



Universidad de Navarra
Escuela Superior de Ingenieros

FUNDAMENTOS DE
CONTROL AUTOMÁTICO
DE SISTEMAS CONTINUOS Y MUESTREADOS

Dr. Jorge Juan Gil Nobajas
Dr. Ángel Rubio Díaz-Cordovés
San Sebastián, 15 de agosto de 2010

Fundamentos de Control Automático de Sistemas Continuos y Muestreados
©2009 Jorge Juan Gil Nobajas y Ángel Rubio Díaz-Cordovés

ISBN 978-84-613-4618-9
Depósito Legal SS-1094-2009

Reservados todos los derechos.
Queda prohibida la reproducción total o parcial sin autorización previa.

Impreso en España
Imprime: Unicopia, C.B.
Paseo Manuel Lardizábal, 13
20018 San Sebastián (Guipúzcoa) ESPAÑA

Índice general

I	Control de sistemas continuos	9
1.	Introducción	11
1.1.	Definiciones	11
1.2.	Tipos de sistemas de control	12
2.	La transformada de Laplace	15
2.1.	Definición y propiedades	15
2.2.	Transformada de Laplace de funciones elementales	16
2.3.	Transformada inversa de Laplace	17
2.4.	Resolución de ecuaciones diferenciales	18
2.5.	Ejercicios resueltos	18
3.	Representación de los sistemas	21
3.1.	Generalidades	21
3.2.	Función de transferencia de un sistema	22
3.3.	Modelos de sistemas físicos	23
3.3.1.	Sistemas mecánicos	23
3.3.2.	Sistemas eléctricos	26
3.3.3.	Sistemas electromecánicos	28
3.3.4.	Sistemas hidráulicos	29
3.3.5.	Sistemas térmicos	30
3.4.	Diagrama de bloques de un sistema	31
3.4.1.	Reglas para la simplificación de diagramas de bloques	31
3.4.2.	Ejemplo de circuito con dos mallas	32
3.4.3.	Ejemplo de motor de corriente continua	33
3.5.	Sistema de realimentación negativa no unitaria	34
3.6.	Sistema de realimentación negativa unitaria	36
3.7.	Ejercicios propuestos	36
4.	Respuesta temporal	39
4.1.	Sistemas de primer orden	39
4.1.1.	Respuesta ante entrada impulso	39
4.1.2.	Respuesta ante entrada escalón	40
4.1.3.	Respuesta ante entrada sinusoidal	41
4.1.4.	Ejemplos de sistemas de primer orden	41
4.2.	Sistemas de segundo orden	42
4.2.1.	Respuesta subamortiguada ante entrada escalón	43
4.2.2.	Respuesta sobreamortiguada ante entrada escalón	44
4.2.3.	Respuesta críticamente amortiguada ante entrada escalón	45
4.2.4.	Respuesta oscilatoria ante entrada escalón	45
4.2.5.	Respuesta ante entrada impulso	45
4.3.	Sistemas de orden superior	46
4.4.	Influencia de los ceros	47
4.5.	Ejercicios propuestos	48

5. Error en régimen permanente	51
5.1. Definición de error en régimen permanente	51
5.2. Error en sistemas con realimentación negativa unitaria	51
5.2.1. Error de posición	52
5.2.2. Error de velocidad	52
5.2.3. Error de aceleración	52
5.2.4. Resumen de errores	52
5.3. Magnitud y unidades del error	53
5.4. Error en sistemas con realimentación no unitaria	54
5.5. Error en sistemas con varias entradas	54
5.6. Ejercicios resueltos	55
5.7. Ejercicios propuestos	56
6. Estabilidad	61
6.1. Definición de estabilidad	61
6.2. Criterio de Routh-Hurwitz	61
6.2.1. Estabilidad de los sistemas de segundo orden	62
6.2.2. Estabilidad de los sistemas de tercer orden	63
6.2.3. Ejemplo numérico de sistema de cuarto orden	63
6.3. Casos especiales del criterio de Routh-Hurwitz	63
6.3.1. Se anula el primer coeficiente de una fila	64
6.3.2. Se anula toda una fila	64
6.4. Ejercicios resueltos	65
6.5. Ejercicios propuestos	66
7. Lugar de las raíces	67
7.1. Introducción	67
7.2. Generalidades del método	68
7.3. Método para dibujar el lugar de las raíces	68
7.3.1. Polos y ceros en lazo abierto	68
7.3.2. Asíntotas	69
7.3.3. Puntos del eje real que pertenecen al lugar de las raíces	69
7.3.4. Puntos de ruptura	69
7.3.5. Puntos de corte con el eje imaginario	69
7.3.6. Ángulos de salida y llegada	70
7.4. Cálculo de la ganancia	70
7.5. Ejemplos de lugares de las raíces	70
7.5.1. Sistema de tercer orden	70
7.5.2. Sistema de segundo orden con un cero	71
7.6. Estabilidad relativa	74
7.6.1. Margen de ganancia	74
7.6.2. Margen de fase	75
7.7. Lugar de las raíces en función de otros parámetros	75
7.8. Ejercicios propuestos	76
8. Respuesta en frecuencia	79
8.1. Respuesta a una entrada sinusoidal	79
8.2. El diagrama de Bode	79
8.3. Diagramas de Bode de sistemas elementales	80
8.3.1. Ganancia	80
8.3.2. Retraso en el tiempo	80
8.3.3. Integrador	81
8.3.4. Derivador	81
8.3.5. Polo simple estable	82
8.3.6. Cero simple con parte real negativa	83
8.3.7. Polos estables complejos conjugados	84
8.3.8. Ceros complejo conjugados	85
8.3.9. Polo simple con parte real positiva	86
8.3.10. Cero simple con parte real positiva	86

8.4.	Diagrama de Bode de cualquier función de transferencia	87
8.5.	Diagrama de Bode de un sistema en lazo cerrado	88
8.5.1.	Ancho de banda	88
8.5.2.	Margen de fase y margen de ganancia	89
8.6.	Ejercicios propuestos	89
9.	Compensadores de adelanto y de retraso de fase	91
9.1.	Generalidades	91
9.1.1.	Especificaciones	91
9.1.2.	Tipos de compensación	91
9.1.3.	Método de ajuste	92
9.2.	Compensador de adelanto de fase	92
9.2.1.	Ajuste por el lugar de las raíces	93
9.2.2.	Ajuste por el diagrama de Bode	97
9.3.	Compensador de retraso de fase	101
9.3.1.	Ajuste por el diagrama de Bode	101
9.3.2.	Ajuste por el lugar de las raíces	104
9.4.	Compensador de adelanto-retraso	105
9.5.	Ejercicios propuestos	106
10.	Controladores PID	111
10.1.	Expresión general	111
10.1.1.	Forma estándar	111
10.1.2.	Forma paralela	112
10.1.3.	Forma serie	112
10.2.	Sentido físico de la actuación de un PID	112
10.2.1.	Actuación proporcional	113
10.2.2.	Actuación proporcional-derivativa	113
10.2.3.	Actuación proporcional-integral	114
10.3.	Ajuste experimental de PID	115
10.3.1.	Ajuste de Ziegler-Nichols	115
10.3.2.	Otros ajustes experimentales	117
10.3.3.	Ejemplo comparativo	117
10.4.	Ajuste analítico de PID's por asignación de polos	118
10.5.	Control con dos grados libertad	119
10.6.	Modificaciones del PID	120
10.6.1.	Supresión del efecto kick-off	120
10.6.2.	Set-point weighting	121
10.6.3.	Filtro de la derivada	121
10.6.4.	Prevención del efecto windup integral	122
10.7.	Ejercicios propuestos	123
11.	Control en espacio de estado	127
11.1.	Introducción	127
11.2.	Tipos de variables de estado	128
11.2.1.	Variables de fase	128
11.2.2.	Variables canónicas o normales	129
11.2.3.	Variables físicas	130
11.3.	Controlabilidad y observabilidad	131
11.4.	Realimentación completa de estados	131
11.4.1.	Asignación de polos	132
11.4.2.	Método de Ackermann	132
11.4.3.	Controlador óptimo cuadrático	132
11.5.	Realimentación parcial de estados	133
11.6.	Observadores de estado	133
11.7.	Realimentación completa de estados observados	134
11.8.	Ejercicios propuestos	134

II	Control de sistemas muestreados	137
12.	Introducción	139
12.1.	Ejemplo de implementación analógica	139
12.2.	Ejemplo de implementación digital	140
12.3.	Concepto de muestreo	140
12.4.	Concepto de cuantización	141
12.5.	Clasificación de los sistemas	141
13.	Tratamiento matemático de la señal muestreada	143
13.1.	Definición de muestreo periódico	143
13.1.1.	Función portadora	143
13.1.2.	Función temporal muestreada	145
13.2.	Transformada de Fourier de la función muestreada	145
13.3.	El problema del <i>aliasing</i>	146
13.3.1.	Teorema de Shannon	146
13.3.2.	<i>Aliasing</i> y reconstrucción de la señal original	147
13.3.3.	<i>Aliasing</i> y ruido en la medida de la señal	148
14.	El muestreo ideal	149
14.1.	Definición de muestreo ideal	149
14.1.1.	Función portadora	149
14.1.2.	Función temporal muestreada	150
14.2.	Transformada de Fourier de la función muestreada	150
14.3.	Transformada de Laplace de la función muestreada	151
14.3.1.	Forma cerrada y región de convergencia	151
14.3.2.	Forma alternativa para la transformada de Laplace	152
14.3.3.	Periodicidad de la transformada de Laplace	152
14.3.4.	Franjas primaria y complementarias	152
15.	Reconstrucción de la función continua original	153
15.1.	Filtro ideal	153
15.1.1.	Características del filtro ideal	153
15.1.2.	Imposibilidad física de construcción del filtro ideal	153
15.1.3.	Reconstrucción de la señal con el filtro ideal	154
15.2.	Retenedor de orden cero	155
15.2.1.	Características del retenedor de orden cero	155
15.2.2.	Expresión de Laplace del retenedor de orden cero	156
15.2.3.	Respuesta en frecuencia del retenedor de orden cero	156
15.3.	Retenedor de primer orden	157
15.3.1.	Características del retenedor de primer orden	157
15.3.2.	Expresión de Laplace del retenedor de primer orden	158
15.3.3.	Respuesta en frecuencia del retenedor de primer orden	158
15.4.	Retenedor polinomial	159
15.4.1.	Características del retenedor polinomial	159
15.4.2.	Expresión de Laplace del retenedor polinomial	159
16.	La transformada Zeta	161
16.1.	Cálculo de la transformada Zeta	161
16.2.	Tabla de la transformada Zeta de funciones elementales	162
16.3.	Teoremas de la transformada Zeta	162
16.4.	Cálculo de la transformada inversa de Zeta	162
16.4.1.	Método directo o de la expansión de potencia	163
16.4.2.	Método de la expansión en fracciones	164
16.5.	Función de transferencia Zeta	164
16.6.	Ecuaciones en diferencias	165

17. Diagramas de bloques en Zeta	167
17.1. Generalidades	167
17.2. Bloques en cascada con muestreadores	167
17.2.1. Un único bloque continuo	167
17.2.2. Bloques continuos con muestreador intermedio	168
17.2.3. Bloques continuos sin muestreador intermedio: el problema de la convolución	169
17.2.4. Sistemas en lazo cerrado	169
17.3. Método de simplificación	171
17.4. Sistemas con bloques continuos y discretos	172
18. Correspondencia entre el plano S y el plano Z	173
18.1. Franja primaria y círculo unitario	173
18.2. Líneas de parámetros constantes	174
18.3. Variación de la posición de los polos y ceros con T	174
18.4. Cálculo del número de muestras por ciclo	176
19. Análisis de estabilidad	179
19.1. Criterio general	179
19.2. Criterio de Jury	179
19.3. Transformación bilineal y criterio de Routh-Hurwitz	180
19.4. Ejemplo	181
20. Respuesta transitoria y régimen permanente	183
20.1. Respuesta transitoria	183
20.2. Régimen permanente	185
20.3. Error en régimen permanente	185
20.4. Tipo de sistema	186
21. Lugar de las raíces	187
21.1. Definición	187
21.2. Punto de partida	187
21.3. Método gráfico	189
21.4. Diseño de compensadores de adelanto de fase	190
21.5. Ejemplo de diseño	190
21.6. Ejercicios propuestos	194
22. Métodos de digitalización	197
22.1. Generalidades de los métodos de digitalización	197
22.2. Integración numérica	197
22.2.1. Método trapezoidal o de Tustin	198
22.2.2. Método de Euler implícito	198
22.2.3. Método de Euler explícito	199
22.2.4. Otros métodos numéricos de integración	200
22.2.5. Ejemplo de digitalización usando integración numérica	200
22.3. Derivación numérica	201
22.3.1. Método de <i>backwards</i>	201
22.3.2. Otros métodos de derivación	202
22.3.3. Ejemplos de digitalización de PID	203
22.3.4. Ejemplos de digitalización de filtros	203
22.4. Método de equiparación de polos y ceros	204
22.4.1. Caso particular	204
22.4.2. Método de equiparación modificado	205
22.4.3. Ejemplo	205
22.5. Método de la equivalencia del retenedor	206
23. Respuesta en frecuencia	207
23.1. Aproximación de la respuesta en frecuencia	207
23.2. Ejemplo numérico	208
23.3. Respuesta en frecuencia exacta	210

24. Espacio de estado muestreado	211
24.1. Introducción	211
24.2. Ejemplo de modelización	211
24.3. Control mediante realimentación completa de estados	212
A. Ampliación de espacio de estado	215
A.1. Matriz de transición de estados	215

Parte I

Control de sistemas continuos

Capítulo 1

Introducción

La *ingeniería de control* formula leyes matemáticas para el gobierno de sistemas físicos conforme a una serie de requerimientos o especificaciones. La aplicación de estas leyes convierte al sistema físico en un *sistema controlado* que, o bien posee una dinámica mejorada, o bien se ha convertido en un automatismo, es decir, un sistema que es capaz de “auto-conducirse” siguiendo una consigna de referencia.

Esta disciplina es esencial para la automatización de procesos industriales y brinda los medios adecuados para lograr el funcionamiento óptimo de cualquier sistema dinámico. Resulta muy conveniente que los ingenieros posean un amplio conocimiento de esta materia.

La Parte I del presente libro describe las herramientas clásicas para el control de sistemas continuos en el tiempo, es decir, aquellos sistemas en los que se puede medir y actuar en todo instante. En electrónica, este tipo de sistemas se llaman analógicos. La Parte II estudiará los discretos o digitales, utilizando muchas de las herramientas que se describieron en la primera parte.

1.1. Definiciones

En esta disciplina se emplea mucho la palabra **sistema**, que se puede definir como una combinación de elementos que actúan conjuntamente y cumplen un determinado objetivo. A veces se usará la expresión sistema dinámico, es decir, que evoluciona a lo largo del tiempo.

Sistema es un término muy general, que puede aplicarse casi a cualquier realidad física. Así, un vehículo impulsado por un motor de combustión interna es un sistema. Pero también el conjunto de elementos que regulan la temperatura de un edificio es un sistema. El primer ejemplo (el vehículo) sería un sistema que goza de una cierta unidad. En cambio el segundo ejemplo (el sistema de calefacción de un edificio) es un sistema distribuido, es decir, sus elementos están repartidos en distintos lugares (los sensores de temperatura, la caldera, los radiadores e incluso el volumen de aire que se pretende calentar forman parte del sistema).

Por otro lado, todos y cada uno de los elementos que componen un sistema pueden ser considerados, en sí mismos, como sistemas. El vehículo antes citado es un sistema de locomoción, pero el motor de combustión interna que lo impulsa puede ser considerado, en sí mismo, como un sistema al margen del vehículo.

Para evitar equívocos, se reservará el término **planta** para el designar el sistema que se desea controlar. Desde el punto de vista del control automático, es muy importante identificar la planta, es decir, el sistema físico que se pretende controlar. Así, y siguiendo con el ejemplo del vehículo, es muy distinto pretender controlar la velocidad del vehículo para que sea 120 km/h constante (cualquiera que sea la pendiente de la carretera), que controlar la temperatura interior del automóvil para que sea 22°C constante (cualquiera que sea la temperatura exterior). De alguna forma, previamente a abordar el problema de control, hay que formular matemáticamente las ecuaciones diferenciales que describen la dinámica de la planta. El apartado 3.3 se dedica precisamente a la formulación de ecuaciones diferenciales de algunos sistemas físicos elementales.

El concepto de **perturbación** está incluido también en los ejemplos de control apuntados anteriormente. La pendiente de la carretera para el control de la velocidad y la temperatura exterior para el control de la temperatura interior son perturbaciones de dichos sistemas, es decir, agentes físicos que el ingeniero no puede controlar o modificar pero sí influyen en la variable física que se quiere gobernar. La propia planta podría contener elementos variables que modifiquen la respuesta del sistema. En este caso, en lugar de perturbaciones se denominarán **incertidumbres**. Por ejemplo, la masa del vehículo dismi-

nuye conforme se consume el combustible, influyendo en la velocidad del vehículo. Es una incertidumbre de la planta, no una perturbación exterior.

También se han mencionado en los ejemplos una **referencia** para el sistema controlado (los 120 km/h para la velocidad o los 22°C para la temperatura). Evidentemente se trata de una señal o magnitud física que se desea para el sistema controlado. Un sistema bien controlado seguirá con fidelidad la consigna de la referencia. Un sistema mal controlado no será capaz de alcanzar la referencia o la seguirá con un error considerable.

El **error** que existe entre la referencia deseada y la respuesta real del sistema también es un elemento de interés en los sistemas controlados. De hecho, el error puede ser objeto de especificación. Por ejemplo, el ingeniero puede dar como satisfactorio un sistema controlado que sea capaz de alcanzar la referencia con un error menor que un determinado umbral.

El elemento central del sistema de control es lo que denomina **controlador**. A veces se le llama compensador o regulador. Este elemento es el encargado de actuar en la planta en función de la referencia o del error, para conseguir que dicha planta se comporte de acuerdo con las especificaciones de diseño. En el ejemplo del climatizador del vehículo, en función de la temperatura interior y la referencia deseada, el controlador introducirá aire frío o caliente en el habitáculo. En el ejemplo del control de la velocidad de crucero, en función de la velocidad actual del vehículo y de la referencia deseada, el controlador inyectará en el carburador más o menos caudal de gasolina o incluso puede activar el sistema de frenado. Hablando impropriadamente, el controlador es el elemento “inteligente” del sistema. La existencia de este elemento de gobierno es lo que hace que un sistema sea *automático* y no *manual*. Cuando un controlador consigue su objetivo a pesar de las incertidumbres de la planta, se dice que el controlador es robusto.

1.2. Tipos de sistemas de control

Con los elementos enunciados en el apartado anterior, es posible dibujar —de forma cualitativa— cómo funciona un sistema de control. De momento, los distintos elementos del sistema se representarán con nubes o “cajas negras” y las señales con flechas.

Un sistema controlado **en lazo abierto** es aquel cuyo controlador actúa sólo en función de la referencia deseada para la respuesta del sistema (Fig. 1.1). Este tipo de control se puede emplear si las perturbaciones sobre el sistema son pequeñas y se tiene un buen modelo de planta. También se utiliza este tipo de control si la señal de salida del sistema es imposible o muy difícil de medir.

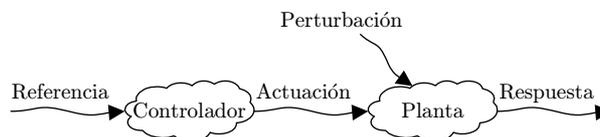


Figura 1.1: Sistema controlado en lazo abierto

En los sistemas controlados **en lazo cerrado** la variable controlada se mide y se utiliza para modificar la actuación sobre la planta. Evidentemente, para realizar esta medida se necesita un sensor. A priori, un sistema controlado en lazo cerrado es más complejo y caro que un sistema controlado en lazo abierto. La forma más habitual de “cerrar el lazo” en los sistemas de control es calcular el error entre la referencia y la respuesta actual del sistema. El controlador actúa entonces en función del error (Fig. 1.2). El concepto de **realimentación** (*feedback*) es uno de los principios básicos del control automático.

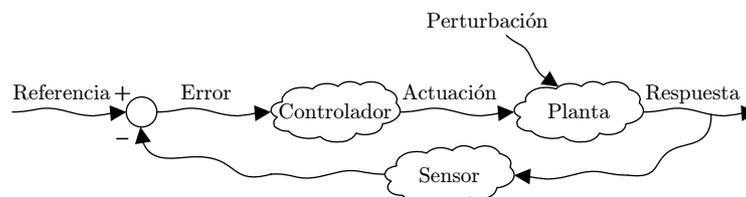


Figura 1.2: Sistema controlado en lazo cerrado

Al margen de la arquitectura que tenga la ley de control, el sistema de control en el que se hace especial hincapié a la capacidad del sistema de seguir los cambios de la referencia se le denomina **servosistema**. En cambio, el sistema de control en el que se hace especial hincapié a la capacidad del sistema de rechazar las perturbaciones exteriores y mantener una referencia constante, o que cambia muy lentamente, se le denomina **regulador**.

Sin pretender realizar una clasificación exhaustiva de los tipos de sistemas de control, merece la pena señalar algunos de ellos, y especificar cuáles se tratarán en este libro. Así por ejemplo, si se atiende a la variación en el tiempo de la ley de control se puede distinguir entre:

- **Control fijo o estándar:** La ley de control no varía en el tiempo. Es interesante la planta es fija. Como ya se ha apuntado en el apartado anterior, se llama control robusto a aquel que funciona correctamente ante errores en la modelización o incertidumbres de la planta.
- **Control adaptable (*gain scheduling*):** La ley de la planta cambia, y se puede decidir para cada ley un controlador distinto. Aquí se selecciona una ley de control como se ve en la Fig. 1.3.
- **Control adaptativo (*adaptive control*):** El controlador cambia gradualmente en función de una estimación de la planta que se actualiza en todo momento. Este tipo de estrategia es adecuada para aquellos sistemas en los que el modelo de la planta varía mucho a lo largo del tiempo.

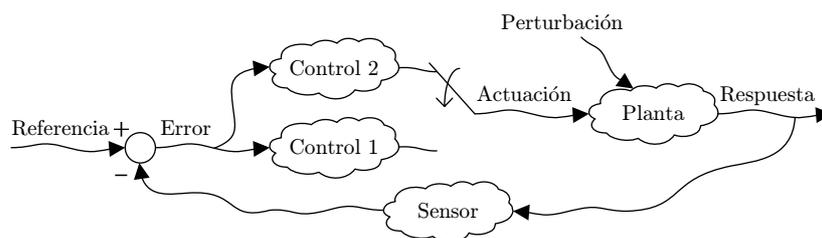


Figura 1.3: Sistema de control adaptable

Si se atiende al número de entradas y de salidas que posee el sistema, se denominan sistemas **SISO** (*single input, single output*) los que poseen una única entrada y una salida, y sistemas **MIMO** (*multiple input, multiple output*) si poseen varias entradas y varias salidas.

Si las ecuaciones en diferenciales que describen el sistema, tanto la planta como el controlador, son lineales entonces todo el sistema de control se denomina **lineal**. En cambio, si falta la propiedad de la linealidad en la planta o el controlador, todo el sistema será **no lineal**.

Por otro lado, la división de este libro en dos grandes partes alude a otra posible clasificación de los sistemas de control: los sistemas **continuos**, en los que la ley de control posee información de la planta y actúa en todo instante de tiempo. Y los sistemas **muestreados o discretos** en los que la ley de control recibe información y actúa en determinados instantes que suele imponer un reloj. La Parte II del presente manual se dedica al estudio de este último tipo de sistemas.

Finalmente, si el sistema está descrito por ecuaciones diferenciales ordinarias se llama sistema de **parámetros concentrados**, pero si está descrito por medio de ecuaciones diferenciales en derivadas parciales se llama de **parámetros distribuidos**. Un ejemplo de este último tipo de sistemas puede ser el control de la transmisión de calor a través de una superficie, o el control de la vibración de un punto de una membrana.

Capítulo 2

La transformada de Laplace

En el año 1782 Pierre Simon Laplace estudió la transformación integral que lleva su nombre. Sin embargo, no es hasta el periodo de 1880-1887 cuando Oliver Heaviside la aplica para la resolución de ecuaciones diferenciales. Dado que en ingeniería de control se usa mucho esta herramienta matemática, el presente capítulo resume sus principales propiedades.

2.1. Definición y propiedades

Se define la transformada de Laplace $F(s)$ de una determinada función temporal $f(t)$ como:

$$F(s) = \mathcal{L}[f(t)] = \int_0^{\infty} f(t)e^{-ts} dt \quad (2.1)$$

Donde $f(t)$ es una función real de variable real, generalmente el tiempo, y su transformada de Laplace $F(s)$ es una función compleja de variable compleja. Para que exista la transformada de Laplace es suficiente que la integral exista para algún valor s complejo. Se reservarán las letras minúsculas para las funciones temporales y las mayúsculas para sus transformadas de Laplace.

$$f(t) \xrightarrow{\mathcal{L}} F(s) \quad (2.2)$$

La transformada de Laplace no existe para cualquier función temporal $f(t)$. Una condición suficiente —pero no necesaria— de existencia, es que $f(t)$ sea seccionalmente continua en $[0, T]$, $\forall T > 0$, y que sea de orden exponencial cuando $t \rightarrow \infty$, es decir, que $\exists M, T > 0$ y $\alpha \in \mathbb{R} / |f(t)| < Me^{\alpha t} \forall t > T$. Las funciones que cumplen esta condición suficiente se suelen decir que pertenecen al conjunto \mathcal{A} , es decir, $f(t) \in \mathcal{A}$.

Como la integral (2.1) se extiende desde cero hasta infinito, dos funciones cualesquiera que difieran únicamente en valores de tiempo negativos, poseen la misma transformada de Laplace. Es decir, los valores de $f(t)$ para t negativos, no influyen en la transformada de Laplace. Para que la relación entre una función y su transformada de Laplace sea biunívoca, a partir de ahora sólo se considerarán *funciones causales*, es decir, aquellas que son nulas para tiempos negativos, $f(t) = 0 \forall t < 0$, y toman valores finitos en tiempos positivos. Para funciones $f(t)$ causales y continuas¹ para $t > 0$, entonces la relación entre $f(t)$ y $F(s)$ es biunívoca, es decir, que para toda $f(t)$ existe una única $F(s)$ y viceversa.

La variable compleja s tiene en módulo unidades de rad/s. Pero si el número complejo lo dividimos en parte real y parte imaginaria, se puede considerar que tiene unidades de rad/s sobre el eje imaginario y de s^{-1} sobre el eje real. Se observa en la definición de la transformada de Laplace,

$$e^{-ts} = e^{-t(a+bj)} = e^{-ta} | \underline{-tb} \quad (2.3)$$

que el exponente del módulo del número complejo es adimensional si consideramos que a , que es la parte real de la variable compleja s , tiene unidades de s^{-1} . El argumento tendrá unidades de radianes si b , que es la parte imaginaria de la variable compleja s , tiene unidades de rad/s.

En la Tabla 2.1 se resumen las principales propiedades de la transformada de Laplace. La propiedad de la linealidad existe si $f(t)$ y $g(t)$ poseen transformada de Laplace. La propiedad de la derivación real

¹Para Laplace no sería estrictamente necesario que fueran continuas, porque afirma que dos funciones que difieran en un número finito o infinito de puntos aislados deben considerarse iguales.

se da si $f(t)$ es continua en el intervalo $(0, \infty)$, $f(t)$ es de orden exponencial cuando $t \rightarrow \infty$ y la $f'(t)$ es seccionalmente continua en $[0, T]$, $\forall T > 0$. El resto de propiedades se dan simplemente si $f(t) \in \mathcal{A}$. Por lo general, estas condiciones rara vez se tienen en cuenta ya que las variables físicas que se manejan en ingeniería de control son casi siempre funciones causales.

Tabla 2.1: Propiedades de la transformada de Laplace

Propiedad	Expresión
Linealidad	$\mathcal{L}[\alpha f(t) + \beta g(t)] = \alpha F(s) + \beta G(s)$
Integración real	$\mathcal{L}[\int_0^t f(\tau) d\tau] = \frac{F(s)}{s}$
Derivación real	$\mathcal{L}\left[\frac{df(t)}{dt}\right] = sF(s) - f(0^+)$
Valor final	$\lim_{t \rightarrow \infty} f(t) = \lim_{s \rightarrow 0} sF(s)$
Valor inicial	$\lim_{t \rightarrow 0^+} f(t) = \lim_{s \rightarrow \infty} sF(s)$
Traslación en el tiempo	$\mathcal{L}[f(t - a)] = e^{-as} F(s)$
Traslación en Laplace	$\mathcal{L}[e^{-as} f(t)] = F(s + a)$
Convolución	$\mathcal{L}[f(t) \otimes g(t)] = F(s)G(s)$
Escalado en el tiempo	$\mathcal{L}[f(\frac{t}{\alpha})] = \alpha F(\alpha s)$

Conviene señalar que la traslación de una función en el tiempo hace que aparezcan los valores nulos de la función causal en tiempos positivos (ver Fig. 2.1). Este hecho se suele olvidar y es fuente de importantes errores.

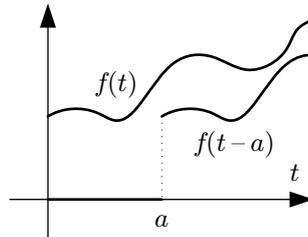


Figura 2.1: Función trasladada en el tiempo

2.2. Transformada de Laplace de funciones elementales

En este apartado se calculan las transformadas de Laplace de algunas funciones elementales. La función *escalón unidad* $u(t)$ se define como:

$$u(t) = \begin{cases} 1 & \text{para } t \geq 0 \\ 0 & \text{para } t < 0 \end{cases} \quad (2.4)$$

Su transformada de Laplace se obtiene por definición:

$$U(s) = \mathcal{L}[u(t)] = \int_0^{\infty} e^{-ts} dt = \left[\frac{e^{-ts}}{-s} \right]_0^{\infty} = \frac{1}{s} \quad (2.5)$$

Para el caso de la función *pulso de área unidad* $p(t)$, también por definición:

$$p(t) = \begin{cases} \frac{1}{\alpha} & \text{para } 0 \leq t < \alpha \\ 0 & \text{resto} \end{cases} \quad (2.6)$$

$$P(s) = \mathcal{L}[p(t)] = \int_0^{\alpha} \frac{1}{\alpha} e^{-ts} dt = \frac{1}{\alpha} \left[\frac{e^{-ts}}{-s} \right]_0^{\alpha} = \frac{1 - e^{-\alpha s}}{\alpha s} \quad (2.7)$$

La función *impulso unidad* $\delta(t)$ se define como:

$$\delta(t) = \begin{cases} \infty & \text{para } t = 0 \\ 0 & \text{resto} \end{cases}, \text{ siendo } \int_{-\infty}^{\infty} \delta(t) dt = 1 \quad (2.8)$$

En este caso, su transformada de Laplace se puede obtener como límite de la función pulso de área unidad, cuando el parámetro α tiende a cero, es decir,

$$\Delta(s) = \mathcal{L}[\delta(t)] = \lim_{\alpha \rightarrow 0} \frac{1 - e^{-\alpha s}}{\alpha s} = 1, \quad (2.9)$$

donde se ha empleado el teorema de l'Hôpital para el cálculo del límite. Otra forma de obtener este mismo resultado es considerar función escalón unidad se obtiene integrando la función impulso unidad:

$$U(s) = \mathcal{L}[u(t)] = \frac{1}{s} = \mathcal{L}\left[\int_0^t \delta(\tau) d\tau\right] = \frac{\Delta(s)}{s} \implies \Delta(s) = 1 \quad (2.10)$$

En cualquier caso, la función impulso unidad $\delta(t)$ es un poco especial, y el camino inverso al que se ha usado en la demostración anterior no funciona:

$$\Delta(s) = \mathcal{L}[\delta(t)] = \mathcal{L}\left[\frac{du(t)}{dt}\right] = s\frac{1}{s} - u(0^+) = 1 - 1 = 0 \text{ ¡falso!} \quad (2.11)$$

Este hecho no puede sorprender ya que la función impulso es la derivada de la función escalón en un sentido impropio. Por otro lado, no se cumplen las condiciones señaladas para poder aplicar la propiedad de la derivación real.

Tabla 2.2: Transformadas de las entradas habituales en los sistemas

Función	$f(t)$	$F(s)$
Impulso unidad	$\delta(t)$	1
Escalón unidad	$u(t) = 1$, para $t > 0$	$\frac{1}{s}$
Rampa unidad	$r(t) = t$, para $t > 0$	$\frac{1}{s^2}$
Aceleración unidad	$a(t) = \frac{1}{2}t^2$, para $t > 0$	$\frac{1}{s^3}$

Señaladas estas advertencias matemáticas, las funciones que más se emplean como entradas en los sistemas controlados son precisamente aquellas que se obtienen al ir integrando sucesivamente la función impulso unidad, como se observa en la Tabla 2.2. En la Tabla 2.3 se muestran las transformadas de otras funciones, definidas para tiempos positivos.

Tabla 2.3: Transformadas de Laplace de diversas funciones

$f(t)$	$F(s)$	$f(t)$	$F(s)$
e^{-at}	$\frac{1}{s+a}$	t^{k-1}	$\frac{(k-1)!}{s^k}$
te^{-at}	$\frac{1}{(s+a)^2}$	$e^{-at} - e^{-bt}$	$\frac{b-a}{(s+a)(s+b)}$
$t^{k-1}e^{-at}$	$\frac{(k-1)!}{(s+a)^k}$	$\sin at$	$\frac{a}{s^2+a^2}$
$1 - e^{-at}$	$\frac{a}{s(s+a)}$	$\cos at$	$\frac{s}{s^2+a^2}$
$t - \frac{1-e^{-at}}{a}$	$\frac{a}{s^2(s+a)}$	$e^{-at} \sin bt$	$\frac{b}{(s+a)^2+b^2}$
$1 - (1+at)e^{-at}$	$\frac{a^2}{s(s+a)^2}$	$e^{-at} \cos bt$	$\frac{s+a}{(s+a)^2+b^2}$

2.3. Transformada inversa de Laplace

El proceso matemático de pasar de la expresión matemática en el dominio de Laplace a la expresión en el dominio del tiempo se denomina transformada inversa de Laplace. Su expresión es:

$$f(t) = \mathcal{L}^{-1}[F(s)] = \lim_{\omega \rightarrow \infty} \left[\frac{1}{2\pi j} \int_{c-j\omega}^{c+j\omega} F(s)e^{ts} ds \right], \quad (2.12)$$

donde c es cualquier constante real mayor que la parte real de cualquier polo de $F(s)$.

Evaluar la integral (2.12) puede ser bastante complicado por lo que se suele calcular acudiendo a la Tabla 2.3. Si en la tabla no se encuentra una determinada función $F(s)$, se recomienda descomponerla en funciones simples en s , de las cuales sí se conozcan sus transformadas inversas. Como las funciones de Laplace que se van a utilizar suelen ser fracciones de polinomios en s , el cálculo de transformadas inversas se reduce a dividir estas expresiones en fracciones simples.

$$F(s) = \frac{s^2 + 2s + 3}{(s + 1)^3} \quad (2.13)$$

Como ejemplo, se va a calcular la función temporal de la función de Laplace $F(s)$ de la ecuación (2.13). Lo primero que se hace es dividir la única fracción en tres simples:

$$F(s) = \frac{A}{(s + 1)^3} + \frac{B}{(s + 1)^2} + \frac{C}{s + 1} = \frac{A + B(s + 1) + C(s + 1)^2}{(s + 1)^3} \quad (2.14)$$

Las constantes A , B y C se calculan igualando coeficientes de los polinomios del numerador. También es posible obtenerlos igualando los numeradores después de dar un valor numérico a la variable s . Los valores numéricos más adecuados son las raíces de distintos monomios. De esta forma es posible determinar más rápidamente las constantes. Para el caso anterior los valores de A , B y C son respectivamente 2, 0 y 1. Entonces:

$$F(s) = \frac{2}{(s + 1)^3} + \frac{1}{s + 1} \quad (2.15)$$

$$f(t) = t^2 e^{-t} + e^{-t} = e^{-t}(1 + t^2), \text{ para } t > 0 \quad (2.16)$$

2.4. Resolución de ecuaciones diferenciales

En este apartado se utiliza la transformada de Laplace para resolver ecuaciones diferenciales lineales. Sea la siguiente ecuación diferencial

$$a_0 f(t) + a_1 \frac{df(t)}{dt} + a_2 \frac{d^2 f(t)}{dt^2} = b_0 r(t) + b_1 \frac{dr(t)}{dt}, \quad (2.17)$$

donde las condiciones iniciales son:

$$f(0^+) = c_0, \quad \frac{df(0^+)}{dt} = c_1, \quad r(0^+) = d_0 \quad (2.18)$$

Se aplica la transformada de Laplace a los dos miembros de la ecuación:

$$a_0 F(s) + a_1 [sF(s) - c_0] + a_2 [s^2 F(s) - c_0 s - c_1] = b_0 R(s) + b_1 [sR(s) - d_0] \quad (2.19)$$

$$F(s) = \frac{b_0 + b_1 s}{a_0 + a_1 s + a_2 s^2} R(s) + \frac{a_1 c_0 + a_2 c_1 - b_1 d_0 + a_2 c_0 s}{a_0 + a_1 s + a_2 s^2} \quad (2.20)$$

La ecuación diferencial (2.17) se convierte en una ecuación algebraica en el dominio de Laplace. De esta forma es muy sencillo obtener la solución (2.20) a la ecuación diferencial, también en el dominio de Laplace. La solución en el dominio del tiempo se puede obtener calculando la transformada inversa de Laplace de $F(s)$, conocida la función $r(t)$.

2.5. Ejercicios resueltos

- **Ejercicio 1:** Obtener la función $x(t)$ que cumple la ecuación diferencial con condiciones iniciales no nulas:

$$\frac{d^2 x(t)}{dt^2} + 3 \frac{dx(t)}{dt} + 2x(t) = 0, \quad \begin{cases} x(0^+) = a \\ \frac{dx(0^+)}{dt} = b \end{cases} \quad (2.21)$$

Aplicando la transformada de Laplace a la ecuación diferencial,

$$s^2 X(s) - sa - b + 3[sX(s) - a] + 2X(s) = 0, \quad (2.22)$$

la solución en el dominio de Laplace es:

$$X(s) = \frac{sa + 3a + b}{s^2 + 3s + 2}. \quad (2.23)$$

La solución en el dominio del tiempo es:

$$x(t) = \mathcal{L}^{-1} \left[\frac{sa + 3a + b}{s^2 + 3s + 2} \right] = \mathcal{L}^{-1} \left[\frac{2a + b}{s + 1} - \frac{a + b}{s + 2} \right] \quad (2.24)$$

$$x(t) = (2a + b)e^{-t} - (a + b)e^{-2t}, \text{ para } t > 0 \quad (2.25)$$

- **Ejercicio 2:** Obtener la función $x(t)$ que cumple la ecuación diferencial con condiciones iniciales nulas:

$$\frac{d^2x(t)}{dt^2} + 2\frac{dx(t)}{dt} + 5x(t) = 3u(t) \quad (2.26)$$

Aplicando la transformada de Laplace a la ecuación diferencial,

$$s^2X(s) + 2sX(s) + 5X(s) = \frac{3}{s}, \quad (2.27)$$

la solución en el dominio de Laplace es:

$$X(s) = \frac{3}{s(s^2 + 2s + 5)}. \quad (2.28)$$

La solución en el dominio del tiempo es:

$$x(t) = \mathcal{L}^{-1} \left[\frac{3}{s(s^2 + 2s + 5)} \right] = \mathcal{L}^{-1} \left[\frac{A}{s} + \frac{Bs + C}{s^2 + 2s + 5} \right] \quad (2.29)$$

$$x(t) = \frac{3}{5} \mathcal{L}^{-1} \left[\frac{1}{s} \right] + \frac{3}{5} \mathcal{L}^{-1} \left[\frac{s + 1}{(s + 1)^2 + 2^2} + \frac{1}{2} \frac{2}{(s + 1)^2 + 2^2} \right] \quad (2.30)$$

$$x(t) = \frac{3}{5} \left(1 - e^{-t} \cos 2t - \frac{1}{2} e^{-t} \sin 2t \right), \text{ para } t > 0 \quad (2.31)$$

Capítulo 3

Representación de los sistemas

Los sistemas de control se pueden representar gráficamente de diversas formas, por ejemplo, mediante diagramas de flujo o diagramas de Bond-Graph. Sin embargo, en este libro sólo se emplearán los diagramas de bloques. Previamente, se definirá el concepto de función de transferencia y se aplicará a distintos tipos de sistemas físicos.

3.1. Generalidades

En la Fig. 3.1 se muestra la forma gráfica más elemental de representar un sistema. En dicha figura aparecen tres elementos: 1) la variable física de la entrada, que se representa con una flecha apuntando al sistema, 2) la variable física de la salida, que es la flecha dirigida del sistema al exterior, y 3) el propio sistema, representado aquí como una nube o “caja negra” del que se desconoce a priori su funcionamiento interno.

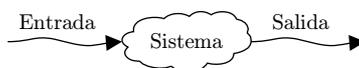


Figura 3.1: Diagrama de un sistema cualquiera

Esta forma tan elemental de representar un sistema permite al ingeniero establecer una primera descripción del mismo, sus posibles partes, así como las diferentes líneas de causalidad que se dan. Por ejemplo, en la Fig. 3.2 se muestra esquemáticamente el sistema “central hidroeléctrica”, que tiene como entrada el caudal de agua y como salida la tensión eléctrica. Este sistema se puede dividir en dos subsistemas: 1) la turbina que transforma el caudal de agua entrante en una velocidad de giro en su eje, y 2) la dinamo o alternador que convierte el giro mecánico en tensión eléctrica.

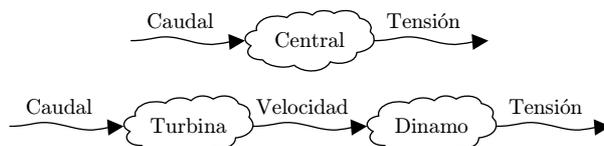


Figura 3.2: Diagramas equivalentes de una central hidroeléctrica

Evidentemente se podría haber dividido el sistema completo en muchas otras partes, conservando todas las representaciones igual validez. También se podrían haber encontrado otras variables intermedias entre la entrada y la salida, que unieran los distintos subsistemas, y que harían referencia a otras realidades físicas o incluso sin sentido físico, pero coherentes desde el punto de vista matemático.

El ingeniero evitará por todos los medios cambiar el orden natural de la causalidad. En el ejemplo anterior, no debe definir la tensión como entrada y el caudal de agua como salida. Y esto aunque se pueda encontrar la relación matemática inversa que deduce la segunda variable a partir de la primera. Por otro lado, existen infinitas formas de representar gráficamente un sistema cualquiera. Sin embargo, hay formas más adecuadas que otras (por ejemplo, porque muestren las variables físicas más importantes que intervienen en el sistema).

Encontrar el esquema o modelo más adecuado para un sistema físico es uno de los principales retos a los que se enfrenta el ingeniero, pero no es el cometido de esta asignatura. Además, cada ejemplo

mecánico, eléctrico, hidráulico, térmico, etc. podría requerir mucho tiempo de análisis y simplificación. Sin embargo, en el apartado 3.3 se estudian algunos sistemas físicos cuyas ecuaciones diferenciales se pueden obtener fácilmente aplicando las leyes fundamentales de cada disciplina.

Lo habitual es que las leyes matemáticas aparezcan directamente en el enunciado del ejercicio. En esta asignatura se supondrán correctas y se tomarán como punto de partida para la resolución del problema. Sin embargo, en el ejercicio de su profesión, el ingeniero no suele aceptar de forma acrítica cualquier ley matemática que se le sugiera. Muchas veces deberá obtenerlas a partir de ensayos o incluso deducirlas teóricamente.

Uno de los principales inconvenientes de trabajar con las leyes matemáticas que describen un sistema es que ingeniero puede perder fácilmente el sentido físico de los problemas de control. Conviene aquí recordar que detrás de una ley matemática se esconde un sistema físico real, del que se trabaja sólo con un modelo.

3.2. Función de transferencia de un sistema

En los esquemas propuestos en el apartado anterior el funcionamiento interno del sistema o subsistemas era desconocido. Una forma de ofrecer esa información es escribir la ecuación diferencial que relaciona la entrada con la salida. Sin embargo, lo habitual es trabajar en el dominio de Laplace (Fig. 3.3) definiendo la *función de transferencia* del sistema.

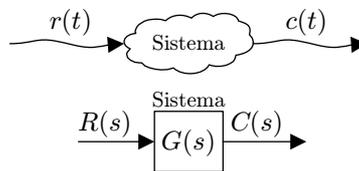


Figura 3.3: Diagramas generales de un sistema

La función de transferencia, en general $G(s)$, de un determinado proceso o sistema es la relación en el dominio de Laplace entre la función de salida $c(t)$ y su correspondiente entrada $r(t)$, con condiciones iniciales nulas para ambas funciones. La función de transferencia es un invariante del sistema, es decir, para cualquier entrada que se introduzca en el sistema, la salida que se obtiene siempre está relacionada con la entrada a través de la función de transferencia.

$$G(s) = \frac{\mathcal{L}[c(t)]}{\mathcal{L}[r(t)]} = \frac{C(s)}{R(s)} \quad (3.1)$$

Como la función de transferencia es un invariante del sistema, se puede obtener experimentalmente introduciendo una función temporal conocida y midiendo la salida. Aplicando la transformada de Laplace a las dos señales y calculando su cociente, se consigue la función de transferencia.

Si es posible introducir en el sistema una función impulso en la entrada, $\delta(t)$, la función de transferencia es directamente la transformada de Laplace de la función temporal de salida del sistema.

$$G(s) = \frac{\mathcal{L}[c(t)]}{\mathcal{L}[\delta(t)]} = \frac{\mathcal{L}[c(t)]}{1} = \mathcal{L}[c(t)] \quad (3.2)$$

También es posible obtener de forma teórica la función de transferencia de un sistema, mediante las ecuaciones diferenciales de su modelo matemático. Por ejemplo, según la segunda ley de Newton, un cuerpo con masa m experimenta en el vacío una aceleración $a(t)$ proporcional a la fuerza $f(t)$ que se le aplique, de acuerdo con la ecuación:

$$f(t) = ma(t) \quad (3.3)$$

La entrada o causa en el sistema es la fuerza, mientras que la salida o consecuencia es la aceleración del cuerpo. La función de transferencia se puede obtener muy fácilmente aplicando la transformada de Laplace a la ecuación (3.3) suponiendo condiciones iniciales nulas:

$$\mathcal{L}[f(t) = ma(t)] \quad (3.4)$$

$$\mathcal{L}[f(t)] = m\mathcal{L}[a(t)] \quad (3.5)$$

$$F(s) = mA(s) \quad (3.6)$$

$$G(s) = \frac{A(s)}{F(s)} = \frac{1}{m} \quad (3.7)$$

En este ejemplo, la función de transferencia es una constante. Sin embargo, si el ingeniero quisiera estudiar otros efectos de este sistema, como el desplazamiento del cuerpo, la ecuación diferencial que debería plantear es:

$$f(t) = m \frac{d^2 x(t)}{dt^2} \quad (3.8)$$

Y al aplicar la transformada de Laplace, suponiendo condiciones iniciales nulas, obtendría:

$$\mathcal{L} \left[f(t) = m \frac{d^2 x(t)}{dt^2} \right] \quad (3.9)$$

$$\mathcal{L} [f(t)] = m \mathcal{L} \left[\frac{d^2 x(t)}{dt^2} \right] \quad (3.10)$$

$$F(s) = ms^2 X(s) \quad (3.11)$$

$$G(s) = \frac{X(s)}{F(s)} = \frac{1}{ms^2} \quad (3.12)$$

Ahora la función de transferencia no es una constante, sino una función de la variable de Laplace. Hay que remarcar que el mismo sistema físico puede tener distintas funciones de transferencia dependiendo de las variables que se tomen como entradas y salidas.

A partir de este momento, una expresión como $H(s)$ puede corresponder tanto a la transformada de Laplace de una función temporal, $H(s) = \mathcal{L}[h(t)]$, como a la función de transferencia de un sistema. Normalmente por el contexto es posible deducir a qué se refiere en cada caso. Conviene resaltar que:

- La función de transferencia es una propiedad intrínseca del sistema. Conocida la función de transferencia de un sistema, se puede conocer el comportamiento del mismo ante cualquier entrada.
- La función de transferencia responde a la ecuación diferencial que gobierna un sistema pero no ofrece información acerca de su configuración interna. Dos sistemas físicos diferentes pueden poseer idénticas funciones de transferencia.

3.3. Modelos de sistemas físicos

Ya se ha dicho que no es sencillo obtener un modelo matemático que caracterice de forma adecuada el comportamiento de un sistema real. De hecho, ningún modelo matemático puede abarcar toda la realidad de un sistema. Sin embargo, para que un modelo sea útil no es necesario que sea excesivamente complicado. Basta con que represente los aspectos esenciales del mismo y que las predicciones sobre el comportamiento del sistema, basadas en dicho modelo, sean lo suficientemente precisas.

Los modelos de los sistemas suelen ser ecuaciones diferenciales. Normalmente se buscan ecuaciones diferenciales lineales de coeficientes constantes. De hecho, cuando aparecen ecuaciones diferenciales no lineales, lo habitual es linealizarlas en un punto de operación. En los siguientes apartados se estudian los modelos más elementales de cuatro tipos de sistemas: mecánicos, eléctricos, hidráulicos y térmicos. Se plantean las ecuaciones diferenciales elementales que gobiernan dichos sistemas y, teniendo en cuenta cuál es la entrada y cuál es la salida, se hallarán sus funciones de transferencia.

3.3.1. Sistemas mecánicos

Los sistemas mecánicos describen el movimiento en el espacio de cuerpos sometidos a fuerzas o pares. Aunque dichos cuerpos poseen dimensiones y propiedades físicas distribuidas en el espacio, la forma más sencilla de analizarlos obtener un modelo de parámetros concentrados. Los elementos básicos para construir un modelo con parámetros concentrados son la masa, el muelle y el amortiguador.

La ecuación diferencial que rige el comportamiento de una *masa* es la segunda ley de Newton:

$$f(t) = m \frac{d^2 x(t)}{dt^2} \quad (3.13)$$

Donde $f(t)$ es la suma de las fuerzas exteriores aplicadas a la masa y $x(t)$ es su desplazamiento. El parámetro m es la masa y su unidad fundamental en el Sistema Internacional es el kilogramo, kg (siempre con minúsculas). Si el cuerpo gira en lugar de desplazarse, la ecuación que gobierna su movimiento es:

$$\tau(t) = J \frac{d^2 \theta(t)}{dt^2} \quad (3.14)$$

Donde $\tau(t)$ es la suma de los pares exteriores aplicados al sistema y $\theta(t)$ su giro. El parámetro constante J es la inercia del sistema y su unidad es el $\text{kg}\cdot\text{m}^2$.

La fuerza $f(t)$ que restituye un *amortiguador* cuando se comprime es proporcional a la velocidad con que se aproximan sus extremos. La ecuación diferencial que rige su comportamiento es:

$$f(t) = c \frac{dx(t)}{dt} \quad (3.15)$$

El parámetro c es la constante del amortiguador o viscosidad, y su unidad es el Ns/m . Si una masa se desplaza dentro de un medio viscoso (al aire, el agua, etc.), además de su propia inercia debe vencer una fuerza viscosa proporcional a la velocidad con que se desplaza dicha masa. Este efecto se puede modelizar matemáticamente con un amortiguador cuyos extremos estuvieran anclados uno en el centro de gravedad de la masa y otro en un punto exterior fijo del medio. Evidentemente, este efecto no aparece en el vacío o en el espacio exterior, fuera de la atmósfera.

La fuerza $f(t)$ que restituye un *muelle* o resorte cuando se comprime es proporcional a la distancia $x(t)$ que se han acercado sus extremos desde su longitud natural. Es la llamada ley de Hooke (3.16), donde la constante k representa la rigidez del muelle y su unidad es el N/m .

$$f(t) = kx(t) \quad (3.16)$$

Si un sistema posee varios elementos combinados hay que acudir a los conocimientos de Mecánica para obtener la ecuación diferencial que gobierna el sistema. Por ejemplo, aislando las masas o los nodos, y poniendo las fuerzas que actúan en ellos. Después se aplica la segunda ley de Newton si son masas, o sumatorio de fuerzas igual a cero si son nodos.

Ejemplo mecánico 1

El sistema masa-muelle-amortiguador (Fig. 3.4) es un ejemplo típico de sistema mecánico. Se rige por la ecuación diferencial (3.17), donde la entrada es la fuerza $f(t)$ y la salida el desplazamiento $x(t)$. Aplicando la transformada de Laplace se puede obtener la función de transferencia del sistema.

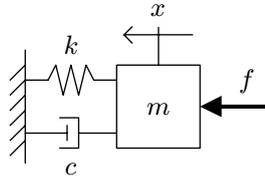


Figura 3.4: Sistema mecánico masa-muelle-amortiguador

$$f(t) = m \frac{d^2x(t)}{dt^2} + c \frac{dx(t)}{dt} + kx(t) \quad (3.17)$$

$$F(s) = ms^2X(s) + csX(s) + kX(s) \quad (3.18)$$

$$G(s) = \frac{X(s)}{F(s)} = \frac{1}{ms^2 + cs + k} \quad (3.19)$$

El diagrama de la Fig. 3.5 representa el sistema masa-muelle-amortiguador. Se puede comprobar cómo la función de transferencia (3.19) posee las unidades de m/N , es decir, precisamente las que relacionan la salida con la entrada. Asimismo, los sumandos del denominador son dimensionalmente coherentes.

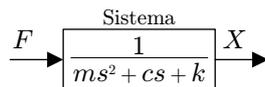


Figura 3.5: Diagrama del sistema masa-muelle-amortiguador

Ejemplo mecánico 2

La entrada de un sistema mecánico puede ser un desplazamiento en lugar de una fuerza, como ocurre en el caso de la Fig. 3.6. El desplazamiento $u(t)$ puede representar, por ejemplo, el desplazamiento de un vástago neumático.

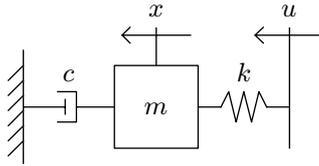


Figura 3.6: Sistema mecánico masa-muelle-amortiguador

La ecuación diferencial que gobierna este nuevo sistema es:

$$ku(t) = m \frac{d^2x(t)}{dt^2} + c \frac{dx(t)}{dt} + kx(t) \quad (3.20)$$

Mientras que su función de transferencia es:

$$G(s) = \frac{X(s)}{U(s)} = \frac{k}{ms^2 + cs + k} \quad (3.21)$$

Ejemplo mecánico 3

También es posible que el sistema pueda modelizarse despreciando la masa de los elementos móviles. Este es el caso del sistema de la Fig. 3.7, regido por la ecuación diferencial (3.22).

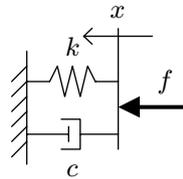


Figura 3.7: Sistema mecánico muelle-amortiguador

$$f(t) = c \frac{dx(t)}{dt} + kx(t) \quad (3.22)$$

$$G(s) = \frac{X(s)}{F(s)} = \frac{1}{cs + k} \quad (3.23)$$

Ejemplo mecánico 4

En el sistema de la Fig. 3.8 ante una única entrada $u(t)$ existen dos variables temporales de salida, los desplazamientos de las masas $x_1(t)$ y $x_2(t)$. Este sistema puede servir para modelizar el comportamiento del sistema de amortiguación de un vehículo. La masa m_2 representa la parte amortiguada del vehículo, mientras que m_1 es el conjunto de la rueda y el eje. El desplazamiento de entrada $u(t)$ es el perfil de la carretera que actúa sobre la rueda a través de la rigidez del neumático k_1 .

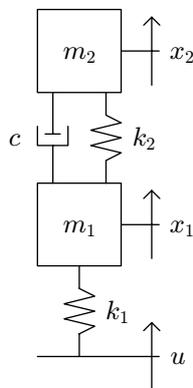


Figura 3.8: Modelo de un sistema de amortiguación

Para obtener las ecuaciones del sistema, habría que tener en cuenta las fuerzas debidas a la gravedad (el peso). Sin embargo, es posible llegar más rápido a la solución definiendo la referencia de los desplazamientos $x_1(t)$ y $x_2(t)$ en el punto de equilibrio estático (no con la longitud natural de los resortes).

De esta forma, el conjunto se rige por el sistema de ecuaciones (3.24), donde —por simplicidad— no se especifica la variable temporal de los desplazamientos.

$$\left. \begin{aligned} k_1(u - x_1) &= m_1 \frac{d^2 x_1}{dt^2} + k_2(x_1 - x_2) + c \left(\frac{dx_1}{dt} - \frac{dx_2}{dt} \right) \\ k_2(x_1 - x_2) + c \left(\frac{dx_1}{dt} - \frac{dx_2}{dt} \right) &= m_2 \frac{d^2 x_2}{dt^2} \end{aligned} \right\} \quad (3.24)$$

Si lo único que interesa del sistema es el desplazamiento de la masa amortiguada, sin importar cómo se mueva la rueda, habría eliminar del sistema de dos ecuaciones y dos incógnitas (3.24) la variable $x_1(t)$. El objetivo sería obtener una única ecuación que relacione la entrada $u(t)$ con la variable $x_2(t)$. Esto es difícil de hacer con las ecuaciones diferenciales en el dominio temporal, pero muy sencillo gracias a la transformada de Laplace:

$$\left. \begin{aligned} k_1(U - X_1) &= m_1 s^2 X_1 + k_2(X_1 - X_2) + cs(X_1 - X_2) \\ k_2(X_1 - X_2) + cs(X_1 - X_2) &= m_2 s^2 X_2 \end{aligned} \right\} \quad (3.25)$$

Queda un sistema de ecuaciones algebraico donde se puede eliminar la variable X_1 . De esta forma la función de transferencia del sistema es:

$$\frac{X_2}{U} = \frac{k_1(k_2 + cs)}{(m_1 s^2 + cs + k_1 + k_2)(m_2 s^2 + cs + k_2) - (k_2 + cs)^2} \quad (3.26)$$

Evidentemente, también se podría haber tomado X_1 como salida del sistema y eliminar variable X_2 . En este caso, la función de transferencia que se obtendría es:

$$\frac{X_1}{U} = \frac{k_1(m_2 s^2 + cs + k_2)}{(m_1 s^2 + cs + k_1 + k_2)(m_2 s^2 + cs + k_2) - (k_2 + cs)^2} \quad (3.27)$$

El sistema mecánico completo, se podría representar gráficamente como un sistema con una entrada y dos salidas como se muestra en la Fig. 3.9. Es interesante ver que, normalmente, las funciones de transferencia son fracciones de polinomios en s .

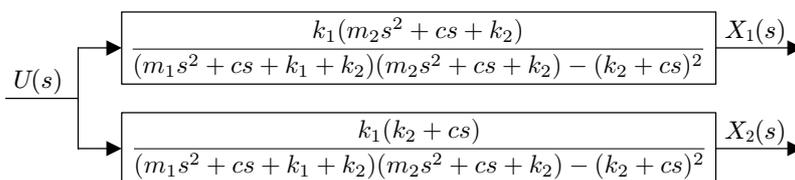


Figura 3.9: Diagrama del sistema de amortiguación

Las raíces del denominador se llaman *polos*, y las raíces del numerador se llaman *ceros*. Habitualmente el número de polos es mayor o igual que el número de ceros. Y el *orden* del sistema lo determina el número de polos. En el ejemplo del sistema de amortiguación, el sistema es de orden 4 (o de 4º orden).

Para un sistema físico que posea una entrada y n salidas, se podrán definir n funciones de transferencia. Y todas las funciones de transferencia tendrán —por norma general— el mismo denominador. Por tanto, los polos del sistema constituyen una caracteriza esencial e invariante del propio sistema. De hecho, el Capítulo 4 analiza la respuesta temporal de los sistemas atendiendo a la posición y número de los polos del sistema.

3.3.2. Sistemas eléctricos

Los sistemas eléctricos, como los mecánicos, también se suelen describir por medio de parámetros concentrados donde los tres elementos fundamentales son las resistencias, los condensadores y las bobinas. La tensión que aparece sobre los extremos de una *resistencia* es proporcional a la intensidad de corriente que circula a través de ella. La constante proporcional se llama igualmente resistencia y su unidad en el Sistema Internacional es el ohmio, Ω .

$$v(t) = Ri(t) \quad (3.28)$$

La tensión que aparece sobre los extremos de una *bobina* es proporcional a la derivada de la intensidad que circula a través de ella respecto del tiempo. La constante proporcional se llama inductancia y su unidad es el henrio, H.

$$v(t) = L \frac{di(t)}{dt} \quad (3.29)$$

La tensión que aparece sobre los extremos de un *condensador* es proporcional a la integral de la intensidad que circula a través de ella a lo largo del tiempo. Desde otro punto de vista, también se puede decir que la intensidad que circula a través de un condensador es proporcional a la variación de la tensión entre sus bornes. Esta última constante proporcional es la que se llama capacidad y su unidad es el faradio, F.

$$i(t) = C \frac{dv(t)}{dt} \quad (3.30)$$

En un circuito en el que existan resistencias, bobinas y condensadores, las ecuaciones diferenciales que lo gobiernan se obtienen aplicando las leyes de Kirchoff en las mallas o en los nudos. A continuación se muestran algunos casos en los que se da una combinación de estos tres elementos y sus respectivas ecuaciones diferenciales.

Ejemplo eléctrico 1

En el sistema de la Fig. 3.10, la entrada en el circuito es la tensión $v_i(t)$ y la salida es la tensión $v_o(t)$ suponiendo que la corriente de salida es nula, o lo que es lo mismo, el circuito se conecta a un dispositivo de alta impedancia de entrada.

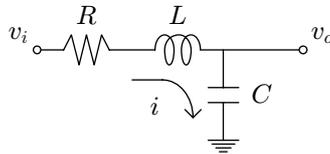


Figura 3.10: Sistema eléctrico resistencia-bobina-condensador

$$\left. \begin{aligned} v_i(t) &= Ri(t) + L \frac{di(t)}{dt} + \frac{1}{C} \int_0^t i(\tau) d\tau \\ v_o(t) &= \frac{1}{C} \int_0^t i(\tau) d\tau \end{aligned} \right\} \quad (3.31)$$

En el sistema de ecuaciones diferenciales (3.31) interviene una variable intermedia: la intensidad $i(t)$. Como ocurría anteriormente en los sistemas mecánicos, en el dominio de Laplace se pueden eliminar aquellas variables que se consideren innecesarias, y obtener una única expresión de la salida del sistema en función de la entrada. Así, aplicando la transformada de Laplace al sistema de ecuaciones que gobierna el sistema, suponiendo condiciones iniciales nulas:

$$\left. \begin{aligned} V_i(s) &= RI(s) + LsI(s) + \frac{I(s)}{sC} \\ V_o(s) &= \frac{I(s)}{sC} \end{aligned} \right\} V_o(s) = \frac{1}{1 + RCs + LCs^2} V_i(s) \quad (3.32)$$

Se ha conseguido expresar la tensión de salida del circuito en función de la tensión de entrada, es decir, la función de transferencia, independientemente de la otra variable, que es la intensidad que circula por la malla.

En este caso particular, como la tensión de entrada y la tensión de salida tienen las mismas unidades, la función de transferencia es adimensional y cada uno de los sumandos del denominador también lo será: ohmio por faradio entre segundo es adimensional y henrio por faradio entre segundo al cuadrado es adimensional. Comprobar las unidades puede ayudar a detectar errores en la resolución de ejercicios.

Si la tensión de entrada en el sistema resistencia-bobina-condensador es un escalón de valor 3 voltios, es posible encontrar el valor que alcanza la tensión en la capacidad cuando el tiempo tiende a infinito a través del teorema del valor final:

$$\lim_{t \rightarrow \infty} v_o(t) = \lim_{s \rightarrow 0} sV_o(s) = \lim_{s \rightarrow 0} s \frac{1}{1 + RCs + LCs^2} \frac{3}{s} = 3 \text{ voltios} \quad (3.33)$$

Con este ejemplo, queda patente cómo es posible conocer algunas características de la respuesta temporal del sistema sin haber calculado la expresión general de la tensión $v_o(t)$ en función del tiempo a través de la transformada inversa de Laplace. Con los teoremas del valor inicial y final es posible conocer el valor en régimen permanente, el valor inicial de la función y las sucesivas derivadas de la función en el origen.

Ejemplo eléctrico 2

En el sistema de la Fig. 3.11 existen dos mallas, por tanto se obtienen dos variables intermedias entre las tensiones de salida y de entrada: las intensidades $i_1(t)$ e $i_2(t)$.

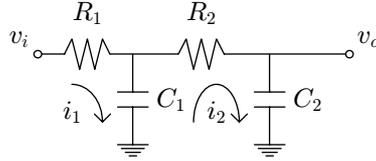


Figura 3.11: Sistema eléctrico con dos mallas

$$\left. \begin{aligned} v_i &= R_1 i_1 + \frac{1}{C_1} \int_0^t (i_1 - i_2) d\tau \\ \frac{1}{C_1} \int_0^t (i_1 - i_2) d\tau &= R_2 i_2 + \frac{1}{C_2} \int_0^t i_2 d\tau \\ v_o &= \frac{1}{C_2} \int_0^t i_2 d\tau \end{aligned} \right\} \quad (3.34)$$

Ejemplo eléctrico 3

En el sistema de la Fig. 3.12 se muestra un ejemplo donde la entrada es una corriente en lugar de una tensión. La entrada es la corriente $i(t)$ de la fuente, la salida es la corriente $i_2(t)$ en la resistencia de carga R_L y existe una variable intermedia que es la corriente $i_1(t)$ de la malla intermedia.

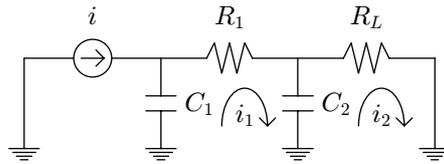


Figura 3.12: Sistema eléctrico con fuente de corriente

$$\left. \begin{aligned} \frac{1}{C_1} \int_0^t (i - i_1) d\tau &= R_1 i_1 + \frac{1}{C_2} \int_0^t (i_1 - i_2) d\tau \\ \frac{1}{C_2} \int_0^t (i_1 - i_2) d\tau &= R_L i_2 \end{aligned} \right\} \quad (3.35)$$

3.3.3. Sistemas electromecánicos

Los sistemas electromecánicos o mecatrónicos, combinan elementos mecánicos y eléctricos. Un ejemplo es el motor de corriente continua que hace girar un objeto con inercia y viscosidad, Fig. 3.13. La entrada es la tensión $v(t)$ y la salida es el giro $\theta(t)$.

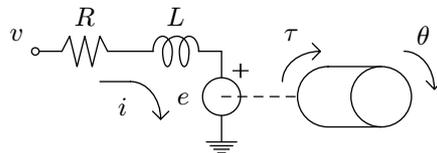


Figura 3.13: Modelo de un motor de corriente continua arrastrando un objeto

$$\left. \begin{aligned} v(t) &= Ri(t) + L \frac{di(t)}{dt} + e(t) \\ e(t) &= K \frac{d\theta(t)}{dt} \\ \tau(t) &= Ki(t) \\ \tau(t) &= J \frac{d^2\theta(t)}{dt^2} + B \frac{d\theta(t)}{dt} \end{aligned} \right\} \quad (3.36)$$

La primera ecuación del sistema (3.36) responde a la única malla del circuito. La tensión $e(t)$ que aparece en el motor es proporcional a la velocidad de giro del mismo. El par $\tau(t)$ que ejerce el motor es proporcional a la intensidad que circula por él. Las constantes de velocidad y de par son la misma K , donde es posible demostrar que tienen las mismas unidades. La última ecuación del sistema es la del modelo mecánico de inercia J y viscosidad B .

3.3.4. Sistemas hidráulicos

Los sistemas hidráulicos pueden incluir muy diferentes elementos (depósitos, válvulas, etc.). En este apartado se muestra un ejemplo de ecuación diferencial que gobierna la altura h de fluido contenido en un depósito.

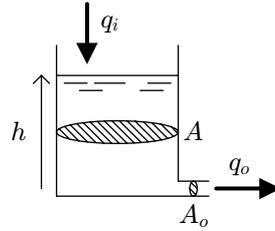


Figura 3.14: Depósito con conducto de desagüe

Las ecuaciones que se pueden plantar en el depósito de la Fig. 3.14 son la conservación de la masa y el caudal de salida,

$$\left. \begin{aligned} q_i - q_o &= A \frac{dh}{dt} \\ q_o &= A_o v_o = A_o f \sqrt{2gh} = K \sqrt{h} \end{aligned} \right\} \quad (3.37)$$

donde q_i y q_o son los caudales de entrada y salida, mientras que A y A_o son las superficies de la sección del depósito y del conducto de salida. Eliminando la variable del caudal de salida q_o resulta:

$$A \frac{dh}{dt} + K \sqrt{h} = q_i \quad (3.38)$$

Lo primero que conviene resaltar es que esta ecuación diferencial no es lineal. En lugar de aparecer una función temporal y sucesivas derivadas temporales, aparece la raíz cuadrada de la función. Un modo de estudiar este tipo de sistemas consiste en linealizar su ecuación diferencial en algún punto de operación. Lo más sencillo es estudiar este comportamiento en el punto de equilibrio del sistema: para una altura H de fluido existe un caudal de entrada Q_i tal que el caudal de salida $Q_o = K \sqrt{H}$ es igual al de entrada. El punto (Q_i, H) es el punto de equilibrio en el que se linealizará este sistema.

Se aplicará el desarrollo en serie de Taylor de primer orden a la función,

$$f(q_i, h) = \frac{dh}{dt}, \quad (3.39)$$

por tanto,

$$f(q_i, h) \approx f(Q_i, H) + \left[\frac{\partial f}{\partial q_i} \right]_{(Q_i, H)} (q_i - Q_i) + \left[\frac{\partial f}{\partial h} \right]_{(Q_i, H)} (h - H) \quad (3.40)$$

$$\frac{dh}{dt} \approx \frac{1}{A} (Q_i - K \sqrt{H}) + \frac{1}{A} (q_i - Q_i) - \frac{K}{2A \sqrt{H}} (h - H) \quad (3.41)$$

$$\frac{dh}{dt} \approx \frac{1}{A} (q_i - Q_i) - \frac{1}{\tau} (h - H) \quad (3.42)$$

Si se define el siguiente cambio de variables:

$$\left. \begin{aligned} \bar{h} &= h - H \\ \bar{q}_i &= q_i - Q_i \end{aligned} \right\} \quad (3.43)$$

la ecuación diferencial lineal en torno al punto de equilibrio es:

$$\frac{d\bar{h}}{dt} + \frac{1}{\tau} \bar{h} = \frac{1}{A} \bar{q}_i \quad (3.44)$$

Para hallar la función de transferencia del sistema en torno al punto de equilibrio, habrá que suponer condiciones iniciales nulas en las variables relativas: $\bar{h}(0^+) = 0$ y $\bar{q}_i(0^+) = 0$, lo que equivale a decir que las condiciones iniciales de las variables absolutas no son nulas: $h(0^+) = H$ y $q_i(0^+) = Q_i$.

Esta linealización se puede realizar en otros puntos distintos al de equilibrio. Para depósitos como el del ejemplo, una formulación aproximada bastante extendida es considerar el caudal de salida proporcional a la altura del depósito,

$$q_o = \frac{h}{R}, \quad (3.45)$$

donde R equivale a una *resistencia al flujo* de salida de caudal por la boquilla. Esta aproximación es un símil eléctrico del flujo del fluido y conduce a ecuaciones diferenciales lineales. Por ejemplo, la conservación de la masa en el depósito conduce a la ecuación,

$$q_i - \frac{h}{R} = A \frac{dh}{dt}, \quad (3.46)$$

que es comparable al resultado que daba la linealización anterior,

$$\frac{dh}{dt} + \frac{1}{AR}h = \frac{1}{A}q_i, \quad (3.47)$$

pero empleando esta vez la altura h y el caudal de entrada q_i absolutos en lugar de sus diferencias respecto al punto de linealización. En definitiva, la suposición (3.45) equivale a linealizar el sistema en todas las alturas del depósito (no sólo en un punto determinado). Siguiendo el símil eléctrico, el caudal se comporta como la intensidad de corriente, la altura del depósito como el potencial y el área de la sección del depósito como una capacidad. Por esto último, en algunos manuales al área A del depósito se le llama *capacitancia del tanque*.

3.3.5. Sistemas térmicos

Los sistemas térmicos describen el calentamiento de los objetos con el flujo de calor. Su descripción matemática es similar a los sistemas eléctricos. En este caso la *resistencia térmica* R es la oposición al flujo de calor $\dot{q}(t)$ entre dos cuerpos que posean temperaturas distintas. En el Sistema Internacional, las unidades de esta resistencia térmica es K/W.

$$\dot{q}(t) = \frac{T_1(t) - T_2(t)}{R} \quad (3.48)$$

La *capacidad calorífica* C se define como el calor almacenado o desprendido por un cuerpo cuando cambia de temperatura. Esta capacidad se suele dar en forma de calor específico, es decir, por unidad de masa. En el Sistema Internacional, las unidades del calor específico c_e es J/kgK.

$$\dot{q}(t) = C \frac{dT(t)}{dt} \quad (3.49)$$

$$\dot{q}(t) = mc_e \frac{dT(t)}{dt} \quad (3.50)$$

Con estas definiciones se puede modelizar el comportamiento térmico de muchos sistemas. Por ejemplo en la Fig. 3.15 se muestra una habitación con un radiador que introduce un flujo de calor $\dot{q}_r(t)$ en presencia de una ventana de resistencia térmica R por la que se pierde un flujo de calor de $\dot{q}_s(t)$. La temperatura exterior $T_e(t)$, aunque se suele suponer constante, en realidad es también una variable temporal.

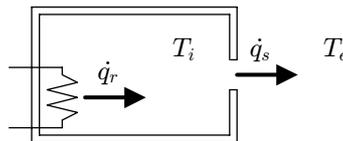


Figura 3.15: Habitación con un radiador y una ventana

Las ecuaciones del sistema son:

$$\left. \begin{aligned} \dot{q}_r(t) - \dot{q}_s(t) &= mc_e \frac{dT_i(t)}{dt} \\ \dot{q}_s(t) &= \frac{T_i(t) - T_e(t)}{R} \end{aligned} \right\} \quad (3.51)$$

Eliminado la variable \dot{q}_s queda una ecuación diferencial lineal de primer orden:

$$\dot{q}_r(t) = mc_e \frac{dT_i(t)}{dt} + \frac{T_i(t) - T_e(t)}{R} \quad (3.52)$$

Para obtener la función de transferencia de un sistema térmico hay que tener especial precaución. El concepto de función de transferencia requiere suponer condiciones iniciales nulas, cosa que sólo ocurre por excepción en los sistemas térmicos. Es difícil que la temperatura de la habitación sea 0 K (-273°C) en el momento de encender el radiador (origen de tiempos). Si en el origen de tiempos la habitación tenía una temperatura igual a T_i^0 , la transformada de Laplace de la ecuación diferencial del sistema es:

$$\dot{Q}_r(s) = mc_e[sT_i(s) - T_i^0] + \frac{1}{R}T_i(s) - \frac{1}{R}T_e(s) \quad (3.53)$$

La salida del sistema depende de dos entradas (por medio de funciones de transferencia) y las condiciones iniciales (a través de un sumando que no es una función de transferencia):

$$T_i(s) = \frac{R}{mc_eRs + 1} \dot{Q}_r(s) + \frac{1}{mc_eRs + 1} T_e(s) + \frac{mc_eRT_i^0}{mc_eRs + 1} \quad (3.54)$$

El flujo de calor que introduce el radiador $\dot{Q}_r(s)$ es una entrada controlable del sistema. La temperatura exterior $T_e(s)$ es un agente que influye en la salida pero no es controlable por el sistema de calefacción, por tanto se trata de una perturbación. Por otro lado, se puede apreciar que el sistema es de primer orden. Este sistema térmico se puede definir por medio de funciones de transferencia si se emplean temperaturas relativas respecto a la condición inicial de la temperatura interior: $T_i(t) = T_i^0 + \bar{T}_i(t)$ y $T_e(t) = T_e^0 + \bar{T}_e(t)$.

3.4. Diagrama de bloques de un sistema

Los diagramas de bloques aparecen cuando el sistema se divide en varios subsistemas. En este caso, en lugar de hallar de función de transferencia del sistema completo se deben encontrar las funciones de transferencia de cada uno de los subsistemas. En este diagrama, cada subsistema es un *bloque* del sistema completo. En las uniones entre bloques pueden aparecer puntos de bifurcación y de suma (Fig. 3.16). Los puntos de bifurcación se emplean en las señales que atacan varias funciones de transferencia. Los puntos de suma se representan con círculos a los que llegan las señales que se combinan para dar el resultado. En la línea de llegada al punto de suma se debe especificar el signo.

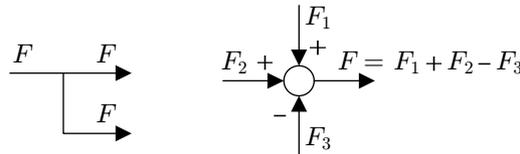


Figura 3.16: Punto de bifurcación (izquierda) y punto de suma (derecha)

El diagrama de bloques de un sistema se puede construir a partir de las ecuaciones diferenciales que lo gobiernan. Primero se toman las transformadas de Laplace de las ecuaciones, suponiendo condiciones iniciales nulas. Luego cada ecuación en el dominio de Laplace se representa en forma de bloque. Finalmente se unen los bloques para formar un único diagrama. Este procedimiento se sigue en los ejemplos del siguiente apartado.

3.4.1. Reglas para la simplificación de diagramas de bloques

Simplificar un diagrama de bloques significa encontrar la función de transferencia equivalente a todo el diagrama. Esto se puede hacer analíticamente, planteando las ecuaciones del diagrama, o gráficamente, aprendiendo unas reglas de simplificación de bloques.

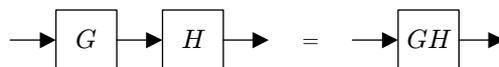


Figura 3.17: Multiplicación de bloques

En las Figuras 3.17-3.20 se muestran algunas las simplificaciones gráficas más útiles. La característica fundamental es que el sistema simplificado es equivalente al anterior.

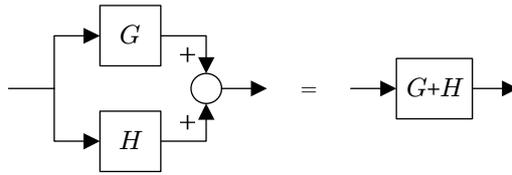


Figura 3.18: Suma de bloques

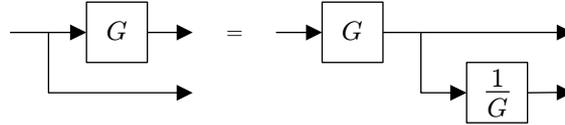


Figura 3.19: Translación de un punto de bifurcación

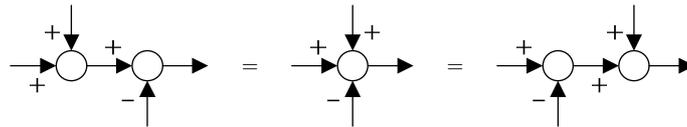


Figura 3.20: Cambio de orden de los sumandos

3.4.2. Ejemplo de circuito con dos mallas

El sistema (3.55) contiene las ecuaciones diferenciales que gobiernan el circuito representado en la Fig. 3.11, después de aplicar la transformada de Laplace. El diagrama de la Fig. 3.21 corresponde a dichas ecuaciones, donde se ha señalado con puntos el conjunto de bloques que corresponde a cada una de ellas.

$$\left. \begin{aligned} V_i &= R_1 I_1 + \frac{I_1 - I_2}{sC_1} \\ \frac{I_1 - I_2}{sC_1} &= R_2 I_2 + \frac{I_2}{sC_2} \\ V_o &= \frac{I_2}{sC_2} \end{aligned} \right\} \quad (3.55)$$

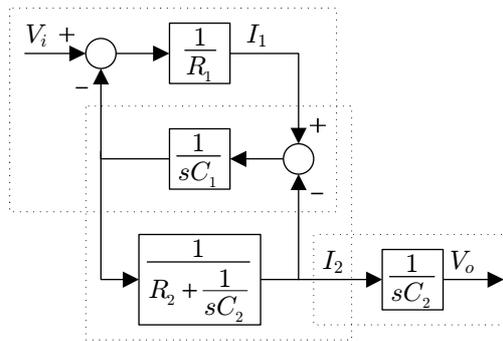


Figura 3.21: Diagrama del circuito con dos mallas

Usando mayor número de variables intermedias, el sistema de ecuaciones aumentaría en número, pero el diagrama de bloques puede resultar más sencillo de representar. En el ejemplo, si se incluye la tensión en un nudo intermedio v_1 y la corriente i_c diferencia de las corrientes en la mallas:

$$\left. \begin{aligned} v_i - v_1 &= R_1 i_1 \\ v_1 &= \frac{1}{C_1} \int_0^t i_c d\tau \\ v_1 - v_o &= R_2 i_2 \\ v_o &= \frac{1}{C_2} \int_0^t i_2 d\tau \\ i_c &= i_1 - i_2 \end{aligned} \right\} \xrightarrow{\mathcal{L}} \left. \begin{aligned} V_i - V_1 &= R_1 I_1 \\ V_1 &= \frac{I_c}{sC_1} \\ V_1 - V_o &= R_2 I_2 \\ V_o &= \frac{I_2}{sC_2} \\ I_c &= I_1 - I_2 \end{aligned} \right\} \quad (3.56)$$

El diagrama de bloques de la Fig. 3.22 a) corresponde al sistema de ecuaciones (3.56) y es equivalente al diagrama anterior. Si se eliminan las variables intermedias de forma analítica en el sistema de ecuaciones, para expresar la tensión de salida en función de la tensión de entrada, se obtiene la función de transferencia equivalente del sistema. Con esta función de transferencia se puede representar el sistema con un único bloque, Fig. 3.22 b). La función de transferencia equivalente del sistema también se puede obtener simplificando de forma gráfica cualquiera de los diagramas de bloques presentados previamente.

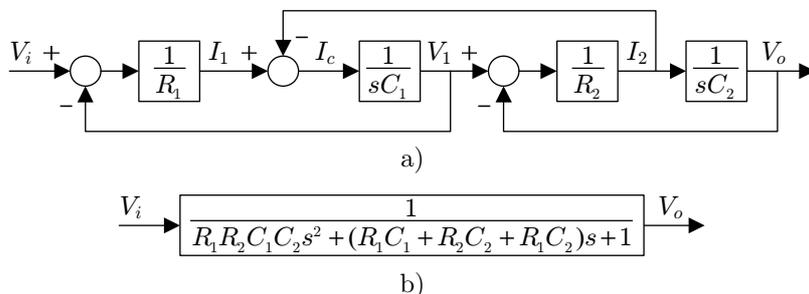


Figura 3.22: Diagramas equivalentes del circuito con dos mallas

3.4.3. Ejemplo de motor de corriente continua

El sistema (3.57) son las ecuaciones diferenciales que gobiernan un motor de corriente continua que arrastra una inercia (3.36) una vez aplicada la transformada de Laplace.

$$\left. \begin{aligned} V &= (R + Ls)I + E \\ E &= K\Omega \\ T &= KI \\ T &= (Js + B)\Omega \end{aligned} \right\} \quad (3.57)$$

El diagrama de bloques que corresponde a estas ecuaciones y su simplificación aparecen en la Fig. 3.23.

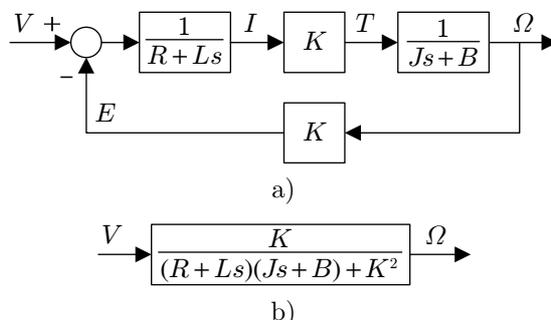


Figura 3.23: Diagrama del motor de corriente continua

Si se eliminan las variables intermedias del sistema de ecuaciones de Laplace se obtiene la función de transferencia equivalente. Aunque el sistema tiene forma de lazo cerrado con realimentación no unitaria, hay que hacer notar que no es propiamente un sistema controlado. La velocidad Ω está impuesta por la tensión V y la magnitud de la inercia J . En este ejemplo se observa cómo la función de transferencia equivalente posee las unidades que relacionan la magnitud de salida con la de entrada.

En el caso de que el giro de la inercia se vea frenado por un muelle torsor de rigidez K_t , las ecuaciones que hay que considerar son las siguientes:

$$\left. \begin{aligned} v &= Ri + L \frac{di}{dt} + e \\ e &= K \frac{d\theta}{dt} \\ \tau &= Ki \\ \tau &= J \frac{d^2\theta}{dt^2} + B \frac{d\theta}{dt} + K_t \theta \end{aligned} \right\} \xrightarrow{\mathcal{L}} \left. \begin{aligned} V &= (R + Ls)I + E \\ E &= Ks\Theta \\ T &= KI \\ T &= (Js^2 + Bs + K_t)\Theta \end{aligned} \right\} \quad (3.58)$$

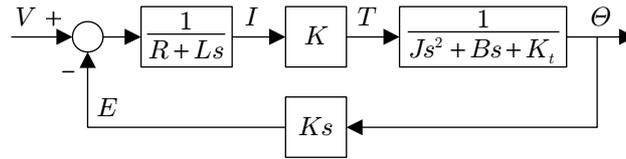


Figura 3.24: Diagrama del motor de corriente continua con muelle torsor

Una posible representación en diagrama de bloques se presenta en la Fig. 3.24, sin embargo no es una buena elección incluir bloques derivadores, es decir, aquellos cuya salida es proporcional a la derivada de la entrada. Este tipo de bloques tienen el inconveniente de que amplifican enormemente el ruido de alta frecuencia que reciban en la entrada. También se comportan mal a la hora de evaluar numéricamente las respuestas temporales del sistema, por ejemplo utilizando Simulink®.

Para evitar el bloque derivador, se propone como alternativa el diagrama de la Fig. 3.25 a). En ambos casos la función de transferencia equivalente de todo el sistema es la misma, Fig. 3.25 b).

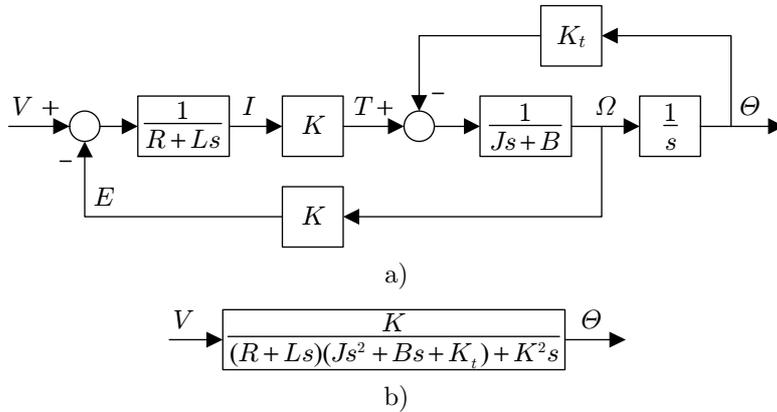


Figura 3.25: Diagramas equivalentes del motor de corriente continua con muelle torsor

3.5. Sistema de realimentación negativa no unitaria

Los sistemas de realimentación negativa son los más extendidos para el control de sistemas, por eso su estructura se estudia de forma pormenorizada. En la Fig. 3.26 se representa el caso más simple de sistema de realimentación negativa no unitaria. Hay que en cuenta que las funciones de transferencia $G(s)$ y $H(s)$ pueden ser el resultado del producto de varias funciones de transferencia.

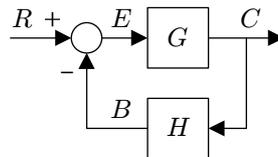


Figura 3.26: Sistema de realimentación negativa no unitaria

En (3.59) se muestra solución del sistema de ecuaciones de Laplace de la realimentación negativa no unitaria, es decir, la salida en función de la entrada.

$$\left. \begin{aligned} C &= GE \\ E &= R - B \\ B &= HC \end{aligned} \right\} C = \frac{G}{1 + GH} R \quad (3.59)$$

Habitualmente se emplea el convenio de usar la letra $C(s)$ para nombrar a la transformada de Laplace de la función de salida y $R(s)$ para la entrada. A la señal $E(s)$ se le llama error y a $B(s)$ señal de realimentación. Las funciones de transferencia que intervienen en el sistema son:

- **Función de transferencia directa:** es la que relaciona la señal de error y la salida.

$$G_d = \frac{C}{E} = G \quad (3.60)$$

- **Función de transferencia en lazo abierto:** es la que relaciona la señal de error y la realimentación. Es el producto de todas las funciones de transferencia que se encuentran dentro del lazo de control.

$$G_{la} = \frac{B}{E} = GH \quad (3.61)$$

- **Función de transferencia en lazo cerrado:** es la que relaciona la señal de entrada y la salida. Es igual a la función de transferencia directa entre uno más la función de transferencia en lazo abierto.

$$G_{lc} = \frac{C}{R} = \frac{G}{1 + GH} = \frac{G_d}{1 + G_{la}} \quad (3.62)$$

Con la función de transferencia en lazo cerrado se puede representar el sistema de la Fig. 3.27 con un único bloque:

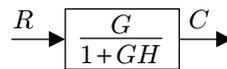


Figura 3.27: Sistema equivalente en lazo cerrado

Para el diseño de controladores son especialmente importantes las expresiones de las funciones de transferencia en lazo abierto y cerrado. El sistema controlado responde a la función de transferencia en lazo cerrado, sin embargo, muchas de las características del sistema controlado se deducen a partir de la función de transferencia en lazo abierto, como se irá mostrando en los sucesivos apartados y capítulos.

En la Fig. 3.28 se representa el caso de sistema realimentación negativa no unitaria en presencia de perturbaciones. Para la señal de perturbación se suele emplear la letra N y su signo puede ser positivo o negativo.

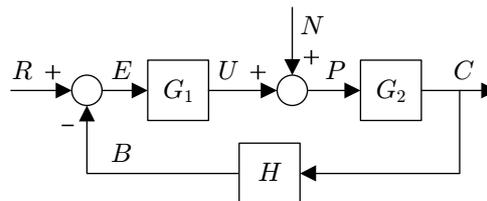


Figura 3.28: Sistema de realimentación con perturbaciones

En (3.63) se muestra la solución del nuevo sistema de ecuaciones de Laplace. En este caso es la salida en función de las dos entradas al sistema: la referencia R y la perturbación N .

$$\left. \begin{array}{l} C = G_2 P \\ P = U + N \\ U = G_1 E \\ E = R - B \\ B = HC \end{array} \right\} C = \frac{G_1 G_2}{1 + G_1 G_2 H} R + \frac{G_2}{1 + G_1 G_2 H} N \quad (3.63)$$

La solución (3.63) se puede obtener por superposición, es decir, sumando las salidas que se producen con entrada R y N nula más la salida con entrada N y R nula. La entrada propiamente dicha en el sistema es la señal R y se llama *referencia* porque se desea que el sistema controlado la siga fielmente. Observando la ecuación (3.63), es posible deducir que el seguimiento se consigue de forma exacta, $C = R$, cuando la función de transferencia que multiplica a R se asemeja a la unidad, y la que multiplica a N se asemeja a cero.

Una forma de conseguir las dos cosas es hacer G_1 todo lo grande que sea posible y H igual a la unidad. Por esta razón es habitual estudiar los sistemas de control de realimentación negativa unitaria, donde el controlador se coloca inmediatamente después del cálculo del error, es decir, el controlador actúa en función de la señal del error. A la actuación del controlador se añaden las perturbaciones que puedan existir sobre la planta.

Los sistema servo busca sobre todo el seguimiento de la señal, es decir, que la función de transferencia de la R sea lo más parecida a la unidad, mientras que un sistema regulador busca sobre todo el rechazo a las perturbaciones, es decir, anular la función de transferencia que multiplica a la perturbación.

También hay que notar que el denominador de las dos funciones de transferencia es idéntico. Este denominador es una característica esencial del sistema, como se verá en los siguientes capítulos.

3.6. Sistema de realimentación negativa unitaria

En la Fig. 3.29 se representa el caso de sistema realimentación negativa unitaria con perturbaciones. Que la realimentación sea unitaria implica que el sensor que mide la salida es ideal, es decir, no modifica en absoluto dicha señal. La función de transferencia G_1 incluye el controlador y la etapa final de amplificación, mientras que G_2 es la planta que se desea controlar.

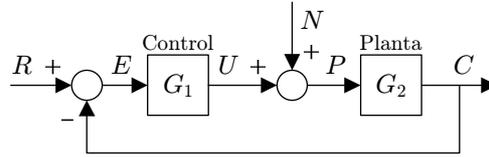


Figura 3.29: Sistema de realimentación negativa unitaria

Hay que resalta que, en el caso de realimentación negativa unitaria las funciones de transferencia directa y de lazo abierto coinciden y es el producto de G_1 y G_2 . Si no se especifica otra cosa, cuando se desee controlar un sistema, se entenderá que se le introduce en un lazo de control similar al de la Fig. 3.29. Las perturbaciones, también si no se especifica otra cosa, se supondrán nulas.

3.7. Ejercicios propuestos

- **Ejercicio 1:** Sea un sistema mecánico que compuesto de una muelle y un amortiguador en serie (Fig. 3.30) cuya entrada es el desplazamiento u y cuya salida es el desplazamiento x . Hallar la ecuación diferencial que gobierna el sistema, así como la solución analítica de la respuesta ante una entrada escalón unidad.

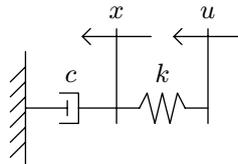


Figura 3.30: Sistema mecánico muelle-amortiguador

La ecuación diferencial del sistema es:

$$ku(t) = kx(t) + c \frac{dx(t)}{dt} \tag{3.64}$$

La solución general en el dominio de Laplace es:

$$X(s) = \frac{k}{k + cs} U(s) \tag{3.65}$$

Con una entrada escalón unidad:

$$X(s) = \frac{k}{(k + cs)s} \tag{3.66}$$

Acudiendo a las tablas de la transformada de Laplace, la solución en el dominio del tiempo es:

$$x(t) = 1 - e^{-\frac{k}{c}t}, \text{ para } t > 0 \tag{3.67}$$

Y se puede demostrar que, tanto sustituyendo valores en la solución temporal, como aplicando los teoremas de valor inicial y final en la solución del dominio de Laplace, se cumple que:

$$x(\infty) = 1 \tag{3.68}$$

$$\frac{dx(0^+)}{dt} = \frac{k}{c} \tag{3.69}$$

- **Ejercicio 2:** Un sistema mecánico similar al anterior, pero en el que un muelle de rigidez k_2 está en serie con un conjunto paralelo muelle-amortiguador (Fig. 3.31). La entrada sigue siendo

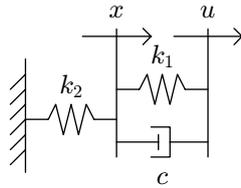


Figura 3.31: Sistema mecánico muelle-muelle-amortiguador

el desplazamiento u y la salida el desplazamiento x . Hallar la ecuación diferencial que gobierna el sistema.

Solución:

$$k_1 u + c \frac{du}{dt} = (k_1 + k_2)x + c \frac{dx}{dt} \quad (3.70)$$

- **Ejercicio 3:** Un cuerpo de masa m unido a dos amortiguadores, uno de los cuales está amarrado al suelo (ver Fig. 3.32). La entrada es el desplazamiento u en el extremo del amortiguador libre y la salida el desplazamiento x de la masa. Hallar la ecuación diferencial que gobierna el sistema.

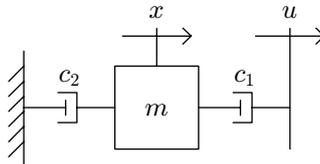


Figura 3.32: Sistema mecánico amortiguador-masa-amortiguador

Solución:

$$c_1 \frac{du}{dt} = m \frac{d^2 x}{dt^2} + (c_1 + c_2) \frac{dx}{dt} \quad (3.71)$$

- **Ejercicio 4:** Dos cuerpos unidos entre sí por un resorte (ver Fig. 3.33). Uno de ellos está amarrado al suelo a través de otro muelle y en el otro actúa una fuerza f . La salida son los desplazamientos de los dos cuerpos. Hallar el sistema de ecuaciones diferenciales que gobierna el sistema.

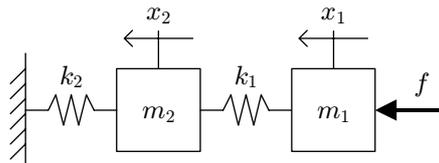


Figura 3.33: Sistema mecánico masa-muelle-masa-muelle

Solución:

$$\left. \begin{aligned} f &= m_1 \frac{d^2 x_1}{dt^2} + k_1(x_1 - x_2) \\ k_1(x_1 - x_2) &= m_2 \frac{d^2 x_2}{dt^2} + k_2 x_2 \end{aligned} \right\} \quad (3.72)$$

- **Ejercicio 5:** Dos cuerpos unidos entre sí por un resorte (ver Fig. 3.34). La entrada es la fuerza f que actúa sobre un cuerpo y la salida son los desplazamientos de los dos cuerpos.

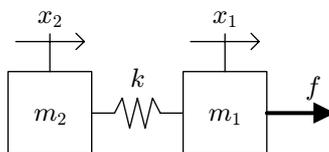


Figura 3.34: Sistema mecánico masa-muelle-masa

Solución:

$$\left. \begin{aligned} f &= m_1 \frac{d^2 x_1}{dt^2} + k(x_1 - x_2) \\ k(x_1 - x_2) &= m_2 \frac{d^2 x_2}{dt^2} \end{aligned} \right\} \quad (3.73)$$

- **Ejercicio 6:** Dos resistencias conectadas en serie a través de una capacidad (ver Fig. 3.35). La entrada es la tensión v_i aplicada sobre una resistencia y poniendo a tierra la otra. La salida es la tensión v_o en un extremo de la capacidad.

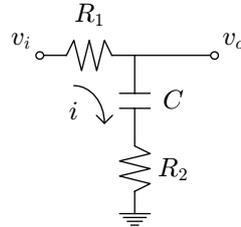


Figura 3.35: Sistema eléctrico resistencia-capacidad-resistencia

Solución:

$$R_1 C \frac{dv_i}{dt} + v_i = R_2 C \frac{dv_o}{dt} + v_o \quad (3.74)$$

- **Ejercicio 7:** Simplificar de forma gráfica o analítica el siguiente diagrama de bloques:

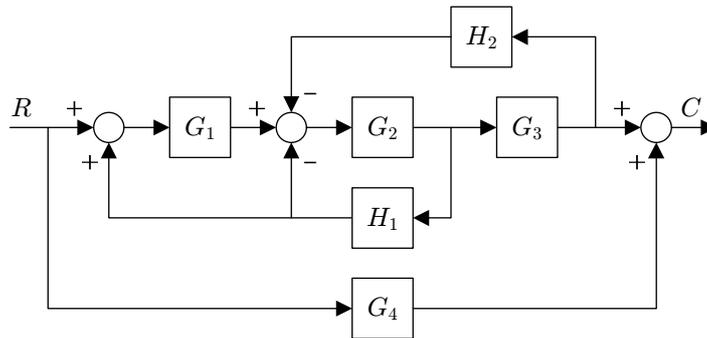


Figura 3.36: Diagrama de bloques

Solución: $\frac{C(s)}{R(s)} = \frac{G_1 G_2 G_3 + G_4 + G_2 G_4 H_1 + G_2 G_3 G_4 H_2 - G_1 G_2 G_4 H_1}{1 + G_2 H_1 + G_2 G_3 H_2 - G_1 G_2 H_1}$

- **Ejercicio 8:** Simplificar el diagrama de bloques de la Fig. 3.37 y escribir la función de transferencia que relaciona la salida con la entrada de la forma más compacta posible.

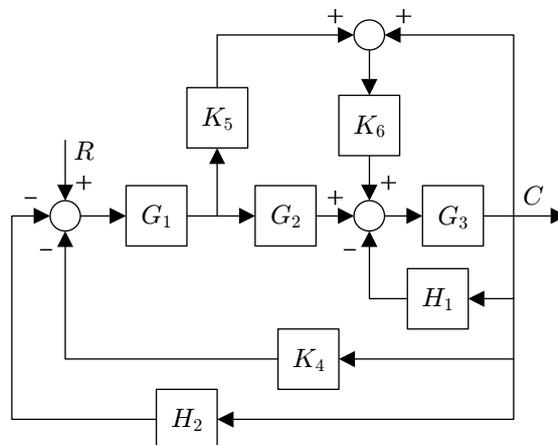


Figura 3.37: Diagrama de bloques

Solución: $\frac{C(s)}{R(s)} = \frac{G_1 G_3 (G_2 + K_5 K_6)}{1 + G_3 (H_1 - K_6) + G_1 G_3 (G_2 + K_5 K_6) (H_2 + K_4)}$

Capítulo 4

Respuesta temporal

Para analizar el comportamiento de un sistema se toma como punto de partida la representación matemática del mismo. Esta modelización, Fig. 4.1, es su función de transferencia $G(s)$.

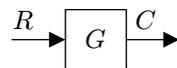


Figura 4.1: Respuesta del sistema ante una entrada

El sistema puede ser excitado con distintas señales de entrada $r(t)$. Las más utilizadas son las funciones impulso unidad, escalón unidad, rampa unidad y sinusoidal de amplitud unidad, Fig. 4.2. La respuesta del sistema ante las distintas entradas suele tener un régimen transitorio y otro permanente, aunque este último puede no darse y depende de la estabilidad del sistema.

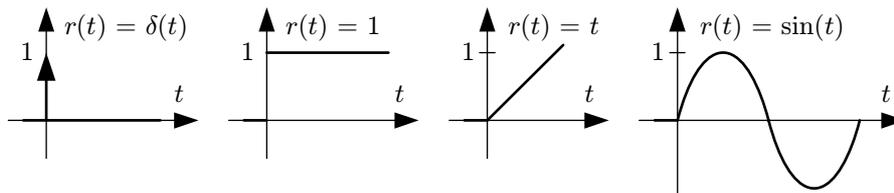


Figura 4.2: Tipos de entradas a los sistemas

En este capítulo se analizará la respuesta temporal de un sistema en función de la entrada que se imponga y de las propias características de su función de transferencia.

4.1. Sistemas de primer orden

Por lo general, la función de transferencia $G(s)$ de un sistema es una expresión racional de polinomios en s . Las raíces del denominador se llaman polos y las raíces del numerador se llaman ceros. Un sistema de primer orden se define como aquel que posee un único polo.

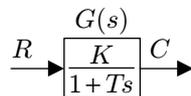


Figura 4.3: Sistema de primer orden

En la Fig. 4.3 se muestra la representación general de un sistema de primer orden. A la constante K se le llamará ganancia estática del sistema y a T constante de tiempo del sistema.

4.1.1. Respuesta ante entrada impulso

La salida temporal $c(t)$ del sistema de primer orden ante una entrada impulso unidad es:

$$c(t) = \mathcal{L}^{-1}[C(s)] = \mathcal{L}^{-1}[G(s)R(s)] = \mathcal{L}^{-1}\left[\frac{K}{1+Ts}\right] = \frac{K}{T}e^{-\frac{t}{T}} \implies \begin{cases} c(0^+) = \frac{K}{T} \\ c(\infty) = 0 \end{cases}, \quad (4.1)$$

donde se han calculado los valores inicial y final de dicha salida. La pendiente inicial de la curva se puede calcular a partir de la expresión general de la derivada:

$$\dot{c}(t) = \frac{dc(t)}{dt} = -\frac{K}{T^2}e^{-\frac{t}{T}} \implies \dot{c}(0^+) = -\frac{K}{T^2} \quad (4.2)$$

Estos resultados se pueden obtener a través de las propiedades de las transformadas de Laplace, sin necesidad de obtener la salida temporal del sistema:

$$c(0^+) = \lim_{t \rightarrow 0^+} c(t) = \lim_{s \rightarrow \infty} sC(s) = \lim_{s \rightarrow \infty} s \frac{K}{1+Ts} = \frac{K}{T} \quad (4.3)$$

$$c(\infty) = \lim_{t \rightarrow \infty} c(t) = \lim_{s \rightarrow 0} sC(s) = \lim_{s \rightarrow 0} s \frac{K}{1+Ts} = 0 \quad (4.4)$$

$$\dot{c}(0^+) = \lim_{t \rightarrow 0^+} \dot{c}(t) = \lim_{s \rightarrow \infty} s[sC(s) - c(0^+)] = \lim_{s \rightarrow \infty} s \left(s \frac{K}{1+Ts} - \frac{K}{T} \right) = -\frac{K}{T^2} \quad (4.5)$$

En la Fig. 4.4 se muestra un ejemplo de respuesta ante entrada impulso. Aparece también la recta que comienza en $\frac{K}{T}$ con pendiente $-\frac{K}{T^2}$. Se observa que dicha recta pasa por cero para $t = T$. En el ejemplo, $T = 0.33$ s.

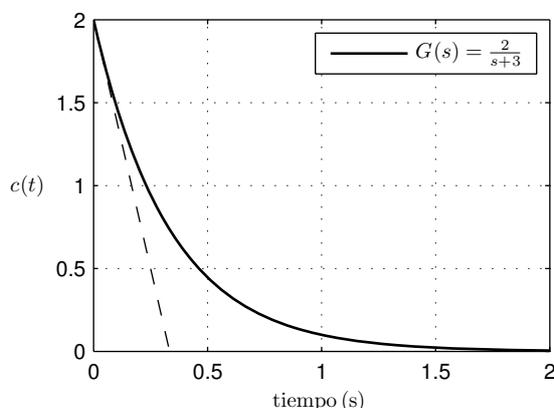


Figura 4.4: Respuesta de un sistema de primer orden ante entrada impulso

4.1.2. Respuesta ante entrada escalón

La salida temporal del sistema de primer orden ante una entrada escalón unidad es:

$$c(t) = \mathcal{L}^{-1}[C(s)] = \mathcal{L}^{-1} \left[\frac{K}{s(1+Ts)} \right] = K(1 - e^{-\frac{t}{T}}) \implies \begin{cases} c(0^+) = 0 \\ c(\infty) = K \end{cases}, \quad (4.6)$$

donde se han calculado los valores inicial y final de dicha salida. La pendiente inicial de la curva es:

$$\dot{c}(t) = \frac{dc(t)}{dt} = \frac{K}{T}e^{-\frac{t}{T}} \implies \dot{c}(0^+) = \frac{K}{T} \quad (4.7)$$

También es posible obtener estos resultados a partir de las propiedades de la transformada de Laplace. En la Fig. 4.5 se muestra la respuesta ante entrada escalón unidad del mismo ejemplo que el apartado anterior. Ahora el valor final es K , mientras que recta que sale del origen con pendiente $\frac{K}{T}$ toma el valor K para $t = T$. Estas líneas pueden usarse como referencias para dibujar la respuesta de un sistema a mano alzada.

Por tanto, el valor de la respuesta en régimen permanente coincide con la ganancia estática K . Cuanto menor sea la constante de tiempo T más rápidamente tiende la respuesta del sistema a su valor en régimen permanente. La constante de tiempo da una idea de la duración del régimen transitorio del sistema. Aproximadamente la salida llega al 62% del régimen permanente en el instante de tiempo igual a la constante de tiempo del sistema:

$$c(T) \approx 0.62K \quad (4.8)$$

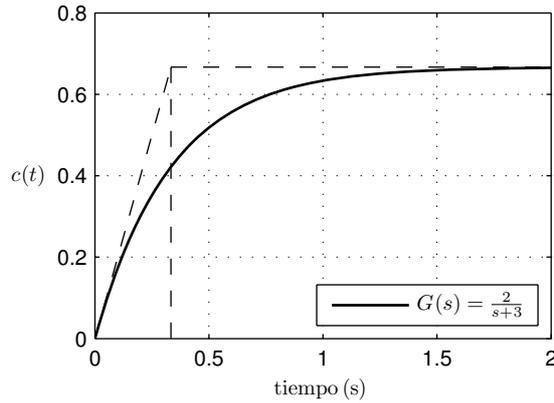


Figura 4.5: Respuesta de un sistema de primer orden ante entrada escalón

4.1.3. Respuesta ante entrada sinusoidal

La salida temporal $c(t)$ del sistema de primer orden ante una entrada sinusoidal de amplitud unidad y frecuencia ω es:

$$c(t) = \mathcal{L}^{-1}[C(s)] = \mathcal{L}^{-1} \left[\frac{K\omega}{(1+Ts)(s^2 + \omega^2)} \right] \quad (4.9)$$

$$c(t) = \underbrace{\frac{K\omega T}{1 + (\omega T)^2} e^{-\frac{t}{T}}}_{\text{transitorio}} + \underbrace{\sqrt{\frac{K^2}{1 + (\omega T)^2}} \sin[(\omega t) - \arctan(\omega T)]}_{\text{permanente}} \quad (4.10)$$

Se observa que la salida $c(t)$ posee dos sumandos: el primero es transitorio, desaparece prácticamente después de T segundos, y el segundo es una sinusoidal de frecuencia igual a la de la señal de entrada, pero con una amplitud y un retraso que dependen tanto de la frecuencia ω de entrada como de las características del sistema de primer orden.

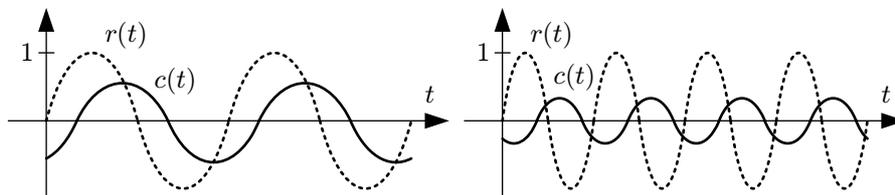


Figura 4.6: Respuestas del sistema de primer orden ante entrada sinusoidal

Si la frecuencia ω de la sinusoidal de entrada aumenta, la sinusoidal de salida poseerá una amplitud cada vez menor y un retraso cada vez mayor (cualitativamente en la Fig. 4.6). En definitiva, un sistema de primer orden actúa en el dominio de las frecuencias como un filtro pasa-baja, es decir, atenúa las frecuencias elevadas.

Una forma de obtener la amplitud de la salida y su retraso en función de la planta y la frecuencia de entrada consiste en tomar la función de transferencia de la planta y sustituir la variable s por $j\omega$. El resultado es un número complejo cuyo módulo es la amplitud de salida y cuya fase es el retraso de la salida respecto de la entrada.

$$G(s) = \frac{K}{1+Ts} \xrightarrow{s=j\omega} G(j\omega) = \frac{K}{1+j\omega T} \begin{cases} |G(j\omega)| = \sqrt{\frac{K^2}{1+(\omega T)^2}} \\ \angle G(j\omega) = -\arctan(\omega T) \end{cases} \quad (4.11)$$

Esta propiedad no es exclusiva de los sistemas de primer orden. Se cumple siempre que la entrada es sinusoidal, cualquiera que sea la expresión $G(s)$ de la función de transferencia de la planta.

4.1.4. Ejemplos de sistemas de primer orden

- **Ejemplo 1:** En la Fig. 4.7 se puede ver un ejemplo de un sistema físico de primer orden, y en la ecuación (4.12) la función de transferencia de dicho sistema.

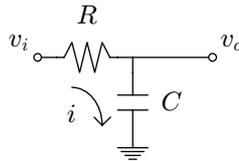


Figura 4.7: Sistema eléctrico de primer orden

$$G(s) = \frac{V_o}{V_i} = \frac{1}{1 + RCs} \quad (4.12)$$

- **Ejemplo 2:** En la Fig. 4.8 se puede ver un ejemplo de un sistema físico de primer orden, y en la ecuación (4.13) la función de transferencia de dicho sistema.

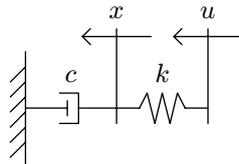


Figura 4.8: Sistema mecánico de primer orden

$$G(s) = \frac{X}{U} = \frac{k}{k + cs} \quad (4.13)$$

4.2. Sistemas de segundo orden

Un sistema de segundo orden es aquel que posee dos polos. Este tipo de sistemas se suele representar de la siguiente forma:

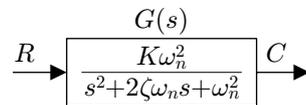


Figura 4.9: Sistema de segundo orden

La constante K es la ganancia estática del sistema, ζ es el amortiguamiento y ω_n es la frecuencia natural. Dependiendo del carácter de los polos, el sistema de segundo orden puede ser:

- **Sistema subamortiguado.** El amortiguamiento posee un valor entre 0 y 1 y los polos del sistema de segundo orden son complejo-conjugados. Su posición aparece en la siguiente ecuación:

$$p_{1,2} = -\zeta\omega_n \pm \omega_n\sqrt{1 - \zeta^2}j = -\sigma \pm \omega_d j \quad (4.14)$$

La constante σ es la atenuación del sistema y ω_d la frecuencia natural amortiguada. En la Fig. 4.10 se define el ángulo ϕ que forman los polos complejo-conjugados en el plano complejo S con el origen.

- **Sistema sobreamortiguado.** El amortiguamiento es mayor que la unidad y los polos del sistema de segundo orden son reales localizados en:

$$p_{1,2} = -\zeta\omega_n \pm \omega_n\sqrt{\zeta^2 - 1} \quad (4.15)$$

- **Sistema críticamente amortiguado.** El amortiguamiento es igual a la unidad y los polos son reales e iguales:

$$p_{1,2} = -\omega_n \text{ doble} \quad (4.16)$$

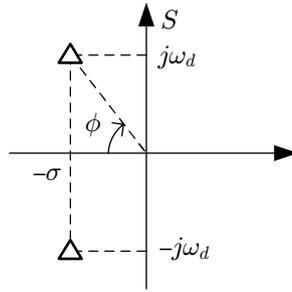


Figura 4.10: Localización de los polos de un sistema de segundo orden subamortiguado

Cualquiera que sea el amortiguamiento del sistema, existen tres puntos clave de la respuesta temporal que siempre cumplen los sistemas de segundo orden ante una entrada escalón unidad:

$$c(0^+) = \lim_{t \rightarrow 0^+} c(t) = \lim_{s \rightarrow \infty} sC(s) = \lim_{s \rightarrow \infty} s \frac{K\omega_n^2}{(s^2 + 2\zeta\omega_n s + \omega_n^2)s} = 0 \quad (4.17)$$

$$c(\infty) = \lim_{t \rightarrow \infty} c(t) = \lim_{s \rightarrow 0} sC(s) = \lim_{s \rightarrow 0} s \frac{K\omega_n^2}{(s^2 + 2\zeta\omega_n s + \omega_n^2)s} = K \quad (4.18)$$

$$\dot{c}(0^+) = \lim_{t \rightarrow 0^+} \dot{c}(t) = \lim_{s \rightarrow \infty} s[sC(s) - c(0^+)] = \lim_{s \rightarrow \infty} s \left[s \frac{K\omega_n^2}{(s^2 + 2\zeta\omega_n s + \omega_n^2)s} - 0 \right] = 0 \quad (4.19)$$

Es decir, la respuesta temporal de todos los sistemas de segundo orden comienzan en el origen con pendiente nula, y alcanzan en régimen permanente el valor de la ganancia estática K .

- **Sistema oscilatorio.** El amortiguamiento es cero y los polos del sistema de segundo orden son complejo conjugados imaginarios puros localizados en:

$$p_{1,2} = \pm j\omega_n \quad (4.20)$$

En este último caso no existe ningún valor de régimen permanente ante entrada escalón unidad.

4.2.1. Respuesta subamortiguada ante entrada escalón

La respuesta de un sistema subamortiguado ($\zeta < 1$) ante una entrada escalón unidad es:

$$c(t) = K \left[1 - \frac{e^{-\sigma t}}{\sqrt{1 - \zeta^2}} \sin(\omega_d t + \phi) \right] \quad (4.21)$$

En la Fig. 4.11 se muestra un ejemplo de respuesta temporal de un sistema subamortiguado. Se trata de una señal sinusoidal cuya amplitud se va atenuando según un patrón exponencial.

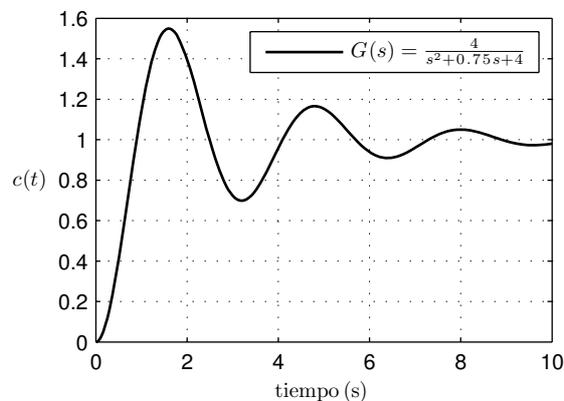


Figura 4.11: Respuesta de un sistema de segundo orden subamortiguado ante entrada escalón

Existen varios puntos clave en la respuesta temporal. El primero es el tiempo de levantamiento t_r , y es el instante en el que la salida pasa por primera vez por el valor de su régimen permanente.

$$t_r = \frac{\pi - \phi}{\omega_d} \quad (4.22)$$

El tiempo de tipo t_p es el instante en el que la salida temporal alcanza su primer máximo. A la diferencia entre el valor del máximo y el valor en régimen permanente, expresada en por unidad respecto del valor en régimen permanente, se le llama sobreimpulso máximo M_p .

$$t_p = \frac{\pi}{\omega_d} \quad (4.23)$$

$$M_p = \frac{c(t_p) - c(\infty)}{c(\infty)} = e^{\frac{-\pi\zeta}{\sqrt{1-\zeta^2}}} = e^{\frac{-\pi}{\tan\phi}} \quad (4.24)$$

El tiempo de establecimiento se define como el instante a partir del cual la respuesta temporal queda circunscrita en una banda del 2% ó del 5% en torno al valor en régimen permanente.

$$t_s(2\%) \approx \frac{4}{\zeta\omega_n} = \frac{4}{\sigma} \quad (4.25)$$

$$t_s(5\%) \approx \frac{3}{\zeta\omega_n} = \frac{3}{\sigma} \quad (4.26)$$

En los sistemas de control no es deseable que exista una respuesta con mucho sobreimpulso ni muy oscilatoria. Se suele buscar que el sistema controlado posea un sobreimpulso entre el 0% y el 20% con el menor tiempo de establecimiento posible.

4.2.2. Respuesta sobreamortiguada ante entrada escalón

La respuesta de un sistema sobreamortiguado ($\zeta > 1$) ante una entrada escalón unidad es:

$$c(t) = K \left[1 + \frac{\omega_n}{2\sqrt{\zeta^2 - 1}} \left(\frac{e^{-p_1 t}}{p_1} - \frac{e^{-p_2 t}}{p_2} \right) \right] \quad (4.27)$$

donde p_1 y p_2 son las raíces reales definidas en (4.15). En la Fig. 4.12 se muestra un ejemplo de respuesta temporal de un sistema sobreamortiguado. Los sistemas sobreamortiguados no poseen sobreimpulso.

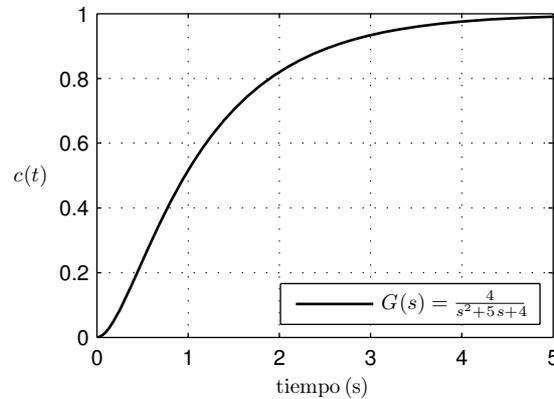


Figura 4.12: Respuesta de un sistema de segundo orden sobreamortiguado ante entrada escalón

En este caso existen pocas referencias para dibujar a mano alzada la respuesta temporal. Una forma aproximada de representar la respuesta consiste en dividir la función de transferencia en dos sistemas de primer orden, Fig. 4.13, y determinar cuál de los dos es más lento, es decir, el que tiene la constante de tiempo más grande.

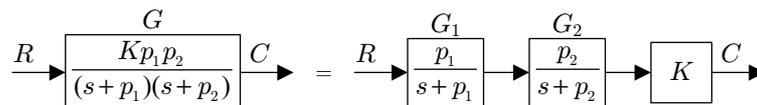


Figura 4.13: Respuesta de un sistema de segundo orden sobreamortiguado

La respuesta ante una entrada escalón se asemejará a la respuesta del sistema de primer orden más lento añadiendo un pequeño retraso, aproximadamente igual a la constante de tiempo del sistema de primer orden más rápido, y haciendo la pendiente de salida nula. En la Fig. 4.14 se compara la salida del

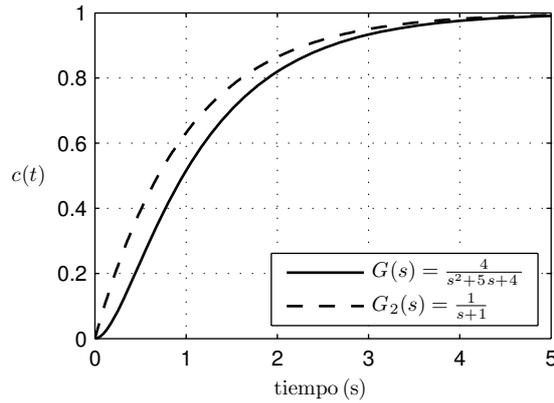


Figura 4.14: Respuesta de un sistema de segundo orden sobreamortiguado ante entrada escalón

sistema sobreamortiguado de la anterior Fig. 4.12 con la del sistema de primer orden que impone el polo más lento.

Una forma intuitiva de explicar este comportamiento consiste en imaginar en la Fig. 4.13 cuál será la entrada de la función de transferencia G_2 suponiendo que sea ésta la que posee el polo más lento. Su entrada será una exponencial con una constante de tiempo pequeña, es decir, aproximadamente un escalón unidad retrasado la constante de tiempo de G_1 .

4.2.3. Respuesta críticamente amortiguada ante entrada escalón

La respuesta de un sistema críticamente amortiguado ($\zeta = 1$) ante una entrada escalón unidad es:

$$c(t) = K[1 + e^{-\omega_n t}(1 + \omega_n t)] \quad (4.28)$$

En la Fig. 4.15 se muestra un ejemplo de respuesta temporal de un sistema críticamente amortiguado. Este tipo de sistemas tampoco poseen sobreimpulso.

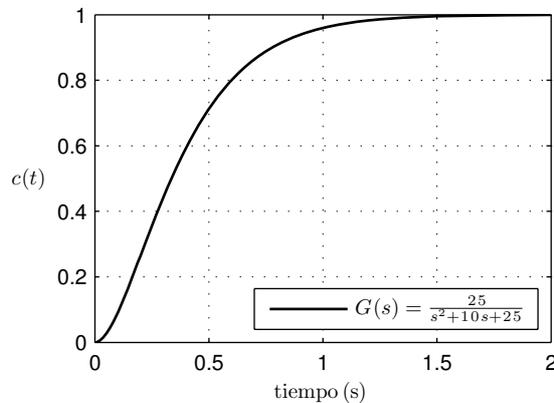


Figura 4.15: Respuesta de un sistema de segundo orden críticamente amortiguado con entrada escalón

4.2.4. Respuesta oscilatoria ante entrada escalón

La salida de un sistema oscilatorio es:

$$c(t) = K[1 - \cos(\omega_n t)] \quad (4.29)$$

En la Fig. 4.16 se muestra un ejemplo de respuesta temporal de un sistema oscilatorio. En este tipo de sistemas el sobreimpulso es del 100%.

4.2.5. Respuesta ante entrada impulso

La respuesta de un sistema ante una entrada impulso se puede obtener a partir de la respuesta ante una entrada escalón. En la Fig. 4.17 se observa cómo la respuesta ante una entrada impulso se puede conseguir derivando directamente la respuesta del sistema con entrada escalón unidad.

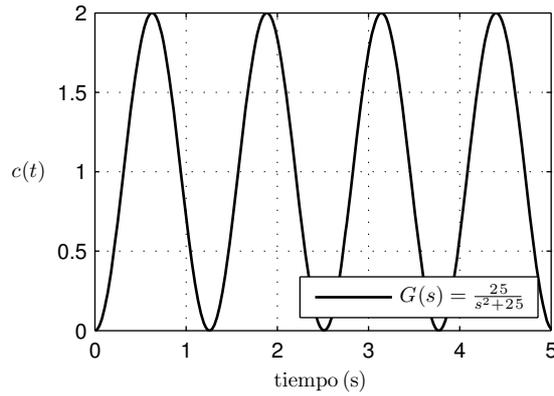


Figura 4.16: Respuesta de un sistema de segundo orden oscilatorio con entrada escalón

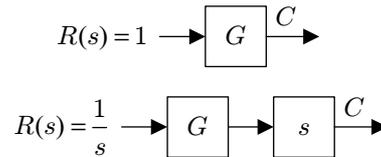


Figura 4.17: Respuesta de un sistema ante entrada impulso

Por tanto, basta con derivar las respuestas temporales de los apartados anteriores para obtener la respuesta del sistema ante entrada impulso. Como los sistemas de segundo orden, cualquiera que sea su amortiguamiento, comienzan y acaban con derivada nula, la respuesta ante entrada impulso comienza y acaba con valor nulo.

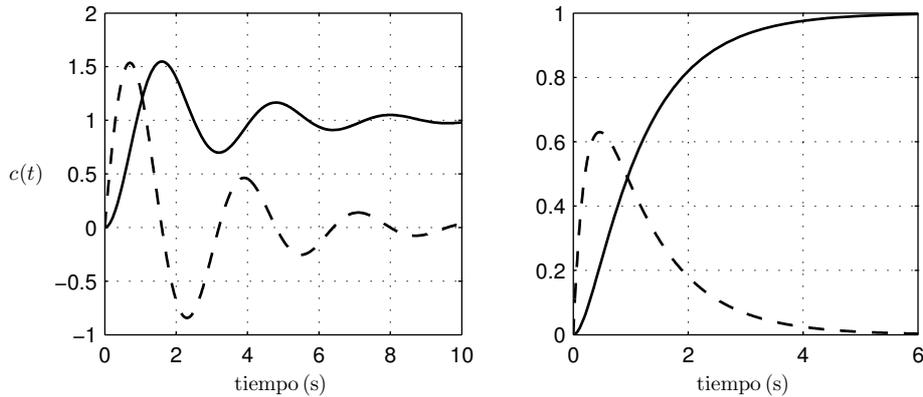


Figura 4.18: Respuestas de sistemas de segundo orden ante entrada escalón (línea continua) y ante entrada impulso (línea discontinua)

4.3. Sistemas de orden superior

El comportamiento de los sistemas de orden superior, es decir, de aquellos que poseen tres o más polos, depende fundamentalmente del carácter de los polos más lentos del sistema. Como se ha visto en el apartado anterior, el polo más lento es el que posee la constante de tiempo más grande, es decir, aquel polo se encuentran más cerca del origen en el plano complejo S .

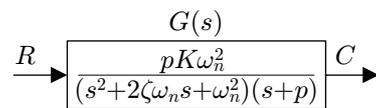


Figura 4.19: Sistema de tercer orden

Sea un sistema de tercer orden, Fig. 4.19, en el que existe un polo real y dos complejo-conjugados. La respuesta temporal, depende de la posición relativa de los tres polos del sistema. La Fig. 4.20 muestra el

caso particular de que los polos complejo-conjugados sean los más lentos. La respuesta se asemeja a la del sistema de segundo orden subamortiguado, pero está un poco retrasada en el tiempo y tiene un menor sobreimpulso. Ese retraso en el tiempo es aproximadamente igual a la constante de tiempo del polo real.

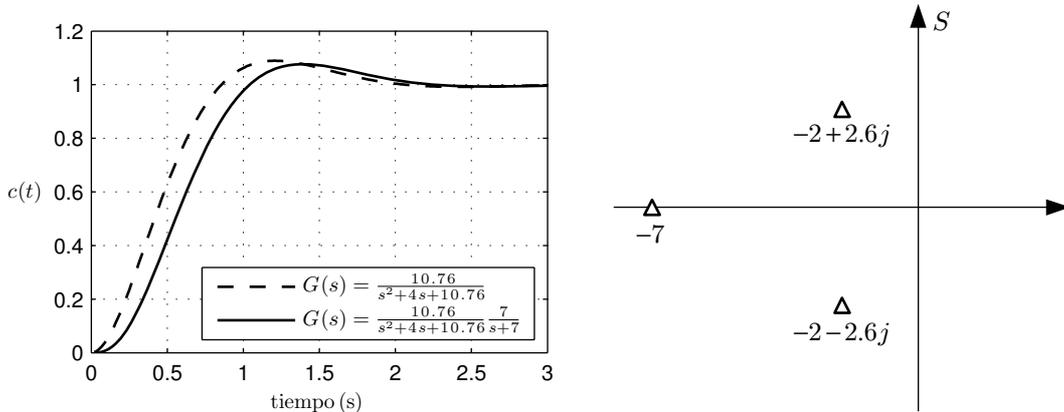


Figura 4.20: Respuesta temporal y posición de los polos de un sistema de tercer orden

Por tanto, la inclusión de polos adicionales a un determinado sistema no influye en la respuesta temporal del mismo mientras los nuevos polos se encuentren suficientemente alejados del eje imaginario del plano complejo S respecto a los que ya tenía el sistema. Por norma general se puede admitir que los polos que se encuentren más alejados que cinco veces la distancia de los polos más lentos al eje imaginario, tienen una influencia en la respuesta temporal del sistema prácticamente despreciable. Por esta razón, los polos lentos se llaman también polos dominantes del sistema.

En la Fig. 4.21 se muestra un caso particular en el que el polo real es el más lento. La respuesta se asemeja a la del sistema de primer orden, con un retraso adicional y pendiente inicial nula.

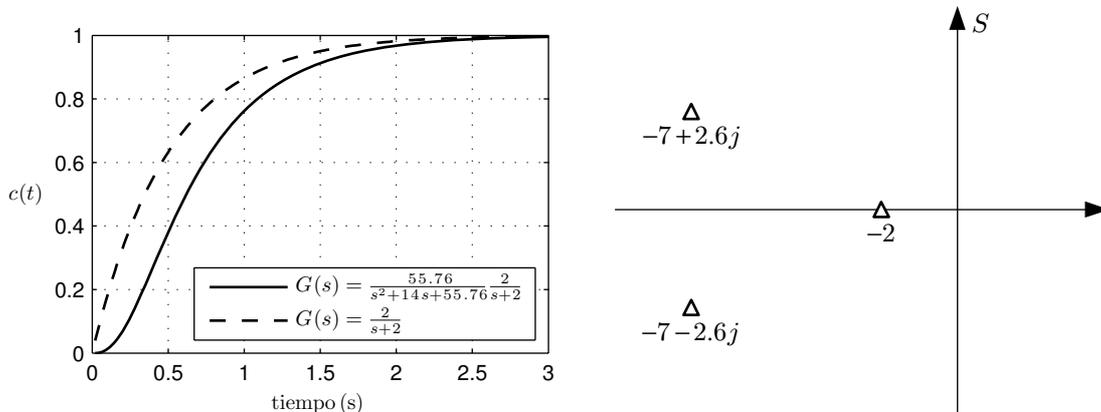


Figura 4.21: Respuesta temporal y posición de los polos de un sistema de tercer orden

4.4. Influencia de los ceros

Los ceros del sistema son las raíces del numerador de la función de transferencia. La presencia de ceros en la función de transferencia, modifica la respuesta que se podría esperar del sistema atendiendo a la posición de los polos. Se va a mostrar la con el ejemplo de la Fig. 4.22.

La presencia de un cero real negativo hace el efecto contrario un polo, es decir, adelanta la respuesta temporal en lugar de retrasarla. Además, modifica las condiciones iniciales de la respuesta temporal. Si el sistema tenía dos polos, la pendiente inicial del sistema pasa de ser nula a no nula. Si el sistema tenía tres polos, la derivada segunda en el instante inicial para ser nula a no nula. Y así sucesivamente.

En el caso concreto de sistema de segundo orden con cero, como es el caso de la Fig. 4.22, se puede calcular la pendiente inicial de forma sencilla:

$$\dot{c}(0^+) = \lim_{s \rightarrow \infty} s[sC(s) - c(0^+)] = \lim_{s \rightarrow \infty} s \left[s \frac{K\omega_n^2(s+z)}{z(s^2 + 2\zeta\omega_n s + \omega_n^2)s} - 0 \right] = \frac{K\omega_n^2}{z} \quad (4.30)$$

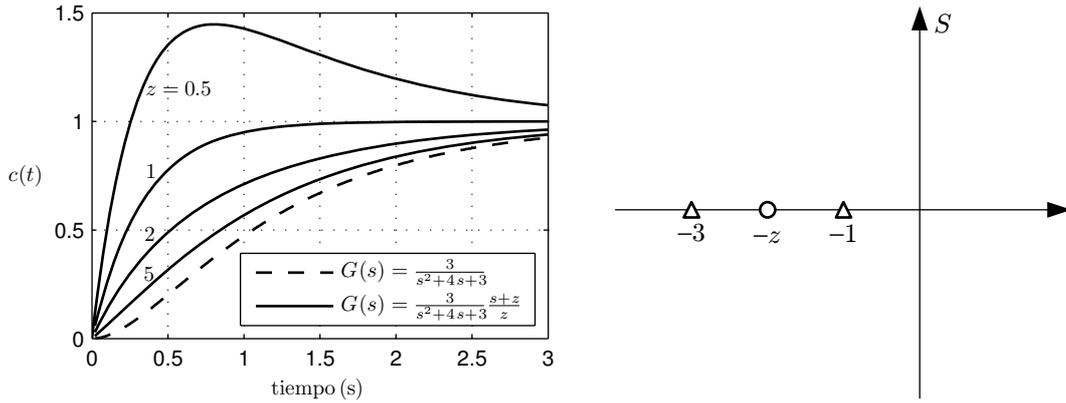


Figura 4.22: Influencia del cero en la respuesta temporal del sistema

Por tanto, conforme el cero está más cerca del origen mayor es el valor de la pendiente inicial.

Otro efecto muy interesante que se puede dar en un sistema es la cancelación de un polo con un cero. Esto ocurre en el ejemplo de la Fig. 4.22 cuando z toma el valor 1. En ese caso, el sistema disminuye su orden en una unidad y pasa de ser de orden dos a ser de orden uno. Se puede comprobar que la respuesta temporal en ese caso es efectivamente una exponencial con constante de tiempo igual a la que fija el polo que permanece en el sistema.

En la práctica, para cancelar un polo con un cero no es necesario que ambos se encuentren exactamente en la misma posición. Basta con que estén muy próximos para que el efecto de uno se anule con el del otro.

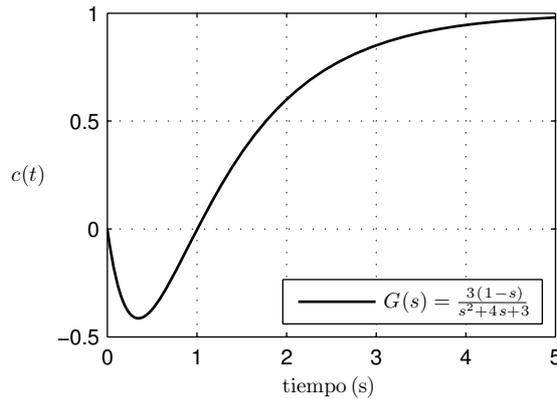


Figura 4.23: Respuesta temporal de un sistema de fase no mínima

Los sistemas de fase no mínima son aquellos que poseen un cero real positivo. La respuesta temporal de este tipo de sistemas tiene la característica de que comienza evolucionando en la dirección contraria al valor en régimen permanente, Fig. 4.23.

4.5. Ejercicios propuestos

- **Ejercicio 1:** Hallar la respuesta temporal del sistema de la Fig. 4.24 para una entrada escalón unidad para los valores de la ganancia $K = 1, 2$ y 5 .

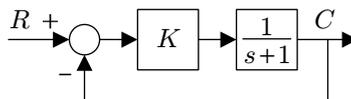


Figura 4.24: Sistema de control

- **Ejercicio 2:** Dibujar la respuesta temporal del sistema de la Fig. 4.25 para una entrada escalón unidad y valores de $K = 10, 50$ y 100 . Calcular el valor de K que consigue que el sistema sea críticamente amortiguado.

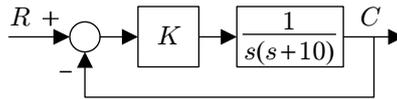


Figura 4.25: Sistema de control

- **Ejercicio 3:** La respuesta temporal de un sistema cuya función de transferencia se desconoce presenta un sobreimpulso del 20% a los 413 ms. En régimen estacionario se alcanza el valor exacto de la señal de referencia. Se pide deducir la función de transferencia del sistema, la frecuencia de las oscilaciones, el tiempo de establecimiento del sistema y la posición de las raíces en el plano S .

Soluciones: $G(s) = \frac{73.1}{s^2 + 7.8s + 73.1}$, $\omega_d = 7.6$ rad/s, $t_s = 1.026$ s y $p_{1,2} = -3.9 \pm 7.6j$ ($\zeta = 0.456$)

- **Ejercicio 4:** Dibujar la respuesta temporal del sistema de la Fig. 4.26 para una entrada escalón unidad y $K = 1$. Calcular los valores de K que consiguen una respuesta temporal bien amortiguada.

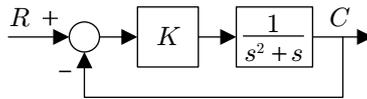


Figura 4.26: Sistema de control

Solución: Tomando $1 > \zeta > 0.5$, entonces $0.25 < K < 1$

- **Ejercicio 5:** Obtener el tiempo de crecimiento y el sobreimpulso del sistema de la Fig. 4.27.

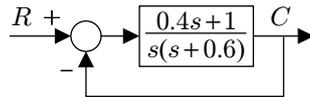


Figura 4.27: Sistema de control

Solución: $t_r = 2.42$ s y $M_p = 16.3\%$

- **Ejercicio 6:** Dibujar, en una gráfica, la respuesta al escalón unidad de tres plantas cuyas funciones de transferencia son, respectivamente:

$$G_1(s) = \frac{5}{s^2 + 2s + 5} \quad (4.31)$$

$$G_2(s) = \frac{20}{(s^2 + 2s + 5)(s + 4)} \quad (4.32)$$

$$G_3(s) = \frac{20(s + 4)}{4s^2 + 8s + 20} \quad (4.33)$$

- **Ejercicio 7:** La Fig. 4.28 muestra un péndulo invertido de 1 m de longitud cuya masa m de 1 kg se puede considerar concentrada en un punto de su extremo, moviéndose en un medio sin viscosidad.

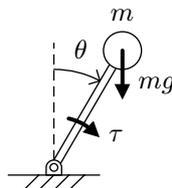


Figura 4.28: Sistema de un péndulo invertido

Su eje de giro está actuado por un motor de corriente continua con inductancia despreciable, constante de par 5 Nm/A y resistencia 1 Ω . Se pide representar el diagrama de bloques del sistema con entrada tensión $V(s)$ y salida ángulo de giro $\Theta(s)$. Determinar la función de transferencia del sistema. Para el sistema de realimentación negativa unitaria con controlador proporcional obtener la ganancia K del controlador que hace que el sistema esté críticamente amortiguado.

Solución:

Las ecuaciones que gobiernan el sistema de la Fig. 4.28 son:

$$\left. \begin{array}{l} v = Ri + e \\ e = K_m \frac{d\theta}{dt} \\ \tau = K_m i \\ \tau + mgl \sin \theta = ml^2 \frac{d^2\theta}{dt^2} \end{array} \right\} \xrightarrow{\sin \theta \approx \theta} \left. \begin{array}{l} v = Ri + e \\ e = K_m \frac{d\theta}{dt} \\ \tau = K_m i \\ \tau + mgl\theta = ml^2 \frac{d^2\theta}{dt^2} \end{array} \right\} \xrightarrow{\mathcal{L}} \left. \begin{array}{l} V = RI + E \\ E = K_m s\Theta \\ T = K_m I \\ T + mgl\Theta = ml^2 s^2 \Theta \end{array} \right\} \quad (4.34)$$

Sustituyendo los valores numéricos de las constantes:

$$\left. \begin{array}{l} V = I + E \\ E = 5s\Theta \\ T = 5I \\ T + 10\Theta = s^2\Theta \end{array} \right\} \Rightarrow \frac{\Theta(s)}{V(s)} = \frac{5}{s^2 + 25s - 10} \quad (4.35)$$

El sistema controlado por medio de un controlador proporcional es:

$$\frac{\Theta(s)}{\Theta_r(s)} = \frac{5K}{s^2 + 25s + 5K - 10} \quad (4.36)$$

Identificando coeficientes del denominador:

$$\left. \begin{array}{l} 5K - 10 = \omega_n^2 \\ 25 = 2\zeta\omega_n \end{array} \right\} \xrightarrow{\zeta=1} \begin{array}{l} \omega_n = 12.5 \text{ rad/s} \\ K = 33.25 \text{ V/rad} \end{array} \quad (4.37)$$

Es interesante ver en este problema cómo un sistema intrínsecamente inestable como es un péndulo invertido puede hacerse estable al controlarlo. Por otro lado, el valor del ángulo $\theta(t)$ en régimen permanente es mayor que cualquier ángulo de referencia θ_r que se introduzca. Esto es así porque la diferencia de posiciones —el error— genera el par motor que contrarresta el peso de la masa.

Capítulo 5

Error en régimen permanente

Un aspecto importante a tener en cuenta es el comportamiento de un sistema ante diversas entradas en régimen permanente. En cualquier sistema físico de control existe un error inherente, que es el error en estado estacionario en respuesta a determinados tipos de entradas. Puede ocurrir que un sistema presente o no error en régimen permanente ante diferentes entradas. El error en régimen permanente de los sistemas en lazo cerrado viene determinado por la función de transferencia en lazo abierto, como se verá a continuación.

5.1. Definición de error en régimen permanente

Generalmente el error en régimen permanente sólo se suele definir en los sistemas controlados donde la referencia de entrada y la salida controlada son dimensionalmente coherentes, es decir, tienen las mismas unidades. Así, el error a lo largo del tiempo se define como la diferencia entre la entrada y la salida:

$$e(t) = r(t) - c(t) \quad (5.1)$$

Así, el error en régimen permanente e_{ss} (del inglés *steady-state error*) se define como el límite cuando el tiempo tiene a infinito de la señal temporal del error:

$$e_{ss} = \lim_{t \rightarrow \infty} e(t) = \lim_{t \rightarrow \infty} [r(t) - c(t)] \quad (5.2)$$

Al error en régimen permanente se le suele llamar muchas veces simplemente *error*, y puede ser nulo, finito o infinito. Evidentemente es un valor con dimensión, y sus unidades son las mismas que las de la entrada y la salida.

De cara al control de sistemas, en general, lo deseable es que el error tienda a cero (o a una cantidad insignificante) lo más rápidamente posible. Así se garantiza que la salida controlada coincide con la referencia. En este capítulo se deja de lado la rapidez con que se alcanza el régimen permanente y simplemente se analiza el valor que alcanza el error.

5.2. Error en sistemas con realimentación negativa unitaria

La definición anterior del error se aplica frecuentemente a los sistemas de realimentación negativa unitaria (Fig. 5.1). En este caso particular, la señal del error se puede identificar a la salida del comparador de la referencia y la salida.

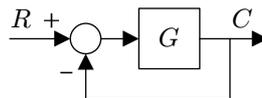


Figura 5.1: Sistema de control en lazo cerrado

La función de transferencia en lazo abierto del sistema, en este caso $G(s)$, puede tener diversos polos y ceros. En este capítulo la constante K representa la ganancia estática de la función de transferencia en lazo abierto (5.3) y N es el número de polos de dicha función de transferencia que se encuentran en el origen del plano S .

$$G(s) = K \frac{(1 + T_1 s) \dots (1 + \alpha_1 s + \beta_1 s^2)}{s^N (1 + T_2 s) \dots (1 + \alpha_2 s + \beta_2 s^2)} \quad (5.3)$$

El valor de N en la función de transferencia en lazo abierto, determina el *tipo* del sistema. Así, si un sistema posee en lazo abierto dos polos en el origen (es decir, $N = 2$), entonces se dice que dicho sistema es de tipo II. El tipo es una característica que se puede decir de cualquier función de transferencia, pero es mejor referirla a todo el sistema controlado (a través de su función de transferencia en lazo abierto).

El error en régimen permanente definido en el dominio temporal se puede pasar al dominio de Laplace aplicando el teorema del valor final:

$$e_{ss} = \lim_{t \rightarrow \infty} [r(t) - c(t)] = \lim_{s \rightarrow 0} s[R(s) - C(s)] \quad (5.4)$$

Para el caso del sistema de la Fig. 5.1, se puede desarrollar la expresión del error:

$$e_{ss} = \lim_{s \rightarrow 0} s[R(s) - C(s)] = \lim_{s \rightarrow 0} s \left[R(s) - \frac{G(s)}{1 + G(s)} R(s) \right] = \lim_{s \rightarrow 0} \frac{sR(s)}{1 + G(s)} \quad (5.5)$$

Por tanto, el valor del error depende de la forma de la entrada y del tipo del sistema. En los siguientes apartados se particulariza este resultado para distintas formas de entrada (escalón, rampa, etc.).

5.2.1. Error de posición

El error de posición en estado estacionario es el que se produce en el sistema ante una entrada escalón. Para el caso de escalón unidad, su valor es:

$$e_p = \lim_{s \rightarrow 0} \frac{1}{1 + G(s)} = \frac{1}{1 + \lim_{s \rightarrow 0} G(s)} = \frac{1}{1 + K_p} = \begin{cases} \frac{1}{1+K} & \text{tipo 0} \\ 0 & \geq \text{tipo I} \end{cases} \quad (5.6)$$

La constante K_p se llama *coeficiente de error de posición* y su valor es igual a la ganancia estática de la función de transferencia en lazo abierto del sistema o infinito, dependiendo del tipo del sistema, como se puede deducir de la expresión (5.3).

5.2.2. Error de velocidad

El error de velocidad en estado estacionario es el que se produce en el sistema ante una entrada rampa. Para el caso de rampa con pendiente unidad, su valor es:

$$e_v = \lim_{s \rightarrow 0} \frac{1}{s + sG(s)} = \frac{1}{\lim_{s \rightarrow 0} sG(s)} = \frac{1}{K_v} = \begin{cases} \infty & \text{tipo 0} \\ \frac{1}{K} & \text{tipo I} \\ 0 & \geq \text{tipo II} \end{cases} \quad (5.7)$$

La constante K_v se llama *coeficiente de error de velocidad*. Su valor es igual a cero, o bien igual a la ganancia estática de la función de transferencia en lazo abierto del sistema, o bien infinito, dependiendo del tipo del sistema.

5.2.3. Error de aceleración

El error de aceleración en estado estacionario es el que se produce en el sistema ante una entrada parabólica. Para el caso de parábola unidad, su valor es:

$$e_a = \lim_{s \rightarrow 0} \frac{1}{s^2 + s^2G(s)} = \frac{1}{\lim_{s \rightarrow 0} s^2G(s)} = \frac{1}{K_a} = \begin{cases} \infty & \leq \text{tipo I} \\ \frac{1}{K} & \text{tipo II} \\ 0 & \geq \text{tipo III} \end{cases} \quad (5.8)$$

La constante K_a se llama *coeficiente de error de aceleración*. Su valor es igual a cero, o bien igual a la ganancia estática de la función de transferencia en lazo abierto del sistema, o bien infinito, dependiendo del tipo del sistema.

5.2.4. Resumen de errores

La Tabla 5.1 resume los errores en estado estacionario en función de la entrada y el tipo del sistema. Se observa que se puede incrementar el tipo de sistema para mejorar el comportamiento de estado estacionario, sin embargo esto hace que sea más difícil controlar el sistema y puede introducir inestabilidades. Hay que recordar que esta tabla de errores sólo es aplicable a los sistemas de control con realimentación negativa unitaria.

Tabla 5.1: Errores en régimen permanente

Entrada	$\frac{1}{s}$	$\frac{1}{s^2}$	$\frac{1}{s^3}$	$\frac{1}{s^4}$
Tipo 0	$\frac{1}{1+K_p}$	∞	∞	∞
Tipo I	0	$\frac{1}{K_v}$	∞	∞
Tipo II	0	0	$\frac{1}{K_a}$	∞

5.3. Magnitud y unidades del error

Por el convenio de signos de esta expresión, un error positivo significa que la salida no ha alcanzado la referencia de entrada, mientras que un error negativo —que se puede dar— significa que la salida es mayor que la entrada. Como ya se ha dicho, el error tiene siempre las mismas unidades que la entrada y la salida. Evidentemente, cuando el error es nulo o infinito, se obvian las unidades. Cuando el error es finito puede ser mayor o menor en función de la amplitud de la entrada al sistema.

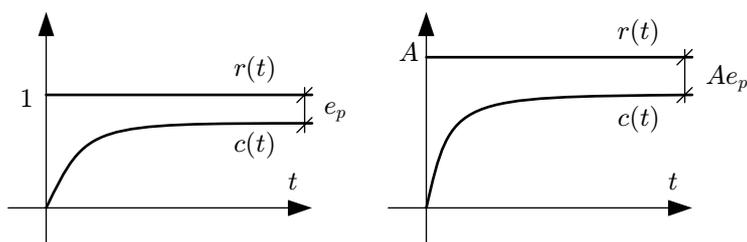


Figura 5.2: Respuestas ante entradas escalón

En la Fig. 5.2 se muestra cómo, en el caso del error de posición, el error es directamente proporcional a la amplitud del escalón de entrada. Para un escalón de amplitud A , el error es A veces mayor que error ante entrada escalón unidad. Esto se puede demostrar con la expresión (5.9) que calcula el error ante un escalón de amplitud A y compararlo con el que se calculó en (5.6).

$$e_p = \lim_{s \rightarrow 0} s \frac{1}{1+G(s)} \frac{A}{s} = \lim_{s \rightarrow 0} \frac{A}{1+G(s)} = \frac{A}{1+\lim_{s \rightarrow 0} G(s)} = \frac{A}{1+K_p} \quad (5.9)$$

Una forma de expresar el error de posición para cualquier amplitud que posea el escalón de entrada es darlo como un porcentaje respecto de la entrada (5.10). En el caso de que la salida sea mayor que la referencia no se habla de “porcentajes negativos”, se debe dar el porcentaje siempre en positivo y, como mucho, especificar si es por encima o por debajo de la referencia. El error ante entrada escalón unidad sería el valor del error en por unidad.

$$e_p = \frac{100}{1+K_p} \% \quad (5.10)$$

Para el caso de entradas dependientes del tiempo, como rampas, parábolas, etc., la forma de expresar el valor del error para cualquier entrada es diferente. En la Fig. 5.3 se observa cómo el valor del error también es proporcional a la pendiente de la entrada rampa. Sin embargo, existe algo invariante en ambos casos, a saber, el retraso en el tiempo entre las dos señales, e_t .

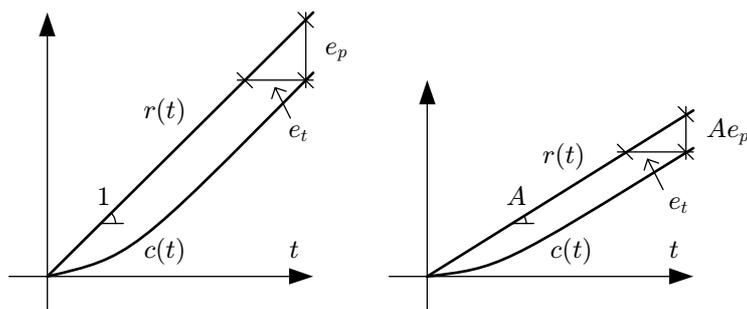


Figura 5.3: Respuestas ante entradas rampa

Por esta razón, se utiliza la medida del error en el tiempo, en segundos, para expresarlo de forma independiente a la pendiente de entrada. Como se puede deducir en la Fig. 5.3, la magnitud numérica de los errores e_t y e_p coincide si la rampa de entrada tiene pendiente unidad. Por tanto, se puede emplear el valor numérico del error en segundos como resultado del límite de la ecuación (5.7), aunque no se conozcan las unidades de la entrada ni de la salida. También por esta razón, es habitual dar el valor del coeficiente de error de velocidad K_v con unidades de s^{-1} .

5.4. Error en sistemas con realimentación no unitaria

Cuando la realimentación de un sistema de control en lazo cerrado no es unitaria, se plantea un problema en la definición del error que se dio en la ecuación (5.2), porque la señal de referencia puede tener unidades diferentes que la señal de salida. Como no se puede definir la diferencia entre la entrada y la salida, se calcula el límite cuando el tiempo tiende a infinito de la señal $e(t)$ de error del lazo:

$$e_{ss} = \lim_{t \rightarrow \infty} e(t) = \lim_{t \rightarrow \infty} [r(t) - b(t)] \quad (5.11)$$

$$e_{ss} = \lim_{s \rightarrow 0} sE(s) = \lim_{s \rightarrow 0} s[R(s) - B(s)] = \lim_{s \rightarrow 0} \frac{sR(s)}{1 + G(s)H(s)} \quad (5.12)$$

De algún modo la variable de salida “verdadera” del sistema es $B(s)$ y no $C(s)$, pero el observador sólo es capaz de ver $C(s)$. Esta idea se ve reforzada por el hecho de que muchas veces el bloque $H(s)$ corresponde exclusivamente a la función de transferencia del sensor de la señal $C(s)$, por lo que el sistema controlado es capaz de “observar” la señal $B(s)$ en lugar de $C(s)$.

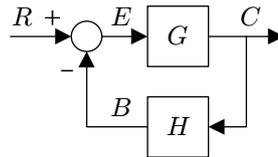


Figura 5.4: Sistema de control en lazo cerrado

Se observa en la ecuación (5.12), que el error en régimen permanente vuelve a depender del tipo de la función de transferencia en lazo abierto, en este caso es el producto $G(s)H(s)$. Por tanto, la Tabla 5.1 sigue sirviendo para el cálculo de los errores, tomando el tipo de la función de transferencia en lazo abierto y calculando los coeficientes de error también con la expresión de la función de transferencia en lazo abierto. Este hecho corrobora la afirmación que se apuntó en la introducción del capítulo de que la magnitud del error en régimen permanente de los sistemas de control en lazo cerrado viene determinada por el tipo y la ganancia estática de la función de transferencia en lazo abierto.

5.5. Error en sistemas con varias entradas

En un sistema controlado, las perturbaciones son entradas al mismo en puntos distintos al de la referencia. Estas entradas afectan a la salida del sistema en régimen permanente, y por tanto, modifican la magnitud del error.

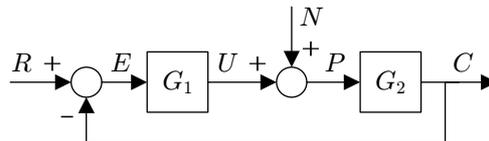


Figura 5.5: Sistema de control en lazo cerrado con perturbaciones

Si se desarrolla la expresión del error para el sistema de la Fig. 5.5, se obtiene:

$$e_{ss} = \lim_{s \rightarrow 0} sE = \lim_{s \rightarrow 0} s(R - C) = \lim_{s \rightarrow 0} s \left[R - \frac{G_1 G_2}{1 + G_1 G_2} R - \frac{G_2}{1 + G_1 G_2} N \right] \quad (5.13)$$

La expresión (5.13) se puede dividir en dos partes, la primera es el error del sistema ante la entrada referencia con perturbación nula y la segunda es el error del sistema ante la perturbación con referencia

nula.

$$e_{ss} = e_{ss}^R + e_{ss}^N \begin{cases} e_{ss}^R = \lim_{s \rightarrow 0} s E^R = \lim_{s \rightarrow 0} s(R - C^R) \\ e_{ss}^N = \lim_{s \rightarrow 0} s E^N = \lim_{s \rightarrow 0} s(0 - C^N) \end{cases} \quad (5.14)$$

Para calcular la magnitud del error en régimen permanente del sistema ante la entrada referencia con perturbación nula, se puede emplear de nuevo la Tabla 5.1. Sin embargo, el error ante perturbación con referencia nula, siempre hay que calcularlo analíticamente.

El error ante perturbaciones también se mide en las unidades de la variable de salida y, evidentemente, su valor dependerá de la magnitud de la perturbación. Si, por ejemplo, ante una perturbación escalón de amplitud 10 N se produce un error de -5 cm y se desea expresar este error en por unidad, se puede decir que es -0.5 cm por unidad de perturbación, es decir, un valor de unidades de salida [cm] por unidad de perturbación [N]. Nunca hay que perder de vista que se están relacionando variables con unidades distintas y que la referencia es nula, es decir, nunca se trata de un porcentaje respecto la referencia.

5.6. Ejercicios resueltos

- **Ejercicio 1:** El sistema de la Fig. 5.6 representa una inercia sin viscosidad, es decir en el espacio exterior, a la que se impone un movimiento con velocidad constante. La ley de control consiste en una actuación en par puramente proporcional al error de velocidades.

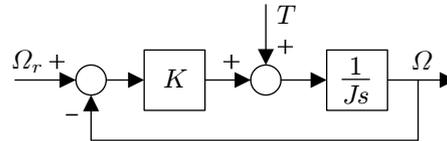


Figura 5.6: Sistema de control en velocidad de una inercia sin viscosidad

El sistema es de tipo I por lo que tendrá error nulo ante entrada referencia escalón, sin embargo, se pide calcular el error ante entradas escalones en la referencia y en la perturbación, por lo que habrá que calcularlo analíticamente:

$$e_{ss} = \lim_{s \rightarrow 0} s(\Omega_r - \Omega) = \lim_{s \rightarrow 0} s \left[\Omega_r - \frac{K}{Js + K} \Omega_r - \frac{1}{Js + K} T \right] \quad (5.15)$$

Sustituyendo las entradas por escalones de amplitudes ω_r y τ :

$$e_{ss} = \lim_{s \rightarrow 0} s \left[\frac{Js}{Js + K} \frac{\omega_r}{s} - \frac{1}{Js + K} \frac{\tau}{s} \right] = -\frac{\tau}{K} \text{ rad/s} \quad (5.16)$$

Como era de esperar por el tipo del sistema, ante la entrada referencia el error es nulo. Sin embargo, la perturbación modifica ese error, haciendo que la velocidad de la inercia sea mayor —el error es negativo— si se introduce un par extra positivo τ .

- **Ejercicio 2:** Determinar cuál debe ser el tipo del controlador $G_c(s)$ para que el sistema de la Fig. 5.7 posea error nulo ante entradas escalón en referencia y perturbación. El sistema pretende el control de la velocidad lineal de un vehículo.

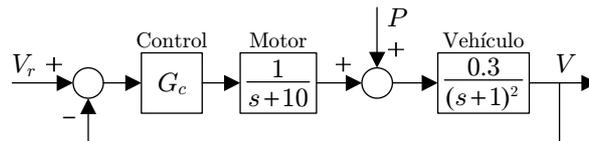


Figura 5.7: Sistema de control en velocidad de un vehículo

La expresión del error ante entradas referencia y perturbación:

$$e_{ss} = \lim_{s \rightarrow 0} s(V_r - V) = \lim_{s \rightarrow 0} s \left[V_r - \frac{0.3G_c V_r}{(s+10)(s+1)^2 + 0.3G_c} - \frac{P}{(s+10)(s+1)^2 + 0.3G_c} \right] \quad (5.17)$$

Sustituyendo las entradas por escalones unidad:

$$e_{ss} = \lim_{s \rightarrow 0} s \left[\frac{(s+10)(s+1)^2}{(s+10)(s+1)^2 + 0.3G_c} - \frac{0.3(s+10)}{(s+10)(s+1)^2 + 0.3G_c} \right] \frac{1}{s} \quad (5.18)$$

$$e_{ss} = \frac{10}{10 + 0.3G_c(0)} - \frac{3}{10 + 0.3G_c(0)} \quad (5.19)$$

El error sólo puede ser nulo si la función de transferencia del controlador se hace infinito para $s = 0$, es decir, posee al menos un polo en el origen. Por tanto, el controlador debe ser de tipo I o superior:

$$e_{ss} = \frac{10}{10 + 0.3\infty} - \frac{3}{10 + 0.3\infty} = 0 \quad (5.20)$$

Si se compara este resultado con el del ejercicio del apartado anterior, se observa cómo la posición del bloque integrador dentro del lazo es crucial para conseguir que error sea nulo ante la entrada perturbación o no. Si el integrador está antes de la perturbación el error ante escalones perturbación será nulo; si el integrador está después, en la planta a controlar, el error será finito.

- **Ejercicio 3:** Calcular la respuesta temporal ante una entrada escalón unidad del sistema de la Fig. 5.8. ¿Tiene sentido el valor el valor numérico del error entre la entrada y la salida que ofrece el límite en el dominio de Laplace?

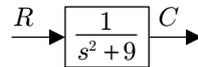


Figura 5.8: Sistema oscilatorio

La respuesta temporal es:

$$c(t) = \mathcal{L}^{-1}[C(s)] = \mathcal{L}^{-1} \left[\frac{1}{(s^2 + 9)s} \right] = \frac{1}{9} \mathcal{L}^{-1} \left[\frac{1}{s} \right] - \frac{1}{9} \mathcal{L}^{-1} \left[\frac{s}{s^2 + 9} \right] = \frac{1}{9} - \frac{1}{9} \cos 3t \quad (5.21)$$

El error en régimen permanente del sistema es:

$$e_{ss} = \lim_{s \rightarrow 0} s(R - C) = \lim_{s \rightarrow 0} s \left(R - \frac{R}{s^2 + 9} \right) = \lim_{s \rightarrow 0} \frac{s^2 + 8}{s^2 + 9} = \frac{8}{9} \text{ ¡falso!} \quad (5.22)$$

Aunque se llega a una solución numérica, el resultado es falso porque —aunque existe el límite en el dominio de Laplace— no existe el límite en el dominio temporal. Por tanto no se puede hablar propiamente de error. En la Fig. 5.9 se puede observar cómo en realidad se ha encontrado el valor en torno al cual la respuesta temporal oscila continuamente.

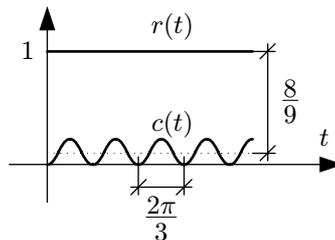


Figura 5.9: Sistema oscilatorio

5.7. Ejercicios propuestos

- **Ejercicio 1:** Sea el sistema de la Fig. 5.10, se pide:
 - Calcular el valor de K para que el sobreimpulso sea del 5%. ¿Cuál es el error en régimen permanente para dicha K ?
 - Calcular el valor de K que consigue la mitad del error del apartado anterior. ¿Cuál es el valor del sobreimpulso para esa nueva K ?

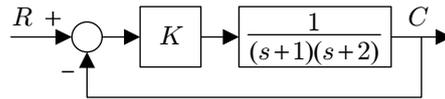


Figura 5.10: Sistema controlado de segundo orden

Soluciones:

- a) $K = 2.72$ y $e_{ss} = 42.33\%$
- b) $K = 7.44$ y $M_p = 17.26\%$

- **Ejercicio 2:** Se pretende que una pequeña central hidroeléctrica produzca una tensión de valor V_r , sea cual sea la carga que se le conecte P . Un modelo simplificado de dicho sistema aparece en la Fig. 5.11, donde lo único que puede regular el ingeniero es la apertura de la válvula que gobierna el caudal de entrada en la turbina. Calcular el error en régimen permanente del sistema de ante referencia nula y perturbación escalón de amplitud τ . Dar la solución en voltios y en rad/s.

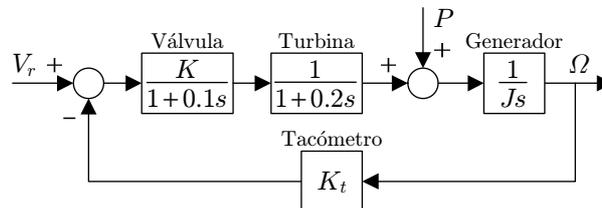


Figura 5.11: Sistema de control de la velocidad de un turbo-generador

Solución:

$$e_{ss} = -\frac{\tau}{K} \text{ V} \quad (5.23)$$

$$e_{ss} = -\frac{\tau}{KK_t} \text{ rad/s} \quad (5.24)$$

- **Ejercicio 3:** Para el sistema de la Fig. 5.12, se pide:
 - a) Error con controlador constante de valor K , perturbación nula y referencia escalón unidad.
 - b) Determinar el tipo del controlador $G_c(s)$ para que el error sea nulo ante referencia rampa unidad y N nula.
 - c) Determinar el tipo del controlador $G_c(s)$ para que el error sea nulo ante referencia escalón unidad y perturbación escalón de amplitud p .

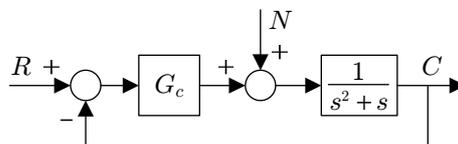


Figura 5.12: Sistema de control

Soluciones:

- a) $e_{ss} = 0$
- b) $G_c(s)$ tipo I o superior.
- c) $G_c(s)$ tipo I o superior.

- **Ejercicio 4:** En modelo de la Fig. 5.13 representa un sistema de control con doble lazo, uno exterior de posición y otro interior de velocidad, muy usado para el control de máquina-herramienta. Calcular el tipo del controlador $G_c(s)$ para que el error sea nulo ante entradas escalón.

Solución: $G_c(s)$ tipo I o superior.

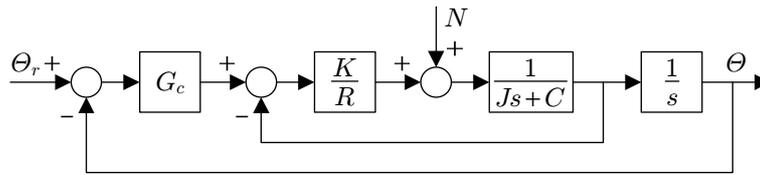


Figura 5.13: Sistema de control de doble lazo

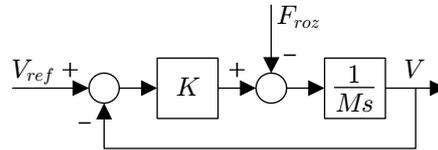


Figura 5.14: Sistema de control de velocidad de un fórmula 1

- **Ejercicio 5:** El sistema de la figura representa la velocidad V que alcanza un fórmula 1 de masa M impulsado por una fuerza controlada para alcanzar una determinada velocidad de referencia V_{ref} , en presencia de posibles perturbaciones de la fuerza de rozamiento F_{roz} .

- Determinar el error del sistema ante una entrada V_{ref} escalón de 100 m/s.
- Determinar el error del sistema ante una entrada V_{ref} rampa de pendiente 9.8 m/s². Expresar el valor del error en forma de e_v en m/s y en forma de K_v en s⁻¹.
- Determinar el error del sistema ante una entrada escalón V_{ref} de 60 m/s y una perturbación fuerza de rozamiento F_{roz} escalón de valor 50 N.
- Demostrar que si el rozamiento del sistema es viscoso, es decir, proporcional a la velocidad del tren, el sistema pasa a ser de tipo cero. Determinar en este caso, de nuevo, el error del sistema ante una entrada V_{ref} escalón de 100 m/s, llamando B a la nueva constante proporcional del rozamiento viscoso.

Soluciones:

- 0
- $e_v = \frac{9.8M}{K}$ m/s, $K_v = \frac{K}{M}$ s⁻¹
- $\frac{50}{K}$ m/s
- $\frac{100B}{B+K}$ m/s

- **Ejercicio 6:** El diagrama de bloques de un determinado sistema es:

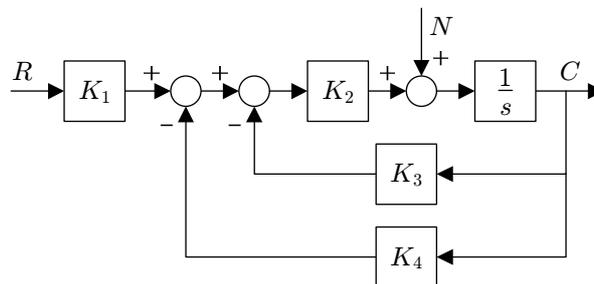


Figura 5.15: Sistema de control

- Obtener la función de transferencia que relaciona la salida $C(s)$ con la referencia $R(s)$.
- Obtener la función de transferencia que relaciona la salida $C(s)$ con la perturbación $N(s)$.
- Calcular el error ante referencia escalón unidad y perturbación nula. ¿Se puede anular el error con las ganancias?
- Calcular el error ante perturbación escalón unidad y referencia nula. ¿Se puede anular el error con las ganancias?

Soluciones:

- a) $\frac{C(s)}{R(s)} = \frac{K_1 K_2}{s + (K_3 + K_4) K_2}$
 b) $\frac{C(s)}{N(s)} = \frac{1}{s + (K_3 + K_4) K_2}$
 c) $e_{ss}^R = \frac{K_3 + K_4 - K_1}{K_3 + K_4}$ y se puede anular el error ante entrada referencia haciendo $K_1 = K_3 + K_4$
 d) $e_{ss}^N = \frac{-1}{(K_3 + K_4) K_2}$ y no se puede anular el error ante entrada perturbación.

- **Ejercicio 7:** La inercia J del sistema de control de la Fig. 5.16 vale $25 \text{ kg}\cdot\text{m}^2$. Las ganancias del sistema las puede ajustar el ingeniero. Se pide:

- a) Dar un valor para K_3 de forma que el error del sistema sea de 1 cm para una entrada $x_{ref}(t) = t$, es decir, una rampa de pendiente 1 m/s.
 b) Con la ganancia del apartado anterior, dar un valor para el producto $K_1 K_2$ para que el sobreimpulso máximo sea del orden del 10%. ¿Cuál es el valor del tiempo de establecimiento del sistema?

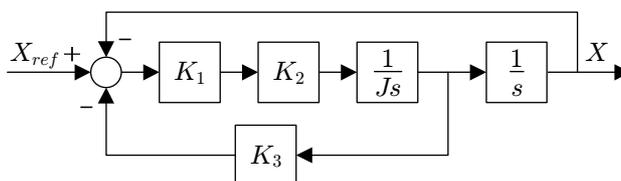


Figura 5.16: Sistema de control de doble lazo

Soluciones:

- a) $K_3 = 0.01$
 b) $K_1 K_2 = 360000$ y $t_s = 0.055 \text{ s}$

- **Ejercicio 8:** Sea el sistema de la figura:

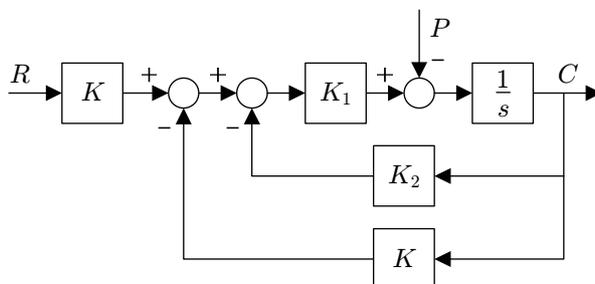


Figura 5.17: Sistema de control

- a) Determinar la función de transferencia que relaciona la salida con la entrada de referencia.
 b) Calcular el valor del error cuando la entrada perturbación es un escalón unidad.
 c) Dibujar aproximadamente la respuesta temporal del sistema $c(t)$, si $K = K_2 = 1$, $1 < K_1 < 10$ y la entrada referencia es un escalón unidad. Elegir K_1 para que la respuesta temporal sea lo más rápida posible.

- **Ejercicio 9:** El sistema de calefacción de una casa está regido por un conjunto de ecuaciones diferenciales que se explican a continuación. Calor generado por los radiadores es:

$$q_r(t) = K_1 x(t - t_r), \quad (5.25)$$

donde $x(t)$ es la posición de la válvula de paso de agua caliente y t_r el retraso debido al transporte del agua. el calor perdido por las paredes es:

$$q_p(t) = K_2 [T_i(t) - T_e(t)], \quad (5.26)$$

donde T_i y T_e son la temperatura interior y exterior respectivamente. La temperatura interior es:

$$T_i(t) = K_3 \int_0^t [q_r(\tau) - q_p(\tau)] d\tau, \quad (5.27)$$

donde K_3 es la constante que tiene en cuenta la masa del edificio y su calor específico. Y la ley de control del termostato es:

$$x(t) = K_c [T_{ref}(t) - T_i(t)], \quad (5.28)$$

donde T_{ref} es la temperatura de referencia deseada y K_c es la ganancia del controlador del termostato.

- a) Dibujar el diagrama de bloques de esta planta considerando que la entrada es la temperatura de referencia y la salida es la temperatura interior. La temperatura se tratará dentro del diagrama como si fuera una perturbación.
- b) Suponiendo que la constante del controlador K_c vale 5 que el resto de las constantes son la unidad, que el retraso de transporte es nulo, que la temperatura exterior es de 0°C , calcular la función de transferencia que relaciona la temperatura deseada con la real.
- c) Determinar el error en régimen permanente, es decir, la diferencia entre la temperatura deseada y la real, en caso de que la primera sea de 20°C suponiendo las mismas condiciones que en el apartado anterior.

Soluciones:

- La función de transferencia en lazo cerrado es $G_{lc}(s) = \frac{5}{s+6}$.
- El error es 3.33°C .

Capítulo 6

Estabilidad

Para que un sistema de control sea útil, lo primero que debe cumplir es que sea estable. Si el sistema es inestable no existe régimen permanente aunque numéricamente se puedan encontrar los valores de los límites en el dominio de Laplace que se plantearon en el capítulo anterior. Por tanto, asegurar la estabilidad del sistema debe ser un paso previo al cálculo numérico de los errores en régimen permanente, y por este motivo en algunos libros de texto este capítulo preceda al anterior.

6.1. Definición de estabilidad

Existen varias definiciones para la estabilidad, que no tienen necesariamente por qué coincidir. Así por ejemplo, se dice que un sistema es estable cuando:

- La respuesta del sistema a un impulso tiende a cero cuando el tiempo tiende a infinito.
- Ante una entrada acotada le corresponde una salida también acotada.

A la segunda situación, se le conoce como estabilidad BIBO (de *Bounded-Input Bounded-Output*). Para sistemas continuos y lineales que se pueden definir por medio de funciones de transferencia racionales propias, la estabilidad BIBO se cumple si y sólo si todos los polos del sistema tienen parte real negativa. Esto equivale a decir que todos los polos deben estar localizados en el semiplano izquierdo de S .

No es de extrañar que la estabilidad del sistema dependa de la posición de los polos del sistema, ya que se ha estudiado en capítulos anteriores cómo la localización de los polos de la función de transferencia resulta crucial en el régimen transitorio. Incluso es posible intuir el resultado anterior atendiendo a la tabla de transformadas de Laplace elementales.

Los polos del sistema son las raíces de la ecuación que resulta de igualar a cero el denominador de la función de transferencia del sistema. Esa ecuación se conoce con el nombre de ecuación característica del sistema. Por tanto, las raíces de la ecuación característica ofrecen información no sólo del transitorio del sistema, sino también de su estabilidad.

6.2. Criterio de Routh-Hurwitz

Conocer las raíces de la ecuación característica, para comprobar si las partes reales de todas ellas son negativas y asegurar así que el sistema es estable, es difícil cuando el orden del sistema es superior a dos. El problema se incrementa si además, los coeficientes de la ecuación no son valores numéricos, sino que dependen de algún parámetro variable.

El criterio de Routh-Hurwitz¹ aplicado a la ecuación característica de un sistema permite conocer si es estable o no, sin necesidad de calcular las raíces de dicha ecuación característica.

Sea la función de transferencia,

$$\frac{C(s)}{R(s)} = \frac{p(s)}{a_n s^n + a_{n-1} s^{n-1} + \dots + a_1 s + a_0}, \quad (6.1)$$

¹Edward John Routh (1831-1907) ganó en 1877 el premio Adams por su método de analizar la estabilidad de sistemas dinámicos de cualquier orden, siguiendo estudios de James Clerk Maxwell. El matemático alemán Adolf Hurwitz (1859-1919) propuso su método para asegurar la estabilidad de turbinas de vapor en 1895, siguiendo estudios de Aurel Boreoslav Stodola. En 1911 se demostró que ambos criterios de estabilidad eran equivalentes.

su ecuación característica posee $n + 1$ coeficientes a_i reales:

$$a_n s^n + a_{n-1} s^{n-1} + \dots + a_1 s + a_0 = 0, \quad (6.2)$$

Primero se comprueba que todos los coeficientes a_i sean positivos. Si hubiese algún coeficiente nulo o negativo, el sistema no sería estable. Si se cumple la condición anterior, que se conoce como condición de Cardano-Viète, el sistema puede ser estable o no. Para comprobar si es estable, se disponen los coeficientes a_i de forma que sigan el patrón impuesto por la siguiente tabla:

$$\begin{array}{l|lllll} s^n & a_n & a_{n-2} & a_{n-4} & a_{n-6} & \dots \\ s^{n-1} & a_{n-1} & a_{n-3} & a_{n-5} & a_{n-7} & \dots \\ s^{n-2} & b_1 & b_2 & b_3 & \dots & \\ s^{n-3} & c_1 & c_2 & \dots & & \\ \dots & \dots & \dots & & & \\ s^0 & d & & & & \end{array} \quad (6.3)$$

Donde los coeficientes a_i se distribuyen en las dos primeras columnas. Los coeficientes de las sucesivas filas se calculan empleando los coeficientes de las dos columnas inmediatamente superiores. Así los coeficientes b_i se calculan como sigue:

$$b_1 = \frac{a_{n-1}a_{n-2} - a_n a_{n-3}}{a_{n-1}} \quad (6.4)$$

$$b_2 = \frac{a_{n-1}a_{n-4} - a_n a_{n-5}}{a_{n-1}} \quad (6.5)$$

$$b_3 = \frac{a_{n-1}a_{n-6} - a_n a_{n-7}}{a_{n-1}} \quad (6.6)$$

Análogamente, los coeficientes c_i se calculan:

$$c_1 = \frac{b_1 a_{n-3} - a_{n-1} b_2}{b_1} \quad (6.7)$$

$$c_2 = \frac{b_1 a_{n-5} - a_{n-1} b_3}{b_1} \quad (6.8)$$

A partir de un momento, los coeficientes de las filas valen sucesivamente cero. Estos ceros a veces son necesarios para calcular coeficientes posteriores. Se puede observar que el cálculo de los coeficientes sigue un patrón que se puede memorizar. El denominador siempre es el primer coeficiente de la fila inmediatamente superior. El numerador depende de los coeficientes de las dos filas inmediatamente superiores y es la diferencia de dos productos cuyos términos poseen una posición cruzada. Para sucesivos coeficientes, los dos primeros términos siempre se emplean en el producto cruzado, mientras que los otros dos van avanzando.

El proceso acaba cuando se calcula la fila de coeficientes en s^0 , que sólo posee un coeficiente no nulo, d en la expresión (6.3). El criterio afirma que el sistema es estable si y sólo si todos los coeficientes de la primera columna de Routh-Hurwitz son positivos. Es, por tanto, una condición necesaria y suficiente. La primera columna la forman los primeros coeficientes de todas las filas. Aunque el criterio sólo se fije en los primeros coeficientes, las filas hay que completarlas enteras, porque todos los coeficientes son necesarios para calcular los inferiores.

Cuando no se cumple el criterio de Routh-Hurwitz, es posible conocer el número de polos del sistema que están en el semiplano de parte real positiva. Existen tantos polos con parte real positiva como cambios de signo aparecen a la largo de la primera columna de Routh-Hurwitz.

Es importante recalcar que criterio de Routh-Hurwitz informa sobre la estabilidad absoluta, es decir, se limita a mostrar si el sistema es estable o no, sin indicar el grado de estabilidad o inestabilidad, lo próximo o lo alejado que se está de volverse inestable o estable.

6.2.1. Estabilidad de los sistemas de segundo orden

En el caso de sistemas de segundo orden, discernir la estabilidad del sistema es especialmente sencillo. Sea la ecuación característica general de segundo orden:

$$a_2 s^2 + a_1 s + a_0 = 0 \quad (6.9)$$

Si los tres coeficientes a_1 , a_2 y a_3 son positivos no nulos, se cumple la condición de Cardano-Viète y el sistema puede ser estable. Se construye entonces la tabla de Routh-Hurwitz:

$$\begin{array}{c|cc} s^2 & a_2 & a_0 \\ s^1 & a_1 & \\ s^0 & & a_0 \end{array} \quad (6.10)$$

En este caso concreto, los coeficientes de la primera columna coinciden exactamente con los coeficientes del polinomio de la ecuación característica. Por tanto, basta con observar si los tres coeficientes de la ecuación característica son positivos para que se cumpla tanto la condición de Cardano-Viète como el criterio de Routh-Hurwitz y se pueda afirmar que el sistema es estable.

6.2.2. Estabilidad de los sistemas de tercer orden

En el caso de sistemas de tercer orden, también resulta relativamente sencillo predecir la estabilidad o no del mismo. Sea la ecuación característica general de tercer orden:

$$a_3s^3 + a_2s^2 + a_1s + a_0 = 0 \quad (6.11)$$

Si los coeficientes son positivos no nulos, el sistema puede ser estable. Se construye entonces la tabla de Routh-Hurwitz:

$$\begin{array}{c|ccc} s^3 & a_3 & a_1 & \\ s^2 & a_2 & a_0 & \\ s^1 & \frac{a_2a_1 - a_3a_0}{a_2} & & \\ s^0 & a_0 & & \end{array} \quad (6.12)$$

Para que todos los coeficientes de la primera columna sean positivos, la única condición que se debe cumplir es:

$$a_2a_1 > a_3a_0 \quad (6.13)$$

6.2.3. Ejemplo numérico de sistema de cuarto orden

El número de condiciones que se deben cumplir para casos genéricos de sistemas de orden elevado es cada vez mayor. En este apartado se resuelve el caso concreto de un sistema de cuarto orden con la siguiente ecuación característica:

$$s^4 + 2s^3 + 3s^2 + 4s + 5 = 0 \quad (6.14)$$

Todos los coeficientes son positivos no nulos, por lo que se construye la tabla de Routh-Hurwitz:

$$\begin{array}{c|cccc} s^4 & 1 & 3 & 5 & \\ s^3 & 2 & 4 & 0 & \\ s^2 & \frac{2 \cdot 3 - 1 \cdot 4}{2} = 1 & \frac{2 \cdot 5 - 1 \cdot 0}{2} = 5 & & \\ s^1 & \frac{1 \cdot 4 - 2 \cdot 5}{1} = -6 & 0 & & \\ s^0 & \frac{-6 \cdot 5 - 1 \cdot 0}{-6} = 5 & & & \end{array} \quad (6.15)$$

Todos los coeficientes de la primera columna son positivos menos uno que es negativo, por tanto el sistema es inestable. Asimismo, existen dos cambios de signo en la columna, por tanto existen dos raíces con parte real positiva. Las raíces de la ecuación característica, es decir, los polos del sistema son:

$$\begin{cases} p_1 = 0.29 + 1.41j \\ p_2 = 0.29 - 1.41j \\ p_3 = 1.29 + 0.86j \\ p_4 = 1.29 - 0.86j \end{cases} \quad (6.16)$$

Efectivamente, como predice el criterio, existen dos polos de parte real positiva.

6.3. Casos especiales del criterio de Routh-Hurwitz

La confección de la tabla de Routh-Hurwitz puede ser imposible en varios casos, por ejemplo cuando alguno de los denominadores de los coeficientes se hace nulo. En los siguientes apartados se presenta el modo de actuar para el caso de las dos situaciones especiales más frecuentes.

6.3.1. Se anula el primer coeficiente de una fila

Si existe un cero en la primera posición de una fila, todos los coeficientes de la fila inmediatamente inferior se hacen infinitos. Para evitar esta situación, se puede sustituir el coeficiente nulo por una constante positiva ϵ muy próxima a cero. Esta constante se arrastra en el cálculo de los siguientes coeficientes y permite estudiar el signo de todos ellos.

Como ejemplo se puede observar qué ocurre cuando la ecuación característica es:

$$s^3 - 3s + 2 = (s - 1)^2(s + 2) = 0 \quad (6.17)$$

No cumple la condición de Cardano-Viète, por la que ya se puede afirmar que el sistema es inestable. Si se construye la tabla de Routh-Hurwitz:

$$\begin{array}{l|ll} s^3 & 1 & -3 \\ s^2 & \epsilon & 2 \\ s^1 & -3 - \frac{2}{\epsilon} & \\ s^0 & 2 & \end{array} \quad (6.18)$$

Si ϵ toma un valor positivo muy pequeño, el siguiente coeficiente de la columna es negativo muy grande, por lo que el sistema es inestable. Existen dos cambios de signo, de la segunda a la tercera fila y de la tercera a la cuarta, por tanto existen dos polos de parte real positiva. En efecto, $s = 1$ es un polo doble con parte real positiva, como ya se mostró en la definición de la ecuación característica.

6.3.2. Se anula toda una fila

Cuando se anula toda una fila de la tabla de Routh-Hurwitz significa que existen raíces simétricas respecto un eje y situadas encima del otro. Es decir, serán raíces imaginarias puras conjugadas o reales de signo contrario. También pueden existir raíces en el origen. Estas raíces peculiares, se obtienen resolviendo la ecuación que se construye con la fila superior a la nula, es decir, el último renglón no nulo.

Como ejemplo, se puede observar cómo se obtienen esas raíces peculiares en la siguiente ecuación característica:

$$s^3 + 2s^2 + s + 2 = 0 \quad (6.19)$$

Su tabla de Routh-Hurwitz es:

$$\begin{array}{l|ll} s^3 & 1 & 1 \\ s^2 & 2 & 2 \\ s^1 & 0 & \\ s^0 & \text{indeterminado} & \end{array} \quad (6.20)$$

Toda la fila en s^1 es nula. Con los coeficientes de la fila inmediatamente superior no nula, la fila en s^2 , se construye la ecuación (6.21), llamada ecuación auxiliar.

$$2s^2 + 2 = 0 \quad (6.21)$$

Las raíces de la ecuación auxiliar, en este caso $\pm j$, son también raíces de la ecuación característica.

Para poder seguir construyendo la tabla de Routh-Hurwitz, se realiza de la forma habitual una vez que se ha sustituido la fila de ceros por los coeficientes que resultan de derivar el polinomio de la ecuación auxiliar respecto de s . Como ejemplo se puede observar qué ocurre cuando la ecuación característica es:

$$s^5 + 2s^4 + 24s^3 + 48s^2 - 25s - 50 = 0 \quad (6.22)$$

Su tabla de Routh-Hurwitz es:

$$\begin{array}{l|lll} s^5 & 1 & 24 & -25 \\ s^4 & 2 & 48 & -50 \\ s^3 & 0 & 0 & \\ \text{nueva } s^3 & 8 & 96 & \\ s^2 & 24 & -50 & \\ s^1 & 112.7 & & \\ s^0 & -50 & & \end{array} \quad (6.23)$$

Donde la nueva fila en s^3 se obtiene derivando respecto de s el polinomio de la ecuación auxiliar:

$$\frac{d}{ds}(2s^4 + 48s^2 - 50) = 8s^3 + 96s \quad (6.24)$$

La anterior fila en s^3 no se tiene en cuenta para la construcción de la tabla. En este caso, como era de esperar, por Cardano-Viète, el sistema es inestable y, por tener un único cambio de signo, le corresponde una sola raíz con parte real positiva.

6.4. Ejercicios resueltos

- **Ejercicio 1:** El criterio de Routh-Hurwitz permite estudiar la estabilidad de un sistema de control respecto uno o dos parámetros que puedan intervenir en el sistema. Esto permite establecer las condiciones que deben cumplir dichos parámetros para que el sistema sea estable. En el sistema de la Fig. 6.1, se desea determinar el rango de valores de K para que el sistema sea estable.

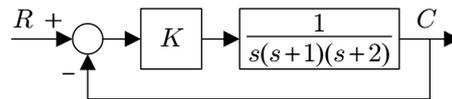


Figura 6.1: Sistema de control

La función de transferencia del sistema es:

$$\frac{C(s)}{R(s)} = \frac{K}{s(s+1)(s+2) + K} = \frac{K}{s^3 + 3s^2 + 2s + K} \quad (6.25)$$

Su tabla de Routh-Hurwitz es:

$$\begin{array}{c|cc} s^3 & 1 & 2 \\ s^2 & 3 & K \\ s^1 & \frac{6-K}{3} & \\ s^0 & K & \end{array} \quad (6.26)$$

Para que el sistema sea estable, K debe pertenecer al intervalo $(0,6)$.

- **Ejercicio 2:** Estudiar la estabilidad del sistema en función del parámetro K si la ecuación característica del mismo es:

$$s^5 + 15s^4 + 85s^3 + 225s^2 + 274s + 120 + K = 0 \quad (6.27)$$

Su tabla de Routh-Hurwitz es:

$$\begin{array}{c|ccc} s^5 & 1 & 85 & 274 \\ s^4 & 15 & 225 & 120 + K \\ s^3 & 70 & \frac{3990-K}{15} & \\ s^2 & \frac{11760+K}{70} & 120 + K & \\ s^1 & a & & \\ s^0 & 120 + K & & \end{array} \quad (6.28)$$

$$a = \frac{36288 - \frac{387K}{5} - \frac{K^2}{1050}}{\frac{11760+K}{70}} \quad (6.29)$$

Establecer las condiciones para que todos los coeficientes de la primera columna de la tabla sean positivos es un razonamiento puramente algebraico. El primer coeficiente de la fila en s^2 es positivo si se cumple que:

$$K > -11760 \quad (6.30)$$

El primer coeficiente de la fila en s^1 es positivo en los siguientes intervalos de K :

$$K \in (-\infty, -81736.16) \cup (-11760, 466.16) \quad (6.31)$$

El primer coeficiente de la fila en s^0 es positivo si se cumple que:

$$K > -120 \quad (6.32)$$

Combinando las tres condiciones para que todos ellos a la vez sean positivos, se tiene que el sistema sea estable si K pertenece al intervalo:

$$-120 < K < 466.16 \quad (6.33)$$

6.5. Ejercicios propuestos

- **Ejercicio 1:** Sea el sistema de la Fig. 6.2, de dos entradas y dos salidas:

- Determinar la función de transferencia $\frac{X}{X_{ref}}$ con $F_{ref} = 0$
- Determinar la función de transferencia $\frac{X}{F_{ref}}$ con $X_{ref} = 0$
- Determinar la función de transferencia $\frac{F}{F_{ref}}$ con $X_{ref} = 0$
- Determinar la función de transferencia $\frac{F}{X_{ref}}$ con $F_{ref} = 0$
- Identificar la ecuación característica del sistema.
- ¿Qué condición se debe cumplir para que el sistema sea estable?

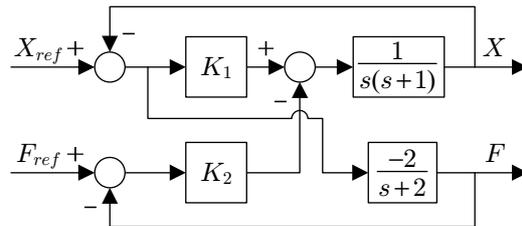


Figura 6.2: Sistema de dos entradas y dos salidas

Soluciones:

- $\frac{X}{X_{ref}} = \frac{K_1(s+2) - 2K_2}{s(s+1)(s+2) + K_1(s+2) - 2K_2}$
- $\frac{X}{F_{ref}} = \frac{-K_2(s+2)}{s(s+1)(s+2) + K_1(s+2) - 2K_2}$
- $\frac{F}{F_{ref}} = \frac{-2K_2}{s(s+1)(s+2) + K_1(s+2) - 2K_2}$
- $\frac{F}{X_{ref}} = \frac{-2s(s+1)}{s(s+1)(s+2) + K_1(s+2) - 2K_2}$
- $s(s+1)(s+2) + K_1(s+2) - 2K_2 = 0$
- $K_1 > K_2$

Capítulo 7

Lugar de las raíces

7.1. Introducción

En labores de diseño es interesante conocer cómo varía la ubicación de los polos del sistema en función de un parámetro que el ingeniero pueda modificar a su arbitrio. Con esta información se puede saber qué especificaciones de régimen transitorio se pueden imponer en la respuesta del sistema. Habitualmente el parámetro de diseño es una ganancia proporcional dentro del lazo de control. El cálculo de los polos es sencillo con una función de transferencia de segundo orden:

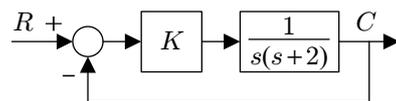


Figura 7.1: Sistema controlado de segundo orden

$$\frac{C(s)}{R(s)} = \frac{K}{s^2 + 2s + K} \quad (7.1)$$

Los polos del sistema son las raíces del denominador, es decir:

$$p_{1,2} = -1 \pm \sqrt{1 - K} \quad (7.2)$$

La Fig. 7.2 muestra el lugar que van ocupando estos polos, variando K entre cero e infinito. El sistema es siempre estable para cualquier valor del parámetro K . Además, su respuesta se puede imponer subamortiguada, con cualquier valor de amortiguamiento.

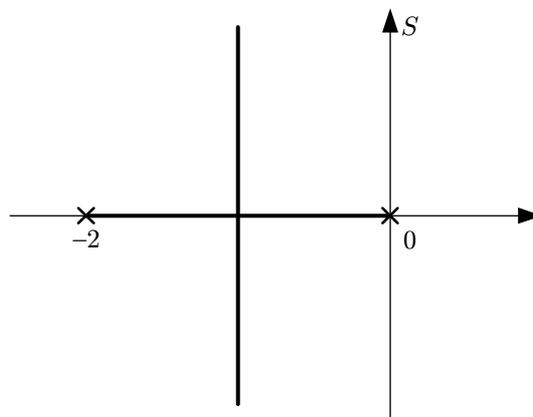


Figura 7.2: Lugar de las raíces del sistema de segundo orden

El estudio analítico de sistemas de tercer orden o superior es más complicado. El método del lugar de las raíces ideado en 1948 por el norteamericano Walter Richard Evans (1920-1999), permite obtener de forma gráfica la localización de los polos del sistema sin tener que realizar su cálculo numérico.

7.2. Generalidades del método

Se llama lugar de las raíces de un sistema al lugar geométrico de los puntos del plano S que satisfacen su ecuación característica cuando se va modificando un determinado parámetro. Se trata por tanto de los polos del sistema en función de dicho parámetro.

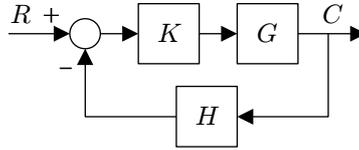


Figura 7.3: Sistema de control en lazo cerrado

Si el parámetro es una ganancia proporcional K dentro de un lazo de control, con realimentación negativa no unitaria, Fig. 7.3, el lugar de las raíces es la ubicación de los polos en lazo cerrado del sistema. La función de transferencia en lazo cerrado de este sistema es:

$$G_{lc} = \frac{C(s)}{R(s)} = \frac{KG(s)}{1 + KG(s)H(s)} = \frac{KG_d(s)}{1 + KG_{la}(s)} \quad (7.3)$$

Los puntos q del lugar de las raíces anulan el denominador de la función de transferencia:

$$1 + KG_{la}(q) = 0 \quad (7.4)$$

Dado que $G_{la}(q)$ es un número complejo, la ecuación característica (7.4) se puede descomponer en su módulo y argumento:

$$KG_{la}(q) = -1 \Rightarrow \begin{cases} |KG_{la}(q)| = 1 \\ \angle G_{la}(q) = \pm 180^\circ(2k + 1), \text{ con } k \in \mathbb{N} \end{cases} \quad (7.5)$$

Se observa que la condición del argumento es independiente del parámetro K variable. Por tanto, los puntos q del plano complejo S que cumplen la condición del argumento son aquellos que pertenecen al lugar de las raíces. En cambio, la condición del módulo sirve para calcular el valor del parámetro K necesario para que un determinado punto q sea polo del sistema en lazo cerrado.

Si se escribe la función de transferencia en lazo abierto factorizada en sus n polos y m ceros, la ecuación característica del sistema es:

$$1 + K \frac{K_{la}(q - z_1)(q - z_2)\dots(q - z_m)}{(q - p_1)(q - p_2)\dots(q - p_n)} = 0 \quad (7.6)$$

Las nuevas condiciones de módulo y argumento son:

$$\begin{cases} |KK_{la}| = \frac{\prod_{j=1}^n |q - p_j|}{\prod_{i=1}^m |q - z_i|} \\ \sum_{i=1}^m \angle(q - z_i) - \sum_{j=1}^n \angle(q - p_j) = \pm 180^\circ(2k + 1), \text{ con } k \in \mathbb{N} \end{cases} \quad (7.7)$$

Se hace notorio que para trazar el lugar de las raíces de un sistema es necesario conocer la localización de los polos y los ceros de la función de transferencia en lazo abierto.

7.3. Método para dibujar el lugar de las raíces

Para dibujar el lugar de las raíces no es necesario evaluar la condición del argumento en todos los puntos del plano S . Se propone seguir los siguientes pasos:

7.3.1. Polos y ceros en lazo abierto

Se calculan los polos y los ceros de la función de transferencia en lazo abierto y se sitúan en el plano S . El lugar de las raíces es un conjunto de líneas simétricas respecto al eje real que nacen en los polos en lazo abierto y acaban en los ceros en lazo abierto (si el número de polos y ceros coincide).

La característica de simetría se debe a que si un número complejo es raíz de la ecuación característica, también lo es su conjugado. En cuando a los puntos de salida, son los polos del sistema cuando $K = 0$:

$$1 + KG_{la}(s) = 1 + K \frac{N(s)}{D(s)} = 0 \quad (7.8)$$

$$D(s) + KN(s) = 0 \quad (7.9)$$

$$D(s) = 0, \text{ si } K = 0 \quad (7.10)$$

Se observa que los polos del sistema en lazo cerrado cuando $K = 0$ coinciden con los polos del sistema en lazo abierto. Los puntos de llegada del lugar de las raíces se obtienen cuando $K = \infty$ y coinciden con los ceros en lazo abierto:

$$\frac{D(s)}{K} + N(s) = 0 \quad (7.11)$$

$$N(s) = 0, \text{ si } K = \infty \quad (7.12)$$

Por tanto, si hay tantos polos como ceros, el lugar de las raíces describe un conjunto de curvas cerradas, que parten de los polos y acaban en los ceros. Si no hay suficientes ceros, algunas ramas no describen curvas cerradas y van hacia el infinito.

7.3.2. Asíntotas

Si el número de polos no coincide con el número de ceros, aparecen en el lugar de las raíces tantas asíntotas como diferencia entre polos y ceros. Es decir, el número de asíntotas es $n - m$. Todas las asíntotas se cortan en un único punto del eje real, que se calcula de la forma:

$$\sigma_a = \frac{\sum_{j=1}^n p_j - \sum_{i=1}^m z_i}{n - m} \quad (7.13)$$

Los ángulos que forman las asíntotas respecto al eje real son:

$$\theta_a = \frac{180^\circ}{n - m} (2k + 1), \text{ con } k = 0, 1, \dots, n - m - 1 \quad (7.14)$$

7.3.3. Puntos del eje real que pertenecen al lugar de las raíces

Los intervalos de la parte del eje real que pertenecen al lugar de las raíces, son aquellos que dejan a su derecha un número impar de ceros y polos. Esto se puede justificar aplicando la condición angular.

7.3.4. Puntos de ruptura

En un punto del lugar de las raíces se pueden juntar varios polos del sistema. Se juntan siguiendo una dirección y se separan siguiendo otra diferente. Son los llamados puntos de ruptura y se buscan entre las raíces de la ecuación:

$$\frac{dG_{la}(s)}{ds} = 0 \quad (7.15)$$

Sólo las raíces de la ecuación (7.15) que pertenezcan al lugar de las raíces son puntos de ruptura del mismo. Como lo habitual es que los puntos de ruptura aparezcan en el eje real, es suficiente comprobar si las raíces se encuentran dentro de los tramos del eje real que pertenecen al lugar de las raíces.

7.3.5. Puntos de corte con el eje imaginario

Para encontrar los puntos del lugar de las raíces que cortan el eje imaginario se emplea el criterio de Routh-Hurwitz. Primero se encuentra la ganancia crítica del sistema, si existe. Posteriormente se sustituye este valor en la tabla del método de Routh-Hurwitz. Se anulará una de las filas de la tabla y, con la fila inmediatamente superior, se construye el polinomio auxiliar en s , cuyas raíces son los puntos de corte con el eje imaginario.

7.3.6. Ángulos de salida y llegada

Es interesante conocer el ángulo con que salen las ramas del lugar de las raíces desde los polos en lazo abierto y el ángulo con que llegan a los ceros. Para ello se aplica la condición del argumento en un punto q muy próximo al polo o cero objeto de estudio.

$$\sum_{i=1}^m \widehat{z}_i q - \sum_{j=1}^n \widehat{p}_j q = \pm 180^\circ (2k + 1), \text{ con } k \in \mathbb{N} \quad (7.16)$$

Esta condición se puede leer como: “la suma de los ángulos vistos desde los ceros menos la suma de los ángulos vistos desde los polos es igual a 180° o un ángulo equivalente”.

7.4. Cálculo de la ganancia

Con los pasos del apartado anterior es posible dibujar fácilmente el lugar de las raíces del sistema. Con dicha representación se eligen los polos en lazo cerrado más adecuados para satisfacer una determinada condición de diseño. En este caso el requerimiento del diseño suele ser un determinado comportamiento transitorio de la salida controlada del sistema. La ganancia que hace que los polos del sistema sean los puntos elegidos del lugar de las raíces, se calcula con la condición del módulo:

$$KK_{la} = \frac{\prod_{j=1}^n \overline{p}_j q}{\prod_{i=1}^m \overline{z}_i q} \quad (7.17)$$

Esta condición se puede leer como: “la ganancia total del sistema en lazo abierto, la que introduce el controlador proporcional y la propia del sistema, es igual al producto de las distancias desde los polos en lazo abierto hasta el polo objetivo dividido por el producto de las distancias desde los ceros”. En el caso de que la función de transferencia en lazo abierto no tenga ceros, la condición del módulo no se divide por nada:

$$KK_{la} = \prod_{j=1}^n \overline{p}_j q \quad (7.18)$$

7.5. Ejemplos de lugares de las raíces

En los siguientes apartados se muestran varios ejemplos de lugares de las raíces.

7.5.1. Sistema de tercer orden

Dado el sistema de la Fig. 7.4, dibujar su lugar de las raíces y obtener el valor de la ganancia K para que la salida controlada posea un amortiguamiento de 0.5.

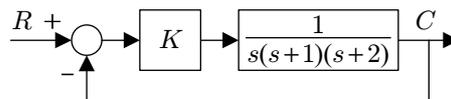


Figura 7.4: Sistema controlado de tercer orden

El tramo del eje real $(-\infty, -2] \cup [-1, 0]$ pertenece al lugar de las raíces. Existen tres asíntotas de ángulos 60° , 180° y 300° . El punto de cruce de las asíntotas se encuentra en:

$$\sigma_a = \frac{-1 - 2}{3} = -1 \quad (7.19)$$

Los puntos de ruptura de buscan entre las raíces de la ecuación:

$$3s^2 + 6s + 2 = 0 \rightarrow \begin{cases} r_1 = 0.4226 \\ r_2 = -1.5774 \end{cases} \quad (7.20)$$

El único punto de ruptura es r_1 porque r_2 no pertenece al lugar de las raíces. Para calcular los puntos de corte con el eje imaginario, se aplica el criterio de Routh-Hurwitz a la ecuación característica:

$$\begin{array}{c|cc} s^3 & 1 & 2 \\ s^2 & 3 & K \\ s^1 & \frac{6-K}{3} & \\ s^0 & K & \end{array} \quad (7.21)$$

En el rango $0 < K < 6$ el sistema es estable. Para la ganancia crítica de valor $K_{CR} = 6$ se anula la fila en s^1 y con la fila en s^2 se construye la ecuación (7.22), cuyas raíces son los puntos de corte con el eje imaginario.

$$3s^2 + 6 = 0 \rightarrow s = \pm\sqrt{2}j \quad (7.22)$$

Con esta información se puede dibujar el lugar de las raíces, Fig. 7.5.

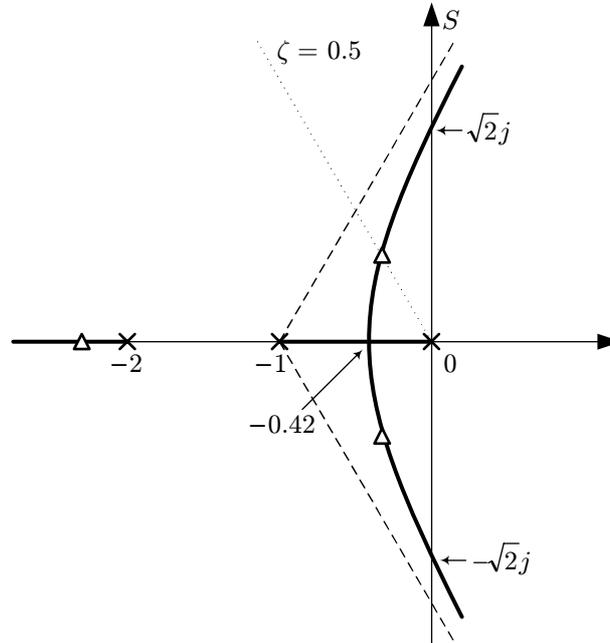


Figura 7.5: Lugar de las raíces del sistema de tercer orden

Para que la salida del sistema posea un amortiguamiento de 0.5 los polos deben formar un ángulo de 60° con el origen de coordenadas. El único punto del lugar de las raíces que cumple esta condición se ha representado en la Fig. 7.5 con un triángulo. Se aplica la condición del módulo en ese polo y resulta:

$$K = 0.68 \cdot 0.87 \cdot 1.75 = 1.03 \quad (7.23)$$

Como es una resolución gráfica, no es aconsejable dar el resultado con más de dos decimales. El polo objetivo señalado como un triángulo no es el único polo en lazo cerrado. Es un sistema de tercer orden, por lo que hay tres polos en lazo cerrado. Los tres polos se pueden calcular asignando un valor $K = 1.03$ a la ganancia en la ecuación característica del sistema:

$$s^3 + 3s^2 + 2s + 1.03 = 0 \rightarrow \begin{cases} q_1 = -2.3317 \\ q_{2,3} = -0.3341 \pm 0.5745j \end{cases} \quad (7.24)$$

El polo objetivo q_2 forma 59.82° con el origen, es decir, posee un amortiguamiento casi exactamente de 0.5. Se comprueba que el método gráfico es lo suficientemente robusto como para ser válido a pesar de la variabilidad que puede suponer la representación gráfica a mano alzada, o la medición aproximada de las distancias.

También se observa que el polo objetivo complejo-conjugado es dominante frente al polo real dentro del sistema es de tercer orden. Por tanto, se puede afirmar que la respuesta transitoria del sistema posee el amortiguamiento deseado. Para asegurar que se cumplen las especificaciones transitorias deseadas, siempre es conveniente comprobar que los polos objetivos son los dominantes en el sistema controlado.

7.5.2. Sistema de segundo orden con un cero

Dado el sistema de la Fig. 7.6, dibujar su lugar de las raíces.

Los polos en lazo abierto están en:

$$s^2 + 2s + 3 = 0 \rightarrow p_{1,2} = -1 \pm \sqrt{2}j \quad (7.25)$$

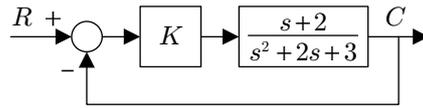


Figura 7.6: Sistema controlado de segundo orden con cero

Hay un cero en lazo abierto en -2 . El tramo del eje real $(-\infty, -2]$ pertenece al lugar de las raíces. Existe una única asíntota que forma 180° con la horizontal. Los puntos de ruptura se buscan entre las raíces de la ecuación:

$$s^2 + 4s + 1 = 0 \rightarrow \begin{cases} r_1 = -0.2679 \\ r_2 = -3.7321 \end{cases} \quad (7.26)$$

El único punto de ruptura es -3.7321 porque el otro no pertenece al lugar de las raíces. Para calcular los puntos de corte con el eje imaginario, se observa la ecuación característica del sistema:

$$s^2 + (2 + K)s + (3 + 2K) = 0 \quad (7.27)$$

El sistema es estable para cualquier valor de K , por lo que no existen cortes con el eje imaginario. Los puntos de salida de los polos imaginarios se calculan aplicando la condición del ángulo:

$$\arctan \sqrt{2} - \phi_{\text{salida}} - 90^\circ = -180^\circ \quad (7.28)$$

$$\phi_{\text{salida}} = 144.73^\circ \quad (7.29)$$

El ángulo de salida del polo conjugado será -144.73° para que el lugar de las raíces sea simétrico respecto del eje real. Con estos datos, ya es posible dibujar el lugar de las raíces del sistema, Fig. 7.7.

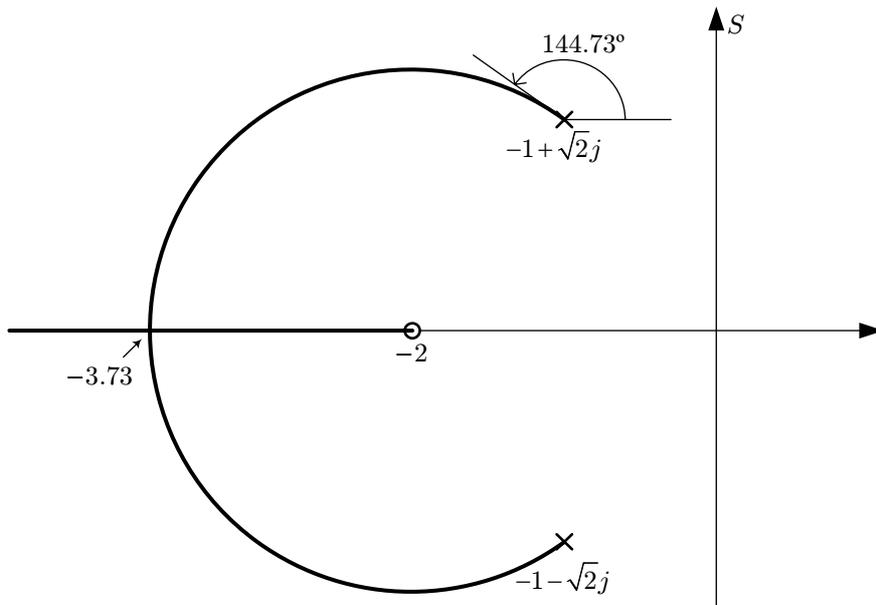


Figura 7.7: Lugar de las raíces del sistema de segundo orden con un cero

El tramo que va desde los polos del sistema hasta el punto de ruptura es una circunferencia centrada en el cero. Esto sólo ocurre en los sistemas segundo orden con un único cero. Si se desea conocer algún punto del lugar de las raíces entre los polos de salida y el punto de ruptura, es posible hacer un cambio de variable en s y encontrar los puntos de corte con los nuevos ejes imaginarios.

Por ejemplo, con el cambio de variable $\hat{s} = s + 2$, la ecuación característica en la nueva variable \hat{s} es:

$$\hat{s}^2 + (K - 2)\hat{s} + 3 = 0 \quad (7.30)$$

En los nuevos ejes el sistema es estable para valores de K mayores que 2. Para $K = 2$, los puntos de corte con los nuevos ejes son:

$$\hat{s}^2 + 3 = 0 \rightarrow \hat{s} = \pm\sqrt{3}j \quad (7.31)$$

En la Fig. 7.8 se muestran los nuevos ejes tras el cambio de variable y los puntos de corte del lugar de las raíces con dichos ejes. Se puede ver que el radio de la circunferencia mide 1.7321.

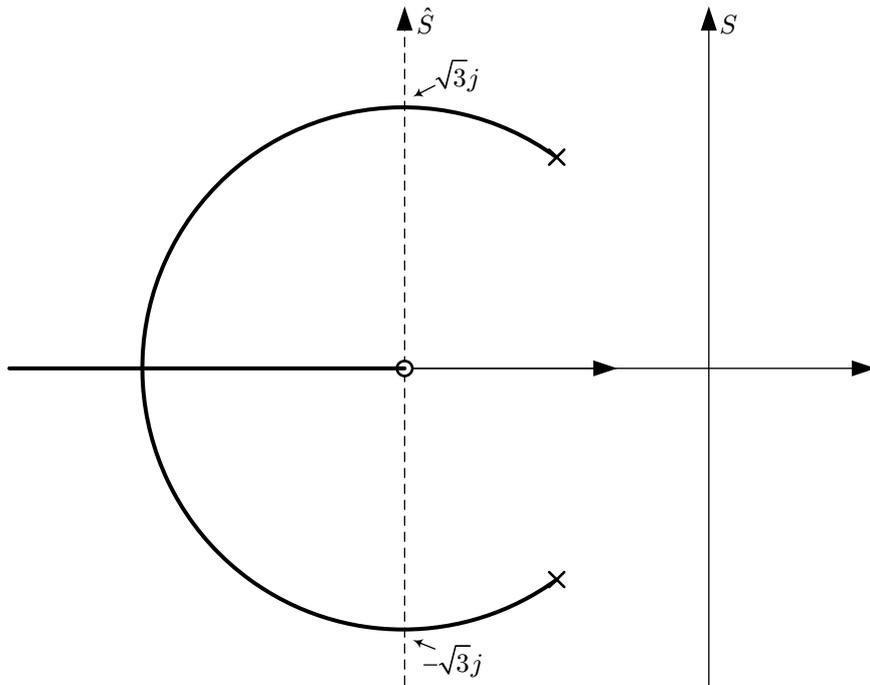


Figura 7.8: Lugar de las raíces del sistema de segundo orden con un cero

Se pueden efectuar cuantos cambios de variable se desee para determinar la posición de otros polos en lazo cerrado. Si se buscan gráficamente los puntos del lugar de las raíces que poseen un amortiguamiento de 0.7, señalados con triángulos en la Fig. 7.9, la ganancia que hace que el sistema posea dichos polos en lazo cerrado es $K = 1.39$.

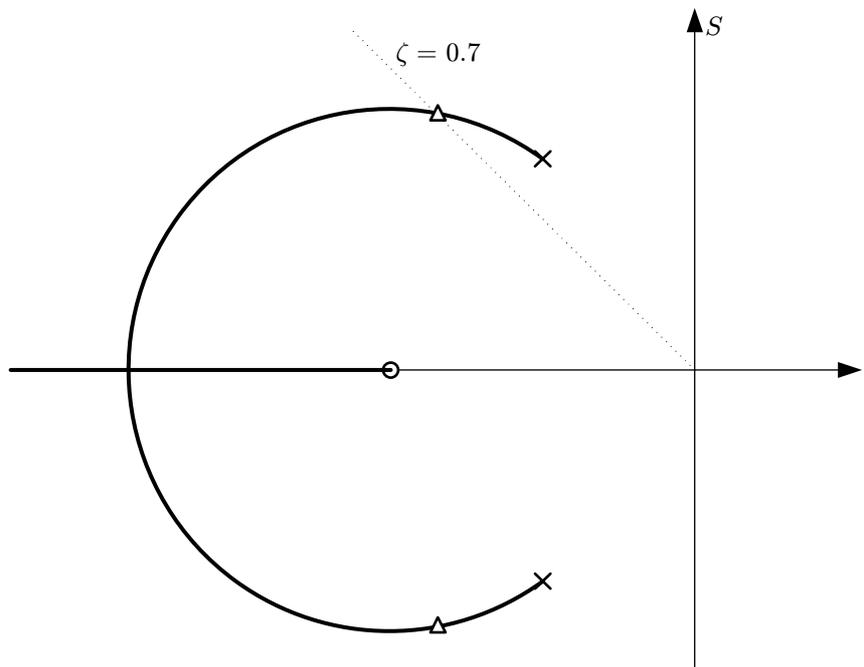


Figura 7.9: Lugar de las raíces del sistema de segundo orden con un cero

$$K = \frac{0.76 \cdot 3.18}{1.73} = 1.39 \quad (7.32)$$

Sin embargo, a pesar de que los polos en lazo cerrado tienen ese amortiguamiento, la respuesta transitoria no es exactamente subamortiguada con esa característica. La razón es que existe un cero en el sistema tan cercano al eje imaginario como los dos polos. En la ecuación (7.33) se muestra cómo en un

sistema de control con realimentación negativa unitaria los polos en lazo cerrado varían con la ganancia K mientras que los ceros en lazo cerrado no se modifican y coinciden con los ceros de la función de transferencia en lazo abierto.

$$\frac{C(s)}{R(s)} = \frac{K \frac{N(s)}{D(s)}}{1 + K \frac{N(s)}{D(s)}} = \frac{KN(s)}{D(s) + KN(s)} \quad (7.33)$$

En el ejemplo de este apartado, con la ganancia $K = 1.39$ calculada, la función de transferencia en lazo cerrado queda:

$$\frac{C(s)}{R(s)} = \frac{1.39(s+2)}{s^2 + 2s + 3 + 1.39(s+2)} \quad (7.34)$$

Los polos del sistema en lazo cerrado están en $-1.6950 \pm 1.7049j$ y el cero en -2 . En la Fig. 7.10 se compara la respuesta de este sistema ante una entrada escalón unidad, con la un sistema de segundo orden de amortiguamiento 0.5, sin cero e igual régimen permanente. La presencia del cero hace que la respuesta del sistema sea más rápida, con pendiente inicial no nula. El primer sobreimpulso se adelanta y es mayor al que impone un amortiguamiento de 0.5.

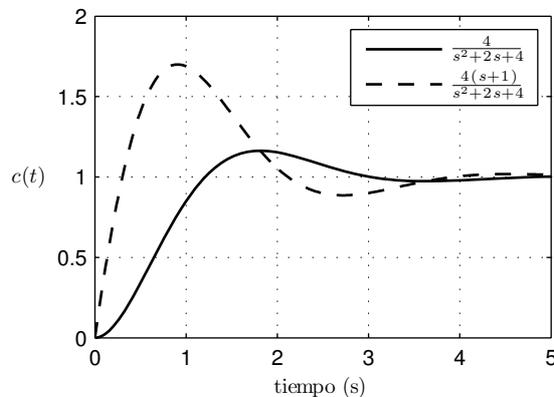


Figura 7.10: Respuesta temporal del sistema controlado de segundo orden con cero y sin cero

Con este ejemplo queda patente existen circunstancias en las que con una ley de control puramente proporcional no es posible cumplir las especificaciones de régimen transitorio. Observando la respuesta temporal del sistema, también es posible observar otra limitación del controlador puramente proporcional, y es que no se tiene ningún control sobre el error en régimen permanente del sistema. Cuando se emplea la condición del módulo para calcular la ganancia K , además de situar los polos en lazo cerrado en una determinada posición para escoger un comportamiento transitorio, también se está fijando el error en régimen permanente que va a tener el sistema.

7.6. Estabilidad relativa

El criterio de Routh-Hurwitz informa sobre la estabilidad absoluta de un sistema, pero no dice nada sobre la estabilidad relativa del mismo, es decir, cuánto de cerca o de lejos está de volverse inestable. Para dar una medida de la estabilidad relativa de un sistema se definen los conceptos de *margen de ganancia* y *margen de fase*.

7.6.1. Margen de ganancia

El margen de la ganancia es el factor proporcional que se debe introducir dentro del lazo de control para que el sistema se vuelva críticamente estable. El margen de ganancia es, por tanto, el cociente de la ganancia crítica del sistema entre la ganancia actual de la función de transferencia en lazo abierto.

En la Fig. 7.11 se muestra el lugar de las raíces de un sistema. Se ha impuesto un comportamiento transitorio con amortiguamiento 0.5 y la ganancia necesaria para ello es $K = 1.03$. El sistema se hace inestable para una ganancia crítica $K_{CR} = 6$. El margen de ganancia es $MG = 5.82$.

Si el sistema fuera estable para cualquier valor de K , el margen de ganancia del sistema sería infinito.

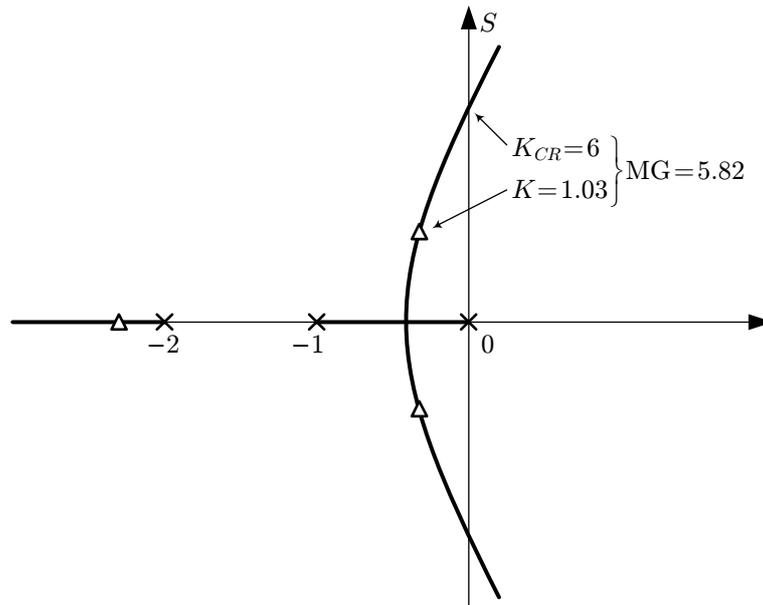


Figura 7.11: Margen de ganancia de un sistema

7.6.2. Margen de fase

El margen de fase se define como el ángulo que se puede sustraer al sistema para dejarlo en el límite de estabilidad, manteniendo constante la ganancia del mismo. En la Fig. 7.12 se señalan los puntos del plano S en los que se obtiene una ganancia en lazo abierto de $K = 2$ aplicando la condición del módulo.

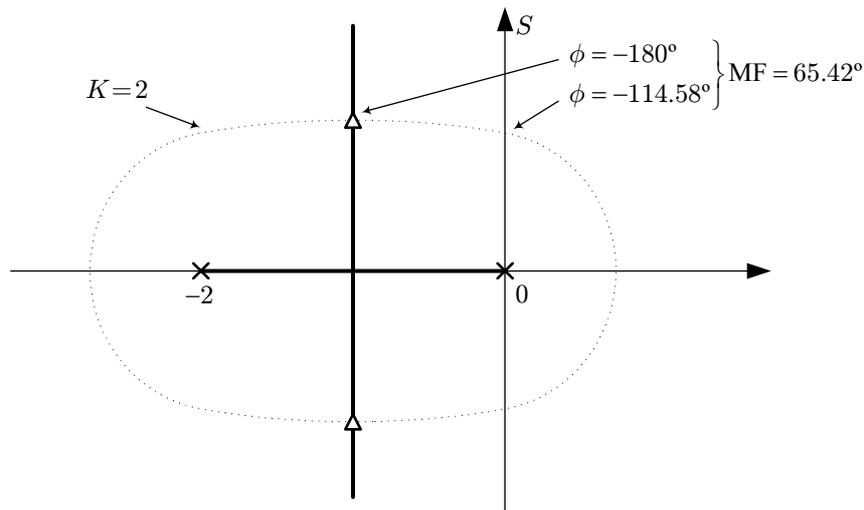


Figura 7.12: Margen de fase de un sistema

Sólo los dos puntos en los que se obtiene una fase de -180° aplicando la condición del ángulo pertenecen al lugar de las raíces. La diferencia entre la fase de esos puntos y la fase de los puntos con la misma ganancia que están sobre el eje imaginario es el margen de fase del sistema.

Los retrasos en el dominio temporal producen precisamente una pérdida de fase en el sistema que los pueden hacer inestables.

7.7. Lugar de las raíces en función de otros parámetros

La técnica del lugar de las raíces se puede emplear también si el parámetro variable no es una ganancia proporcional dentro del lazo de control. Para no contradecir el desarrollo teórico empleado, es necesario expresar la ecuación característica del sistema de la forma siguiente:

$$1 + \alpha F(s) = 0 \tag{7.35}$$

Donde $F(s)$ es una función de transferencia cualquiera. Se dibuja el lugar de las raíces como si $F(s)$ fuera la función de transferencia en lazo abierto del sistema.

Hay que señalar que la función de transferencia $F(s)$ no responde la modelización matemática de ningún elemento físico del sistema. Puede tener, por ejemplo, más ceros que polos, cosa que nunca ocurre con funciones de transferencia de elementos físicos.

7.8. Ejercicios propuestos

- **Ejercicio 1:** Dibujar el lugar de las raíces de los sistemas cuyas funciones de transferencia en lazo abierto son las siguientes:

$$G(s) = \frac{K(s+5)}{s(s+3)^3} \quad (7.36)$$

$$G(s) = \frac{3Ks(s-6)}{(s+3)[(s+4)^2+1]} \quad (7.37)$$

$$G(s) = \frac{K}{s(s+3)[(s+3)^2+49]} \quad (7.38)$$

$$G(s) = \frac{K(s^2+4s+5)}{s(s^2+2s+10)} \quad (7.39)$$

$$G(s) = \frac{K(s-1)^2}{(s+2)(s+3)[(s+2)^2+9]} \quad (7.40)$$

- **Ejercicio 2:** Sea el sistema de la figura:

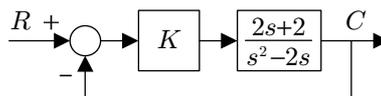


Figura 7.13: Sistema de control

Se pide determinar la ganancia K , por el método del lugar de las raíces, tal que los polos en lazo cerrado posean un amortiguamiento de 0.5.

Solución: $K = 2$

- **Ejercicio 3:** Sea el sistema:

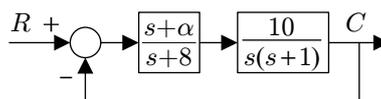


Figura 7.14: Sistema de control

- Dibujar el lugar de las raíces del sistema, en función del parámetro α .
- Elegir de forma adecuada el parámetro α para el sistema.
- ¿Qué polos y/o ceros influyen en la respuesta transitoria del sistema compensado con el parámetro elegido en el apartado anterior?

- **Ejercicio 4:** Dibujar el lugar de las raíces del sistema de la Fig. 7.15, en función del parámetro K .

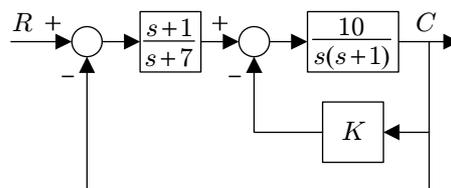


Figura 7.15: Sistema de control

- **Ejercicio 5:** Un sistema controlado posee el diagrama de bloques de la Fig. 7.16. El ingeniero puede elegir el factor constante K_2 que multiplica a la referencia R antes de entrar en el lazo, y el coeficiente K_1 del controlador. La salida del sistema se mide a través de un sensor con función de transferencia no unitaria.

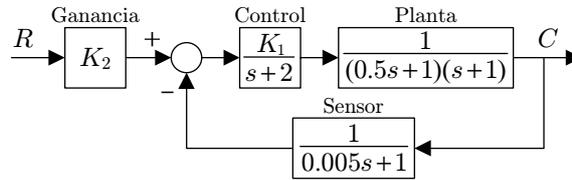


Figura 7.16: Sistema de control

Se pide asignar unos valores adecuados para las constantes K_1 y K_2 de forma que el sistema posea error nulo ante entrada referencia escalón de amplitud 50 unidades y los polos dominantes en lazo cerrado posean aproximadamente un amortiguamiento de 0.7.

Nota: Los puntos de ruptura se deben buscar entre las siguientes raíces: -150.41 , -2 y -1.33

Solución: $K_1 = 1.24$ y $K_2 = 2.61$

- **Ejercicio 6:** Dibujar el lugar de las raíces del siguiente sistema en lazo abierto:

$$G(s) = \frac{4(s+z)}{s(s+1)(s+3)} \quad (7.41)$$

Donde el parámetro z varía desde cero hasta infinito. Se pide contestar razonadamente:

- ¿Se puede dejar el sistema con amortiguamiento 0.5?
- ¿Existe algún cero que afecte la respuesta temporal? En caso positivo, ¿se podría intentar anular su efecto?

Nota: No hay puntos de ruptura.

Soluciones: Operando, la ecuación característica del sistema es $1 + z \frac{4}{s(s^2+4s+7)} = 0$.

- Sí se pueden dejar los polos complejo-conjugados del sistema en lazo cerrado con un ángulo de 60° con el origen.
- La respuesta temporal siempre está afectada por el cero en $-z$ que se puede anular con el polo real en lazo cerrado.

- **Ejercicio 7:** La Fig. 7.17 muestra el diagrama de bloques de un *extender*, que es un mecanismo que sirve para amplificar la fuerza de un operador humano.

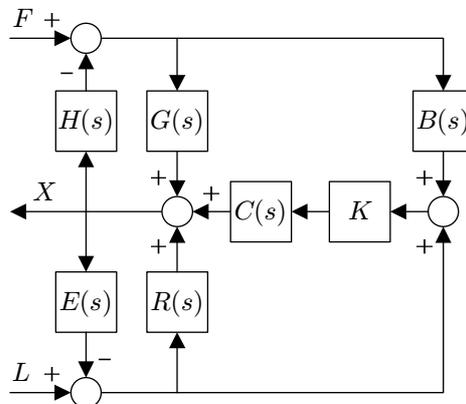


Figura 7.17: Extender

Hallar la función de transferencia general que relaciona la fuerza de entrada F y la posición de salida X .

Si las funciones de transferencia que intervienen en la figura toman los siguientes valores:

$$H(s) = E(s) = B(s) = 1, G(s) = R(s) = \frac{1}{s(s-1)}, C(s) = \frac{1}{s} \quad (7.42)$$

Entonces la función de transferencia que relaciona la fuerza F y la posición X es:

$$\frac{X(s)}{F(s)} = \frac{K(s-1)+1}{s^2+(2K-1)s+2(1-K)} \quad (7.43)$$

- Determinar el rango de ganancias K para los que dicha función de transferencia es estable.
- Dibujar el lugar de las raíces de dicha función de transferencia en función de la ganancia K .
- Elegir el valor de la ganancia K que deja los polos de la función de transferencia con amortiguamiento 1 y describir cómo será la respuesta temporal ante una entrada escalón unidad. Calcular el valor en régimen permanente, el valor y la pendiente de salida de la respuesta temporal ante entrada escalón unidad.

Soluciones: $\frac{X(s)}{F(s)} = \frac{G+KBC}{1+E(R+KC)+H(G+KBC)}$

- Estable para $0.5 < K < 1$
- Dos polos en lazo abierto ficticio en $0.5 \pm 1.32j$ y un cero en 1, un punto de ruptura en -0.41 , cortes con el eje imaginario en el origen y $\pm j$ y ángulos de salida $\pm 200.7^\circ$
- $K = 0.92$

- **Ejercicio 8:** La función de transferencia de una planta es:

$$G(s) = \frac{1}{(s+1)(s+2)(s+5)} \quad (7.44)$$

- Dibujar el lugar de las raíces.
- Elegir una ganancia de forma que el amortiguamiento de la planta sea de 0.5.
- Indicar la posición de todos los polos en lazo cerrado e indicar cuáles son los dominantes.
- Calcular el error en régimen permanente del sistema en lazo cerrado ante una entrada escalón.
- Dibujar de forma aproximada la respuesta de la planta en lazo cerrado con la ganancia anteriormente calculada. Indicar claramente cómo afectan los polos adicionales con respecto a la respuesta de la planta considerando sólo los polos dominantes.
- Se desea que el error ante una entrada escalón sea nulo (conservando en lo posible el comportamiento anteriormente calculado). Para ello se sustituye la ganancia por un cierto controlador. Indicar qué controlador es el adecuado, cuál es su ecuación y calcular los valores de sus parámetros para esta planta particular.

Nota: Para hacer algunos apartados no hace falta conocer los anteriores.

Capítulo 8

Respuesta en frecuencia

En este capítulo se estudia la respuesta del sistema ante una entrada sinusoidal. Esto permite hallar la función de transferencia de un sistema con una planta compleja mediante un método práctico sencillo.

La representación de la respuesta en frecuencia de un sistema sirve para dar una medida de su estabilidad relativa, completando la información que puede dar el criterio de Routh-Hurwitz.

8.1. Respuesta a una entrada sinusoidal

Sea $G(s)$ la función de transferencia de un sistema y $R(s)$ una entrada sinusoidal. La salida del sistema en el dominio temporal y régimen permanente es:

$$r(t) = \sin(\omega t) \implies c(t) = |G(j\omega)| \sin[\omega t + \angle G(j\omega)] \quad (8.1)$$

Por tanto el sistema amplifica o atenúa en función de la frecuencia de la señal de entrada. Lo mismo ocurre con el adelanto o retraso de la señal de salida respecto de la entrada. Existen varias formas de representar esos cambios en función de la frecuencia, Fig. 8.1. En esta asignatura, sin embargo, sólo se emplearán los diagramas de Bode.

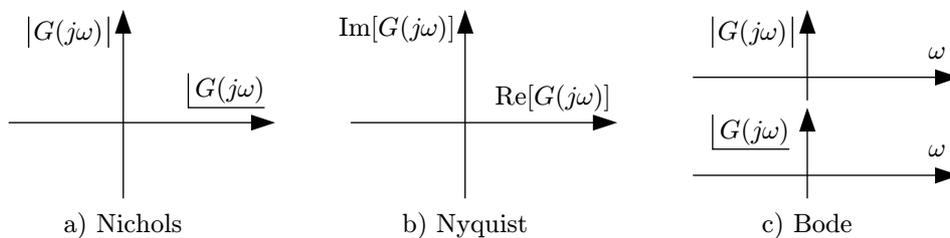


Figura 8.1: Representaciones gráficas de la respuesta en frecuencia

8.2. El diagrama de Bode

El norteamericano Hendrik Wade Bode (1905-1982) usó por primera vez en 1938 el diagrama que lleva su nombre para el estudio de estabilidad de sistemas en lazo cerrado. Durante la Segunda Guerra Mundial contribuyó al rápido desarrollo de servomecanismos para dispositivos de control de disparo. Su uso se extendió ampliamente en el estudio de los circuitos electrónicos.

Un diagrama de Bode consta de dos gráficas, una para la amplitud de salida y otra para el desfase de salida. Se los denominará respectivamente diagrama de ganancias y diagrama de fases. Los dos diagramas representan las frecuencias de forma logarítmica en el eje de abscisas empleando rad/s.

El diagrama de ganancias representa en el eje de ordenadas la amplitud de la señal de salida transformados a decibelios, ecuación (8.2). El diagrama de fases representa en el eje de ordenadas el desfase de la señal de salida en grados.

$$20 \log |G(j\omega)| \quad (8.2)$$

En realidad, el uso de los decibelios como unidad de medida es una forma solapada de representar la amplitud de salida en escala logarítmica. Conviene resaltar que los logaritmos son siempre decimales, no

neperianos. El factor 20 de la definición (8.2) se debe en parte al uso de la fracción del belio y en parte al empleo de la potencia de la señal —lo que hace que haya que elevar al cuadrado la amplitud dentro del logaritmo y salga fuera de él como un factor de dos—.

En el eje logarítmico de frecuencias se denomina *década* a cualquier intervalo que va desde una determinada frecuencia hasta otra diez veces mayor. Se denomina *octava* a cualquier intervalo que va desde una frecuencia hasta su doble. El nombre no tiene nada que ver con un hipotético factor de ocho, sino que proviene de las notas musicales. La escala musical tiene siete notas, al hacer sonar la octava nota se obtiene la misma nota que la inicial con el doble de frecuencia. Tanto la década como la octava son distancias constantes en una escala logarítmica.

8.3. Diagramas de Bode de sistemas elementales

Para poder dibujar el diagrama de Bode de una función de transferencia cualquiera, es necesario conocer la forma que adopta dicho diagrama es el caso de las funciones de transferencia más elementales. Las funciones de transferencia más complicadas se obtendrán como combinación de las elementales. Las funciones de transferencia que se tomarán como elementales son: una ganancia, un retraso en el tiempo, un integrador, un derivador, un polo, un cero, un polo doble y un cero doble.

8.3.1. Ganancia

Una ganancia se limita a amplificar o a atenuar la entrada sin introducir retrasos o adelantos en la señal de salida. Por tanto, es de esperar que el diagrama de Bode de una ganancia sea nulo en fases y no nulo en amplitud.

$$G(s) = K \xrightarrow{s=j\omega} G(j\omega) = K \begin{cases} 20 \log |G(j\omega)| = 20 \log K \\ \angle G(j\omega) = 0^\circ \end{cases} \quad (8.3)$$

Como se observa el diagrama de Bode de la Fig. 8.2, es lógico que una ganancia amplifique o atenúe siempre el mismo factor cualquiera que sea la frecuencia de la señal de entrada, es decir, adopte una forma constante con ω .

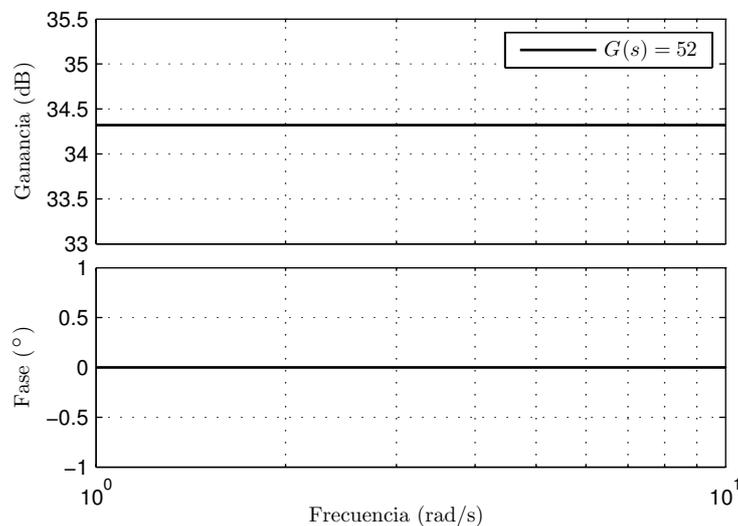


Figura 8.2: Diagrama de Bode de una ganancia

Si K es menor que la unidad, la ganancia atenúa y se obtiene un nivel de decibelios negativo. Si K es mayor que la unidad, la ganancia amplifica y se obtiene un nivel de decibelios positivo. Por tanto, en el tramo del diagrama de Bode que los decibelios sean positivos, quiere decir que la señal de entrada se amplifica, mientras que en los tramos de decibelios negativos, la señal de entrada se atenúa.

8.3.2. Retraso en el tiempo

Un retraso ni amplifica ni atenúa. La forma de la salida es exactamente igual a la de la entrada, aunque la salida está retrasada T segundos respecto de la entrada. Dicho esto, es de esperar que sea nulo

el diagrama de ganancias y negativo el de fases.

$$G(s) = e^{-Ts} \xrightarrow{s=j\omega} G(j\omega) = e^{-j\omega T} \begin{cases} 20 \log |G(j\omega)| = 0 \text{ dB} \\ \angle G(j\omega) = -\omega T \text{ rad} \end{cases} \quad (8.4)$$

Para una frecuencia en rad/s igual a la inversa del tiempo T de retraso, el diagrama de fases toma un valor de -1 rad. Una década después -10 rad. Dos décadas después -100 rad. Así sucesivamente.

El diagrama de Bode, Fig. 8.3, muestra que la función de transferencia genera desfases cada vez mayores con la frecuencia. El desfase es directamente proporcional a la frecuencia, por tanto, la gráfica es una línea recta con la frecuencia en escala lineal y queda con forma exponencial con la frecuencia en escala logarítmica.

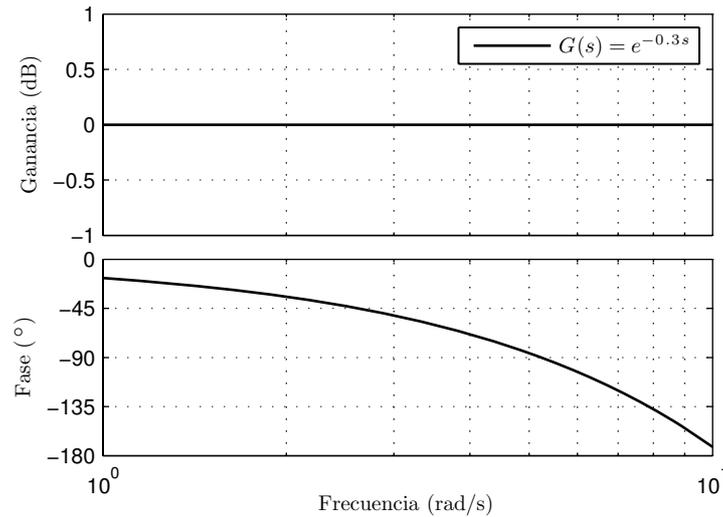


Figura 8.3: Diagrama de Bode de un retraso

La salida se adelanta respecto de la entrada, hecho poco probable en el mundo real, la función de transferencia sería una exponencial con exponente positivo. Su diagrama de Bode en fases tendría ángulos positivos. Por este motivo, cuando en el diagrama de fases aparezcan ángulos negativos se hablará de retrasos de fases y, al revés, con ángulos positivos se hablará de adelanto de fases.

8.3.3. Integrador

Un integrador tiene por salida la integral de la función de entrada. El diagrama de Bode de fases es constante en -90° . Esto es lógico ya que la integral de un seno es un menos coseno, que está retrasado 90° respecto el seno.

$$G(s) = \frac{1}{s} \xrightarrow{s=j\omega} G(j\omega) = \frac{1}{j\omega} = \frac{1}{\omega} \angle -90^\circ \begin{cases} 20 \log |G(j\omega)| = -20 \log \omega \\ \angle G(j\omega) = -90^\circ \end{cases} \quad (8.5)$$

El diagrama de Bode de ganancias es una recta con pendiente -20 dB/década. La recta pasa por 0 dB en la frecuencia de 1 rad/s.

8.3.4. Derivador

Un derivador tiene por salida la derivada de la función de entrada. El diagrama de Bode de fases es constante en 90° . Esto es lógico ya que la derivada de un seno es un coseno, que está adelantado 90° respecto el seno.

$$G(s) = s \xrightarrow{s=j\omega} G(j\omega) = j\omega = \omega \angle 90^\circ \begin{cases} 20 \log |G(j\omega)| = 20 \log \omega \\ \angle G(j\omega) = 90^\circ \end{cases} \quad (8.6)$$

El diagrama de Bode de ganancias es una recta con pendiente 20 dB/década. Igual que los integradores, la recta pasa por 0 dB en la frecuencia de 1 rad/s.

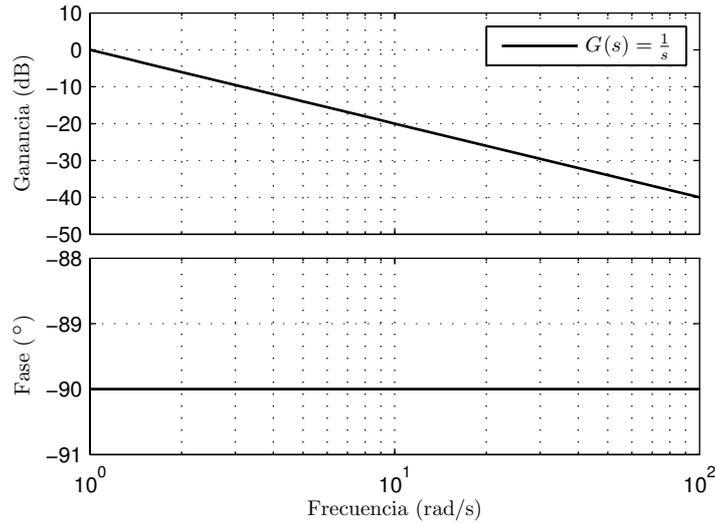


Figura 8.4: Diagrama de Bode de un integrador

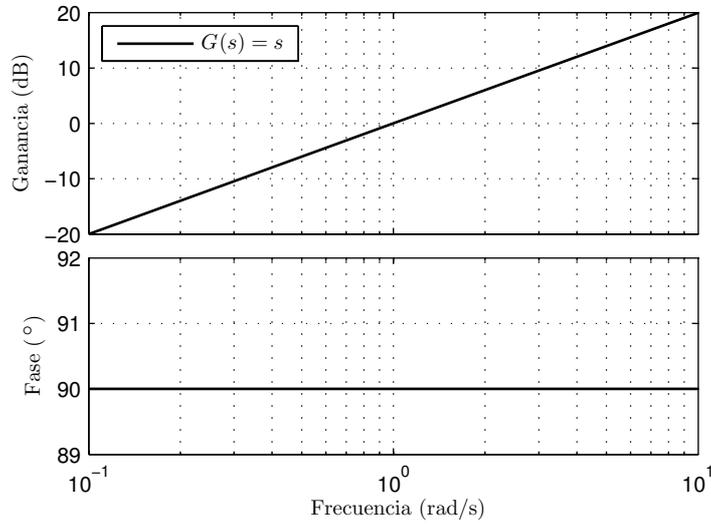


Figura 8.5: Diagrama de Bode de un derivador

8.3.5. Polo simple estable

Se estudiará en este apartado un polo simple estable con ganancia estática igual a la unidad. Es un sistema ideal de primer orden.

$$G(s) = \frac{1}{1 + Ts} \xrightarrow{s=j\omega} G(j\omega) = \frac{1}{1 + j\omega T} \quad (8.7)$$

Las ganancias y fases de la ecuación (8.7) se van a particularizar para distintos casos. En la ecuación (8.8) se observa cómo la ganancia a bajas frecuencias es aproximadamente 0 dB. Para altas frecuencias la ganancia se parece a un integrador, una recta de pendiente -20 dB/década, que pasa por 0 dB en la frecuencia igual a la inversa de la constante de tiempo.

$$20 \log |G(j\omega)| = -20 \log \sqrt{1 + (\omega T)^2} \begin{cases} \omega \rightarrow 0 \implies 0 \text{ dB} \\ \omega = \frac{1}{T} \implies -20 \log \sqrt{2} \approx -3 \text{ dB} \\ \omega \rightarrow \infty \implies -20 \log(\omega T) \text{ dB} \end{cases} \quad (8.8)$$

En la ecuación (8.9) se observa cómo la fase para bajas frecuencias es aproximadamente 0° y para altas frecuencias -90° . El diagrama de Bode de la Fig. 8.6 corrobora el comportamiento en ganancias y en fases.

$$\angle G(j\omega) = -\arctan(\omega T) \begin{cases} \omega \rightarrow 0 \implies 0^\circ \\ \omega = \frac{1}{T} \implies -45^\circ \\ \omega \rightarrow \infty \implies -90^\circ \end{cases} \quad (8.9)$$

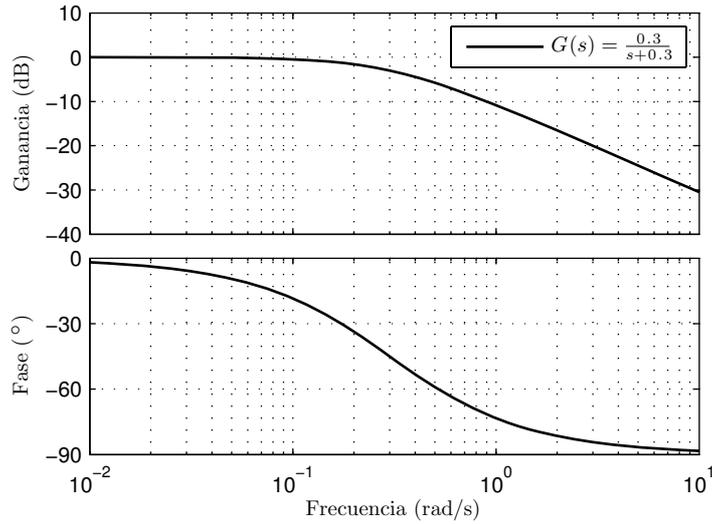


Figura 8.6: Diagrama de Bode de un polo

8.3.6. Cero simple con parte real negativa

En este apartado se estudia el caso inverso del apartado anterior.

$$G(s) = 1 + Ts \xrightarrow{s=j\omega} G(j\omega) = 1 + j\omega T \quad (8.10)$$

La ganancia a bajas frecuencias también comienzan en 0 dB, ecuación (8.11). En cambio, para altas frecuencias la ganancia se comporta como un derivador; una recta de pendiente 20 dB/década, que pasa por 0 dB en la frecuencia igual a la inversa de la constante de tiempo.

$$20 \log |G(j\omega)| = 20 \log \sqrt{1 + (\omega T)^2} \begin{cases} \omega \rightarrow 0 \Rightarrow 0 \text{ dB} \\ \omega = \frac{1}{T} \Rightarrow 20 \log \sqrt{2} \approx 3 \text{ dB} \\ \omega \rightarrow \infty \Rightarrow 20 \log(\omega T) \text{ dB} \end{cases} \quad (8.11)$$

En fases, ecuación (8.12), para bajas frecuencias toma valores próximos a 0° y para altas frecuencias aproximadamente 90° .

$$\angle G(j\omega) = \arctan(\omega T) \begin{cases} \omega \rightarrow 0 \Rightarrow 0^\circ \\ \omega = \frac{1}{T} \Rightarrow 45^\circ \\ \omega \rightarrow \infty \Rightarrow 90^\circ \end{cases} \quad (8.12)$$

El diagrama de Bode de la Fig. 8.7 muestra el comportamiento en ganancias y en fases.

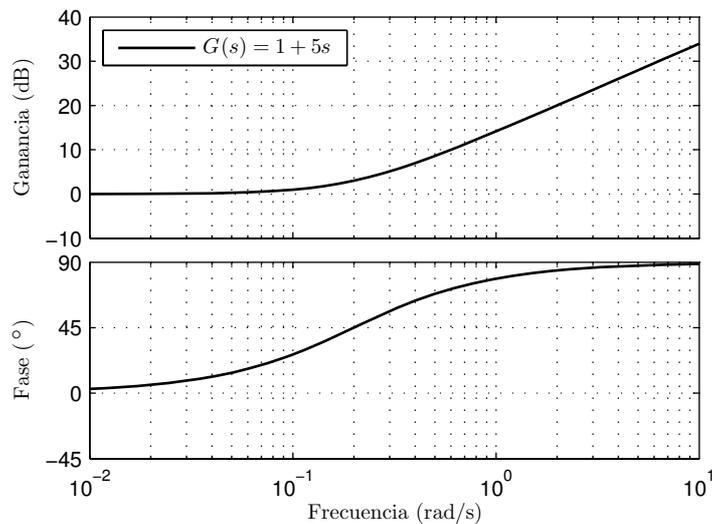


Figura 8.7: Diagrama de Bode de un cero

8.3.7. Polos estables complejos conjugados

Es este apartado se analiza un sistema de segundo orden con polos complejo conjugados estables y ganancia estática igual a la unidad.

$$G(s) = \frac{\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2} \xrightarrow{s=j\omega} G(j\omega) = \frac{\omega_n^2}{\omega_n^2 - \omega^2 + j2\zeta\omega\omega_n} \quad (8.13)$$

En la ecuación (8.14) se observa cómo la ganancia para bajas frecuencias es aproximadamente 0 dB y para altas frecuencias es una recta de pendiente -40 dB/década que pasa por 0 dB en la frecuencia igual a la frecuencia natural.

$$20 \log |G(j\omega)| = 20 \log \frac{\omega_n^2}{\sqrt{(\omega_n^2 - \omega^2)^2 + (2\zeta\omega\omega_n)^2}} \begin{cases} \omega \rightarrow 0 \implies 0 \text{ dB} \\ \omega \rightarrow \infty \implies -40 \log \frac{\omega}{\omega_n} \text{ dB} \end{cases} \quad (8.14)$$

En la ecuación (8.15) se observa cómo la fase para bajas frecuencias es aproximadamente 0° y para altas frecuencias -180° . El diagrama de la Fig. 8.8 muestra el comportamiento en ganancias y en fases.

$$\angle G(j\omega) = -\arctan \frac{2\zeta\omega\omega_n}{\omega_n^2 - \omega^2} \begin{cases} \omega \rightarrow 0 \implies 0^\circ \\ \omega = \omega_n \implies -90^\circ \\ \omega \rightarrow \infty \implies -180^\circ \end{cases} \quad (8.15)$$

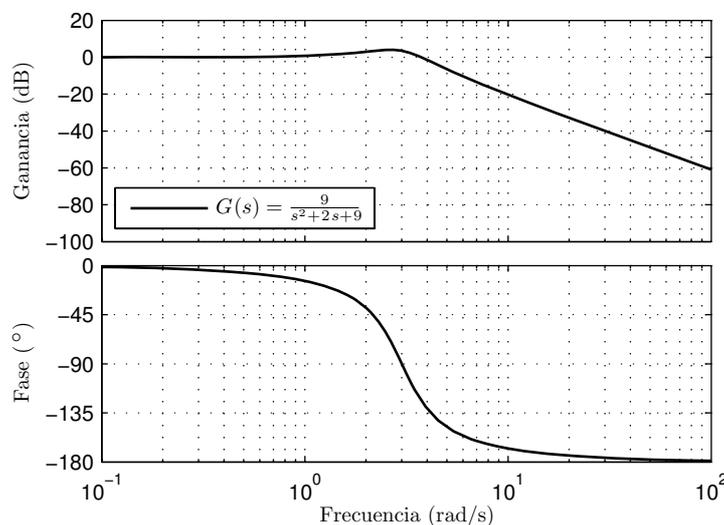


Figura 8.8: Diagrama de Bode de dos polos complejo conjugados

En un rango de frecuencias próximo a la frecuencia natural, el diagrama de Bode se comporta de forma distinta en función del amortiguamiento. En la Fig. 8.8 se observa cómo aparece un máximo en el diagrama de ganancias. Se va a emplear la expresión (8.14) para determinar la magnitud de dicho máximo y a qué frecuencia se produce. En concreto, se va a derivar la expresión del denominador para buscar un mínimo —que será un máximo de la ganancia—.

$$\frac{d}{d\omega} [(\omega_n^2 - \omega^2)^2 + (2\zeta\omega\omega_n)^2] = -4\omega[(1 - 2\zeta^2)\omega_n^2 - \omega^2] = 0 \begin{cases} \omega = 0 \\ \omega = \omega_r = \omega_n \sqrt{1 - 2\zeta^2} \end{cases} \quad (8.16)$$

Si se estudiara el signo de la segunda derivada, se comprobaría cómo la primera solución, para frecuencia nula, se trata de un máximo del denominador y por tanto un mínimo del diagrama de Bode. La otra solución ω_r , llamada *frecuencia de resonancia*, es el máximo del diagrama de Bode que se había observado. Si se sustituye el valor de la frecuencia de resonancia en la expresión de la ganancia (8.14), se obtiene la magnitud del máximo, ecuación (8.17), que se suele denominar *pico de resonancia*.

$$M_r = 20 \log \frac{1}{2\zeta\sqrt{1 - \zeta^2}} \quad (8.17)$$

El fenómeno de la resonancia no siempre existe. Sólo se da para un determinado rango de amortiguamientos. En concreto, aquellos amortiguamientos que hacen positivo el discriminante de la raíz cuadrada de la expresión (8.16).

$$\exists M_r \Leftrightarrow 0 \leq \zeta \leq \frac{1}{\sqrt{2}} = 0.7071 \quad (8.18)$$

En la ecuación (8.18) aparece el rango de amortiguamientos con pico de resonancia. Se trata siempre de sistemas subamortiguados, aunque no todos los sistemas subamortiguados poseen pico de resonancia. En la Fig. 8.9 se muestra cómo varía el diagrama de Bode con el amortiguamiento. Cuanto menor es el amortiguamiento mayor es el pico de resonancia y más próximo está a la frecuencia natural. También cuanto menor es el amortiguamiento más brusco es el cambio de fases en torno a la frecuencia natural.

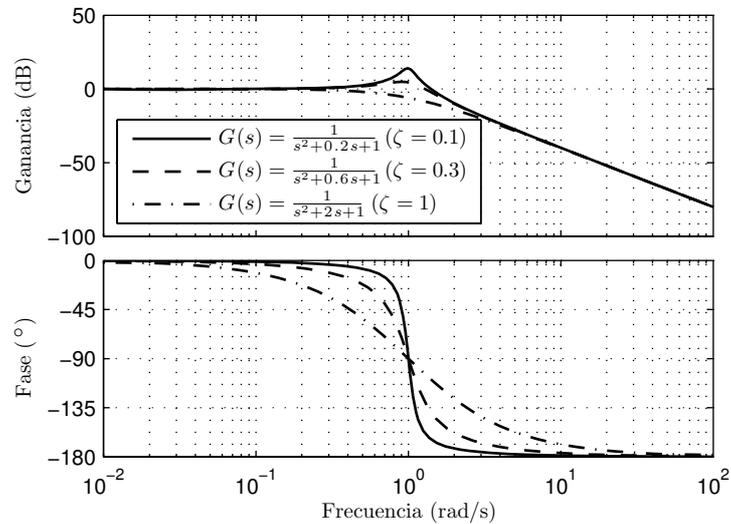


Figura 8.9: Diagrama de Bode de un sistema de segundo orden variando el amortiguamiento

En el caso concreto de amortiguamiento igual a uno, es decir, críticamente amortiguado y polo real doble, el diagrama de ganancias pasa por -6 dB en la frecuencia natural.

8.3.8. Ceros complejo conjugados

En este apartado se estudia el caso inverso al del apartado anterior.

$$G(s) = \frac{s^2 + 2\zeta\omega_n s + \omega_n^2}{\omega_n^2} \xrightarrow{s=j\omega} G(j\omega) = \frac{\omega_n^2 - \omega^2 + j2\zeta\omega\omega_n}{\omega_n^2} \quad (8.19)$$

El desarrollo matemático es análogo, por lo que no se va a repetir. En el caso de los ceros complejo conjugados también existe el fenómeno de la resonancia, sólo que se manifiesta en forma de mínimo en lugar de un máximo en el diagrama de ganancias.

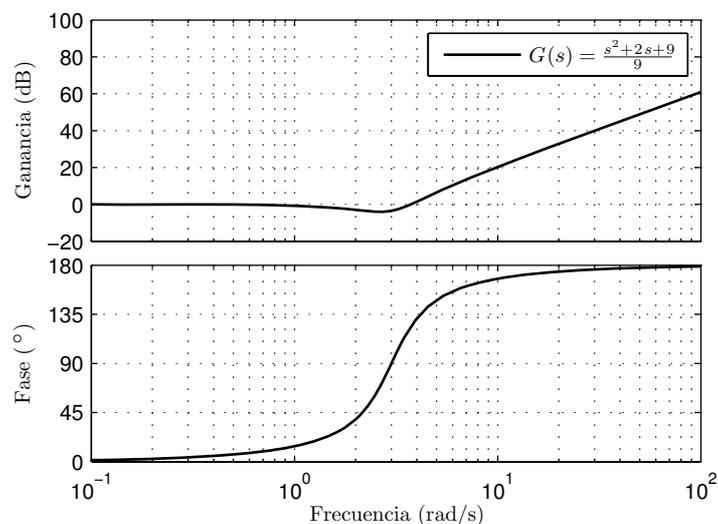


Figura 8.10: Diagrama de Bode de dos ceros complejo conjugados

8.3.9. Polo simple con parte real positiva

Un polo con parte real positiva constituye, en sí mismo, un sistema inestable. Estrictamente hablando, un sistema inestable no tiene diagrama de Bode. Sin embargo, se puede considerar que forma parte de un sistema estable. Esta situación se da, por ejemplo, cuando en lazo abierto la función de transferencia tiene un polo inestable, pero la función de transferencia en lazo cerrado es estable. Es posible, en este caso, buscar la respuesta matemática y el diagrama de Bode que resultan de considerar las entradas y salidas particulares de ese polo “inestable”.

$$G(s) = \frac{1}{1 - Ts} \xrightarrow{s=j\omega} G(j\omega) = \frac{1}{1 - j\omega T} \quad (8.20)$$

$$20 \log |G(j\omega)| = -20 \log \sqrt{1 + (\omega T)^2} \begin{cases} \omega \rightarrow 0 \Rightarrow 0 \text{ dB} \\ \omega = \frac{1}{T} \Rightarrow -20 \log \sqrt{2} \approx -3 \text{ dB} \\ \omega \rightarrow \infty \Rightarrow -20 \log(\omega T) \text{ dB} \end{cases} \quad (8.21)$$

$$\angle G(j\omega) = \arctan(\omega T) \begin{cases} \omega \rightarrow 0 \Rightarrow 0^\circ \\ \omega = \frac{1}{T} \Rightarrow 45^\circ \\ \omega \rightarrow \infty \Rightarrow 90^\circ \end{cases} \quad (8.22)$$

Su comportamiento en ganancias es igual que un polo con parte real negativa. Esto se debe a que el cambio de signo no afecta al módulo del número complejo. Sin embargo, su comportamiento en fases es igual que un cero simple.

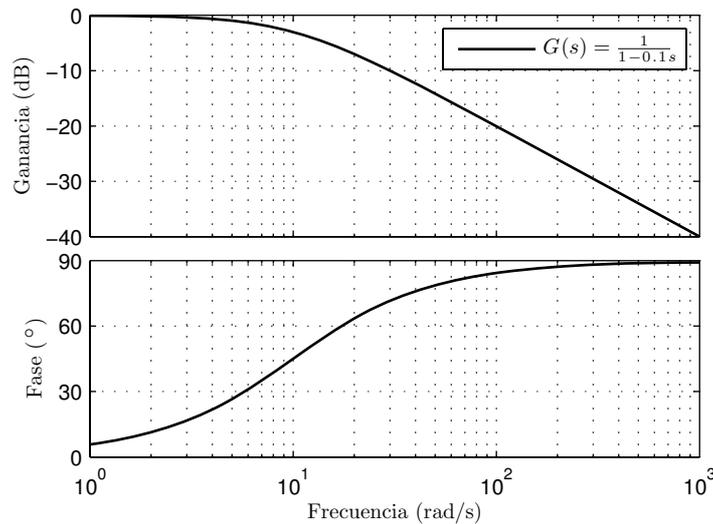


Figura 8.11: Diagrama de Bode de un polo con parte real positiva

8.3.10. Cero simple con parte real positiva

Un cero con parte real positiva puede darse en un sistema estable, pero su presencia ya se mostró cómo convertía a una planta en un sistema de fase no mínima.

$$G(s) = 1 - Ts \xrightarrow{s=j\omega} G(j\omega) = 1 - j\omega T \quad (8.23)$$

$$20 \log |G(j\omega)| = 20 \log \sqrt{1 + (\omega T)^2} \begin{cases} \omega \rightarrow 0 \Rightarrow 0 \text{ dB} \\ \omega = \frac{1}{T} \Rightarrow 20 \log \sqrt{2} \approx 3 \text{ dB} \\ \omega \rightarrow \infty \Rightarrow 20 \log(\omega T) \text{ dB} \end{cases} \quad (8.24)$$

$$\angle G(j\omega) = -\arctan(\omega T) \begin{cases} \omega \rightarrow 0 \Rightarrow 0^\circ \\ \omega = \frac{1}{T} \Rightarrow -45^\circ \\ \omega \rightarrow \infty \Rightarrow -90^\circ \end{cases} \quad (8.25)$$

Su comportamiento en ganancias es exactamente igual que un cero con parte real negativa, mientras que su comportamiento en fases es igual que un polo simple.

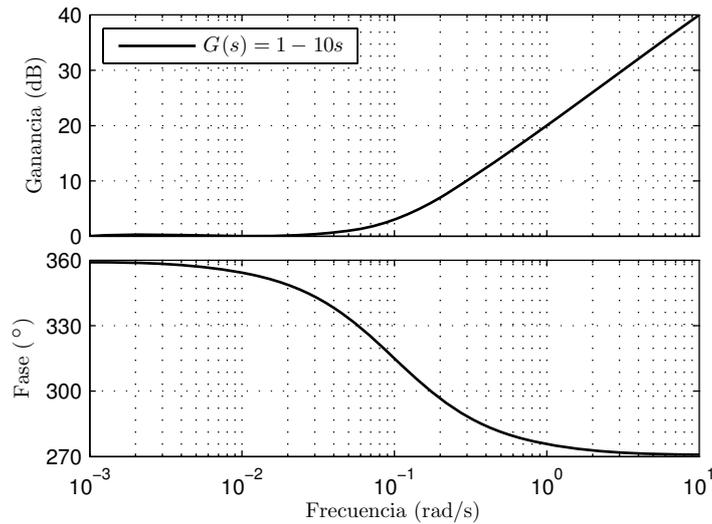


Figura 8.12: Diagrama de Bode de un cero con parte real positiva

8.4. Diagrama de Bode de cualquier función de transferencia

Para dibujar el diagrama de Bode de una función de transferencia cualquiera, se suman las aportaciones de cada una de las funciones elementales en las que se puede desglosar. Para el caso general de que la función de transferencia se pueda dividir en dos funciones de transferencia elementales:

$$G(s) = G_1(s)G_2(s) \xrightarrow{s=j\omega} G(j\omega) = G_1(j\omega)G_2(j\omega) \quad (8.26)$$

$$G(j\omega) = G_1(j\omega)G_2(j\omega) \begin{cases} 20 \log |G(j\omega)| = 20 \log |G_1(j\omega)| + 20 \log |G_2(j\omega)| \\ \angle G(j\omega) = \angle G_1(j\omega) + \angle G_2(j\omega) \end{cases} \quad (8.27)$$

Cualquiera que sea número de funciones elementales que compongan un sistema, basta con sumar para cada frecuencia los diagramas de Bode de cada una ellas. A continuación se muestra el ejemplo la función de transferencia (8.28), que se ha dividido en cuatro funciones de transferencia elementales: una ganancia, un integrador, un cero y un polo.

$$G(s) = \frac{25(s+3)}{s(s+50)} = 1.5 \frac{1}{s} \frac{s+3}{3} \frac{50}{s+50} \quad (8.28)$$

El diagrama de Bode de esta función de transferencia se muestra en la Fig. 8.13.

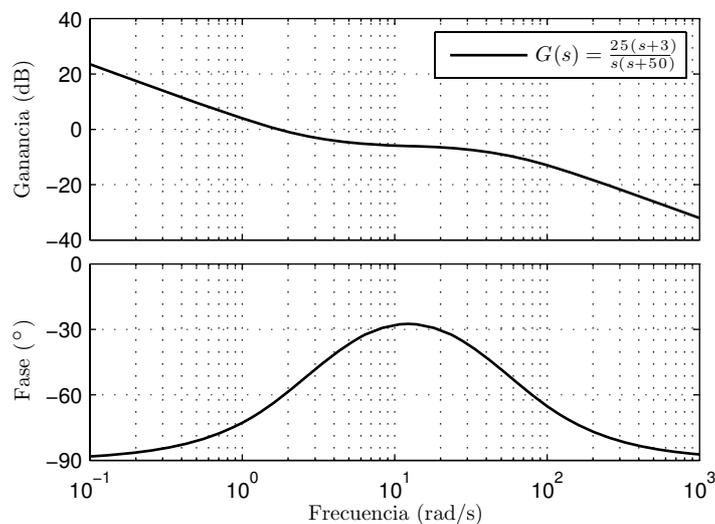


Figura 8.13: Diagrama de Bode de una función de transferencia compuesta

8.5. Diagrama de Bode de un sistema en lazo cerrado

En los sistemas controlados es habitual introducir una planta dentro de un lazo de realimentación negativa. Para dibujar el diagrama de Bode de la función de transferencia del sistema en lazo cerrado se puede emplear el de la función de transferencia en lazo abierto.

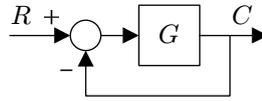


Figura 8.14: Sistema de control en lazo cerrado

Para el sistema de control en lazo cerrado de la Fig. 8.14 se desea dibujar el diagrama de Bode de la función de transferencia en lazo cerrado, ecuación (8.29).

$$G_{lc}(s) = \frac{G(s)}{1 + G(s)} \quad (8.29)$$

Se supone conocido el diagrama de Bode de la función de transferencia en lazo abierto, Fig. 8.15, y se dibuja directamente el de la función de transferencia en lazo cerrado.

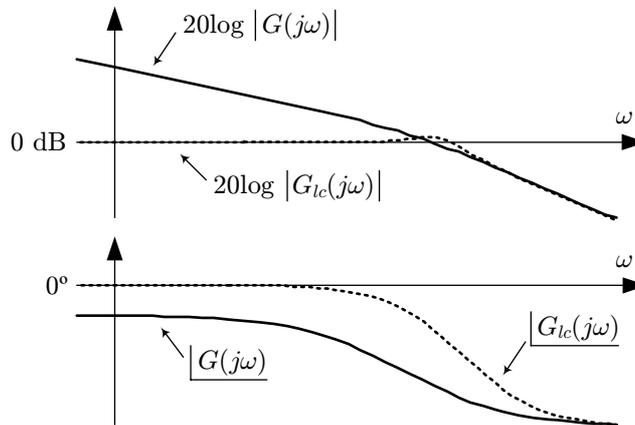


Figura 8.15: Diagrama de Bode del sistema en lazo abierto y en lazo cerrado

Para bajas frecuencias el sistema en lazo cerrado es la unidad, 0 dB y 0°, y a frecuencias elevadas es igual que el diagrama de Bode el lazo abierto. La frecuencia que marca la zona intermedia es la frecuencia de cruce de ganancias, es decir, la frecuencia que toma 0 dB el diagrama en lazo abierto.

$$G_{lc}(j\omega) = \frac{G(j\omega)}{1 + G(j\omega)} \begin{cases} |G(j\omega)| \gg 1 \implies G_{lc}(j\omega) \approx 1 \\ |G(j\omega)| \ll 1 \implies G_{lc}(j\omega) \approx G(j\omega) \end{cases} \quad (8.30)$$

En torno a la frecuencia de cruce de ganancias, el diagrama de Bode en lazo cerrado puede presentar un pico de resonancia o no, en función de la fase que posea en dicha frecuencia el diagrama en lazo abierto.

8.5.1. Ancho de banda

Es interesante apreciar cómo es la respuesta en frecuencia de un sistema controlado, es decir, en lazo cerrado. Sin importar la forma que posea la planta en lazo abierto, el sistema controlado en lazo cerrado es capaz de seguir fielmente a la entrada de referencia del sistema hasta un determinado valor de frecuencia. A partir de ese valor, que es prácticamente la frecuencia de cruce de ganancias del diagrama de Bode en lazo abierto, el sistema controlado empieza a atenuar y a retrasar la referencia: la salida del sistema no es capaz de seguir a la entrada.

Este fenómeno es una explicación intuitiva del concepto de *ancho de banda* de un sistema controlado. Matemáticamente se define como el valor de la frecuencia en rad/s en el que la ganancia del diagrama de Bode en lazo cerrado toma el valor de -3 dB. En la práctica este valor exacto se puede aproximar al valor de la frecuencia de cruce de fases del diagrama en lazo abierto.

8.5.2. Margen de fase y margen de ganancia

En el capítulo anterior se definieron los márgenes de fase y ganancia y se relacionaron con la estabilidad relativa del sistema. Es importante resaltar que los márgenes de fase y ganancia de un sistema controlado, y por tanto en lazo cerrado, se miden sobre el diagrama de Bode de la función de transferencia en lazo abierto.

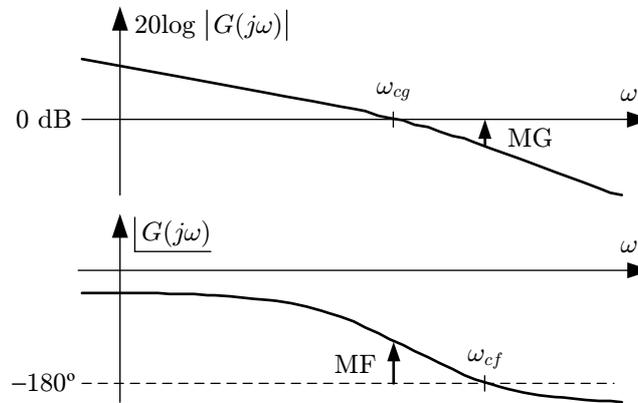


Figura 8.16: Margen de fase y de ganancia

El margen de ganancia MG se mide en la frecuencia de cruce de fases del diagrama de Bode en lazo abierto. Es el valor en decibelios faltan al sistema para alcanzar los 0 dB. Si el diagrama está por encima de 0 dB a esa frecuencia, se dice que el margen de ganancia es negativo.

El margen de fase MF se mide en la frecuencia de cruce de ganancias fases del diagrama de Bode en lazo abierto. Es el valor en grados que hay por encima de -180° hasta el diagrama de fases. Si el diagrama de fases está por debajo de -180° a esa frecuencia, se dice que el margen de fase es negativo.

Un sistema en lazo cerrado es estable cuando sus márgenes de fase y ganancia son ambos positivos.

8.6. Ejercicios propuestos

- **Ejercicio 1:** Dibujar el diagrama de Bode de las siguientes funciones de transferencia:

$$G(s) = 10 \frac{(s + 3)}{(s + 1)(s^2 + 80s + 10000)} \quad (8.31)$$

$$G(s) = \frac{400}{s(s^2 + 80s + 40000)} \quad (8.32)$$

$$G(s) = 700 \frac{s - 5}{(s + 1)(s + 100)} \quad (8.33)$$

$$G(s) = 10 \frac{e^{-0.3s}}{s + 1} \quad (8.34)$$

$$G(s) = 0.02 \frac{s - 1}{s(s + 20)(s + 0.1)} \quad (8.35)$$

Capítulo 9

Compensadores de adelanto y de retraso de fase

Controlar un sistema dado es establecer en su funcionamiento de acuerdo con unos requisitos o especificaciones. En este capítulo se analizan los cambios que se pueden lograr en un dispositivo al añadir un nuevo polo y un nuevo cero al sistema, que se introduce en un lazo de control.

Estos nuevos elementos no tienen por qué ser de las mismas características que la planta. Si la planta es un sistema mecánico, el controlador puede estar constituido por nuevos elementos mecánicos o un dispositivo eléctrico. La implementación final del compensador no importa mientras su función de transferencia, la ecuación diferencial que gobierna su comportamiento, sea la misma.

Se puede definir compensación como la modificación de la dinámica de un sistema para cumplir unas especificaciones determinadas.

9.1. Generalidades

El compensador objeto de estudio en este capítulo actúa sobre la planta en función del error que existe entre la salida del sistema y la referencia que se desea seguir, Fig. 9.1, y posee una función de transferencia con una ganancia, un polo y un cero.

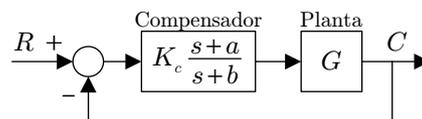


Figura 9.1: Sistema compensado

Con estos presupuestos, se da por supuesto que el sistema se controla con un lazo de realimentación negativa unitaria. Sin embargo, en ocasiones pueden diseñarse compensadores que se coloquen en otras posiciones del lazo de control, resultando una realimentación negativa no unitaria. Este tipo de compensadores no se estudiarán en la presente asignatura.

9.1.1. Especificaciones

Las especificaciones exigibles a un sistema pueden ser de muy distinta índole. Habitualmente se clasifican como restricciones del sistema en el dominio temporal:

- **Régimen transitorio:** tiempo de establecimiento, tiempo de crecimiento, tiempo de pico, sobreimpulso máximo, ancho de banda, etc.
- **Régimen permanente:** error nulo o limitado ante un tipo determinado de entrada, rechazo a las perturbaciones, etc.

9.1.2. Tipos de compensación

Es posible lograr que un sistema cumpla una serie de especificaciones con distintos compensadores. Una posible clasificación para los tipos de compensador puede ser:

- **Compensador de adelanto de fase.** El cero produce un adelanto de fase a bajas frecuencias respecto el polo. El resultado final es que el compensador adelanta fase en un determinado rango de frecuencias.
- **Compensador de retraso de fase.** En este caso el polo produce un retraso de fase a más bajas frecuencias respecto el cero. El resultado final es que el compensador retrasa fase en un determinado rango de frecuencias.
- **Compensador de adelanto-retraso de fase:** el compensador consiste el producto de las funciones de transferencia de un compensador de adelanto de fase y uno de retraso. El compensador de retraso se coloca a menores frecuencias que el de adelanto.

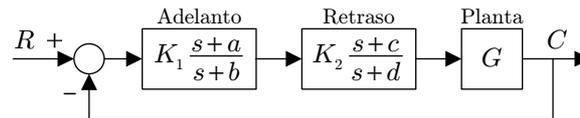


Figura 9.2: Sistema con compensador de adelanto-retraso

9.1.3. Método de ajuste

El diseño o ajuste de los parámetros de un compensador se puede realizar de varias formas. Dependiendo de las características de la planta la resolución podrá ser más sencilla con uno de los dos métodos más empleados en el control clásico:

- **Lugar de las raíces:** En el plano S se añade el nuevo polo y cero del compensador. Se modifica el lugar de las raíces del sistema de forma que los polos dominantes de la planta en lazo cerrado queden ubicados donde se desee.
- **Diagrama de Bode:** Sobre el diagrama de Bode de la planta se añade la contribución del compensador. Teniendo como grados de libertad la ganancia junto con la posición del polo y el cero, es posible modificar la respuesta en frecuencia del sistema en lazo abierto de forma que se consiguieran cumplir una amplia gama de especificaciones, tanto de régimen transitorio como de régimen permanente.

9.2. Compensador de adelanto de fase

Un compensador de adelanto de fase tiene la siguiente expresión:

$$D(s) = K \frac{1 + Ts}{1 + \alpha Ts} = \frac{K}{\alpha} \frac{s + \frac{1}{T}}{s + \frac{1}{\alpha T}} = K_c \frac{s + a}{s + b}, \text{ con } \begin{cases} 0 < \alpha < 1 \\ a < b \end{cases} \quad (9.1)$$

La primera expresión es más útil cuando se trabaja en el diagrama de Bode, mientras que la segunda es preferible cuando se trabaja en el lugar de las raíces.

En los siguientes apartados se explicarán los procedimientos de ajuste de los tres parámetros del compensador. Para todos los métodos se usará siempre el mismo ejemplo de la Fig. 9.3.

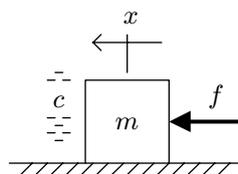


Figura 9.3: Sistema que se desea compensar

Se trata de un sistema mecánico en el que un actuador mueve una compuerta de masa m dentro de un medio con viscosidad c . El actuador debe colocar la compuerta siguiendo la referencia de posición que se le comande de tal forma que el tiempo de establecimiento sea menor o igual que 2 segundos y que el sobreimpulso máximo respecto la referencia sea del 20% o menor. La masa de la compuerta es de 0.25 kg y el coeficiente viscoso es 0.5 Ns/m.

La función de transferencia que relaciona la fuerza aplicada a la masa y su desplazamiento aparece en (9.2), sustituyendo también los valores numéricos.

$$G(s) = \frac{1}{ms^2 + cs} = \frac{4}{s(s+2)} \quad (9.2)$$

9.2.1. Ajuste por el lugar de las raíces

La introducción de un polo y un cero en el sistema puede beneficiar al régimen transitorio, es decir, puede hacer el sistema más rápido si el nuevo cero está más cerca del origen que el nuevo polo. En ese caso, las ramas del lugar de las raíces se alejan del eje imaginario del plano S . Por tanto, se puede elegir una ganancia para el sistema que deje los polos dominantes del sistema con un tiempo de establecimiento menor.

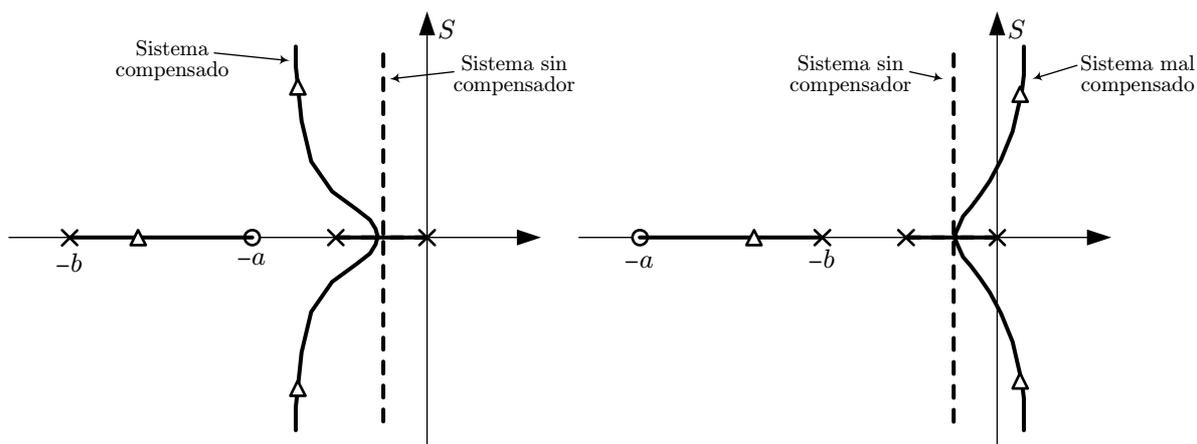


Figura 9.4: Efecto de la adición de un polo y un cero en el lugar de las raíces

Este efecto se muestra cualitativamente en la Fig. 9.4 y justifica la definición del compensador de adelanto de fase de la ecuación (9.1), donde se decía $a < b$.

Evidentemente, la localización de polo y del cero del compensador dependerá de cuánto se quiera alejar las ramas del lugar de las raíces del eje imaginario. Son las especificaciones que se deseen conseguir las que van a marcar la separación relativa del nuevo polo y cero. Resulta casi inmediato, a partir de los requerimientos de tiempo de establecimiento y sobreimpulso calcular los polos dominantes objeto del diseño.

$$M_p \leq 0.2 \Rightarrow \zeta \geq 0.45, \text{ se elige } \zeta = 0.5 \quad (9.3)$$

$$t_s \leq 2 \text{ segundos} \Rightarrow \omega_n \geq 4 \text{ rad/s}, \text{ se elige } \omega_n = 4 \text{ rad/s} \quad (9.4)$$

$$p_\Delta = -\zeta\omega_n \pm \omega_n\sqrt{1-\zeta^2}j = -2 \pm 3.46j \quad (9.5)$$

Con el valor de sobreimpulso se elige un valor para el amortiguamiento de los polos objetivo, ecuación (9.3), y con éste y el tiempo de establecimiento se calcula la frecuencia natural, ecuación (9.4). La posición final de los polos objetivo, ecuación (9.5), depende de las elecciones que haya realizado el ingeniero. Es difícil que dos ingenieros obtengan exactamente la misma solución numérica para un mismo problema. Por tanto, las soluciones que se proponen en estos apuntes no son las únicas que se pueden dar correctamente para cada problema.

Obtenida la localización de los polos objetivo, se aplica en ese punto la condición del ángulo. Esto se realiza precisamente porque se quiere garantizar que esos puntos pertenecen al nuevo lugar de las raíces del sistema. Evidentemente, la condición del ángulo se debe aplicar teniendo en cuenta el nuevo polo y cero que introduce el compensador, cuya posición todavía está por determinar, Fig. 9.5.

$$\sum_{i=1}^m \widehat{z_i p_\Delta} - \sum_{j=1}^n \widehat{p_j p_\Delta} = \theta_z - \theta_p - \theta_1 - \theta_2 = \phi_c - \theta_1 - \theta_2 = -180^\circ \quad (9.6)$$

En la ecuación (9.6) los ángulos vistos desde los polos o ceros de la planta son conocidos. La única incógnita es la diferencia de ángulos del cero y polo del compensador. A esa diferencia se le llamará ϕ_c

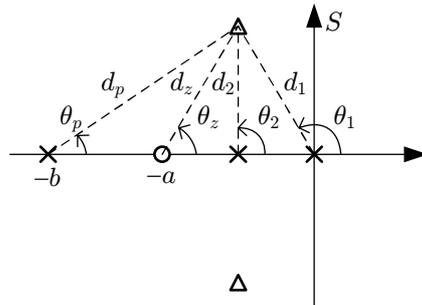


Figura 9.5: Ángulos y distancias vistas al polo objetivo

y es el ángulo con que el polo objetivo “ve” al polo y cero del compensador. Cualquier pareja polo-cero que el polo objetivo vea con el mismo ángulo ϕ_c es una posible solución al problema.

$$\phi_c = -180^\circ + \theta_1 + \theta_2 = -180^\circ + 120^\circ + 90^\circ = 30^\circ \quad (9.7)$$

En el caso concreto del ejemplo ϕ_c vale 30° . Siempre que se trate de un compensador de adelanto de fase este ángulo debe dar *positivo*, ya que el cero está más cercano al origen y verá al polo objetivo con mayor ángulo que el polo del compensador. Cualquier pareja polo-cero que sea vista por el polo objetivo con 30° es una hipotética solución al problema, Fig. 9.6.

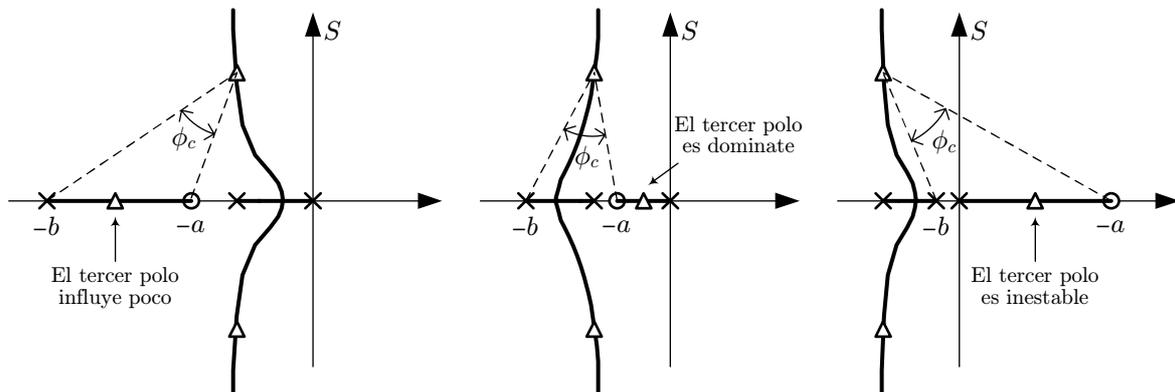


Figura 9.6: Posiciones del par polo-cero del compensador con ϕ_c constante

Sin embargo, la elección de la posición del par polo-cero no es absolutamente arbitraria. Efectivamente se puede conseguir que el lugar de las raíces pase por el polo objetivo, pero puede ser que en la configuración final del sistema, el polo dominante sea otro. Esta posibilidad hay que evitarla. Si se busca que el régimen transitorio lo caracterice la posición de los polos objetivo, éstos deben quedar como dominantes del sistema.

Incluso es posible colocar el par polo-cero en la zona del plano S de parte real positiva y cumplir el deseo de que el polo objetivo pertenezca al lugar de las raíces, pero dejar el tercer polo del sistema de forma que lo hace inestable.

En conclusión, la posición del polo y del cero del compensador es libre mientras el polo objetivo los vea con el ángulo ϕ_c adecuado y éstos queden como dominantes del sistema.

Dejando claro que el ingeniero puede elegir la posición del polo y del cero del compensador, cumpliendo lo que se ha enunciado en el párrafo anterior, en los siguientes apartados se explican algunos criterios que se han propuesto para su colocación. Dependiendo de la planta es posible que alguno de los criterios sea inviable o no consiga que el polo objetivo quede dominante en el sistema compensado. Siempre habrá que comprobar este último extremo, aunque sea de forma cualitativa.

En cuanto al cálculo de la ganancia del compensador, se debe aplicar la condición del módulo en el polo objetivo, Fig. 9.5, empleando lógicamente la posición del polo y el cero del compensador que se haya elegido.

$$K_c K_{la} = \frac{\prod_{j=1}^n \overline{p_j p_\Delta}}{\prod_{i=1}^m \overline{z_i p_\Delta}} = \frac{d_1 d_2 d_p}{d_z} \quad (9.8)$$

La única incógnita de la ecuación (9.8) es la ganancia K_c del compensador. En el ejemplo, la ganancia K_{la} de la planta en lazo abierto expresando los polos y ceros en monomios es 4, la distancia d_1 es 4, la

distancia d_2 es 3.46 y las distancias d_z y d_p se miden en el lugar de las raíces y dependen de la posición elegida para el polo y el cero del compensador.

Criterio de la bisectriz

La posición del polo y del cero del compensador se elige centrándolos sobre la bisectriz del arco que forma la recta que une el polo objetivo con el origen del plano S , y una recta horizontal que pasa por el polo objetivo, Fig. 9.7.

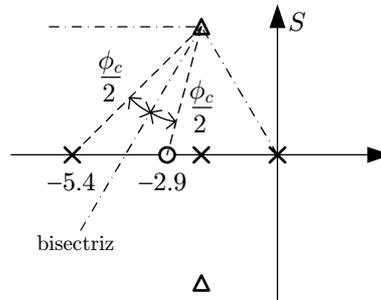


Figura 9.7: Criterio de la bisectriz

En el ejemplo propuesto, el polo del compensador queda aproximadamente en -5.4 y el cero en -2.9 . Al aplicar la condición del módulo en el polo objetivo, la ganancia del compensador resulta ser aproximadamente 4.7, por tanto, el compensador tiene la siguiente expresión:

$$D(s) = 4.7 \frac{s + 2.9}{s + 5.4} \quad (9.9)$$

Se puede observar en la Fig. 9.7 que en el caso del ejemplo, este criterio no hubiera sido recomendable si el ángulo ϕ_c hubiera sido mayor que 60° . Si así fuera, el cero del compensador quedaría entre los dos polos de la planta y el lugar de las raíces tendría un tercer polo real en lazo cerrado más dominante que los polos objetivo.

Criterio de anular un polo dominante de la planta

El cero del compensador se sitúa sobre un polo de la planta y el polo del compensador donde quede, manteniendo el ángulo ϕ_c . La única limitación de este método, es que nunca se debe anular un polo de la planta que esté en $s = 0$, es decir, disminuir el tipo del sistema. El resultado es un sistema compensado con igual número de polos que sin compensar: no aparece ningún polo nuevo.

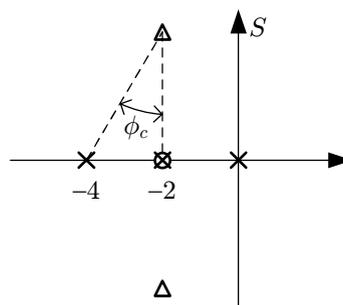


Figura 9.8: Criterio de anular el segundo polo dominante de la planta

En el ejemplo propuesto, el cero del compensador queda en -2 y el polo en -4 . Al aplicar la condición del módulo en el polo objetivo, la ganancia del compensador resulta ser 4, por tanto, el compensador tiene la siguiente expresión:

$$D(s) = 4 \frac{s + 2}{s + 4} \quad (9.10)$$

La limitación de uso de este criterio dependerá evidentemente de la posición del segundo polo dominante de la planta en lazo abierto. En el caso del ejemplo, este criterio no se puede emplear para un ángulo ϕ_c mayor que 90° .

Criterio del cero bajo el polo objetivo

El cero del compensador se sitúa justo debajo del polo objetivo y el polo donde quede, manteniendo el ángulo ϕ_c . En este ejemplo, el resultado es el mismo que en el apartado anterior, porque ha coincidido que el segundo polo dominante de la planta en lazo abierto queda debajo del polo objetivo. Sin embargo, lo habitual es que den resultados distintos.

Este método se puede emplear sólo en el caso de que queden a la derecha del cero del compensador al menos dos polos de la planta en lazo abierto. Se puede observar en la Fig. 9.7 que el ejemplo se encuentra en el caso límite de uso, ya que si el segundo polo de la planta hubiese estado más a la izquierda, el tercer polo en lazo cerrado del sistema quedaría más dominante que los polos objetivo.

Si se cumple lo dicho en el párrafo anterior, para cualquier planta que se desee compensar, este criterio se puede emplear siempre que el ángulo ϕ_c sea igual o menor que 90° .

Criterio de un compensador proporcional-derivativo

El polo del compensador se sitúa en el $-\infty$ del plano S y el cero del compensador donde quede, manteniendo el ángulo ϕ_c . El compensador, por tanto, está compuesto por un único cero y ningún polo, Fig. 9.9.

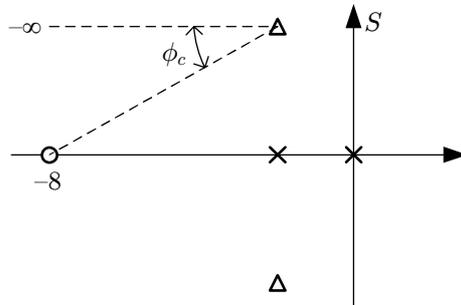


Figura 9.9: Criterio de un compensador proporcional-derivativo

El ángulo ϕ_c coincide exactamente con el ángulo del que ve el cero al polo objetivo. En el ejemplo propuesto, el cero del compensador queda en -8 y la ganancia resulta ser 0.5 , por tanto, el compensador tiene la siguiente expresión:

$$D(s) = 0.5(s + 8) \quad (9.11)$$

Este criterio, como la anulación de un polo de la planta en lazo abierto, no añade ningún polo nuevo en el sistema en lazo cerrado. El nombre de compensador proporcional-derivativo es coherente si se calcula en el dominio del tiempo cuál es la actuación sobre la planta $u(t)$ en función del error $e(t)$, ecuación (9.12), donde se observa que hay una componente proporcional al error y otra proporcional a la derivada del error.

$$U(s) = K_c(s + a)E(s) = K_c s E(s) + a K_c E(s) \xrightarrow{\mathcal{L}^{-1}} u(t) = K_c \frac{de(t)}{dt} + a K_c e(t) \quad (9.12)$$

Un inconveniente de los compensadores PD es que amplifican el ruido de alta frecuencia, Fig. 9.10. Por este motivo, en la práctica se desaconseja elegir este tipo de compensadores a no ser que se pida de forma explícita.

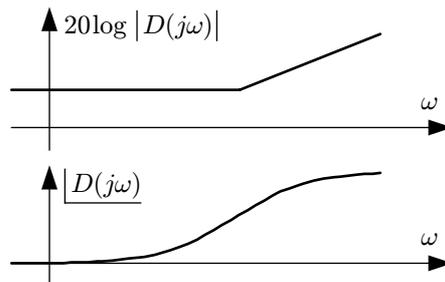


Figura 9.10: Diagrama de Bode de un compensador proporcional-derivativo

Comparación de la respuesta temporal de los distintos criterios

Resulta interesante comparar la respuesta temporal que consiguen los distintos criterios que se han enunciado. En la Fig. 9.11 se muestra la respuesta ante una entrada escalón unidad del sistema en lazo cerrado con los distintos compensadores que se han ido calculando.

Evidentemente, todos los compensadores consiguen cumplir las especificaciones de partida: tiempo de establecimiento menor que 2 segundos y sobreimpulso del orden del 20%. El compensador PD es fácilmente reconocible en la Fig. 9.11 porque la pendiente inicial no es nula. El compensador del criterio de la bisectriz es un poco más rápido que el que anula el polo de la planta porque tiene la influencia del cero del compensador —también por eso tiene un poco más de sobreimpulso—.

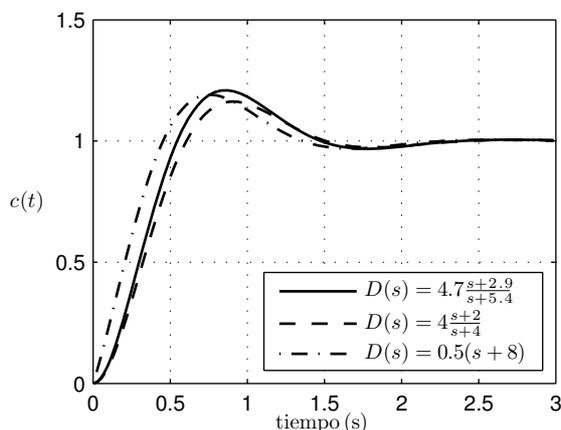


Figura 9.11: Respuesta temporal del sistema compensado ante una entrada escalón unidad

Antes de continuar con el diseño de compensadores de adelanto de fase por métodos frecuencias, conviene señalar que el método del lugar de las raíces que se ha descrito no permite el cumplimiento de especificaciones de error en régimen permanente. Los parámetros K_c , a y b del compensador se emplean para localizar los polos del sistema en lazo cerrado. Por tanto sólo puede lograr especificaciones de régimen transitorio.

Ha quedado patente en el ejemplo que se ha empleado que existe más de un compensador válido para unas especificaciones dadas. Sin embargo, a la hora de resolver un problema hay que decir explícitamente qué elecciones se están tomando en cada momento de forma justificada.

Se puede diseñar el compensador sin llegar a dibujar exactamente el lugar de las raíces del sistema, compensado o no, porque todas las operaciones se hacen con la posición de los polos y ceros de la planta y compensador en lazo abierto. El único punto que se conoce de forma exacta del lugar de las raíces puede ser el polo objetivo. Sin embargo, es muy conveniente al menos imaginar cómo va a quedar el lugar de las raíces del sistema compensado, de forma cualitativa.

Siempre que se pretendan mejorar el régimen transitorio, en el sentido de hacer el sistema más rápido de lo que de suyo es, el ángulo ϕ_c va a dar positivo. Si da un ángulo pequeño, por ejemplo menos de 10° , significa que el lugar de las raíces del sistema sin compensador pasa bastante cerca de los polos objetivo. En este caso, es suficiente con diseñar un compensador puramente proporcional que deje los polos en lazo cerrado lo más cerca posible de los polos objetivo.

Si ϕ_c da un ángulo muy grande, por ejemplo mayor que 70° , significa que el polo objetivo queda muy lejos del lugar de las raíces del sistema sin compensador. Quizá se pueden modificar las elecciones realizadas para el cálculo de los polos objetivo para que de un ángulo algo menor. Si no es así, no es recomendable elegir el compensador de adelanto de fase para cumplir las especificaciones exigidas. Habría que ir a otro tipo de controlador o afirmar que no se pueden cumplir dichas especificaciones.

9.2.2. Ajuste por el diagrama de Bode

El diseño de un compensador de adelanto de fase con el diagrama de Bode es muy sencillo si se conoce bien la forma que posee el compensador en dicho diagrama, Fig. 9.12. Como ya se ha indicado, cuando se trabaje con el diagrama de Bode se empleará la expresión del compensador (9.1) con constantes de tiempo.

Las ganancias comienzan y acaban con pendiente nula. Siempre el aporte de ganancia a altas frecuencias es mayor que a bajas frecuencias. Sin embargo, dependiendo de K y α ese aporte puede ser finalmente positivo, negativo o nulo para alguna de esas frecuencias.

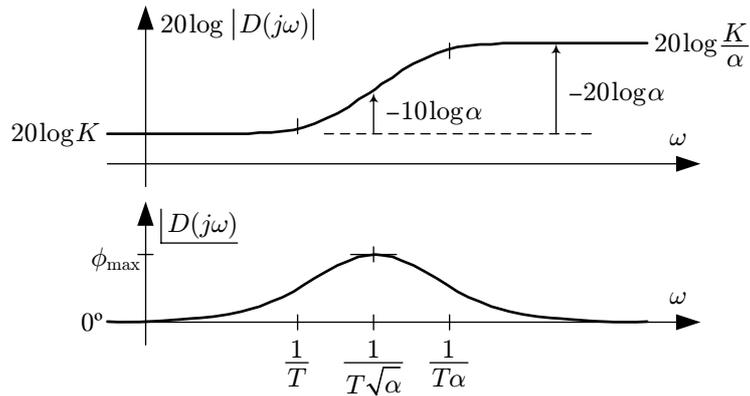


Figura 9.12: Diagrama de Bode de un compensador de adelanto de fase

El compensador no aporta fase para bajas y elevadas frecuencias. Sin embargo, entre el cero y el polo, aporta siempre fase positiva. Por tanto, añadirá fase a la planta en ese rango de frecuencias. El ángulo $\phi_{\text{máx}}$ de adelanto máximo, depende de lo separados que están el cero y el polo del compensador. Depende por tanto exclusivamente del parámetro α según la expresión:

$$\sin \phi_{\text{máx}} = \frac{1 - \alpha}{1 + \alpha} \quad (9.13)$$

Evidentemente, como máximo el valor del adelanto máximo será de 90° en el caso de que estén muy alejados el cero y el polo del compensador. La frecuencia $\omega_{\text{máx}}$ a la que se da el adelanto máximo es la media geométrica de la posición del cero y el polo. En la Fig. 9.12 también se aparece el valor del aumento de ganancias que se produce en la frecuencia de adelanto de fases máximo.

Traducción de las especificaciones al dominio de la frecuencia

Una vez que se conoce la forma del compensador de adelanto de fase, hay que saber cuáles van a ser los objetivos del diseño en el dominio de la frecuencia. Para el caso del lugar es directa la traducción de sobreimpulso máximo de la salida en amortiguamiento de los polos dominantes en lazo cerrado, el tiempo de establecimiento en atenuación de dichos polos, etc. La traducción que se recomienda en el dominio de la frecuencia es la siguiente:

- El margen de fase del sistema en lazo abierto en grados es aproximadamente igual a 100 veces el amortiguamiento de los polos dominantes en lazo cerrado.

$$\text{MF} \approx 100\zeta \quad (9.14)$$

- La frecuencia de cruce de ganancias del sistema en lazo abierto es aproximadamente igual a la frecuencia natural de los polos dominantes en lazo cerrado.
- La frecuencia de cruce de ganancias del sistema en lazo abierto es aproximadamente igual al ancho de banda del sistema.

También se puede trasladar fácilmente al dominio de la frecuencia una especificación de régimen permanente. El error en régimen permanente es igual a la inversa de los coeficientes de error. Por tanto, si se observa la definición del coeficiente de error de un sistema de tipo 0 compensado:

$$K_p = \lim_{s \rightarrow 0} K \frac{1 + Ts}{1 + \alpha Ts} G(s) = KG(0) \quad (9.15)$$

Es decir, el coeficiente de error del sistema compensado es K veces mayor que el del sistema sin compensador. El coeficiente de error del sistema sin compensador, $G(0)$, puede verse en el diagrama de Bode directamente como la magnitud de las ganancias a bajas frecuencias. Por tanto:

- Para reducir el error en régimen permanente de un sistema compensado en lazo cerrado, se debe aumentar la magnitud de las ganancias a bajas frecuencias del sistema en lazo abierto.

Como el efecto de la ganancia K del compensador es elevar todo el diagrama de ganancias la magnitud de K en decibelios. Se puede aprovechar ese aumento para disminuir el error en régimen permanente.

En el diagrama de Bode se puede también deducir si el error en régimen permanente será nulo o no en función del tipo del sistema. Si el diagrama de ganancias comienza horizontal para bajas frecuencias, el sistema es de tipo 0. Si a bajas frecuencias comienza con -20 dB por década, el sistema es de tipo I y tendrá error nulo ante entrada escalón. Si comienza con -40 dB por década, el sistema es de tipo II y tendrá error nulo ante entrada escalón y rampa. Y así sucesivamente.

Ejemplo de diseño

Se va a diseñar un compensador de adelanto de fase para el mismo ejemplo que se utilizó en el método del lugar de las raíces. Lo primero que se hace es traducir las especificaciones al dominio de la frecuencia:

$$M_p \leq 0.2 \Rightarrow \zeta \geq 0.45, \text{ se elige } \zeta = 0.5 \Rightarrow \text{MF} \approx 50^\circ \quad (9.16)$$

$$t_s \leq 2 \text{ segundos} \Rightarrow \omega_n \geq 4 \text{ rad/s}, \text{ se elige } \omega_n = 4 \text{ rad/s} \Rightarrow \omega_{cg} \approx 4 \text{ rad/s} \quad (9.17)$$

A continuación se representa el diagrama de Bode de la planta sin el compensador:

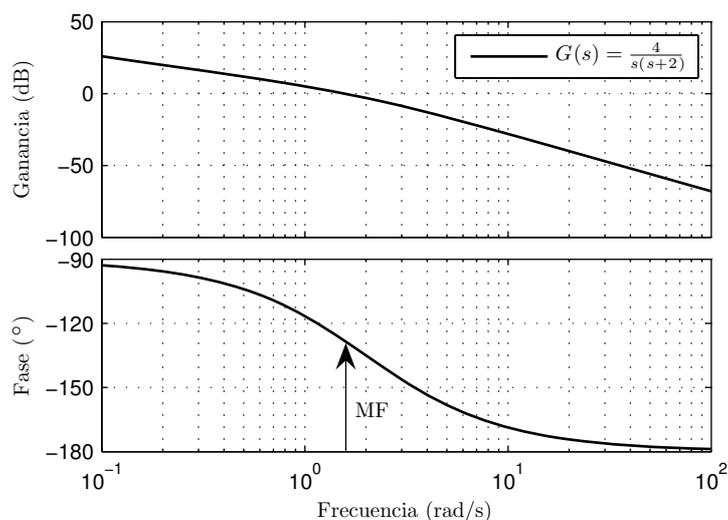


Figura 9.13: Diagrama de Bode de la planta en lazo abierto sin compensador

Se observa que actualmente el MF es aproximadamente 52° y la frecuencia de cruce de ganancias es de 1.6 rad/s. Por tanto, el sistema sin compensador en lazo cerrado aproximadamente cumple la especificación de amortiguamiento, y por tanto de sobreimpulso máximo, pero no cumple la de tiempo de establecimiento.

Se podría introducir un compensador proporcional para aumentar la frecuencia de cruce de ganancias y cumplir la especificación de tiempo de establecimiento, pero el margen de fase disminuiría, el amortiguamiento también disminuiría y el sobreimpulso aumentaría. Para cumplir las dos especificaciones a la vez no queda más remedio que emplear un compensador de adelanto de fase que haga las dos cosas: aumentar la frecuencia de cruce de ganancias y conservar el margen de fase.

Lo más difícil del diseño es decidir cuál de los parámetros se ajusta primero. Con las especificaciones del ejemplo, el dato clave es que al sistema compensado se le va a obligar a que tenga su frecuencia de cruce de ganancias en 4 rad/s. A esa frecuencia el diagrama de Bode de la planta sólo tiene 28° . Se puede introducir el compensador cuyo adelanto máximo sea de 24° y se dé en 4 rad/s.

$$\sin 24^\circ = \frac{1 - \alpha}{1 + \alpha} \Rightarrow \alpha = 0.4217 \quad (9.18)$$

$$\frac{1}{T\sqrt{\alpha}} = 4 \text{ rad/s} \Rightarrow T = 0.385 \text{ s} \quad (9.19)$$

De esa forma, si la frecuencia de cruce de ganancias es realmente 4 rad/s, se conseguirá un margen de fase de 52° , porque se ha diseñado un compensador que añade justo la fase necesaria para cumplirlo. La ganancia del compensador sube o baja el diagrama de ganancias y se elegirá la que sea necesaria para conseguir 0 dB en 4 rad/s. La ganancia necesaria no se mide directamente del diagrama de Bode de ganancias de la planta, porque el compensador de adelanto modifica tanto las fases como las ganancias. En concreto, en la frecuencia de adelanto máximo, 4 rad/s, introduce una subida de ganancias de:

$$-10 \log \alpha = 3.75 \text{ dB} \quad (9.20)$$

Por tanto, si la planta necesitaba en 4 rad/s una subida de 12 dB y el compensador de adelanto ya ha introducido una subida de 3.75 dB, la ganancia sólo debe añadir 8.25 dB. Es decir:

$$20 \log K = 8.25 \text{ dB} \Rightarrow K = 2.58 \quad (9.21)$$

Con todo lo dicho ya están fijados los tres parámetros del compensador, que tiene la forma:

$$D(s) = 2.58 \frac{1 + 0.385s}{1 + 0.162s} = 6.11 \frac{s + 2.59}{s + 6.16} \quad (9.22)$$

Se observa que es similar, pero no exactamente igual, a los que se diseñaron en el lugar de las raíces.

Ejemplo de diseño con especificación de error

Se supone que las especificaciones de la planta del ejemplo del apartado anterior son obtener un MF de 50° y un coeficiente de error de velocidad de 20 s^{-1} . La forma de diseñar el compensador de adelanto de fase es completamente distinta. Primero se utiliza la ganancia K del compensador para lograr la especificación de error en régimen permanente:

$$K_v = \lim_{s \rightarrow 0} sK \frac{1 + Ts}{1 + \alpha Ts} \frac{4}{s(s+2)} = 2K = 10 \Rightarrow K = 10 \quad (9.23)$$

Se decide empezar por la ganancia K porque es el único elemento del compensador que interviene en la mejora del error, pero además porque modifica el MF de la planta. En la Fig. 9.14 se muestra cómo al introducir la ganancia de 10, el margen de fase del sistema disminuye de 52° hasta 18° . El cero y el polo del compensador colocados en torno a la nueva frecuencia de cruce de ganancias, aproximadamente 6 rad/s, conseguirán aumentar el margen de fase hasta los 50° deseados.

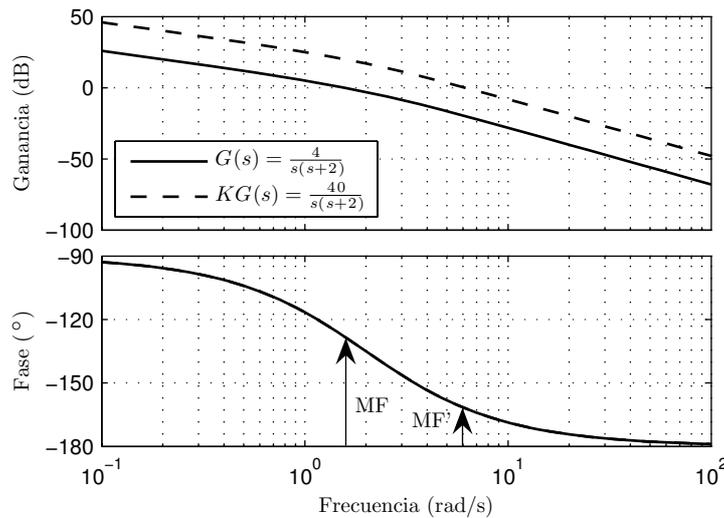


Figura 9.14: Diagrama de Bode de la planta en lazo abierto sin compensador y con ganancia

El adelanto de fases máximo del compensador debe ser unos 32° si se quieren conseguir 50° de MF. Sin embargo, el compensador también introduce una pequeña subida de ganancias, por lo que la frecuencia de cruce de ganancias definitiva se va a desplazar un poco a la derecha. Por ese motivo, en lugar de diseñar un compensador que adelante exactamente lo necesario para dar el MF de la especificación se suelen sumar entre 5° y 12° de más. El margen se va a medir al final en un punto a una frecuencia un poco más elevada que donde está MF'. Se elige por tanto, diseñar un controlador cuyo adelanto máximo sea un poco más de 32° , por ejemplo 38° .

$$\sin 38^\circ = \frac{1 - \alpha}{1 + \alpha} \Rightarrow \alpha = 0.24 \quad (9.24)$$

Hecha esta elección, ya se puede saber cuánto va a subir en dB el diagrama de Bode para la frecuencia de adelanto máximo:

$$-10 \log \alpha = 6.2 \text{ dB} \quad (9.25)$$

Por tanto, se elige para la frecuencia de adelanto máximo el punto del diagrama de Bode con ganancia que tenga -6.2 dB , para que al añadir el polo y el cero el adelanto máximo se dé en la frecuencia de

cruce de ganancias definitiva. En la Fig. 9.14 se pueden medir aproximadamente -6.2 dB en la planta con ganancia en la frecuencia de 9 rad/s. Por tanto:

$$\frac{1}{T\sqrt{\alpha}} = 9 \text{ rad/s} \Rightarrow T = 0.227 \text{ s} \quad (9.26)$$

Con esto ya están definidos los tres parámetros del compensador de adelanto de fase que queda así:

$$D(s) = 10 \frac{1 + 0.227s}{1 + 0.0544s} = 41.66 \frac{s + 4.4}{s + 18.36} \quad (9.27)$$

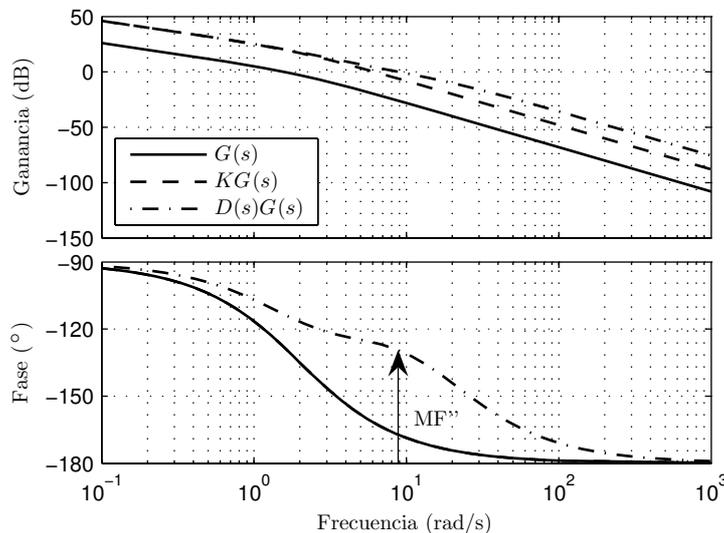


Figura 9.15: Diagrama de Bode de la planta en lazo abierto sin y con compensador

Como las especificaciones en este apartado son distintas que en el apartado anterior, la expresión del compensador es bastante diferente. En la Fig. 9.15 se muestra el diagrama de Bode del sistema en lazo abierto, el producto $D(s)G(s)$, y se observa cómo efectivamente al final se consigue un margen de fase aproximadamente de 50° .

Como recomendación final para el diseño, no es conveniente implementar un compensador de adelanto de fase que en frecuencias posea un adelanto de fase máximo $\phi_{\text{máx}}$ superior a 60° . En el caso de que sea necesario añadir tanta fase se pueden encontrar soluciones más satisfactorias, por ejemplo con compensadores de adelanto-retraso.

9.3. Compensador de retraso de fase

Un compensador de retraso de fase tiene la siguiente forma:

$$D(s) = K \frac{1 + Ts}{1 + \beta Ts} = \frac{K}{\beta} \frac{s + \frac{1}{T}}{s + \frac{1}{\beta T}} = K_c \frac{s + a}{s + b}, \text{ con } \begin{cases} \beta > 1 \\ a > b \end{cases} \quad (9.28)$$

La primera expresión es más útil cuando se trabaja en el diagrama de Bode, mientras que la última es preferible cuando se trabaja en el lugar de las raíces.

9.3.1. Ajuste por el diagrama de Bode

Para el diseño de un compensador de retraso de fase con el diagrama de Bode se debe estudiar en primer lugar qué forma adopta dicho compensador en el diagrama, Fig. 9.16.

Las ganancias comienzan y acaban con pendiente nula. Siempre el aporte de ganancia a altas frecuencias es menor que a bajas frecuencias. Dependiendo de K y β ese aporte puede ser finalmente positivo, negativo o nulo para alguna de esas frecuencias. Aunque en los compensadores de retraso tiene menos interés, la fase de retraso máximo es:

$$\sin \phi_{\text{mín}} = \frac{1 - \beta}{1 + \beta} \quad (9.29)$$

El ángulo $\phi_{\text{mín}}$ depende exclusivamente de β y como mucho puede ser de -90° cuando el polo y el cero están muy separados.

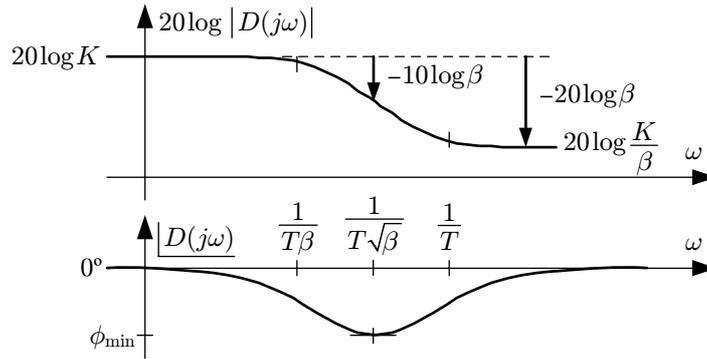


Figura 9.16: Diagrama de Bode de un compensador de retraso de fase

Compensador de retraso que mejora el error

Un compensador de retraso de fase puede utilizarse para disminuir el error en régimen permanente sin empeorar el régimen transitorio del sistema. Haciendo K igual a β , ecuación (9.30), la ganancia y fase del compensador valen ambos 0 para altas frecuencias y por tanto no se modifica el diagrama de Bode de la planta, Fig. 9.17. Sin embargo, a bajas frecuencias se produce un aumento de ganancia, por lo que aumenta el coeficiente de error del sistema y por tanto disminuye el error del mismo.

$$D(s) = \beta \frac{1 + Ts}{1 + \beta Ts} = \frac{s + \frac{1}{T}}{s + \frac{1}{\beta T}} = \frac{s + a}{s + b}, \text{ con } \begin{cases} \beta > 1 \\ a > b \end{cases} \quad (9.30)$$

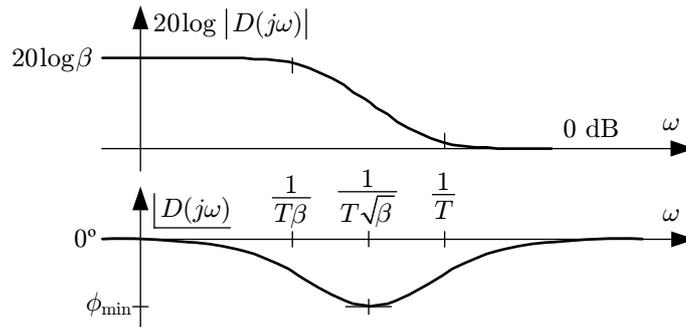


Figura 9.17: Diagrama de Bode de un compensador de retraso de fase que disminuye el error

Si se emplea el compensador de retraso de fase exclusivamente para disminuir el error en régimen transitorio, lo único que hay que hacer es: calcular β con la expresión (9.31) del coeficiente de error, igualar la ganancia K a β para que el compensador no modifique el Bode a altas frecuencias y elegir T de tal forma que el cero del compensador quede entre una década y una octava antes de la frecuencia de cruce de ganancias del sistema. Esto último se hace para asegurar precisamente que el compensador no modifica el Bode en la frecuencia que rige el régimen transitorio.

$$K_p = \lim_{s \rightarrow 0} \beta \frac{1 + Ts}{1 + \beta Ts} G(s) = \beta G(0) \quad (9.31)$$

Compensador de retraso que anula el error o compensador proporcional-integral

Se trata de la misma idea que en el apartado anterior, pero en lugar de disminuir el error del sistema, se anula. Para ello se incrementa el tipo del sistema en una unidad, haciendo que el polo del compensador se sitúe en 0 rad/s en el diagrama de Bode, o lo que es lo mismo, en el origen del plano S .

$$D(s) = \frac{1 + Ts}{Ts} = \frac{s + \frac{1}{T}}{s} = \frac{s + a}{s} \quad (9.32)$$

Este tipo de compensador se llama proporcional-integral porque en el dominio del tiempo la actuación en la planta es la suma de una parte proporcional al error y otra proporcional a la integral del error.

$$U(s) = \frac{s + a}{s} E(s) = E(s) + a \frac{E(s)}{s} \xrightarrow{\mathcal{L}^{-1}} u(t) = e(t) + a \int_0^t e(\tau) d\tau \quad (9.33)$$

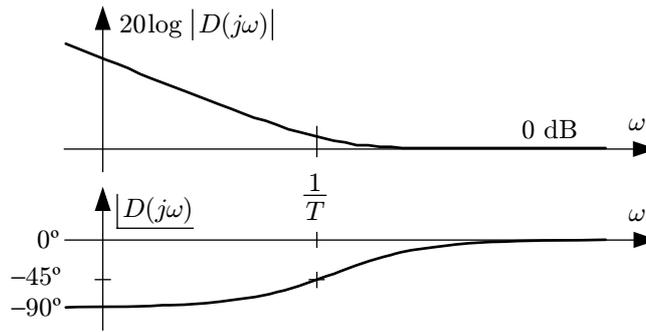


Figura 9.18: Diagrama de Bode de un compensador de retraso de fase proporcional-integral

El método de diseño es todavía más sencillo que el caso anterior porque sólo haya que elegir la constante de tiempo T del cero del compensador de forma que éste quede entre una década y una octava antes de la frecuencia de cruce de ganancias de la planta cuyo transitorio no se quiere modificar.

Compensador de retraso que mejora el error y el transitorio

Los compensadores de retraso que se han explicado en los apartados anteriores pueden incluir una mejora en las especificaciones de régimen transitorio. Lo que hay que hacer es añadir una ganancia K que modifique el Bode de la planta antes de incorporar los compensadores de retraso de los apartados anteriores.

Con la ganancia K se puede modificar la frecuencia de cruce de ganancias para conseguir una de estas dos cosas: un ancho de banda determinado —que es lo mismo que un tiempo de establecimiento determinado— o un margen de fase determinado —que es lo mismo que un amortiguamiento determinado—. Pero con una ganancia no es posible conseguir las dos cosas a la vez.

Si, como se acaba de decir, se pueden cumplir una especificación de error y una especificación de transitorio, entonces el ejemplo usado para el apartado de compensador de adelanto de fase se puede solucionar tanto con un compensador de adelanto de fase como con uno de retraso.

Se va resolver aquí ese mismo ejemplo con un compensador de retraso de fase. Las especificaciones eran obtener un margen de fase de 50° y un coeficiente de error de velocidad de 20 s^{-1} . El diagrama de Bode de la planta daba un margen de fase de 52° , por lo que se puede introducir una ganancia muy poco menor que la unidad para conseguir exactamente 50° de margen de fase. La diferencia en este caso es tan pequeña, que se ha optado por no introducir ninguna ganancia K , Fig. 9.19.

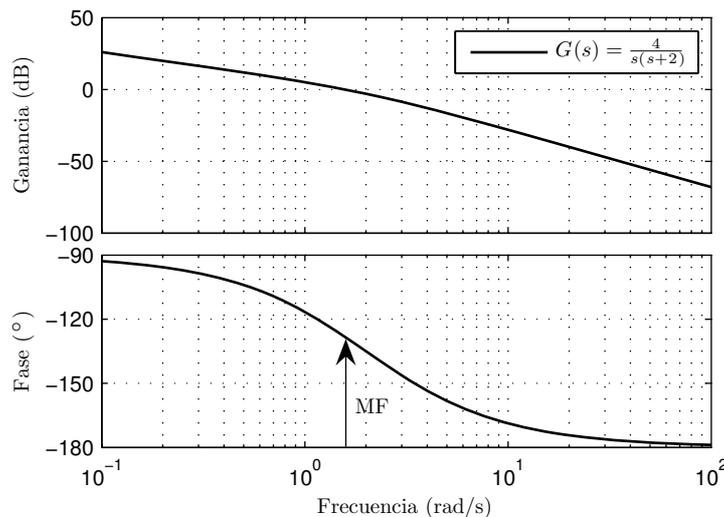


Figura 9.19: Diagrama de Bode de la planta en lazo abierto sin compensador

El compensador de retraso de fase que mejora sólo el error debe cumplir:

$$K_v = \lim_{s \rightarrow 0} s\beta \frac{1 + Ts}{1 + \beta Ts} \frac{4}{s(s+2)} = 2\beta = 10 \Rightarrow \beta = 10 \quad (9.34)$$

En cuanto a la posición del cero de cero del compensador, se propuso que estuviera entre una década y una octava antes que la frecuencia de cruce de ganancias. En la ecuación (9.35) se ha elegido colocar el cero una década antes.

$$\frac{1}{T} = \frac{1.6}{10} \text{ rad/s} \Rightarrow T = 6.25 \text{ s} \quad (9.35)$$

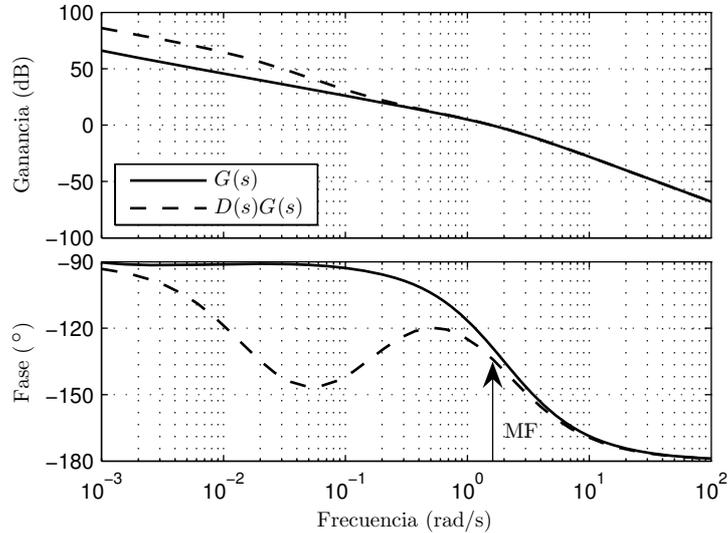


Figura 9.20: Diagrama de Bode de la planta con compensador de retraso de fase

En la Fig. 9.20 se muestra el diagrama de Bode del sistema en lazo abierto compensado. La expresión final del compensador de retraso de fase que cumple las especificaciones dadas es:

$$D(s) = 10 \frac{1 + 6.25s}{1 + 62.5s} = \frac{s + 0.16}{s + 0.016} \quad (9.36)$$

En la Fig. 9.21 se compara el sistema compensado de las dos maneras, con el compensador de retraso de fase y con el compensador de adelanto de fase. Como era de esperar, el sistema compensado con el de adelanto es más rápido porque su frecuencia de cruce de ganancias es mucho mayor. Sin embargo, como el margen de fase es aproximadamente igual en los dos casos, el amortiguamiento —y por tanto el sobreimpulso máximo— son similares.

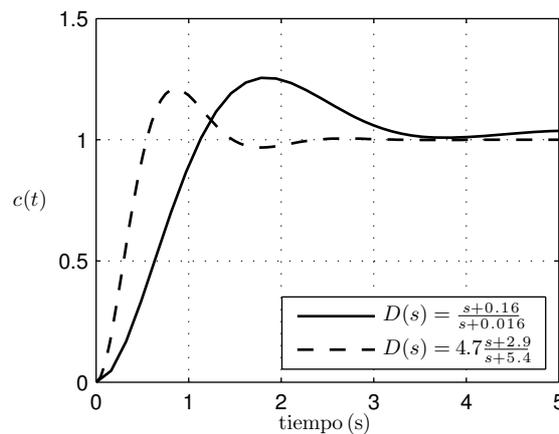


Figura 9.21: Respuesta temporal del sistema con compensador de retraso y de adelanto

9.3.2. Ajuste por el lugar de las raíces

En el lugar de las raíces se diseñan sólo compensadores de retraso de fase que mejoren el error en régimen permanente pero no modifiquen el régimen transitorio del sistema. Por tanto, la expresión del

compensador siempre tendrá K igual a β , es decir, en monomios será un polo y un cero sin ganancia aparente como ya se demostró en la ecuación (9.30).

$$D(s) = \frac{s+a}{s+b}, \text{ con } a > b \quad (9.37)$$

La única forma de que un nuevo polo y cero en el lugar de las raíces no modifiquen la localización de los polos objetivo que caracterizan el régimen transitorio y, a la vez, puedan mejorar sustancialmente el régimen permanente del sistema, es que ambos estén muy cerca del origen.

Un polo y un cero exactamente en un mismo lugar se anulan, pero si están muy cerca del uno del otro, a y b muy cercanos, su influencia es casi despreciable. Cualquier punto del plano S suficientemente alejado de ellos los verá con el mismo ángulo y distancia, por lo que no cambiará el lugar de las raíces, es decir, el hecho de que pase o no por ese punto.

Si ese polo y cero muy próximos entre sí, están además muy lejos del origen del plano S , tampoco pueden modificar la expresión del error en régimen permanente, porque el cociente $\frac{a}{b}$ será muy próximo a la unidad.

$$K_p = \lim_{s \rightarrow 0} \frac{s+a}{s+b} G(s) = \frac{a}{b} G(0) \quad (9.38)$$

Sin embargo, si el polo y el cero están muy próximos entre sí, y a la vez están muy cerca del origen, es posible mejorar cuanto se desee el error en régimen permanente sin modificar el lugar de las raíces del sistema y, por tanto, conservando los mismo polos objetivo que sin ellos. Precisamente porque se quiere mejorar el error en régimen permanente se debe escoger $a > b$.

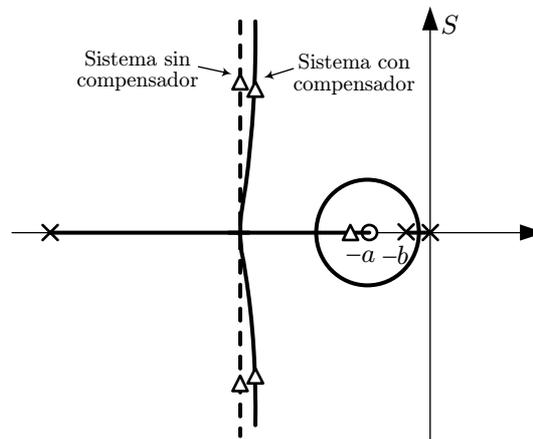


Figura 9.22: Lugar de las raíces al introducir un compensador de retraso

Como se observa en el ejemplo de sistema de segundo orden de la Fig. 9.22, aparece un tercer polo en lazo cerrado debido a la introducción del compensador, pero apenas influirá en la respuesta transitoria del sistema porque está muy cercano al cero del compensador y sus actuaciones se cancelan prácticamente.

Para diseñar el compensador con el lugar de las raíces, lo único que hay que hacer es conocer cuál debe ser la separación relativa entre el polo y el cero del compensador para que con la ecuación (9.38) —u otra equivalente dependiendo del tipo de error— se cumpla la especificación deseada, y posteriormente colocar el cero del compensador, que es el más cercano al polo objetivo, entre tres y diez veces más cercano al origen.

Si el compensador suficientemente alejado del polo objetivo, éste último variará su posición muy poco, por lo que no se recomienda introducir ninguna ganancia extra que modifique de nuevo su posición, aunque algunos autores incluyen este último paso.

9.4. Compensador de adelanto-retraso

Un compensador de adelanto-retraso es el producto de uno de adelanto y uno de retraso:

$$D(s) = K \frac{1+Ts}{1+\alpha Ts} \frac{1+T's}{1+\beta T's}, \text{ con } \begin{cases} 0 < \alpha < 1 \\ \beta > 1 \end{cases} \quad (9.39)$$

Los compensadores de adelanto-retraso son muy útiles, por ejemplo, cuando se deben cumplir tres especificaciones en un sistema. Hasta ahora en los ejemplos que se han resuelto sólo existían dos especificaciones, sin embargo, puede darse el caso que se pidan dos especificaciones de régimen transitorio y

otra de error que hagan necesario contar con más grados de libertad que los que ofrecen por separado los compensadores de adelanto de fase y los de retraso.

En este caso, lo que se recomienda es diseñar primero la parte de adelanto del compensador que haga cumplir las dos especificaciones de régimen transitorio y después la parte de retraso que no modifique las altas frecuencias y mejore el error en régimen permanente. Aquí se recomienda situar el cero del compensador de retraso una década antes de frecuencia de cruce ganancias que resulte después de ajustar el compensador de adelanto.

También es posible que un sistema con sólo dos especificaciones lleve al ingeniero a elegir un compensador de adelanto-retraso. Por ejemplo, cuando resulte que un compensador de adelanto de fase debe tener un adelanto de fase máximo $\phi_{\text{máx}}$ mayor que los 60° recomendados o incluso mayor que su límite superior de 90° .

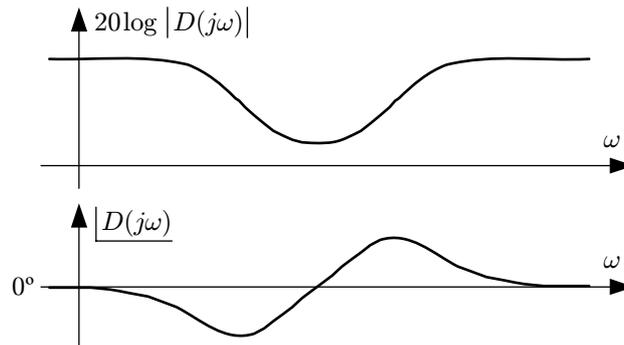


Figura 9.23: Diagrama de Bode de un compensador de adelanto-retraso

En cualquiera de los dos casos mencionados, el compensador de retraso se coloca a bajas frecuencias y el de adelanto a más elevadas frecuencias, Fig. 9.23. Nunca se diseñan compensadores de adelanto-retraso en los que esta situación esté invertida.

9.5. Ejercicios propuestos

- **Ejercicio 1:** Diseñar un compensador que, en lazo cerrado, consiga un coeficiente K_v de 5 s^{-1} para la siguiente planta:

$$G(s) = \frac{1}{s(s+1)(s+2)} \quad (9.40)$$

- **Ejercicio 2:** Diseñar un compensador que, en lazo cerrado, consiga un coeficiente K_v mayor o igual que 100 s^{-1} y un margen de fase mayor o igual que 30° para la siguiente planta:

$$G(s) = \frac{4}{s(1+0.1s)(1+0.2s)} \quad (9.41)$$

- **Ejercicio 3:** Diseñar un compensador que, en lazo cerrado, consiga un amortiguamiento de 0.5, una frecuencia natural de 5 rad/s y coeficiente K_v de 80 s^{-1} para la siguiente planta:

$$G(s) = \frac{4}{s(s+0.5)} \quad (9.42)$$

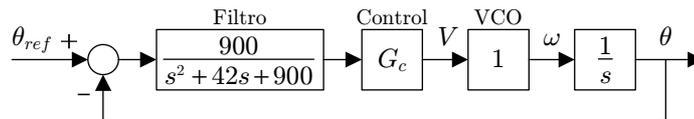
- **Ejercicio 4:** Un PLL (*Phase Locked Loop*) es un circuito electrónico utilizado para sincronizar la frecuencia y la fase de los relojes de dos sistemas que desean intercambiar información. Es necesario utilizarlo en sistemas digitales que envíen la onda modulada. Su utilización, por tanto, es amplísima: teléfonos móviles, acceso en banda ancha, etc. Técnicamente el sistema no es lineal, pero bajo determinadas hipótesis se puede considerar que es un sistema lineal como el que muestra el diagrama de bloques:

- En primer lugar se estima la diferencia entre el ángulo de la referencia.
- En el cálculo de esa diferencia se introduce términos de altas frecuencias que se filtran con un filtro pasa-baja.
- La salida del filtro pasa-baja es la entrada del controlador que se pide diseñar.

- La salida de este controlador, es la entrada de un oscilador controlado por voltaje (VCO), es decir, un oscilador que genera una señal de frecuencia proporcional al voltaje que se aplica sobre él. Aquí se ha supuesto una ganancia unidad.
- El ángulo de la senoide es directamente la integral de la frecuencia.

Se pide diseñar un compensador que cumpla las siguientes propiedades:

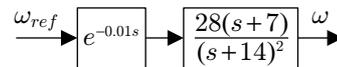
- Error nulo ante una entrada rampa (esto es equivalente a, por ejemplo, sincronizar los relojes de un teléfono móvil en un coche y la estación base).
- Sobreimpulso menor del 10 %.
- Ancho de banda de unos 20 rad/s. (Nota: los valores habituales son entre 50 y 100 Hz, pero se ha puesto éste para facilitar los cálculos).



Nota: Aunque el diagrama de Bode que se da puede ser útil, no es estrictamente necesario para la resolución de este ejercicio.

Solución: $G_c(s) = 10.3 \frac{1+0.107s}{1+0.0233s} \frac{1+1.07s}{1.07s}$

- **Ejercicio 5:** Se quiere controlar de forma rápida y precisa la *posición* de un robot que dispone de un controlador para la velocidad angular independiente en cada una de las articulaciones. El siguiente diagrama de bloques muestra un modelo estimado de una articulación con su control de velocidad:



Se requiere diseñar un controlador de posición adecuado para dicha articulación en los siguientes casos:

- Controlador proporcional.
- Compensador de adelanto de fase. Por consideraciones de diseño, el máximo avance de fase del compensador es de 45°.

Nota: Se recomienda dibujar primero el diagrama de bloques del sistema con el controlador de posición, teniendo en cuenta que la entrada es la referencia de posición deseada y la salida es la posición θ de la articulación del robot (no su velocidad ω).

Soluciones:

- $G_c(s) = 10$
- $G_c(s) = 20.6 \frac{1+0.069s}{1+0.0117s} = 121 \frac{s+5.88}{s+85.25}$

- **Ejercicio 6:** Ajustar, por el método del lugar de las raíces el compensador de la Fig. 9.24, de tal forma que los polos en lazo cerrado queden situados en $-2 \pm \sqrt{2}j$.

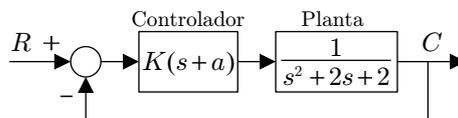


Figura 9.24: Sistema de control

Solución: $a = 2$ y $K = 2$

- **Ejercicio 7:** Una máquina de control numérico (Fig. 9.25) debe seguir una posición de referencia con un error inferior a 2 mm. La velocidad máxima de referencia de posición es de 1 m/s (Fig. 9.26). Los parámetros de la máquina son: $M = 1$ kg y $C = 6$ Ns/m. Determinar un compensador de retraso de fase adecuado para el sistema. El sistema debe tener un comportamiento transitorio razonable.

Nota 1: La referencia de posición se puede tomar como una rampa.

Nota 2: Por seguridad, la ganancia del controlador a bajas frecuencias debe ser finita, es decir, no puede tener un polo en el origen.

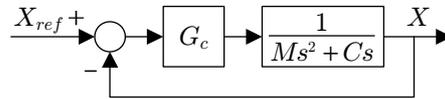


Figura 9.25: Máquina de control numérico

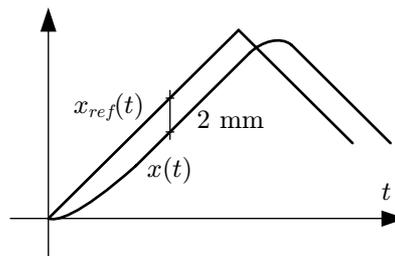


Figura 9.26: Señales de referencia y salida

Solución:

$$G_c(s) = 18 \frac{s+1}{s+0.006} \text{ usando el lugar de las raíces.}$$

$$G_c(s) = 23.71 \frac{s+1}{s+0.0079} \text{ usando el diagrama de Bode.}$$

- **Ejercicio 8:** Sea el sistema de la Fig. 9.27, diseñar un controlador de forma que el coeficiente de error ante entrada rampa sea de 28 s^{-1} y el margen de fase sea al menos de 45° .

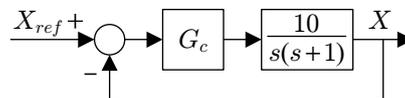


Figura 9.27: Sistema de control

Solución: Se pueden cumplir especificaciones con un compensador de retraso, sin embargo, lo mejor es diseñar un compensador de adelanto:

$$G_c(s) = 2.8 \frac{1 + 0.268s}{1 + 0.0582s} \quad (9.43)$$

- **Ejercicio 9:** Sea el sistema de la Fig. 9.28, calcular un compensador de adelanto de fase que consiga que el sistema posea un sobreimpulso menor del 10% y un error en régimen permanente menor del 5%. ¿Cuál es el ancho de banda del sistema compensado?

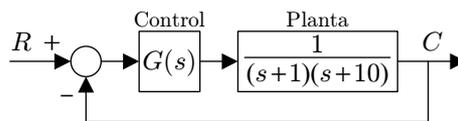


Figura 9.28: Sistema de control

Solución: $G(s) = 200 \frac{1+0.119s}{1+0.0583s}$, con 12 rad/s de ancho de banda aproximadamente.

- **Ejercicio 10:** Se desea diseñar un compensador de adelanto de fase que controle la posición de un satélite en el espacio. El diagrama de bloques del sistema se muestra en la Fig. 9.29. Se requiere un tiempo de establecimiento menor que 2 segundos con un amortiguamiento adecuado.

- Diseñar el compensador por el método del lugar de las raíces.

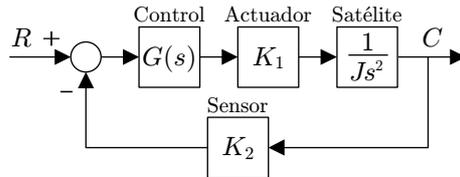


Figura 9.29: Sistema de control

b) Diseñar el compensador por el método de la respuesta en frecuencia.

Recomendaciones:

1. Expresar la ganancia del compensador en función de las constantes J , K_1 y K_2 del sistema.
2. La frecuencia de cruce de ganancias es aproximadamente igual a la frecuencia natural de los polos dominantes del sistema.

Soluciones: Si se elige $\zeta = 0.45$,

$$\begin{aligned} \text{a) } G(s) &= \frac{34J}{K_1 K_2} \frac{s+2}{s+7.5} \\ \text{b) } G(s) &= \frac{48.22J}{K_1 K_2} \frac{s+1.82}{s+10.79} \end{aligned}$$

- **Ejercicio 11:** Diseñar un compensador de retraso de fase que deje el sistema con un error del 4% y un margen de fase de 45° , donde la planta es:

$$G(s) = \frac{1}{(s+3)^2} \quad (9.44)$$

Solución: $G_c(s) = 216.5 \frac{1+s}{1+1.93s}$

- **Ejercicio 12:** Sea la siguiente planta en lazo abierto:

$$G(s) = K \frac{e^{-0.2s}}{s} \quad (9.45)$$

- a) Determinar la ganancia K que deja al sistema con margen de fase de 60° .
- b) Determinar la ganancia crítica del sistema.
- c) Dibujar en unos ejes de coordenadas los valores que va tomando el margen de fase en función de la ganancia K , si dicha ganancia varía entre 1 y la ganancia crítica. Explicar el resultado.

Soluciones:

- a) $K = 2.61$
- b) $K_{CR} = 7.74$
- c) El margen de fase frente a la ganancia K , sale una línea recta decreciente. Esto es lógico porque pasar la ganancia de dB a K es pasar a escala logarítmica y el margen de fase también cae de forma logarítmica por el retraso.

- **Ejercicio 13:** Diseñar un compensador de adelanto de fase tal que el error sea de 5 mm ante entrada rampa de pendiente unidad y el margen de fase sea de 45° , donde la planta (que tiene unidades en el SI) es:

$$G(s) = \frac{2500}{s(s+25)} \quad (9.46)$$

Soluciones: $K = 2$, $\omega'_{cg} = 70$ rad/s, $\phi_m = 33^\circ$, $\alpha = 0.29$ y $T = 0.02$ s.

- **Ejercicio 14:** Se desea tener el mínimo error, asegurando un comportamiento transitorio razonable, en la siguiente planta:

$$G(s) = \frac{e^{-0.5s}}{s+5} \quad (9.47)$$

- a) Con un controlador puramente proporcional, ¿qué magnitud de error puede dejar en el sistema?

- b) ¿Y con un compensador de adelanto de fase? ¿Puede mejorar el error manteniendo el comportamiento transitorio?

- **Ejercicio 15:** Sea el sistema de la Fig. 9.30:

- a) Elegir una constante K adecuada para el sistema y determinar el valor del margen de fase y ancho de banda del sistema compensado.
- b) ¿Cuánto vale el error ante una entrada rampa $r(t) = t$?
- c) Dibujar aproximadamente la respuesta temporal del sistema $c(t)$ ante una entrada escalón y una entrada rampa.

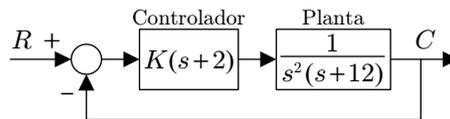


Figura 9.30: Sistema de control

Capítulo 10

Controladores PID

Actualmente los dispositivos de control son de uso común en las empresas. Las técnicas de fabricación en serie han hecho que no se implementen compensadores para una función particular, sino dispositivos genéricos que sean capaces de ajustarse para una labor específica, según las características del sistema.

El controlador proporcional-integral-derivativo, o controlador PID, es un dispositivo de control genérico donde el diseñador sólo tiene que dar valores adecuados, según lo requiera la situación, a los distintos parámetros que contiene. Por tanto, se elude la necesidad de fabricar el compensador que se desea implementar.

10.1. Expresión general

Como su propio nombre indica, un controlador PID es un caso particular de compensador de adelanto-retraso en el que el compensador de adelanto es proporcional-derivativo y el compensador de retraso es proporcional-integral. Del producto de ambos compensadores, se obtiene un controlador con dos ceros —que en general pueden ser reales o no—, un polo en el origen y una ganancia.

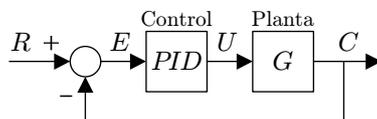


Figura 10.1: Sistema controlado con PID

$$G_{\text{PID}}(s) = G_{\text{PD}}(s)G_{\text{PI}}(s) = K_1(s+a)K_2\frac{s+b}{s} = K\frac{s^2 + \alpha s + \beta}{s} \quad (10.1)$$

Un controlador PID, por tanto, tiene tres parámetros que se pueden elegir: la posición de los dos ceros y el valor de la ganancia.

10.1.1. Forma estándar

Una expresión equivalente a la ecuación (10.1) es la que se presenta en (10.2), también llamada forma estándar del controlador PID.

$$U(s) = K_p \left(1 + \frac{1}{T_i s} + sT_d \right) E(s) \quad (10.2)$$

En (10.3) se observa la actuación temporal del controlador en la planta, que tiene tres sumandos: uno proporcional al error, otro proporcional a la integral del error y otro proporcional a la derivada del error.

$$u(t) = K_p \left[e(t) + \frac{1}{T_i} \int_0^t e(\tau) d\tau + T_d \frac{de(t)}{dt} \right] \quad (10.3)$$

A la constante K_p se le llama ganancia proporcional y posee las unidades que relacionan la actuación con el error, T_i es la constante de tiempo integral y tiene unidades de segundos, y T_d la constante de tiempo derivativa y también tiene unidades de segundos.

10.1.2. Forma paralela

La actuación del controlador se puede separar en forma de tres sumandos diferentes. Cada uno de ellos acapara respectivamente la actuación proporcional, integral y derivativa:

$$U(s) = \left(K_P + \frac{K_I}{s} + sK_D \right) E(s) \quad (10.4)$$

Las constantes K_P , K_I y K_D se obtienen fácilmente conocidos los parámetros estándar K_p , T_i y T_d . Esta forma de expresar el controlador PID se conoce como paralela porque se puede representar como aparece en la Fig. 10.2.

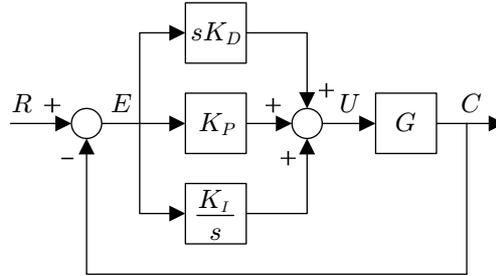


Figura 10.2: Sistema de control con PID en forma paralela

10.1.3. Forma serie

En el caso de que los dos ceros del controlador sean reales, se puede encontrar la forma serie o clásica del PID. La actuación del controlador PID serie se expresa como:

$$U(s) = K'_p \left(1 + \frac{1}{T'_i s} \right) (1 + sT'_d) E(s) \quad (10.5)$$

Los nuevos parámetros serie K'_p , T'_i y T'_d se pueden obtener también a partir de los parámetros estándar. La condición que deben cumplir los parámetros estándar para que los dos ceros del controlador sean reales es que el tiempo de integración sea mayor que cuatro veces el tiempo de derivación: $T_i > 4T_d$. Entonces, el controlador serie PID se puede representar en serie como aparece en la Fig. 10.3.

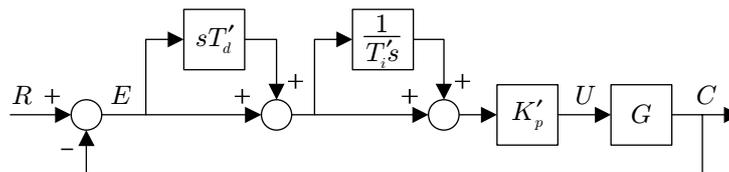


Figura 10.3: Sistema de control con PID en forma serie

La forma serie se llama también clásica porque los primeros actuadores PID neumáticos que se lograron implementar, siempre resultaban tener ceros reales. Es bueno conocer las tres formas de expresar los PID, porque el ingeniero puede manejar controladores comerciales que permitan introducir las constantes de alguna de estas tres maneras. En este capítulo, a partir de este momento, siempre se emplearán los parámetros de la forma estándar, que son los que tienen sentido físico más evidente.

10.2. Sentido físico de la actuación de un PID

Es posible ajustar los parámetros de un controlador PID sin un conocimiento preciso del tipo de actuación que realiza cada una de las partes del mismo. Sin embargo, resulta muy conveniente para poder predecir cómo afecta al sistema la modificación de cada uno de ellos.

10.2.1. Actuación proporcional

Si el tiempo de integración se hace infinito y el de derivación nulo, el controlador PID se transforma en una ley de control puramente proporcional al error entre la referencia y la salida.

$$u(t) = K_p e(t) \quad (10.6)$$

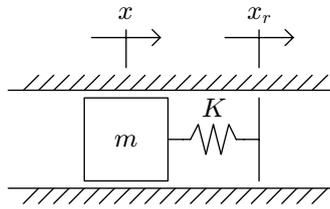


Figura 10.4: Actuación proporcional

Un símil mecánico de esta actuación, Fig. 10.4, es la que haría un muelle de rigidez $K = K_p$ que uniera la referencia con una masa que se deseara mover hasta dicha referencia. Se puede hacer más rápido el sistema aumentando la rigidez del muelle, es decir, aumentando la ganancia proporcional del controlador. Como contrapartida, también es previsible que el sistema se oscile más en torno a la referencia.

10.2.2. Actuación proporcional-derivativa

Una forma de evitar las fuertes oscilaciones que se pueden producir en torno a la referencia es añadir a la actuación proporcional otra actuación proporcional a la derivada del error. Esto es lo mismo que dotar al sistema una cierta capacidad de “anticipación” porque la inclusión del término derivativo es equivalente a actuar proporcionalmente al error que existirá dentro de T_d segundos.

$$u(t) = K_p \left[e(t) + T_d \frac{de(t)}{dt} \right] \approx K_p e(t + T_d) \quad (10.7)$$

Esta antelación es beneficiosa porque el sistema es capaz de “frenar” antes de llegar a la referencia. En la Fig. 10.5 se muestra cómo en el instante t_1 el error todavía es positivo, por lo que el control proporcional seguirá actuando en la planta para acercar la masa a la referencia, aunque sea una fuerza pequeña. Pero un usuario previsor, deduciría que con la elevada velocidad que lleva la masa en breves instantes se rebasará la posición de referencia por lo que en ese instante introduciría una fuerza contraria o “de frenado”. Es decir, actuar en t_1 con la fuerza que se estima para t_2 .

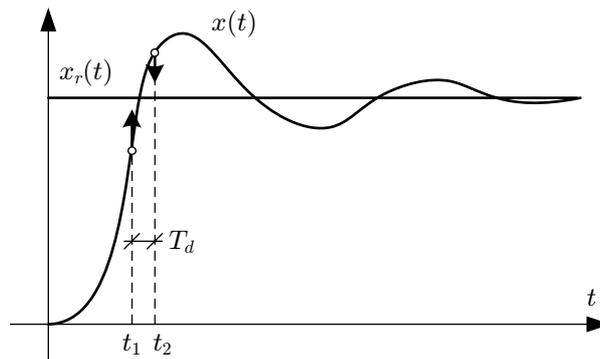


Figura 10.5: Actuación proporcional-derivativa

Un símil mecánico para la actuación proporcional-derivativa es la de imaginar que la posición de la masa y la referencia se encuentran unidas por un muelle y un amortiguador en paralelo. La rigidez del muelle sigue siendo igual a la ganancia proporcional, mientras que el coeficiente de amortiguamiento es el producto de la ganancia por la constante de tiempo de derivación.

$$u(t) = K_p \left[e(t) + T_d \frac{de(t)}{dt} \right] = K_p e(t) + T_d K_p \frac{de(t)}{dt} = K e(t) + B v(t) \quad (10.8)$$

Con esta comparación se muestra de forma más evidente cómo la actuación derivativa puede frenar el sistema haciéndolo menos oscilatorio, más amortiguado. Además, también se observa cómo al aumentar el valor de T_d se incrementa el “amortiguamiento” del sistema.

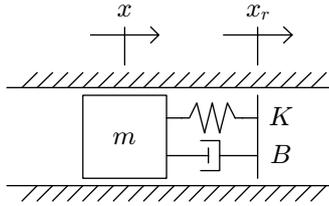


Figura 10.6: Actuación proporcional-derivativa

Con estas consideraciones, es posible aventurar unos valores adecuados para el tiempo T_d de la parte derivativa. No parece razonable asignar a T_d un valor muy elevado, en concreto superior al periodo de oscilación que posee el sistema sin acción derivativa. Parece lógico pensar que la estimación del error en T_d segundos sólo es buena mientras T_d se encuentre entre cero y, como mucho, un cuarto del periodo de oscilación del sistema.

10.2.3. Actuación proporcional-integral

Una característica común de la actuación proporcional y la proporcional-derivativa es que se hace cero cuando el error desaparece. Sin embargo, en algunos casos puede ser necesario que esto no sea así. En la Fig. 10.7 se muestra el mismo ejemplo de los apartados anteriores con la masa moviéndose verticalmente.

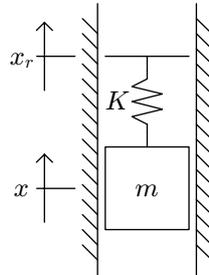


Figura 10.7: Sistema con error no nulo ante actuación proporcional

Si la actuación es puramente proporcional, la masa en régimen permanente no alcanzaría la referencia sino que se colocará en el lugar donde la acción del muelle contrarreste la fuerza del peso.

$$K_p e_{ss} = mg \rightarrow e_{ss} = \frac{mg}{K_p} \quad (10.9)$$

Si se desea que no exista error en régimen permanente a la acción proporcional hay que añadir una actuación extra u_0 , (10.10). En los primeros sistemas controlados, la actuación u_0 se añadía de forma manual, incrementando esa especie de offset hasta que desaparecía el error.

$$u(t) = K_p e(t) + u_0 \quad (10.10)$$

En el caso del sistema de la Fig. 10.7, es evidente que, consiguiendo error nulo, el controlador debe introducir una fuerza $u(t) = u_0 = mg$. La solución automatizada más adecuada a este problema es ir aumentando el valor de u_0 de forma proporcional a la integral del error. La función integral del error aumenta paulatinamente mientras exista error no nulo hasta alcanzar, con error nulo, un valor finito.

$$u(t) = K_p e(t) + \frac{K_p}{T_i} \int_0^t e(\tau) d\tau \quad (10.11)$$

La constante de tiempo de integración T_i da una idea del tiempo que se tarda en anular el error de forma automática. Esto se puede mostrar, de forma aproximada, calculando el valor de u_0 si el error en régimen permanente permanece constante:

$$u_0 = \frac{K_p}{T_i} \int_0^t e(\tau) d\tau = \frac{K_p}{T_i} e_{ss} \int_0^t d\tau = \frac{K_p}{T_i} e_{ss} t \quad (10.12)$$

Sustituyendo en la ecuación (10.12) el valor del error que se obtuvo en (10.9), resulta que cuando el tiempo es igual a la constante de tiempo integral, $t = T_i$, el valor de u_0 alcanza el valor deseado mg .

$$u_0 = \frac{K_p}{T_i} e_{ss} t = \frac{K_p}{T_i} \frac{mg}{K_p} t = mg \frac{t}{T_i} \quad (10.13)$$

Por tanto, la constante de tiempo T_i da una idea del momento en que se anula el error en régimen permanente. Si se elige una T_i muy elevada, el sistema tarda mucho en alcanzar la referencia, Fig. 10.8.

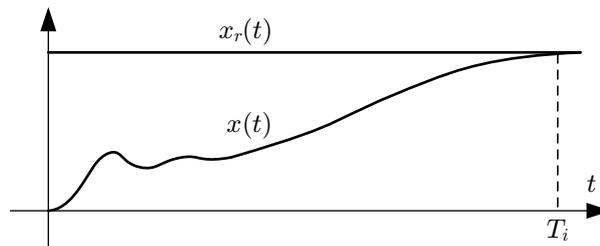


Figura 10.8: Actuación proporcional-integral con T_i muy grande

La interpretación física que se acaba de dar a la actuación integral concuerda con el hecho de que cuando T_i se hace infinito entonces el sistema no tiene actuación integral. Un valor adecuado para T_i puede ser el periodo de oscilación del sistema, o un tiempo algo menor.

La actuación integral puede darse también en sistemas que, de suyo, carezcan de error en régimen permanente ante un determinado tipo de entrada. Lo que consigue entonces la parte integral es elevar el tipo del sistema en una unidad y anular el error ante entradas más severas.

10.3. Ajuste experimental de PID

Las ideas enunciadas en el apartado anterior ayudan a conocer cómo cambia la respuesta del sistema modificando alguno de los parámetros del controlador, pero resultan insuficientes para poder asignar de forma adecuada sus valores numéricos.

Para asignar valores a los parámetros del controlador sin conocer la función de transferencia de la planta que se desea controlar, se han propuesto una serie de tablas que utilizan varios parámetros que se obtienen de forma experimental sobre la planta.

10.3.1. Ajuste de Ziegler-Nichols

El método más utilizado es el que propusieron en 1942 John G. Ziegler y Nataniel B. Nichols [1] para el control de servomecanismos hidráulicos en baterías antiaéreas empleadas en la Segunda Guerra Mundial. El ajuste de Ziegler-Nichols propone unos parámetros para el PID de forma que el sistema controlado posea un buen rechazo a las perturbaciones que se puedan introducir en el sistema. Esto quiere decir que el seguimiento que hace el sistema a la referencia puede ser poco amortiguado, con demasiado sobreimpulso. Pero esto se considera intrascendente comparado con la especificación mencionada.

Tabla 10.1: Ajuste de Ziegler-Nichols

Tipo	K_p	T_i	T_d	K_p	T_i	T_d
P	$\frac{1}{a}$	∞	0	$0.5K_{CR}$	∞	0
PI	$\frac{0.9}{a}$	$3L$	0	$0.45K_{CR}$	$\frac{T_{CR}}{1.2}$	0
PID	$\frac{1.2}{a}$	$2L$	$0.5L$	$0.6K_{CR}$	$\frac{T_{CR}}{2}$	$\frac{T_{CR}}{8}$

En muchos procesos industriales un buen rechazo a las perturbaciones es mucho más interesante que un buen seguimiento a la referencia. Por ejemplo, en una planta de elaboración de objetos plásticos, es muy importante que la temperatura del fluido permanezca constante en la referencia a pesar de las perturbaciones que suponen la entrada y la salida de material. El proceso inicial de calentamiento, o régimen transitorio, no es muy importante de cara a la producción. Puede ser más o menos largo, con mayor o menor sobreimpulso, pero lo importante es que una vez que se llega a la temperatura de régimen permanente, las perturbaciones no hagan variar la temperatura dentro de un rango permisible.

En concreto, la especificación que se pretende con Ziegler-Nichols es obtener una relación de caída de sobreimpulsos de un cuarto, es decir, que ante la entrada de una perturbación los sucesivos rebases en torno a la referencia sean sucesivamente cada uno cuatro veces inferior al anterior.

Los valores para los parámetros del PID se obtienen con la Tabla 10.1. Existen dos formas de ajuste: uno emplea los parámetros a y L de la respuesta de la planta ante una entrada escalón unidad y otro que emplea los parámetros de ganancia crítica K_{CR} y periodo de oscilación crítico T_{CR} de la planta.

En la Fig. 10.9 se muestra cómo se obtienen los parámetros a y L de la respuesta de la planta ante una entrada escalón unidad. El valor de la salida en régimen permanente K se relaciona por trigonometría con el tiempo muerto L y la constante de tiempo T , según la ecuación (10.14).

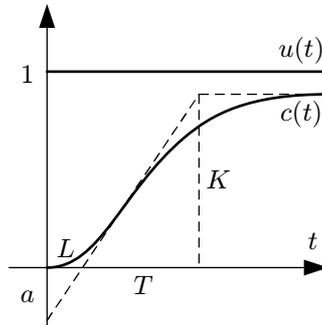


Figura 10.9: Respuesta de la planta a un escalón unidad

$$a = \frac{KL}{T} \quad (10.14)$$

La función de transferencia de la planta no se conoce, pero una aproximación de la misma puede ser (10.15), es decir, un sistema de primer orden con constante de tiempo T , un tiempo muerto L y una ganancia estática K .

$$G(s) = \frac{Ke^{-Ls}}{1 + Ts} \quad (10.15)$$

Los parámetros de ganancia crítica K_{CR} y periodo de oscilación crítico T_{CR} de la planta se pueden obtener experimentalmente de varias formas. Una posibilidad es introducir la planta en un sistema de control proporcional y aumentar la ganancia hasta volver la salida del sistema oscilatoria ante una entrada escalón, es decir, en el límite de estabilidad. La ganancia que deja el sistema en el límite de estabilidad es la ganancia crítica K_{CR} , mientras que el periodo de oscilación que se observe en la salida del sistema es el crítico T_{CR} .

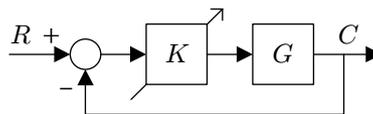


Figura 10.10: Cálculo de la ganancia crítica

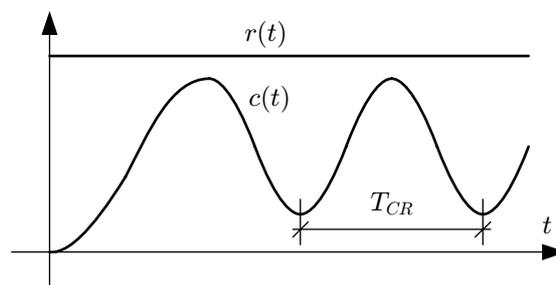


Figura 10.11: Medida del periodo de oscilación crítico

Si se conoce la función de transferencia de la planta, es posible obtener la ganancia crítica del sistema analíticamente por medio del criterio de Routh-Hurwitz. El periodo de oscilación crítico se puede obtener sustituyendo K por la K_{CR} , a través de los polos en lazo cerrado del sistema de la Fig. 10.10, que se sitúan sobre el eje imaginario del plano S , ecuación (10.16).

$$p_{CR} = \pm \frac{2\pi}{T_{CR}}j \quad (10.16)$$

El ajuste de Ziegler-Nichols cumple una peculiaridad, y es que el tiempo de integración siempre es cuatro veces mayor que el tiempo de derivación. Esta elección es razonable, como se vio en el sentido físico de cada uno de las partes del PID. Matemáticamente, se puede demostrar que, en este caso concreto, los dos ceros del PID son reales y doble (10.17).

$$U(s) = \frac{K_p T_d}{s} \left(s + \frac{1}{2T_d} \right)^2 E(s) \quad (10.17)$$

10.3.2. Otros ajustes experimentales

Varios autores han propuesto tablas de ajuste para los parámetros del PID. A continuación se presentan las que propusieron Chien-Hrones-Rewick. Estos autores ofrecen los parámetros que consiguen sobreimpulsos del 20 % o del 0 %, tanto ante entradas referencia, como ante entrada perturbación.

Tabla 10.2: Ajuste de Chien-Hrones-Rewick para entrada perturbación N

Tipo	$M_p = 0\%$			$M_p = 20\%$		
	K_p	T_i	T_d	K_p	T_i	T_d
P	$\frac{0.3}{a}$	∞	0	$\frac{0.7}{a}$	∞	0
PI	$\frac{0.6}{a}$	$4L$	0	$\frac{0.7}{a}$	$2.3L$	0
PID	$\frac{0.95}{a}$	$2.4L$	$0.42L$	$\frac{1.2}{a}$	$2L$	$0.42L$

Tabla 10.3: Ajuste de Chien-Hrones-Rewick para entrada referencia R

Tipo	$M_p = 0\%$			$M_p = 20\%$		
	K_p	T_i	T_d	K_p	T_i	T_d
P	$\frac{0.3}{a}$	∞	0	$\frac{0.7}{a}$	∞	0
PI	$\frac{0.35}{a}$	$1.2T$	0	$\frac{0.6}{a}$	T	0
PID	$\frac{0.6}{a}$	T	$0.5L$	$\frac{0.95}{a}$	$1.4T$	$0.47L$

Tabla 10.4: Ajuste de Cohen-Coon

Tipo	K_p	T_i	T_d
P	$\frac{1}{a} \left(1 + \frac{L}{3T} \right)$	∞	0
PI	$\frac{1}{a} \left(\frac{9}{10} + \frac{L}{12T} \right)$	$\frac{30T+3L}{9T+20L} L$	0
PID	$\frac{1}{a} \left(\frac{4}{3} + \frac{L}{4T} \right)$	$\frac{32T+6L}{13T+8L} L$	$\frac{4T}{11T+2L} L$

Hay que destacar que, los parámetros Chien-Hrones-Rewick no mantienen la relación cuatro a uno entre los tiempos de integración y derivación que tenían los de Ziegler-Nichols. También se puede observar cómo los parámetros Chien-Hrones-Rewick de la Tabla 10.2 se asemejan a los de Ziegler-Nichols si aumenta el sobreimpulso permitido. El ajuste de Cohen-Coon, Tabla 10.4, es otra propuesta para conseguir un buen seguimiento ante la entrada referencia.

10.3.3. Ejemplo comparativo

En este apartado se calculan los parámetros del PID para el sistema de la Fig. 10.12, por los métodos de Ziegler-Nichols y de Chien-Hrones-Rewick para referencia y 20 % de sobreimpulso.

La respuesta de la planta a una entra escalón unidad se observa en la Fig. 10.13, donde se puede medir un tiempo muerto L de 0.7826 segundos y un coeficiente a de 0.2092.

Con L y a , acudiendo a la Tabla 10.1 los parámetros del PID por Ziegler-Nichols son 5.7361 para la K_p , 1.5652 segundos para la T_i y 0.3913 segundos para la T_d . Si se acude a la Tabla 10.3, los valores por Chien-Hrones-Rewick son 4.5411 para la K_p , 5.1738 segundos para la T_i y 0.3678 segundos para la T_d . En

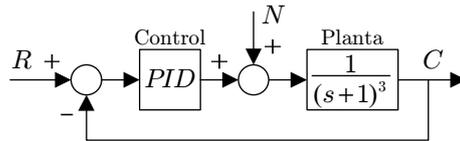


Figura 10.12: Sistema controlado por un PID

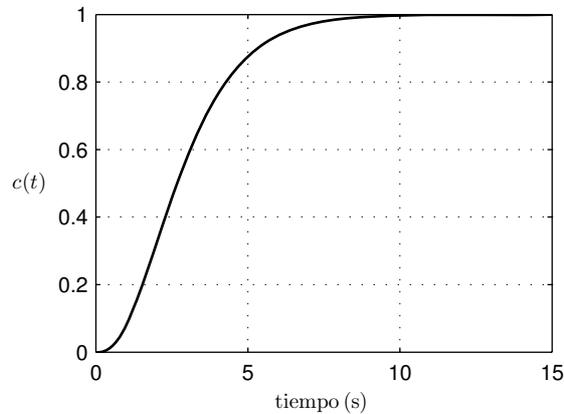


Figura 10.13: Respuesta temporal ante entrada escalón unidad

la Fig. 10.14 se comparan las respuestas del sistema con los dos PID ante una entrada referencia escalón unidad en el instante inicial y una entrada perturbación, también escalón unidad, a los 20 segundos.

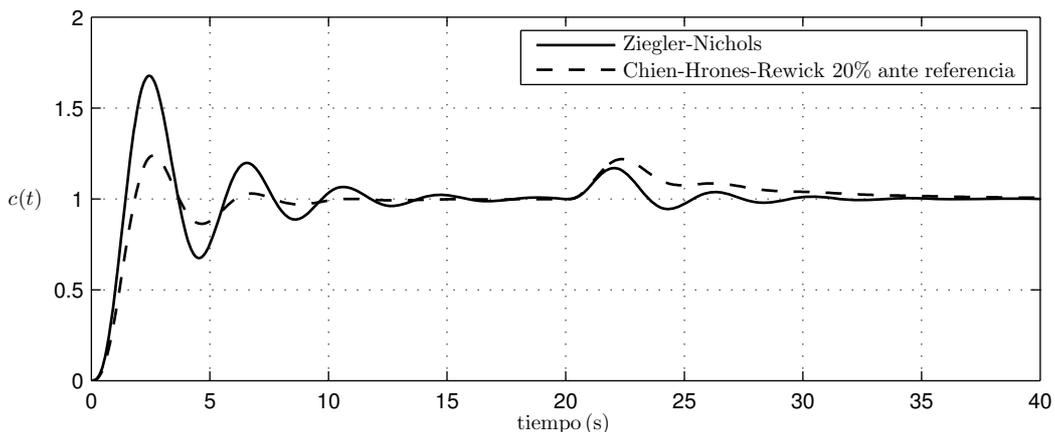


Figura 10.14: Respuesta temporal del sistema

Como era de esperar, los parámetros de Ziegler-Nichols propician un rechazo a las perturbaciones mejor que los de Chien-Hrones-Rewick, pero ante la entrada referencia el sobreimpulso es muy elevado, en este caso casi del 70%. También se observa cómo los parámetros de Chien-Hrones-Rewick consiguen el sobreimpulso aproximadamente de 20% para los que están diseñados.

10.4. Ajuste analítico de PID por asignación de polos

Existen muchas formas de ajustar los controladores PID. En este apartado sólo se mencionará un método analítico de ajuste que persigue colocar los polos del sistema en lazo cerrado en aquellas posiciones que garantizan un comportamiento adecuado en régimen transitorio.

El método se va a explicar con el ejemplo de la Fig. 10.15, en el que se desea ajustar un controlador PI para un sistema de primer orden. La función de transferencia en lazo cerrado es:

$$\frac{C(s)}{R(s)} = \frac{KK_p(1 + T_i s)}{TT_i s^2 + T_i(1 + KK_p)s + KK_p} \quad (10.18)$$

Por tanto, queda un sistema de segundo orden con un cero. La localización de los polos en lazo cerrado

se pueden colocar analíticamente donde se desee. Basta resolver los valores de K_p y T_i que satisfacen:

$$TT_i s^2 + T_i(1 + KK_p)s + KK_p = \alpha(s^2 + 2\zeta\omega_n s + \omega_n^2) \quad (10.19)$$

La posición del cero viene dada con la asignación de polos, pero es posible eliminar su efecto con un prefiltro de la referencia. El diseño de prefiltros no se explica en estos apuntes.

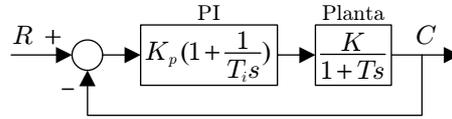


Figura 10.15: Control PI sobre una planta de primer orden

10.5. Control con dos grados libertad

Anteriormente se ha mostrado cómo el ingeniero a la hora de ajustar el PID debe elegir qué tipo de especificación desea obtener: si un buen comportamiento ante cambios en la referencia o un buen comportamiento ante presencia de perturbaciones. Sin embargo, es posible encontrar arquitecturas de control que permitan satisfacer especificaciones diferentes ante distintas entradas. Se llamará *control con dos grados libertad* aquella estrategia de control que permite ajustar a la vez especificaciones ante los cambios de referencia y la presencia de perturbaciones.

Las arquitecturas de control con dos grados de libertad se podrían estudiar en un capítulo distinto que el de los controladores PID. Sin embargo, los controladores PID más avanzados se entenderán mejor una vez que se han explicado los fundamentos del control con dos grados de libertad.

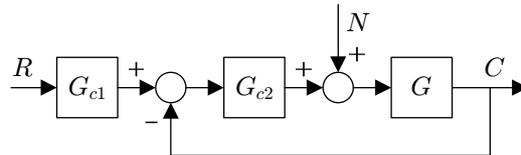


Figura 10.16: Arquitectura de control con dos grados de libertad

Como ejemplo de control con dos grados de libertad, se propone el diagrama de bloques de la Fig. 10.16. Antes de escribir las funciones de transferencia que relacionan la salida $C(s)$ con la referencia $R(s)$ o la perturbación $N(s)$, es razonable imaginar que la función de transferencia del controlador externo al lazo $G_{c1}(s)$ sí va a influir en el “camino” que va de $R(s)$ a $C(s)$, pero no en el que va de $N(s)$ a $C(s)$.

$$C(s) = \frac{G_{c1}(s)G_{c2}(s)G(s)}{1 + G_{c2}(s)G(s)}R(s) + \frac{G_{c2}(s)G(s)}{1 + G_{c2}(s)G(s)}N(s) \quad (10.20)$$

Por tanto, el ingeniero puede diseñar a voluntad la función de transferencia del controlador interno al lazo $G_{c2}(s)$ para fijar un comportamiento adecuado ante la entrada perturbación. A continuación puede determinar la función de transferencia del controlador exterior al lazo $G_{c1}(s)$ para definir el comportamiento —en general distinto— del sistema ante cambio de referencia. De ahí que $G_{c1}(s)$ y $G_{c2}(s)$ son los dos grados de libertad de la arquitectura de control de la Fig. 10.16. A la función de transferencia $G_{c1}(s)$ de esta estrategia se le suele llamar *filtro de la referencia*.

Es interesante notar que las arquitecturas de control con dos grados de libertad poseen funciones de transferencia con idéntico denominador y, por tanto, la estabilidad es una característica común de toda la estrategia.

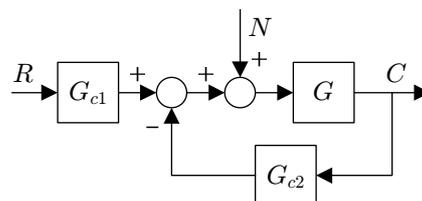


Figura 10.17: Variante de la estrategia con filtro de referencia

Una posible variante de arquitectura de control con dos grados de libertad es la que se presenta en la Fig. 10.17. En este caso, el controlador interno al lazo se sitúa en la realimentación. La salida del sistema es:

$$C(s) = \frac{G_{c1}(s)G(s)}{1 + G_{c2}(s)G(s)}R(s) + \frac{G(s)}{1 + G_{c2}(s)G(s)}N(s) \quad (10.21)$$

En las siguientes variantes es posible identificar en el diagrama de bloques la señal del error $E(s)$, cosa que no se podía hacer en las anteriores arquitecturas. En la estrategia de la Fig. 10.18 el filtro de la referencia se sustituye por un lazo de prealimentación.

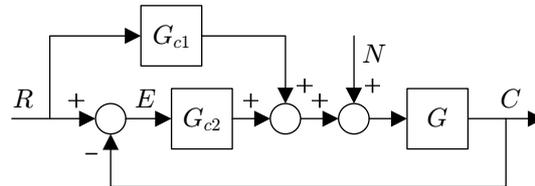


Figura 10.18: Estrategia con lazo de prealimentación

En la estrategia de la Fig. 10.19 existen en cambio dos caminos de realimentación. Se deja como ejercicio la obtención de los grados de libertad de esta arquitectura.

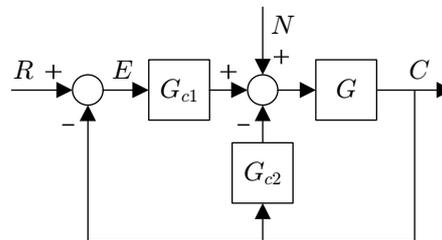


Figura 10.19: Estrategia con dos caminos de realimentación

10.6. Modificaciones del PID

La expresión estándar del PID se suele modificar para obtener mejores prestaciones. En los siguientes apartados se explican algunas de las mejoras más habituales.

10.6.1. Supresión del efecto kick-off

El efecto *kick-off* se produce en todos los sistemas controlados en los que actúa de forma proporcional a la derivada del error y la referencia es una entrada escalón. En el momento del cambio finito de la referencia la derivada del error se hace infinita, por lo que la actuación también se hace muy grande —teóricamente infinita— provocando la saturación de los actuadores. Esta actuación puede resultar nociva para la planta.

Una forma de solucionar este problema es hacer que la parte derivativa del PID no actúe sobre la señal del error $E(s)$, sino exclusivamente sobre la señal de salida $C(s)$ del sistema. Es fácil demostrar que esta estrategia no altera el comportamiento del sistema controlado ante entrada escalón y sólo suprime los impulsos infinitos que introduce la derivación del salto finito. Estos impulsos infinitos producirían una especie de golpes bruscos en el sistema: es el efecto *kick-off* que evita esta estrategia.

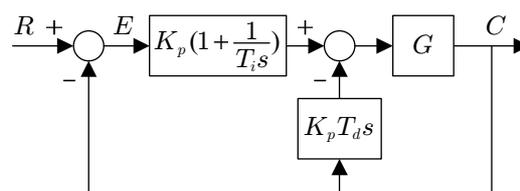


Figura 10.20: Controlador PI-D

En la Fig. 10.20 se muestra la estrategia descrita. Es evidente que se trata de una arquitectura de control con dos grados de libertad. Algunos autores llaman a esta arquitectura controlador PI-D.

En aplicaciones en las que interesa saturar el dispositivo con rapidez, se conserva la derivada de la referencia, e incluso se multiplica con un factor c , de la forma:

$$U(s) = K_p \left\{ E(s) + \frac{1}{T_i s} E(s) + T_d s [cR(s) - C(s)] \right\}. \quad (10.22)$$

Si ingeniero posee control sobre el parámetro c , se asignar un valor finito o incluso superior a la unidad para conseguir el efecto *kick-off* o hacer $c = 0$ para cancelarlo y obtener un controlador PI-D. Muchos fabricantes de controladores PID no dejan posibilidad de elección y hacen $c = 0$ por hardware.

10.6.2. Set-point weighting

Otro modo muy habitual de mejorar el controlador PID es introducir un factor b a la referencia de la parte proporcional:

$$U(s) = K_p \left\{ bR(s) - C(s) + \frac{1}{T_i s} E(s) + T_d s E(s) \right\}. \quad (10.23)$$

Para controladores P con mucho error en régimen permanente es habitual elegir valores de b mayores que la unidad, porque consiguen disminuir el error. En caso de controladores PI es habitual elegir valores entre 0 y 1, con lo que se consigue disminuir el primer sobreimpulso de la salida ante cambios bruscos de la referencia. En cualquier caso, el valor más habitual es $b = 1$.

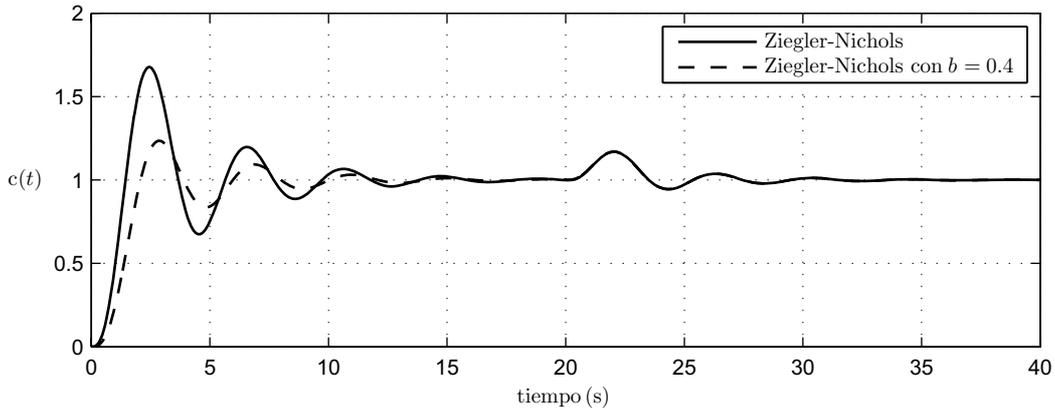


Figura 10.21: Respuesta temporal del sistema

En la Fig. 10.21 se observa cómo haciendo $b = 0.4$ en el sistema ejemplo de la Fig. 10.12, se ha reducido la magnitud del primer sobreimpulso hasta dejarlo del orden del 20%. De esta forma se consigue mejorar el comportamiento de los parámetros de Ziegler-Nichols ante la entrada referencia, conservando intacto el buen rechazo de las perturbaciones.

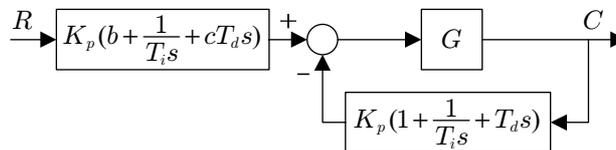


Figura 10.22: Controlador PID modificado con parámetros b y c

El PID modificado puede incluir los parámetros b y c a la vez. En la Fig. 10.22 se representa el diagrama de bloques de dicho PID modificado y se comprueba cómo esta ley de control también esconde una arquitectura con dos grados de libertad. El caso particular de PID con coeficientes $c = 0$ y $b = 0$ se le conoce como controlador I-PD y se muestra en la Fig. 10.23.

10.6.3. Filtro de la derivada

Las señales que pasan por el controlador suelen ir cargadas de ruido. Como la parte derivativa amplifica sin ningún tipo de limitación el ruido de alta frecuencia, es habitual introducir un filtro de primer orden

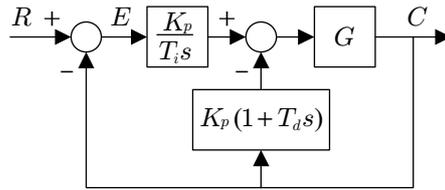


Figura 10.23: Controlador I-PD

en la parte derivativa, de tal forma que posea ganancia finita para altas frecuencias. Esto se consigue añadiendo un polo —es un filtro pasa-baja— a la parte derivativa:

$$U(s) = K_p \left\{ bR(s) - C(s) + \frac{1}{T_i s} E(s) + \frac{T_d s}{1 + \frac{T_d}{N} s} [cR(s) - C(s)] \right\} \quad (10.24)$$

El nuevo polo cuya constante de tiempo igual al tiempo derivativo dividido por N . Normalmente el valor de N está comprendido entre 2 y 20, siendo valores típicos 8 y 10. Como se puede deducir de la expresión (10.24), cuanto mayor es N menor es el efecto el filtro.

10.6.4. Prevención del efecto windup integral

La operación integral de un PID convencional también puede ocasionar problemas en los sistemas de control. Para hacerse una idea cualitativa de qué tipo de problemas pueden ocurrir, en la Fig. 10.24 se muestra el comportamiento típico de un sistema con *windup* integral.

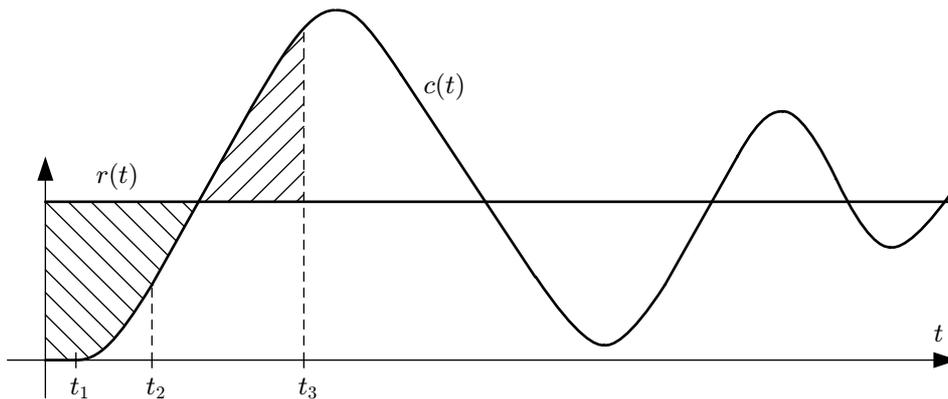


Figura 10.24: Efecto *windup* integral

En este ejemplo cualitativo existe un intervalo de tiempo, desde $t = 0$ hasta $t = t_1$, en el que el controlador está calculando fuerzas pero el actuador del sistema está apagado y por esa razón no se ha movido de su posición inicial. Este fenómeno puede darse muy fácilmente si el sistema cuenta, por ejemplo, con una seta de emergencia, porque es lo último que normalmente se suele accionar.

Como consecuencia de este hecho —o simplemente porque la referencia esté muy alejada de la posición inicial del sistema— el actuador alcanza el punto de saturación. Aunque podía haber ocurrido incluso antes que $t = t_1$, supóngase que la saturación se alcanza en el instante $t = t_2$. A partir de t_2 , a pesar de que el controlador demande una actuación mayor —la parte integral proporcional al área rayada sigue aumentando— el actuador introduce a la planta una actuación constante, la máxima que puede dar. Por esa razón, el crecimiento de la salida pocos instantes de $t = t_2$ se asemeja a una recta de pendiente constante.

Cuando la salida $c(t)$ del sistema sobrepasa la referencia $r(t)$, no se detiene en su crecimiento, porque —aunque la parte proporcional del PID comienza a actuar en contra de dicho movimiento— la parte integral es muy elevada y no se anulará y cambiará de dirección hasta mucho después, en torno a $t = t_3$ donde las áreas rayadas son parecidas.

Comienza entonces el efecto más típico de *windup* integral: una serie de oscilaciones poco amortiguadas en torno a la referencia, con formas muy triangulares hasta que —al cabo del tiempo— desaparezca la saturación de los actuadores.

El efecto *windup* integral es especialmente peligroso en aquellos sistemas con espacio de trabajo limitado. Es muy típico en aplicaciones de robótica que se imponga al robot una posición de referencia

muy alejada a la posición inicial del robot, pero cerca del extremo de recorrido del robot. Si el ingeniero no se percató de la existencia de este fenómeno, el robot sobrepasará la referencia sin disminuir su velocidad y llegará a los toques del mecanismo provocando eventualmente su rotura.

Existen varias maneras de prevenir el efecto *windup* integral. Si el PID está implementado en código, no mediante un circuito analógico, es posible realizar una integración condicionada, es decir, sólo se integra mientras no se alcance la saturación. También es relativamente fácil asegurar por software que no se exigen cambios de referencia demasiado bruscos, que son los que provocan fácilmente la saturación de los actuadores. Así por ejemplo, se puede reemplazar las referencias escalón grandes por rampas cuya pendiente no sobrepase la velocidad máxima que puede alcanzar el sistema (como se muestra en la Fig. 10.25).

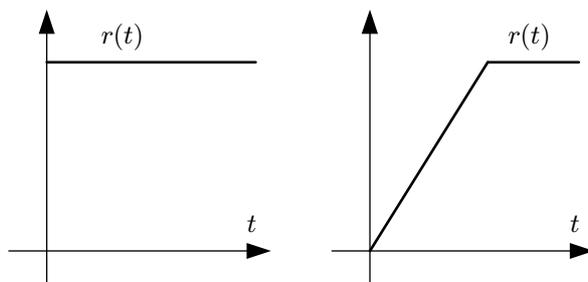


Figura 10.25: Cambio de referencia escalón y cambio de referencia con velocidad limitada

Existen otras formas de evitar el efecto *windup* integral usando el tracking de la señal de actuación saturada. Como son estrategias no lineales no se estudiarán en esta asignatura.

10.7. Ejercicios propuestos

- **Ejercicio 1:** Dados los siguientes sistemas controlados:

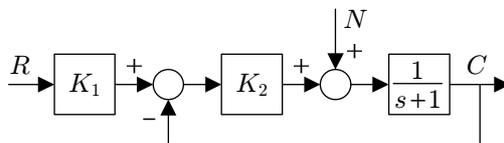


Figura 10.26: Sistema 1

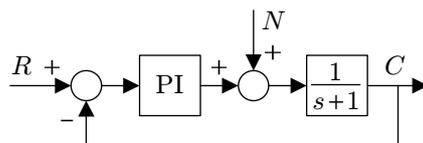


Figura 10.27: Sistema 2

- Determinar los valores de K_1 y K_2 para que el Sistema 1 tenga error nulo en régimen permanente y una constante de tiempo de 0.1 segundos
- Determinar las constantes del PI, en el Sistema 2, para que la respuesta sea la misma que en el apartado anterior. La ecuación de este controlador es:

$$U(s) = K_p \left(1 + \frac{1}{T_i s} \right) E(s) \quad (10.25)$$

Nota: Se recomienda cancelar el polo de la planta con el cero del PI, y ajustar la ganancia del sistema resultante.

- Al cabo de 10 segundos hay una perturbación de tipo escalón de amplitud 3. Determinar el error en régimen permanente para cada sistema.

Soluciones:

- $K_2 = 9$ y $K_1 = \frac{10}{9}$

- b) $T_i = 1$ y $K_p = 10$
- c) Error de $-\frac{3}{10}$ en el Sistema 1 y nulo en el Sistema 2.

- **Ejercicio 2:** Sea el sistema de la figura:

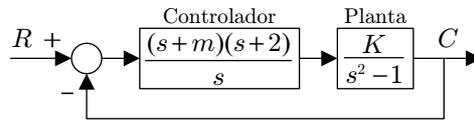


Figura 10.28: Sistema de control

- a) Determinar la relación que deben cumplir K y m , positivos, para que el sistema sea estable.
- b) Elegir el valor de m que deja el sistema en lazo cerrado de segundo orden para cualquier K y a continuación elegir el valor de K en la planta para obtener un comportamiento transitorio adecuado en el sistema compensado, por el método del lugar de las raíces. Calcular los puntos de ruptura y cortes con el eje imaginario.
- c) Calcular el error en régimen permanente del sistema compensado del apartado b) para una entrada rampa de pendiente unidad. Dibujar cualitativamente a mano alzada la entrada y la salida en el mismo eje de tiempos.
- d) Con el valor de m elegido en el apartado b) el bloque del controlador de la figura representa un PID, ¿cuáles son los valores estándar K_p , T_i y T_d de dicho PID?

Soluciones:

- a) $K > \frac{2m+1}{m+2}$
- b) $m = 1$ para anular un polo de la planta y $K = 5$ para dejar los dos polos en lazo cerrado sobre el único cero, puntos de ruptura en -4.449 y 0.449 , puntos de corte con en eje imaginario en $1.414j$ y $-1.414j$ para $K = 1$
- c) $e_{ss} = -0.1$ para $K = 5$
- d) $K_p = 3$, $T_i = 1.5$ segundos y $T_d = 0.33$ segundos para $m = 1$.

- **Ejercicio 3:** Diseñar un controlador PI para el siguiente sistema indicando la función de transferencia del controlador.

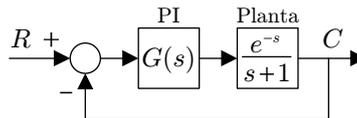


Figura 10.29: Sistema de control

- a) Usando el primer método de Ziegler-Nichols.
- b) Usando el segundo método de Ziegler-Nichols.

- **Ejercicio 4:** El desplazamiento de una herramienta se controla mediante dos lazos de retroalimentación, uno de velocidad y otro de posición. La masa m vale 50 kg y el amortiguamiento b es de 10 Ns/m.

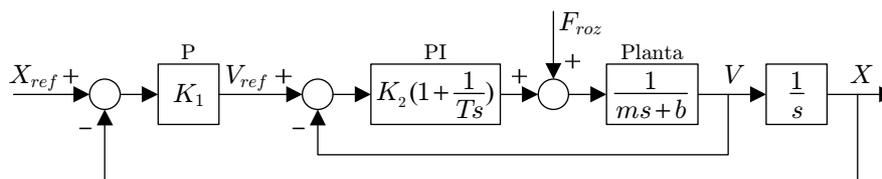


Figura 10.30: Herramienta con doble lazo de control

1. Determinar el valor de T para que el cero del controlador PI anule al polo de la herramienta.

2. Usando el valor de T del apartado anterior determinar:
 - a) La inecuación que deben cumplir el resto de los parámetros del sistema para que el error de posición ante una entrada rampa de 10 cm/s sea menor que 1 mm.
 - b) La inecuación que deben cumplir el resto de los parámetros del sistema para que el ancho de banda del lazo de control de velocidad sea mayor que 2 rad/s.
 - c) La inecuación que deben cumplir el resto de los parámetros del sistema para que el tiempo de establecimiento del control en posición sea menor que 2 décimas de segundo.
 - d) La inecuación que deben cumplir el resto de los parámetros del sistema para que el amortiguamiento del sistema de control en posición sea menor que 1 y mayor que 0.5.
3. El PI del lazo de control en velocidad, ¿anula el error ante entrada escalón en la fuerza de rozamiento de cualquier magnitud?

- **Ejercicio 5:** Dibujar el diagrama de bloques correspondiente al siguiente controlador PID:

$$u(t) = K_p \left\{ by_{ref}(t) - y(t) + \frac{1}{T_i} \int_0^t [y_{ref}(\tau) - y(\tau)] d\tau - T_d \frac{dy(t)}{dt} \right\} \quad (10.26)$$

Capítulo 11

Control en espacio de estado

11.1. Introducción

El comportamiento de un sistema lineal de orden n se puede definir por medio de n ecuaciones diferenciales de primer orden. Una forma matricial de expresar esas ecuaciones diferenciales es el llamado modelo en *espacio de estado*, sistema (11.1), donde además de las entradas $\mathbf{u}(t)$ y las salidas $\mathbf{y}(t)$ aparecen un conjunto de variables $\mathbf{x}(t)$ que se denominan *estados del sistema*.

$$\begin{cases} \dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) & \text{ecuación de estado} \\ \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t) & \text{ecuación de salida} \end{cases} \quad (11.1)$$

Los sistemas lineales de coeficientes constantes poseen matrices \mathbf{A} , \mathbf{B} , \mathbf{C} y \mathbf{D} con elementos constantes. Habitualmente la matriz \mathbf{D} es nula. En adelante se estudiarán exclusivamente los sistemas *single input single output*, cuyas ecuaciones en espacio de estado son:

$$\begin{cases} \dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{b}u(t) \\ y(t) = \mathbf{c}^T\mathbf{x}(t) \end{cases} \quad (11.2)$$

Por ejemplo, para un sistema de tercer orden las ecuaciones de estado y de salida son:

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \\ \dot{x}_3(t) \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} u(t) \quad (11.3)$$

$$y(t) = [c_1 \quad c_2 \quad c_3] \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix} \quad (11.4)$$

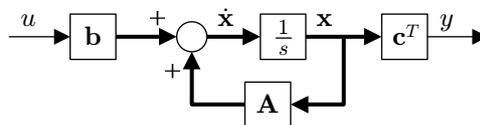


Figura 11.1: Modelo en espacio de estado de un sistema

El modelo en espacio de estado es muy útil para obtener la respuesta temporal de un sistema de forma numérica iterativa. La Fig. 11.1 muestra el diagrama de bloques del modelo. El único inconveniente es que no deja patente el orden de las multiplicaciones matriciales.

Es posible trasladar la ecuación (11.2) al dominio de Laplace,

$$\begin{cases} s\mathbf{x}(s) = \mathbf{A}\mathbf{x}(s) + \mathbf{b}u(s) \\ y(s) = \mathbf{c}^T\mathbf{x}(s) \end{cases} \quad (11.5)$$

y encontrar la expresión de la función de transferencia del sistema:

$$\frac{y(s)}{u(s)} = \mathbf{c}^T(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{b} \quad (11.6)$$

El paso de espacio de estado a función de transferencia no siempre es posible. En cambio, el paso de función de transferencia a espacio de estado siempre es posible y, además, de infinitas maneras distintas.

Con la misma limitación que la enunciada para la función de transferencia, se puede definir la ecuación característica como

$$q(s) = \det(s\mathbf{I} - \mathbf{A}) = 0, \quad (11.7)$$

donde se observa que la posición de los polos del sistema depende exclusivamente de la *matriz del sistema* \mathbf{A} .

11.2. Tipos de variables de estado

Un mismo sistema se puede modelizar de infinitas maneras en espacio de estado. El número n de estados puede variar para un mismo sistema dependiendo del modelo, pero conviene elegir el mínimo número posible para que no existan estados redundantes. En sistemas lineales de coeficientes constantes, el número mínimo de estados coincide con el orden del sistema.

11.2.1. Variables de fase

Cuando los estados de un sistema lo forman una variable temporal y sus sucesivas derivadas, se dice que el sistema posee variables de fase. Las variables de fase de un sistema de tercer orden son:

$$\mathbf{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix} = \begin{bmatrix} x_1(t) \\ \dot{x}_1(t) \\ \ddot{x}_1(t) \end{bmatrix} \quad (11.8)$$

Considérese un sistema de tercer orden con más polos que ceros¹:

$$\frac{y(s)}{u(s)} = \frac{K(\beta_2 s^2 + \beta_1 s + \beta_0)}{s^3 + \alpha_2 s^2 + \alpha_1 s + \alpha_0} \quad (11.9)$$

Definiendo una variable intermedia x_1 , la función de transferencia se puede dividir la en dos partes:

$$\frac{y(s)}{x_1(s)} \frac{x_1(s)}{u(s)} = (\beta_2 s^2 + \beta_1 s + \beta_0) \frac{K}{s^3 + \alpha_2 s^2 + \alpha_1 s + \alpha_0} \quad (11.10)$$

Con la primera parte, que recoge los ceros de la función de transferencia, se puede deducir la ecuación de salida:

$$\frac{y(s)}{x_1(s)} = \beta_2 s^2 + \beta_1 s + \beta_0 \quad (11.11)$$

$$y(s) = \beta_2 s^2 x_1(s) + \beta_1 s x_1(s) + \beta_0 x_1(s) \quad (11.12)$$

$$y(s) = \beta_2 x_3(s) + \beta_1 x_2(s) + \beta_0 x_1(s) \quad (11.13)$$

Con la segunda parte, en la que aparecen sólo los polos, se puede deducir una de las ecuaciones de estado:

$$\frac{x_1(s)}{u(s)} = \frac{K}{s^3 + \alpha_2 s^2 + \alpha_1 s + \alpha_0} \quad (11.14)$$

$$s^3 x_1(s) + \alpha_2 s^2 x_1(s) + \alpha_1 s x_1(s) + \alpha_0 x_1(s) = K u(s) \quad (11.15)$$

$$s x_3(s) + \alpha_2 x_3(s) + \alpha_1 x_2(s) + \alpha_0 x_1(s) = K u(s) \quad (11.16)$$

$$s x_3(s) = -\alpha_2 x_3(s) - \alpha_1 x_2(s) - \alpha_0 x_1(s) + K u(s) \quad (11.17)$$

El modelo completo en espacio de estado usando variables de fase es:

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \\ \dot{x}_3(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -\alpha_0 & -\alpha_1 & -\alpha_2 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ K \end{bmatrix} u(t) \quad (11.18)$$

$$y(t) = [\beta_0 \quad \beta_1 \quad \beta_2] \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix} \quad (11.19)$$

¹Todos los procesos físicos cumplen esta propiedad. En caso contrario, la matriz \mathbf{D} de la ecuación (11.1) no sería nula. Las funciones de transferencia que cumplen esta propiedad se llaman funciones *estrictamente propias*. Si tienen igual número de polos que de ceros, entonces se les conoce simplemente como funciones *propias*.

La ganancia K se puede trasladar a la ecuación de salida sin que el método pierda generalidad. La Fig. 11.2 muestra el diagrama de bloques del sistema usando variables de fase.

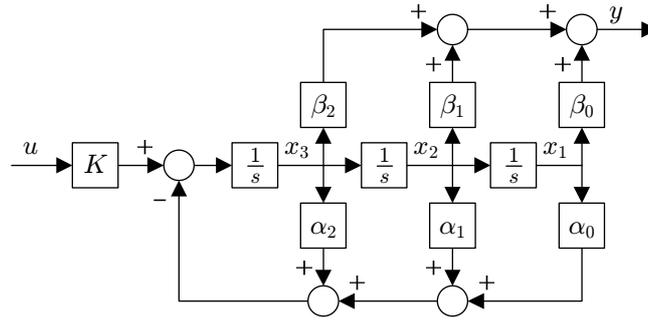


Figura 11.2: Modelo en espacio de estado usando variables de fase

Otra forma de usar las mismas variables de fase es:

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \\ \dot{x}_3(t) \end{bmatrix} = \begin{bmatrix} -\alpha_2 & 1 & 0 \\ -\alpha_1 & 0 & 1 \\ -\alpha_0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix} + \begin{bmatrix} K\beta_2 \\ K\beta_1 \\ K\beta_0 \end{bmatrix} u(t) \quad (11.20)$$

$$y(t) = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix} \quad (11.21)$$

11.2.2. Variables canónicas o normales

Las variables canónicas son adecuadas cuando la función de transferencia del sistema se puede descomponer en fracciones simples. Por ejemplo, la función de transferencia de un sistema de tercer orden con polos reales distintos es:

$$\frac{y(s)}{u(s)} = \frac{d_1}{s + \lambda_1} + \frac{d_2}{s + \lambda_2} + \frac{d_3}{s + \lambda_3} \quad (11.22)$$

La ecuación de salida se obtiene definiendo las variables canónicas z_i :

$$y(s) = d_1 \frac{u(s)}{s + \lambda_1} + d_2 \frac{u(s)}{s + \lambda_2} + d_3 \frac{u(s)}{s + \lambda_3} \quad (11.23)$$

$$y(s) = d_1 z_1(s) + d_2 z_2(s) + d_3 z_3(s) \quad (11.24)$$

Las ecuaciones de estado se obtienen de la propia definición de cada una de las variables canónicas:

$$z_i(s) = \frac{u(s)}{s + \lambda_i} \quad (11.25)$$

$$s z_i(s) = -\lambda_i z_i(s) + u(s) \quad (11.26)$$

Por tanto, las ecuaciones de estado usando variables canónicas queda:

$$\begin{bmatrix} \dot{z}_1(t) \\ \dot{z}_2(t) \\ \dot{z}_3(t) \end{bmatrix} = \begin{bmatrix} -\lambda_1 & 0 & 0 \\ 0 & -\lambda_2 & 0 \\ 0 & 0 & -\lambda_3 \end{bmatrix} \begin{bmatrix} z_1(t) \\ z_2(t) \\ z_3(t) \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} u(t) \quad (11.27)$$

$$y(t) = \begin{bmatrix} d_1 & d_2 & d_3 \end{bmatrix} \begin{bmatrix} z_1(t) \\ z_2(t) \\ z_3(t) \end{bmatrix} \quad (11.28)$$

La Fig. 11.3 muestra un modelo de sistema usando variables canónicas. Todos los estados están desacoplados. Este modo de representación se asemeja a la técnica de programación en paralelo para la simulación de funciones de transferencia.

La matriz \mathbf{A} del sistema es diagonal y sus elementos $-\lambda_i$ son los polos del sistema, como queda patente observando la ecuación característica:

$$q(s) = \det(s\mathbf{I} - \mathbf{A}) = \prod_{i=1}^n (s + \lambda_i) = 0, \quad (11.29)$$

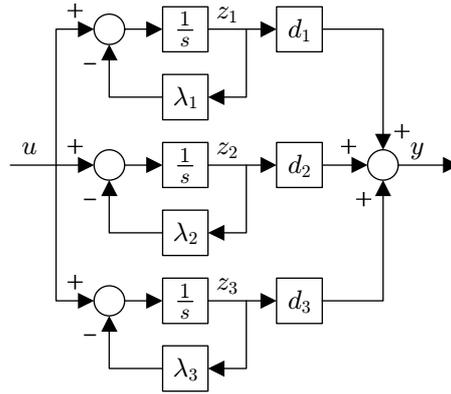


Figura 11.3: Modelo en espacio de estado usando variables canónicas

11.2.3. Variables físicas

Es posible usar variables físicas del sistema en el modelo de espacio de estado. Así por ejemplo, las ecuaciones diferenciales que gobiernan un motor de corriente continua son:

$$\begin{cases} v(t) = Ri(t) + Li'(t) + e(t) \\ e(t) = K\theta'(t) \\ \tau(t) = Ki(t) \\ \tau(t) = J\theta''(t) + B\theta'(t) \end{cases} \quad (11.30)$$

En este ejemplo se sigue la notación de Lagrange en lugar de la de Newton, es decir se ponen primas en lugar de puntos para señalar las derivadas, para usar sin confusión la variable $i(t)$.

Tomando como variables de estado el ángulo $\theta(t)$ del motor, la velocidad de giro $\theta'(t)$ y la intensidad $i(t)$, se puede construir el siguiente modelo:

$$\begin{bmatrix} \theta'(t) \\ \theta''(t) \\ i'(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & -\frac{B}{J} & \frac{K}{J} \\ 0 & -\frac{K}{L} & -\frac{R}{L} \end{bmatrix} \begin{bmatrix} \theta(t) \\ \theta'(t) \\ i(t) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \frac{1}{L} \end{bmatrix} v(t) \quad (11.31)$$

$$\theta(t) = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \theta(t) \\ \theta'(t) \\ i(t) \end{bmatrix} \quad (11.32)$$

La principal ventaja de usar variables físicas es que existen más probabilidades de poder medir los estados del sistema por medio de sensores.

A veces las variables físicas coinciden con las variables de fase o las canónicas. Por ejemplo, un sistema mecánico compuesto por una masa y un amortiguador se rige por la ecuación diferencial:

$$f(t) = m\ddot{x}(t) + b\dot{x}(t) \quad (11.33)$$

Tomando como estados las variables físicas de posición x y velocidad \dot{x} , se puede obtener como modelo físico del sistema en espacio de estado:

$$\begin{bmatrix} \dot{x}(t) \\ \ddot{x}(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & -\frac{b}{m} \end{bmatrix} \begin{bmatrix} x(t) \\ \dot{x}(t) \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{1}{m} \end{bmatrix} f(t) \quad (11.34)$$

$$x(t) = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} x(t) \\ \dot{x}(t) \end{bmatrix} \quad (11.35)$$

Como la velocidad es la derivada de la posición, se puede obtener el mismo resultado considerando la representación con variables de fase. También es posible comprobar que:

$$s\mathbf{I} - \mathbf{A} = \begin{bmatrix} s & -1 \\ 0 & s + \frac{b}{m} \end{bmatrix} \quad (11.36)$$

$$\frac{x(s)}{f(s)} = \mathbf{c}^T (s\mathbf{I} - \mathbf{A})^{-1} \mathbf{b} = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{s} & \frac{m}{ms^2 + bs} \\ 0 & \frac{m}{ms + b} \end{bmatrix} \begin{bmatrix} 0 \\ \frac{1}{m} \end{bmatrix} = \frac{1}{ms^2 + bs} \quad (11.37)$$

11.3. Controlabilidad y observabilidad

Para diseñar controladores en espacio de estado, primero hay que averiguar si el sistema es observable y controlable. Se dice que un sistema es *controlable* si existe una entrada u capaz de variar el estado inicial \mathbf{x}_0 a cualquier otro estado deseado \mathbf{x}_t en un tiempo finito. Matemáticamente, un sistema es controlable si su *matriz de controlabilidad* \mathbf{P}_c no es singular. La matriz de controlabilidad se define como

$$\mathbf{P}_c = [\mathbf{b} \quad \mathbf{A}\mathbf{b} \quad \mathbf{A}^2\mathbf{b} \quad \dots \quad \mathbf{A}^{n-1}\mathbf{b}], \quad (11.38)$$

y el sistema es controlable si y sólo si el $\det(\mathbf{P}_c) \neq 0$.

Un sistema controlable es siempre *estabilizable*. Por otro lado, se dice que un sistema es *observable* si cualquier estado \mathbf{x}_0 puede ser determinado mediante la salida del sistema y durante un tiempo finito. Matemáticamente, un sistema es observable si su *matriz de observabilidad* \mathbf{P}_o no es singular. La matriz de observabilidad se define como

$$\mathbf{P}_o = \begin{bmatrix} \mathbf{c}^T \\ \mathbf{c}^T \mathbf{A} \\ \mathbf{c}^T \mathbf{A}^2 \\ \dots \\ \mathbf{c}^T \mathbf{A}^{n-1} \end{bmatrix}, \quad (11.39)$$

y el sistema es observable si y sólo si el $\det(\mathbf{P}_o) \neq 0$. Se puede demostrar que el paso de espacio de estado a función de transferencia sólo es posible si el sistema es controlable y observable.

El sistema con masa y amortiguador, que apareció anteriormente como ejemplo, es controlable y observable. Sus matrices de controlabilidad y observabilidad son:

$$\mathbf{b} = \begin{bmatrix} 0 \\ \frac{1}{m} \end{bmatrix} \quad \text{y} \quad \mathbf{A}\mathbf{b} = \begin{bmatrix} 0 & 1 \\ 0 & -\frac{b}{m} \end{bmatrix} \begin{bmatrix} 0 \\ \frac{1}{m} \end{bmatrix} = \begin{bmatrix} \frac{1}{m} \\ -\frac{b}{m^2} \end{bmatrix} \quad (11.40)$$

$$\mathbf{P}_c = \begin{bmatrix} 0 & \frac{1}{m} \\ \frac{1}{m} & -\frac{b}{m^2} \end{bmatrix} \quad (11.41)$$

$$\mathbf{c}^T = [1 \quad 0] \quad \text{y} \quad \mathbf{c}^T \mathbf{A} = [1 \quad 0] \begin{bmatrix} 0 & 1 \\ 0 & -\frac{b}{m} \end{bmatrix} = [0 \quad 1] \quad (11.42)$$

$$\mathbf{P}_o = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad (11.43)$$

El sistema es controlable porque el $\det(\mathbf{P}_c) = -\frac{1}{m^2} \neq 0$; y es observable porque el $\det(\mathbf{P}_o) = 1 \neq 0$.

11.4. Realimentación completa de estados

Cuando son medibles todos los estados de un sistema, Fig. 11.4, se puede definir una realimentación $\mathbf{k}^T = [k_1 \ k_2 \ \dots \ k_n]$ que asigna una ganancia a cada estado. El objetivo de obtener un comportamiento específico entre la referencia r y la salida y . Las ecuaciones del sistema controlado quedan de forma:

$$\begin{cases} \dot{\mathbf{x}}(t) = (\mathbf{A} - \mathbf{b}\mathbf{k}^T)\mathbf{x}(t) + k_r \mathbf{b}r(t) \\ y(t) = \mathbf{c}^T \mathbf{x}(t) \end{cases} \quad (11.44)$$

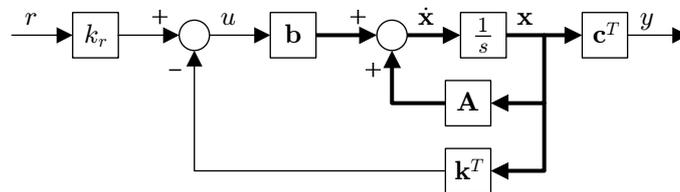


Figura 11.4: Realimentación completa de estados

La nueva ecuación característica del sistema es

$$q_d(s) = \det(s\mathbf{I} - \mathbf{A} + \mathbf{b}\mathbf{k}^T) = 0, \quad (11.45)$$

donde la posición de los polos del sistema controlado depende de la nueva matriz del sistema $\mathbf{A} - \mathbf{b}\mathbf{k}^T$. La ganancia k_r no modifica la posición de los polos, y por tanto tampoco el régimen transitorio, pero sí el régimen permanente. Para poder colocar todos los polos del sistema controlado en posiciones arbitrarias, la condición necesaria y suficiente que el sistema sea controlable.

11.4.1. Asignación de polos

El ajuste de las ganancias de realimentación de estados se puede hacer de distintos modos. En este apartado se realiza una asignación de polos en lazo cerrado de acuerdo con unas especificaciones de diseño.

Si en el ejemplo de la masa y amortiguador se desea un comportamiento críticamente amortiguado y frecuencia natural ω_n rad/s, los dos polos del sistema compensado deben ser reales en $s = -\omega_n$. Usando variables de fase, una matriz que cumpliría este requerimiento es

$$\mathbf{A} - \mathbf{b}\mathbf{k}^T = \begin{bmatrix} 0 & 1 \\ -\omega_n^2 & -2\omega_n \end{bmatrix}. \quad (11.46)$$

Por tanto:

$$\begin{bmatrix} 0 & 1 \\ 0 & -\frac{b}{m} \end{bmatrix} - \begin{bmatrix} 0 \\ \frac{1}{m} \end{bmatrix} [k_1 \quad k_2] = \begin{bmatrix} 0 & 1 \\ -\omega_n^2 & -2\omega_n \end{bmatrix} \quad (11.47)$$

$$\begin{bmatrix} 0 & 1 \\ -\frac{k_1}{m} & -\frac{b+k_2}{m} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\omega_n^2 & -2\omega_n \end{bmatrix} \quad (11.48)$$

Las ganancias del vector de realimentación \mathbf{k}^T deben tomar los valores $k_1 = m\omega_n^2$ y $k_2 = 2m\omega_n - b$. Este método se podría hacer de forma sistemática en el caso de que $\mathbf{b} = [0 \quad 0 \quad \dots \quad 1]^T$ y conociendo los coeficientes de la ecuación característica del sistema sin compensar:

$$q(s) = s^n + a_{n-1}s^{n-1} + a_{n-2}s^{n-2} + \dots + a_2s^2 + a_1s + a_0 = 0, \quad (11.49)$$

y los de la ecuación característica del sistema compensado:

$$q_d(s) = s^n + b_{n-1}s^{n-1} + b_{n-2}s^{n-2} + \dots + b_2s^2 + b_1s + b_0 = 0. \quad (11.50)$$

En ese caso, el vector de realimentación es:

$$\mathbf{k}^T = [k_1 \quad k_2 \quad \dots \quad k_n] = [b_0 - a_0 \quad b_1 - a_1 \quad \dots \quad b_{n-1} - a_{n-1}] \quad (11.51)$$

11.4.2. Método de Ackermann

Otro método de asignación de polos es usar la *fórmula de Ackermann* [2]:

$$\mathbf{k}^T = [k_1 \quad k_2 \quad \dots \quad k_n] = [0 \quad 0 \quad \dots \quad 1] \mathbf{P}_c^{-1} q_d(\mathbf{A}), \quad (11.52)$$

donde \mathbf{P}_c es la matriz de controlabilidad y

$$q_d(\mathbf{A}) = \mathbf{A}^n + b_{n-1}\mathbf{A}^{n-1} + b_{n-2}\mathbf{A}^{n-2} + \dots + b_2\mathbf{A}^2 + b_1\mathbf{A} + b_0\mathbf{I}. \quad (11.53)$$

11.4.3. Controlador óptimo cuadrático

Aunque las especificaciones de diseño suelen referirse habitualmente en esta asignatura a determinados regímenes transitorios, y por tanto a localización de polos en lazo cerrado, a veces el objetivo del controlador es obtener una respuesta estable que minimice una función de coste con múltiples interpretaciones físicas (consumo de combustible, potencia de entrada, etc.). Para sistemas MIMO, la función de coste se suele expresar como:

$$J = \int_0^\infty (\mathbf{x}^T \mathbf{Q} \mathbf{x} + \mathbf{u}^T \mathbf{R} \mathbf{u}) dt \quad (11.54)$$

En los sistemas SISO que nos ocupan en este manual, la función de coste se reduce:

$$J = \int_0^\infty (\mathbf{x}^T \mathbf{Q} \mathbf{x} + ru^2) dt \quad (11.55)$$

Como parámetro, la función de coste consta de dos elementos: el primer sumando penaliza las desviaciones de los estados con respecto a ciertos niveles que se consideren deseables, mientras que el segundo sumando penaliza la energía de la señal de control. Evidentemente, dependiendo de los valores numéricos de \mathbf{Q} y r , tendrá más peso uno u otro sumando.

El vector de realimentación completa de estados \mathbf{k}^T que minimiza la función de coste se conoce como controlador óptimo del sistema o *regulador lineal cuadrático* (LQR). Su expresión es:

$$\mathbf{k}^T = r^{-1} \mathbf{b}^T \mathbf{P}, \quad (11.56)$$

donde \mathbf{P} es una matriz simétrica real que proviene de resolver la ecuación matricial de Riccati:

$$\mathbf{A}^T \mathbf{P} + \mathbf{P} \mathbf{A} - \mathbf{P} \mathbf{b} r^{-1} \mathbf{b}^T \mathbf{P} + \mathbf{Q} = \mathbf{0}. \quad (11.57)$$

11.5. Realimentación parcial de estados

En muchas ocasiones no es posible medir todos los estados de un sistema. En lugar de los n estados \mathbf{x} , sólo son accesibles m medidas \mathbf{z} (con $m < n$), que serán combinación lineal de los estados,

$$\mathbf{z} = \mathbf{E}\mathbf{x}, \quad (11.58)$$

donde la matriz \mathbf{E} no es cuadrada, sino de orden $m \times n$.

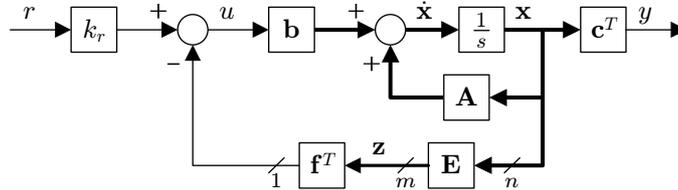


Figura 11.5: Realimentación parcial de estados

La ley de control mediante realimentación parcial de estados tiene la forma que se muestra en la Fig. 11.5. Ahora, el vector de realimentación \mathbf{f}^T sólo posee m ganancias: f_1, f_2, \dots, f_m . Por tanto, no es posible situar todos los polos del sistema controlado en posiciones arbitrarias y hay que buscar una solución de compromiso. La nueva ecuación característica del sistema es:

$$q_d(s) = \det(s\mathbf{I} - \mathbf{A} + \mathbf{b}\mathbf{f}^T\mathbf{E}) = 0. \quad (11.59)$$

11.6. Observadores de estado

La limitación que posee la realimentación parcial de estados, llevó al desarrollo de herramientas matemáticas que estimaran de forma precisa todos los estados del sistema de cara a su posterior empleo en una realimentación completa de estados observados. Aunque estos *observadores de estado* pueden desearse en sí mismos para la monitorización del sistema.

En este apartado se describe el observador de estados que conoce la entrada u y la salida y del sistema. Dicho observador intenta anular el error que existe entre la salida real y la estimada mediante un lazo de realimentación, Fig. 11.6. El sistema de ecuaciones que gobierna el observador es:

$$\begin{cases} \dot{\hat{\mathbf{x}}} = \hat{\mathbf{A}}\hat{\mathbf{x}} + \hat{\mathbf{b}}u + \mathbf{l}(y - \hat{y}) \\ \hat{y} = \mathbf{c}^T\hat{\mathbf{x}} \end{cases} \quad (11.60)$$

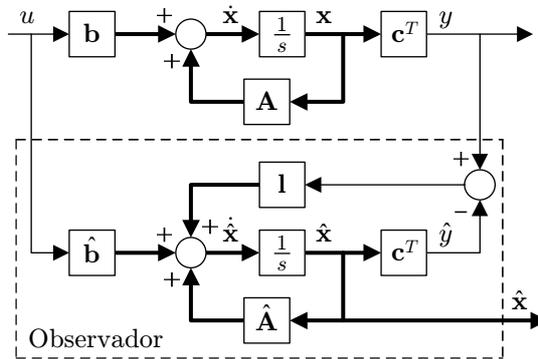


Figura 11.6: Observador de estados

El lazo de realimentación posee un vector \mathbf{l} de ganancias proporcionales. Se define la variación del error de la estimación como:

$$\dot{\mathbf{e}} = \dot{\mathbf{x}} - \dot{\hat{\mathbf{x}}} = (\mathbf{b} - \hat{\mathbf{b}})u + \mathbf{A}\mathbf{x} - \hat{\mathbf{A}}\hat{\mathbf{x}} - \mathbf{l}(y - \hat{y}) \quad (11.61)$$

$$\dot{\mathbf{e}} = (\mathbf{b} - \hat{\mathbf{b}})u + \mathbf{A}\mathbf{x} - \hat{\mathbf{A}}\hat{\mathbf{x}} - \mathbf{l}\mathbf{c}^T(\mathbf{x} - \hat{\mathbf{x}}) \quad (11.62)$$

Suponiendo $\mathbf{A} \approx \hat{\mathbf{A}}$ y $\mathbf{b} \approx \hat{\mathbf{b}}$:

$$\dot{\mathbf{e}} \approx (\mathbf{A} - \mathbf{l}\mathbf{c}^T)(\mathbf{x} - \hat{\mathbf{x}}) \quad (11.63)$$

$$\dot{\mathbf{e}} \approx (\mathbf{A} - \mathbf{l}\mathbf{c}^T)\mathbf{e} \quad (11.64)$$

El error tiende a cero si los valores propios de la matriz $\mathbf{A} - \mathbf{l}\mathbf{c}^T$ tienen parte real negativa. Y dicho error tenderá a cero de forma más o menos rápida de acuerdo con la posición de los polos de la siguiente ecuación característica:

$$q_e(s) = \det(s\mathbf{I} - \mathbf{A} + \mathbf{l}\mathbf{c}^T) = 0. \quad (11.65)$$

Ackermann propuso otra fórmula para asignar los polos del observador:

$$\mathbf{l} = [l_1 \ l_2 \ \dots \ l_n]^T = q_e(\mathbf{A})\mathbf{P}_o^{-1} [0 \ 0 \ \dots \ 1]^T, \quad (11.66)$$

Otros autores han propuesto diferentes formas de encontrar los valores para el vector \mathbf{l} . Kalman [3] [4] propone su valor óptimo teniendo en cuenta la varianza del ruido en la medida de la actuación u y la salida y .

$$\begin{cases} \dot{\hat{\mathbf{x}}} = \hat{\mathbf{A}}\hat{\mathbf{x}} + \hat{\mathbf{b}}(u + w) + \mathbf{l}(y + v - \hat{y}) \\ \hat{y} = \mathbf{c}^T\hat{\mathbf{x}} \end{cases} \quad (11.67)$$

La señal w es el ruido que se solapa a la actuación y v es el ruido que se solapa a la salida. Se supone que ambos ruidos están desacoplados y siguen distribuciones normales de media cero. El filtro Kalman se usa a veces para estimar la derivada de la salida sin introducir una diferenciación.

11.7. Realimentación completa de estados observados

Los observadores se pueden usar para cerrar el lazo de control usando todos los estados aunque sólo se pueda medir la entrada u la salida y . La Fig. 11.7 muestra la estrategia de control. Diseñando adecuadamente el observador, con las reglas que se han propuesto en el apartado anterior, se asegura que el controlador actúa prácticamente sobre los estados del sistema.

Se puede aplicar el principio de superposición: diseñar el observador como si no existiera el controlador y el controlador como si actuara directamente sobre la planta. Sólo hay que tener cuidado de colocar los polos del observador más rápidos que los de la planta; por ejemplo 10 veces más rápidos.

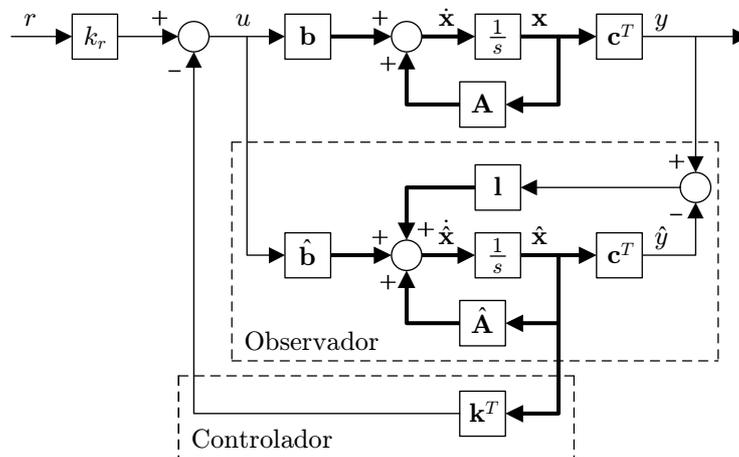


Figura 11.7: Control mediante realimentación completa de estados

Evidentemente, el comportamiento del sistema será mejor si el controlador actúa directamente sobre los estados del sistema, porque se pueden medir, que si actúa sobre las estimaciones que calcula el observador.

11.8. Ejercicios propuestos

- **Ejercicio 1:** Un determinado sistema se comporta de acuerdo con la siguiente ecuación diferencial:

$$\ddot{y}(t) - 0.03y(t) = 0.9u(t) \quad (11.68)$$

Donde la señal $u(t)$ es la entrada y la señal $y(t)$ es la salida del sistema. Se pide:

- a) Determinar el modelo en espacio de estado del sistema —matrices \mathbf{A} , \mathbf{b} y \mathbf{c} — usando variables de fase.
- b) Calcular la matriz de controlabilidad \mathbf{P}_c del sistema. ¿Es el sistema controlable?
- c) Calcular la matriz de observabilidad \mathbf{P}_o del sistema. ¿Es el sistema observable?
- d) Determinar —si es posible— la función de transferencia del sistema. ¿Es estable el sistema?
- e) Determinar el vector de ganancias \mathbf{k} —utilizando la fórmula de Ackermann— que se debe usar en la realimentación de estados, para que los polos del sistema controlado sean las raíces de la ecuación:

$$q_d(s) = s^2 + 6s + 18 \quad (11.69)$$

- **Ejercicio 2:** Un determinado sistema se comporta de acuerdo con la siguiente ecuación diferencial:

$$\left. \begin{aligned} \dot{x}(t) + x(t) + y(t) &= u(t) \\ \dot{y}(t) + 2y(t) &= 0 \end{aligned} \right\} \quad (11.70)$$

Donde los estados del sistema son $y(t)$ y $x(t)$, la señal $u(t)$ es la entrada y la señal $y(t)$ es la salida del sistema. Se pide:

- a) Determinar el modelo en espacio de estado del sistema, es decir, las matrices \mathbf{A} , \mathbf{b} y \mathbf{c} .
- b) Calcular la matriz de controlabilidad \mathbf{P}_c del sistema. ¿Es el sistema controlable?
- c) Calcular la matriz de observabilidad \mathbf{P}_o del sistema. ¿Es el sistema observable?
- d) Determinar —si es posible— la función de transferencia del sistema. ¿Es estable el sistema?
- e) Determinar el vector de ganancias \mathbf{k} —utilizando la fórmula de Ackermann— que se debe usar en la realimentación de estados, para que los polos del sistema controlado sean las raíces de la ecuación:

$$q_d(s) = s^2 + 6s + 18 \quad (11.71)$$

- **Ejercicio 3:** El movimiento de un sistema viene definido por la siguiente función de transferencia:

$$G(s) = \frac{2}{s(s+1)} \quad (11.72)$$

Se pide:

- a) Escribir el modelo en espacio de estado, usando variables de fase.
- b) Diseñar un controlador por la fórmula de Ackermann de tal forma que los polos del sistema controlado estén en $-2 \pm 2j$.

- **Ejercicio 4:** El comportamiento de un sistema viene definido por la siguiente función de transferencia:

$$G(s) = \frac{2}{(s+1)(s+2)} \quad (11.73)$$

Se pide:

- a) Escribir el modelo en espacio de estado, usando variables canónicas.
- b) Diseñar un controlador de realimentación completa de estados usando la fórmula de Ackermann de forma que los polos del sistema controlado estén en $-3 \pm 3j$.

Parte II

Control de sistemas muestreados

Capítulo 12

Introducción

En la Parte I de estos apuntes se han descrito las herramientas clásicas de control para sistemas continuos en el tiempo. Todas ellas se pueden aplicar directamente, por ejemplo, en el diseño de circuitos analógicos que gobiernan cualquier tipo de sistema físico. Sin embargo, la aparición de los microprocesadores supuso una auténtica revolución en el mundo de la ingeniería de control. Con su uso, el ingeniero es capaz de ajustar o cambiar la ley de control de mucha mucho más rápida. Además, la potencia cálculo de estos dispositivos ha permitido desarrollar nuevas estrategias de control.

La única dificultad que plantea el uso de microprocesadores es que es necesario trasladar todos los conceptos clásicos de control a un nuevo escenario en el que las señales no son conocidas en todo instante de tiempo. A este tipo de sistemas se les llamará sistemas de tiempo discreto.

El presente libro de texto describe las herramientas básicas para el control de sistemas discretos en el tiempo. Se introducirá el concepto de muestreo y la herramienta matemática que se usará para manejar las señales muestreadas es la transformada Z .

12.1. Ejemplo de implementación analógica

Sea el sistema de control de la Fig. 12.1, en el que el controlador es un compensador de adelanto-retraso de fase. Si las señales de error y actuación son señales eléctricas, el ingeniero puede implementar físicamente el circuito de la Fig. 12.2 para introducir su ley de control. Para ajustar los parámetros del controlador a los deseados, debe sintonizarse de forma adecuada los potenciómetros del circuito.

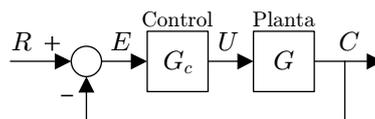


Figura 12.1: Sistema de control

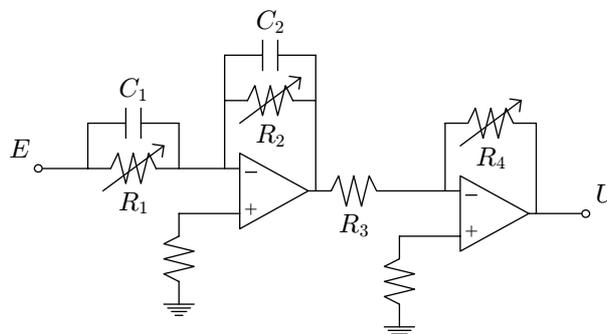


Figura 12.2: Circuito analógico para un compensador de adelanto

Los potenciómetros del circuito permiten al ingeniero modificar, dentro de unos límites, la ley del compensador. Sin embargo, si el ingeniero desea probar otro tipo de compensador, tendrá que soldar un nuevo circuito. Esta forma de diseñar controladores analógicos no es muy rápida y suele ser bastante tediosa.

12.2. Ejemplo de implementación digital

En la Fig. 12.3 se muestra la alternativa digital para el mismo problema. El controlador se sustituye en este caso por un microprocesador que es capaz de adquirir la magnitud del error por los puertos de entrada (generalmente son convertidores A/D, pero también pueden ser contadores de pulsos, de encoder, etc.) y puede comandar la actuación en la planta a través de los puertos de salida (convertidores D/A).

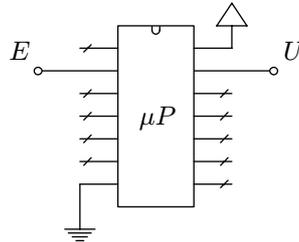


Figura 12.3: Microprocesador que controla el sistema

La operación del microprocesador está gobernada por un reloj interno (aunque algunos microprocesadores lo tienen externo), que marca los instantes en los que se ejecutan las sentencias del programa que puede introducir el ingeniero (de nuevo este programa puede estar almacenado en una memoria interna o externa). El reloj es quien señala también con qué frecuencia se produce la lectura de los convertidores A/D y el comando de las salidas D/A.

Si el ingeniero tiene implementado un determinado programa de control y desea cambiar el algoritmo, sólo debe modificar unas líneas de código y volverlo a compilar. No es necesario que cambie el hardware ni las conexiones. Esto hace que, a la hora del diseño, el control digital sea mucho más versátil que el analógico.

El ejemplo sirve también para plantearse la siguiente pregunta: ¿qué datos puede usar el ingeniero en su programa para calcular la salida o actuación del controlador? Evidentemente en cada instante que marque el reloj puede usar la lectura actual de la entrada. También puede disponer de las entradas y las salidas en instantes anteriores, si ha tenido el cuidado de almacenarlas en variables. Nunca podrá usar valores futuros de esas señales. Y como la memoria del dispositivo es finita, tampoco podrá disponer de infinitos valores de entrada y salida anteriores en el tiempo. Tendrá que decidir si guarda dos, tres, diez, valores pasados... pero no puede disponer de un "histórico infinito". Por tanto, el valor de la salida (es decir, la actuación del controlador) en cada instante será una combinación de un número finito de valores, que corresponderán a entradas actuales o anteriores en el tiempo, y salidas anteriores en el tiempo.

¿Es posible con esa información controlar el sistema de la misma manera que se hacía con el circuito anterior? Y en el caso de que sea posible, ¿cuál es la combinación matemática de todos esos valores para se controle de forma equivalente el sistema? Evidentemente la respuesta a la primera pregunta dependerá de la frecuencia del reloj que gobierna el controlador. No se puede esperar que un microprocesador sea capaz de controlar el movimiento de un cabezal del disco duro, si el reloj del mismo ordena la ejecución del programa cada minuto. Pero para controlar la temperatura del interior de un edificio sí puede ser suficiente con esa frecuencia... Se ve cómo la frecuencia de ejecución del programa de control es una decisión clave del ingeniero. En esta asignatura se darán las herramientas necesarias para controlar los sistemas de estas características.

12.3. Concepto de muestreo

Una señal continua en el tiempo se encuentra *muestreada* por un observador exterior cuando éste conoce sólo información de dicha señal en determinados momentos de tiempo. En la Fig. 12.4 a) se representa una señal continua, mientras que en Fig. 12.4 b) se ha representado esa misma señal muestreada a intervalos.

Las señales pueden ser muestreadas de muy diversas maneras. Lo más habitual será que el observador exterior posea información puntual de la señal continua a intervalos regulares de tiempo. A la duración de esos intervalos se le llama *periodo de muestreo* y se designará con la variable T_s o simplemente T . Para la frecuencia de muestreo se usará f_s (en Hz) o bien ω_s (en rad/s).

El periodo de muestreo va a ser crucial en esta asignatura. El comportamiento de los sistemas cambia con dicho periodo de muestreo. De hecho, un mismo sistema puede ser estable o no, dependiendo de qué periodo de muestreo se haya elegido.

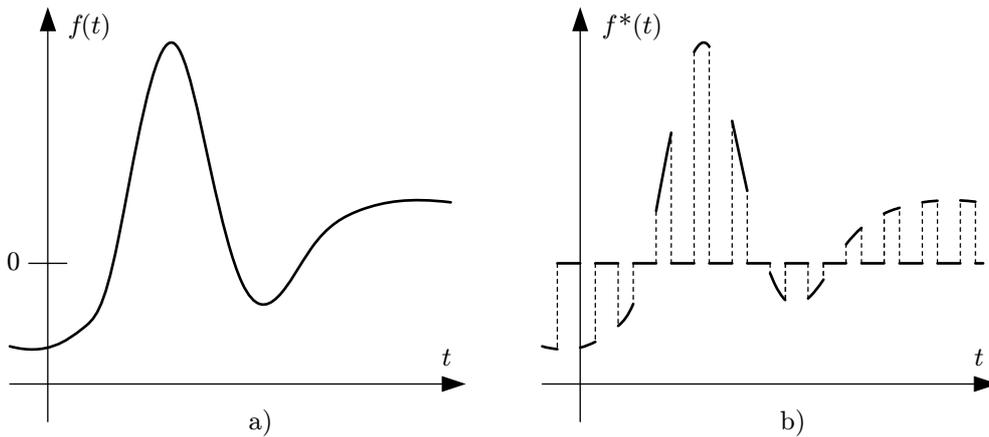


Figura 12.4: Función continua (a) y muestreada (b)

12.4. Concepto de cuantización

Los sistemas de medición que llevan incorporados los microprocesadores (convertidores A/D, contadores de pulsos, etc.) conllevan intrínsecamente la cuantización de la señal. Es decir, no se pueden conocer infinitos decimales de la señal física que se pretende medir. En algún momento se debe truncar la magnitud medida. El momento en que se produce el truncamiento determina la resolución del sistema sensor.

Por ejemplo, si se desea medir el voltaje de una señal eléctrica con un convertidor de 14 bits, la resolución que se obtiene cuando el fondo de escala es de -10 a $+10$ V, es 1.22 mV. Cambios en la tensión por debajo de este umbral no serán detectables por el sistema. Consecuencia de la cuantización es que la señal medida ya no es continua en el tiempo. Es continua a intervalos y “escalonada”.

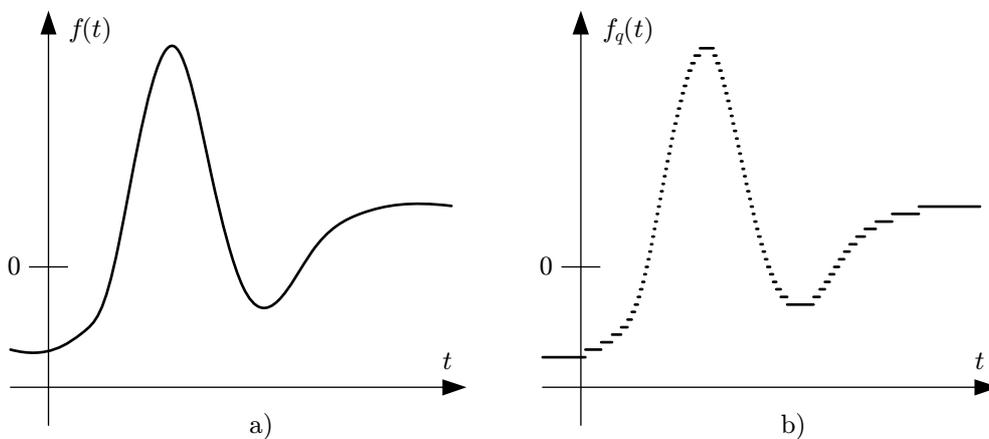


Figura 12.5: Función continua (a) y cuantizada (b)

Aunque el fenómeno de la cuantización puede ser crítico en determinadas circunstancias, lo habitual es que se pueda despreciar. Es decir, si el controlador diseñado es robusto y consigue que el sistema siga la referencia a pesar la existencia del ruido o las perturbaciones, mantendrá esa característica a pesar de la existencia de la cuantización. De hecho, en muchos manuales, la cuantización se estudia como un ruido aleatorio que se suma a la señal original.

12.5. Clasificación de los sistemas

Una vez introducidos los conceptos de muestreo y cuantización, es posible clasificar los sistemas en:

- **Sistemas continuos.** Son aquellos en los que sólo intervienen señales continuas en el tiempo.
- **Sistemas muestreados o discretos en el tiempo.** Son aquellos en los que existe al menos un proceso de muestreo en alguna de las señales que intervienen en el mismo.

- **Sistemas cuantizados.** Son aquellos en los que existe al menos un proceso de cuantización de alguna de las señales que intervienen en el mismo.
- **Sistemas digitales.** Son aquellos en los que existe a la vez al menos un proceso de muestreo y cuantización de alguna de las señales que intervienen en el mismo.

Evidentemente, un sistema se puede dividir en partes que posean distintas características. Así, en un sistema muestreado se puede definir un subsistema que sea estrictamente continuo. Y lo mismo puede decirse de los sistemas digitales y cuantizados.

Hay que advertir que los adjetivos “muestreado”, “discreto” y “digital” se suelen usar como sinónimos. Esto se debe al hecho antes citado de que el fenómeno de la cuantización se pueda despreciar en la mayoría de los casos.

Capítulo 13

Tratamiento matemático de la señal muestreada

En este capítulo se presenta el tratamiento matemático que se dará a la señal muestreada en el dominio temporal y, mucho más importante, cuál es su transformada de Fourier. Todo lo que se diga de la transformada de Fourier se podrá trasladar más adelante a la transformada de Laplace.

13.1. Definición de muestreo periódico

Sea $f(t)$ una función continua en el tiempo. De dicha función, Fig. 13.1 a), un observador sólo conoce su evolución en pequeños intervalos de duración a a cada cierto tiempo T , con $a < T$, como se muestra en la Fig. 13.1 b).

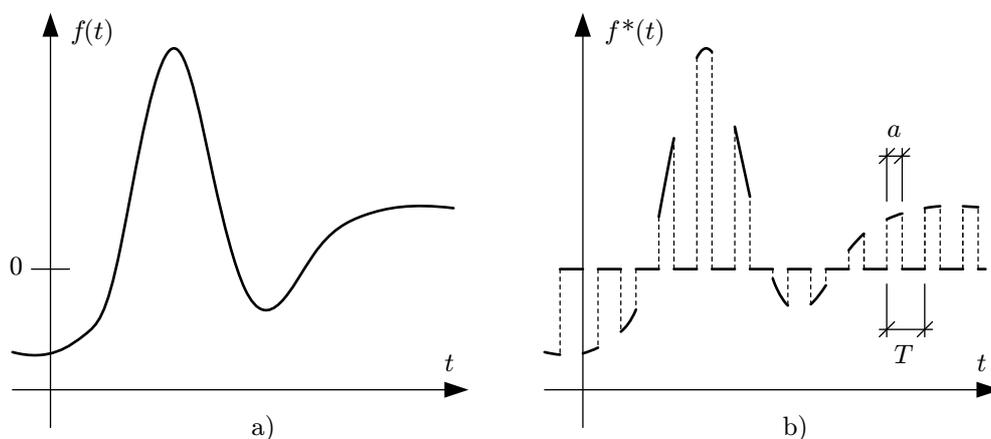


Figura 13.1: Función continua (a) y muestreada (b)

13.1.1. Función portadora

El resultado final se puede considerar como resultado de multiplicar la función original $f(t)$ por otra función $p(t)$ que se define como se muestra en la Fig. 13.2. Así pues, desde un punto de vista matemático el muestreo de una función consiste simplemente en un producto de funciones. A semejanza de un modulador, $f(t)$ sería la señal moduladora y $p(t)$ la señal portadora.

Dado que $p(t)$ es una función periódica se puede definir de dos maneras. En la primera (13.1) se dan los valores de la función dentro de su periodo, mientras que la segunda (13.2) es su desarrollo en serie de Fourier.

$$p(t) = \begin{cases} 1 & \text{si } 0 \leq t < a \\ 0 & \text{si } a \leq t < T \end{cases} \quad (13.1)$$

$$p(t) = \sum_{n=-\infty}^{\infty} c_n e^{jn\omega_s t} \quad (13.2)$$

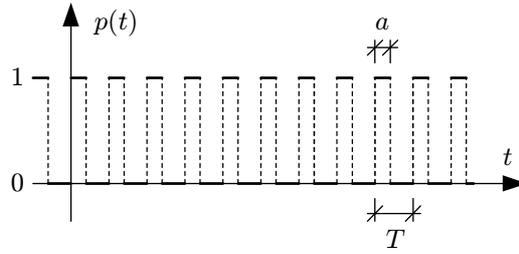


Figura 13.2: Función periódica que produce el muestreo

Es posible calcular el valor de los coeficientes de Fourier para la función $p(t)$ elegida:

$$c_n = \frac{1}{T} \int_0^T p(t) e^{-jn\omega_s t} dt = \frac{1}{T} \int_0^a e^{-jn\omega_s t} dt = \frac{1}{jn\omega_s T} [e^{-jn\omega_s t}]_0^a = \frac{1 - e^{-jn\omega_s a}}{jn\omega_s T} \quad (13.3)$$

Para ver qué valores van tomando los distintos coeficientes en función de n , es interesante modificar convenientemente la expresión (13.3). Operando con números complejos, se sabe que:

$$2j \sin \alpha = e^{j\alpha} (1 - e^{-2j\alpha}) \quad (13.4)$$

Por tanto se puede afirmar que:

$$1 - e^{-jn\omega_s a} = \frac{2j \sin \frac{n\omega_s a}{2}}{e^{j\frac{n\omega_s a}{2}}} = 2j e^{-j\frac{n\omega_s a}{2}} \sin \frac{n\omega_s a}{2} \quad (13.5)$$

Y la expresión de los coeficientes de Fourier queda así:

$$c_n = \frac{1}{jn\omega_s T} 2j e^{-j\frac{n\omega_s a}{2}} \sin \frac{n\omega_s a}{2} = \frac{a \sin \frac{n\omega_s a}{2}}{T \frac{n\omega_s a}{2}} e^{-j\frac{n\omega_s a}{2}} \quad (13.6)$$

A la función seno de un ángulo dividido por el mismo ángulo se le llama función *sinus cardinalis* (abreviado “sinc”) o función de interpolación:

$$\frac{\sin \alpha}{\alpha} = \text{sinc } \alpha \quad (13.7)$$

La representación de dicha función, dependiendo de la variable, es:

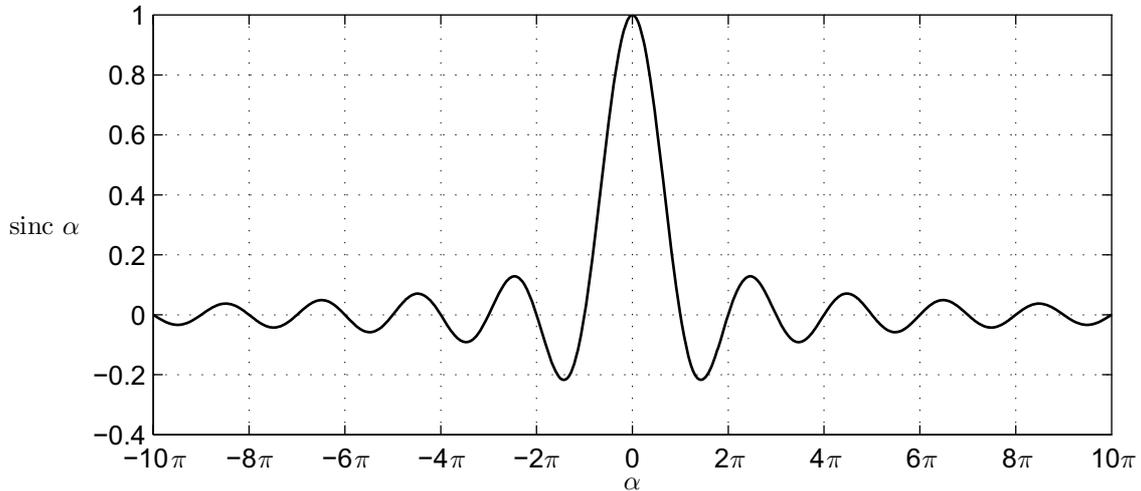


Figura 13.3: Función *sinus cardinalis*

En definitiva, los coeficientes de Fourier de la función $p(t)$ son números complejos cuyo módulo y argumento se pueden obtener fácilmente de la expresión (13.6). Usando otra notación habitual para los números complejos, ecuación (13.8), se advierte que el argumento da saltos finitos de 180° cada vez que la función *sinus cardinalis* cambia de signo.

$$c_n = \frac{a}{T} e^{-j\frac{n\omega_s a}{2}} \text{sinc} \frac{n\omega_s a}{2} \quad (13.8)$$

En la Fig. 13.4 se señalan con círculos los valores de módulo y argumento que van tomando de los coeficientes c_n de Fourier.

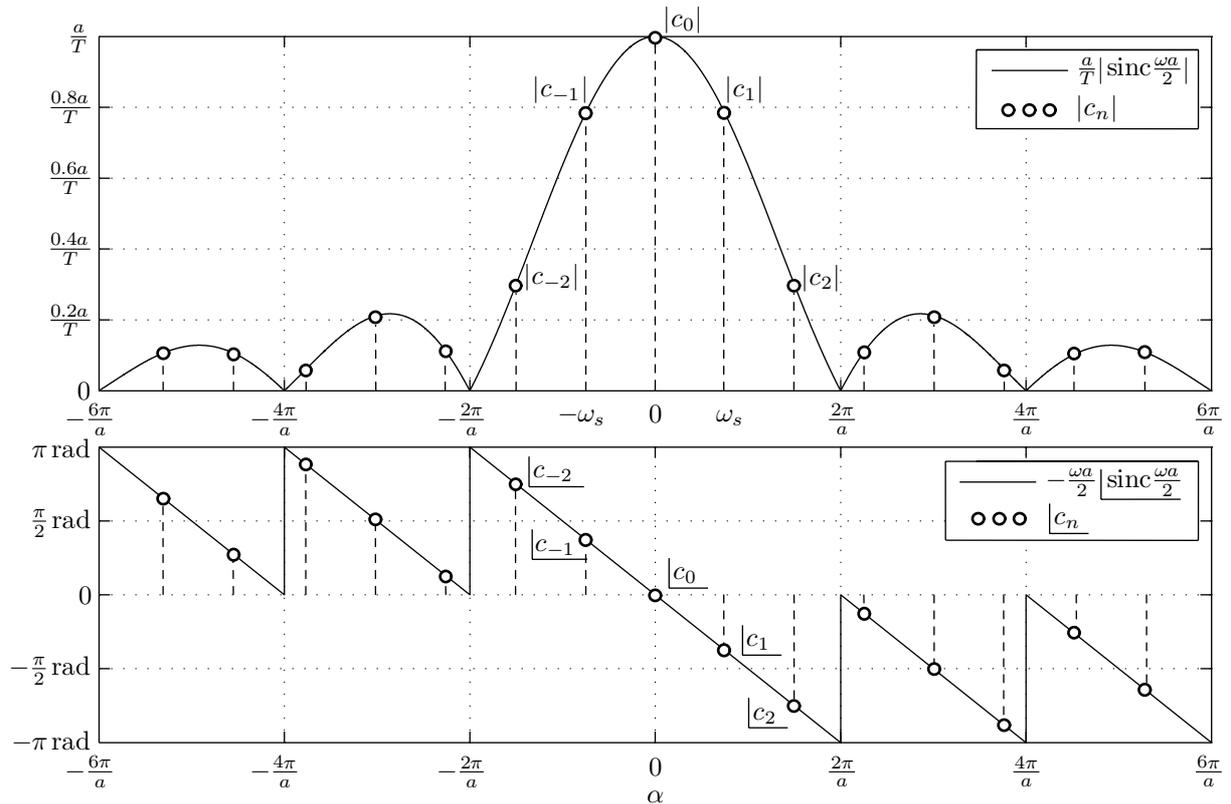


Figura 13.4: Módulo y argumento de los coeficientes de Fourier c_n de la función portadora

13.1.2. Función temporal muestreada

Mediante la función periódica $p(t)$ definida en el apartado anterior, se puede dar una expresión matemática para la función temporal muestreada:

$$f^*(t) = f(t)p(t) = f(t) \sum_{n=-\infty}^{\infty} c_n e^{jn\omega_s t} = \sum_{n=-\infty}^{\infty} c_n f(t) e^{jn\omega_s t} \quad (13.9)$$

Siempre se usará el símbolo asterisco para denotar la función resultante después del muestreo. La expresión (13.9) será de mucha utilidad para calcular en el siguiente apartado su transformada de Fourier.

13.2. Transformada de Fourier de la función muestreada

Por definición, la transformada de Fourier de una función es:

$$F(\omega) = \mathcal{F} [f(t)] = \int_{-\infty}^{\infty} f(t) e^{-j\omega t} dt \quad (13.10)$$

Es una función compleja de variable real que ofrece información del contenido en frecuencia de la función original $f(t)$. Si la variable t tiene unidades de segundos, la variable ω tiene unidades de rad/s. Las dos variables reales, t y ω , se extienden desde $-\infty$ a ∞ .

La propiedad de traslación de la transformada de Fourier dice que:

$$\mathcal{F} [f(t)e^{j\omega_1 t}] = F(\omega - \omega_1) \quad (13.11)$$

Por tanto, la transformada de Fourier de la función muestreada es:

$$F^*(\omega) = \mathcal{F} [f^*(t)] = \mathcal{F} \left[\sum_{n=-\infty}^{\infty} c_n f(t) e^{jn\omega_s t} \right] = \sum_{n=-\infty}^{\infty} c_n F(\omega - n\omega_s) \quad (13.12)$$

Es decir, la transformada de Fourier de una función muestreada es igual a una suma con infinitos sumandos, y cada sumando es una copia de la transformada de Fourier de la función original trasladada en frecuencia (un múltiplo de la frecuencia de muestreo) y ponderada por el coeficiente c_n .

Es interesante hacer notar que este resultado es independiente del tipo o la forma de muestreo que se haya llevado a cabo. En otras palabras, la forma de la función portadora $p(t)$ sólo influye en los coeficientes complejos c_n .

13.3. El problema del *aliasing*

13.3.1. Teorema de Shannon

El fenómeno del *aliasing* es uno de los conceptos más importantes de la asignatura. Una forma de entender qué es el *aliasing* consiste en intentar dibujar la transformada de Fourier de una función muestreada, cuya expresión se ha obtenido en el apartado anterior.

Para mayor simplicidad, se intentará dibujar sólo el módulo de la transformada de Fourier (recuérdese que es un número complejo) y además se supondrá que la función original $f(t)$ tiene ancho de banda limitado, es decir, su transformada de Fourier tomará valores nulos a partir de una determinada frecuencia m que señala el máximo contenido en frecuencia de la señal $f(t)$. En la Fig. 13.5 se muestra cualitativamente cómo podría representarse el módulo de la transformada de Fourier de la señal, aunque podría adoptar cualquier otra forma mientras sea nula fuera del intervalo $-\omega_m < \omega < \omega_m$.

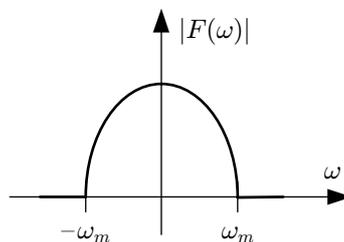


Figura 13.5: Transformada de Fourier de una función con ancho de banda limitado

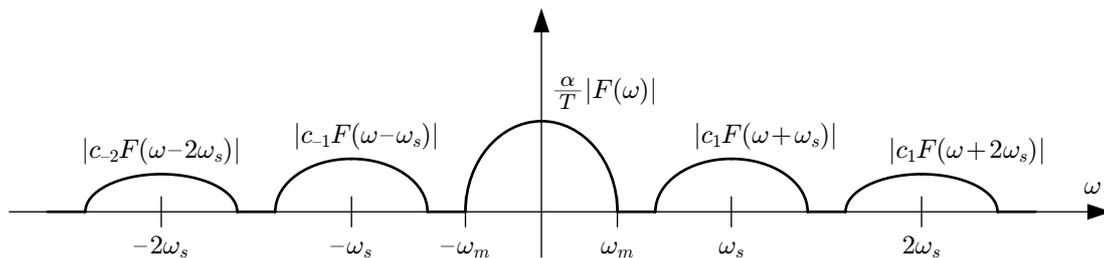


Figura 13.6: Transformada de Fourier de la función muestreada

En la Fig. 13.6 se muestra la forma que adopta la suma de la ecuación (13.12) que origina la transformada de Fourier de la función muestreada. Cada sumando es una réplica escalada de la señal original trasladada en frecuencia. La condición para que no se produzca solapamiento entre los sumandos es:

$$\omega_s > 2\omega_m \tag{13.13}$$

Y esta condición es independiente de valores tomen los coeficientes c_n , por tanto es también independiente del tipo de muestreo que se haya efectuado. Esta condición general es el *teorema de Shannon del muestreo* (1948). Cuando no se cumple el teorema de Shannon, existe solapamiento en los sumandos y aparece el fenómeno del *aliasing*. En la Fig. 13.7 se muestra la forma de la transformada de Fourier de la señal muestreada cuando existe *aliasing*.

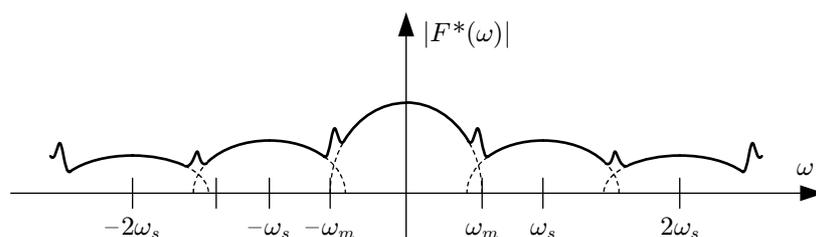


Figura 13.7: Transformada de Fourier de la función muestreada con *aliasing*

13.3.2. *Aliasing* y reconstrucción de la señal original

Comparando la Fig. 13.6 con la Fig. 13.7 es posible observar la ventaja que supone para el ingeniero que no exista *aliasing*. Cuando no hay *aliasing* el ingeniero podría ser capaz de obtener la señal original no muestreada a partir de la muestreada. Bastaría que fuera capaz de aislar el sumando no trasladado en frecuencia entre $-\omega_N$ y ω_N con un filtro pasa-baja y deshacer el escalado del coeficiente c_0 . Así hallaría la transformada de Fourier de la señal continua en el tiempo y podría encontrar su expresión exacta en el dominio del tiempo.

En cambio, si existe *aliasing*, no es posible obtener la transformada de Fourier de la señal original. El solapamiento de los sumandos, hace imposible este paso.

Ahora se puede entender mejor porqué en el apartado anterior se hizo el análisis para el caso de que la función continua tuviera ancho de banda limitado. Si el contenido en frecuencia de la señal continua no está limitado, siempre existirá *aliasing* por muy rápido que se realice el muestreo. En otras palabras, habrá información de alta frecuencia de la señal original que se perderá y no será posible reconstruir dicha señal a partir de la función muestreada. En cambio, si el ancho de banda es limitado, es posible muestrear rápidamente la señal sin perder información.

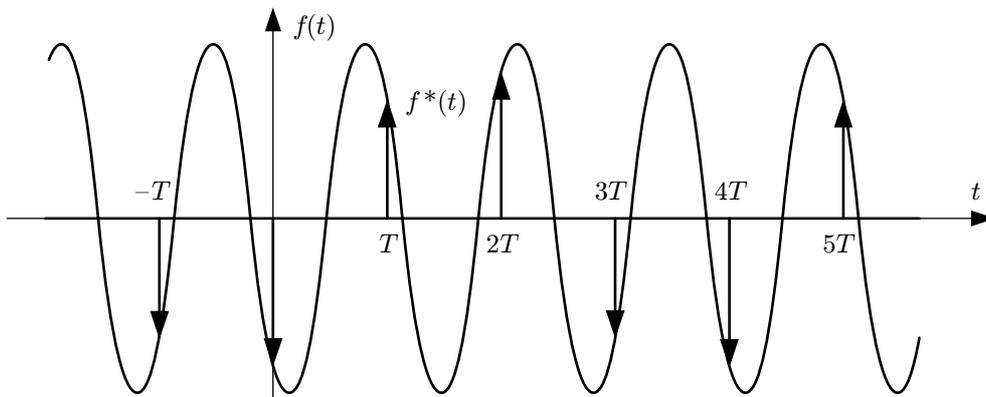


Figura 13.8: Señal continua y muestreada que no cumple el teorema de Shannon

En la Fig. 13.8 se muestra un ejemplo de muestreo de una señal a una frecuencia inferior al umbral que marca el teorema de Shannon. Con sólo la información de la señal muestreada, Fig. 13.9, no es posible llegar a la señal original.

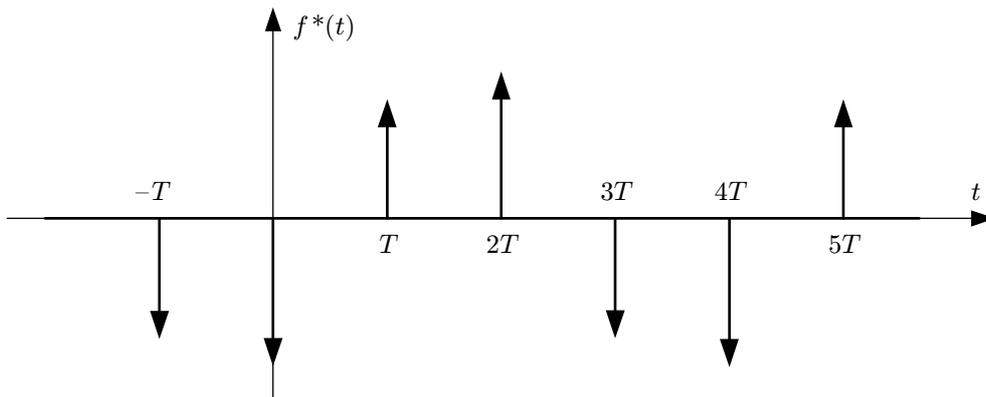


Figura 13.9: Señal muestreada

En más, la señal sinusoidal que se podría sugerir para como la original podría ser errónea, como la que se presenta en la Fig. 13.10.

Por tanto, muestrear por debajo del umbral que marca el teorema de Shannon, no sólo implica que se pierda información del contenido en alta frecuencia; también implica que las señales de alta frecuencia pueden ser observadas con una frecuencia inferior a la que realmente tienen. En eso consiste en la práctica el solapamiento de los sumandos en frecuencia.

Por último cabría preguntarse si tiene sentido suponer que las señales continuas que se van a estudiar tienen ancho de banda limitado o no. Para conocerlo y asegurar que no se va a producir *aliasing*, habría que calcular la transformada de Fourier de la señal, o una herramienta análoga como es el espectro

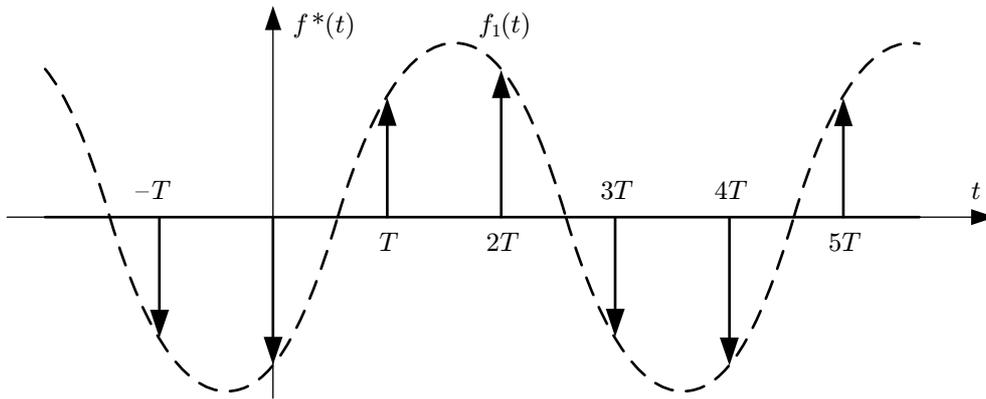


Figura 13.10: Señal muestreada y continua que no corresponde a la original

de frecuencia. En sistemas mecánicos es muy corriente suponer que su respuesta en posición ante una entrada fuerza tiene esta característica. En cualquier caso, llega un momento en que la respuesta de alta frecuencia posee una amplitud tan pequeña que está por debajo de la resolución del sensor de medida y, por lo menos a partir de ese momento su respuesta se puede suponer nula.

13.3.3. *Aliasing* y ruido en la medida de la señal

Un problema a la hora de tratar las señales muestreadas es el ruido que se puede introducir en la lectura del sensor de medida. Ese ruido se puede malinterpretar como presente en la función continua original y podría ser causante de *aliasing* si se intenta llevar a cabo una reconstrucción de la misma.

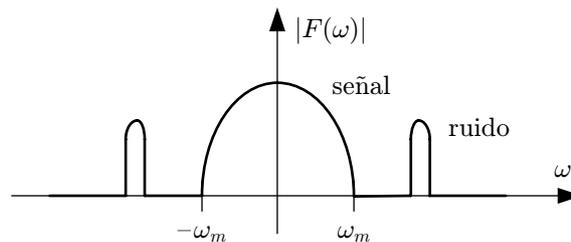


Figura 13.11: Señal de ancho de banda limitado y ruido de alta frecuencia

En la Fig. 13.11 se muestra la transformada de Fourier de una señal continua con ancho de banda limitado y ruido de alta frecuencia. En la Fig. 13.12 se muestra el *aliasing* que se puede producir debido al ruido.

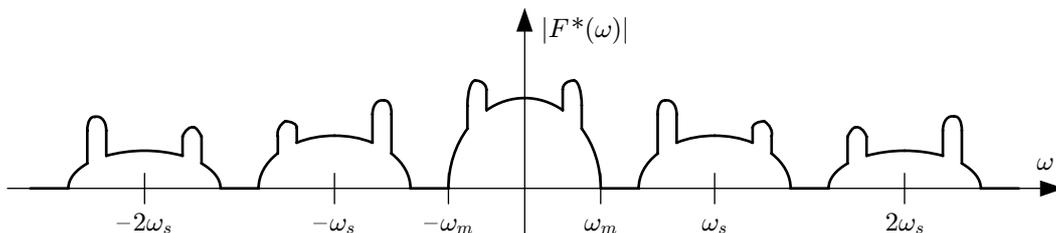


Figura 13.12: *Aliasing* producido por el ruido

Por este motivo, muchos sensores llevan incorporado un filtro de ruido que se suele denominar como filtro anti-*aliasing*. Y, si no los lleva, el ingeniero debe filtrar convenientemente la señal medida para prevenir este problema.

Capítulo 14

El muestreo ideal

En este capítulo se introduce el concepto de muestreo ideal. Aunque muchas de las conclusiones que se obtuvieron en el capítulo anterior son independientes del tipo de muestreo que se aplique, la realidad es que los convertidores A/D no graban la señal original en un intervalo de tiempo a , sino que conocen un valor puntual de la función. Por tanto, habrá que definir una nueva función portadora que tome sólo valores puntuales de la función continua.

14.1. Definición de muestreo ideal

14.1.1. Función portadora

El muestreo ideal se consigue cuando se multiplica la función continua $f(t)$ con la función portadora $p(t)$ que se define como:

$$p(t) = \begin{cases} \infty & \text{si } t = 0 \\ 0 & \text{si } 0 < t < T \end{cases} \quad \text{con} \quad \int_{-\frac{T}{2}}^{\frac{T}{2}} p(t) dt = 1 \quad (14.1)$$

$$p(t) = \sum_{n=-\infty}^{\infty} \delta(t + nT) \quad (14.2)$$

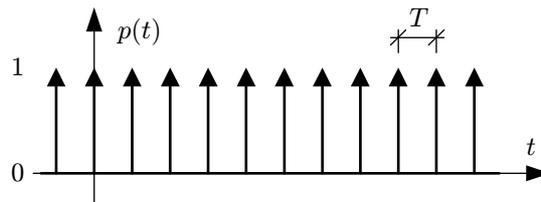


Figura 14.1: Función portadora para el muestreo ideal

Como se observa en la Fig. 14.1, la función portadora es una secuencia o tren de impulsos unidad que se extiende a lo largo del tiempo, desde $-\infty$ hasta ∞ , separados cada T segundos. Como la nueva función portadora $p(t)$ sigue siendo periódica, es posible expresarla también mediante una serie de Fourier con coeficientes:

$$c_n = \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} p(t) e^{-jn\omega_s t} dt = \frac{1}{T} \quad (14.3)$$

Se obtiene que todos los coeficientes son iguales y reales. Por tanto:

$$p(t) = \sum_{n=-\infty}^{\infty} c_n e^{-jn\omega_s t} = \frac{1}{T} \sum_{n=-\infty}^{\infty} e^{-jn\omega_s t} \quad (14.4)$$

Las expresiones (14.2) y (14.4) son completamente equivalentes y se usará una u otra indistintamente dependiendo de cuál ofrezca mayores ventajas.

14.1.2. Función temporal muestreada

Con las dos expresiones equivalentes de la función periódica $p(t)$ definida en el apartado anterior, se pueden hallar otras dos para la función temporal de muestreada:

$$f^*(t) = f(t)p(t) = f(t) \sum_{n=-\infty}^{\infty} \delta(t + nT) = \sum_{n=-\infty}^{\infty} f(nT)\delta(t + nT) \quad (14.5)$$

$$f^*(t) = f(t)p(t) = f(t) \frac{1}{T} \sum_{n=-\infty}^{\infty} e^{-jn\omega_s t} = \frac{1}{T} \sum_{n=-\infty}^{\infty} f(nT)e^{-jn\omega_s t} \quad (14.6)$$

También estas expresiones son equivalentes, aunque queda más patente en la ecuación (14.5) que en la función muestreada sólo se conocen los valores de la función original en cada periodo de muestreo. La función muestreada es una sucesión de impulsos, que se extienden en el tiempo desde $-\infty$ hasta ∞ , y el área encerrada por cada uno de ellos es igual al valor de la función original en ese instante.

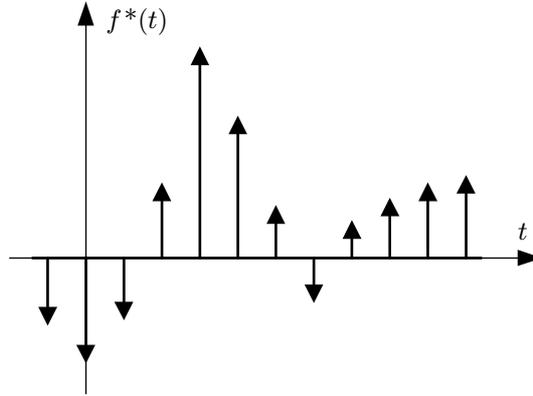


Figura 14.2: Función muestreada de forma ideal

14.2. Transformada de Fourier de la función muestreada

Para calcular la transformada de Fourier de la función muestreada a partir de la ecuación (14.5), hay que recordar previamente que la transformada de Fourier del impulso unidad en el origen de tiempos, es la unidad:

$$\mathcal{F}[\delta(t)] = 1 \quad (14.7)$$

La propiedad de traslación en el tiempo de la transformada de Fourier dice que:

$$\mathcal{F}[f(t + t_1)] = e^{-j\omega t_1} \mathcal{F}[f(t)] \quad (14.8)$$

Conociendo estas propiedades, es posible calcular la transformada de Fourier de la función muestreada, que se denotará como $F^*(\omega) = \mathcal{F}[f^*(t)]$, de la siguiente forma:

$$F^*(\omega) = \mathcal{F} \left[\sum_{n=-\infty}^{\infty} f(nT)\delta(t + nT) \right] = \sum_{n=-\infty}^{\infty} f(nT)\mathcal{F}[\delta(t + nT)] = \sum_{n=-\infty}^{\infty} f(nT)e^{-j\omega nT} \quad (14.9)$$

También se puede obtener una transformada de Fourier completamente equivalente a partir de la ecuación (13.12) de forma más inmediata:

$$F^*(\omega) = \sum_{n=-\infty}^{\infty} c_n F(\omega - n\omega_s) = \frac{1}{T} \sum_{n=-\infty}^{\infty} F(\omega - n\omega_s) \quad (14.10)$$

De nuevo las expresiones (14.9) y (14.10) son completamente equivalentes. Es interesante hacer notar que en el caso del muestreo ideal todos los “alias” son exactamente iguales a la transformada de Fourier original, sólo que se encuentran multiplicados por una misma constante, Fig. 14.3. También en el caso de muestreo ideal se debe cumplir el teorema de Shannon para que se pueda reconstruir la señal original a partir de la muestreada.

A partir de ahora, mientras no se especifique lo contrario, se entenderá siempre que se hable de muestreo que se trata del muestreo ideal.

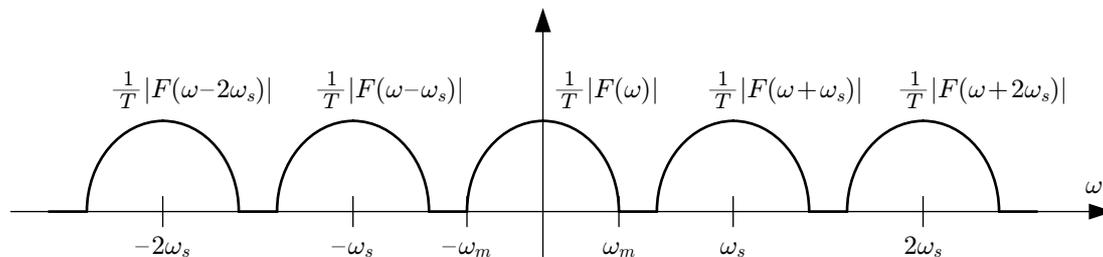


Figura 14.3: Transformada de Fourier de la señal muestreada

14.3. Transformada de Laplace de la función muestreada

Con la transformada de Fourier se pudo relacionar el máximo contenido en frecuencia de una señal con el periodo de muestreo adecuado para dicha señal, de cara a su posterior reconstrucción, por medio del teorema de Shannon. Sin embargo, la herramienta matemática que más se emplea en esta asignatura es la transformada de Laplace. En este apartado se estudia si es posible obtener la transformada de Laplace de una señal muestreada.

Aplicando la definición y propiedades de la transformada de Laplace que se contienen en el Capítulo 2 a una función muestreada:

$$F^*(s) = \mathcal{L} \left[\sum_{n=-\infty}^{\infty} f(nT) \delta(t + nT) \right] = \sum_{n=-\infty}^{\infty} f(nT) \mathcal{L}[\delta(t + nT)] = \sum_{n=0}^{\infty} f(nT) e^{-snT} \quad (14.11)$$

Se obtiene un sumatorio donde no aparecen valores de n negativos. La razón es que los impulsos que se encuentran en la zona de “tiempos negativos” tienen transformada de Laplace nula. Es fácil recordar la expresión (14.11) ya que cada valor muestreado de la función temporal va multiplicado por una exponencial que define precisamente el momento en que ha sido muestreado, o mejor aún, el retraso de tiempo que se debe esperar hasta que se dé dicho valor en la función.

14.3.1. Forma cerrada y región de convergencia

En el siguiente ejemplo, se muestra la transformada de Laplace de un escalón unidad muestreado, Fig. 14.4:

$$U^*(s) = \mathcal{L}[u^*(t)] = \sum_{n=0}^{\infty} e^{-snT} = 1 + e^{-sT} + e^{-s2T} + e^{-s3T} + \dots \quad (14.12)$$

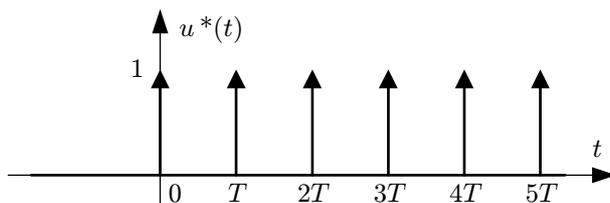


Figura 14.4: Escalón unidad muestreado

Resulta una suma de infinitos términos, bastante más difícil de manejar que su equivalente para la función no muestreada. Sin embargo, si se cumplen ciertas condiciones es posible encontrar una forma más compacta para la expresión de la transformada de Laplace de la función muestreada. En concreto, en el ejemplo anterior:

$$U^*(s) = 1 + e^{-sT} + e^{-s2T} + e^{-s3T} + \dots = \frac{1}{1 - e^{-sT}} \text{ si } |e^{-sT}| < 1 \quad (14.13)$$

El sumatorio de esta serie geométrica converge si la razón de dicha serie es menor que la unidad. La condición de convergencia de la forma cerrada es llamada habitualmente con la abreviatura ROC (del inglés *region of convergence*) y en algunos casos puede llegar a tener importancia. Sin embargo, en esta asignatura las regiones de convergencia no suelen tener especial interés.

14.3.2. Forma alternativa para la transformada de Laplace

Se puede dar una expresión equivalente de la transformada de Laplace usando de la función muestreada como serie de Fourier:

$$F^*(s) = \mathcal{L}[f^*(t)] = \mathcal{F}\left[\frac{1}{T} \sum_{n=-\infty}^{\infty} f(t)e^{jn\omega_s t}\right] = \frac{1}{T} \sum_{n=-\infty}^{\infty} F(s - jn\omega_s) \quad (14.14)$$

Esta expresión es equivalente a la ecuación (14.11). Hay que advertir que ahora el sumatorio se extiende desde $-\infty$ hasta ∞ . Aquí no ocurre que se puedan despreciar los sumandos para valores de n negativos, debido a que esta suma se extiende en el dominio de Laplace. Esto suele generar no pocas confusiones, por lo que se recomienda un estudio riguroso de las dos formas de la transformada de Laplace.

14.3.3. Periodicidad de la transformada de Laplace

Una propiedad muy importante de la transformada de Laplace de la función muestreada es que es una función compleja y periódica con *periodo* $j\omega_s$. Como por lo general los alumnos están acostumbrados a funciones periódicas reales con periodos reales, se va a mostrar a continuación porqué la transformada de Laplace es una función compleja periódica y que su periodo es el número complejo imaginario puro que se ha citado anteriormente:

$$\begin{aligned} f(t) \text{ es periódica de periodo } T \text{ si } f(t + mT) &= f(t) \quad \forall m \\ F^*(s) \text{ es periódica de periodo } j\omega_s \text{ si } F^*(s + mj\omega_s) &= F^*(s) \quad \forall m \end{aligned}$$

$$F^*(s + mj\omega_s) = \sum_{n=0}^{\infty} f(nT)e^{-(s+mj\omega_s)nT} = \sum_{n=0}^{\infty} f(nT)e^{-snT}e^{-jmn2\pi} = F^*(s) \quad (14.15)$$

El término exponencial $e^{-jmn2\pi}$ es un “giro” múltiplo entero de 2π rad del número complejo previo, por lo que no lo modifica. Hay que advertir que la función muestreada es periódica en el dominio de Laplace, es decir, su transformada de Laplace. Evidentemente la propia función muestreada no es periódica en el dominio temporal.

14.3.4. Franjas primaria y complementarias

Una consecuencia muy importante de la propiedad de periodicidad de la transformada de Laplace es que se pueden definir dentro del plano complejo S en el que está definida la variable de Laplace una serie de franjas o bandas como las que se muestran en la Fig. 14.5.

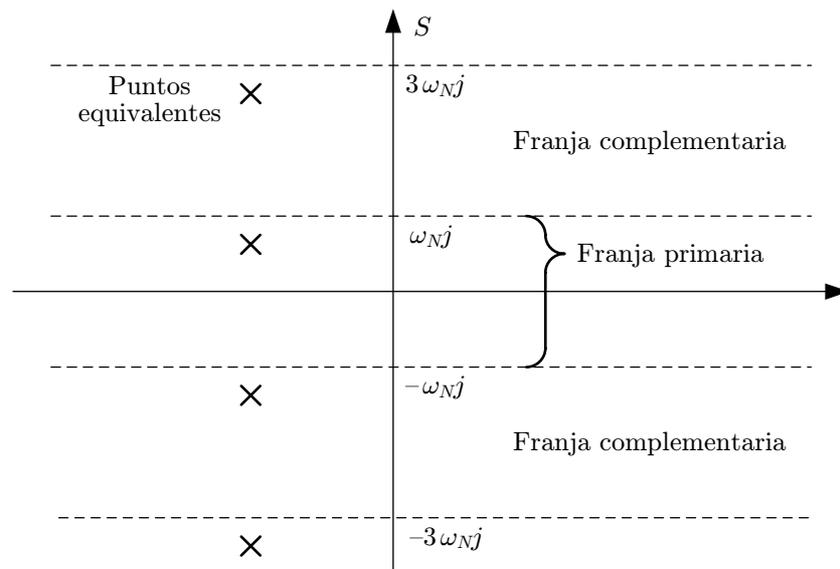


Figura 14.5: Franjas primaria y secundarias

Todo lo que ocurra en la franja primaria se repite en las franjas complementarias de forma periódica. Como se verá más adelante, en los sistemas muestreados si s_1 es un polo del sistema, los infinitos valores $s_1 + jm\omega_s$ que se pueden definir con m cualquier valor entero, también serán polos del sistema.

Capítulo 15

Reconstrucción de la función continua original

Al hablar del fenómeno del *aliasing* ya se mencionó que a partir de una señal muestreada se podría intentar reconstruir la función continua original utilizando un filtro. En este capítulo se explica cómo se puede hacer esto de forma exacta y qué otros filtros —menos exactos— se usan realmente en los sistemas muestreados. A los filtros que reconstruyen funciones temporales a partir de los valores muestreados se les suele llamar también retenedores (en inglés *holders*).

15.1. Filtro ideal

15.1.1. Características del filtro ideal

Suponiendo que se cumple el teorema de Shannon, el espectro de la señal original se repite en el dominio de la frecuencia sin solape ninguno. Por tanto, lo que se necesita es un filtro con la siguiente forma:

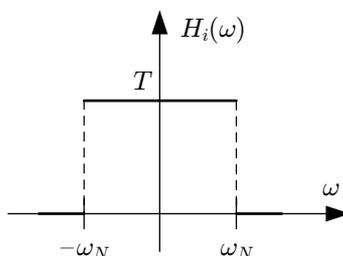


Figura 15.1: Filtro ideal

$$H_i(\omega) = \begin{cases} T & \text{si } |\omega| < \frac{\omega_s}{2} \\ 0 & \text{si } |\omega| \geq \frac{\omega_s}{2} \end{cases} \quad (15.1)$$

A la mitad de la frecuencia de muestreo se le conoce como *frecuencia de Nyquist*:

$$\omega_N = \frac{\pi}{T} = \frac{\omega_s}{2}. \quad (15.2)$$

Si se hubiera empleado una escala logarítmica en la Fig. 15.1 tanto en ganancias como en frecuencias, como es habitual en los diagramas de Bode, quedaría patente que el filtro ideal posee una pendiente de $-\infty$ dB/década a partir de la frecuencia de Nyquist. Este hecho previene ya al ingeniero de que no es posible llevar a la práctica un filtro de tales características.

15.1.2. Imposibilidad física de construcción del filtro ideal

En este apartado se demuestra lo que de alguna manera ya se ha anticipado en el apartado anterior, es decir, por qué es imposible implementar físicamente un filtro ideal.

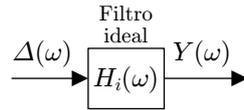


Figura 15.2: Respuesta en frecuencia del filtro ideal a un impulso

Como lo que se conoce es la forma del filtro ideal en el dominio de la frecuencia, o de Fourier, habrá que calcular la transformada inversa de Fourier de la salida del filtro. La transformada inversa de Fourier se define como:

$$F(\omega) \xrightarrow{\mathcal{F}^{-1}} f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) e^{j\omega t} d\omega \quad (15.3)$$

Como la transformada de Fourier del impulso unidad es la propia unidad, $\mathcal{F}[\delta(t)] = \Delta(\omega) = 1$, la transformada de Fourier da la salida coincidirá con la expresión del filtro ideal:

$$Y(\omega) = H_i(\omega)\Delta(\omega) = H_i(\omega) \quad (15.4)$$

Por tanto:

$$y(t) = \frac{1}{2\pi} \int_{-\omega_N}^{\omega_N} T e^{j\omega t} d\omega = \frac{T}{2\pi} \left[\frac{e^{j\omega t}}{jt} \right]_{-\omega_N}^{\omega_N} = \frac{2(e^{j\omega_N t} - e^{-j\omega_N t})}{j\omega_N t} = \frac{\sin \omega_N t}{\omega_N t} = \text{sinc } \omega_N t \quad (15.5)$$

La respuesta del filtro ideal ante un impulso es una función *sinus cardinalis* de frecuencia igual a la frecuencia de Nyquist (Fig. 15.3). Sin embargo, ¿cómo es posible que el filtro responda en tiempos negativos, es decir, *antes de que se produzca la entrada*? Esto viola el principio de causalidad. No se puede observar un efecto antes de que actúe la causa. Así pues, aunque matemáticamente sea posible dar la definición del filtro ideal, físicamente no se puede implementar.

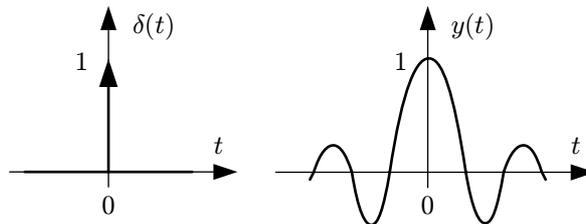


Figura 15.3: Respuesta temporal del filtro ideal a un impulso

15.1.3. Reconstrucción de la señal con el filtro ideal

Otra forma de demostrar que el filtro ideal es irrealizable consiste en observar directamente qué expresión temporal continua resulta al introducir en dicho filtro una función muestreada:

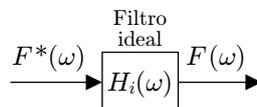


Figura 15.4: Respuesta del filtro ideal a una función muestreada

Por definición la respuesta debe ser la función continua $f(t)$ original, pero se pretende ver cómo el filtro ideal consigue reconstruirla en el dominio del tiempo. La operación producto en el dominio de la frecuencia se puede convertir en el dominio temporal como la operación convolución:

$$F(\omega) = H_i(\omega)F^*(\omega) \quad (15.6)$$

$$f(t) = h_i(t) \otimes f^*(t) \quad (15.7)$$

Para calcular la función temporal $h_i(t)$ se debe hacer la siguiente transformada inversa de Fourier:

$$h_i(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} H_i(\omega) e^{j\omega t} d\omega = \frac{1}{2\pi} \int_{-\omega_N}^{\omega_N} T e^{j\omega t} d\omega = \text{sinc } \omega_N t \quad (15.8)$$

Resulta ser la misma expresión que la anterior ecuación (15.5). Es por eso que la función temporal asociada a una función de transferencia se le suele llamar “función respuesta impulso unidad” o expresiones equivalentes. Volviendo al cálculo de la función temporal original, se puede demostrar que la convolución de la función muestreada con la función *sinus cardinalis* es:

$$f(t) = \sum_{n=-\infty}^{\infty} f(nT) \operatorname{sinc}(\omega_n t + nT) \quad (15.9)$$

Es decir, que para poder calcular el valor de la función $f(t)$ en un instante cualquiera es necesario conocer *todos* los valores muestreados de dicha función. Por tanto, no es posible generar *on-line* dicha función porque no se conocen los valores muestreados futuros (suponiendo que se sea posible memorizar los infinitos valores pasados).

En la fórmula (15.9) se puede ver por qué se llama también a la función *sinus cardinalis* “función de interpolación”. Es decir, en el intervalo de tiempo comprendido entre dos instantes de muestreo, se utiliza esta función para interpolar. Pero se ve también que no es exactamente una interpolación, ya que es un procedimiento que no sólo afecta al periodo de tiempo comprendido entre dos muestreos. Las funciones de interpolación afectan a todos los instantes de tiempo.

15.2. Retenedor de orden cero

15.2.1. Características del retenedor de orden cero

En lugar del filtro ideal, lo que se utiliza en la práctica es algún tipo de retenedor de los que se verán a continuación. Los retenedores más extendidos actúan a intervalos de forma similar a cómo aproxima una función la suma en serie de Taylor.

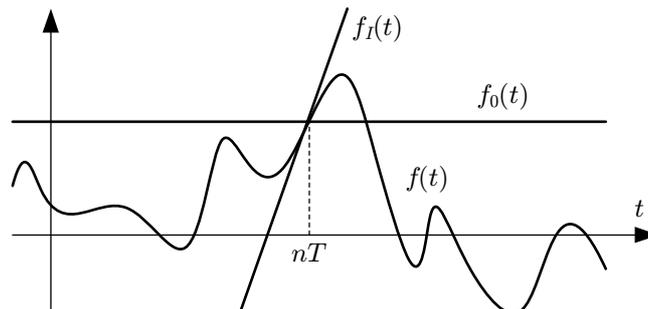


Figura 15.5: Aproximaciones de Taylor en torno a un instante de muestreo

En la Fig. 15.5 se ve cómo se aproximaría por Taylor la función $f(t)$ en los puntos próximos a un instante cualquiera de muestreo, $t = nT$. En dicha gráfica, $f_0(t)$ sería la aproximación de orden cero y $f_1(t)$ la aproximación de primer orden. Formulando esta idea en una ecuación:

$$f(t) \approx f(nt) + \left[\frac{df(t)}{dt} \right]_{t=nT} (t - nT) + \left[\frac{d^2 f(t)}{dt^2} \right]_{t=nT} (t - nT)^2 + \dots \quad (15.10)$$

El *retenedor de orden cero* (en inglés *zero-order holder*) es el “constructor” de la función original que usa sólo el término de orden cero de la serie de Taylor, es decir, un valor constante, entre un periodo de muestreo y el siguiente:

$$f_0(t) = \begin{cases} 0 & \text{si } t < 0 \\ f(0) & \text{si } 0 \leq t < T \\ f(T) & \text{si } T \leq t < 2T \\ f(2T) & \text{si } 2T \leq t < 3T \\ \dots & \dots \\ f(nT) & \text{si } nT \leq t < nT + T \end{cases} \quad (15.11)$$

La reconstrucción que hace el retenedor de orden cero es una función no continua, o continua a tramos, con saltos finitos cada periodo de muestreo (Fig. 15.6).

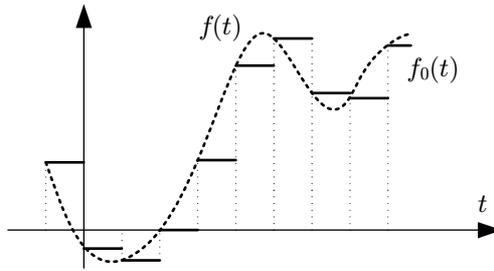


Figura 15.6: Reconstrucción de la función original con el retenedor de orden cero

15.2.2. Expresión de Laplace del retenedor de orden cero

Para encontrar una expresión en el dominio de Laplace del retenedor de orden cero, se acude al concepto de respuesta al impulso unidad al que se hizo mención tras la ecuación (15.8). Con las características que han descrito en el apartado anterior, la respuesta del retenedor de orden cero a un impulso unidad en el origen de tiempos debería ser la que se muestra en la Fig. 15.7.

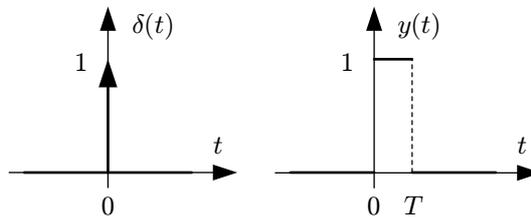


Figura 15.7: Respuesta del retenedor de orden cero a un impulso unidad

Matemáticamente, en el dominio temporal y pasando al dominio de Laplace, se puede escribir:

$$y(t) = u(t) - u(t - T) \quad (15.12)$$

$$Y(s) = \frac{1}{s} - \frac{1}{s} e^{-sT} \quad (15.13)$$

Dado que la transformada de Laplace de la entrada es la unidad, se puede concluir que la expresión en Laplace del retenedor de orden cero es:

$$H_0(s) = ZOH(s) = \frac{Y(s)}{\Delta(s)} = \frac{1 - e^{-sT}}{s} \quad (15.14)$$

Hay que destacar que la respuesta del retenedor de orden cero no viola el principio de causalidad, ya que el efecto (la salida) es posterior a la causa (la entrada).

15.2.3. Respuesta en frecuencia del retenedor de orden cero

En cuanto a la respuesta en frecuencia, se puede calcular la transformada de Fourier del retenedor de orden cero sustituyendo la variable s de Laplace por $j\omega$.

$$H_0(\omega) = ZOH(\omega) = \frac{1 - e^{-j\omega T}}{j\omega} = (\dots) = T e^{-j\frac{\omega T}{2}} \operatorname{sinc} \frac{\omega T}{2} \quad (15.15)$$

Con la expresión (15.15) se puede representar fácilmente la respuesta en frecuencia, separando el módulo del argumento. En la Fig. 15.8 la respuesta en frecuencia en escala lineal.

Como se podría haber deducido observando la Fig. 15.6 el retenedor de orden cero introduce cierto retraso en la señal reconstruida. También se observa en la Fig. 15.8 que el retenedor de orden cero posee, para frecuencias inferiores a la de muestreo la misma caída lineal en fases que un retraso puro de valor mitad que el periodo de muestreo.

$$ZOH(s) \approx T e^{-\frac{T}{2}s} \quad (15.16)$$

Por tanto, la expresión (15.16) es una buena aproximación del retenedor de orden cero para frecuencias inferiores a la de Nyquist. Esta aproximación se suele usar para analizar en el dominio continuo aquellos sistemas discretos cuya dinámica sea mucho menor que la frecuencia de muestreo.

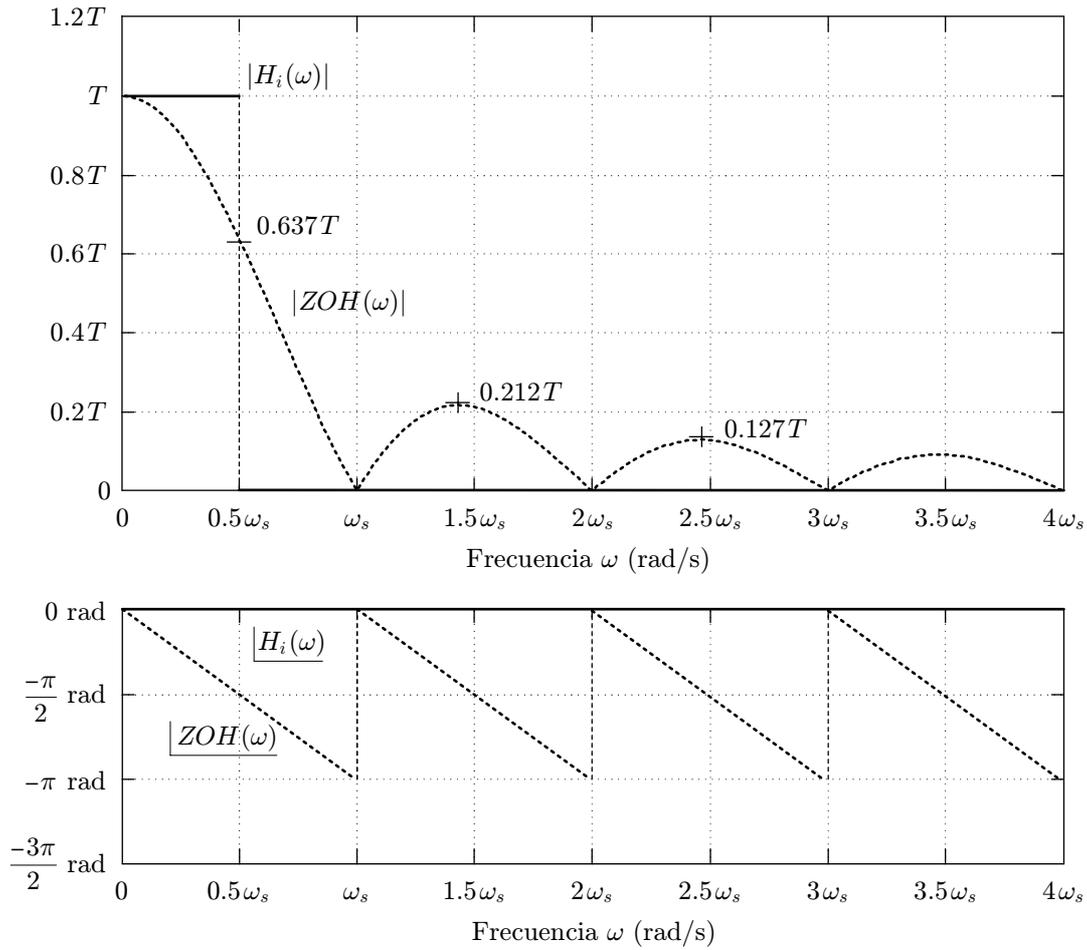


Figura 15.8: Respuesta en frecuencia del retenedor de orden cero

15.3. Retenedor de primer orden

15.3.1. Características del retenedor de primer orden

El retenedor de primer orden es el “constructor” de la función original que entre dos instantes de muestreo usando una serie de Taylor de primer orden:

$$f_1(t) = f(nt) + \left[\frac{df(t)}{dt} \right]_{t=nT} (t - nT) \text{ para } nT \leq t < nT + T \quad (15.17)$$

La mayor dificultad de este retenedor es cómo estima la derivada de la función original a partir de los valores muestreados de la función original. Aunque se puede hacer de distintas maneras, la mejor aproximación de la derivada es calcular la diferencia de la función entre el muestreo actual y el anterior, dividido por el periodo de muestreo.

$$f_1(t) = f(nt) + \frac{f(nt) - f(nt - T)}{T} (t - nT) \text{ para } nT \leq t < nT + T \quad (15.18)$$

Con esto se consigue, por ejemplo, la reconstrucción que se muestra en la Fig. 15.9. Como se puede observar, este modo de reconstrucción es peor que el que se consigue con el retenedor de orden cero. Esto se debe a que la aproximación de la derivada es buena para el intervalo de muestreo precedente, o —como mucho— para el valor de muestreo actual, pero no resulta muy buena para el periodo de muestreo siguiente. Por otro lado, la reconstrucción que hace el retenedor de primer orden también presenta discontinuidades cada muestreo.

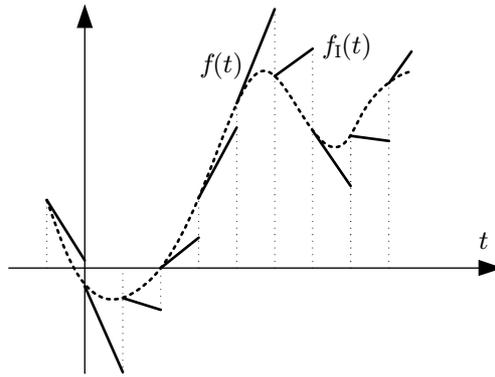


Figura 15.9: Reconstrucción de la función original con el retenedor de primer orden

15.3.2. Expresión de Laplace del retenedor de primer orden

La respuesta del retenedor de primer orden a un impulso unidad en el origen de tiempos es:

$$f_1(t) = \begin{cases} 0 & \text{si } t < 0 \\ 1 + \frac{t}{T} & \text{si } 0 \leq t < T \\ 1 - \frac{t}{T} & \text{si } T \leq t < 2T \\ 0 & \text{si } t \geq 2T \end{cases} \quad (15.19)$$

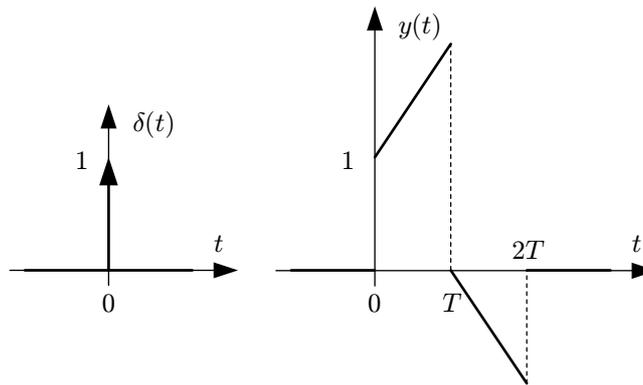


Figura 15.10: Respuesta del retenedor de primer orden a un impulso unidad

Empleando las funciones escalón unidad $u(t)$ y rampa de pendiente unidad $r(t)$ (funciones que son mulas para “tiempos negativos”) se puede escribir como:

$$y(t) = u(t) + \frac{r(t)}{T} - 2u(t-T) - 2\frac{r(t-T)}{T} + u(t-T) + \frac{r(t-2T)}{T} \quad (15.20)$$

Y su transformada de Laplace es:

$$Y(s) = \frac{1}{s} + \frac{1}{Ts^2} - \frac{2}{s}e^{-sT} - \frac{2}{Ts^2}e^{-sT} + \frac{1}{s}e^{-2sT} + \frac{1}{Ts^2}e^{-2sT} \quad (15.21)$$

Por tanto, la función de transferencia del retenedor de primer orden es:

$$H_1(s) = FOH(s) = \frac{Y(s)}{\Delta(s)} = \left(\frac{1 - e^{-sT}}{s}\right)^2 \frac{1 + Ts}{T} = ZOH^2(s) \frac{1 + Ts}{T} \quad (15.22)$$

15.3.3. Respuesta en frecuencia del retenedor de primer orden

En cuanto a la respuesta en frecuencia, en la Fig. 15.11 se comparan los módulos de la respuesta en frecuencia de los tres retenedores vistos hasta el momento.

El retenedor de orden uno no es claramente superior al retenedor de orden cero. Es peor pasa-baja, introduce mayor retraso a alta frecuencia y es mucho más difícil de implementar que el de orden cero. Por estos motivos, el que más se usa es el retenedor de orden cero.

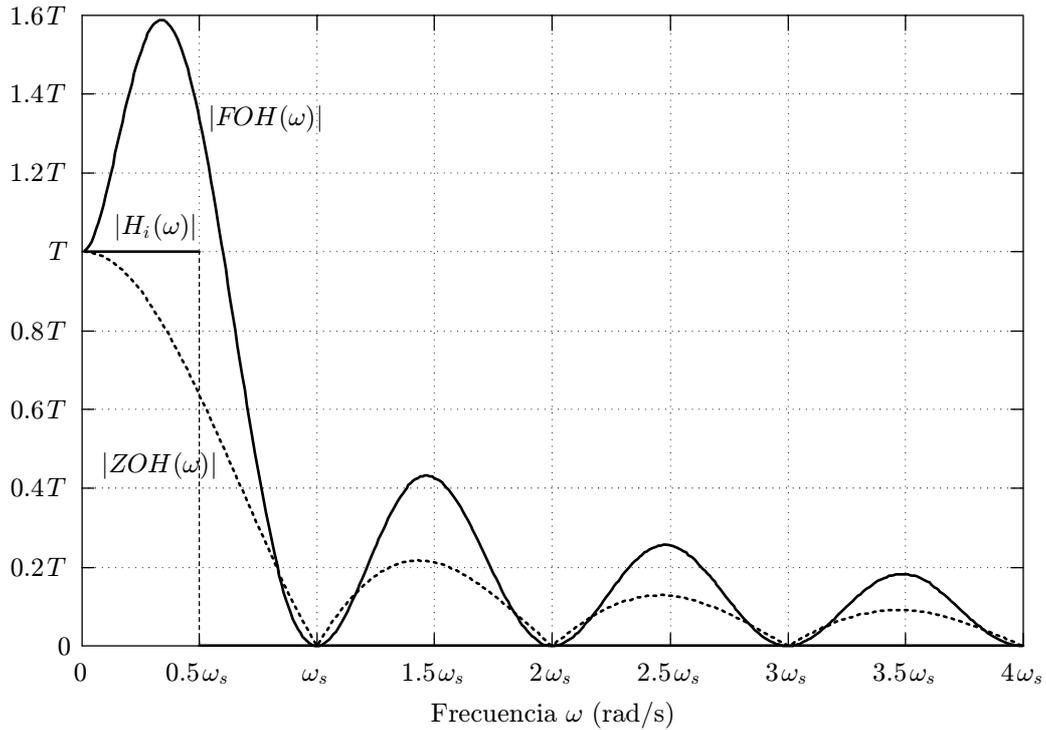


Figura 15.11: Módulo de la respuesta en frecuencia del retenedor de primer orden

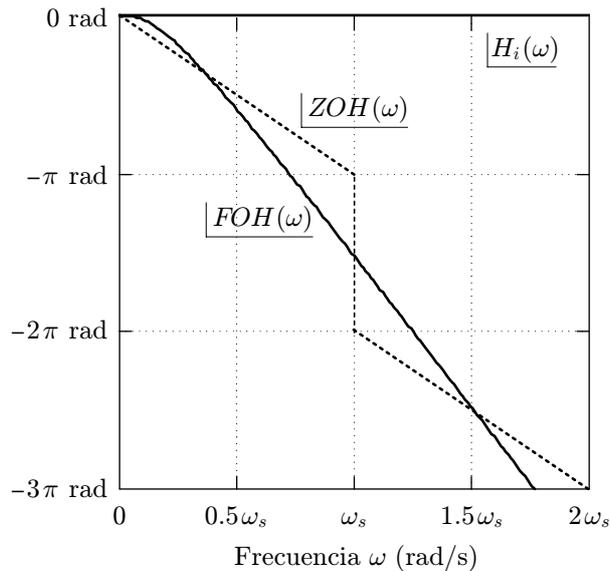


Figura 15.12: Argumento de la respuesta en frecuencia del retenedor de primer orden

15.4. Retenedor polinomial

15.4.1. Características del retenedor polinomial

El retenedor polinomial es un retenedor de primer orden retrasado un periodo de muestreo, con lo que se consigue una reconstruir una función continua (aunque con puntos angulosos) que tiene la forma del polinomio que une los valores muestreados de la función. Pero como en el instante actual no se puede conocer el valor inmediatamente posterior, es por eso que sólo se puede conseguir retrasando la respuesta un periodo de muestreo (Fig. 15.13).

15.4.2. Expresión de Laplace del retenedor polinomial

La respuesta del retenedor polinomial a un impulso unidad se muestra en la Fig. 15.14.

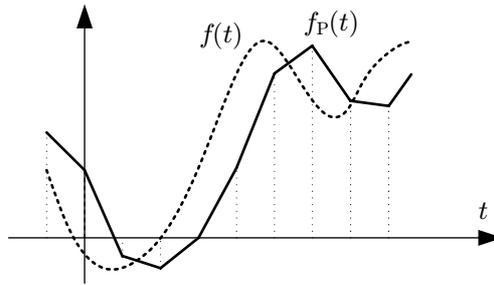


Figura 15.13: Reconstrucción de la función original con el retenedor polinomial

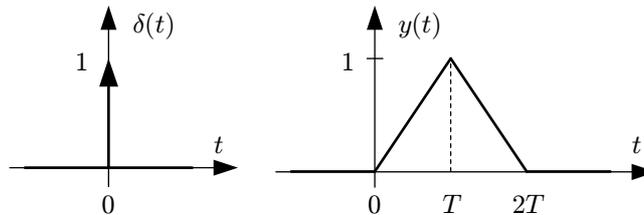


Figura 15.14: Respuesta del retenedor polinomial a un impulso unidad

Matemáticamente:

$$y(t) = \frac{r(t)}{T} - 2\frac{r(t-T)}{T} + \frac{r(t-2T)}{T} \quad (15.23)$$

Y su transformada de Laplace es:

$$Y(s) = \frac{1}{Ts^2} - \frac{2}{Ts^2}e^{-sT} + \frac{1}{Ts^2}e^{-2sT} \quad (15.24)$$

Por tanto, la función de transferencia del retenedor polinomial es:

$$H_P(s) = \frac{Y(s)}{\Delta(s)} = \frac{1}{T} \left(\frac{1 - e^{-sT}}{s} \right)^2 = \frac{1}{T} ZOH^2(s) \quad (15.25)$$

Con este retenedor se consigue un mejor filtro pasa-baja que con el retenedor de orden cero (el doble mejor), pero también un mayor retraso (el doble de retraso). Otro inconveniente del retenedor polinomial es que, si se vuelve a muestrear el resultado de la reconstrucción, no se obtiene la misma señal muestreada de la que se ha partido, cosa que sí ocurriría con los retenedores anteriores.

En general, en cada ejercicio se tendría que especificar qué retenedor se está empleando. Si no se dice nada se supondrá que se usa el retenedor de orden cero, porque es la respuesta que ofrecen normalmente los convertidores digital-analógicos.

También hay que advertir que los retenedores que se han explicado en este capítulo sólo tienen sentido con el muestreo ideal (con impulsos unidad). No se pueden mezclar estos retenedores con otras formas de muestreo.

Capítulo 16

La transformada Zeta

La transformada \mathcal{Z} es la parte esencial de la teoría de control discreto. Algunos manuales de control discreto comienzan precisamente con la definición de esta transformada. Sin embargo, se ha considerado muy conveniente esperar hasta este momento para definirla, porque sólo si se ha entendido el sentido de los desarrollos matemáticos anteriores se puede alcanzar una idea acertada de por qué se usa la transformada \mathcal{Z} en esta asignatura.

16.1. Cálculo de la transformada Zeta

La transformada \mathcal{Z} de una función temporal es igual a la transformada de Laplace de esa función muestreada idealmente, pero con un último cambio de variable. Este cambio de variable permitirá precisamente obtener transformadas \mathcal{Z} en forma de fracciones de polinomios en la variable compleja, como ocurría en el caso de las transformadas de Laplace.

$$f(t) \xrightarrow{\mathcal{Z}} F(z) \quad (16.1)$$

La función $f(t)$ es una función real de variable t real, y $F(z)$ es una función compleja de variable z compleja. La notación que se suele emplear es $F(z) = \mathcal{Z}[f(t)]$ y el modo de obtener esta función $F(z)$ compleja es:

$$f(t) \xrightarrow{\text{muestreo ideal}} f^*(t) \xrightarrow{\mathcal{L}} F^*(s) \xrightarrow{z=e^{Ts}} F(z) \quad (16.2)$$

Por tanto, dado que la transformada de Laplace de una función muestreada es

$$F^*(s) = \sum_{n=0}^{\infty} f(nT)e^{-snT}, \quad (16.3)$$

haciendo el siguiente cambio de variable compleja:

$$z = e^{Ts}, \quad (16.4)$$

se obtiene:

$$F(z) = \sum_{n=0}^{\infty} f(nT)z^{-n}. \quad (16.5)$$

Por las características de la transformada de Laplace, existe una relación biunívoca entre $f^*(t)$ y $F(z)$. Pero no así entre $f(t)$ y $F(z)$, puesto que ya se sabe que dos funciones temporales podrían tener la misma función muestreada.

La expresión (16.5) es la forma más habitual de definir la transformada \mathcal{Z} de cualquier función $f(t)$. Sabiendo que la variable z representa una traslación en el tiempo, como se aprecia en su definición (16.4), se puede emplear la misma regla nemotécnica que se mencionó para la transformada de Laplace: cada valor muestreado de la función está multiplicado por la variable z elevada al exponente que dice en qué periodo de muestreo respecto del origen de tiempos se da dicho valor.

Evidentemente, se puede definir el cambio de variable de z a s como:

$$s = \frac{1}{T} \ln z \quad (16.6)$$

16.2. Tabla de la transformada Zeta de funciones elementales

La transformada \mathcal{Z} se puede calcular siguiendo al proceso descrito en el apartado anterior. Por ejemplo, para el caso del escalón unidad $f(t) = u(t)$, se había encontrado la siguiente forma cerrada:

$$U^*(s) = \frac{1}{1 - e^{-sT}} \text{ si } |e^{-sT}| < 1 \quad (16.7)$$

Sustituyendo $z = e^{Ts}$, la transformada \mathcal{Z} del escalón unidad sería entonces:

$$U(z) = \frac{1}{1 - z^{-1}} = \frac{z}{z - 1} \text{ si } |z^{-1}| < 1 \quad (16.8)$$

Se observa que el ingeniero puede elegir usar fracciones de polinomios en z con exponentes negativos o positivos.

Tabla 16.1: Transformadas Zeta de diversas funciones

$f(t)$	$F(s)$	$F(z)$
$\delta(t)$	1	1
$u(t)$	$\frac{1}{s}$	$\frac{z}{z-1}$
t	$\frac{1}{s^2}$	$\frac{Tz}{(z-1)^2}$
$\frac{t^2}{2}$	$\frac{1}{s^3}$	$\frac{T^2 z(z+1)}{2(z-1)^3}$
e^{-at}	$\frac{1}{s+a}$	$\frac{z}{z-e^{-aT}}$
te^{-at}	$\frac{1}{(s+a)^2}$	$\frac{Tze^{-aT}}{(z-e^{-aT})^2}$
$\sin at$	$\frac{a}{s^2+a^2}$	$\frac{z \sin aT}{z^2 - 2z \cos aT + 1}$
$\cos at$	$\frac{s}{s^2+a^2}$	$\frac{z(z - \cos aT)}{z^2 - 2z \cos aT + 1}$
$e^{-at} - e^{-bt}$	$\frac{b-a}{(s+a)(s+b)}$	$\frac{(e^{-aT} - e^{-bT})z}{(z-e^{-aT})(z-e^{-bT})}$
$1 - e^{-at}$	$\frac{a}{s(s+a)}$	$\frac{z(1-e^{-aT})}{(z-1)(z-e^{-aT})}$
$t - \frac{1-e^{-at}}{a}$	$\frac{a}{s^2(s+a)}$	$\frac{z[(aT-1+e^{-aT})z + (1-e^{-aT}-aTe^{-aT})]}{a(z-1)^2(z-e^{-aT})}$
$e^{-at} \sin bt$	$\frac{b}{(s+a)^2+b^2}$	$\frac{ze^{-aT} \sin bT}{z^2 - 2ze^{-aT} \cos bT + e^{-2aT}}$
$e^{-at} \cos bt$	$\frac{s+a}{(s+a)^2+b^2}$	$\frac{z^2 - ze^{-aT} \cos bT}{z^2 - 2ze^{-aT} \cos bT + e^{-2aT}}$
$1 - (1+at)e^{-at}$	$\frac{a^2}{s(s+a)^2}$	$\frac{z}{z-1} - \frac{z}{z-e^{-aT}} - \frac{aTze^{-aT}}{(z-e^{-aT})^2}$

La Tabla 16.1 es un resumen de las transformadas \mathcal{Z} de funciones habituales, donde se supone que las funciones temporales $f(t)$ son nulas para “tiempos negativos”. Las regiones de convergencia de las transformadas \mathcal{Z} no se muestran por mayor simplicidad. Las expresiones de las transformadas \mathcal{Z} se asemejan a las transformadas de Laplace, en el sentido que son fracciones de polinomios en la variable compleja cuyas raíces son los polos y los ceros de tales expresiones.

16.3. Teoremas de la transformada Zeta

Si $F(z)$ es la transformada \mathcal{Z} de $f(t)$ y $G(z)$ la correspondiente a $g(t)$, donde $f(t)$ y $g(t)$ son funciones causales, se pueden demostrar los teoremas que se muestran en la Tabla 16.2.

16.4. Cálculo de la transformada inversa de Zeta

Lo primero que hay que hacer notar es que, dado un periodo de muestreo, la transformada inversa de \mathcal{Z} no es única. Es decir, dada una expresión $F(z)$ existen infinitas funciones continuas causales $f(t)$ que poseen esa transformada \mathcal{Z} . Todas las funciones continuas que posean valores idénticos en los instantes de muestreo tendrán, evidentemente, la misma transformada \mathcal{Z} .

Tabla 16.2: Propiedades de la transformada Zeta

Propiedad	Expresión
Linealidad	$\mathcal{Z}[\alpha f(t) + \beta g(t)] = \alpha F(z) + \beta G(z)$
Valor final	$\lim_{n \rightarrow \infty} f(nT) = \lim_{z \rightarrow 1} (1 - z^{-1})F(z)$ si el sistema es estable
Valor inicial	$f(0) = \lim_{z \rightarrow \infty} F(z)$ si existe el límite
Retraso en el tiempo	$\mathcal{Z}[f(t - nT)] = z^{-n}F(z)$
Adelanto en el tiempo	$\mathcal{Z}[f(t + nT)] = z^n[F(z) - \sum_{k=0}^{n-1} f(kT)z^{-k}]$
Convolución	$F(z)G(z) = \mathcal{Z}[f(nT) \otimes g(nT)] = \mathcal{Z}[\sum_{k=0}^{\infty} f(kT)g(nT - kT)]$

En la Fig. 16.1 se muestran dos funciones continuas que poseen igual transformada \mathcal{Z} porque su función muestreada es idéntica. Por tanto, a la hora de calcular la transformada inversa de \mathcal{Z} , sólo podrán dar de forma exacta los valores temporales de la función continua en todos los instantes de muestreo. En forma matemática, se podrán dar los valores de $f(nT)$, no la entera expresión de $f(t)$.

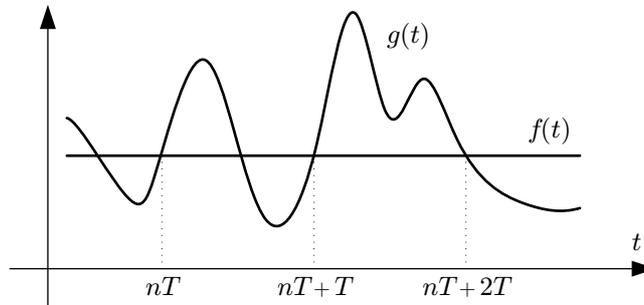


Figura 16.1: Funciones con igual transformada \mathcal{Z}

16.4.1. Método directo o de la expansión de potencia

La forma directa de obtener los valores de $f(nT)$ consiste simplemente en escribir la función $F(z)$ como fracción de polinomios en z con exponentes negativos y efectuar la división de los polinomios. El resultado de la división es directamente la expresión de la transformada \mathcal{Z} de la forma:

$$F(z) = \sum_{n=0}^{\infty} f(nT)z^{-n}. \quad (16.9)$$

Con esta expresión se pueden identificar directamente los valores de la expresión temporal en los instantes de muestreo. En el siguiente ejemplo, se aplica el método descrito:

$$F(z) = \frac{10z + 5}{(z - 1)(z - 0.2)} \quad (16.10)$$

En primer lugar se reescribe la transformada \mathcal{Z} como:

$$F(z) = \frac{10z^{-1} + 5z^{-2}}{1 - 1.2z^{-1} + 0.2z^{-2}}, \quad (16.11)$$

y se efectúa la división:

$$\begin{array}{r}
 \begin{array}{r}
 10z^{-1} \quad +5z^{-2} \\
 -10z^{-1} \quad +12z^{-2} \quad -2z^{-3} \\
 \hline
 0 \quad +17z^{-2} \quad -2z^{-3} \\
 \quad -17z^{-2} \quad +20.4z^{-3} \quad -3.4z^{-4} \\
 \hline
 \quad \quad 0 \quad +18.4z^{-3} \quad -3.4z^{-4} \\
 \quad \quad \quad \dots \quad \dots
 \end{array}
 &
 \left|
 \begin{array}{l}
 1 - 1.2z^{-1} + 0.2z^{-2} \\
 10z^{-1} + 17z^{-2} + 18.4z^{-3} + \dots
 \end{array}
 \right.
 \end{array}
 \quad (16.12)$$

Por tanto, los valores de la función temporal en los instantes de muestreo son:

$$\begin{aligned} f(0) &= 0 \\ f(T) &= 10 \\ f(2T) &= 17 \\ f(3T) &= 18.4 \\ f(4T) &= \dots \end{aligned} \tag{16.13}$$

Con estos valores es posible ir representando a mano alzada la función temporal.

16.4.2. Método de la expansión en fracciones

Este método trata de dividir la función $F(z)$ en una combinación de fracciones elementales de las que se conozca su función temporal correspondiente, es decir, que se puedan encontrar en la Tabla 16.1. Si se puede, es preferible buscar las fracciones simples de $F(z)$ dividido por z , porque suele ser más sencillo. Por ejemplo, aunque la función:

$$F(z) = \frac{(1 - e^{-aT})z}{(z - 1)(z - e^{-aT})}, \tag{16.14}$$

ya posea una solución directamente en la Tabla 16.1, se va a calcular con otras expresiones más elementales. En primer lugar se divide en fracciones simple de la forma:

$$\frac{F(z)}{z} = \frac{1 - e^{-aT}}{(z - 1)(z - e^{-aT})} = \frac{1}{z - 1} + \frac{1}{z - e^{-aT}}, \tag{16.15}$$

y por tanto:

$$F(z) = \frac{z}{z - 1} + \frac{z}{z - e^{-aT}}. \tag{16.16}$$

Se puede concluir que:

$$f(nT) = 1 - e^{-anT}, \tag{16.17}$$

donde se ha escrito nT en lugar de la variable temporal t porque, como ya se ha dicho, cuando se calcula la transformada inversa de \mathcal{Z} sólo se pueden dar con certeza los valores de la función en los instantes de muestreo. Con esta fórmula se podrían ir calculando los sucesivos valores que toma la función, de forma análoga a como se escribieron en (16.13) y representar a mano alzada la forma de la función temporal.

Recordar que cuando en la función \mathcal{Z} aparecen términos $(z - a)^n$, hay que obtener tantas fracciones como indica la potencia, con denominadores: $(z - a)$, $(z - a)^2$, ..., $(z - a)^n$.

16.5. Función de transferencia Zeta

En el dominio continuo se definía funciones de transferencia como una expresión matemática función de la variable s que era precisamente la relación de las transformadas de Laplace de la señal de salida y la de entrada, en un determinado sistema, con condiciones iniciales nulas.

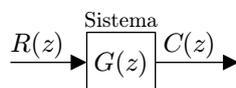


Figura 16.2: Función de transferencia \mathcal{Z}

También es posible definir una expresión matemática en z que sea la relación de las transformadas \mathcal{Z} de la salida respecto de la entrada, asumiendo condiciones iniciales nulas. Como habitualmente las transformadas \mathcal{Z} de las señales temporales son fracciones de polinomios en z , las funciones de transferencia \mathcal{Z} también serán fracciones de polinomios en z . Las raíces del polinomio numerador serán los ceros de la función de transferencia \mathcal{Z} y las raíces del polinomio denominador serán los polos.

$$G(z) = \frac{C(z)}{R(z)} = \frac{a_0 + a_1z + a_2z^2 + a_3z^3 + \dots}{b_0 + b_1z + b_2z^2 + b_3z^3 + \dots} \tag{16.18}$$

Igual que ocurría en el dominio continuo de Laplace, las funciones de transferencia \mathcal{Z} de sistemas físicos suelen poseer más polos que ceros, o como mucho, igual número de ceros que de polos.

16.6. Ecuaciones en diferencias

Saber escribir la *ecuación en diferencias* equivalente a una función \mathcal{Z} es una de las herramientas más interesantes de la asignatura, de cara a poder implementar los controladores digitales en lenguajes de programación como C o Matlab®. El objetivo es traducir la función de transferencia \mathcal{Z} que relaciona la salida con la entrada en operaciones elementales. Primero se mostrará un ejemplo. El punto de partida es la función de transferencia \mathcal{Z} en donde se ha dividido numerador y denominador por z las veces necesarias para que los términos en z de los polinomios tengan exponentes negativos. Considérese la función de transferencia:

$$G(z) = \frac{C(z)}{R(z)} = \frac{a_0 + a_1 z^{-1}}{b_0 + b_1 z^{-1} + b_2 z^{-2}}, \quad (16.19)$$

donde:

$$C(z)[b_0 + b_1 z^{-1} + b_2 z^{-2}] = R(z)[a_0 + a_1 z^{-1}] \quad (16.20)$$

$$b_0 C(z) + b_1 C(z)z^{-1} + b_2 C(z)z^{-2} = a_0 R(z) + a_1 R(z)z^{-1} \quad (16.21)$$

Aplicando la transformada inversa de \mathcal{Z} y el teorema del desplazamiento en el tiempo:

$$b_0 c(nT) + b_1 c(nT - T) + b_2 c(nT - 2T) = a_0 r(nT) + a_1 r(nT - T) \quad (16.22)$$

$$c(nT) = \frac{1}{b_0} [a_0 r(nT) + a_1 r(nT - T) - b_1 c(nT - T) - b_2 c(nT - 2T)] \quad (16.23)$$

Es decir, el valor de la salida en un instante arbitrario nT se puede calcular en función de la entrada en dicho instante, entradas en instantes anteriores y salidas en instantes anteriores. Se trata de una combinación lineal muy sencilla de implementar en código C o Matlab®. Además el número de términos que intervienen es finito, es decir, no hay que guardar infinitos valores anteriores de entrada y la salida.

Este tipo de ecuaciones se llaman *ecuaciones en diferencias* porque son una traducción muy sencilla de las ecuaciones diferenciales que gobiernan el comportamiento del sistema, pero con la ventaja de que no aparecen derivadas ni integrales. Sólo hay que realizar sumas, restas, multiplicaciones y/o divisiones.

Para cambiar un controlador en el dominio discreto, basta con modificar la ley que se haya implementado con ecuaciones en diferencias. Hay que advertir que el número de decimales que se use en los coeficientes es muy importante. Normalmente el ingeniero está acostumbrado a despreciar los decimales a partir de la segunda o tercera cifra significativa. En control digital es mejor no hacer eso. Cuantos más decimales se empleen, mejor adecuación habrá entre el comportamiento real del sistema y el esperado. Incluso es posible que un pequeño error a la hora de introducir los coeficientes (aparentemente insignificante, como cambiar el orden entre el cuarto y quinto decimal), puede hacer el sistema inestable.

Las ecuaciones en diferencias nos permiten proponer un tercer método a los anteriormente enunciados para calcular la transformada inversa de \mathcal{Z} . Se le llamará “método computacional” y consiste en determinar la ecuación diferencia asociada a la $F(z)$ y obtener la salida aplicando como entrada un impulso unidad. Esta metodología responde al concepto ya citado de considerar la transformada de una señal como la respuesta a un impulso unidad de la función de transferencia que posee la misma expresión.

En el siguiente ejemplo se calcula por medio de las ecuaciones en diferencias la transformada inversa \mathcal{Z} de la función (16.24) que, según la Tabla 16.1, debería corresponder a la unidad en todo instante de muestreo.

$$F(z) = \frac{1}{1 - z^{-1}} = \frac{F(z)}{1} = \frac{F(z)}{\Delta(z)} \quad (16.24)$$

En la ecuación (16.24) ya se especifica que se va a considerar como entrada el impulso unidad en el origen de tiempos. La ecuación diferencia asociada a dicha función es:

$$f(nT) = \delta(nT) + f(nT - T), \quad (16.25)$$

donde:

$$\delta(nT) = \begin{cases} 1 & \text{para } n = 0 \\ 0 & \text{para } n > 0 \end{cases}, \quad (16.26)$$

por tanto:

$$\begin{aligned} f(0) &= \delta(0) + f(-T) = 1 - 0 = 1 \\ f(T) &= \delta(T) + f(0) = 0 + 1 = 1 \\ f(2T) &= \delta(2T) + f(T) = 0 + 1 = 1 \\ f(3T) &= \delta(3T) + f(2T) = 0 + 1 = 1 \\ f(4T) &= \dots \end{aligned} \quad (16.27)$$

Con lo que se comprueba que efectivamente el valor función temporal en los instantes de muestreo es siempre la unidad. Por tanto la transformada inversa \mathcal{Z} es:

$$f(nT) = 1 \quad (16.28)$$

A continuación se escribe en lenguaje Matlab® la ecuación (16.25):

```
T = 0.001;
entrada = [0 1 0 0 0 0 0];
n = length(entrada);
tiempo = -T:T:T*(n-2);
salida = zeros(n,1);
for k = 2:1:n
    salida(k) = entrada(k) + salida(k-1);
end
plot(tiempo,salida,'o');
```

Se advierte que el programador ha decidido definir las variables desde un valor negativo de tiempos ($t = -T$) porque la ecuación diferencia necesita en el instante $t = 0$ el valor “anterior”. Hay que tener cuidado al definir los valores que debe tomar la variable del bucle para dar valores a los elementos de los vectores correctos. Por otro lado, al programador sólo le ha interesado calcular la salida hasta el instante $t = 5T$, pero eso se puede cambiar alargando la definición de la variable de entrada.

Capítulo 17

Diagramas de bloques en Zeta

En el capítulo anterior se vio cómo es posible definir funciones de transferencia \mathcal{Z} , en el presente este capítulo se conectarán distintos sistemas de forma gráfica a través de diagramas de bloques de forma análoga a como se hacía en el dominio continuo. Aunque la forma de trabajar es similar al dominio continuo, aparecen nuevos elementos en los diagramas (los retenedores y los muestreadores) y pueden haber “mezclados” bloques continuos (en S) y discretos (en \mathcal{Z}).

17.1. Generalidades

Los diagramas de bloques en \mathcal{Z} poseen análogas reglas de simplificación gráfica que los bloques continuos en el dominio de Laplace.

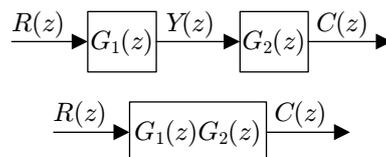


Figura 17.1: Simplificación de bloques en \mathcal{Z}

Dos bloques en serie, Fig. 17.1, se pueden sustituir por un único bloque cuya función de transferencia sea el producto de las funciones de transferencia originales. Como siempre que se hace este tipo de expresiones más compactas, se pierde la información las señales intermedias entre la entrada y la salida.

La única dificultad que plantea el uso de diagramas de bloques es que lo normal es los sistemas de control tengan partes continuas y partes discretas. Así pues, hay que poner especial atención al modo en que se “conectan” ambas partes. Lo normal será que para pasar del dominio continuo al discreto habrá que introducir un elemento muestreador, mientras que para pasar del dominio discreto al continuo habrá que introducir un elemento retenedor.

Antes de usar diagramas de bloques en \mathcal{Z} , es muy importante estudiar en el siguiente apartado cómo se simplifican diagramas de bloques continuos en presencia de muestreadores.

17.2. Bloques en cascada con muestreadores

17.2.1. Un único bloque continuo

En la Fig. 17.2 se muestra un sistema continuo precedido de un muestreador. El objetivo será determinar la salida del sistema.

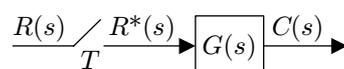


Figura 17.2: Bloque continuo después de un muestreador

La salida del sistema es:

$$C(s) = R^*(s)G(s) \tag{17.1}$$

Si existiera un nuevo muestreador después de la salida (entonces el usuario del sistema sólo sería capaz de conocer la salida en los instantes de muestreo), la señal $C(s)$ muestreada sería, por definición:

$$C^*(s) = \frac{1}{T} \sum_{n=-\infty}^{\infty} C(s - jn\omega_s) \quad (17.2)$$

Usando la expresión (17.1), sería:

$$C^*(s) = \frac{1}{T} \sum_{n=-\infty}^{\infty} R^*(s - jn\omega_s)G(s - jn\omega_s) \quad (17.3)$$

Como las funciones muestreadas son periódica, se cumple que:

$$R^*(s) = R^*(s - jn\omega_s), \quad (17.4)$$

y por tanto:

$$C^*(s) = R^*(s) \frac{1}{T} \sum_{n=-\infty}^{\infty} G(s - jn\omega_s) \quad (17.5)$$

$$C^*(s) = R^*(s)G^*(s) \quad (17.6)$$

Comparando las ecuaciones (17.1) y (17.6) se puede definir la “operación muestreo” en una ecuación en el dominio de Laplace y pasar directamente de una a otra. También se puede dar la expresión equivalente en \mathcal{Z} de la ecuación (17.6):

$$C(z) = R(z)G(z) \quad (17.7)$$

17.2.2. Bloques continuos con muestreador intermedio

En la Fig. 17.3 se muestra un sistema con dos bloques que representan sistemas continuos con una muestreador intermedio. Se supondrá que todos los muestreadores poseen igual periodo de muestreo y están “sincronizados”.

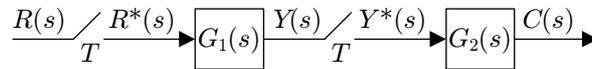


Figura 17.3: Bloques continuos con muestreador intermedio

Las ecuaciones que se pueden plantear para el sistema son:

$$\left. \begin{aligned} C(s) &= Y^*(s)G_2(s) \\ Y(s) &= R^*(s)G_1(s) \end{aligned} \right\}, \quad (17.8)$$

donde no es posible sustituir la segunda en la primera. Si aplicamos la “operación muestreo” a las dos ecuaciones (a partir de ahora se dirá “muestrear una ecuación”) resulta:

$$\left. \begin{aligned} C^*(s) &= Y^*(s)G_2^*(s) \\ Y^*(s) &= R^*(s)G_1^*(s) \end{aligned} \right\}, \quad (17.9)$$

y ahora ya se puede reducir el sistema de dos ecuaciones en:

$$C^*(s) = R^*(s)G_1^*(s)G_2^*(s). \quad (17.10)$$

Estrictamente hablando, no hubiera sido necesario muestrear la primera ecuación para reducir las dos ecuaciones a una sola. Muestreando sólo la segunda ecuación se hubiera llegado a la siguiente expresión para la salida:

$$C(s) = R^*(s)G_1^*(s)G_2(s). \quad (17.11)$$

Y con (17.11) se puede obtener (17.10) aplicando a su vez la operación muestreo. En cualquier caso, en los sistemas discretos que se pretende estudiar, siempre se presta atención a la salida muestreada en lugar de a la salida continua (por tanto, aunque realmente no lo haya, siempre se supone que después de la salida continua hay un muestreador final).

17.2.3. Bloques continuos sin muestreador intermedio: el problema de la convolución

En la Fig. 17.4 se muestra un sistema con dos bloques continuos sin una muestreador intermedio. Las ecuaciones que se pueden plantear para el sistema son:

$$\left. \begin{aligned} C^*(s) &= Y(s)G_2(s) \\ Y(s) &= R^*(s)G_1(s) \end{aligned} \right\}, \quad (17.12)$$

en definitiva:

$$C(s) = R^*(s)G_1(s)G_2(s). \quad (17.13)$$

Muestreando la ecuación:

$$C^*(s) = R^*(s)[G_1(s)G_2(s)]^* \quad (17.14)$$

Evidentemente este caso se reduce al presentado en el apartado 17.2.1 si se simplifican previamente los dos bloques continuos por su producto. Lo que se quiere destacar es que al muestrear una ecuación sólo se pueden “sacar fuera de la operación muestreo” las señales que ya estén muestreadas, y calcular la transformada de Laplace muestreada del producto de todas las que sean continuas. Esto equivale a decir que en el anterior paso de (17.3) a (17.5) sólo pueden salir del sumatorio las señales muestreadas.

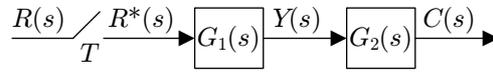


Figura 17.4: Bloques continuos sin muestreador intermedio

Escribiendo la expresión equivalente en el dominio \mathcal{Z} de la ecuación (17.14):

$$C(z) = R(z)\mathcal{Z}[G_1(s)G_2(s)] \quad (17.15)$$

Dado que:

$$\mathcal{Z}[G_1(s)G_2(s)] \neq \mathcal{Z}[G_1(s)]\mathcal{Z}[G_2(s)] = G_1(z)G_2(z), \quad (17.16)$$

por tanto:

$$C(z) = R(z)\mathcal{Z}[G_1(s)G_2(s)] \neq R(z)G_1(z)G_2(z) \quad (17.17)$$

En el siguiente ejemplo se muestran en distintos pasos la *forma correcta* de aplicar la operación muestreo a una ecuación en el dominio de Laplace:

$$A(s) = B(s)C^*(s)D(s)E(s)F^*(s) \quad (17.18)$$

$$[A(s)]^* = [B(s)C^*(s)D(s)E(s)F^*(s)]^* \quad (17.19)$$

$$A^*(s) = C^*(s)F^*(s)[B(s)D(s)E(s)]^* \quad (17.20)$$

Con el mismo ejemplo se muestra la *forma incorrecta* de muestrear la ecuación:

$$A(s) = B(s)C^*(s)D(s)E(s)F^*(s) \quad (17.21)$$

$$[A(s)]^* = [B(s)C^*(s)D(s)E(s)F^*(s)]^* \quad (17.22)$$

$$A^*(s) \neq B^*(s)C^*(s)[D(s)E(s)]^*F^*(s) \quad (17.23)$$

Es muy importante manejar muy bien la operación muestreo para conseguir simplificar correctamente los diagramas de bloques que mezclan partes continuas y partes discretas.

17.2.4. Sistemas en lazo cerrado

Con las indicaciones de los apartados anteriores, se calculan en este apartado las salidas muestreadas de varios sistemas de control realimentados con muestreadores en algún punto del lazo.

Para el caso de la Fig. 17.5 donde se muestrea sólo el error, la salida es:

$$C^*(s) = \frac{G^*(s)}{1 + [G(s)H(s)]^*} R^*(s) \quad (17.24)$$

Como se ve en la ecuación (17.24), es posible obtener una función de transferencia en lazo cerrado que relacione la referencia muestreada con la salida muestreada.

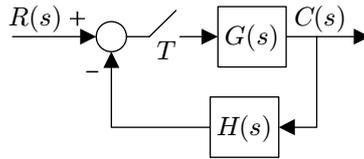


Figura 17.5: Ejemplo 1 de sistema en lazo cerrado

La salida para sistema de la Fig. 17.6, donde se supone que los muestreadores están sincronizados, es:

$$C^*(s) = \frac{G^*(s)}{1 + G^*(s)H^*(s)} R^*(s) \quad (17.25)$$

De nuevo es posible obtener una función de transferencia en lazo cerrado que relacione la referencia muestreada con la salida muestreada. Sin embargo, la ecuación característica es distinta.

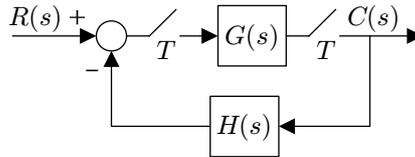


Figura 17.6: Ejemplo 2 de sistema en lazo cerrado

En el caso del sistema de la Fig. 17.7, la salida es:

$$C^*(s) = \frac{[G(s)R(s)]^*}{1 + [G(s)H(s)]^*} \quad (17.26)$$

donde no es posible encontrar una función de transferencia en lazo cerrado que relacione la referencia muestreada con la salida muestreada. Sin embargo, esto no es una dificultad para el ingeniero ya que, en cualquier caso, es posible identificar la ecuación característica del sistema. Dicha ecuación característica es el denominador de la salida muestreada $C^*(s)$ igualado a cero. Es interesante señalar que tanto el sistema del ejemplo 1 como el del ejemplo 3 tienen la misma ecuación característica.

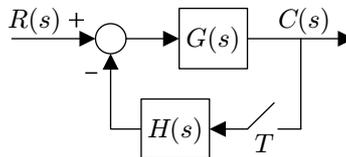


Figura 17.7: Ejemplo 3 de sistema en lazo cerrado

La salida del cuarto ejemplo de sistema en lazo cerrado que aparece en la Fig. 17.8 es:

$$C^*(s) = \frac{G_2^*(s)[G_1(s)R(s)]^*}{1 + [G_1(s)G_2(s)H(s)]^*}, \quad (17.27)$$

donde tampoco es posible encontrar la función de transferencia muestreada en lazo cerrado, pero sí la ecuación característica del sistema.

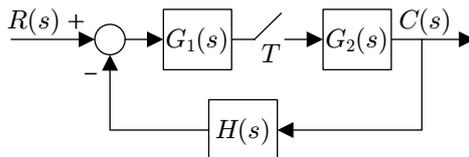


Figura 17.8: Ejemplo 4 de sistema en lazo cerrado

En cambio, calculando la salida del sistema de la Fig. 17.9, sé es posible identificar la función de transferencia en lazo cerrado del sistema de control:

$$C^*(s) = \frac{G_1^*(s)G_2^*(s)}{1 + G_1^*(s)[G_2(s)H(s)]^*} R^*(s), \quad (17.28)$$

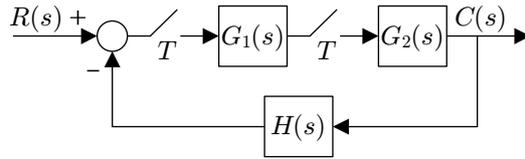


Figura 17.9: Ejemplo 5 de sistema en lazo cerrado

Realizando los ejemplos de este apartado es posible encontrar muchas dificultades a la hora de simplificar las ecuaciones de un sistema muestreado. En concreto, la clave es saber cuándo hay que usar la operación muestreo sobre una ecuación y cuándo no.

En el siguiente apartado se describe un método que evita todos estos problemas, es decir, siguiendo sus pasos siempre será posible encontrar la salida muestreada del sistema.

17.3. Método de simplificación

Al método de simplificación de sistemas muestreados que se describe en este apartado se le denominará “método de Phillips-Nagle” [5]. Por muy difícil que sea el diagrama de bloques que se pretenda resolver y no importa el número de muestreadores que contenga, siempre será posible encontrar la salida muestreada del sistema siguiendo los pasos de este método.

Los pasos que se deben seguir son los siguientes:

1. Tomar como “entradas” las entradas verdaderas y las salidas de los muestreadores.
2. Tomar como “salidas” las salidas verdaderas y las entradas en los muestreadores.
3. Expresar las “salidas” en función de las “entradas”.
4. Muestrear todas las “salidas”.
5. Despejar la salida real del sistema resolviendo el sistema de ecuaciones.

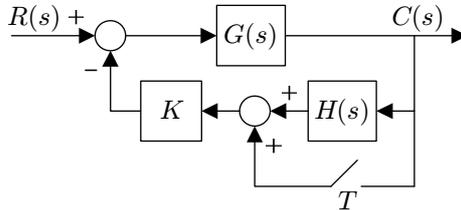


Figura 17.10: Ejemplo de sistema de control

A continuación se aplica el método al sistema de la Fig. 17.10:

1. Se toman como “entradas” del sistema: $R(s)$ y $C^*(s)$.
2. Se toma como “salidas” del sistema: $C(s)$
3. Se expresan las “salidas” en función de las “entradas”:

$$C(s) = G(s)\{R(s) - K[H(s)C(s) + C^*(s)]\} \quad (17.29)$$

$$C(s)[1 + KG(s)H(s)] = G(s)R(s) - KG(s)C^*(s) \quad (17.30)$$

$$C(s) = \frac{G(s)R(s)}{1 + KG(s)H(s)} - \frac{KG(s)C^*(s)}{1 + KG(s)H(s)} \quad (17.31)$$

4. Se muestrean todas las “salidas” (en este caso es única):

$$C^*(s) = \left[\frac{G(s)R(s)}{1 + KG(s)H(s)} \right]^* - K \left[\frac{G(s)}{1 + KG(s)H(s)} \right]^* C^*(s) \quad (17.32)$$

5. Se despeja la salida real del sistema:

$$C^*(s) = \frac{\left[\frac{G(s)R(s)}{1 + KG(s)H(s)} \right]^*}{1 + K \left[\frac{G(s)}{1 + KG(s)H(s)} \right]^*} \quad (17.33)$$

17.4. Sistemas con bloques continuos y discretos

Es difícil encontrar sistemas cuyos diagramas de bloques posean sólo funciones de transferencia continuas y muestreadores. Lo habitual será que existan también bloques discretos y retenedores.

En la Fig. 17.11 se muestra la arquitectura de control típica en tiempo discreto. El error se mide con un convertidor A/D que muestrea la señal. El controlador se implementa por medio de ecuaciones en diferencias que responden a un bloque en \mathcal{Z} . La actuación del controlador es una señal muestreada que se convierte en continua a través de un convertidor D/A que actúa como un retenedor de orden cero. La planta que se desea controlar es continua así como la señal de salida del sistema.

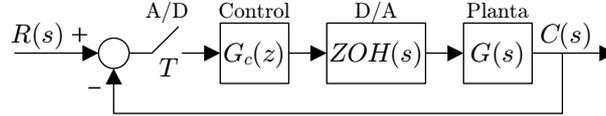


Figura 17.11: Sistema de control discreto

La salida del sistema de la Fig. 17.11 es:

$$C^*(s) = \frac{G_c(z)[ZOH(s)G(s)]^*}{1 + G_c(z)[ZOH(s)G(s)]^*} R^*(s), \quad (17.34)$$

en la variable compleja z :

$$C(z) = \frac{G_c(z)\mathcal{Z}[ZOH(s)G(s)]}{1 + G_c(z)\mathcal{Z}[ZOH(s)G(s)]} R(z). \quad (17.35)$$

Por tanto, la ecuación característica del sistema en z es:

$$1 + G_c(z)\mathcal{Z}[ZOH(s)G(s)] = 0 \quad (17.36)$$

Es muy parecida a la que se obtenía en continuo pero ahora aparece un nuevo bloque que es el del retenedor que parece modificar la planta que se pretende controlar. En realidad, *el retenedor de orden cero no modifica el orden de la planta*. Si la planta tenía dos polos en s , la ecuación característica en z que se obtenga al operar en (17.36) será un polinomio de orden 2, es decir, tendrá el mismo número de polos que tenía la planta en continuo.

La transformada \mathcal{Z} que aparece en la ecuación (17.36) se puede operar de la forma:

$$\mathcal{Z}[ZOH(s)G(s)] = \mathcal{Z}\left[\frac{1 - e^{-Ts}}{s}G(s)\right] = (1 - z^{-1})\mathcal{Z}\left[\frac{G(s)}{s}\right] \quad (17.37)$$

Es decir, lo que habrá que buscar en la tabla de las transformadas \mathcal{Z} la función $G(s)$ dividida por s , y no directamente $G(s)$. Es tan habitual este sistema de control y el uso del retenedor de orden cero, que algunos autores definen:

$$G(z) = \mathcal{Z}[ZOH(s)G(s)] = (1 - z^{-1})\mathcal{Z}\left[\frac{G(s)}{s}\right] \quad (17.38)$$

Otros autores emplean un símbolo especial:

$$\bar{G}(z) = \mathcal{Z}[ZOH(s)G(s)] = (1 - z^{-1})\mathcal{Z}\left[\frac{G(s)}{s}\right] \quad (17.39)$$

En este libro nunca se usará la definición (17.38) y siempre se ha entendido que:

$$G(z) = \mathcal{Z}[G(s)] \quad (17.40)$$

En cualquier caso el ingeniero deberá estar prevenido sobre qué tipo de nomenclatura se está usando en cada manual.

Capítulo 18

Correspondencia entre el plano S y el plano Z

En el capítulo anterior se han obtenido las primeras ecuaciones características en la variable compleja z . La posición de los polos en lazo cerrado dentro del plano complejo S se utiliza mucho en control continuo, por ejemplo en el método del lugar de las raíces. El presente capítulo pretende estudiar qué características posee el plano complejo Z y cómo trasladar a él las propiedades que se conocían en S .

18.1. Franja primaria y círculo unitario

La relación entre los planos S y Z se deduce de la expresión que relaciona las dos variables complejas:

$$z = e^{Ts} \tag{18.1}$$

Aplicando esa transformación a los puntos que delimitan la franja primaria del plano complejo S :

$$s = 0 \implies z = 1 \tag{18.2}$$

$$s = j\omega_d \implies z = 1 \angle T\omega_d \tag{18.3}$$

$$s = j\omega_N \implies z = 1 \angle \pi \tag{18.4}$$

$$s = -\sigma + j\omega_N \implies z = e^{-\sigma T} \angle \pi \tag{18.5}$$

$$s = -\infty + j\omega_N \implies z = 0 \angle \pi \tag{18.6}$$

$$s = -\infty - j\omega_N \implies z = 0 \angle -\pi \tag{18.7}$$

$$s = -\sigma - j\omega_N \implies z = e^{-\sigma T} \angle -\pi \tag{18.8}$$

$$s = -j\omega_N \implies z = 1 \angle -\pi \tag{18.9}$$

$$s = -j\omega_d \implies z = 1 \angle -T\omega_d \tag{18.10}$$

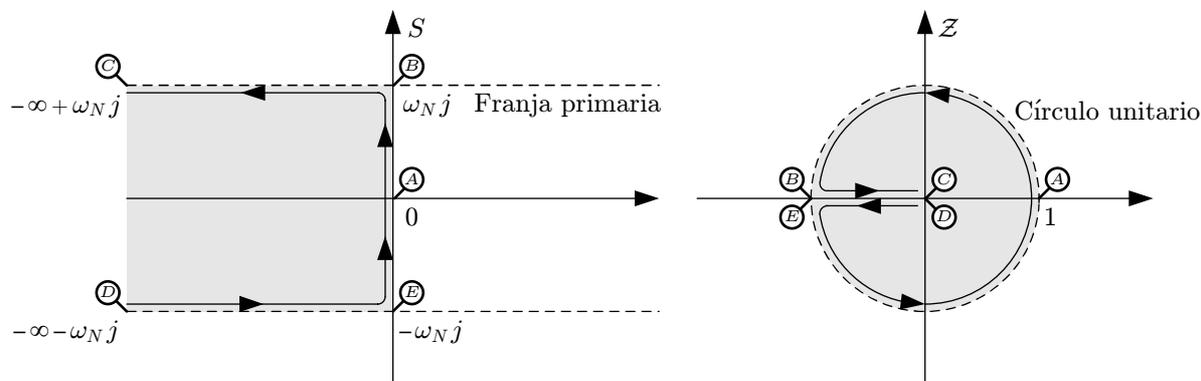


Figura 18.1: Transformación del plano S al plano Z

Representando los resultados, Fig. 18.1, se observa que la semi-franja primaria de parte real negativa se transforma en la región del plano complejo Z que está en el interior de la circunferencia de radio

unidad. La semi-franja primaria de parte real positiva se transforma en la región exterior al círculo unitario. Las franjas complementarias se superponen en estas mismas regiones. Por tanto, un punto en el plano complejo Z posee infinitos puntos equivalentes en el plano complejo S . Esto es congruente con la propiedad de la periodicidad de la transformada de Laplace de una función muestreada. En realidad, el plano complejo Z posee la ventaja de que el ingeniero puede centrar su estudio en los polos que caen dentro de la franja primaria y dar por supuesto que se repite lo mismo en las franjas complementarias.

Con este resultado, ya se puede adelantar que la condición de estabilidad en sistemas discretos será que todos los polos de la función de transferencia discreta se encuentren en el interior del círculo unitario.

18.2. Líneas de parámetros constantes

Con la expresión (18.1) también se pueden deducir las formas que poseen las líneas con características de frecuencia natural, amortiguamiento, atenuación y frecuencia natural amortiguada constante. En la Fig. 18.2 se muestran las líneas de frecuencia natural y amortiguamiento constantes. Estas líneas son perpendiculares entre sí tanto en el plano S como en el Z .

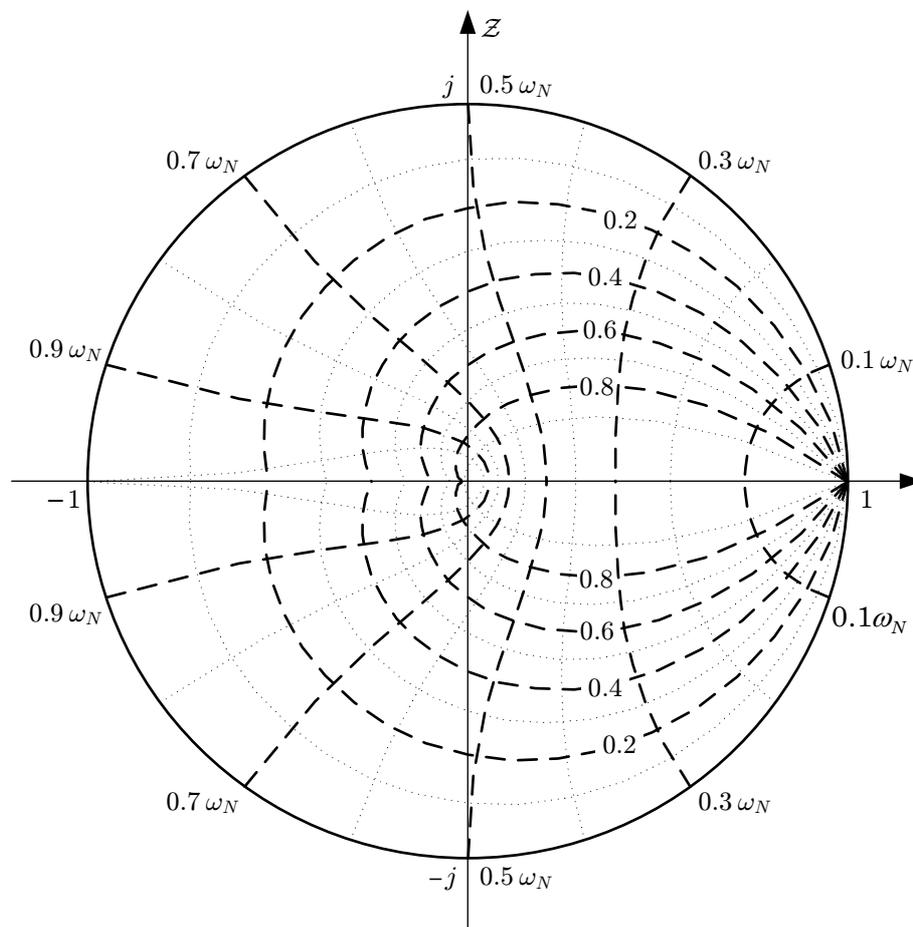


Figura 18.2: Líneas de frecuencia natural y amortiguamiento constantes

En la Fig. 18.3 se muestran las líneas de frecuencia natural amortiguada y atenuación constantes. También estas líneas son perpendiculares entre sí tanto en el plano S como en el Z . No se han dado valores a las líneas de atenuación constante porque se suelen usar poco.

18.3. Variación de la posición de los polos y ceros con T

Las líneas de parámetros constantes del apartado anterior se pueden usar para elegir la localización de los polos en lazo cerrado de un sistema de control para asegurar un determinado comportamiento en régimen transitorio de igual forma que se hacía en control en tiempo continuo.

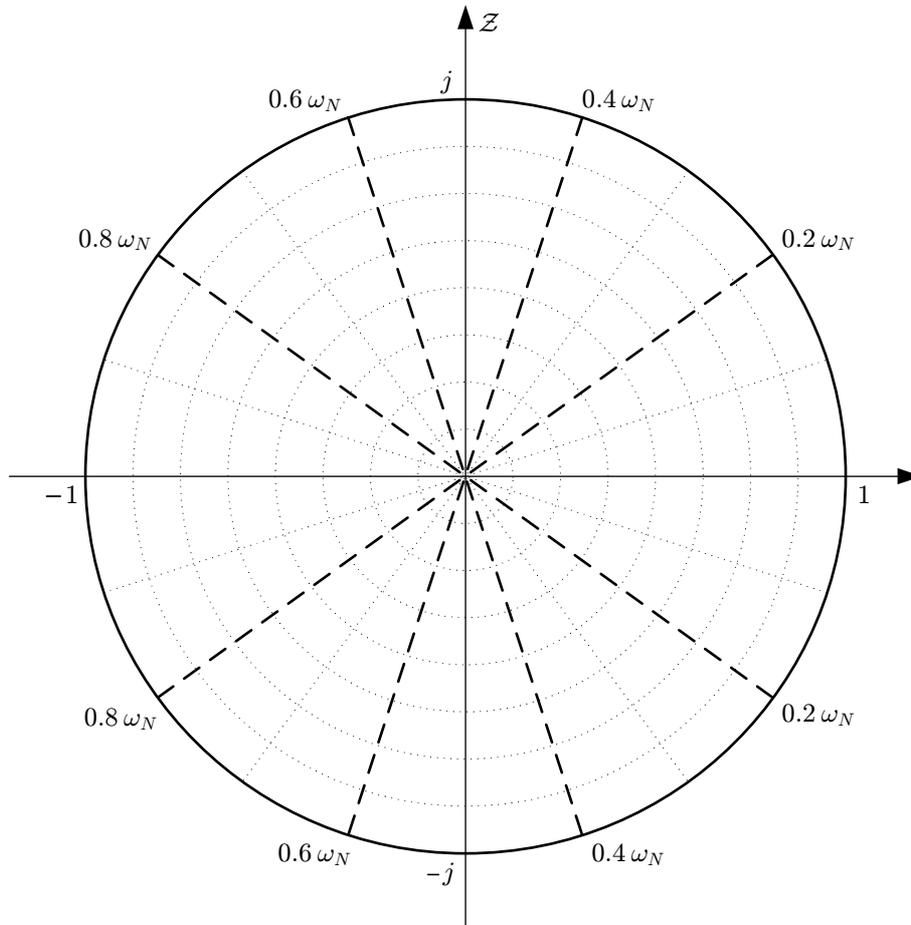


Figura 18.3: Líneas de frecuencia natural amortiguada y atenuación constantes

Así pues, si s_o es la posición de los “polos objetivo” en continuo, la posición dentro del plano complejo \mathcal{Z} de esos mismos polos objetivo es:

$$z_o = e^{Ts_o} \quad (18.11)$$

La fórmula (18.11) previene al ingeniero de un hecho muy importante: la posición de los “polos objetivo” dentro del plano complejo \mathcal{Z} depende del periodo de muestreo T que tenga o que se haya elegido para el sistema.

Como ejercicio se puede representar en el plano complejo \mathcal{Z} la posición del polo doble complejo conjugado $-1 \pm 1j$ cuando el periodo de muestreo es 1 segundo y 0.1 segundos, Fig. 18.4. Estas posiciones son:

$$s_{1,2} = -1 \pm 1j \begin{cases} z_{1,2} = 0.1987 \pm 0.3095j \text{ para } T = 1 \text{ s} \\ z_{1,2} = 0.9003 \pm 0.0903j \text{ para } T = 0.1 \text{ s} \end{cases} \quad (18.12)$$

Usando módulo y argumento:

$$s_{1,2} = -1 \pm 1j \begin{cases} z_{1,2} = 0.3678 \angle \pm 1 \text{ rad para } T = 1 \text{ s} \\ z_{1,2} = 0.9048 \angle \pm 0.1 \text{ rad para } T = 0.1 \text{ s} \end{cases} \quad (18.13)$$

Esta última forma es más útil para comprobar si un polo está dentro o fuera del círculo unitario. Lo normal en este libro de texto es que se usen siempre cuatro decimales para la posición de los polos y ceros en el plano complejo \mathcal{Z} .

También se puede observar que, cuando se usan periodos de muestreo muy pequeños (es lo habitual), los polos se van a encontrar siempre muy próximos de la unidad (que corresponde al origen de coordenadas del plano complejo \mathcal{S}). Por este motivo es tan importante usar varios decimales, de lo contrario se pierde mucha información.

Al cambiar el periodo de muestreo, automáticamente cambia la “anchura” de la franja primaria. Por tanto la posición relativa del polo, en el plano complejo \mathcal{S} , cambia respecto a dicha franja. Es por eso que cambia la posición del polo en el plano complejo \mathcal{Z} . Sin embargo, el ángulo que forma respecto al origen no

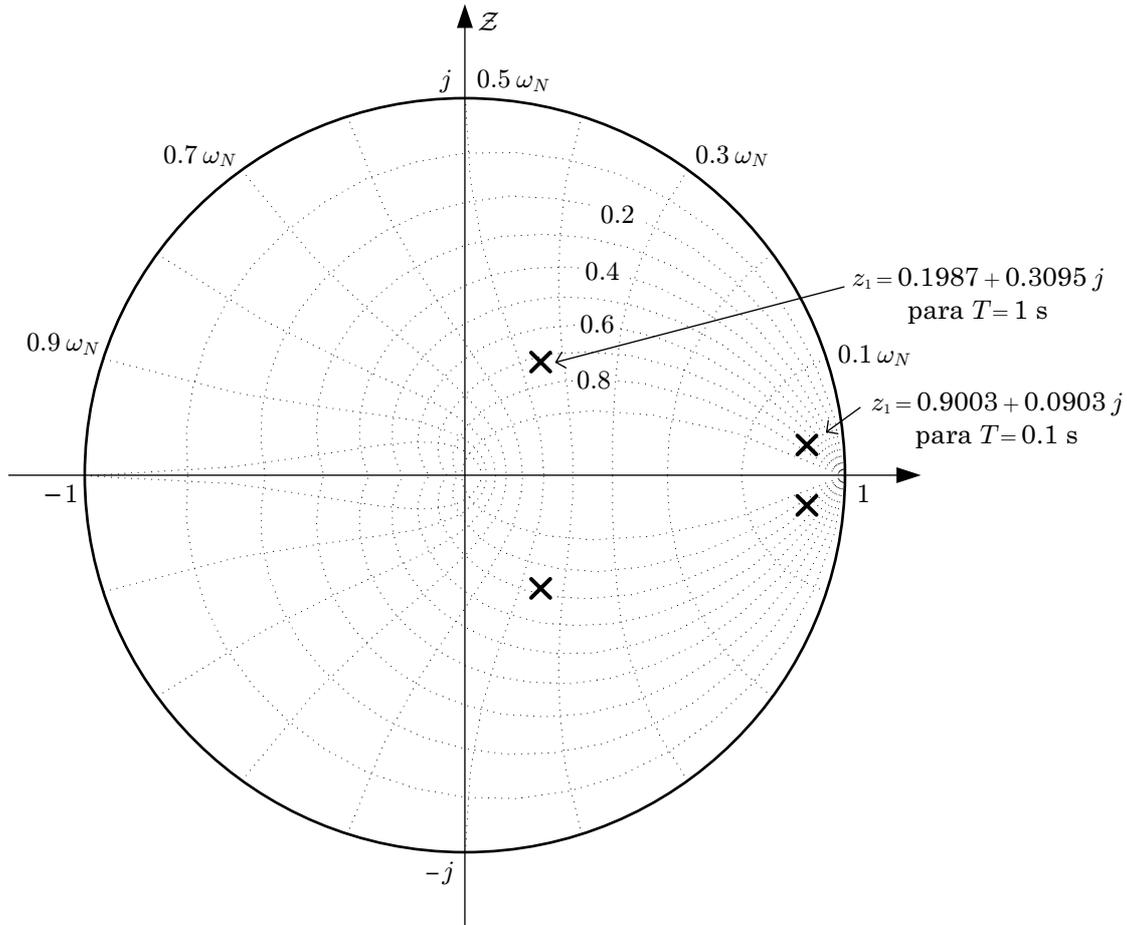


Figura 18.4: Posición de los polos en función del periodo de muestreo

cambia con el periodo de muestreo. Por ese motivo los polos y ceros en el plano complejo Z se “mueven” o cambian de posición con diferentes periodos de muestreo según trayectorias de amortiguamiento constante. Compruébese en el ejemplo de la Fig. 18.4.

Para el mapeo entre S y Z se pueden usar las siguientes relaciones:

$$s = -\sigma \pm \omega_d j = -\zeta \omega_n \pm \omega_n \sqrt{1 - \zeta^2} j \quad (18.14)$$

$$|z| = e^{-T\zeta\omega_n} \implies \ln |z| = -T\zeta\omega_n = -T\sigma \quad (18.15)$$

$$\angle z = T\omega_n \sqrt{1 - \zeta^2} \text{ rad} = T\omega_d \text{ rad}, \quad (18.16)$$

que se deducen aplicando directamente la relación (18.1).

18.4. Cálculo del número de muestras por ciclo

El hecho antes citado de que la franja primaria varíe su anchura con el periodo de muestreo, va a servir de ayuda de cara a la elección del periodo de muestreo. En definitiva, se trata de conseguir que todos los polos y ceros del sistema queden dentro de la franja primaria.

El periodo de muestreo de la Fig. 18.5 a) es menor que en el caso b), porque cuanto mayor es el periodo de muestreo menor es la anchura de la franja primaria. En el caso a) los polos del sistema quedan dentro de la franja primaria, mientras que en el caso b) caen fuera. También se observan las posiciones que tendrán las réplicas de los polos del sistema debido a al muestreo. Es muy interesante destacar que cuando se trabaje en el plano Z , en el caso a) sólo se verán los polos del sistema y ninguna de las réplicas debidas al muestreo; mientras que en el caso b) se trabajará con unos polos que no son los del sistema.

Este error es la traducción en el plano complejo S del efecto del *aliasing* que se describió en el apartado 13.3. Las réplicas de los polos se pueden interpretar las señales que pueden engañar al ingeniero, que instintivamente se queda con la de menor frecuencia (aquellas que caen dentro de la franja primaria).

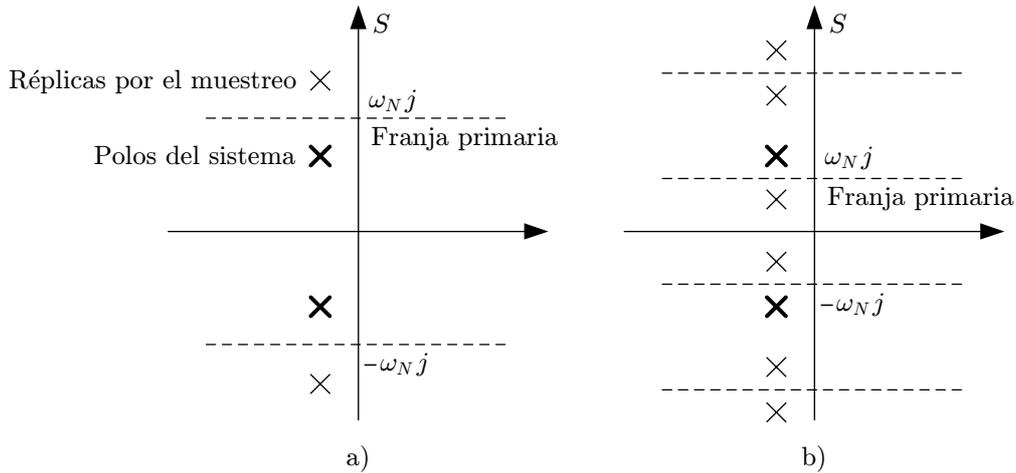


Figura 18.5: Réplicas de los polos del sistema en función del periodo de muestreo

Por tanto, dado un sistema concreto la manera de evitar el fenómeno del *aliasing* es elegir un periodo de muestreo tal que la frecuencia de Nyquist sea mayor que la frecuencia natural amortiguada asociada a todos sus polos.

$$\omega_N > \omega_d \quad (18.17)$$

$$\omega_s > 2\omega_d \quad (18.18)$$

$$T < \frac{\pi}{\omega_d} \quad (18.19)$$

Si lo que se desea es ajustar un sistema en lazo cerrado, lo que habrá que asegurar es que la frecuencia natural amortiguada de los polos en lazo cerrado (y en particular los “polos objetivo”) es menor que la de Nyquist.

Estos valores son los “mínimos” para que no se produzca el *aliasing*. Lo normal es que se elija un periodo de muestreo bastante menor. Aquí cada ingeniero puede determinar qué margen utiliza. Lo normal es pensar en términos de “muestras por ciclo”. La frecuencia natural amortiguada está asociada al periodo de oscilación de la salida del sistema ante una entrada escalón, por tanto, se puede elegir el periodo de muestreo de cara a asegurar un número determinado de muestreos dentro de un ciclo de oscilación de la salida.

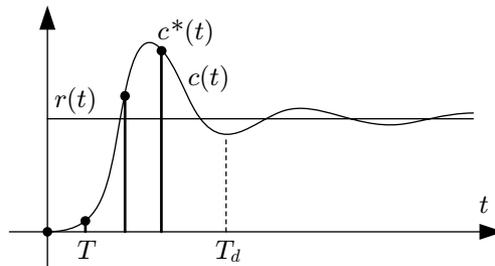


Figura 18.6: Número de muestras por ciclo dependiendo del periodo de muestreo

Por ejemplo en la Fig. 18.6 se ha elegido un periodo de muestreo que asegura al menos cuatro muestras por ciclo de oscilación. El mínimo que establece el teorema de Shannon para que no exista *aliasing* es dos muestras por ciclo de oscilación. Algunos autores sugieren elegir un periodo de muestreo que asegure entre 8 y 10 muestras por ciclo. Con este criterio el periodo de muestreo debería ser:

$$\frac{\pi}{4\omega_d} \leq T \leq \frac{\pi}{5\omega_d} \quad (18.20)$$

Lo que se propone en este libro de texto es elegir el periodo de muestreo más “redondo” que se obtenga dentro del intervalo de la inecuación (18.20).

Capítulo 19

Análisis de estabilidad

La estabilidad es una característica muy importante en Ingeniería de Control. Un sistema es estable cuando ante una entrada de amplitud finita responde el sistema con una salida también de amplitud finita. Por el contrario, un sistema inestable presentará oscilaciones que tienden a infinito ante entradas finitas.

Este tipo de oscilaciones arbitrariamente grandes ante entradas finitas pueden causar serios daños a las personas y al propio sistema (que normalmente posee un espacio de trabajo limitado). Por tanto, siempre hay que asegurar la estabilidad de los sistemas y tomar ciertas medidas de protección durante los experimentos de ajuste (setas de interrupción de corriente, botones de “hombre muerto”, etc.).

19.1. Criterio general

El criterio de estabilidad en el plano complejo S era que todos los polos del sistema tuvieran parte real negativa. Con la transformación del plano S a Z vista en el capítulo anterior, se puede afirmar como criterio general que un *sistema discreto es estable si todos sus polos se encuentran dentro del círculo unitario*. La posición de los ceros no afecta a la estabilidad.

Los polos que se encuentren justo sobre la circunferencia de radio unidad harán que la salida posea una oscilación permanente no amortiguada y de frecuencia igual a la frecuencia de Nyquist.

En la ecuación característica escrita en términos de z no se pueden aplicar ni la condición Cardano-Viète ni el criterio de Routh-Hurwitz. Así por ejemplo, un sistema con la siguiente ecuación característica:

$$z - 0.5 = 0 \tag{19.1}$$

Es evidente que su único polo está dentro del círculo unitario y por tanto el sistema es estable. Sin embargo, no se cumple la condición de Cardano-Viète de que todos los coeficientes del polinomio deben ser positivos y no nulos.

Para ecuaciones características con más de dos polos, es necesario contar con un criterio de estabilidad equivalente al de Routh-Hurwitz. Es decir, que diga si el sistema es estable o no sin necesidad de obtener analíticamente todas las raíces de dicha ecuación.

19.2. Criterio de Jury

El criterio de estabilidad de Jury permite deducir la estabilidad o no de un sistema discreto sin necesidad de calcular todas las raíces de la ecuación característica.

Sea la ecuación característica:

$$P(z) = a_n z^n + a_{n-1} z^{n-1} + \dots + a_1 z + a_0 = 0, \tag{19.2}$$

donde $a_n > 0$. El sistema es estable si y sólo si, se cumplen las siguientes condiciones:

$$|a_0| < a_n \tag{19.3}$$

$$P(1) > 0 \tag{19.4}$$

$$P(-1) \begin{cases} > 0 \text{ si } n \text{ es par} \\ < 0 \text{ si } n \text{ es impar} \end{cases} \tag{19.5}$$

Cuando $n > 2$, se deben cumplir más condiciones que se obtienen construyendo la tabla de Jury. A continuación se muestra la forma de la tabla de Jury para el caso de que $P(z)$ sea de cuarto orden:

$$\begin{array}{ccccc}
 z^0 & z^1 & z^2 & z^3 & z^4 \\
 \hline
 a_0 & a_1 & a_2 & a_3 & a_4 \\
 a_4 & a_3 & a_2 & a_1 & a_0 \\
 \hline
 b_0 & b_1 & b_2 & b_3 & \\
 b_3 & b_2 & b_1 & b_0 & \\
 \hline
 c_0 & c_1 & c_2 & &
 \end{array} \tag{19.6}$$

Las sucesivas filas se construyen, hasta que sólo haya tres términos en la última fila, de la siguiente forma:

$$b_0 = \begin{vmatrix} a_0 & a_4 \\ a_4 & a_0 \end{vmatrix} \tag{19.7}$$

$$b_1 = \begin{vmatrix} a_0 & a_3 \\ a_4 & a_1 \end{vmatrix} \tag{19.8}$$

$$b_2 = \begin{vmatrix} a_0 & a_2 \\ a_4 & a_2 \end{vmatrix} \tag{19.9}$$

$$b_3 = \begin{vmatrix} a_0 & a_1 \\ a_4 & a_3 \end{vmatrix} \tag{19.10}$$

Donde se observa que los sucesivos determinantes poseen siempre la misma primera columna. Las nuevas condiciones que se tienen que cumplir, para que el sistema sean estable, son:

$$|b_0| > |b_3| \tag{19.11}$$

$$|c_0| > |c_2| \tag{19.12}$$

En general, en cada nueva fila resultante, el valor absoluto del primer valor debe ser mayor que el valor absoluto del último.

Como ejercicios, se propone determinar la estabilidad de las siguientes ecuaciones características mediante el criterio de Jury:

$$P(z) = z^3 - 1.1z^2 - 0.1z + 0.2 = 0 \tag{19.13}$$

$$P(z) = z^3 - 1.3z^2 - 0.08z + 0.24 = 0 \tag{19.14}$$

$$P(z) = z^4 - 1.2z^3 + 0.07z^2 + 0.3z - 0.08 = 0 \tag{19.15}$$

19.3. Transformación bilineal y criterio de Routh-Hurwitz

El método de Jury requiere habitualmente más operaciones que el criterio de Routh-Hurwitz. Además, cuando alguno de los coeficientes contiene algún parámetro alfanumérico, trabajar con valores absolutos es especialmente complicado. Por este motivo, se han buscado formas para cambiar la ecuación característica en z de forma que se pueda emplear el criterio de Routh-Hurwitz ya conocido.

Las transformaciones bilineales o de Möbius (se puede escribir Moebius) es una herramienta que permite realizar cambios de variables muy interesantes con la siguiente expresión general:

$$z = \frac{aw + b}{cw + d} \text{ con } a, b, c, d \in \mathbb{C}, ad - bc \neq 0 \tag{19.16}$$

La variable z definida en el plano complejo \mathcal{Z} se transforma en la nueva variable w definida en el plano complejo \mathcal{W} . Estas transformaciones pueden ser traslaciones, dilataciones, rotaciones e inversiones. En el caso que nos ocupa, se usará la siguiente transformación bilineal:

$$z = \frac{w + 1}{w - 1} \tag{19.17}$$

Este cambio de variable, transforma el interior del círculo unitario en \mathcal{Z} el semiplano de parte real negativa en \mathcal{W} , Fig. 19.1. Si se comparan el plano S y el plano W , se pueden anotar algunas importantes diferencias. En el plano W sólo aparecerán los polos del plano S contenidos en la franja primaria. Las réplicas debidas al muestreo no se tendrán en cuenta. La posición de los polos dentro del plano W no

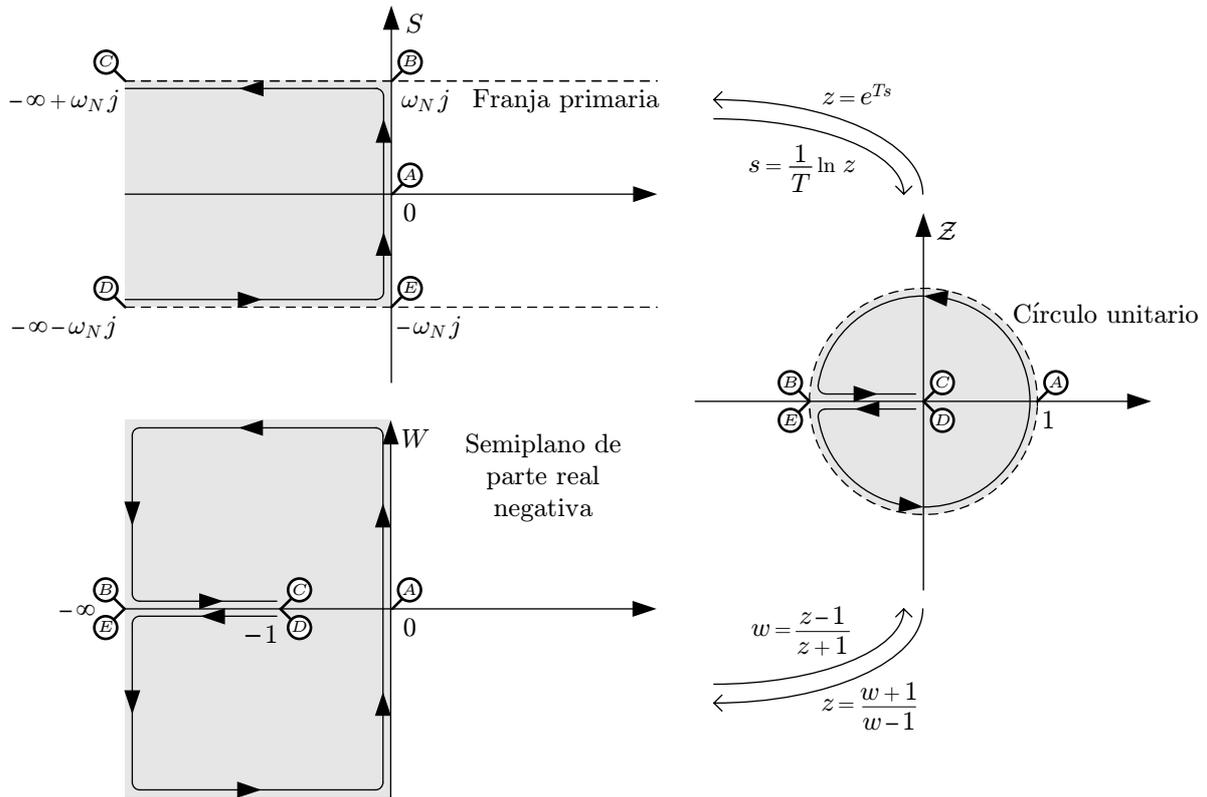


Figura 19.1: Cambios de variable entre el plano S , el plano Z y el plano W

ofrece información relacionada con la respuesta transitoria del sistema como ocurría dentro del plano S . Así por ejemplo, la parte real negativa no se puede asociar a la atenuación, ni la parte imaginaria con la frecuencia natural amortiguada del sistema.

Sin embargo, el carácter positivo a negativo de la parte real de los polos dentro del plano W sí que se conserva con las sucesivas transformaciones. Además, la transformación bilineal conduce a una ecuación característica que será un polinomio en w , por lo que se podrá aplicar el criterio de Routh-Hurwitz para determinar la estabilidad de los polos. Recuérdese que la variable s parece en las potencias de los términos exponenciales, por lo que no es posible aplicar en el dominio S el criterio de Routh-Hurwitz aún cuando se usen expresiones cerradas.

19.4. Ejemplo

Determinar la estabilidad del siguiente sistema en función del parámetro K por cualquiera de los métodos estudiados hasta el momento:

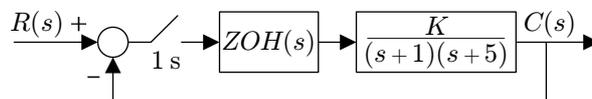


Figura 19.2: Ejemplo de sistema de discreto

En primer lugar se identifica la ecuación característica del sistema:

$$1 + \mathcal{Z} \left[\text{ZOH}(s) \frac{K}{(s+1)(s+5)} \right] = 0 \quad (19.18)$$

$$1 + (1 - z^{-1}) \mathcal{Z} \left[\frac{K}{s(s+1)(s+5)} \right] = 0 \quad (19.19)$$

Es interesante recordar que el sistema es de segundo orden y —aunque parezca en la ecuación anterior que hay tres polos— la ecuación característica en z tiene que quedar de segundo orden después de calcular la transformada \mathcal{Z} .

Descomponiendo en fracciones simples y acudiendo a la Tabla 16.1:

$$1 + (1 - z^{-1})KZ \left[\frac{1}{5} \frac{z}{z-1} - \frac{1}{4} \frac{z}{z-0.367} + \frac{1}{20} \frac{z}{z-0.00673} \right] = 0 \quad (19.20)$$

$$1 + \frac{K}{5} - \frac{K}{4} \frac{z-1}{z-0.367} + \frac{K}{20} \frac{z-1}{z-0.00673} = 0 \quad (19.21)$$

$$20z^2 + (2.1718K - 7.474)z + 0.3432K + 0.0494 = 0 \quad (19.22)$$

Efectivamente queda una ecuación de segundo orden. En este caso, es relativamente sencillo aplicar el criterio de Jury. Las tres condiciones que debe cumplir para la estabilidad son:

$$\begin{cases} |0.3432K + 0.0494| < 20 \\ 20 + 2.1718K - 7.474 + 0.3432K + 0.0494 > 0 \\ 20 - 2.1718K + 7.474 + 0.3432K + 0.0494 > 0 \end{cases} \quad (19.23)$$

Operando:

$$\begin{cases} -58.4189 < K < 58.1311 \\ K > -5.0001 \\ K < 15.0516 \end{cases} \quad (19.24)$$

Por lo que la condición de estabilidad es:

$$-5.0001 < K < 15.0516 \quad (19.25)$$

Se puede comprobar por el criterio de Routh-Hurwitz que el resultado es correcto. Aplicando a la ecuación característica (9.22) la transformación bilineal (9.17), queda:

$$(12.5754 + 2.515K)w^2 + (39.9012 - 0.6864K)w + 27.5234 - 1.8286K = 0 \quad (19.26)$$

En los sistemas de segundo orden, la columna principal del criterio de Routh-Hurwitz coincide con los coeficientes del polinomio de la ecuación característica, por tanto, las condición que se debe cumplir es que los tres coeficientes sean positivos:

$$\begin{cases} 12.5754 + 2.515K > 0 \\ 39.9012 - 0.6864K > 0 \\ 27.5234 - 1.8286K > 0 \end{cases} \quad (19.27)$$

Denuevo, el rango que valores de K que satisface las tres inecuaciones es:

$$-5.0001 < K < 15.0516 \quad (19.28)$$

Por tanto, como no podía ser de otra forma, los dos métodos obtienen el mismo resultado. En este caso concreto, el más rápido es el criterio de Jury, porque el cambio de variable de la transformación bilineal requiere muchas operaciones, por lo que pasar de la ecuación (19.22) a la ecuación (19.26) puede llevar mucho tiempo.

Capítulo 20

Respuesta transitoria y régimen permanente

En el presente capítulo se muestra cómo se puede calcular el valor en régimen permanente de un sistema y la relación entre la posición de los polos y ceros de una función de transferencia y el régimen transitorio. Todo el capítulo es una traslación directa de todo lo que se conoce ya en sistemas continuos.

20.1. Respuesta transitoria

En la Fig. 20.1 se muestra una representación cualitativa de las diferentes respuestas ante entrada impulso que se pueden obtener en función de la posición de los polos de una función de transferencia.

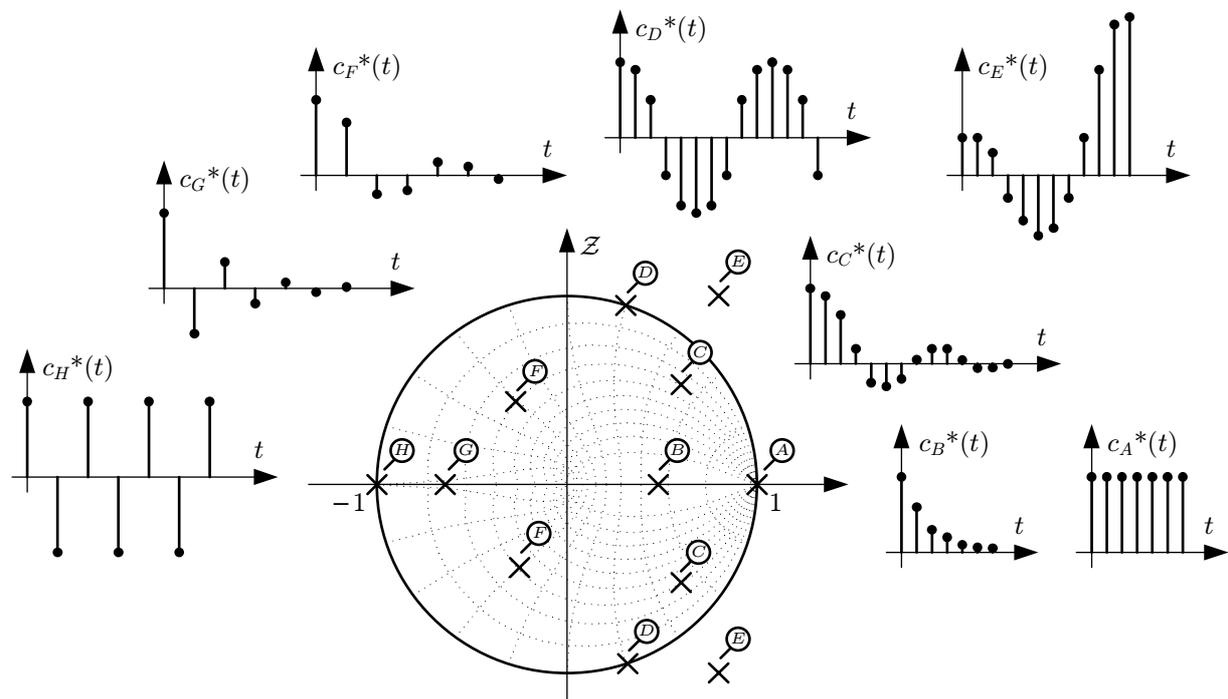


Figura 20.1: Respuesta la a entrada impulso de distintos polos

Ya se vio en el capítulo 18 cual era la relación numérica entre la posición de los polos en el plano Z y algunas especificaciones de régimen transitorio. Por ejemplo:

$$|z| = e^{-T\zeta\omega_n} \implies \ln |z| = -T\zeta\omega_n = -T\sigma \quad (20.1)$$

$$\angle z = T\omega_n\sqrt{1 - \zeta^2} \text{ rad} = T\omega_d \text{ rad} \quad (20.2)$$

Con las ecuaciones (20.1) y (20.2) se pueden calcular los valores de amortiguamiento, frecuencia natural, frecuencia natural amortiguada y atenuación que producen dos polos complejo-conjugados. Si

dichos polos son los dominantes de la función de transferencia, esos parámetros serán los que determinen la respuesta transitoria del sistema.

La influencia de nuevos polos y ceros es exactamente igual que en sistemas continuos, simplemente que ahora la referencia cambia. Un polo no es dominante cuando se encuentra cerca del origen de coordenadas (como ocurría en el caso continuo) sino cuando esté cerca de la unidad, es decir, el punto $z = 1$. También se puede utilizar la fórmula conocida para el tiempo de establecimiento del sistema (20.3), el tiempo de levantamiento, el tiempo de pico, etc.

$$t_s(\pm 2\%) \approx \frac{4}{\sigma} = \frac{4}{\zeta\omega_n} \quad (20.3)$$

NOTA: En el plano \mathcal{Z} , los polos simples de parte real negativa producen respuestas oscilatorias, con periodo siempre dos veces el periodo de muestreo. Este hecho puede llamar la atención, porque el ingeniero puede estar acostumbrado en sistemas continuos a asociar las respuestas no oscilatorias (exponenciales que convergen o no a un punto) a polos simples y respuestas oscilatorias (amortiguadas o no) a polos dobles.

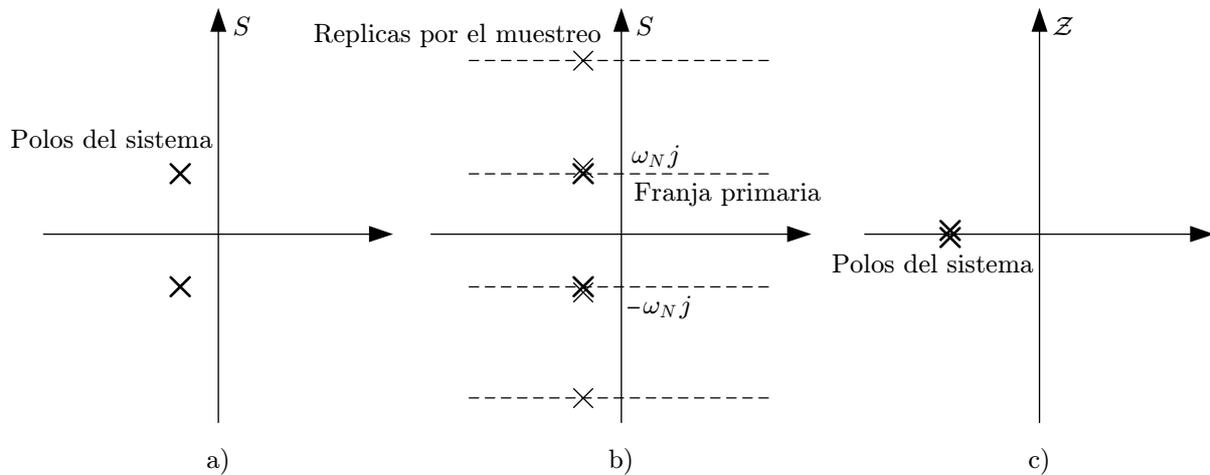


Figura 20.2: Los polos a) complejo-conjugados de la planta en S , con un determinado muestreo b), se convierten en c) polos reales dobles en el plano \mathcal{Z}

Los polos con sentido físico de la planta que caen sobre el eje \mathcal{Z} de parte real negativa necesariamente serán reales dobles. Esto se debe a que en continuo eran complejo-conjugados y la frontera de la franja primaria pasa justo por encima suyo, Fig. 20.2. Por tanto, la posición de los dos polos complejo-conjugados en S coincide en el mismo punto en el plano \mathcal{Z} . Este hecho es coherente con lo que se ha afirmado que el periodo de oscilación es dos veces el periodo de muestreo.

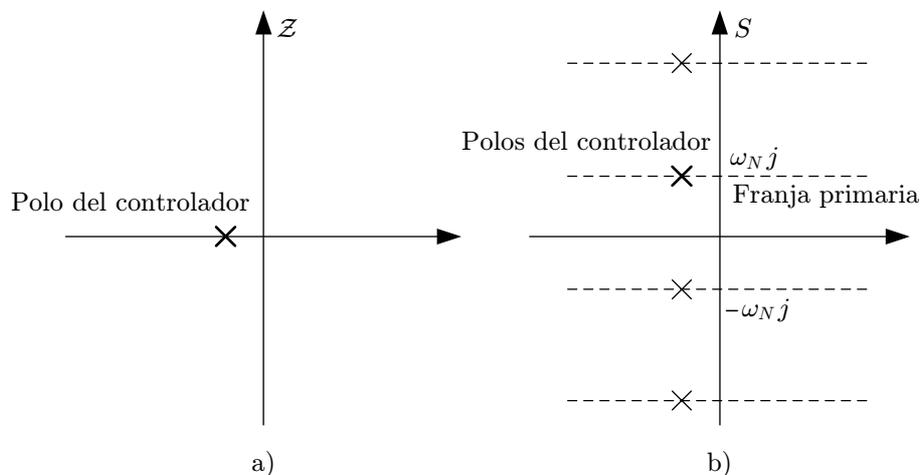


Figura 20.3: Un polo simple del controlador en \mathcal{Z} a) se convierte en polos complejo-conjugados en S

Sólo se pueden conseguir polos simples de parte real negativa, programándolos explícitamente en un microprocesador (a través de su ley por medio de ecuaciones en diferencias). Por tanto sólo pueden

pertenecer al controlador, no a la planta. En realidad, lo que está introduciendo en el dominio S son polos complejo-conjugados. Por tanto, aunque “aparentemente” es simple, en realidad es complejo-conjugado, Fig. 20.3. Por tanto, ya no puede extrañar que los polos simples de parte real negativa en \mathcal{Z} posean respuestas oscilatorias.

20.2. Régimen permanente

El régimen permanente de un sistema se calcula evaluando el límite cuando el tiempo tiende a infinito de la señal de salida. Ese límite se puede trasladar al dominio \mathcal{Z} con la propiedad correspondiente de la transformada \mathcal{Z} :

$$c_{ss} = \lim_{t \rightarrow \infty} c(t) = \lim_{n \rightarrow \infty} c(nT) = \lim_{z \rightarrow 1} [(1 - z^{-1})C(z)] \quad (20.4)$$

Y conviene recordar que sólo tiene sentido hablar de régimen permanente en los sistemas estables. Si un sistema no es estable puede existir un valor numérico para el límite definido en (20.4), pero no quiere decir que la respuesta converja hacia ese valor. Antes habrá que asegurar la estabilidad del sistema.

20.3. Error en régimen permanente

Quizá más interesante que el valor numérico del régimen permanente de un sistema, es definir cuánto vale el error respecto a la referencia en un sistema de control realimentado:

$$e_{ss} = \lim_{z \rightarrow 1} \{(1 - z^{-1})[R(z) - C(z)]\} \quad (20.5)$$

Como se vio a lo largo del capítulo 17, la respuesta $C(z)$ podrá ponerse o no en función de la referencia muestreada $R(z)$. Por este motivo tal vez no sea posible sacar factor común de la expresión (20.5) la referencia $R(z)$ como se hacía en los sistemas continuos usando la función de transferencia en lazo cerrado. Ya se ha visto que en el dominio \mathcal{Z} a veces no se puede encontrar la función de transferencia en lazo cerrado de un sistema.

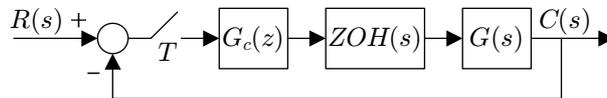


Figura 20.4: Sistema de control discreto

Como *caso particular*, se puede ver las expresiones que se obtienen con el sistema de control de la Fig. 20.4. En este caso particular —pero que se utiliza mucho— sí es posible determinar la función de transferencia en lazo cerrado y seguir operando en la ecuación (20.5).

$$e_{ss} = \lim_{z \rightarrow 1} \left\{ (1 - z^{-1}) \left[R(z) - \frac{G_c(z)\mathcal{Z}[ZOH(s)G(s)]}{1 + G_c(z)\mathcal{Z}[ZOH(s)G(s)]} R(z) \right] \right\} \quad (20.6)$$

$$e_{ss} = \lim_{z \rightarrow 1} \left\{ (1 - z^{-1}) \left[1 - \frac{G_c(z)\mathcal{Z}[ZOH(s)G(s)]}{1 + G_c(z)\mathcal{Z}[ZOH(s)G(s)]} \right] R(z) \right\} \quad (20.7)$$

$$e_{ss} = \lim_{z \rightarrow 1} \left\{ \frac{(1 - z^{-1})R(z)}{1 + G_c(z)\mathcal{Z}[ZOH(s)G(s)]} \right\} \quad (20.8)$$

Ahora se puede particularizar la expresión (20.8) para distintos tipos de entrada (escalón, rampa, parábola, etc). Así pues, para una entrada *referencia escalón de amplitud A*, el error es:

$$e_{ss} = \lim_{z \rightarrow 1} \left\{ \frac{(1 - z^{-1})}{1 + G_c(z)\mathcal{Z}[ZOH(s)G(s)]} \frac{Az}{z - 1} \right\} = \lim_{z \rightarrow 1} \left\{ \frac{A}{1 + G_c(z)\mathcal{Z}[ZOH(s)G(s)]} \right\} \quad (20.9)$$

$$e_{ss} = \frac{A}{1 + \lim_{z \rightarrow 1} \{G_c(z)\mathcal{Z}[ZOH(s)G(s)]\}} = \frac{A}{1 + K_p} \quad (20.10)$$

Donde se ve que se puede definir algo similar al coeficiente de error en posición K_p que se vio en sistemas continuos. También en este caso el error es proporcional a la amplitud del escalón de entrada y se podría definir de forma general como en porcentaje o en por unidad respecto a la entrada. Y en este caso el error será nulo o finito en función del número de polos en el punto $z = 1$ que posea el sistema.

Para una entrada *referencia rampa de pendiente A*, el error es:

$$e_{ss} = \lim_{z \rightarrow 1} \left\{ \frac{(1 - z^{-1})}{1 + G_c(z) \mathcal{Z}[ZOH(s)G(s)]} \frac{ATz}{(z - 1)^2} \right\} \quad (20.11)$$

$$e_{ss} = \lim_{z \rightarrow 1} \left\{ \frac{AT}{(z - 1) + (z - 1)G_c(z) \mathcal{Z}[ZOH(s)G(s)]} \right\} \quad (20.12)$$

$$e_{ss} = \frac{A}{\lim_{z \rightarrow 1} \left\{ \frac{z-1}{T} G_c(z) \mathcal{Z}[ZOH(s)G(s)] \right\}} = \frac{A}{K_v} \quad (20.13)$$

Donde de nuevo se puede definir un coeficiente similar al coeficiente de error en velocidad K_v , y de nuevo el error será nulo o finito o infinito en función del número de polos en el punto $z = 1$ que posea el sistema.

Análogamente para una entrada *referencia parábola de aceleración A*, el error es:

$$e_{ss} = \frac{A}{\lim_{z \rightarrow 1} \left\{ \frac{2(z-1)^2}{T^2 z(z+1)} G_c(z) \mathcal{Z}[ZOH(s)G(s)] \right\}} \quad (20.14)$$

$$e_{ss} = \frac{A}{\lim_{z \rightarrow 1} \left\{ \frac{(z-1)^2}{T^2} G_c(z) \mathcal{Z}[ZOH(s)G(s)] \right\}} = \frac{A}{K_a} \quad (20.15)$$

Donde se define el coeficiente de error en aceleración K_a . Como el valor numérico del límite no cambia si se introduce varias veces el factor z (que equivale a la unidad cuando se evalúa el límite), se pueden definir los coeficientes de error también como:

$$K_v = \lim_{z \rightarrow 1} \left\{ \frac{1 - z^{-1}}{T} G_c(z) \mathcal{Z}[ZOH(s)G(s)] \right\} \quad (20.16)$$

$$K_a = \lim_{z \rightarrow 1} \left\{ \left(\frac{1 - z^{-1}}{T} \right)^2 G_c(z) \mathcal{Z}[ZOH(s)G(s)] \right\} \quad (20.17)$$

Quizá estas expresiones sean de más fácil memorización ya que se parecen más a las definiciones que tenían en el dominio continuo cambiando la variable s por su aproximación de *backwards*.

20.4. Tipo de sistema

El tipo de sistema N se define como el número de polos en el punto $z = 1$ que posee la función de transferencia en lazo abierto del sistema de la Fig. 20.4.

Como ocurría en sistemas analógicos, cuanto mayor sea el tipo, el error del sistema será nulo ante mayor número de entradas distintas. También comparte con ellos el inconveniente de que cuantos más polos haya en $z = 1$ más difícil será hacer estable el sistema, porque es más fácil que las ramas del lugar de las raíces salgan del círculo unitario.

Capítulo 21

Lugar de las raíces

El lugar de las raíces es una herramienta gráfica que se demostró muy útil en el dominio continuo de Laplace. En este capítulo se describe cómo se puede utilizar en el plano complejo \mathcal{Z} .

21.1. Definición

El lugar de las raíces, como su propio nombre indica, es el lugar geométrico de los puntos del plano complejo \mathcal{Z} que son polos de un sistema controlado en función de un parámetro que suele ser una ganancia proporcional dentro del lazo de control.

Por tanto, es un método que sólo se puede utilizar para ajustar un único parámetro de diseño. Si existe más de una variable en el sistema de control, el uso del lugar de las raíces no es sencillo porque habría que aplicarlo de forma iterativa. Es decir, habría que fijar todos los parámetros de diseño menos uno, y ver si se llega a una solución satisfactoria. En caso contrario, se toman otros valores numéricos y se vuelve a comenzar.

El lugar de las raíces sólo da información de la posición de los polos en lazo cerrado del sistema compensado. Cuando el parámetro de variación es una ganancia proporcional dentro de un lazo de realimentación simple, los ceros no cambian de posición con ella. En cualquier caso para diseñar correctamente el comportamiento transitorio del sistema hay que tener en cuenta la posición de los polos y ceros dominantes del sistema.

Con estas premisas, se entiende que sea un método muy conveniente para ajustar controladores puramente proporcionales.

21.2. Punto de partida

El método del lugar de las raíces siempre parte de la ecuación característica del sistema, porque las raíces de la misma serán los polos en lazo cerrado del sistema. En sistemas muestreados, no importa que no se pueda a veces encontrar la función de transferencia del sistema porque siempre será posible determinar, al menos, la ecuación característica.

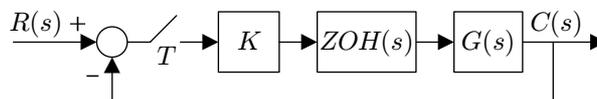


Figura 21.1: Sistema de control discreto con controlador proporcional

En el caso concreto de que el sistema de control posea simplemente una realimentación negativa unitaria y un controlador proporcional, como el que se muestra en la Fig. 21.1, la ecuación característica del mismo es:

$$1 + KZ[ZOH(s)G(s)] = 0, \quad (21.1)$$

en definitiva:

$$1 + KF(z) = 0. \quad (21.2)$$

Por tanto, a la función de transferencia $F(z)$ se le pueden aplicar las mismas reglas que se definieron en sistemas continuos para representar el lugar de las raíces.

Si en la parte muestreada existe una función de transferencia determinada (un filtro, una red de adelanto o retraso que ya tiene valores numéricos, etc.) entonces el sistema se podría describir como:

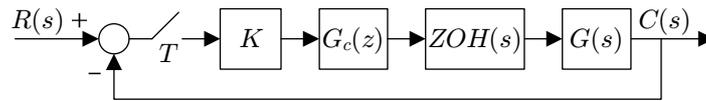


Figura 21.2: Sistema de control discreto

Y la ecuación característica:

$$1 + KG_c(z)\mathcal{Z}[ZOH(s)G(s)] = 0, \quad (21.3)$$

donde, de nuevo:

$$1 + KF(z) = 0. \quad (21.4)$$

Ahora bien, si el parámetro de variación no es la ganancia proporcional dentro del lazo (nótese que no importa si está en la parte muestreada o en la parte continua), la ecuación característica podría tomar una forma distinta de la presentada en la ecuación (21.4). Por ejemplo, sea el sistema:

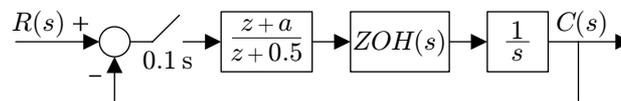


Figura 21.3: Ejemplo de sistema de control discreto

La ecuación característica es:

$$1 + \frac{z+a}{z+0.5}\mathcal{Z}\left[ZOH(s)\frac{1}{s}\right] = 0 \quad (21.5)$$

$$1 + \frac{z+a}{z+0.5}\mathcal{Z}\left[\frac{1-e^{-0.1s}}{s}\frac{1}{s}\right] = 0 \quad (21.6)$$

$$1 + \frac{z+a}{z+0.5}(1-z^{-1})\mathcal{Z}\left[\frac{1}{s^2}\right] = 0 \quad (21.7)$$

$$1 + \frac{z+a}{z+0.5}(1-z^{-1})\frac{0.1z}{(z-1)^2} = 0 \quad (21.8)$$

$$1 + \frac{0.1(z+a)}{(z+0.5)(z-1)} = 0 \quad (21.9)$$

En primer lugar se constata que el retenedor no modifica el orden del sistema. Existe un único polo en la parte muestreada y otro más en la parte continua, por lo que en la ecuación característica se han obtenido dos polos como era de esperar.

La dificultad en este caso consiste en que el parámetro que hace variar los polos del sistema —las raíces de la ecuación (21.5)— es a , y a no se encuentra multiplicando ninguna función de transferencia \mathcal{Z} . Sin embargo, es posible operar en la ecuación característica para dejarlo de esa forma:

$$(z+0.5)(z-1) + 0.1(z+a) = 0 \quad (21.10)$$

$$(z+0.5)(z-1) + 0.1z + 0.1a = 0 \quad (21.11)$$

$$1 + a\frac{0.1}{(z+0.5)(z-1) + 0.1z} = 0 \quad (21.12)$$

En la ecuación (21.12) sí se puede encontrar una $F(z)$ a la que aplicar el método gráfico del lugar de las raíces. El gráfico del lugar de las raíces será correcto a pesar de que la nueva $F(z)$ no tiene nada que ver con la función de transferencia en lazo abierto del sistema.

El mismo ejemplo un poco modificado, sirve para mostrar cómo el parámetro de variación puede ser incluso el periodo de muestreo del sistema:

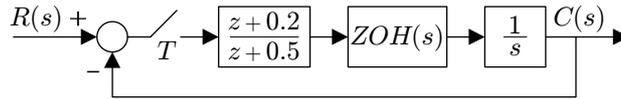


Figura 21.4: Ejemplo de sistema de control discreto

La ecuación característica es:

$$1 + \frac{z + 0.2}{z + 0.5} \mathcal{Z} \left[ZOH(s) \frac{1}{s} \right] = 0 \quad (21.13)$$

$$1 + \frac{z + 0.2}{z + 0.5} \mathcal{Z} \left[\frac{1 - e^{-Ts}}{s} \frac{1}{s} \right] = 0 \quad (21.14)$$

$$1 + \frac{z + 0.2}{z + 0.5} (1 - z^{-1}) \mathcal{Z} \left[\frac{1}{s^2} \right] = 0 \quad (21.15)$$

En este punto podría parecer que el parámetro de variación ha desaparecido, pero:

$$1 + \frac{z + 0.2}{z + 0.5} (1 - z^{-1}) \frac{Tz}{(z - 1)^2} = 0 \quad (21.16)$$

$$1 + \frac{T(z + 0.2)}{(z + 0.5)(z - 1)} = 0 \quad (21.17)$$

Y en este caso, identificar la $F(z)$ es inmediato. Pero en los dos ejemplos, en el sistema sólo existía un parámetro variable y los demás estaban dados.

21.3. Método gráfico

En este apartado se resumen los pasos estudiados el apartado 7.3 para dibujar el lugar de las raíces. Se hacen exactamente los mismo pasos. Sólo hay que cambiar la variable de Laplace s por la variable z :

1. Señalar en el plano \mathcal{Z} los n polos $p_1, p_2, \dots, p_j, \dots, p_n$ y los m ceros $z_1, z_2, \dots, z_i, \dots, z_m$ de la función de transferencia $F(z)$, que se considerará la “función de transferencia en lazo abierto”.

2. Número de asíntotas: $n - m$

3. Punto de corte de las asíntotas:

$$\sigma_a = \frac{\sum_{j=1}^n p_j - \sum_{i=1}^m z_i}{n - m} \quad (21.18)$$

4. Ángulos de las asíntotas:

$$\theta_a = \frac{180^\circ}{n - m} (2k + 1), \text{ con } k = 0, 1, \dots, n - m - 1 \quad (21.19)$$

5. Puntos del eje real \mathcal{Z} que pertenecen al lugar de las raíces: los que dejan a su derecha un número impar de polos y ceros.

6. Puntos de ruptura: son las raíces de la siguiente ecuación que pertenecen al lugar de las raíces:

$$\frac{dF(z)}{dz} = 0 \quad (21.20)$$

7. Cortes que con eje imaginario: aplicando el criterio de Routh-Hurwitz directamente a la ecuación característica (sin ninguna transformación bilineal previa). Ahora la ganancia crítica que se obtenga no será tal, porque ahora la estabilidad no depende del signo de la parte real de los polos, sino si están dentro o fuera del círculo unitario.

8. Ángulos de salida y de llegada: aplicar la condición del ángulo en un punto q del plano \mathcal{Z} muy cercano del polo o cero objeto de estudio:

$$\sum_{i=1}^m \widehat{z_i q} - \sum_{j=1}^n \widehat{p_j q} = \pm 180^\circ (2k + 1), \text{ con } k \in \mathbb{N} \quad (21.21)$$

La condición del módulo, aplicada en cualquier punto q del plano \mathcal{Z} , se definía como:

$$KK_{la} = \frac{\prod_{j=1}^n \overline{p_j q}}{\prod_{i=1}^m z_i q}, \quad (21.22)$$

donde no se dividía por nada si no existían ceros.

21.4. Diseño de compensadores de adelanto de fase

En general, con un controlador puramente proporcional no es posible establecer los polos del sistema compensado en cualquier lugar que se desee. Para ello se necesitan más parámetros de diseño (añadir nuevos polos y ceros al sistema) de forma que se “fuerce” al lugar de las raíces del sistema compensado a pasar por lo que se denominarán los “polos objetivo” de diseño.

En el dominio de Laplace se utilizó el lugar de las raíces para diseñar compensadores de adelanto de fase que hicieran precisamente eso. En el plano \mathcal{Z} se puede utilizar la misma técnica, pero con algunas advertencias.

En primer lugar hay que resaltar que el método de la bisectriz a veces falla en sistemas muestreados. La razón es que fácilmente puede quedar el nuevo polo del compensador fuera del círculo unitario, por lo que se puede estar creando una nueva rama que siempre queda fuera del mismo y por tanto el sistema sea siempre inestable. En los sistemas continuos daba igual que aparecieran nuevas ramas en la parte real negativa lejos del origen. Ahora lejos del origen siempre significa inestabilidad.

A continuación se resumen los pasos que se estudiaron en el dominio de Laplace:

1. Con las especificaciones de diseño se calculan los “polos objetivo” del sistema.
2. Se sitúan en el plano \mathcal{Z} los polos objetivo y los polos y ceros de $F(z)$, que son los equivalentes a los polos y ceros en lazo abierto del sistema.
3. Se aplica la condición del ángulo a uno de los polos objetivo (teniendo en cuenta el nuevo polo y cero del controlador) y se calcula el ángulo con que el polo objetivo “ve” al nuevo polo y cero del controlador.
4. Se elige la posición del polo y cero del controlador de forma que los “polos objetivo” queden dominantes sin que ningún polo del sistema en lazo cerrado quede fuera del círculo unitario.
5. Una vez situados el polo y cero del controlador, se aplica la condición del módulo en el polo objetivo para calcular la ganancia que hace que los “polos objetivo” sean precisamente los polos en lazo cerrado del sistema.

La dificultad del método radica en elegir correctamente la posición del polo y del cero del compensador. En el siguiente ejemplo se muestra cómo puede razonar en el diseño.

21.5. Ejemplo de diseño

Diseñar un controlador digital para que el sistema de la Fig. 21.5 posea una frecuencia natural de 3.5 rad/s y un amortiguamiento de 0.45.

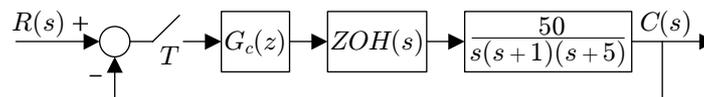


Figura 21.5: Ejemplo de sistema de control discreto

Se observa que la planta es de tercer orden, por lo que se deben encontrar tres polos en el dominio \mathcal{Z} . También se observa que el periodo de muestreo no está fijado en el enunciado, por lo que lo primero que se debe hacer es elegir un periodo de muestreo adecuado a las especificaciones que se desean obtener. Dependiendo de la elección que se haga, el problema sale muy distinto (al menos en los valores numéricos).

Las especificaciones de diseño imponen la siguiente frecuencia natural amortiguada en la respuesta transitoria del sistema compensado:

$$\omega_d = \omega_n \sqrt{1 - \zeta^2} \text{ rad} \approx 3.1 \text{ rad} \quad (21.23)$$

Usando la ecuación (18.20) para medir entre 8 y 10 muestras por ciclo, se obtiene:

$$0.25 \leq T \leq 0.20 \quad (21.24)$$

Con un periodo de muestreo de 0.2 segundos sería suficiente. En este ejemplo se tomará todavía más pequeño, T igual a 0.1 segundos, con lo que se medirán unas 20 muestras por ciclo. Con esta elección se puede encontrar la ecuación característica del sistema:

$$1 + K \frac{z+a}{z+b} \mathcal{Z} \left[ZOH(s) \frac{50}{s(s+1)(s+5)} \right] = 0, \quad (21.25)$$

donde ya se ha puesto la expresión del controlador de adelanto de fase. Operando:

$$1 + K \frac{z+a}{z+b} (1-z^{-1}) \mathcal{Z} \left[\frac{50}{s^2(s+1)(s+5)} \right] = 0. \quad (21.26)$$

Para calcular la transformada \mathcal{Z} , se divide la expresión racional en fracciones simples:

$$\frac{50}{s^2(s+1)(s+5)} = \frac{A}{s^2} + \frac{B}{s} + \frac{C}{s+1} + \frac{D}{s+5} = 0, \quad (21.27)$$

donde el numerador es:

$$50 = A(s+1)(s+5) + Bs(s+1)(s+5) + Cs^2(s+5) + Ds^2(s+1) \quad (21.28)$$

Se obtiene fácilmente que:

$$s = 0 \implies 50 = 5A \implies A = 10 \quad (21.29)$$

$$s = -1 \implies 50 = 4C \implies C = 12.5 \quad (21.30)$$

$$s = -5 \implies 50 = -100D \implies D = -0.5 \quad (21.31)$$

$$\text{término en } s \implies 0 = 6A + 5B \implies B = -12 \quad (21.32)$$

Por tanto, se sigue operando en la ecuación característica:

$$1 + K \frac{z+a}{z+b} (1-z^{-1}) \mathcal{Z} \left[\frac{10}{s^2} - \frac{12}{s} + \frac{12.5}{s+1} - \frac{0.5}{s+5} \right] = 0 \quad (21.33)$$

$$1 + K \frac{z+a}{z+b} (1-z^{-1}) \left[\frac{z}{(z-1)^2} - \frac{12z}{z-1} + \frac{12.5z}{z-e^{-0.1}} - \frac{0.5z}{z-e^{-0.5}} \right] = 0 \quad (21.34)$$

$$1 + K \frac{z+a}{z+b} (z-1) \left[\frac{1}{(z-1)^2} - \frac{12}{z-1} + \frac{12.5}{z-e^{-0.1}} - \frac{0.5}{z-e^{-0.5}} \right] = 0 \quad (21.35)$$

Sumando las fracciones queda:

$$1 + K \frac{z+a}{z+b} (z-1) \frac{0.0068z^2 + 0.0264z + 0.0043}{(z-1)^2(z-0.9048)(z-0.6065)} \quad (21.36)$$

$$1 + 0.0068K \frac{(z+a)(z+3.7119)(z+0.1703)}{(z+b)(z-1)(z-0.9048)(z-0.6065)} \quad (21.37)$$

Viendo la ecuación (21.37) se comprueba, como era de esperar, que la planta añade tres polos porque era de tercer orden. Lo que no era de esperar, pero ocurre con frecuencia al pasar al dominio \mathcal{Z} , es que añade también dos ceros, cuando en continuo no tenía ninguno. Este hecho no debe sorprender al ingeniero a partir de ahora. El polo en el origen se ha transformado, también como era de esperar, en un polo en $z = 1$. Los tres polos de la planta tienen parte real positiva en el dominio \mathcal{Z} pero esto no significa que se “hayan hecho inestables” ya que los tres están dentro del círculo unitario. Todas estas consideraciones ayudan al ingeniero a localizar posibles fallos durante el desarrollo numérico.

Se calculan ahora los “polos objetivo” para la posición de los polos en lazo cerrado del sistema compensado. Usando las especificaciones de diseño se obtiene:

$$q = e^{Ts} = e^{-T\zeta\omega_n \pm T\omega_n \sqrt{1-\zeta^2}j} = e^{-0.1575 \pm 0.3125j} = 0.8128 \pm 0.2626j = 0.8542 \angle \pm 17.9083^\circ \quad (21.38)$$

Ya se pueden dibujar en el plano complejo \mathcal{Z} los polos y ceros “en lazo abierto” de la planta y los “polos objetivo”, y discutir cuál es la mejor posición para el polo y el cero del controlador. En la Fig 21.6 aparecen representados todos ellos.

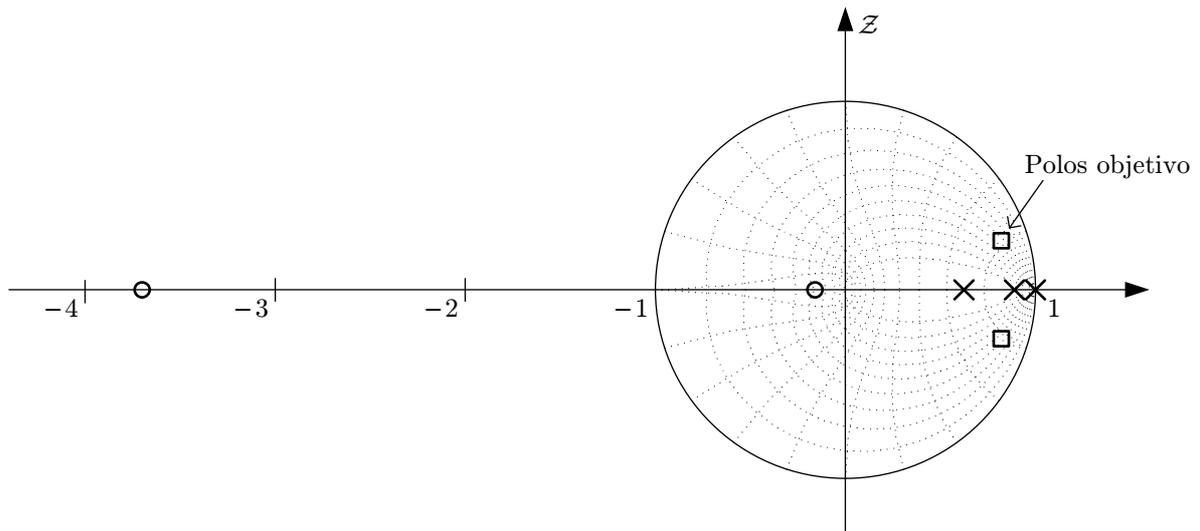


Figura 21.6: Polos y ceros “en lazo abierto” de la planta y “polos objetivo”

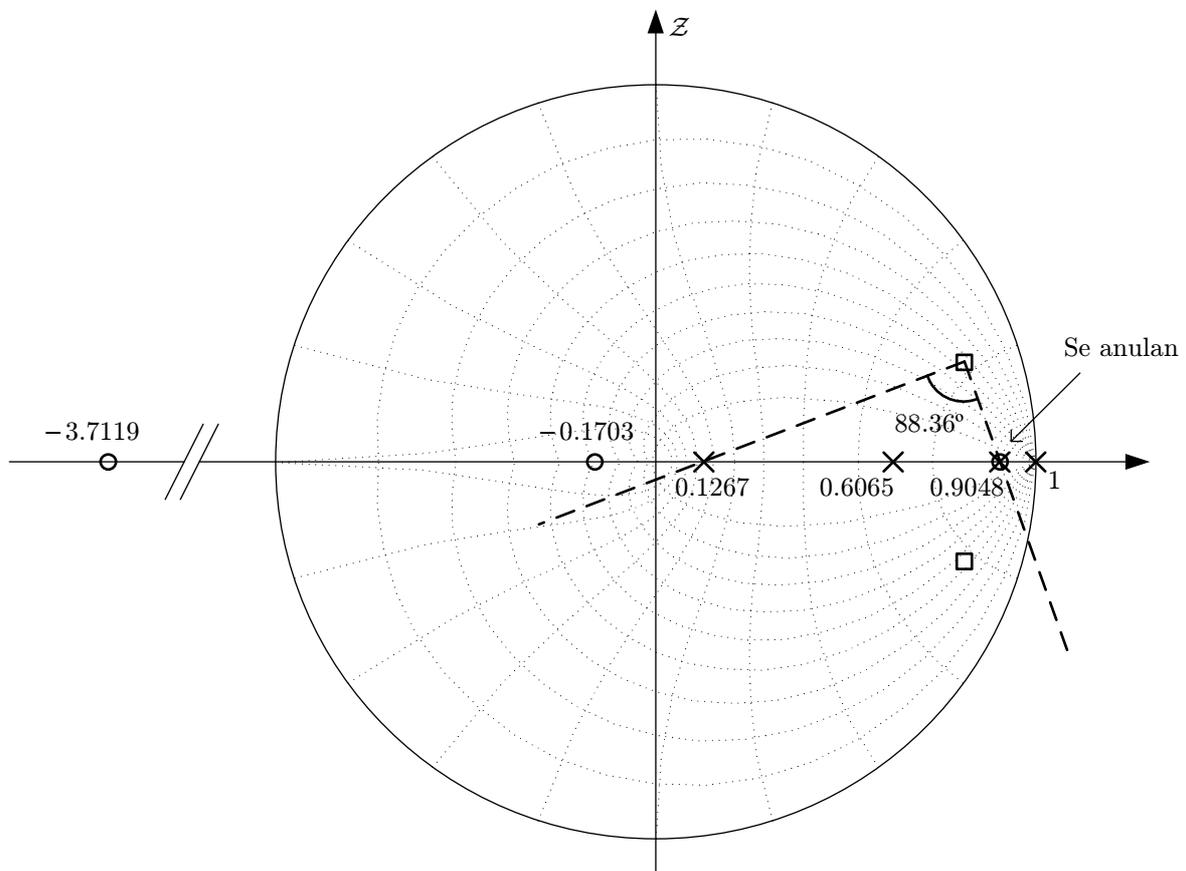


Figura 21.7: Compensador que anula el segundo polo de la planta

Para decidir la posición del polo y cero del controlador, lo primero que hay que hacer es calcular el ángulo con que un “polo objetivo” ve al controlador. Para ello, a su vez, hay que aplicar la condición del ángulo en uno de los “polos objetivo”:

$$\sum_{i=1}^m \widehat{z}_i q - \sum_{j=1}^n \widehat{p}_j q = -180^\circ \quad (21.39)$$

$$3.3215^\circ + 14.9553^\circ - 125.4839^\circ - 109.3075^\circ - 51.8466^\circ + \phi_c = -180 \quad (21.40)$$

$$\phi_c = 88.3612^\circ \quad (21.41)$$

oscilatoria no amortiguada. Es posible que dicha oscilación no se aprecie en la salida del sistema compensado, porque la planta actúe a modo de filtro, pero en cualquier caso es mejor evitarlo. Por tanto, el polo de compensador cuanto más cerca del origen mejor.

Con la expresión (21.44) es posible que se dé por terminado el problema. Sin embargo, a veces se pedirá que se dibuje el lugar de las raíces del sistema compensado. O que se exprese el controlador en forma de ecuaciones en diferencias. Esto último es casi inmediato:

$$G_c(z) = \frac{U(z)}{E(z)} = 2.52 \frac{z - 0.9048}{z - 0.1267} \quad (21.45)$$

$$\frac{U(z)}{E(z)} = \frac{2.52 - 2.2801z^{-1}}{1 - 0.1267z^{-1}} \quad (21.46)$$

$$u(nT) = 0.1267u(nT - T) + 2.52e(nT) - 2.2801e(nT - T) \quad (21.47)$$

También se puede pedir el cálculo numérico de la evolución temporal de la salida del sistema compensado ante una determinada entrada. Para ello hay que hallar la expresión de la salida en función de la entrada. En el ejemplo es:

$$C(z) = \frac{G_c(z)\mathcal{Z}[ZOH(s)G(s)]}{1 + G_c(z)\mathcal{Z}[ZOH(s)G(s)]}R(z) \quad (21.48)$$

$$C(z) = \frac{0.001713 \frac{(z+3.7119)(z+0.1703)}{(z+0.1267)(z-1)(z-0.6065)}}{1 + 0.001713 \frac{(z+3.7119)(z+0.1703)}{(z+0.1267)(z-1)(z-0.6065)}}R(z) \quad (21.49)$$

$$C(z) = \frac{0.001713(z + 3.7119)(z + 0.1703)}{(z + 0.1267)(z - 1)(z - 0.6065) + 0.001713(z + 3.7119)(z + 0.1703)}R(z) \quad (21.50)$$

$$C(z) = \frac{0.001713z^2 + 0.06652z + 0.01083}{z^3 - 1.7160z^2 + 0.8765z - 0.06597}R(z) \quad (21.51)$$

Y se puede usar cualquiera de los tres métodos para calcular la transformada inversa de \mathcal{Z} . Con entrada escalón unidad lo que se obtiene es:

$$\begin{array}{lll} c(0) = 0 & c(1.0) = 1.19087 & c(2.0) = 0.93899 \\ c(0.1) = 0.01713 & c(1.1) = 1.19288 & c(2.1) = 0.93691 \\ c(0.2) = 0.11304 & c(1.2) = 1.17297 & c(2.2) = 0.93947 \\ c(0.3) = 0.27343 & c(1.3) = 1.13900 & c(2.3) = \dots \\ c(0.4) = 0.46571 & c(1.4) = 1.09819 & \\ c(0.5) = 0.66138 & c(1.5) = 1.05652 & \\ c(0.6) = 0.83915 & c(1.6) = 1.01845 & \\ c(0.7) = 0.98481 & c(1.7) = 0.98678 & \\ c(0.8) = 1.09177 & c(1.8) = 0.96296 & \\ c(0.9) = 1.15927 & c(1.9) = 0.94725 & \end{array} \quad (21.52)$$

La salida presenta un “máximo local” para $t = 1.1$ segundos, donde se puede identificar el primer sobreimpulso como de 19.28 % aproximadamente. Presenta un “mínimo local” cerca del doble de tiempo, $t = 2.1$ segundos, que sería aproximadamente el periodo de oscilación amortiguada. Como se comentó a la hora de la elección del periodo de muestreo, se han conseguido poco más de 20 muestras por ciclo de oscilación. También se puede deducir que el régimen permanente será la unidad, por tanto error nulo, coherente con el hecho de que el sistema sea de tipo $N = 1$ (un polo en $z = 1$).

En la Fig. 21.9 se muestra la respuesta que se obtiene simulando en Simulink® el sistema con el compensador elegido. A pesar de pequeños errores numéricos que se pueden arrastrar en (21.52), se ha obtenido el mismo resultado.

21.6. Ejercicios propuestos

- **Ejercicio 1:** Diseñar un controlador digital para que el sistema de la Fig. 21.10 posea un tiempo de establecimiento de 2 segundos y un amortiguamiento de 0.5.

Suponer que el periodo de muestreo T es 0.2 segundos. Calcular la respuesta ante entrada escalón unidad. Calcular el coeficiente de error de velocidad.

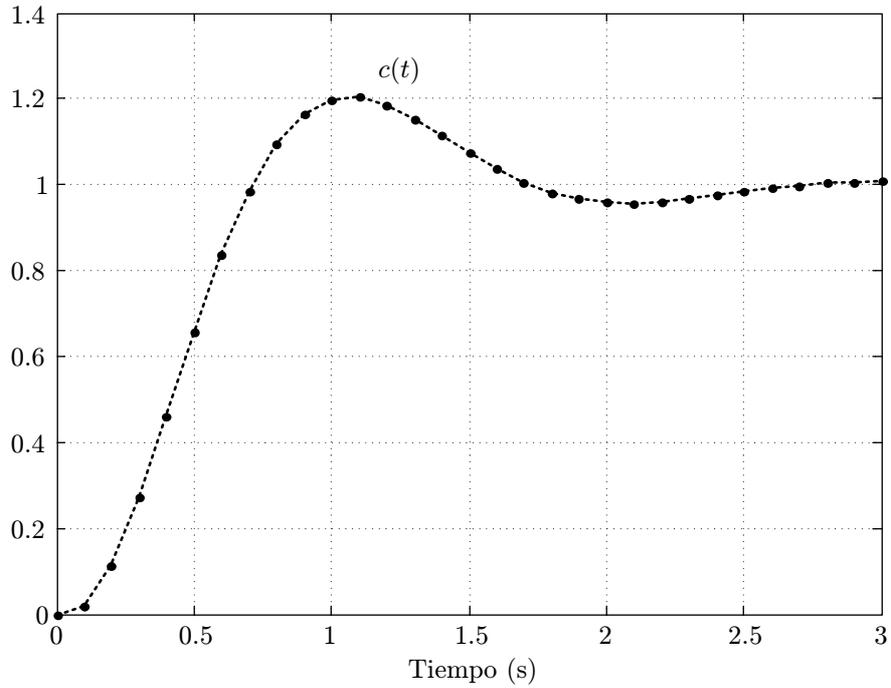


Figura 21.9: Salida del sistema compensado

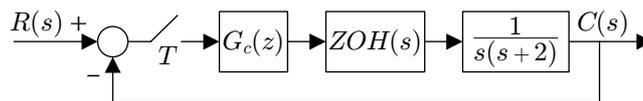


Figura 21.10: Primer ejercicio propuesto

- **Ejercicio 2:** Determinar la función de transferencia del sistema de la Fig. 21.11 cuando el controlador es igual a la unidad y el periodo de muestreo T es igual a 0.5 segundos. Diseñar un controlador digital que consiga una frecuencia natural de 1.5 rad/s y un amortiguamiento de 0.707.

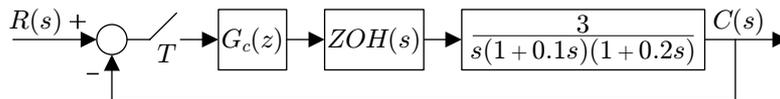


Figura 21.11: Segundo ejercicio propuesto

- **Ejercicio 3:** Diseñar un compensador digital que posea la siguiente forma:

$$G_c(z) = K \frac{z - a}{z} \quad (21.53)$$

De forma que la frecuencia natural del sistema de la Fig. 21.12 sea 0.3 rad/s y el amortiguamiento de 0.5.

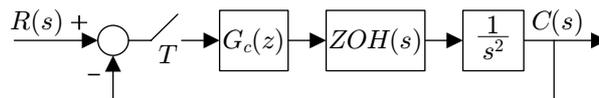


Figura 21.12: Tercer ejercicio propuesto

Capítulo 22

Métodos de digitalización

Este capítulo describe una forma “nueva” de diseñar controladores digitales. Se trata de realizar el diseño en el dominio continuo de Laplace (con cualquiera de las herramientas que se vieron en la primera parte de este libro) y posteriormente “traducirlo” o “pasarlo” al dominio discreto. A este proceso de “traducción” es lo que se conoce como digitalización de controladores analógicos.

22.1. Generalidades de los métodos de digitalización

Los métodos de diseño en el dominio analógico-continuo se presentaron en la primera parte de este libro. Ahora sólo se explicarán los distintos métodos de digitalización y se tomará como punto de partida el controlador en la variable s .

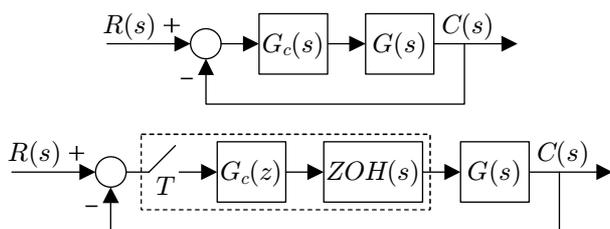


Figura 22.1: Digitalización o búsqueda del equivalente digital

Gráficamente, Fig. 22.1, lo que se pretende hacer es encontrar el controlador $G_c(z)$ equivalente a $G_c(s)$, es decir, que cumpla aproximadamente las mismas especificaciones de régimen transitorio y permanente. Como se puede apreciar, el periodo de muestreo no se tiene en cuenta a la hora de calcular el controlador analógico. Por tanto, la primera cosa que hay que decidir en el proceso de digitalización es un periodo de muestreo adecuado, que regirá el muestreo y el retenedor, y que también establecerá la relación que existirá entre las variables s y z .

La elección del periodo de muestreo se puede realizar conforme a lo explicado en el apartado 18.4, es decir, eligiendo un número de muestras adecuado por ciclo de oscilación de la respuesta transitoria (evidentemente, de la respuesta transitoria del sistema controlado).

22.2. Integración numérica

El primer método de digitalización está basado en la búsqueda de una aproximación numérica para la operación integral (22.1), o el bloque integrador Fig. 22.2.

$$y(t) = \int_0^t u(\tau) d\tau \tag{22.1}$$

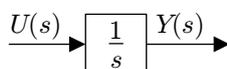


Figura 22.2: Operación integral

Si se considera la integral hasta un instante múltiplo del periodo de muestreo:

$$y(nT) = \int_0^{nT} u(\tau) d\tau = \int_0^{nT-T} u(\tau) d\tau + \int_{nT-T}^{nT} u(\tau) d\tau = y(nT-T) + \int_{nT-T}^{nT} u(\tau) d\tau \quad (22.2)$$

El valor de la integral entre dos periodos de muestreo se puede aproximar de diversas formas. En los siguientes apartados se explican las más habituales.

22.2.1. Método trapezoidal o de Tustin

El método de Tustin define la integral entre dos periodos de muestreo como el trapecio que forman el valor actual de la función y el anterior, Fig. 22.3.

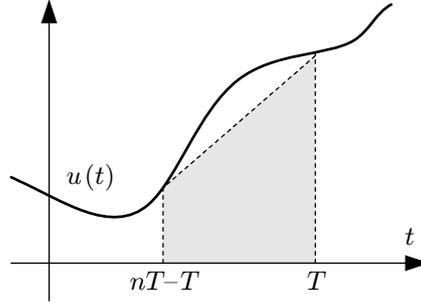


Figura 22.3: Aproximación de Tustin

Matemáticamente esto se puede escribir como:

$$\int_{nT-T}^{nT} u(\tau) d\tau = T \frac{u(nT) + u(nT-T)}{2} \quad (22.3)$$

De esta forma, la función integral (22.2) se puede completar como:

$$y(nT) = y(nT-T) + T \frac{u(nT) + u(nT-T)}{2}, \quad (22.4)$$

y por tanto:

$$Y(z)(1 - z^{-1}) = U(z) \frac{T}{2}(1 + z^{-1}), \quad (22.5)$$

$$\frac{Y(z)}{U(z)} = \frac{T}{2} \frac{1 + z^{-1}}{1 - z^{-1}} = \frac{T}{2} \frac{z + 1}{z - 1}. \quad (22.6)$$

En el dominio de Laplace, la operación integral es:

$$\frac{Y(s)}{U(s)} = \frac{1}{s} \quad (22.7)$$

Por tanto, se puede usar la ecuación (22.8) para definir un cambio de variable entre s y z (que recibe el nombre de aproximación trapezoidal, o bilineal, o de Tustin) y transformar una función de transferencia continua en discreta.

$$\frac{1}{s} \approx \frac{T}{2} \frac{z + 1}{z - 1} \quad (22.8)$$

22.2.2. Método de Euler implícito

El método de Euler implícito define la integral entre dos periodos de muestreo como el rectángulo de altura igual al valor actual de la función, Fig. 22.4.

Matemáticamente esto se puede escribir como:

$$\int_{nT-T}^{nT} u(\tau) d\tau = T u(nT) \quad (22.9)$$

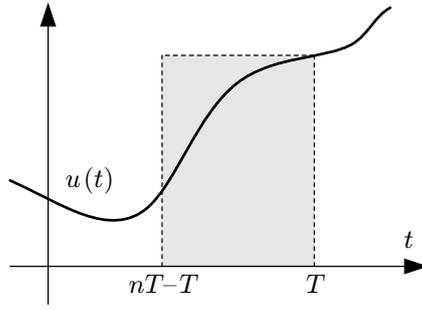


Figura 22.4: Aproximación de Euler implícito

De esta forma, la función integral (22.2) se puede completar como:

$$y(nT) = y(nT - T) + Tu(nT), \quad (22.10)$$

y por tanto:

$$Y(z)(1 - z^{-1}) = TU(z), \quad (22.11)$$

$$\frac{Y(z)}{U(z)} = \frac{T}{1 - z^{-1}} = \frac{Tz}{z - 1}. \quad (22.12)$$

Esto define un nuevo cambio de variable (conocido como método de integración de Euler implícito) para realizar la digitalización de una función de transferencia:

$$\frac{1}{s} \approx \frac{Tz}{z - 1} \quad (22.13)$$

22.2.3. Método de Euler explícito

El método de Euler explícito define la integral entre dos periodos de muestro como el rectángulo de altura igual al valor anterior de la función, Fig. 22.5.

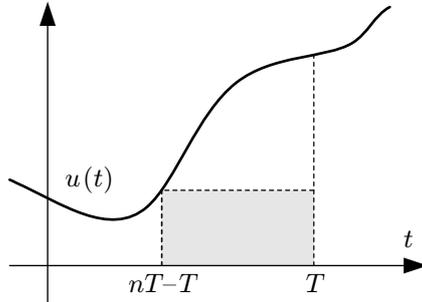


Figura 22.5: Aproximación de Euler explícito

Matemáticamente esto se puede escribir como:

$$\int_{nT-T}^{nT} u(\tau) d\tau = Tu(nT - T) \quad (22.14)$$

De esta forma, la función integral (22.2) se puede completar como:

$$y(nT) = y(nT - T) + Tu(nT - T), \quad (22.15)$$

y por tanto:

$$Y(z)(1 - z^{-1}) = TU(z)z^{-1}, \quad (22.16)$$

$$\frac{Y(z)}{U(z)} = \frac{Tz^{-1}}{1 - z^{-1}} = \frac{T}{z - 1}. \quad (22.17)$$

Esto define un nuevo cambio de variable (conocido como método de integración de Euler explícito) para realizar la digitalización de una función de transferencia:

$$\frac{1}{s} \approx \frac{T}{z - 1} \quad (22.18)$$

22.2.4. Otros métodos numéricos de integración

En la literatura científica existen muchos métodos de integración numérica —como son Bogacki-Shampine, Runge-Kutta, Dormand-Prince, etc.— y se conocen como *solvers* de paso fijo. Por citar otro método en forma de función de transferencia \mathcal{Z} , la regla de integración de Simpson se define como:

$$\frac{1}{s} \approx \frac{T}{3} \frac{z^2 + 4z + 1}{z^2 - 1} \quad (22.19)$$

Los beneficios que se obtienen son relativamente pequeños comparados con la complicación en el cálculo manual que supone añadir nuevos polos y ceros. Sin embargo, en simulaciones con ordenador es usual tomar el de Runge-Kutta. Casi todos los métodos citados están disponibles en los parámetros de simulación de Simulink®.

22.2.5. Ejemplo de digitalización usando integración numérica

En este apartado se utilizará alguno de los métodos descritos anteriormente en un ejemplo habitual de integración numérica. A priori no es posible decir qué tipo de integración es mejor de las anteriores. Si se comparan las tres expresiones (22.8), (22.13) y (22.18), se ve que tienen en común el polo en $z = 1$, hecho completamente lógico ya que aproximan un polo en el origen del plano S , cuyo equivalente en el plano \mathcal{Z} es la unidad real pura. En cambio, en el número y posición de ceros no coinciden.

El ejemplo que servirá de comparación es determinar la posición de un objeto de masa m y viscosidad b conocida la fuerza que se le aplica. Este ejemplo, aunque sencillo, es muy interesante ya que, a veces, medir la posición de objetos es técnicamente imposible y hay que recurrir a este tipo de estrategias (medir la fuerza y estimar la posición). Este ejemplo, pero con ecuaciones más complicadas, también es muy habitual en simulación de mecanismos. Se conocen las ecuaciones que gobiernan un mecanismo u objeto “virtual” y se desea calcular o simplemente visualizar su movimiento ante determinadas entradas.

$$\frac{X(s)}{F(s)} = \frac{1}{ms^2 + bs} \quad (22.20)$$

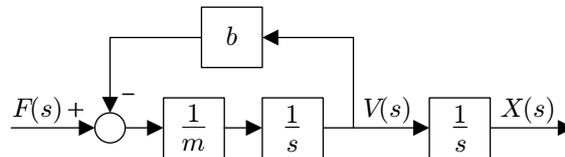


Figura 22.6: Aproximación de Euler explícito

En la ecuación (22.20) y Fig. 22.6 se muestra el proceso matemático del ejemplo. En realidad la ecuación del movimiento de un objeto encierra dos integraciones. Una pasa de la fuerza a la velocidad y la otra de la velocidad a la posición.

Se escribe la primera integración como:

$$[F(s) - bV(s)] \frac{1}{m} \frac{1}{s} = V(s) \quad (22.21)$$

Usando la aproximación de Euler implícito se obtiene:

$$[F(z) - bV(z)] \frac{1}{m} \frac{Tz}{z-1} = V(z) \quad (22.22)$$

$$\frac{V(z)}{F(z)} = \frac{Tz}{m(z-1) + bTz} \quad (22.23)$$

Expresado por medio de ecuaciones en diferencias:

$$v(nT) = \frac{T}{m + bT} f(nT) + \frac{m}{m + bT} v(nT - T) \quad (22.24)$$

La integración de velocidad a posición y usando la misma aproximación es:

$$V(s) \frac{1}{s} X(s) \quad (22.25)$$

$$V(z) \frac{Tz}{z-1} X(z) \quad (22.26)$$

Por tanto la función de transferencia discreta final es:

$$\frac{X(z)}{F(z)} = \frac{T^2 z^2}{m(z-1)^2 + bTz(z-1)} \quad (22.27)$$

De donde se pueden deducir las ecuaciones en diferencias que calculan directamente la posición a partir de la fuerza y posiciones anteriores:

$$x(nT) = \frac{T^2}{m+bT}f(nT) + \frac{2m+Tb}{m+bT}x(nT-T) - \frac{m}{m+bT}x(nT-2T) \quad (22.28)$$

Es importante hacer notar que la ecuación (22.27) se puede obtener directamente de la continua (22.20) sustituyendo la s por el cambio de variable que define Euler implícito, sin necesidad de identificar las dos integraciones que esconde la función de transferencia.

También es posible usar distintas aproximaciones para las diferentes integraciones. Por ejemplo, si se utiliza primero Euler implícito (para pasar de fuerzas a velocidades) y después Tustin (para pasar de velocidades a posiciones), se obtiene:

$$\frac{X(z)}{F(z)} = \frac{T^2 z(z+1)}{2m(z-1)^2 + 2bTz(z-1)} \quad (22.29)$$

Y también es importante resaltar que usar aproximaciones distintas para las diferentes integrales es distinto que sustituir cada una de las variables s por distintas aproximaciones.

Como ejercicio se pide representar la respuesta de las funciones de transferencia (22.29) y (22.27) ante una entrada fuerza escalón unidad (es decir 1 N), la masa es $m = 5$ kg y la viscosidad b es 1 Ns/m. La diferencia de las respuestas será relativamente pequeñas (del orden de los μm) por lo que cualquiera de las dos digitalizaciones es buena. En general, muchos autores recomiendan usar Euler implícito para integrar porque es más estable de cara a las simulaciones.

22.3. Derivación numérica

La operación contraria a la integración es la derivación. Como se sabe por lo estudiado en el dominio continuo, hay que intentar evitar la operación derivada. Sin embargo, en algunos casos no quedará más remedio que emplearla (Fig. 22.7). Cabe preguntarse si el análisis realizado en el apartado anterior es válido para una supuesta derivación numérica. Al fin y al cabo, invirtiendo las ecuaciones (22.8), (22.13) y (22.18), se obtendrían unas aproximaciones para la variable s que se pueden interpretar como aproximaciones de la función de transferencia que define la operación derivada.

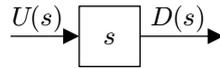


Figura 22.7: Operación derivada

Desarrollando las inversas de las ecuaciones (22.8), (22.13) y (22.18) en términos de ecuaciones en diferencias donde la variable $d(t)$ es la derivada de $u(t)$, queda para el caso del método de Tustin:

$$d(nT) = 2 \frac{u(nT) - u(nT-T)}{T} - d(nT-T) \quad (22.30)$$

Para Euler implícito:

$$d(nt) = \frac{u(nT) - u(nT-T)}{T} \quad (22.31)$$

Para Euler explícito:

$$d(nt) = \frac{u(nT+T) - u(nT)}{T} \quad (22.32)$$

22.3.1. Método de *backwards*

Observando las ecuaciones (22.30), (22.31) y (22.32) se puede concluir que:

1. El método de Euler explícito (22.32) *no se puede implementar* físicamente para la operación derivada ya que requiere el conocimiento del valor de la función en el siguiente periodo de muestreo, cosa que es imposible en cada instante.

2. El método de Euler implícito (22.31) pensado como método de derivación es al que uno puede llegar usando la “lógica”, es decir, definiendo la derivada como el último cambio de la función dividido entre el tiempo que ha transcurrido. Por ejemplo si la función de entrada es posición y la salida es velocidad, esta operación es el último incremento de posición dividido por el incremento de tiempo. El método de Euler implícito cuando se usa para aproximar la derivada recibe el nombre de método *backwards* o de “paso atrás”.

$$s \approx \frac{1 - z^{-1}}{T} \quad (22.33)$$

3. El método de Tustin se puede implementar para calcular la derivada, pero *no se recomienda su uso*. La razón es que introduce oscilaciones no deseables en la señal derivada. A modo de ejemplo, se puede comparar la derivada que calculan los métodos de Tustin y de *backwards* ante una entrada rampa de pendiente unidad en la Fig. 22.8.

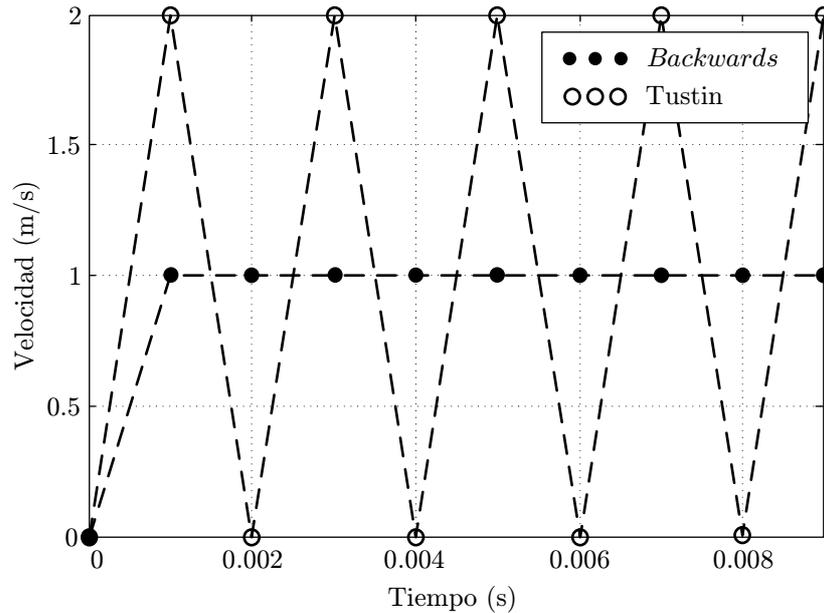


Figura 22.8: Derivada de una entrada rampa de pendiente unidad

La aproximación de Tustin introduce oscilaciones no amortiguadas en torno al verdadero valor de la velocidad. Por esta razón no se recomienda su uso en el caso de la operación derivación.

22.3.2. Otros métodos de derivación

Un método de derivación análogo al de *backwards*, pero que además introduce un cierto filtrado de la señal de entrada, es tomar la diferencia no respecto al periodo de muestreo interior, sino n periodos de muestreo antes:

$$s \approx \frac{1 - z^{-n}}{nT} \quad (22.34)$$

Evidentemente, esta forma de diferenciar introduce retraso en el sistema (como todos los filtros), pero su comportamiento puede ser mejor que aplicar primero *backwards* y luego introducir un filtro ordinario de primer orden.

Se ha resaltado que el método de Euler explícito no se puede usar *online* (en tiempo de ejecución) porque no se puede conocer valores futuros de la variable. Sin embargo, ese método sí que se puede usar *offline*, es decir, después de grabar todos los datos de una variable a lo largo del tiempo. Este caso se da, por ejemplo, cuando se quiere determinar *offline* después de ciertos experimentos algunos parámetros físicos de la planta. Para estos casos también se ha propuesto otro método de estimación de la derivada que reduce el retraso de la estimación de la derivada. Se trata de la diferencia “centrada”:

$$d(nt) = \frac{u(nT + T) - u(nT - T)}{2T}, \quad (22.35)$$

es decir:

$$s \approx \frac{z - z^{-1}}{2T} \quad (22.36)$$

22.3.3. Ejemplos de digitalización de PID

Las consideraciones de este apartado se aplican sólo a operación derivada. En el caso de que se quiera digitalizar una función de transferencia estrictamente propia, es decir, que posee más polos que ceros, no existen las dificultades antes mencionadas y se puedan cualquiera de las aproximaciones conocidas. La diferencia que se apreciará en la salida, al usar una u otra, será relativamente pequeña.

En el caso se que se pretenda digitalizar la expresión de un controlador PI, PD o PID analógico, sí se pueden “distinguir” las operaciones integral y derivada. Por ejemplo se puede usar Tustin para la integral y *backwards* para la derivada:

$$\text{PID}(s) = \frac{U(s)}{E(s)} = K_P + K_I \frac{1}{s} + K_D s \quad (22.37)$$

$$\text{PID}(z) = \frac{U(z)}{E(z)} = K_P + K_I \frac{Tz + 1}{2z - 1} + K_D \frac{z - 1}{Tz} \quad (22.38)$$

$$u_n = u_{n-1} + \left(K_P + \frac{K_I T}{2} + \frac{K_D}{T} \right) e_n - \left(K_P - \frac{K_I T}{2} + \frac{2K_D}{T} \right) e_{n-1} + \frac{K_D}{T} e_{n-2} \quad (22.39)$$

Donde se ha preferido una notación de subíndices que significa: $u_n = u(nT)$, $u_{n-1} = u(nT - T)$, y así sucesivamente. La actuación del PID (22.39) es igual a la actuación en el instante anterior más una combinación lineal en la que están involucrados el error actual, el error en el instante anterior y el error hace dos periodos de muestreo.

Pero se pueden realizar otras combinaciones. Por ejemplo, usando la misma aproximación las dos operaciones (Euler implícito para la integral y *backwards* para la derivada), quedaría:

$$\text{PID}(z) = \frac{U(z)}{E(z)} = K_P + K_I \frac{Tz}{z - 1} + K_D \frac{z - 1}{Tz} \quad (22.40)$$

$$u_n = u_{n-1} + \left(K_P + K_I T + \frac{K_D}{T} \right) e_n - \left(K_P + \frac{2K_D}{T} \right) e_{n-1} + \frac{K_D}{T} e_{n-2} \quad (22.41)$$

Aunque intervienen las mismas variables que antes, los coeficientes han cambiado.

Para obtener la forma del controlador PD digital, se parte de su expresión analógica:

$$\text{PD}(s) = \frac{U(s)}{E(s)} = K_P + K_D s, \quad (22.42)$$

y se aplica la aproximación *backwards* de la derivada:

$$\text{PD}(z) = \frac{U(z)}{E(z)} = K_P + K_D \frac{z - 1}{Tz} \quad (22.43)$$

$$\text{PD}(z) = \frac{K_P T + K_D}{T} \frac{z - \frac{K_D}{K_P T + K_D}}{z} \quad (22.44)$$

$$\text{PD}(z) = K \frac{z - a}{z} \quad (22.45)$$

Esta expresión justifica el uso del polo en el origen del plano \mathcal{Z} y cero de parte real positiva que se usó en el apartado 21.5. Para el caso de controlador PI existen más posibilidades, porque se conocen distintas aproximaciones para la integral, si bien lo común a todas ellas será que tendrán el polo en $z = 1$.

22.3.4. Ejemplos de digitalización de filtros

Es interesante también ver qué tipo de expresiones se obtienen en \mathcal{Z} para las funciones de transferencia que responden a filtros analógicos de distintos tipo. Así por ejemplo, la ecuación (22.46) responde a un filtro pasa-baja de primer orden, es decir, con una atenuación de 20 dB por década a partir de la frecuencia de corte (ω_c rad/s).

$$G_f(s) = \frac{Y(s)}{X(s)} = \frac{\omega_c}{s + \omega_c} \quad (22.46)$$

Aplicando la aproximación *backwards* se obtiene:

$$G_f(z) = \frac{Y(z)}{X(z)} = \frac{\frac{T\omega_c}{1+T\omega_c}}{1 - \frac{1}{1+T\omega_c} z^{-1}} = \frac{\alpha}{1 - (1 - \alpha)z^{-1}}, \quad (22.47)$$

Donde se supone que $\omega_s \gg \omega_c$. Este resultado expresado con ecuaciones en diferencias se escribe:

$$y_n = \alpha u_n + (1 - \alpha)y_{n-1} \quad (22.48)$$

Es una expresión muy fácil de implementar en código C o en Matlab®. Para el caso de un filtro de segundo orden de la forma:

$$G_f(s) = \frac{Y(s)}{X(s)} = \frac{\omega_c^2}{(s + \omega_c)^2} \quad (22.49)$$

Usando la misma aproximación resulta:

$$G_f(z) = \frac{Y(z)}{X(z)} = \frac{\alpha^2}{1 - 2(1 - \alpha)z^{-1} + (1 - \alpha)^2 z^{-2}} \quad (22.50)$$

$$y_n = \alpha^2 u_n + 2(1 - \alpha)y_{n-1} - (1 - \alpha)^2 y_{n-2} \quad (22.51)$$

donde el parámetro α se define de igual forma que en el caso anterior. Se observa que en los dos filtros desarrollados la salida actual del filtro depende de valores de salida anteriores en el tiempo. A este tipo de filtros se les conoce como filtros IIR (de *infinite impulse response*). El nombre responde al hecho de que, ante una entrada impulso la salida del filtro es siempre no nula en cualquier instante de tiempo posterior al impulso.

Existen otro tipo de filtros cuya salida depende exclusivamente del valor de un número finito de entradas. A este tipo de filtros se les conoce como filtros FIR (*finite impulse response*). Pertenecen a este grupo los que hacen distintas “medias” de los últimos valores de las entradas. Por ejemplo:

$$y_n = \frac{1}{2}u_n + \frac{1}{2}u_{n-1} \quad (22.52)$$

$$y_n = \frac{1}{3}u_n + \frac{1}{3}u_{n-1} + \frac{1}{3}u_{n-2} \quad (22.53)$$

$$y_n = \frac{1}{4}u_n + \frac{1}{4}u_{n-1} + \frac{1}{4}u_{n-2} \quad (22.54)$$

Los filtros FIR son siempre estables, mientras que los IIR pueden presentar respuestas inestables. Observando las expresiones (22.48), (22.51), (22.52), (22.53) y (22.54) parece lógico esperar que los coeficientes que aparecen en la ecuación en diferencias de un filtro sumen la unidad.

22.4. Método de equiparación de polos y ceros

El método de equiparación de polos y ceros (MPZ) utiliza el cambio de variable exacto entre s y z (es decir, $z = e^{Ts}$) para digitalizar una función de transferencia (12.55).

$$G_c(s) = \frac{s + a}{s + b} \longrightarrow G_c(z) = K \frac{z - e^{-aT}}{z - e^{-bT}} \quad (22.55)$$

El método MPZ toma directamente la función de transferencia digital que posee los polos y ceros en los lugares “correspondientes” en el plano \mathcal{Z} . La ganancia K de la función de transferencia discreta es tal que hace:

$$\lim_{s \rightarrow 0} G_c(s) = \lim_{z \rightarrow 1} G_c(z) \quad (22.56)$$

Es decir, se iguala la ganancia estática (o para bajas frecuencias) de las dos funciones de transferencia. En el ejemplo de la ecuación es:

$$K = \frac{a(1 - e^{-bT})}{b(1 - e^{-aT})} \quad (22.57)$$

22.4.1. Caso particular

Con lo dicho hasta ahora del método MPZ es suficiente para digitalizar la mayoría de los controladores analógicos. Especialmente los de adelanto de fase, retraso de fase, y adelanto-retraso, que poseen funciones de transferencia con igual número de polos que de ceros. Sin embargo, en caso de que se desee digitalizar una función de transferencia que posea más polos que ceros, es usual añadir tantos ceros en $z = -1$ como sean necesarios para igualar el número de polos y de ceros en \mathcal{Z} .

Por ejemplo, en la siguiente digitalización se añade un cero que antes no existía en la función de transferencia discreta:

$$G_c(s) = \frac{s+a}{s(s+b)} \longrightarrow G_c(z) = K \frac{(z - e^{-aT})(z+1)}{(z-1)(z - e^{-bT})} \quad (22.58)$$

El ajuste de la ganancia se hace como en el caso general.

22.4.2. Método de equiparación modificado

El método modificado de equiparación de polos y ceros (MMPZ) se suele aplicar sólo en el caso de que no sea posible utilizar el valor actual del error para el cálculo de la actuación del controlador.

Es decir, cuando un controlador en \mathcal{Z} posee igual número de polos que de ceros, si se escribe la ecuación diferencia de la actuación del controlador en el instante actual, dicha actuación será función de actuaciones pasadas y valores del error actuales y pasados:

$$u_n = f(u_{n-1}, u_{n-2}, u_{n-3}, \dots, e_n, e_{n-1}, e_{n-2}, \dots) \quad (22.59)$$

Pero en algunos sistemas de control no es posible contar con la lectura actual del error para comandar (en ese mismo instante) la actuación en función de ella. Por ejemplo, un convertidor A/D puede necesitar “mucho tiempo” (comparado con el periodo de muestreo) y en ese caso lo que se suele hacer es solicitar una lectura en la interrupción actual pero no usarla hasta la interrupción siguiente. Con este tipo de estrategia de lectura, sólo es posible implementar ecuaciones en diferencias en las que no intervenga el error actual:

$$u_n = f(u_{n-1}, u_{n-2}, u_{n-3}, \dots, e_{n-1}, e_{n-2}, e_{n-3}, \dots) \quad (22.60)$$

Una forma de conseguirlo para el caso particular del apartado anterior es no introducir tantos ceros ficticios como los necesarios para igualar el número de polos y de ceros; sino dejar la función de transferencia con un polo más que ceros. Esta forma de actuar es lo que se llama MMPZ.

NOTA: En caso de que la función de transferencia en \mathcal{Z} ya posea igual número de polos que de ceros, otra solución es simplemente introducir un retraso de un periodo de muestreo en la función de transferencia (es decir, multiplicarla por z^{-1}). Esto también hace que la función de transferencia final tiene un polo más que ceros, pero esta vez la estrategia ha sido introducir un nuevo polo en el origen del plano \mathcal{Z} .

22.4.3. Ejemplo

Se va a diseñar un compensador para el mismo ejemplo del apartado 21.5 y las mismas especificaciones. Primero se diseña el compensador de adelanto de fase con continuo por el método del lugar de las raíces:

$$G_c(s) = 2.94 \frac{s+1}{s+10} \quad (22.61)$$

Como se ve, se ha elegido también anular el segundo polo de la planta. Ahora se debe elegir el periodo de muestreo. Para comparar los resultados se va a elegir $T = 0.1$ como en el caso anterior. Ahora ya se puede digitalizar el compensador por el método MPZ:

$$G_c(z) = 1.96 \frac{z - 0.9048}{z - 0.3678} \quad (22.62)$$

Como es lógico, el resultado es un poco diferente al que se propuso cuando se ajustó directamente en el lugar de las raíces del plano \mathcal{Z} . Sin embargo, las diferencias son pequeñas y se cumplen las especificaciones de diseño.

Como valor de referencia, el método de digitalización MPZ suele elegir un periodo de muestreo más pequeño que el propuesto para conseguir 8 ó 10 muestras por ciclo. En concreto se suele elegir una frecuencia de muestreo 20 veces mayor que la frecuencia natural del sistema compensado. Es decir:

$$\omega_s \geq 20\omega_n \quad (22.63)$$

$$T \leq \frac{\pi}{10\omega_n} \quad (22.64)$$

22.5. Método de la equivalencia del retenedor

La equivalencia del retenedor realiza la digitalización (también una vez que se posee la función de transferencia en s y se ha elegido el periodo de muestreo) simplemente calculando:

$$G_c(s) \longrightarrow G_c(z) = \mathcal{Z}[ZOH(s)G_c(s)] \quad (22.65)$$

El resultado de este método para el mismo ejemplo que el apartado anterior es:

$$G_c(z) = \mathcal{Z} \left[\frac{1 - e^{-Ts}}{s} 2.94 \frac{s+1}{s+10} \right] \quad (22.66)$$

$$G_c(z) = 2.94(1 - z^{-1}) \mathcal{Z} \left[\frac{s+1}{s(s+10)} \right] \quad (22.67)$$

$$G_c(z) = 2.94 \frac{z - 0.9367}{z - 0.3678} \quad (22.68)$$

También el resultado es parecido, aunque la ganancia es un poco grande, por lo que la respuesta será menos amortiguada que en el caso anterior. Por ejemplo en este ejemplo el sobreimpulso alcanza el 40% cuando en el controlador digital lo dejaba en torno al 20%.

Capítulo 23

Respuesta en frecuencia

El diagrama de Bode también tiene sentido para funciones de transferencia discreta. En este capítulo se estudia cómo dibujarlo de forma sencilla y usarlo para ajustar compensadores de adelanto o de retraso de fase directamente en el dominio \mathcal{Z} .

23.1. Aproximación de la respuesta en frecuencia

Para poder aplicar el criterio de estabilidad de Routh-Hurwitz se mostró una transformación bilineal que conseguía transformar el círculo unitario en el semiplano de parte real negativa del nuevo plano W . Se dijo entonces que la posición de los polos dentro del plano W no ofrecía especial información sobre la respuesta transitoria del sistema.

En este apartado se investigará si es posible encontrar una transformación bilineal que sí permita de alguna manera realizar este tipo de análisis. El estudio centrará su atención en cómo transforma el cambio de variable bilineal la posición de un polo que esté situado sobre el círculo unitario:

$$z_1 = 1 \angle \theta \quad (23.1)$$

$$z_1 = e^{\theta j} \quad (23.2)$$

Después de la transformación bilineal, la posición de este polo en el plano W es:

$$w_1 = \frac{z_1 - 1}{z_1 + 1} = \frac{e^{\theta j} - 1}{e^{\theta j} + 1}, \quad (23.3)$$

y para el caso de $\theta = \frac{\pi}{2}$ es:

$$w_1 = \frac{e^{\frac{\pi}{2}j} - 1}{e^{\frac{\pi}{2}j} + 1} = \frac{j - 1}{j + 1} = j. \quad (23.4)$$

Como era de esperar se encuentra sobre el eje imaginario. Como ya se explicó en el capítulo 19, los valores de frecuencias que se miden en el plano W no corresponden con los valores del plano S . Se va a medir a continuación cuánto vale esa discrepancia. El cambio de variable de w a s es:

$$w = \frac{z - 1}{z + 1} = \frac{e^{Ts} - 1}{e^{Ts} + 1} = \frac{e^{\frac{T}{2}s} - e^{-\frac{T}{2}s}}{e^{\frac{T}{2}s} + e^{-\frac{T}{2}s}} = \tanh\left(\frac{Ts}{2}\right) \quad (23.5)$$

Por tanto los puntos imaginarios puros del plano W , $w = \nu j$, se relacionan con los puntos imaginarios puros del plano S , $s = \omega j$, de la siguiente forma:

$$\nu j = \tanh\left(\frac{T\omega j}{2}\right) = \tanh\left(\frac{T\omega}{2}\right) j \quad (23.6)$$

Cuando el ángulo $T\omega$ es pequeño, la expresión anterior se puede aproximar como:

$$\nu j \approx \frac{T\omega}{2} j \quad (23.7)$$

Por tanto, las frecuencias naturales amortiguadas que se miden en el plano W es igual a las frecuencias naturales amortiguadas que se miden en el plano S multiplicadas por la una constante igual al periodo de muestreo partido por dos:

$$\nu \approx \frac{T}{2} \omega \quad (23.8)$$

Se podría definir un nuevo cambio de variable que elimine esta diferencia:

$$x = \frac{2}{T} \frac{z-1}{z+1} = \frac{2}{T} \tanh\left(\frac{Ts}{2}\right) \quad (23.9)$$

Los puntos imaginarios puros del nuevo plano X , $x = \xi j$, se corresponden bien con los puntos imaginarios puros del plano S , $s = \omega j$, mientras el producto $T\omega$ sea relativamente pequeño:

$$\xi j = \frac{2}{T} \tanh\left(\frac{T\omega j}{2}\right) = \frac{2}{T} \tanh\left(\frac{T\omega}{2}\right) j \approx \omega j \quad (23.10)$$

La nueva transformación bilineal definida en (23.9) también se puede usar como paso previo a la aplicación del criterio de Routh-Hurwitz, ya que sólo “escala” las frecuencias el plano W . También es interesante señalar que la nueva transformación coincide exactamente con la definición del método de integración de Tustin. Por tanto, la nueva variable x es una aproximación muy buena de la variable original s de Laplace. Por esta razón también se dice que el método de Tustin asegura la transformación de una función de transferencia continua estable en otra función de transferencia en la variable z que es también estable.

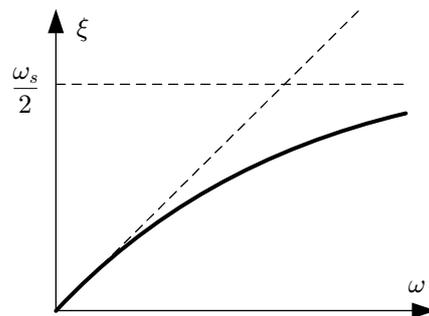


Figura 23.1: Frecuencia real ξ en función de la frecuencia aproximada ω

En la Fig. 23.1 se comparan cualitativamente la frecuencia real ξ con la aproximada ω . La gráfica cambia dependiendo del periodo de muestreo, pero en general para pequeñas frecuencias ω , la correspondencia con ξ es exacta, mientras que para altas frecuencias la correspondencia no se mantiene, sino que ξ tiende a la frecuencia de Nyquist.

Como consecuencia, los análisis en el dominio de la frecuencia que se estudiaron para sistemas continuos y en particular el diagrama de Bode de una función de transferencia $G(s)$, se pueden realizar, de forma suficientemente aproximada con las funciones de transferencia equivalentes $G(x)$. Por tanto, el procedimiento que se propone para diseñar compensadores en el dominio en frecuencia es el siguiente:

1. Tomar la función de transferencia “en lazo abierto” del sistema de control $G(z)$.
2. Calcular la función de transferencia equivalente $G(x)$ con la nueva transformación bilineal, es decir, con la aproximación de Tustin. El cambio de z a x es:

$$z = \frac{1 + \frac{T}{2}x}{1 - \frac{T}{2}x} \quad (23.11)$$

3. Representar el diagrama de Bode de $G(x)$ y diseñar un controlador $G_c(x)$ utilizando dicho diagrama.
4. Deshacer el cambio de variable para traducir el controlador $G_c(x)$ al dominio \mathcal{Z} , $G_c(z)$.

$$x = \frac{2}{T} \frac{z-1}{z+1} \quad (23.12)$$

23.2. Ejemplo numérico

A continuación se dibujan la respuesta en frecuencia exacta y distintas aproximaciones aplicando los la fórmula de Tustin para diferentes periodos de muestreo. La función de transferencia de la planta es:

$$G(s) = \frac{1.8}{1 + 0.2s} = 1.8 \frac{5}{s + 5} \quad (23.13)$$

Para diseñar un controlador en el dominio de la frecuencia hay que dibujar el diagrama de Bode de la función de transferencia en lazo abierto. Si la planta (23.13) se introduce en un lazo de control con retroalimentación unitaria, muestro, controlador digital, retenedor y planta, es evidente que se debe representar el diagrama de Bode de la planta más el retenedor, para posteriormente añadir la influencia del controlador (lazo abierto). Por tanto, en primer lugar se calcula la transformada \mathcal{Z} de la planta más el retenedor, que es el lazo abierto sin controlador:

$$G_{la}(z) = 1.8\mathcal{Z}\left[\frac{1 - e^{-Ts}}{s} \frac{5}{s + 5}\right] = 1.8\frac{1 - e^{-5T}}{z - e^{-5T}} \quad (23.14)$$

En este momento es cuando se debe elegir un periodo de muestreo adecuado para el sistema, en función de las especificaciones de diseño. Se va a particularizar para $T_1 = \frac{1}{3}$ s y para $T_2 = \frac{1}{15}$ s.

$$G_{la1}(z) = \frac{1.46}{z - 0.188} \quad (23.15)$$

$$G_{la2}(z) = \frac{0.51}{z - 0.716} \quad (23.16)$$

Se comprueba una vez más que el retenedor no modifica el orden de la planta. Aplicando la transformación de Tustin (23.11), que es distinta en función del periodo de muestreo, las aproximaciones de la función de transferencia en lazo abierto son:

$$G_{la1}(z) = 1.2289\frac{6 - x}{4.1101 + x} \quad (23.17)$$

$$G_{la2}(z) = 0.2972\frac{30 - x}{4.965 + x} \quad (23.18)$$

Se observa que la aproximación de Tustin ha introducido un cero de fase no mínima en la función de transferencia. También se puede observar cómo la ganancia estática de la función de transferencia (tanto exacta con $z \rightarrow 1$ como aproximada con $x \rightarrow 0$) permanece constante e igual al valor original de (en este caso 1.8 cuando $s \rightarrow 0$).

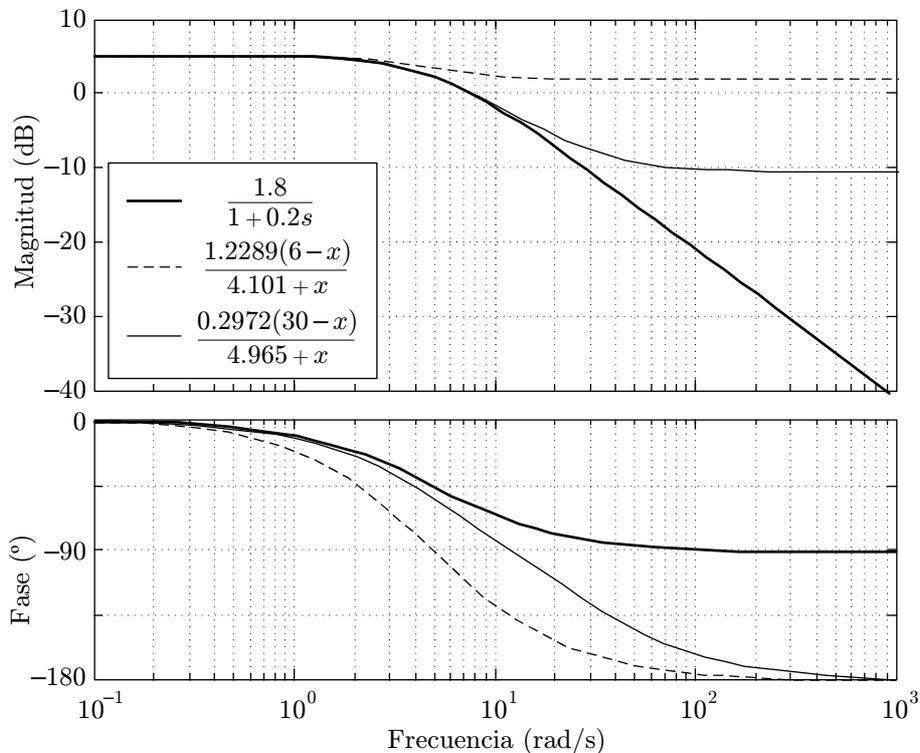


Figura 23.2: Diagramas de Bode de la función de transferencia en lazo abierto continua y aproximadas

En la Fig. 23.2 se representan los diagramas de Bode de las funciones de transferencia (23.13), (23.17) y (23.18), tomando el mismo el eje de abscisas para la frecuencia real y aproximada. En el diagrama se observa que las aproximaciones se asemejan a la original para bajas frecuencias, como sugiere la ecuación

(23.10). Atendiendo a esta misma ecuación resulta también lógico que la aproximación con $T = \frac{1}{15}$ s sea más similar al Bode de la función de transferencia continua que la que se obtiene con $T = \frac{1}{3}$ s. En general un sistema digital se comporta tanto más parecido al continuo original cuanto menor es el periodo de muestreo (mayor la frecuencia de muestreo).

23.3. Respuesta en frecuencia exacta

En este apartado se resalta que la respuesta en frecuencia del sistema digital se parece siempre más a la correspondiente aproximación en el dominio X que a la que se obtiene con la función de transferencia continua. De otra forma no tendría sentido hallar las aproximaciones que se están estudiando. En la Fig. 23.3 se comparan los diagramas de Bode de la función de transferencia continua, digital y aproximada para $T = \frac{1}{3}$ s del ejemplo del apartado anterior.

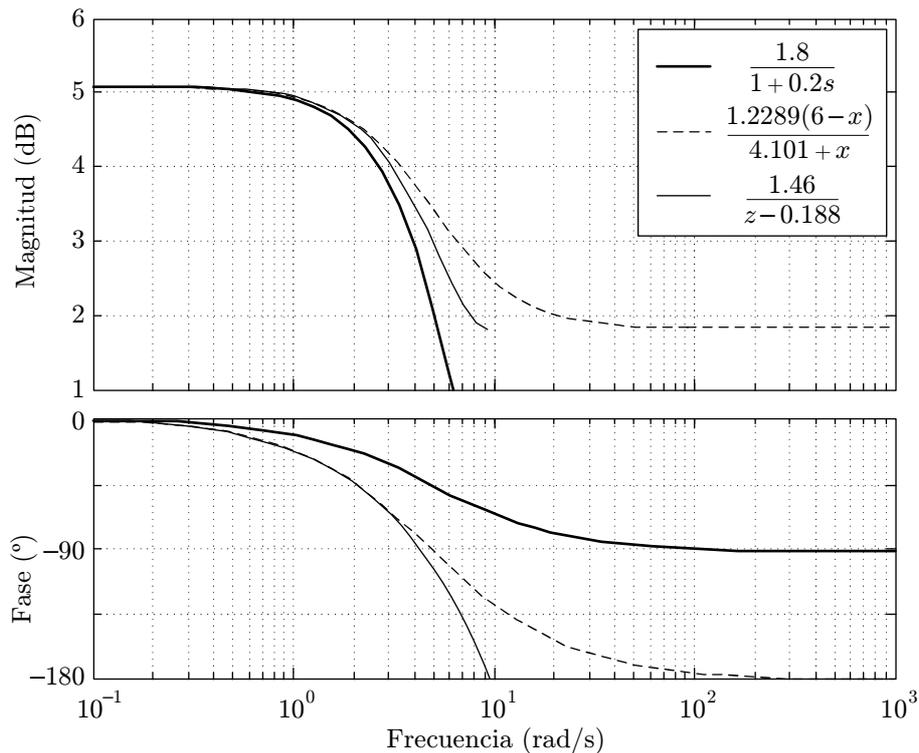


Figura 23.3: Diagramas de Bode de la función de transferencia continua, digital y aproximada

La respuesta en frecuencia exacta de la función de transferencia muestreada se ha obtenido directamente con Matlab® y sólo toma valores hasta la frecuencia de Nyquist (en este caso 9.42 rad/s) porque no tiene sentido excitar el sistema con frecuencias superiores, ya que se produciría aliasing. Atendiendo a la Fig. 23.1, se podría decir que la respuesta en frecuencia exacta en Z es parecida a la aproximada en X pero “apelmazando” en torno a la frecuencia de Nyquist lo que tienda a infinito.

Con la Fig. 23.3 queda patente la utilidad de las funciones de transferencia aproximadas en el dominio X , ya que se asemejan al Bode real del sistema muestreado en un rango mucho mayor de frecuencias.

Capítulo 24

Espacio de estado muestreado

24.1. Introducción

Los modelos de los sistemas en espacio de estado también se pueden introducir el concepto de muestreo. Como se observa en la definición general (24.1), los valores de los estados en el periodo de muestreo siguiente $\mathbf{x}(nT+T)$, dependen de los valores actuales de las entradas, $\mathbf{u}(nT)$ y los propios estados $\mathbf{x}(nT)$; mientras que las salidas en el instante actual dependen de los valores actuales de los estados $\mathbf{x}(nT)$.

$$\begin{cases} \mathbf{x}(nT+T) = \mathbf{G}\mathbf{x}(nT) + \mathbf{H}\mathbf{u}(nT) & \text{ecuación de estado} \\ \mathbf{y}(nT) = \mathbf{C}\mathbf{x}(nT) + \mathbf{D}\mathbf{u}(nT) & \text{ecuación de salida} \end{cases} \quad (24.1)$$

En espacio de estado es habitual usar valores naturales, k , para referirse a los instantes de muestreo en lugar de poner múltiplos del periodo de muestreo, nT . Por otro lado, en este capítulo sólo se consideran sistemas estrictamente propios de una entrada y una salida:

$$\begin{cases} \mathbf{x}(k+1) = \mathbf{G}\mathbf{x}(k) + \mathbf{h}u(k) \\ y(k) = \mathbf{c}^T\mathbf{x}(k) \end{cases} \quad (24.2)$$

Para no confundir el dominio continuo del muestreo, la ecuación de estado emplea letras distintas para sus matrices: \mathbf{G} y \mathbf{h} . En cualquier caso, los valores propios de la matriz \mathbf{G} son los polos del sistema muestreado.

24.2. Ejemplo de modelización

Considérese un sistema cuya dinámica puede caracterizarse mediante la función de transferencia (24.3). Se pide un modelo en espacio de estado muestreado, para el caso de $T = 0.01$ segundos.

$$\frac{Y(s)}{U(s)} = \frac{1}{5s^2 + 3s} \quad (24.3)$$

Dado que el periodo de muestreo es bastante inferior a la constante de tiempo propia de la respuesta del sistema, se propone —como primera aproximación muestreada de la dinámica del sistema— la transformada \mathcal{Z} de la función de transferencia del sistema continuo:

$$\frac{Y(z)}{U(z)} = \mathcal{Z} \left[ZOH(s) \frac{1}{5s^2 + 3s} \right] \quad (24.4)$$

$$\frac{Y(z)}{U(z)} = 10^{-5} \frac{2.994z + 2.988}{(z-1)(z-0.994)} \quad (24.5)$$

$$\frac{Y(z)}{U(z)} = 10^{-5} \frac{2.994z + 2.988}{z^2 - 1.994z + 0.994} = \frac{Y(z)}{Y_1(z)} \frac{Y_1(z)}{U(z)} \quad (24.6)$$

Como se hizo con las variables de fase en el apartado 11.2.1, se utiliza una variable intermedia $Y_1(z)$ para dividir la función de transferencia en dos partes: por un lado se considerará el numerador y por otro el denominador:

$$Y(z) = 10^{-5} [2.994zY_1(z) + 2.988Y_1(z)] \quad (24.7)$$

$$U(z) = z^2Y_1(z) - 1.994zY_1(z) + 0.994Y_1(z) \quad (24.8)$$

Siendo el sistema de segundo orden, hay que definir dos variables como estados. En este caso, se toman como estados del sistema: $Y_1(z)$ e $Y_2(z) = zY_1(z)$, de forma que se puede escribir:

$$\begin{bmatrix} zY_1(z) \\ zY_2(z) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -0.994 & 1.994 \end{bmatrix} \begin{bmatrix} Y_1(z) \\ Y_2(z) \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} U(z) \quad (24.9)$$

$$Y(z) = 10^{-5} [2.988 \quad 2.994] \begin{bmatrix} Y_1(z) \\ Y_2(z) \end{bmatrix} \quad (24.10)$$

Como era de esperar, los valores propios de \mathbf{G} son 1 y 0.994 que coinciden con los polos del sistema muestreado (que, por cierto, es de tipo I).

Es interesante hacer notar que, siguiendo el método descrito para obtener variables de fase, los distintos estados del sistema no son las sucesivas derivadas del estado inicial elegido. La variable $Y_2(z)$ en el dominio temporal es la misma variable física que $Y_1(z)$, pero en el instante siguiente: $y_2(nT) = y_1(nT + T)$. Por tanto, en este ejemplo, para medir todos los estados de cara a una realimentación completa de estados, podría bastar medir una única variable física del sistema, teniendo cuidado de guardar un “histórico” de las lecturas en periodos de muestreo anteriores. Tantos valores como sea el orden del sistema.

Por otro lado, la variable física medida en el instante actual habría que identificarla en el ejemplo con $Y_2(z)$. Nunca con $Y_1(z)$. De la primera forma, la correcta, y_2 es la variable física actual e y_1 es la misma variable en el periodo de muestreo anterior. Mientras que siguiendo la forma incorrecta, y_1 sería la variable física actual e y_2 sería el valor de esa variable en el periodo de muestreo siguiente. Por ser un tiempo futuro no se podría implementar físicamente.

En este punto también es interesante hacer notar que ninguno de los estados que se han propuesto en el ejemplo coinciden con la salida $Y(z)$ del sistema. Esto es muy habitual en los sistemas muestreados, en los que normalmente se obtienen no sólo polos, sino también ceros. Por tanto, el modelo en espacio de estado muestreado que se ha propuesto en este ejemplo no es muy útil. Habría que encontrar uno equivalente —en definitiva proponer un cambio de variable a través de una matriz de transformación— buscando que \mathbf{c}^T sea $[0 \ 1]$. Así el estado en el periodo de muestreo actual coincide con la salida actual.

24.3. Control mediante realimentación completa de estados

De forma análoga a como se hizo en control continuo en espacio de estado, se propone ahora un control mediante realimentación completa de estados. Supóngase, en el ejemplo del apartado anterior, que los estado $Y_1(z)$ y $Y_2(z)$ son ambos conocidos por el ingeniero y, por tanto, se pueden realimentar los dos. Se pide que el sistema controlado posea ante una entrada escalón un tiempo de establecimiento de 1 segundo y un amortiguamiento de 0.7.

Con dichas especificaciones, los polos del sistema en el plano complejo S deberían estar en $-4 \pm 4j$, pero en el plano \mathcal{Z} deben estar en $0.96 \pm 0.03842j$, teniendo en cuenta que el periodo de muestreo eran 0.01 segundos. Mediante la realimentación completa de estados a través del vector \mathbf{k}^T , la nueva matriz de estados del sistema compensado es $\mathbf{G} - \mathbf{h}\mathbf{k}^T$, es decir:

$$\mathbf{G} - \mathbf{h}\mathbf{k}^T = \begin{bmatrix} 0 & 1 \\ -0.994 & 1.994 \end{bmatrix} - \begin{bmatrix} 0 \\ 1 \end{bmatrix} [k_1 \quad k_2] = \begin{bmatrix} 0 & 1 \\ -0.994 - k_1 & 1.994 - k_2 \end{bmatrix} \quad (24.11)$$

Por tanto, la ecuación característica del sistema compensado es:

$$\det(z\mathbf{I} - \mathbf{G} + \mathbf{h}\mathbf{k}^T) = 0 \quad (24.12)$$

$$z^2 + (k_2 - 1.994)z + k_1 + 0.994 = 0 \quad (24.13)$$

De forma sencilla se puede calcular que, para que las raíces de esta ecuación sean $0.96 \pm 0.03842j$, los valores de las ganancias deben ser: $k_1 = -0.07093$ y $k_2 = 0.074$. Con estas ganancias de realimentación se consiguen las especificaciones de régimen transitorio deseadas. Para conseguir error nulo ante una determinada entrada escalón, todavía había que proponer un valor para la ganancia k_{ref} que multiplica a la referencia.

Referencias

- [1] John G. Ziegler and Nataniel B. Nichols, “Optimum Settings for Automatic Controllers”, *Transactions of the ASME*, 1942, vol. 64, pp 759–768.
- [2] Jürgen Ackermann, *Abtastregelung*, Springer-Verlag, Berlin, 1972.
- [3] Rudolf Emil Kalman, “A New Approach to Linear Filtering and Prediction Problems”, *Journal of Basic Engineering*, 1960, pp 35–45.
- [4] Rudolf Emil Kalman and Richard S. Bucy, “New Results in Linear Filtering and Prediction Theory”, *Journal of Basic Engineering*, 1961, pp 95–108.
- [5] Charles L. Phillips and H. Troy Nagle, *Digital Control System: Analysis and Design*, Prentice-Hall International, New Jersey, USA, 1990.

Apéndice A

Ampliación de espacio de estado

A.1. Matriz de transición de estados

El cálculo de la solución del siguiente sistema de ecuaciones homogéneo,

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t), \quad (\text{A.1})$$

es relativamente sencillo en el dominio de Laplace, suponiendo que los estados del sistema parten de un estado inicial \mathbf{x}_0 no nulo:

$$s\mathbf{x}(s) - \mathbf{x}_0 = \mathbf{A}\mathbf{x}(s) \quad (\text{A.2})$$

$$(s\mathbf{I} - \mathbf{A})\mathbf{x}(s) = \mathbf{x}_0 \quad (\text{A.3})$$

$$\mathbf{x}(s) = (s\mathbf{I} - \mathbf{A})^{-1}\mathbf{x}_0 \quad (\text{A.4})$$

$$\mathbf{x}(t) = \mathcal{L}^{-1} [(s\mathbf{I} - \mathbf{A})^{-1}] \mathbf{x}_0, \quad (\text{A.5})$$

$$\mathbf{x}(t) = \Phi(t)\mathbf{x}_0, \quad (\text{A.6})$$

donde $\Phi(t)$ es una matriz que contiene funciones temporales y se conoce como *matriz de transición de estados*. Esta matriz $\Phi(t)$ también se puede calcular como:

$$\Phi(t) = \mathcal{L}^{-1} [(s\mathbf{I} - \mathbf{A})^{-1}] = \mathbf{I} + \mathbf{A}t + \mathbf{A}^2 \frac{t^2}{2} + \mathbf{A}^3 \frac{t^3}{3!} + \dots = \sum_{k=0}^{\infty} \frac{\mathbf{A}^k t^k}{k!} = e^{\mathbf{A}t} \quad (\text{A.7})$$

Se puede demostrar fácilmente la definición de $\Phi(t) = e^{\mathbf{A}t}$, comprobando que se cumple el sistema de ecuaciones homogéneo (A.1):

$$\dot{\mathbf{x}}(t) = \frac{d\mathbf{x}(t)}{dt} = \frac{d(e^{\mathbf{A}t}\mathbf{x}_0)}{dt} = \frac{de^{\mathbf{A}t}}{dt}\mathbf{x}_0 \quad (\text{A.8})$$

$$\dot{\mathbf{x}}(t) = \frac{d}{dt} \left(\mathbf{I} + \mathbf{A}t + \mathbf{A}^2 \frac{t^2}{2} + \mathbf{A}^3 \frac{t^3}{3!} + \dots \right) \mathbf{x}_0 \quad (\text{A.9})$$

$$\dot{\mathbf{x}}(t) = \left(\mathbf{A} + \mathbf{A}^2 t + \mathbf{A}^3 \frac{t^2}{2!} + \dots \right) \mathbf{x}_0 = \mathbf{A}e^{\mathbf{A}t}\mathbf{x}_0 = \mathbf{A}\mathbf{x}(t) \quad (\text{A.10})$$

A continuación se presenta una tabla resumen de las propiedades de la matriz de transición de estados:

Tabla A.1: Propiedades de la matriz de transición de estados

Propiedad	Expresión
Valor inicial	$\Phi(0) = \mathbf{I}$
Inversa	$\Phi^{-1}(t) = \Phi(-t)$
Translación	$\Phi(t_1 + t_2) = \Phi(t_1)\Phi(t_2)$
Potencia	$\Phi^n(t) = \Phi(nt)$