_0 n)u(n)$

3.3

3.2.2 Localización de los polos y el comportamiento en el dominio de n para señales causales

Si una señal real tiene una transformada z con un solo polo, este tiene que ser real, y la señal correspondiente es entonces la exponencial real:

$$X(z) = \frac{1}{1 - az^{-1}} = \frac{z}{z - a}, \quad \text{ROC: } |z| > |a| \quad \bullet \rightarrow \circ x(n) = a^n u(n)$$

que tiene además un cero en $z = 0$. Nótese que $X_2(z) = \frac{1}{z-a}$, corresponde a $X(z)z^{-1}$ y por tanto la señal correspondiente a $X_2(z) \bullet \rightarrow \circ x_2(n) = x(n-1)$ que sigue siendo una exponencial real.

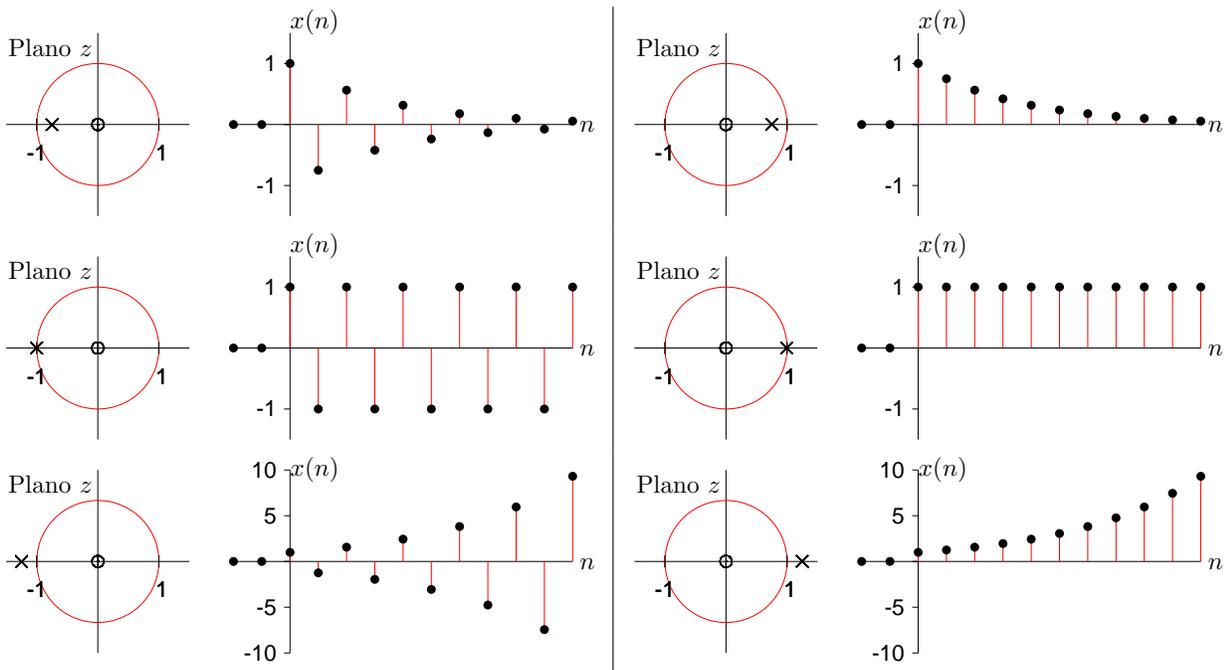


Figura 3.8: Comportamiento de la señal en relación con la posición de un polo real.

La figura 3.8 muestra el comportamiento de la señal en relación con la posición del polo real, incluyendo su magnitud y signo.

De la tabla 3.1, se obtiene que:

$$X(z) = \frac{az^{-1}}{(1 - az^{-1})^2} = \frac{az^{-1}}{z^{-2}(z - a)^2} = \frac{az}{(z - a)^2}, \quad \text{ROC: } |z| > |a| \quad \bullet \rightarrow \circ x(n) = na^n u(n)$$

es una función con un polo en $z = a$ de multiplicidad dos y un cero en $z = 0$. Su comportamiento se muestra en la figura 3.9.

Nótese que si $|a| = 1$ y el polo es simple, la señal corresponde a la respuesta al impulso de un sistema, éste es entonces condicionalmente estable, pero no así si el polo es de mayor multiplicidad.

Un par de polos complejos conjugados simples conducen a una señal discreta sinusoidal, multiplicada por una envolvente exponencial discreta cuya forma está determinada por la

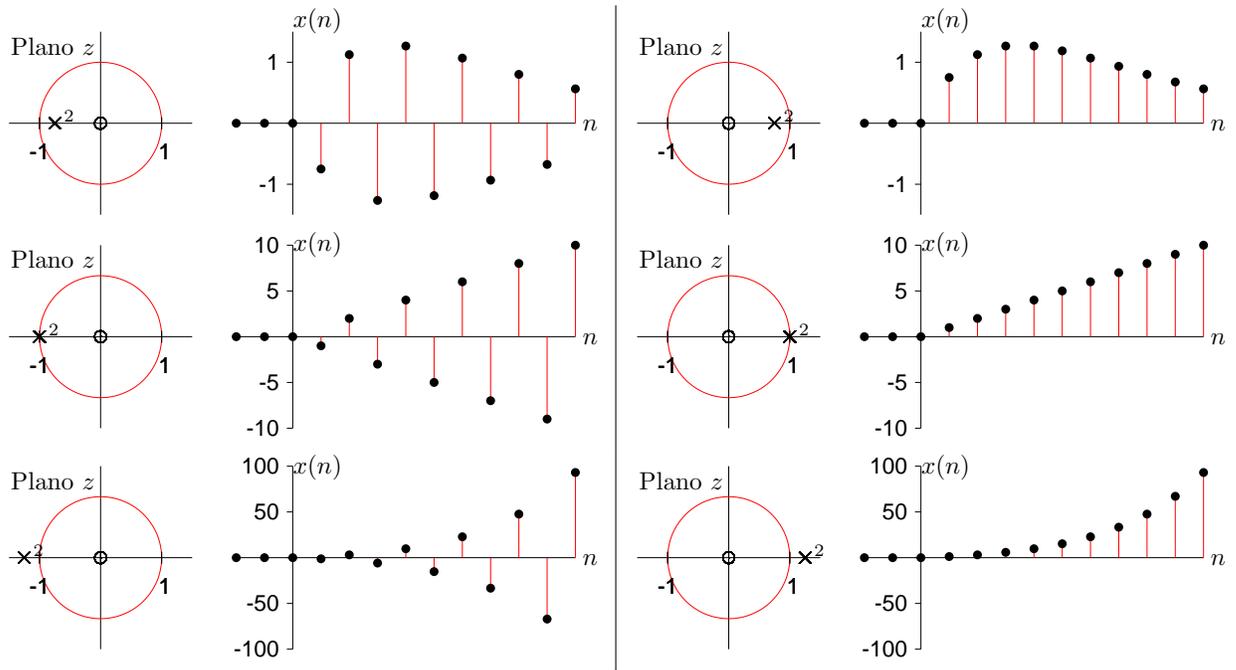


Figura 3.9: Comportamiento de la señal $x(n)$ si su transformada tiene un polo doble real, de acuerdo con la posición del polo.

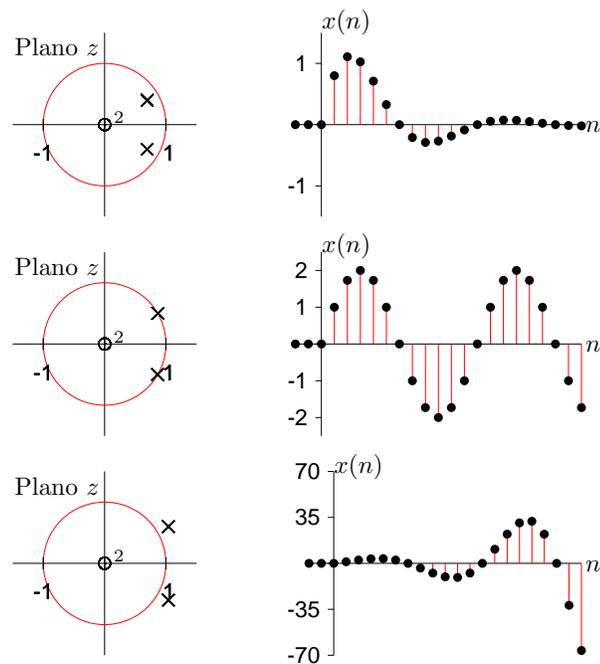


Figura 3.10: Señales en relación con la ubicación de un par de polos simples conjugados en su transformada z .

magnitud de los polos (figura 3.10). Las señales reales causales con polos reales simples o pares de polos simples complejos conjugados que están dentro de la circunferencia unitaria son acotadas en magnitud. Una señal con polos cerca del origen decrece más rápidamente que otra con polos más cercanos a la circunferencia unitaria. El efecto de la posición de los polos no es tan determinante como la posición de los polos. Por ejemplo, en el caso de las

señales senoidales, sólo afectan la fase.

3.2.3 La función de transferencia de un sistema LTI

En el capítulo 2 se encontró que la salida $y(n)$ de un sistema LTI con respuesta impulsional $h(n)$ y entrada $x(n)$ está dada por la convolución $y(n) = h(n) * x(n)$, que ya se demostró en el dominio z ser equivalente a $Y(z) = H(z)X(z)$, con $y(n) \circ \bullet Y(z)$, $x(n) \circ \bullet X(z)$ y $h(n) \circ \bullet H(z)$. A $H(z)$, la transformada z de la respuesta impulsional, se le denomina *función de transferencia* del sistema. Nótese que conociendo $H(z)$ y la transformada z de la entrada $X(z)$ es posible calcular fácilmente la transformada z de la salida del sistema, por medio de su producto. Sin embargo, es también posible, conociendo $Y(z)$ y $X(z)$, obtener la función de transferencia con:

$$H(z) = \frac{Y(z)}{X(z)}$$

a partir de la cual se puede determinar la respuesta impulsional por medio de la transformada z inversa.

Con la ecuación de diferencias del sistema

$$y(n) = - \sum_{k=1}^N a_k y(n-k) + \sum_{k=0}^M b_k x(n-k)$$

se obtiene transformando al dominio z ambos lados:

$$\begin{aligned} Y(z) &= - \sum_{k=1}^N a_k Y(z) z^{-k} + \sum_{k=0}^M b_k X(z) z^{-k} \\ &= -Y(z) \sum_{k=1}^N a_k z^{-k} + X(z) \sum_{k=0}^M b_k z^{-k} \\ Y(z) \left(1 + \sum_{k=1}^N a_k z^{-k} \right) &= X(z) \sum_{k=0}^M b_k z^{-k} \\ H(z) = \frac{Y(z)}{X(z)} &= \frac{\sum_{k=0}^M b_k z^{-k}}{1 + \sum_{k=1}^N a_k z^{-k}} \end{aligned}$$

Lo que implica que un sistema lineal descrito por una ecuación de diferencias lineal con coeficientes constantes tiene una función de transferencia racional.

Si $a_k = 0$ para $k \in [1, N]$ (sistema no recursivo), entonces

$$H(z) = \sum_{k=0}^M b_k z^{-k} = \frac{1}{z^M} \sum_{k=0}^M b_k z^{M-k}$$

es decir, el sistema tiene un polo en el origen de multiplicidad M y M ceros determinados por los parámetros $\{b_k\}$. Dado que el sistema contiene M polos triviales en $z = 0$ y M ceros

no triviales, se le denomina *sistema de todos ceros*. Un sistema así es evidentemente FIR y se le denomina sistema de media ponderada móvil.

Si $b_k = 0$ para $k \in [1, M]$ el sistema es recursivo y:

$$H(z) = \frac{b_0}{1 + \sum_{k=1}^N a_k z^{-k}} = \frac{b_0 z^N}{\sum_{k=0}^N a_k z^{N-k}}, \quad a_0 \stackrel{!}{=} 1$$

que tiene un cero trivial en $z = 0$ de orden N y N polos determinados por los coeficientes $\{a_k\}$. A este sistema se le denomina *sistema de todos polos*, que siempre corresponde a un sistema IIR.

La forma general se denomina *sistema de polos y ceros*, con N polos y M ceros. Los polos y ceros en $z = 0$ y $z = \infty$, usualmente no se consideran de forma explícita. Por la presencia de polos, este sistema es también IIR.

Ejemplo 3.4 Determine la función de transferencia y la respuesta impulsional del sistema descrito por:

$$y(n) = \frac{1}{2}y(n-1) + 2x(n)$$

Solución:

$$\begin{aligned} Y(z) &= \frac{1}{2}Y(z)z^{-1} + 2X(z) \\ Y(z) \left(1 - \frac{1}{2}z^{-1}\right) &= 2X(z) \\ H(z) = \frac{Y(z)}{X(z)} &= \frac{2}{1 - \frac{1}{2}z^{-1}} = \frac{2z}{z - \frac{1}{2}} \end{aligned}$$

que tiene un polo en $z = \frac{1}{2}$ y un cero en $z = 0$. Utilizando la tabla de transformadas se obtiene:

$$H(z) \bullet \circ h(n) = 2 \left(\frac{1}{2}\right)^n u(n)$$

3.4

Ejemplo 3.5 Asíumase que un sistema se rige por el diagrama de polos y ceros mostrado en la figura 3.7. Encuentre la ecuación de diferencias que lo realiza.

Solución:

En el diagrama se observa una región de convergencia externa al círculo de radio r , lo que implica que el sistema es causal.

Utilizando los resultados del ejemplo 3.3 se obtiene que la función de transferencia equivalente al diagrama de polos y ceros corresponde a:

$$\begin{aligned}
 H(z) &= \frac{Y(z)}{X(z)} = G \frac{(z - z_1)(z - z_2)}{(z - p_1)(z - p_2)} \\
 &= G \frac{1 - r \cos \omega_0 z^{-1}}{1 - 2r \cos \omega_0 z^{-1} + r^2 z^{-2}}, \quad \text{ROC: } |z| > r
 \end{aligned}$$

y por tanto

$$\begin{aligned}
 Y(z) (1 - 2r \cos \omega_0 z^{-1} + r^2 z^{-2}) &= X(z) G (1 - r \cos \omega_0 z^{-1}) \\
 Y(z) - 2r \cos \omega_0 Y(z) z^{-1} + r^2 Y(z) z^{-2} &= G X(z) - G r \cos \omega_0 X(z) z^{-1}
 \end{aligned}$$

•
|
○

$$y(n) - 2r \cos \omega_0 y(n-1) + r^2 y(n-2) = G x(n) - G r \cos \omega_0 x(n-1)$$

y finalmente

$$y(n) = 2r \cos \omega_0 y(n-1) - r^2 y(n-2) + G x(n) - G r \cos \omega_0 x(n-1)$$

3.5

3.3 Análisis de sistemas LTI en el dominio z

3.3.1 Respuesta de sistemas con función de transferencia racional

Sea un sistema dado por la ecuación de diferencias:

$$y(n) = - \sum_{k=1}^N a_k y(n-k) + \sum_{k=0}^M b_k x(n-k)$$

Su función de transferencia $H(z)$ es racional de la forma $H(z) = \frac{B(z)}{A(z)}$, donde el polinomio $B(z)$ determina los ceros y $A(z)$ los polos de $H(z)$.

Si se asume además que el sistema está en reposo, esto es, $y(-1) = y(-2) = \dots = y(-N) = 0$, y que la entrada $X(z)$ también puede representarse en forma racional como $X(z) = \frac{N(z)}{Q(z)}$, entonces se tiene que la transformada de la función de salida es de la forma:

$$Y(z) = \frac{B(z)N(z)}{A(z)Q(z)}$$

Supóngase ahora que $A(z) = \prod_{k=1}^N (1 - p_k z^{-1})$, $Q(z) = \prod_{k=1}^L (1 - q_k z^{-1})$ y $p_k \neq q_m, \forall k, m$, y que ningún cero coincide con los polos, entonces se cumple que:

$$Y(z) = \sum_{k=1}^N \frac{A_k}{1 - p_k z^{-1}} + \sum_{k=1}^L \frac{Q_k}{1 - q_k z^{-1}}$$

y por lo tanto su transformada inversa será:

$$y(n) = \underbrace{\sum_{k=1}^N A_k p_k^n u(n)}_{\text{respuesta natural}} + \underbrace{\sum_{k=1}^L Q_k q_k^n u(n)}_{\text{respuesta forzada}}$$

Esta secuencia se compone de dos partes: las respuestas natural y forzada. La *respuesta natural* depende de los polos $\{p_k\}$ del sistema y es influenciada por la entrada $x(n)$ a través de los coeficientes $\{A_k\}$. La *respuesta forzada* depende de los polos de la entrada $\{q_k\}$ y es influenciada por el sistema a través de los coeficientes $\{Q_k\}$.

Si $X(z)$ ó $H(z)$ tienen polos múltiples o polos en común, entonces la respuesta $Y(z)$ tendrá términos de la forma $\sum_{l=1}^m \frac{1}{(1-p_k z^{-1})^l}$, donde m es el orden del polo, lo que conducirá a términos en $y(n)$ de la forma $n^{l-1} p_k^n u(n)$.

3.3.2 Condiciones iniciales no nulas

Para calcular la respuesta de un sistema a una entrada causal $x(n)$ con condiciones iniciales no nulas, se utiliza la transformada z unilateral:

$$y(n) = - \sum_{k=1}^N a_k y(n-k) + \sum_{k=1}^M b_k x(n-k)$$

○
●

$$Y^+(z) = - \sum_{k=1}^N a_k z^{-k} \left[Y^+(z) + \sum_{n=1}^k y(-n) z^n \right] + \sum_{k=1}^M b_k z^{-k} \left[X^+(z) + \underbrace{\sum_{n=1}^k x(-n) z^n}_{=0 \text{ por causalidad}} \right]$$

Puesto que $x(n)$ es causal, entonces $X^+(z) = X(z)$ y se tiene:

$$Y^+(z) \left(1 + \sum_{k=1}^N a_k z^{-k} \right) = X^+(z) \sum_{k=0}^M b_k z^{-k} - \sum_{k=1}^N a_k z^{-k} \sum_{n=1}^k y(-n) z^n,$$

de donde se despeja:

$$\begin{aligned} Y^+(z) &= \frac{\sum_{k=0}^M b_k z^{-k}}{1 + \sum_{k=1}^N a_k z^{-k}} X(z) - \frac{\sum_{k=1}^N a_k z^{-k} \sum_{n=1}^k y(-n) z^n}{1 + \sum_{k=1}^N a_k z^{-k}} \\ &= H(z) X(z) + \frac{N_0(z)}{A(z)} \end{aligned}$$

La salida del sistema consta entonces de dos partes, la respuesta de estado nulo

$$Y_{zs}(z) = H(z) X(z)$$

y la respuesta de entrada nula que depende de las condiciones iniciales

$$Y_{zi}^+(z) = \frac{N_0(z)}{A(z)}$$

puesto que $Y^+(z) = Y_{zs}(z) + Y_{zi}^+(z) \bullet \xrightarrow{zu} y(n) = y_{zs}(n) + y_{zi}(n)$. Al ser $A(z)$ el denominador de $Y_{zi}^+(z)$ sus polos serán $\{p_k\}$, y la respuesta a entrada nula tendrá la forma:

$$y_{zi} = \sum_{k=1}^N D_k p_k^n u(n)$$

lo que puede añadirse a la respuesta natural y combinarse para dar:

$$y(n) = \sum_{k=1}^N A'_k p_k^n u(n) + \sum_{k=1}^L Q_k q_k^n u(n), \text{ con } A'_k = A_k + D_k$$

En resumen, las condiciones iniciales alteran la respuesta natural del sistema modificando los factores de escala $\{A_k\}$. No se introducen nuevos polos ni se modifica la respuesta forzada.

Ejemplo 3.6 Determine la respuesta al escalón $u(n)$ del sistema dado por

$$y(n) = 0,9y(n-1) - 0,81y(n-2) + x(n)$$

para las condiciones iniciales:

1. $y(-1) = y(-2) = 0$
2. $y(-1) = y(-2) = 1$

Solución: La función de transferencia está dada por:

$$H(z) = \frac{1}{1 - 0,9z^{-1} + 0,81z^{-2}}$$

cuyos polos son complejos conjugados: $p_{1,2} = 0,9e^{\pm j\frac{\pi}{3}}$

$$x(n) = u(n) \circ \bullet \frac{1}{1 - z^{-1}}$$

Con lo que se obtiene directamente:

$$\begin{aligned} Y_{zs}(z) &= \frac{1}{(1 - 0,9e^{j\frac{\pi}{3}}z^{-1})(1 - 0,9e^{-j\frac{\pi}{3}}z^{-1})(1 - z^{-1})} \\ &= \frac{0,542 - j0,049}{1 - 0,9e^{j\frac{\pi}{3}}z^{-1}} + \frac{0,542 + j0,049}{1 - 0,9e^{-j\frac{\pi}{3}}z^{-1}} + \frac{1,099}{1 - z^{-1}} \end{aligned}$$

Por lo que:

$$y_{zs}(n) = \left[1,099 + 1,088(0,9)^n \cos\left(\frac{\pi}{3}n - 5,2^\circ\right) \right] u(n)$$

Para 1. $y(n) = y_{zs}(n)$

Para 2., con las condiciones iniciales se obtiene:

$$\begin{aligned}
 N_0 &= -\sum_{k=1}^N a_k z^{-k} \sum_{n=1}^k y(-n) z^n \\
 &= -(a_1 z^{-1} \{y(-1)z\} + a_2 z^{-2} \{y(-1)z + y(-2)z^2\}) \\
 &= -(a_1 y(-1) + a_2 y(-1)z^{-1} + a_2 y(-2)) \\
 &= -(-0,9 + 0,81 + 0,81z^{-1}) \\
 &= 0,09 - 0,81z^{-1} \\
 Y_{zi}(z) &= \frac{N_0(z)}{A(z)} = \frac{0,09 - 0,81z^{-1}}{1 - 0,9z^{-1} + 0,81z^{-2}} \\
 &= \frac{0,026 + j0,4936}{1 - 0,9ze^{j\frac{\pi}{3}}z^{-1}} + \frac{0,026 - j0,4936}{1 - 0,9ze^{-j\frac{\pi}{3}}z^{-1}}
 \end{aligned}$$

y por tanto:

$$y_{zi}(n) = 0,988(0,9)^n \cos\left(\frac{\pi}{3}n + 87^\circ\right) u(n)$$

y $Y(z) = Y_{zs}(z) + Y_{zi}(z)$

$$Y(z) = \frac{1,099}{1 - z^{-1}} + \frac{0,568 + j0,445}{1 - 0,9ze^{j\frac{\pi}{3}}z^{-1}} + \frac{0,568 - j0,442}{1 - 0,9ze^{-j\frac{\pi}{3}}z^{-1}}$$

que equivale a:

$$y(n) = 1,099u(n) + 1,44(0,9)^n \cos\left(\frac{\pi}{3}n + 38^\circ\right) u(n)$$

3.6

3.3.3 Respuesta transitoria y en régimen permanente

Si para todos los polos del sistema $\{p_k\}$ se cumple que $|p_k| < 1$, entonces la respuesta natural del sistema:

$$y_{nr}(n) = \sum_{k=1}^N A_k p_k^n u(n)$$

decae si n tiende a infinito, y se le denomina *respuesta transitoria*. Su tasa de decaimiento dependerá de qué tan cerca estén los polos de la circunferencia unidad. Si los polos $\{q_k\}$ de la respuesta forzada tienen magnitudes $|q_k| < 1$, la respuesta a la entrada también decae. Si la entrada es sinusoidal implica que hay polos en la circunferencia unitaria, y la respuesta forzada también tendrá esos polos y será sinusoidal, y la respuesta se denomina *de estado permanente*.

3.3.4 Causalidad y Estabilidad

Un sistema es causal si $h(n) = 0$ para $n < 0$, lo que implica que su ROC es el exterior de una circunferencia. El sistema es estable si:

$$\sum_{n=-\infty}^{\infty} |h(n)| < \infty$$

lo que ocurre sólo si $|z| = 1$ está dentro de la ROC, o de otro modo, si los polos se encuentran dentro de la circunferencia unitaria, puesto que

$$H(z) = \sum_{n=-\infty}^{\infty} h(n)z^{-n}$$

$$|H(z)| \leq \sum_{n=-\infty}^{\infty} |h(n)||z^{-n}|$$

y evaluando en $|z| = 1$

$$|H(z)|_{|z|=1} \leq \sum_{n=-\infty}^{\infty} |h(n)|$$

Un sistema LTI será entonces estable sólo si $|z| = 1$ está en la ROC de $H(z)$. Esta condición también se cumple para sistemas anticausales, es decir, la ROC deberá ser el interior de una circunferencia que contenga $|z| = 1$.

3.3.5 Cancelación polo-cero

Los efectos de un polo del sistema pueden cancelarse o reducirse con un cero ya sea del mismo sistema o de la entrada. Sin embargo, posibles efectos como la estabilización de un sistema inestable por medio de ceros en la entrada deben evitarse, porque la posición de los polos y ceros no puede alcanzarse con precisión arbitraria en sistemas digitales reales.

3.3.6 Polos de orden múltiple y estabilidad

Un sistema con polos en la circunferencia unitaria es inestable pues basta encontrar una entrada con un polo en el mismo sitio para producir una salida no acotada.

Ejemplo 3.7 Determine la respuesta al escalón de:

$$y(n) = y(n-1) + x(n)$$

Solución:

Con $H(z) = \frac{1}{1-z^{-1}}$ y $X(z) = \frac{1}{1-z^{-1}}$ se obtiene:

$$Y(z) = H(z)X(z) = \frac{1}{(1-z^{-1})^2} \bullet \circ y(n) = (n+1)u(n+1)$$

que contiene un polo doble en $z = 1$, y equivale a una rampa no acotada.

Esto se explica por el hecho de que si $Y(z)$ contiene polos múltiples, entonces $y(n)$ tendrá términos de la forma:

$$A_k n^b (p_k)^n u(n)$$

que son acotados sólo si $|p_k| < 1$.

3.7

3.3.7 Estabilidad de sistemas de segundo orden

Los sistemas de segundo orden (con dos polos) se utilizan como bloque básico para la construcción de sistemas de mayor orden, y por ello requieren de un poco de atención.

Considérese entonces el sistema:

$$y(n) = -a_1 y(n-1) - a_2 y(n-2) + b_0 x(n)$$

con $a_1, a_2 \in \mathbb{R}$. Su función de transferencia es

$$\begin{aligned} H(z) &= \frac{Y(z)}{X(z)} = \frac{b_0}{1 + a_1 z^{-1} + a_2 z^{-2}} = \frac{b_0 z^2}{z^2 + a_1 z + a_2} \\ &= \frac{b_0 z^2}{(z - p_1)(z - p_2)} = \frac{b_0 z^2}{z^2 - (p_1 + p_2)z + p_1 p_2} \end{aligned}$$

Este sistema tiene dos ceros en el origen y dos polos en

$$p_{1,2} = -\frac{a_1}{2} \pm \sqrt{\frac{a_1^2 - 4a_2}{4}}$$

El sistema es estable BIBO si $|p_{1,2}| < 1$. Puesto que se cumple que $a_1 = -(p_1 + p_2)$ y $a_2 = p_1 p_2$, se debe cumplir que

$$|a_2| = |p_1 p_2| = |p_1| |p_2| < 1$$

y

$$|a_1| < 1 + a_2 \Rightarrow a_2 > |a_1| - 1$$

En el plano a_1, a_2 estas inecuaciones delimitan una región triangular (figura 3.11).

Para obtener estas ecuaciones se sabe que $|p_{1,2}| < 1$. Primero asúmase que los polos son reales, es decir, $\left(\frac{a_1}{2}\right)^2 \geq a_2$, que equivale a la parte inferior de la parábola $a_2 = \left(\frac{a_1}{2}\right)^2$. En esta región, los polos están centrados en $-\frac{a_1}{2}$, con un desplazamiento hacia la izquierda y derecha de $\sqrt{\left(\frac{a_1}{2}\right)^2 - a_2}$ (figura 3.12).

Por lo tanto, debe cumplirse que:

$$1 - \frac{|a_1|}{2} > \sqrt{\left(\frac{a_1}{2}\right)^2 - a_2}$$

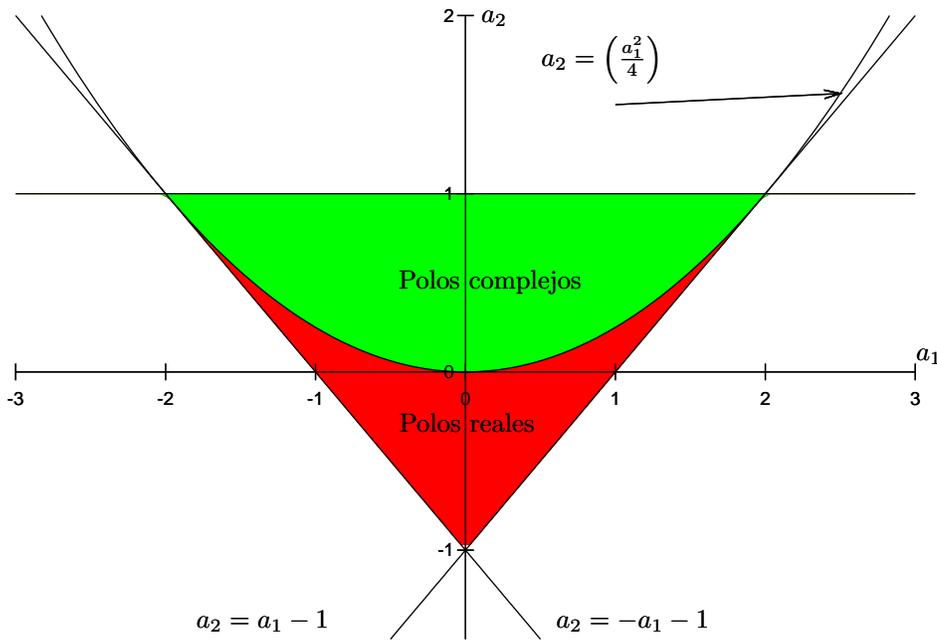


Figura 3.11: Triángulo de estabilidad en el plano de coeficientes (a_1, a_2) para un sistema de segundo orden.

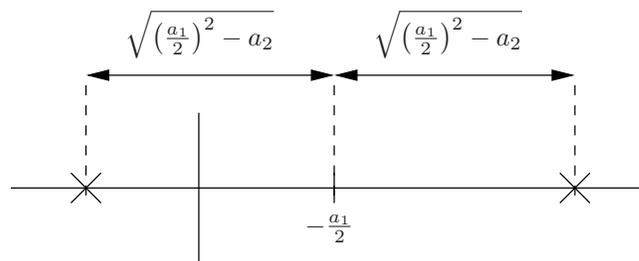


Figura 3.12: Desplazamiento de los polos a la izquierda y a la derecha.

Elevando al cuadrado ambos lados y considerando que $a_1 \in \mathbb{R}$ se obtiene:

$$1 - |a_1| + \left(\frac{|a_1|}{2}\right)^2 > \left(\frac{a_1}{2}\right)^2 - a_2 \quad \Rightarrow \quad 1 + a_2 > |a_1|$$

En caso de que $p_{1,2}$ sean complejos (nótese que a_2 es positivo puesto que debe ser mayor que $(a_1/2)^2$), entonces:

$$|p_{1,2}|^2 = \left| -\frac{a_1}{2} \pm j\sqrt{a_2 - \left(\frac{a_1}{2}\right)^2} \right|^2 = \left(\frac{a_1}{2}\right)^2 + a_2 - \left(\frac{a_1}{2}\right)^2 = a_2$$

y por tanto, para que $|p_{1,2}| < 1$, entonces $0 < a_2 < 1$, lo que confirma la anterior condición $|a_2| < 1$.

Los pares (a_1, a_2) en la parte superior a la parábola conducen entonces a polos complejos conjugados, y bajo la parábola a polos reales y distintos. Los pares (a_1, a_2) en la parábola $\left(\left(\frac{a_1}{2}\right)^2 = a_2\right)$ conducen a un polo de orden 2.

Las funciones de transferencia y sus equivalentes en el dominio del tiempo se resumen en la tabla 3.2, con $p = re^{j\omega_0}$, por lo que $a_1 = -2r \cos \omega_0$, $a_2 = r^2$.

Tabla 3.2: Funciones de transferencia de segundo orden y equivalentes temporales

Polos	Condición	$H(z)$	$h(n)$
Reales y distintos	$a_1^2 > 4a_2$	$\frac{b_0}{p_1 - p_2} \left(\frac{p_1}{1 - p_1 z^{-1}} - \frac{p_2}{1 - p_2 z^{-1}} \right)$	$\frac{b_0}{p_1 - p_2} (p_1^{n+1} - p_2^{n+1}) u(n)$
Reales e iguales	$a_1^2 = 4a_2$	$\frac{b_0}{(1 - pz^{-1})^2}$	$b_0(n+1)p^n u(n)$
Complejos conjugados	$a_1^2 < 4a_2$	$\frac{b_0}{j2 \operatorname{sen} \omega_0} \left(\frac{e^{j\omega_0}}{1 - re^{j\omega_0} z^{-1}} - \frac{e^{-j\omega_0}}{1 + re^{-j\omega_0} z^{-1}} \right)$	$\frac{b_0 r^n}{\operatorname{sen} \omega_0} \operatorname{sen}[(n+1)\omega_0] u(n)$

3.4 Problemas

Problema 3.1. Grafique las siguientes funciones e indique cualitativamente qué regiones de convergencia (ROC) tiene su transformada z :

1. $x(n) = \text{sen}(\omega n)u(n)$
2. $x(n) = u(n+4) - u(n-2)$
3. $x(n) = u(-n-2)$
4. $x(n) = u_r(n) - 2u_r(n-5) + u_r(n-10)$
5. $x(n) = \left(-\frac{1}{2}\right)^{-|n|}$
6. $x(n) = u_r(n+5)u(-n-5)$

Problema 3.2. Encuentre las regiones del plano z donde las siguientes series convergen:

1. $\sum_{n=-2}^{\infty} \left(\frac{1}{3}\right)^{n+2} z^{-n}$
2. $\sum_{n=0}^{\infty} \left[\frac{1+(-1)^n}{2}\right] z^{-n}$
3. $\sum_{n=2}^{\infty} \left(\frac{1}{3}\right)^{-n+2} z^n$
4. $\sum_{n=-\infty}^{\infty} \left(\frac{1}{3}\right)^{|n|} \cos\left(\frac{\pi}{4}n\right) z^n$

Problema 3.3. Encuentre la transformada z de

$$x(n) = \frac{u(n-2)}{4^n}$$

con su correspondiente ROC.

Problema 3.4. Sea

$$x(n) = (-1)^n u(n) + \alpha^n u(-n - n_0)$$

Encuentre para qué valores de α y n_0 es la ROC de la transformada z de $x(n)$

$$1 < |z| < 2$$

Problema 3.5. Encuentre la transformada z de

$$x(n) = \begin{cases} \left(\frac{1}{2}\right)^n \cos\left(\frac{\pi}{4}n\right) & n \leq 0 \\ 0 & n > 0 \end{cases}$$

Indique los polos, ceros y ROC.

Problema 3.6. Para las siguientes expresiones identifique los ceros y polos finitos e infinitos.

1. $\frac{z^{-1} \left(1 - \frac{1}{2}z^{-1}\right)}{\left(1 - \frac{1}{3}z^{-1}\right) \left(1 - \frac{1}{4}z^{-1}\right)}$
2. $\frac{(1 - z^{-1})(1 - 2z^{-1})}{(1 - 3z^{-1})(1 - 4z^{-1})}$
3. $\frac{z^{-2} (1 - z^{-1})}{\left(1 - \frac{1}{4}z^{-1}\right) \left(1 + \frac{1}{4}z^{-1}\right)}$

Problema 3.7. Si $x(n)$ es absolutamente sumable y tiene transformada z racional, con un polo en $1/2$, entonces ¿podría $x(n)$ ser

1. una señal finita?
2. una señal izquierda?
3. una señal derecha?
4. una señal bilateral?

Problema 3.8. Sea

$$X(z) = \frac{1 - \frac{1}{4}z^{-2}}{\left(1 + \frac{1}{4}z^{-2}\right) \left(1 + \frac{5}{4}z^{-1} + \frac{3}{8}z^{-2}\right)}$$

Indique cuántas y cuáles regiones de convergencia son posibles para $X(z)$.

Problema 3.9. Encuentre para todas las señales discretas $x(n)$ mostradas en la tabla 3.1 la transformada z correspondiente utilizando la definición.

Problema 3.10. Sea $x(n)$ una señal con transformada z racional $X(z)$, que tiene un polo en $z = 1/2$. Se sabe además que

$$x_1(n) = \left(\frac{1}{4}\right)^n x(n)$$

es absolutamente sumable, pero

$$x_2(n) = \left(\frac{1}{8}\right)^n x(n)$$

no es absolutamente sumable. Con esta información indique si $x(n)$ es izquierda, derecha, bilateral o finita.

Problema 3.11. Encuentre las funciones en tiempo discreto equivalentes a las transformadas z indicadas en la tabla 3.1 utilizando la definición integral de la transformada z inversa.

Problema 3.12. Utilizando la definición de la transformada z inversa, encuentre la secuencia en el tiempo discreto equivalente a

$$X(z) = \frac{1 - \frac{1}{3}z^{-1}}{(1 - z^{-1})(1 + 2z^{-1})}, \quad \text{ROC: } |z| > 2$$

Problema 3.13. Encuentre por división polinomial la transformada z inversa de

$$X(z) = \frac{1 + z^{-1}}{1 + \frac{1}{3}z^{-1}}$$

para ROC: $|z| > 1/3$ y para ROC: $|z| < 1/3$.

Problema 3.14. Encuentre la transformada inversa de

$$X(z) = \frac{1 - \frac{1}{3}z^{-1}}{(1 - z^{-1})(1 + 2z^{-1})}$$

para todas las posibles regiones de convergencia por medio de descomposición en fracciones parciales.

Problema 3.15. Encuentre la transformada z inversa de

$$X(z) = \frac{1}{256} \left[\frac{256 - z^{-8}}{1 - \frac{1}{2}z^{-1}} \right], \quad \text{ROC: } |z| > 0$$

Problema 3.16. Para la ventana rectangular

$$x(n) = \begin{cases} 1 & 0 \leq n \leq k \\ 0 & \text{en el resto} \end{cases}$$

sea

$$g(n) = x(n) - x(n - 1)$$

1. Encuentre una expresión para $g(n)$ y su transformada z .
2. Encuentre la transformada z de $x(n)$ considerando que

$$x(n) = \sum_{k=-\infty}^n g(k)$$

Problema 3.17. Para las siguientes funciones de transferencia de sistemas discretos, si se sabe que estos son estables indique si además son causales:

1. $\frac{1 - \frac{4}{3}z^{-1} + \frac{1}{2}z^{-2}}{z^{-1} \left(1 - \frac{1}{2}z^{-1}\right) \left(1 - \frac{1}{3}z^{-1}\right)}$
2. $\frac{z - \frac{1}{2}}{z^2 + \frac{1}{2}z - \frac{3}{16}}$
3. $\frac{z + 1}{z + \frac{4}{3} - \frac{1}{2}z^{-2} - \frac{2}{3}z^{-3}}$

Problema 3.18. Un sistema LTI tiene función de transferencia $H(z)$ y respuesta al impulso $h(n)$. Se sabe

1. $h(n)$ es real
2. $h(n)$ es derecha
3. $\lim_{z \rightarrow \infty} H(z) = 1$

4. $H(z)$ tiene dos ceros
5. $H(z)$ tiene uno de sus polos en una ubicación no real en el círculo $|z| = 3/4$

¿Es el sistema causal? ¿Es estable?

Problema 3.19. Encuentre la transformada z unilateral de las siguientes señales.

1. $x_1(n) = \left(\frac{1}{4}\right)^n u(n+5)$
2. $x_2(n) = \delta(n+3) + \delta(n) + 2^n u(-n)$
3. $x_3(n) = \left(\frac{1}{2}\right)^{|n|}$

Problema 3.20. Un sistema de entrada $x(n)$ y salida $y(n)$ se rige por la ecuación de diferencias:

$$y(n-1) + 2y(n) = x(n)$$

1. Determine la respuesta de entrada cero al sistema si su condición inicial es $y(-1) = 2$.
2. Encuentra la respuesta de estado cero si su entrada es $x(n) = (1/4)^n u(n)$.
3. Determine la salida del sistema para $n \geq 0$ si $y(-1) = 2$ y $x(n) = (1/4)^n u(n)$

Capítulo 4

Análisis frecuencial

El cursos anteriores se derivó la descomposición de una función periódica en componentes sinusoidales (o exponenciales complejos). La importancia de esta descomposición es que en sistemas LTI únicamente las señales sinusoidales y exponenciales complejas conservan su forma y su frecuencia (son *funciones propias*). Tan solo magnitud y fase de cada componente son alteradas por estos sistemas.

El análisis frecuencial proporciona entonces mecanismos para la extracción de esas componentes. El término *espectro* se utiliza para referirse al contenido de frecuencias de una señal. En la práctica, puesto que no se conoce toda la señal en su extensión infinita, pasada y futura, debe utilizarse la *estimación espectral*, para la que se utilizan instrumentos y algoritmos denominados *analizadores espectrales*.

4.1 Espectro de señales continuas

4.1.1 Espectro de señales continuas periódicas

Las señales continuas periódicas se analizan por medio de la serie de Fourier. Para la síntesis de la señal se utiliza

$$x(t) = \sum_{k=-\infty}^{\infty} c_k e^{j\Omega_0 kt}$$

y para el análisis:

$$c_k = \frac{1}{T_p} \int_{t_0}^{t_0+T_p} x(t) e^{-j\Omega_0 kt} dt$$

Si la señal periódica tiene potencia media finita P_x entonces

$$P_x = \frac{1}{T_p} \int_{t_0}^{t_0+T_p} |x(t)|^2 dt$$

y puesto que $|x(t)|^2 = x(t)x^*(t)$ se cumple que

$$\begin{aligned}
 P_x &= \frac{1}{T_p} \int_{t_0}^{t_0+T_p} |x(t)|^2 dt = \frac{1}{T_p} \int_{t_0}^{t_0+T_p} x(t)x^*(t) dt \\
 &= \frac{1}{T_p} \int_{t_0}^{t_0+T_p} x(t) \sum_{k=-\infty}^{\infty} c_k^* e^{-j\Omega_0 kt} dt \\
 &= \sum_{k=-\infty}^{\infty} c_k^* \left[\frac{1}{T_p} \int_{t_0}^{t_0+T_p} x(t) e^{-j\Omega_0 kt} dt \right] \\
 &= \sum_{k=-\infty}^{\infty} c_k^* c_k = \sum_{k=-\infty}^{\infty} |c_k|^2
 \end{aligned}$$

que se conoce como la relación de Parseval, e implica que la potencia media de la señal equivale a la suma de las potencias medias de cada uno de sus componentes frecuenciales, también llamados *armónicos*. A la gráfica de $|c_k|^2$ para cada frecuencia kF_0 (con la frecuencia fundamental $F_0 = \Omega_0/2\pi$) se le denomina densidad espectral de potencia, espectro de la densidad de potencia o simplemente espectro de potencia. Si $x(t)$ es real, puesto que $|c_{-k}| = |c_k^*| = |c_k|$ entonces la densidad espectral de potencia es una función de simetría par.

Al par de gráficas de $|c_k|$ vs. kF_0 y $\theta_k = \angle c_k$ vs. kF_0 se le denomina espectro de tensión. Puesto que para funciones reales $\theta_{-k} = \angle c_k^* = -\angle c_k = -\theta_k$, la fase es impar.

Ejemplo 4.1 La serie de Fourier de un tren periódico de pulsos rectangulares de ancho τ en tiempo continuo (figura 4.1a) tiene como coeficientes

$$c_k = \frac{A\tau \operatorname{sen}(\pi k F_0 \tau)}{T_p \pi k F_0 \tau}.$$

Analice el efecto del periodo T_p y el ancho del pulso τ en el espectro de la señal.

Solución: La distancia entre dos líneas espectrales correspondientes a c_k y c_{k+1} es $F_0 = 1/T_p$ y el término τ , además de determinar el ancho del pulso temporal, indica qué tan extensa resulta la función $\operatorname{sen}(x)/x$. Las figuras 4.1(b) y (c) permiten comparar que si se modifica el ancho del pulso temporal manteniendo el periodo, la distancia entre cada línea espectral se mantiene, mientras que el espectro se contrae (si τ aumenta). Por otro lado, si T_p aumenta (o lo que es equivalente, la frecuencia fundamental F_0 baja), pero se mantiene el ancho del pulso τ , la misma función en la frecuencia $\operatorname{sen}(\pi k F_0 \tau)/(\pi k F_0 \tau)$ es muestreada más a menudo (la densidad de líneas espectrales aumenta).

4.1

La tabla 4.1 resume algunas propiedades de la serie de Fourier.

4.1.2 Espectro de señales continuas aperiódicas

La distancia entre dos líneas espectrales en el caso de señales periódicas es $F_0 = 1/T_p$, lo que implica que a mayor periodo, las líneas espectrales se acercan (compare por ejemplo

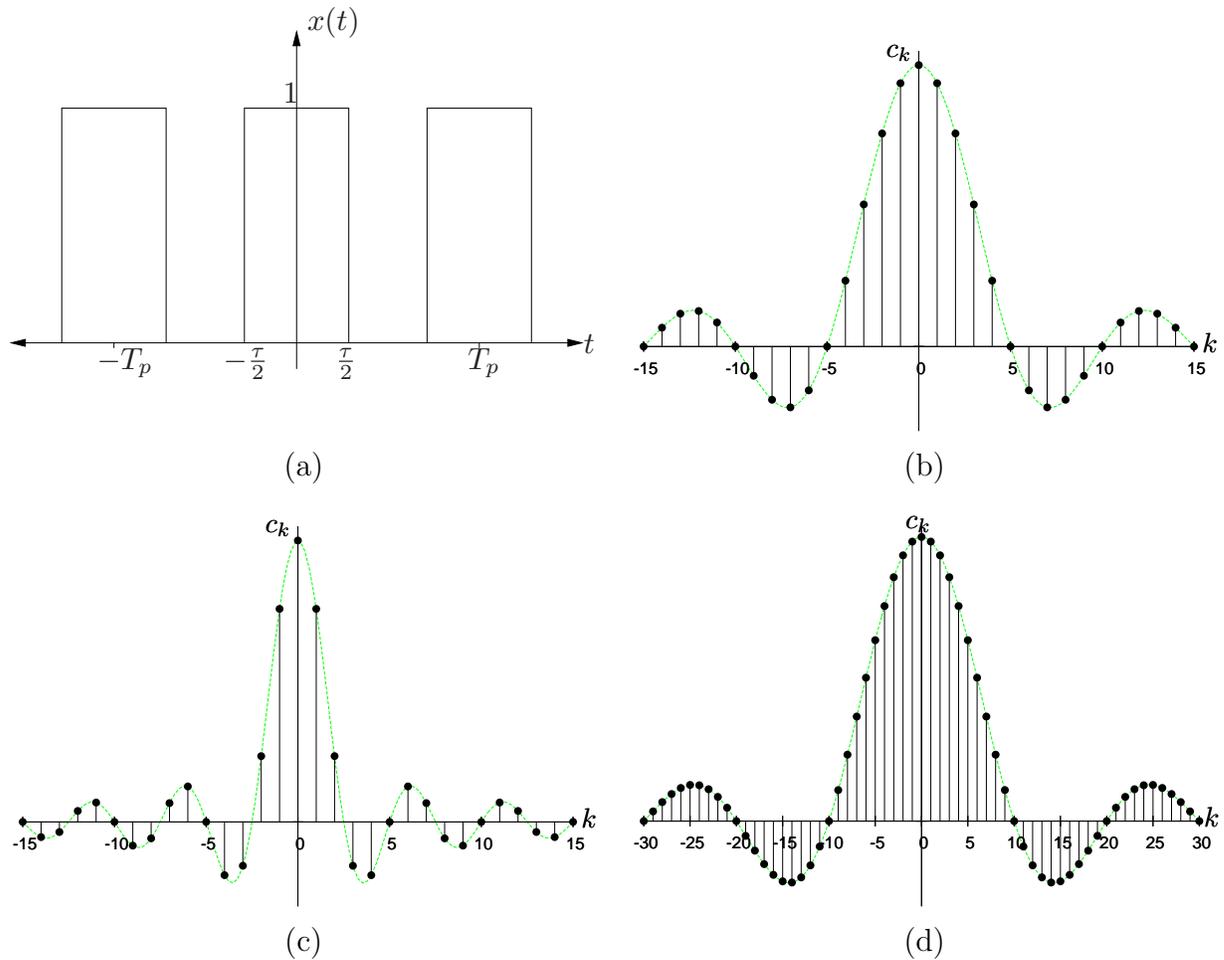


Figura 4.1: Secuencia periódica de pulsos rectangulares en tiempo continuo. (a) Representación temporal. (b) Líneas espectro de (a) con $\tau = 1$ y $T_p = 5$. (c) Líneas espectrales de (a) con $\tau = 2$ y $T_p = 5$. (d) Líneas espectrales de (a) con $\tau = 1$ y $T_p = 10$.

la figura 4.1b con 4.1d). Retomando el cálculo de los coeficientes de la serie de Fourier utilizando como $t_0 = T_p/2$ se tiene que

$$c_k = \frac{1}{T_p} \int_{-T_p/2}^{T_p/2} x(t) e^{-j2\pi F_0 k t} dt. \quad (4.1)$$

Dejando crecer el periodo T_p se puede alcanzar una frecuencia F_0 arbitrariamente pequeña, que puede entonces reemplazarse por ΔF_0 de tal forma que si $T_p \rightarrow \infty$ (señal aperiódica) entonces ΔF_0 se transforma en un diferencial dF y $k\Delta F_0$ se convierte en una variable de valor continuo F . La *transformada de Fourier* se obtiene con estos conceptos eliminando el factor $\frac{1}{T_p}$ y haciendo $T_p \rightarrow \infty$:

$$X(F) = \int_{-\infty}^{\infty} x(t) e^{-j2\pi F t} dt \quad (4.2)$$

y está relacionada directamente por lo tanto con los coeficientes de la serie a través de

$$X(F) = \lim_{T_p \rightarrow \infty} T_p c_k$$

Tabla 4.1: Propiedades de la Serie de Fourier

Propiedad	Señal en el tiempo	Coefficientes
	$x(t)$	c_k
	$x_1(t)$	c_{1k}
	$x_2(t)$	c_{2k}
Linealidad	$\alpha_1 x_1(t) + \alpha_2 x_2(t)$	$\alpha_1 c_{1k} + \alpha_2 c_{2k}$
Simetría par	$x(t) = x(-t)$	$c_k = \frac{2}{T_p} \int_0^{\frac{T_p}{2}} x(t) \cos(\omega_0 kt) dt$ $c_k \in \mathbb{R}$
Simetría impar	$x(t) = -x(-t)$	$c_k = -\frac{2j}{T_p} \int_0^{\frac{T_p}{2}} x(t) \text{sen}(\omega_0 kt) dt$ $c_k \in j\mathbb{R}$
Función real	$x(t) \in \mathbb{R}$	$c_k = c_{-k}^*$
Desplazamiento temporal	$x(t - \tau)$	$e^{-j\omega_0 k\tau} c_k$
Conjugación	$x^*(t)$	c_{-k}^*
Inversión en el tiempo	$x(-t)$	c_{-k}
Escalamiento en el tiempo	$x(\alpha t), \alpha > 0$	c_k
Convolución periódica	$\int_{T_p} x_1(\tau) x_2(t - \tau) d\tau$	$T_p c_{1k} c_{2k}$
Multiplicación	$x_1(t) x_2(t)$	$\sum_{l=-\infty}^{\infty} c_{1l} c_{2k-l}$
Diferenciación	$\frac{dx(t)}{dt}$	$jk\omega_0 c_k$
Integración	$\int_{-\infty}^t x(t) dt, c_0 = 0$	$\frac{c_k}{jk\omega_0}$
Relación de Parseval	$\frac{1}{T_p} \int_{t_0}^{t_0+T_p} x(t) ^2 dt = \sum_{k=-\infty}^{\infty} c_k ^2$	

En términos de la frecuencia angular $\Omega = 2\pi F$ esta transformación equivale a

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

Puede demostrarse además [1] que la transformada inversa de Fourier está dada por:

$$x(t) = \int_{-\infty}^{\infty} X(F) e^{j2\pi Ft} dF$$

o con $d\Omega = 2\pi dF$

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

Si $x(t)$ es una señal de extensión finita entonces existe una extensión periódica $x_p(t)$ de periodo T_p mayor a la extensión de $x(t)$, como lo muestra la figura 4.2. En este caso se

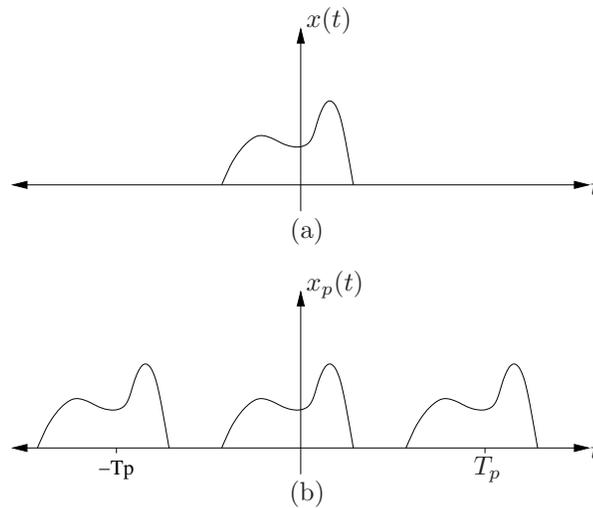


Figura 4.2: Una función finita $x(t)$ y su extensión periódica $x_p(t)$

pueden comparar las ecuaciones (4.1) y (4.2) y se obtiene que

$$c_k = \frac{1}{T_p} X(kF_0)$$

es decir, la extensión periódica de una señal aperiódica conduce al muestreo del espectro de la transformada de Fourier con una tasa F_0 .

Las condiciones de Dirichlet para la existencia de la transformada de Fourier son equivalentes a las ya mencionadas en las series de Fourier de señales periódicas continuas, excepto que la señal debe ser ahora absolutamente integrable en todo t y no solo en un periodo:

$$\int_{-\infty}^{\infty} |x(t)| dt < \infty$$

lo que se deduce del hecho que

$$|X(F)| = \left| \int_{-\infty}^{\infty} x(t) e^{-j2\pi Ft} dt \right| \leq \int_{-\infty}^{\infty} |x(t)| |e^{-j2\pi Ft}| dt = \int_{-\infty}^{\infty} |x(t)| dt < \infty$$

De forma similar a las señales periódicas puede demostrarse que

$$E_x = \int_{-\infty}^{\infty} |x(t)|^2 dt = \int_{-\infty}^{\infty} |X(F)|^2 dF$$

que es la relación de Parseval para señales aperiódicas de energía finita. La cantidad $S_{xx}(F) = |X(F)|^2$ representa la distribución de energía de la señal en función de la frecuencia, por lo que a S_{xx} se le denomina *densidad espectral de energía*.

Para señales reales se tiene que $|X(-F)| = |X(F)|$ y $\angle X(-F) = -\angle X(F)$ por lo que $S_{xx}(-F) = S_{xx}(F)$.

Ejemplo 4.2 Calcule la transformada de Fourier de (figura 4.3a)

$$x(t) = \begin{cases} A & |t| \leq \tau/2 \\ 0 & |t| > \tau/2 \end{cases}$$

Utilizando la ecuación (4.2) se obtiene (figura 4.3b):

$$X(F) = \int_{-\tau/2}^{\tau/2} A e^{-j2\pi Ft} dt = \frac{A\tau \operatorname{sen} \pi F\tau}{\pi F\tau}$$

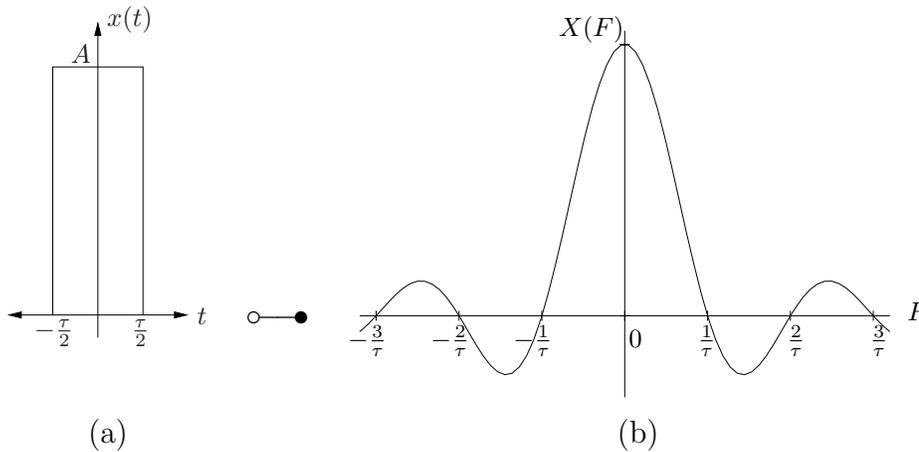


Figura 4.3: Pulso de ancho temporal τ (a) y su transformada de Fourier (b).

4.2

Nótese que mientras $x(t)$ sea más localizada en el tiempo, es decir, mientras más pequeño sea τ , mucho más amplio es su espectro puesto que $1/\tau$ crece. Este llamado *principio de incertidumbre* es general para todas las señales, e indica que algo muy localizado en el tiempo no puede ser localizado con precisión en el dominio de la frecuencia, y algo con frecuencias muy puntuales no tiene una ubicación específica en el tiempo.

La tabla 4.2 resume algunas propiedades de la Transformada de Fourier.

4.2 Espectro de señales en tiempo discreto

En la sección anterior se presentó que el espectro de señales continuas abarca frecuencias de $-\infty$ a ∞ , donde en el caso de señales periódicas solo se encontrarán frecuencias “discretas” múltiplos de $F_0 = 1/T_p$, donde T_p es el periodo fundamental de la señal. En capítulos anteriores se demostró que, sin embargo, para señales de naturaleza discreta solo existe unicidad para frecuencias normalizadas en el intervalo $] -1/2; 1/2]$, ó $[0; 1[$.

4.2.1 Espectro de señales discretas periódicas

Una señal discreta periódica con periodo fundamental N puede tener componentes frecuenciales separadas por $\omega = 2\pi/N$ radianes o $f = 1/N$ ciclos, que se representan por los

Tabla 4.2: Propiedades de la Transformada de Fourier

Propiedad	Señal en el tiempo	Transformada
	$x(t)$	$X(j\omega)$
	$x_1(t)$	$X_1(j\omega)$
	$x_2(t)$	$X_2(j\omega)$
Linealidad	$\alpha_1 x_1(t) + \alpha_2 x_2(t)$	$\alpha_1 X_1(j\omega) + \alpha_2 X_2(j\omega)$
Simetría par	$x(t) = x(-t)$	$2 \int_0^{\infty} x(t) \cos(\omega t) dt$ $X(j\omega) \in \mathbb{R}$
Simetría impar	$x(t) = -x(-t)$	$-2j \int_0^{\infty} x(t) \sen(\omega t) dt$ $X(j\omega) \in j\mathbb{R}$
Función real	$x(t) \in \mathbb{R}$	$X(j\omega) = X^*(-j\omega)$
Dualidad	$X(jt)$	$2\pi x(-\omega)$
Desplazamiento temporal	$x(t - \tau)$	$e^{-j\omega\tau} X(j\omega)$
Desplazamiento en frecuencia	$e^{j\omega_0 t} x(t)$	$X(j\omega - j\omega_0)$
Modulación	$\cos(\omega_0 t) x(t)$	$\frac{1}{2} X(j\omega - j\omega_0) + \frac{1}{2} X(j\omega + j\omega_0)$
Conjugación	$x^*(t)$	$X^*(-j\omega)$
Inversión en el tiempo	$x(-t)$	$X(-j\omega)$
Escalamiento en el tiempo	$x(at)$	$\frac{1}{ a } X\left(\frac{j\omega}{a}\right)$
Convolución	$x_1(t) * x_2(t)$	$X_1(j\omega) X_2(j\omega)$
Multiplicación	$x_1(t) x_2(t)$	$\frac{1}{2\pi} X_1(j\omega) * X_2(j\omega)$
Diferenciación	$\frac{dx(t)}{dt}$	$j\omega X(j\omega)$
	$tx(t)$	$j \frac{d}{d\omega} X(j\omega)$
Integración	$\int_{-\infty}^t x(t) dt$	$\frac{1}{j\omega} X(j\omega) + \pi X(0) \delta(\omega)$
Relación de Parseval	$\int_{-\infty}^{\infty} x(t) ^2 dt =$	$\frac{1}{2\pi} \int_{-\infty}^{\infty} X(j\omega) ^2 d\omega$

exponenciales complejos armónicamente relacionados:

$$s_k(n) = e^{j2\pi kn/N}, \quad k = 0 \dots N - 1$$

Utilizando la serie generalizada de Fourier con estas señales como base se obtiene para la síntesis de una señal:

$$x(n) = \sum_{k=0}^{N-1} c_k e^{j2\pi kn/N} \quad (4.3)$$

y considerando que

$$\sum_{n=0}^{N-1} a^n = \begin{cases} N & a = 1 \\ \frac{1 - a^N}{1 - a} & a \neq 1 \end{cases}$$

se puede afirmar para las exponenciales complejas

$$\sum_{n=0}^{N-1} e^{j2\pi kn/N} = \begin{cases} N & k = 0, \pm N, \pm 2N, \dots \\ 0 & \text{en el resto} \end{cases}$$

tomando $a = e^{j2\pi k/N}$ por lo que $a^N = 1$.

Multiplicando (4.3) por $e^{-j2\pi ln/N}$ y sumando de 0 a $N - 1$.

$$\begin{aligned} \sum_{n=0}^{N-1} x(n)e^{-j2\pi ln/N} &= \sum_{n=0}^{N-1} \sum_{k=0}^{N-1} c_k e^{j2\pi kn/N} e^{-j2\pi ln/N} \\ &= \sum_{k=0}^{N-1} c_k \sum_{n=0}^{N-1} e^{j2\pi(k-l)n/N} \\ &= Nc_l \end{aligned}$$

y finalmente

$$c_l = \frac{1}{N} \sum_{n=0}^{N-1} x(n)e^{-j2\pi ln/N}, \quad l = 0, 1, \dots, N - 1$$

Nótese que en esta serie de Fourier en tiempo discreto (DTFS, discrete time Fourier series) el coeficiente c_k representa la amplitud y fase de la componente frecuencial $s_k(n) = e^{j2\pi kn/N} = e^{j\omega_k n}$ con frecuencia angular normalizada $\omega_k = 2\pi k/N$ (o lo que es equivalente $f_k = k/N$). Puesto que $s_k(n)$ es periódica al ser f_k un número racional ($s_k(n) = s_k(n + N)$), entonces los coeficientes c_k también son periódicos con periodo fundamental N :

$$c_{k+N} = \frac{1}{N} \sum_{n=0}^{N-1} x(n)e^{-j2\pi(k+N)n/N} = \frac{1}{N} \sum_{n=0}^{N-1} x(n)e^{-j2\pi kn/N} e^{-j2\pi n} = c_k$$

Por conveniencia se utiliza para c_k normalmente el intervalo $k = 0 \dots N - 1$.

La relación de Parseval es en este caso

$$\frac{1}{N} \sum_{n=0}^{N-1} |x(n)|^2 = \sum_{k=0}^{N-1} |c_k|^2$$

es decir, la potencia media de la señal es igual a la suma de las potencias medias de cada componente en el dominio de la frecuencia.

4.2.2 Espectro de señales discretas aperiódicas

La transformada de Fourier de una señal de energía finita en tiempo discreto $x(n)$ se define como:

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

Nótese que $X(\omega) = X(\omega + 2\pi k)$, lo que implica que el rango de frecuencias únicas se limita a $[0, 2\pi[$, o de forma equivalente a $]-\pi, \pi]$, que contrasta con el rango $[-\infty, \infty]$ de las señales continuas.

Como la señal $x(n)$ es discreta, una sumatoria reemplaza la integral del caso continuo. Puede demostrarse además que

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

La transformada de Fourier converge si

$$|X(\omega)| = \left| \sum_{n=-\infty}^{\infty} x(n) e^{-j\omega n} \right| \leq \sum_{n=-\infty}^{\infty} |x(n)| < \infty$$

es decir, si la señal es absolutamente sumable.

La relación de Parseval en este caso es:

$$E_x = \sum_{n=-\infty}^{\infty} |x(n)|^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} |X(\omega)|^2 dt$$

Para señales reales, el espectro $X(\omega)$ tiene magnitud par y fase impar, es decir, se cumple que $X(\omega) = X^*(-\omega)$, por lo que usualmente basta determinar el espectro para $\omega \in [0, \pi]$, pues la contraparte puede generarse por simetría.

Debe notarse en los cuatro casos anteriores la correspondencia entre periodicidad en un dominio y la naturaleza discreta de la representación en el dominio complementario. Es decir, si la señal es periódica, su espectro será discreto y si el espectro es periódico, es porque la señal es discreta. Si la distancia entre las muestras o la líneas espectrales es α entonces la periodicidad en el otro dominio será $1/\alpha$.

periódico \longleftrightarrow discreto

aperiódico \longleftrightarrow continuo

Por ejemplo, una señal continua periódica tendrá un espectro aperiódico discreto; una señal discreta aperiódica tendrá un espectro periódico continuo.

4.2.3 Relación entre las transformadas de Fourier y z

La transformada z de la secuencia $x(n)$ se ha definido como

$$X(z) = \sum_{n=-\infty}^{\infty} x(n) z^{-n}, \quad \text{ROC} : r_2 < |z| < r_1 \quad (4.4)$$

Expresando z como $z = r e^{j\omega}$ ($r = |z|$, $\omega = \angle z$), entonces (4.4) se puede reescribir como

$$X(z)|_{z=r e^{j\omega}} = \sum_{n=-\infty}^{\infty} [x(n) r^{-n}] e^{-j\omega n}$$

que equivale a la transformada de Fourier de la secuencia $x(n)r^{-n}$. Para el caso especial de $r = 1$ se obtiene

$$X(z)|_{z=e^{j\omega}} = X(\omega) = \sum_{n=-\infty}^{\infty} x(n)e^{-j\omega n}$$

es decir, la transformada de Fourier de $x(n)$ equivale a su transformada z evaluada en la circunferencia unitaria $z = e^{j\omega}$. La figura 4.4 muestra como ejemplo la magnitud de la transformada z de una secuencia $x(n)$ y a la derecha la misma función “recortada” en $|z| = 1$. La silueta formada en $|z| = 1$ corresponde entonces a $X(\omega)$.

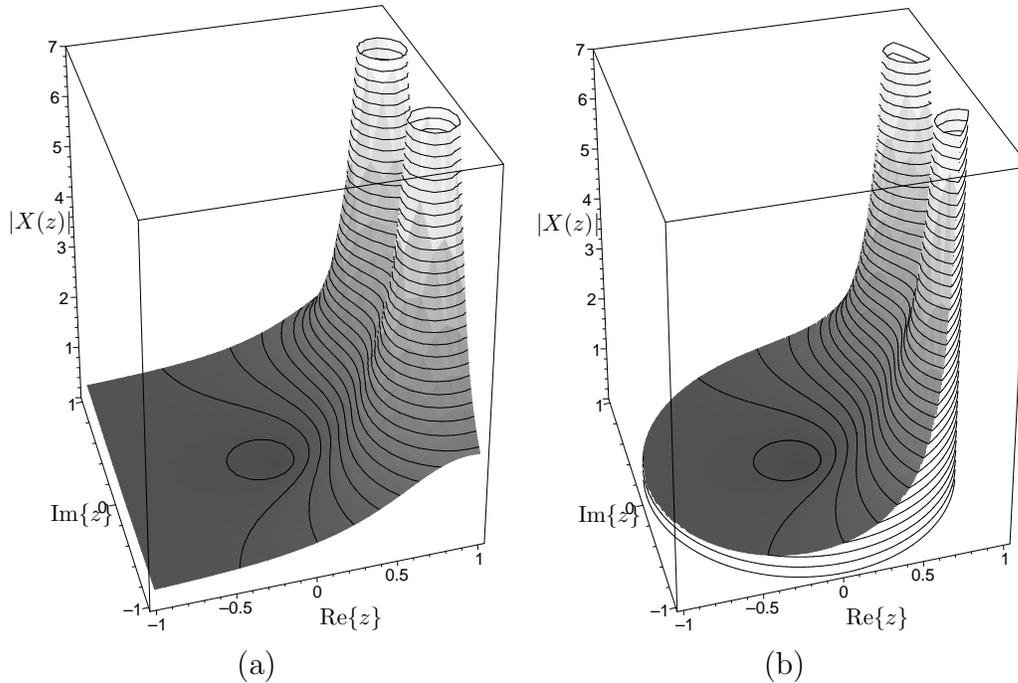


Figura 4.4: Correspondencia entre $X(\omega)$ y $X(z)$ para una función con un par de polos complejos conjugados en $z = 0,95e^{\pm j30^\circ}$ y un cero en $z = 0$. (a) $|X(z)|$. (b) $|X(z)|$ para $|z| < 1$.

Si $X(z)$ no converge en $|z| = 1$ entonces la transformada de Fourier no existe. Por otro lado, existen funciones con transformada de Fourier, que no tienen transformada z . Por ejemplo, $x(n) = \frac{\text{sen } \omega_c n}{\pi n}$ tiene transformada de Fourier

$$X(\omega) = \begin{cases} 1 & |\omega| \leq \omega_c \\ 0 & \omega_c < |\omega| < \pi \end{cases}$$

mas no posee transformada z , puesto que para la serie $\sum_{n=-\infty}^{\infty} \frac{\text{sen } \omega_c n}{\pi n} z^{-n}$ no existe ninguna región de convergencia.

Algunas secuencias con polos en $|z| = 1$ en su transformada z pueden tener una transformada de Fourier si se extiende la definición de transformada para utilizar impulsos de Dirac

$$\delta(\omega) = \begin{cases} \infty & \omega = 0 \\ 0 & \omega \neq 0 \end{cases} \quad \int_{-\infty}^{\infty} \delta(\omega) d\omega = 1$$

Ejemplo 4.3 Determine la transformada de Fourier de las señales

1. $x_1(n) = u(n)$
2. $x_2(n) = (-1)^n u(n)$
3. $x_3(n) = \cos(\omega_0 n) u(n)$

1. La transformada de $x_1(n)$ es $X_1(z) = \frac{1}{1 - z^{-1}} = \frac{z}{z - 1}$ con una ROC $|z| > 1$, pues tiene un polo en $z = 1$. Con $z = e^{j\omega}$ se obtiene

$$X_1(\omega) = \frac{e^{j\omega}}{e^{j\omega} - 1} = \frac{e^{j\omega/2}}{e^{j\omega/2} - e^{-j\omega/2}} = \frac{1}{2 \operatorname{sen}(\omega/2)} e^{j\frac{\omega - \pi}{2}}; \quad \omega \neq 2\pi k, k \in \mathbb{Z}$$

cuya magnitud y fase se muestran en la figura 4.5.

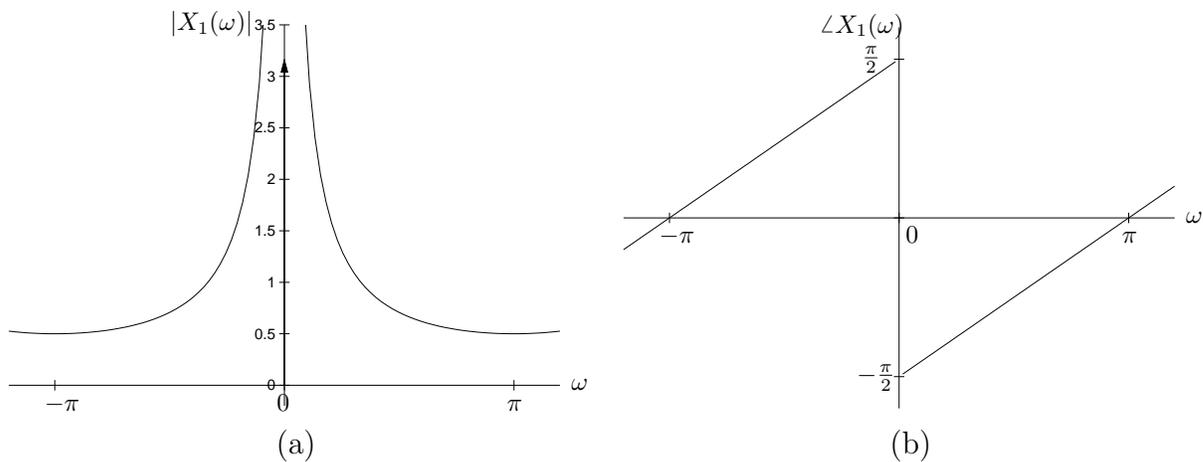


Figura 4.5: Espectros de magnitud y fase para la función escalón $u(n)$. (a) Espectro de magnitud. (b) Espectro de fase.

Puesto que solo en $\omega = 2k\pi$ $X_1(\omega)$ está indefinido, se sustituye allí la transformada por $\pi\delta(\omega)$, donde la constante π se obtiene de un desarrollo matemático riguroso que excede el contexto de estas notas.¹

2. $X_2(z) = \frac{1}{1 + z^{-1}} = \frac{z}{z + 1}$ con un polo en $z = -1 = e^{j\pi}$. La transformada de Fourier es entonces $X_2(\omega) = \frac{e^{j\omega/2}}{2 \cos(\omega/2)}$, con $\omega \neq 2\pi k + \pi, k \in \mathbb{Z}$. Donde se agregan los impulsos en las frecuencias $\omega = -\pi$ y $\omega = \pi$ (figura 4.6).

3. $X_3(z) = \frac{1 - z^{-1} \cos \omega_0}{1 - 2z^{-1} \cos \omega_0 + z^{-2}}$, ROC $|z| > 1$ puesto que tiene dos polos complejos conjugados en $e^{\pm j\omega_0}$. Por lo tanto, la transformada de Fourier es

$$X_3(\omega) = \frac{1 - e^{-j\omega} \cos \omega_0}{(1 - e^{-j(\omega - \omega_0)})(1 - e^{-j(\omega + \omega_0)})}$$

con $\omega \neq \pm\omega_0 + 2\pi k$ (figura 4.7).

¹El desarrollo parte del hecho que $\int_{-\infty}^{\infty} e^{j\omega t} d\omega = 2\pi\delta(t)$, o haciendo un cambio de variable $\int_{-\infty}^{\infty} e^{j\omega t} dt = 2\pi\delta(\omega)$.

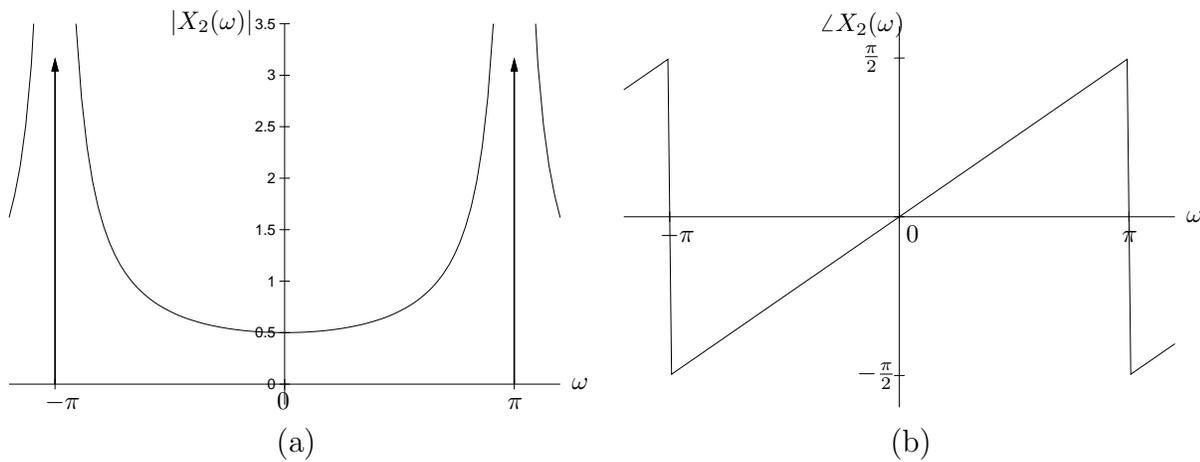


Figura 4.6: Espectros de magnitud y fase para la función $(-1)^n u(n)$. (a) Espectro de magnitud. (b) Espectro de fase.

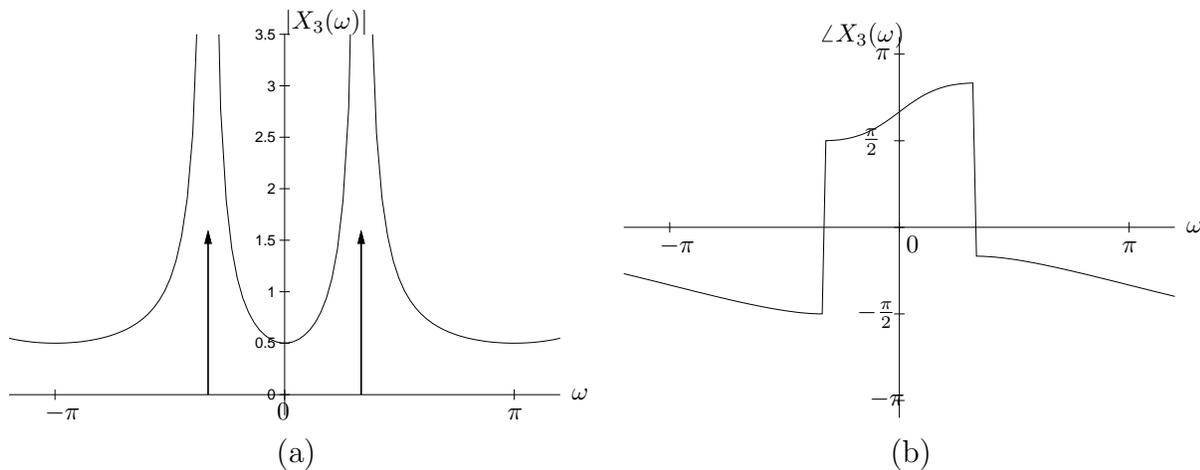


Figura 4.7: Espectros de magnitud y fase para la función $\cos(\pi n/3)$. (a) Espectro de magnitud. (b) Espectro de fase.

4.3

4.2.4 El teorema del muestreo

Derivación conceptual por análisis de Fourier

La función puente $p_T(t)$ mostrada en la figura 4.8 se utiliza para modelar la toma de muestras con el intervalo T de la señal analógica $x_a(t)$. La señal muestreada $x_s(t)$ se modela a través del producto de la señal $x_a(t)$ y $p_T(t)$ (figura 4.9):

$$x_s(t) = x_a(t)p_T(t)$$

Puesto que la función puente $p_T(t)$ es periódica, se puede expresar por medio de una serie

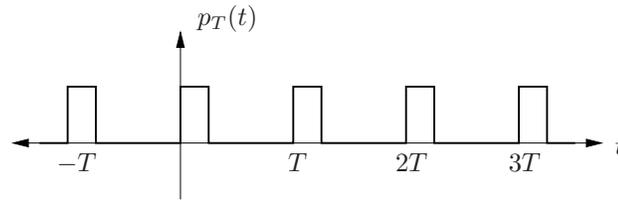


Figura 4.8: Función puente utilizada para tomar muestras periódicas en el tiempo de señales continuas.

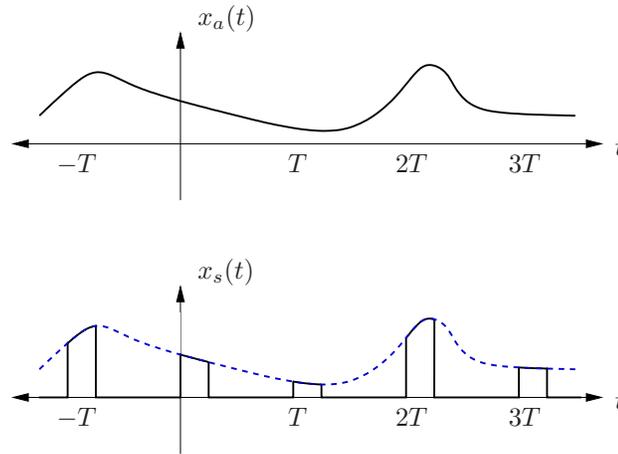


Figura 4.9: Muestreo de señal utilizando la función puente.

de Fourier con coeficientes P_n :

$$p_T(t) = \sum_{n=-\infty}^{\infty} P_n e^{jn\Omega_0 t}; \quad \Omega_0 = \frac{2\pi}{T} = 2\pi F$$

de modo que se cumple

$$x_s(t) = x_a(t) \sum_{n=-\infty}^{\infty} P_n e^{jn\Omega_0 t}$$

Aplicando la transformada de Fourier a ambos lados de la ecuación y utilizando la propiedad de linealidad se obtiene

$$X_s(j\Omega) = \mathcal{F} \{x_s(t)\} = \mathcal{F} \left\{ x_a(t) \sum_{n=-\infty}^{\infty} P_n e^{jn\Omega_0 t} \right\} = \sum_{n=-\infty}^{\infty} P_n \mathcal{F} \{x_a(t) e^{jn\Omega_0 t}\}$$

y considerando la propiedad de desplazamiento en frecuencia, lo anterior equivale a

$$\begin{aligned} X_s(j\Omega) &= \sum_{n=-\infty}^{\infty} P_n X_a(j\Omega - jn\Omega_0) \\ &= P_0 X_a(j\omega) + \sum_{\substack{n=-\infty \\ n \neq 0}}^{\infty} P_n X_a(j\Omega - jn\Omega_0) \end{aligned}$$

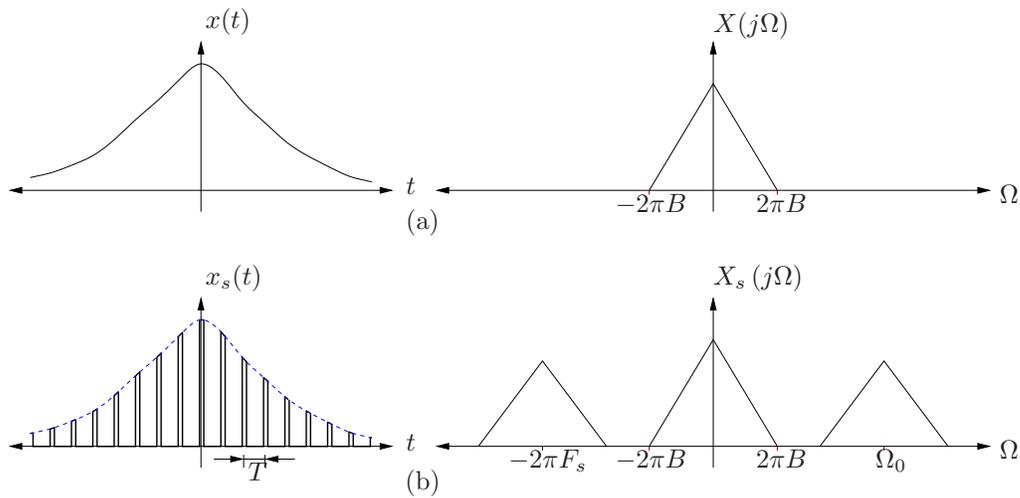


Figura 4.10: Replicación del espectro por medio del muestreo.

de donde se deduce que con el muestreo se producen réplicas del espectro original de la señal $X_a(j\Omega)$, separadas en el dominio de la frecuencia por $\Omega_0 = 2\pi/T$ y ponderadas por los coeficientes de la representación de la función puente como serie de Fourier (figura 4.10).

Si el espectro original $X_a(j\Omega)$ es de banda limitada, es decir, si a partir de cierta frecuencia $2\pi B$ (donde B es el llamado *ancho de banda*) se cumple $X_a(j\omega) = 0$, entonces eligiendo el periodo de muestreo suficientemente alto es posible evitar que las réplicas se traslapen con el espectro original. Si $x_a(t)$ es real, entonces la magnitud de su espectro es par, y por tanto para evitar el traslape la siguiente réplica se deberá posicionar al menos en $\Omega_0 > 4\pi B$ lo que es equivalente a afirmar que la frecuencia de muestreo f_s debe ser al menos el doble del ancho de banda B del espectro $X_a(j\Omega)$ para evitar que las replicas se traslapen.

Se puede demostrar que si no hay traslape entonces es posible rescatar a partir de la señal muestreada $x_s(t)$ la señal original $x_a(t)$ utilizando un filtro paso-bajos que permita pasar únicamente el rango de frecuencias angulares desde $-2\pi B$ hasta $2\pi B$.

Estos principios se resumen en el llamado *teorema del muestreo* que establece que para muestrear una señal analógica sin pérdidas de información se requiere una tasa de muestreo de al menos el doble del ancho de banda del espectro de la señal.

En los últimos años han tomado fuerza nuevas propuestas² para requerir aún menos muestras que las establecidas por este teorema, aunque su funcionamiento parte de requisitos adicionales impuestos a la función $x_a(t)$, donde la posición de las muestras deja de ser regular, y el comportamiento matemático pasa a ser más basado en procesos estocásticos que deterministas.

²El lector interesado puede buscar entre otros los temas de *compressive sampling* y el *finite rate of innovation*.

Derivación algebraica

En la sección 2.2.2 se revisó el hecho de que la máxima frecuencia normalizada representable en una señal muestreada es $f_{max} = \frac{F_{max}}{F_s} = \frac{1}{2}$. Puesto que la representación de Fourier de la señal discreta es

$$x(n) \circ \bullet X(\omega) = \sum_{n=-\infty}^{\infty} x(n)e^{-j\omega n}, \quad \omega \in]-\pi, \pi]$$

y $X(\omega)$ es periódico con periodo fundamental 2π , se obtiene que la frecuencia mínima de muestreo F_s debe ser elegida de tal modo que sea al menos el doble de la frecuencia máxima de la señal (figura 4.11).

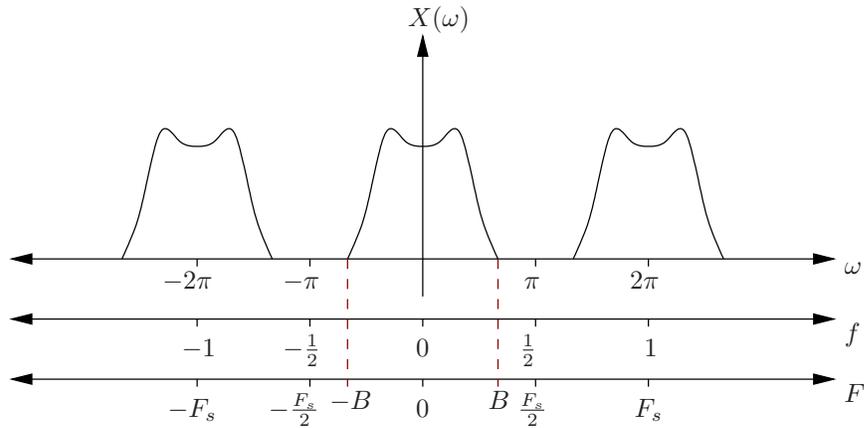


Figura 4.11: Espectro de señal discreta con banda limitada sin aliasing.

De otro modo existirá entre las repeticiones periódicas del espectro de la señal analógica un traslape, denominado *aliasing*. Sea $x_a(t)$ la señal analógica muestreada periódicamente cada T segundos para producir la señal en tiempo discreto $x(n)$:

$$x(n) = x_a(nT), \quad -\infty < n < \infty.$$

Si $x_a(t)$ es aperiódica de energía finita, entonces se cumple para su espectro:

$$X_a(F) = \int_{-\infty}^{\infty} x_a(t)e^{-j2\pi Ft} dt \bullet \circ x_a(t) = \int_{-\infty}^{\infty} X_a(F)e^{j2\pi Ft} dF$$

En el dominio discreto se tiene:

$$X(f) = \sum_{n=-\infty}^{\infty} x(n)e^{-j2\pi fn} \bullet \circ x(n) = \int_{-\frac{1}{2}}^{\frac{1}{2}} X(f)e^{j2\pi fn} df$$

Puesto que para la señal muestreada se tiene

$$t = nT = \frac{n}{F_s}$$

entonces se cumple que

$$x(n) = x_a(nT) = \int_{-\infty}^{\infty} X_a(F) e^{j2\pi nF/F_s} dF \stackrel{!}{=} \int_{-1/2}^{1/2} X(f) e^{j2\pi f n} df$$

Considerando que $f = F/F_s$ y $df = dF/F_s$ se obtiene entonces

$$\frac{1}{F_s} \int_{-F_s/2}^{F_s/2} X(F/F_s) e^{j2\pi nF/F_s} dF = \int_{-\infty}^{\infty} X_a(F) e^{j2\pi nF/F_s} dF$$

El lado derecho se puede segmentar en bloques de ancho F_s de la siguiente manera:

$$\int_{-\infty}^{\infty} X_a(F) e^{j2\pi nF/F_s} dF = \sum_{k=-\infty}^{\infty} \int_{(k-1/2)F_s}^{(k+1/2)F_s} X_a(F) e^{j2\pi nF/F_s} dF$$

Con un cambio de variable en la integral $F' = F - kF_s$, $dF' = dF$, se obtiene un desplazamiento del k -ésimo bloque al intervalo $[-F_s/2, F_s/2]$:

$$\begin{aligned} \int_{(k-1/2)F_s}^{(k+1/2)F_s} X_a(F) e^{j2\pi nF/F_s} dF &= \int_{-F_s/2}^{F_s/2} X_a(F' + kF_s) e^{j2\pi n \frac{F'+kF_s}{F_s}} dF' \\ &= \int_{-F_s/2}^{F_s/2} X_a(F + kF_s) e^{j2\pi nF/F_s} dF \end{aligned}$$

por lo que

$$\begin{aligned} \frac{1}{F_s} \int_{-F_s/2}^{F_s/2} X\left(\frac{F}{F_s}\right) e^{j2\pi nF/F_s} dF &= \sum_{k=-\infty}^{\infty} \int_{-F_s/2}^{F_s/2} X_a(F + kF_s) e^{j2\pi nF/F_s} dF \\ &= \int_{-F_s/2}^{F_s/2} \left[\sum_{k=-\infty}^{\infty} X_a(F + kF_s) \right] e^{j2\pi nF/F_s} dF \end{aligned}$$

es decir

$$X\left(\frac{F}{F_s}\right) = X(f) = F_s \sum_{k=-\infty}^{\infty} X_a((f - k)F_s)$$

Nótese que el espectro de la señal discreta $X(f)$ es igual a la repetición periódica con periodo F_s del espectro escalado $F_s X_a(F)$.

Si el espectro de la señal analógica es de banda limitada, es decir, $X_a(F) = 0$ si $|F| \geq B$, entonces si se elige F_s mayor a $2B$ se cumple para el rango de frecuencias únicas $|F| \leq F_s/2$:

$$X\left(\frac{F}{F_s}\right) = F_s X_a(F), \quad |F| \leq F_s/2.$$

En este caso no existe aliasing y los espectros de las señales continua y discreta son idénticos en el intervalo de frecuencias $|F| \leq F_s/2$, exceptuando el factor de escala F_s (figura 4.12).

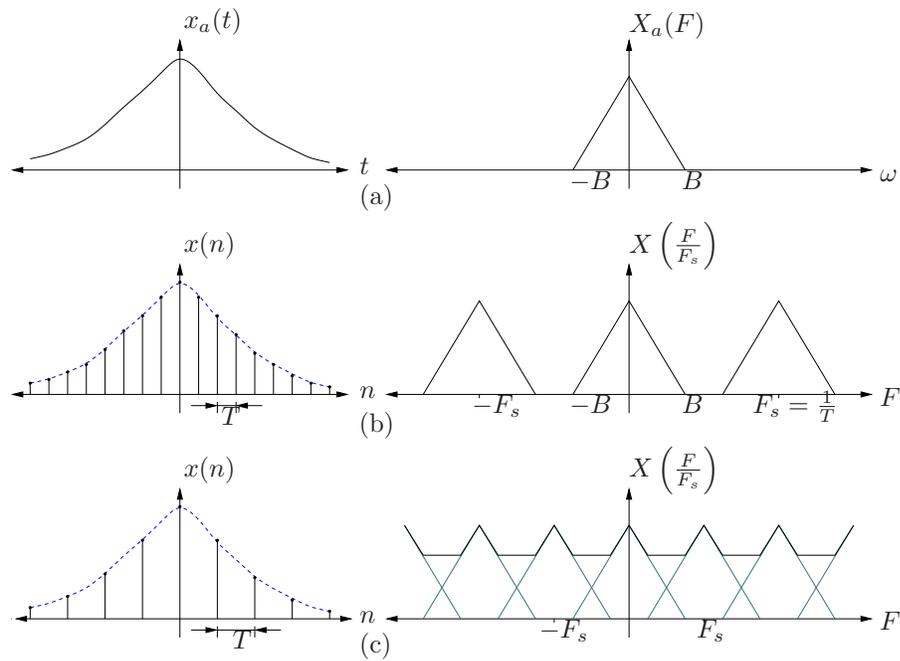


Figura 4.12: Señal analógica de espectro finito y los efectos de diferentes frecuencias de muestreo. (a) Señal analógica $x_a(t)$ y su espectro. (b) Muestreo con $F_s > 2B$. (c) Muestreo con $F_s < 2B$.

Si $F_s < 2B$ entonces el solapamiento espectral impide que la señal original pueda ser recuperada a partir de las muestras. Si no hay solapamiento, entonces

$$X_a(F) = \begin{cases} \frac{1}{F_s} X\left(\frac{F}{F_s}\right) & |F| \leq F_s/2 \\ 0 & |F| > F_s/2 \end{cases}$$

y puesto que

$$X\left(\frac{F}{F_s}\right) = \sum_{n=-\infty}^{\infty} x(n) e^{-j2\pi F n / F_s}$$

y además

$$x_a(t) = \int_{-F_s/2}^{F_s/2} X_a(F) e^{j2\pi F t} dF$$

se tiene que

$$\begin{aligned} x_a(t) &= \frac{1}{F_s} \int_{-F_s/2}^{F_s/2} \left[\sum_{n=-\infty}^{\infty} x(n) e^{-j2\pi F n / F_s} \right] e^{j2\pi F t} dF \\ &= \frac{1}{F_s} \sum_{n=-\infty}^{\infty} x(n) \int_{-F_s/2}^{F_s/2} e^{j2\pi F (t - \frac{n}{F_s})} dF \end{aligned}$$

Para $t = n/F_s$ se cumple

$$\int_{-F_s/2}^{F_s/2} e^{j2\pi F (t - \frac{n}{F_s})} dF = \int_{-F_s/2}^{F_s/2} dF = F_s$$

y puesto que para $t \neq n/F_s$ se tiene

$$\begin{aligned} \int_{-F_s/2}^{F_s/2} e^{j2\pi F(t-\frac{n}{F_s})} dF &= \frac{e^{j2\pi F(t-\frac{n}{F_s})} \Big|_{-F_s/2}^{F_s/2}}{j2\pi \left(t - \frac{n}{F_s}\right)} = \frac{e^{j\pi F_s(t-\frac{n}{F_s})} - e^{-j\pi F_s(t-\frac{n}{F_s})}}{j2\pi \left(t - \frac{n}{F_s}\right)} \\ &= \frac{\text{sen} \left(\pi F_s \left(t - \frac{n}{F_s} \right) \right)}{\pi \left(t - \frac{n}{F_s} \right)} = \frac{F_s \text{sen} \left(\pi F_s \left(t - \frac{n}{F_s} \right) \right)}{\pi F_s \left(t - \frac{n}{F_s} \right)} \end{aligned}$$

con lo que se cumple para todo t

$$\int_{-F_s/2}^{F_s/2} e^{j2\pi F(t-\frac{n}{F_s})} dF = F_s \text{sa} \left(\pi F_s \left[t - \frac{n}{F_s} \right] \right)$$

entonces

$$\begin{aligned} x_a(t) &= \sum_{n=-\infty}^{\infty} x(n) \text{sa} \left(\pi F_s \left[t - \frac{n}{F_s} \right] \right) \\ &= \sum_{n=-\infty}^{\infty} x_a(nT) \text{sa} (\pi F_s [t - nT]) \end{aligned}$$

que es la interpolación de las muestras utilizando el interpolador ideal

$$g(t) = \frac{\text{sen} \left(\pi \frac{t}{T} \right)}{\pi \frac{t}{T}} = \text{sa} \left(\pi \frac{t}{T} \right). \quad (4.5)$$

Nótese que $g(t)$ es cero para $t = kT$, $k \in \mathbb{Z} \setminus 0$. En otras palabras, si $x_a(t)$ es de banda limitada B y $x(n) = x_a(nT)$ es muestreada con $F_s \geq 2B$ entonces $x_a(t)$ puede ser reconstruida completamente y de forma única utilizando el interpolador $g(t)$ dado en (4.5). Debe advertirse, sin embargo, que $g(t)$ tiene extensión infinita y no es causal, por lo que en aplicaciones prácticas suele aproximarse con interpoladores finitos. La figura 4.13 muestra un ejemplo de interpolación de una secuencia finita de muestras $x(n) = \{0, 3, 2, -1, 1, 0\}$. Se aprecia que la función interpolada atraviesa todas las muestras, mientras que para valores de t no enteros todas las muestras contribuyen al valor interpolado.

4.3 Propiedades de la transformada de Fourier de señales discretas

Simetría: Anteriormente se discutió el hecho de que cualquier señal puede representarse como la suma de una componente par y otra impar. Esto se puede aplicar tanto a la parte real como a la imaginaria del espectro y de su señal. Se puede demostrar que la relación entre estas componentes es la siguiente:

$$\begin{array}{ccccccccc} x(n) & = & [x_R^e(n) & + & jx_I^e(n)] & + & [x_R^o(n) & + & jx_I^o(n)] \\ \circ & & \circ & & \circ & & \circ & & \circ \\ \bullet & & \bullet & & \bullet & & \bullet & & \bullet \\ X(\omega) & = & [X_R^e(\omega) & + & jX_I^e(\omega)] & + & [jX_I^o(\omega) & + & X_R^o(\omega)] \end{array}$$

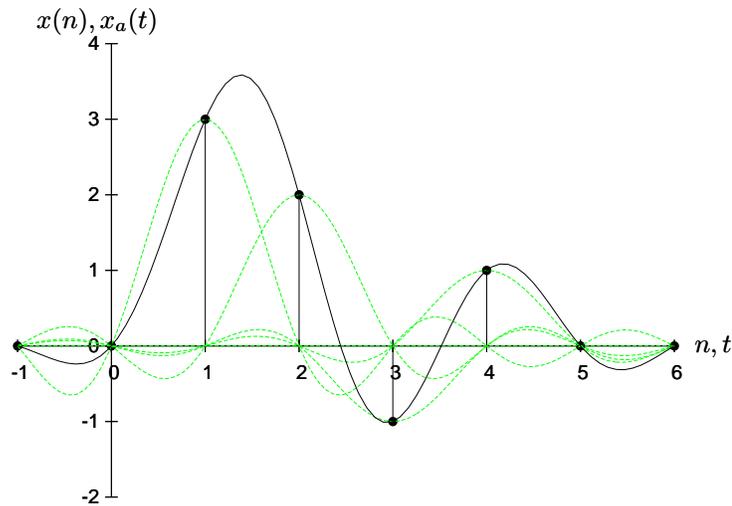


Figura 4.13: Ejemplo de interpolación ideal para una secuencia de muestras finita.

Si $x(n)$ es real, entonces $X(-\omega) = X^*(\omega)$

Linealidad: Si $x_1(n) \circ \bullet X_1(\omega)$ y $x_2(n) \circ \bullet X_2(\omega)$, entonces:

$$a_1x_1(n) + a_2x_2(n) \circ \bullet a_1X_1(\omega) + a_2X_2(\omega)$$

Desplazamiento temporal:

$$x(n) \circ \bullet X(\omega) \Rightarrow x(n - k) \circ \bullet e^{-j\omega k} X(\omega)$$

Reflexión temporal:

$$x(n) \circ \bullet X(\omega) \Rightarrow x(-n) \circ \bullet X(-\omega)$$

Teorema de la Convolución:

$$\begin{aligned} x_1(n) \circ \bullet X_1(\omega), x_2(n) \circ \bullet X_2(\omega) \\ x_1(n) * x_2(n) \circ \bullet X_1(\omega)X_2(\omega) \end{aligned}$$

Teorema de la Correlación:

$$\begin{aligned} x_1(n) \circ \bullet X_1(\omega), x_2(n) \circ \bullet X_2(\omega) \\ r_{x_1x_2}(n) \circ \bullet S_{x_1x_2}(\omega) = X_1(\omega)X_2(-\omega) \end{aligned}$$

Si la señal es real, puesto que $X(-\omega) = X^*(\omega)$ entonces:

$$r_{xx}(n) \circ \bullet X(\omega)X(-\omega) = X(\omega)X^*(\omega) = |X(\omega)|^2 = S_{xx}(\omega)$$

que se conoce como el teorema de Wiener-Khinchin.

Desplazamiento frecuencial:

$$x(n) \circ \bullet X(\omega) \Rightarrow e^{j\omega_0 n} x(n) \circ \bullet X(\omega - \omega_0)$$

Teorema de modulación:

$$x(n) \circ \bullet X(\omega) \Rightarrow x(n) \cos(\omega_0 n) \circ \bullet \frac{1}{2} \{X(\omega + \omega_0) + X(\omega - \omega_0)\}$$

Teorema de Parseval:

$$\begin{aligned} x_1(n) \circ \bullet X_1(\omega), x_2(n) \circ \bullet X_2(\omega) \\ \Rightarrow \sum_{n=-\infty}^{\infty} x_1(n)x_2^*(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} X_1(\omega)X_2^*(\omega)d\omega \end{aligned}$$

Demostración:

$$\begin{aligned} \frac{1}{2\pi} \int_{-\pi}^{\pi} X_1(\omega)X_2^*(\omega)d\omega &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \left(\sum_{n=-\infty}^{\infty} x_1(n)e^{-j\omega n} \right) X_2^*(\omega)d\omega \\ &= \sum_{n=-\infty}^{\infty} x_1(n) \frac{1}{2\pi} \int_{-\pi}^{\pi} X_2^*(\omega)e^{-j\omega n}d\omega \\ &= \sum_{n=-\infty}^{\infty} x_1(n)x_2^*(n) \end{aligned}$$

Para el caso $x_1(n) = x_2(n) = x(n) \circ \bullet X(\omega)$:

$$E_x = \sum_{n=-\infty}^{\infty} |x(n)|^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} |X(\omega)|^2 d\omega = \frac{1}{2\pi} \int_{-\pi}^{\pi} S_{xx}(\omega) d\omega$$

Teorema del ententado (multiplicación de secuencias):

$$\begin{aligned} x_1(n) \circ \bullet X_1(\omega), x_2(n) \circ \bullet X_2(\omega) \\ \Rightarrow x_3(n) = x_1(n)x_2(n) \circ \bullet X_3(\omega) = \frac{1}{2\pi} \int_{-\pi}^{\pi} X_1(\lambda)X_2(\omega - \lambda)d\lambda = X_1(\omega) * X_2(\omega) \end{aligned}$$

donde el símbolo * a la derecha, denota la convolución periódica en el dominio de la frecuencia.

Diferenciación en el dominio de la frecuencia:

$$x(n) \circ \bullet X(\omega) \Rightarrow nx(n) \circ \bullet j \frac{dX(\omega)}{d\omega}$$

4.4 Sistemas LTI en el dominio de la frecuencia

4.4.1 La función de respuesta en frecuencia

En el capítulo 2 se demostró que la respuesta de cualquier sistema LTI en reposo a una señal de entrada $x(n)$, está determinada por la convolución con la respuesta impulsional $h(n)$:

$$y(n) = \sum_{k=-\infty}^{\infty} h(k)x(n-k)$$

Para analizar este sistema en el dominio de la frecuencia se analiza su respuesta a una señal exponencial compleja $x(n) = Ae^{j\omega n}$. Se obtiene como salida:

$$y(n) = \sum_{k=-\infty}^{\infty} h(k)Ae^{j\omega(n-k)} = Ae^{j\omega n} \left[\sum_{k=-\infty}^{\infty} h(k)e^{-j\omega k} \right] = AH(\omega)e^{j\omega n}$$

Nótese que $H(\omega)$ existe si el sistema es estable, puesto que $\sum_{n=-\infty}^{\infty} |h(n)| < \infty$.

La respuesta del sistema LTI a una entrada exponencial compleja es entonces otra señal exponencial compleja de la misma frecuencia, pero con una nueva amplitud y fase determinadas por $H(\omega)$. Puesto que $H(\omega) = |H(\omega)|e^{j\angle H(\omega)}$ entonces es claro que la magnitud cambia de acuerdo a $|H(\omega)|$ y la fase de acuerdo a $\angle H(\omega)$.

Puesto que $h(n)$ es discreta, entonces $H(\omega) \bullet \circ h(n)$ es periódica y se cumple entonces:

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

Además, por las propiedades de simetría, si $h(n)$ es real, entonces $H(\omega)$ tiene simetría hermítica, es decir $H(-\omega) = H^*(\omega)$, lo que implica que $|H(\omega)| = |H(-\omega)|$ (simetría par), $\angle H(\omega) = -\angle H(-\omega)$ (impar), $\text{Re}\{H(\omega)\} = \text{Re}\{H(-\omega)\}$ (par), $\text{Im}\{H(\omega)\} = -\text{Im}\{H(-\omega)\}$ (simetría impar).

A $H(\omega)$ se le denomina *respuesta en frecuencia* del sistema. Esta función determina la amplitud y fase para cada señal exponencial compleja de entrada. Además, puesto que $\cos(\omega n) = \frac{e^{j\omega n} + e^{-j\omega n}}{2}$, entonces la respuesta del sistema LTI con respuesta impulsional $h(n)$ ante la entrada $x(n) = A \cos(\omega n)$, será:

$$\begin{aligned} y(n) &= \frac{A}{2} (H(\omega)e^{j\omega n} + H(-\omega)e^{-j\omega n}) \\ &= \frac{A}{2} (|H(\omega)|e^{j(\omega n + \angle H(\omega))} + |H(-\omega)|e^{-j(\omega n - \angle H(-\omega))}) \\ &= \frac{A}{2} |H(\omega)| (e^{j(\omega n + \angle H(\omega))} + e^{-j(\omega n + \angle H(\omega))}) \\ &= A|H(\omega)| \cos(\omega n + \angle H(\omega)) \end{aligned}$$

Es decir, $|H(\omega)|$ determina la atenuación o amplificación de la componente frecuencial de la entrada con frecuencia angular ω , y por ende se le denomina *respuesta en magnitud*, mientras que $\angle H(\omega)$ determina la fase a la salida de la misma componente (*respuesta en fase*).

Ejemplo 4.4 Para el sistema descrito por la ecuación:

$$y(n) = ay(n-1) + bx(n), \quad 0 < a < 1 \quad \wedge \quad a, b \in \mathbb{R}$$

1. Determine $H(\omega)$
2. Elija b de tal modo que $|H(\omega)|$ sea a lo sumo 1 y grafique $|H(\omega)|$ y $\angle H(\omega)$ para $a = 0,9$
3. Determine la salida para $x(n) = 5 + 12 \sin(\frac{\pi}{2}n) - 20 \cos(\pi n + \frac{\pi}{4})$

1. Puesto que $Y(z) = aY(z)z^{-1} + bX(z) \Rightarrow \frac{Y(z)}{X(z)} = \frac{b}{1-az^{-1}} = H(z) \bullet \rightarrow h(n) = ba^n u(n)$

$$\begin{aligned} H(\omega) &= \frac{b}{1 - ae^{-j\omega}} \\ |H(\omega)| &= \frac{|b|}{|1 - ae^{-j\omega}|} = \frac{|b|}{|1 - a(\cos \omega - j \sin \omega)|} \\ &= \frac{|b|}{\sqrt{(1 - a \cos \omega)^2 + (a \sin \omega)^2}} \\ &= \frac{|b|}{\sqrt{1 - 2a \cos \omega + a^2 \cos^2 \omega + a^2 \sin^2 \omega}} = \frac{|b|}{\sqrt{1 - 2a \cos \omega + a^2}} \\ \angle H(\omega) &= \angle b - \arctan\left(\frac{a \sin \omega}{1 - a \cos \omega}\right) \end{aligned}$$

2. $|H(\omega)|$ es máximo si el denominador es mínimo, y esto ocurre, puesto que $a \in [0, 1]$ cuando $\cos \omega$ es máximo, es decir, cuando $\omega = 0$:

$$|H(0)| = \frac{|b|}{\sqrt{1 - 2a + a^2}} = \frac{|b|}{1 - a} \Rightarrow |b| = 1 - a$$

Elijiendo $b = 1 - a$ se obtienen las gráficas en la figura 4.14.

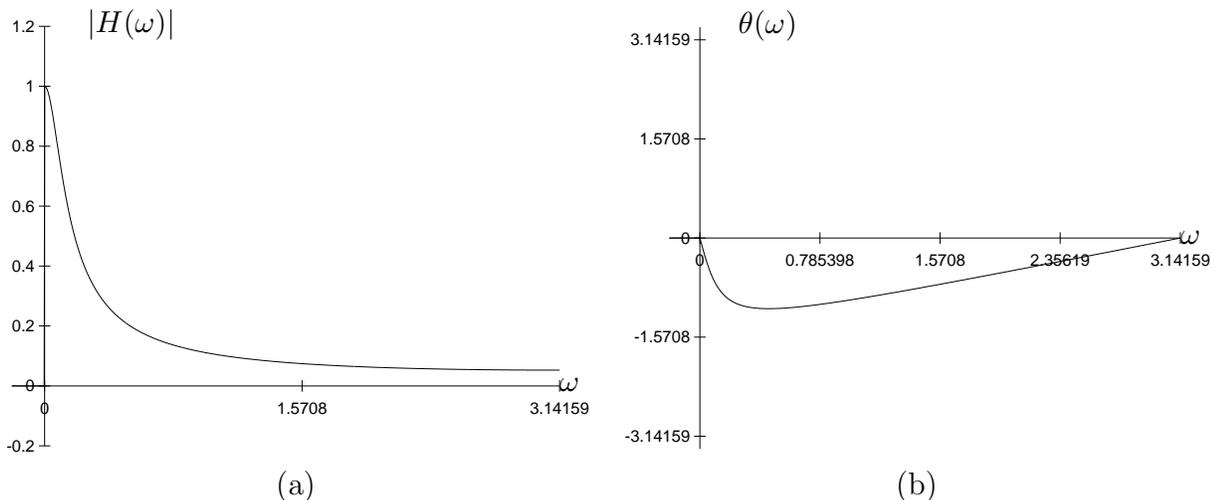


Figura 4.14: Respuestas de magnitud y fase para el sistema en el ejemplo (4.4).

3. Asumiendo $a = 0,9$, se requiere $H(0)$, $H(\frac{\pi}{2})$, $H(\pi)$

$$\begin{aligned} H(0) &= 1\angle 0^\circ \\ \left| H\left(\frac{\pi}{2}\right) \right| &= \frac{1-a}{\sqrt{1+a^2}} = \frac{0,1}{\sqrt{1,81}} = 0,074 \\ \angle H\left(\frac{\pi}{2}\right) &= -\arctan a = -42^\circ \\ |H(\pi)| &= \frac{1-a}{1+a} = \frac{0,1}{0,9} = 0,053 \quad \angle H(\pi) = 0 \end{aligned}$$

entonces

$$\begin{aligned} y(n) &= 5 + 12 \left| H\left(\frac{\pi}{2}\right) \right| \sin \left[\frac{\pi}{2}n + \angle H\left(\frac{\pi}{2}\right) \right] \\ &\quad - 20 |H(\pi)| \cos \left[\pi n + \frac{\pi}{4} + \angle H(\pi) \right] \\ &= 5 + 0,888 \sin \left(\frac{\pi}{2}n - 42^\circ \right) - 1,06 \cos \left(\pi n + \frac{\pi}{4} \right) \end{aligned}$$

4.4

4.4.2 Respuesta transitoria y en régimen permanente a entradas sinusoidales

La sección anterior analizó la respuesta de sistemas LTI a una entrada cosenoidal o exponencial compleja “eterna”, es decir, aplicadas al sistema en $n = -\infty$. La respuesta observada es entonces en régimen permanente, y no hay respuesta transitoria.

Todo sistema LTI estable BIBO excitado por una señal exponencial compleja en $n = 0$ u otro tiempo finito, tendrá una respuesta transitoria que tiende a cero a medida que n tiende a infinito.

Ejemplo 4.5 Determine las respuestas transitoria y de estado permanente para:

$$x(n) = Ae^{j\omega n}u(n)$$

del sistema

$$y(n) = ay(n-1) + x(n)$$

En la sección 2.5.2 se demostró que la respuesta de este sistema con condiciones iniciales no nulas es:

$$y(n) = a^{n+1}y(-1) + \sum_{k=0}^n a^k x(n-k), n \geq 0$$

Sustituyendo $x(n) = Ae^{j\omega n}u(n)$ se obtiene:

$$\begin{aligned}
 y(n) &= a^{n+1}y(-1) + A \sum_{k=0}^n a^k e^{j\omega(n-k)} \\
 &= a^{n+1}y(-1) + A \left[\sum_{k=0}^n a^k e^{-j\omega k} \right] e^{j\omega n} \\
 &= a^{n+1}y(-1) + A \frac{1 - a^{n+1}e^{-j\omega(n+1)}}{1 - ae^{-j\omega}} e^{j\omega n} \\
 &= a^{n+1}y(-1) - A \frac{a^{n+1}e^{-j\omega(n+1)}}{1 - ae^{-j\omega}} e^{j\omega n} + A \frac{e^{j\omega n}}{1 - ae^{-j\omega}}
 \end{aligned}$$

El sistema es estable si $a < 1$, con lo que se obtiene para $n \rightarrow \infty$ la respuesta de estado permanente (*steady state*):

$$y_{ss}(n) = \lim_{n \rightarrow \infty} y(n) = \frac{A}{1 - ae^{-j\omega}} e^{j\omega n}$$

La respuesta transitoria está dada por los otros términos:

$$\begin{aligned}
 y_{tr}(n) &= a^{n+1} \left\{ y(-1) - \frac{Ae^{-j\omega(n+1)}}{1 - ae^{-j\omega}} e^{j\omega n} \right\} \\
 &= a^{n+1} \left\{ y(-1) - \frac{Ae^{-j\omega}}{1 - ae^{-j\omega}} \right\}
 \end{aligned}$$

y tiende a cero para $n \rightarrow \infty$.

4.5

4.4.3 Respuesta en régimen permanente a señales de entrada periódicas

Sea $x(n)$ una señal periódica, con periodo fundamental N , definida en el rango $-\infty < n < \infty$, por lo que la respuesta de un sistema LTI estable será la respuesta en régimen permanente. La señal $x(n)$ puede entonces expresarse por medio de una serie de Fourier:

$$x(n) = \sum_{k=0}^{N-1} c_k e^{j2\pi k \frac{n}{N}}$$

Puesto que la respuesta a $x_k(n) = c_k e^{j2\pi k \frac{n}{N}}$ es:

$$y_k(n) = c_k H \left(\frac{2\pi}{N} k \right) e^{j2\pi k \frac{n}{N}}$$

entonces, utilizando la linealidad del sistema se obtiene:

$$y(n) = \sum_{k=0}^{N-1} c_k H \left(\frac{2\pi}{N} k \right) e^{j2\pi k \frac{n}{N}} = \sum_{k=0}^{N-1} d_k e^{j2\pi k \frac{n}{N}}$$

con $d_k = c_k H \left(\frac{2\pi}{N} k \right)$. La respuesta tiene entonces el mismo periodo de la entrada, pero otra forma de señal debido a los cambios de amplitud y fase aportados por los coeficientes d_k .

4.4.4 Respuesta a señales de entrada aperiódicas

Puesto que la salida de un sistema LTI en tiempo discreto es:

$$y(n) = h(n) * x(n)$$

se obtiene en el dominio de la frecuencia:

$$Y(\omega) = H(\omega)X(\omega)$$

donde se deriva:

$$\begin{aligned} |Y(\omega)| &= |H(\omega)||X(\omega)| \\ \angle Y(\omega) &= \angle H(\omega) + \angle X(\omega) \end{aligned}$$

La respuesta en magnitud $|H(\omega)|$ indica cuáles frecuencias se amplifican y cuáles se atenúan. Nótese además que la salida de un sistema LTI no puede tener frecuencias que no estén presentes en la entrada y en la respuesta en frecuencia; sólo sistemas no lineales o variantes en el tiempo pueden generar nuevas frecuencias.

A pesar de que la salida $y(n)$ se puede obtener de la respuesta en frecuencia:

$$y(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} Y(\omega) e^{j\omega n} d\omega$$

no es usual emplear la transformada inversa de Fourier por ser relativamente más simple el tratamiento con la transformada z . La relación entre las densidades espectrales de la entrada $S_{xx}(\omega)$ y de la salida $S_{yy}(\omega)$ se obtiene fácilmente con:

$$S_{yy}(\omega) = |Y(\omega)|^2 = |H(\omega)|^2 |X(\omega)|^2 = |H(\omega)|^2 S_{xx}(\omega)$$

La energía de la salida es entonces:

$$E_y = \frac{1}{2\pi} \int_{-\pi}^{\pi} S_{yy}(\omega) d\omega = \frac{1}{2\pi} \int_{-\pi}^{\pi} |H(\omega)|^2 S_{xx}(\omega) d\omega$$

4.4.5 Relaciones entre la función de transferencia y la respuesta en frecuencia

Si la ROC de la función de transferencia contiene a la circunferencia unitaria, entonces:

$$H(\omega) = H(z)|_{z=e^{j\omega}} = \sum_{n=-\infty}^{\infty} h(n) e^{-j\omega n}$$

Si $H(z)$ es racional, $H(z) = \frac{B(z)}{A(z)}$ y asumiendo que $h(n)$ es real, entonces:

$$H(\omega) = \frac{B(\omega)}{A(\omega)} = \frac{\sum_{k=0}^M b_k e^{-j\omega k}}{1 + \sum_{k=1}^N a_k e^{-j\omega k}} = b_0 \frac{\prod_{k=1}^M (1 - z_k e^{-j\omega})}{\prod_{k=1}^N (1 - p_k e^{-j\omega})}$$

con los coeficientes reales $\{a_k\}$ y $\{b_k\}$, y los ceros $\{z_k\}$ y polos $\{p_k\}$ que pueden ser complejos. Si $h(n)$ es real, entonces $H(-\omega) = H^*(\omega)$ y así:

$$|H(\omega)|^2 = H(\omega)H^*(\omega) = H(\omega)H(-\omega) = H(z)H(z^{-1})\Big|_{z=e^{j\omega}} \quad (4.6)$$

$$= \mathcal{Z}\{r_{hh}(n)\}\Big|_{z=e^{j\omega}} \quad (4.7)$$

lo que confirma el teorema de Wiener-Khinchin $|H(\omega)|^2 \bullet \circ r_{hh}(n)$.

Sea $D(z) = B(z)B(z^{-1}) \bullet \circ b(n) * b(-n)$ y $C(z) = A(z)A(z^{-1}) \bullet \circ a(n) * a(-n)$. Puede demostrarse que:

$$c(n) = \sum_{k=0}^{N-|n|} a(k)a(k+n), \quad -N \leq n \leq N, a(k) = a_k$$

$$d(n) = \sum_{k=0}^{M-|n|} b(k)b(k+n), \quad -M \leq n \leq M, b(k) = b_k$$

que corresponden a la autocorrelación de $a(n)$ y $b(n)$, los coeficientes de denominador y numerador de la función de transferencia. Puesto que $A(\omega) = \sum_{k=0}^N a(k)e^{-j\omega k}$ y $B(\omega) = \sum_{k=0}^N b(k)e^{-j\omega k}$, $c_k = c_{-k}$ y $d_k = d_{-k}$:

$$|H(\omega)|^2 = \frac{d_0 + 2 \sum_{k=1}^M d_k \cos k\omega}{c_0 + 2 \sum_{k=1}^N c_k \cos k\omega}$$

donde con $\cos k\omega = \sum_{m=0}^k \beta_m (\cos \omega)^m$, entonces $|H(\omega)|^2$ es una función racional de polinomios de $\cos \omega$. Esto implica que la densidad espectral de energía es continua, suave y si se indefinire, o se hace cero, lo hace en puntos finitos aislados.

4.4.6 Cálculo de la respuesta en frecuencia

Para evaluar el efecto de los polos y ceros de la función de transferencia en la respuesta en frecuencia, evalúese entonces:

$$H(\omega) = b_0 \frac{\prod_{k=1}^M (1 - z_k e^{-j\omega})}{\prod_{k=1}^N (1 - p_k e^{-j\omega})} = b_0 e^{j\omega(N-M)} \frac{\prod_{k=1}^M (e^{j\omega} - z_k)}{\prod_{k=1}^N (e^{j\omega} - p_k)}$$

y con $e^{j\omega} - z_k = V_k(\omega)e^{j\theta_k(\omega)}$, $e^{j\omega} - p_k = U_k(\omega)e^{j\phi_k(\omega)}$, $V_k(\omega) = |e^{j\omega} - z_k|$, $\theta_k(\omega) = \angle(e^{j\omega} - z_k)$, $U_k(\omega) = |e^{j\omega} - p_k|$, $\phi_k(\omega) = \angle(e^{j\omega} - p_k)$, se obtiene:

$$|H(\omega)| = |b_0| \frac{V_1(\omega) \dots V_M(\omega)}{U_1(\omega) \dots U_M(\omega)}$$

$$\angle H(\omega) = \angle b_0 + \omega(N - M) + \theta_1(\omega) + \dots + \theta_M(\omega) - [\phi_1(\omega) + \dots + \phi_N(\omega)]$$

Nótese que $(e^{j\omega} - p_k)$ puede interpretarse geoméricamente como el fasor que va del polo p_k al punto $e^{j\omega}$ que se encuentra en la circunferencia unitaria con el ángulo ω . El mismo concepto se aplica a los ceros.

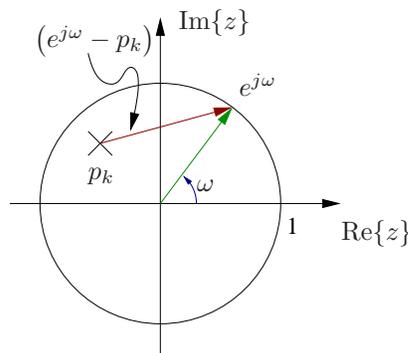


Figura 4.15: Interpretación geométrica del factor $(e^{j\omega} - p_k)$.

La respuesta en frecuencia $H(\omega)$ se hará muy pequeña si $z = e^{j\omega}$ pasa cerca de un cero, o se hará muy grande si pasa cerca de un polo.

Nótese que si la magnitud de $H(\omega)$ se expresa en decibelios, entonces:

$$|H(\omega)|_{dB} = 20 \log_{10} |b_0| + 20 \sum_{k=1}^M \log_{10} V_k(\omega) - 20 \sum_{k=1}^N \log_{10} U_k(\omega)$$

4.5 Sistemas LTI como filtros selectivos en frecuencia

Filtro es un caso particular de sistema que discrimina algún atributo de los objetos o señales aplicados a su entrada, y permite o suprime el paso de dichos objetos o señales si y solo si estos presentan ese atributo. Así, filtros de aire y aceite permiten que estas sustancias los atraviesen, evitando que otras impurezas como polvo alcancen su salida; filtros de luz (por ejemplo, luz infrarroja) permiten el paso de sólo una parte del espectro electromagnético. En las secciones anteriores se presentó la propiedad de los sistemas LTI de filtrar diferentes componentes en frecuencia de su señal de entrada, caracterizada por la respuesta en frecuencia $H(\omega)$. Es por ello que los términos *sistema LTI* y *filtro* se utilizan con frecuencia como sinónimos.

Eligiendo la posición de los polos y ceros en la función de transferencia se eligen los coeficientes a_k y b_k para obtener una determinada forma de la respuesta en frecuencia $H(\omega)$, que puede considerarse como función de ponderación o conformación espectral, pues la salida del sistema es $Y(\omega) = H(\omega)X(\omega)$.

Los sistemas LTI en estas funciones de filtrado selectivo de frecuencias o conformación espectral se utilizan de varias maneras: como eliminadores de ruido, conformación espectral para igualación de canales de comunicación, detección de señales en radar y sonar, análisis espectral, etc. En esta sección se revisaran algunos conceptos básicos, que serán retomados en el capítulo 7 con mayor detalle.

4.5.1 Filtros ideales

Como filtro ideal se conoce un sistema LTI con respuesta en frecuencia

$$H(\omega) = \begin{cases} Ce^{-j\omega n_0} & \omega_1 < \omega < \omega_2 \\ 0 & \text{en caso contrario} \end{cases}$$

es decir, una ganancia constante C en la banda de paso, y cero en la banda eliminada. Algunos filtros selectivos se muestran en la figura 4.16.

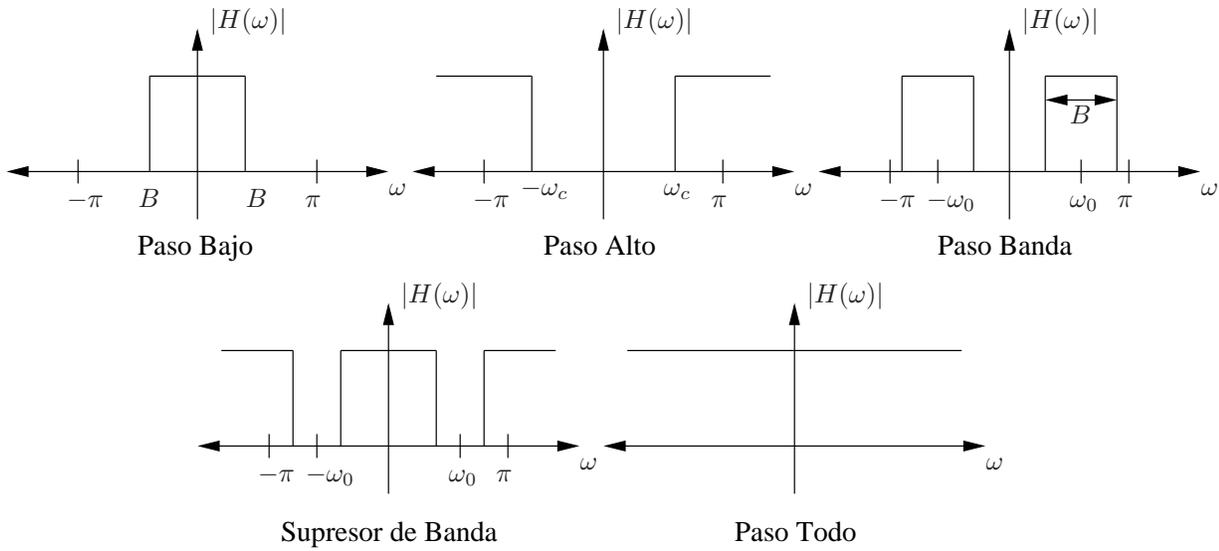


Figura 4.16: Respuesta en magnitud de algunos filtros ideales selectivos en frecuencia.

La respuesta de fase lineal permite que señales de entrada con componentes frecuenciales confinadas en el intervalo $\omega \in [\omega_1, \omega_2]$ sean simplemente escaladas y trasladadas en el tiempo, pues:

$$\begin{aligned} Y(\omega) &= H(\omega)X(\omega) \\ &= Ce^{-j\omega n_0}X(\omega) \\ &= CX(\omega)e^{-j\omega n_0} \bullet \circ y(n) = Cx(n - n_0) \end{aligned}$$

El retardo no es considerado como una distorsión grave de la señal y es usualmente tolerado, tal y como el escalado en amplitud. En estos filtros la derivada de la fase tiene unidades de retardo, y a la magnitud

$$\tau_g(\omega) = -\frac{d\theta(\omega)}{d\omega}, \quad \theta(\omega) = \angle H(\omega)$$

se le denomina *retardo de envolvente* o *retardo de grupo* del filtro. En el caso de los filtros ideales, puesto que $\theta(\omega) = -\omega n_0$, entonces $\tau_g(\omega) = n_0$ y es constante. En general, $\tau_g(\omega)$ indica el retardo temporal que experimenta la componente de frecuencia ω cuando pasa a través del sistema. Entonces, en los filtros ideales todas las componentes se retrasan.

Los filtros ideales no son realizables. Por ejemplo, la respuesta impulsional del filtro paso bajo es:

$$h_{lp}(n) = \frac{\text{sen}(\omega_c \pi n)}{\pi n}, \quad -\infty < n < \infty$$

que es anticausal, infinita y no es absolutamente sumable, por lo que el sistema es inestable.

Algunos filtros digitales básicos pueden diseñarse colocando polos y ceros en el plano z , cerca de la circunferencia unitaria, de tal forma que se obtenga la forma de la respuesta en magnitud deseada, donde los polos deben estar en el interior de la circunferencia unidad, y si hay polos o ceros complejos, deben presentarse entonces en pares conjugados.

4.5.2 Filtros paso alto, paso bajo y paso banda

En el diseño de filtros paso bajo, los polos deben situarse cerca de los puntos de la circunferencia unidad correspondientes a las bajas frecuencias (cerca de $\omega = 0$) y los ceros cerca de los puntos de la circunferencia unidad correspondientes a las altas frecuencias (cerca de $\omega = \pi$). El caso contrario es para filtros paso alto.

Ejemplo 4.6 Diseñe un filtro paso bajo con $|H(0)| = 1$ y

1. un solo polo en $p_1 = 0,9$
2. un polo y un cero en $p_1 = 0,9$ y $z_1 = -1$

1. En general, un sistema LTI tiene la forma:

$$H(z) = \frac{\sum_{k=0}^M b_k z^{-k}}{1 + \sum_{k=0}^N a_k z^{-k}} = b_0 \frac{\prod_{k=1}^M (1 - z_k z^{-1})}{\prod_{k=1}^N (1 - p_k z^{-1})}$$

Un filtro de primer orden sin ceros tendrá la función de transferencia

$$H_1(z) = \frac{b_0}{1 - p_1 z^{-1}} = \frac{b_0}{1 - 0,9z^{-1}} \Rightarrow H_1(\omega) = \frac{b_0}{1 - 0,9e^{-j\omega}}$$

Puesto que $|H_1(0)| = 1 = \frac{b_0}{0,1} \Rightarrow b_0 = 0,1$; $H_1(\pi) = \frac{0,1}{1,9} = 0,0526$.

2. En este caso

$$H_2(\omega) = \frac{b_0(1 - z_1 z^{-1})}{1 - p_1 z^{-1}} = b_0 \frac{1 + z^{-1}}{1 - 0,9z^{-1}} \Rightarrow H_2(\omega) = b_0 \frac{1 + e^{-j\omega}}{1 - 0,9e^{-j\omega}}$$

y $H_2(0) = b_0 \frac{1+1}{0,1} = 20b_0 \Rightarrow b_0 = \frac{1}{20}$, $H_2(\pi) = 0$

4.6

Reflejando los polos y ceros con respecto al eje imaginario, se obtienen filtros paso alto:

$$H_1(\omega) = \frac{b_0}{1 + 0,9e^{-j\omega}}, \quad H_2(\omega) = \frac{b_0(1 - e^{-j\omega})}{1 + 0,9e^{-j\omega}}$$

Para filtros paso banda, se debe escoger un par de polos complejos conjugados cerca de la circunferencia unidad en la vecindad de la banda de paso.

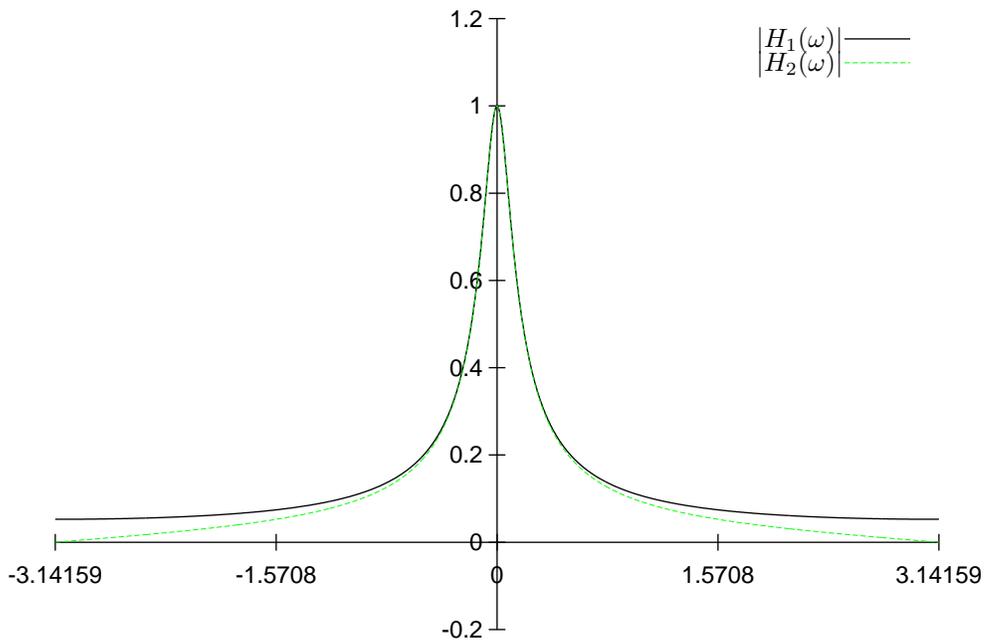


Figura 4.17: Respuesta en magnitud de (1) filtro paso bajo de un polo, (2) filtro paso bajo de un polo y un cero; $H_1(z) = (1-a)/(1-az^{-1})$, $H_2(z) = [(1-a)/2][(1+z^{-1})/(1-az^{-1})]$ y $a = 0,9$.

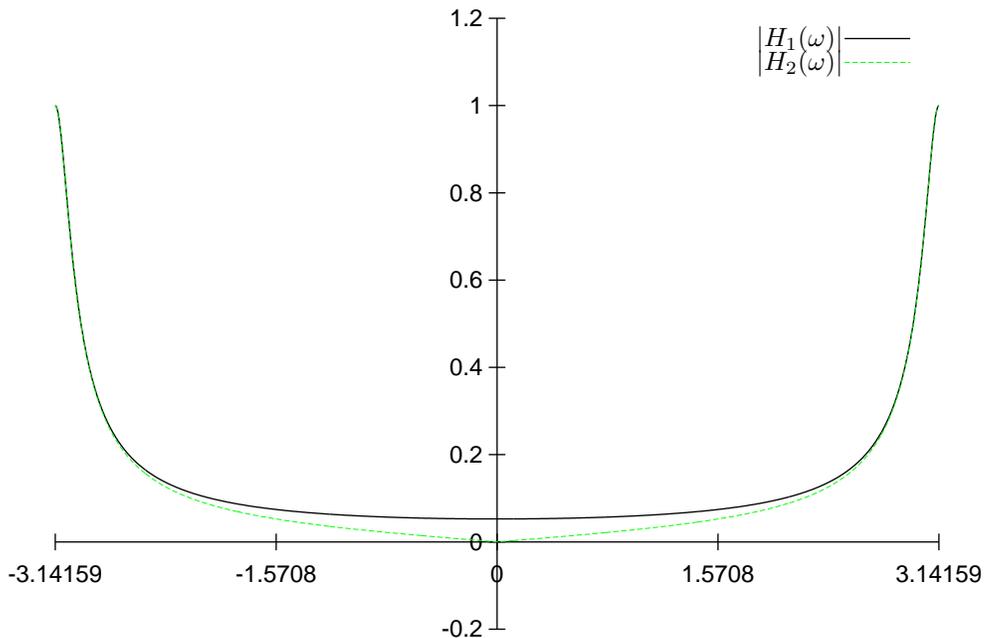


Figura 4.18: Respuesta en magnitud de un filtro paso alto $H_1(z) = (1-a)/(1+az^{-1})$, $H_2(z) = [(1-a)/2][(1-z^{-1})/(1+az^{-1})]$ y $a = 0,9$.

Ejemplo 4.7 Diseñe un filtro paso banda de segundo orden con banda de paso en $\omega = \frac{\pi}{2}$, con $|H(0)| = |H(\pi)| = 0$, $|H(\frac{4\pi}{9})| = \frac{1}{\sqrt{2}}$, $|H(\frac{\pi}{2})| = 1$.

Solución: Los polos deben estar en $p_{1,2} = re^{\pm j\frac{\pi}{2}} = \pm rj$, $0 < r < 1$, y los ceros en $z_1 = 1$ y

$z_2 = -1$, por lo que:

$$H(z) = b_0 \frac{(z-1)(z+1)}{(z-jr)(z+jr)} = b_0 \frac{z^2-1}{z^2+r^2} \Rightarrow H(\omega) = b_0 \frac{e^{2j\omega}-1}{e^{2j\omega}+r^2}$$

$$\left| H\left(\frac{\pi}{2}\right) \right| = \left| b_0 \frac{-2}{1-r^2} \right| \stackrel{!}{=} 1 \Rightarrow b_0 \frac{2}{1-r^2} = 1 \Rightarrow b_0 = \frac{1-r^2}{2}$$

$$\left| H\left(\frac{4\pi}{9}\right) \right|^2 = \left(\frac{1-r^2}{2} \right)^2 \frac{2-2\cos\left(\frac{8\pi}{9}\right)}{1+r^4+2r^2\cos\left(\frac{8\pi}{9}\right)} \stackrel{!}{=} \frac{1}{2}$$

La solución resulta en $r^2 = 0,7$ (figura 4.19).

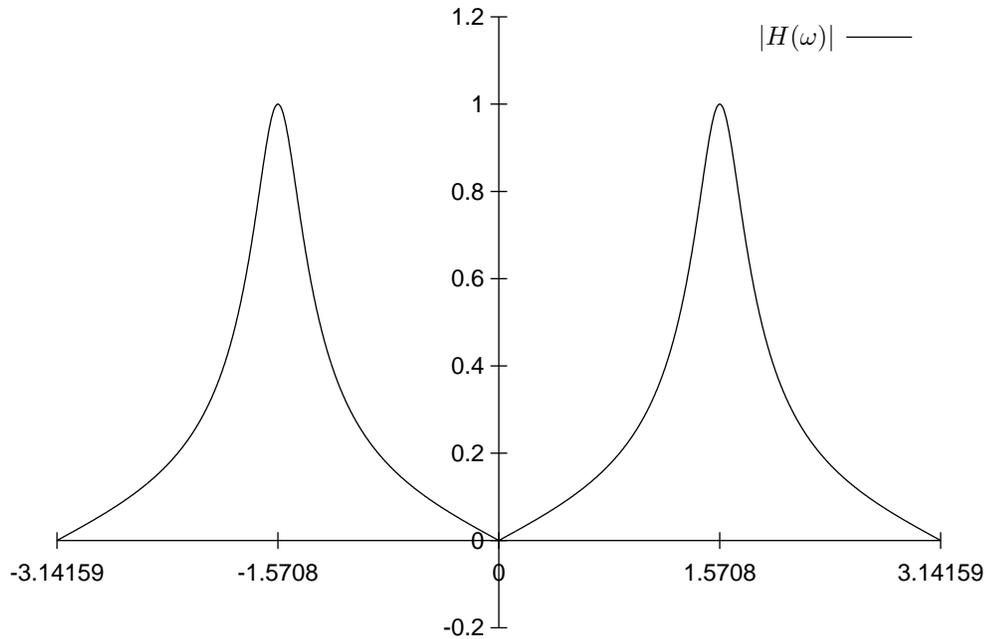


Figura 4.19: Respuesta en magnitud de filtro paso banda; $0,15[(1-z^{-2})/(1+0,7z^{-2})]$.

4.7

Otra manera de convertir un filtro paso bajos en uno paso altos es aplicar el desplazamiento en frecuencia la respuesta $H(\omega)$ π radianes:

$$H_{hp} = H_{lp}(\omega - \pi)$$

con $H_{hp}(\omega)$ la respuesta del filtro paso alto (hp: high pass) y $H_{lp}(\omega)$ la respuesta del filtro paso bajo (lp: low pass). Nótese que la respuesta impulsional sería

$$h_{hp}(n) = (e^{j\pi})^n h_{lp}(n) = (-1)^n h_{lp}(n)$$

Ahora interesa observar qué le ocurre a los coeficientes de la ecuación de diferencias con este desplazamiento en frecuencia. Puesto que:

$$y(n) = -\sum_{k=1}^N a_k y(n-k) + \sum_{k=0}^M b_k x(n-k)$$

$$H_{lp}(\omega) = \frac{\sum_{k=0}^M b_k e^{-j\omega k}}{1 + \sum_{k=1}^N a_k e^{-j\omega k}}$$

Sustituyendo ω por $\omega + \pi$ se obtiene:

$$H_{hp}(\omega) = \frac{\sum_{k=0}^M (-1)^k b_k e^{-j\omega k}}{1 + \sum_{k=1}^N (-1)^k a_k e^{-j\omega k}}$$

que equivale a:

$$y(n) = - \sum_{k=1}^N (-1)^k a_k y(n-k) + \sum_{k=0}^M (-1)^k b_k x(n-k)$$

y por lo tanto el único cambio debe hacerse a los coeficientes a_k y b_k con k impar, quienes cambian su signo.

4.5.3 Resonadores digitales

Un *resonador digital* es un filtro paso banda con un par de polos complejos conjugados, que determinan con su posición angular la frecuencia de resonancia del filtro y con un máximo de dos ceros que se sitúan en el origen o en $z = \pm 1$ (figura 4.20).

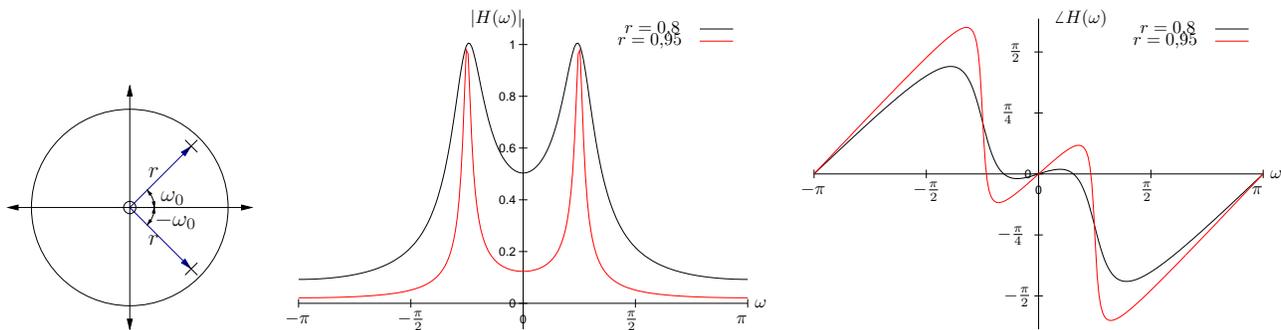


Figura 4.20: Resonadores digitales todo-polos.

Para el caso con los ceros en el origen, se obtiene como función de transferencia:

$$H(z) = \frac{b_0}{(1 - re^{j\omega_0} z^{-1})(1 - re^{-j\omega_0} z^{-1})} = \frac{b_0}{1 - (2r \cos \omega_0) z^{-1} + r^2 z^{-2}}$$

Dado que las bandas de paso se encuentran centradas en ω_0 o cerca de ω_0 , b_0 se elige de tal modo que $|H(\omega_0)| = 1$

$$\begin{aligned} H(\omega_0) &= \frac{b_0}{(1 - re^{j\omega_0} e^{-j\omega_0})(1 - re^{-j\omega_0} e^{j\omega_0})} = \frac{b_0}{(1-r)(1-re^{-j2\omega_0})} \\ |H(\omega_0)| &= \frac{b_0}{(1-r)\sqrt{1+r^2-2r\cos 2\omega_0}} = 1 \\ \Rightarrow b_0 &= (1-r)\sqrt{1+r^2-2r\cos 2\omega_0} \end{aligned}$$

Si se expresa $|H(\omega)|$ como $\frac{b_0}{U_1(\omega)U_2(\omega)}$ y $\theta(\omega)$ como $2\omega - \phi_1(\omega) - \phi_2(\omega)$, donde $(1 - p_1 z^{-1}) = (1 - re^{j\omega_0} z^{-1}) = U_1(\omega)e^{j\phi_1(\omega)}$, $(1 - p_2 z^{-1}) = (1 - re^{-j\omega_0} z^{-1}) = U_2(\omega)e^{j\phi_2(\omega)}$ entonces

$$U_1(\omega) = \sqrt{1 + r^2 - 2r \cos(\omega - \omega_0)}, \quad U_2(\omega) = \sqrt{1 + r^2 - 2r \cos(\omega + \omega_0)}$$

U_1 alcanza su mínimo en $\omega = \omega_0$ ($U_1(\omega_0) = 1 - r$) y U_2 en $\omega = -\omega_0$ ($U_2(-\omega_0) = 1 - r$).

Sin embargo, su producto $U_1(\omega)U_2(\omega)$ alcanza su mínimo en (figura 4.21):

$$\omega_r = \cos^{-1} \left(\frac{1 + r^2}{2r} \cos \omega_0 \right), \quad \lim_{r \rightarrow 1} \omega_r = \omega_0$$

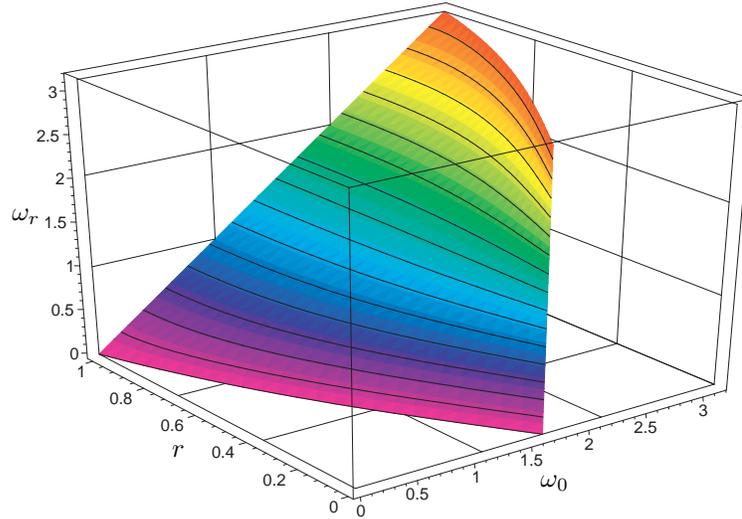


Figura 4.21: Posición de la frecuencia de resonancia con respecto al radio y al ángulo de los polos.

Puede demostrarse que el ancho de banda de -3 dB es aproximadamente $\Delta\omega = 2(1 - r)$ para $r \rightarrow 1$.

Si los ceros se sitúan en $\omega = 0$ y $\omega = \pi$, la función de transferencia es:

$$H(z) = G \frac{(1 - z^{-1})(1 + z^{-1})}{(1 - re^{j\omega_0}z^{-1})(1 - re^{-j\omega_0}z^{-1})} = G \frac{(1 - z^{-2})}{1 - (2r \cos \omega_0)z^{-1} + r^2z^{-2}}$$

La presencia de ceros cambia la posición de la frecuencia de resonancia (figura 4.22): es ahora más cercana a ω_0 . Además, su ancho de banda es usualmente menor.

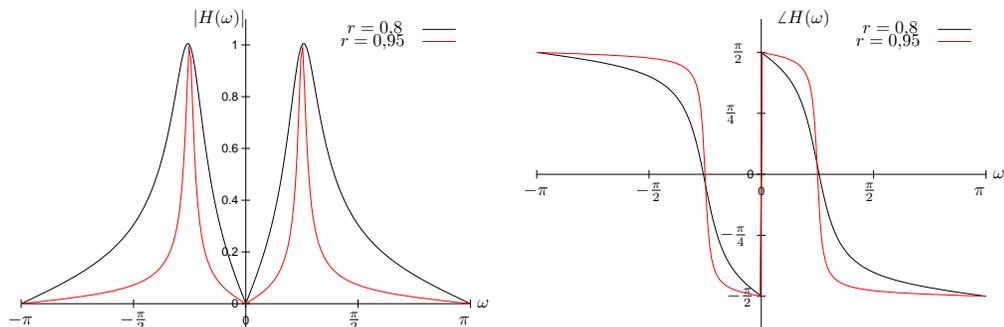


Figura 4.22: Resonadores digitales con ceros en $\omega = 0$ y $\omega = \pi$.

4.5.4 Filtros ranura

Un *filtro ranura o muesca*, es un filtro con uno o más cortes profundos, idealmente ceros perfectos en su respuesta en frecuencia. Para crear éstos en la frecuencia ω_0 , se introduce un par de ceros complejos conjugados en la circunferencia unitaria con ángulo ω_0 . Es decir, $z_{1,2} = e^{\pm j\omega_0}$, lo que resulta en:

$$H(z) = b_0(1 - e^{j\omega_0} z^{-1})(1 - e^{-j\omega_0} z^{-1}) = b_0(1 - 2 \cos \omega_0 z^{-1} + z^{-2})$$

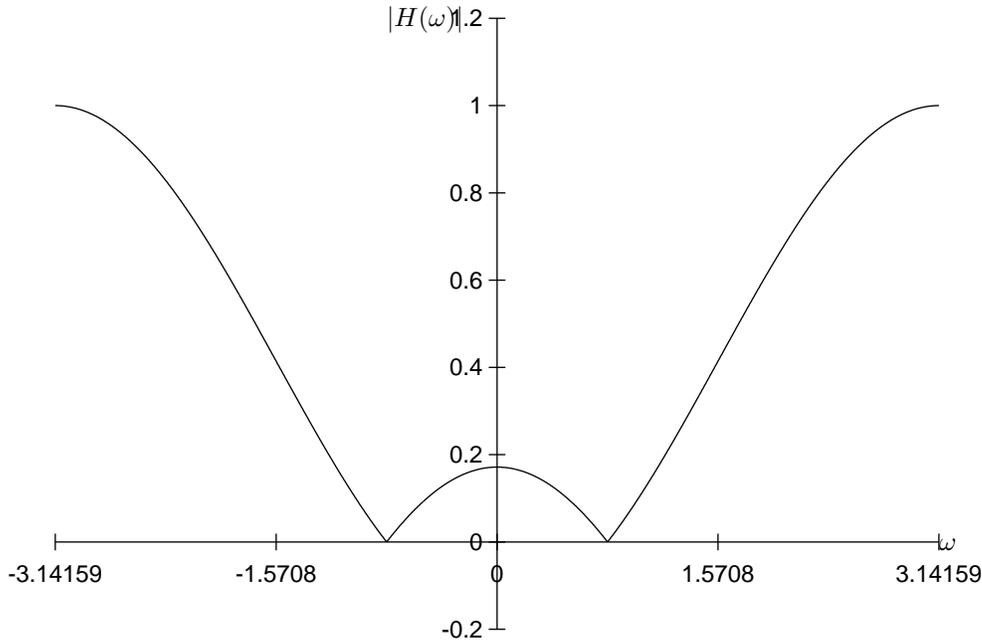


Figura 4.23: Característica de respuesta en frecuencia de un filtro ranura.

El problema de esta configuración es que el corte tiene un ancho de banda relativamente grande. Esto puede ser atenuado situando polos complejos conjugados en $p_{1,2} = re^{\pm j\omega_0}$, que introducen resonancia en la vecindad del cero, y por tanto reducen el ancho de banda de la banda suprimida. La función de transferencia resultante será:

$$H(z) = b_0 \frac{1 - 2 \cos \omega_0 z^{-1} + z^{-2}}{1 - 2r \cos \omega_0 z^{-1} + r^2 z^{-2}}$$

4.5.5 Filtros peine

Un *filtro peine* puede verse como un filtro ranura en el que los ceros aparecen periódicamente en toda la banda de frecuencias. Se utilizan por ejemplo en el rechazo de armónicos en líneas de potencia y en la separación de componentes solares y lunares de las medidas ionosféricas de la concentración de electrones.

La forma más sencilla de filtro peine es el filtro de media móvil (FIR):

$$y(n) = \frac{1}{M+1} \sum_{k=0}^M x(n-k)$$

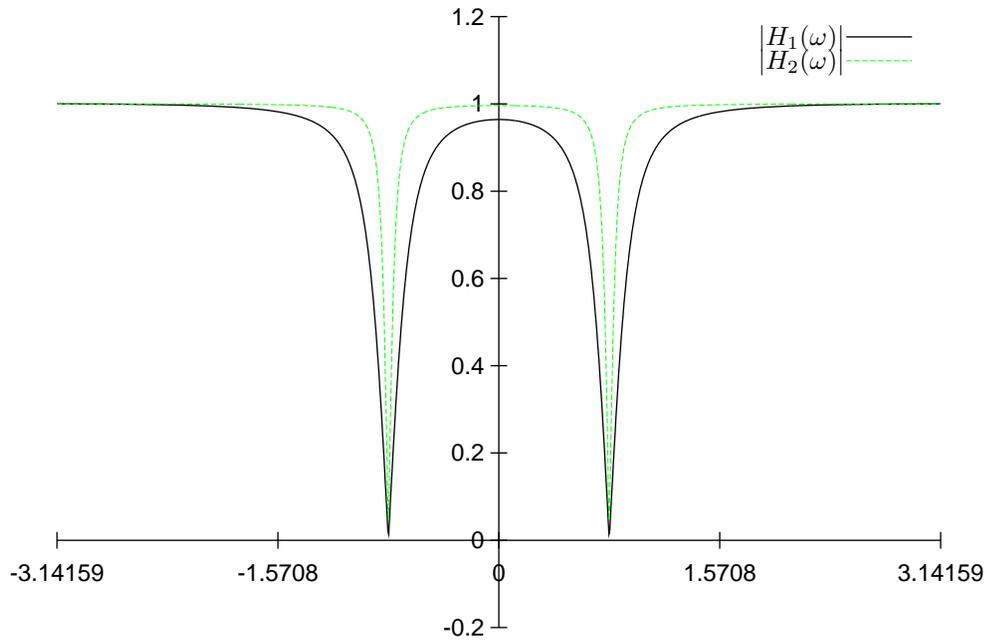


Figura 4.24: Característica de respuesta en frecuencia de un filtro ranura con un cero en $\omega = \frac{\pi}{4}$;

$$H_i(z) = G_i \frac{1 - 2 \cos \omega_0 z^{-1} + z^{-2}}{1 - 2r_i \cos \omega_0 z^{-1} + r_i^2 z^{-2}}, \text{ para } r_1 = 0,85 \text{ y } r_2 = 0,95$$

con función de transferencia

$$H(z) = \frac{1}{M+1} \sum_{k=0}^M z^{-k} = \frac{1 - z^{-(M+1)}}{1 - z^{-1}}$$

y respuesta de frecuencia

$$H(\omega) = \frac{e^{-j\omega \frac{M}{2}} \operatorname{sen} \left(\omega \frac{M+1}{2} \right)}{M+1 \operatorname{sen} \left(\frac{\omega}{2} \right)}$$

De $H(z)$ se deduce con $1 - z^{-(M+1)} = 0 \Rightarrow 1 = e^{2k\pi} = z^{M+1} \Rightarrow z = e^{\frac{2k\pi}{M+1}}$, $k = 1, 2, \dots, M$. Nótese que el polo en $z = 1$ se cancela con el cero en $z = 1$ ($k = 0$), por lo que el único polo está en $z = 0$.

Otra manera de generar filtros peine es a partir de cualquier filtro FIR con al menos un cero en cierta frecuencia ω_0 . Si este filtro tiene función de transferencia $H(z) = \sum_{k=0}^M h(k)z^{-k}$ y se reemplaza z por z^L , con $L \in \mathbb{Z}$, $L > 0$, entonces el nuevo filtro FIR tiene función de transferencia

$$H_L(z) = \sum_{k=0}^M h(k)z^{-kL}$$

Si la respuesta en frecuencia correspondiente a $H(z)$ es $H(\omega)$, entonces:

$$H_L(\omega) = \sum_{k=0}^M h(k)e^{-jkL\omega} = H(L\omega)$$

es decir, una repetición de orden L de $H(\omega)$ en el rango $0 \leq \omega \leq \pi$, donde $H(\omega)$ se comprime para que todas las repeticiones ocupen dicho intervalo.

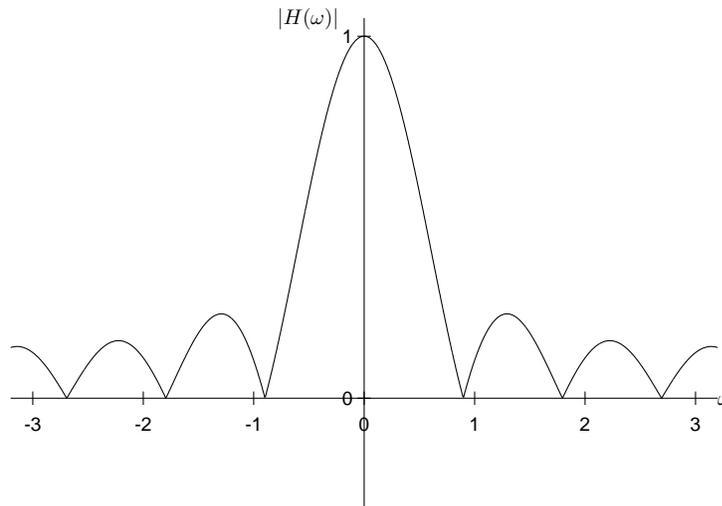


Figura 4.25: Ejemplo de filtro peine FIR con $M = 6$.

Ejemplo 4.8 Diseñe un filtro peine a partir del filtro FIR

$$y(n) = \frac{1}{2}(x(n) + x(n-1])$$

tal que suprima las frecuencias $\omega = \frac{\pi}{4} + k\frac{\pi}{2}$.

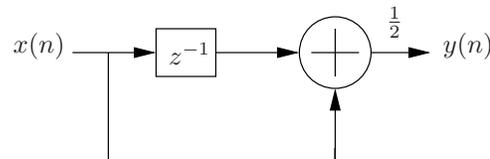


Figura 4.26: Diagrama de bloques del ejemplo (4.8).

La función de transferencia es:

$$H(z) = \frac{1 + z^{-1}}{2}$$

y por lo tanto (figura 4.27)

$$H(\omega) = \frac{1 + e^{-j\omega}}{2} = \frac{e^{-j\frac{\omega}{2}}(e^{+j\frac{\omega}{2}} + e^{-j\frac{\omega}{2}})}{2} = e^{-j\frac{\omega}{2}} \cos\left(\frac{\omega}{2}\right)$$

Utilizando $L = 4$ se obtiene $H_L(\omega) = e^{-j2\omega} \cos(2\omega)$, proveniente de $H(z) = \frac{1+z^{-4}}{2}$ (figura 4.28).

4.8

Nótese entonces que si $H(\omega)$ tiene un cero en ω_0 , $H_L(\omega)$ tendrá ceros en $\frac{\omega_0 + 2\pi k}{L}$, $k = 0, 1, 2, \dots, L-1$.

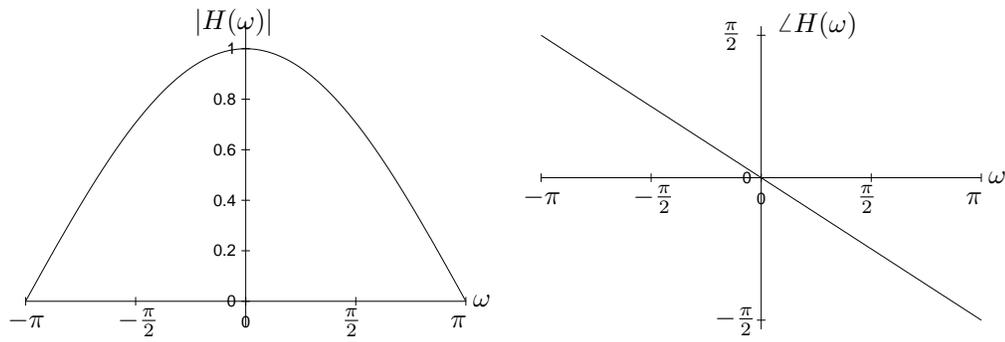


Figura 4.27: Respuesta en frecuencia de $y(n) = (x(n) + x(n-1))/2$.

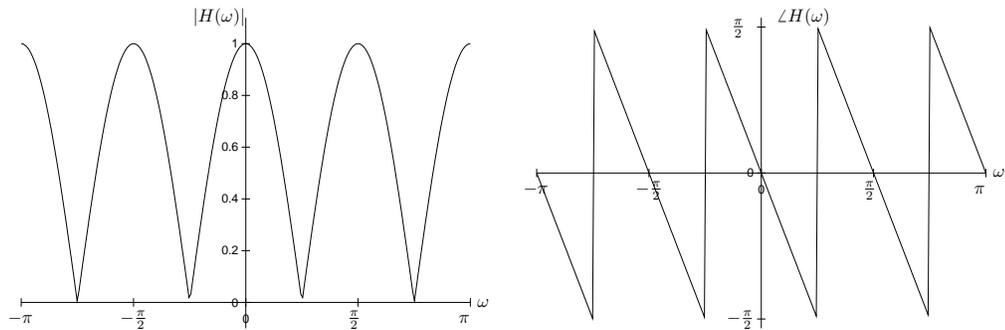


Figura 4.28: Respuesta en frecuencia de $y(n) = (x(n) + x(n-1))/2$.

4.5.6 Filtros paso todo

Un *filtro paso todo* tiene una respuesta en magnitud constante para todas las frecuencias:

$$|H(\omega)| = 1, \quad 0 \leq \omega \leq \pi$$

El ejemplo más simple es el registro de desplazamiento:

$$H(z) = z^{-k} \Rightarrow H(\omega) = e^{-jk\omega} \Rightarrow \begin{cases} |H(\omega)| = 1 \\ \angle H(\omega) = -k\omega \end{cases}$$

Otro caso está dado por filtros con funciones de transferencia:

$$H(z) = z^{-N} \frac{A(z^{-1})}{A(z)}, \quad \text{con el polinomio } A(z) = \sum_{k=0}^N a_k z^{-k}, \quad a_0 = 1,$$

es decir,

$$H(z) = \frac{a_N + a_{N-1}z^{-1} + \dots + a_1z^{-N+1} + z^{-N}}{1 + a_1z^{-1} + \dots + a_Nz^{-N}}$$

con el resultado de (4.6) en la página 116,

$$|H(\omega)|^2 = H(z)H(z^{-1}) \Big|_{z=e^{j\omega}} = z^{-N+N} \frac{A(z^{-1})}{A(z)} \frac{A(z)}{A(z^{-1})} \Big|_{z=e^{j\omega}} = 1.$$

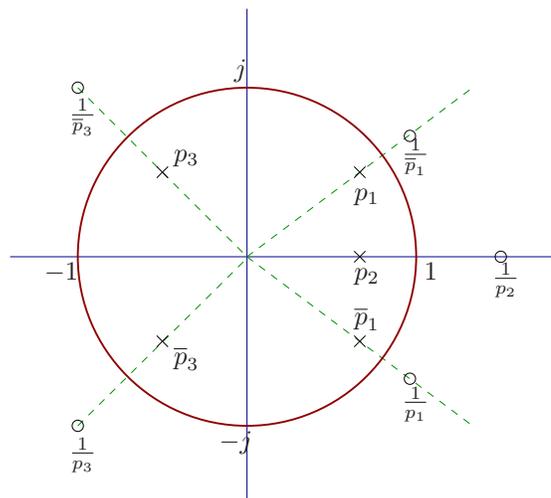


Figura 4.29: Correspondencia entre polos y ceros de un filtro pasa-todo

Nótese que si z_0 es un cero, entonces $\frac{1}{z_0}$ es un polo, lo que quiere decir que si algún polo o cero está dentro del círculo unitario, la contraparte estará fuera (figura 4.29).

Si este filtro paso todo tiene polos reales α_k y polos complejos β_k entonces puede expresarse como:

$$H_{ap}(z) = \prod_{k=1}^{N_R} \frac{z^{-1} - \alpha_k}{1 - \alpha_k z^{-1}} \prod_{k=1}^{N_C} \frac{(z^{-1} - \beta_k)(z^{-1} - \beta_k^*)}{(1 - \beta_k z^{-1})(1 - \beta_k^* z^{-1})}$$

donde hay N_R polos y ceros reales, y N_C pares de polos y ceros complejos conjugados. Si el sistema es causal y estable, entonces $|\alpha_k| < 1$ y $|\beta_k| < 1$. Para analizar el comportamiento de la fase y retardo de grupo puede tomarse uno de los términos anteriores y con $\alpha_k = r e^{j\theta}$:

$$\begin{aligned} H_{ap_k}(z) &= \frac{z^{-1} - \alpha_k^*}{1 - \alpha_k z^{-1}} \Rightarrow \\ H_{ap_k}(\omega) &= \frac{e^{-j\omega} - r e^{-j\theta}}{1 - r e^{j\theta} e^{-j\omega}} = e^{-j\omega} \frac{1 - r e^{j(\omega - \theta)}}{1 - r e^{j(\theta - \omega)}} \\ &= e^{-j\omega} \frac{1 - r \cos(\omega - \theta) - jr \sin(\omega - \theta)}{1 - r \cos(\omega - \theta) + jr \sin(\omega - \theta)} \\ \theta_{ap}(\omega) &= \angle(H_{ap_k}(z)) = -\omega - 2 \arctan\left(\frac{r \sin(\omega - \theta)}{1 - r \cos(\omega - \theta)}\right) \end{aligned}$$

y con $\frac{\partial}{\partial x} \arctan(x) = \frac{1}{1+x^2}$ se obtiene para el retardo de grupo:

$$\begin{aligned} \tau(\omega) &= -\frac{\partial \theta_{ap}(\omega)}{\partial \omega} \\ &= +1 + 2 \frac{1}{1 + \frac{r^2 \sin^2(\omega - \theta)}{(1 - r \cos(\omega - \theta))^2}} \frac{r \cos(\omega - \theta)(1 - r \cos(\omega - \theta)) - r \sin(\omega - \theta)r \sin(\omega - \theta)}{(1 - r \cos(\omega - \theta))^2} \\ &= +1 + 2 \frac{(r \cos(\omega - \theta) - r^2 \cos^2(\omega - \theta) - r^2 \sin^2(\omega - \theta))}{1 - 2r \cos(\omega - \theta) + r^2 \cos^2(\omega - \theta) + r^2 \sin^2(\omega - \theta)} \end{aligned}$$

$$\begin{aligned}
&= +1 + 2 \frac{(r \cos(\omega - \theta) - r^2)}{1 - 2r \cos(\omega - \theta) + r^2} \\
&= \frac{1 + r^2 - 2r \cos(\omega - \theta) + 2r \cos(\omega - \theta) - 2r^2}{1 + r^2 - 2r \cos(\omega - \theta)} \\
&= \frac{1 - r^2}{1 + r^2 - 2r \cos(\omega - \theta)} > 0, \forall r < 1, r > 0
\end{aligned}$$

Un sistema pasotodo con varios polos y ceros tendrá siempre un retardo de grupo positivo, por estar conformado el retardo total por la suma de términos como el anterior, que serán siempre positivos.

Este tipo de filtros se utilizan como “igualadores de fase”, colocándolos en cascada con un sistema de respuesta en fase no deseada, de tal modo que se compensa la fase total del sistema para obtener un comportamiento lineal.

4.5.7 Osciladores digitales

Los *osciladores* son el caso extremo de un resonador digital, donde los polos complejos conjugados se encuentran sobre la circunferencia unidad. El sistema de segundo orden:

$$H(z) = \frac{b_0}{1 + a_1 z^{-1} + a_2 z^{-2}}$$

con $a_1 = -2r \cos \omega_0$, $a_2 = r^2$, tiene polos complejos conjugados en $p = r e^{\pm j\omega_0}$ y respuesta impulsional:

$$h(n) = \frac{b_0 r^n}{\sin \omega_0} \sin((n+1)\omega_0) u(n)$$

Para el caso $r = 1$ entonces:

$$h(n) = \frac{b_0 r^n}{\sin \omega_0} \sin((n+1)\omega_0) u(n) = A \sin((n+1)\omega_0) u(n)$$

con $A = \frac{b_0}{\sin \omega_0}$. Es decir, cuando el sistema se alimenta con $\delta(n)$, oscila indefinidamente. La ecuación de diferencias:

$$y(n) = -a_1 y(n-1) - y(n-2) + b_0 \delta(n) = 2 \cos \omega_0 y(n-1) - y(n-2) + A \sin \omega_0 \delta(n)$$

se puede realizar con el sistema mostrado en la figura 4.30.

Nótese que el mismo resultado puede obtenerse con entrada cero, pero con $y(n-1) = 0$ y $y(n-2) = -A \sin \omega_0$.

4.6 Sistemas inversos

Por medio de la convolución, los sistemas LTI transforman una señal $x(n)$ en la salida $y(n)$. En ciertos casos se desconocen exactamente las características del sistema y se desea,

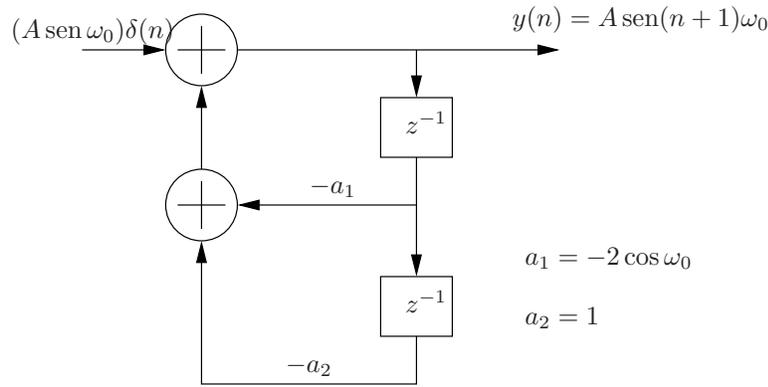


Figura 4.30: Generador digital de sinusoides.

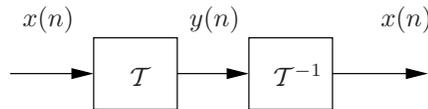


Figura 4.31: Sistema \mathcal{T} en cascada con su inverso \mathcal{T}^{-1} .

teniendo $y(n)$, determinar $x(n)$. Esto ocurre por ejemplo cuando se desea contrarrestar los efectos de distorsión que produce un canal de transmisión de datos. Se desea entonces diseñar un sistema que en cascada con el sistema original, produzca la señal de entrada.

A este segundo sistema se le denomina *sistema inverso* y a la operación que toma $y(n)$ y produce $x(n)$ se le denomina *deconvolución*.

La determinación del sistema inverso requiere conocimiento sobre la respuesta impulsional $h(n)$ o la respuesta en frecuencia $H(\omega)$, que se obtienen en un proceso llamado “identificación del sistema”. Para esto se alimenta al sistema desconocido con señales de entrada conocidas (por ejemplo, senoidales de frecuencia y amplitud conocidas) y se revisan las salidas.

4.6.1 Invertibilidad de sistemas LTI

Un sistema es *invertible* si existe una correspondencia biunívoca entre sus señales de entrada y de salida. Esto quiere decir que si se conoce la salida $y(n)$ para todo n , entonces se puede inferir la entrada $x(n)$. El sistema inverso se denota con \mathcal{T}^{-1} y se cumple que:

$$y(n) = \mathcal{T}[x(n)], \quad x(n) = \mathcal{T}^{-1}[y(n)] = \mathcal{T}^{-1}[\mathcal{T}[x(n)]]$$

Para sistemas LTI, asúmase que el sistema \mathcal{T} tiene una respuesta impulsional $h(n)$, y $h_I(n)$ es la respuesta del sistema inverso \mathcal{T}^{-1} . Se debe cumplir entonces que:

$$h(n) * h_I(n) = \delta(n) \quad \text{---} \bullet \quad H(z)H_I(z) = 1$$

de donde se obtiene

$$H_I(z) = \frac{1}{H(z)}$$

Para un sistema con $H(z)$ racional, si:

$$H(z) = \frac{B(z)}{A(z)}$$

entonces

$$H_I(z) = \frac{A(z)}{B(z)}$$

lo que quiere decir que los ceros de $H(z)$ se convierten en los polos de $H_I(z)$ y viceversa. Esto a su vez implica que si $H(z)$ es todo ceros (FIR), entonces $H_I(z)$ es todo polos (IIR) y viceversa.

Ejemplo 4.9 Determine el sistema inverso de $h(n) = u(n)$.

$$H(z) = \frac{1}{1 - z^{-1}} \Rightarrow H_I(z) = 1 - z^{-1} \bullet \circ h_I(n) = \delta(n) - \delta(n - 1)$$

4.9

Ejemplo 4.10 Determine el sistema inverso del sistema con respuesta impulsional $h(n) = \delta(n) - \delta(n - 1)$. Con

$$H(z) = 1 - z^{-1} \Rightarrow H_I(z) = \frac{1}{1 - z^{-1}}$$

que tiene dos posibles soluciones

$$h_I(n) = u(n) \quad \text{ó} \quad h_I(n) = -u(-n - 1)$$

4.10

Otra manera más “algorítmica” de obtener la respuesta impulsional $h_I(n)$ se basa en la convolución. Si se asume que $h(n)$ y su inverso son causales, entonces:

$$h(n) * h_I(n) = \sum_{k=0}^n h(k)h_I(n - k) = \delta(n) \quad (4.8)$$

Para $n = 0$ se cumple $h(0)h_I(0) = 1 \Rightarrow h_I(0) = \frac{1}{h(0)}$, y puesto que para $n > 0$:

$$\begin{aligned} \sum_{k=0}^n h(k)h_I(n - k) &= h(0)h_I(n) + \sum_{k=1}^n h(k)h_I(n - k) = 0 \\ \Rightarrow h_I(n) &= -\frac{\sum_{k=1}^n h(k)h_I(n - k)}{h(0)} \end{aligned}$$

lo que se puede programar directamente. Si $h(0)$ es cero, entonces debe introducirse un retardo en (4.8) sustituyendo $\delta(n)$ por $\delta(n - m)$, donde para $n = m$, $h(n) \neq 0$ y $h(n) = 0$ si $n < m$. Este método tiene, sin embargo, el problema de que conforme n crece, el error numérico acumulado crece.

Ejemplo 4.11 Determine el inverso causal del sistema de respuesta impulsional $h(n) = \delta(n) - \alpha\delta(n-1)$. Con $h(0) = 1$, $h(1) = -\alpha$ y $h(n) = 0$ para $n \geq 2$, entonces:

$$\begin{aligned} h_I(0) &= \frac{1}{h(0)} = 1 \\ h_I(1) &= -\frac{h(1)h(0)}{h(0)} = \alpha \\ h_I(2) &= -(h(1)h_I(1) + h(2)h_I(0)) = -(-\alpha\alpha) = \alpha^2 \\ h_I(3) &= -(h(1)h_I(2) + h(2)h_I(1) + h(3)h_I(0)) = \alpha^3 \\ &\vdots \\ h_I(n) &= \alpha^n \end{aligned}$$

4.11

4.6.2 Sistemas de fase mínima, fase máxima y fase mixta

Sean los sistemas

$$H_1(z) = 1 + \alpha_1 z^{-1} \quad \text{y} \quad H_2(z) = \alpha_2 + z^{-1}$$

$H_1(z)$ tiene un cero en $z = -\alpha_1$ y $H_2(z)$ en $z = -\frac{1}{\alpha_2}$. Puede demostrarse que si $\alpha_1 = \alpha_2 = \alpha$, entonces las respuestas en magnitud $|H_1(\omega)| = |H_2(\omega)|$ son ambas idénticas a $\sqrt{1 + \alpha^2 + 2\alpha \cos \omega}$. Para la fase de $H_1(\omega)$ se tiene que:

$$\begin{aligned} H_1(z) = 1 + \alpha_1 z^{-1} &= z^{-1}(z + \alpha_1) \Rightarrow \angle H_1(\omega) = \arg(e^{-j\omega}(e^{j\omega} + \alpha_1)) \\ \angle H_1(\omega) &= -\omega + \arctan\left(\frac{\text{sen}(\omega)}{\alpha_1 + \cos \omega}\right) \end{aligned}$$

y para $\angle H_2(\omega)$ se tiene que:

$$\begin{aligned} H_2(z) = \alpha_2 + z^{-1} &= z^{-1}(\alpha_2 z + 1) \Rightarrow \angle H_2(\omega) = \arg(e^{-j\omega}(\alpha_2 e^{j\omega} + 1)) \\ \angle H_2(\omega) &= -\omega + \arctan\left(\frac{\alpha_2 \text{sen}(\omega)}{1 + \alpha_2 \cos \omega}\right) = -\omega + \arctan\left(\frac{\text{sen} \omega}{\frac{1}{\alpha_2} + \cos \omega}\right) \end{aligned}$$

El término $\arctan\left(\frac{\text{sen}(\omega)}{\alpha + \cos \omega}\right)$ se comporta de la siguiente manera:

$$\arctan\left(\frac{\text{sen} \omega}{\alpha + \cos \omega}\right) \approx \begin{cases} \omega, & \text{si } \alpha \rightarrow 0 \\ \frac{\omega}{2}, & \text{si } \alpha \rightarrow 1 \\ \frac{\text{sen} \omega}{\alpha}, & \text{si } \alpha \gg 1 \end{cases}$$

La figura 4.32 muestra este último componente en función de α y ω . El lado derecho de la misma figura presenta la fase $\angle H_i(z)$, que es igual al lado izquierdo menos el aporte igual a ω del factor exponencial.

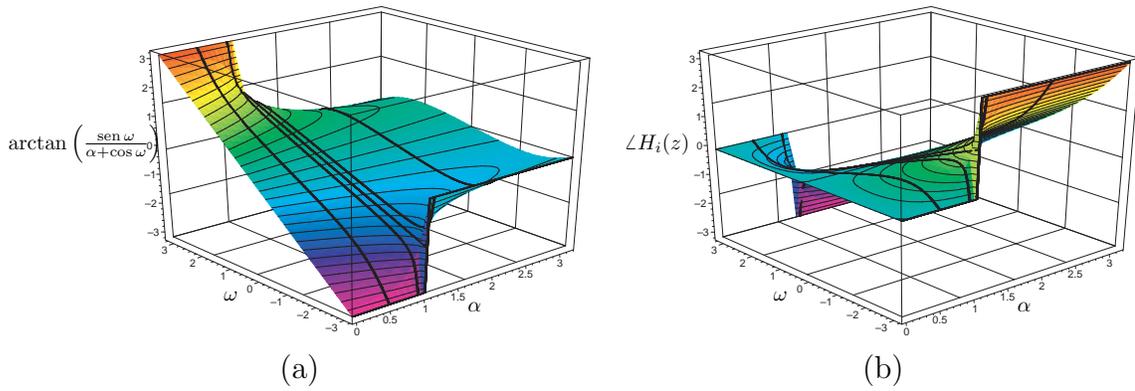


Figura 4.32: Comportamiento de la fase de $H_i(z) = 1 + \alpha z^{-1}$. (a) Componente $\arctan\left(\frac{\text{sen } \omega}{\alpha + \cos \omega}\right)$.
 (b) Fase $\angle H(\omega) = -\omega + \arctan\left(\frac{\text{sen } \omega}{\alpha + \cos \omega}\right)$

En las expresiones anteriores el denominador $\alpha + \cos \omega$ corresponde a la parte real del segundo factor de la fase ($\alpha_1 + e^{j\omega}$ o $1 + \alpha_2 e^{j\omega}$). Si se asumen ahora frecuencias cercanas pero menores a $\omega = \pi$ entonces $\text{sen } \omega / (\alpha + \cos \omega)$ tiende a cero y puesto que para $x \approx 0$ se cumple $\arctan x \approx x$ entonces para $\alpha > 1$ se cumple que el aporte en fase de este término aproximadamente cero. Si la parte real es negativa, es decir, si $\alpha < 1$, entonces el término aportará una fase cercana a π .

Para un valor constante de α con $\alpha < 1$, se observa entonces que la fase inicia con cero en $\omega = -\pi$, luego oscila cruzando una sola vez cero, para nuevamente alcanzarlo en $\omega = \pi$. Por otro lado, si $\alpha > 1$ entonces la fase inicia con un valor de π en $\omega = -\pi$ y baja monótonicamente hasta alcanzar $-\pi$ en $\omega = \pi$. Esto quiere decir que el cambio neto de fase $\theta_1(\pi) - \theta_1(0)$ para un α particular es cero si $\alpha_1 < 1$, mientras que es π si $\alpha_1 > 1$. El primer caso se conoce entonces como *sistema de fase mínima* y el segundo como *sistema de fase máxima*.

Un sistema FIR de longitud $M + 1$ tiene M ceros y su respuesta en frecuencia se puede expresar como:

$$H(\omega) = b_0(1 - z_1 e^{-j\omega})(1 - z_2 e^{-j\omega}) \dots (1 - z_M e^{-j\omega})$$

Si todos los ceros se encuentran dentro de la circunferencia unitaria, el cambio de fase neto será entonces también de cero entre $\omega = 0$ y $\omega = \pi$, por lo que al sistema se le denomina de fase mínima. Si todos los ceros se encuentran fuera de la circunferencia unidad, el cambio neto de fase será $\angle H(\pi) - \angle H(0) = M\pi$, y el sistema se denomina de fase máxima.

Un sistema es entonces de fase mixta si algunos de los ceros están dentro de la circunferencia unitaria y otros fuera, y si el sistema tiene en total M ceros se cumple que el cambio de fase neto será menor a $M\pi$ radianes, entre $\omega = 0$ y $\omega = \pi$.

Puesto que la derivada de la función de fase representa en cierta medida el retraso de cada componente frecuencial, un sistema de fase mínima implica una función de retardo mínima, mientras que un sistema de fase máxima implica un retardo máximo.

Un sistema IIR descrito por la función de transferencia racional:

$$H(z) = \frac{B(z)}{A(z)}$$

se denomina de fase mínima si todos sus polos y ceros se encuentran dentro de la circunferencia unitaria, lo que implica además que el sistema es estable y causal. Si en un sistema causal y estable todos los ceros están fuera de $|z| = 1$, entonces el sistema es de fase máxima, y si sólo algunos de sus ceros están fuera de $|z| = 1$, entonces el sistema es de fase mixta.

Puesto que el inverso de $H(z)$ es $H^{-1}(z) = \frac{A(z)}{B(z)}$, entonces se deduce de lo anterior que el inverso de un sistema de fase mínima estable es también un sistema de fase mínima estable. En general, la fase mínima de $H(z)$ asegura la estabilidad de su sistema inverso $H^{-1}(z)$, y la estabilidad de $H(z)$ asegura la fase mínima de $H^{-1}(z)$. Los sistemas de fase mixta o máxima tienen entonces sistemas inversos inestables.

Descomposición de sistemas de polos y ceros de fase no mínima

Cualquier sistema de función de transferencia racional y de fase no mínima puede descomponerse en la cascada de un sistema de fase mínima y un filtro pasa todos.

$$H(z) = H_{min}(z)H_{ap}(z)$$

Sea $H(z) = \frac{B(z)}{A(z)}$. El numerador $B(z)$ se puede factorizar como $B(z) = B_1(z)B_2(z)$, donde $B_1(z)$ tiene todas sus raíces dentro de $|z| = 1$ y $B_2(z)$ todas sus raíces fuera. Esto a su vez implica que $B_2(z^{-1})$ tiene todas sus raíces dentro de $|z| = 1$. De esta forma:

$$H_{min}(z) = \frac{B_1(z)B_2(z^{-1})}{A(z)}, \quad H_{ap}(z) = \frac{B_2(z)}{B_2(z^{-1})}$$

Nótese que $H_{ap}(z)$ es estable puesto que todos los polos de $B_2(z^{-1})$ se encuentran dentro de la circunferencia unitaria. Además, es un filtro paso todo como los descritos anteriormente, y es de fase máxima al estar todos sus ceros fuera de la circunferencia unitaria.

Este sistema $H(z)$ de fase no mínima tiene el retardo de grupo

$$\tau_g(\omega) = \tau_g^{min}(\omega) + \tau_g^{ap}(\omega)$$

Puesto que $\tau_g^{ap}(\omega) \geq 0$ para $0 \leq \omega \leq \pi$, entonces se puede concluir que entre todos los sistemas con igual respuesta en magnitud, el menor retardo de grupo lo tiene el sistema de fase mínima.

4.6.3 Identificación de sistemas

Supóngase que se observa una secuencia de salida $y(n)$ cuando se excita a un sistema de respuesta desconocida con la entrada $x(n)$. Se desea ahora determinar la respuesta impulsional del sistema.

Si existen formas cerradas en el dominio z para $x(n)$ y $y(n)$, entonces $h(n)$ se obtiene fácilmente con

$$h(n) \circ \bullet H(z) = \frac{Y(z)}{X(z)}$$

Otro método trata directamente la convolución en el dominio del tiempo, asumiendo que $h(n)$ es causal:

$$\begin{aligned} y(n) &= \sum_{k=0}^n h(k)x(n-k), n \geq 0 \\ \Rightarrow y(0) &= h(0)x(0) \Rightarrow h(0) = \frac{y(0)}{x(0)} \\ y(n) &= \sum_{k=0}^{n-1} h(k)x(n-k) + h(n)x(0) \\ \Rightarrow h(n) &= \frac{y(n) - \sum_{k=0}^{n-1} h(k)x(n-k)}{x(0)}, n \geq 1 \end{aligned} \quad (4.9)$$

Aquí se ha asumido que $x(0) \neq 0$. El problema se torna numéricamente inestable para $n \gg 1$.

Un tercer método utiliza la correlación cruzada entrada-salida:

$$\begin{aligned} r_{yx}(n) &= y(n) * x(-n) = h(n) * x(n) * x(-n) = h(n) * r_{xx}(n) \\ &= \sum_{k=0}^{\infty} h(k)r_{xx}(n-k) \\ &\quad \circ \bullet \\ S_{yx}(\omega) &= H(\omega)S_{xx}(\omega) = H(\omega)|X(\omega)|^2 \\ \Rightarrow H(\omega) &= \frac{S_{yx}(\omega)}{S_{xx}(\omega)} = \frac{S_{yx}(\omega)}{|X(\omega)|^2} \end{aligned}$$

Se puede aplicar entonces la fórmula de deconvolución (4.9) para determinar $h(n)$ o la transformada inversa de Fourier de $H(\omega)$. Este método tiene la ventaja de que el uso de las correlaciones puede eliminar términos de ruido en la identificación, que podrían alterar los resultados de los otros métodos.

4.7 Transformada Discreta de Fourier

En las secciones anteriores se observó que señales discretas aperiódicas, como señales digitales de audio, señales digitales sísmicas, señales médicas, etc., tienen un espectro periódico pero continuo, que no tiene una representación directa para ser manejado por medios digitales. En esta sección se estudia la transformada discreta de Fourier (DFT, *Discrete Fourier Transform*) que representa un mecanismo para el estudio frecuencial por medios digitales para las mencionadas señales.

4.7.1 Muestreo en el dominio de la frecuencia

Toda señal aperiódica $x(n)$ de energía finita tiene un espectro continuo

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

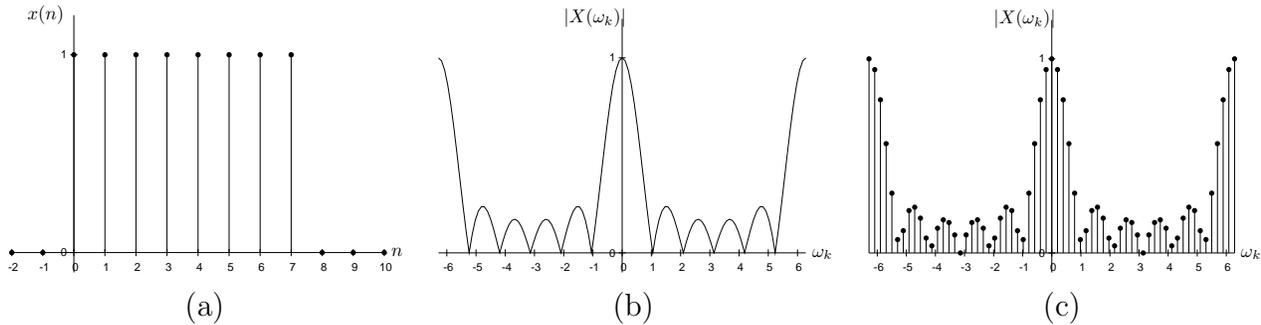


Figura 4.33: Señal discreta aperiódica, su espectro continuo, y su espectro muestreado.

Si $X(\omega)$ se muestrea periódicamente (figura 4.33) con un espaciado frecuencial $\delta\omega = \frac{2\pi}{N}$ solo se necesitarán N muestras puesto que el espectro, por corresponder a una señal discreta, es a su vez periódico con periodo 2π . Las muestras en $\omega = k\delta\omega = \frac{2\pi}{N}k$ son entonces

$$X\left(\frac{2\pi}{N}k\right) = \sum_{n=-\infty}^{\infty} x(n)e^{-j2\pi kn/N} \quad k = 0, 1, \dots, N-1$$

donde descomponiendo la sumatoria en subintervalos de N elementos y haciendo un cambio de variable para trasladar cada subintervalo al origen se modifica en

$$X\left(\frac{2\pi}{N}k\right) = \sum_{l=-\infty}^{\infty} \sum_{n=lN}^{lN+N-1} x(n)e^{-j2\pi kn/N} = \sum_{n=0}^{N-1} \underbrace{\left[\sum_{l=-\infty}^{\infty} x(n-lN) \right]}_{x_p(n)} e^{-j2\pi kn/N}$$

con $k = 0, 1, \dots, N-1$.

La señal

$$x_p(n) = \sum_{l=-\infty}^{\infty} x(n-lN)$$

se obtiene repitiendo periódicamente $x(n)$ cada N muestras, y por ser periódica puede calcularse su serie de Fourier como

$$x_p(n) = \sum_{k=0}^{N-1} c_k e^{j2\pi kn/N}, \quad n = 0, 1, \dots, N-1$$

con los coeficientes

$$c_k = \frac{1}{N} \sum_{n=0}^{N-1} x_p(n) e^{-j2\pi kn/N} = \frac{1}{N} X\left(\frac{2\pi}{N}k\right)$$

Con esta serie, $x_p(n)$ puede ser reconstruido del espectro muestreado con

$$x_p(n) = \frac{1}{N} \sum_{k=0}^{N-1} X \left(\frac{2\pi}{N} k \right) e^{j2\pi kn/N}, \quad n = 0, 1, \dots, N-1$$

La señal $x(n)$ puede recuperarse de $x_p(n)$ si y solo si no hay *aliasing* en el dominio del tiempo, es decir, si la longitud L de $x(n)$ es menor que el periodo N de $x_p(n)$ (figura 4.34):

$$x(n) = \begin{cases} x_p(n), & 0 \leq n \leq N-1 \\ 0, & \text{en el resto} \end{cases}$$

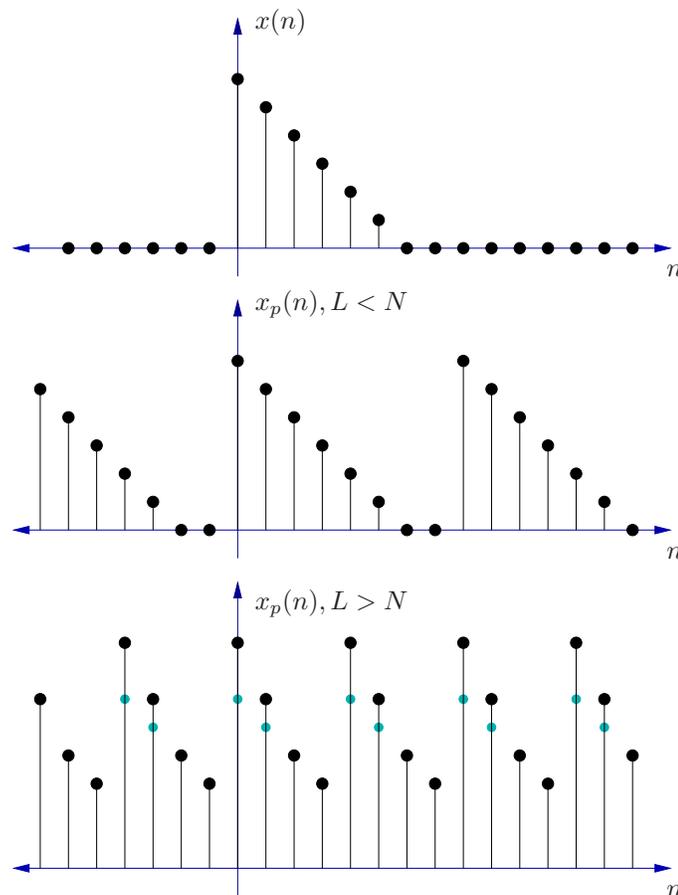


Figura 4.34: Extensión periódica por muestreo del espectro. a) Señal original, b) Extensión con $L < N$, c) Extensión con $L > N$.

Si no hay aliasing en el tiempo ($L < N$) entonces

$$x(n) = x_p(n) = \frac{1}{N} \sum_{k=0}^{N-1} X \left(\frac{2\pi}{N} k \right) e^{j2\pi kn/N}, \quad 0 \leq n \leq N-1$$

con espectro

$$\begin{aligned}
 X(\omega) &= \sum_{n=0}^{N-1} x(n)e^{-j\omega n} = \sum_{n=0}^{N-1} \left[\frac{1}{N} \sum_{k=0}^{N-1} X\left(\frac{2\pi}{N}k\right) e^{j2\pi kn/N} \right] e^{-j\omega n} \\
 &= \sum_{k=0}^{N-1} X\left(\frac{2\pi}{N}k\right) \left[\frac{1}{N} \sum_{n=0}^{N-1} e^{-j\left(\omega - \frac{2\pi k}{N}\right)n} \right] \\
 &= \sum_{k=0}^{N-1} X\left(\frac{2\pi}{N}k\right) P\left(\omega - \frac{2\pi k}{N}\right)
 \end{aligned}$$

donde

$$\begin{aligned}
 P(\omega) &= \frac{1}{N} \sum_{n=0}^{N-1} e^{-j\omega n} = \frac{1}{N} \frac{1 - e^{-j\omega N}}{1 - e^{-j\omega}} \\
 &= \frac{1}{N} \frac{\text{sen}(\omega N/2)}{\text{sen}(\omega/2)} e^{-j\omega(N-1)/2}
 \end{aligned} \tag{4.10}$$

es la función de interpolación que sustituye en este caso a $\text{sen}(\theta)/\theta$ con una versión de propiedades similares pero periódica (figura 4.35):

$$P\left(\frac{2\pi}{N}k\right) = \begin{cases} 1 & k = 0 \\ 0 & k = 1, 2, \dots, N-1 \end{cases}$$

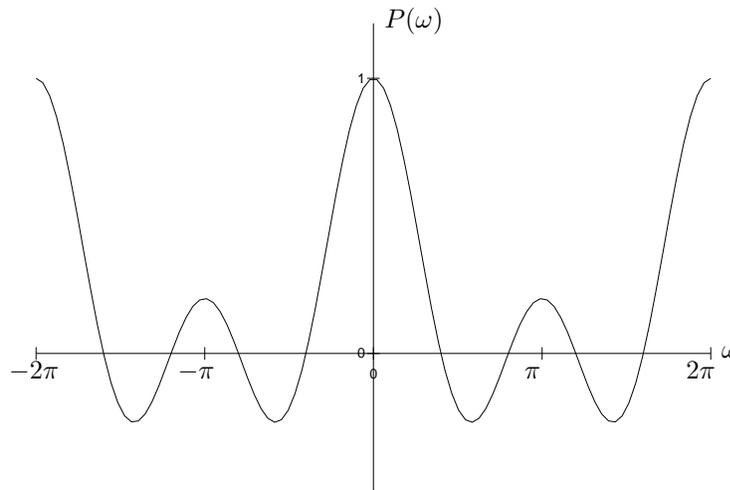


Figura 4.35: Interpolador ideal $P(\omega)$ para espectro muestreado con $N = 5$.

Ejemplo 4.12 Determine el aliasing de la secuencia $x(n) = a^n u(n)$, $|a| < 1$, si el espectro se muestrea a las frecuencias $\omega_k = 2\pi k/N$.

El espectro es

$$X(\omega) = \sum_{n=0}^{\infty} a^n e^{-j\omega n} = \frac{1}{1 - ae^{-j\omega}}$$

Las muestras están dadas por $X(\omega_k) = X\left(\frac{2\pi}{N}k\right) = \frac{1}{1 - ae^{-j2\pi k/N}}$

La secuencia periódica $x_p(n)$ representada por este espectro discreto es

$$x_p(n) = \sum_{l=-\infty}^{\infty} x(n - lN) = \sum_{l=-\infty}^0 a^{n-lN} = a^n \sum_{l=0}^{\infty} a^{lN} = a^n \frac{1}{1 - a^N}, \quad 0 \leq n \leq N - 1$$

donde la constante $\frac{1}{1 - a^N}$ representa el aliasing y tiende a uno si $N \rightarrow \infty$, como es de esperar.

4.12

4.7.2 La transformada discreta de Fourier

En la sección anterior se demostró que el espectro muestreado de una señal $x(n)$ de longitud finita L no representa a $x(n)$ directamente sino a su extensión periódica $x_p(n)$ obtenida de

$$x_p(n) = \sum_{l=-\infty}^{\infty} x(n - lN)$$

Si no hay aliasing temporal, entonces $x_p(n)$ es simplemente la concatenación periódica de $x(n)$ de longitud L , que se expande con ceros para alcanzar la longitud $N > L$:

$$x_p(n - lN) = \begin{cases} x(n) & 0 \leq n \leq L - 1 \\ 0 & L \leq n \leq N - 1 \end{cases}$$

Nótese que el efecto de rellenar con ceros $x(n)$ produce un muestreo más detallado de su espectro $X(\omega)$ que tendrá N y no L muestras. Para todo $N > L$ las muestras no proveen mayor información que el caso $N = L$, solo una “mejor” representación (figura 4.36).

La transformación de $x(n)$ hacia su espectro muestreado

$$\begin{aligned} X(k) &= X\left(\frac{2\pi}{N}k\right) = \sum_{n=0}^{L-1} x(n)e^{-j2\pi kn/N} \\ &= \sum_{n=0}^{N-1} x(n)e^{-j2\pi kn/N}, \quad k = 0, 1, \dots, N - 1 \end{aligned}$$

se conoce como transformada discreta de Fourier DFT y a la reconstrucción de $x(n)$ de las muestras espectrales

$$x(n) = \frac{1}{N} \sum_{k=0}^{N-1} X(k)e^{j2\pi kn/N}, \quad n = 0, 1, \dots, N - 1$$

se le conoce como transformada discreta de Fourier inversa (IDFT).

Estas transformaciones se expresan a menudo como

$$\begin{aligned} \text{DFT: } X(k) &= \sum_{n=0}^{N-1} x(n)W_N^{kn} & k = 0, 1, \dots, N - 1 \\ \text{IDFT: } x(n) &= \frac{1}{N} \sum_{k=0}^{N-1} X(k)W_N^{-kn} & n = 0, 1, \dots, N - 1 \end{aligned}$$

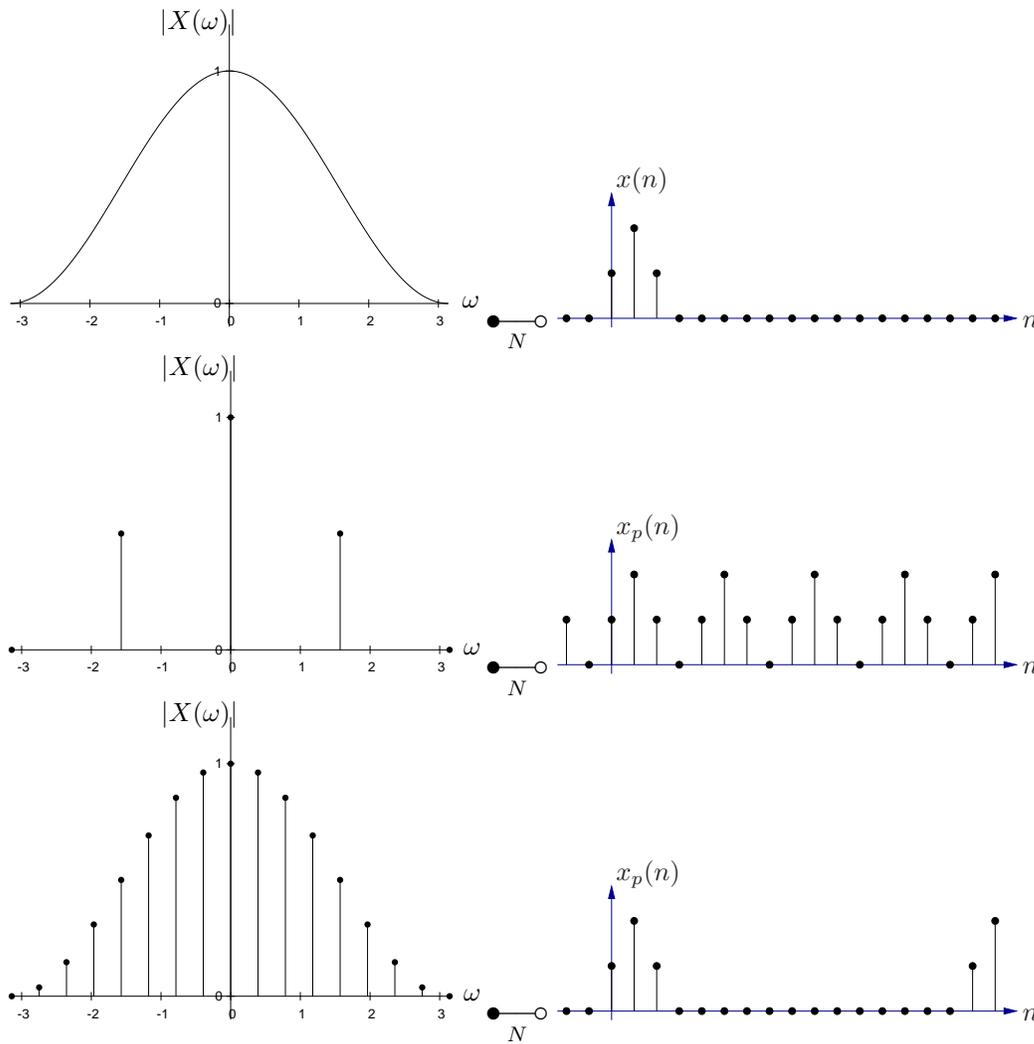


Figura 4.36: Efecto de extensión de señal con ceros y la DFT resultante.

donde $W_N = e^{-j2\pi/N}$ es una raíz N -ésima de la unidad. Nótese que para cada k , $X(k)$ se puede expresar como el producto punto de la entrada $x(n)$ y un vector que consiste en $[1, W_N^k, W_N^{2k}, \dots, W_N^{(N-1)k}]$, o, si se interpreta $X(k)$ también como vector, entonces

$$X(k) \equiv \underline{\mathbf{X}}_N = \begin{bmatrix} X(0) \\ X(1) \\ \vdots \\ X(N-1) \end{bmatrix} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & W_N & \cdots & W_N^{N-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & W_N^{N-1} & \cdots & W_N^{(N-1)(N-1)} \end{bmatrix} \begin{bmatrix} x(0) \\ x(1) \\ \vdots \\ x(N-1) \end{bmatrix}$$

o en forma compacta

$$\underline{\mathbf{X}}_N = \mathbf{W}_N \underline{\mathbf{x}}_N$$

de donde se deduce

$$\underline{\mathbf{x}}_N = \mathbf{W}_N^{-1} \underline{\mathbf{X}}_N = \frac{1}{N} \mathbf{W}_N^* \underline{\mathbf{X}}_N$$

4.7.3 Relación de la DFT con otras transformadas

Series de Fourier

La secuencia $x_p(n)$ por ser periódica con periodo N tiene una representación en series de Fourier

$$x_p(n) = \sum_{k=0}^{N-1} c_k e^{j2\pi nk/N}, \quad -\infty < n < \infty$$

$$c_k = \frac{1}{N} \sum_{n=0}^{N-1} x_p(n) e^{-j2\pi nk/N} \quad k = 0, 1, \dots, N-1$$

Esto es $X(k) = Nc_k$.

Transformada de Fourier de secuencias aperiódicas

Ya se verificó la relación entre la DFT $X(k)$ como muestras del espectro continuo $X(\omega)$

$$X(k) = X(\omega)|_{\omega=\frac{2\pi}{N}k}$$

que a su vez corresponden con los coeficientes de la DFT de $x_p(n)$, la extensión periódica de $x(n)$. Ambas secuencias son idénticas para $n = 0, 1, \dots, L-1$ si no hay aliasing temporal, es decir, si $x(n)$ es finita de longitud L y $N > L$.

Transformada z

Si la región de convergencia de la transformada z de $x(n)$ incluye la circunferencia unitaria, entonces

$$X(k) = X(z)|_{z=e^{j2\pi k/N}} = \sum_{n=-\infty}^{\infty} x(n) e^{-j2\pi nk/N}, \quad k = 0, 1, \dots, N-1$$

que representa al muestreo de la transformada z sobre el círculo unitario con ángulos distribuidos homogéneamente. Sin aliasing temporal se cumple entonces, con $x(n)$ de longitud finita N

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

$$= \frac{1}{N} \sum_{k=0}^{N-1} X(k) \sum_{n=0}^{N-1} (e^{j2\pi k/N} z^{-1})^n$$

$$X(z) = \frac{1 - z^{-N}}{N} \sum_{k=0}^{N-1} \frac{X(k)}{1 - e^{j2\pi k/N} z^{-1}}$$

que es la reconstrucción de $X(z)$ a partir de las muestras $X(k)$ de la DFT. Evaluando en la circunferencia unitaria se obtiene para la transformada de Fourier

$$X(\omega) = \frac{1 - e^{-j\omega N}}{N} \sum_{k=0}^{N-1} \frac{X(k)}{1 - e^{-j(\omega - \frac{2\pi k}{N})}}$$

que es otra forma de interpolación de las muestras $X(k)$ equivalente a la expresada anteriormente con $P(\omega)$ (ver ecuación 4.10 en la página 138).

4.7.4 Propiedades de la DFT

Sea $x(n)$ una secuencia de longitud $L < N$ y $X(k)$ su correspondiente DFT de longitud N . Se cumple que

$$\begin{aligned} \text{DFT: } X(k) &= \sum_{n=0}^{N-1} x(n) W_N^{kn} & k = 0, 1, \dots, N-1 \\ \text{IDFT: } x(n) &= \frac{1}{N} \sum_{k=0}^{N-1} X(k) W_N^{-kn} & n = 0, 1, \dots, N-1 \end{aligned}$$

con $W_N = e^{-j2\pi/N}$.

La relación entre $x(n)$ y su DFT $X(k)$ se denota con

$$x(n) \xleftrightarrow[N]{DFT} X(k) \quad \text{ó} \quad x(n) \circ\text{---}\bullet_N X(k)$$

Linealidad

Si $x_1(n) \circ\text{---}\bullet_N X_1(k)$ y $x_2(n) \circ\text{---}\bullet_N X_2(k)$ entonces

$$a_1 x_1(n) + a_2 x_2(n) \circ\text{---}\bullet_N a_1 X_1(k) + a_2 X_2(k)$$

Simetría circular

Puesto que la DFT de N puntos de una secuencia finita $x(n)$ de longitud $L \leq N$ es equivalente a la DFT de N puntos de una secuencia periódica $x_p(n)$ de periodo N obtenida como (figura 4.37)

$$x_p(n) = \sum_{l=-\infty}^{\infty} x(n - lN)$$

entonces un desplazamiento de $x_p(n)$ k unidades hacia la derecha equivale a un desplazamiento circular (figura 4.38).

$$x_p(n - k) \circ\text{---}\bullet_N \cdot \bullet_N \circ\text{---}\bullet_N x(n - k \pmod N) \equiv x((n - k))_N$$

Una secuencia es *circularmente par* si es simétrica con respecto al punto cero de la circunferencia (figura 4.39):

$$x(N - n) = x(n), \quad 1 \leq n \leq N - 1$$

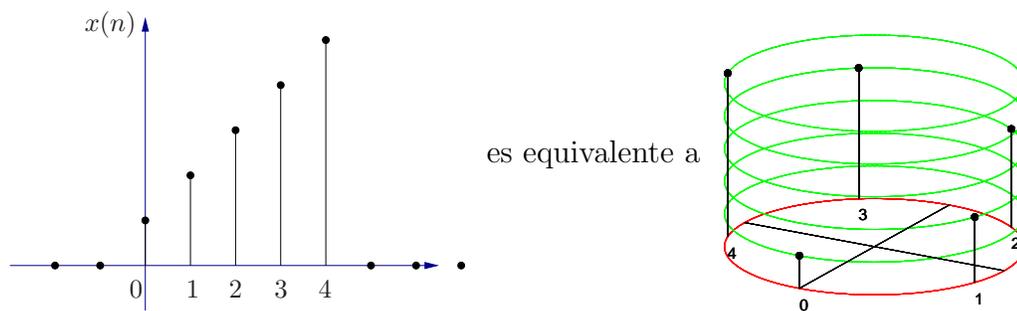


Figura 4.37: Secuencia circular equivalente a $x(n)$.

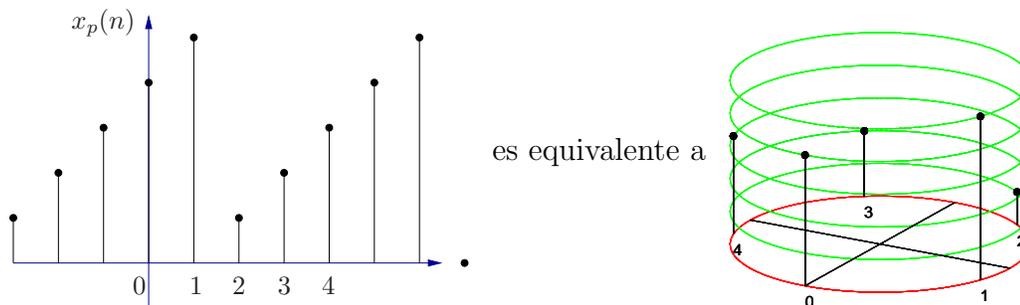


Figura 4.38: Secuencia circular equivalente a $x_p(n - 2)$.

La secuencia es *circularmente impar* si es antisimétrica con respecto al punto cero (figura 4.40):

$$x(N - n) = -x(n), \quad 1 \leq n \leq N - 1$$

La reflexión temporal circular se obtiene con

$$x((-n))_N = x(N - n), \quad 0 \leq n \leq N - 1$$

Utilizando $x_p(n)$ en vez de $x(n)$ se tiene

par: $x_p(n) = x_p(-n) = x_p(N - n)$

impar: $x_p(n) = -x_p(-n) = -x_p(N - n)$

hermítica: $x_p(n) = x_p^*(N - n)$ (conjugada par)

anti-hermítica: $x_p(n) = -x_p^*(N - n)$ (conjugada impar)

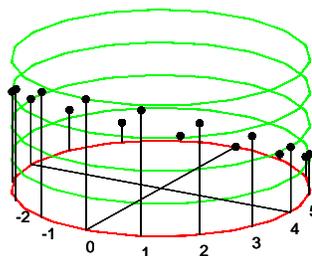


Figura 4.39: Simetría circular par

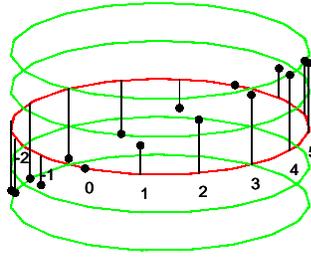


Figura 4.40: Simetría circular impar

La secuencia se puede descomponer en

$$\begin{aligned}x_p(n) &= x_{pe}(n) + x_{po}(n) \\x_{pe}(n) &= \frac{1}{2} [x_p(n) + x_p^*(N-n)] \\x_{po}(n) &= \frac{1}{2} [x_p(n) - x_p^*(N-n)]\end{aligned}$$

Si $x(n) = x_R(n) + jx_I(n)$, $0 \leq n \leq N-1$, y $X(k) = X_R(k) + jX_I(k)$ entonces se obtiene con la definición de la DFT.

$$\begin{aligned}X_R(k) &= \sum_{n=0}^{N-1} \left[x_R(n) \cos\left(\frac{2\pi kn}{N}\right) + x_I(n) \sen\left(\frac{2\pi kn}{N}\right) \right] \\X_I(k) &= -\sum_{n=0}^{N-1} \left[x_R(n) \sen\left(\frac{2\pi kn}{N}\right) - x_I(n) \cos\left(\frac{2\pi kn}{N}\right) \right] \\x_R(n) &= \frac{1}{N} \sum_{k=0}^{N-1} \left[X_R(k) \cos\left(\frac{2\pi kn}{N}\right) - X_I(k) \sen\left(\frac{2\pi kn}{N}\right) \right] \\x_I(n) &= \frac{1}{N} \sum_{k=0}^{N-1} \left[X_R(k) \sen\left(\frac{2\pi kn}{N}\right) + X_I(k) \cos\left(\frac{2\pi kn}{N}\right) \right]\end{aligned}$$

Si $x(n)$ es real, entonces

$$X(N-k) = X^*(k) = X(-k) \Rightarrow |X(N-k)| = |X(k)|, \angle X(N-k) = -\angle X(k)$$

Si $x(n)$ es real y par entonces $x(n) = x(N-n)$, para $0 \leq n \leq N-1$, y la DFT se torna

$$X(k) = \sum_{n=0}^{N-1} x(n) \cos\left(\frac{2\pi kn}{N}\right), 0 \leq k \leq N-1$$

que también es real y par. Puesto que $X_I(k) = 0$ entonces

$$x(n) = \frac{1}{N} \sum_{k=0}^{N-1} X(k) \cos\left(\frac{2\pi kn}{N}\right), 0 \leq n \leq N-1$$

Si $x(n)$ es real e impar entonces

$$x(n) = -x(N - n), \quad 0 \leq n \leq N - 1$$

$$X(k) = -j \sum_{n=0}^{N-1} x(n) \operatorname{sen} \left(\frac{2\pi kn}{N} \right), \quad 0 \leq k \leq N - 1$$

y puesto que $X_R(k) = 0$ entonces

$$x(n) = \frac{j}{N} \sum_{k=0}^{N-1} X(k) \operatorname{sen} \left(\frac{2\pi kn}{N} \right), \quad 0 \leq n \leq N - 1$$

Convolución circular

Sean $x_1(n)$ y $x_2(n)$ dos secuencias de duración finita, ambas de longitud N , y sus respectivas DFT de N puntos

$$X_1(k) = \sum_{n=0}^{N-1} x_1(n) e^{-j2\pi nk/N}, \quad k = 0, 1, \dots, N - 1$$

$$X_2(k) = \sum_{n=0}^{N-1} x_2(n) e^{-j2\pi nk/N}, \quad k = 0, 1, \dots, N - 1$$

Si $X_3 = X_1(k)X_2(k)$, $k = 0, 1, \dots, N - 1$ entonces

$$\begin{aligned} x_3(m) &= \frac{1}{N} \sum_{k=0}^{N-1} X_3(k) e^{j2\pi km/N} \\ &= \frac{1}{N} \sum_{k=0}^{N-1} X_1(k) X_2(k) e^{j2\pi km/N} \\ &= \frac{1}{N} \sum_{k=0}^{N-1} \left[\sum_{n=0}^{N-1} x_1(n) e^{-j2\pi kn/N} \right] \left[\sum_{l=0}^{N-1} x_2(l) e^{-j2\pi kl/N} \right] e^{j2\pi km/N} \\ &= \frac{1}{N} \sum_{n=0}^{N-1} x_1(n) \sum_{l=0}^{N-1} x_2(l) \left[\sum_{k=0}^{N-1} e^{j2\pi k(m-n-l)/N} \right] \end{aligned}$$

y con $a = e^{j2\pi(m-n-l)/N}$ en

$$\sum_{k=0}^{N-1} a^k = \begin{cases} N & \text{si } a = 1 \Leftrightarrow m - n - l = pN \Rightarrow l = ((m - n))_N \\ \frac{1-a^N}{1-a} = 0 & \text{si } a \neq 1, \text{ puesto que } a^N = 1 \end{cases}$$

por lo que

$$x_3(m) = \sum_{n=0}^{N-1} x_1(n) x_2((m - n))_N, \quad m = 0, 1, \dots, N - 1 = x_1(m) \circledast x_2(m)$$

operación denominada convolución circular. Nótese que esta función es para los elementos $x_3(m)$ con $m = 0, 1, \dots, N-1$ equivalente a la convolución de $x_1(n)$ con la extensión periódica con periodo N de $x_2(n)$.

Otras propiedades se resumen en la tabla 4.3.

Tabla 4.3: Propiedades de la DFT

Propiedad	Dominio temporal	Dominio frecuencial
Notación	$x(n), y(n)$	$X(k), Y(k)$
Periodicidad	$x(n) = x(n + N)$	$X(k) = X(k + N)$
Linealidad	$a_1x_1(n) + a_2x_2(n)$	$a_1X_1(k) + a_2X_2(k)$
Reflexión temporal	$x(N - n)$	$X(N - k)$
Desplazamiento temporal circular	$x((n - l))_N$	$X(k) e^{-j2\pi kl/N}$
Desplazamiento frecuencial circular	$x(n)e^{j2\pi ln/N}$	$X((k - l))_N$
Conjugación compleja	$x^*(n)$	$X^*(N - k)$
Convolución circular	$x_1(n) \otimes x_2(n)$	$X_1(k)X_2(k)$
Correlación circular	$x(n) \otimes y^*(-n)$	$X(k)Y^*(k)$
Multiplicación de dos secuencias	$x_1(n)x_2(n)$	$\frac{1}{N}X_1(k) \otimes X_2(k)$
Teorema de Parseval	$\sum_{n=0}^{N-1} x(n)y^*(n)$	$\frac{1}{N} \sum_{k=0}^{N-1} X(k)Y^*(k)$

4.7.5 Filtrado lineal basado en la DFT

La existencia de algoritmos eficientes para extraer la DFT (llamados FFT del inglés *Fast Fourier Transform*) hace posible que sea en ocasiones más eficiente calcular el efecto de un filtro en el dominio de la frecuencia que directamente en el dominio del tiempo.

Sea la entrada $x(n)$ de longitud L y un filtro FIR con respuesta impulsional $h(n)$ de longitud M . Asíumase además que ambas secuencias son causales, es decir

$$\begin{aligned} x(n) &= 0, & n < 0, n \geq L \\ h(n) &= 0, & n < 0, n \geq M \end{aligned}$$

La salida del filtro puede calcularse en el dominio del tiempo con la convolución

$$y(n) = \sum_{k=0}^{M-1} h(k)x(n - k) = h(n) * x(n)$$

que por la naturaleza finita de $x(n)$ y $h(n)$ también es finita. En el ejemplo 2.12 (pág. 37) se determinó que la salida tendrá longitud $M + L - 1$.

En el dominio de la frecuencia la salida es

$$Y(\omega) = X(\omega)H(\omega)$$

Si se desea representar $y(n)$ a través de muestras de $Y(\omega)$ se necesitan entonces al menos $N \geq M + L - 1$ muestras para evitar así el aliasing temporal y se obtiene con la DFT:

$$Y(k) = Y(\omega)|_{\omega=2\pi k/N} = X(\omega)H(\omega)|_{\omega=2\pi k/N} = X(k)H(k)$$

donde $X(k)$ y $H(k)$ representan entonces la DFT de las secuencias $x(n)$ y $h(n)$ que han sido extendidas con ceros para alcanzar la longitud N .

Nótese que la compensación de longitud de $x(n)$ y $y(n)$ logra que la convolución circular sea equivalente a la convolución lineal.

Ejemplo 4.13 Calcule la respuesta del filtro con respuesta impulsional $h(n) = \{\frac{1}{4}, \frac{1}{2}, \frac{1}{4}\}$ ante la entrada $x(n) = \{4, 2, 2, 4\}$.

La longitud de $x(n)$ es $L = 4$ y la de $h(n)$ es $M = 3$, por lo que la salida necesita al menos $L + M - 1 = 6$ muestras. Se utilizará $N = 8$ por simplicidad y porque los algoritmos de FFT usualmente requieren $N = 2^k$ muestras, con $k \in \mathbb{N}$. Así

$$\begin{aligned} X(k) &= \sum_{n=0}^7 x(n)e^{j2\pi kn/8} \\ &= x(0) + x(1)e^{-j\frac{\pi}{4}k} + x(2)e^{-j\frac{\pi}{2}k} + x(3)e^{-j\frac{3\pi}{4}k} \\ &= 4(1 + e^{-j\frac{3\pi}{4}k}) + 2(e^{-j\frac{\pi}{4}k} + e^{-j\frac{\pi}{2}k}) \end{aligned}$$

con lo que se obtiene

$$\begin{aligned} X(0) &= 12 \\ X(1) &= 4 - \sqrt{2} - j(2 + 3\sqrt{2}) \\ X(2) &= 2 + j2 \\ X(3) &= 4 + \sqrt{2} + j(2 - 3\sqrt{2}) \\ X(4) &= 0 \\ X(5) &= 4 + \sqrt{2} - j(2 - 3\sqrt{2}) = X^*(3) \\ X(6) &= 2 - j2 = X^*(2) \\ X(7) &= 4 - \sqrt{2} + j(2 + 3\sqrt{2}) = X^*(1) \end{aligned}$$

De forma similar

$$\begin{aligned} H(k) &= \sum_{n=0}^7 h(n)e^{-j2\pi kn/8} \\ &= h(0) + h(1)e^{-j\frac{\pi}{4}k} + h(2)e^{-j\frac{\pi}{2}k} \\ &= \frac{1}{4} (1 + e^{-j\frac{\pi}{2}k}) + \frac{1}{2}e^{-j\frac{\pi}{4}k} \end{aligned}$$

con lo que se obtiene

$$\begin{aligned} H(0) &= 1 \\ H(1) &= \frac{1 + \sqrt{2}}{4} - j \frac{1 + \sqrt{2}}{4} = H^*(7) \\ H(2) &= -\frac{j}{2} = H^*(6) \\ H(3) &= \frac{1 - \sqrt{2}}{4} + j \frac{1 - \sqrt{2}}{4} = H^*(5) \\ H(4) &= 0 \end{aligned}$$

El producto de ambas secuencias es

$$\begin{aligned} Y(0) &= 12 \\ Y(1) &= -\frac{3 + \sqrt{2}}{2} - j \frac{5 + 2\sqrt{2}}{2} = Y^*(7) \\ Y(2) &= 1 - j = Y^*(6) \\ Y(3) &= \frac{\sqrt{2} - 3}{2} + j \frac{5 - 2\sqrt{2}}{2} = Y^*(5) \\ Y(4) &= 0 \end{aligned}$$

Y con la IDFT se obtiene finalmente

$$\begin{aligned} y(n) &= \sum_{k=0}^7 Y(k) e^{j2\pi kn/8}, \quad n = 0, 1, \dots, 7 \\ &= \left\{ 1, \frac{5}{2}, \frac{5}{2}, \frac{5}{2}, \frac{5}{2}, 1, 0, 0 \right\} \end{aligned}$$

4.13

4.7.6 Filtrado de secuencias de larga duración

En aplicaciones reales, la secuencia de entrada puede ser muy larga o incluso de longitud impredecible, por lo que el tratamiento, tanto por limitantes de espacio de memoria como por limitaciones en el retardo de la respuesta, debe realizarse separando en bloques esa señal de entrada.

Aquí se analizarán dos métodos que dividen la señal de entrada en bloques de longitud L y aplican por medio de la DFT el filtro $h(n)$ de longitud M en el dominio de la frecuencia. La salida, de forma equivalente a la aplicación del filtro al bloque completo de la entrada, se obtiene por medio de la concatenación adecuada de cada bloque obtenido para cada bloque de entrada. Sin pérdida de generalidad se asume la longitud de la entrada mucho mayor que la del filtro ($L \gg M$).

Método de solapamiento y almacenamiento

En el método de solapamiento y almacenamiento (*overlap-save*) la entrada se separa en bloques de $N = L + M - 1$ muestras donde cada bloque contiene $M - 1$ muestras del bloque de entrada anterior seguidas por L muestras nuevas. La respuesta impulsional del filtro se aumenta con $L - 1$ ceros para alcanzar así la longitud N . La longitud de las DFT e IDFT es entonces N , y la salida para cada bloque se calcula a través de su producto. Puesto que el bloque de entrada tiene longitud N y el filtro longitud M , en el tiempo discreto la salida de la convolución debería tener $N + M - 1$ muestras (ver ejemplo 2.12), por lo que es de esperar que las primeras $M - 1$ muestras de la salida estén distorsionadas por aliasing, por lo que se descartan (figura 4.41). La salida para el m -ésimo bloque se obtiene con

$$\hat{y}_m(n) \underset{N}{\circlearrowleft} \bullet \hat{Y}_m(k) = H(k)X_m(k), \quad k = 0, 1, \dots, N - 1$$

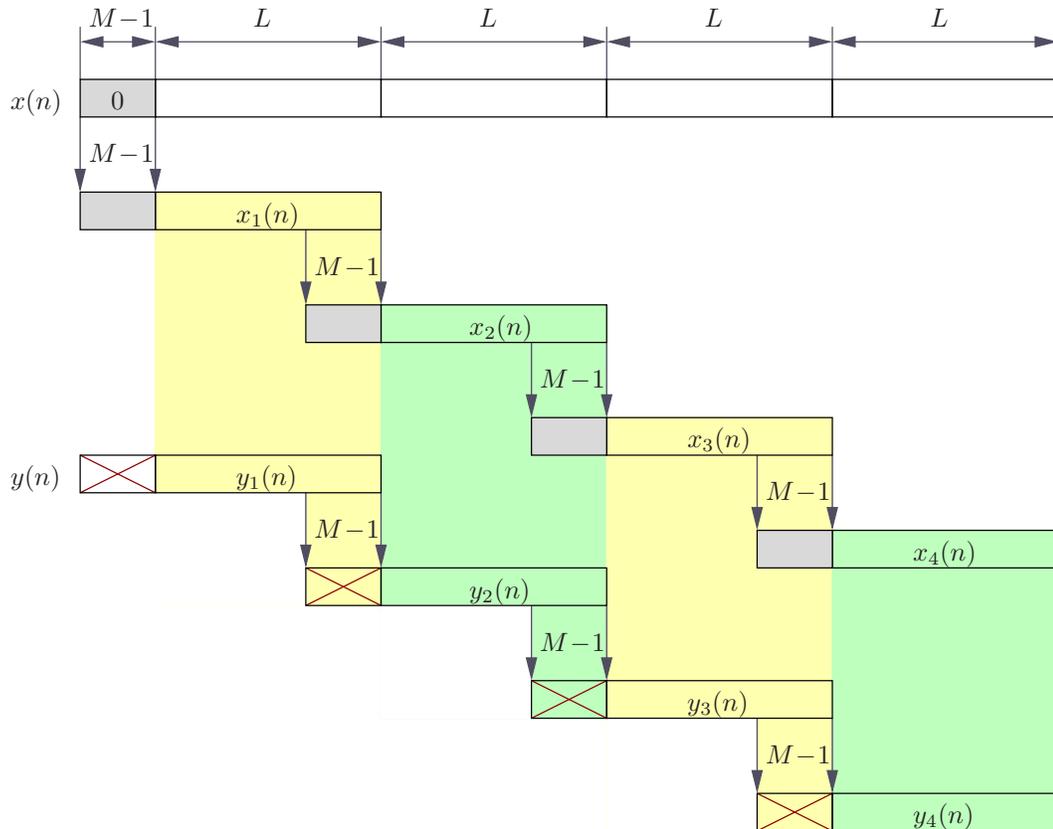


Figura 4.41: Método de solapamiento y almacenamiento.

La salida total se obtiene concatenando las últimas L muestras de cada bloque de salida.

Método de solapamiento y suma

En este método la DFT se aplica a L muestras de la entrada, a las que se concatenan $M - 1$ ceros. Los bloques de salida deben entonces traslaparse y sumarse para obtener el resultado final (figura 4.42).

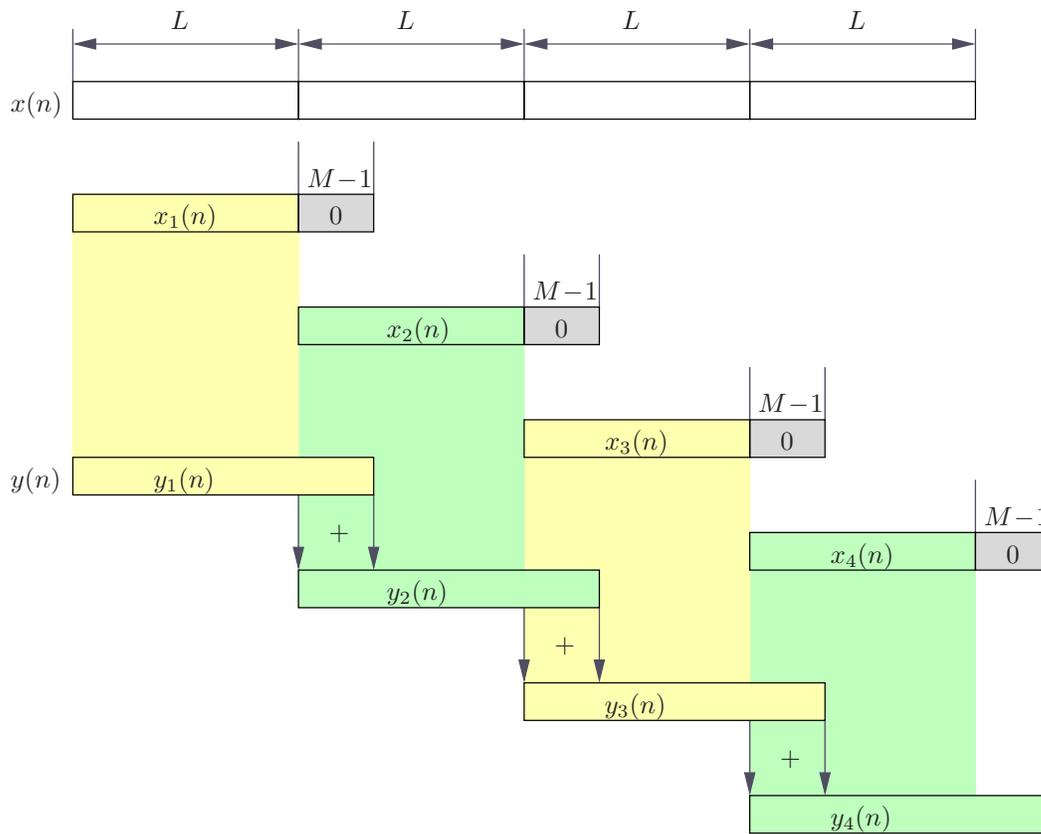


Figura 4.42: Método de solapamiento y suma.

4.7.7 Análisis espectral de señales usando la DFT

Para calcular el espectro exacto de una señal, indiferentemente si esta es continua o discreta, es necesario contar con información de todo instante de tiempo, en el intervalo $t \in]-\infty, \infty[$. En aplicaciones reales esto es imposible y se debe utilizar un intervalo finito de muestras que tendrá repercusiones en la medición espectral.

El procedimiento para el análisis espectral de una señal analógica $x_a(t)$ involucra primero limitar el ancho de banda B con un filtro *antialias* seguido por el muestreo a una tasa $F_s \geq 2B$, que asegura que la mayor frecuencia será $F_s/2$, equivalente a la frecuencia normalizada $f = 1/2$. Además, es necesario limitar la longitud temporal de la señal a L muestras, lo que implica que $T_0 = LT$ es la duración del bloque de señal analizado. No se estudia entonces la señal $x(n)$ sino otra obtenida por enventanado

$$\hat{x}(n) = x(n)w(n)$$

con la ventana de longitud L

$$w(n) = \begin{cases} 1 & 0 \leq n \leq L-1 \\ 0 & \text{en el resto} \end{cases}$$

Para el caso particular de una componente espectral $x(n) = \cos(\omega_0 n)$ con espectro $X(\omega) =$

$\frac{1}{2} [\delta(\omega - \omega_0) + \delta(\omega + \omega_0)]$ la señal enventanada tendrá espectro:

$$\hat{x}(n) \circ \bullet \hat{X}(\omega) = \frac{1}{2} [W(\omega - \omega_0) + W(\omega + \omega_0)]$$

donde $W(\omega)$ es el espectro de la ventana que para el caso de la ventana rectangular es (figura 4.43)

$$W(\omega) = \frac{\text{sen}\left(\frac{\omega L}{2}\right)}{\text{sen}\left(\frac{\omega}{2}\right)} e^{-j\omega(L-1)/2}$$

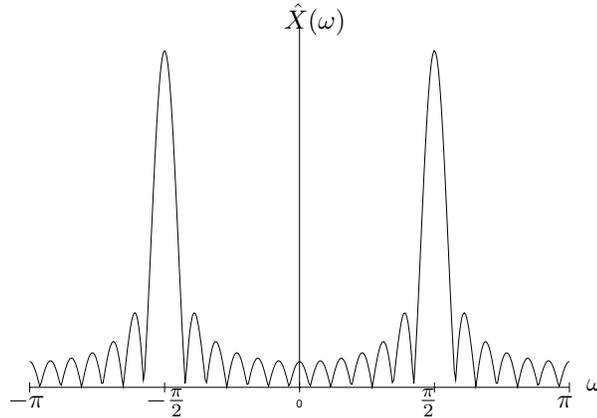


Figura 4.43: Espectro del producto de una señal cosenoidal de frecuencia $\omega = \pi/2$ con una ventana rectangular de longitud $L = 25$.

Nótese que el espectro no se concentra en $\pm\omega_0$, como la señal original, sino que ha ocurrido un *derrame* o *fuga* de la potencia de la señal a todo el intervalo de frecuencias (en inglés *leakage*). Este derrame impide además distinguir líneas espectrales que se encuentren muy cerca. Por ejemplo, con

$$x(n) = \cos(\omega_0 n) + \cos(\omega_1 n) + \cos(\omega_2 n)$$

y $\omega_0 = 0,2\pi$, $\omega_1 = 0,22\pi$ y $\omega_2 = 0,6\pi$ se obtienen las respuestas en magnitud de la figura 4.44.

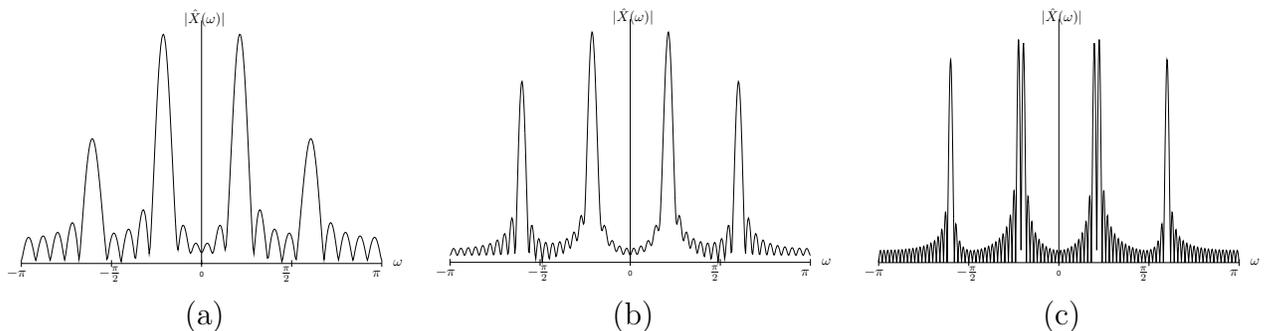


Figura 4.44: Respuestas en magnitud para una superposición de tres señales cosenoidales con (a) $L = 25$, (b) $L = 50$ y (c) $L = 100$.

Los lóbulos laterales pueden reducirse utilizando otras ventanas diferentes a las rectangulares, como la de Hanning:

$$w(n) = \begin{cases} \frac{1}{2} \left(1 - \cos \frac{2\pi}{L-1}n\right) & 0 \leq n \leq L-1 \\ 0 & \text{en el resto} \end{cases}$$

a costa del mayor ancho del lóbulo central.

Capítulo 5

Conversión analógica/digital y digital/analógica

Tal y como se describió en la figura 1.2, para tratar señales analógicas por medios digitales es necesario poder convertir entre representaciones analógicas y digitales.

Conceptualmente en la conversión de una señal analógica a una representación digital intervienen tres pasos (figura 5.1):

1. *Muestreo* es la conversión de una señal de variable continua a otra de variable discreta que es el resultado de tomar “muestras” de la señal de variable continua en ciertos instantes. Si $x_a(t)$ es la entrada al bloque de muestreo, entonces la salida es $x_a(nT)$, donde a T se le denomina el *intervalo de muestreo*.
2. *Cuantificación* es la conversión de la señal de variable discreta y valores continuos a otra señal de variable discreta pero con valores discretos. El valor de cada muestra es aproximado entonces con un valor de un conjunto finito de posibles valores. A la diferencia entre el valor continuo y su aproximación se le denomina *error de cuantificación*.
3. *Codificación* consiste en la asignación de una representación usualmente binaria para los valores cuantificados.

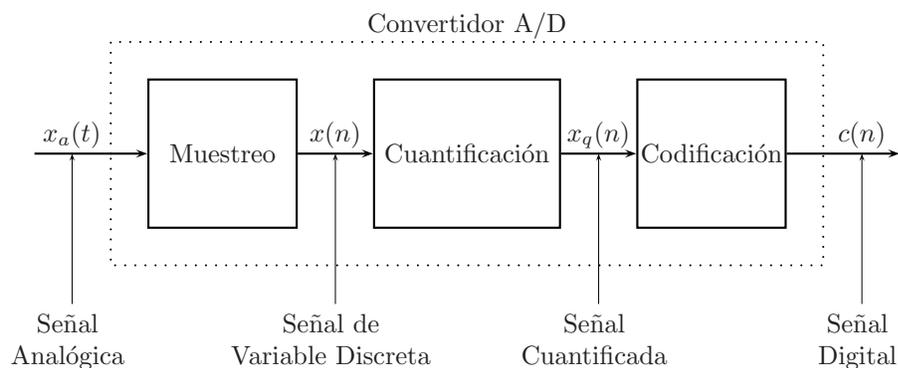


Figura 5.1: Pasos básicos en la conversión analógica/digital.

Estos pasos en la práctica se realizan en un solo bloque operacional.

La conversión digital/analógica es necesaria en todos aquellos casos en los que la representación propiamente digital no tiene significado para el usuario final. Por ejemplo, la salida digital de un decodificador de voz debe ser convertida a una señal analógica acústica para poder ser comprendida por los usuarios. Este proceso de conversión debe *interpolarse* la señal discreta.

5.1 Muestreo de señales analógicas

Existen diversas posibilidades de seleccionar las muestras de una señal analógica en la fase de muestreo. En el análisis clásico se utiliza el llamado *muestreo periódico* o *uniforme* por las facilidades que este brinda al análisis matemático, lo cual ha hecho que sea el tipo de muestreo más utilizado en aplicaciones reales. En el muestreo periódico la relación entre la señal analógica y la señal de variable discreta está dada por

$$x(n) = x_a(nT) \quad -\infty < n < \infty$$

donde la secuencia $x(n)$ contiene entonces muestras de la señal analógica $x_a(t)$ separadas por un intervalo T (figura 5.2).

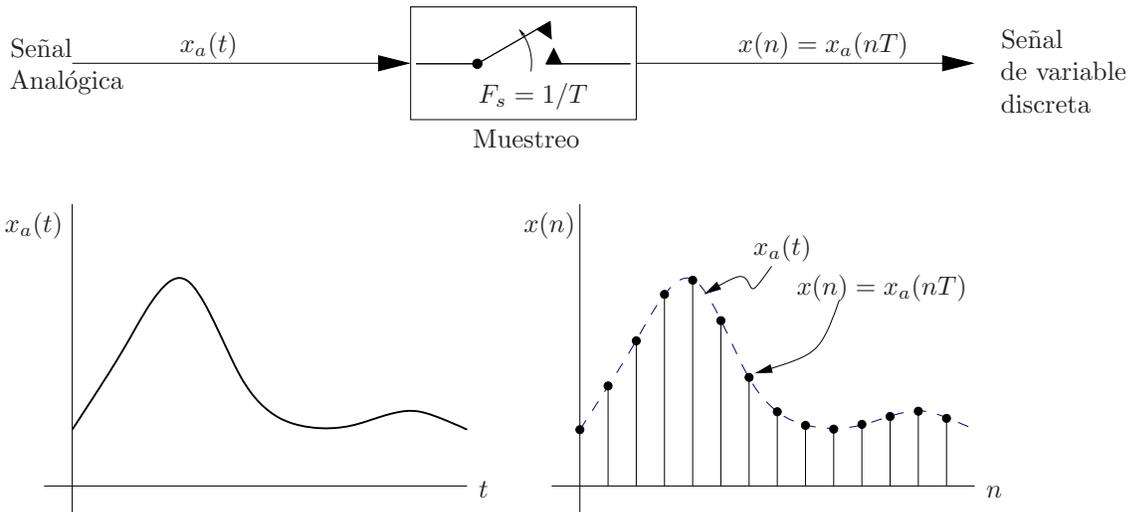


Figura 5.2: Muestreo periódico de una señal analógica.

Las variables t y n de las señales de variable continua y discreta respectivamente están relacionadas a través del intervalo de muestreo T

$$t = nT = n/F_s$$

donde a F_s se le denomina tasa de muestreo.

En la sección 2.2.2 se demostró que la relación entre las frecuencias de las señales en los dominios continuo y discreto está dada por

$$\omega = \Omega T$$

o lo que es equivalente

$$f = \frac{F}{F_s} \quad (5.1)$$

por lo que a f se le denomina también *frecuencia normalizada*. Con esto queda claro que la frecuencia f puede corresponder con las unidades de frecuencia de F (por ejemplo, en Hertz) si y solo si se conoce la frecuencia de muestreo F_s .

En las secciones anteriores se obtuvo además que el rango para la frecuencia F es prácticamente infinito, mientras que el rango de la frecuencia f es finito ($[-\frac{1}{2}, \frac{1}{2}]$). Utilizando (5.1) se obtiene

$$\begin{aligned} -\frac{1}{2} < f &\leq \frac{1}{2} \\ -\frac{1}{2} < \frac{F}{F_s} &\leq \frac{1}{2} \\ -\frac{F_s}{2} < F &\leq \frac{F_s}{2} \end{aligned}$$

El muestreo periódico de la señal de variable continua representa entonces un mapeo del rango infinito de frecuencias F a un rango finito de frecuencias f . Puesto que la mayor frecuencia de f es $f = \frac{1}{2}$ puede concluirse que con la tasa de muestreo F_s como máximo se podrá representar

$$\begin{aligned} F_{max} &= \frac{F_s}{2} = \frac{1}{2T} \\ \Omega_{max} &= \pi F_s = \frac{\pi}{T} \end{aligned}$$

En capítulos anteriores se demostró que las frecuencias f son únicas dentro de un rango finito: el rango fundamental, y que todas las frecuencias fuera de ese rango son alias de alguna frecuencia fundamental. Se demostró además que para una señal sinusoidal la frecuencia angular ω y $\omega + 2k\pi$ son equivalentes para todo $k \in \mathbb{Z}$, lo que implica que $f + k$ es un alias de f . Con (5.1) se tiene que

$$f + k = \frac{F}{F_s} \Rightarrow fF_s + kF_s \equiv F$$

o en otras palabras $fF_s + kF_s$ es un alias de $F = fF_s$. Con esto, la relación entre las frecuencias de las señales de variable continua y discreta puede graficarse como lo muestra la figura 5.3.

La figura 5.4 muestra un ejemplo con las frecuencias $F_1 = 2/7$ Hz y su alias $F_2 = -5/7$ Hz para una tasa de muestreo de $F_s = 1$ Hz. Aquí se aprecia con claridad que las muestras tomadas para las dos señales coinciden, por lo que estas muestras por si solas no pueden determinar de cuál de las dos señales fue muestreada.

La frecuencia $F_s/2$ corresponde entonces a la frecuencia máxima representable en forma discreta con una tasa de muestreo F_s , que corresponde entonces a $f = \frac{1}{2}$. La frecuencia fundamental correspondiente a un alias de frecuencia mayor a $F_s/2$ se puede terminar utilizando $F_s/2$ como pivote para reflejar el alias dentro del rango fundamental. A $F_s/2$ (o $\omega = \pi$) se le denomina entonces *frecuencia de plegado*, o en inglés *folding frequency*.

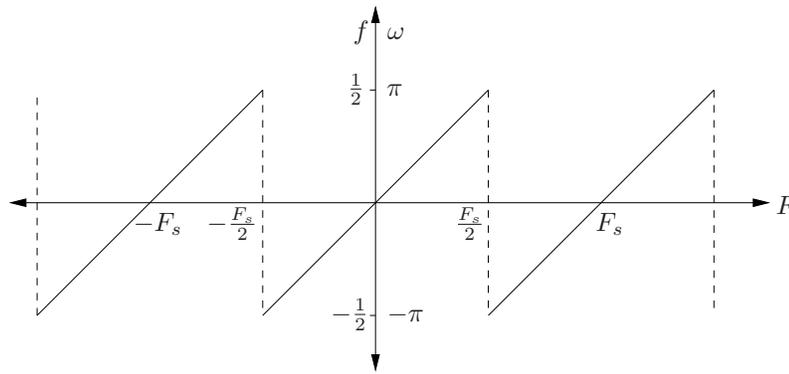


Figura 5.3: Relación entre las frecuencias de señales de variable discreta y continua para el caso de muestreo uniforme.

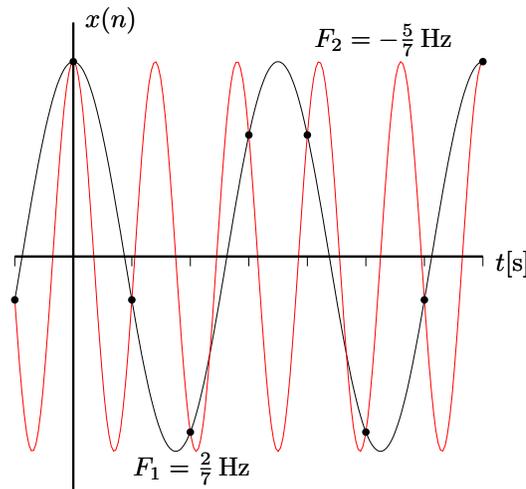


Figura 5.4: Muestreo de una señal sinusoidal y de un alias.

5.1.1 Muestreo de señales pasa-bajos

Como señal pasabajos se considera aquella cuyo espectro tiene componentes de frecuencia diferentes de cero para un valor dado de frecuencia $F < B$. La figura 4.11 ilustra un espectro de señal pasabajos.

Tal y como se ha demostrado en la sección 4.2.4, no hay pérdida de información si una señal se muestrea con $F_s > 2B$, pues las réplicas del espectro original inducidas por el muestreo, y ubicadas cada F_s , no se traslapan y permiten la recuperación del espectro original por medio de un interpolador ideal aplicado a las muestras.

5.1.2 Muestreo de señales pasa-banda

Una señal pasa-banda es aquella cuyo espectro es cero para $|F| < F_c - B/2$ y para $|F| > F_c + B/2$. A F_c se le conoce como frecuencia central de la banda o frecuencia *portadora* y a B como ancho de la banda. La figura 5.5 ilustra un ejemplo de espectro paso-banda de

una señal real. Según el teorema de muestreo es posible reconstruir la señal si se utiliza

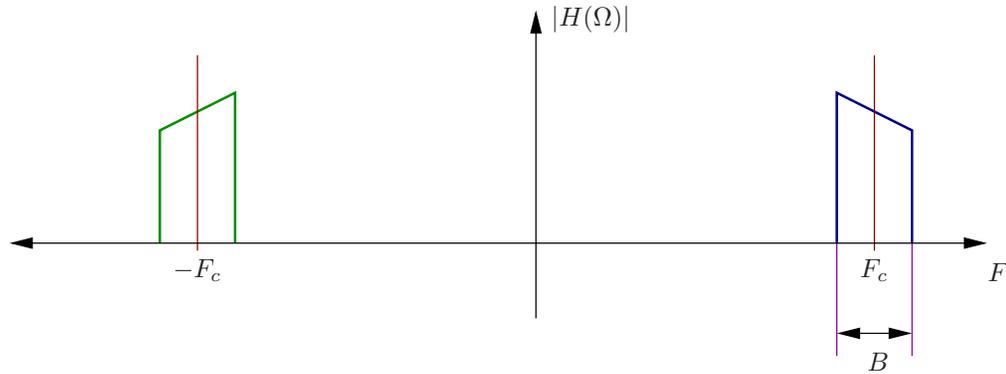


Figura 5.5: Señal pasa-banda

como frecuencia de muestreo al menos el doble de la frecuencia máxima encontrada en la señal, es decir, $F_s > 2F_c + B$. En aplicaciones donde la frecuencia portadora es superior a la disponibilidad de muestreo de convertidores analógico digital comerciales, o la capacidad de procesamiento requerida no da abasto con el número de muestras capturado por intervalo de tiempo, es posible re-interpretar el teorema de muestreo para utilizar réplicas espectrales en secciones de menor frecuencia y así utilizar *submuestreo*, o muestreo sub-Nyquist.

La figura 5.6 ilustra un caso de muestreo de la señal con ancho de banda B montada sobre una portadora en F_c utilizando como frecuencia de muestreo $F_s = F_c - B/2$.

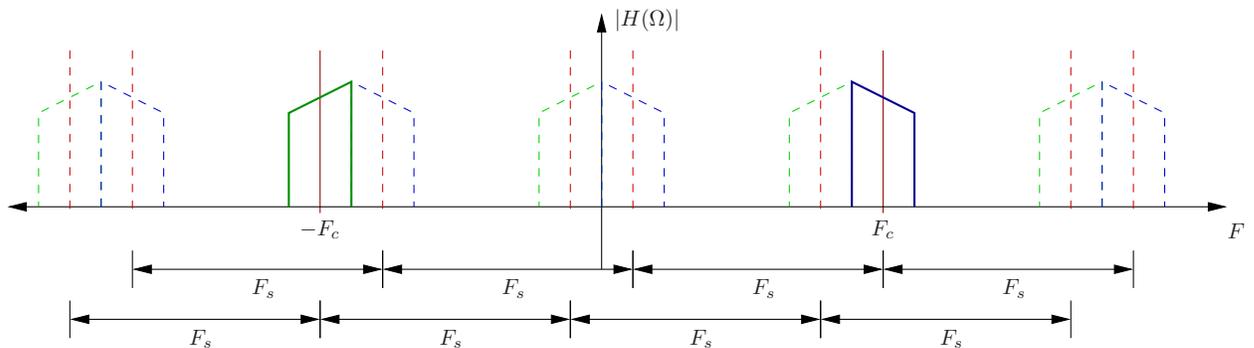


Figura 5.6: Muestreo de señal pasa-banda sin aliasing

Respetando el teorema del muestreo, enunciado ahora como $F_s > 2B$, es posible elegir la frecuencia de muestreo de tal modo que las réplicas de las bandas no se traslapen, evitando así que ocurran efectos de aliasing.

La figura 5.7 ilustra un caso particular en el que se logran situar seis réplicas entre las bandas de una señal analógica.

Obsérvese que en el rango $2F_c - B$, existente entre las dos bandas de la señal, deben caber exactamente m réplicas, por lo que se debe cumplir

$$mF_s = 2F_c - B$$

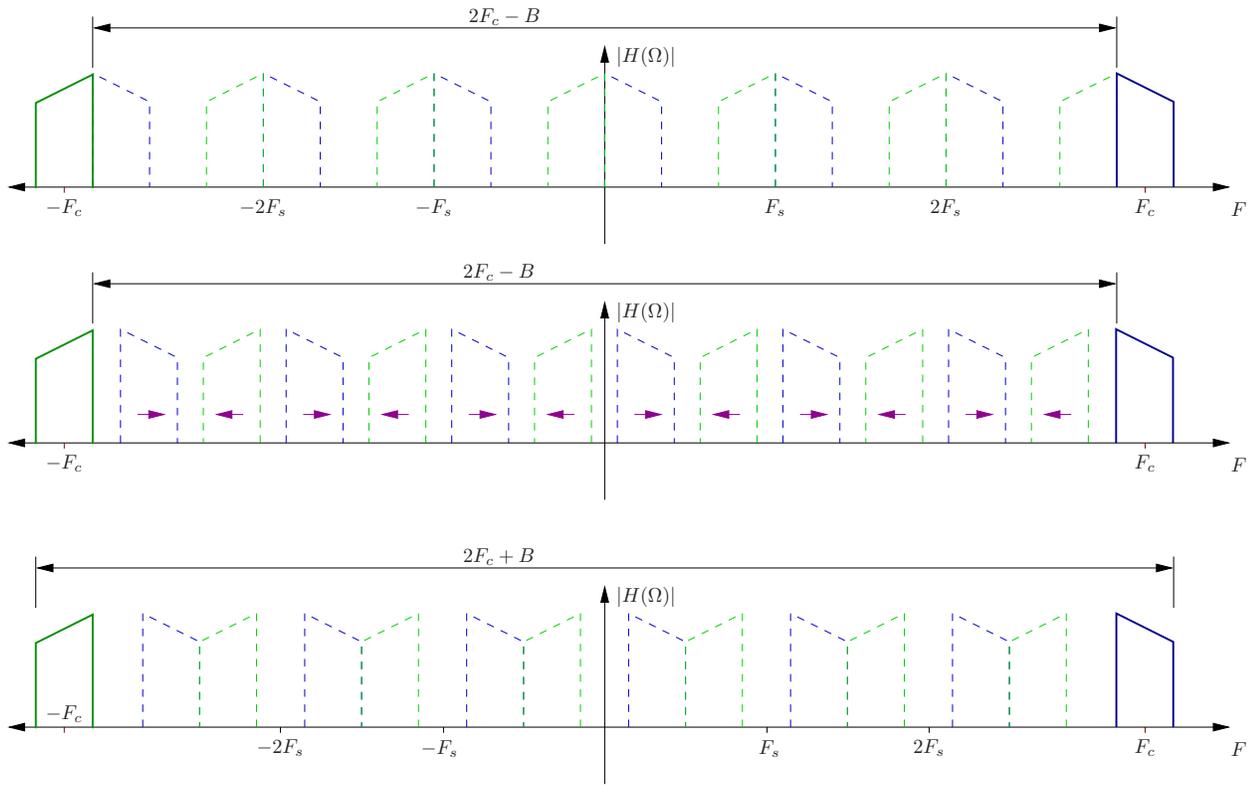


Figura 5.7: Muestreo de señal pasa-banda sin aliasing. Arriba se utiliza $F_s = (2F_c/B)/6$; al medio se muestrea con $F'_s < F_s$; abajo se utiliza la menor frecuencia posible.

y por tanto

$$F_s = \frac{2F_c - B}{m} \quad (5.2)$$

donde el número natural m estará limitado además por la restricción $F_s > 2B$.

Si la frecuencia de muestreo F_s se incrementa, las réplicas se separarán entre sí, causando un traslape y por tanto aliasing en la señal muestreada. Si F_s se reduce, las réplicas se acercarán entre sí, y F_s puede reducirse hasta que ahora en el rango extendido que cubre a las dos bandas originales ($2F_c + B$) quepan entonces $m + 1$ réplicas distribuidas de forma homogénea:

$$(m + 1)F_s = 2F_c + B$$

y por tanto

$$F_s = \frac{2F_c + B}{m + 1} \quad (5.3)$$

Si se reduce F_s aun más, de nuevo ocurrirá un traslape entre las bandas.

Resumiendo lo anterior, se debe cumplir para la frecuencia de muestreo F_s

$$\frac{2F_c + B}{m + 1} \leq F_s \leq \frac{2F_c - B}{m} \quad (5.4)$$

donde, de nuevo, m es un número natural que debe además asegurar que $F_s > 2B$.

Defínase r como la razón entre la mayor componente de frecuencia $F_c + B/2$ y el ancho de banda B de la señal a muestrear. Reescribiendo (5.4) se obtiene

$$\begin{aligned} \frac{2F_c + B}{m + 1} &\leq F_s \leq \frac{2F_c - B}{m} \\ 2\frac{(F_c + \frac{B}{2})}{m + 1} &\leq F_s \leq \frac{2F_c - B + B - B}{m} \\ 2\frac{(F_c + \frac{B}{2})\frac{B}{B}}{m + 1} &\leq F_s \leq 2\frac{(F_c + \frac{B}{2} - B)}{m} \\ \frac{2Br}{m + 1} &\leq F_s \leq 2\frac{Br - B}{m} \\ \frac{2r}{m + 1} &\leq \frac{F_s}{B} \leq \frac{2(r - 1)}{m} \end{aligned}$$

La figura 5.8 ilustra la frecuencia de muestreo normalizada con el ancho de banda B en función del radio r . La zona sombreada representa configuraciones inválidas.

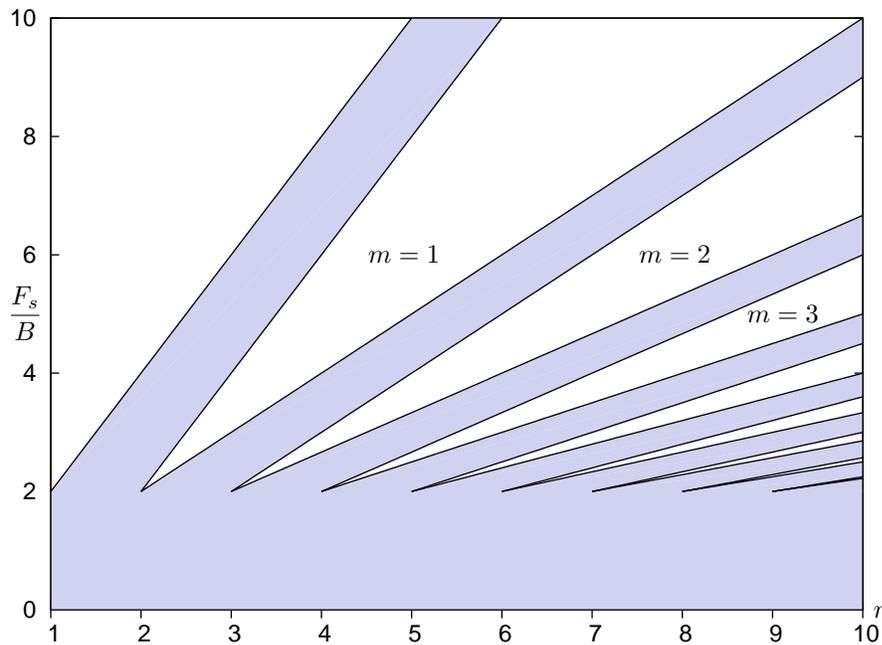


Figura 5.8: Rangos válidos de frecuencias de sub-muestreo, en función del radio entre la frecuencia espectral máxima y el ancho de banda.

Debe notarse que si m es impar, la réplicas cercanas a cero se invertirán, en el sentido que la réplica de la banda original en frecuencias positivas se encontrará en frecuencias negativas y la réplica de la banda original en frecuencias negativas se encontrará en frecuencias positivas. Si la aplicación requiere contar con el espectro sin invertir, es necesario utilizar m par.

Al igual que en el caso de señales pasa-bajo, para aplicaciones reales es necesario dejar un margen entre las bandas con el fin de poder utilizar un filtros que rescaten las réplicas inferiores, con las cuales usualmente se desea trabajar.

5.2 Cuantificación de señales de amplitud continua

La cuantificación de una señal de amplitud continua consiste en asignar a cada valor continuo otro valor tomado de un conjunto finito. El reemplazar los valores continuos por los valores cuantificados introduce el llamado *error de cuantificación* o *ruido de cuantificación*. Si $x(n)$ es la secuencia de valores continuos entonces la secuencia de valores cuantificados será $x_q(n) = Q[x(n)]$ donde Q representa un operador de cuantificación. La secuencia de errores de cuantificación está dada entonces por

$$e_q(n) = x_q(n) - x(n)$$

Para analizar el error de cuantificación considérese la señal sinusoidal de variable continua mostrada en la figura 5.9. Es posible apreciar que alrededor de cada nivel de cuantificación

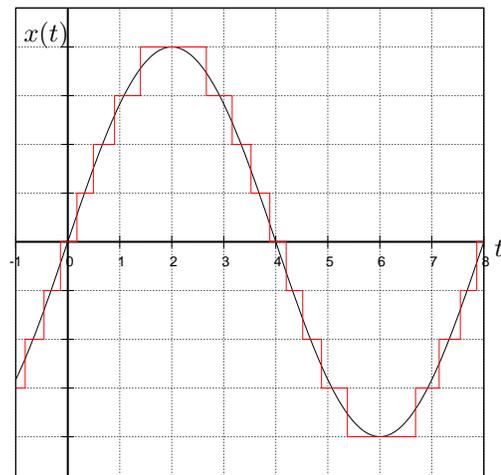


Figura 5.9: Señal sinusoidal de variable continua cuantificada.

la señal analógica se puede aproximar de forma lineal. En términos estadísticos este hecho se describe diciendo que el ruido de cuantificación sigue una distribución uniforme, donde los valores de error estarán distribuidos con las mismas probabilidades entre $-\Delta/2$ y $\Delta/2$, donde Δ es igual al *cuanto*, es decir la distancia entre dos niveles de cuantificación.

Esta aproximación lineal se esquematiza en la figura 5.10, donde se ha asumido que los “saltos” hacia los otros cuantos de esta señal ocurren en $-\tau$ y τ .

La potencia media del error cuadrático P_q es

$$P_q = \frac{1}{2\tau} \int_{-\tau}^{\tau} e_q^2(t) dt = \frac{1}{\tau} \int_0^{\tau} e_q^2(t) dt$$

que considerando la aproximación lineal del error $e_q(t) \approx (\Delta/2\tau)t$ para $t \in [-\tau, \tau]$ resulta en

$$P_q = \frac{1}{\tau} \int_0^{\tau} \left(\frac{\Delta}{2\tau} \right)^2 t^2 dt = \frac{\Delta^2}{12} \quad (5.5)$$

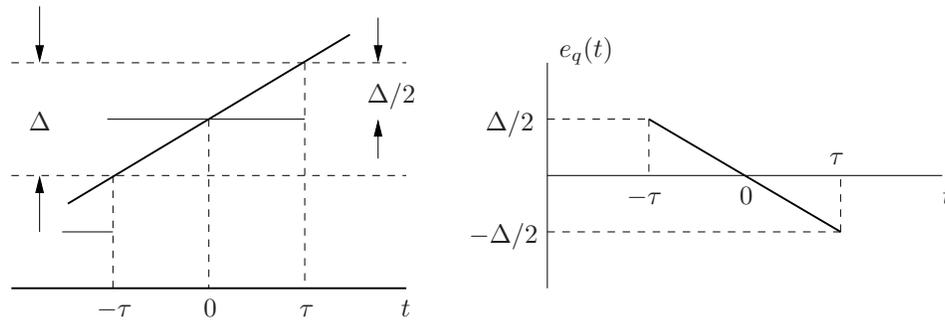


Figura 5.10: Error de cuantificación $e_q(t) = x_a(t) - x_q(t)$.

Si el cuantificador utiliza b bits de precisión y todo el rango posible con ellos para representar la amplitud de $2A$, entonces se puede calcular el cuanto como $\Delta = 2A/2^b$, lo que implica que

$$P_q = \frac{A^2/3}{2^{2b}} \quad (5.6)$$

La potencia media de la señal $x_a(t)$ se calcula como

$$P_x = \frac{1}{T_p} \int_0^{T_p} (A \cos(\Omega_0 t))^2 dt = \frac{A^2}{2}.$$

Con estos dos términos es posible entonces aproximar la relación señal a ruido de cuantificación (*SQNR* signal to quantization noise ratio) como:

$$\text{SQNR} = \frac{P_x}{P_q} = \frac{3}{2} \cdot 2^{2b}$$

y expresada en decibels (dB)

$$\text{SQNR(dB)} = 10 \log_{10} \text{SQNR} = (1,76 + 6,02b) \text{ dB}$$

lo que implica que por cada bit añadido a la palabra, y si se utiliza todo el rango, es decir, si se duplica la cantidad de cuantos utilizados, la calidad de la señal mejora en 6 dB. El factor $3/2$ en la razón de potencias es específico para señales sinusoidales. En un caso más general, si se asume que la señal de entrada es generada por un proceso aleatorio de media cero y con potencia $P_x = \sigma_x^2$, entonces la razón SQNR es

$$\text{SQNR} = \frac{P_x}{P_q} = \frac{\sigma_x^2}{\frac{A^2/3}{2^{2b}}} = \frac{\sigma_x^2}{A^2} 3 \cdot 2^{2b}$$

y en decibels

$$\begin{aligned} \text{SQNR(dB)} &= 10 \log_{10} \left(\frac{\sigma_x^2}{A^2} \right) + 10 \log_{10} 3 + 2b \log_{10} 2 \\ &= 10 \log_{10} \left(\frac{\sigma_x^2}{A^2} \right) + 4,77 \text{ dB} + 6,02b \text{ dB} \end{aligned} \quad (5.7)$$

de donde se observa que el término $(6,02b)$ dB se mantiene.

Ejemplo 5.1 El sistema de audio digital utilizado en CD comerciales utiliza una frecuencia de muestreo de 44,1 kHz con 24 bits por muestra. Determine la frecuencia máxima representable por este sistema y el valor máximo de la relación de señal a ruido de cuantificación SQNR alcanzable para un tono puro.

¿Cuántos bits son estrictamente necesarios por muestra si la aplicación requiere para un tono puro una SQNR de al menos 100 dB?

Solución:

Con 44,1 kHz de frecuencia de muestreo es posible representar a lo sumo frecuencias de 22,05 kHz.

Los 24 bits por muestra permiten, si el convertidor A/D cubre exactamente el rango dinámico de la señal, un $SQNR = (1,76 + 6,02 \times 24) = 146,24$ dB, asumiendo un tono sinusoidal.

Si se requirieran 100 dB entonces bastarían para reproducir un tono puro (señal sinusoidal) 17 bits por muestra. 5.1

5.3 Codificación de los valores cuantificados

El proceso de codificación asigna un número binario único a cada nivel de cuantificación. Si se disponen de L niveles entonces serán necesarios al menos $b = \lceil \log_2(L) \rceil$ bits. Los convertidores A/D disponibles comercialmente usualmente utilizan precisiones de 8 a 16 bits.

Dos familias de codificación se utilizan con frecuencia: coma fija y coma flotante.

5.3.1 Números codificados con coma fija

En las representaciones con coma fija el peso de cada bit en la representación es constante. Dos codificaciones frecuentemente utilizadas consisten en los enteros sin signo y el complemento a dos.

Enteros sin signo

Sea x un número entero sin signo de N -bits

$$x = \sum_{n=0}^{N-1} b_n 2^n$$

donde $b_n \in \{0, 1\}$ es el n -ésimo dígito de x . El rango representable es entonces desde 0 hasta $2^N - 1$. El dígito b_0 es el menos significativo (LSB, *least significant bit*) y su peso relativo es igual a uno. El dígito b_{N-1} es el más significativo (MSB, *most significant bit*) y tiene un peso relativo de 2^{N-1}

Coma fija sin signo

Sea x un número sin signo de N -bits

$$x = \frac{1}{M} \sum_{n=0}^{N-1} b_n 2^n$$

donde $b_n \in \{0, 1\}$ es el n -ésimo dígito de x y M es una constante de normalización elegida usualmente como 2^m . El rango representable es entonces desde 0 hasta $(2^N - 1)/M$. El dígito b_0 es el menos significativo (LSB, *least significant bit*) y su peso relativo es igual a $1/M$. El dígito b_{N-1} es el más significativo (MSB, *most significant bit*) y tiene un peso relativo de $2^{N-1}/M$.

Complemento a dos

La representación de N bits de un número entero con signo en complemento a dos está dada por

$$x = -b_{N-1} 2^{N-1} + \sum_{n=0}^{N-2} b_n 2^n$$

lo que permite representar números en el rango desde -2^{N-1} hasta $2^{N-1} - 1$. El dígito b_0 es el menos significativo (LSB, *least significant bit*) y su peso relativo es igual a uno. El dígito b_{N-2} es el más significativo (MSB, *most significant bit*) y tiene un peso relativo de 2^{N-2} . El último bit b_{N-1} codifica al signo.

El uso del complemento a dos es quizá el más difundido de todas las representaciones de números con signo. Esto se debe a que es posible sumar varios números con signo, y siempre que el resultado final se encuentre en el rango de representación, es irrelevante si resultados intermedios producen desbordamiento. Por ejemplo, supóngase que se debe hacer la secuencia de operaciones $2 + 3 - 2$ con números de 3 bits. La secuencia de adiciones es entonces

$(x_i)_{10}$	$(x_i)_2$	$(\sum x_i)_2$	$(\sum x_i)_{10}$
2_{10}	010	010	2_{10}
3_{10}	011	101	-3_{10}
-2_{10}	110	011	3_{10}

donde el resultado intermedio 5 fue representado por el número -3 , sin afectar el resultado final. Otra ventaja de esta representación es que permite encontrar el módulo con respecto a un número 2^L utilizando únicamente los L bits menos significativos, lo que simplifica la implementación de búferes anulares. Por ejemplo, para el caso $N = 4$ y $L = 2$ se obtiene

$(x_i)_{10}$	$(x_i)_2$	$(x_i \bmod 4)_2$	$(x_i \bmod 4)_{10}$
-8 ₁₀	1000	00	0 ₁₀
-7 ₁₀	1001	01	1 ₁₀
-6 ₁₀	1010	10	2 ₁₀
-5 ₁₀	1011	11	3 ₁₀
-4 ₁₀	1100	00	0 ₁₀
-3 ₁₀	1101	01	1 ₁₀
-2 ₁₀	1110	10	2 ₁₀
-1 ₁₀	1111	11	3 ₁₀
0 ₁₀	0000	00	0 ₁₀
1 ₁₀	0001	01	1 ₁₀
2 ₁₀	0010	10	2 ₁₀
3 ₁₀	0011	11	3 ₁₀
4 ₁₀	0100	00	0 ₁₀
5 ₁₀	0101	01	1 ₁₀
6 ₁₀	0110	10	2 ₁₀
7 ₁₀	0111	11	3 ₁₀

Coma fija con signo

La representación de N bits de un número con signo en complemento a dos está dada por

$$x = \frac{1}{M} \left(-b_{N-1}2^{N-1} + \sum_{n=0}^{N-2} b_n 2^n \right)$$

con la constante de normalización M elegida usualmente como 2^m . El rango representable será entonces desde $-2^{N-1}/M$ hasta $(2^{N-1} - 1)/M$.

Un caso frecuentemente utilizado permite representar números de valor absoluto igual o inferior a uno empleando $M = 2^{N-1}$ con lo que se obtiene

$$x = -b_{N-1} + \sum_{n=0}^{N-2} b_n 2^{n-N+1}$$

5.3.2 Números codificados con coma flotante

La representación en coma flotante permite ampliar el rango de representación numérica. Las representaciones de 32 y 64 bits más frecuentemente utilizadas han sido estandarizadas por la IEEE (estándar 754-1985). Las unidades de manejo de coma flotante integradas en algunos procesadores Digitales de Señal avanzados (por ejemplo, la familia C674x de Texas Instruments) así como algunos módulos disponibles gratuita y/o comercialmente en HDL (Verilog o VHDL) dan soporte a dicho estándar. La flexibilidad del hardware reconfigurable permite sin embargo utilizar formatos de coma flotante adecuados para cada aplicación específica sin afectar el desempeño.

Un número codificado con el estándar consiste en un bit de signo s , el exponente e con E bits y la mantisa m normalizada (fraccionaria) de M bits, y se codifica como

s	Exponente e	Mantisa m
-----	---------------	-------------

De forma algebraica, el número representado es

$$x = (-1)^s \times 1, m \times 2^{e-\text{bias}}$$

con

$$\text{bias} = 2^{E-1} - 1$$

Nótese que la mantisa se completa con un 1 *oculto* (en el sentido de que no se indica explícitamente en la representación), mientras que los bits especificados en la mantisa representan solo la parte fraccionaria.

Ejemplo 5.2 Indique cuál es la representación en coma flotante del número $10,125_{10}$ en un formato de 14 bits que utiliza $E = 6$ bits y $M = 7$ bits.

Solución:

Primero, el bias está dado por

$$\text{bias} = 2^{E-1} - 1 = 2^5 - 1 = 31$$

y para la mantisa

$$10,125_{10} = 1010,0010_2 = 1,0100010_2 \times 2^3$$

El exponente corregido se obtiene con

$$e = 3 + \text{bias} = 34_{10} = 100010_2$$

Finalmente, la representación del número es:

s	Exponente e	Mantisa m
0	100010_2	0100010_2

5.2

Ejemplo 5.3 Encuentre qué número decimal es representado por el código de coma flotante con $E = 6$ bits y $M = 7$ bits:

s	Exponente e	Mantisa m
1	011110_2	1000000_2

Tabla 5.1: Estándar de coma flotante IEEE 754-1985

	Simple	Doble
Ancho de palabra	32	64
Mantisa	23	52
Exponente	8	11
Bias	127	1023
Rango	$2^{128} \approx 3,8 \times 10^{38}$	$2^{1024} \approx 1,8 \times 10^{308}$

Solución:

El número representado está dado por

$$-1 \times 1,1000000_2 \times 2^{30-\text{bias}} = -1,1_2 \times 2^{-1} = -0,11_2 = -0,75_{10}$$

5.3

La tabla 5.1 muestra las especificaciones del estándar IEEE 754. Además de la estructura de representación de los números reales, el estándar especifica además detalles como modos de redondeo, denormalizaciones, y representaciones de números inválidos o infinitos.

Por hacer:

Agregar cómo se representa el cero, que no es trivial por el hidden bit 1 y otros números

5.4 Circuitos de conversión analógica digital

Diferentes estrategias para la conversión analógica a digital se han propuesto, cada una de ellas adecuada a factores como potencia, velocidad o precisión, que caracterizarán a cada aplicación en particular. La selección del convertidor analógico digital es parte del diseño de un sistema de procesamiento digital.

5.4.1 Circuito de muestreo y retención

El circuito de muestreo y retención (o *sample and hold*) es utilizado como etapa previa al circuito convertidor propiamente dicho para estabilizar el valor a ser muestreado. El periodo total de muestreo $T_s = 1/F_s$ se divide en el tiempo de muestreo t_s en el que la salida de tensión U_2 del retenedor se estabiliza siguiendo a la tensión de entrada U_1 (figura 5.11), y el tiempo de retención t_H donde se mantiene el último valor de tensión de la entrada del tiempo t_s . Durante el tiempo de retención la salida U_2 es constante y es convertida a una palabra digital en el proceso de cuantificación.

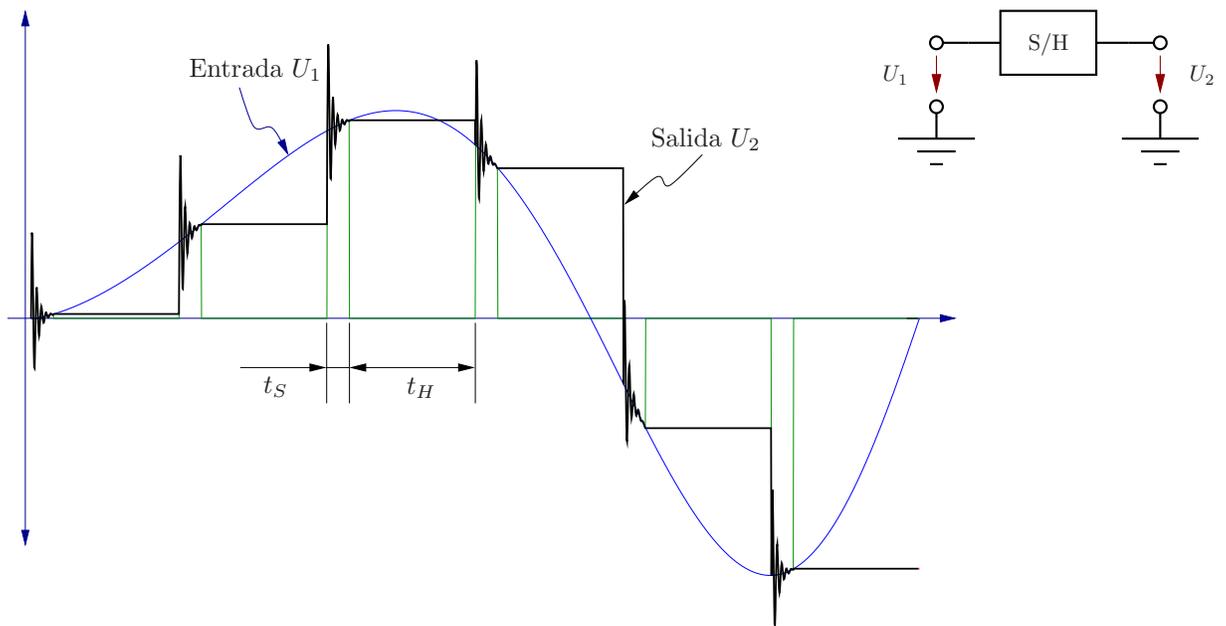


Figura 5.11: Muestreo y retención

5.4.2 Contador

Una primera estrategia de conversión analógica a digital utiliza el esquema ilustrado en la figura 5.12. Durante cada fase de retención del circuito S/H, una señal con forma de

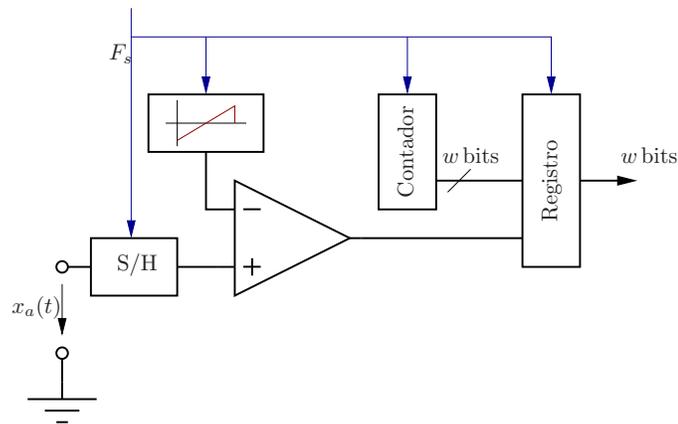


Figura 5.12: Conversión con contador

de sierra $s(t)$ recorre todo el rango de valores analógicos representables y en paralelo un contador digital es reinicializado e inicia una cuenta ascendente sincronizada con la señal $s(t)$. En el instante en que $s(t)$ supere a la salida del retenedor, el valor del contador es almacenado en un registro y utilizado como representación digital de la señal.

Esta técnica es utilizada para bajas frecuencias de muestreo (menores a 50 kHz) debido a la cantidad de ciclos de reloj requeridos por el contador para recorrer todo el rango de valores. Por las mismas razones, tampoco es posible utilizar un número elevado de bits.

La precisión de este método está limitada por la precisión de generación de la señal $s(t)$, así

como la precisión de la sincronización de esta señal con los contadores.

5.4.3 Aproximaciones sucesivas

En la conversión por aproximaciones sucesivas se sigue la estrategia ilustrada en la figura 5.13. El registro de aproximación sucesiva (RAS) consiste en un sistema que verifica e inicializa

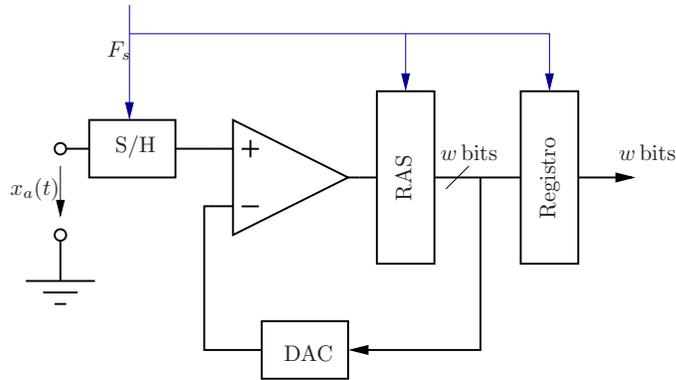


Figura 5.13: Conversión por aproximaciones sucesivas

los bits del código digital del más significativo al menos significativo. En un primer paso, a dicho bit se le asigna el valor de 1, y esa representación se convierte a un valor analógico con el DAC. Si la tensión a la salida del DAC es menor que la salida del retenedor, eso indica que el bit más significativo debe ser uno, o de otra manera debe ser cero. El valor de dicho bit se almacena y se procede de igual forma con el siguiente bit de menor peso relativo. De esta manera, en un número de pasos w igual al número de bits de la representación digital se realiza la conversión. Este proceso comparativo se ilustra en la figura 5.14.

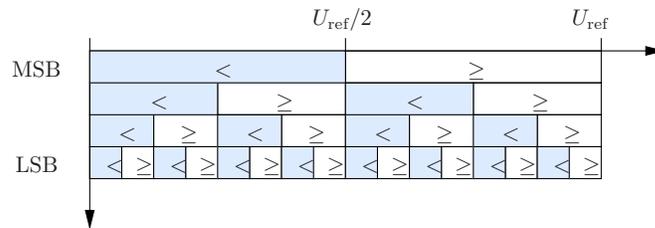


Figura 5.14: Registro de aproximaciones sucesivas

Con una resolución de 12 bits se alcanzan frecuencias de muestreo de hasta 1 MHz. Mayores resoluciones de hasta 16 bits se encuentran disponibles pero a menores frecuencias de muestreo.

Obsérvese que la precisión de este método depende de qué tan exacto sea el convertidor digital analógico (DAC).

5.4.4 Convertidor paralelo

La figura 5.15 muestra un método directo de conversión denominado convertidor *flash* o convertidor paralelo. En este método, la salida del retenedor es comparada con $2^w - 1$ tensiones de referencia obtenidas por una red resistiva. La salida de los comparadores se interpreta como una secuencia de $2^w - 1$ bits que pueden ser codificados para obtener un código de w bits.

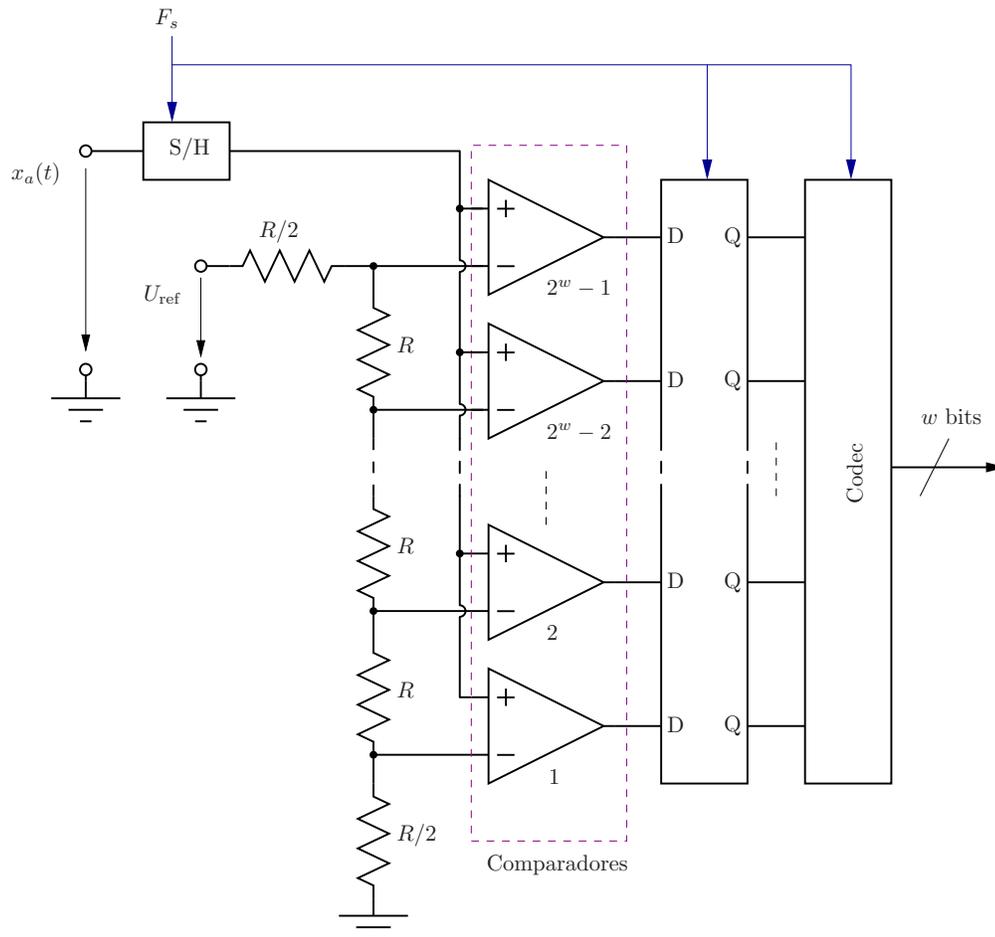


Figura 5.15: Convertidor paralelo

Las tasas de muestreo obtenibles con estos convertidores oscilan entre 1 y 500 MHz para resoluciones de hasta 10 bits. Sin embargo, las tolerancias alcanzables en la fabricación de la red resistiva (por ejemplo 1023 resistencias en un convertidor de 10 bits) afectará directamente la linealidad del convertidor, o en otras palabras, la homogeneidad del ancho analógico de tensión asignado al bit menos significativo en todo el rango de conversión es difícil de alcanzar con convertidores de este tipo.

5.4.5 Convertidor en subrangos

Una manera de reducir el número de resistencias requeridas en el convertidor anterior consiste en partir el problema en dos subproblemas con la mitad del número de bits, tal y como lo

ilustra la figura 5.16. En un primer paso un convertidor *flash* se utiliza para obtener los

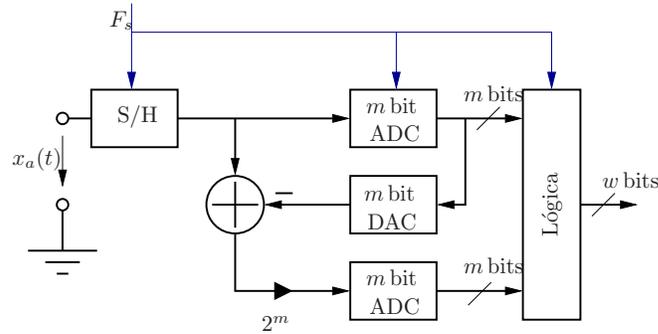


Figura 5.16: Convertidor half-flash

m bits más significativos de la tensión a la salida del retenedor. A partir de ellos se obtiene una tensión que aproxima a la entrada redondeada a solo m bits. La salida de convertidor DAC utilizado para generar dicha aproximación se sustrae de la salida del retenedor para obtener la fracción faltante, la cual es amplificada por un factor 2^m y transformada con otro convertidor *flash* para obtener los m bits menos significativos.

El esquema anterior (denominado *half-flash*) puede ser simplificado en una estructura de conversión en subrangos como se ilustra en la figura 5.17. Con este esquema, solo se utiliza un convertidor flash. Las tasas de conversión alcanzan entre 100 kHz y 40 MHz con resoluciones

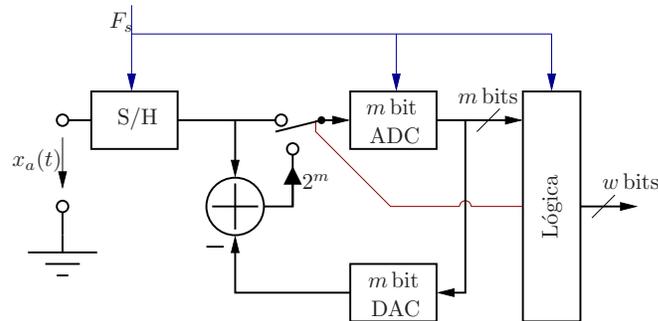


Figura 5.17: Convertidor en subrangos

de hasta 16 bits.

5.4.6 Convertidor delta-sigma

El convertidor delta-sigma ($\Delta\Sigma$) utiliza una estrategia completamente diferente a todos los circuitos ilustrados anteriormente. Los sistemas anteriores trabajan en frecuencias cercanas al límite de Nyquist ($F_s > 2B$) y por eso se les denomina convertidores ADC en tasa de Nyquist. Éstos requieren estructuras electrónicas complejas que determinan la linealidad de la conversión, y su resolución está determinada por el número de niveles en el cuantificador. Los sistemas $\Delta\Sigma$ operan por el contrario con sobre-muestreo ($F_s \gg 2B$) utilizando una resolución baja, típicamente de un único bit, por lo que se les denomina “convertidores de

un bit". Para reducir el error de cuantificación se utilizan técnicas del procesamiento digital de señales.

La figura 5.18 ilustra cualitativamente el compromiso entre ancho de banda y resolución de los diferentes tipos de conversión, donde se aprecia que a pesar de utilizar un único bit

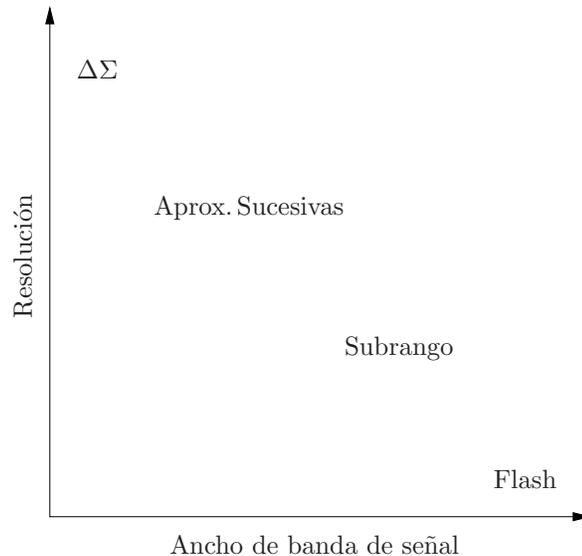


Figura 5.18: Compromiso entre ancho de banda y resolución.

para la conversión, esta estrategia es la que brinda mayor resolución. Los convertidores $\Delta\Sigma$ son utilizados por tanto en aplicaciones de voz ($B = 4\text{ kHz}$) y audio ($B = 20 - 24\text{ kHz}$) con resoluciones de hasta 18 bits. En vídeo se encuentran con más frecuencia convertidores flash para $B = 5\text{ MHz}$ y resoluciones de 8 bits.

Sobremuestreo

Se demostró anteriormente que un cuantificador de b bits introduce un ruido de cuantificación uniformemente distribuido con potencia

$$P_q = \frac{\Delta}{12} = \frac{A^2/3}{2^{2b}}$$

Es evidente que dicha potencia no cambia si b se mantiene constante, aunque la frecuencia de muestreo F_s aumente considerable sobre el límite de Nyquist $F_{sn} = 2B$. Sin embargo, al ser el ruido de cuantificación *blanco*, su energía se distribuye en todo el rango válido de frecuencias. Por otro lado, la relación de Parseval indica que la potencia en el tiempo y en la frecuencia son equivalentes, por lo que el área bajo la curva de la densidad espectral de potencia debe mantenerse constante independientemente de la frecuencia de muestreo utilizada (figura 5.19), de modo que

$$\int_{-F_{sn}/2}^{F_{sn}/2} P_{qn}(F) dF = \int_{-F_{so}/2}^{F_{so}/2} P_{qo}(F) dF = P_q$$

de donde se deriva

$$P_{qn}(F) = \frac{P_q}{F_{sn}} = \frac{P_q}{2B}$$

$$P_{qo}(F) = \frac{P_q}{F_{so}}$$

Por lo anterior, solo una fracción de la potencia total del error de cuantificación coincide con la banda de frecuencia de la señal de entrada $[-B, B]$. La potencia fuera de ese rango puede ser atenuada utilizando un filtro paso-bajos, diseñado para dejar pasar únicamente la señal de entrada. Siguiendo a este filtro, y debido a que su salida está limitada en banda, es posible sub-muestrear o *diezmar* la señal. La figura 5.20 ilustra el concepto de sobremuestreo, donde $F_{so} = 1/T_{so} = L \cdot 2B$.

La potencia del ruido de cuantificación P_{qo} correspondiente al intervalo de frecuencias $[-B, B]$ de la señal sobremuestreada se obtiene con

$$P_{qo} = \int_{-B}^B P_{qo}(F) dF = P_q \frac{2B}{F_{so}} = \frac{P_q}{L}$$

donde $L = F_{so}/F_{sn}$ y $B = F_{sn}/2$.

Si se asume que la señal de entrada tiene potencia $P_x = \sigma_x^2$ y pasa por el filtro pasa-bajos inalterada, entonces su potencia seguirá siendo P_x a la salida del filtro y la relación señal a ruido de la señal sobremuestreada y filtrada es

$$SNRQ = 10 \log_{10} \left(\frac{P_x}{P_{qo}} \right) = 10 \log_{10} \left(\frac{P_x}{P_q} \right) + 10 \log_{10} \left(\frac{F_{so}}{2B} \right) \quad (5.8)$$

donde el primer término ya fue analizado en (5.7) y el segundo término otorga mayor calidad de señal conforme mayor sea la frecuencia de sobremuestreo con respecto al ancho de banda B de la señal. Si se elige por ejemplo el muestreo de Nyquist con $F_{so} = 2B$ entonces (5.8) se reduce a (5.7). Si se elige $L = 2^r$ entonces (5.8) se puede expresar como

$$SNRQ = 10 \log_{10} \left(\frac{P_x}{P_q} \right) + 3,01r \text{ dB} \quad (5.9)$$

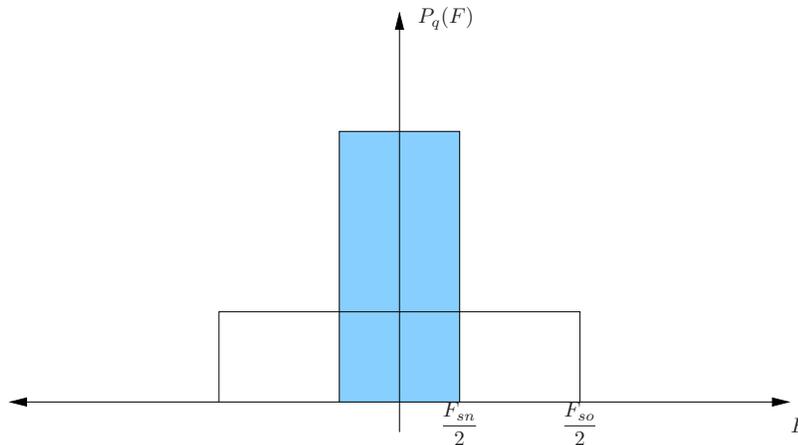


Figura 5.19: Densidad espectral de potencia del ruido de cuantificación para la tasa Nyquist F_{sn} y una tasa de sobremuestreo $F_{so} = 3F_{sn}$.

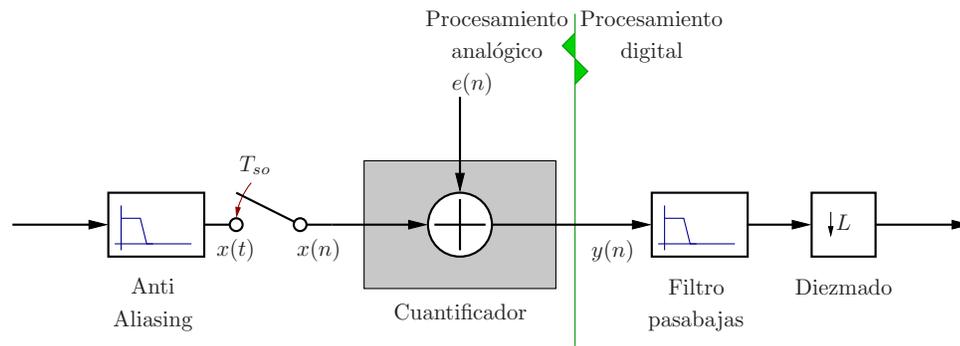


Figura 5.20: Sistema de conversión con sobremuestreo.

de donde se observa que cada duplicación de la frecuencia de muestreo equivale a una mejora de medio bit.

Si bien es cierto este esquema de sobre-muestreo permite obtener más resolución aumentando la frecuencia de muestreo, la tasa requerida para alcanzar resoluciones altas es muy elevada para las aplicaciones prácticas.

Modulación delta-sigma

El modelo del cuantificador como combinación lineal del ruido de cuantificación $e(n)$ con la señal de entrada (figura 5.20) se ha expresado matemáticamente en el dominio del tiempo discreto como

$$y(n) = x(n) + e(n)$$

que transformado al dominio z equivale a

$$Y(z) = X(z) + E(z) \quad (5.10)$$

Dicho modelo puede generalizarse utilizando subsistemas para modificar la entrada y el ruido de cuantificación de la siguiente forma:

$$Y(z) = H_x(z)X(z) + H_e(z)E(z) \quad (5.11)$$

donde $H_x(z)$ es la función de transferencia de la señal y $H_e(z)$ es la función de transferencia del ruido de cuantificación. En la sección anterior se utilizó $H_x(z) = H_e(z) = 1$, pero eligiendo otras configuraciones es posible manipular el ruido de cuantificación tal que se puedan alcanzar mejores resoluciones que con el simple sobremuestreo.

La estrategia a utilizar será entonces modificar el espectro del ruido de cuantificación de tal modo que la mayor parte de la energía se coloque fuera del ancho de banda de la señal y sea eliminado al aplicar el filtro pasobajas.

La figura 5.21 muestra un convertidor ADC delta-sigma de primer orden, que consiste en un *modulador* delta-sigma seguido por una etapa de filtrado y diezmado.

La señal a cuantificar no es la entrada $x(n)$ sino la salida de un integrador discreto cuya entrada es la diferencia entre $x(n)$ y una representación analógica $y_a(n)$ de la salida. La

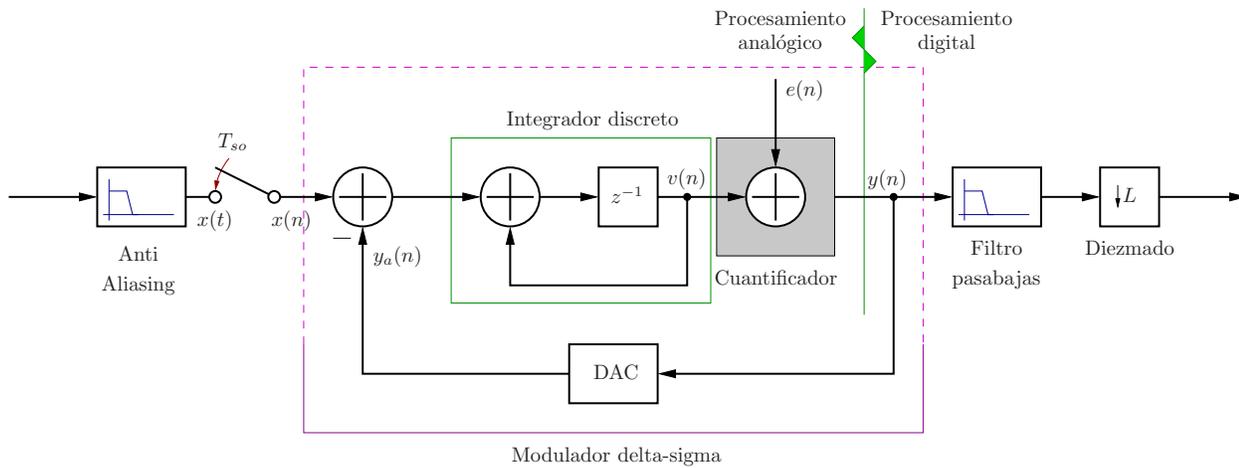


Figura 5.21: Conversión delta-sigma de primer orden.

integración se realiza con la función de transferencia $z^{-1}/(1-z^{-1})$. Dicho integrador se realiza en circuitos reales con tecnología de capacitores conmutados, que permiten implementar sistemas analógicos en tiempo discreto.

A la salida del cuantificador, usualmente de 1 bit, se cuenta con las representaciones digitales de máximo y mínimo valor representables.

Assumiendo que el convertidor DAC es ideal, su función de transferencia es unitaria, y se cumple para la salida $Y(z)$ del modulador

$$Y(z) = X(z)z^{-1} + E(z)(1 - z^{-1})$$

de modo que $H_x(z) = z^{-1}$ y $H_e(z) = 1 - z^{-1}$. La salida del modulador equivale a la entrada retrasada una muestra más el ruido modificado con un filtro pasa altos (o diferenciador discreto). La respuesta en frecuencia de $H_e(f)$ se obtiene con $z = e^{j\omega} = e^{j2\pi f}$ y se ilustra en la figura 5.22. Se observa que el ruido de cuantificación para el nivel CC es incluso cero.

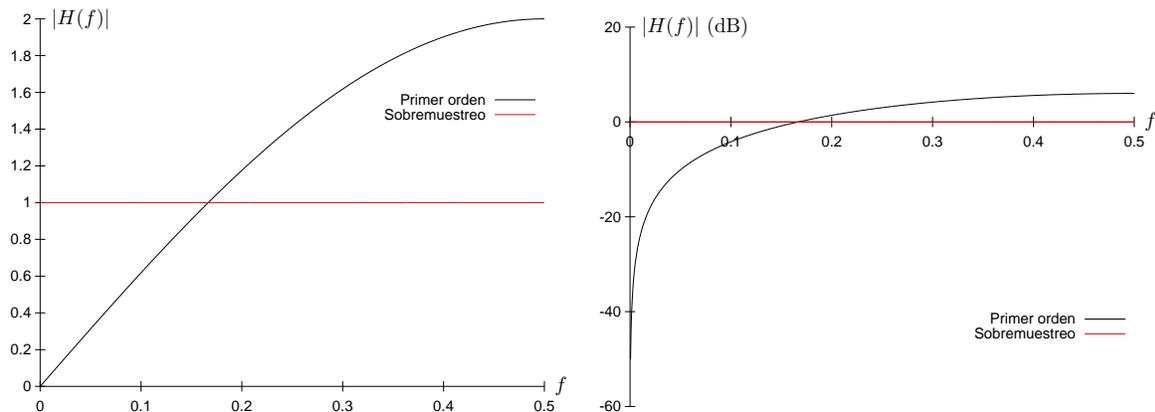


Figura 5.22: Respuesta en frecuencia de la función de transferencia de ruido $H_e(z)$ en función de la frecuencia normalizada $f = F/F_{so}$.

Para que el sistema anterior trabaje correctamente el DAC debe ser tan lineal como sea posible. Puesto que convertidores de 1 bit son perfectamente lineales, son los utilizados

preferentemente en este tipo de circuitos, con la ventaja adicional que su implementación consiste en un comparador, que se acopla directamente a la cuantificación de 1 bit generalmente utilizada.

Utilizando la relación de Parseval se obtiene que la potencia del ruido de cuantificación luego de pasar por el filtro $H_e(z)$ y por el filtro pasabajas estará dada por

$$\begin{aligned} P_e &= \int_{-B}^B P_{qo}(F) |H_e(F)|^2 dF \\ &= \frac{P_q}{F_{so}} \int_{-B}^B |H_e(F)|^2 dF = 2 \frac{P_q}{F_{so}} \int_0^B |H_e(F)|^2 dF \\ &= 2 \frac{P_q}{F_{so}} \int_0^B |1 - e^{j2\pi F/F_{so}}|^2 dF \\ &= 2 \frac{P_q}{F_{so}} \int_0^B 4 \operatorname{sen}^2 \left(\frac{\pi F}{F_{so}} \right) dF \end{aligned}$$

y asumiendo que $F_{so} \gg B$ y que $\operatorname{sen}(x) \approx x$ para $x \approx 0$ se tiene

$$P_e = 8 \frac{P_q}{F_{so}} \int_0^B \left(\frac{\pi F}{F_{so}} \right)^2 dF = P_q \frac{\pi^2}{3} \left(\frac{2B}{F_{so}} \right)^3$$

En decibelios

$$\begin{aligned} SNRQ(\text{dB}) &= 10 \log_{10} P_x - 10 \log_{10} P_q - 10 \log_{10} \frac{\pi^2}{3} + 30 \log_{10} \frac{F_{so}}{2B} \\ &= 10 \log_{10} P_x - 10 \log_{10} P_q - 5,17 \text{ dB} + 30 \log_{10} L \end{aligned}$$

y si $L = 2^r$ entonces

$$SNRQ(\text{dB}) = 10 \log_{10} P_x - 10 \log_{10} P_q - 5,17 \text{ dB} + 9,03r \text{ dB}$$

lo que equivale a 1,5 bits por cada duplicación de la frecuencia de sobremuestreo.

A pesar de que ha mejorado con este modulador la ganancia obtenida por cada duplicación de frecuencia, no es aún suficiente para aplicaciones reales. Por ello, se utilizan moduladores de mayor orden, como el ilustrado en la figura 5.23. El lector puede demostrar que la función de transferencia para el ruido es en este caso $H_e(z) = (1 - z^{-1})^2$. Siguiendo un desarrollo similar al anterior se obtiene que

$$\begin{aligned} SNRQ(\text{dB}) &= 10 \log_{10} P_x - 10 \log_{10} P_q - 10 \log_{10} \frac{\pi^4}{5} + 50 \log_{10} \frac{F_{so}}{2B} \\ &= 10 \log_{10} P_x - 10 \log_{10} P_q - 12,89 \text{ dB} + 50 \log_{10} L \end{aligned}$$

y con $L = 2^r$

$$SNRQ(\text{dB}) = 10 \log_{10} P_x - 10 \log_{10} P_q - 12,89 \text{ dB} + 15,05r \text{ dB}$$

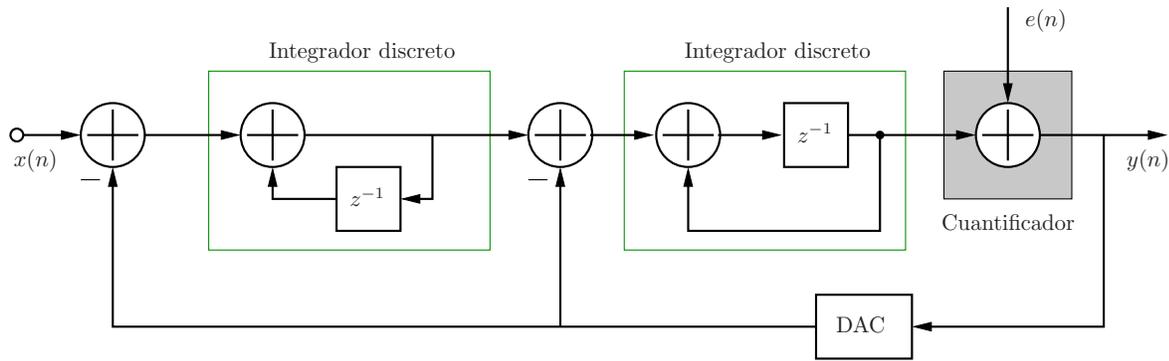


Figura 5.23: Modulador delta-sigma de segundo orden.

lo que implica que con este modulador se obtienen 2,5 bits por cada duplicación de la frecuencia de muestreo.

La figura 5.24 presenta la respuesta en frecuencia de la función de transferencia del error de cuantificación para el sistema de sobremuestreo, y los sistemas de primer, segundo y tercer orden de modulación delta-sigma. Se aprecia como la atenuación energética en bajas frecuencias es compensada con una amplificación a altas frecuencias.

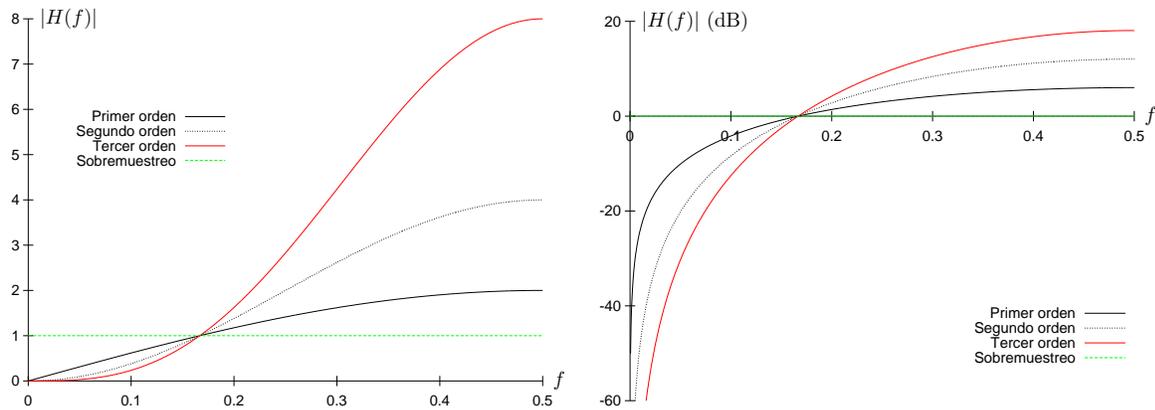


Figura 5.24: Respuesta en frecuencia de la función de transferencia de ruido $H_e(z)$ en función de la frecuencia normalizada $f = F/F_{so}$ para tres órdenes de modulación delta-sigma.

5.5 Conversión Digital/Analógica

Los principios utilizados para la conversión de una representación digital a una señal analógica se basan en técnicas de conversión directas, lo que permite altas velocidades de conversión.

En secciones anteriores se determinó que en la conversión ideal de una señal digital a una analógica interviene el interpolador ideal $sa(t) = \text{sen}(\pi t)/\pi t$. Este interpolador tiene la propiedad de reconstruir a la perfección señales con un espectro limitado en banda muestreadas con una tasa superior a dos veces dicho ancho de banda; sin embargo, la no causalidad y la infinita extensión de dicho interpolador obliga a utilizar simplificaciones, como el retenedor de orden cero presentado anteriormente (circuito S/H).

5.5.1 Fuentes de tensión conmutadas

La conversión con fuentes de tensión conmutadas se ilustra en la figura 5.25. Ésta se realiza

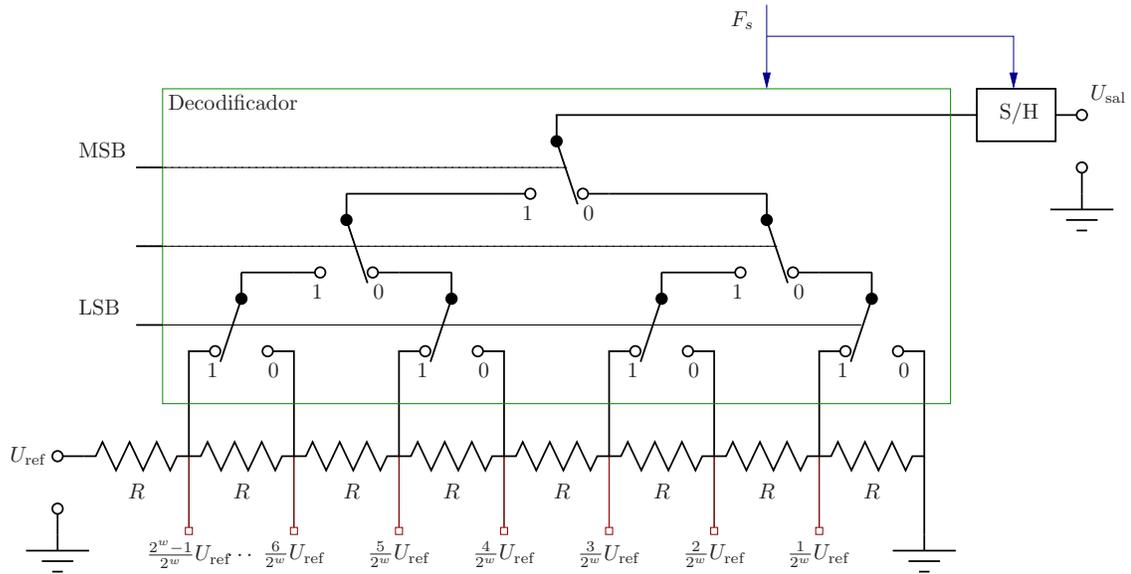


Figura 5.25: Fuentes de tensión conmutadas

utilizando una tensión de referencia U_{ref} conectada a una red resistiva de 2^w resistencias de igual valor y conmutada por etapas a través de un decodificador que logra llevar a la salida del retenedor la tensión equivalente.

Una obvia desventaja de esta estrategia está en el tamaño de la red resistiva, y del decodificador, que crecen exponencialmente con el número de bits a convertir.

5.5.2 Resistencias conmutadas

La figura 5.26 ilustra otra estrategia de conversión basada en la conmutación de corrientes, donde se observa que la corriente correspondiente al bit b_i es siempre el doble de la corriente del bit b_{i-1} . Sea U_1 la tensión a la salida del amplificador operacional. Ésta tensión se obtiene con

$$\begin{aligned} U_1 = RI &= R \left(\frac{b_0}{2^w R} + \frac{b_1}{2^{w-1} R} + \dots + \frac{b_{w-1}}{2R} \right) U_{\text{ref}} \\ &= (b_0 2^{-w} + b_1 2^{-w+1} + \dots + b_{w-1} 2^{-1}) U_{\text{ref}} \end{aligned}$$

Si bien es cierto esta estrategia utiliza menos resistencias que las fuentes de tensión conmutadas, su exacta calibración es un reto en el diseño del circuito integrado. Otra desventaja es claramente el consumo de potencia presente en todas las resistencias. Si bien es cierto éste puede reducirse utilizando valores mayores de R , usualmente esto irá acompañado de una reducción en la razón señal a ruido.

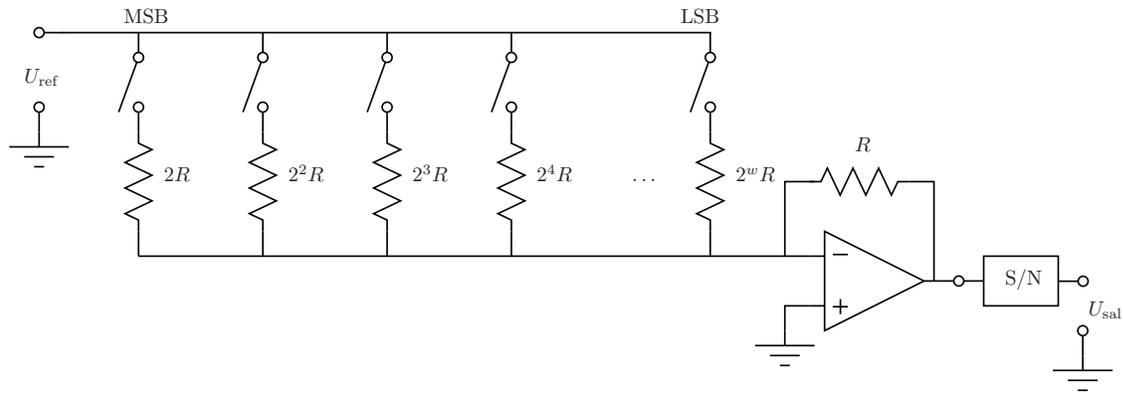


Figura 5.26: Conversión digital-analógica con resistores ponderados

5.5.3 Condensadores conmutados

La conversión digital analógica con condensadores conmutados se muestra en la figura 5.27. En este caso la conversión requiere varios pasos. En el primer paso (denotado con 1 en

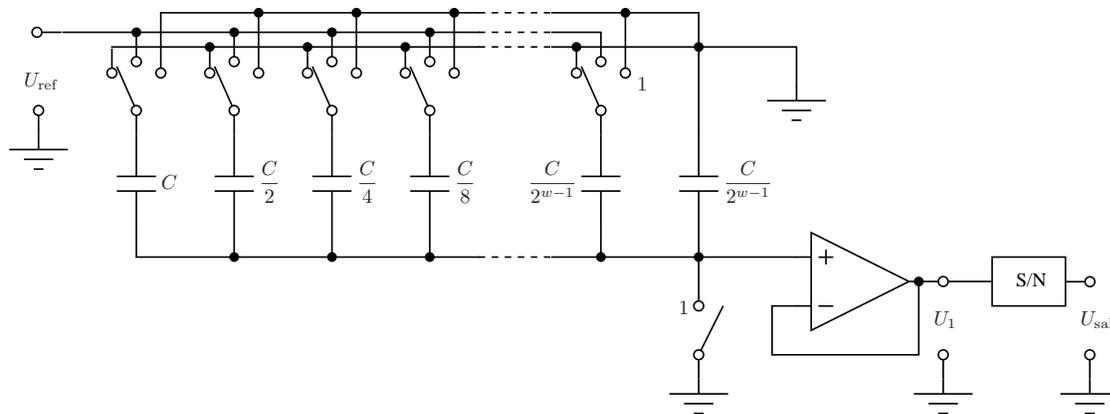


Figura 5.27: Conversión digital-analógica con condensadores ponderados

la figura), todos los condensadores se descargan conectando los bornes de todos ellos a referencia. En el segundo paso, el terminal positivo del seguidor de tensión se desconecta de tierra y cada uno de los bits con valor 1 se conectan a la tensión de referencia U_{ref} , mientras que los bits 0 permanecen conectados a la tierra. La carga acumulada es entonces

$$U_{\text{ref}}Q = \left(b_{w-1}C + b_{w-2}\frac{C}{2} + \dots + b_0\frac{C}{2^{w-1}} \right)$$

En el último paso de la conversión, la carga Q acumulada en los condensadores correspondientes a bits 1 es repartida entre la capacitancia total del circuito conmutando los interruptores a la referencia, como se ilustra en la figura 5.27. Dicha capacitancia es $2C$ gracias al condensador terminal de valor $C/2^{w-1}$, por lo que

$$Q = 2CU_1$$

con U_1 la tensión a la salida del seguidor de tensión. Combinando las dos ecuaciones anteriores se deriva

$$U_1 = (b_{w-1}2^{-1} + b_{w-2}2^{-2} + \dots + b_12^{-w+1} + b_02^{-w}) U_{\text{ref}}$$

5.5.4 Redes resistivas $R - 2R$

Otra manera de lograr conmutación de corrientes es a través de redes resistivas $R - 2R$ como la ilustrada en la figura 5.28. A diferencia de la estructura de resistencias conmutadas, solo se requieren aquí dos valores de resistencias con una razón de valor 2:1. Obsérvese que en

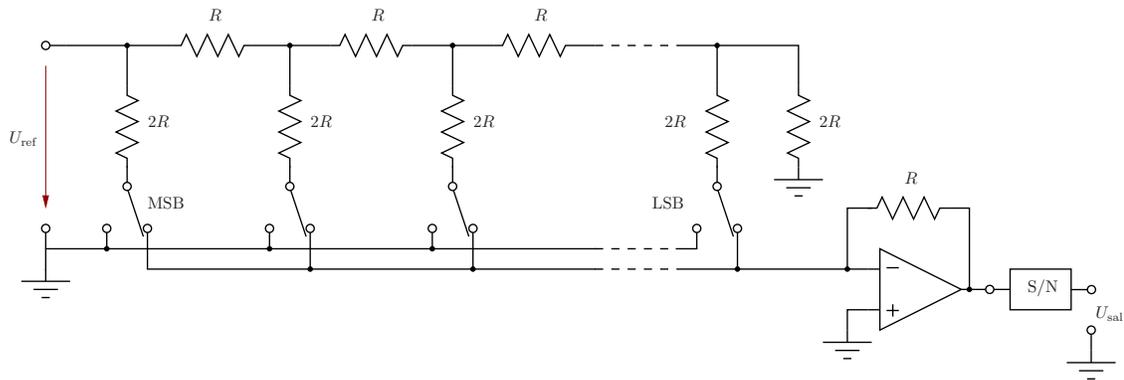


Figura 5.28: Conversión digital-analógica con redes $R-2R$

cada nodo la corriente que entra divide en dos tantos iguales, de modo que la tensión U_1 a la salida del amplificador operacional es

$$\begin{aligned} U_1 &= -RI = -R \left(\frac{b_{w-1}}{2R} + \frac{b_{w-2}}{4R} + \dots + \frac{b_1}{2^{w-1}R} + \frac{b_0}{2^wR} \right) U_{\text{ref}} \\ &= -U_{\text{ref}} (b_{w-1}2^{-1} + b_{w-2}2^{-2} + \dots + b_12^{-w+1} + b_02^{-w}) \end{aligned}$$

5.5.5 Conversión DAC delta-sigma

Los principios de modulación y conversión delta sigma se pueden aplicar en sentido inverso, como lo muestra la figura 5.29. En este caso se traslada la señal digital con códigos de w bits a una secuencia con LF_s de sobremuestreo de bits simples. Un filtro pasabajos analógico recupera la señal analógica.

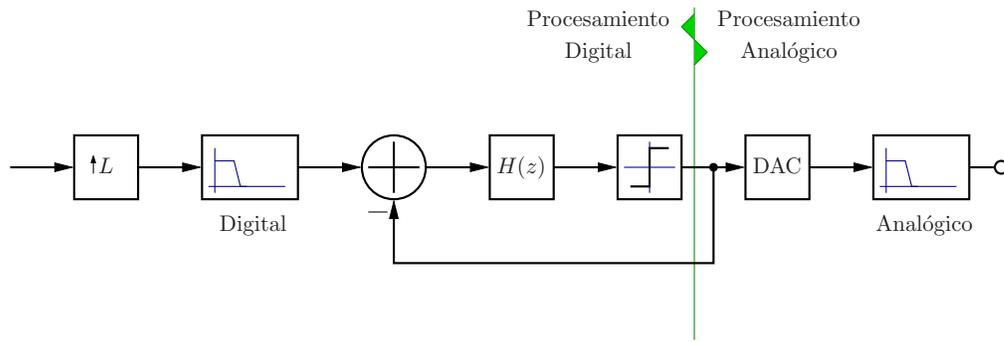


Figura 5.29: Conversión DAC delta-sigma.

5.6 Problemas

Problema 5.1. Investigue qué circuitos se utilizan para realizar la conversión de una señal digital a una analógica.

Capítulo 6

Implementación de sistemas discretos

Hasta ahora se ha visto el análisis de sistemas LTI descritos mediante ecuaciones de diferencias lineales de coeficientes constantes y el análisis por convolución, en el dominio del tiempo (o de variable) discreto y en el dominio de la frecuencia.

En la práctica, el diseño y la implementación de sistemas digitales se tratan conjuntamente, y se deben considerar aspectos de costo, tamaño, hardware, potencia, etc. Algunos esquemas básicos de implementación se revisarán aquí.

6.1 Número de Condición

Los sistemas de procesamiento digital son implementados en computadores digitales que utilizan representaciones numéricas de precisión limitada. Estas limitaciones afectan el desempeño de los sistemas y obliga a revisar con detalle su comportamiento numérico.

En el área de análisis numérico, el *número de condición* de una función con respecto a un argumento indica, para el peor caso posible, cuánto puede cambiar el valor de la función con respecto a un cambio dado del argumento. Se dice entonces que un problema es *mal condicionado* (en inglés *ill-conditioned*) si el número de condición es alto, es decir, si una pequeña perturbación del argumento produce cambios grandes en la salida. Si el número de condición es pequeño (cercano a uno) entonces se habla de un problema bien condicionado (en inglés, *well-conditioned*).

El número de condición es una propiedad de cada problema particular, y por tanto un problema puede ser inherentemente bien condicionado o mal condicionado. Por ejemplo, el pronóstico del tiempo es un problema inherentemente mal condicionado al depender no solo de un elevado número de variables en tiempo y espacio (presión atmosférica, humedad relativa, temperatura, velocidades de viento, etc.), sino que el resultado es sensible a cambios leves en los valores de dichas variables.

La solución de un sistema de ecuaciones lineales expresado como el sistema matricial

$$\mathbf{Ax} = \mathbf{b} \tag{6.1}$$

donde \mathbf{A} es una matriz cuadrada de $N \times N$ y $\underline{\mathbf{x}}$ y $\underline{\mathbf{b}}$ son vectores N -dimensionales, tendrá un número de condición dependiente totalmente de la matriz \mathbf{A} que determina cuánto afecta un error en los valores $\underline{\mathbf{b}}$ los valores encontrados para $\underline{\mathbf{x}}$. Si la determinante de \mathbf{A} es un número pequeño, entonces el problema es mal condicionado.

No solo el problema es bien o mal condicionado, sino también el algoritmo particular desarrollado para dar solución al problema. Si un problema es mal condicionado, ningún algoritmo podrá cambiar esa condición. El cuidado debe tenerse en que un algoritmo mal condicionado puede producir errores considerables en la solución de problemas inherentemente bien condicionados. Para el caso de algoritmos, el número de condición se redefine entonces como la razón de salida/entrada pero utilizando el algoritmo concreto. Por ejemplo, en la solución del sistema lineal de ecuaciones (6.1) pueden utilizarse algoritmos como la eliminación gaussiana (mal condicionado), la eliminación de Gauss-Jordan, descomposición de valores singulares (SVD), o la descomposición QR (bien condicionados), entre otros.

El problema que compete directamente en éste capítulo es encontrar las raíces z_i de un polinomio expresado a través de sus coeficientes c_i :

$$P(z) = \sum_{i=0}^N c_i z^i = c_0 + c_1 z + c_2 z^2 + \dots + c_N z^N \quad (6.2)$$

es decir, encontrar los valores z_i para los que se cumple $P(z_i) = 0$, para poder expresar al polinomio de la forma

$$P(z) = \prod_{i=1}^N (z - z_i) = (z - z_1)(z - z_2) \dots (z - z_N) \quad (6.3)$$

Este problema es en general mal condicionado, en el sentido de que pequeños cambios en los valores de c_i pueden producir grandes cambios en la posición de las raíces z_i , hecho que empeora conforme aumenta el orden del polinomio. Las consecuencias del mal condicionamiento de éste problema en el procesamiento digital tienen que ver con el posicionamiento de polos y ceros en la implementación de un sistema utilizando ecuaciones de diferencias que tratan ya sea (6.2) o (6.3).

6.2 Estructuras directas

Considérese el sistema de primer orden:

$$y(n) = -a_1 y(n-1) + b_0 x(n) + b_1 x(n-1)$$

y su “realización directa” (figura 6.1).

que se puede interpretar como la serie de dos sistemas LTI:

$$v(n) = b_0 x(n) + b_1 x(n-1)$$

y

$$y(n) = -a_1 y(n-1) + v(n).$$

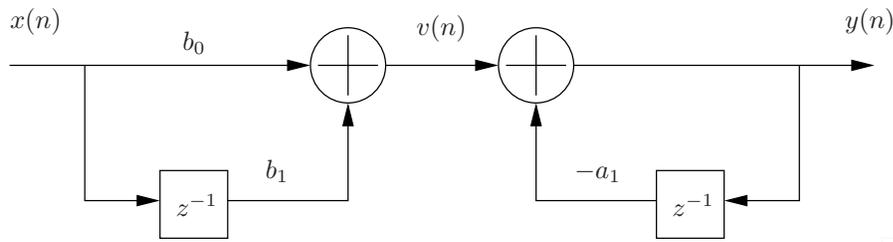


Figura 6.1: Realización directa del sistema de primer orden $y(n) = -a_1y(n-1) + b_0x(n) + b_1x(n-1)$.

Como la conexión en serie de sistemas LTI es conmutativa, se puede reemplazar el sistema anterior por el indicado en la figura 6.2, donde se cumplen las ecuaciones:

$$\begin{aligned}\omega(n) &= x(n) - a_1\omega(n-1) \\ y(n) &= b_0\omega(n) + b_1\omega(n-1)\end{aligned}$$

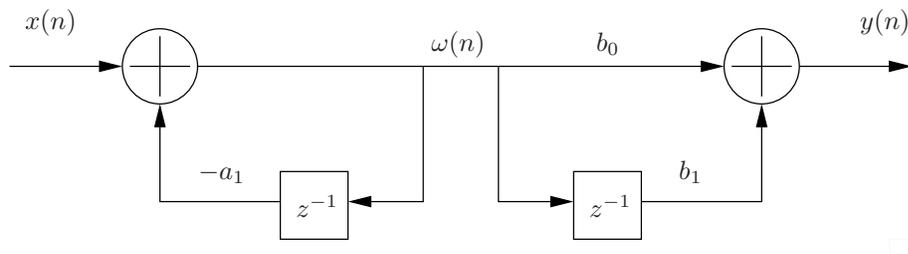


Figura 6.2: Reemplazo del sistema de la figura 6.1.

Adicionalmente, se observa que los dos retardadores tienen la misma entrada $\omega(n)$ y por lo tanto la misma salida $\omega(n-1)$, y se pueden agrupar en uno solo según se muestra en la figura 6.3.

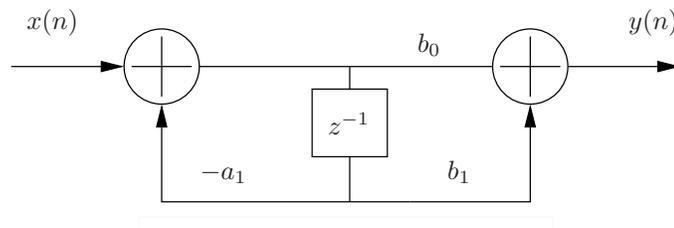


Figura 6.3: Simplificación del sistema de la figura 6.2.

A esta estructura se le denomina *forma directa II* y utiliza sólo un elemento retardador, en lugar de los dos utilizados en la forma directa I (figura 6.1).

Nótese que esto puede generalizarse para la ecuación de diferencias:

$$y(n) = -\sum_{k=1}^N a_k y(n-k) + \sum_{k=0}^M b_k x(n-k) \quad (6.4)$$

compuesta de un sistema no recursivo:

$$v(n) = \sum_{k=0}^M b_k x(n-k)$$

y un sistema recursivo:

$$y(n) = - \sum_{k=1}^N a_k y(n-k) + v(n)$$

mostrados en la figura 6.4. Invirtiendo el orden y combinando los retardadores se obtiene,

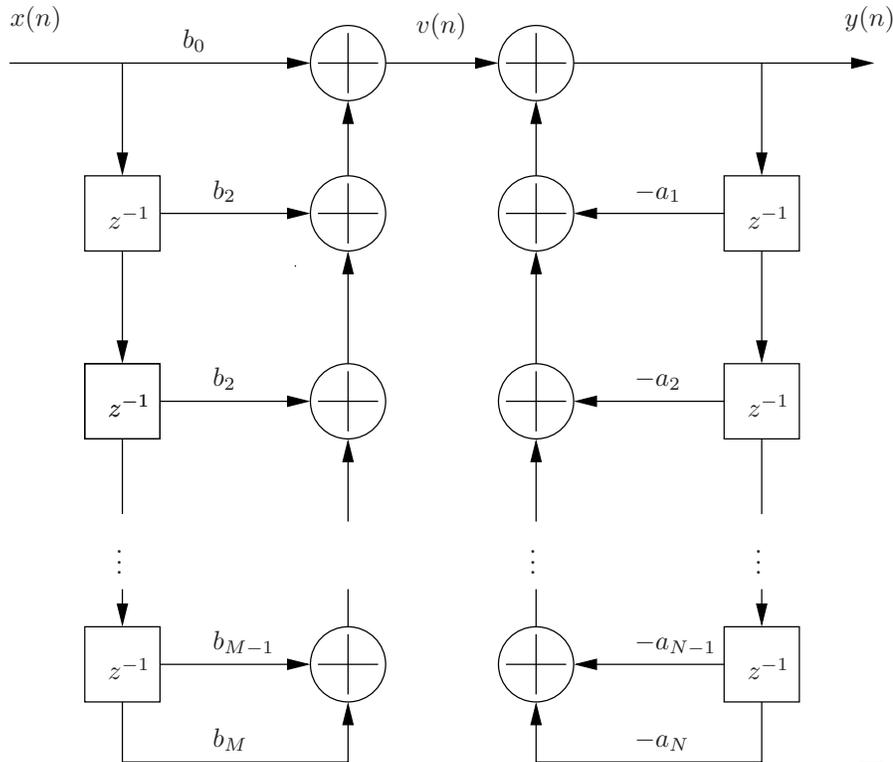


Figura 6.4: Forma directa I.

con $M < N$, el diagrama mostrado en la figura 6.5. Nótese que esta estructura requiere $\max\{N, M\}$ retardadores y $M + N + 1$ multiplicaciones, que contrastan con los $N + M$ retardadores de la forma directa, aunque el número de multiplicaciones es el mismo. A esta forma II, por poseer el mínimo de retardadores, se le denomina también *forma canónica*.

Al caso especial no recursivo con $a_k = 0, k = 1 \dots N$:

$$y(n) = \sum_{k=0}^M b_k x(n-k)$$

se le denomina *sistema de media ponderada móvil* o sistema MWA (*moving weighted average*), que es un sistema FIR de respuesta impulsional:

$$h(k) = \begin{cases} b_k, & 0 \leq k \leq M \\ 0, & \text{en otro caso} \end{cases}$$

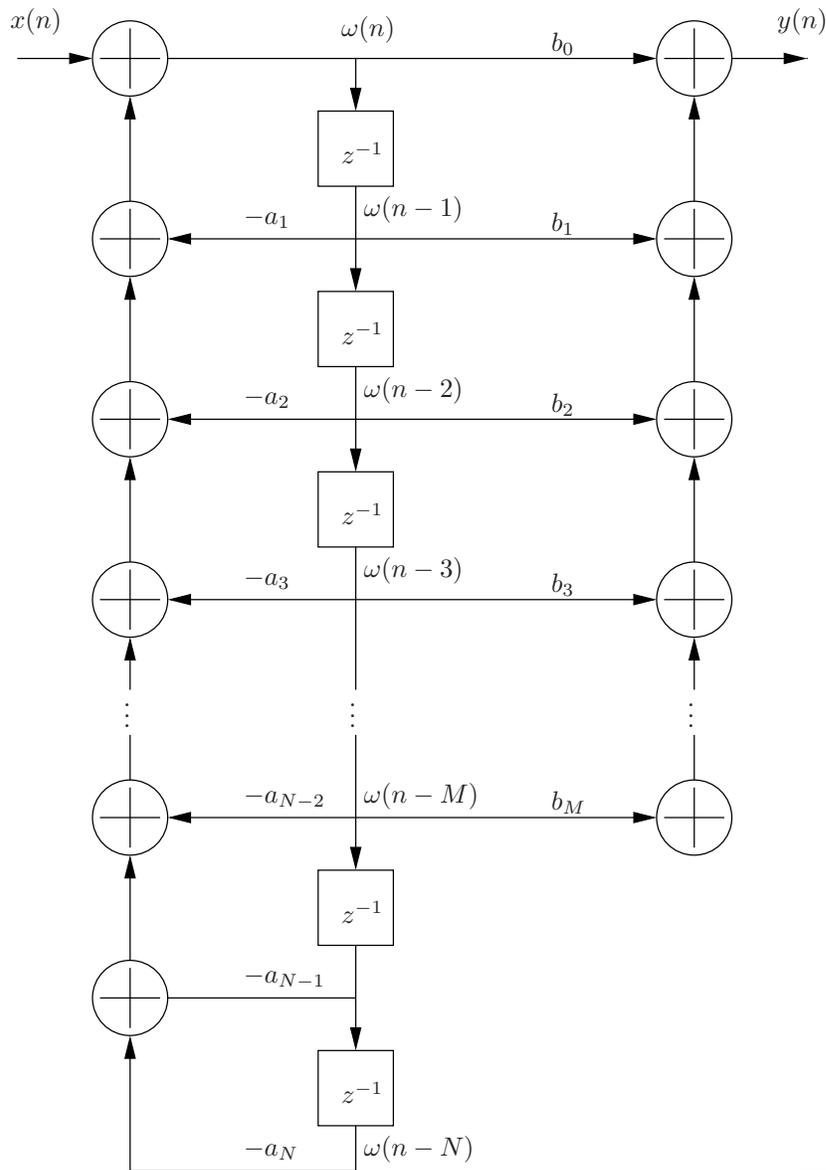


Figura 6.5: Forma directa II.

Con $M = 0$, el sistema de la ecuación (6.4) se torna puramente recursivo:

$$y(n) = - \sum_{k=1}^N a_k y(n-k) + b_0 x(n)$$

y calcula la combinación lineal de las últimas N salidas y la entrada actual.

6.2.1 Estructuras directas para filtros FIR simétricos

El caso particular de filtros FIR de longitud M cuyos coeficientes presentan condiciones de simetría o antisimetría

$$h(n) = \pm h(M-1-n)$$

tiene aplicaciones particularmente en el caso de filtros de fase lineal. Este tipo de simetría permite reducir el número de productos requeridos de una implementación convencional de M a $M/2$ o a $(M - 1)/2$ para M par o impar respectivamente.

Las figuras 6.6 y 6.7 ilustran las estructuras en forma directa para estos filtros.

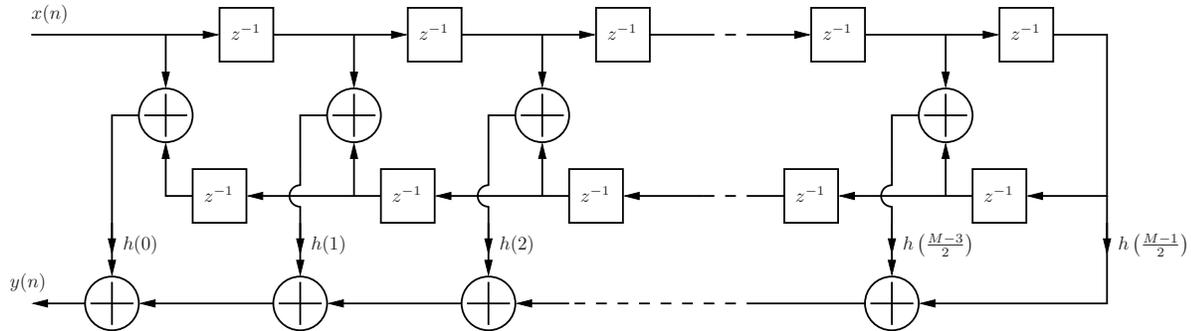


Figura 6.6: Filtro FIR simétrico con M impar.

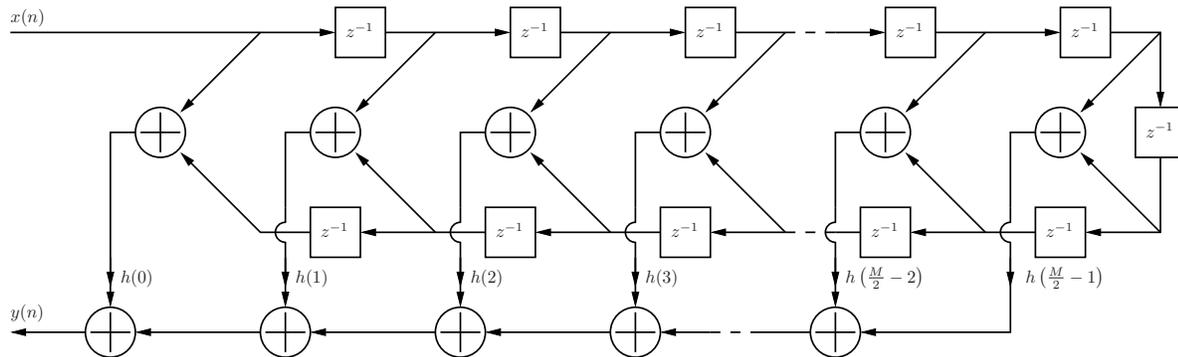


Figura 6.7: Filtro FIR simétrico con M par.

6.3 Grafos de flujo de señal y estructuras transpuestas

Una representación alternativa de los diagramas de bloques utilizados hasta ahora son los diagramas de flujo de señal, caracterizados por ramas con transmitancias específicas y nodos en donde convergen y se distribuyen las señales. Los nodos suman todas las señales que convergen en ellos y distribuyen la misma señal en todas las ramas que salen de ellos. Un nodo que solo distribuye la misma señal (es decir, un nodo en el que no converge ninguna señal) recibe el nombre de *nodo fuente*, y uno a donde solo convergen señales (es decir, no sale ninguna señal) se denomina *nodo sumidero*. La figura 6.8 ilustra el grafo de flujo de señal correspondiente al diagrama de bloques de un sistema de segundo orden ilustrado en la parte superior, en donde se muestra que $y(n)$ llega a un nodo sumidero y $x(n)$ parte de un nodo fuente.

La representación de sistemas con una topología de grafos permite aplicar principios como el llamado *teorema de transposición*, que especifica una forma de transformar el grafo de modo

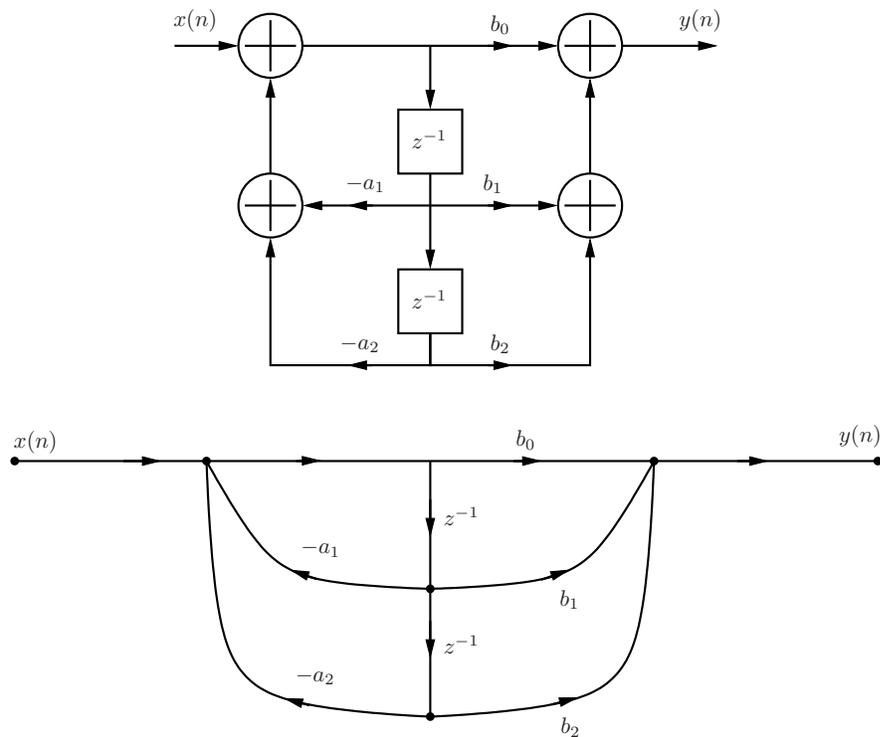


Figura 6.8: Diagrama de bloques de sistema de segundo orden y su correspondiente grafo de flujo de señal.

que la relación entre entrada y salida no cambie. El teorema de transposición especifica que si se invierten las direcciones de todas las transmitancias de rama z se intercambia la entrada con la salida en el grafo, la función de transferencia del sistema no cambia. La transposición del sistema de segundo orden de la figura 6.8 se ilustra en la figura 6.9.

Obsérvese que en la transposición los nodos se convierten en sumadores y los sumadores en nodos.

Ejemplo 6.1 Encuentre las formas directas I y II y la forma transpuesta II para el sistema

$$H(z) = \frac{b_0 + b_1 z^{-1} + b_2 z^{-2}}{1 + a_1 z^{-1} + a_2 z^{-2}}$$

Solución:

Por hacer:

Solución. Tabla 9.1 PM

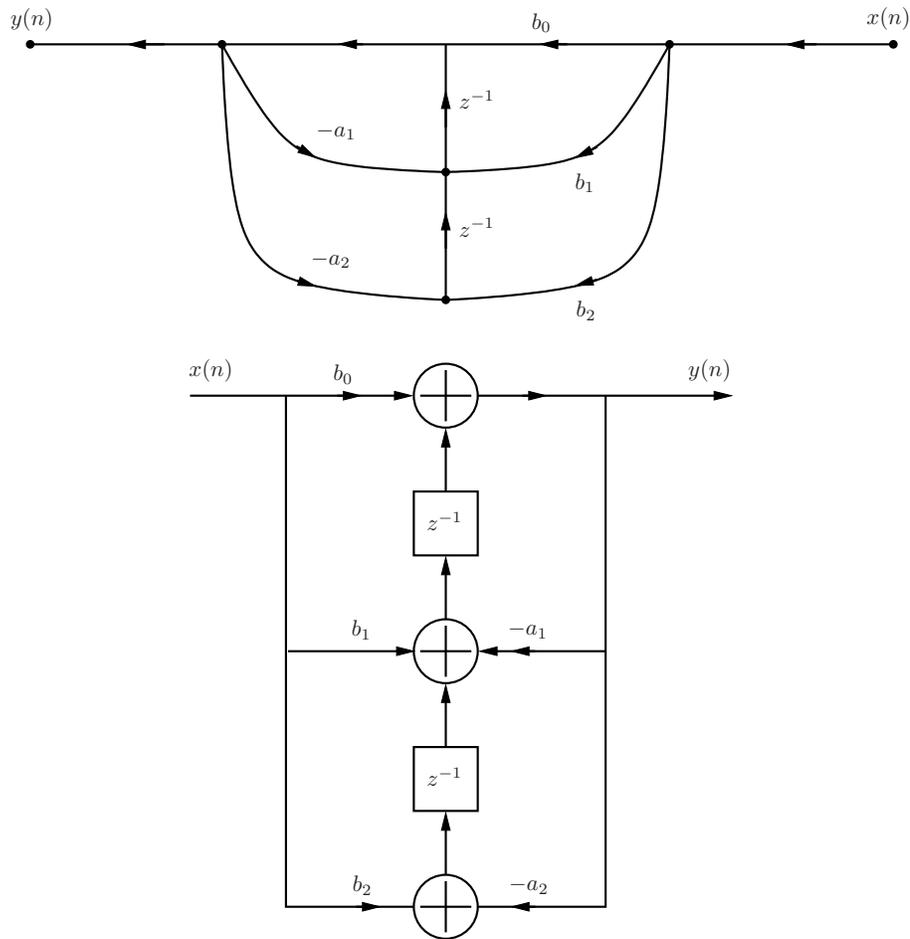


Figura 6.9: Grafo de flujo de señal transpuesto y el diagrama de bloques derivado.

6.4 Muestreo en frecuencia

Dado un conjunto de frecuencias ω_k espaciadas de forma homogénea

$$\omega_k = \frac{2\pi}{M}(k + \alpha), \quad k = 0, 1, \dots, \frac{M-1}{2}, \quad M \text{ impar}$$

$$k = 0, 1, \dots, \frac{M}{2} - 1, \quad M \text{ par}$$

$$\alpha = 0 \text{ ó } \frac{1}{2}$$

Se sabe que la respuesta en frecuencia $H(\omega)$ está dada por la transformada de Fourier en tiempo discreto de la respuesta al impulso $h(n)$:

$$H(\omega) = \sum_{n=0}^{M-1} h(n)e^{j\omega n}$$

y los valores de $H(\omega)$ en las frecuencias ω_k son entonces

$$H(k + \alpha) \stackrel{!}{=} H\left(\frac{2\pi}{M}(k + \alpha)\right) = \sum_{n=0}^{M-1} h(n)e^{-j2\pi(k+\alpha)n/M}, \quad k = 1, 2, \dots, M-1$$

Obsérvese que si $\alpha = 0$ entonces la ecuación anterior corresponde a la DFT de M puntos de $h(n)$. Utilizando los mismos principios de derivación de la IDFT se despeja:

$$h(n) = \frac{1}{M} \sum_{k=0}^{M-1} H(k + \alpha) e^{j2\pi(k+\alpha)n/M}, \quad k = 1, 2, \dots, M - 1 \quad (6.5)$$

que equivale a la IDFT si $\alpha = 0$.

Se sabe que la función de transferencia del filtro está dada por

$$H(z) = \sum_{n=0}^{M-1} h(n) z^{-n}$$

y utilizando (6.5) se obtiene

$$H(z) = \sum_{n=0}^{M-1} \left[\frac{1}{M} \sum_{k=0}^{M-1} H(k + \alpha) e^{j2\pi(k+\alpha)n/M} \right] z^{-n}$$

e intercambiando el orden de las sumatorias

$$\begin{aligned} H(z) &= \sum_{k=0}^{M-1} H(k + \alpha) \left[\frac{1}{M} \sum_{n=0}^{M-1} (e^{j2\pi(k+\alpha)/M} z^{-1})^n \right] \\ &= \frac{1 - z^{-M} e^{j2\pi\alpha}}{M} \sum_{k=0}^{M-1} \frac{H(k + \alpha)}{1 - e^{j2\pi(k+\alpha)/M} z^{-1}} \end{aligned}$$

que está expresada enteramente por las muestras del espectro $H(\omega)$. La anterior ecuación se interpreta como la cascada de dos sistemas: un filtro todo ceros del tipo peine caracterizado por

$$H_1(z) = \frac{1}{M} (1 - z^{-M} e^{j2\pi\alpha})$$

donde los ceros se encuentran equiespaciados sobre la circunferencia unitaria en

$$z_k = e^{j2\pi(k+\alpha)/M}, \quad k = 0, 1, \dots, M - 1$$

El otro elemento de la cascada es un banco de filtros de primer orden, cada uno con un solo polo en

$$p_k = e^{j2\pi(k+\alpha)/M}, \quad k = 0, 1, \dots, M - 1$$

Obsérvese que los polos coinciden en su posición con los ceros.

La figura 6.10 muestra la realización de un filtro FIR por muestreo en frecuencia.

Cuando la respuesta en frecuencia deseada es de banda angosta, la mayor parte de los coeficientes $H(\omega_k)$ son cero, y la estructura se simplifica. Además, si se requiere una respuesta impulsional real, entonces debido a la simetría hermítica de $H(\omega)$ la estructura puede simplificarse aun más combinando pares de filtros de primer orden en filtros de segundo orden.

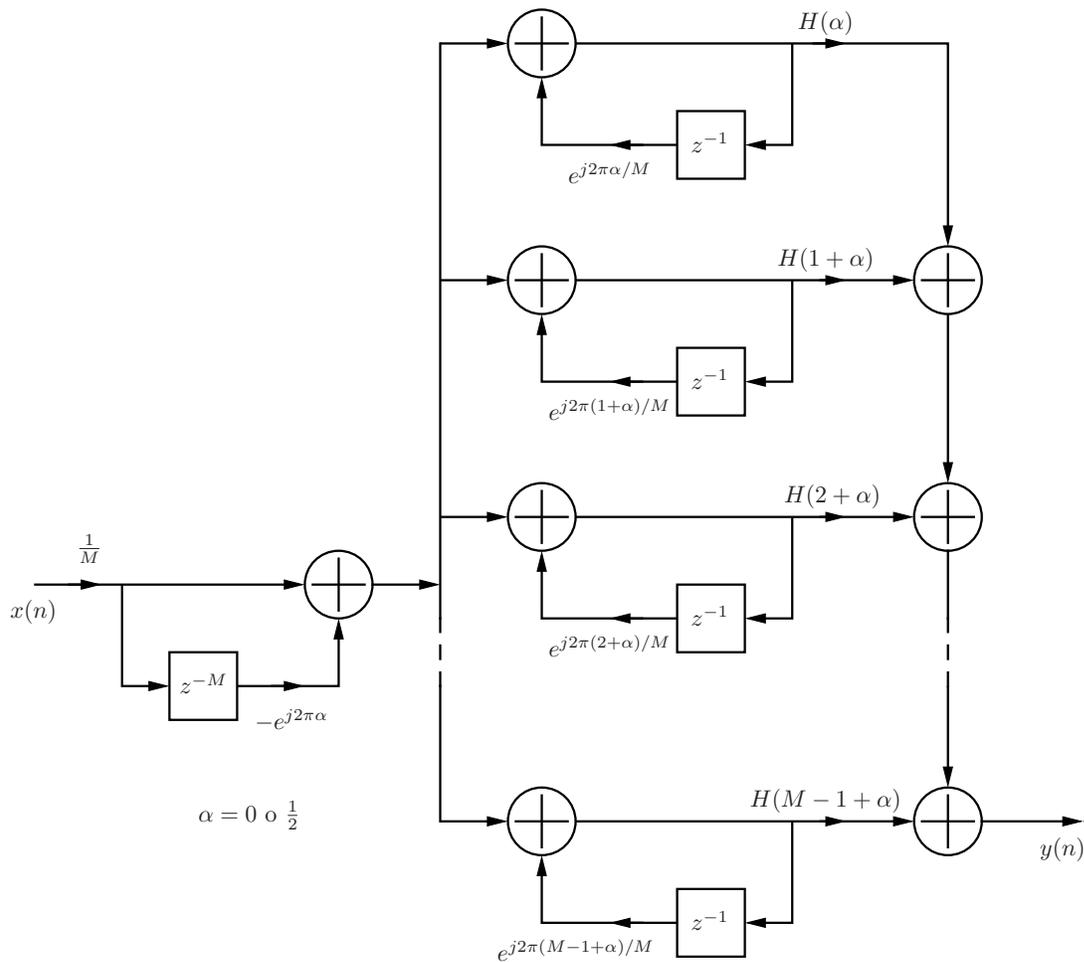


Figura 6.10: Realización de filtro FIR por muestreo en frecuencia

6.5 Sistemas en cascada

Las formas directas I y II utilizan los coeficientes de los polinomios en el numerador y denominador de la función de transferencia racional que implementan. Puesto que el posicionamiento de las raíces polinomiales es un problema mal condicionado, la ubicación de los polos y ceros, considerando las precisiones finitas con las que los coeficientes pueden ser representados, pueden variar considerablemente, convirtiendo por ejemplo sistemas estables en inestables.

Los sistemas en cascada persiguen una utilización directa de los valores de polos y ceros, que hacen más predecible el efecto de cuantificación en sus representaciones digitales.

Para ello se parte de una factorización de los polinomios en el numerador y denominador de la función de transferencia del sistema

$$H(z) = \frac{\sum_{k=0}^M b_k z^{-k}}{\sum_{k=0}^N a_k z^{-k}} = A \frac{\prod_{k=1}^{M_1} (1 - f_k z^{-1}) \prod_{k=1}^{M_2} (1 - g_k z^{-1})(1 - g_k^* z^{-1})}{\prod_{k=1}^{N_1} (1 - c_k z^{-1}) \prod_{k=1}^{N_2} (1 - d_k z^{-1})(1 - d_k^* z^{-1})}$$

donde $M = M_1 + 2M_2$ y $N = N_1 + 2N_2$. Si a_k y b_k son reales entonces f_k y c_k también lo son, mientras que g_k y d_k son complejos y aparecen junto a sus pares complejos conjugados

g_k^* y d_k^* .

Usualmente se combinan pares de factores reales o complejos conjugados en estructuras de segundo orden, de modo que

$$H(z) = \prod_{k=1}^{N_s} \frac{b_{0k} + b_{1k}z^{-1} + b_{2k}z^{-2}}{1 - a_{1k}z^{-1} + a_{2k}z^{-2}} \quad (6.6)$$

con $N_s = \lfloor (N + 1)/2 \rfloor$. Aquí se ha asumido que $M \leq N$. En caso de que hubiese un número impar de ceros (o polos) reales, entonces uno de los términos b_{2k} (o a_{2k}) será cero. Para cada uno de estos términos puede utilizarse la forma directa II, aunque se acostumbra utilizar una reducción en el número de multiplicadores necesarios utilizando la factorización

$$H(z) = b_0 \prod_{k=1}^{N_s} \frac{1 + \tilde{b}_{1k}z^{-1} + \tilde{b}_{2k}z^{-2}}{1 - a_{1k}z^{-1} + a_{2k}z^{-2}}$$

La forma (6.6) tiene sin embargo la ventaja sobre la forma (6.5) de que el valor de b_0 queda distribuido en todos los términos de la cascada, lo que es conveniente en implementaciones en punto fijo.

Por hacer:

Figura de estructura en cascada

6.6 Sistemas paralelos

De forma alternativa a la factorización en términos de segundo orden, utilizando descomposición en fracciones parciales es posible expresar la función de transferencia como

$$H(z) = \sum_{k=0}^{N_p} C_k z^{-k} + \sum_{k=1}^{N_1} \frac{A_k}{1 - c_k z^{-1}} + \sum_{k=1}^{N_2} \frac{B_k(1 - e_k z^{-1})}{(1 - d_k z^{-1})(1 - d_k^* z^{-1})}$$

donde $N = N_1 + 2N_2$. El primer término está presente solo si $H(z)$ es una función racional impropia, es decir, si $M \geq N$ en cuyo caso $N_p = M - N$. Si los coeficientes a_k y b_k son reales, también lo son A_k , B_k , e_k y d_k . Esta expresión puede interpretarse como la combinación en paralelo de sistemas IIR de primer y segundo orden, más una cadena de N_p retardadores. Agrupando también los polos reales en pares, la función de transferencia se puede expresar como

$$H(z) = \sum_{k=0}^{N_p} C_k z^{-k} + \sum_{k=1}^{N_2} \frac{e_{0k} - e_{1k}z^{-1}}{1 - a_{1k}z^{-1} - a_{2k}z^{-2}}$$

Por hacer:

Figura de configuración paralela

Capítulo 7

Introducción al diseño de filtros digitales

7.1 Causalidad y sus implicaciones

Sea $h(n)$ la respuesta impulsional de un filtro paso bajo ideal con respuesta en frecuencia

$$H(\omega) = \begin{cases} 1 & |\omega| \leq \omega_C \\ 0 & \omega_C < \omega < \pi \end{cases}$$

•
|
○

$$h(n) = \begin{cases} \frac{\omega_C}{\pi} & n = 0 \\ \frac{\omega_C}{\pi} \frac{\text{sen}(\omega_C n)}{\omega_C n} & \text{en otro caso} \end{cases}$$

que obviamente no es causal y por tanto no realizable. Pero ¿cómo debe ser $H(\omega)$ para que $h(n)$ sea causal? Esta pregunta la responde el Teorema de Paley-Wiener que afirma: si $h(n)$ tiene energía finita y es causal ($h(n) = 0, n < 0$) entonces

$$\int_{-\pi}^{\pi} |\ln |H(\omega)| | d\omega < \infty$$

Si esta ecuación se cumple para $|H(\omega)|$ entonces puede buscarse una respuesta de fase asociada $\Theta(\omega)$ tal que $H(\omega) = |H(\omega)|e^{j\Theta(\omega)}$ represente una señal causal.

Nótese que de acuerdo a este teorema la función $|H(\omega)|$ puede ser cero en frecuencias puntuales aisladas, pero no en una banda finita, puesto que la integral se haría infinita. Por lo tanto, ningún filtro ideal es causal.

Puesto que $h(n)$ se puede separar en componentes par e impar:

$$h(n) = h_e(n) + h_o(n)$$

$$h_e(n) = \frac{1}{2} [h(n) + h(-n)]$$

$$h_o(n) = \frac{1}{2} [h(n) - h(-n)]$$

si $h(n)$ es causal entonces se tiene además

$$h(n) = 2h_e(n)u(n) - h_e(0)\delta(n) = 2h_o(n)u(n) + h(0)\delta(n)$$

$$h_e(n) = h_o(n), \quad n > 0$$

y si $h(n)$ es absolutamente sumable (estable BIBO) entonces

$$H(\omega) = H_R(\omega) + jH_I(\omega)$$

y puesto que $h(n)$ es causal y real entonces

$$h_e(n) \circ \bullet H_R(\omega)$$

$$h_o(n) \circ \bullet H_I(\omega)$$

con lo que se deduce que $H_R(\omega)$ es suficiente para establecer $H(\omega)$. En otras palabras $H_R(\omega)$ y $H_I(\omega)$ son interdependientes y no se pueden especificar libremente para sistemas causales.

Para establecer la relación entre las partes real e imaginaria de la respuesta en frecuencia se plantea utilizando el teorema del emventanado

$$H(\omega) = H_R(\omega) + jH_I(\omega) = \mathcal{F} \{2h_e(n)u(n) - h_e(0)\delta(n)\}$$

$$= 2 [H_R(\omega) * U(\omega)] - h_e(0)$$

$$= \frac{1}{\pi} \int_{-\pi}^{\pi} H_R(\lambda) U(\omega - \lambda) d\lambda - h_e(0)$$

que con

$$u(n) \circ \bullet U(\omega) = \pi\delta(\omega) + \frac{1}{1 - e^{-j\omega}}$$

$$= \pi\delta(\omega) + \frac{1}{2} - j\frac{1}{2} \cot \frac{\omega}{2}, \quad -\pi < \omega < \pi$$

es equivalente a

$$H(\omega) = \underbrace{\frac{1}{\pi} \int_{-\pi}^{\pi} H_R(\lambda) \pi\delta(\omega - \lambda) d\lambda}_{H_R(\omega)} + \underbrace{\frac{1}{\pi} \int_{-\pi}^{\pi} H_R(\lambda) \frac{1}{2} d\lambda}_{h_e(0)}$$

$$+ \frac{1}{\pi} \int_{-\pi}^{\pi} H_R(\lambda) \frac{j}{2} \cot \left(\frac{\omega - \lambda}{2} \right) d\lambda - h_e(0)$$

y puesto que

$$H_R(\omega) + jH_I(\omega) = H_R(\omega) - \frac{j}{2\pi} \int_{-\pi}^{\pi} H_R(\lambda) \cot \left(\frac{\omega - \lambda}{2} \right) d\lambda$$

$$H_I(\omega) = -\frac{1}{2\pi} \int_{-\pi}^{\pi} H_R(\lambda) \cot \left(\frac{\omega - \lambda}{2} \right) d\lambda$$

denominada Transformada de Hilbert discreta de $H_R(\omega)$.

La causalidad además tiene otras consecuencias aquí no demostradas, como por ejemplo que $|H(\omega)|$ no puede ser constante en ningún rango finito de frecuencias y la transición de la banda de paso a la banda de rechazo no puede ser infinitamente abrupta. Por estas razones, las respuestas en magnitud de filtros reales solo pueden ser aproximaciones de las versiones ideales, tal y como lo muestra la figura 7.1. La frecuencia angular ω_P define el límite superior

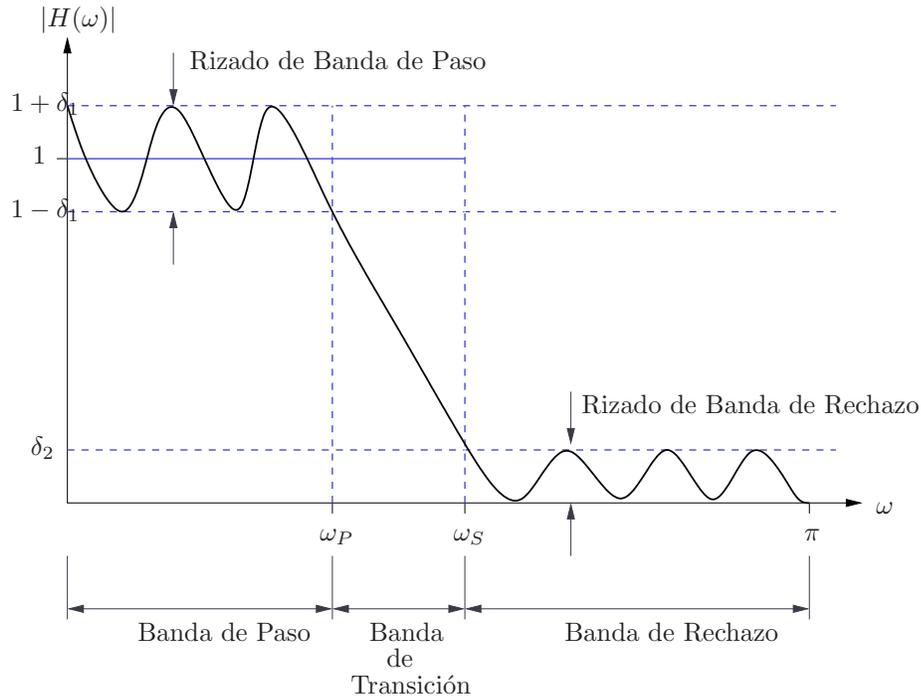


Figura 7.1: Respuesta real de filtro pasa bajas.

de la banda de paso y así el ancho de banda del filtro. La frecuencia angular ω_S indica el inicio de la banda de rechazo. El ancho de banda de transición es entonces $\omega_S - \omega_P$. Los rizados de las bandas de paso y rechazo son δ_1 y δ_2 respectivamente.

En aplicaciones reales se debe especificar entonces antes de diseñar el filtro

1. Máximo rizado permitido en la banda de paso δ_1
2. Máximo rizado permitido en la banda de rechazo δ_2
3. Frecuencia de corte de la banda de paso ω_P
4. Frecuencia de corte de la banda de rechazo ω_S .

donde las dos últimas se escogen usualmente en términos de potencia mitad.

El grado en que $H(\omega)$ acata las especificaciones depende en parte de los criterios utilizados para seleccionar $\{a_k\}$ y $\{b_k\}$, así como del número de polos y ceros utilizados.

Generalmente el diseño de filtros se concentra en el tipo pasa bajos, para lo que existen gran variedad de técnicas. Algunas de ellas se muestran en la figura 7.2 y se ilustran con más

detalle en [15]. Existen métodos para transformar los filtros pasa bajos a otros tipos como paso altos o paso banda.

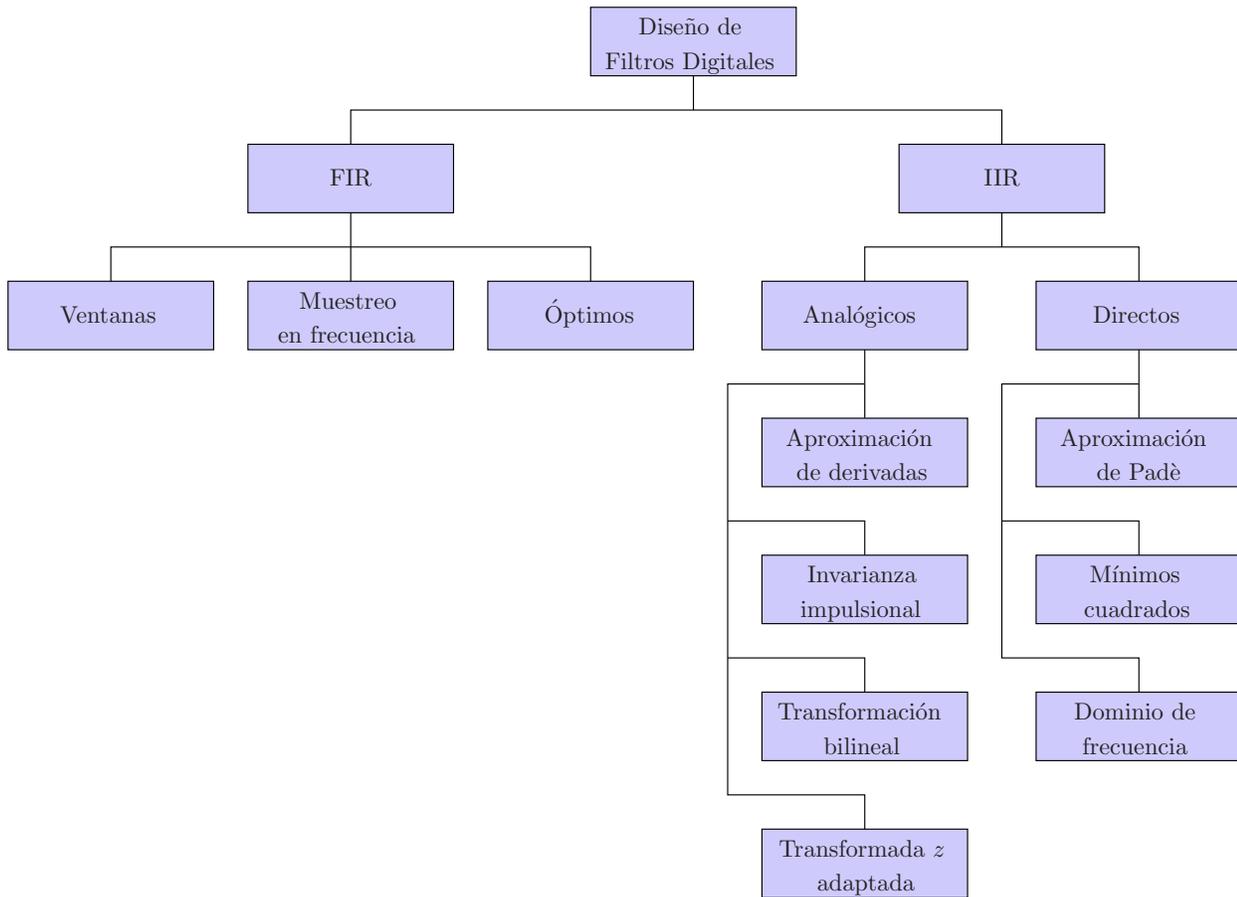


Figura 7.2: Taxonomía de métodos para el diseño de filtros.

7.2 Filtros de respuesta de impulso finita

Para un filtro FIR de longitud M la relación de entrada salida está dada por la convolución y es

$$y(n) = \sum_{k=0}^{M-1} b_k x(n-k) = \sum_{k=0}^{M-1} h(k)x(n-k) = h(n) * x(n)$$

y la función de transferencia es

$$H(z) = \sum_{k=0}^{M-1} h(k)z^{-k} = \frac{1}{z^{M-1}} \sum_{k=0}^{M-1} h(k)z^{M-1-k}$$

cuyas raíces son los ceros del filtro. Este filtro tiene fase lineal si su respuesta impulsional satisface las simetrías:

$$h(n) = \pm h(M-1-n), \quad n = 0, 1, \dots, M-1$$

que para la transformada z de $h(n)$, $H(z)$ implica los siguientes casos.

- Para M par

$$H(z) = z^{-(M-1)/2} \sum_{k=0}^{\frac{M}{2}-1} h(k) \left[z^{(M-1-2k)/2} \pm z^{-(M-1-2k)/2} \right]$$

y con $z = e^{j\omega}$ y asumiendo simetría circular par $h(n) = h(M-1-n)$ entonces

$$\begin{aligned} H(\omega) &= e^{-j\omega(M-1)/2} \sum_{k=0}^{\frac{M}{2}-1} h(k) \left[e^{j\omega(M-1-2k)/2} + e^{-j\omega(M-1-2k)/2} \right] \frac{2}{2} \\ &= e^{-j\omega(M-1)/2} 2 \sum_{k=0}^{\frac{M}{2}-1} h(k) \cos\left(\frac{\omega}{2}(M-1-2k)\right) \\ &= e^{-j\omega(M-1)/2} H_r(\omega) \end{aligned}$$

donde $H_r(\omega)$ es una función real por corresponder a una suma ponderada con coeficientes reales $h(k)$ y términos cosenoidales con argumentos reales. De forma equivalente, para el caso antisimétrico circular $h(n) = -h(M-1-n)$:

$$\begin{aligned} H(\omega) &= e^{-j\omega \frac{M-1}{2}} \sum_{k=0}^{\frac{M}{2}-1} h(k) \left[e^{j\omega(M-1-2k)/2} - e^{-j\omega(M-1-2k)/2} \right] \frac{2j}{2j} \\ &= e^{-j\left(\omega \frac{M-1}{2} - \frac{\pi}{2}\right)} 2 \sum_{k=0}^{\frac{M}{2}-1} h(k) \operatorname{sen}\left(\frac{\omega}{2}(M-1-2k)\right) \\ &= e^{-j\left(\omega \frac{M-1}{2} - \frac{\pi}{2}\right)} H_r(\omega) \end{aligned}$$

- Para M impar:

$$H(z) = z^{-(M-1)/2} \left\{ h\left(\frac{M-1}{2}\right) + \sum_{k=0}^{\frac{M-3}{2}} h(k) \left[z^{\frac{M-1-2k}{2}} \pm z^{-\frac{M-1-2k}{2}} \right] \right\}$$

Con simetría circular par $h(n) = h(M-1-n)$ lo anterior, en el dominio de la frecuencia, conduce a

$$\begin{aligned} H(\omega) &= e^{-j\omega(M-1)/2} \left[h\left(\frac{M-1}{2}\right) + 2 \sum_{k=0}^{\frac{M-3}{2}} h(k) \cos\left(\omega \frac{M-1-2k}{2}\right) \right] \\ &= e^{-j\omega(M-1)/2} H_r(\omega) \end{aligned}$$

y con antisimetría circular $h(n) = -h(M-1-n)$, que además implica $h\left(\frac{M-1}{2}\right) = 0$:

$$\begin{aligned} H(\omega) &= e^{-j\left(\omega \frac{M-1}{2} - \frac{\pi}{2}\right)} 2 \sum_{k=0}^{\frac{M-3}{2}} h(k) \operatorname{sen}\left(\omega \frac{M-1-2k}{2}\right) \\ &= e^{-j\left(\omega \frac{M-1}{2} - \frac{\pi}{2}\right)} H_r(\omega) \end{aligned}$$

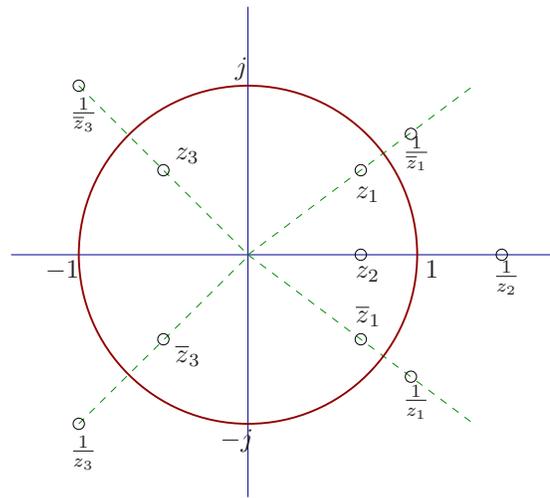


Figura 7.3: Posición de los ceros en un filtro FIR de fase lineal.

Nótese que se cumple además

$$z^{-(M-1)}H(z^{-1}) = \pm H(z)$$

lo que implica que si z_k es un cero, entonces z_k^* , $1/z_k$ y $1/z_k^*$ también lo son (figura 7.3).

La tabla 7.1 resume las posibilidades brindadas por las diferentes simetrías para filtros FIR de fase lineal.

Tabla 7.1: Simetrías en filtros FIR de fase lineal

Simetría	Simétrica $h(n) = h(M - 1 - n)$	Antisimétrica $h(n) = -h(M - 1 - n)$
M par	$H_r(0) = 2 \sum_{k=0}^{\frac{M}{2}-1} h(k)$	$H_r(0) = 0$ no apto como filtro paso bajos
M impar	$H_r(0) = h\left(\frac{M-1}{2}\right) + 2 \sum_{k=0}^{\frac{M-3}{2}} h(k)$	$H_r(0) = H_r(\pi) = 0$ no apto como filtro paso bajos o altos

7.2.1 Diseño de filtros FIR por el método de ventanas

Sea $H_d(\omega)$ la respuesta en frecuencia deseada, que usualmente sigue la forma de un filtro ideal. En general, la respuesta impulsional correspondiente $h_d(n)$ es infinita y dada por la transformada inversa de Fourier

$$h_d(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} H_d(\omega) e^{j\omega n} d\omega$$

El filtro FIR se obtiene truncando $h_d(n)$ por medio de una ventana

$$w(n) = \begin{cases} 1 & n = 0, 1, \dots, M-1 \\ 0 & \text{en otro caso} \end{cases}$$

es decir,

$$h(n) = h_d(n)w(n) = \begin{cases} h_d(n) & n = 0, 1, \dots, M-1 \\ 0 & \text{en otro caso} \end{cases}$$

En el dominio de la frecuencia $H(\omega) = H_d(\omega) * W(\omega)$. Puesto que $W(\omega)$ tiene generalmente un lóbulo central y lóbulos laterales, el efecto de la convolución es suavizar la respuesta $H_d(\omega)$. Para reducir el efecto de dichos lóbulos laterales se utilizan ventanas diferentes a la rectangular, caracterizadas por no tener cambios abruptos en el dominio del tiempo, lo que conduce a lóbulos menores en el dominio de la frecuencia. Algunas ventanas típicas y sus características se presentan en la tabla 7.2.

Tabla 7.2: Funciones utilizadas como ventanas.

Ventana	$h(n), 0 \leq n \leq M-1$	Ancho lobular	Pico lóbulo lateral [dB]
Rectangular	1	$4\pi/M$	-13
Bartlett (triangular)	$1 - \frac{2 n - \frac{M-1}{2} }{M-1}$	$8\pi/M$	-27
Hamming	$0,54 - 0,46 \cos \frac{2\pi n}{M-1}$	$8\pi/M$	-32
Hanning	$\frac{1}{2} \left(1 - \cos \frac{2\pi n}{M-1} \right)$	$8\pi/M$	-43

Las ventanas se ilustran en el dominio temporal en la figura 7.4, y en el dominio de la frecuencia, con su magnitud espectral en las figuras 7.5 y 7.6.

Ejemplo 7.1 Diseñe un filtro pasa bajos de longitud M que aproxime

$$H_d(\omega) = \begin{cases} e^{j\omega(M-1)/2} & 0 \leq |\omega| \leq \omega_C \\ 0 & \text{en el resto} \end{cases}$$

Se obtiene con la transformada inversa de Fourier

$$h_d(n) = \frac{1}{2\pi} \int_{-\omega_C}^{\omega_C} e^{j\omega(n - \frac{M-1}{2})} d\omega = \frac{\text{sen } \omega_C \left(n - \frac{M-1}{2} \right)}{\pi \left(n - \frac{M-1}{2} \right)}$$

Para el filtro FIR se toman solo M muestras:

$$h(n) = \frac{\text{sen } \omega_C \left(n - \frac{M-1}{2} \right)}{\pi \left(n - \frac{M-1}{2} \right)}, \quad 0 \leq n \leq M-1$$

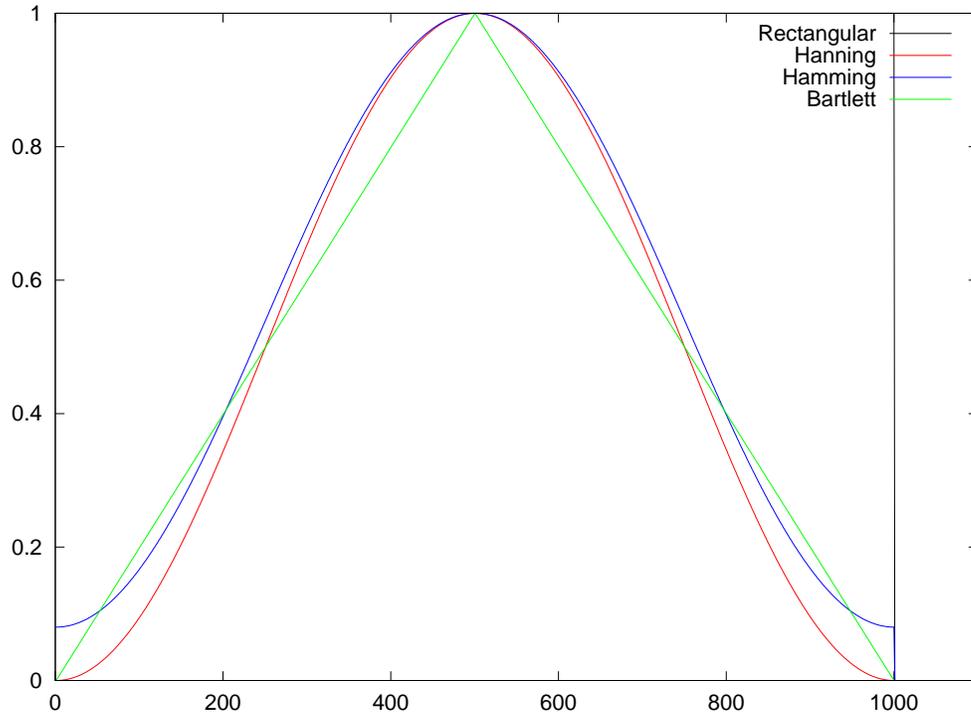


Figura 7.4: Ventanas en el dominio del tiempo.

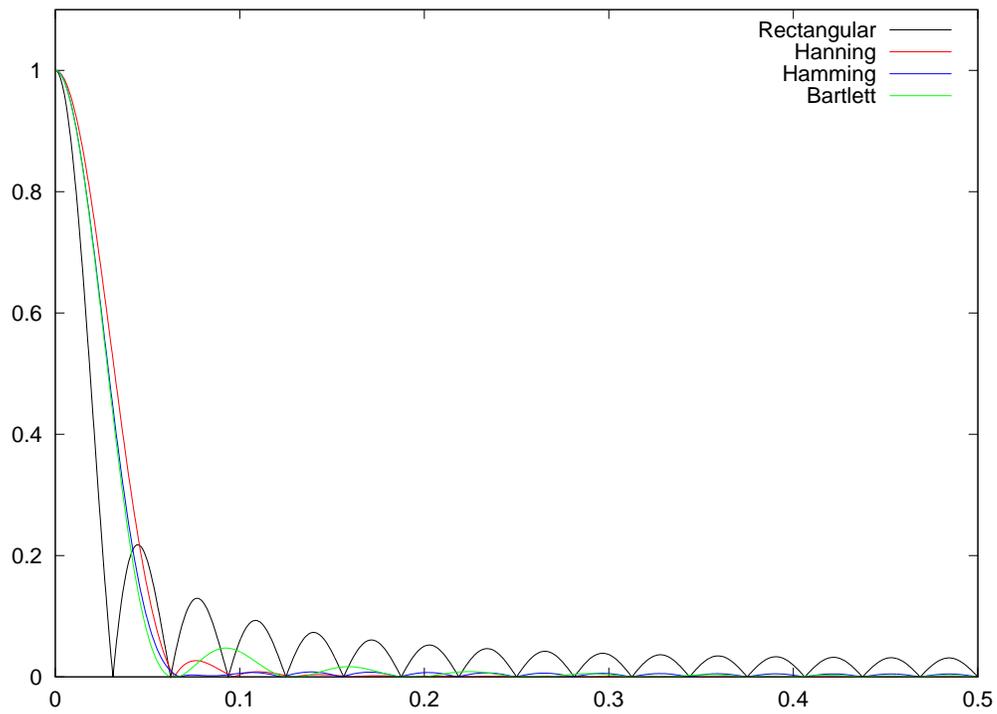


Figura 7.5: Ventanas en el dominio de la frecuencia (normalizado).

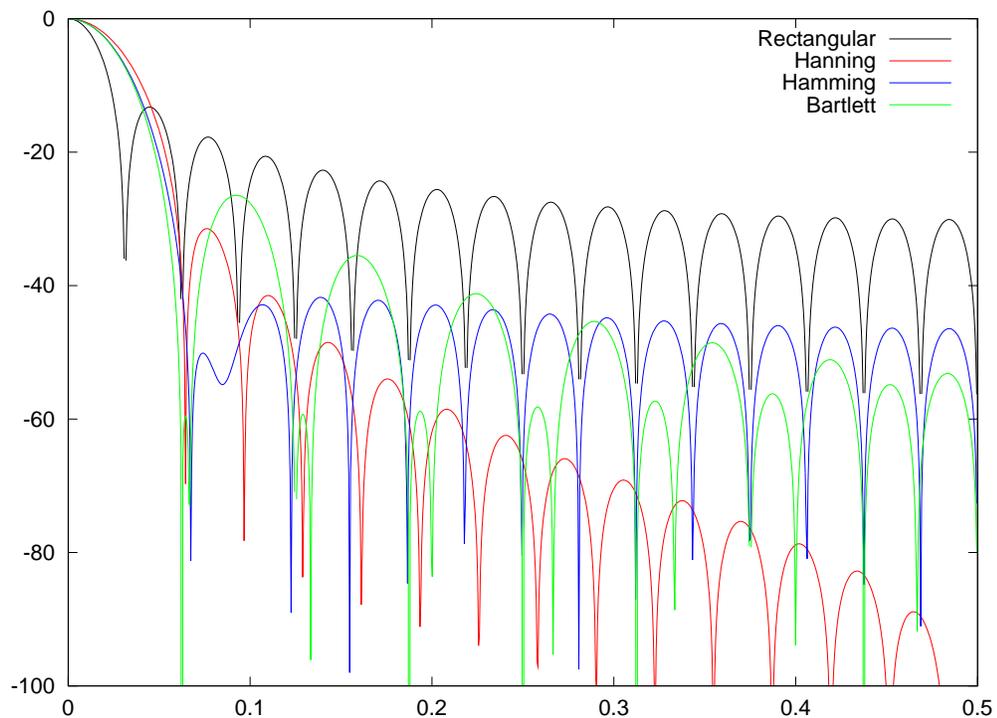


Figura 7.6: Ventanas en el dominio de la frecuencia (en dB).

7.1

El truncamiento de $h_d(n)$ conduce a un comportamiento oscilatorio cerca del límite de la banda de paso denominado *fenómeno de Gibbs*. Nótese que este método no permite mucho control sobre el valor de los parámetros δ_1 , δ_2 , ω_S y ω_P resultantes.

Diseño de filtros FIR de fase lineal por el método de muestreo de frecuencia

En este método se muestrea la respuesta en frecuencia deseada en $\lfloor \frac{M}{2} \rfloor$ puntos equiespaciados

$$\omega_k = \frac{2\pi}{M}(k + \alpha) \quad \begin{array}{l} k = 0, 1, \dots, \frac{M-1}{2} \quad M \text{ impar} \\ k = 0, 1, \dots, \frac{M}{2} - 1 \quad M \text{ par} \end{array}$$

$$\alpha = 0 \text{ ó } \frac{1}{2}$$

y se calcula la respuesta impulsional $h(n)$ correspondiente, donde para reducir los lóbulos laterales se utilizan métodos numéricos de optimización que sitúan las muestras en la banda de transición.

Puesto que

$$H(k + \alpha) = H\left(\frac{2\pi}{M}(k + \alpha)\right) = \sum_{n=0}^{M-1} h(n)e^{-j2\pi(k+\alpha)n/M}, \quad k = 0, 1, \dots, M-1$$

y haciendo uso de la ortogonalidad de la exponencial compleja se cumple además

$$h(n) = \frac{1}{M} \sum_{k=0}^{M-1} H(k + \alpha) e^{j2\pi(k+\alpha)n/M}, \quad n = 0, 1, \dots, M-1$$

Con $\alpha = 0$ estas ecuaciones equivalen a la DFT e IDFT respectivamente. Si $h(n)$ es real, entonces

$$H(k + \alpha) = H^*(M - k - \alpha)$$

lo que justifica la utilización de $\lfloor \frac{M}{2} \rfloor$ muestras en vez de M . Utilizando la simetría se obtiene

$$H(k + \alpha) = H_r \left(\frac{2\pi}{M}(k + \alpha) \right) e^{j[\beta\pi/2 - 2\pi(k+\alpha)(M+1)/2M]}$$

donde $\beta = 0$ si $h(n)$ es simétrica y $\beta = 1$ si $h(n)$ es antisimétrica. Con

$$G(k + \alpha) = (-1)^k H_r \left(\frac{2\pi}{M}(k + \alpha) \right) \quad k = 0, 1, \dots, M-1$$

se obtiene

$$H(k + \alpha) = G(k + \alpha) e^{j\pi k} e^{j[\beta\pi/2 - 2\pi(k+\alpha)(M+1)/2M]}$$

Respuesta impulsional $h(n) = \pm h(M-1-n)$

Simetría circular par

$$a = 0: \begin{cases} H(k) = G(k) e^{j\pi k/M} & k = 0, 1, \dots, M-1 \\ G(k) = (-1)^k H_r \left(\frac{2\pi k}{M} \right) & G(k) = -G(M-k) \\ h(n) = \frac{1}{M} \left\{ G(0) + 2 \sum_{k=1}^{\lfloor \frac{M-1}{2} \rfloor} G(k) \cos \left(\frac{2\pi k}{M} \left(n + \frac{1}{2} \right) \right) \right\} \end{cases}$$

$$a = \frac{1}{2}: \begin{cases} H \left(k + \frac{1}{2} \right) = G \left(k + \frac{1}{2} \right) e^{-j\frac{\pi}{2}} e^{j\pi(2k+1)/2M} \\ G \left(k + \frac{1}{2} \right) = G \left(M - k - \frac{1}{2} \right) \\ h(n) = \frac{2}{M} \sum_{k=0}^{\lfloor \frac{M-1}{2} \rfloor} G \left(k + \frac{1}{2} \right) \text{sen} \frac{2\pi}{M} \left(k + \frac{1}{2} \right) \left(n + \frac{1}{2} \right) \end{cases}$$

Simetría circular impar

$$a = 0: \begin{cases} H(k) = G(k) e^{j\pi/2} e^{j\pi k/M} & k = 0, 1, \dots, M-1 \\ G(k) = (-1)^k H_r \left(\frac{2\pi k}{M} \right) & G(k) = G(M-k) \\ h(n) = -\frac{2}{M} \sum_{k=1}^{\frac{M-1}{2}} G(k) \text{sen} \left(\frac{2\pi k}{M} \left(n + \frac{1}{2} \right) \right) & M \text{ impar} \\ h(n) = \frac{1}{M} \left\{ (-1)^{n+1} G(M/2) - 2 \sum_{k=1}^{\frac{M-1}{2}} G(k) \text{sen} \left(\frac{2\pi k}{M} \left(n + \frac{1}{2} \right) \right) \right\} & M \text{ par} \end{cases}$$

$$a = \frac{1}{2} : \begin{cases} H\left(k + \frac{1}{2}\right) = G\left(k + \frac{1}{2}\right) e^{j\pi(2k+1)/2M} \\ G\left(k + \frac{1}{2}\right) = (-1)^k H_r\left(\frac{2\pi}{M}\left(k + \frac{1}{2}\right)\right) \\ G\left(k + \frac{1}{2}\right) = -G\left(M - k - \frac{1}{2}\right); \quad G(M/2) = 0 \text{ para } M \text{ impar} \\ h(n) = \frac{2}{M} \sum_{k=0}^{\lfloor \frac{M}{2}-1 \rfloor} G\left(k + \frac{1}{2}\right) \cos \frac{2\pi}{M}\left(k + \frac{1}{2}\right)\left(n + \frac{1}{2}\right) \end{cases}$$

Ejemplo 7.2 Diseñe un filtro de longitud $M = 15$ con respuesta impulsional $h(n)$ simétrica y respuesta en frecuencia condicionada a

$$H_r\left(\frac{2\pi k}{15}\right) = \begin{cases} 1 & k = 0, 1, 2, 3 \\ 0,4 & k = 4 \\ 0 & k = 5, 6, 7 \end{cases}$$

Solución: $h(n)$ es simétrica y en este caso $\alpha = 0$. Con $G(k) = (-1)^k H_r\left(\frac{2\pi k}{15}\right)$, $k = 0, 1, \dots, 7$ y las ecuaciones anteriores se obtiene

$$\begin{aligned} h(0) &= h(14) = -0,014112893 \\ h(1) &= h(13) = -0,001945309 \\ h(2) &= h(12) = 0,04000004 \\ h(3) &= h(11) = 0,01223454 \\ h(4) &= h(10) = -0,09138802 \\ h(5) &= h(9) = -0,1808986 \\ h(6) &= h(8) = 0,3133176 \\ h(7) &= 0,52 \end{aligned}$$

7.2

7.2.2 Diseño de filtros óptimos

El diseño de filtros se denomina *óptimo* si el error de aproximación entre la respuesta en frecuencia deseada y la actual se distribuye equitativamente a lo largo de las bandas de paso y rechazo.

Dadas las frecuencias de corte para las bandas de paso y rechazo ω_P y ω_S respectivamente, el filtro debe satisfacer además

$$\begin{aligned} 1 - \delta_1 &\leq H_r(\omega) \leq 1 + \delta_1 & |\omega| \leq \omega_P \\ -\delta_2 &\leq H_r(\omega) \leq \delta_2 & |\omega| > \omega_S \end{aligned}$$

por lo que δ_1 representa el rizado de la banda de paso y δ_2 la atenuación o rizado en la banda de rechazo.

Utilizando los filtros simétricos y antisimétricos de fase lineal puede demostrarse que

$$H_r(\omega) = Q(\omega)P(\omega)$$

con $Q(\omega)$ y $P(\omega)$ dados en la tabla 7.3.

El error de aproximación se define entonces como

$$\begin{aligned} E(\omega) &= W(\omega)[H_{dr}(\omega) - H_r(\omega)] \\ &= W(\omega)[H_{dr}(\omega) - Q(\omega)P(\omega)] \\ &= W(\omega)Q(\omega) \left[\frac{H_{dr}(\omega)}{Q(\omega)} - P(\omega) \right] \end{aligned}$$

donde $W(\omega)$ es una función de ponderación de error que usualmente se define como

$$W(\omega) = \begin{cases} \frac{\delta_1}{\delta_2} & \omega \text{ en la banda de paso} \\ 1 & \omega \text{ en la banda de rechazo} \end{cases}$$

y $H_{dr}(\omega)$ la respuesta ideal del filtro deseada que es 1 en la banda de paso y cero en la de rechazo. Con

$$\begin{aligned} \hat{W}(\omega) &= W(\omega)Q(\omega) \\ \hat{H}_{dr}(\omega) &= \frac{H_{dr}(\omega)}{Q(\omega)} \end{aligned}$$

el error se puede reescribir como

$$E(\omega) = \hat{W}(\omega) \left[\hat{H}_{dr}(\omega) - P(\omega) \right].$$

El diseño del filtro consiste entonces en buscar los parámetros $\{\alpha(k)\} = \{a(k)\}, \{b(k)\}, \{c(k)\}$ ó $\{d(k)\}$ que conducen al menor valor máximo de $|E(\omega)|$.

$$\arg \min_{\{\alpha(k)\}} \left[\max_{\omega \in S} |E(\omega)| \right] = \arg \min_{\{\alpha(k)\}} \left[\max_{\omega \in S} \left| \hat{W}(\omega) \left[\hat{H}_{dr} - \sum_{k=0}^L \alpha(k) \cos \omega k \right] \right| \right]$$

donde S representa el conjunto (disjunto) de bandas sobre las que se realiza la optimización. Este problema de optimización requiere también métodos numéricos de optimización (por ejemplo el *algoritmo de intercambio* de Remez) basados en el teorema de la alterternancia, que establece que la función de error $E(\omega)$ debe exhibir al menos $L + 2$ frecuencias extremas en S , es decir, deben existir al menos $L + 2$ frecuencias $\{\omega_i\}$ en S tales que $\omega_1 < \omega_2 < \dots < \omega_{L+2}$, $E(\omega_i) = -E(\omega_{i-1})$ y $|E(\omega_i)| = \max_{\omega \in S} |E(\omega)|$, $i = 1, 2, \dots, L + 2$ para que $P(\omega) = \sum_{k=0}^L \alpha(k) \cos \omega k$ sea la mejor y única aproximación ponderada de $\hat{H}_{dr}(\omega)$.

El nombre del teorema se debe a la alternancia de signo entre dos extremos consecutivos.

Tabla 7.3: Descomposición de filtros FIR en $P(\omega)$ y $Q(\omega)$.

	Simétrico	Antisimétrico
	$h(n) = h(M-1-n)$	$h(n) = -h(M-1-n)$
M Impar	Caso 1	Caso 3
	$Q(\omega) = 1$	$Q(\omega) = \text{sen}(\omega)$
	$P(\omega) = \sum_{k=0}^{(M-1)/2} a(k) \cos \omega k$	$P(\omega) = \sum_{k=0}^{(M-3)/2} \tilde{c}(k) \cos \omega k$
	$a(k) = \begin{cases} h\left(\frac{M-1}{2}\right) & k=0 \\ 2h\left(\frac{M-1}{2}-k\right) & k=1, 2, \dots, \frac{M-1}{2} \end{cases}$	$\tilde{c}\left(\frac{M-3}{2}\right) = 2h(0)$ $\tilde{c}\left(\frac{M-5}{2}\right) = 4h(1)$ $\tilde{c}(k-1) - \tilde{c}(k+1) = 4h\left(\frac{M-1}{2}-k\right), \quad 2 \leq k \leq \frac{M-5}{2}$ $\tilde{c}(0) + \frac{1}{2}\tilde{c}(2) = 2h\left(\frac{M-3}{2}\right)$
M Par	Caso 2	Caso 4
	$Q(\omega) = \cos\left(\frac{\omega}{2}\right)$	$Q(\omega) = \text{sen}\left(\frac{\omega}{2}\right)$
	$P(\omega) = \sum_{k=0}^{\frac{M}{2}-1} \tilde{b}(k) \cos(\omega k)$	$P(\omega) = \sum_{k=0}^{\frac{M}{2}-1} \tilde{d}(k) \cos(\omega k)$
	$\tilde{b}(0) = h\left(\frac{M}{2}-1\right)$	$\tilde{d}\left(\frac{M}{2}-1\right) = 4h(0)$
	$\tilde{b}(k) = 4h\left(\frac{M}{2}-k\right) - \tilde{b}(k-1), \quad 1 \leq k \leq \frac{M}{2}-2$	$\tilde{d}(k-1) - \tilde{d}(k) = 4h\left(\frac{M}{2}-k\right), \quad 2 \leq k \leq \frac{M}{2}-1$
	$\tilde{b}\left(\frac{M}{2}-1\right) = 4h(0)$	$\tilde{d}(0) - \frac{1}{2}\tilde{d}(1) = 2h\left(\frac{M}{2}-1\right)$

7.3 Diseño de filtros de respuesta impulsional infinita a partir de filtros analógicos

Puesto que el diseño de filtros analógicos es un campo maduro y bien desarrollado, puede hacerse uso de sus resultados transformando filtros analógicos, descritos en el dominio de frecuencia compleja $s = \sigma + j\Omega$ como

$$H_a(s) = \frac{B(s)}{A(s)} = \frac{\sum_{k=0}^M \beta_k s^k}{\sum_{k=0}^N \alpha_k s^k} \quad (7.1)$$

a filtros digitales utilizando algún mapeo adecuado entre las variables s y z . La función de transferencia $H_a(s)$ se relaciona con la respuesta impulsional $h(t)$ a través de la Transformada de Laplace

$$H_a(s) = \int_{-\infty}^{\infty} h(t)e^{-st} dt$$

que teniendo una forma racional como en (7.1) está relacionada con la ecuación diferencial

$$\sum_{k=0}^N \alpha_k \frac{d^k y(t)}{dt^k} = \sum_{k=0}^M \beta_k \frac{d^k x(t)}{dt^k} \quad (7.2)$$

donde $x(t)$ representa la entrada del filtro y $y(t)$ su salida.

Puesto que el sistema descrito por $H_a(s)$ es estable si sus polos se encuentran al lado izquierdo del eje $\sigma = 0$ entonces el mapeo de s a z debe

1. Transformar $s = j\Omega$ en la circunferencia unitaria en el plano z .
2. El semiplano izquierdo (LHP) de s debe corresponder con el interior de la circunferencia unitaria.

En la sección anterior se mencionó que un filtro tiene fase lineal si

$$H(z) = \pm z^{-N} H(z^{-1}).$$

Con filtros FIR esto se utiliza para ubicar los ceros, pero con filtros IIR esto implica que cada polo dentro de la circunferencia unitaria en z tiene otro polo correspondiente fuera de ella, lo que implica que el filtro será inestable. En otras palabras, si se requiere un filtro de fase lineal, debe utilizarse un filtro FIR como los tratados en la sección anterior.

Con filtros IIR usualmente se enfoca el diseño a la respuesta en magnitud, dejando que las interrelaciones existentes entre fase y magnitud determinen la fase, según corresponda con la metodología de diseño utilizada.

7.3.1 Diseño por aproximación de derivadas

Si se cuenta con la ecuación diferencial del filtro (7.2), se pueden aproximar las derivadas haciendo uso de

$$\left. \frac{dy(t)}{dt} \right|_{t=nT} \approx \frac{y(nT) - y(nT - T)}{T} = \frac{y(n) - y(n-1)}{T}$$

con el intervalo de muestreo T . Transformando al dominio s y z se tiene

$$s = \frac{1 - z^{-1}}{T}.$$

Puede deducirse de forma similar que la k -ésima derivada resulta en la relación

$$s^k = \left(\frac{1 - z^{-1}}{T} \right)^k$$

por lo que la aproximación para el filtro IIR digital es

$$H(z) = H_a(s) \Big|_{s=\frac{1-z^{-1}}{T}}$$

La equivalencia entre s y z se puede derivar de

$$z = \frac{1}{1 - sT}$$

mapeo que se ilustra en la figura 7.7.

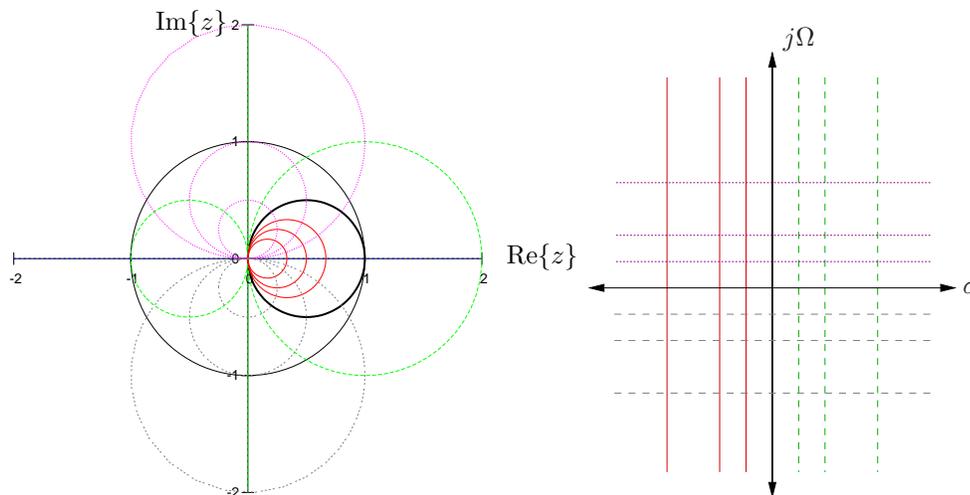


Figura 7.7: Mapeo entre el plano s y el plano z para la aproximación de derivadas.

Nótese que los polos del filtro digital solo se ubicarán en frecuencias pequeñas, lo que restringe la utilidad de este método a solo filtros paso bajo y paso banda con frecuencias resonantes relativamente bajas.

7.3.2 Diseño por invarianza impulsional

Este método consiste en buscar una respuesta impulsional del filtro digital que corresponda a la respuesta impulsional muestreada del filtro analógico:

$$h(n) = h(t)|_{t=nT}, \quad n = 0, 1, \dots$$

En capítulos anteriores se analizó el hecho de que el muestreo en el tiempo conduce a una extensión periódica que consiste en la superposición del espectro analógico desplazado por múltiplos de la frecuencia de muestreo $F_s = 1/T$.

$$H(\omega) = F_s \sum_{k=-\infty}^{\infty} H_a(\omega F_s - 2\pi F_s k)$$

El aliasing que ocurre a altas frecuencias hace este método inapropiado para el diseño de filtros paso alto.

En este caso puede demostrarse que la equivalencia entre los dos dominios es $z = e^{sT} = e^{\sigma + j\Omega T} = e^{\sigma T} e^{j\Omega T} = r e^{j\omega}$ con $r = e^{\sigma T}$ y $\omega = \Omega T$, lo que implica que el semiplano izquierdo (LHP) corresponde al interior de la circunferencia unitaria ($r < 1$). Nótese que todos los intervalos de frecuencia $(2k - 1)\frac{\pi}{T} \leq \Omega \leq (2k + 1)\frac{\pi}{T}$ se proyectan en la circunferencia, lo que hace de esta correspondencia una relación inyectiva que proviene del efecto de aliasing debido al muestreo (ver sección 2.2.2).

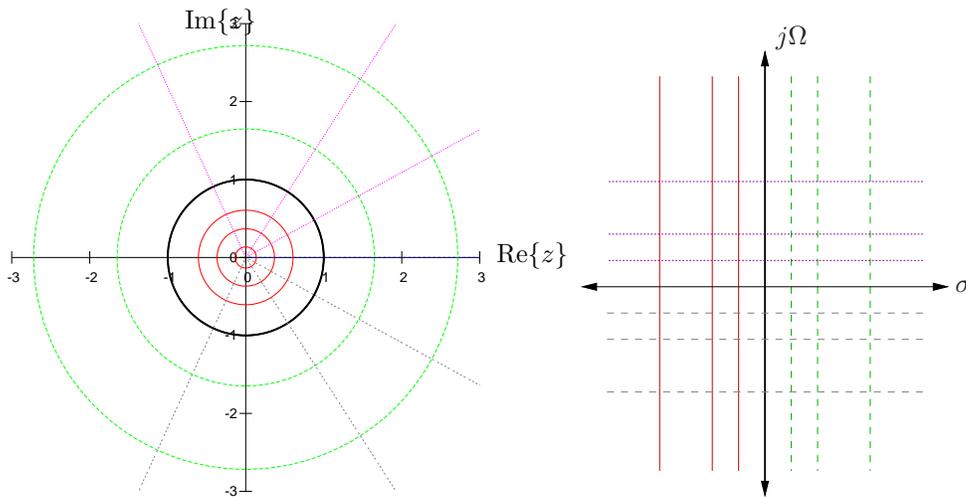


Figura 7.8: Mapeo entre el plano s y el plano z para la invarianza impulsional.

Puede demostrarse que el filtro analógico

$$H_a(s) = \sum_{k=1}^M \frac{c_k}{s - p_k} \text{ equivale a } H(z) = \sum_{k=1}^M \frac{c_k}{1 - e^{p_k T} z^{-1}}$$

y aunque los polos siguen la relación $z_k = e^{p_k T}$, esto no se satisface para los ceros.

7.3.3 La transformada z adaptada

De forma similar al diseño por invarianza impulsional, se utiliza aquí el mapeo $z = e^{sT}$, pero tanto para polos como para ceros. Así, transformación z adaptada se le denomina a la equivalencia entre $(s - a)$ y $(1 - e^{aT}z^{-1})$. De esta forma

$$H(s) = \frac{\prod_{k=1}^M (s - z_k)}{\prod_{k=1}^N (s - p_k)} \text{ es equivalente a } H(z) = \frac{\prod_{k=1}^M (1 - e^{z_k T} z^{-1})}{\prod_{k=1}^N (1 - e^{p_k T} z^{-1})}$$

Con este método debe seleccionarse T lo suficientemente pequeño para evitar el aliasing.

7.3.4 Diseño por transformación bilineal

En este método se utiliza la relación

$$s = \frac{2}{T} \left(\frac{1 - z^{-1}}{1 + z^{-1}} \right)$$

denominada relación bilineal y derivada de la fórmula trapezoidal para integración numérica.

Este método es apto para transformar filtros paso alto y paso banda por tener una representación biyectiva entre ω y Ω . Con

$$z = -\frac{Ts + 2}{Ts - 2}$$

se obtiene el mapeo mostrado en la figura 7.9.

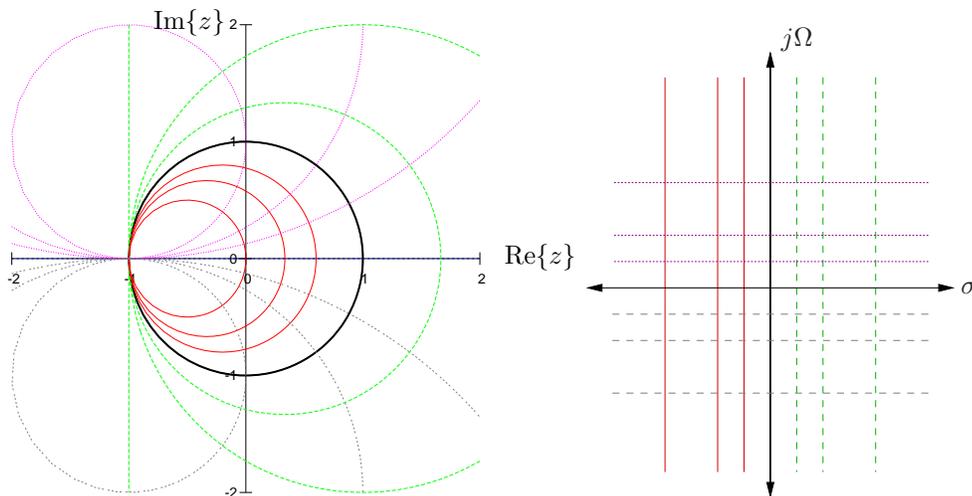


Figura 7.9: Mapeo bilineal entre el plano s y el plano z .

7.3.5 Filtros Analógicos

Estas transformaciones pueden aplicarse para derivar filtros digitales, por ejemplo, filtros Butterworth, filtros Chebyshev (tipos I y II), filtros elípticos y filtros Bessel (tabla 7.4).

Tabla 7.4: Características de Filtros Paso Bajo Analógicos

Filtro	Función de Transferencia	Características
Butterworth (todo-polos)	$ H(\Omega) ^2 = \frac{1}{1 + (\Omega/\Omega_C)^{2N}}$ Ω_C : frecuencia de corte	$ H(\Omega) $ monótona en las bandas de paso y rechazo.
Chebyshev, Tipo I (todo-polos)	$ H(\Omega) ^2 = \frac{1}{1 + \epsilon^2 T_N^2\left(\frac{\Omega}{\Omega_P}\right)}$ T_N : polinomio de Chebyshev $T_0(x) = 1$ $T_1(x) = x$ $T_{N+1}(x) = 2xT_N(x) - T_{N-1}(x)$	Rizado constante en la banda de paso y característica monótona en la banda de rechazo.
Chebyshev, Tipo II (polos y ceros)	$ H(\Omega) ^2 = \frac{1}{1 + \epsilon^2 \left[\frac{T_N^2\left(\frac{\Omega_S}{\Omega_P}\right)}{T_N^2\left(\frac{\Omega}{\Omega}\right)} \right]}$ T_N : polinomio de Chebyshev	Rizado constante en la banda de rechazo y característica monótona en la banda de paso.
Elíptico ó Cauer (polos y ceros)	$ H(\Omega) ^2 = \frac{1}{1 + \epsilon^2 U_N\left(\frac{\Omega}{\Omega_P}\right)}$ U_N : función elíptica Jacobiana de orden N	Rizado constante tanto en la banda de paso como en la de rechazo.
Bessel (todo-polos)	$H(s) = \frac{1}{B_N(s)}$ $B_N(s)$: funciones de Bessel $B_0(s) = 1$ $B_1(s) = s + 1$ $B_N(s) = (2N - 1)B_{N-1}(s) + s^2 B_{N-2}(s)$	Respuesta de fase lineal en la banda de paso (aunque se destruye en la conversión a digital).

7.4 Transformación de Frecuencia

Hay dos maneras de obtener un filtro digital paso bajo, paso alto, paso banda o supresor de banda a partir de un filtro paso bajo analógico:

1. Transformar en el dominio analógico al filtro deseado, y luego usar uno de los mapeos

de s a z presentados anteriormente para transformarlo al dominio digital.

2. Transformar el filtro paso bajo a un equivalente paso bajo digital, para luego hacer la transformación al filtro deseado en el dominio digital.

Para el primer caso se utilizan las transformaciones mostradas en la tabla 7.5. Para el segundo se presentan las modificaciones en la tabla 7.6.

Tabla 7.5: Transformaciones de frecuencia para filtros analógicos.

Tipo de filtro deseado	Transformación	Nuevas frecuencias de corte
Paso bajo	$s \rightarrow \frac{\Omega_P}{\Omega'_P} s$	Ω'_P
Paso alto	$s \rightarrow \frac{\Omega_P \Omega'_P}{s^2 + \Omega_l \Omega_u}$	Ω'_P
Paso banda	$s \rightarrow \Omega_P \frac{s^2 + \Omega_l \Omega_u}{s(\Omega_u - \Omega_l)}$	Ω_u, Ω_l
Supresor de banda	$s \rightarrow \Omega_P \frac{s(\Omega_u - \Omega_l)}{s^2 + \Omega_l \Omega_u}$	Ω_u, Ω_l

Ω_P : frecuencia de corte del prototipo
 Ω_l : frecuencia de corte inferior
 Ω_u : frecuencia de corte superior

Tabla 7.6: Transformaciones de frecuencia para filtros digitales.

Tipo de filtro deseado	Transformación	Constantes	Nuevas frecuencias de corte
Paso bajo	$z^{-1} \rightarrow \frac{z^{-1} - a}{1 - az^{-1}}$	$a = \frac{\operatorname{sen}\left(\frac{\omega_P - \omega'_P}{2}\right)}{\operatorname{sen}\left(\frac{\omega_P + \omega'_P}{2}\right)}$	ω'_P
Paso alto	$z^{-1} \rightarrow \frac{z^{-1} + a}{1 + az^{-1}}$	$a = \frac{\operatorname{cos}\left(\frac{\omega_P + \omega'_P}{2}\right)}{\operatorname{cos}\left(\frac{\omega_P - \omega'_P}{2}\right)}$	ω'_P
Paso banda	$z^{-1} \rightarrow \frac{z^{-2} - a_1 z^{-1} + a_2}{a_2 z^{-2} - a_1 z^{-1} + 1}$	$a_1 = -2\alpha \frac{K}{K+1}$ $a_2 = \frac{K-1}{K+1}$ $\alpha = \frac{\operatorname{cos}\left(\frac{\omega_u + \omega_l}{2}\right)}{\operatorname{cos}\left(\frac{\omega_u - \omega_l}{2}\right)}$ $K = \cot \frac{\omega_u - \omega_l}{2} \tan \frac{\omega_P}{2}$	ω_u, ω_l
Supresor de banda	$z^{-1} \rightarrow \frac{z^{-2} - a_1 z^{-1} + a_2}{a_2 z^{-2} - a_1 z^{-1} + 1}$	$a_1 = -2\alpha \frac{1}{K+1}$ $a_2 = \frac{1-K}{1+K}$ $\alpha = \frac{\operatorname{cos}\left(\frac{\omega_u + \omega_l}{2}\right)}{\operatorname{cos}\left(\frac{\omega_u - \omega_l}{2}\right)}$ $K = \tan \frac{\omega_u - \omega_l}{2} \tan \frac{\omega_P}{2}$	ω_u, ω_l

ω_P : frecuencia de corte normalizada del prototipo

ω_l : frecuencia de corte normalizada inferior

ω_u : frecuencia de corte normalizada superior

Bibliografía

- [1] P. Alvarado. Modelos de sistemas. Notas de clase, 2005. URL <http://www.ie.itcr.ac.cr/palvarado/Modelos/modelos.pdf>. 94
- [2] P.M. Aziz, H. V. Sorensen, and J. van der Spiegel. An overview of sigma-delta converters. *IEEE Signal Processing Magazine*, pages 61–84, January 1996.
- [3] C. S. Burrus, J. H. McClellan, A. V. Oppenheim, T. W. Parks, R. W. Schafer, and H. W. Schuessler. *Ejercicios de Tratamiento de la Señal. Un enfoque práctico*. Prentice Hall, 1998.
- [4] H. F. Davis. *Fourier series and orthogonal functions*. Dover Publications, Inc., 1963.
- [5] John W. Eaton. Octave [online]. 1998 [visitado el 22 de noviembre de 2010]. URL <http://www.octave.org>. 14, 215
- [6] John W. Eaton. Octave repository [online]. 1998 [visitado el 22 de noviembre de 2010]. URL <http://octave.sourceforge.net/afunclist.html>. 215
- [7] S. Haykin and B. van Veen. *Señales y sistemas*. Limusa Wiley, 2001.
- [8] G. James. *Matemáticas Avanzadas para Ingeniería*. Prentice Hall, 2da edition, 2002.
- [9] E. Kreyszig. *Matemáticas Avanzadas para Ingeniería*, volume I. Limusa Wiley, 3ra edition, 2000.
- [10] E. Kreyszig. *Matemáticas Avanzadas para Ingeniería*, volume II. Limusa Wiley, 3ra edition, 2000.
- [11] D. Lindner. *Introducción a las señales y los sistemas*. McGraw Hill, 2002.
- [12] MathWorks. Matlab [online]. 1994 [visitado el 22 de noviembre de 2010]. URL <http://www.matlab.com>. 14, 215
- [13] U. Meyer-Baese. *Digital Signal Processing with Field Programmable Gate Arrays*. Signals and Communication Technology. Springer, 3ra edición edition, 2007.
- [14] A. Oppenheim, A. Willsky, and S. H. Nawab. *Señales y Sistemas*. Prentice Hall, 2da edition, 1998.

-
- [15] J. G. Proakis and D. G. Manolakis. *Tratamiento Digital de Señales*. Prentice Hall, 1998. c, 4, 22, 23, 70, 196
- [16] M. J. Roberts. *Señales y Sistemas. Análisis mediante métodos de transformada y MatLab*. McGraw Hill, 2005.
- [17] G. E. Shilov. *Elementary Real and Complex Analysis*. Dover Publications, Inc., 1973.
- [18] E. Soria Olivas, M. Martínez Sober, J. V. Francés Villora, and G. Camps Valls. *Tratamiento Digital de Señales. Problemas y ejercicios resueltos*. Prentice Hall, Madrid, 2003.
- [19] F. G. Stremler. *Introducción a los sistemas de comunicación*. Addison Wesley Longman, 1993.
- [20] Wikimedia. Wikipedia [online]. Diciembre 2005 [visitado el 22 de noviembre de 2010]. URL <http://en.wikipedia.org/wiki>.
- [21] U. Zölzer. *Digital Audio Signal Processing*. John Wiley and Sons, 2008.

Apéndice A

Octave y Matlab

MatLab[12] y Octave[5] son programas utilizados en procesamiento numérico de datos, dentro de lo cual figura el área de procesamiento digital de señales. Este apéndice es una brevísima introducción a la utilización del intérprete de ambos sistemas. Muchas de las funciones disponibles para MatLab deben ser instaladas por aparte para el Octave, y pueden ser encontradas en la base de datos del grupo que lo desarrolla [6].

La representación gráfica de señales de variable discreta en MatLab se realiza con la función `stem`. Por ejemplo:

```
nn = 0:30;           % genere un vector de 31 elementos: 0, 1, 2 .. 30
sinus = sin(nn/2+1); % para cada x en nn genere sin(x/2+1)
stem(nn,sinus);     % muestre valores en 'sinus' para las muestras 'nn'
```

El punto y coma ';' le indica al MatLab que no debe mostrar el resultado de la operación. Si es omitido, el resultado se mostrará directamente.

Con los comandos `hold` y `hold off` se puede controlar si se debe graficar sobre la misma imagen o no.

La función `zeros(m,n)` genera una matriz de `m` filas por `n` columnas. Así que para generar un impulso puede utilizarse

```
imp=zeros(20,1);
imp(5)=1;
```

Dependerá de los índices generados para esta secuencia el número de muestra para el impulso. Por ejemplo, con `nn=-5:14` el retardo del impulso es -1 pero con `nn=0:19` el retardo sería 4.

Para generar señales periódicas puede utilizarse el siguiente concepto:

```
x = [0;1;1;0;0;0] * ones(1,8);
x = x(:);
```

Lo que ocurre en la primera línea es generar una matriz con ocho filas todas idénticas al vector dado antes del símbolo '*'. La segunda línea convierte la matriz a un vector concatenando todas las filas, lo que equivale una señal periódica con $N = 6$ periodos.

Especial cuidado debe tenerse con la programación eficiente en Octave. El uso de ciclos debe ser evitado al máximo, y en vez de ello deben utilizarse las operaciones vectoriales en tanto se pueda.

Por ejemplo, supóngase que necesita generar una secuencia $x(n)$ con n desde -10 hasta 40 , que contenga un escalón unitario. Una implementación ingenua e ineficiente es:

```
n=[-10:40];
x=n;
for i=1:length(n)
    if n(i)<0
        x(i)=0;
    else
        x(i)=1;
    end
end
end
```

La versión anterior funciona, pero requiere mucho más tiempo que la versión utilizando los conceptos del lenguaje en cuestión:

```
n=[-10:40];
x=(n>=0);
```

Lo anterior genera un vector con 0 si el elemento correspondiente en n es menor que cero y 1 en caso contrario.

Un concepto fundamental para la programación eficiente es conocer la diferencia de sintaxis entre los operadores (por ejemplo "*"), y los operadores precedidos por un punto (por ejemplo ".*"). El primero representa una operación matricial, y requiere que las dimensiones de los operandos a la izquierda y derecha sean compatibles con la operación. El segundo caso aplica el operador elemento por elemento a las dos matrices, que deben para ello tener exactamente las mismas dimensiones.

Ejemplos:

```
[1 2 3]*[1 -1 -2] % Retorna un error, porque el primero o el segundo vector
                  % deberían ser un vector columna.
[1 2 3]*[1 -1 -2]' % El apóstrofe en el segundo vector indica que éste debe
                  % transponerse y por tanto es equivalente a un producto
                  % interno que devuelve el valor escalar -7
[1 2 3]'*[1 -1 -2] % Ahora esto equivale al producto externo y retorna una
                  % matriz 3x3.
[1 2 3].*[1 -1 -2] % Multiplique elemento por elemento para retornar [1 -2 -6]
```

Índice alfabético

- alias, 24, 155
- aliasing, 105
 - temporal, 137
- amplitud, 22
- análisis, 3
- analógica, 2
- analógico, 4
- analizador espectral, 91
- ancho de banda, 104
- anticausalidad, 42
- armónica
 - relación exponencial, 26
- armónicos, 92

- BIBO, 34

- canales, 2
- cascada, 34
- cero, 69
- cifrado, 8
- codificación, 153
- compresión, 8
- condición inicial, 28
- convolución, 37
- correlación, 57
 - autocorrelación, 57
 - cruzada, 57
 - normalizada, 58
- cuantificación
 - definición, 153
 - error, 153, 160
 - ruido, 160
- cuanto, 160

- deconvolución, 130
- densidad espectral
 - de energía, 95
 - de potencia, 92
 - de tensión, 92
- derrame, 151
- desplazamiento, 37
- DFT, 135
- digital, 2
- DTFS, 98

- ecuación
 - de diferencias, 45
- energía, 17
- espectro, 91
- estimación espectral, 91
- estructura
 - forma directa I, 183
 - forma directa II, 183
- exponencial
 - relacionadas armónicamente, 26

- fase, 22
- fenómeno de Gibbs, 201
- filtrado, 8
- filtro, 8, 117
 - paso todo, 127
 - peine, 124
 - ranura, 124
 - solapamiento y almacenamiento, 149
 - solapamiento y suma, 149
- forma
 - canónica, 184
 - directa I, 183
 - directa II, 183
- frecuencia, 22
 - angular, 22
 - de plegado, 155

- fundamental, 24
- normalizada, 25, 155
- fuga, 151
- función
 - de energía, 17
 - de potencia, 17
 - de transferencia, 76
- funcion
 - propia, 91
- Gibbs, 201
- Hilbert, 195
- identidad, 27
- identificación, 8
- interconexión
 - casca, 34
 - paralela, 34
- interpolación, 154
- memoria, 31
- muestra
 - número, 22
- muestreo
 - definición, 153
 - intervalo, 153
 - periódico, 154
 - teorema del, 104
 - uniforme, 154
- multiplicación, 37
- número de condición, 181
- osciladores, 129
- Paley-Wiener
 - Teorema, 193
- paralelo, 34
- Parseval
 - señales continuas periódicas, 92
- periodo
 - fundamental, 19, 23
- polinomio
 - característico, 51
- polo, 69
- potencia promedio, 17
- principio de incertidumbre, 96
- procesamiento, 3
- rango fundamental, 26
- reconocimiento, 8
- reflexión, 37
- reposo, 28
- resonador digital, 122
- respuesta
 - de estado permanente, 81
 - en fase, 111
 - en frecuencia, 111
 - en magnitud, 111
 - entrada cero, 48
 - estado nulo, 48
 - forzada, 48, 79
 - natural, 79
 - transitoria, 81
- retardo
 - de grupo, 118
- ruido, 8
- señal, 1
 - acotada, 18
 - anticausal, 34
 - asimétrica, 19
 - de energía, 17
 - de potencia, 17
 - impar, 19
 - multicanal, 2
 - par, 19
 - periódica, 19
 - simétrica, 19
 - tiempo discreto, 2
- simetría circular
 - impar, 143
 - par, 142
- sintetización, 8
- sistema, 3
 - de fase mínima, 133
 - de fase máxima, 133
 - de polos y ceros, 77
 - de todos ceros, 77

- de todos polos, 77
- discreto, 27
- estable, 34
- inestable, 34
- inverso, 130
- invertible, 130
- LTI, 117
- MWA, 184
- no recursivo, 46
- recursivo, 45
- solución
 - homogénea, 50
 - particular, 50
 - total, 54
- SQNR, 161
- submuestreo, 15
- suma, 37
- teorema
 - del muestreo, 104
 - Wiener-Khinchin, 109
- transformada
 - de Fourier, 93
- transformada z , 67
 - directa, 67
- transformada de Hilbert, 195
- vector, 2
- Wiener-Khinchin, 109