

Control PID avanzado

Karl J. Åström
Tore Hägglund

PEARSON
Prentice
Hall

Control PID avanzado

Control PID avanzado

Karl J. Åström Tore Hägglund
*Department of Automatic Control
Lund Institute of Technology
Lund University*

TRADUCCIÓN Y REVISIÓN TÉCNICA
Sebastián Dormido Bencomo
*Dpto. Informática y Automática
UNED*

José Luis Guzmán Sánchez
*Dpto. Lenguajes y Computación
Universidad de Almería*



Madrid · México · Santafé de Bogotá · Buenos Aires · Caracas · Lima · Montevideo
San Juan · San José · Santiago · São Paulo · White Plains

CONTROL PID AVANZADO
Karl J. Åström, Tore Hägglund

PEARSON EDUCACIÓN, S.A. 2009
ISBN: 978-84-8322-511-0

Materia: 681.5, Ingeniería

Formato: 195 x 250 mm Páginas: 488

Cualquier forma de reproducción, distribución, comunicación pública o transformación de esta obra sólo puede ser realizada con la autorización de sus titulares, salvo excepción prevista por la ley. La infracción de los derechos mencionados puede ser constitutiva de delito contra la propiedad intelectual (*arts. 270 y sgts. Código penal*).

Diríjase a CEDRO (Centro Español de Derechos Reprográficos: www.cedro.org), si necesita fotocopiar o escanear algún fragmento de esta obra.

DERECHOS RESERVADOS
© 2009, PEARSON EDUCACIÓN S. A.
Ribera del Loira, 28
28042 Madrid (España)

ISBN: 978-84-8322-511-0

Depósito Legal:

Equipo editorial:

Editor: Miguel Martín-Romo

Técnico Editorial:

Equipo de producción:

Director: José A. Clares

Técnico: Diego Marín

Diseño de cubierta: Equipo de diseño de Pearson Educación S. A.

Composición:

Impreso por:

IMPRESO EN ESPAÑA - PRINTED IN SPAIN

Este libro ha sido impreso con papel y tintas ecológicos.

Nota sobre enlaces a páginas web ajena: Este libro puede incluir enlaces a sitios web gestionados por terceros y ajenos a PEARSON EDUCACIÓN S.A. que se incluyen sólo con finalidad informativa. PEARSON EDUCACIÓN S.A. no asume ningún tipo de responsabilidad por los daños y perjuicios derivados del uso de los datos personales que pueda hacer un tercero encargado del mantenimiento de las páginas web ajenas a PEARSON EDUCACIÓN S. A y del funcionamiento, accesibilidad o mantenimiento de los sitios web no gestionados por PEARSON EDUCACIÓN S.A. Las referencias se proporcionan en el estado en que se encuentran en el momento de publicación sin garantías, expresas o implícitas, sobre la información que se proporcione en ellas.

Prefacio

El controlador PID es la solución más común a los problemas prácticos de control. Aunque controladores con acción proporcional e integral han sido utilizados desde la época en que los molinos de viento y las máquinas de vapor eran las tecnologías dominantes, la forma actual del controlador PID emergió con los controladores neumáticos en los años 30 del siglo pasado. Una razón fue que las realizaciones con computadores hizo posible añadir características tales como capacidad de autosintonía y diagnóstico, que son muy beneficiosas para los usuarios. Desde una perspectiva de ingeniería, resulta particularmente interesante analizar lo que sucedió con las sucesivas tecnologías que se fueron incorporando, que hicieron que algunas funcionalidades importantes fueran redescubiertas y otras fueran añadidas.

Este libro se ha elaborado durante más de 25 años de desarrollo de los auto sintonizadores para los controladores PID en estrecha colaboración con la industria. A través de este trabajo hemos abordado un gran número de problemas reales de control industrial. Nos hemos beneficiado mucho de haber podido participar en el desarrollo, puesta en marcha y problemas de los controladores industriales. El trabajo práctico ha inspirado también la investigación.

Este texto es la última parte de una trilogía. El primero de ellos, *Automatic Tuning of PID Controllers, 1988*, que tenía 6 capítulos, daba una breve descripción de nuestras primeras experiencias en el desarrollo de controladores autosintonizados basados en el método del relé. El segundo, *PID Controllers: Theory, Design, and Tuning, 1995*, que tiene 7 capítulos, surgió de la necesidad de tratar de forma más amplia muchos aspectos del control PID. En particular, revisa muchos métodos de diseño para controladores PID que investigamos en conexión con nuestro trabajo sobre los controladores autosintonizados.

El conocimiento que se tenía en 1995 del control PID todavía no era satisfactorio para el diseño de los controladores autosintonizados. Una desventaja era que el usuario tenía que proporcionar al controlador con ciertas elecciones de diseño. Resulta particularmente difícil para un usuario evaluar si la dinámica está dominada por el retardo o por la constante de tiempo. Esta cuestión estimuló más investigación sobre el tema. Debido al aumento drástico en la potencia de cálculo, fue también posible emplear algoritmos de diseño que requieren más computación.

La sintonía y diseño de controladores PID se ha basado tradicionalmente en técnicas especiales. El control robusto fue un gran desarrollo de la teoría de con-

trol que maduró a finales de los 90, y que dió como resultado potentes métodos de diseño basados en la deformación robusta de la función de transferencia en lazo abierto. Esto nos estimuló a iniciar un programa de investigación para adaptar estos métodos al control PID. Al mismo tiempo parecía natural acercar el control PID a la corriente principal de ideas en control. Cuando trabajamos con auto sintonizadores industriales, también vimos que había una gran necesidad de incluir diagnósticos en el controlador, ya que no se suele sintonizar un controlador si el proceso tiene activas alarmas graves. El presente libro, *Control PID Avanzado*, es el resultado de este esfuerzo.

Con un total de 13 capítulos, este nuevo libro amplia sustancialmente algunos de los temas cubiertos en las versiones previas y proporciona algunos nuevos capítulos que tratan del diseño del controlador, diseño por anticipación (feedforward), sustitución de las reglas de sintonía de Ziegler-Nichols, control predictivo, evaluación del comportamiento del lazo e interacción. En este punto en nuestra trilogía de libros, asumimos que el lector está familiarizado con la teoría del control.

Nuestra investigación ha dado como resultado una comprensión más profunda de los compromisos entre atenuación de la perturbación en la carga, inyección de ruido en la medida y respuesta a cambios en el punto de consigna. Hemos sido también capaces de responder a cuestiones tales como: ¿Debería un controlador sintonizarse para dar respuesta a perturbaciones en la carga o a cambios en el punto de consigna?, ¿Qué información se requiere para diseñar un controlador PID?, ¿Cuándo puede la acción derivativa dar unas mejoras significativas?, ¿Cuándo se justifica el empleo de controladores más sofisticados? Con el conocimiento desarrollado, ahora es posible diseñar controladores autosintonizados que pueden hacer estas evaluaciones autónomamente. Además, hemos desarrollado nuevos métodos sencillos para diseñar controladores PID.

Como un ejemplo del conocimiento que se ha ganado podemos mencionar que la teoría de control dice que no es necesario efectuar un compromiso entre sintonía para respuesta a perturbaciones en la carga y respuesta a cambios en el punto de consigna. Ambos requisitos se pueden cumplir utilizando un controlador con dos grados de libertad, que combina realimentación y acción feedforward. Las ganancias de la realimentación se deberían escoger de forma que se satisfieran los requerimientos sobre la atenuación de la perturbación y la robustez. La respuesta deseada a cambios en el punto de consigna se puede entonces obtener mediante un uso apropiado de la acción feedforward. La ponderación del punto de consigna es una forma sencilla de acción feedforward para control PID. En algunos casos, está justificado utilizar acciones feedforward más elaboradas. Por esta razón, hemos incluido un capítulo sobre diseño del controlador y otro capítulo sobre feedforward en el nuevo libro.

El análisis de robustez también muestra la ventaja de tener una baja ganancia del controlador en altas frecuencias y una gran disminución de la ganancia en altas frecuencias. Esto se puede conseguir filtrando la salida del proceso con un filtro de segundo orden. Basándonos en el conocimiento obtenido, se recomienda un uso generalizado de la ponderación del punto de consigna o una acción feedforward más avanzada. También recomendamos que se filtre la salida del proceso utilizando un filtro de segundo orden.

Nos gustaría agradecer a muchas personas que han compartido con nosotros

sus conocimientos, ideas e inspiración. Nuestro interés en el control PID tuvo su motivación a comienzos de los 80 en nuestros encuentros con Axel Westrelius y Mike Somerville de Eurotherm. Hemos aprendido mucho del trabajo con nuestros estudiantes; en particular damos las gracias a Lars Göran Elfgren (Eurotherm), Göran Grönhammar (LTH), Ari Ingimundarson (UPC), Oskar Nordin (Volvo), Helene Panagopoulos (Volvo), Per Persson (Volvo), Mikael Petersson (ABB), Ola Slättke (ABB), and Anders Wallén (Ericsson Mobile Platforms), quienes continúan proporcionándonos ideas muy valiosas aunque ahora prosiguen sus carreras en la industria.

Estamos muy agradecidos a Sune Larsson y Lars Bááth, anteriormente en NAF Controls, con quienes desarrollamos el primer controlador autosintonizado industrial basado en el método del relé. La compañía NAF Controls se fusionó algunas veces y ahora forma parte de ABB, donde hemos disfrutado de interacciones con Göran Arinder, Alf Isaksson, Per Erik Modén, Lars Pernebo, y Thomas Vonheim. Hemos compartido el placer y los retos de transferir técnicas de auto sintonía y diagnóstico a un amplio abanico de productos industriales. Muchas conversaciones estimulantes con nuestros colegas Anton Cervin (LTH), Sebastián Dormido (UNED), Guy Dumont (UBC), Chang Chieh Hang (NUS), Karl Henrik Johansson (KTH), Birgitta Kristiansson (CTH), Bengt Lennartsson (CTH), Manfred Morari (ETH), Dale Seborg (UCSB), Sigurd Skogestad (NTNU), Björn Wittenmark (LTH), and Karl-Erik Årzén (LTH) del mundo académico son también muy apreciadas.

Nuestros amigos en la industria Bill Bialkowski, Terry Blevins, Greg McMillan, y Willy Wojsznis de Emerson, Edgar Bristol, Sigifredo Niño, y Greg Shinskey de Foxboro, Börje Eriksson (M-real), Krister Forsman (Perstorp), Ken Goff (Leeds and Northrup), Niklas Karlsson (Evolution Robotics), Joseph Lu (Honeywell), Tor Steinar Schei (Cybernetica), Stefan Rönnbäck (Optimation), han compartido generosamente sus conocimientos e ideas con nosotros. Estamos particularmente agradecidos a Peter Hansen, anteriormente en Foxboro, quien leyó el manuscrito completo y nos proporcionó muy buena realimentación

Estamos muy agradecidos a Leif Andersson quien hizo el formateado del texto y nos ayudó mucho con TeX, a Agneta Tuszyński que tradujo una gran parte del texto a LATEX, y a Eva Dagnegård que dibujó algunas de las figuras.

Finalmente, nos gustaría agradecer al Consejo de Investigación Sueco (VR), a la Agencia Sueca para la Innovación de Sistemas (VINNOVA), y a la Fundación Sueca para la Investigación Estratégica (SSF) quienes han financiado nuestra investigación durante muchos años.

KARL JOHAN ÅSTRÖM
TORE HÄGGLUND

Departamento de Control Automático
Instituto de Tecnología de Lund
Box 118, SE-221 00 Lund, Suecia

karl_johan.astrom@control.lth.se
tore.hagglund@control.lth.se

Índice general

Prefacio	V
1. Introducción	1
1.1 Introducción	1
1.2 Realimentación	3
1.3 Formas simples de realimentación	4
1.4 Cómo se desarrolló el controlador PID	6
1.5 Cambios de tecnología y transferencia de conocimiento	7
1.6 Perfil del contenido del libro	9
1.7 Resumen	11
1.8 Notas y referencias	11
2. Modelos de procesos	13
2.1 Introducción	13
2.2 Modelos estáticos	14
2.3 Modelos dinámicos	15
2.4 Modelos basados en características	24
2.5 Modelos de procesos típicos	30
2.6 Modelos para perturbaciones	47
2.7 Cómo obtener los modelos	50
2.8 Reducción del modelo	60
2.9 Resumen	64
2.10 Notas y referencias	65
3. Control PID	67
3.1 Introducción	67
3.2 El controlador PID	67
3.3 Filtraje de la acción derivativa	76
3.4 Ponderación del punto de consigna	77
3.5 Integrador windup	80
3.6 ¿Cuándo se puede utilizar el control PID?	91
3.7 Resumen	97
3.8 Notas y referencias	97
4. Diseño del controlador	101
4.1 Introducción	101
4.2 Una rica variedad de problemas de control	102

Índice general

4.3	Fundamentos de la realimentación	102
4.4	Estabilidad	108
4.5	Polos y ceros en lazo cerrado	116
4.6	Las funciones de sensibilidad	118
4.7	Robustez frente a variaciones del proceso	125
4.8	Cuantificando los requisitos	129
4.9	Especificaciones clásicas	135
4.10	Resumen	144
4.11	Notas y referencias	145
5.	Diseño feedforward	147
5.1	Introducción	147
5.2	Respuesta mejorada del punto de consigna	147
5.3	Ponderación del punto de consigna	152
5.4	Feedforward neutro	154
5.5	Respuesta rápida del punto de consigna	158
5.6	Atenuación de la perturbación	162
5.7	Resumen	165
5.8	Notas y referencias	166
6.	Diseño PID	167
6.1	Introducción	167
6.2	Método de Ziegler-Nichols y procedimientos relacionados	168
6.3	Sintonía empírica basada en reglas	179
6.4	Asignación de polos	184
6.5	Sintonía lambda	197
6.6	Diseño algebraico	200
6.7	Métodos de optimización	207
6.8	Deformación robusta del lazo	219
6.9	Resumen	234
6.10	Notas y referencias	235
7.	Una sustitución a Ziegler-Nichols	239
7.1	Introducción	239
7.2	El conjunto de prueba	241
7.3	Control PI	242
7.4	Control PID	244
7.5	Métodos de respuesta en frecuencia	253
7.6	Control PID basado en un modelo de segundo orden	257
7.7	Comparación de los métodos	262
7.8	Ruido de medida y filtrado	267
7.9	Desintonía	269
7.10	Resumen	278
7.11	Notas y referencias	282
8.	Control predictivo	283
8.1	Introducción	283
8.2	El predictor de Smith	284
8.3	Ánalisis del control del predictor de Smith	289
8.4	El controlador PPI	296
8.5	Predictores para procesos integradores	301

8.6	Control predictivo basado en modelo	303
8.7	Resumen	310
8.8	Notas y referencias	310
9.	Sintonía automática y adaptación	313
9.1	Introducción	313
9.2	Conocimiento del proceso	314
9.3	Técnicas adaptativas	315
9.4	Métodos basados en modelo	319
9.5	Métodos basados en reglas	322
9.6	Supervisión de controladores adaptativos	325
9.7	Sintonía de realimentación iterativa	334
9.8	Productos comerciales	337
9.9	Resumen	349
9.10	Notas y referencias	350
10.	Evaluación del lazo y del rendimiento	353
10.1	Introducción	353
10.2	Válvulas	353
10.3	Evaluación del lazo	359
10.4	Evaluación del rendimiento	360
10.5	Sintonía y diagnosis integradas	370
10.6	Resumen	370
10.7	Notas y referencias	371
11.	Interacción	373
11.1	Introducción	373
11.2	Interacción de lazos sencillos	374
11.3	Desacoplo	380
11.4	Sistemas paralelos	387
11.5	Resumen	391
11.6	Notas y referencias	392
12.	Paradigmas de control	393
12.1	Introducción	393
12.2	Enfoques ascendente y descendente	394
12.3	Control repetitivo	395
12.4	Control en cascada	399
12.5	Control de rango medio y de rango partido	406
12.6	Elementos no lineales	409
12.7	Control de red neuronal	417
12.8	Control borroso	421
12.9	Estructurando el sistema	427
12.10	Resumen	434
12.11	Notas y referencias	435
13.	Implementación	436
13.1	Introducción	436
13.2	Implementaciones analógicas	437
13.3	Implementaciones en computador	442
13.4	Algoritmos de velocidad	450
13.5	Aspectos operacionales	452

Índice general

13.6	Salidas del controlador	458
13.7	Resumen	462
13.8	Notas y referencias	462
Bibliografía	464

1

Introducción

1.1 Introducción

La idea de la realimentación es engañosamente simple y sin embargo extremadamente potente. La realimentación puede reducir los efectos de las perturbaciones, puede hacer que un sistema sea insensible a las variaciones del proceso y puede lograr que un sistema siga fielmente a sus señales de entrada. La realimentación ha tenido también una influencia profunda sobre la tecnología. La aplicación del principio de realimentación ha producido grandes avances en los campos del control, comunicación e instrumentación. Muchas patentes han sido concedidas a esta idea.

El controlador PID es una implementación simple de la idea de realimentación. Tiene la capacidad de eliminar errores en estado estacionario mediante la acción integral, y puede anticipar el futuro con la acción derivativa. Los controladores PID, o incluso los controladores PI, son suficientes para muchos problemas de control, particularmente cuando las dinámicas del proceso son benignas y los requisitos de comportamiento son modestos. Los controladores PID se encuentran en un gran número en todas las industrias y se presentan de muchas formas diferentes. Hay sistemas integrados en un dispositivo apto para uno o unos pocos lazos de control. El controlador PID es un elemento clave de los sistemas para control de motores. El controlador PID es un ingrediente importante de los sistemas distribuidos para control de procesos. Los controladores están también embebidos en muchos sistemas de control de propósito especial. Se encuentran en sistemas tan diversos como los reproductores de CD y DVD, el control de velocidad de los coches, y los microscopios de fuerza atómica. En control de procesos, más del 95 por ciento de los lazos de control son del tipo PID; la mayoría de los lazos son realmente control PI. Muchas características útiles del control PID no han sido ampliamente disseminadas porque han sido consideradas secretos comerciales. Ejemplos típicos son las técnicas para conmutación de modos y la desaturación del término integral (anti windup).

El control PID se combina a menudo con cierta capacidad lógica, funciones secuenciales, selectores y bloques de funciones sencillos para construir así los complicados sistemas de automatización utilizados en la producción de energía, transporte y procesos de fabricación. Muchas estrategias de control sofisticadas,

tales como el control predictivo, se organizan también jerárquicamente. El control PID se emplea en el nivel inferior de la jerarquía; el controlador multivariable da los puntos de consigna a los controladores en el nivel inferior. Se puede decir pues que el controlador PID es el «pan y mantequilla» de la ingeniería de control. Es un componente importante en la caja de herramientas de todo ingeniero de control.

Los controladores PID han sobrevivido a muchos cambios de tecnología, que van desde la neumática a los microprocesadores pasando por los tubos de vacío, los transistores, y los circuitos integrados. El microprocesador ha tenido una influencia crucial sobre el controlador PID. Prácticamente todos los controladores PID fabricados hoy día se basan en microprocesadores. Esto ha creado oportunidades para proporcionar características adicionales tales como sintonía automática, planificación de ganancia, adaptación continua y diagnóstico. La mayoría de los nuevos controladores PID que se producen en la actualidad tienen alguna capacidad para la sintonía automática. Sintonía y adaptación se pueden hacer de muchas formas diferentes. El controlador PID se ha convertido de hecho en un banco de pruebas para muchas nuevas ideas en control. Ha habido también un renacimiento de la implementación analógica en sistemas micro-mecánicos ya que requieren menos superficie de silicio que las realizaciones digitales. El controlador PID se realiza también utilizando FPGA (field programmable gate arrays) en aplicaciones donde se precisa una acción de control muy rápida.

Un gran número de ingenieros de instrumentación y proceso están familiarizados con el control PID. Hay una práctica bien establecida de instalar, sintonizar y utilizar estos controladores. A pesar de esto existe aún un potencial sustancial para mejorar el control PID. Una prueba de esto se puede encontrar en las salas de control de cualquier industria. Muchos controladores se ponen en modo manual, y entre los controladores que están operando en modo automático, la acción derivativa se desconecta frecuentemente por la simple razón de que es difícil de sintonizar adecuadamente. Las razones claves para un pobre comportamiento son problemas en la válvula y los sensores del equipo, restricciones en el proceso y una mala sintonía. Los problemas de la válvula incluyen un dimensionamiento inadecuado, histéresis y fricción. Los problemas en la medida se deben a utilizar un filtro anti aliasing mal diseñado o incluso a no emplearlo; a una excesiva acción de filtrado en sensores «inteligentes», a la existencia de mucho ruido, o a una mala calibración. Es posible conseguir mejoras sustanciales. El incentivo para seguir progresando se incrementa por las demandas de mejora en la calidad, que se impone por estándares tales como ISO 9000. Entendimiento y comprensión son los elementos claves para un comportamiento óptimo del lazo de control. Se requiere un conocimiento específico del proceso así como del control PID.

Basándonos en nuestra experiencia creemos que está emergiendo una nueva era del control PID. Este libro tendrá en cuenta el desarrollo realizado, evaluará su potencial, e intentará acelerarlo al compartir nuestras experiencias en este excitante y útil campo del control automático. El objetivo del libro es proporcionar el fundamento técnico necesario para comprender el control PID.

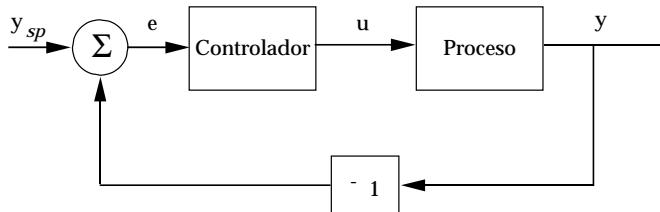


Figura 1.1 Diagrama de bloques de un proceso con un controlador por realimentación.

1.2 Realimentación

En la Figura 1.1 se muestra un sistema de realimentación sencillo mediante un diagrama de bloques. El sistema tiene dos grandes componentes, el proceso y el controlador, representados como cajas con flechas que denotan la relación causal entre entradas y salidas. El proceso tiene una entrada, la variable manipulada (MV), también llamada variable de control. Se denota por u . La variable de control influye sobre el proceso vía un actuador, que suele ser una válvula o un motor. La salida del proceso se llama la variable de proceso (PV) y se representa por y . Esta variable se mide con un sensor. En la Figura 1.1 el actuador y el sensor se consideran parte del bloque etiquetado «Proceso». El valor deseado de la variable de proceso se llama el punto de consigna (SP) o valor de referencia. Se denota por y_{sp} . El error de control e es la diferencia entre el punto de consigna y la variable de proceso, i.e., $e = y_{sp} - y$.

Supóngase por simplicidad que el proceso es tal que la variable de proceso aumenta cuando se incrementa la variable manipulada. El principio de realimentación se puede expresar como sigue:

Aumentar la variable manipulada cuando el error es positivo, y disminuirla cuando el error es negativo.

Este tipo de realimentación se llama *realimentación negativa* porque la variable manipulada se mueve en dirección opuesta a la variable de proceso $e = y_{sp} - y$.

El controlador PID es con diferencia la forma más común de realimentación. Este tipo de controlador ha sido desarrollado a lo largo de un gran período de tiempo y ha sobrevivido a muchos cambios de tecnología, de la mecánica y la neumática a la basada en la electrónica y en los computadores. Conocer algo de esto es útil con el fin de comprender sus propiedades básicas tal como se analiza en la Sección 1.4.

Algunas propiedades de la realimentación se pueden comprender intuitivamente a partir de la Figura 1.1. Si la realimentación funciona bien el error será pequeño, e idealmente será cero. Cuando el error es pequeño la variable de proceso está también próxima al punto de consigna independientemente de las propiedades del proceso. Para conseguir realimentación es necesario tener sensores y actuadores apropiados que efectúen las acciones de control.

La realimentación tiene algunas propiedades interesantes y útiles.

- La realimentación puede reducir los efectos de las perturbaciones
- La realimentación puede hacer que un sistema sea insensible a las varia-

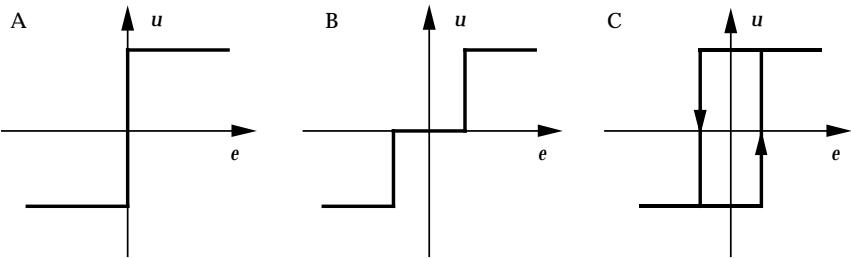


Figura 1.2 Características del controlador para control on-off ideal (A), y modificaciones con zona muerta (B) e histéresis (C).

ciones del proceso

- La realimentación puede crear relaciones bien definidas entre variables en un sistema

1.3 Formas simples de realimentación

Muchas de las propiedades agradables de la realimentación se pueden lograr con controladores simples. En esta sección presentaremos algunas formas simples de realimentación, es decir, control on-off, control proporcional, control integral, y control PID.

Control on-off

La realimentación se puede realizar de muchas formas diferentes. Un mecanismo de realimentación simple se puede describir como

$$u = \begin{cases} u_{\max}, & \text{if } e > 0 \\ u_{\min}, & \text{if } e < 0, \end{cases} \quad (1.1)$$

donde $e = y_{sp} - y$ es el error de control. Esta ley de control implica que siempre se utiliza una acción correctora máxima. Este tipo de realimentación se llama *control on-off*. Es simple y no hay parámetros que elegir. El control on-off a menudo funciona al mantener la variable de proceso próxima al punto de consigna, pero generalmente resultará en un sistema donde las variables oscilan. Obsérvese que en la Ecuación 1.1 la variable de control no está definida cuando el error es cero. Es común realizar algunas modificaciones bien introduciendo una histéresis o una zona muerta (ver Figura 1.2).

Control proporcional

La razón de por qué el control on-off a menudo da lugar a oscilaciones es que el sistema sobreactúa, ya que un pequeño cambio en el error hará que la variable manipulada varíe entre los valores máximos. Este efecto se evita en el control proporcional, donde la característica del controlador es proporcional al error de

control para pequeños errores. Esto se puede conseguir haciendo la señal de control proporcional al error

$$u = K(y_{sp} - y) = Ke, \quad (1.2)$$

donde K es la ganancia del controlador.

Control integral

El control proporcional tiene la desventaja de que la variable de proceso a menudo se desvía del punto de consigna. esto se puede evitar haciendo la acción de control proporcional a la integral del error

$$u(t) = k_i \int_0^t e(\tau) d\tau, \quad (1.3)$$

donde k_i es la ganancia integral. Esta estrategia se llama control integral. El control integral tiene una propiedad sorprendente. Supóngase que hay un estado estacionario con un error constante e_0 y una señal de control constante u_0 . Se sigue de la ecuación anterior que

$$u_0 = k_i e_0 t.$$

Como u_0 es una constante se sigue que e_0 debe ser cero. Encontramos así que si hay un estado estacionario y un controlador tiene acción integral, el error en estado estacionario es siempre cero. Esto es también cierto para el controlador PI

$$u(t) = Ke(t) + k_i \int_0^t e(\tau) d\tau. \quad (1.4)$$

Esta es una de las razones de por qué los controladores PI son tan comunes.

Control PID

Un refinamiento adicional consiste en dotar al controlador de una capacidad anticipativa utilizando una predicción de la salida basada en una extrapolación lineal. Ver Figura 1.3. Esto se puede expresar matemáticamente como

$$u(t) = K \left(e(t) + \frac{1}{T_i} \int_0^t e(\tau) d\tau + T_d \frac{de(t)}{dt} \right). \quad (1.5)$$

La acción de control es así una suma de tres términos que representan el pasado por la acción integral del error (el término-I), el presente (el término-P) y el futuro por una extrapolación lineal del error (el término-D). El término $e + T_d \frac{de}{dt}$ es una predicción lineal del error T_d unidades de tiempo en el futuro. Los parámetros del controlador se llaman: ganancia proporcional K , tiempo integral T_i , y tiempo derivativo T_d .

Se ha comprobado empíricamente que el controlador PID es capaz de resolver un amplio espectro de problemas de control. Hay controladores más complejos que difieren del controlador PID porque utilizan métodos más sofisticados para la predicción.

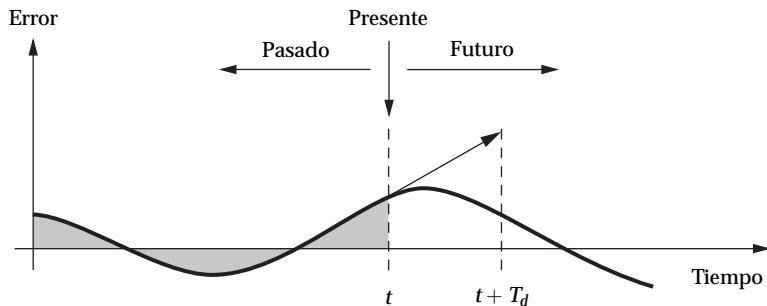


Figura 1.3 Un controlador PID calcula su acción de control basándose en los errores de control pasados, presentes, y futuros.

1.4 Cómo se desarrolló el controlador PID

El controlador PID se ha desarrollado a lo largo de un período de tiempo que se extiende al menos 250 años. Es útil tener alguna perspectiva de este desarrollo en orden a comprender muchos de los temas. La tecnología utilizada para implementarlos, naturalmente, ha cambiado significativamente a lo largo de los años. Los primeros controladores fueron dispositivos mecánicos (reguladores centrífugos) empleados para controlar molinos de viento y máquinas de vapor. La medida de la velocidad angular se combinaba con la actuación de las válvulas. Una gran dosis de inteligencia se empleó en idear la acción integral.

Un cambio importante ocurrió en conexión con el desarrollo del control de procesos industriales. Las funciones de medida, control, y actuación fueron entonces separadas y se construyeron dispositivos especiales que efectuaban las acciones de control representadas por la Ecuación 1.5. Una característica interesante era que la transmisión de señal y el cálculo se hacían neumáticamente. Un gran avance ocurrió cuando los tubos utilizados para transmitir la presión y los niveles de presión se normalizaron a 3–15 PSI. Esto permitió combinar sensores, controladores, y actuadores de diferentes fabricantes. También hizo posible concentrar los controladores en salas de control que están localizadas lejos de los sensores y actuadores. Mucho ingenio se utilizó otra vez para desarrollar estos controladores. El uso de la realimentación en los propios controladores fue un gran logro. De esta forma era posible obtener acción lineal fuera del rango de los componentes que tenían unas características fuertemente no lineales.

Comenzando en los años 50, pronto estuvieron disponibles las versiones electrónicas de los controladores PID. Las acciones de control representadas por la Ecuación 1.5 se obtuvieron mediante un simple computador analógico basado en amplificadores operacionales. La transmisión de señal se normalizó también en el rango de 4–20 mA. Representar el cero mediante una corriente no nula era útil para realizar diagnósticos.

Todavía sucedió otro avance cuando se utilizaron computadores digitales para implementar los controladores. En primer lugar cuando emergió el control por computador se emplearon sistemas fuertemente centralizados, ya que el cálculo digital está justificado desde un punto de vista de coste económico en grandes

sistemas. Con la aparición de los microprocesadores en los años 70 incluso los controladores simples se implementaron utilizando computadores. Cuando se emplea un computador digital es también factible añadir muchas funciones tales como sintonía automática, adaptación y diagnóstico. Esta es un área de un desarrollo muy activo.

Hoy día estamos experimentando otro desplazamiento en tecnología. Las realizaciones analógicas están volviendo a reaparecer en sistemas eléctricos micromecánicos (MEMS), y los controladores digitales se implementan también usando FPGA (field programmable gate arrays), que admiten períodos de muestreo muy cortos. Las FPGA difieren significativamente de los computadores digitales ya que son altamente paralelas.

En la actualidad encontramos controladores PID en muchos formatos. Hay controladores dedicados que pueden controlar uno o unos pocos lazos. Las funciones PID se encuentran en los automátas programables que fueron originalmente diseñados para sustituir a los relés. Hay sistemas que contienen muchos controladores PID implementados en computadores que van desde pequeños sistemas para unas pocas docenas de lazos a grandes sistemas distribuidos para control de procesos. Los controladores PID se usan normalmente en sistemas dedicados para control de movimiento. Hay también un abanico de controladores especiales tales como autopilotos y sistemas de control para los reproductores de CD y DVD y memorias ópticas que se basan en el control PID.

1.5 Cambios de tecnología y transferencia de conocimiento

El controlador PID es un caso de estudio interesante para el control de la tecnología, debido a su larga historia y por los muchos cambios que ha experimentado. Puesto que hemos tenido experiencias personales de algunos de estos cambios de tecnología, presentaremos algunas reflexiones personales donde analizamos la creación y destrucción de conocimiento y el papel de personas y documentos que han sido claves.

Las transferencias de tecnología son a menudo abruptas y no planificadas. La razón de por qué una empresa decide cambiar de tecnología puede ser una reducción drástica en los costos de los equipos o la presión de la competencia y de los clientes. Un cambio en la tecnología con frecuencia significa que el grupo de I+D tiene que ser sustituido por nuevas personas que son familiares con la nueva tecnología, pero que no suelen serlo con la antigua. Esto significa que se corre un alto riesgo de que durante la transición se pierda información. Como la transferencia de tecnología debe hacerse con rapidez, hay también un riesgo elevado de no utilizar todo el potencial de la nueva tecnología.

Los primeros controladores de temperatura fueron del tipo on-off. Los controladores on-off eran sencillos y baratos, pero tienen oscilaciones que son inevitables. La amplitud de las oscilaciones se puede mantener en niveles razonables, ya que la dinámica de muchos sistemas térmicos está dominada por la constante de tiempo del proceso. Cuando la electrónica tuvo precios competitivos, hubo una transición del control on-off al control PID continuo. El desarrollo de los controladores PID analógicos está bien documentado en material público accesible de Eurotherm, que fue comenzado por investigadores de la universidad de

Manchester. Los controladores se desarrollaron basándose en un conocimiento sólido de modelado y control. La teoría ayuda, porque muchas aplicaciones de temperatura admiten ganancias elevadas y la acción derivativa puede ser muy beneficiosa. Se proporcionaron también reglas de sintonía y se desarrollaron esquemas de protección para el windup con el nombre de *desaturación del integrador* y *constantes de tiempo cruzadas*. El resultado del desarrollo fue una mejora drástica del comportamiento de los controladores de temperatura. Es interesante observar que pasó mucho tiempo antes de que el fenómeno interesante e importante de la saturación del integrador recibiese alguna atención por parte del mundo académico.

Cuando emergió el control de procesos basado en computador a comienzos de los años 60, el foco de interés estaba inicialmente en las funciones de control del nivel superior. Los controladores PID analógicos se utilizaban en el nivel de base y el computador suministraba los puntos de consigna a los controladores. Cuando se desarrollaron los sistemas, la atención se centró otra vez en el control PID donde muchos lazos PID se implementaban en un único computador, lo que se denominó control digital directo. El desarrollo técnico se dedicó a la discretización del algoritmo PID, una razón fue que los recursos de cálculo eran un cuello de botella. Sin embargo se prestó escasa atención al problema de la saturación del integrador y se dedicó alguna al filtraje de las señales de los sensores.

La aparición del microprocesador hizo que el cálculo digital estuviese disponible de forma muy barata para sistemas pequeños, un desarrollo que tuvo un gran impacto sobre el controlador PID. Esto dio como resultado controladores de un solo lazo, controladores para unos pocos lazos, y grandes sistemas distribuidos. El desarrollo fue lento por dos razones. Muchas personas nuevas sin experiencia previa del control analógico entraron en la escena, y muchos de los antiguos no estaban interesados en aprender la nueva tecnología. Aspectos importantes como la saturación del integrador y el filtrado no se documentaron de forma que estuviesen fácilmente accesibles. Por lo tanto, pasó algún tiempo antes de que se volviese a recuperar el conocimiento y la experiencia apropiada. Hubo también una tendencia a simplemente implementar viejas ideas en la nueva tecnología sin considerar las oportunidades que esta última ofrecía. Gradualmente, las capacidades potenciales del computador digital fueron explotadas al incorporar características tales como, autosintonía, adaptación y diagnóstico en los sistemas.

Cuando los sistemas de control distribuido (DCS) sustituyeron a los sistemas analógicos, se retuvo la arquitectura distribuida. Los controladores analógicos y los módulos de función fueron representados como bloques en los programas DCS. Esto fue probablemente una buena idea, pero no se utilizaban las oportunidades que se tenían por el hecho de que todas las señales estaban disponibles en un computador. Se tardó una década para que se presentasen los sistemas DCS que manejaban la saturación del término integral.

Un par de conclusiones que se pueden sacar son que es muy importante la documentación de los sistemas y tener personas de mentes abiertas que puedan reducir la brecha entre diferentes tecnologías. Cuando las nuevas tecnologías están ya disponibles es también útil parar y pensar para analizar cómo la nueva tecnología se puede explotar en lugar de rápidamente implementar viejas ideas con la nueva tecnología. Es también bueno filtrar la esencia de los sistemas an-

tiguos de forma que no se pierdan buenas características. Finalmente, es importante documentar las ideas, escribir libros, y asegurar que la información no sólo se transfiere de persona a persona, sino que tiene una amplia diseminación.

1.6 Perfil del contenido del libro

Se aconseja al lector que examine la tabla de contenido para ver la estructura global del libro. La dinámica del proceso es clave para comprender cualquier problema de control. El Capítulo 2 presenta conceptos que son útiles para describir el comportamiento de los procesos. Se mencionan brevemente los modelos estáticos, pero el interés principal del capítulo es sobre la dinámica de los procesos. Se dan representaciones en términos de la respuesta en el tiempo y en la frecuencia. Estas visiones duales son muy útiles para lograr una buena comprensión de la dinámica. A lo largo del texto se utilizan las nociones de respuesta en escalón y funciones de transferencia. Se analizan con detalle una serie de modelos típicos que se usan en el control PID. También se tratan los modelos para las perturbaciones así como las técnicas para la determinación experimental de los modelos.

En el Capítulo 3 se da una presentación en profundidad del controlador PID. Esto incluye, los fundamentos así como muchos detalles de su implementación, tales como el filtrado para proporcionar una atenuación rápida en altas frecuencias, la desaturación del término integral (anti windup), la mejora en la respuesta del punto de consigna, etc. El controlador PID se puede estructurar de diferentes formas. Las formas comúnmente utilizadas son las estructuras serie y paralelo. Se analizan con detalle las diferencias entre estas y los parámetros del controlador empleados. Se describen también las limitaciones del control PID. Casos típicos donde tienen interés controladores más complejos son los sistemas con gran retraso y los sistemas oscilatorios. Se exponen de forma breve las extensiones del control PID para tratar con tales sistemas.

El Capítulo 4 trata el diseño del controlador en general. Hay una rica variedad de problemas de control con objetivos muy diversos. El capítulo muestra una revisión de ideas y conceptos que son relevantes para el control PID. Se intenta así llevar el diseño de los controladores PID dentro de la corriente principal de diseño de sistemas de control. Se tratan temas tales como limitaciones fundamentales, estabilidad, robustez y especificaciones.

En el Capítulo 5 se trata el control feedforward, una técnica simple y potente que complementa a la realimentación. Se da un diseño sistemático de control feedforward para mejorar las respuestas a cambios en el punto de consigna así como una presentación del diseño de sistemas que siguen a un modelo. Se analiza con detalle el caso especial de la ponderación del punto de consigna y métodos para determinar los pesos del punto de consigna. El capítulo también muestra cómo se puede emplear el control feedforward para reducir el efecto de las perturbaciones que se pueden medir.

El Capítulo 6 describe métodos para el diseño de controladores PID. Se presentan muchos métodos diferentes para la sintonía de controladores PID que se han desarrollado a lo largo de los años. Se ha intentado mantener un equilibrio al proporcionar una perspectiva histórica y presentar métodos potentes.

Un método de diseño razonable debería considerar las perturbaciones de la

carga, las incertidumbres del modelo, el ruido en la medida y la respuesta frente a cambios en el punto de consigna. Una desventaja de muchas de las reglas de sintonía tradicionales para el control PID es que tales reglas no consideran todos estos aspectos de una forma equilibrada. En el Capítulo 7 se estudian nuevas técnicas de sintonía que sí consideran todos estos criterios.

El Capítulo 8 trata los controladores predictivos basados en modelos. En primer lugar se presenta y analiza el predictor de Smith, y modificaciones al mismo para tratar el caso de procesos con integrador. A continuación se analizan otros tipos de controladores predictivos basados en modelos, tales como el controlador MPC, el controlador de Dahlin-Higham, el controlador dinámico matricial y el control de varianza mínima.

En el Capítulo 9 discutimos algunas técnicas para la adaptación y la sintonía automática de controladores PID. Esto incluye métodos basados en modelos paramétricos y técnicas no paramétricas. También se exponen la supervisión de controladores adaptativos y la sintonía por realimentación iterativa. Se describen una serie de controladores comerciales para ilustrar las diferentes técnicas.

El Capítulo 10 trata de los métodos para la puesta en marcha, supervisión y diagnóstico de los lazos de control. Procedimientos para evaluar los lazos se emplean para investigar las propiedades del lazo de control, e.g. niveles de señales, no linealidades, y condiciones del equipo. Procedimientos de evaluación del comportamiento se usan para supervisar los lazos de control durante la operación y asegurar que cumplen las especificaciones.

El controlador PID se emplea normalmente como un controlador de lazo simple. En la práctica hay a menudo interacciones entre los lazos. En el Capítulo 11 se estudian algunas cuestiones claves acerca de lazos que interactúan que son de particular relevancia para el control. En particular se muestra que los parámetros del controlador en un lazo pueden tener un comportamiento significativo sobre la dinámica de otros lazos. Se introduce también el array de ganancias relativas de Bristol que es una forma sencilla de caracterizar las interacciones. Se expone el problema del emparejamiento de entradas y salidas y un método de diseño basado en el desacoplo, que es una extensión natural de los métodos de sintonía para sistemas de una entrada y una salida.

En el Capítulo 12 se muestra cómo se pueden resolver problemas de control complejos combinando diferentes formas controladores sencillos. Se analizan los paradigmas de control repetitivo, control en cascada, control de rango medio y de rango partido, control de relación y control con selectores. También se cubren de forma breve el uso de técnicas actualmente populares como las redes neuronales y el control borroso.

El Capítulo 13 presenta cuestiones de implementación relacionadas con el control PID. En primer lugar se da una breve revisión de las implementaciones neumáticas y con electrónica analógica. A continuación se expone una presentación detallada de aspectos de la implementación en un computador tales como muestreo, prefiltrado y discretización del algoritmo PID. Aspectos operacionales tales como la transferencia sin salto entre los modos manual y automático. El capítulo finaliza con un análisis acerca de las diferentes salidas del controlador que se tienen que usar dependiendo de cuál es el dispositivo de actuación que se emplea.

1.7 Resumen

En esta sección se ha dado una breve descripción del concepto de realimentación. La aplicación de la realimentación ha tenido un impacto muy interesante y algunas veces revolucionario. Se han expuesto algunas de sus propiedades útiles, su capacidad para reducir las perturbaciones, su insensibilidad a las variaciones del proceso, la linealidad entre el punto de consigna y la variable del proceso. Hemos descrito también brevemente algunas formas simples de realimentación tales como el control on-off y el control PID. Se ha analizado de forma sucinta el desarrollo de los controladores y finalmente se ha reseñado el contenido del libro.

1.8 Notas y referencias

Los controladores PID se utilizaron ampliamente en los primeros desarrollos del control desde 1870 hasta 1920. La forma moderna del controlador PID emergió con el desarrollo del control de procesos en los años 30 y 40, tal como se analiza en [45] y [46]. El controlador PID es todavía la herramienta estándar para resolver problemas de control industrial. En un estudio detallado del estado del arte en control de procesos industriales llevado a cabo por Electric Measuring Instrument Manufacturer en Japón en 1989 se encontró que más del 90 por ciento de los lazos de control eran del tipo PID; ver [365]. El trabajo [94] revisó la industria en EE.UU. Se ha descubierto que hay más de 8 millones de instalaciones en las industrias petroquímicas, papelera, energía, y del metal. Cada instalación tiene entre 500 y 5.000 lazos de control regulatorios, el 97 por ciento de ellos son del tipo PID. Los controladores PID se utilizan también en grandes cantidades en otras industrias. Las memorias ópticas para los reproductores de CD y DVD contienen tres lazos PID para el control de la velocidad de rotación, la focalización y el seguimiento de la pista. En el año 2002 se fabricaron cerca de 140 millones de unidades; ver [3]. Además, hay un gran número de controladores PID en el control de motores y en los sistemas de posicionamiento. Se puede pues decir con absoluta seguridad que el controlador PID es una de las herramientas más comunes en control.

El Control PID se estudia en la mayoría de los libros de texto sobre control de procesos tales como [221, 313, 229, 48, 302], y hay también libros que se focalizan sobre el control PID [239, 90, 329, 354, 287, 356, 257, 242].

La teoría de los controladores PID se basó durante mucho tiempo en técnicas especiales. Últimamente han habido esfuerzos para traer el control PID dentro de la corriente principal de la teoría de control. Un esfuerzo notable fue hecho en el año 2000 cuando la International Federation of Automatic Control (IFAC) preparó una reunión científica sobre el pasado, presente, y futuro del control PID; ver [287]. Una selección de trabajos de esta reunión se publicó también como un número especial de la revista Control Engineering and Practice. Los trabajos [47] y [21] dan una perspectiva del desarrollo del control PID.

Debido al gran número de controladores PID y a su uso extendido hay todavía beneficios significativos en mejorar la práctica del control PID. Tal mejora requiere prestar atención al lazo de control completo y no simplemente al propio controlador tal como se demuestra en el trabajo [50] que describe las auditorías de papeleras en Canadá. Una papelera típica tiene más de 2.000 lazos de control,

Capítulo 1. Introducción

el 97 por ciento de los lazos se basan en control PI. Se encontró que sólamente el 20 por ciento de los lazos de control operaban bien y disminuían la variabilidad del proceso. Las razones de por qué el comportamiento es pobre son una mala sintonía (30 por ciento) y problemas en las válvulas (30 por ciento). El restante 20 por ciento de los controladores no funcionaban correctamente a causa de una serie de razones tales como; problemas con el sensor, mala elección del período de muestreo, y pobre o incluso no utilización de filtros de anti aliasing. Observaciones similares se dan en [107], donde se estima que el 30 por ciento de los controladores de proceso operan en modo manual, que el 20 por ciento de los lazos utilizan los parámetros por defecto fijados por el fabricante del controlador (la denominada «sintonía de fábrica»), y que el 30 por ciento de los lazos funcionan pobemente a causa de problemas con el equipo en las válvulas y los sensores.

2

Modelos de procesos

2.1 Introducción

Para describir la conducta de los procesos se suelen utilizar modelos matemáticos. Los modelos proporcionan una forma unificada para tratar con sistemas de tipos muy diferentes, que hacen posible introducir una serie de conceptos útiles. Los modelos son también esenciales para la simulación y el diseño del control. En este capítulo revisaremos algunos de los modelos que se emplean normalmente en el control PID. Los modelos intentan capturar algunos aspectos del proceso que son relevantes para el control. Se usan muchos tipos diferentes de modelos.

La conducta en estado estacionario de un proceso se puede obtener mediante una función que proporciona el valor en estado estacionario de la variable de proceso para valores dados de la variable manipulada. Tales modelos se analizan en la Sección 2.2.

Para controlar un sistema es necesario conocer la dinámica del proceso. Para la finalidad del control es a menudo suficiente describir pequeñas desviaciones respecto de un estado de equilibrio. En este caso la conducta se puede modelar como un sistema dinámico lineal. Este es un campo muy rico con muchos conceptos y herramientas útiles, que forman el núcleo de la teoría de control. En la Sección 2.3 se exponen diferentes formas de describir la dinámica de un proceso. Se introducen las ideas de respuesta transitoria y respuesta en frecuencia así como los conceptos importantes de respuesta escalón, respuesta impulsional y función de transferencia.

En el control PID han sido utilizadas tradicionalmente técnicas especiales para modelar la dinámica del proceso. La idea es representar la dinámica del proceso mediante unas pocas características. Esto se analiza en la Sección 2.4, donde se introducen características tales como tiempo de residencia medio, retardo aparente, constante de tiempo aparente, retardo normalizado, ganancia última, frecuencia última y razón de ganancia.

En la Sección 2.5 estudiamos algunos modelos particulares que son ampliamente utilizados en el control PID. Estos modelos se introducen en términos de sus funciones de transferencia. En esa sección se analiza también el concepto importante de normalización, así como las no linealidades. Los ejemplos tratados en la Sección 2.5 serán ampliamente usados en el libro.

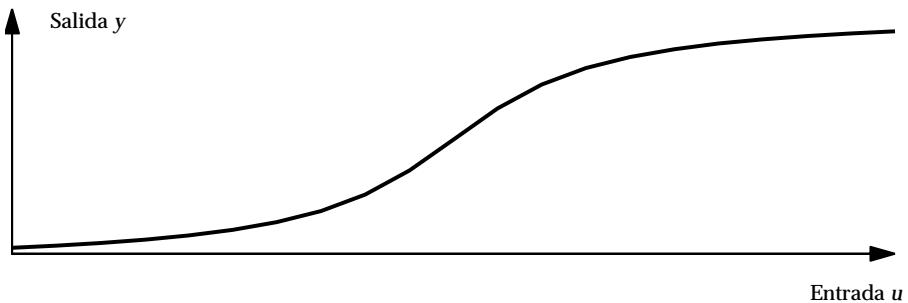


Figura 2.1 Característica de proceso estática, que muestra la salida del proceso y como función de la entrada del proceso u bajo condiciones de estado estacionario.

Las perturbaciones son un aspecto importante de un problema de control. En la Sección 2.6 presentamos algunos modelos que se emplean para describir perturbaciones. La Sección 2.7 describe métodos simples para obtener modelos, y la Sección 2.8 expone algunas técnicas usadas para simplificar un modelo complicado. En la Sección 2.9 se resume el contenido del capítulo, y en la Sección 2.10 se dan las referencias.

2.2 Modelos estáticos

Es natural comenzar describiendo la conducta estacionaria del proceso. Esto se puede hacer mediante una curva que muestra el valor en estado estacionario de la variable del proceso y (la salida) para diferentes valores de la variable manipulada u (la entrada); ver Figura 2.1. Esta curva se llama un modelo estático o una característica de proceso estática. Todas las investigaciones del proceso deberían comenzar con una determinación del modelo del proceso estático. Puede emplearse para determinar el rango de señales de control requeridas para cambiar la salida del proceso en el rango deseado, para dimensionar los actuadores y para seleccionar la resolución del sensor. La pendiente de la curva en la Figura 2.1 nos dice cuánto cambia la variable de proceso para pequeños cambios en la variable manipulada. Esta pendiente se llama la ganancia estática del proceso. Grandes variaciones en la ganancia indican que el problema de control puede ser difícil.

El modelo estático se puede obtener experimentalmente de algunas formas. Una manera natural es mantener la entrada en un valor constante y medir la salida en estado estacionario. Esto da un punto de la característica del proceso. El experimento se repite para cubrir el rango completo de entradas.

Un procedimiento alternativo es hacer un experimento en lazo cerrado donde la salida del sistema se mantiene constante mediante realimentación y se mide el valor en estado estacionario de la entrada.

Los experimentos requeridos para determinar el modelo del proceso estático a menudo dan una buena idea intuitiva acerca de cuán fácil es el control de un proceso y si hay muchas perturbaciones. Los datos para los modelos en estado estacionario se pueden obtener de medidas tomadas en línea.

Algunas veces las operaciones del proceso no permiten que los experimentos puedan hacerse tal como hemos descrito. Normalmente se permiten pequeñas perturbaciones, pero puede no ser posible mover el proceso a lo largo de su rango completo de operación. En tales casos el experimento se debe hacer durante un largo período de tiempo. Es posible proporcionar un sistema de control con capacidades para automáticamente determinar el modelo estático del proceso durante su operación normal; ver Capítulo 10.

2.3 Modelos dinámicos

Un modelo de proceso estático como el analizado en la sección anterior nos da la relación en estado estacionario entre las señales de entrada y salida. Un modelo dinámico debería proporcionar la relación entre la entrada y la salida durante el transitorio. Es obviamente mucho más difícil capturar la conducta dinámica. Sin embargo, esto es, esencial cuando se trata con problemas de control.

Caracterización cualitativa de dinámicas de procesos

Antes de intentar modelar un sistema con frecuencia es útil obtener una caracterización cruda de su conducta dinámica. Para describir la conducta dinámica simplemente mostraremos la respuesta del sistema a un cambio en escalón en la variable manipulada. Esto se llama la respuesta en escalón del sistema o la curva de reacción del proceso,

Hay que hacer una distinción entre sistemas estables e inestables. La respuesta en escalón de un sistema estable va a un valor constante. Un sistema inestable no alcanzará un estado estacionario después de un cambio en escalón. Los sistemas que tienen un integrador son un ejemplo típico de un sistema inestable. En los primeros textos y trabajos sobre control de procesos, los sistemas estables se llamaban sistemas auto regulados.

A partir de la respuesta en escalón se pueden obtener muchas propiedades de un sistema. La Figura 2.2 muestra las respuestas en escalón que se encuentran típicamente en control de procesos.

En la Figura 2.2a, la salida del proceso se cambia monotónicamente a un nuevo valor estacionario. Éste es el tipo más común de respuesta en escalón que se encuentra en control de procesos. En la Figura 2.2b, la salida del proceso oscila alrededor de su valor final estacionario. Este tipo de proceso es poco común en control de procesos. Un caso donde ocurre es en el control de la concentración de fluidos en recirculación. En diseños mecánicos, sin embargo, son usuales los procesos oscilantes cuando se usan materiales elásticos, e.g., ejes flexibles en servos, construcciones con muelles, etc. Los sistemas de las Figuras 2.2a y 2.2b son estables, mientras que los sistemas que se muestran en las Figuras 2.2c y 2.2d son inestables. El sistema de la Figura 2.2c es un proceso integrador. Ejemplos de procesos integradores son el control de nivel, control de presión en un recipiente cerrado, control de concentración en procesos por lotes y el control de temperatura en cámaras bien aisladas. El factor común en todos estos procesos es que en ellos ocurre alguna clase de almacenamiento. En el control de nivel, presión y concentración se produce un almacenamiento de masa, mientras que en el caso

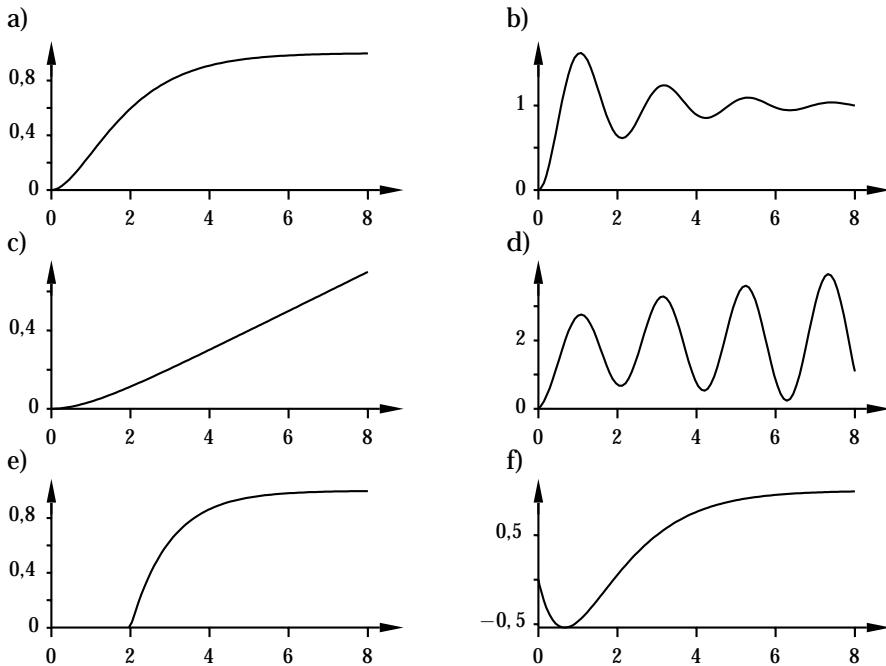


Figura 2.2 Respuestas en escalón en lazo abierto.

del control de temperatura hay un almacenamiento de energía. El sistema de la Figura 2.2e tiene un gran retardo de tiempo. El retardo de tiempo ocurre cuando hay retardo de transporte en el proceso. El sistema en la Figura 2.2f es un sistema de fase no mínima. Obsérvese que la salida inicialmente se mueve en la dirección equivocada. El nivel de agua en las calderas a menudo reacciona de esta forma después de un cambio en escalón en el caudal del agua de alimentación.

Sistemas lineales invariantes en el tiempo

Hay una clase restringida de modelos, llamados sistemas lineales invariantes en el tiempo, que se pueden usar a menudo. Tales modelos describen la conducta de sistemas para pequeñas desviaciones del equilibrio. Invariante en el tiempo significa que la conducta del sistema no cambia con el tiempo. Linealidad quiere decir que se cumple el principio de superposición. Esto significa que si la entrada u_1 da la salida y_1 y la entrada u_2 da la salida y_2 entonces se sigue que la entrada $au_1 + bu_2$ da la salida $ay_1 + by_2$.

Una propiedad agradable de los sistemas lineales invariantes en el tiempo es que su respuesta a una entrada arbitraria se puede caracterizar completamente en términos de la respuesta a una señal simple. Se pueden emplear muchas señales diferentes para caracterizar a un sistema. En términos generales, podemos diferenciar entre las respuestas transitoria y en frecuencia.

En un sistema de control típicamente nos concentraremos solamente sobre dos señales, la señal de control y la variable medida. La dinámica de procesos trata de la relación entre estas señales. Esto significa que se incluye dinámica en los actuadores, proceso y sensores. Estas dinámicas están con frecuencia dominadas

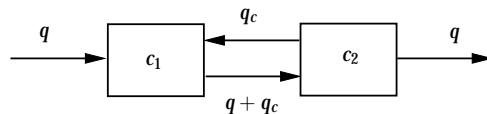


Figura 2.3 Diagrama esquemático de un sistema consistente en dos tanques.

por la dinámica del proceso. En algunos casos, sin embargo, son los sensores y los actuadores los que dan la mayor contribución a la dinámica. Por ejemplo, es muy común que existan filtros con grandes constantes de tiempo en los sensores de temperatura. También pueden haber ruido de medida y otras imperfecciones. Puede haber también una dinámica significativa en los actuadores. Para realizar una buena tarea de control, es necesario tener conocimiento del origen físico de la dinámica del proceso para juzgar si una buena respuesta en la variable medida realmente se corresponde con una buena respuesta de la variable física del proceso. Incluso si la atención se centra sobre la variable medida conviene siempre tener presente que la variable de proceso es la señal que realmente importa.

Modelado físico – Ecuaciones diferenciales

Una forma tradicional de obtener un modelo del proceso es usar las leyes físicas básicas tales como balances de masa, momentos y energía. Tales descripciones típicamente conducen a un modelo matemático en términos de ecuaciones diferenciales. Ilustramos esto con dos ejemplos.

EJEMPLO 2.1—TANQUE AGITADO

Considérese un reactor de un tanque agitado ideal. Sea V el volumen del reactor y q el caudal volumétrico a través del reactor. La variable manipulada es la concentración u del caudal de entrada, y la variable de proceso y es la concentración en el reactor. Un balance de masas para el reactor da

$$V \frac{dy}{dt} = q(u - y).$$

El parámetro $T = V/q$, que tiene dimensiones de tiempo, es el tiempo de residencia medio de las partículas que entran al reactor. Se llama también la constante de tiempo del sistema. \square

El sistema del Ejemplo 2.1 es de primer orden porque sólo se necesita una variable para dar cuenta del almacenamiento en el tanque. Esto es posible porque el tanque está bien agitado de forma que la concentración se puede suponer constante en todo el volumen. En casos más complicados se requieren muchas variables para tomar en cuenta el almacenamiento de masa, energía y momento. Esto se ilustra en el ejemplo siguiente.

EJEMPLO 2.2—TANQUES ACOPLADOS

Sea el sistema que se muestra en la Figura 2.3, que está compuesto de dos tanques bien agitados. Suponga que cada tanque tiene un volumen V , que los caudales de entrada y de salida son q , y que el caudal de recirculación es qc . Más aún, sea la concentración del caudal de entrada la entrada $u = c_{in}$, y la concentración del

caudal de salida la salida, $y = c_{out}$. Cuando los tanques están bien agitados el balance de masa se puede caracterizar por las concentraciones en los tanques. Dicho balance para el tanque es

$$\begin{aligned} V \frac{dc_1}{dt} &= -(q + q_c)c_1 + q_c c_2 + qu \\ V \frac{dc_2}{dt} &= (q + q_c)c_1 - (q + q_c)c_2 \\ y &= c_2. \end{aligned}$$

□

El modelo en el ejemplo consiste en dos ecuaciones diferenciales de primer orden. Hay dos ecuaciones diferenciales porque el sistema se describe completamente mediante balances de masas y el almacenamiento de las masas se puede definir mediante dos variables. Se obtienen descripciones similares para sistemas más complicados, pero el número de ecuaciones aumenta con la complejidad del sistema. La ecuación diferencial puede ser también no lineal si hay fenómenos de transporte no lineales.

El modelo en el Ejemplo 2.2 consiste en un sistema de ecuaciones diferenciales de primer orden. Si sólo estamos interesados en las relaciones entre la entrada u y también se puede describir la salida y un modelo lineal por una ecuación diferencial de orden más alto, i.e.,

$$\frac{d^n y}{dt^n} + a_1 \frac{d^{n-1}y}{dt^{n-1}} + \dots + a_n y = b_1 \frac{d^{n-1}u}{dt^{n-1}} + \dots + b_n u. \quad (2.1)$$

El número n es igual al número de variables necesitadas para dar cuenta del almacenamiento. Éste es uno de los modelos estándar utilizados en control automático.

La ecuación diferencial (2.1) se caracteriza por dos polinomios

$$\begin{aligned} a(s) &= s^n + a_1 s^{n-1} + \dots + a_n \\ b(s) &= b_1 s^{n-1} + \dots + b_n, \end{aligned} \quad (2.2)$$

donde el polinomio $a(s)$ se llama el polinomio característico. Los ceros del polinomio $a(s)$ se llaman los *polos* del sistema, y los ceros del polinomio $b(s)$ se llaman los *ceros* del sistema.

La ecuación diferencial (2.1) tiene una solución de la forma

$$y(t) = \sum k C_k(t) e^{\alpha_k t} + \int_0^t g(t-\tau) u(\tau) d\tau, \quad (2.3)$$

donde α_k son los polos del sistema y $C_k(t)$ son polinomios (constantes si los polos son distintos). El primer término de la ecuación anterior depende de las condiciones iniciales y el segundo de la entrada. La función g tiene la misma forma que el primer término del lado derecho de (2.3). Los polos proporcionan así una idea cualitativa útil de las propiedades del sistema.

En situaciones más complicadas puede ser más difícil tener en cuenta el almacenamiento de masa, momento y energía. Ilustramos esto con un ejemplo sencillo.

EJEMPLO 2.3—RETARDO DE TIEMPO

Sea un sistema donde la masa se transporta a lo largo de una cinta transportadora. Sea la entrada $u(t)$ el caudal mísico de entrada en la cinta, y sea la salida $y(t)$ el caudal mísico de salida en la cinta. La relación entrada-salida para el sistema es entonces

$$y(t) = u(t - L), \quad (2.4)$$

donde L es el tiempo que tarda una partícula en recorrer la cinta. Para tomar en cuenta el almacenamiento de masa en la cinta es necesario especificar la distribución de masa en la cinta. La salida es así simplemente la entrada retardada. Este sistema se llama por lo tanto un retardo de tiempo o un retardo de transporte. Un retardo de tiempo se llama también un tiempo muerto. El modelo (2.4) también describe la concentración en una tubería cuando no hay mezcla. \square

Otros sistemas físicos tales como la conducción de calor y la difusión dan lugar a modelos en términos de ecuaciones en derivadas parciales; en la Sección 2.5 se dan ejemplos de tales modelos.

Una característica atractiva de los modelos físicos es que los parámetros de la ecuación pueden estar relacionados con magnitudes físicas tales como volúmenes, caudales, y constantes de los materiales. Se pueden construir también modelos complicados dividiendo un sistema en subsistemas, deduciendo modelos simples para cada subsistema y combinando estos modelos.

Modelos de estado

La noción de *estado* es un concepto importante en dinámica de sistemas. El estado es un conjunto de variables que resumen la conducta pasada del sistema y admite una predicción del futuro bajo la hipótesis de que se conocen las entradas futuras. Para el sistema en el Ejemplo 2.2, que consiste en dos tanques, el estado es simplemente las concentraciones c_1 y c_2 en los tanques. En general, el estado es la variable requerida para describir el almacenamiento de masa, momento y energía en un sistema. Para el sistema en el Ejemplo 2.3 el estado en el tiempo t son las entradas a lo largo de un intervalo de longitud L , i.e., $\{u(\tau), t - L \leq \tau < t\}$.

Respuesta transitoria

Una alternativa a describir modelos por ecuaciones diferenciales es concentrarse directamente sobre la conducta de entrada-salida. La dinámica puede en principio expresarse mediante una gran tabla de señales de entrada y sus correspondientes señales de salida. Este enfoque, que se llama respuesta transitoria, es quizás la forma más intuitiva de caracterizar dinámicas de procesos. Una propiedad muy agradable de los sistemas lineales invariantes en el tiempo es que la tabla se puede describir por un par de señales. La señal de entrada particular se escoge a menudo de forma que es fácil de generar experimentalmente. Ejemplos típicos son escalones, pulsos e impulsos. Recuerde que en la Figura 2.2 se mostraron respuestas en escalón típicas.

Debido al principio de superposición se pueden normalizar las amplitudes de las señales de entrada. Por simplicidad es práctica normalizar dividiendo la salida por la magnitud del escalón de entrada. Es también usual trasladar la curva de forma que el escalón comience en el tiempo $t = 0$. Es entonces suficiente

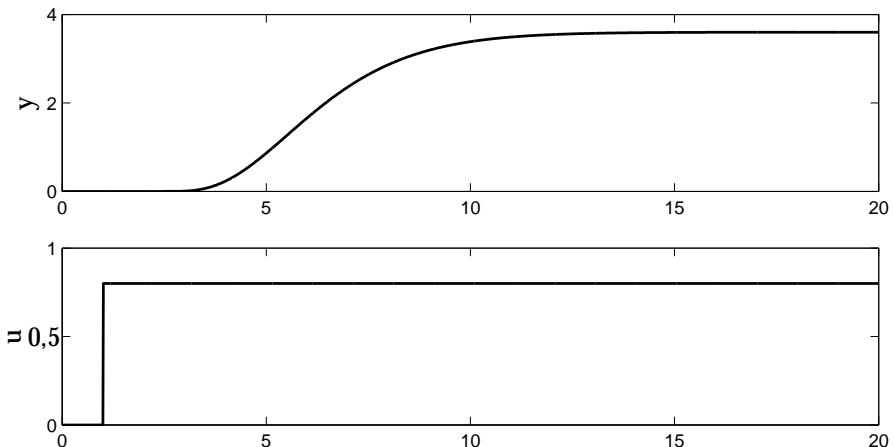


Figura 2.4 La curva inferior muestra una señal de entrada en la forma de un escalón y la curva superior la respuesta del sistema al escalón.

mostrar sólo la salida. En este libro seguiremos esta práctica. Por ejemplo, en la Figura 2.4 la salida se debería dividir por 0,8 y trasladarla una unidad a la izquierda. En las primeras publicaciones sobre control de procesos la respuesta a un escalón se llamó también la curva de reacción.

La salida generada por una entrada arbitraria se puede calcular a partir de la respuesta a un escalón. Sea $h(t)$ la respuesta a un escalón unitario. La salida $y(t)$ a una señal de entrada arbitraria $u(t)$ viene dada por

$$y(t) = \int_0^t u(\tau) \frac{dh(t-\tau)}{dt} d\tau = \int_0^t u(\tau)g(t-\tau)d\tau, \quad (2.5)$$

donde hemos introducido $g(t)$ como la derivada de la respuesta a un escalón $h(t)$. La función $g(t)$ se llama la respuesta impulsional del sistema ya que se puede interpretar como la respuesta del sistema para un impulso muy corto con área unidad.

La función de transferencia

La fórmula (2.5) se puede simplificar bastante al introducir la transformada de Laplace. La transformada de Laplace $F(s)$ de una función del tiempo $f(t)$ se define como

$$F(s) = \int_0^\infty e^{-st} f(t) dt. \quad (2.6)$$

Suponiendo que el sistema está inicialmente en reposo, i.e., $y(t) = 0$ y $u(t) = 0$ para $t \leq 0$, y utilizando la transformada de Laplace, la Ecuación 2.5 se puede escribir como

$$Y(s) = G(s)U(s), \quad (2.7)$$

donde $U(s)$, $Y(s)$, y $G(s)$ son la transformada de Laplace de $u(t)$, $y(t)$, y $g(t)$, respectivamente. La función $G(s)$ se llama la función de transferencia del sistema.

La función de transferencia $G(s)$ es también la transformada de Laplace de la respuesta impulsional $g(t)$.

La fórmula dada por (2.7) tiene una fuerte interpretación intuitiva. La transformada de Laplace de la salida es simplemente la transformada de Laplace de la entrada multiplicada por la función de transferencia del sistema. Ésta es una de las razones principales para utilizar la transformada de Laplace cuando se analizan sistemas lineales. El análisis de sistemas lineales se reduce a pura álgebra. Una característica agradable es que procesos, controladores, y señales se describen de la misma forma.

Se puede utilizar también la Ecuación 2.7 para definir la función de transferencia como la razón de las transformadas de Laplace de la entrada y la salida de un sistema. Como ilustraciones daremos la función de transferencia de algunos sistemas.

EJEMPLO 2.4—TANQUE AGITADO

El tanque agitado en el Ejemplo 2.1 tiene la función de transferencia

$$G(s) = \frac{1}{sV/q + 1} = \frac{1}{sT + 1} \quad (2.8)$$

donde la cantidad $T = V/q$, que tiene dimensiones de tiempo, se llama la constante de tiempo del sistema. \square

EJEMPLO 2.5—RETARDO DE TIEMPO

Considérese el sistema que describe un retardo de transporte en el Ejemplo 2.3. Suponiendo que $u(t) = 0$ para $-L \leq t \leq 0$ encontramos

$$Y(s) = \int_0^\infty e^{-st} y(t) dt = \int_0^\infty e^{-st} u(t - L) dt = e^{-sL} U(s).$$

La función de transferencia de un retardo de transporte es así

$$G(s) = e^{-sL}. \quad (2.9)$$

\square

La Ecuación 2.7 implica que es fácil obtener la función de transferencia de un sistema interconectado. Esto se ilustra por el ejemplo siguiente.

EJEMPLO 2.6—SISTEMA DE PRIMER ORDEN CON RETARDO DE TIEMPO (FOTD)
 Sea un sistema que es un tanque agitado sin mezcla que se alimenta por una tubería. Multiplicando la función de transferencia del tanque en el Ejemplo 2.4 por la función de transferencia de un retardo de tiempo en el Ejemplo 2.5 encontramos que el sistema tiene la función de transferencia

$$G(s) = \frac{1}{1 + sT} e^{-sL}. \quad (2.10)$$

Este modelo es muy común en control de procesos. Se llama sistema de primer orden con un retardo de tiempo o sistema FOTD. \square

Otra buena propiedad de la transformada de Laplace es que la transformada de una derivada está dada por la fórmula

$$\int_0^\infty e^{-st} f'(t) dt = s \int_0^\infty e^{-st} f(t) dt - f(0) = sF(s) - f(0).$$

Si el valor inicial de la función del tiempo es cero, se sigue que la derivada de una función temporal corresponde a la multiplicación de la transformada de Laplace por s . Análogamente, se puede demostrar que la integración de una señal corresponde a dividir la transformada de Laplace por s . Esto da una regla muy simple para manipular ecuaciones diferenciales donde los valores iniciales son ceros. Simplemente hay que sustituir las funciones por sus correspondientes transformadas de Laplace y las derivadas por s . La relación entre señales se obtiene entonces mediante simple álgebra.

EJEMPLO 2.7—ECUACIÓN DIFERENCIAL GENERAL

Sea el sistema descrito por la ecuación diferencial (2.1). Suponiendo que el sistema está inicialmente en reposo y tomando la transformada de Laplace de (2.1) obtenemos

$$(s^n + a_1 s^{n-1} + \dots + a_n) Y(s) = (b_1 s^{n-1} + b_2 s^{n-2} + \dots + b_n) U(s),$$

donde $Y(s)$ es la transformada de Laplace de la salida, y $U(s)$ la transformada de Laplace de la entrada. La función de transferencia del sistema es la razón de las transformadas de Laplace de la salida y la entrada, i.e.,

$$G(s) = \frac{Y(s)}{U(s)} = \frac{b_1 s^{n-1} + b_2 s^{n-2} + \dots + b_n}{s^n + a_1 s^{n-1} + \dots + a_n}. \quad (2.11)$$

\square

EJEMPLO 2.8—CONTROLADOR PID

El controlador PID dado por la Ecuación 1.5 es un sistema dinámico con la función de transferencia

$$C(s) = \frac{U(s)}{E(s)} = K \left(1 + \frac{1}{sT_i} + sT_d \right). \quad (2.12)$$

\square

Los últimos dos ejemplos ilustran que las funciones de transferencia se pueden obtener directamente por inspección a partir de las ecuaciones diferenciales. La regla es simplemente sustituir derivadas por s , integrales por $1/s$, y funciones del tiempo por sus transformadas. Las funciones de transferencia se obtienen entonces como la razón entre señales.

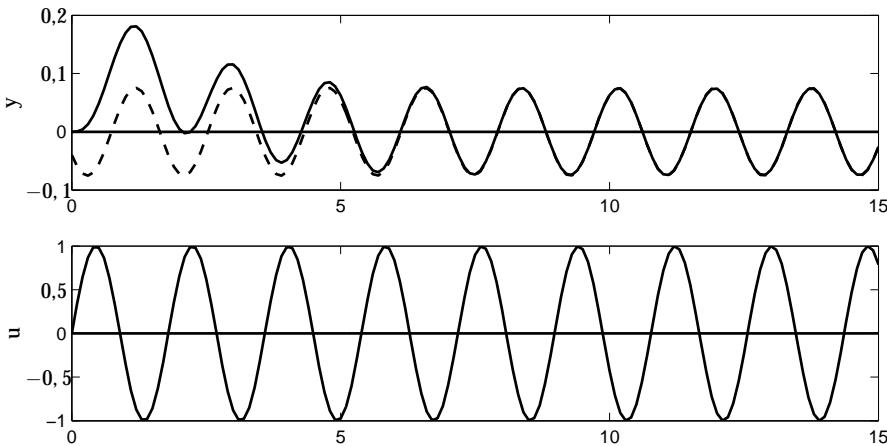


Figura 2.5 Ilustración de la respuesta en frecuencia. La señal de entrada u es una sinusode, y la señal de salida y se hace sinusoidal despues de un transitorio. La línea a trazos muestra la respuesta en estado estacionario a la entrada sinusoidal.

Respuesta en frecuencia

Otra forma de caracterizar la dinámica de un sistema lineal e invariante en el tiempo es investigar la respuesta del sistema a señales de entradas sinusoidales, una idea que se retrotrae al matemático francés Fourier. La respuesta en frecuencia es menos intuitiva que la respuesta transitoria pero da otras perspectivas.

Sea un sistema lineal estable. Si la señal de entrada al sistema es una sinusode, entonces la señal de salida será también una sinusode despues de un cierto transitorio (ver Figura 2.5). La salida tendrá la misma frecuencia que la señal de entrada. Sólo la fase y la amplitud son diferentes. Si la señal de entrada es $u(t) = u_0 \sin \omega t$ la salida en estado estacionario es

$$y(t) = a(\omega) u_0 \sin(\omega t + \varphi(\omega)).$$

Las relaciones en estado estacionario entre la salida y una entrada sinusoidal con frecuencia ω se puede describir por dos números: la razón entre las amplitudes y la fase. La razón entre las amplitudes es la amplitud de salida dividida por la amplitud de entrada, y la fase es el desfase de la salida en relación a la entrada. Las funciones $a(\omega)$ y $\varphi(\omega)$ dan la razón entre las amplitudes y la fase para todas las frecuencias. Las funciones $a(\omega)$ y $\varphi(\omega)$ se relacionan con la función de transferencia de la forma siguiente.

$$G(i\omega) = a(\omega) e^{i\varphi(\omega)}. \quad (2.13)$$

Los valores de la función de transferencia para argumentos imaginarios describen así la transmisión en estado estacionario de señales sinusoidales, y $G(i\omega)$ se llama la función de respuesta en frecuencia del sistema.

El diagrama de Nyquist

Hay muchas ilustraciones gráficas útiles de la respuesta en frecuencia. El número complejo $G(i\omega)$ se puede representar por un vector de longitud $a(\omega)$ que forma

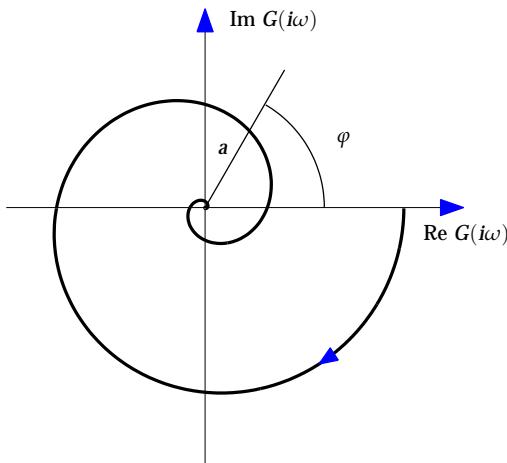


Figura 2.6 El diagrama de Nyquist de un sistema es el lugar del número complejo $G(i\omega)$ cuando ω va desde 0 a ∞ .

un ángulo $\varphi(\omega)$ con el eje real (ver Figura 2.6). Cuando la frecuencia va desde 0 a ∞ , el vector describe una curva en el plano, que se llama la curva de frecuencia o el diagrama de Nyquist.

El diagrama de Nyquist da una descripción completa del sistema. Se puede determinar experimentalmente enviando sinusoides de diferentes frecuencias al sistema. Esto, sin embargo, puede llevar mucho tiempo. También se puede determinar a partir de otras señales.

El diagrama de Bode

El diagrama de Bode es otra representación gráfica de la función de transferencia. El diagrama de Bode de una función de transferencia consiste en dos curvas, la curva de ganancia y la curva de fase; ver Figura 2.7. La curva de amplitud o ganancia muestra la razón de amplitudes $a(\omega) = |G(i\omega)|$ en función de la frecuencia ω . La curva de fase muestra la fase $\varphi(\omega) = \arg G(i\omega)$ en función de la frecuencia ω . La frecuencia, en ambas curvas, se representa en escalas logarítmicas en rad/s o Hz. La ganancia también se expresa en escalas logarítmicas. El ángulo se da en escala lineal. El diagrama de Bode da una buena panorámica de las propiedades de un sistema a lo largo de un amplio rango de frecuencias. Debido a que las escalas de la curva de ganancia también tienen asíntotas lineales.

2.4 Modelos basados en características

Algunas veces es deseable tener una caracterización aproximada de un proceso basándose solamente en unas pocas características. Estas deberían de escogerse de forma que sean significativas y con una buena interpretación física. Deberían también ser fáciles de determinar experimentalmente. Esta forma de describir la dinámica tiene una larga tradición en control de procesos. Es útil comenzar

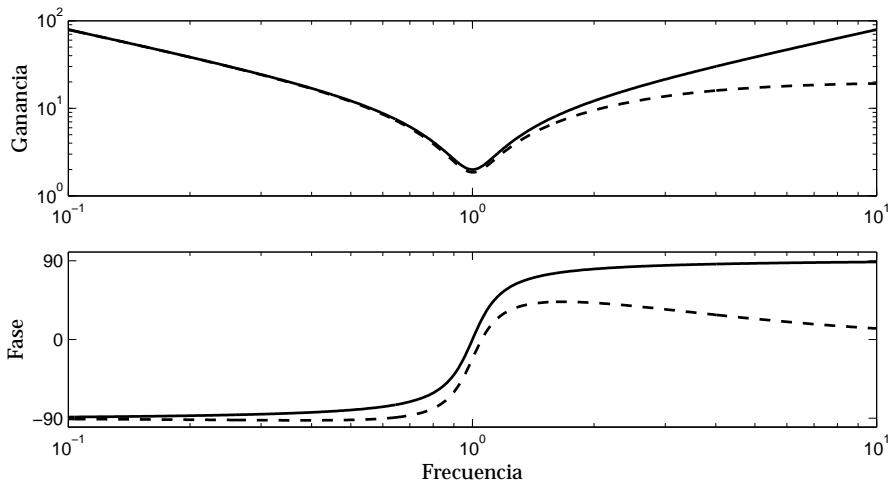


Figura 2.7 El diagrama de Bode de un controlador PID ideal (línea continua) y un controlador con un filtro (línea a tramos). La curva superior muestra la curva de ganancia $|G(i\omega)|$, y el diagrama inferior la curva de fase $\arg G(i\omega)$. El controlador tiene una ganancia elevada para bajas frecuencias, y la fase es -90° . El controlador ideal tiene también una elevada ganancia en altas frecuencias y la fase es 90° . El controlador con un filtro tiene ganancia constante en altas frecuencias.

con una clasificación cruda de la respuesta a un escalón tal como se ilustra en la Figura 2.2.

Ganancia del proceso

Para procesos estables se puede describir la conducta en estado estacionario mediante un parámetro, la ganancia del proceso K_p . Para procesos con integración una entrada constante da en estado estacionario una salida que cambia con una velocidad constante. Esta conducta se puede expresar haciendo uso de la constante de velocidad K_v .

Tiempo de residencia medio

Es también útil encontrar unos pocos parámetros para caracterizar la dinámica del proceso. La conducta temporal de un sistema estable con respuesta impulsional positiva se puede caracterizar por el parámetro

$$T_{ar} = \frac{\int_0^\infty t g(t) dt}{\int_0^\infty g(t) dt}, \quad (2.14)$$

que se llama el tiempo de residencia medio. El tiempo de residencia medio es una medida aproximada de cuánto tarda la entrada en tener una influencia significativa sobre la salida. Obsérvese que la función $g(t) / \int g(t) dt$ se puede interpretar como una densidad de probabilidad si $g(t) \geq 0$.

El tiempo de residencia medio se puede calcular a partir de la respuesta a un

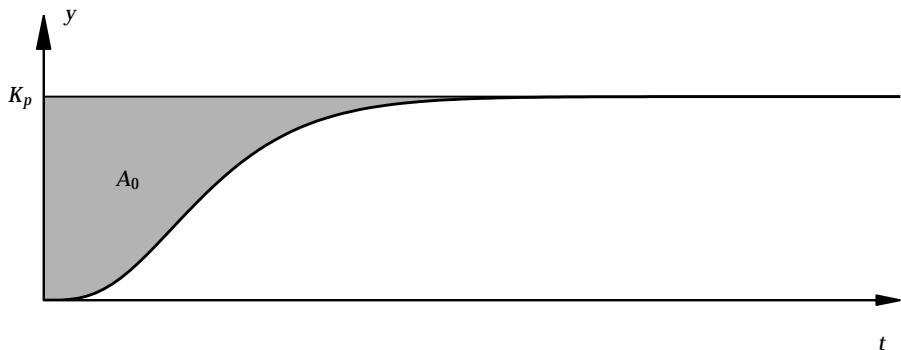


Figura 2.8 Ilustra el método del área para determinar el tiempo de residencia medio.

escalón de la forma siguiente

$$T_{ar} = \frac{\int_0^{\infty} (h(\infty) - h(t)) dt}{K_p} = \frac{A_0}{K_p}, \quad (2.15)$$

donde $h(t)$ es la respuesta a un escalón y $K_p = G(0)$ es la ganancia estática del proceso. Obsérvese que $K_p = h(\infty)$ y que A_0 es el área sombreada en la Figura 2.8.

Tiempo de residencia medio y funciones de transferencia

El tiempo de residencia medio se puede calcular de forma muy conveniente a partir de la función de transferencia. Como la función de transferencia es la transformada de Laplace de la respuesta impulsional tenemos

$$G(s) = \int_0^{\infty} e^{-st} g(t) dt.$$

Derivando esta expresión con respecto a s da

$$G'(s) = - \int_0^{\infty} e^{-st} t g(t) dt.$$

Haciendo $s = 0$ en estas expresiones se sigue entonces de la definición del tiempo de residencia medio (2.14) que

$$T_{ar} = - \frac{G'(0)}{G(0)}. \quad (2.16)$$

Esta fórmula se ilustrará ahora mediante unos pocos ejemplos.

EJEMPLO 2.9—TIEMPO DE RESIDENCIA MEDIO PARA EL TANQUE AGITADO
La función de transferencia para el tanque agitado en el Ejemplo 2.4 es

$$G(s) = \frac{1}{1+sT}$$

Tenemos

$$G'(s) = -\frac{T}{(1+sT)^2},$$

y de (2.16) se sigue que el tiempo de residencia medio es

$$T_{ar} = T = \frac{V}{q}.$$

El tiempo de residencia medio es así la razón entre el volumen y el caudal a través del tanque. \square

EJEMPLO 2.10—TIEMPO DE RESIDENCIA MEDIO PARA EL RETARDO DE TIEMPO
La función de transferencia para el retardo de tiempo en el Ejemplo 2.5 es

$$G(s) = e^{-sL}.$$

tenemos

$$G'(s) = -Le^{-sL},$$

y de (2.16) se sigue que el tiempo de residencia medio es

$$T_{ar} = L.$$

El tiempo de residencia medio es así igual al retardo de tiempo. \square

EJEMPLO 2.11—TIEMPO DE RESIDENCIA MEDIO PARA SISTEMAS EN CASCADA
Un sistema que es la combinación en cascada de dos sistemas lineales estables con funciones de transferencia $G_1(s)$ y $G_2(s)$ tiene la función de transferencia

$$G(s) = G_1(s)G_2(s).$$

Tomando la derivada da

$$G'(s) = G'_1(s)G_2(s) + G_1(s)G'_2(s).$$

Se sigue de (2.16) que el tiempo de residencia medio es

$$T_{ar} = -\frac{G'_1(0)G_2(0) + G_1(0)G'_2(0)}{G_1(0)G_2(0)} = -\frac{G'_1(0)}{G_1(0)} - \frac{G'_2(0)}{G_2(0)}.$$

El tiempo de residencia medio es la suma de los tiempos de residencia de cada sistema. \square

Este ejemplo muestra que el tiempo de residencia medio para el modelo FOTD del Ejemplo 2.6 es $T_{ar} = L + T$.

Un sistema con la función de transferencia

$$G(s) = \frac{K_p(1+sT_1)(1+sT_2)}{(1+sT_3)(1+sT_4)(1+sT_5)} e^{-sL}.$$

tiene el tiempo de residencia medio $T_{ar} = T_3 + T_4 + T_5 + L - T_1 - T_2$.

Modelos con dos parámetros

Una forma muy simple de caracterizar la dinámica de un proceso estable es utilizar la ganancia K_p y el tiempo de residencia medio T_{ar} . Esto da el modelo siguiente

$$\begin{aligned} G(s) &= \frac{K_p}{1+sT_{ar}} \\ G(s) &= K_p e^{-sT_{ar}}, \end{aligned} \tag{2.17}$$

donde la dinámica se representa por una constante de tiempo o un retardo de tiempo.

Retardo de tiempo aparente y constante de tiempo aparente

Los sistemas con respuestas a un escalón que son esencialmente monótonas son muy comunes en control de procesos. Tales sistemas se pueden modelar como sistemas de primer orden con retardo de tiempo mediante la función de transferencia

$$G(s) = \frac{K_p}{1+sT} e^{-sL}. \tag{2.18}$$

Para poner de manifiesto que los parámetros L y T son aproximados nos referimos a ellos como el *retardo de tiempo aparente* y la *constante de tiempo aparente*, o el *retardo aparente*, respectivamente. El tiempo de residencia medio es $T_{ar} = L + T$. El parámetro

$$\tau = \frac{L}{T_{ar}} = \frac{L}{L+T}, \tag{2.19}$$

que tiene la propiedad $0 \leq \tau \leq 1$, se llama el *retardo de tiempo normalizado* o el *tiempo muerto normalizado*. Este parámetro se puede usar para caracterizar la dificultad de controlar un proceso. Se llama algunas veces la *razón de controlabilidad*. En términos generales se puede decir que procesos con τ pequeña son fáciles de controlar, y la dificultad en controlar el sistema aumenta cuando τ aumenta. Sistemas con $\tau = 1$ corresponden a procesos con retardo de tiempo puro, que son también difíciles de controlar.

Ganancia última y período último

Hasta aquí hemos utilizado características que se basan en la respuesta transitoria. Es también posible emplear características de la respuesta en frecuencia. Los

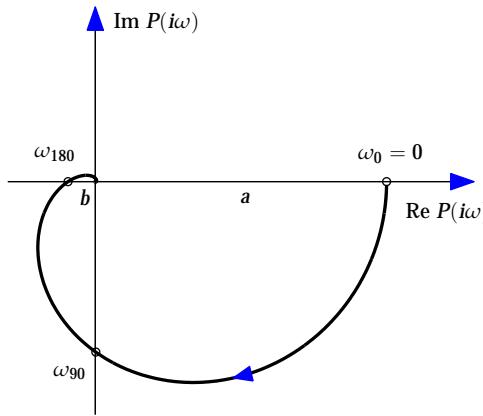


Figura 2.9 Diagrama de Nyquist con los puntos ω_0 , ω_{90} y ω_{180} . La razón de ganancia κ es la relación entre las distancias a y b .

modelos se pueden expresar en términos de sus retardos de fases y la frecuencia, donde los sistemas tienen un retardo de fase dado. Para este objetivo, introducimos ω_φ para denotar la frecuencia donde el retardo de fase es de φ grados, e introducimos $K_\varphi = |G(i\omega_\varphi)|$ para denotar la ganancia del proceso en ω_φ . Las frecuencias ω_{90} y ω_{180} y las correspondientes ganancias del proceso K_{90} y K_{180} son de interés particular para el control PID. Estas frecuencias corresponden a las intersecciones del diagrama de Nyquist con los ejes reales e imaginarios negativos; ver Figura 2.9. También tienen interpretaciones físicas agradables. Considérese un proceso con control proporcional puro. Si se aumenta la ganancia del controlador el proceso comenzará a oscilar, y alcanzará el límite de estabilidad cuando la ganancia del controlador es $K_u = 1/K_{180}$. La oscilación tendrá la frecuencia ω_{180} . Esta frecuencia se llama la frecuencia última. El parámetro K_u se llama la ganancia última o la ganancia crítica. Los parámetros K_{90} y ω_{90} tienen interpretaciones similares para un proceso con control integral puro.

La razón de ganancia

La razón de ganancia es un parámetro adicional que da una información útil acerca del sistema. Este parámetro se define como

$$\kappa = \frac{K_{180}}{K_p} = \frac{|G(i\omega_{180})|}{G(0)}. \quad (2.20)$$

Es un indicador de cuán difícil es controlar el proceso. Procesos con un pequeño κ son fáciles de controlar. La dificultad aumenta cuando κ aumenta. El parámetro es también la razón entre las distancias a y b en el diagrama de Nyquist; ver Figura 2.9.

El parámetro κ está también relacionado con el retardo de tiempo normalizado τ . Para el modelo FOTD dado por la Ecuación 2.18 los parámetros τ y κ están

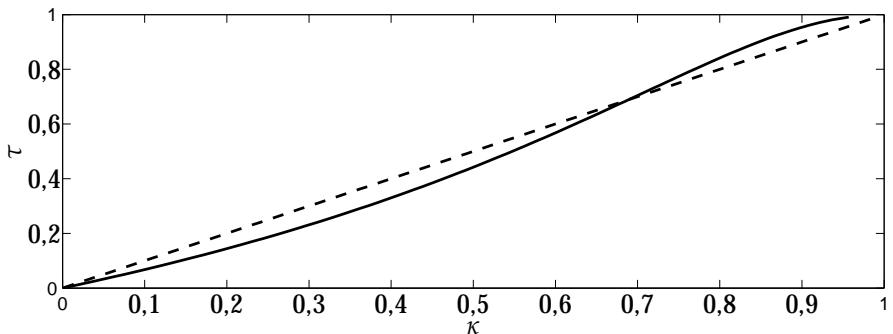


Figura 2.10 El retardo de tiempo normalizado τ como función de la razón de ganancia κ para el sistema (2.18). La línea a tramos muestra la aproximación de línea recta $\kappa = \tau$.

relacionados de la forma siguiente:

$$\tau = \frac{\pi - \arctan \sqrt{1/\kappa^2 - 1}}{\pi - \arctan \sqrt{1/\kappa^2 - 1} + \sqrt{1/\kappa^2 - 1}}. \quad (2.21)$$

Esta relación es casi lineal tal como se muestra en la Figura 2.10, y se mantiene aproximadamente para muchos otros sistemas. Como una aproximación bruta podemos igualar κ y τ . Para valores pequeños se obtiene una mejor aproximación con $\kappa = 1,6\tau$. Para el modelo FOTD es posible también encontrar los parámetros L y T de κ y ω_{180} utilizando las siguientes ecuaciones

$$\begin{aligned} T &= \frac{1}{\omega_{180}} \sqrt{\kappa^{-2} - 1} \\ L &= \frac{1}{\omega_{180}} (\pi - \arctan \kappa^{-2} - 1) \\ K &= \frac{|G(i\omega_{180})|}{\kappa}. \end{aligned} \quad (2.22)$$

2.5 Modelos de procesos típicos

Gran parte de la conducta dinámica encontrada en control es relativamente simple. Los procesos se diseñan de tal forma que sean fáciles de controlar. Si se utiliza control PID es natural que se empleen modelos de procesos sencillos. En esta sección presentaremos algunos de los modelos que se suelen utilizar en conexión con el control PID. La mayoría de estos modelos vienen caracterizados solamente por unos pocos parámetros.

El modelo FOTD

Un modelo de proceso que se suele utilizar en control de procesos tiene la función de transferencia (2.18). Es simple y describe aproximadamente la dinámica de muchos procesos industriales. Una comparación con los Ejemplos 2.3 y 2.4

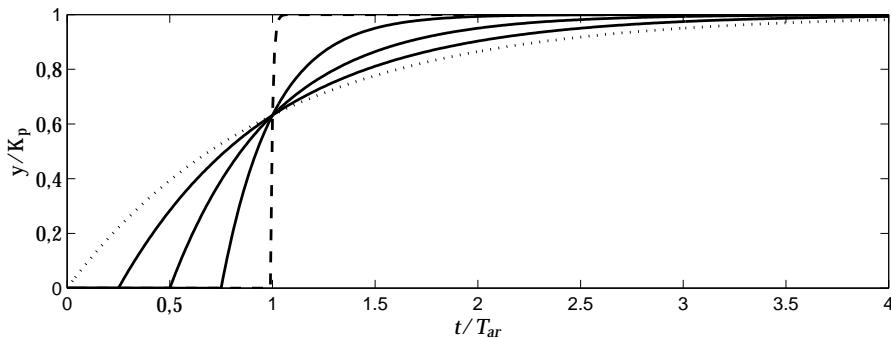


Figura 2.11 Respuesta a un escalón normalizado del modelo FOTD (2.18) para diferentes valores del retardo de tiempo normalizado. El retardo de tiempo normalizado es $\tau = 0$ (punteada), $0,25$, $0,5$, $0,75$ y $0,99$ (a tramos).

muestra que pueden representar la dinámica de un tanque agitado sin mezcla con una tubería. El modelo se caracteriza por tres parámetros: la ganancia (estática) K_p , la constante de tiempo T , y el retardo de tiempo L . La constante de tiempo T se llama también el retardo. La respuesta en escalón del modelo (2.18) es

$$h(t) = K_p \left(1 - e^{-(t-L)/T} \right).$$

Como el tiempo de residencia medio es $T_{ar} = L + T$, el valor de la respuesta en escalón en este tiempo es

$$h(T_{ar}) = K_p \left(1 - e^{-1} \right) \approx 0,63K_p.$$

El tiempo de residencia medio se puede así determinar como el tiempo en el que la respuesta a un escalón ha alcanzado el 63 por ciento de su valor en estado estacionario.

Dos parámetros del modelo (2.18) corresponden al escalado de los ejes que se puede reducir por normalización. Se pueden escoger como la ganancia y el tiempo de residencia medio. Esto significa que si se escala la salida por la ganancia $K_p = G(0)$ y el tiempo por el tiempo de residencia medio T_{ar} la respuesta está completamente caracterizada por un parámetro, el tiempo muerto normalizado τ . El sistema es un retardo de tiempo puro para $\tau = 1$ y un sistema de primer orden para $\tau = 0$.

La Figura 2.11 muestra la respuesta en escalón normalizada para diferentes valores de τ . Obsérvese que todas las curvas intersectan en un punto $t = T_{ar}$ a causa de la normalización.

Tanques no interactuantes o retardos múltiples

La función de transferencia (2.8) representa la dinámica de un tanque simple. La parte superior de la Figura 2.12 muestra un sistema que es una combinación en cascada de n tanques. Este sistema tiene la función de transferencia

$$G_n(s) = \frac{K_p}{(1+sT)^n}. \quad (2.23)$$

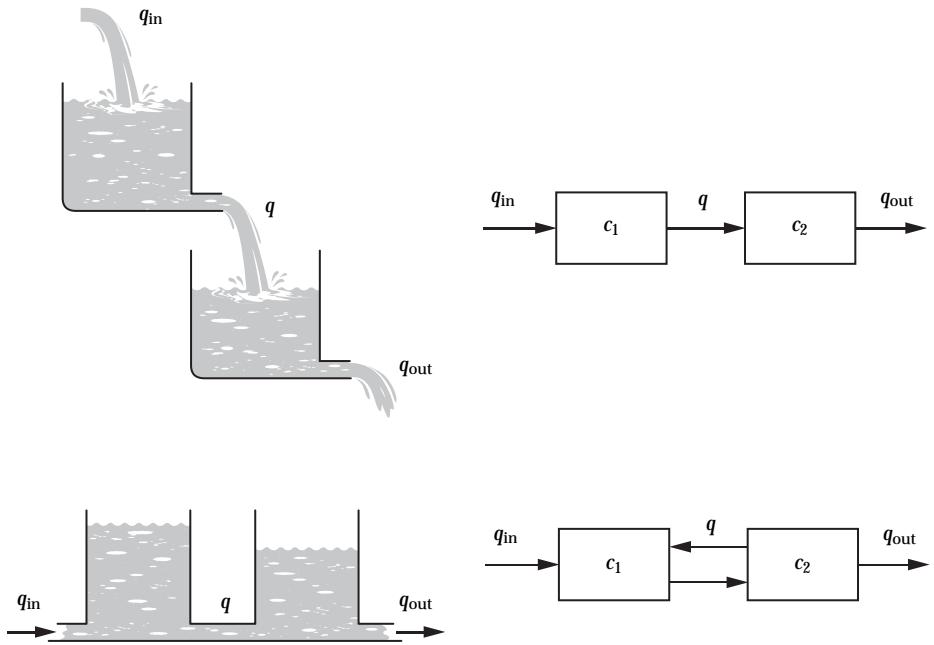


Figura 2.12 Tanques en cascada y sus correspondientes representaciones en diagrama de bloques. Los tanques superiores son no interactuantes, y los inferiores son interactuantes.

donde n es el número de tanques. Como un sistema de primer orden se llama también un retraso, el sistema se denomina también un sistema con retraso múltiple. Obsérvese que esta fórmula se cumple si el caudal de cada tanque sólo depende de su nivel. Esto significa que no hay interacción entre los tanques.

El tiempo de residencia medio es

$$T_{ar} = -\frac{G'(0)}{G(0)} = nT.$$

El modelo (2.23) tiene la respuesta impulsional

$$g(t) = \frac{K_p}{(n-1)!} \frac{t^{n-1}}{T^n} e^{-t/T}, \quad (2.24)$$

que tiene su máximo

$$\max g(t) = \frac{K_p(n-1)^{n-2}}{T(n-2)!} e^{-n+1},$$

para $t = (n-1)T$. La respuesta a un escalón unitario es

$$h(t) = K_p \left(1 - \left(1 + \frac{t}{T} + \frac{t^2}{2T^2} + \dots + \frac{t^{n-1}}{(n-1)! T^{n-1}} \right) \right) e^{-t/T}. \quad (2.25)$$

Cuadro 2.1 Constante de tiempo aparente T_e , retardo de tiempo aparente L_e , tiempo de residencia medio T_{ar} , y retardo de tiempo normalizado τ para el proceso (2.23).

n	2	3	4	5	6	7	8	16	32
T_e	1,86	2,44	2,91	3,32	3,68	4,01	4,31	6,23	8,90
L_e	0,28	0,81	1,43	2,10	2,81	3,55	4,31	10,78	24,67
T_{ar}	2,14	3,25	4,34	5,42	6,49	7,56	8,62	17,02	33,57
τ	0,13	0,25	0,33	0,39	0,43	0,47	0,50	0,63	0,73

La respuesta a un escalón se caracteriza por tres parámetros, K_p , n , and T . El número de parámetros se puede reducir por normalización. Los parámetros K_p y T sólo influyen sobre el escalado de los ejes. La forma de la respuesta en escalón está así únicamente dada por el parámetro n . En la Figura 2.13 se muestran las respuestas normalizadas a un escalón para diferentes valores de n . Las respuestas a un escalón están próximas pero no son iguales en $t = T_{ar}$. Cuando n va a infinito tenemos

$$\lim_{n \rightarrow \infty} G_n(s) = K_p e^{-t/T_{ar}}.$$

Para grandes n el sistema se aproxima a un retardo de tiempo puro. La Figura 2.13 muestra, sin embargo, que se requieren muy grandes valores de n para obtener una buena aproximación de la respuesta a un escalón de un retardo de tiempo.

La función de transferencia $G_n(s)$ se puede aproximar por un sistema FOTD. En la Tabla 2.1 se dan las constantes y retardos de tiempo aparente para la aproximación.

Tanques interactuantes múltiples – Retrasos distribuidos

La dinámica de los tanques en cascada es muy diferente si los tanques son interactuantes. En el sistema que se muestra en la parte inferior de la Figura 2.12 el caudal de un tanque depende de los niveles de los tanques vecinos. Sea x_k el niv-

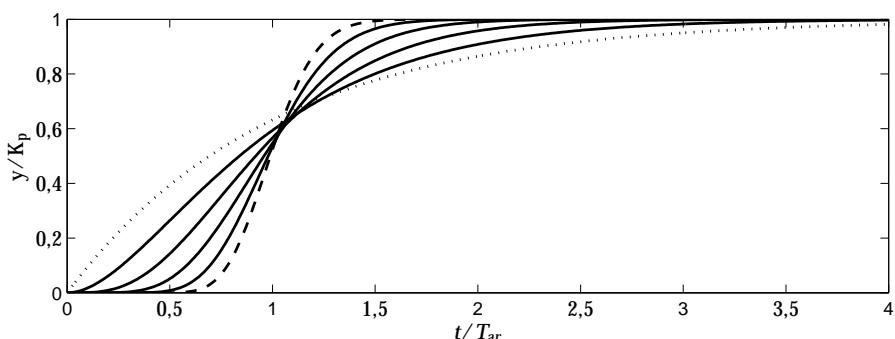


Figura 2.13 Respuestas normalizadas a un escalón para los procesos $G_n(s) = 1/(1 + sT)^n$ para $n = 1$ (punteada), 2, 4, 8, 16, y 32 (a tramos).

el del tanque k . La variable de control u es el caudal de entrada al primer tanque, y la salida es el caudal del tanque n . Supóngase que los tanques tienen una sección recta unitaria, y que el caudal del tanque k al tanque $k+1$ es $x_k - x_{k-1}$. Los balances de masa para los tanques son

$$\begin{aligned}\frac{dx_1}{dt} &= -x_1 + x_2 + u \\ &\vdots \\ \frac{dx_k}{dt} &= x_{k-1} - 2x_k + x_{k+1} \\ &\vdots \\ \frac{dx_n}{dt} &= x_{n-1} - 2x_n.\end{aligned}\tag{2.26}$$

Éste es un modelo de estado con n estados. Las variables de estado representan los niveles en los diferentes tanques. El sistema se denomina también un retraso distribuido. Con una entrada en escalón unitaria los valores de equilibrio de los estados son $x_k = n - k + 1$. Los polinomios característicos de los sistemas que tienen diferentes órdenes son

$$\begin{aligned}d_1 &= s + 1 \\ d_2 &= s^2 + 3s + 1 \\ d_n &= (s + 2)d_{n-1} - d_{n-2}.\end{aligned}$$

Las funciones de transferencia para algunos valores de n están dados por

$$\begin{aligned}G_1(s) &= \frac{1}{s + 1} \\ G_2(s) &= \frac{1}{s^2 + 3s + 1} \\ G_4(s) &= \frac{1}{s^4 + 7s^3 + 15s^2 + 10s + 1} \\ G_8(s) &= \frac{1}{s^8 + 15s^7 + 91s^6 + 286s^5 + 495s^4 + 462s^3 + 210s^2 + 36s + 1}.\end{aligned}$$

El tiempo de residencia medio es la razón entre el volumen total en estado estacionario y el caudal, por lo tanto

$$T_{ar} = \frac{n(n+1)}{2}.$$

Éste es también el coeficiente del término- s en el denominador de la función de transferencia. Cuando el número de tanques aumenta tenemos asintóticamente para grandes n

$$G_n(s) \approx \frac{1}{\cosh \sqrt{2T_{ar}}s}.$$

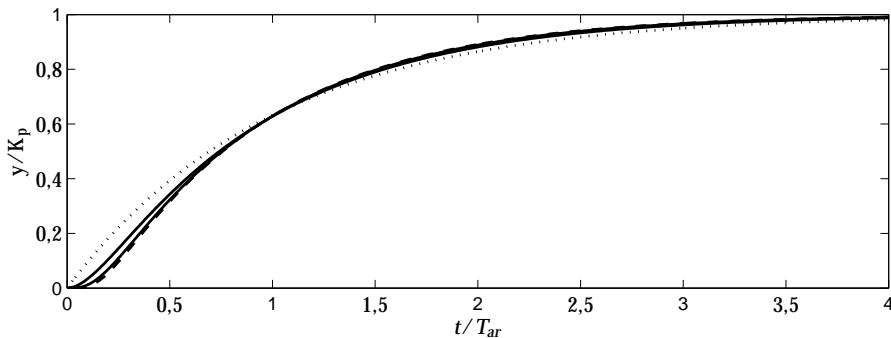


Figura 2.14 Respuestas normalizadas a un escalón para tanques interactuantes, (2.26), para $n = 1$ (punteada), 2, 4, y 8 (a tramos).

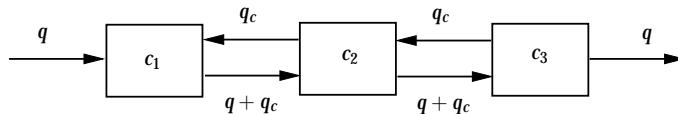


Figura 2.15 Diagrama esquemático de tres tanques en cascada con recirculación.

Estas funciones de transferencias son muy diferentes de la función de transferencia (2.23) de los tanques no interactuantes.

La Figura 2.14 muestra las respuestas a un escalón normalizadas para tanques interactuantes. Observése que las respuestas son muy similares para valores grandes de n . Una comparación con la Figura 2.13 muestra que hay una diferencia significativa entre los tanques interactuantes y no interactuantes.

Otra versión de tanques interactuantes

El modelo (2.26) no es la única forma de interconectar tanques. En la Figura 2.15 se muestra otra configuración. Por simplicidad hemos mostrado un sistema con tres tanques. El sistema consiste en tanques agitados idénticos con un caudal directo q y un caudal de recirculación q_c . Sea V el volumen del tanque, u la concentración del caudal de entrada, c_k la concentración en el tanque k , e $y = c_n$ la concentración en el caudal de salida. Los balances de masa para un sistema con n tanques son

$$\begin{aligned} V \frac{dc_1}{dt} &= -(q + q_c)c_1 + q_c c_2 + qu \\ &\vdots \\ V \frac{dc_k}{dt} &= (q + q_c)c_{k-1} - (q + 2q_c)c_k + q_c c_{k+1} \\ &\vdots \\ V \frac{dc_n}{dt} &= (q + q_c)c_{n-1} - (q + q_c)c_n \\ y &= c_n. \end{aligned}$$

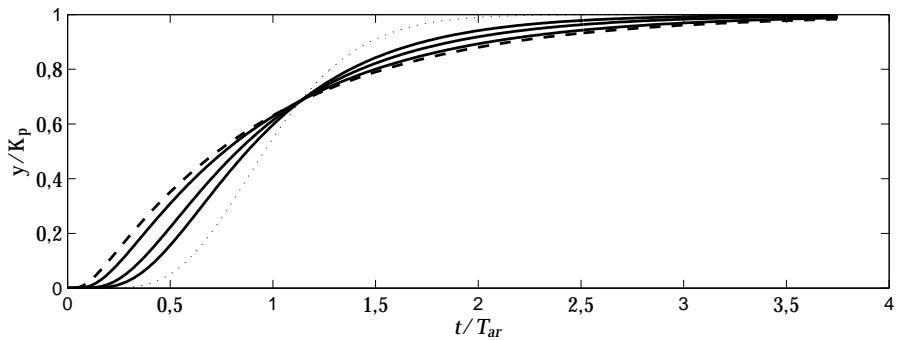


Figura 2.16 Respuestas normalizadas a un escalón para ocho tanques con recirculación. La razón de recirculación es $q_c/q = 0$ (punteada), 1, 2, 5 y 10 (a tramos).

Esto es también un modelo de estado donde los estados son las concentraciones en los diferentes tanques. Las funciones de transferencia para algunos valores de n son

$$\begin{aligned} G_1(s) &= \frac{q}{Vs + q} \\ G_2(s) &= \frac{q(q + q_c)}{Vs^2 + 2V(q + q_c)s + q(q + q_c)} \\ G_3(s) &= \frac{2q_c(q + q_c)^2}{(Vs + q + q_c)(V^2s^2 + (2q + 3q_c)s + 2q_c(q + q_c))} \\ &= \frac{2q_c(q + q_c)^2}{V^3s^3 + (3q + 4q_c)V^2s^2 + (q + q_c)(q + 3q_c)Vs + 2q_c(q + q_c)^2}. \end{aligned}$$

La ganancia estática es $K_p = 1$ y el tiempo de residencia medio es

$$T_{ar} = \frac{nV}{q}.$$

El caudal de recirculación tiene un gran impacto sobre la dinámica. Para $q_c = 0$ no hay interacción, y el sistema es equivalente al modelo dado por (2.23). Cuando $q_c/q \rightarrow \infty$ el modelo es equivalente al modelo dado por (2.26). El modelo con recirculación permite así interpolar entre los modelos con retrasos no interactuantes y distribuidos.

La Figura 2.16 muestra la respuesta a un escalón de un sistema de orden n para diferentes valores del caudal de recirculación.

Sistemas oscilatorios

El modelo (2.18) no puede describir sistemas con respuestas oscilatorias. Un modelo simple para tales sistemas está dado por la función de transferencia

$$G(s) = \frac{K_p}{1 + 2\zeta sT + (sT)^2}. \quad (2.27)$$

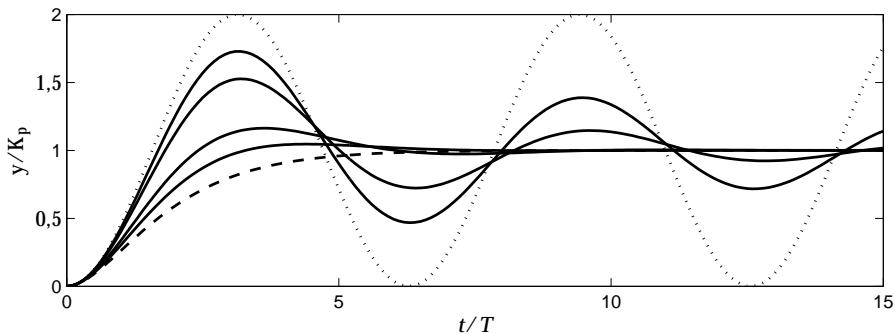


Figura 2.17 Respuestas normalizadas a un escalón de sistemas oscilatorios (2.27) con $\zeta = 0$ (punteada), 0,1, 0,2, 0,5, 0,7, y 1,0 (a tramos).

Este modelo tiene tres parámetros: ganancia estática K_p , constante de tiempo T , y amortiguamiento relativo ζ . El parámetro $1/T$ se llama también frecuencia natural no amortiguada. Las respuestas a un escalón se pueden normalizar por la ganancia y la constante de tiempo. Su forma está entonces determinada por un único parámetro. En la Figura 2.17 se muestran las respuestas a un escalón. Para $\zeta < 1$ la respuesta a un escalón tiene su máximo

$$M = K_p e^{-\frac{\pi \zeta}{\sqrt{1-\zeta^2}}},$$

que ocurre en

$$t_{max} = \frac{2\pi T}{\sqrt{1-\zeta^2}}.$$

La posición del máximo aumenta cuando ζ aumenta, y se hace infinita para $\zeta = 1$ cuando desaparece la sobreelongación. La función de transferencia es entonces

$$G(s) = \frac{K_p}{(1+sT)^2},$$

y la respuesta a un escalón es

$$h(t) = K_p \left(1 - e^{-t/T} - \frac{t}{T} e^{-t/T} \right).$$

En la Figura 2.18 se muestran los diagramas.

Procesos con integración

Los procesos con integración no alcanzarán el estado estacionario durante operaciones en lazo abierto. En la práctica, lo mismo es cierto para procesos con constantes de tiempo muy grandes. Asintóticamente, la salida cambiará a velocidad constante después de un cambio en escalón en la señal de control. En los primeros trabajos sobre control de proceso a estos sistemas se los denominaba sistemas

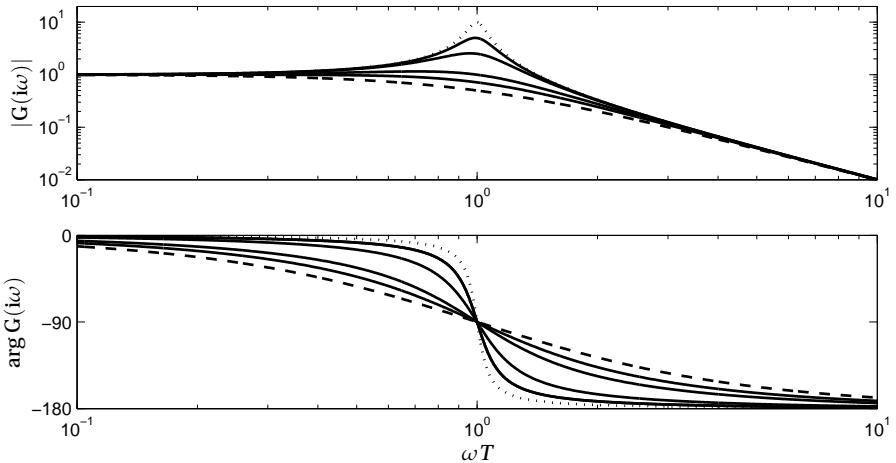


Figura 2.18 Diagramas de Bode de sistemas oscilatorios (2.27) con $\zeta = 0,05$ (punteada), 0,1, 0,2, 0,5, 0,7, y 1,0 (a tramos).

sin auto regulación, ya que la variable de proceso no alcanzaba un estado estacionario después de una perturbación. Muchos métodos para la sintonía de controladores PID también tratan a estos sistemas separadamente. Los modelos para tales sistemas se obtienen simplemente al dividir la función de transferencia de un proceso con auto-regulación por s .

Una combinación de un integrador y un retardo de tiempo es un modelo muy común. La función de transferencia es

$$G(s) = \frac{K_v}{s} e^{-sL}. \quad (2.28)$$

Este modelo se caracteriza por dos parámetros, una ganancia y un retardo de tiempo. La ganancia de integración se representa mediante un símbolo especial K_v , que indica cuán rápido aumenta la salida en estado estacionario después de un cambio a la entrada en un escalón unitario. El parámetro K_v tiene dimensión de frecuencia.

Una combinación de un retardo y un integrador es un modelo que se suele emplear para describir servosistemas sencillos. Este modelo tiene la función de transferencia

$$G(s) = \frac{K_v}{s(1+sT)}. \quad (2.29)$$

La función de transferencia $K_v/(1+sT)$ representa la función de transferencia de la tensión del servosistema a la velocidad de rotación, y el integrador representa la relación entre la velocidad angular y el ángulo.

Un modelo ligeramente más complicado se obtiene añadiendo integración al modelo estándar (2.18).

$$G(s) = \frac{K_v}{s(1+sT)} e^{-sL}. \quad (2.30)$$

A esto lo llamamos el modelo FOTD con integración o FOTDI de forma breve. Este proceso se puede normalizar de la misma forma que el modelo (2.18) al in-

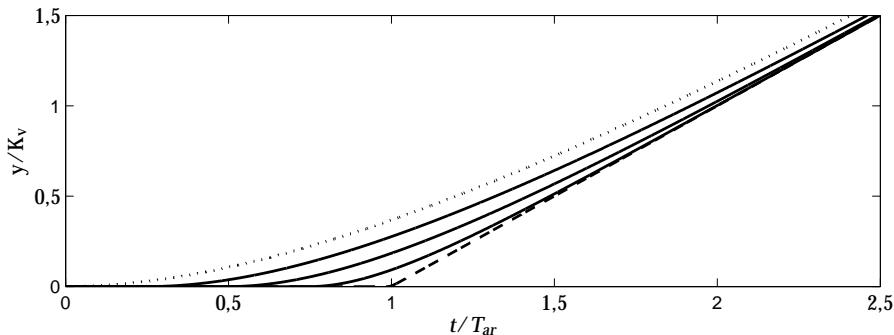


Figura 2.19 Respuestas normalizadas a un escalón del modelo FOTDI (2.30) para $\tau = 0$ (punteada), 0,25, 0,5, 0,75, y 0,99 (a tramos).

troducir el retardo de tiempo normalizado dado por (2.19). En la Figura 2.19 se muestra la respuesta a un escalón del modelo FOTDI normalizado.

Sistemas con respuesta inversa

Los sistemas analizados hasta ahora no tienen ningún cero. Los sistemas que se representan como una conexión en paralelo de algunos sistemas pueden tener funciones de transferencia del tipo

$$G(s) = \frac{1+sT}{s^2 + 1,4s + 1}. \quad (2.31)$$

Este sistema tiene un cero en $s = -1/T$, que puede tener una influencia significativa sobre la respuesta del sistema. La Figura 2.20 muestra la respuesta en escalón de este sistema para $T = -2, -1, 0, 1, y 2$. Obsérvese que la sobreelongación de la respuesta en escalón aumenta con valores positivos crecientes de T . También se ve que la señal de salida inicialmente se mueve en la dirección equivocada cuando T es negativo. Tales sistemas se dice que tienen respuestas inversas. Esta clase de sistemas son difíciles de controlar. Ejemplos de tales sistemas son la dinámica de nivel en los generadores de vapor, la dinámica de las centrales de energía hidroeléctrica, la dinámica de los coches de choque.

Conducción del calor

El control de temperatura es una aplicación muy común del control PID. Ahora analizaremos algunos modelos que se basan directamente en relaciones físicas. Sea una barra infinitamente larga con difusividad térmica λ . Supóngase que no hay transferencia de calor radial y que la entrada es la temperatura en el extremo izquierdo de la barra. La función de transferencia en un punto a la distancia a desde el extremo izquierdo es

$$G(s) = e^{-\sqrt{sT}}, \quad (2.32)$$

donde $T = a^2/\lambda$. La respuesta impulsional del sistema viene dada por

$$h(t) = \frac{\sqrt{T}}{2\sqrt{\pi}t^{3/2}} e^{-\frac{T}{4t}}. \quad (2.33)$$

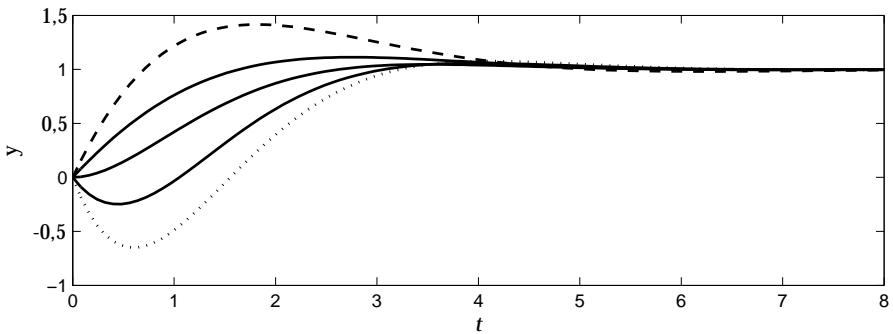


Figura 2.20 Respuestas a un escalón del modelo (2.31) para $T = -2$ (punteada), $-1, 0, 1$ y 2 (a tramos).

Esta respuesta impulsional tiene la propiedad de que todas sus derivadas son cero para $t = 0$, lo que significa que la respuesta inicial del sistema es muy lenta. La respuesta impulsional tiene un máximo en $t = T/6$. Para grandes valores de t la respuesta impulsional decae muy lentamente como $t^{-1,5}$. La respuesta a un escalón del sistema es

$$y(t) = 1 - \operatorname{erf} \sqrt{\frac{T}{4t}} = 1 - \frac{2}{\sqrt{\pi}} \int_{-\infty}^{\sqrt{T/4t}} e^{-x^2} dx. \quad (2.34)$$

En la Figura 2.21 se muestran las respuestas en escalón e impulsional. Obsérvese que la temperatura inicialmente comienza a elevarse muy lentamente. Después de un aumento rápido se aproxima también al estado estacionario muy lentamente.

Vamos a considerar ahora la situación cuando se aísla el lado derecho. La función de transferencia entonces se hace

$$G(s) = \frac{1}{\cosh \sqrt{sT}}. \quad (2.35)$$

Este es un sistema con infinitos retrasos con constantes de tiempo $4T/\pi^2, T/9\pi^2, 4T/25\pi^2, 4T/49\pi^2, \dots$. Esta función de transferencia se llama también un retraso distribuido.

En la Figura 2.22 se muestra la respuesta a un escalón de esta función de transferencia. Obsérvese que la respuesta se aproxima al estado estacionario más rápidamente que el sistema (2.32). La respuesta a un escalón de los sistemas es así bastante diferente. El aislamiento del lado derecho de la barra facilita mucho la transferencia de calor en el sistema. Un sencillo cálculo muestra que el tiempo de residencia medio para el sistema es

$$T_{ar} = -G'(0) = \frac{T}{2}. \quad (2.36)$$

El sistema (2.32) con la función de transferencia $e^{-\sqrt{sT}}$ tiene un tiempo de residencia infinito que refleja el hecho de que la respuesta impulsional decae muy lentamente; comparar con la Figura 2.21.

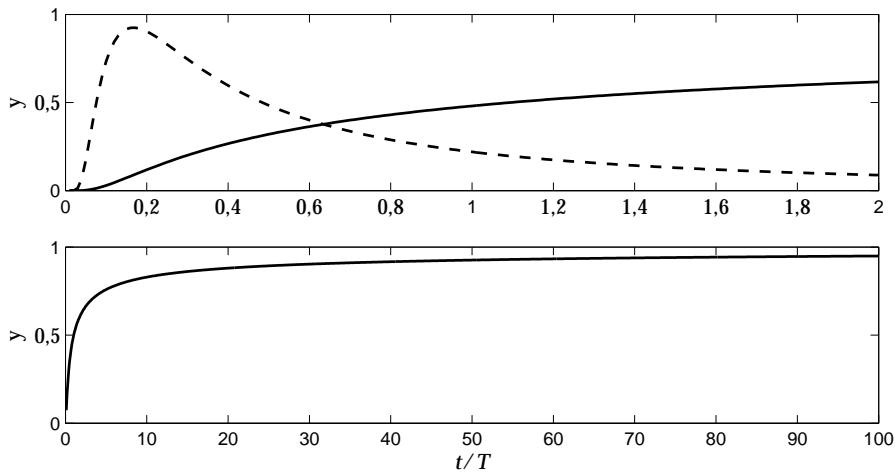


Figura 2.21 Respuesta a un escalón (continua) y respuesta impulsional (a tramos) para un sistema con la función de transferencia $e^{-\sqrt{sT}}$. Las curvas superiores muestran las respuestas a un escalón e impulsional. La curva inferior muestra la respuesta a un escalón en una escala de tiempo diferente.

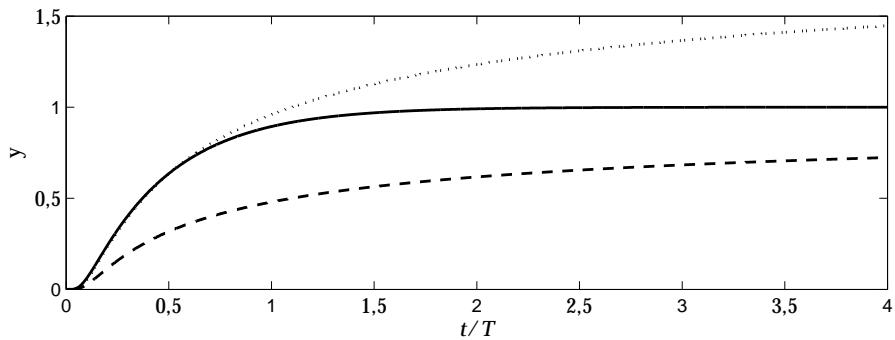


Figura 2.22 Respuestas a un escalón para la función de transferencia $1/\cosh \sqrt{sT}$ (continua), $e^{-\sqrt{sT}}$ (a tramos) y $2e^{-\sqrt{sT}}$ (punteada).

Investigaremos ahora las respuestas en frecuencia de los sistemas (2.32) y (2.35). Se puede demostrar que ambas funciones de transferencias tienen un desfase de 180° en la misma frecuencia.

$$\omega_{180} = \frac{2\pi^2}{T}. \quad (2.37)$$

La magnitud de las funciones de transferencias en ω_{180} están dadas por

$$\begin{aligned} |e^{-\sqrt{i\omega_{180} T}}| &= e^{-\pi} \approx 0,04321 \\ \frac{1}{|\cosh \sqrt{i\omega_{180} T}|} &= \frac{2e^{-\pi}}{1+e^{-2\pi}} \approx 0,08627. \end{aligned}$$

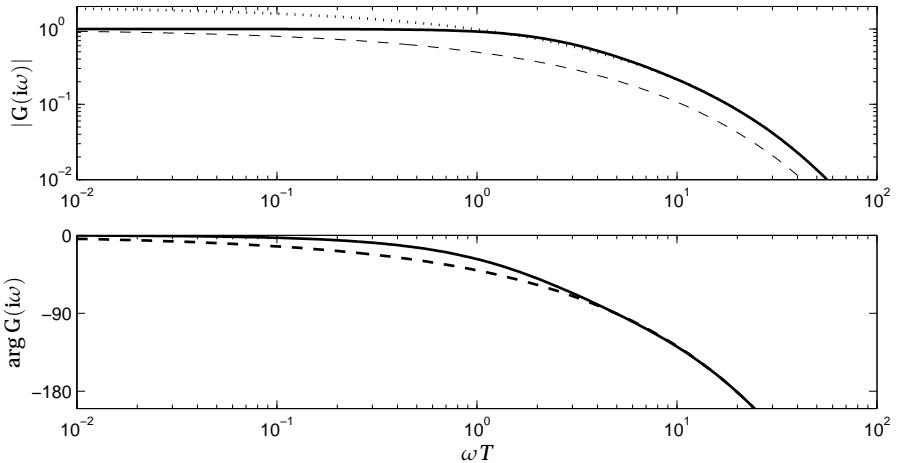


Figura 2.23 Diagrama de Bode para las funciones de transferencia $1/\cosh \sqrt{sT}$ (líneas continuas), $e^{-\sqrt{sT}}$ (a tramos). Para comparación también mostramos la curva de ganancia para la función de transferencia $2e^{-\sqrt{s}}$ (punteada).

En la frecuencia donde el desfase es 180° la ganancia del sistema (2.35) es así muy próxima al doble de la ganancia para el sistema (2.32). En la Figura 2.23 se muestran los diagramas de Bode del sistema. Obsérvese que para frecuencias por encima de 2 rad/s hay muy pequeñas diferencias entre las funciones de transferencia $2e^{-\sqrt{s}}$ y $1/\cosh \sqrt{sT}$ incluso si las diferencias de la respuesta en escalón difieren significativamente. Esta observación es muy importante para el diseño de sistemas de control. La Figura 2.22 también muestra que las respuestas a un escalón para las funciones de transferencias $2e^{-\sqrt{s}}$ y $1/\cosh \sqrt{sT}$ están muy próximas.

Un intercambiador de calor

la función de transferencia de la temperatura de entrada a la temperatura de salida de un intercambiador de calor ideal es

$$G(s) = \frac{1}{sT}(1 - e^{-sT}). \quad (2.38)$$

En la Figura 2.24 se muestran las respuestas a un escalón e impulsional de este sistema. Obsérvese que la respuesta a un escalón se asienta en el valor final en el tiempo $t = T$ y que la respuesta impulsional es cero después de ese tiempo. Esto refleja el hecho de que una vez que el líquido ha pasado a través del intercambiador de calor su temperatura no se ve ya influenciada. El tiempo de residencia medio del sistema es

$$T_{ar} = \frac{T}{2}.$$

La respuesta en frecuencia del sistema es

$$G(i\omega) = \frac{1}{i\omega T}(1 - e^{-i\omega T}).$$

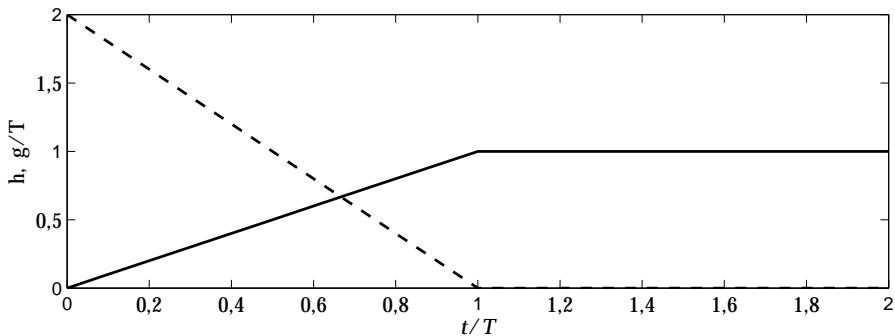


Figura 2.24 Respuesta a un escalón (continua) e impulsional (a tramos) normalizadas para la función de transferencia (2.38) de un intercambiador de calor ideal.

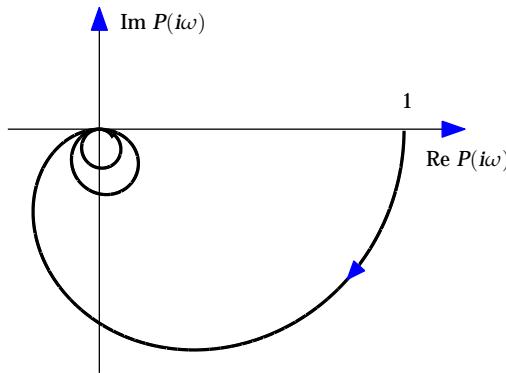


Figura 2.25 Diagrama de Nyquist de la función de transferencia $G(s)$ (2.38) de un intercambiador de calor ideal.

La función de transferencia es cero para $\omega T = 2n\pi$. Esto se ve claramente en el diagrama de Nyquist de la función de transferencia en la Figura 2.25. Una propiedad interesante de esta función de transferencia es que

$$\arg G(i\omega) = -\frac{\omega T}{2}, \quad \text{for } \omega T < 2\pi.$$

Un reactor de tanque agitado continuo

Considérese un reactor de tanque agitado de tiempo continuo donde tiene lugar la reacción $\mathcal{A} \rightarrow \mathcal{B}$. La reacción es exotérmica, y el calor de reacción se elimina mediante un refrigerante. El sistema se modela por balances de masa y energía. El balance de masa es

$$\frac{dc}{dt} = \frac{q}{V}(c_f - c) - k(T)c, \quad (2.39)$$

donde $c [\text{kmol}/\text{m}^3]$ es la concentración de las especies \mathcal{A} , c_f la concentración de \mathcal{A} en la alimentación, $q [\text{m}^3/\text{s}]$ el caudal volumétrico, $V [\text{m}^3]$ el volumen del reactor,

Cuadro 2.2 Parámetros del tanque agitado de tiempo continuo exotérmico.

q	$0,002 \text{ m}^3/\text{s}$	T_f, T_c	300 K
V	$0,1 \text{ m}^3$	E/R	8750 K
ρ	1000 kg/m^3	c_f	2 kmol/m^3
C_p	4 kJ/kgK	UA	50 W/K
k_0	$3 \times 10^8 \text{ s}^{-1}$	ΔH	$-5 \times 10^5 \text{ kJ/kmol}$

y $k(T) [\text{s}^{-1}]$ la velocidad de reacción que es una función de la temperatura

$$k(T) = k_0 e^{-E/RT}. \quad (2.40)$$

El primer término del lado derecho representa el caudal másico y el segundo la velocidad de eliminación de A a través de la reacción.

El balance de energía se puede escribir como

$$\frac{dT}{dt} = \frac{q}{V}(T_f - T) + k(T) \frac{-\Delta H}{\rho C_p} c + \frac{UA}{\rho V C_p} (T_c - T), \quad (2.41)$$

donde $\Delta H [\text{kJ/kmol}]$ es el calor de reacción, $\rho [\text{kg/m}^3]$ la densidad de las especies A , $C_p [\text{kJ/kgK}]$ el calor específico, $U [\text{J/min/K/m}^2]$ el coeficiente de transferencia de calor, $A [\text{m}^2]$ el área, $T_c [\text{K}]$ la temperatura de refrigeración, y $T_f [\text{K}]$ la temperatura de alimentación. El primer término del lado derecho representa el caudal de energía del sistema, el segundo término representa la potencia generada por la reacción, y el último término representa la velocidad de eliminación de energía mediante refrigeración. En la Tabla 2.2 se dan parámetros típicos.

En primer lugar analizaremos las soluciones en estado estacionario. En estado estacionario se sigue de (2.40) que

$$c = \frac{1}{1 + V k(T)/q} c_f.$$

La potencia generada por la reacción es

$$P_g = \frac{k(T)}{1 + V k(T)/q} (-\Delta H) c, \quad (2.42)$$

y la velocidad de eliminación de energía es

$$P_r = \frac{q \rho C_p}{V} (T - T_f) + UA(T - T_c). \quad (2.43)$$

Igualando P_g y P_r da una ecuación en una variable para determinar la temperatura de reacción T . Tal como se ilustra en la Figura 2.26, una solución gráfica que muestra P_g y P_r como funciones de la temperatura, da una visión esclarecedora.

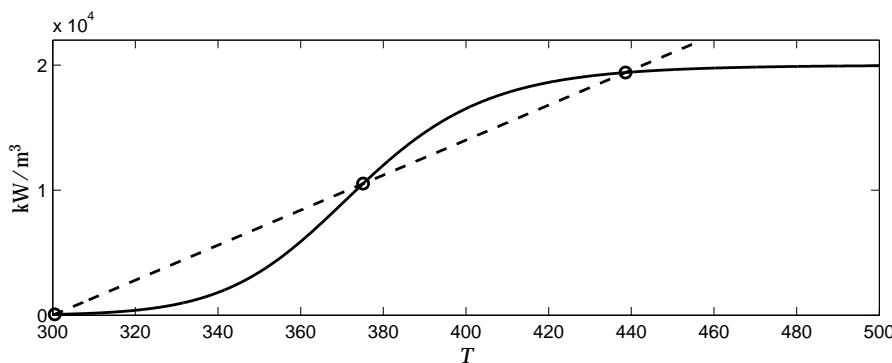


Figura 2.26 Velocidad de generación de calor en estado estacionario (continua) y velocidad de eliminación de calor (a tramos) en función de la temperatura. Los equilibrios se marcan con \circ .

Hay tres equilibrios donde las curvas intersectan en $T = 300,5, 375,1$, y $438,6$. El equilibrio en $T = 375,1$ es inestable porque la velocidad de generación de calor es mayor que la velocidad de eliminación de calor si se aumenta la temperatura. Los otros equilibrios son estables. Aproximando la dinámica en la cercanía del equilibrio inestable da el siguiente modelo lineal del sistema

$$\begin{aligned}\frac{dx_1}{dt} &= -0,0422x_1 + 0,0013x_2 \\ \frac{dx_2}{dt} &= 2,7746x_1 - 0,0064x_2 + 0,15u,\end{aligned}\tag{2.44}$$

donde $x_1 = c - c_0$, $x_2 = T - T_0$, y $u = T_c - T_{c0}$ y c_0 , T_0 , y T_{c0} son los valores de equilibrio. La función de transferencia es

$$P(s) = -\frac{0,15s + 0,0063}{s^2 + 0,048631s - 0,003359} = -0,15 \frac{s + 0,04220}{(s + 0,08717)(s - 0,03854)}$$

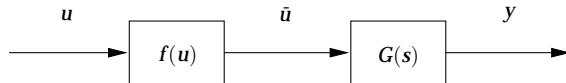
El sistema tiene el polo $s = 0,03854$ en el semiplano derecho.

Modelos negros no lineales

El modelo estático analizado en la Sección 2.2 podría ser no lineal. Sin embargo los modelos dinámicos vistos hasta ahora, han sido lineales. Como en la práctica las no linealidades son usuales resulta muy deseable tener modelos no lineales. Las válvulas, los actuadores los sensores, y la propia dinámica del propio proceso pueden también ser no lineales. Modelos generales para dinámica no lineal son complicados, y no hay buenos métodos para diseñar controladores PID para tales sistemas.

Afortunadamente, hay clases especiales de modelos que están bien adaptados para control PID. Un sistema se puede representar como una combinación de una no linealidad estática y un sistema dinámico lineal. Tales modelos son bastante simples, y se adecúan muy bien para el control PID, pero hay sistemas no lineales que no se pueden modelar bien utilizando este enfoque.

a)



b)

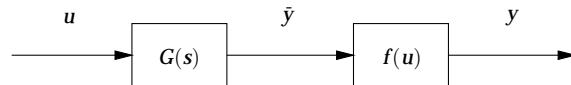


Figura 2.27 Un modelo de Hammerstein a), y un modelo de Wiener b).

La no linealidad puede estar delante de la parte lineal tal como se muestra en la Figura 2.27a. Este modelo se llama un modelo Hammerstein. Es un buen modelo para un sistema con un actuador no lineal, por ejemplo, una válvula no lineal.

La no linealidad se puede colocar también después del sistema dinámico lineal. Esto da un modelo de Wiener, que se ilustra en el diagrama de bloques en la Figura 2.27b. El modelo de Wiener es una buena representación para un sistema con un sensor no lineal, por ejemplo, un electrodo de pH.

Si el proceso es no lineal la dinámica está variando con las condiciones operativas. Idealmente, el controlador debería sintonizarse con respecto a estas variaciones. Un enfoque conservador es sintonizar el controlador para el caso más desfavorable y aceptar un comportamiento degradado en otras condiciones de operación. Otro enfoque es encontrar una variable medible que esté bien correlacionada con la no linealidad del proceso. Tal variable se denomina una variable de planificación. El controlador se sintoniza entonces para unos pocos valores de la variable de planificación. Los parámetros del controlador para valores intermedios se pueden obtener por interpolación. Este enfoque para generar un controlador no lineal se llama planificación de ganancia. Se analizará con más detalle en la Sección 9.3.

Es fácil compensar la no linealidad para un sistema que está descrito por un modelo de Wiener o un modelo de Hammerstein utilizando un controlador no lineal compuesto de un controlador PID y una no linealidad estática. El controlador PID lineal se diseña como si el sistema fuese lineal. Cuando el proceso tiene una no linealidad en la entrada simplemente pasamos la señal de control a través de la inversa de la no linealidad. Si la no linealidad está en la salida, como sucede en el modelo de Wiener, pasamos entonces la señal del sensor por la inversa de la no linealidad antes de alimentar la señal de medida al controlador. Muchos controladores PID tienen la capacidad de poder introducir una no linealidad caracterizada como una función lineal a tramos.

2.6 Modelos para perturbaciones

Hasta ahora, sólo hemos analizado modelos de dinámicas de procesos. Las perturbaciones son otro aspecto importante del problema de control. De hecho, sin perturbaciones ni incertidumbres de procesos no habría necesidad de realimentación. Hay una rama especial del control, la teoría de control estocástico, que trata explícitamente con las perturbaciones. Esto ha tenido poco impacto sobre la sintonía y diseño de controladores PID. En el control PID, las perturbaciones se han considerado en la mayoría de los casos de forma indirecta, e.g., al introducir acción integral. Cuando aumentan nuestras exigencias y nos empeñamos en controlar los sistemas con mejores prestaciones será útil considerar las perturbaciones explícitamente. En esta sección, por lo tanto, presentaremos algunos modelos que se pueden utilizar con esta finalidad. Los modelos para perturbaciones son útiles para simulación, diagnóstico y evaluación de comportamiento.

La naturaleza de las perturbaciones

Distinguimos entre tres tipos de perturbaciones, es decir, cambios en el punto de consigna, perturbaciones en la carga, y ruido de medida. En control de procesos, la mayoría de los lazos de control tienen puntos de consigna que permanecen constante durante largos períodos de tiempo con cambios ocasionales. Un modelo apropiado es por lo tanto una señal constante a tramos. Los cambios en el punto de consigna típicamente se conocen de antemano. Buena respuesta a los cambios en el punto de consigna es uno de los temas fundamentales en el caso de los servosistemas.

Las perturbaciones de carga son perturbaciones que entran en el lazo de control en alguna parte del proceso y llevan al sistema fuera de su punto de operación deseado. Las perturbaciones de carga normalmente tienen baja frecuencia. La reducción eficiente de las perturbaciones de carga es un tema clave en los sistemas de control de procesos.

El ruido de medida representa perturbaciones que distorsionan la información acerca de las variables del proceso obtenidas de los sensores. El ruido de medida es a menudo una perturbación de alta frecuencia. Con frecuencia se intenta filtrar las señales medidas para reducir el ruido de medida. El filtraje, sin embargo, añade dinámica al sistema.

El carácter de las perturbaciones

Una forma de obtener una estimación de las perturbaciones es registrar la variable medida. La señal medida tiene ambas contribuciones de la perturbación de carga y del ruido de medida. Si hay grandes variaciones es a menudo útil investigar el sensor para reducir algo el ruido de medida. El filtraje también puede ser útil. El filtraje debería hacerse de tal forma que no deteriore el control.

Las variaciones del proceso pueden tener un carácter muy diferente. En la Figura 2.28 se dan algunos ejemplos. Las perturbaciones se pueden clasificar como pulsos (a), escalones (b), rampas (c), y periodica (d). Es útil calcular estadísticas tales como valores medios, varianzas, y desviación máxima. Es también útil representar un histograma de la distribución de amplitud de las perturbaciones.

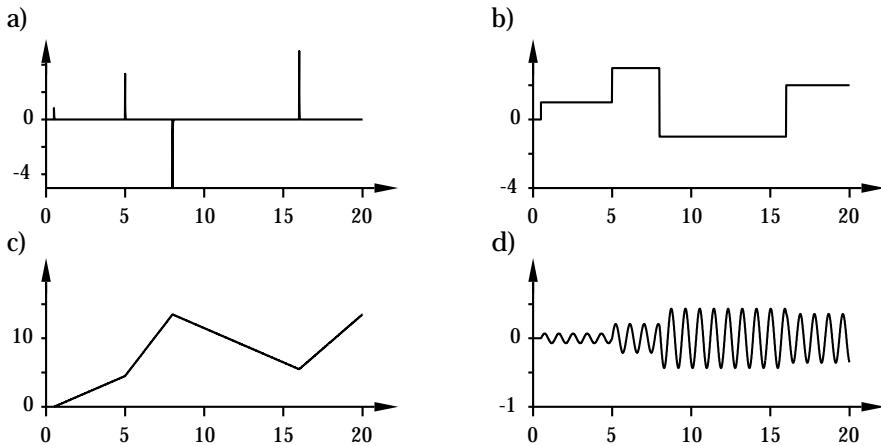


Figura 2.28 Diferentes tipos de perturbaciones: a) impulsos, b) escalones, c) rampas, y d) sinusoides.

Modelos sencillos

Es útil tener modelos sencillos de las perturbaciones para simulación y evaluación de estrategias de control. En la Figura 2.28 se muestran modelos que se emplean típicamente. El impulso es una idealización matemática de un pulso cuya duración es corta en comparación con la escala de tiempo. Las señales son esencialmente determinísticas. Los únicos elementos inciertos en el impulso, escalón, y rampa son los tiempos de comienzo y la amplitud de la señal. Los elementos inciertos de la sinusode son frecuencia, amplitud, y fase.

Fluctuaciones aleatorias

Las perturbaciones se pueden hacer también más irregulares tal como se muestra en la Figura 2.29. Hay conceptos y técnicas bien desarrolladas para tratar con fluctuaciones aleatorias que se describen como procesos estocásticos. Existen ambas características en el dominio del tiempo y en el dominio de la frecuencia. En el dominio de la frecuencia las perturbaciones aleatorias se caracterizan por la función densidad espectral $\phi(\omega)$. La varianza de la señal está dada por

$$\sigma^2 = \int_{-\infty}^{\infty} \phi(\omega) d\omega.$$

La densidad espectral dice cómo se distribuye en diferentes frecuencias la variación de la señal. El valor

$$2\phi(\omega)\Delta\omega$$

es la energía media en una banda estrecha de anchura $\Delta\omega$ centrada alrededor de ω . Una señal donde $\phi(\omega)$ es constante se llama ruido blanco. Tal señal tiene su energía igualmente distribuida en todas las frecuencias .

Hay técnicas eficientes para calcular la densidad espectral de una función dada. Si se conoce la densidad espectral es posible evaluar cómo las variaciones

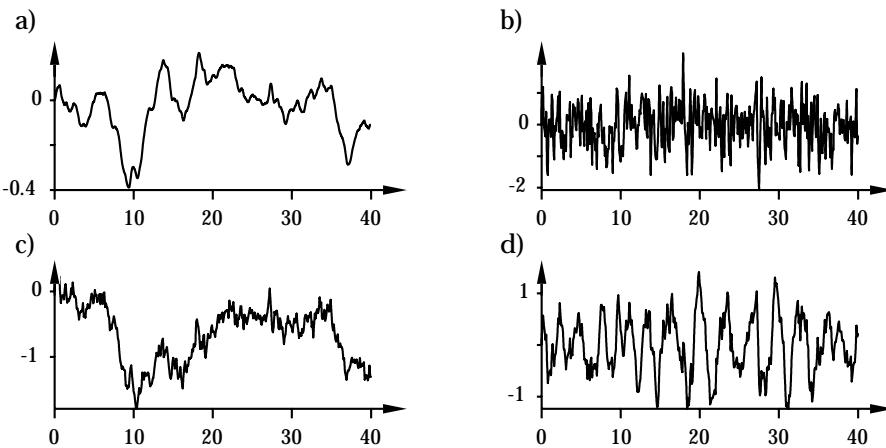
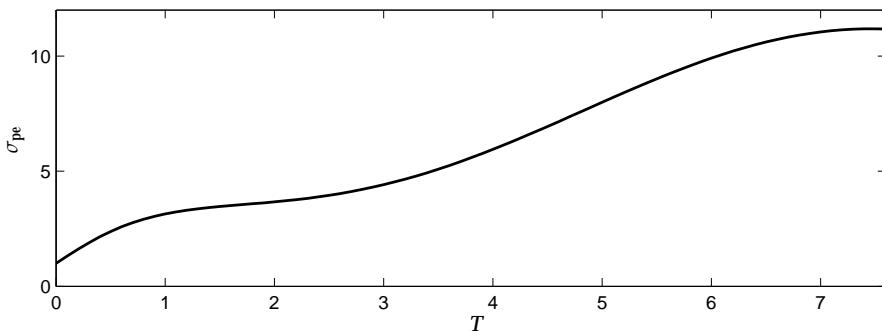


Figura 2.29 Ejemplos de perturbaciones estocásticas.

Figura 2.30 Error de predicción σ_{pe} en función del tiempo de predicción T_p .

en la variable del proceso se ven influenciadas por las diferentes estrategias de control.

Predicción de perturbaciones

Cuando se controlan variables de calidad importantes en un proceso es a menudo de interés evaluar las mejoras que se pueden alcanzar y determinar si una estrategia de control en particular tiene un comportamiento que está próximo a los límites alcanzables. Esto se puede hacer como sigue. Se registra la variable de proceso $y(t)$ durante la operación normal con o sin control. Al analizar las fluctuaciones es posible determinar con cuánta precisión se puede predecir la variable de proceso T_p unidades de tiempo en el futuro basado en los valores pasados y presente de y . Sea $\hat{y}(t+T_p|t)$ la mejor predicción de $y(t+T_p)$ basada en $y(\tau)$ para todo $\tau < t$. Al representar la varianza del error de predicción $y(t+T_p) - \hat{y}(t+T_p|t)$ en función del tiempo de predicción obtenemos la curva que se muestra en la Figura 2.30. Para grandes tiempos de predicción el error de predicción es igual a la varianza de la variable de proceso, aproximadamente $\sigma_{pe} = 12$ en la figura. El mejor error de control que se puede alcanzar es el error de

predicción en un tiempo de predicción T_p que corresponde al retardo de tiempo del proceso y al tiempo de muestreo del controlador. Esto se puede lograr con el denominado controlador de varianza mínima. Ver Sección 8.6. La figura muestra que se pueden lograr varianzas menores que 5 si T_p es menor que 3.4. Más aún son posibles reducciones para T_p más pequeños, pero varianzas menores que 1 no se pueden lograr incluso si T_p es muy corto. Al comparar esto con la varianza actual tenemos una evaluación del comportamiento alcanzable. Esto se analiza con más detalle en el Capítulo 10. Hay software eficiente para calcular el error de predicción y su varianza a partir de los datos del proceso.

2.7 Cómo obtener los modelos

En secciones previas hemos mencionado brevemente cómo se pueden obtener los modelos. En esta sección daremos una explicación más detallada de los métodos para determinar los modelos. Hay dos grandes tipos de métodos que se pueden usar. Uno es el modelado físico, y el otro es el modelado a partir de los datos.

El modelado físico usa primeros principios para derivar las ecuaciones que describen el sistema. Las leyes físicas expresan conservación de masa, momento, y energía. Se combinan con las ecuaciones constitutivas que describen las propiedades del material. Cuando se derivan modelos físicos un sistema se divide típicamente en subsistemas. Se determinan las ecuaciones de cada subsistema, y se combinan los resultados para obtener un modelo para el sistema completo. En la Sección 2.3 se dan ejemplos sencillos. El modelado físico es a menudo muy costoso en tiempo. Con frecuencia hay que tomar decisiones difíciles sobre las aproximaciones que son apropiadas. Los modelos obtenidos pueden sin embargo, ser muy útiles ya que tienen una base física sólida. También dan una idea considerable sobre la dependencia del modelo respecto de los parámetros físicos. Una forma sencilla para comenzar es modelar las dinámicas como sistemas de primer orden donde las constantes tiempo son la razón de almacenamiento y caudal.

El modelado de los datos es un procedimiento experimental. Los datos se generan perturbando la señal de entrada (la variable manipulada) y registrando la salida del sistema. El experimento se puede efectuar también bajo condiciones de lazo cerrado, por ejemplo, perturbando el punto de consigna del controlador o la salida del controlador. Se intenta entonces encontrar un modelo que se ajuste bien a los datos. Hay algunas cuestiones importantes que se deben considerar; selección de las señales de entrada, selección de una estructura apropiada del modelo, ajuste de los parámetros, y validación del modelo. Idealmente, las condiciones experimentales deberían escogerse de forma que sean lo más similares posibles al uso que se pretende del modelo. El ajuste de parámetros se puede hacer manualmente para modelos aproximados o utilizando técnicas de optimización.

Modelos estáticos

Los modelos estáticos son muy fáciles de obtener observando la relación entre la entrada y la salida en estado estacionario. Para procesos estables bien amortiguados, se puede obtener la relación fijando la entrada a un valor constante y observando la salida en el estado estacionario. El procedimiento se repite entonces

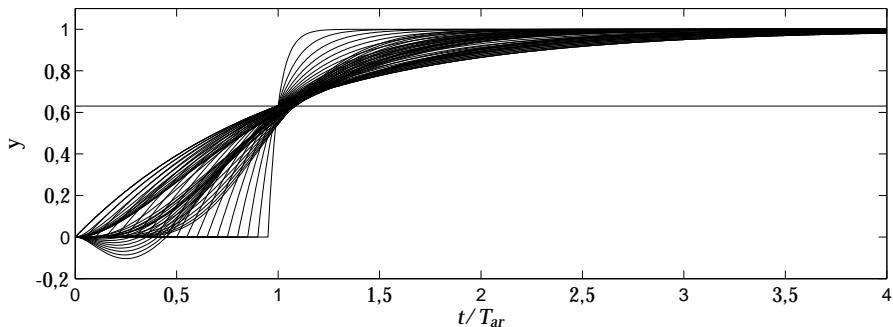


Figura 2.31 Respuestas a un escalón para un gran lote de sistemas estables. Las respuestas han sido normalizadas para dar el mismo tiempo de residencia medio.

para diferentes valores de la entrada hasta que se cubre todo el rango. Para los sistemas que tienen un integrador es conveniente usar un controlador para mantener la salida en un valor constante. Se cambia entonces el punto de consigna de manera que se cubra el rango completo. Los efectos de las perturbaciones se pueden reducir tomando medias.

El test del salto en escalón

El test del salto en escalón es un procedimiento sencillo que se usa normalmente en control de procesos. Se basa en una determinación experimental de la respuesta en escalón. Para realizar el experimento el sistema se lleva primero al estado estacionario. La variable manipulada se cambia rápidamente a un nuevo valor constante y se registra la salida. Los datos medidos se escalan para corresponder a un escalón unitario. El cambio en la variable manipulada debería ser grande con el fin de obtener una buena razón señal-ruido, pero no debería ser tanto como para que la conducta del proceso sea no lineal. La magnitud permitida está también limitada por la operación del proceso. Es también útil registrar las fluctuaciones en la señal medida cuando la señal de control es constante. Esto da información acerca del ruido del proceso. Es buena práctica repetir el experimento para diferentes amplitudes de la señal de entrada y en distintas condiciones de operación. Esto proporciona una indicación de los rangos de señal cuando el modelo es lineal. También indica si el proceso cambia con las condiciones de operación.

Por inspección de la respuesta en escalón es posible hacer una clasificación cruda de la dinámica del sistema en las categorías que se muestran en la Figura 2.2. Se ajusta entonces un modelo con pocos parámetros a los datos.

El tiempo de residencia medio

El tiempo de residencia medio es una forma sencilla de caracterizar el tiempo de respuesta de sistemas con respuestas a un escalón esencialmente monótonas. La Figura 2.31 muestra las respuestas a un escalón para un gran lote de sistemas que se normalizan para dar el mismo tiempo de residencia medio. (En la Sección 7.1 se dan las funciones de transferencia de los sistemas.) La figura muestra que todas las respuestas a un escalón están próximas para $t = T_{ar}$. Para todos los proce-

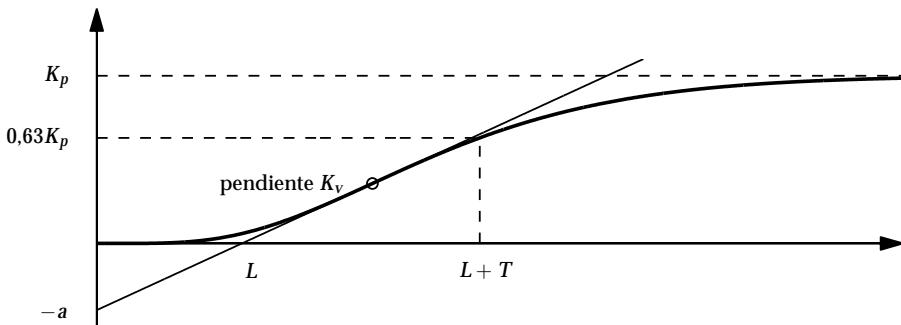


Figura 2.32 Respuesta a un escalón unitario de un proceso y un procedimiento usado para determinar los parámetros del proceso K_p , L , T , y K_v de un modelo FOTD. El punto de mayor pendiente se denota por \circ .

sos en el lote del test tenemos $0,99 < T_{63}/T_{ar} < 1,08$. El tiempo de residencia medio se puede así estimar como el tiempo T_{63} donde la respuesta a un escalón ha alcanzado el 63 por ciento de su valor final.

El modelo FOTD

Los parámetros del modelo FOTD dados por la Ecuación 2.18 se pueden determinar a partir de un test en escalón tal como se ilustra en la Figura 2.32. La ganancia estática K_p se determina simplemente a partir de los valores en estado estacionario de las señales antes y después del cambio en escalón. El retardo de tiempo aparente L está dado por el punto donde la tangente de mayor pendiente interseca al nivel de estado estacionario antes del cambio en escalón. El tiempo de residencia medio $T_{ar} = T + L$ se determina como el tiempo T_{63} donde la respuesta a un escalón ha alcanzado el 63 por ciento de su valor final en estado estacionario. Esto da los resultados correctos para el modelo-FOTD (ver Figura 2.11) y resultados aproximados para muchos otros modelos (ver Figuras 2.13, 2.14, 2.16, y 2.31). La ganancia de velocidad K_v es la tangente de mayor pendiente.

Se pueden utilizar métodos similares cuando la señal de entrada es un pulso en lugar de un escalón. Los pulsos se pueden emplear cuando no está permitido usar un escalón. Esto es común en aplicaciones médicas y biológicas y es menos usual en control de procesos. El análisis de respuesta en rampa es normal cuando se estudian servosistemas y sistemas hidráulicos.

La aproximación del integrador y el retardo de tiempo

Los parámetros del modelo (2.28), que tiene la función de transferencia

$$G(s) = \frac{K_v}{s} e^{-sL} = \frac{a}{sL} e^{-sL}, \quad (2.45)$$

se puede determinar también de un test en escalón tal como se indica en la Figura 2.32. La constante de velocidad K_v es la mayor pendiente de la respuesta en escalón y las intersecciones de esta tangente con los ejes vertical y horizontal dan a y L , respectivamente. El modelo dado por la Ecuación 2.45 es la base del procedimiento de sintonía de Ziegler-Nichols analizado en el Capítulo 6.

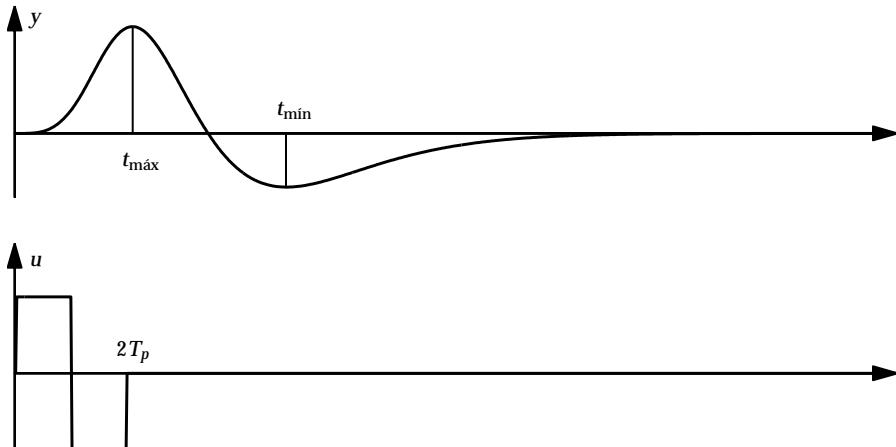


Figura 2.33 Determinación de los parámetros de un modelo FOTD excitando el proceso por un pulso doble.

El método del doble pulso

Una variación del test en escalón es excitar el proceso por un pulso doble tal como se ilustra en la Figura 2.33. La amplitud del pulso a se escoge de forma que la respuesta está bastante por encima del nivel de ruido, y la anchura del pulso T_p se selecciona un poco más grande que el retardo de tiempo del proceso. Se determinan el máximo y_{\max} y el mínimo y_{\min} y los tiempo t_{\max} y t_{\min} cuando ocurren. Cálculos sencillos muestran que para un sistema FOTD con la función de transferencia (2.10) tenemos

$$\begin{aligned} y_{\max} &= aK_p(1 - e^{-T_p/T}) \\ y_{\min} &= -aK_p(1 - e^{-T_p/T})^2 \\ t_{\max} &= L + T_p \\ t_{\min} &= L + 2T_p. \end{aligned}$$

Se sigue de estas ecuaciones que

$$\begin{aligned} \frac{y_{\min}}{y_{\max}} &= -1 + e^{-T_p/T} \\ \frac{y_{\max}^2}{y_{\min}} &= -aK_p, \end{aligned}$$

y obtenemos las siguientes ecuaciones sencillas para los parámetros del modelo

$$\begin{aligned} K_p &= -\frac{y_{\max}^2}{ay_{\min}} \\ T &= \frac{T_p}{\log(1 + y_{\max}/y_{\min})} \\ L &= t_{\max} - T_p \\ L &= t_{\min} - 2T_p. \end{aligned} \tag{2.46}$$

El hecho de que el retardo de tiempo L se pueda estimar de dos formas se puede emplear para evaluar si un proceso se puede modelar por un modelo FOTD.

La selección del tiempo del pulso T_p se puede determinar automáticamente, por ejemplo, como el tiempo en el que la salida ha cambiado una cantidad especificada. El método se puede aplicar a modelos SOTD, pero las fórmulas son más complicadas.

Las ventajas principales de usar un pulso doble son que la salida del proceso retorna a su valor original después de la perturbación, y que el tiempo requerido para determinar la dinámica es corto porque no es necesario esperar al estado estacionario como sucede con el test en escalón. Las desventajas del método son que es difícil determinar de forma precisa los tiempos en los que ocurren los extremos y que la estima de la ganancia no es buena principalmente porque la excitación del pulso está en el régimen de alta frecuencia. Otra desventaja es que el método no se puede aplicar a sistemas oscilatorios.

El modelo SOTD

El modelo

$$G(s) = \frac{K_p}{(1+sT_1)(1+sT_2)} e^{-sL_1}, \quad (2.47)$$

que es una generalización natural del modelo FOTD (2.18), se llama el modelo de segundo orden con retardo de tiempo o modelo SOTD. Sin pérdida de generalidad se puede suponer que $T_2 \leq T_1$. La respuesta a un escalón del sistema (2.47) es

$$y(t) = \begin{cases} K_p \left(1 - \frac{T_1}{T_1 - T_2} e^{-(t-L_1)/T_1} - \frac{T_2}{T_2 - T_1} e^{-(t-L_1)/T_2} \right) & \text{si } T_1 \neq T_2 \\ K_p \left(1 - e^{-(t-L_1)/T_1} - \frac{t}{T_1} e^{-(t-L_1)/T_1} \right) & \text{si } T_1 = T_2. \end{cases} \quad (2.48)$$

En la Figura 2.34 se muestran las respuestas a un escalón normalizadas para diferentes razones de T_2/T_1 . Las respuestas han sido normalizadas de forma que todos los sistemas tienen el mismo tiempo de residencia medio. Todas las respuestas a un escalón están bastante próximas y son casi idénticas para $t/(T_1 + T_2) \approx 1.3$. Como la separación de las curvas es tan pequeña es difícil determinar los parámetros T_1 y T_2 robustamente a partir de la respuesta a un escalón, particularmente si hay una pequeña cantidad de ruido. Son necesarias otras entradas que excitan al sistema mejor para determinar fiablemente los parámetros. La figura muestra que sería más fácil determinar los parámetros basándose en una respuesta impulsional, que se podría obtener diferenciando la respuesta a un escalón.

Un sistema oscilatorio

El modelo (2.18) no puede modelar sistemas con respuestas oscilatorias. Un modelo sencillo para tales sistemas está dado por la función de transferencia (2.27), que tiene tres parámetros: la ganancia estática K_p , la frecuencia natural no amortiguada $1/T$, y el amortiguamiento relativo ζ . Estos parámetros se pueden determinar aproximadamente de la respuesta en escalón tal como se muestra en la

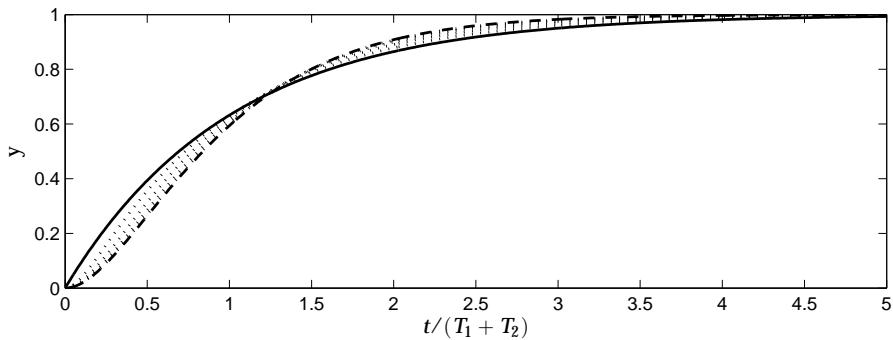


Figura 2.34 Respuestas a un escalón normalizadas para el sistema (2.48) para $T_2/T_1 = 0.1, \dots, 1$.

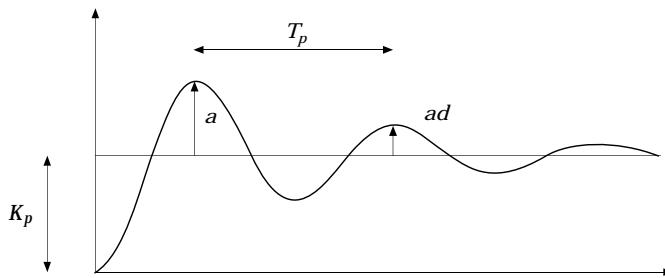


Figura 2.35 Determinación gráfica de modelos matemáticos para sistemas con una respuesta en escalón oscilatoria.

Figura 2.35. Los parámetros T y ζ están relacionados con el período de tiempo T_p y la razón de decaimiento d como sigue.

$$d = e^{-2\zeta\pi/\sqrt{1-\zeta^2}} \quad T_p = \frac{2\pi T}{\sqrt{1-\zeta^2}} \quad (2.49)$$

o

$$\zeta = \frac{1}{\sqrt{1 + (2\pi/\log d)^2}} \quad T = \frac{\sqrt{1-\zeta^2}}{2\pi} T_p. \quad (2.50)$$

La precisión del modelo está condicionada por la limitada excitación obtenida con un escalón o un pulso. Los errores de medida y la dificultad en obtener estados estacionarios son otros factores que limitan la precisión. Se pueden hacer algunas mejoras utilizando optimización para ajustar los parámetros. Típicamente, es difícil determinar más de tres parámetros de una respuesta en escalón a menos que las condiciones experimentales sean excepcionales.

Respuesta en frecuencia

En el análisis de la respuesta en frecuencia se introduce una señal sinusoidal, y se analiza la respuesta en estado estacionario. Una ventaja con el análisis de la respuesta en frecuencia es que se pueden hacer medidas muy precisas empleando

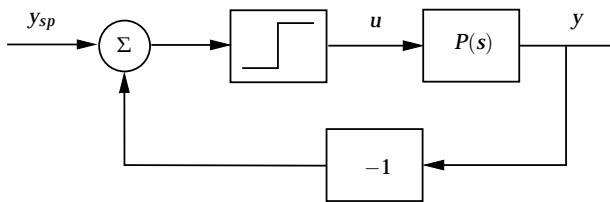


Figura 2.36 Diagrama de bloque de un proceso con realimentación con un relé.

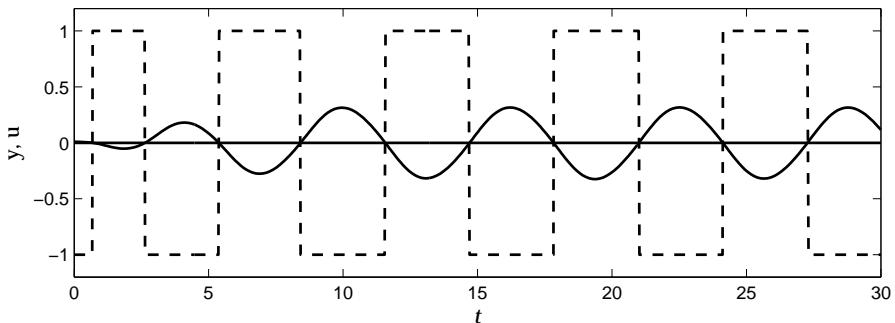


Figura 2.37 Salida del relé u (a tramos) y salida del proceso y (continua) para un sistema con realimentación con relé.

técnicas de correlación. El excesivo tiempo de experimentación es una desventaja.

Es también posible introducir una señal arbitraria como una perturbación. La respuesta en frecuencia se puede obtener como la razón de las transformadas de Fourier de las señales de salida y entrada. Así mismo se pueden ajustar los parámetros de un modelo con una estructura dada a los datos.

Una característica agradable de usar señales distintas a los escalones es que se puede alcanzar un compromiso entre la amplitud de la señal y su duración.

Relimentación con relé

Hay una técnica muy especial que está particularmente adaptada para determinar ω_{180} y K_{180} . Esto ha sido empleado muy efectivamente para sintonizar controladores PID. La idea es la observación de que es posible crear una oscilación con la frecuencia última de forma automática utilizando una realimentación con un relé.

Para hacer el experimento el sistema se conecta en un lazo de realimentación con un relé tal como se muestra en la Figura 2.36. Para muchos sistemas habrá entonces una oscilación (tal como se muestra en la Figura 2.37) donde la señal de control es una onda cuadrada y la salida del proceso está próxima a una sinusoidal. Obsérvese que la entrada y salida del proceso tienen fases opuestas.

Para explicar cómo trabaja el sistema, suponga que la salida del relé se desarrolla en una serie de Fourier y que el proceso atenúa de forma efectiva los armónicos superiores. Es entonces suficiente considerar solamente el componente del primer armónico de la entrada. La entrada y la salida tienen entonces fase

opuesta, lo que significa que la frecuencia de la oscilación es igual a ω_{180} . Si d es la amplitud del relé, el primer armónico de la onda cuadrada de entrada tiene una amplitud de $4d/\pi$. Sea a la amplitud de la salida del proceso. La ganancia del proceso en ω_{180} viene entonces dada por

$$K_{180} = \frac{\pi a}{4d}. \quad (2.51)$$

Obsérvese que el experimento del relé se automatiza fácilmente. Como la amplitud de la oscilación es proporcional a la salida del relé, es fácil controlar ajustando la salida del relé. También se observa en la Figura 2.37 que se establece muy rápidamente una oscilación estable. La amplitud y el período se pueden determinar después de solo unos 20 s, a pesar del hecho de que el sistema comienza lejos del equilibrio lo que le lleva cerca de 8 s para alcanzar el nivel correcto. El tiempo de residencia medio del sistema es de 12 s, lo que significa que tardaría unos 40 s para que la respuesta a un escalón alcanzase el estado estacionario.

El modelo SOTD – Respuesta en frecuencia y a un escalón combinada

Fue mencionado previamente que los parámetros del modelo SOTD no se pueden determinar fiablemente de los datos de la respuesta a un escalón. Sin embargo se pueden obtener buenas estimaciones combinando los datos de respuesta en frecuencia y de respuesta a un escalón. La idea es que la respuesta a un escalón da K_p y T_{63} y el método de respuesta en frecuencia da la frecuencia última $\omega_u = \omega_{180}$ y la ganancia última $K_u = 1/K_{180}$. Esto da las ecuaciones

$$\begin{aligned} K_p^2 K_u^2 &= (1 + \omega_u^2 T_1^2)(1 + \omega_u^2 T_2^2) \\ \pi &= \arctan \omega_u T_1 + \arctan \omega_u T_2 + \omega_u L_1. \end{aligned} \quad (2.52)$$

Combinado con los datos K_p y T_{63} los parámetros están entonces dados por las Ecuaciones 2.48 y 2.52 que proporcionan cuatro ecuaciones con cuatro incógnitas.

$$\begin{aligned} 0 &= \begin{cases} 0,37 - \frac{T_1}{T_1 - T_2} e^{-(T_{63}-L_1)/T_1} - \frac{T_2}{T_2 - T_1} e^{-(T_{63}-L_1)/T_2} & \text{si } T_1 \neq T_2, \\ 1 - e^{-(T_{63}-L_1)/T_1} - \frac{T_{63}}{T_1} e^{-(T_{63}-L_1)/T_1} - 0,63 & \text{si } T_1 = T_2 \end{cases} \\ 0 &= (1 + \omega_u^2 T_1^2)(1 + \omega_u^2 T_2^2) - K_p^2 K_u^2 \\ 0 &= \arctan \omega_u T_1 + \arctan \omega_u T_2 + \omega_u L_1 - \pi. \end{aligned} \quad (2.53)$$

Estas ecuaciones se pueden resolver iterativamente, pero esto es complicado ya que tenemos que tener cuidado con los casos especiales cuando los parámetros T_1 y T_2 son iguales o cero.

Un método alternativo es iterar la razón $a = T_2/T_1$ hasta que las ecuaciones coincidan. El parámetro K_p se determina como la ganancia estática de la respuesta a un escalón. La ecuación (2.52) para la ganancia última se hace entonces

$$(1 + \omega_u^2 T_1^2)(1 + a^2 \omega_u^2 T_1^2) = K_p^2 K_u^2.$$

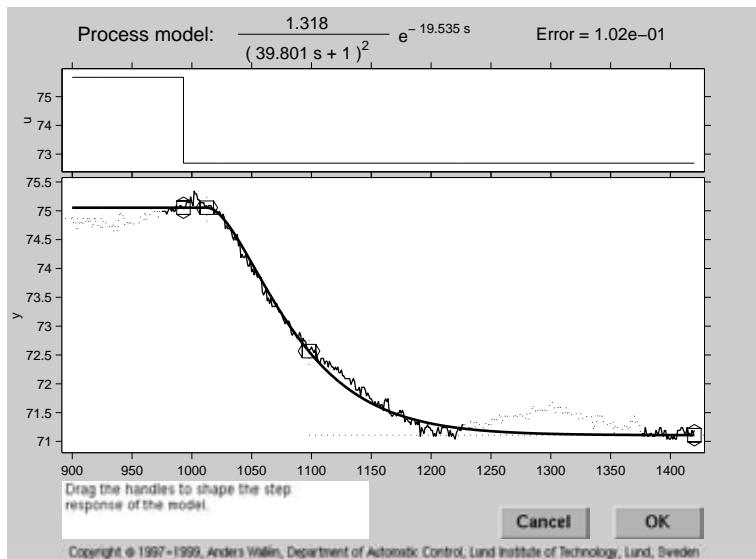


Figura 2.38 Pantalla de computador de una herramienta para modelado de procesos. De [353]

Esta ecuación tiene la solución

$$T_1 = \frac{1}{a\omega_u \sqrt{2}} \sqrt{\sqrt{4a^2 K_p^2 K_u^2 + (1 - a^2)^2} - 1 - a^2}.$$

Los parámetros T_2 y L_1 están entonces dados por

$$T_2 = aT_1$$

$$L_1 = \frac{\pi - \arctan \omega_u T_1 - \arctan \omega_u T_2}{\omega_u}.$$

La respuesta a un escalón (2.48) se puede calcular entonces como una función de a , y el parámetro a se puede iterar para ajustarse al valor de T_{63} .

Herramientas de modelado

Hay algunas herramientas de modelado que son muy útiles. Permiten introducir los datos del proceso en la forma de secuencias de datos de entrada-salida obtenidos de los tests en escalón o de otros experimentos del proceso. Se pueden seleccionar modelos con diferentes estructuras y sus parámetros se pueden ajustar a los datos utilizando algún procedimiento de optimización. Las herramientas también permiten la selección de parte de las secuencias de datos usadas en el análisis.

La Figura 2.38 muestra la pantalla de un computador para un sistema particular. Se puede escoger una estructura de modelo de un menú. Cuando se han introducido los datos se puede ajustar de forma manual un modelo preliminar arrastrando los marcadores que se muestran en la figura. Los marcadores representan el comienzo del escalón, el nivel inicial, el nivel final, y el tiempo cuando

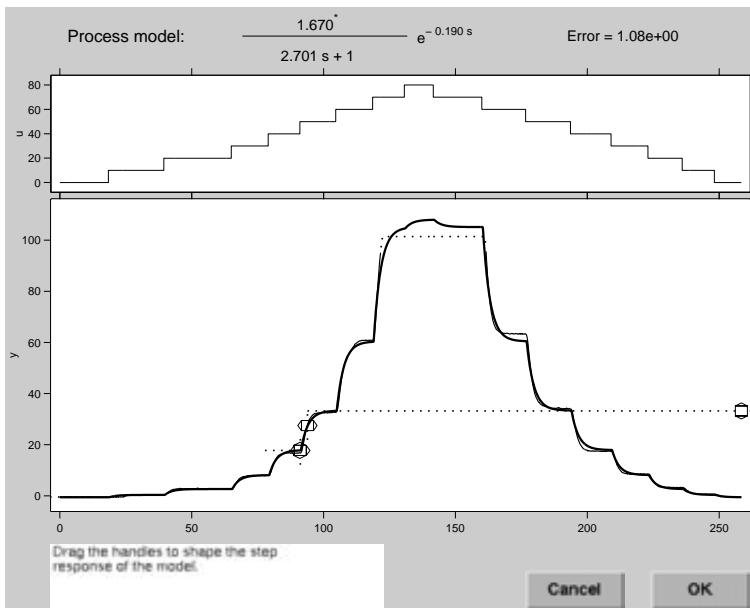


Figura 2.39 Ilustra el modelado no lineal basado en computador. De [353]

la respuesta ha alcanzado el 63 por ciento de su valor final. Se visualizan los parámetros del modelo. Se puede emplear optimización para mejorar el ajuste.

La herramienta particular que se ilustra en la figura también permite usar un modelo no lineal tal como se muestra en la Figura 2.39. En este caso se ajusta en primer lugar un modelo estático a los datos de entrada-salida obtenidos a partir de un experimento estático. A continuación se ajusta un modelo dinámico tal como se indica en la Figura 2.39. Se intentan ambos modelos de Wiener y Hammerstein para ver cuál es el que da el mejor ajuste. El ejemplo particular es un tanque donde el caudal de salida es una función no lineal del nivel. En este caso el modelo de Wiener da el mejor ajuste porque la no linealidad aparece en la salida del sistema. Obsérvese en la Figura 2.39 que los escalones de entrada son de igual tamaño, pero la magnitud de la respuesta de salida cambia significativamente. Estos datos no se pueden ajustar bien mediante un modelo lineal.

Las herramientas interactivas proporcionan una idea muy buena de las relaciones entre los parámetros y la respuesta y la sensibilidad de los parámetros. Es también muy efectiva para combinar un ajuste manual sencillo con optimización numérica. La mayoría de las herramientas permiten ajustar un test en escalón simple, pero hay también herramientas que permiten señales de entrada generales. También es posible determinar las características del ruido y la curva de predicción que se muestra en la Figura 2.30.

2.8 Reducción del modelo

Muchos métodos para la sintonía de controladores PID se basan en modelos sencillos de la dinámica del proceso. Para usar tales métodos es necesario disponer de métodos para simplificar un modelo complicado. Un caso típico es cuando se obtiene un modelo combinando los modelos de subsistemas. Para encontrar aproximaciones adecuadas es necesario especificar el objetivo del modelo. La sintonía de controladores PID puede realizarse especificando el rango de frecuencia de interés. Esto se puede hacer simplemente especificando la frecuencia más alta ω^* donde el modelo es válido. Para control PI la frecuencia ω^* es ω_{145} , la frecuencia donde el desfase del proceso es 145° . La razón para esto es que un controlador PI siempre tiene un desfase. Para un controlador PID, que puede proporcionar adelanto de fase, la frecuencia ω^* se puede escoger como ω_{180} .

La reducción de modelo comienza con un modelo representado por la función de transferencia $G(s)$. La función de transferencia es primero factorizada como

$$G(s) = G_l(s) \frac{1}{1+sT_s} G_h(s). \quad (2.54)$$

El factor de baja frecuencia $G_l(s)$ tiene todos sus polos y ceros y retardos de tiempo en frecuencias alrededor de ω^* o en frecuencias inferiores. El factor de alta frecuencia $G_h(s)$ tiene dinámica en frecuencias superiores a ω^* . La constante de tiempo T_s representa un polo intermedio. La factorización puede siempre hacerse de tal forma que el factor de alta frecuencia $G_h(s)$ tiene la propiedad $G_h(0) = 1$.

Para el diseño de controladores PID el modelo (2.54) será simplificado a

$$\begin{aligned} G(s) &= \frac{K_p}{1+sT} e^{-sL} \\ G(s) &= \frac{K_p}{(1+sT_1)(1+sT_2)} e^{-sL}. \end{aligned} \quad (2.55)$$

La razón para estas elecciones es que hay métodos para diseñar controladores PID para modelos de este tipo. Estos modelos están particularmente adaptados para problemas de control de procesos típicos donde la dinámica tiene esencialmente respuestas a un escalón monótonas.

El factor de baja frecuencia

El factor de baja frecuencia normalmente solo contendrá uno o dos modos. Si el sistema tiene polos múltiples pueden aproximarse por la función de transferencia

$$G_l(s) = \frac{K_p}{1+sT_e} e^{-sL_e}.$$

donde T_e y L_e se obtienen de la Tabla 2.1. De esta forma obtenemos un factor de baja frecuencia de primer o segundo orden, que se requiere para control PID. Si el modelo es más complejo es necesario reducir ω^* o usar un controlador más complejo.

Aproximación de modos rápidos

Hay algunas formas de aproximar los modos rápidos. Una forma sencilla es caracterizar la parte de alta frecuencia por su tiempo de residencia medio T_{arh} . Esto se ilustra por el ejemplo siguiente.

EJEMPLO 2.12—APROXIMACIÓN DE MODOS RÁPIDOS

Sea un sistema donde el factor de alta frecuencia es

$$G_h(s) = \frac{(1+sT_1)(1+sT_2)}{(1+sT_3)(1+sT_4)(1+sT_5)} e^{-sL}.$$

Este sistema tiene el tiempo de residencia medio

$$T_{arh} = T_3 + T_4 + T_5 + L - T_1 - T_2.$$

□

Comparar esto con la Sección 2.4, que muestra cómo se calculan los tiempos de residencia medio. Cuando se utiliza control digital habría que añadir medio período de muestreo a T_{arh} .

Regla de la mitad de Skogestad

Habiendo simplificado los factores de baja y alta frecuencia hemos obtenido un factor de baja frecuencia de la forma dada por (2.55) o (2.58) y una caracterización del factor de alta frecuencia por su tiempo de residencia medio T_{arh} . Skogestad ha sugerido que la constante de tiempo intermedia T_s en (2.54) se aproxima añadiendo $T_s/2$ al retardo de tiempo del modelo y $T_s/2$ a su constante de tiempo. El modelo reducido entonces se hace

$$\begin{aligned} G(s) &= \frac{K_p}{1+s(T+T_s/2)} e^{-s(L+T_{arh}+T_s/2)} \\ G(s) &= \frac{K_p}{(1+sT_1)(1+s(T_2+T_{arh}/2))} e^{-s(L+T_{arh}/2)}. \end{aligned} \quad (2.56)$$

El error del modelo se caracteriza por $T_{arh} + T_s/2$, lo que significa que se debe requerir que $\omega^*(T_{arh} + T_s/2)$ es suficientemente pequeño. Un valor razonable es que es menor que 0,1 o 0,2, lo que quiere decir que la dinámica despreciada tiene un desfase de 6 a 12 grados.

Aproximar modos lentos por integradores

Los modos que son mucho más lentos que ω^* se pueden aproximar por integradores. Por ejemplo, si $\omega^* T$ o $\omega^* T_1$ son mayores que 5 a 10, el modelo (2.56) se puede aproximar por

$$\begin{aligned} G(s) &= \frac{K_p}{1+s(T+T_{arh}/2)} e^{-s(L+T_{arh}/2)} \approx \frac{K_p}{s(T+T_{arh}/2)} e^{-s(L+T_{arh}/2)} \\ G(s) &= \frac{K_p}{(1+sT_1)(1+s(T_2+T_{arh}/2))} e^{-s(L+T_{arh}/2)} \\ &\approx \frac{K_p}{sT_1(1+s(T_2+T_{arh}/2))} e^{-s(L+T_{arh}/2)}. \end{aligned} \quad (2.57)$$

Otra representación del modelo

Para algunas técnicas de diseño es deseable tener modelos de la forma

$$\begin{aligned} G(s) &= \frac{b}{s+a} \\ G(s) &= \frac{b_1 s + b_2}{s^2 + a_1 s + a_2}, \end{aligned} \quad (2.58)$$

que no tienen ningún retardo de tiempo. Estas formas pueden también usarse para sistemas oscilatorios. Los modelos dados por (2.55) se pueden convertir a la forma (2.58) utilizando la aproximación

$$e^{-sT} \approx \frac{1 - sT/2}{1 + sT/2}. \quad (2.59)$$

Retardos de tiempo y ceros en el semiplano derecho son las características que finalmente limitan el comportamiento alcanzable. Estas propiedades las preserva la aproximación anterior.

Ejemplos

A continuación se ilustrará la reducción de modelos con algunos ejemplos.

EJEMPLO 2.13—REDUCCIÓN DE MODELOS

Sea un sistema descrito por la función de transferencia

$$G(s) = \frac{K_p}{(1+s)(1+0,1s)(1+0,01s)(1+0,001s)}. \quad (2.60)$$

Tenemos $\omega_{90} = 3$ y $\omega_{180} = 31,6$, que da los rangos de ω^* . Consideremos en primer lugar la reducción de modelos para un diseño con $\omega^* = 3$. El factor de baja frecuencia es

$$G_l(s) = \frac{K_p}{1+s},$$

el factor de frecuencias medias es $T_s = 0,1$, y el tiempo de residencia medio de la parte de alta frecuencia es $T_{arh} = 0,011$. La regla de la mitad de Skogestad da el modelo

$$\tilde{G}(s) = \frac{K_p}{1+1,05s} e^{-0,061s}.$$

Requeriendo que $\omega^*(T_{arh} + T_s/2) < 0,2$ encontramos que se puede utilizar el modelo para diseños con $\omega^* < 3,3$.

Para $\omega^* = 31,6$ el factor de baja frecuencia se hace

$$G_l(s) = \frac{K_p}{(1+s)(1+0,1s)},$$

la constante de tiempo a frecuencias medias es $T_s = 0,01$, y el tiempo de residencia medio de la parte de alta frecuencia es entonces $T_{arh} = 0,001$. La regla de la mitad de Skogestad da el modelo

$$\tilde{G}(s) = \frac{K_p}{(1+s)(1+0,105s)} e^{-0,006s}.$$

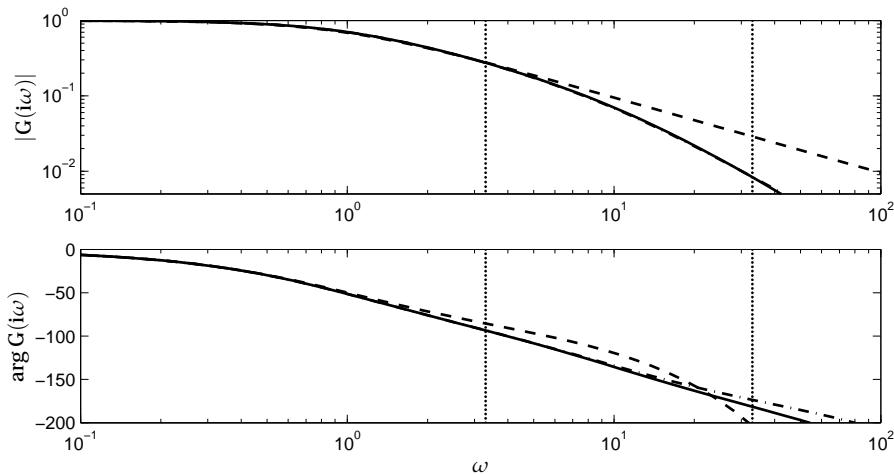


Figura 2.40 Diagramas de Bode del sistema original (línea continua) y las aproximaciones para frecuencias $\omega^* < 3,3$ (a tramos) y $\omega^* < 33$ (punto-tramo).

Requeriendo que $\omega^*(T_{arh} + T_s/2) < 0,2$ encontramos que se puede emplear el modelo para diseños con $\omega^* < 33$.

En la Figura 2.40 se ilustran las aproximaciones. □

Un aviso

El hecho de que las respuestas a un escalón de dos sistemas sean similares no implican que los sistemas sean similares también bajo control por realimentación. Esto se ilustra por el ejemplo siguiente.

EJEMPLO 2.14—LAZO ABIERTO SIMILAR – LAZO CERRADO DIFERENTE Sistemas con las funciones de transferencia

$$G_1(s) = \frac{100}{s+1}, \quad G_2(s) = \frac{100}{(s+1)(1+0,025s)^2}$$

tienen respuestas en lazo abierto muy similares tal como se ilustra en la Figura 2.41. Las diferencias entre las respuestas a un escalón apenas se notan en la figura. Los sistemas en lazo cerrado obtenidos con realimentación unitaria tienen las funciones de transferencia

$$G_{1cl} = \frac{100}{s+101}, \quad G_{2cl} = \frac{100}{(1+0,01192s)(1-0,001519s+0,0005193s^2)}.$$

Los sistemas en lazo cerrado son muy diferentes ya que G_{2cl} es inestable. □

Es posible también tener la situación opuesta, es decir, sistemas cuya conducta en lazo cerrado son muy similares aunque su conducta en lazo abierto son muy diferentes.

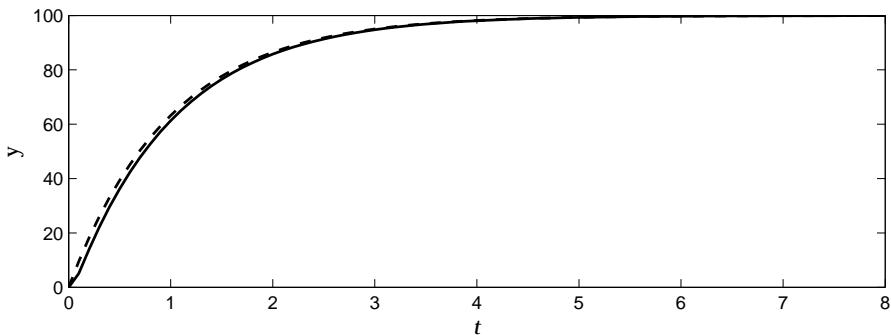


Figura 2.41 Respuestas en escalón para sistemas con funciones de transferencia $G_1(s) = 100/(s+1)$ (a tramos) y $G_2(s) = 100/((s+1)(1+0,025s)^2)$ (continua).

EJEMPLO 2.15—LAZO ABIERTO DIFERENTE – LAZO CERRADO SIMILAR

Los sistemas con las funciones de transferencia

$$P_1(s) = \frac{100}{s+1}, \quad P_2(s) = \frac{100}{s-1}$$

tienen propiedades en lazo abierto muy diferentes porque un sistema es inestable y el otro es estable. Los sistemas en lazo cerrado obtenidos con realimentación unitaria son sin embargo,

$$P_{1cl}(s) = \frac{100}{s+101} \quad P_{2cl}(s) = \frac{100}{s+99},$$

que están muy próximos. □

Las paradojas en los ejemplos se pueden resolver al considerar los rangos de frecuencia que son relevantes para el control en lazo cerrado. En el Ejemplo 2.14 el ancho de banda de relevancia del sistema en lazo cerrado es de 100 rad/s. Esto corresponde a las constantes de tiempo de 0,01 s. Un examen más detallado muestra que las respuestas en escalón en la Figura 2.41 son en realidad bastante diferentes en esa escala de tiempo incluso aunque la apariencia general de las respuestas a un escalón son muy similares. En el Ejemplo 2.15 el ancho de banda en lazo cerrado es también de 100 rad/s, que corresponde a una escala de tiempo de 0,01 s. En esa escala de tiempo los sistemas en lazo abierto son muy similares incluso aunque un modelo es estable y el otro inestable. Es una buena regla conocer los rangos de frecuencia relevantes y analizar los diagramas de Bode. Esta es una de las principales razones para emplear la respuesta en frecuencia.

2.9 Resumen

El modelado es un aspecto importante de la sintonía de un controlador. Los modelos que necesitamos deberían describir como reacciona el proceso a las señales de control. También deberían describir las propiedades de las perturbaciones que

entran al sistema. La mayor parte del trabajo sobre sintonía de controladores se ha centrado sobre la dinámica del proceso, que se refleja también en la presentación en este capítulo.

En este capítulo se han presentado una serie de métodos para determinar la dinámica de un proceso. Algunos son muy sencillos: se basan en una medida directa de la respuesta a un escalón y en construcciones gráficas simples. Otras se basan en la respuesta en frecuencia. Se ha demostrado que se puede generar información muy útil de experimentos de realimentación con un relé. Tales experimentos son particularmente útiles porque se lleva al proceso a una auto-oscilación en la frecuencia última, que es de considerable interés para el diseño de controladores.

Los métodos sencillos son útiles en el trabajo de campo cuando se tiene que sintonizar un controlador y hay pocas herramientas disponibles. Los métodos son también útiles para proporcionar una comprensión del sistema así como porque sirven de referencia cuando se evalúan métodos más complicados que requieren bastantes cálculos.

Han sido presentados modelos de diferente complejidad. Muchos modelos se han caracterizado por unos pocos parámetros. Tales modelos son útiles para muchos objetivos y se analizan en el Capítulo 6. Cuando se utilizan modelos se debería tener presente que son aproximaciones.

Al deducir los modelos también se introdujeron dos parámetros adimensionales, el retardo de tiempo normalizado τ y la razón de ganancia κ . Estos parámetros permiten hacer una evaluación cruda de la dificultad de controlar el proceso. Procesos con pequeños valores de κ y τ son fáciles de controlar. La dificultad aumenta cuando los valores se aproximan a 1. En el Capítulo 7 se proporcionan reglas de sintonía basadas en τ y κ .

Para resumir: cuando se deduce un modelo sencillo para la sintonía de un controlador PID, es importante asegurar que el modelo describe bien al proceso para las señales de entrada típicas obtenidas durante las operaciones del proceso. La amplitud y la distribución en frecuencia de la señal es de importancia. La precisión del modelo puede ser pobre si el proceso es no lineal o variable con el tiempo. La calidad del control se puede mejorar mediante planificación de ganancia o control adaptativo. Es también importante conocer qué clase de perturbaciones están actuando sobre el sistema y qué limitación imponen.

2.10 Notas y referencias

Los primeros esfuerzos en modelado utilizando ecuaciones diferenciales fueron hechos independientemente por [235] y [350] en conexión con el análisis de máquinas de vapor con reguladores centrífugos. La idea de modelar un proceso por su curva de reacción (respuesta a un escalón) emergió en los 1930. La curva de reacción se approximó por un modelo FOTD (2.18) en [72]. La curva de reacción se empleó también en [378]. Los argumentos de la respuesta en frecuencia fueron utilizados en [186] quien investigó un lazo de control de temperatura usando el modelo dado por (2.32). La respuesta en frecuencia fue usada también por [378]. Una referencia temprana a la noción de diagrama de bloque se encuentra en [232].

El modelado de procesos es un elemento clave en comprender y resolver un

problema de control. Buenas presentaciones sobre modelado se encuentran en libros de texto estándar sobre control, tales como [104, 69, 74, 321, 221, 305]. Los libros [261, 229, 48, 291, 302] son de interés particular para control de procesos. Estos libros tienen gran cantidad de material sobre muchas técnicas diferentes de modelado. Presentaciones similares se dan en [121, 157, 260, 334, 95, 314, 326, 141]. Los libros [343] y [216] están escritos por personas de formación práctica que trabajaban en compañías de control. Hay también libros que se especializan en modelado para el diseño de sistemas de control; ver [359, 252, 253, 85].

A mediados de 1950 la respuesta en frecuencia estaba muy bien establecida tal como se puso de manifiesto en un simposio organizado como parte del congreso anual de la American Society of Mechanical Engineering en 1953. Las actas del simposio fueron publicadas en el libro [258]. Una agradable revisión de los métodos de respuesta a un escalón y en frecuencia se dan en el trabajo [290]. Detalles adicionales se encuentran en [328, 8, 7]. El método del doble pulso se analiza en [313], y el método de los momentos se describe en [120].

El método del relé se introduce en [16], y se elabora en [136, 300, 147]. El método de la función descriptiva está bien documentado en [42] y [117]. Un método para estimar lo que hoy se llama un modelo ARX se desarrolló en [12] y se aplicó al modelado y control de una papelera en [31]. Hay muchos libros sobre estimación de parámetros: [214, 215, 324, 56, 188]. Muchos aspectos prácticos útiles sobre identificación de sistemas se dan en [184].

El modelado se ha mejorado mucho utilizando la simulación. La primera simulación de un sistema de control con control PID se hizo en la Universidad de Manchester utilizando una copia del analizador diferencial desarrollado por Vannevar Bush; ver [72]. El analizador diferencial se empleó también en [378] para desarrollar reglas de sintonía. Los primeros simuladores neumáticos construidos de componentes neumáticos se utilizaron por los fabricantes de equipos. El primer computador electrónico analógico desarrollado por Philbrick tuvo un gran impacto, y el uso de la simulación aumentó drásticamente. El rápido desarrollo del cálculo digital ha hecho posible que cada ingeniero tenga herramientas de simulación en su computador personal; ver [13]. Muchos de los programas actuales de simulación mimetizan los diagramas empleados para programar los primeros computadores analógicos en los 1950. Están en marcha grandes esfuerzos para combinar experiencias de modelado de procesos con avances en las ciencias de la computación para desarrollar una nueva generación de lenguajes y herramientas para el modelado de procesos; ver [106] y [339].

Hay muchos métodos para reducción de modelos. Los primeros trabajos se reportan en [379]. Una agradable revisión está en [126]. Un método que se adapta al control PID se presenta en [112]. La regla de la mitad se desarrolló en [318] como un método sencillo que funciona bien para el objetivo de sintonizar controladores PID.

3

Control PID

3.1 Introducción

El controlador PID es con mucho el algoritmo de control más común. La mayoría de los lazos de realimentación se controlan mediante este algoritmo u otro con pequeñas variaciones. Se implementa de muchas formas diferentes, como un controlador único o como parte de un paquete DDC (Control Digital Directo) o un sistema de control de procesos distribuido jerárquico. Muchos miles de ingenieros de instrumentación y control en todo el mundo están usando tales controladores en su trabajo diario. El algoritmo PID se puede contemplar desde muchas direcciones diferentes. Se puede ver como un dispositivo que puede operarse con unas pocas reglas heurísticas, pero que también puede estudiarse analíticamente.

Este capítulo da una introducción al control PID. Se presentan con detalle el algoritmo básico y varias representaciones. Se da una descripción de las propiedades del controlador en un lazo cerrado basado en argumentos intuitivos. Se estudia el fenómeno de la saturación del término integral, que ocurre cuando un controlador con acción integral se conecta a un proceso con un actuador con saturación, incluyendo algunos métodos para evitarlo. También se proporcionan filtros para reducir la influencia del ruido y medios para mejorar la respuesta a cambios en el punto de consigna.

En el Capítulo 13 se presentan aspectos de implementación del controlador PID.

3.2 El controlador PID

La versión de «libro de texto» del algoritmo PID se puede describir como:

$$u(t) = K \left(e(t) + \frac{1}{T_i} \int_0^t e(\tau) d\tau + T_d \frac{de(t)}{dt} \right), \quad (3.1)$$

donde u es la señal de control y e es el error de control ($e = y_{sp} - y$). La señal de control es así una suma de tres términos: el término-P (que es proporcional al

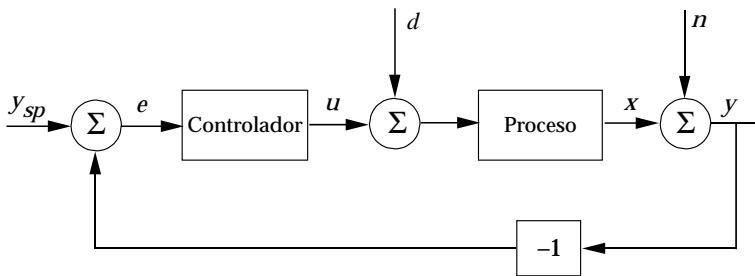


Figura 3.1 Diagrama de bloques de un lazo de realimentación simple.

error), el término-I (que es proporcional a la integral del error), y el término-D (que es proporcional a la derivada del error). Los parámetros del controlador son la ganancia proporcional K , el tiempo integral T_i , y el tiempo derivativo T_d .

Acción proporcional

En el caso del control proporcional puro, la ley de control dada por la Ecuación 3.1 se reduce a

$$u(t) = K e(t) + u_b. \quad (3.2)$$

la acción de control es simplemente proporcional al error de control. La variable u_b es una polarización o un reset. Cuando el error de control e es cero, la señal de control toma el valor $u(t) = u_b$. La polarización u_b a menudo se fija a $(u_{\max} + u_{\min})/2$, pero puede algunas veces ajustarse manualmente de forma que el error de control estacionario es cero en un punto de consigna dado.

Análisis estático Algunas propiedades del control proporcional se pueden comprender por el siguiente argumento, que se basa en consideraciones puramente estáticas. Sea el lazo de realimentación simple, que se muestra en la Figura 3.1, y que se compone de un proceso y un controlador. Suponga que el controlador tiene acción proporcional y que el proceso se modela por el modelo estático

$$x = K_p(u + d), \quad (3.3)$$

donde x es la variable de proceso, u es la variable de control, d es una perturbación de carga, y K_p es la ganancia estática del proceso. Del diagrama de bloques se obtienen las ecuaciones siguientes.

$$\begin{aligned} y &= x + n \\ x &= K_p(u + d) \\ u &= K(y_{sp} - y) + u_b. \end{aligned} \quad (3.4)$$

donde n es el ruido de medida. la eliminación de las variables intermedias da la siguiente relación entre la variable de proceso x , el punto de consigna y_{sp} , la perturbación de carga d , y el ruido de medida n :

$$x = \frac{KK_p}{1 + KK_p} (y_{sp} - n) + \frac{K_p}{1 + KK_p} (d + u_b). \quad (3.5)$$

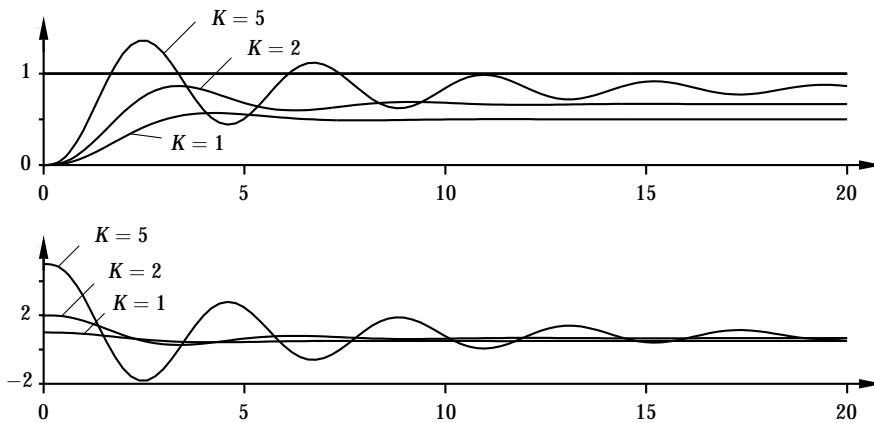


Figura 3.2 Simulación de un sistema en lazo cerrado con control proporcional. La función de transferencia del proceso es $G(s) = (s + 1)^{-3}$. El diagrama superior muestra el punto de consigna $y_{sp} = 1$ y la salida del proceso y para diferentes valores de la ganancia del controlador K . El diagrama inferior muestra la señal de control u para diferentes ganancias del controlador.

El producto KK_p es un número adimensional llamado la *ganancia del lazo*. Algunas propiedades interesantes del sistema en lazo cerrado se pueden deducir de la Ecuación 3.5. En primer lugar supóngase que n y u_b son cero. Entonces la ganancia de lazo debería ser alta con el fin de asegurar que la salida del proceso x está próxima al punto de consigna y_{sp} . Un valor elevado de la ganancia del controlador K también hace al sistema insensible a la perturbación de carga d . Sin embargo, si n es distinto de cero, se sigue de la Ecuación 3.5 que el ruido de medida n influye sobre la salida del proceso de la misma forma que el punto de consigna y_{sp} . Para evitar hacer al sistema sensible al ruido de medida, la ganancia del lazo no debería hacerse demasiado grande. Más aún, la polarización del controlador u_b influye sobre el sistema de la misma manera que una perturbación de carga. Es por lo tanto obvio que el diseño de la ganancia del lazo es un compromiso entre diferentes objetivos de control y que no hay una respuesta simple a cuál es la mejor ganancia de lazo. Esto dependerá del objetivo de control que es más importante.

También se sigue de la Ecuación 3.5 que habrá normalmente un error en estado estacionario con control proporcional. Esto se puede deducir intuitivamente de la siguiente observación de la Ecuación 3.4 que el error de control es cero solamente cuando $u = u_b$ en estacionario. El error, por lo tanto, se puede hacer cero en una condición operativa dada por una elección adecuada de la polarización del controlador u_b .

El análisis estático que se acaba de hacer se basa en la hipótesis de que el proceso se puede describir por un modelo estático. Esto omite algunas propiedades importantes de la dinámica del sistema en lazo cerrado. El más importante es que el sistema en lazo cerrado normalmente será inestable para altas ganancias del lazo si se considera la dinámica del proceso. En la práctica, la máxima ganancia del lazo se determina así por la dinámica del proceso. Una forma de describir

la dinámica del proceso conduce a descripciones como la Ecuación 3.3 donde la ganancia del proceso es dependiente de la frecuencia. Esto se analizó ya en el Capítulo 2.

En la Figura 3.2 se ilustra un ejemplo típico de control proporcional. La figura muestra la conducta de la salida del proceso y la señal de control después de un cambio en escalón en el punto de consigna. El error en estado estacionario se puede calcular de la Ecuación 3.5. El término de polarización u_b , la carga d , y el ruido n son todos cero en la simulación. Con una ganancia del controlador $K = 1$ y una ganancia estática del proceso $K_p = 1$, el error es por lo tanto del 50 por ciento. La figura muestra que el error en estado estacionario disminuye cuando la ganancia del controlador aumenta tal como lo predice la Ecuación 3.5. Obsérvese también que la respuesta se hace más oscilatoria cuando aumenta la ganancia del controlador. Esto se debe a la dinámica del proceso.

Acción integral

La función principal de la acción integral es asegurarse de que la salida del proceso coincide con el punto de consigna en estado estacionario. Con control proporcional, hay normalmente un error de control en estado estacionario. Con acción integral, un pequeño error positivo conducirá siempre a una señal de control creciente, y un error negativo dará una señal de control decreciente sin tener en cuenta lo pequeño que sea el error.

El siguiente argumento simple muestra que el error en estado estacionario siempre será cero con acción integral. Suponga que el sistema está en estado estacionario con una señal de control constante (u_0) y un error constante (e_0). Se sigue de la Ecuación 3.1 que la señal de control viene entonces dada por

$$u_0 = K \left(e_0 + \frac{e_0}{T_i} t \right).$$

Mientras $e_0 \neq 0$, esto claramente contradice la hipótesis de que la señal de control u_0 es constante. Un controlador con acción integral siempre dará error cero en estado estacionario.

La acción integral se puede también visualizar como un dispositivo que automáticamente resetea el término de polarización u_b de un controlador proporcional. Esto se ilustra en el diagrama de bloques de la Figura 3.3, que muestra un controlador proporcional con un reset que se ajusta automáticamente. El ajuste se hace realimentando una señal, que es un valor filtrado de la salida, al punto de suma del controlador. Esta fue realmente una de las primeras invenciones de la acción integral, o «reset automático», tal como se llamó. La implementación que se muestra en la Figura 3.3 se utiliza todavía por muchos fabricantes.

Cálculos sencillos muestran que el controlador de la Figura 3.3 da los resultados deseados. Del diagrama de bloques se obtiene la siguiente ecuación:

$$u = Ke + I = T_i \frac{dI}{dt} + I.$$

De donde,

$$T_i \frac{dI}{dt} = Ke,$$

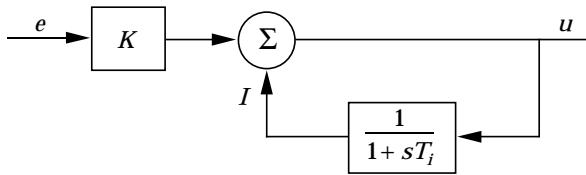


Figura 3.3 Implementación de la acción integral como realimentación positiva alrededor de un sistema de primer orden.

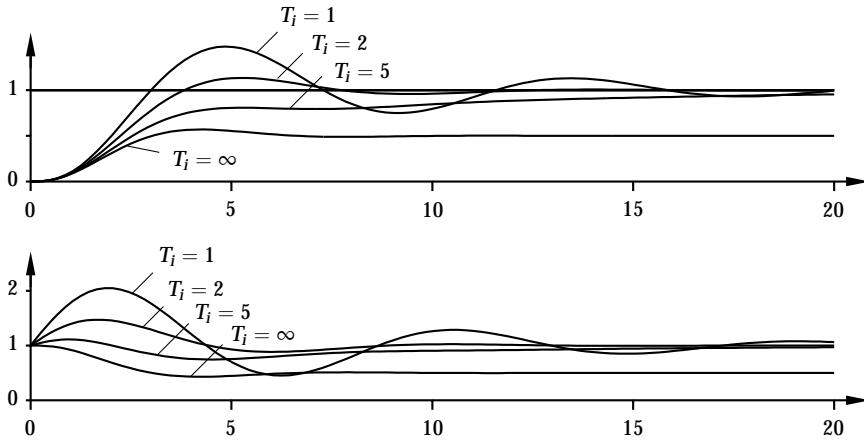


Figura 3.4 Simulación de un sistema en lazo cerrado con control proporcional e integral. La función de transferencia del proceso es $G(s) = (s + 1)^{-3}$, y la ganancia del controlador es $K = 1$. El diagrama superior muestra el punto de consigna $y_{sp} = 1$ y la salida del proceso y para diferentes valores del tiempo integral T_i . El diagrama inferior muestra la señal de control u para diferentes tiempos integrales.

y encontramos que

$$u = K \left(e + \frac{1}{T_i} \int e(\tau) d\tau \right),$$

que es un controlador PI.

Las propiedades de la acción integral se ilustran en la Figura 3.4, que muestra una simulación de un sistema con control PI. La ganancia proporcional es constante, $K = 1$, y se cambia el tiempo integral. El caso $T_i = \infty$ corresponde a control proporcional puro. Este caso es idéntico al caso $K = 1$ de la Figura 3.2, donde el error en estado estacionario es del 50 por ciento. El error en estado estacionario se elimina cuando T_i tiene valores finitos. Para grandes valores del tiempo de integración, la repuesta se mueve lentamente hacia el punto de consigna. La aproximación es aproximadamente exponencial con constante de tiempo T_i / KK_p . Dicha aproximación es más rápida para valores más pequeños de T_i pero es también oscilatoria.

Acción derivativa

El objetivo de la acción derivativa es mejorar la estabilidad en lazo cerrado. El mecanismo de inestabilidad se puede describir intuitivamente como sigue. A

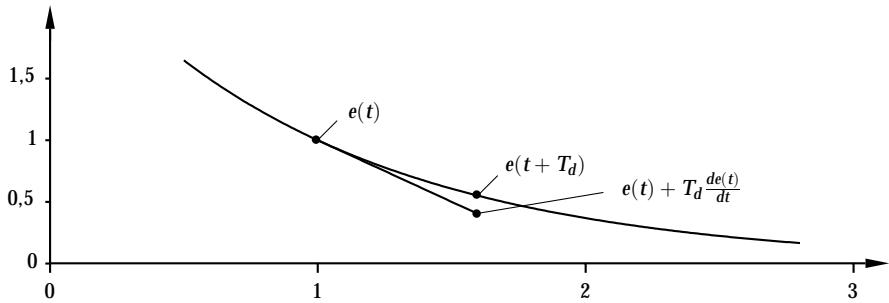


Figura 3.5 Interpretación de la acción derivativa como control predictivo, donde la predicción se obtiene por extrapolación lineal.

causa de la dinámica del proceso, llevará algún tiempo antes de que un cambio en la variable de control sea detectable en la salida del proceso. Así, el sistema de control tardará en corregir un error. La acción de un controlador con acción proporcional y derivativa se puede interpretar como si el control se hiciese proporcional a la salida *predicha* del proceso, donde la predicción se hace extrapolando el error por la tangente a la curva del error (ver Figura 3.5). La estructura básica de un controlador PD es

$$u(t) = K \left(e(t) + T_d \frac{de(t)}{dt} \right).$$

Un desarrollo en serie de Taylor de $e(t + T_d)$ da

$$e(t + T_d) \approx e(t) + T_d \frac{de(t)}{dt}.$$

La señal de control es así proporcional a una estimación del error de control en el tiempo T_d hacia adelante, donde se obtiene la estimación por extrapolación lineal. En la Figura 3.6 se ilustran las propiedades de la acción derivativa, que muestra una simulación de un sistema con control PID.

La ganancia del controlador y el tiempo de integración se mantiene constante, $K = 3$ y $T_i = 2$, y se cambia el tiempo derivativo T_d . Para $T_d = 0$ tenemos control PI puro. El sistema en lazo cerrado es oscilatorio con los parámetros escogidos. Inicialmente el amortiguamiento aumenta cuando el tiempo derivativo aumenta pero disminuye otra vez cuando el tiempo derivativo se hace demasiado grande.

Implementación clásica de la acción derivativa

En la Figura 3.3 se mostró que la acción integral originalmente era implementada mediante realimentación positiva alrededor de un sistema de primer orden. La acción derivativa fue también originalmente implementada utilizando un sistema de primer orden tal como se muestra en el diagrama de bloques de la Figura 3.7. La transformada de Laplace de la salida está dada por

$$U(s) = \left(1 - \frac{1}{1+sT} \right) E(s) = \frac{sT}{1+sT} E(s). \quad (3.6)$$

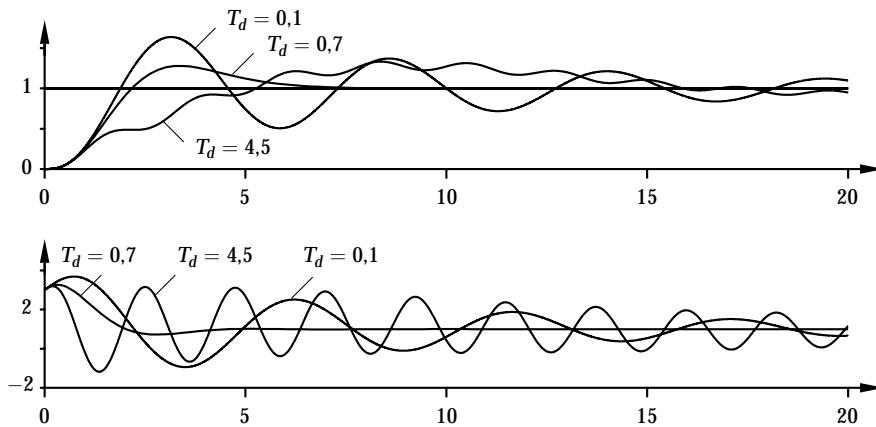


Figura 3.6 Simulación de un sistema en lazo cerrado con control proporcional, integral y derivativo. La función de transferencia del proceso es $G(s) = (s + 1)^{-3}$, la ganancia del controlador es $K = 3$, y el tiempo integral es $T_i = 2$. El diagrama superior muestra el punto de consigna $y_{sp} = 1$ y la salida del proceso y para diferentes valores del tiempo derivativo T_d . El diagrama inferior muestra la señal de control u para diferentes tiempos derivativos.

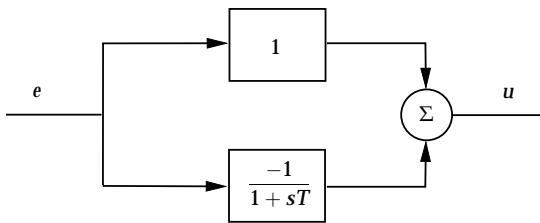


Figura 3.7 Implementación clásica de la acción derivativa.

El sistema así tiene la función de transferencia $G(s) = sT / (1 + sT)$. Obsérvese que el filtraje se obtiene automáticamente con esta implementación.

Representaciones alternativas

El algoritmo PID dado por la Ecuación 3.1 se puede representar por la función de transferencia

$$C(s) = K \left(1 + \frac{1}{sT_i} + sT_d \right). \quad (3.7)$$

una versión ligeramente diferente es más común en controladores comerciales. Este controlador se describe por

$$C'(s) = K' \left(1 + \frac{1}{sT'_i} \right) (1 + sT'_d). \quad (3.8)$$

En la Figura 3.8 se muestran las dos estructuras del controlador en forma de diagrama de bloques. El controlador dado por la Ecuación 3.7 se llama no interactuante, y el dado por la Ecuación 3.8, interactuante. La razón para esta nomenclatura es que en el controlador (3.7) el tiempo integral T_i no influye sobre la parte

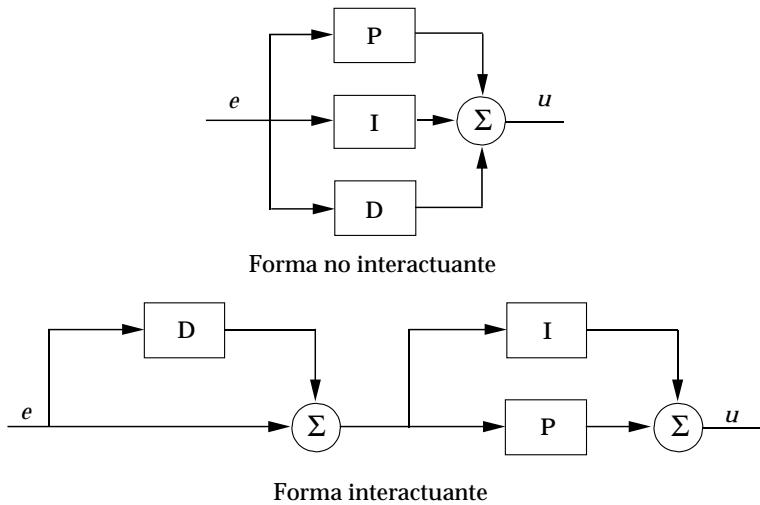


Figura 3.8 Forma interactuante y no interactuante del algoritmo PID.

derivativa, y el tiempo derivativo T_d no influye sobre la parte integral. Las partes son así no interactuantes. En el controlador interactuante, el tiempo derivativo T'_d influye sobre la parte integral. Por lo tanto, las partes son interactuantes.

El controlador interactuante (3.8) puede siempre representarse como un controlador no interactuante cuyos coeficientes están dados por

$$\begin{aligned} K &= K' \frac{T'_i + T'_d}{T'_i} \\ T_i &= T'_i + T'_d \\ T_d &= \frac{T'_i T'_d}{T'_i + T'_d}. \end{aligned} \tag{3.9}$$

Un controlador interactuante de la forma (3.8) que corresponde a un controlador no interactuante se puede encontrar solamente si

$$T_i \geq 4T_d.$$

Entonces,

$$\begin{aligned} K' &= \frac{K}{2} \left(1 + \sqrt{1 - 4T_d/T_i} \right) \\ T'_i &= \frac{T_i}{2} \left(1 + \sqrt{1 - 4T_d/T_i} \right) \\ T'_d &= \frac{T_i}{2} \left(1 - \sqrt{1 - 4T_d/T_i} \right). \end{aligned} \tag{3.10}$$

El controlador no-interactuante dado por la Ecuación 3.7 es más general, y lo utilizaremos en el futuro. Sin embargo, algunas veces se argumenta que el controlador interactuante es más fácil de sintonizar manualmente.

Hay también una razón histórica para preferir al controlador interactuante. Los primeros controladores neumáticos eran más fáciles de construir utilizando la forma interactuante. Ver Capítulo 13. Cuando los fabricantes de controladores cambiaron la tecnología de neumática a eléctrica analógica y, finalmente, a la técnica digital, mantuvieron la forma interactuativa. Por lo tanto, la forma interactuante es más común entre los controladores mono lazo.

Es importante tener presente que diferentes controladores pueden tener diferentes estructuras. Lo que significa que si un controlador en un cierto lazo de control se sustituye por otro tipo de controlador, pueden tener que cambiarse sus parámetros. Obsérvese, sin embargo, que las formas interactuante y no interactuantes son diferentes sólo cuando se usan ambas partes I y D del controlador. Si sólo empleamos el controlador como un controlador P, PI, o PD, las dos formas son equivalentes. Aún otra representación del algoritmo PID esta dada por

$$C''(s) = k + \frac{k_i}{s} + sk_d. \quad (3.11)$$

Los parámetros se relacionan con los parámetros de la forma estándar a través de

$$k = K \quad k_i = \frac{K}{T_i} \quad k_d = KT_d. \quad (3.12)$$

La representación (3.11) es equivalente a la forma estandar, pero los valores de los parámetros son bastante diferentes. Esto puede causar grandes dificultades a quien no es conocedor de esta diferencia, particularmente si al parámetro $1/k_i$ se le llama tiempo integral y al k_d tiempo derivativo. La forma dada por la Ecuación 3.11 es a menudo útil en cálculos analíticos porque los parámetros aparecen linealmente. La representación también tiene la ventaja de que es posible obtener acción proporcional, integral, o derivativa pura mediante valores finitos de los parámetros.

Resumiendo, hemos encontrado así que hay tres formas diferentes del controlador PID.

- La forma estándar o no interactuante dada por la Ecuación 3.7.
- La forma serie o interactuante dada por la Ecuación 3.8.
- La forma paralelo dada por la Ecuación 3.11.

La forma estandar se llama algunas veces el algoritmo ISA, o algoritmo ideal. Las acciones proporcional, integral, y derivativa son no interactuantes en el dominio del tiempo. Este algoritmo admite ceros complejos, lo cual es útil cuando se controlan sistemas con modos oscilatorios.

La forma serie se llama también la forma clásica. Esta representación se obtiene naturalmente cuando se implementa un controlador como un dispositivo analógico basado en un sistema de equilibrio de fuerza neumática. El nombre clásico refleja esto. La forma serie tiene una interpretación atractiva en el dominio de la frecuencia porque los ceros corresponden a los valores inversos de los tiempos derivativo e integral. Todos los ceros del controlador son reales. No

se pueden obtener acciones integral o proporcional puras con valores finitos de los parámetros del controlador.

La forma paralelo es la forma más general porque se puede obtener acción proporcional o integral pura con parámetros finitos. El controlador puede también tener ceros complejos. De esta manera es la forma más flexible. Sin embargo, es también la forma donde los parámetros tienen poca interpretación física.

Resumen

El controlador PID tiene tres términos. El término proporcional P corresponde a control proporcional. El término integral I da una acción de control que es proporcional a la integral del error. Esto asegura que el error en estado estacionario se haga cero. El término derivativo D es proporcional a la derivada del error de control. Este término permite la predicción del error futuro. Hay muchas variaciones del algoritmo PID básico que sustancialmente mejorarán su comportamiento y capacidad de operación. Estos cambios se analizan en las secciones siguientes.

3.3 Filtraje de la acción derivativa

Una desventaja con la acción derivativa es que una derivada ideal tiene una ganancia muy elevada para señales de alta frecuencia. Esto significa que el ruido de medida de alta frecuencia generará grandes variaciones de la señal de control. Para ver esto consideramos una salida medida

$$y = \text{sent} + \text{asen}\omega t,$$

donde el primer término es la señal útil y el segundo término representa el ruido. La razón ruido – señal es así a . El término derivativo del controlador es entonces

$$D = KT_d \frac{dy}{dt} = KT_d(\cos t + a\omega \cos \omega t). \quad (3.13)$$

La amplitud de la señal es KT_d , y la amplitud del ruido es $KT_d a \omega$. La razón ruido – señal es $a\omega$. Esto puede ser arbitrariamente grande incluso si a es pequeña si la frecuencia es suficientemente alta. El efecto del ruido de medida puede de alguna forma reducirse implementando el término derivativo como

$$D = -\frac{sKT_d}{1 + sT_d/N} Y. \quad (3.14)$$

Esto se puede interpretar como una acción derivativa ideal que se filtra usando un sistema de primer orden con la constante de tiempo T_d/N . Para pequeños s la función de transferencia es aproximadamente sKT_d , y para grandes s es igual a KN . La aproximación actúa como una derivada para componentes de señales de baja frecuencia, y la ganancia de alta frecuencia está limitada a KN . El ruido de medida de alta frecuencia se amplifica a lo más por un factor de KN . Valores típicos de N son de 2 a 20. Obsérvese que la implementación de la derivada dada en la Figura 3.7 automáticamente da una limitación de la ganancia de alta frecuencia; ver Ecuación 3.6.

La función de transferencia de un controlador PID con la acción derivativa filtrada es

$$C(s) = K \left(1 + \frac{1}{sT_i} + \frac{sT_d}{1 + sT_d/N} \right). \quad (3.15)$$

La ganancia de alta frecuencia del controlador es $K(1 + N)$. El ruido de medida de alta frecuencia puede así generar variaciones significativas en la señal de control. Es por lo tanto ventajoso utilizar una acción de filtrado más grande.

En lugar de filtrar sólo la derivada es posible filtrar la señal medida y aplicar la señal filtrada a un controlador PID ideal. La función de transferencia del controlador equivalente es

$$C_{eq} = C(s)G_f(s) = K \left(1 + \frac{1}{sT_i} + sT_d \right) \frac{1}{1 + sT_f + (sT_f)^2/2}, \quad (3.16)$$

cuando se usa un filtro de segundo orden con amortiguamiento relativo $\zeta = 1/\sqrt{2}$. La constante de tiempo del filtro T_f se escoge típicamente como T_i/N para control PI o como T_d/N para control PID, donde N va desde 2 a 20.

Se sigue de (3.16) que la ganancia del controlador va a cero para las altas frecuencias. Esta propiedad, que se llama *caída de alta frecuencia*, garantiza que el ruido de medida de alta frecuencia no generará grandes señales de control. La caída de alta frecuencia también aumenta la robustez del sistema en lazo cerrado.

3.4 Ponderación del punto de consigna

El sistema de control de la Figura 3.1 se llama un sistema con *realimentación del error* porque el controlador actúa sobre el error, que es la diferencia entre el punto de consigna y la salida. Una estructura más flexible se obtiene al tratar separadamente el punto de consigna y la salida del proceso. Un controlador PID de esta forma está dado por

$$u(t) = K \left(e_p + \frac{1}{T_i} \int_0^t e(s) ds + T_d \frac{de_d}{dt} \right), \quad (3.17)$$

donde el error en la parte proporcional es

$$e_p = by_{sp} - y, \quad (3.18)$$

y el error en la parte derivativa es

$$e_d = cy_{sp} - y. \quad (3.19)$$

El error en la parte integral debe ser el error de control verdadero

$$e = y_{sp} - y,$$

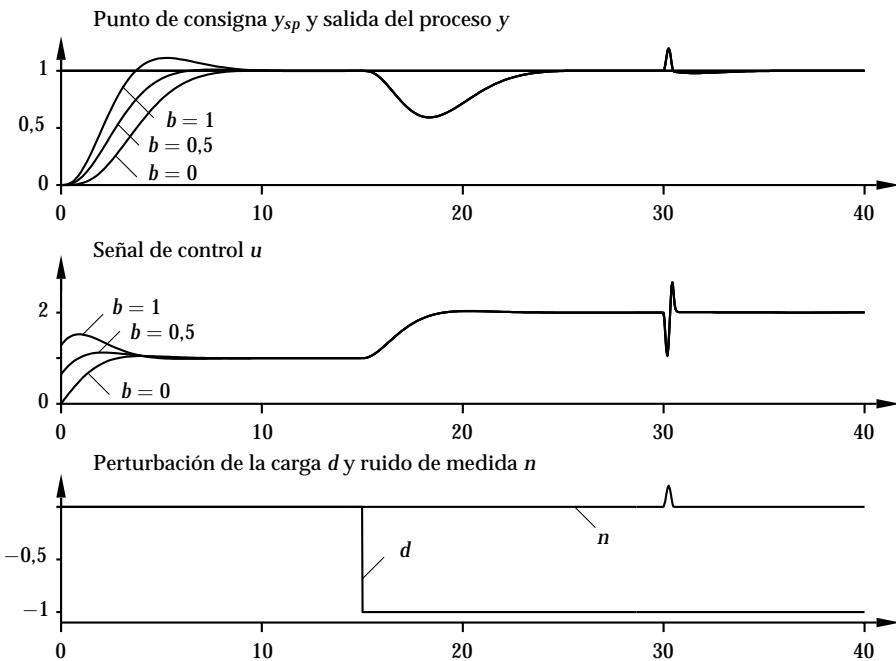


Figura 3.9 La respuesta a cambios en el punto de consigna, perturbaciones de carga, y errores de medida para diferentes valores de la ponderación del punto de consigna b .

con el fin de evitar errores de control en estado estacionario. Los controladores obtenidos para diferentes valores de b y c responderán a las perturbaciones de carga y al ruido de medida de la misma forma. La respuesta a cambios en el punto de consigna dependerá de los valores de b y c , que se llaman *pesos del punto de consigna*.

En la Figura 3.9 se ilustran las propiedades de un sistema donde el controlador tiene ponderación del punto de consigna, que muestra la respuesta de un controlador PID a cambios en el punto de consigna, perturbaciones de carga, y errores de medida para diferentes valores de b . La figura muestra claramente el efecto de cambiar b . La sobreelongación para cambios en el punto de consigna es más pequeña para $b = 0$, que es el caso donde la referencia solo se introduce en el término integral, y aumenta cuando b aumenta. Obsérvese que una simulación como la de la Figura 3.9 es útil con el fin de tener una evaluación rápida de las respuestas de un sistema en lazo cerrado frente a cambios en el punto de consigna, perturbaciones de carga, y errores de medida.

El parámetro c se escoge normalmente igual a cero para evitar grandes transitorios en la señal de control debido a cambios rápidos en el punto de consigna. Una excepción es cuando el controlador es el controlador secundario en un acoplamiento en cascada (ver Sección 12.4). En este caso, en el punto de consigna cambia suavemente ya que está dado por la salida del controlador primario. Obsérvese que si se implementa la acción integral con realimentación positiva alrededor de una constante de tiempo tal como se muestra en la Figura 3.3, el

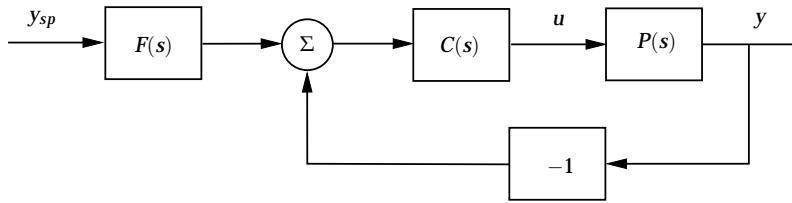


Figura 3.10 Diagrama de bloques de un lazo de realimentación simple con un controlador PID que tiene una estructura de dos grados de libertad. La función de transferencia $P(s)$ es la función de transferencia del proceso.

parámetro b es igual a uno.

El controlador con $b = 0$ y $c = 0$ se llama algunas veces un controlador I-PD, y el controlador $b = 1$ y $c = 0$ es algunas veces denominado un controlador PI-D. Preferimos adherirnos al uso genérico de PID y dar los parámetros b y c , haciendo de este modo una pequeña contribución hacia la reducción de la abreviatura de las tres letras.

En el diagrama de bloques de la Figura 3.1, la salida del controlador se genera del error $e = y_{sp} - y$. Obsérvese que este diagrama no es ya válido cuando se usan la ley de control dada por la Ecuación 3.17 y las definiciones del error (3.18) y (3.19). En este caso un diagrama de bloques para un sistema con control PID viene dado por la Figura 3.10 donde la función de transferencia $C(s)$ es (3.7) y

$$F(s) = \frac{b + \frac{1}{sT_i} + scT_d}{1 + \frac{1}{sT_i} + sT_d} = \frac{cT_i T_d s^2 + bsT_i + 1}{T_i T_d s^2 + sT_i + 1}. \quad (3.20)$$

Sistema con dos grados de libertad

En general, un sistema de control tiene muchos requisitos diferentes. Debería tener buena respuesta transitoria a cambios en el punto de consigna, y debería rechazar las perturbaciones de carga y el ruido de medida. Para un sistema con sólo realimentación del error, se hace un intento de satisfacer todas las demandas con el mismo mecanismo. Tales sistemas se llaman sistemas con un grado de libertad.

El sistema que se muestra en la Figura 3.10 se dice que tiene dos grados de libertad porque el camino de la señal desde el punto de consigna hasta la señal de control es diferente del camino de la señal desde el valor medido hasta la de control.

Hay muchas posibles configuraciones de sistemas con dos grados de libertad. El sistema de la Figura 3.10 es sólo una alternativa. Un uso extendido de estructuras con dos grados de libertad es una extensión muy natural del controlador PID. La idea clave es permitir al controlador C ser un controlador PI o un controlador PID pero usar una acción feedforward de forma más flexible que lo que permite el controlador PID estandar. Esto se analizará completamente en el Capítulo 5.

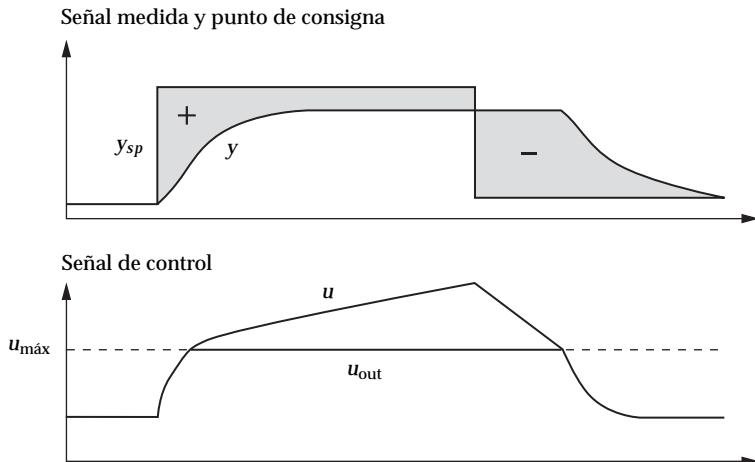


Figura 3.11 Ilustración de la saturación del integrador.

3.5 Integrador windup

Aunque se pueden comprender muchos aspectos de un sistema de control basándose en la teoría lineal, deben tomarse en consideración algunos efectos no lineales. Todos los actuadores tienen limitaciones: un motor tiene una velocidad limitada, una válvula no puede estar más que totalmente abierta o totalmente cerrada, etc. Para un sistema de control con un amplio rango de condiciones operativas, puede suceder que la variable de control alcance los límites del actuador. Cuando esto sucede el lazo de realimentación se rompe y el sistema opera como un sistema en lazo abierto porque el actuador permanecerá en su límite independientemente de la salida del proceso. Si se utiliza un controlador con acción integral, el error puede continuar siendo integrado si el algoritmo no se diseña adecuadamente. Esto significa que el término integral puede hacerse muy grande o, coloquialmente realiza un «windup». Se requiere entonces que el error tenga signo opuesto durante un largo período antes de que las cosas retornen a la normalidad. La consecuencia es que cualquier controlador con acción integral puede sufrir grandes transitorios cuando el actuador se satura.

EJEMPLO 3.1—ILUSTRACIÓN DE LA SATURACIÓN DEL INTEGRADOR

La Figura 3.11 muestra la señal de control, la señal medida, y el punto de consigna en un caso donde la señal de control se satura. Despues del primer cambio en el punto de consigna, la señal de control aumenta a su límite superior u_{\max} . Esta señal de control no es bastante grande para eliminar el error de control. Por lo tanto, la integral del error de control, y la parte integral de la señal de control, aumenta. Como la señal de control deseada u aumenta, hay una diferencia entre esta señal de control y la señal de control verdadera u_{out} .

La Figura 3.11 muestra lo que acontece cuando despues de un cierto tiempo se reduce a un nivel donde el controlador es capaz de eliminar el error de control. Como el signo del error de control se hace negativo, la señal de control comienza a disminuir, pero como la señal de control deseada u está por encima del límite

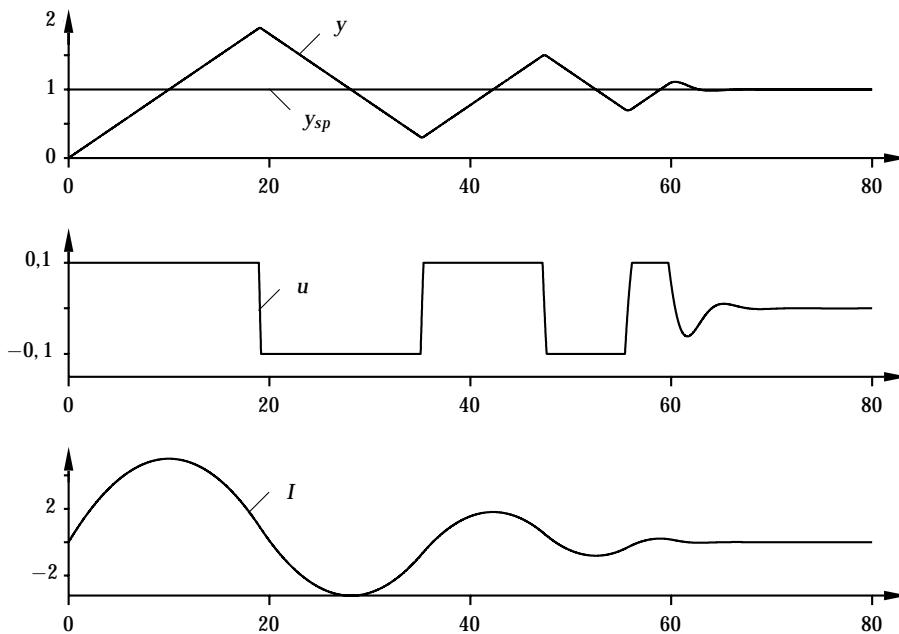


Figura 3.12 Ilustración de la saturación del integrador. El diagrama muestra la salida del proceso y , el punto de consigna y_{sp} , la señal de control u , y la parte integral I .

u_{\max} , la señal de control verdadera u_{out} se mantiene en el límite durante un rato y la respuesta se retarda. \square

El ejemplo siguiente muestra algunos otros efectos que pueden ocurrir debido a la saturación del integrador (windup) cuando el proceso es inestable.

EJEMPLO 3.2—ILUSTRACIÓN DE LA SATURACIÓN DEL INTEGRADOR

El fenómeno del windup se ilustra en la Figura 3.12, que muestra el control de un proceso con integración con un controlador PI. El cambio inicial del punto de consigna es tan grande que el actuador satura en el límite superior. El término integral aumenta inicialmente porque el error es positivo; alcanza su valor más grande en el tiempo $t = 10$ cuando el error pasa por cero. La salida permanece saturada en este punto debido al gran valor del término integral. No deja el límite de saturación hasta que el error se hace negativo durante un tiempo suficientemente grande como para permitir a la parte integral bajar a un nivel inferior. Obsérvese que la señal de control rebota entre sus límites algunas veces. El efecto neto es una gran sobreelongación y una oscilación amortiguada donde la señal de control comuta de un extremo al otro. La salida finalmente viene tan próxima al punto de consigna que el actuador no se satura. El sistema entonces se comporta linealmente y se asienta. \square

La saturación del integrador (windup) puede ocurrir en conexión con grandes cambios en el punto de consigna, o puede ser causado por grandes perturbaciones o mal funcionamiento del equipo. El windup puede ocurrir también cuan-

do se usan selectores de forma que algunos controladores están moviendo un actuador. En el control en cascada, el windup puede ocurrir en el controlador primario cuando el controlador secundario se conmuta a modo manual, usa su punto de consigna local, o si su señal de control se satura. Ver Sección 12.4.

El fenómeno del windup era bien conocido a los fabricantes de controladores analógicos, quienes inventaron algunos trucos para evitarlo. Fueron descritos bajo nombres como precarga, unidad batch, etc. Aunque el problema era bien comprendido, había a menudo límites impuestos a causa de las implementaciones analógicas. Las ideas fueron con frecuencia guardadas como secretos comerciales y no se hablaba mucho de ello. El problema del windup fue redescubierto cuando los controladores se implementaron digitalmente y se presentaron trabajos en revistas proponiendo algunos métodos para evitar el windup. En la sección siguiente describimos algunas de las ideas.

Limitación del punto de consigna

Una manera de intentar evitar el windup del integrador es introducir limitadores sobre las variaciones del punto de consigna de forma que la salida del controlador nunca alcanzará los límites del actuador. Esto con frecuencia conduce a cotas conservadoras y limitaciones sobre el comportamiento del controlador. Más aún, no evita el windup causado por las perturbaciones.

Algoritmos incrementales

En las fases iniciales del control por realimentación, la acción integral era embedida en el actuador al tener este un motor que movía directamente la válvula. En este caso, el windup se maneja automáticamente porque la integración se para cuando la válvula no se mueve. Cuando los controladores fueron implementados por técnicas analógicas, y posteriormente con computadores, muchos fabricantes usaron una configuración que era una analogía del antiguo diseño mecánico. Esto condujo a los algoritmos de velocidad, analizados en el Capítulo 13. En este algoritmo se calcula primero la velocidad de cambio de la señal de control y entonces se alimenta a un integrador. En algunos casos este integrador es un motor directamente conectado al actuador. En otros casos el integrador se implementa internamente en el controlador. Con este enfoque es fácil gestionar cambios de modo y el problema del windup. El windup se evita inhibiendo la integración siempre que se satura la salida. Este método es equivalente al recalculado, que se describe posteriormente. Si la salida del actuador no se mide, se puede emplear un modelo que calcula la salida saturada. Es también fácil limitar la velocidad de cambio de la señal de control.

Recálculo y seguimiento

El recálculo trabaja como sigue. Cuando la salida satura, se recalcula el término integral en el controlador de forma que su nuevo valor da una salida en el límite de la saturation. Es ventajoso no resetear el integrador instantáneamente sino dinámicamente con una constante de tiempo T_t .

La Figura 3.13 muestra un diagrama de bloques de un controlador PID con protección anti-windup basado en el recálculo. El sistema tiene un camino de realimentación extra que se genera midiendo la salida real del actuador, o la salida

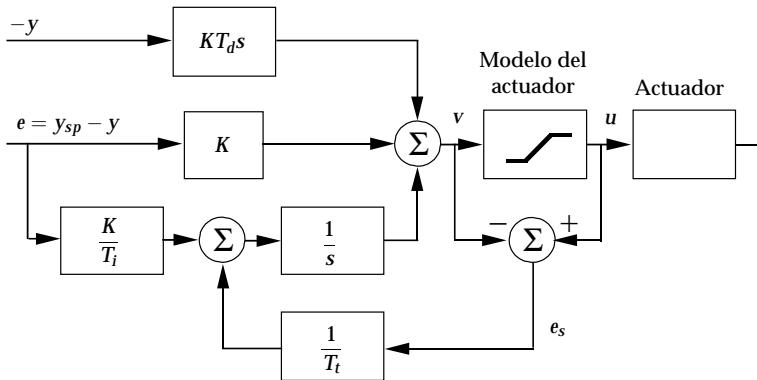


Figura 3.13 Controlador PID con mecanismo de protección anti-windup.

de un modelo matemático del actuador con saturación, y formando una señal de error (e_s) como la diferencia entre la salida del controlador (v) y la salida del actuador (u). La señal e_s se alimenta a la entrada del integrador a través de la ganancia $1/T_t$. La señal es cero cuando no hay saturación. Así pues, no tendrá ningún efecto sobre la operación normal cuando el actuador no se satura. Cuando el actuador se satura, la señal e_s es diferente de cero. Se rompe el camino de realimentación normal alrededor del proceso porque la entrada al mismo permanece constante. Hay, sin embargo, un camino de realimentación alrededor del integrador. A causa de esto, la salida del integrador se mueve hacia un valor tal que la entrada del integrador es

$$\frac{1}{T_t} e_s + \frac{K}{T_i} e,$$

donde e es el error de control. De aquí se sigue,

$$e_s = -\frac{KT_t}{T_i} e$$

en estado estacionario. Como $e_s = u - v$, se deduce que

$$v = u_{\text{lim}} + \frac{KT_t}{T_i} e,$$

donde u_{lim} es el valor de saturación de la variable de control. Como las señales e y u_{lim} tienen el mismo signo, se sigue que v es siempre mayor que u_{lim} en magnitud. Esto previene al integrador del windup. La velocidad con la cual se resetea la salida del controlador está gobernada por la ganancia de realimentación, $1/T_t$, donde T_t se puede interpretar como la constante de tiempo, lo que determina con cuánta rapidez se resetea la integral. Llamamos a esto la constante de tiempo de seguimiento.

La Figura 3.14 muestra qué sucede cuando se aplica un controlador con anti-windup al sistema simulado en la Figura 3.12. Obsérvese que la salida del integrador se resetea rápidamente a un valor tal que la salida del controlador está en

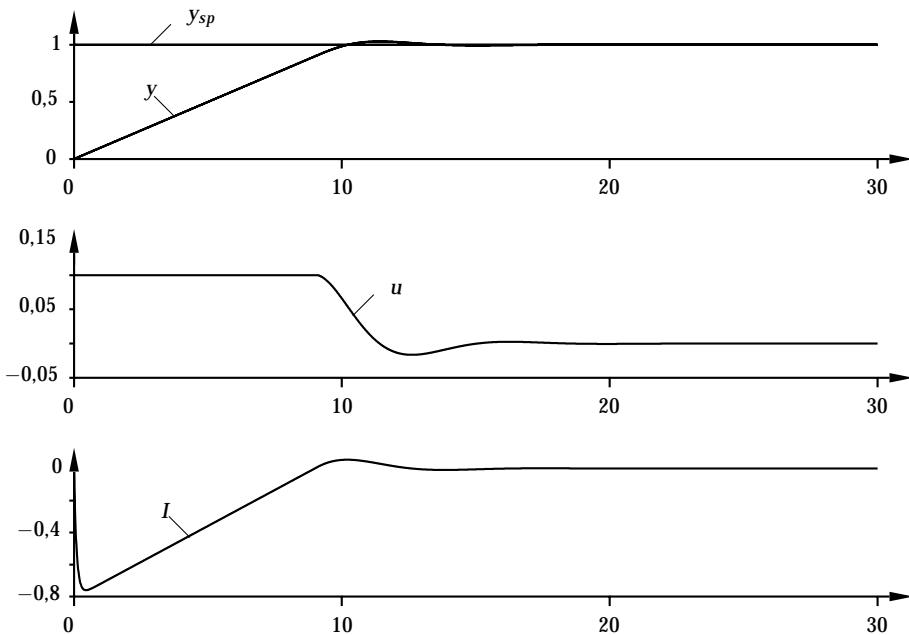


Figura 3.14 Controlador con anti windup aplicado al sistema de la Figura 3.12. Los diagramas muestran la salida del proceso y , el punto de consigna y_{sp} , la señal de control u , y la parte integral I .

el límite de la saturación y la integral tiene un valor negativo durante la fase inicial cuando el actuador se satura. Esta conducta es drásticamente diferente de la de la Figura 3.12, donde la integral tiene un valor positivo durante el transitorio inicial. Nótese también la mejora drástica en comportamiento comparada con el controlador PI ordinario usado en la Figura 3.12.

En la Figura 3.15 se ilustra el efecto de cambiar los valores de la constante de seguimiento. De esta figura, puede así parecer que es una ventaja escoger siempre un valor muy pequeño de la constante de tiempo porque el integrador se resetea entonces rápidamente. Sin embargo, se debe tener algún cuidado al introducir anti windup en sistemas con acción derivativa. Si la constante de tiempo escogida es demasiado pequeña, errores espúreos pueden causar la saturación de la salida, lo cual accidentalmente resetea al integrador. La constante de tiempo de seguimiento T_t debería ser mayor que T_d y más pequeña que T_i . Una regla heurística que ha sido sugerida es escoger $T_t = \sqrt{T_i T_d}$.

Controladores con un modo de seguimiento

Un controlador con recálculo se puede interpretar como que tiene dos modos: el *modo de control* normal, cuando opera como un controlador ordinario, y un *modo de seguimiento*, cuando el controlador está en seguimiento de modo que coinciden las entradas y salidas dadas. Como un controlador con seguimiento puede operar en dos modos, podemos esperar que sea necesario tener una señal lógica para la comutación de modo. Sin embargo, esto no es necesario, porque el seguimiento

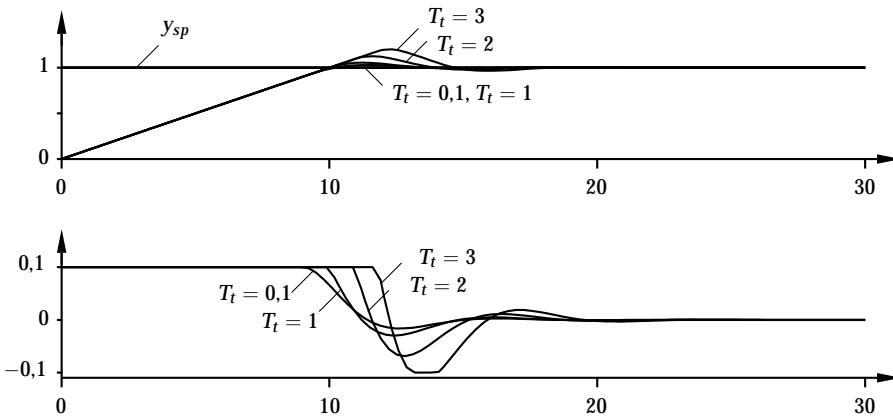


Figura 3.15 La respuesta a un escalón del sistema en la Figura 3.12 para diferentes valores de la constante de tiempo de seguimiento T_t . La curva superior muestra la salida del proceso y y el punto de consigna y_{sp} , y la curva inferior muestra la señal de control u .

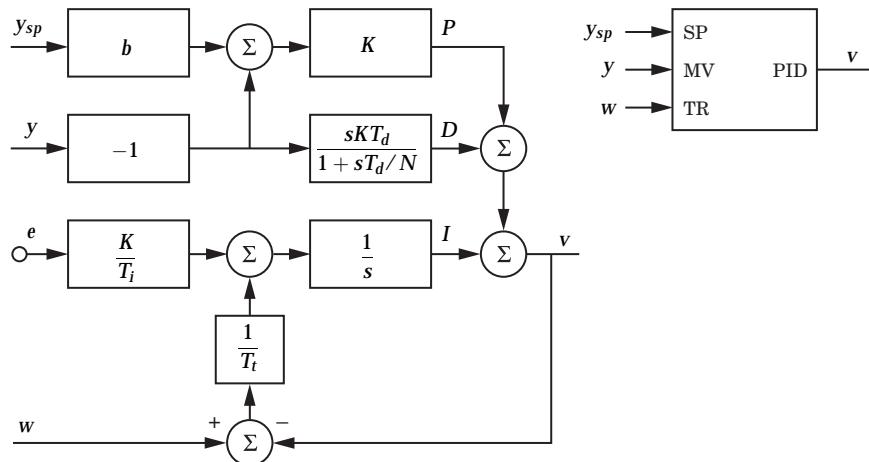


Figura 3.16 Diagrama de bloques y representación simplificada del módulo PID con señal de seguimiento.

se inhibe automáticamente cuando la señal de seguimiento w es igual a la salida del controlador. Esto se puede usar con gran ventaja cuando se construyen sistemas complejos con selectores y control en cascada.

La Figura 3.16 muestra un módulo PID con una señal de seguimiento. El módulo tiene tres entradas: el punto de consigna, la salida medida, y una señal de seguimiento. La nueva entrada TR se llama una señal de seguimiento porque la salida del controlador seguirá a esta señal. Obsérvese que el seguimiento se inhibe cuando $w = v$. Usando el módulo el sistema mostrado en la Figura 3.13 se puede presentar tal como se muestra en la Figura 3.17.

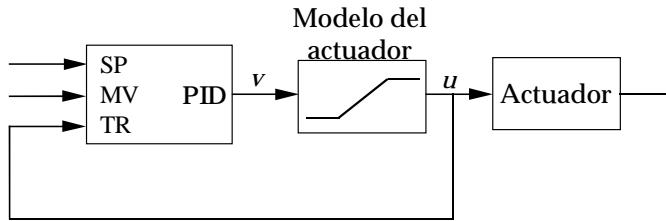


Figura 3.17 Representación del controlador con anti windup de la Figura 3.13 empleando el módulo de control básico con seguimiento mostrado en la Figura 3.16.

La banda proporcional

La noción de banda proporcional es útil para comprender el efecto del windup y explicar esquemas para proporcionar mecanismos de anti windup. La *banda proporcional* es un intervalo tal que el actuador no se satura si el valor instantáneo de la salida del proceso o su valor predicho está en el intervalo. Para control PID sin limitación de la ganancia derivativa, la señal de control está dada por

$$u = K(by_{sp} - y) + I - KT_d \frac{dy}{dt}. \quad (3.21)$$

Resolviendo la salida del proceso predicha

$$y_p = y + T_d \frac{dy}{dt},$$

da la banda proporcional (y_l, y_h) como

$$\begin{aligned} y_l &= by_{sp} + \frac{I - u_{\max}}{K} \\ y_h &= by_{sp} + \frac{I - u_{\min}}{K}. \end{aligned} \quad (3.22)$$

Y u_{\min}, u_{\max} son los valores de la señal de control para el cual satura el actuador. El controlador opera en el modo lineal, si la salida predicha está en la banda proporcional. La señal de control satura cuando la salida predicha está fuera de la banda proporcional. Nótese que la banda proporcional se puede desplazar cambiando el término integral.

Para ilustrar que la banda proporcional es útil para comprender el windup, mostramos la banda proporcional en la Figura 3.18 para el sistema analizado en el Ejemplo 3.2. (Comparar con la Figura 3.12.) La figura muestra que la banda proporcional comienza a moverse hacia arriba porque aumenta el término integral. Esto implica que la salida no alcanza la banda proporcional hasta que es mucho más grande que el punto de consigna. Cuando se alcanza la banda proporcional la señal de control disminuye rápidamente. Sin embargo, este cambio de la banda proporcional es tan rápido, que la salida se mueve con mucha rapidez a través de la banda, y este proceso se repite algunas veces.

La noción de banda proporcional nos ayuda a comprender algunos esquemas de anti windup. La Figura 3.19 muestra la banda proporcional para el sistema

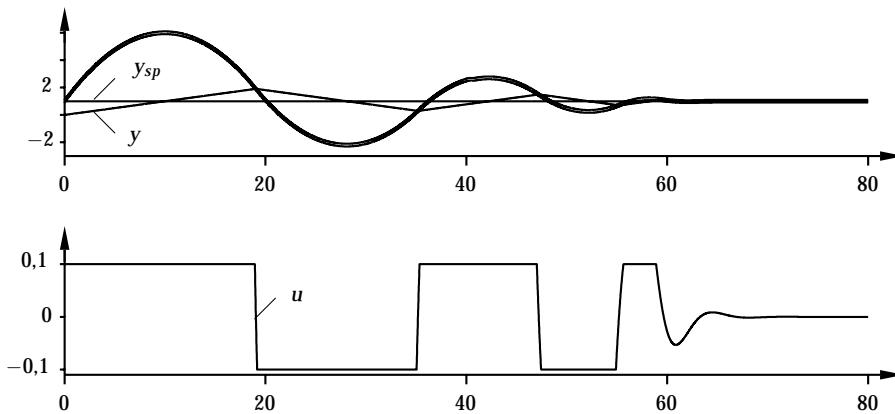


Figura 3.18 La banda proporcional para el sistema del Ejemplo 3.2. El diagrama superior muestra la salida del proceso y y la banda proporcional. El diagrama inferior muestra la señal de control u .

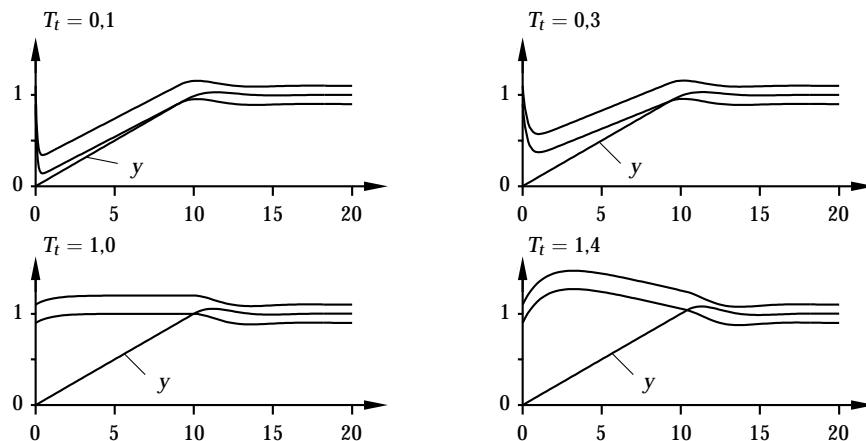


Figura 3.19 La banda proporcional y la salida del proceso y para un sistema con integración condicional y seguimiento con diferentes constantes de tiempo de seguimiento T_t .

con seguimiento para diferentes valores de la constante de tiempo de seguimiento T_t . La figura muestra que la constante de tiempo de seguimiento tiene una influencia significativa sobre la banda proporcional. A causa del seguimiento, la banda proporcional se mueve más próxima a la salida del proceso. Con cuánta rapidez se hace esto viene gobernado por la constante de tiempo de seguimiento T_t . Obsérvese que puede haber una desventaja en moverse demasiado rápidamente, puesto que la salida predicha puede entonces moverse dentro de la banda proporcional debido al ruido y originar así que la señal de control disminuya innecesariamente.

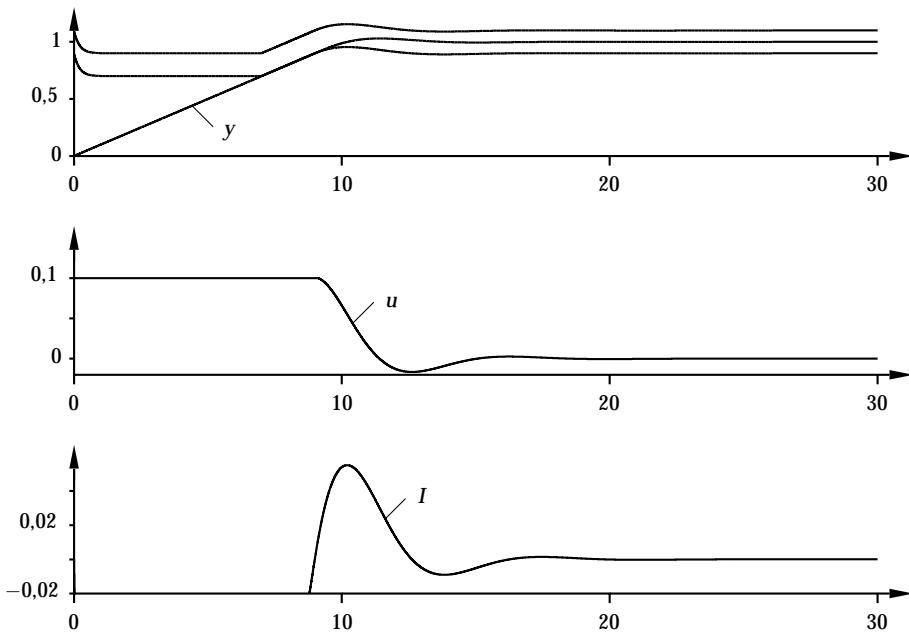


Figura 3.20 Simulación del sistema en el Ejemplo 3.2 con integración condicional. Los diagramas muestran la banda proporcional, la salida del proceso y , la señal de control u , y la parte integral I .

Integración condicional

La integración condicional es una alternativa para el recálculo o seguimiento. En este método la integración se desconecta cuando el control está lejos del estado estacionario. La acción integral se utiliza así solamente cuando se cumplen ciertas condiciones; en otro caso el término integral se mantiene constante. El método se llama también limitación del integrador.

Las condiciones cuando se inhibe la integración se puede expresar de muchas formas diferentes. La Figura 3.20 muestra una simulación del sistema del Ejemplo 3.2 con integración condicional tal que el término integral se mantiene constante durante la saturación. Una comparación con la Figura 3.19 muestra que, en este caso particular, hay muy poca diferencia en comportamiento entre la integración condicional y el seguimiento. Sin embargo, los diversos esquemas de windup, mueven de forma diferente las bandas proporcionales.

Se consideran ahora unas pocas diferentes condiciones de comutación. Un enfoque simple es desconectar la integración cuando el error de control es grande. Otro enfoque es desconectar la integración durante la saturación. Ambos métodos tienen la desventaja de que el controlador puede quedarse encallado en un error de control distinto de cero si el término integral tiene un gran valor en el instante de desconexión.

Un método sin esta desventaja es el siguiente: la integración se desconecta cuando el controlador se satura y la actualización del integrador es tal que origina que la señal de control se hace más saturada. Supóngase, por ejemplo, que el

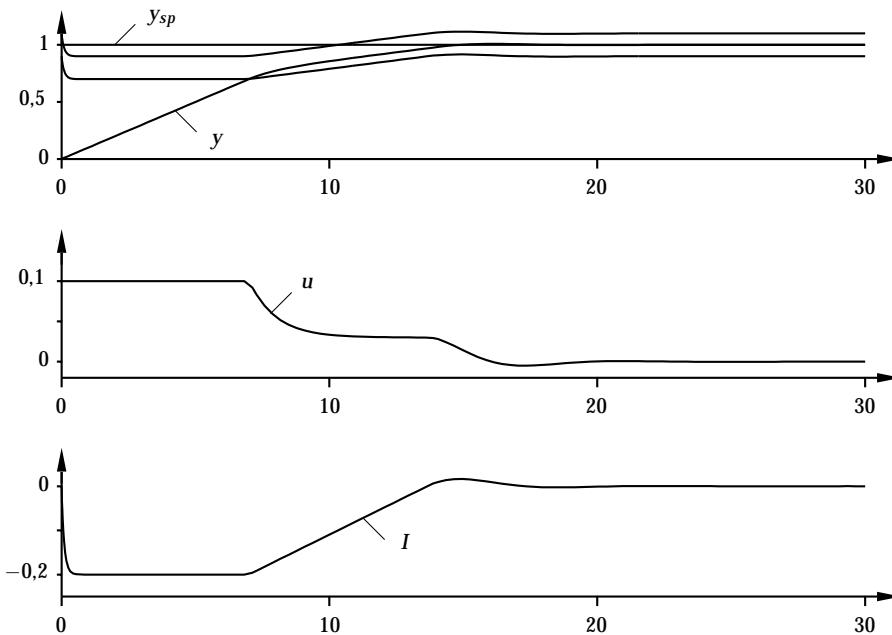


Figura 3.21 Ajuste de la banda proporcional utilizando los parámetros de precarga. Los diagramas muestran la banda proporcional, el punto de consigna y_{sp} , la salida del proceso y , la señal de control u , y la parte integral I .

controlador se satura en el nivel superior de saturación. La integración entonces se desconecta si el error de control es positivo, pero no si es negativo.

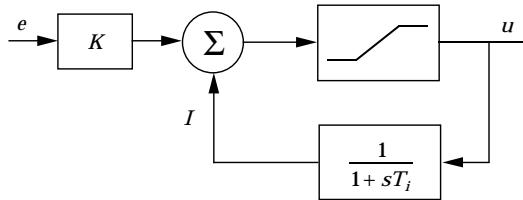
Algunos métodos de integración condicional se utilizan principalmente en arranque de procesos por lotes, cuando pueden haber grandes cambios en el punto de consigna. Una versión particular, utilizada en el control de temperatura, fija la banda proporcional fuera del punto de consigna cuando hay grandes desviaciones de control. Esta desviación se puede emplear para ajustar la respuesta transitoria obtenida durante el arranque del proceso. Los parámetros usados se llaman de vuelta atrás o precarga (ver Figura 3.21). En este sistema la banda proporcional se posiciona con un extremo en el punto de consigna y el otro extremo en el valor medido cuando hay grandes variaciones. Estos métodos pueden dar windup durante las perturbaciones.

Implementación serie

En la Figura 3.3, se mostró una implementación especial de un controlador en forma interactuante. Para evitar windup en este controlador podemos incorporar un modelo de la saturación en el sistema tal como se muestra en la Figura 3.22a. Obsérvese que en esta implementación la constante de tiempo de seguimiento T_t es la misma que el tiempo de integración T_i . Este valor de la constante de tiempo de seguimiento es a menudo demasiado grande.

En la Figura 3.22a, el modelo de la saturación limitará directamente la señal de control. Es importante, por lo tanto, tener un buen modelo de la saturación

a)



b)

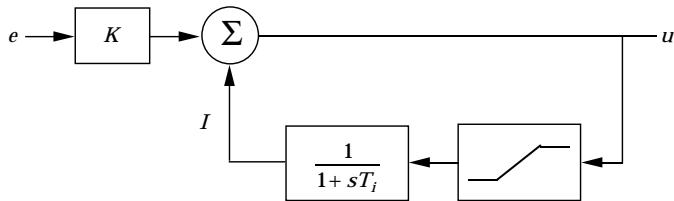


Figura 3.22 Dos formas de proporcionar anti-windup en el controlador de la Figura 3.3 donde la acción integral se genera como reset automático.

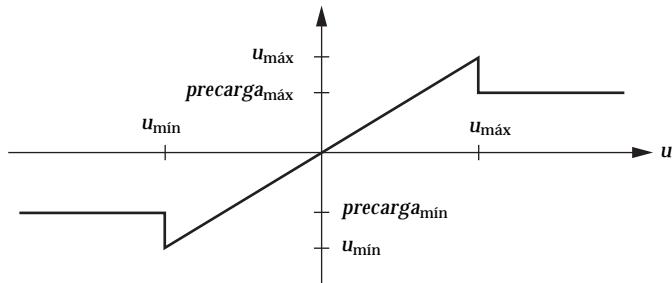


Figura 3.23 Una «unidad batch» utilizada para proporcionar anti windup en el controlador de la Figura 3.3.

física. Una restricción demasiado dura causará una limitación innecesaria de la acción de control. Una limitación demasiado débil producirá windup.

Se proporciona más flexibilidad si la saturación se posiciona como en la Figura 3.22b. En este caso, la saturación no influirá sobre la parte proporcional del controlador. Con esta estructura es posible también forzar la parte integral para asumir otros valores de precarga durante la saturación. Esto se logra sustituyendo la función de saturación por la no linealidad mostrada en la Figura 3.23. Este procedimiento de anti windup se llama algunas veces una «unidad batch» y se puede considerar como un tipo de integración condicional. Se utiliza principalmente para ajustar la *sobreelongación* durante el arranque cuando hay un gran cambio en el punto de consigna. En los primeros controladores mono lazo se proporcionaba la unidad batch como un hardware especial que había que añadir.

Esquemas combinados

El seguimiento y la integración condicional pueden también combinarse. En [177] se sugiere manipular explícitamente la banda proporcional para el control batch. Esto se hace introduciendo los denominados *puntos de corte*. El corte superior está por encima del punto de consigna, y el corte inferior está por debajo. El integrador se fija cuando la salida predicha del proceso está fuera del intervalo de corte. La integración se realiza con una constante de tiempo de seguimiento especificada cuando la salida del proceso está dentro de los puntos de corte. Los puntos de corte se consideran como parámetros del controlador que se ajustan para influir sobre la respuesta frente a grandes cambios en el punto de consigna. Se propone un método similar en [100], donde la integración condicional se combina con un recalcularamiento. En [309], al integrador se le da un valor preestablecido $i = i_0$ durante la saturación. El valor de i_0 se sintoniza para dar una sobrelongación satisfactoria en el arranque. Este enfoque se llama también precarga.

3.6 ¿Cuándo se puede utilizar el control PID?

Hay muchos requisitos sobre un sistema controlado. Debería responder bien a los cambios en el punto de consigna, debería atenuar la perturbaciones de carga, el ruido de medida no debería producir acciones excesivas de control, y el sistema debería ser insensible a las variaciones del proceso. El diseño de un sistema de control también envuelve aspectos de dinámica de proceso, saturación del actuador, y características de la perturbación. Puede parecer sorprendente que un controlador tan sencillo como el controlador PID pueda trabajar tan bien. La observación empírica general es que la mayoría de los procesos industriales se pueden controlar razonablemente bien con control PID a condición de que las demandas sobre el comportamiento del control no sean demasiado exigentes. En los párrafos siguientes profundizaremos más en estas cuestiones aunque primariamente se consideran casos donde el control PID es suficiente y a continuación se analizan algunos problemas genéricos donde es aconsejable un control más sofisticado.

¿Cuándo es el control PI suficiente?

Todos los procesos estables se pueden controlar mediante un controlador I si los requisitos de comportamiento son modestos. La acción proporcional da una mejora adicional de comportamiento. No es por lo tanto sorprendente que el controlador PI sea el controlador más común. Sin tomar en cuenta las saturaciones, a un proceso con dinámica de primer orden se le puede dar un comportamiento deseado utilizando un controlador PI. El control PI se puede utilizar también con procesos que poseen acción integral.

La acción derivativa no se suele utilizar. Es una observación interesante que muchos controladores industriales sólo tienen acción PI y que en otros la acción derivativa puede (y frecuentemente así es) desconectarse. Se puede demostrar que el control PI es adecuado para todos los procesos donde las dinámicas son esencialmente de primer orden (control de nivel en tanques simples, reactores de tanque agitado con mezcla perfecta, etc). Es relativamente fácil descubrir si

éste es el caso midiendo la respuesta a un escalón o la respuesta en frecuencia del proceso. Si la respuesta a un escalón se parece a aquella de un sistema de primer orden o, más precisamente, si la curva de Nyquist permanece sólo en el primer y cuarto cuadrante, entonces el control PI es suficiente. Otra razón es que el proceso ha sido diseñado de forma que su operación no requiere un control muy exigente. Entonces, incluso si el proceso tiene dinámicas de orden superior, lo que se necesita es una acción para proporcionar una desviación nula en estado estacionario y una respuesta transitoria adecuada mediante acción proporcional.

¿Cuándo es útil la acción derivativa?

Un doble integrador no se puede controlar por un controlador PI. La razón es que el proceso tiene un desfase de 180° y que un controlador PI también tiene un desfase. Se necesita la acción derivativa para tales procesos. Sin tomar en cuenta las saturaciones a un proceso con dinámica de segundo orden se le puede dotar de cualquier comportamiento deseado usando un controlador PID.

Similarmente, el control PID es suficiente para procesos donde las dinámicas dominantes son de segundo orden. Para tales procesos no se obtiene ningún beneficio utilizando un controlador más complejo. Un caso típico donde la acción derivativa mejora la respuesta es cuando las dinámicas se caracterizan por constantes de tiempo que difieren en magnitud. La acción derivativa puede usarse entonces beneficiosamente para acelerar la respuesta. El control de temperatura es un caso típico. La acción derivativa es también rentable cuando se requiere un control exigente de un sistema de orden superior. La dinámica de esta clase de sistemas limitaría la cantidad de ganancia proporcional para un buen control. Con una acción derivativa, se proporciona una mejora en el amortiguamiento lo que implica que se puede emplear una ganancia proporcional más elevada para acelerar la respuesta transitoria.

Muchos procesos encontrados en control de procesos tienen dinámicas con respuestas a un escalón esencialmente monótonas, a menudo con retardo de tiempo. Si la dinámica está dominada por el retardo la acción derivativa da unas mejoras de comportamiento modestas comparadas con el control PI, sin embargo la acción derivativa da mejoras significativas en el caso de procesos donde domina la constante de tiempo. Esta cuestión se analiza con más detalle en el Capítulo 7.

Cuándo se necesita un control más sofisticado?

Los beneficios de emplear un controlador más sofisticado que el control PID se demuestra a continuación mediante algunos ejemplos.

Procesos de orden superior Cuando el sistema es de orden superior a dos, el control se puede mejorar usando un controlador más complejo que el controlador PID. Esto se ilustra por el ejemplo siguiente.

EJEMPLO 3.3—SISTEMAS DE ORDEN SUPERIOR

Sea un proceso de orden superior descrito por la siguiente función de transferencia

$$P(s) = \frac{1}{(s+1)^3}.$$

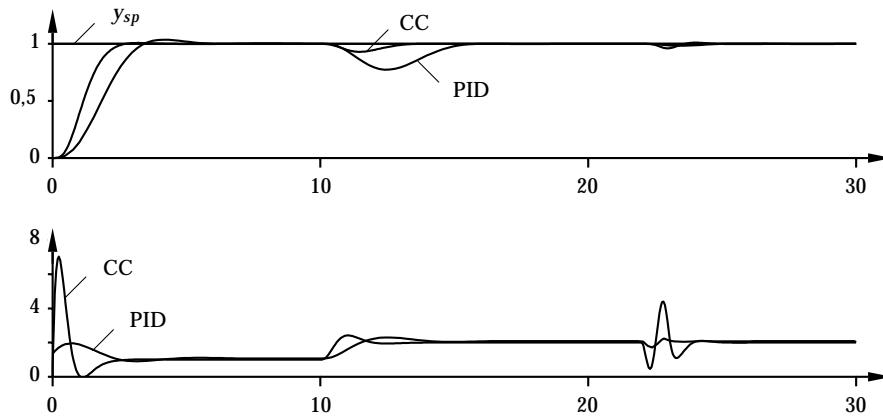


Figura 3.24 Control del sistema de tercer orden en el Ejemplo 3.3 usando un controlador PID (PID) y un controlador más complejo (CC). La figura muestra las respuestas a un cambio en el punto de consigna, una perturbación en la carga, y finalmente una perturbación en la medida. El diagrama superior muestra el punto de consigna y_{sp} y la medida de señal y , y el diagrama inferior muestra la señal de control u .

La Figura 3.24 muestra el control obtenido usando un controlador PID y un controlador más complejo de orden superior.

El controlador PID tiene los parámetros $K = 3,4$, $T_i = 2,0$, y $T_d = 0,6$. El controlador PID se compara con un controlador de la forma

$$R(s)u(t) = -S(s)y(t) + T(s)y_{sp}(t),$$

con los siguientes polinomios del controlador

$$\begin{aligned} R(s) &= s(s^2 + 11,5s + 57,5) \\ S(s) &= 144s^3 + 575s^2 + 870s + 512 \\ T(s) &= 8s^3 + 77s^2 + 309s + 512. \end{aligned}$$

En la figura se demuestran los beneficios de emplear un controlador más complejo en el caso de dinámicas de orden superior. □

Sistemas con grandes retardos de tiempo El control de sistemas con un retardo de tiempo dominante es notoriamente difícil. Es también un tema sobre el que hay muchas opiniones diferentes con respecto a si es ventajoso utilizar el control PID. Parece que hay un acuerdo general en que la acción derivativa no es de mucha utilidad en procesos con retardos de tiempo dominante. Para procesos estables en lazo abierto, la respuesta a señales de entrada se puede mejorar sustancialmente al introducir una compensación para el retardo de tiempo. También en cierta medida se puede mejorar el rechazo a la perturbación de carga porque un compensador del retardo de tiempo permite una ganancia del lazo más elevada que un controlador PID. Los sistemas con retardos de tiempo dominante son así candidatos para un control más sofisticado.

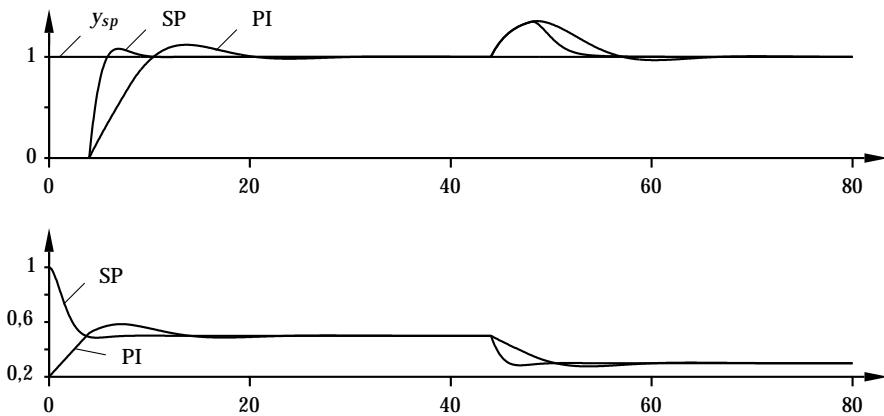


Figura 3.25 Control del sistema en el Ejemplo 3.4 con control PI (PI) y con un predictor de Smith (SP). El diagrama superior muestra el punto de consigna y_{sp} y la señal medida y , y el diagrama inferior muestra la señal de control u .

EJEMPLO 3.4—COMPENSACIÓN DE RETARDOS DE TIEMPO

Sea un proceso con la función de transferencia

$$P(s) = \frac{1}{1+2s} e^{-4s},$$

que tiene un retardo de tiempo significativo. La Figura 3.25 muestra una simulación del sistema en lazo cerrado obtenido con un controlador PI con una ganancia $K = 0.2$ y un tiempo integral de $T_i = 2.5$. Para comparación la figura también muestra el comportamiento con un predictor de Smith, que es un controlador especial para un sistema con retardo de tiempo. Este controlador se analizará en detalle en el Capítulo 8. La respuesta a cambios en el punto de consigna es mucho mejor en el predictor de Smith, pero la mejora en la respuesta a perturbaciones de carga es menor. \square

Sistemas con modos oscilatorios Sistemas con modos oscilatorios ocurren en aplicaciones tales como brazos de robot flexibles, unidades de disco y memorias ópticas, microscopios de fuerza atómica (AFM), sistemas micro-mecánicos (MEMS), estructuras espaciales flexibles, y sistemas de combustión. Hay dificultades particulares cuando el amortiguamiento es muy bajo, así que el sistema es altamente resonante. Aplicaciones típicas están en sistemas micro-mecánicos y en microscopios de fuerza atómica. Los sistemas con modos resonantes no son tan comunes en aplicaciones de control de procesos. La acción derivativa puede dar una mejora drástica en los sistemas oscilatorios tal como se ilustra en el siguiente ejemplo.

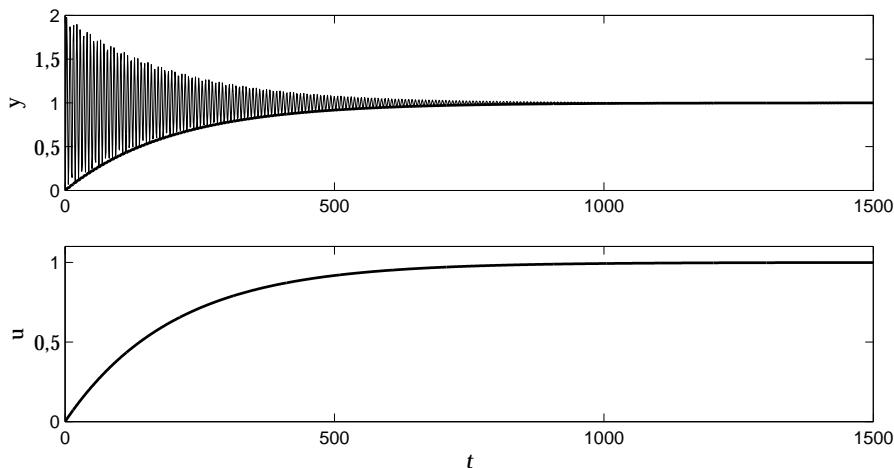


Figura 3.26 Salida y señales de control para control PI del sistema oscilatorio. La señal osilante es la respuesta en lazo abierto.

EJEMPLO 3.5—UN SISTEMA OSCILATORIO CON BAJO AMORTIGUAMIENTO

Sea un sistema con la función de transferencia normalizada

$$P(s) = \frac{a^2}{s^2 + 2\zeta as + a^2}.$$

Consideraremos sistemas con amortiguamiento relativo muy bajo $\zeta = 0,005$. El comportamiento obtenido con un controlador PI está limitado severamente por el bajo amortiguamiento relativo del proceso. Como un controlador PI no puede proporcionar ningún adelanto de fase el amortiguamiento de los modos oscilatorios del sistema en lazo cerrado será más pequeño que la del sistema en lazo abierto. Un requisito clave para que el controlador PI no excite los modos de alta frecuencia.

Un controlador que sea un integrador puro es una elección razonable. La condición de estabilidad para tal controlador es $k_i < 2\zeta a^3$, que implica que $k_i = \zeta a^3$ es un buen valor de la ganancia del controlador. Con esta elección, el sistema en lazo cerrado tiene el mismo tiempo de asentamiento que el sistema en lazo abierto. El tiempo de respuesta sólo se puede mejorar un poco añadiendo acción proporcional. La Figura 3.26 muestra la entrada y la salida para un cambio en escalón en el punto de consigna para un controlador con estos parámetros. La respuesta a un escalón tiene un tiempo de asentamiento de 1500. La razón de por qué el sistema tiene que ser tan lento es que el movimiento oscilatorio no se puede amortiguar por el controlador PI y es, por lo tanto, necesario tener un controlador lento de forma que los modos oscilatorios no se exciten por el controlador. En la figura hemos mostrado también la respuesta a un escalón del proceso.

El comportamiento se puede mejorar drásticamente utilizando un controlador

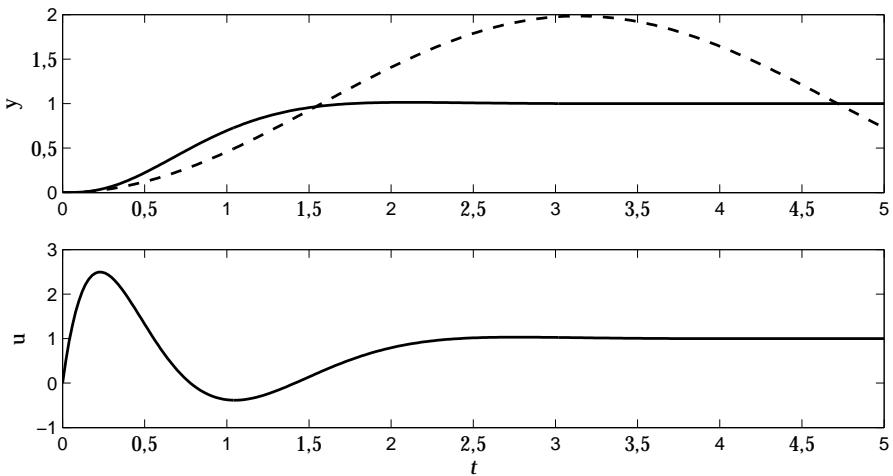


Figura 3.27 Señales de salida y de control para control PID control del sistema oscilatorio. La curva a tramos es la respuesta en lazo abierto.

PID. Una posibilidad es usar un controlador PID con parámetros

$$\begin{aligned} k &= (1 + 2\alpha_0\zeta_0)\frac{\omega_0^2}{a^2} - 1 \\ k_i &= \frac{\alpha_0\omega_0^3}{a^2} \\ k_d &= \frac{(\alpha_0 + 2\zeta_0)\omega_0 - 2\zeta a}{a^2}. \end{aligned} \quad (3.23)$$

Aquí \$\omega_0\$, \$\alpha_0\$, y \$\zeta_0\$ son parámetros de diseño que dan las propiedades del sistema en lazo cerrado. Una elección razonable es \$\omega_0 = 3a\$, \$\alpha_0 = 1\$, \$\zeta_0 = 0.5\$. Las elecciones que son adecuadas para las fórmulas y los parámetros se darán posteriormente en la Sección 6.4. La Figura 3.27 muestra las respuestas de la salida y del controlador frente a un cambio en escalón en el punto de consigna. En este caso el sistema tiene un tiempo de asentamiento de 2 s. ¡Esto es tres órdenes de magnitud mejor que con control PI! La razón para esto es que al utilizar la acción derivativa es posible amortiguar las oscilaciones. Esto se indica en la figura al mostrar la respuesta en lazo abierto del proceso en línea a tramos. También obsérvese la diferencia drástica en las señales de control para control PI y PID. Es también importante usar la ponderación del punto de consigna con \$b = 0\$ para evitar un cambio rápido de la variable de control. Tal cambio excitará los modos oscilatorios pobemente amortiguados. \square

Resumen

Cuando las dinámicas de los procesos que se van a controlar son sencillas, un controlador PID es suficiente. Cuando la dinámica se hace más complicada, el comportamiento se puede mejorar utilizando una estructura de controlador más

sofisticada que el controlador PID. Anteriormente se han dado ejemplos de tales procesos.

Para algunos sistemas con grandes variaciones de parámetros es posible diseñar controladores lineales que permiten operar sobre un rango amplio de parámetros. Sin embargo, tales controladores son con frecuencia de orden superior.

El control de variables de procesos que están relacionadas estrechamente con variables de calidad importante pueden ser de un valor económico significativo. En tales lazos de control es frecuentemente necesario seleccionar el controlador con respecto a las características de la perturbación. Esto a menudo conduce a estrategias que no son del tipo PID. Estos problemas suelen estar asociados con retardos de tiempo.

Un controlador general intenta modelar las perturbaciones actuando sobre el sistema. Como un controlador PID tiene complejidad limitada, no puede en general modelar conducta de perturbaciones complejas ni en particular perturbaciones periódicas.

3.7 Resumen

Se ha realizado una presentación detallada del algoritmo PID. Deben hacerse algunas modificaciones de la versión «libro de texto» para obtener un controlador práctico útil. Los problemas que se deben manejar son, por ejemplo, windup integral y la introducción de valores en el punto de consigna. También se ha efectuado un análisis de las limitaciones del algoritmo PID y una caracterización de procesos en los cuales el controlador PID efectúa el control.

3.8 Notas y referencias

En [47] se presenta un interesante resumen del controlador PID. Se utilizó realimentación proporcional en la forma de un regulador centrífugo para controlar la velocidad de los molinos de viento alrededor de 1750. James Watt empleó un sistema similar para el control de velocidad de máquinas de vapor en 1788. Posteriormente fue descubierta la acción integral por algunos autores. Se explica analíticamente por [235] y [350]. El control por realimentación con acción proporcional e integral fue redescubierto muchas veces después de eso. En las primeras etapas, el desarrollo de controladores estuvo estrechamente relacionado con el desarrollo de sensores y actuadores. Medida, actuación y control eran a menudo combinadas en el mismo dispositivo. Hubo también confusión con las acciones integral y derivativa porque algunos controladores actuaban a través de motores que tenían acción integral.

El controlador PID en la forma que lo conocemos hoy emergió en el período de 1915 a 1940. Los desarrollos fundamentales se hicieron en compañías de instrumentación legendarias tales como Bristol, Fisher, Foxboro, Honeywell, Leeds & Northrup, Mason-Neilan, y Taylor Instrument. La acción integral se llamó reset automático porque reemplazó al reset manual que se utilizaba en los controladores proporcionales para obtener el valor correcto en estado estacionario. El

potencial de un controlador que pudiese anticipar los errores de control en el futuro se analizó en los 1920. Sin embargo, llevó algún tiempo antes de que la idea se pudiese implementar. Un controlador con acción derivativa fue introducido por Ralph Clarridge de Taylor Instrument Company en 1935. En ese tiempo la función se llamó «pre-actuación». Una revisión interesante de la primera historia de los controladores PID se da por [327]. Hay también mucho material interesante en publicaciones de las compañías de instrumentación. La entrevista con Ziegler, quien es uno de los pioneros en nuestro campo, en [54], da una perspectiva sobre los primeros desarrollos; otro material interesante se encuentra en [46].

Es interesante observar que la realimentación fue crucial para la construcción del propio controlador. Los primeros sistemas neumáticos usaron la idea de que se puede obtener un controlador esencialmente lineal mediante un lazo de realimentación compuesto de componentes pasivos lineales, un amplificador no lineal y una válvula de mariposa. Ideas similares se emplearon en los controladores electrónicos con motores eléctricos y relés.

Muchas de las modificaciones de utilidad práctica de los controladores aparecieron primero como funciones especiales de hardware. No fueron expresadas en forma matemática. Un análisis matemático temprano de una máquina de vapor con un regulador fue hecho independientemente por [235] y [350]. Este análisis claramente demostraba la diferencia entre control proporcional e integral. Los trabajos [244, 210, 256, 165] estuvieron disponibles al tiempo que se desarrollaban los controladores PID. Sin embargo, hay pocas pruebas de que los ingenieros en el campo del control de procesos los conociesen. El control de procesos, por lo tanto, se desarrolló de forma independiente. Dos de los primeros trabajos [130] y [186] fueron escritos por ingenieros de la Dow Chemical Company. Hubo también contribuciones de los investigadores de la universidad [72] y [162].

El controlador PID ha seguido un desarrollo interesante debido a los cambios drásticos de tecnología que han acontecido desde 1930. El controlador neumático mejoró sustancialmente al hacer un uso sistemático del principio de balance de fuerza. La neumática fue reemplazada por la electrónica cuando apareció el amplificador operacional en los 1950. La emergencia del control por computador en los 1960 fue un desarrollo importante. En los primeros sistemas de control por computador, el computador calculaba los puntos de consigna de los controladores analógicos. La próxima etapa del desarrollo fue el control digital directo (DDC), donde el computador estaba controlando directamente al actuador; ver [358]. A continuación se utilizó un computador digital para implementar muchos controladores PID. Este desarrollo condujo a una reconsideración de muchos de los fundamentos del control PID; ver [127], [217], [245], y [264]. La aparición de los microprocesadores en los 1970 posibilitó usar control digital para los controladores mono lazos. También condujo al desarrollo de los sistemas de control distribuidos para control de procesos, donde los controladores PID eran un elemento clave; ver [220]. Cuando aumentó la potencia de cálculo de los microprocesadores fue posible introducir sintonía, adaptación, y diagnóstico en los controladores mono lazos. Este desarrollo comenzó en los 1980 y se ha acelerado en los 1990; ver [14].

Es interesante observar que muchos hechos acerca del control PID fueron redescubiertos en conexión con los cambios de tecnología. Siendo una razón que

muchos aspectos prácticos del control PID se consideran como información confidencial que no era fácilmente accesible en las publicaciones. Mucha información útil estuvo también muy diseminada en las publicaciones técnicas.

Se emplearon dos enfoques diferentes para tratar con los cambios en el punto de consigna en los primeros controladores. Algunos controladores utilizaban realimentación del error pero otros introducían el punto de consigna sólo en la parte integral. El efecto de esto es que se puede reducir la sobreelongación que ocurre con los cambios en el punto de consigna. La idea de que se puede realizar una separación de las respuestas a puntos de consigna y perturbaciones de carga usando un controlador con dos grados de libertad se introdujo en [176]. La aplicación al control PID se hizo en [11]. Una primera aplicación industrial se describe en [306], ver también [10, 333]. La ponderación del punto de consigna donde se introduce una fracción ajustable del punto de consigna en las partes proporcional y derivativa es ahora una característica común de los controladores PID.

El fenómeno de la saturación del término integral era bien conocido en las primeras implementaciones analógicas. Las estructuras del controlador utilizadas eran a menudo tales que se evitaban el windup. Los esquemas de anti-windup fueron redescubiertos en conexión con el desarrollo del control digital directo. Esto se analiza en [108]. Mucho trabajo para evitar el windup ha sido hecho desde entonces, y el windup ha encontrado su camino en los libros de texto de control; ver [30]. Hay muchos trabajos escritos sobre el fenómeno del windup; ver [203, 125, 205, 114, 192, 124, 177, 35, 155, 76, 122, 156, 375, 26, 299, 351]. En el libro [123] se da un tratamiento detallado del problema del windup. La conmutación de modo se trata en el trabajo [35].

4

Diseño del controlador

4.1 Introducción

El diseño de sistemas de control es un campo muy rico. En los últimos 50 años ha habido avances sustanciales que han dado como resultado un mejor conocimiento y comprensión así como la aparición de métodos específicos de diseño. Este desarrollo se ha acelerado por los avances en computación y por el desarrollo de herramientas de diseño basadas en computador. Hablando en términos generales, los controladores PID se han diseñado utilizando dos enfoques diferentes; control basado en modelo y sintonía directa. Los enfoques basados en modelos comienzan con un modelo matemático sencillo del proceso. Se han utilizado modelos muy simples, generalmente un sistema de primer orden con un retardo de tiempo. En sintonía directa se aplica un controlador al proceso, y se efectúan algunos experimentos sencillos para obtener los parámetros del controlador. A causa de la simplicidad de los modelos y del controlador se han desarrollado métodos especiales para el control PID. Desde 1990 ha habido un aumento significativo en el interés por el diseño de controladores PID, parcialmente motivado por las necesidades de dispositivos de sintonía automática para tales controladores.

Para desarrollar métodos de diseño es necesario tener presente que hay un rango muy amplio de tipos diferentes de problemas de control incluso si se resiente al controlador PID. Algunos ejemplos típicos son:

- Diseño de un controlador sencillo para una aplicación no crítica.
- Diseño de un controlador para un proceso especial que minimiza las fluctuaciones en variables de control importantes.
- Desarrollo de una técnica de diseño que se pueda utilizar en un auto sintonizador universal para control PID.

Hay también una serie de cuestiones no técnicas importantes que se deberían considerar: ¿Qué tiempo y esfuerzo se requiere para aplicar el método? ¿Cuál es el nivel de conocimiento requerido por parte del usuario? Una solución al problema de diseño debería también dar una comprensión de cuándo es beneficioso añadir acción derivativa a un controlador PI y cuándo incluso se deberían considerar controladores más complejos.

Este capítulo hace una revisión de ideas y conceptos que son relevantes para control PID. Se intenta así llevar el diseño de los controladores PID más en línea con la corriente principal hoy día imperante en el diseño del control.

4.2 Una rica variedad de problemas de control

Antes de analizar métodos de sintonía específicos es útil darse cuenta de que hay un amplio rango de problemas de control con objetivos muy diversos. Algunos ejemplos son: regulación en estado estacionario, seguimiento del punto de consigna y de trayectorias, y control de buffers y tanques de compensación.

El objetivo de la regulación en estado estacionario es mantener las variables del proceso próximas a los valores deseados. Los problemas claves están originados por las perturbaciones de carga, el ruido de medida, y las variaciones del proceso. La regulación en estado estacionario es muy común en control de procesos.

En el seguimiento del punto de consigna se intenta conseguir que las variables del proceso sigan a una función dada del tiempo o a una curva dada. Estos problemas normalmente ocurren en el control de movimiento y en robótica. En algunos casos, por ejemplo, control de máquinas de herramienta o robótica, la demanda sobre la precisión del seguimiento es muy estricta. En otros casos, por ejemplo, robots en movimiento, los requisitos son menos exigentes. Hay una diferencia significativa entre el seguimiento de una curva dada del tiempo y el seguimiento de trayectoria, que normalmente entraña el control de algunas variables.

Los buffers son usuales en la producción industrial. Se utilizan para suavizar las variaciones entre diferentes procesos de producción, tanto en el control de procesos como en la fabricación discreta. En control de procesos se suelen llamar *tanques de compensación*. Los buffers son también comunes en los sistemas de computación. Se emplean en los servidores para suavizar las variaciones en la demanda de los clientes, y también en las redes de computadores para atemperar las variaciones en la carga. Los buffers son también elementos claves en las cadenas de suministros donde el control efectivo del buffer tiene un gran impacto sobre la rentabilidad. Los niveles del buffer deberían fluctuar; de lo contrario el buffer no funciona. Idealmente, ningún control debería aplicarse a menos que haya un riesgo de desbordamiento. Normalmente se suele utilizar un controlador con integración con baja ganancia y un sistema de planificación que proporciona una ganancia más elevada en los límites del buffer.

Los temas claves en muchos de los problemas de control son la atenuación de las perturbaciones de carga, la inyección del ruido de medida, la robustez a las variaciones del proceso, y el seguimiento del punto de consigna. La importancia relativa de estos factores y los requisitos varían según las aplicaciones, pero deben considerarse todos los factores.

4.3 Fundamentos de la realimentación

En la Figura 4.1 se muestra un diagrama de bloques de un lazo de realimentación básico con un controlador que tiene dos grados de libertad. El proceso se repre-

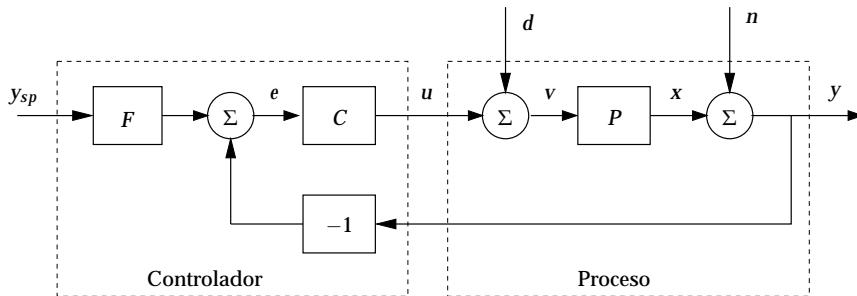


Figura 4.1 Diagrama de bloques de un lazo de realimentación básico que tiene dos grados de libertad.

senta por el bloque P . El controlador se representa por el bloque de realimentación C y la parte feedforward F . Para un controlador PID ideal con ponderación del punto de consigna tenemos

$$\begin{aligned} C(s) &= K \left(1 + \frac{1}{sT_i} + sT_d \right) \\ F(s) &= \frac{b + \frac{1}{sT_i} + csT_d}{1 + \frac{1}{sT_i} + sT_d}. \end{aligned} \quad (4.1)$$

Comparar con (3.7) y (3.20). La señal u es la señal de control, y la señal x es la variable de proceso real. La información acerca de x se obtiene de la señal del sensor y , que está contaminada por el ruido de medida n . La señal d representa las perturbaciones de carga que mueven al sistema fuera de su estado deseado. Esta señal puede entrar al proceso de diferentes formas; en la Figura 4.1 se supone que actúa sobre la entrada del proceso.

El objetivo del diseño del control es determinar las funciones de transferencias C y F tal que la variable del proceso x esté próxima al punto de consigna y_{sp} a pesar de las perturbaciones de carga, el ruido de medida, y las incertidumbres del proceso. La realimentación puede reducir el efecto de las perturbaciones de carga. Debido a la realimentación el ruido de medida se realimenta en el sistema. Es esencial asegurarse de que esto no origina grandes variaciones en la variable del proceso. Como el modelo nunca es totalmente preciso es esencial que la conducta del sistema en lazo cerrado sea insensible a las variaciones en el proceso. La función de transferencia feedforward F se diseña para dar la respuesta deseada a cambios en el punto de consigna.

Relaciones fundamentales

El lazo de realimentación está influenciado por tres señales externas, el punto de consigna y_{sp} , la perturbación de carga d , y el ruido de medida n . Hay al menos tres señales x , y , y u que son de gran interés para el control. Esto significa que hay nueve relaciones entre las señales de entrada y de salida. Como el sistema es lineal estas relaciones se pueden expresar en términos de las funciones de transferencia. Sean X , Y , U , D , N , y Y_{sp} las transformadas de Laplace de x , y , u , d , n ,

e y_{sp} , respectivamente. Del diagrama de bloques en la Figura 4.1 se obtienen las siguientes relaciones:

$$\begin{aligned} X &= \frac{PCF}{1+PC} Y_{sp} + \frac{P}{1+PC} D - \frac{PC}{1+PC} N \\ Y &= \frac{PCF}{1+PC} Y_{sp} + \frac{P}{1+PC} D + \frac{1}{1+PC} N \\ U &= \frac{CF}{1+PC} Y_{sp} - \frac{PC}{1+PC} D - \frac{C}{1+PC} N. \end{aligned} \quad (4.2)$$

Hay algunas conclusiones interesantes que podemos sacar de estas ecuaciones. En primer lugar, podemos observar que algunas funciones de transferencias son las mismas y que todas las relaciones están determinadas por las siguientes seis funciones de transferencias, llamadas en sentido figurado la *banda de los seis*.

$$\begin{array}{ccc} \frac{PCF}{1+PC} & \frac{PC}{1+PC} & \frac{P}{1+PC} \\ \frac{CF}{1+PC} & \frac{C}{1+PC} & \frac{1}{1+PC} \\ \hline \frac{1}{1+PC} & \frac{1}{1+PC} & \frac{1}{1+PC} \end{array} \quad (4.3)$$

Las funciones de transferencia en la primera columna dan la respuesta de la variable del proceso y de la señal de control con respecto al punto de consigna. La segunda columna da las mismas señales en el caso de realimentación pura del error cuando $F = 1$. La función de transferencia $P/(1+PC)$ en la tercera columna muestra cómo la variable del proceso reacciona frente a las perturbaciones de carga, y la función de transferencia $C/(1+PC)$ da la respuesta de la señal de control al ruido de medida.

Obsérvese que sólo se requieren cuatro funciones de transferencia,

$$\begin{array}{ccc} \frac{PC}{1+PC} & \frac{P}{1+PC} \\ \frac{C}{1+PC} & \frac{1}{1+PC} \\ \hline \frac{1}{1+PC} & \frac{1}{1+PC} \end{array} \quad (4.4)$$

para describir cómo el sistema reacciona frente a las perturbaciones de carga y al ruido de medida. Estas funciones de transferencia las denominaremos la *banda de los cuatro*. También capturan la robustez, como se analizará en la Sección 4.6. Se necesitan dos funciones de transferencias adicionales para describir cómo el sistema responde a los cambios en el punto de consigna.

El caso especial cuando $F = 1$ se llama un sistema con realimentación del error (puro). En este caso, todas las acciones de control se basan sólo en la realimentación del error, que queda caracterizado completamente por la banda de los cuatro (4.4).

Con frecuencia estamos interesados en la magnitud de las funciones de transferencias dadas por (4.4). Es importante tener en cuenta que las funciones de transferencias $PC/(1+PC)$ y $1/(1+PC)$ son adimensionales, pero las funciones de transferencias $P/(1+PC)$ y $C/(1+PC)$ no lo son. Para objetivos prácticos es por lo tanto importante normalizar las señales, por ejemplo, escalando las entradas y las salidas del proceso en el intervalo 0 a 1 o -1 a 1 .

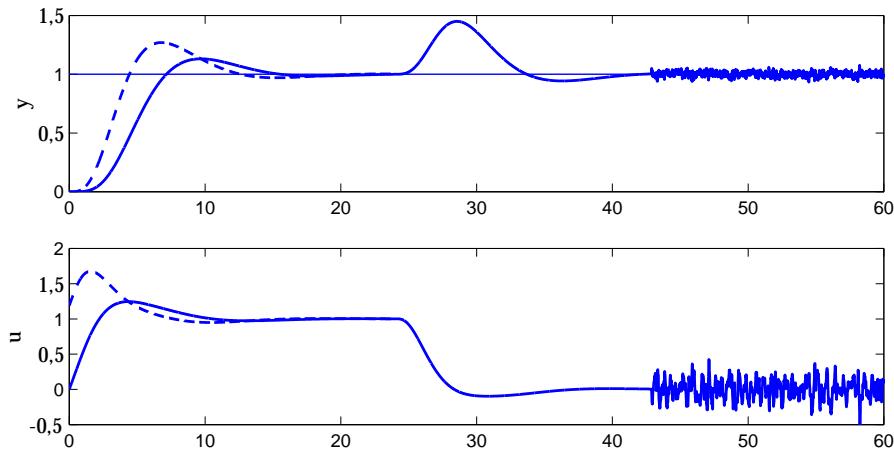


Figura 4.2 Representación de las propiedades de un sistema de realimentación básico mediante respuestas a un escalón en la referencia, un escalón en la perturbación de carga, y ruido de medida. La línea continua es para el caso $b = 0$ y la línea a tramos para $b = 1$ en la ponderación del punto de consigna.

Una consecuencia práctica

El hecho de que se requieran las seis relaciones para capturar las propiedades del lazo de realimentación básico no se considera a menudo en la literatura, particularmente en los artículos sobre control PID. Para describir el sistema adecuadamente es así necesario mostrar la respuesta de todas las seis funciones de transferencia. Las funciones de transferencia se pueden representar de diferentes formas, por sus respuestas a un escalón o por sus respuestas en frecuencia. La mayoría de los trabajos en control solo muestran la respuesta de la variable del proceso frente a cambios en el punto de consigna. Tal curva da solamente información parcial acerca de la conducta del sistema. Para obtener una representación completa del sistema se deberían dar las seis respuestas, por ejemplo, tal como se muestra en la Figura 4.2. Esta figura muestra las respuestas en la variable del proceso y en la señal de control para un experimento con un cambio en escalón en el punto de consigna seguido por un escalón en la perturbación de carga, y el ruido de medida. Las líneas continuas muestran la respuesta cuando $F = 1$ y las líneas a tramos la respuesta cuando se usa feedforward. La Figura 4.2 así da una caracterización completa de las seis funciones de transferencias dadas en (4.3).

Muchas variaciones

El sistema mostrado en la Figura 4.1 es un problema prototípico. Hay muchas variaciones de este problema. En la Figura 4.1 las perturbaciones de carga actúan sobre la entrada del proceso. En la práctica las perturbaciones pueden aparecer en muchos otros lugares en el sistema. El ruido de medida también actúa en la salida del proceso. Pueden haber también dinámica en el sensor, y la señal medida es a menudo filtrada. Todas estas variaciones se pueden estudiar con modificaciones menores del análisis basado en la Figura 4.1. Como una ilustración investigaremos los efectos de un filtro en el sensor. La Figura 4.3 muestra un diagrama de

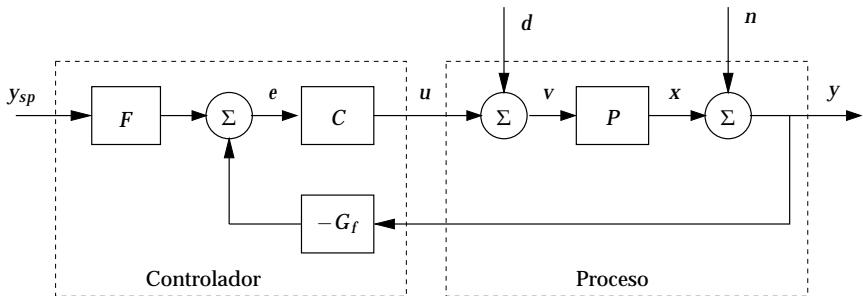


Figura 4.3 Diagrama de bloques de un lazo de realimentación básico que tiene dos grados de libertad y filtraje de la medida.

bloques de tal sistema. Un ejemplo típico es un controlador PID con ponderación del punto de consigna y un filtro de medida de segundo orden. Las funciones de transferencias $F(s)$ y $C(s)$ en la Figura 4.3 están dadas por (4.1) y el filtro de la función de transferencia $G_f(s)$ es

$$G_f(s) = \frac{1}{1 + sT_f + s^2 T_f^2 / 2}. \quad (4.5)$$

Las relaciones entre las señales de entrada y las señales de salida en la Figura 4.3 están dadas por

$$\begin{aligned} X &= \frac{PCF}{1 + PCG_f} Y_{sp} + \frac{P}{1 + PCG_f} D - \frac{PCG_f}{1 + PCG_f} N \\ Y &= \frac{PCF}{1 + PCG_f} Y_{sp} + \frac{P}{1 + PCG_f} D + \frac{1}{1 + PCG_f} N \\ U &= \frac{CF}{1 + PCG_f} Y_{sp} - \frac{PCG_f}{1 + PCG_f} D - \frac{CG_f}{1 + PCG_f} N. \end{aligned} \quad (4.6)$$

La ecuación (4.6) es idéntica a (4.2) si las funciones de transferencia $C(s)$ y $F(s)$ se sustituyen por

$$\bar{C}(s) = C(s) G_f(s), \quad \bar{F}(s) = \frac{F(s)}{G_f(s)}, \quad (4.7)$$

Las modificaciones requeridas para tratar el filtraje son así menores, y es suficiente desarrollar la teoría para la configuración dada en la Figura 4.1.

Separación de las respuestas a perturbaciones y puntos de consigna

En los primeros trabajos sobre control PID era una tradición tener dos reglas de sintonía, una para buena respuesta a cambios en el punto de consigna y otra para atenuación eficiente de las perturbaciones en la carga. Esta práctica todavía continúa. Una gran ventaja de un controlador con dos grados de libertad es que se pueden diseñar separadamente las respuestas a perturbaciones y a cambios en el punto de consigna. Esto se sigue de (4.2), que muestra que la respuesta a las

perturbaciones de carga y al ruido de medida están dadas por $C(s)$, o de (4.6) por $\tilde{C}(s) = C(s)G_f(s)$ cuando se filtra la medida. Un buen procedimiento de diseño es así determinar $C(s)$ para tener en cuenta robustez y perturbaciones. La función de transferencia feedforward $F(s)$ puede entonces seleccionarse para dar la respuesta deseada con respecto a cambios en el punto de consigna. En general, esto requiere que se pueda escoger libremente la función de transferencia feedforward. A menudo simplemente seleccionando los pesos del punto de consigna se obtienen resultados satisfactorios. Obsérvese que hay algunas situaciones donde sólo está disponible la señal de error. No es posible entonces el desacoplo del problema de diseño, y la realimentación tiene así que considerar el compromiso entre perturbaciones, robustez, y respuesta al punto de consigna.

Limitaciones fundamentales

En cualquier problema de diseño es importante conocer las limitaciones fundamentales. Fuentes típicas de limitaciones son

- Dinámica del proceso
- No linealidades
- Perturbaciones
- Incertidumbre del proceso

La dinámica del proceso es, a menudo, un factor limitador. Son factores relevantes los retardos de tiempo y los polos y ceros en el semiplano derecho. Es importante conocer estas limitaciones. Los retardos de tiempo son el factor limitante más común para control PID. Resulta intuitivamente claro que es imposible tener un control exigente de un sistema con un retardo de tiempo. Se puede demostrar que para un proceso con un retardo de tiempo L la ganancia alcanzable en la frecuencia de cruce ω_{gc} , que se define en la Sección 4.4, está limitada por

$$\omega_{gc}L < 1. \quad (4.8)$$

Como

$$e^{-sL} \approx \frac{1 - sL/2}{1 + sL/2},$$

parece también razonable que los ceros en el semiplano derecho también limitan el comportamiento esperable. Se puede demostrar que un cero en el semiplano derecho en $s = b$ limita la ganancia a la frecuencia de cruce a

$$\omega_{gc} < 0.5b. \quad (4.9)$$

Un polo en el semiplano derecho $s = a$ en el proceso limita la ganancia que se puede tener en la frecuencia de cruce ω_{gc} a

$$\omega_{gc} > 2a. \quad (4.10)$$

Obsérvese que retardos de tiempo y ceros en el semiplano derecho dan una cota superior a la ganancia alcanzable a la frecuencia de cruce mientras que los polos en el semiplano derecho dan una cota inferior.

No linealidades del tipo, saturación, y saturación en velocidad son muy comunes; imponen limitaciones sobre cuánto y cuán rápido pueden cambiar las variables del proceso. Las saturaciones combinadas con dinámicas de procesos inestables son particularmente serias porque pueden conducir a situaciones donde no es posible recuperar condiciones de operación estables. Tales situaciones afortunadamente no son comunes en control de procesos.

Las perturbaciones de carga y el ruido de medida limitan con cuánta precisión se puede controlar la variable del proceso. Las limitaciones a menudo interaccionan. La ganancia permitida del controlador está, por ejemplo, limitada por una combinación de ruido de medida y saturación del actuador. El efecto de las perturbaciones de carga depende criticamente del ancho de banda que se alcance.

Los modelos de procesos usados para control son siempre aproximaciones. La dinámica del proceso puede también cambiar durante la operación. La insensibilidad a la incertidumbre del modelo es una de las propiedades esenciales de la realimentación. Hay, sin embargo, un límite a la incertidumbre que se puede tratar. La realimentación no puede ser activa en rangos de frecuencia donde la incertidumbre en la fase del proceso es mayor que $\pm 90^\circ$. Para tener un comportamiento de control razonable la incertidumbre debería ser menor que $\pm 15^\circ$. Si las variaciones del proceso se correlacionan bien con alguna cantidad medida es posible compensar las incertidumbres al cambiar los parámetros del controlador. Esta técnica, que se llama planificación de ganancias, será analizada en la Sección 9.3.

4.4 Estabilidad

La realimentación tiene muchas propiedades útiles. La principal desventaja es que la realimentación puede causar inestabilidad. Es por lo tanto esencial tener una buena comprensión de la estabilidad y el mecanismo que origina la inestabilidad.

Conceptos de estabilidad

La noción de estabilidad es intuitivamente muy sencilla. Dice como se comporta un sistema después de una perturbación. Ya en 1868 Maxwell clasificó la conducta como sigue:

U1: La variable aumenta continuamente

S1: La variable disminuye continuamente

U2: La variable aumenta de forma oscilatoria

S2: La variable disminuye de forma oscilatoria

Estas conductas se ilustran en la Figura 4.4. Maxwell llamó a las conductas etiquetadas S como estables y a las etiquetadas U como inestables. También encontró

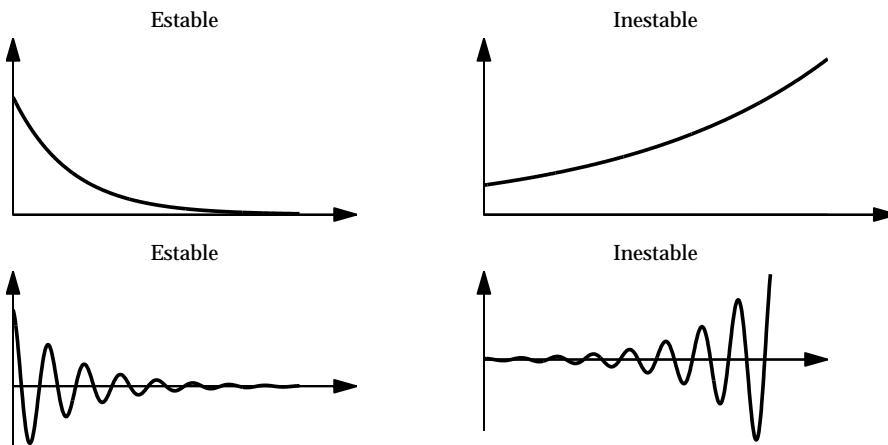


Figura 4.4 Ilustración de diferentes conductas de sistemas para definir estabilidad.

que para sistemas lineales e invariantes en el tiempo la estabilidad estaba relacionada con las raíces de una ecuación algebraica.

Sea un sistema con la función de transferencia

$$G(s) = \frac{b(s)}{a(s)}, \quad (4.11)$$

donde $a(s)$ y $b(s)$ son polinomios. Recuerde que las raíces del polinomio $a(s)$ se llaman los polos del sistema. Como un polo s_i corresponde a una función del tiempo $e^{s_i t}$ se obtienen las relaciones siguientes entre las conductas y las raíces de una ecuación algebraica:

- U1: Corresponde a polos reales con parte real positiva
- S1: Corresponde a polos reales con parte real negativa
- U2: Corresponde a polos complejos con parte real positiva
- S2: Corresponde a polos complejos con parte real negativa

El sistema (4.11) es estable si no tiene polos en el semiplano derecho. La ecuación

$$a(s) = 0 \quad (4.12)$$

se llama la ecuación característica. Un sistema es estable si la ecuación característica no tiene raíces con partes reales positivas. Es práctica común etiquetar los polos sobre el eje imaginario como inestables.

Criterio de estabilidad de Nyquist

La definición algebraica de estabilidad basada en las raíces de la ecuación característica es útil, pero también tiene algunas desventajas. Considérese, por ejemplo, el sistema de realimentación de la Figura 4.5 donde las funciones de transferencias del proceso y del controlador han sido combinadas en un bloque con la función de transferencia $G_I = PC$. La ecuación característica de este sistema es

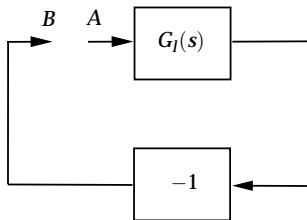


Figura 4.5 Diagrama de bloques de un sistema sencillo de realimentación.

$$1 + G_l(s) = 0. \quad (4.13)$$

La función de transferencia, que es el producto de las funciones de transferencias del proceso y del controlador, describe como se propagan las señales alrededor del lazo de realimentación y se llama la *función de transferencia del lazo*. No es fácil ver como las raíces de (4.13) están influenciadas por las funciones de transferencia del proceso y del controlador. Sin embargo, esto puede hacerse usando una visión totalmente diferente de la estabilidad, que fue desarrollada por Nyquist quien comenzó por investigar las condiciones para mantener una oscilación en el sistema mostrado en la Figura 4.5. Suponga que se rompe el lazo de realimentación como se indica en la figura y que se inyecta la señal $u_A(t) = \sin \omega_0 t$ en el punto A. Después de un transitorio la salida en el punto B viene entonces dada por

$$u_B(t) = -|G_l(i\omega_0)| \sin(\omega_0 t + \arg G_l(i\omega)).$$

Las señales $u_A(t)$ y $u_B(t)$ son idénticas si

$$G_l(i\omega_0) = -1, \quad (4.14)$$

y se mantendrá una oscilación si se cierra el lazo uniendo los puntos A y B. La ecuación 4.14 da así la condición para oscilaciones en el sistema. Se sigue de la (4.13) y (4.14) que la condición para oscilación implica que la ecuación característica del sistema tiene una raíz $s = i\omega_0$. Las frecuencias donde el sistema puede mantener una oscilación se pueden determinar resolviendo (4.14) para ω_0 .

Nyquist desarrolló un criterio de estabilidad basado en la idea de como se propagan las sinusoides alrededor del lazo de realimentación. Nyquist argumenta como sigue. En primer lugar investigó las frecuencias donde las señales u_A y u_B están en fase, i.e., cuando $\arg G_l(i\omega_0) = \pi$. Intuitivamente parece razonable que el sistema es estable si $|G_l(i\omega_0)| < 1$ porque la amplitud se reduce entonces cuando la señal recorre el lazo. La situación es realmente un poco más complicada porque el sistema puede ser estable incluso si $|G_l(i\omega_0)| > 1$. El resultado preciso se puede expresar en términos de la curva de Nyquist introducida en la Section 2.3. Recuerde que la curva de Nyquist es una gráfica de $(\text{Re } G_l(i\omega), \text{Im } G_l(i\omega))$ para $0 \leq \omega \leq \infty$. Cuando la función de transferencia del lazo no tiene polos en el semiplano derecho la condición para estabilidad es que el punto crítico -1 esté a la izquierda de la curva de Nyquist cuando se recorre para ω crecientes.

Una propiedad agradable del criterio de Nyquist es que indica como debería cambiarse un sistema con el fin de mover la curva de Nyquist fuera del punto

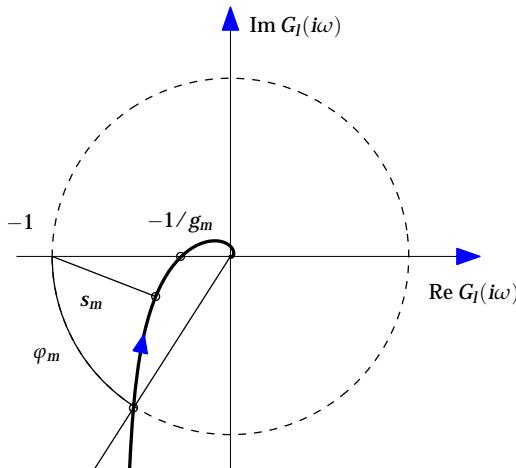


Figura 4.6 Diagrama de Nyquist de la función de transferencia del lazo G_l con margen de ganancia g_m , margen de fase φ_m y margen de estabilidad s_m .

crítico. La Figura 6.4 muestra que la acción derivativa, que introduce adelanto de fase, moldea a la curva fuera del punto crítico. La acción integral, que introduce retardo de fase, mueve a la curva hacia el punto crítico. La idea es modificar el controlador de forma que el diagrama se curve fuera del punto crítico. Esto ha conducido a una clase completa de métodos de diseño llamados de deformación del lazo.

Márgenes de estabilidad

En la práctica no es bastante requerir que el sistema sea estable. Debe de haber también algunos márgenes de estabilidad. Esto significa que la curva de Nyquist no debería estar demasiado próxima al punto crítico. Esto se ilustra en la Figura 4.6, que muestra algunos márgenes de estabilidad. El margen de ganancia g_m dice cuanto se puede aumentar la ganancia del controlador antes de alcanzar el límite de estabilidad. Sea la *frecuencia de cruce de fase* ω_{180} la frecuencia más pequeña donde el desfase de la función de transferencia de lazo $G_l(s)$ es 180° y el margen de ganancia se define como

$$g_m = \frac{1}{|G_l(i\omega_{180})|}. \quad (4.15)$$

El punto donde la curva de Nyquist intersecta al círculo unidad es otro punto interesante. Este punto se puede caracterizar por el ángulo φ_m . Este ángulo llamado *el margen de fase* es también una medida de cuán próximo al punto crítico está la curva de Nyquist. El ángulo φ_m es la cantidad de fase requerida para alcanzar el límite de estabilidad. La *frecuencia de cruce de ganancia* ω_{gc} es la frecuencia más baja donde la función de transferencia del lazo $G_l(s)$ tiene magnitud unidad. El margen de fase se define formalmente como

$$\varphi_m = \pi + \arg G_l(i\omega_{gc}). \quad (4.16)$$

Ambos márgenes de ganancia y fase son medidas clásicas del grado de estabilidad. Ambos valores deben especificarse con el fin de asegurar que la curva de Nyquist está lejos del punto crítico. Pueden sustituirse ambos por un único número, la distancia más corta de la curva de Nyquist al punto crítico -1 , que se llama el *márgen de estabilidad* s_m .

Valores razonables de los márgenes son: margen de fase $\varphi_m = 30^\circ - 60^\circ$, margen de ganancia $g_m = 2 - 5$, margen de estabilidad $s_m = 0,5 - 0,8$.

Los márgenes de ganancia y fase fueron originalmente concebidos para el caso cuando la curva de Nyquist intersecta solamente una vez al círculo unidad y al eje real negativo. Para sistemas más complicados pueden haber muchas intersecciones, y es entonces necesario considerar las intersecciones que están más próximas al punto crítico. Para sistemas más complicados hay también otro número que es altamente relevante, *el margen de retardo*. El margen de retardo se define como el retardo de tiempo más pequeño requerido para hacer al sistema inestable. Para funciones de transferencia de lazo que decaen rápidamente el margen de retardo está estrechamente relacionado con el margen de fase, pero para sistemas donde la magnitud de la función de transferencia del lazo tiene algunos picos de altas frecuencias el margen de retardo es una medida mucho más fiable. Esto es particularmente importante para el predictor de Smith que se analizará en el Capítulo 8.

Estabilidad interna

Hasta ahora sólo hemos analizado el sistema de realimentación sencillo en la Figura 4.5. Para el sistema más general en la Figura 4.1 que se caracteriza por seis funciones de transferencia, es necesario requerir que todas las cuatro funciones de transferencia,

$$\begin{array}{c} \frac{PC}{1+PC} \quad \frac{P}{1+PC} \\ \frac{C}{1+PC} \quad \frac{1}{1+PC} \\ \frac{1}{1+PC} \quad \frac{1}{1+PC'} \end{array} \quad (4.17)$$

sean estables; comparar con (4.3). Esto se llama estabilidad interna. Obsérvese que pueden haber cancelaciones de polos y ceros en el producto PC .

Regiones de estabilidad

Un requisito primario para un controlador PID es que los parámetros del controlador se escogen de tal forma que el sistema en lazo cerrado es estable. Un controlador PID de la forma

$$C(s) = k + \frac{k_i}{s} + k_d s \quad (4.18)$$

tiene sólo tres parámetros, y la región de estabilidad se puede representar por un volumen en tres dimensiones. Para describir este volumen la función de transferencia del proceso se representa como

$$P(i\omega) = r(\omega)e^{i\phi(\omega)} = r(\omega)(\cos(\omega) + i\sin(\omega)),$$

y la condición para oscilación (4.14) entonces se hace

$$P(i\omega)C(i\omega) = r(\omega)(\cos(\omega) + i\sin(\omega))\left(k - i\frac{k_i}{\omega} + ik_d\omega\right) = -1.$$

Separando las partes real e imaginaria encontramos que las fronteras de la región de estabilidad se pueden representar paramétricamente como

$$\begin{aligned} k &= -\frac{\cos \phi(\omega)}{r(\omega)} \\ k_i &= \omega^2 k_d - \frac{\omega \sin \phi(\omega)}{r(\omega)}. \end{aligned} \quad (4.19)$$

Es así directo determinar la región de estabilidad para un valor constante de k_d . Repitiendo los cálculos para un conjunto de k_d -valores se obtiene la región de estabilidad para el controlador PID.

EJEMPLO 4.1—REGIÓN DE ESTABILIDAD PARA $P(s) = 1/(s+1)^4$

La Figura 4.7 muestra la región de estabilidad para un proceso con la función de transferencia $P(s) = 1/(s+1)^4$. El valor $k_d = 0$ corresponde a control PI. La ganancia integral k_i se puede aumentar añadiendo la acción derivativa. La ganancia integral tiene su máximo $k_i = 36$ en la frontera de la región de estabilidad para $k = 8$ y $k_d = 20$. El sistema es inestable para todos los valores de k y k_i si $k_d > 20$.

□

Se pueden sacar algunas conclusiones interesantes del Ejemplo 4.1. Para tener un buen rechazo de las perturbaciones es deseable tener un gran valor de k_i . Esto se muestra en la Sección 4.9. Con control PI, el valor más grande de k_i para un sistema estable es $k_i = 1$. La Figura 4.7 muestra que el valor de k_i se puede aumentar sustancialmente introduciendo la acción derivativa. El valor más alto de k_i que se puede obtener con un sistema estable es $k_i = 36$. Esto será, sin embargo, un controlador muy frágil porque el sistema se puede hacer inestable mediante cambios arbitrariamente pequeños en las ganancias del controlador. Para grandes valores de k_d las curvas tienen extremos afilados en los puntos de máxima ganancia integral. Esta propiedad de la acción derivativa es una razón de por qué la sintonía de controladores con acción derivativa es difícil. Esto se analizará aún más en los Capítulos 6 y 7.

Ganancia proporcional constante

La región de parámetros donde el sistema es estable es un subconjunto de R^3 . Los cálculos efectuados dan las intersecciones bi-dimensional con ganancia derivativa constante. Se pueden obtener propiedades adicionales de otra representación de la región de estabilidad. Para investigar la estabilidad emplearemos el criterio de Nyquist y la gráfica de la función de transferencia del lazo $G_l(s)$. Con control proporcional tenemos $G_l = kP$. Para un valor fijo de la ganancia proporcional $k > 0$ determinamos la frecuencia ω_n donde el diagrama de Nyquist de $kP(i\omega)$ intersecta el círculo con el segmento de línea $(-1, 0)$ como un diámetro; ver Figura 4.8. Primero consideraremos el caso cuando ocurre la intersección del

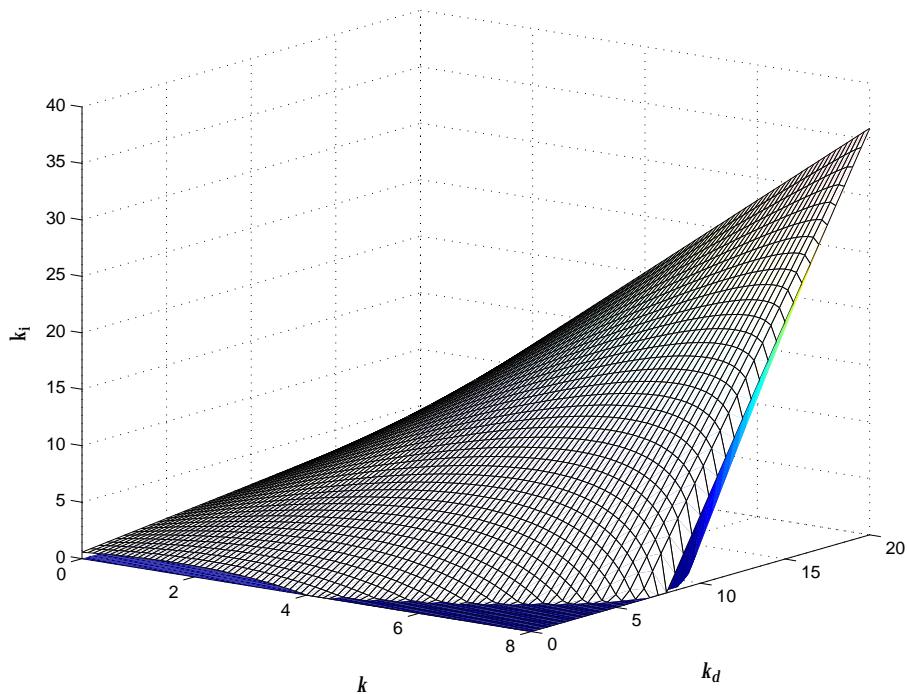


Figura 4.7 Región de estabilidad para el sistema $P(s) = (1 + s)^{-4}$.

diagrama de Nyquist y el círculo en el semiplano inferior tal como se muestra en la Figura 4.8. La función de transferencia del controlador es

$$C(i\omega) = k + i\left(-\frac{k_i}{\omega} + k_d\omega\right) = k - i\left(\frac{k_i}{\omega} - k_d\omega\right),$$

de donde,

$$G_l(i\omega_n) = P(i\omega_n)C(i\omega_n) = kP(i\omega_n) - i\left(\frac{k_i}{\omega_n} - k_d\omega_n\right)P(i\omega_n).$$

Si se fija la ganancia proporcional k el punto $kP(i\omega_n)$ se mueve a $G_l(i\omega_n)$ cuando las ganancias proporcional e integral son diferentes de cero. Para evitar alcanzar el punto crítico se debe requerir que

$$\left(\frac{k_i}{\omega_n} - k_d\omega_n\right)|P(i\omega_n)| < |1 + kP(i\omega_n)|.$$

El mismo análisis se puede hacer cuando ocurre la intersección del diagrama de Nyquist y el círculo en el semiplano superior. Combinando las inigualdades en-

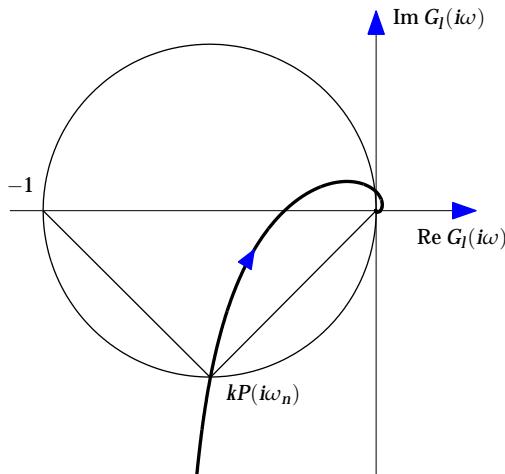


Figura 4.8 La curva de Nyquist de la función de transferencia del lazo $G_l(s) = kP(s)$.

contramos que las regiones de estabilidad están dadas por las condiciones

$$\begin{aligned} k_i &> 0 \\ k_i &< \omega_n^2 k_d + \omega_n \frac{|1 + kP(i\omega_n)|}{|P(i\omega_n)|}, \text{ para } \operatorname{Im} P(i\omega_n) < 0 \\ k_i &> \omega_n^2 k_d - \omega_n \frac{|1 + kP(i\omega_n)|}{|P(i\omega_n)|}, \text{ para } \operatorname{Im} P(i\omega_n) > 0 \end{aligned} \quad (4.20)$$

que debería mantenerse para todo ω_n tal que

$$\left| kP(i\omega_n) + \frac{1}{2} \right| = \frac{1}{2}. \quad (4.21)$$

Podemos así concluir que para ganancia proporcional constante la región de estabilidad se representa por algunos polígonos convexos en el plano k_i - k_d . En general, pueden haber algunos polígonos, y cada uno puede tener muchas superficies. El número de superficies de los polígonos se determina por el número de raíces de la Ecuación 4.21. En muchos casos, los polígonos son también muy simples, tal como se ilustra con el ejemplo siguiente.

EJEMPLO 4.2—CUATRO POLOS IGUALES

Para ilustrar los resultados consideraremos un proceso con la función de transferencia

$$P(s) = \frac{1}{(s+1)^4} = \frac{1}{s^4 + 6s^2 + 1 + 4s(s^2 + 1)}.$$

En este caso, (4.21) se hace

$$\omega^4 - 6\omega^2 + 1 + k = 0.$$

Esta ecuación tiene solamente dos soluciones positivas,

$$\omega^2 = 3 \pm \sqrt{8 - k},$$

y se sigue de (4.20) que la región de estabilidad está dada por las inigualdades

$$\begin{aligned} k_i &> 0 \\ k_i &< (3 - \sqrt{8 - k})k_d + 4k - 56 + 20\sqrt{8 - k} \\ k_i &> (3 + \sqrt{8 - k})k_d + 4k - 56 - 20\sqrt{8 - k}. \end{aligned} \quad (4.22)$$

En la Figura 4.7 se muestra la región de estabilidad. La ganancia integral tiene su máximo $k_i = 36$ en la frontera de la región de estabilidad para $k = 8$ y $k_d = 20$.

□

4.5 Polos y ceros en lazo cerrado

Se pueden obtener muchas propiedades de los sistemas con realimentación de los polos y ceros en lazo cerrado. Para control PID la conducta se caracteriza a menudo por unos pocos polos dominantes, típicamente por aquellos más próximos al origen. Es posible deducir también muchas propiedades del sistema en lazo cerrado de los polos y ceros de la función de sensibilidad complementaria

$$T(s) = \frac{PC(s)}{1 + PC(s)}.$$

Con realimentación del error, $F = 1$ en la Figura 4.1, los ceros en lazo cerrado son los mismos que los ceros de la función de transferencia del lazo $G_l(s)$, y los polos en lazo cerrado son las raíces de la ecuación

$$1 + G_l(s) = 0.$$

Las configuraciones polos-zeros del sistema en lazo cerrado pueden variar considerablemente. Muchos lazos de realimentación sencillos, sin embargo, tendrán una configuración del tipo mostrado en la Figura 4.9, donde las características principales de la respuesta están dadas por un par de polos complejos, p_1 y p_2 , llamados los *polos dominantes*. La respuesta está también influenciada por los polos y ceros reales p_3 y z_1 próximos al origen. Se puede invertir la posición de p_3 y z_1 . Pueden haber también más polos y ceros lejos del origen, que típicamente son de menor influencia. Los polos y ceros a la izquierda de los polos dominantes tienen poca influencia sobre la respuesta transitoria si están suficientemente lejos de los polos dominantes. La influencia de un polo disminuye si tiene un cero próximo.

Los polos complejos se pueden caracterizar en términos de su frecuencia ω_0 , que es la distancia al origen, y su amortiguamiento relativo ζ . Se obtiene una

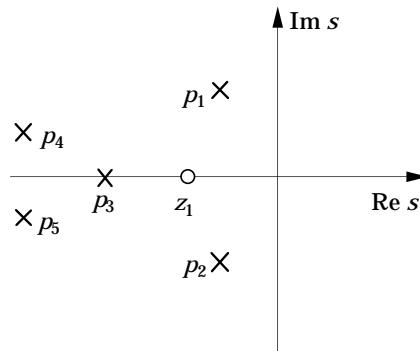


Figura 4.9 Configuración polo-cero de la función de transferencia del punto de consigna a la salida para un sistema de realimentación sencillo.

primera aproximación de la respuesta del sistema de segundo orden equivalente. La respuesta se modifica si hay polos y ceros próximos a los polos dominantes. El control clásico se centró mucho con sistemas en lazo cerrado con la configuración polos-ceros que se muestra en la Figura 4.9.

Aunque muchos sistemas en lazo cerrado tienen una configuración de polos-ceros similar a la que se muestra en la Figura 4.9, hay, no obstante, excepciones. Por ejemplo, sistemas con resonancias mecánicas, que pueden tener polos y ceros próximos al eje imaginario, son ejemplos genéricos de sistemas que no se ajustan a la estructura de polos-ceros de la figura. Otro ejemplo es el de procesos con un gran retardo de tiempo.

El diseño de controladores PID se basa típicamente en modelos de bajo orden, que dan sistemas en lazo cerrado con un número pequeño de polos y ceros.

Polos dominantes de la función de transferencia del lazo

Se expondrá a continuación un método sencillo para aproximar la determinación de los polos dominantes a partir del conocimiento de la curva de Nyquist de la función de transferencia del lazo. Considérese la función de trasnferencia del lazo $G_l(s)$ como una transformación del plano-s al plano- G_l . La transformación del eje imaginario en el plano-s es la curva de Nyquist $G_l(i\omega)$, que se indica en la Figura 4.10. Los polos en lazo cerrado son las raíces de la ecuación característica

$$1 + G_l(s) = 0.$$

La transformación de una línea recta vertical a través de los polos en lazo cerrado dominantes en el plano-s es así una curva a través del punto crítico $G_l = -1$ en el plano- G_l . Esta curva se muestra en la Figura 4.10. Como la transformación es conforme, la línea recta A'C' se transforma en la curva AC, que intersecta ortogonalmente a la curva de Nyquist. El triángulo ABC también se transforma de manera conforme a A'B'C'. Si se puede aproximar ABC por un triángulo, tenemos

$$\frac{G_l(i\omega_2) - G_l(i\omega_1)}{i\omega_2 - i\omega_1} \approx \frac{1 + G_l(i\omega_2)}{\sigma}.$$

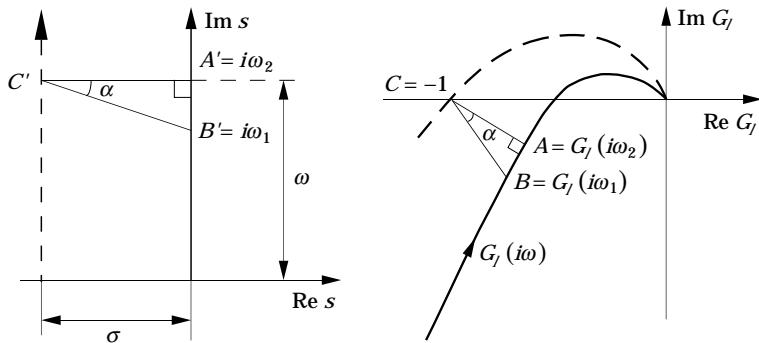


Figura 4.10 Representación de la función de transferencia del lazo $G_l(i\omega)$ como una transformación de planos.

Cuando ω_1 está próximo a ω_2 esto se hace

$$\sigma = (1 + G_l(i\omega_2)) \frac{i\omega_2 - i\omega_1}{G_l(i\omega_2) - G_l(i\omega_1)} \approx \frac{1 + G_l(i\omega_2)}{G'_l(i\omega_2)}, \quad (4.23)$$

donde $G'_l(s) = dG_l(s)/ds$. Para determinar los polos dominantes en primer lugar se calcula el punto A sobre la curva de Nyquist que está próximo al punto crítico -1 . Este punto se caracteriza por la frecuencia ω_2 . A continuación se determina la derivada de la función de transferencia del lazo ω_2 . Los polos dominantes se dan entonces por $s = -\sigma \pm i\omega_2$, donde σ se obtiene de la Ecuación 4.23.

4.6 Las funciones de sensibilidad

Dos de las funciones de transferencia (4.3) son de interés particular, la función de sensibilidad S y la función de sensibilidad complementaria T . Estas funciones se definen por

$$S = \frac{1}{1 + PC} = \frac{1}{1 + G_l}, \quad T = \frac{PC}{1 + PC} = \frac{G_l}{1 + G_l}. \quad (4.24)$$

Las funciones de sensibilidad están únicamente dadas por la función de transferencia del lazo $G_l(s) = P(s)C(s)$ y tienen la propiedad de que $S + T = 1$. Las funciones de transferencias reflejan muchas propiedades interesantes del sistema en lazo cerrado, particularmente la robustez frente a variaciones del proceso.

Pequeñas variaciones del proceso – La función de sensibilidad

Comenzaremos por investigar cuán sensible es la respuesta a cambios en el punto de consigna a pequeñas variaciones del proceso. Se sigue de (4.2) que la función de transferencia desde el punto de consigna a la variable del proceso es

$$G_{xy_{sp}} = G_{yy_{sp}} = \frac{PCF}{1 + PC}.$$

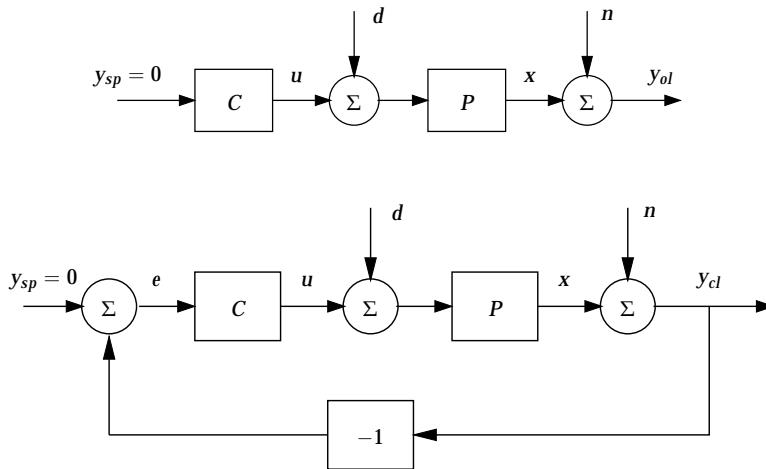


Figura 4.11 Diagramas de bloque de sistemas en lazo abierto y cerrado sujetos a las mismas perturbaciones.

Considérese $G_{xy_{sp}}$ como una función de la función de transferencia del proceso P . Diferenciando con respecto a P se deduce

$$\frac{dG_{xy_{sp}}}{dP} = \frac{CF}{1+PC} - \frac{PC^2F}{(1+PC)^2} = \frac{CF}{(1+PC)^2} = \frac{1}{1+PC} \frac{CF}{1+PC}.$$

De donde,

$$\frac{dG_{xy_{sp}}}{G_{xy_{sp}}} = \frac{1}{1+PC} \frac{dP}{P} = S \frac{dP}{P}. \quad (4.25)$$

Obsérvese que la cantidad dG/G se puede interpretar como la variación relativa en la función de transferencia G . La ecuación 4.25 implica así que el error relativo en la función de transferencia en lazo cerrado $G_{yy_{sp}}$ es igual al producto de la función de sensibilidad por el error relativo en el proceso. Para frecuencias donde la función de sensibilidad es pequeña se sigue así que el sistema en lazo cerrado es muy insensible a las variaciones en el proceso. Éste es realmente uno de los motivos clave para usar realimentación. La fórmula (4.25) es una de las razones de por qué S se llama la función de sensibilidad. La función de sensibilidad tiene también otras propiedades interesantes.

Atenuación de la perturbación

Una cuestión fundamental es en cuánto se ven influenciadas por la realimentación las fluctuaciones en la variable del proceso. Considérese la situación mostrada en la Figura 4.11 donde la misma perturbación de carga actúa sobre un proceso P en lazo abierto y sobre el proceso P en lazo cerrado con el controlador C . Sea y_{ol} la salida del sistema en lazo abierto e y_{cl} la salida del sistema en lazo cerrado. Tenemos la siguiente relación entre las transformadas de Laplace de las señales,

$$\frac{Y_{cl}(s)}{Y_{ol}(s)} = \frac{1}{1 + P(s)C(s)} = S(s). \quad (4.26)$$

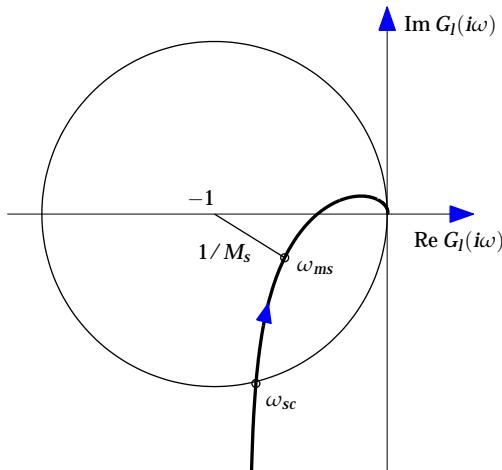


Figura 4.12 Diagrama de Nyquist de la función de transferencia del lazo que muestra la interpretación gráfica de la sensibilidad máxima. En la figura se indica la frecuencia de cruce de sensibilidad ω_{sc} , y la frecuencia ω_{ms} donde la sensibilidad tiene su valor máximo. Todos los puntos dentro del círculo de centro en -1 tienen sensibilidades mayores que 1.

Perturbaciones con frecuencias ω tal que $|S(i\omega)| < 1$ se atenúan así por realimentación, pero perturbaciones tales que $|S(i\omega)| > 1$ se amplifican por la realimentación. Un diagrama de la magnitud de S nos dice inmediatamente el efecto de la realimentación.

Como la sensibilidad sólo depende de la función de transferencia del lazo se puede visualizar gráficamente en el diagrama de Nyquist de dicha función de transferencia. Esto se ilustra en la Figura 4.12. El número complejo $1 + G_l(i\omega)$ se puede representar en el diagrama de Nyquist como el vector que va desde el punto -1 al punto $G_l(i\omega)$. La sensibilidad es así menor que uno para todos los puntos fuera del círculo de radio 1 y centro en -1 . Las perturbaciones de estas frecuencias se atenúan por la realimentación. La frecuencia más baja donde la función de sensibilidad tiene magnitud 1 se llama la *frecuencia de cruce de sensibilidad* ω_{sc} . El valor

$$M_s = \max_{\omega} |S(i\omega)| = \max_{\omega} \left| \frac{1}{1 + P(i\omega)C(i\omega)} \right| = \max_{\omega} \left| \frac{1}{1 + G_l(i\omega)} \right|, \quad (4.27)$$

que se llama la sensibilidad máxima, da el peor caso de amplificación de las perturbaciones.

La sensibilidad no se puede hacer arbitrariamente pequeña. La siguiente relación se mantiene bajo condiciones razonablemente generales para sistemas estables

$$\int_0^{\infty} \log |S(i\omega)| d\omega = 0. \quad (4.28)$$

Esta relación tan importante se llama integral de Bode. Dice que si la sensibilidad se reduce para una frecuencia aumenta en otra frecuencia. La realimentación

puede así redistribuir la atenuación de las perturbaciones para diferentes frecuencias, pero no puede reducir el efecto de las perturbaciones para todas las frecuencias.

En la Sección 2.6 se mencionó que las fluctuaciones aleatorias se pueden modelar por una densidad espectral de potencia. Si la densidad espectral es $\phi(\omega)$ para un sistema sin control se convierte en $|S(i\omega)|^2\phi(\omega)$ para un sistema con control. Las razones de las varianzas bajo lazo abierto y cerrado son así

$$\frac{\sigma_{cl}^2}{\sigma_{ol}^2} = \frac{\int_{-\infty}^{\infty} |S(i\omega)|^2 \phi(\omega) d\omega}{\int_{-\infty}^{\infty} \phi(\omega) d\omega}. \quad (4.29)$$

Márgenes de estabilidad y máxima sensibilidad

Obsérvese que $|1 + G_l(i\omega)|$ es la distancia desde un punto del diagrama de Nyquist de la función de transferencia del lazo al punto -1 . Ver la Figura 4.12. La distancia más corta desde el diagrama de Nyquist de la función de transferencia del lazo al punto crítico -1 es así $1/M_s$, que es igual al margen de estabilidad s_m . Compárense las Figuras 4.12 y 4.6. La máxima sensibilidad puede así servir también como un margen de estabilidad. Un requisito sobre M_s da las siguientes acotaciones para los márgenes de ganancia y de fase

$$g_m \geq \frac{M_s}{M_s - 1}$$

$$\varphi_m \geq 2 \arcsin \left(\frac{1}{2M_s} \right).$$

El requisito $M_s = 2$ implica que $g_m \geq 2$ y $\varphi_m \geq 29^\circ$ y $M_s = 1,4$ implica que $g_m \geq 3,5$ y $\varphi_m \geq 41^\circ$.

No linealidades en el lazo

La condición de que el diagrama de Nyquist de la función de transferencia del lazo esté fuera de un círculo en el punto crítico con radio $1/M_s$ tiene fuertes implicaciones. Se sigue del criterio de estabilidad de Nyquist que el sistema permanece estable incluso si la ganancia se aumenta por el factor $M_s/(M_s - 1)$ o si se disminuye por el factor $M_s/(M_s + 1)$. Más sorprendente es que el lazo cerrado es estable incluso si se inserta una no linealidad estática f en el lazo, a condición de que

$$\frac{M_s}{M_s + 1} < \frac{f(x)}{x} < \frac{M_s}{M_s - 1}. \quad (4.30)$$

Un pequeño valor de M_s así asegura que el sistema permanecerá estable a pesar de la característica no lineal del actuador. Con $M_s = 2$ la función permanece en un sector limitado por líneas rectas a través del origen con pendientes $2/3$ y 2 . Con $M_s = 1,4$ las pendientes están entre 0.28 y 3.5 .

Grandes variaciones

Ahora investigaremos las condiciones para que el sistema permanezca estable cuando hay grandes variaciones en la función de transferencia del proceso. Supongamos que la función de transferencia del proceso cambia de P a $P + \Delta P$, donde ΔP es

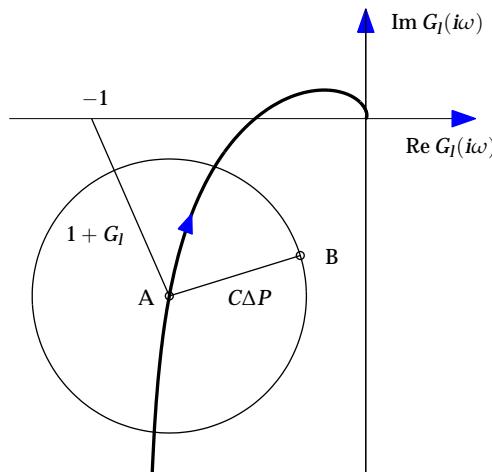


Figura 4.13 Diagrama de Nyquist de una función de transferencia del lazo nominal y su incertidumbre originada por variaciones del proceso ΔP .

una función de transferencia estable. Considérese un punto A sobre el diagrama de Nyquist de la función de transferencia del lazo; ver la Figura 4.13. Este punto entonces cambia de A a B en la figura. La distancia desde el punto crítico -1 al punto A es $|1 + G_l|$. Esto significa que el diagrama de Nyquist perturbado no alcanzará el punto crítico -1 a condición de que

$$|C\Delta P| < |1 + G_l|,$$

que implica que

$$|\Delta P| < \left| \frac{1 + G_l}{C} \right|. \quad (4.31)$$

Obsérvese que esta condición es conservadora porque tal como se sigue de la Figura 4.13 se pueden hacer cambios mucho más grandes en direcciones del punto crítico. La condición (4.31) debe ser válida para todos los puntos del diagrama de Nyquist, i.e., punto a punto para todas las frecuencias. La condición (4.31) para estabilidad se puede entonces describir como

$$\left| \frac{\Delta P(i\omega)}{P(i\omega)} \right| < \frac{1}{|T(i\omega)|}, \quad (4.32)$$

donde T es la función de transferencia complementaria. La inigualdad (4.32) indica que se permiten grandes perturbaciones relativas mientras T sea pequeño. Una estimación conservadora sencilla del error relativo permisible en la función de transferencia del proceso es $1/M_t$ donde

$$M_t = \max_{\omega} |T(i\omega)| = \max_{\omega} \left| \frac{P(i\omega)C(i\omega)}{1 + P(i\omega)C(i\omega)} \right| = \max_{\omega} \left| \frac{G_l(i\omega)}{1 + G_l(i\omega)} \right|, \quad (4.33)$$

es la magnitud más grande de $|T|$. Obsérvese que M_t es también la ganancia más grande de la función de transferencia desde el punto de consigna a la salida para un sistema con realimentación del error.

La ecuación 4.32 puede también escribirse como

$$|\Delta P(i\omega)| < \frac{|P(i\omega)|}{|T(i\omega)|}. \quad (4.34)$$

Se sigue de esta ecuación que la magnitud del error permisible $|\Delta P(i\omega)|$ es pequeña cuando $|P(i\omega)|$ es menor que $|T(i\omega)|$. Se requiere así un modelo de alta precisión para frecuencias donde la ganancia del sistema en lazo cerrado es mayor que la ganancia del sistema en lazo abierto.

Interpretación gráfica de las restricciones sobre las sensibilidades

Los requisitos de que las sensibilidades sean menores que valores dados tienen interesantes interpretaciones geométricas en el diagrama de Nyquist. Como la sensibilidad se define por

$$S(i\omega) = \frac{1}{1 + G_l(i\omega)},$$

se sigue que la sensibilidad tiene magnitud constante en círculos con centro en el punto crítico -1 . La condición de que la sensibilidad más grande es menor que M_s es equivalente a la condición de que el diagrama de Nyquist de la función de transferencia del lazo está fuera de un círculo con centro en -1 y radio $1/M_s$.

Hay una interpretación similar de la sensibilidad complementaria

$$T = \frac{G_l(i\omega)}{1 + G_l(i\omega)}.$$

Introduciendo

$$G_l(i\omega) = \operatorname{Re} G_l(i\omega) + i\operatorname{Im} G_l(i\omega) = x + iy,$$

encontramos que la magnitud de T está dada por

$$|T| = \frac{\sqrt{x^2 + y^2}}{\sqrt{(1+x)^2 + y^2}}.$$

la magnitud de la función de sensibilidad complementaria es constante si

$$x^2 + y^2 = M_t^2((1+x)^2 + y^2) = M_t^2(1 + 2x + x^2 + y^2),$$

o

$$x^2 \frac{M_t^2 - 1}{M_t^2} + 2x + y^2 \frac{M_t^2 - 1}{M_t^2} + 1 = 0.$$

Esta condición se puede escribir como

$$\begin{aligned} x^2 + 2 \frac{M_t^2}{M_t^2 - 1} x + y^2 + \frac{M_t^2}{M_t^2 - 1} \\ = \left(x + \frac{M_t^2}{M_t^2 - 1}\right)^2 + y^2 + \frac{M_t^2}{M_t^2 - 1} - \left(\frac{M_t^2}{M_t^2 - 1}\right)^2 \\ = \left(x + \frac{M_t^2}{M_t^2 - 1}\right)^2 + y^2 - \frac{M_t^2}{(M_t^2 - 1)^2} = 0. \end{aligned}$$

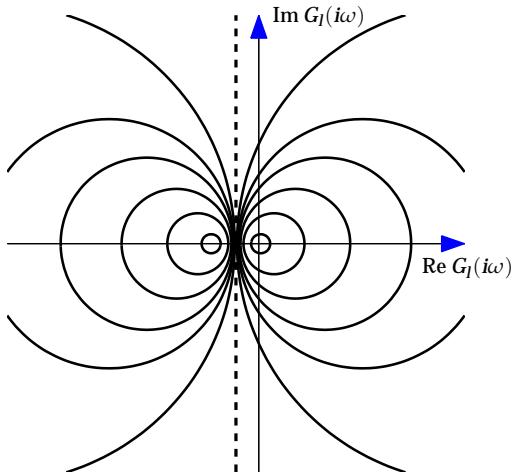


Figura 4.14 Lugar donde la función de sensibilidad complementaria tiene magnitud constante. Las líneas muestran puntos donde la magnitud de la función de sensibilidad es $M_t = 1,1, 1,2, 1,4, 1,5, 2,$ y 5 y los inversos de estos valores. La línea a tramos corresponde a $M_t = 1.$

Éste es un círculo con centro en $x = -M_t^2/(M_t^2 - 1)$ y $y = 0$, y con radio $r = M_t/(M_t^2 - 1).$ Para $M_t = 1$ el círculo degenera en una línea recta con $x = -0,5.$ El requisito de que la función de sensibilidad complementaria sea menor que M_t implica así que el diagrama de Nyquist está fuera del círculo correspondiente. En la Figura 4.14 se muestra el lugar de ganancia constante de la función de sensibilidad complementaria $G_I.$ Obsérvese que los círculos encierran al punto crítico $-1.$ Se deduce también que la función de transferencia en lazo cerrado es insensible a variaciones en frecuencias donde la función de transferencia del lazo está lejos del origen, particularmente si el diagrama de Nyquist está próximo a la línea recta $\text{Re } G_I(i\omega) = -0,5.$ Esto implica que controladores con la propiedad

$$T_i \approx T_{ar} \frac{2KK_p}{1 + 2KK_p} \quad (4.35)$$

son muy robustos. Compárese con la Sección 6.3.

Sensibilidades combinadas

Los requisitos de que la sensibilidad máxima sea menor que M_s y que la sensibilidad complementaria sea menor que M_t implican que el diagrama de Nyquist debería estar fuera de los círculos correspondientes. Es posible encontrar una condición ligeramente más conservadora determinando un círculo que encierre ambos círculos como se ilustra en la Figura 4.15. En la Tabla 4.1 se dan los radios y los centros de los círculos. En esa tabla hemos dado también los círculos que garantizan que ambos M_t y M_s son más pequeños que los valores especificados. Un criterio particular sencillo se obtiene si se requiere que $M_s = M_t.$

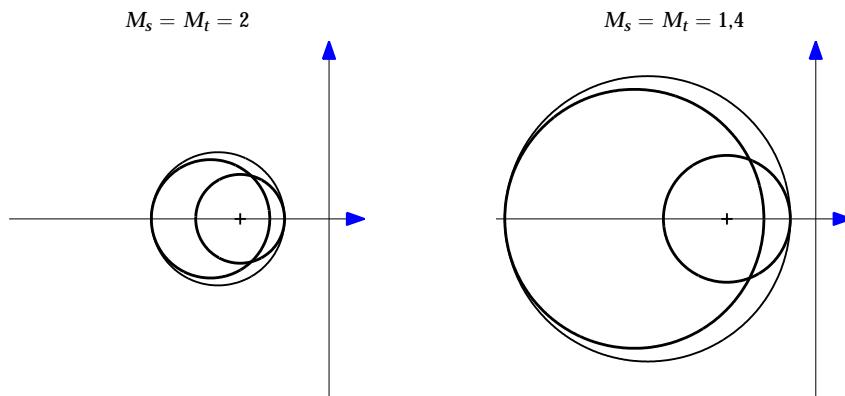


Figura 4.15 Curvas para sensibilidad constante, sensibilidad complementaria constante,y sensibilidad combinada constante.

Cuadro 4.1 Centro y radio de círculos que definen el lugar para sensibilidad constante M_s , sensibilidad complementaria constante M_t , sensibilidad mixta constante, y sensibilidades iguales $M = M_s = M_t$.

Contorno	Centro	Radio
M_s	-1	$1/M_s$
M_t	$-\frac{M_t^2}{M_t^2 - 1}$	$\frac{M_t}{M_t^2 - 1}$
M_s, M_t	$-\frac{M_s(2M_t - 1) - M_t + 1}{2M_s(M_t - 1)}$	$\frac{M_s + M_t - 1}{2M_s(M_t - 1)}$
$M = M_s = M_t$	$-\frac{2M^2 - 2M + 1}{2M(M - 1)}$	$\frac{2M - 1}{2M(M - 1)}$

4.7 Robustez frente a variaciones del proceso

La robustez frente a variaciones del proceso es un tema clave en el diseño de un sistema de control. Los parámetros del proceso pueden cambiar por muchas razones; típicamente dependen de las condiciones de operación. Los retardos de tiempo y las constantes de tiempo a menudo cambian con los niveles de producción. Los parámetros también pueden cambiar a causa del envejecimiento del equipo. Una de las razones claves para utilizar realimentación es que es posible obtener sistemas en lazo cerrado que son insensibles a las variaciones del proceso.

El análisis de las funciones de sensibilidad en la Sección 4.6 arroja luz sobre los efectos de las variaciones del proceso. La ecuación 4.25 muestra el efecto de pequeñas variaciones en el proceso sobre el sistema en lazo cerrado. En particular indica que un sistema en lazo cerrado es insensible a pequeñas variaciones del proceso para frecuencias donde la función de sensibilidad es pequeña.

La inigualdad de robustez dada por (4.32) muestra que un sistema en lazo

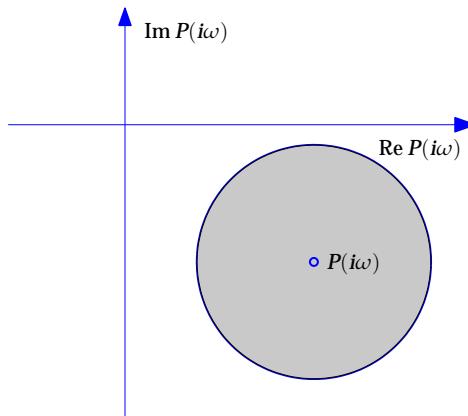


Figura 4.16 El círculo sombreado muestra valores permisibles de $P(i\omega) + \Delta P(i\omega)$ dada por la inigualdad (4.32). El círculo se dibuja para $M_t = 2$.

cerrado es estable cuando se perturba el proceso de $P(s)$ a $P(s) + \Delta P(s)$, donde $\Delta P(i\omega)$ es una función de transferencia estable, si las perturbaciones están acotadas por

$$\frac{|\Delta P(i\omega)|}{|P(i\omega)|} < \frac{1}{|T(i\omega)|}.$$

Esta ecuación es una de las razones de por qué los sistemas realimentados funcionan tan bien en la práctica. Los modelos matemáticos usados para diseñar sistemas de control están con frecuencia fuertemente simplificados. Puede haber errores de modelo y las propiedades de un proceso puede cambiar durante la operación.

La ecuación (4.32) implica que el sistema en lazo cerrado será estable para variaciones sustanciales en la dinámica del proceso. El sistema en lazo cerrado es estable si, para todo ω , la función de transferencia del proceso perturbado $P(i\omega) + \Delta P(i\omega)$ está en un círculo con centro en $P(i\omega)$ y radio $1/|T(i\omega)|$, ver la Figura 4.16. Para un sistema diseñado con $M_t = 2$ es posible cambiar la ganancia del proceso por factores en el rango 0,5 a 1,5 y la fase se puede cambiar 60° . Para un sistema con $M_t = 1,414$ se puede cambiar la ganancia por factores en el rango 0,3 a 1,7, y la fase se puede cambiar 45° .

El problema de la cancelación

Las sensibilidades dependen de la función de transferencia del lazo $G_l = PC$. Criterios de robustez basados en sensibilidades pueden dar resultados erróneos cuando hay factores en las funciones de transferencias del proceso y del controlador que se cancelan entre sí. Ilustraremos lo que acontece con un ejemplo.

EJEMPLO 4.3—CANCELACIONES

Sea un proceso con la función de transferencia

$$P(s) = \frac{1}{s^2 + 2\zeta as + a^2},$$

y un controlador con la función de transferencia

$$C(s) = \frac{50(s^2 + 2\zeta as + a^2)}{s(s^2 + 10s + 50)}.$$

este controlador es una combinación de un controlador PID con un filtro para proporcionar una caída en alta frecuencia y un filtro notch para reducir la excitación del modo oscilatorio de baja-frecuencia. La función de transferencia del lazo es

$$G_l(s) = \frac{50}{s(s^2 + 10s + 50)}.$$

Obsérvese que los modos oscilatorios desaparecen porque el factor aparece en ambos, el controlador y el proceso. Las funciones de sensibilidad son

$$S(s) = \frac{s(s+5)^2}{s^3 + 10s^2 + 50s + 50}$$

$$T(s) = \frac{1}{s^3 + 10s^2 + 50s + 50}.$$

Con los valores numéricos $a = 0,5$ y $\zeta = 0,02$ obtenemos $M_s = 1,2$ y $M_t = 1$. Una aplicación de la inigualdad de robustez (4.32) puede llevarnos a creer que el sistema en lazo cerrado es robusto. Sin embargo, si se diseña un controlador basado en el valor nominal $a = 0,5$ y si los parámetros del proceso se cambian en un 5 por ciento a $a = 0,4775$ el sistema se hace inestable. La razón es que si interpretamos la variación del parámetro como una perturbación aditiva en el modelo del proceso la pequeña perturbación en el parámetro del proceso a se traslada como una perturbación aditiva mucho más grande porque se asocia con un modo resonante con un amortiguamiento relativo muy pequeño. \square

El controlador en el ejemplo no es un buen diseño porque es mala práctica cancelar los polos lentos del proceso.

Otras medidas de robustez

Hay otros resultados de robustez que permiten variaciones más realistas del proceso que la perturbación aditiva estable empleada en la inigualdad de robustez (4.32). Un resultado representa la función de transferencia del proceso como

$$P(s) = \frac{N(s)}{D(s)}$$

donde $N(s)$ y $D(s)$ son funciones de transferencias estables. Los resultados indican que el sistema es estable para variaciones ΔN y ΔD tal que

$$\begin{aligned} \max(|N(i\omega)|, |D(i\omega)|) &= \bar{\sigma} \begin{pmatrix} 1 & P(i\omega) \\ \frac{C(i\omega)}{1 + P(i\omega)C(i\omega)} & \frac{P(i\omega)C(i\omega)}{1 + P(i\omega)C(i\omega)} \end{pmatrix} \\ &= \frac{\sqrt{(1 + |P(i\omega)|^2)(1 + |C(i\omega)|^2)}}{|1 + P(i\omega)C(i\omega)|} = \Sigma(\omega), \end{aligned} \quad (4.36)$$

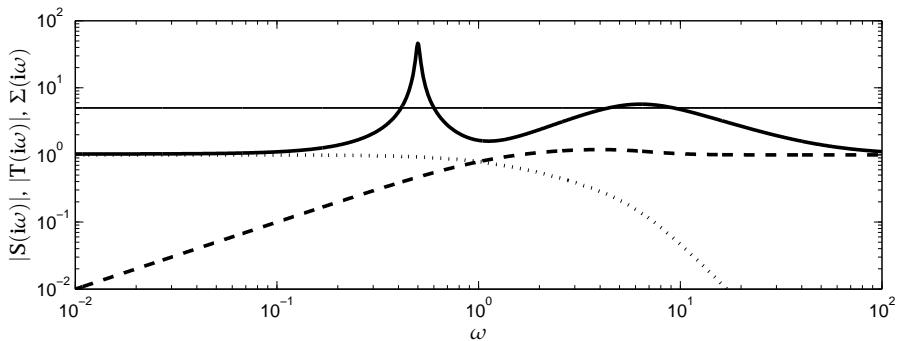


Figura 4.17 Las magnitudes de la función de sensibilidad $|S(i\omega)|$ (punteada), la función de sensibilidad complementaria $|T(i\omega)|$ (a tramos) y el valor singular máximo $\Sigma(i\omega)$ (continua) para el sistema del Ejemplo 4.3.

donde $\bar{\sigma}$ es el valor singular más grande. El parámetro

$$M_\sigma = \max_{\omega} \Sigma(\omega)$$

es una medida de robustez. La condición de robustez (4.32) requiere que la perturbación del proceso $\Delta P(s)$ sea una función de transferencia estable. Criterios basados en M_σ no tienen esta limitación porque permiten perturbaciones más generales del proceso, por ejemplo, cambiar un pequeño polo estable, un integrador, o un polo inestable. También cubre la situación cuando hay cancelaciones de polos y ceros. Para tener buena robustez el parámetro M_σ debería ser menor que 3 a 5. Obsérvese que M_σ es mayor que ambos M_s y M_t .

Para ilustrar la efectividad de M_σ lo aplicamos al Ejemplo 4.3. La Figura 4.17 muestra $|S(i\omega)|$, $|T(i\omega)|$, y $\Sigma(\omega)$ para el sistema nominal del Ejemplo 4.3. Tenemos $M_\sigma = 46$; como esto es mucho mayor que 5 se sigue que el sistema en lazo cerrado tiene una robustez muy pobre.

Otra forma de investigar la robustez es explorar las variaciones en los parámetros del proceso requeridos para hacer inestable al sistema en lazo cerrado. Los cambios en la ganancia y en las constantes de tiempo se pueden obtener sustituyendo $P(s)$ por $\kappa P(\alpha s)$. Las variaciones del proceso que hacen inestable al sistema están dadas por

$$\kappa P(i\alpha\omega)C(i\omega) + 1 = 0.$$

Resolviendo para α y κ para todo ω da las funciones $\kappa(\omega)$ y $\alpha(\omega)$. Peter Hansen ha sugerido el siguiente índice de robustez

$$R_{ph} = \min_{\omega} (\log |\kappa(\omega)| + \log |\alpha(\omega)|). \quad (4.37)$$

Esta medida es una generalización del margen de ganancia y del margen de retraso.

El máximo valor singular M_σ y la medida de robustez R_{ph} son más complicadas que M_s y M_t , y por lo tanto utilizaremos comúnmente M_s y M_t . Debería,

sin embargo, tenerse en cuenta que evaluar la robustez requiere algo de cuidado, particularmente cuando hay cancelaciones y cuando la función de transferencia del lazo tiene picos elevados por encima de la frecuencia de cruce de ganancia. Este es normalmente el caso en el control de movimiento con sistemas que tienen resonancias mecánicas y también para controladores predictivos que se estudian en el Capítulo 8.

4.8 Cuantificando los requisitos

Habiendo comprendido las propiedades fundamentales del lazo de realimentación básico ahora cuantificaremos los requisitos sobre un sistema de control típico. Para hacer esto es necesario tener una clara comprensión del objetivo primario del control. Los problemas de control son muy ricos tal como se analizó en la Sección 4.2. En general, tenemos que considerar

- Atenuación de la perturbación de carga
- Respuesta del ruido de medida
- Robustez frente a incertidumbres del proceso
- Respuesta al punto de consigna

El énfasis sobre los diferentes factores depende del problema de control particular. La robustez es importante para todas las aplicaciones. El seguimiento del punto de consigna es el mayor problema en el control de movimiento, donde se desea que el sistema siga a determinadas trayectorias. En control de procesos, la mayor parte del tiempo el punto de consigna se mantiene normalmente constante; normalmente los cambios se hacen sólo cuando se altera la producción. El rechazo de perturbaciones de carga es en realidad el tema clave en control de procesos. Hay también situaciones donde el objetivo del control no es mantener las variables del proceso en valores especificados. Un ejemplo típico es el control de nivel en tanques de igualación. La razón para utilizar estos tanques es suavizar las variaciones de caudal. En tal caso el nivel del tanque debería fluctuar dentro de ciertos límites. Una buena estrategia es no realizar ninguna acción de control mientras el nivel del tanque permanece dentro de ciertos límites y sólo aplicar control cuando el nivel está próximo a los límites. Esto se llama control promedio o control por tanque de compensación. Hay estrategias especiales desarrolladas para tratar con tales problemas, también han sido aplicadas técnicas tales como planificación de ganancias. Esto se analiza en Sección 9.3.

La conducta lineal del sistema está completamente determinada por la *Banda de los Seis* (4.3). Despreciando la respuesta al punto de consigna es suficiente considerar la *Banda de los Cuatro* (4.4). Las especificaciones se pueden expresar en términos de estas funciones de transferencias.

Una ventaja significativa con una estructura que tiene dos grados de libertad, o ponderación del punto de consigna, es que el problema de la respuesta al punto de consigna se puede desacoplar de la respuesta a las perturbaciones de carga y del ruido de medida. El procedimiento de diseño se puede entonces dividir en dos pasos independientes.

- Primero diseñar el controlador de realimentación C que reduce los efectos de las perturbaciones de carga y la sensibilidad a variaciones en el proceso sin introducir demasiado ruido de medida en el sistema.
- A continuación diseñar el controlador feedforward F para dar la respuesta deseada acambios en el punto de consigna.

A continuación analizaremos cómo se pueden expresar las especificaciones en términos de las propiedades de las funciones de transferencias (4.4).

Respuesta a perturbaciones de carga

Una estimación de la efectividad de un sistema de control para rechazar perturbaciones está dada por (4.26), que compara las salidas de un sistema en lazo abierto y en lazo cerrado cuando las perturbaciones son las mismas. El análisis muestra que se atenúan por la realimentación las perturbaciones con frecuencias menores que la frecuencia de cruce de sensibilidad ω_{sc} y que la amplificación más grande de las perturbaciones es la sensibilidad máxima M_s .

Retornamos ahora específicamente a las perturbaciones de carga que son perturbaciones que mueven las variables del proceso fuera de sus valores deseados. La atenuación de las perturbaciones de carga es de un interés fundamental en control de procesos. Éste es particularmente el caso para problemas de regulación donde los procesos están operando en estado estacionario con punto de consigna constante. Las perturbaciones de carga están a menudo dominadas por las bajas frecuencias. Las señales en escalón son por lo tanto usadas como perturbaciones prototípicas. Las perturbaciones pueden entrar al sistema de muchas formas diferentes. Si nada se conoce, con frecuencia se supone que las perturbaciones entran en la entrada del proceso. La respuesta de la variable de proceso viene entonces dada por la función de transferencia

$$G_{xd} = \frac{P}{1 + PC} = PS = \frac{T}{C}. \quad (4.38)$$

Como las perturbaciones de carga generalmente tienen bajas frecuencias es natural que el criterio realce la conducta de la función de transferencia en bajas frecuencias. El filtrado de la señal de medida tiene sólo un efecto marginal sobre la atenuación de las perturbaciones de carga porque el filtro sólo atenúa las altas frecuencias. Para un sistema con $P(0) \neq 0$ y un controlador con acción integral la ganancia del controlador tiende a infinito para bajas frecuencias, y tenemos la siguiente aproximación para pequeños s :

$$G_{xd} = \frac{T}{C} \approx \frac{1}{C} \approx \frac{s}{k_i}. \quad (4.39)$$

Puesto que las perturbaciones de carga normalmente tienen bajas frecuencias esta ecuación implica que la ganancia integral k_i es una buena medida del rechazo a las perturbaciones de carga.

EJEMPLO 4.4—ATENUACIÓN DE LA PERTURBACIÓN DE CARGA

Sea un proceso con la función de transferencia $P = (s + 1)^{-4}$ y un controlador PI con $k = 0,5$ y $k_i = 0,25$. El sistema tiene $M_s = 1,56$ y $\omega_{ms} = 0,494$. la Figura

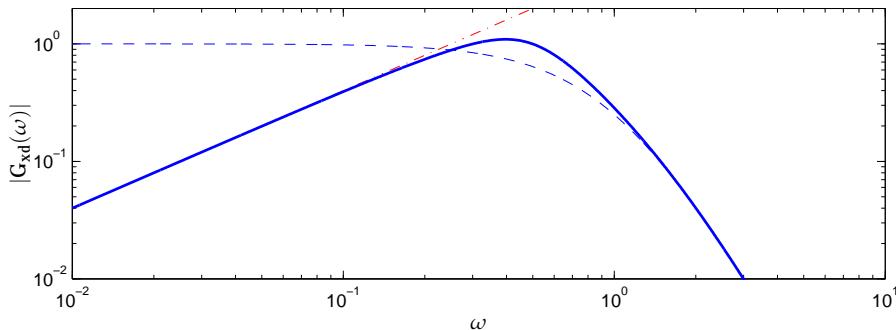


Figura 4.18 La ganancia de la función de transferencia G_{xd} de la perturbación de carga a la variable del proceso para control PI ($k = 0,5$, $T_i = 2,0$) del proceso $P = (s + 1)^{-4}$. La curva con punto y raya muestra la ganancia de s/k_i , y la curva a tramos muestra la ganancia de la función de transferencia del proceso P .

4.18 muestra la curva de magnitud de la función de transferencia (4.38). La figura muestra claramente que la realimentación reduce significativamente la ganancia de baja frecuencia comparado con el sistema en lazo abierto. La línea de punto y raya en la figura muestra la curva de ganancia para la función de transferencia s/k_i . Se observa en la figura que esto es una aproximación muy buena de G_{xd} para bajas frecuencias, aproximadamente hasta ω_{ms} . La ganancia integral k_i es una buena medida de la atenuación de la perturbación de la frecuencia de carga. Para altas frecuencias el rechazo a la perturbación de carga viene dado por la dinámica del proceso; la realimentación no tiene influencia. La frecuencia de cruce de sensibilidad es $\omega_{sc} = 0,25$, que está próxima a k_i .

La atenuación de las perturbaciones de carga puede también caracterizarse en el dominio del tiempo al mostrar la respuesta temporal debida a una perturbación representativa. Esto se ilustra en la Figura 4.19, que muestra la respuesta de la salida del proceso a una perturbación en escalón unitario en la entrada del proceso. La salida tiene su máximo $y_{\max} = 0,66$ para $t_{\max} = 5,62$. Más aun, $t_{\max}\omega_{ms} = 2,76$, la integral del error $IE = 4,00$ y la integral del error absoluto $IAE = 4,26$. □

El error en estado estacionario originado por una perturbación de carga en escalón unitario cuando el control es de tipo proporcional es

$$e_{ss} = \frac{P(0)}{1 + kP(0)}, \quad (4.40)$$

donde k es la ganancia proporcional del controlador. Tal como se indica en la Figura 4.19, el error en estado estacionario para control proporcional se puede emplear como una aproximación del error más grande para control PID. Para el sistema en el Ejemplo 4.4 tenemos $P(0) = 1$ y $k = 0,5$ y (4.40) da la estimación $e_{\max} \approx e_{ss} = 1/1,5 = 0,67$ que está próxima al valor correcto 0.66.

Respuesta a ruido de medida

Una consecuencia inevitable de usar realimentación es que el ruido de medida se realimenta en el sistema. El ruido de medida, que típicamente tiene altas fre-

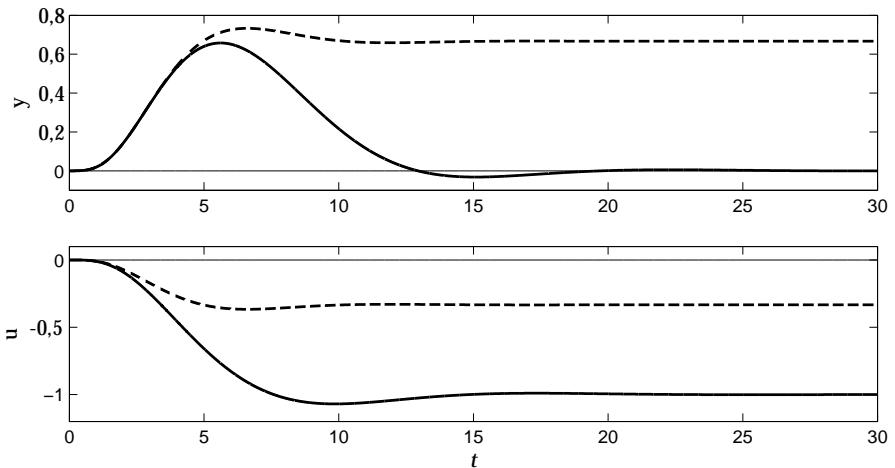


Figura 4.19 Respuesta a una perturbación de carga en la forma de un escalón unitario con un controlador PI con parámetros $k = 0,5$ y $k_i = 0,25$ y el proceso $P = (s + 1)^{-4}$. La curva a tramos muestra la respuesta con un controlador proporcional con ganancia $k = 0,5$.

cuencias, genera acciones de control no deseables y variaciones en la variable del proceso. Las variaciones rápidas en la variable de control son perjudiciales porque originan roturas en válvulas y motores e incluso saturan al actuador. Es importante mantener estas variaciones en un nivel razonable. Un requisito típico es que las variaciones sólo sean una fracción del rango de la señal de control. Las variaciones se pueden ver influidas por el filtraje y por un diseño adecuado de las propiedades de alta frecuencia del controlador.

Los efectos del ruido de medida son así captados por la función de transferencia del ruido de medida a la señal de control

$$G_{un} = \frac{C}{1 + PC} = CS = \frac{T}{P}. \quad (4.41)$$

Para bajas frecuencias (pequeños s) la función de transferencia se approxima a $1/P(0)$ y para altas frecuencias (grandes s) tenemos aproximadamente

$$G_{un} \approx C.$$

Para un controlador PID ideal la función de transferencia G_{un} se hace infinita para grandes s lo que claramente indica la necesidad de filtrar la acción derivativa, tal como se analiza en la Sección 3.3. Lo ilustramos con un ejemplo.

EJEMPLO 4.5—EFFECTO DEL FILTRADO

La Figura 4.20 muestra la curva de ganancia de la función de transferencia (4.41) para control PID del proceso $P = (s + 1)^{-4}$. La línea a tramos es para un controlador con un filtro de primer orden en la acción derivativa y la línea continua para un controlador con un filtro de segundo orden de la señal medida. Las diferencias significativas en las funciones de transferencias para altas frecuencias es

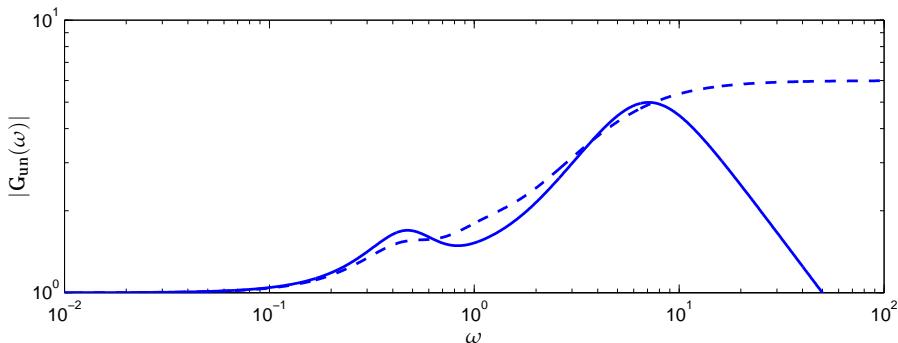


Figura 4.20 La magnitud de la función de transferencia $G_{un} = CS$ para control PID ($k = 1$, $T_i = 2$, $T_d = 1$, $T_f = 0,2$) del proceso $P = (s + 1)^{-4}$. La línea continua representa un controlador con un filtro de ruido de segundo orden de la señal medida (3.16) y la línea a tramos un controlador con un filtro de primer orden de la acción derivativa (3.15).

una buena motivación para preferir el controlador con filtrado de la señal medida. Para bajas frecuencias (pequeños s) la función de transferencia se aproxima a $1/P(0)$. \square

Una medida sencilla del efecto del ruido de medida es la máxima ganancia de la función de transferencia G_{un} ,

$$M_{un} = \max_{\omega} |G_{un}(i\omega)|. \quad (4.42)$$

Para control PI la ganancia de la función de transferencia G_{un} tiene un pico próximo al pico de la función de sensibilidad y tenemos aproximadamente

$$M_{un} \approx M_s K. \quad (4.43)$$

Para control PID la ganancia de la función de transferencia G_{un} tiene normalmente dos máximos locales, uno está próximo al máximo de la función de sensibilidad. El otro pico es mayor

$$M_{un} \approx k_d / T_d, \quad (4.44)$$

y ocurre próximo a la frecuencia $1/T_f$.

Si la desviación estándar del ruido de medida es σ_n , una estimación bruta de las variaciones en la señal de control es $M_{un}\sigma_n$. Puede hacerse una evaluación más precisa si se conoce el espectro en potencia $\phi_n(\omega)$ del ruido de medida. La desviación estándar de la señal de control está entonces dada por

$$\sigma_u^2 = \int_{-\infty}^{\infty} |G_{un}(i\omega)|^2 \phi_n(\omega) d\omega. \quad (4.45)$$

Sin embargo, es raro que se disponga de esta información detallada para aplicaciones típicas.

Robustez a variaciones del proceso

La inversa de la sensibilidad máxima es la distancia más corta desde el punto crítico -1 al digrama de Nyquist de la función de transferencia del lazo.

La sensibilidad a pequeñas variaciones en la dinámica del proceso se deduce de la función de sensibilidad. Tenemos

$$\frac{dT}{T} = S \frac{dP}{P}.$$

Las variaciones en la dinámica del proceso tienen así una influencia pequeña sobre el sistema en lazo cerrado para frecuencias donde la función de sensibilidad es pequeña.

Las variaciones en la dinámica del proceso pueden también conducir a inestabilidad. La condición

$$\frac{|\Delta P(i\omega)|}{|P(i\omega)|} < \frac{1}{|T(i\omega)|}$$

garantiza que una variación $\Delta P(i\omega)$ en la función de transferencia del proceso no haga al sistema inestable. La robustez a variaciones del proceso se obtiene así por las funciones de sensibilidad y de sensibilidad complementaria. Medidas simples son la máxima sensibilidad M_s , el máximo de la sensibilidad complementaria M_t , o la máxima sensibilidad combinada M . Valores típicos de las sensibilidades están en el rango de $1,2 - 2,0$.

Otras medidas son el margen de ganancia g_m (normalmente 2 a 8), el margen de fase φ_m (típicamente 30° a 60°), o el margen de estabilidad $s_m = 1/M_s$ (típicamente 0,5 a 0,8). Compárese con la Sección 4.4.

Compromisos

La atenuación de la perturbación de carga se deduce de la ganancia integral k_i . Se sigue de (4.39) que la atenuación para perturbaciones de baja frecuencia es aproximadamente inversamente proporcional a k_i . La inyección del ruido de medida se determina de la ganancia del ruido M_{un} . Se sigue de (4.42) que M_{un} da la ganancia del ruido de medida a la variable de control. Por lo tanto, el compromiso entre atenuación de la perturbación de carga e inyección de ruido de medida se puede lograr equilibrando las ganancias k_i y M_{un} .

Respuesta del punto de consigna

Utilizando un controlador con dos grados de libertad es posible obtener cualquier respuesta deseada frente a cambios en el punto de consigna. Esto será analizado más aún en el Capítulo 5. Las limitaciones se dan por la magnitud permisible de la señal de control. En algunos casos sólo se mide el error de control. No se puede usar entonces un controlador con dos grados de libertad y la respuesta a cambios en el punto de consigna tiene que manejarse mediante una elección adecuada de la función de transferencia del controlador. Se pueden evitar grandes sobrelongaciones requiriendo valores pequeños de M_t .

Resumen

Resumiendo, encontramos que la conducta del sistema se puede caracterizar de la forma siguiente. la función de transferencia de la perturbación de carga a la

variable del proceso es

$$G_{yd} = \frac{P}{1 + PC} = PS \approx \frac{s}{k_i}, \quad (4.46)$$

donde la aproximación se mantiene para bajas frecuencias.

El efecto del ruido de medida se puede obtener de la ganancia del ruido

$$M_{un} = \max_{\omega} |G_{un}(i\omega)| \approx \begin{cases} kM_s & \text{para control PI} \\ k_d/T_f & \text{para control PID (3.16),} \end{cases} \quad (4.47)$$

que depende fuertemente del filtrado del ruido de medida.

Estabilidad y robustez frente a incertidumbres del proceso se pueden expresar por la función de sensibilidad y la función de sensibilidad complementaria

$$S = \frac{1}{1 + PC}, \quad T = \frac{PC}{1 + PC},$$

donde los valores máximos de las funciones de sensibilidad M_s y M_t son buenas medidas cuantitativas. El parámetro $1/M_s$ es la distancia más corta desde el punto crítico al diagrama de Nyquist de la función de transferencia del lazo.

Características esenciales de la atenuación de la perturbación de carga, la inyección del ruido de medida, y la robustez pueden así obtenerse mediante cuatro parámetros k_i , M_{un} , M_t , y M_s . Una característica atractiva de esta elección de parámetros es que k_i y M_{un} están directamente relacionados con los parámetros del controlador y que hay buenos métodos de diseño que pueden garantizar M_s y M_t dados.

4.9 Especificaciones clásicas

Las especificaciones que hemos dado tienen la ventaja de que determinan la robustez así como las respuestas a perturbaciones de carga, ruido de medida, y cambios en el punto de consigna con sólo cuatro parámetros. Desgraciadamente, ha sido la tradición en control PID juzgar un sistema basándose sólo en una respuesta, típicamente la respuesta de la salida frente a un cambio en el punto de consigna. Esto puede ser muy erróneo tal como hemos analizado previamente. Un gran número de parámetros diferentes han sido también usados para caracterizar las respuestas. Por completitud y para conectar con la literatura clásica sobre control PID en esta sección se resumirán algunas especificaciones clásicas.

Criterios basados en la respuesta temporal

Muchos criterios están relacionados con la respuesta temporal, por ejemplo, la respuesta a un escalón a cambios en el punto de consigna o la respuesta a un escalón a perturbaciones en la carga. Es común usar algunas características del error típicamente valores extremos, asíntotas, áreas, etc.

El error máximo e_{\max} se define como

$$\begin{aligned} e_{\max} &= \max_{0 \leq t < \infty} |e(t)| \\ T_{\max} &= \arg \max |e(t)|. \end{aligned} \quad (4.48)$$

El tiempo T_{\max} donde ocurre el máximo es una medida de la respuesta temporal del sistema. La integral del error absoluto (IAE) se define como

$$IAE = \int_0^{\infty} |e(t)| dt. \quad (4.49)$$

Un criterio relacionado es la integral del error (IE), definida como

$$IE = \int_0^{\infty} e(t) dt. \quad (4.50)$$

Los criterios IE y IAE son el mismo si el error no cambia de signo. Obsérvese que IE puede ser muy pequeño incluso aunque el error no lo sea. Para que IE sea relevante es necesario añadir condiciones que aseguren que el error no es demasiado oscilatorio. El criterio IE es una elección natural para el control de las variables de calidad para un proceso donde el producto se envía a un tanque de mezcla. Sin embargo el criterio puede ser fuertemente erróneo, en otras situaciones. Será cero para un sistema oscilatorio sin amortiguamiento. Será también cero para un lazo de control con dos integradores.

Hay muchos otros criterios, por ejemplo, la integral del tiempo multiplicado por el error absoluto, definido por

$$ITNAE = \int_0^{\infty} t^n |e(t)| dt. \quad (4.51)$$

La integral del error al cuadrado (ISE) se define como

$$ISE = \int_0^{\infty} e(t)^2 dt. \quad (4.52)$$

Hay otros criterios que toman en cuenta ambas señales de entrada y salida, por ejemplo, el criterio cuadrático

$$QE = \int_0^{\infty} (e^2(t) + \rho u^2(t)) dt, \quad (4.53)$$

donde ρ es un factor de ponderación. Los criterios IE y QE se pueden calcular fácilmente de forma analítica, sin embargo se requieren simulaciones para determinar IAE.

Una razón para usar IE es que su valor está directamente relacionado con el parámetro k_i del controlador PID, tal como se ilustra por el ejemplo siguiente.

EJEMPLO 4.6—GANANCIA INTEGRAL E IE PARA PERTURBACIONES DE CARGA
Sea la ley de control

$$u(t) = ke(t) + k_i \int_0^t e(t) dt - k_d \frac{dy}{dt}.$$

Suponga que este controlador da un sistema en lazo cerrado estable. Más aún, suponga que el error es cero inicialmente y que se aplica un escalón unitario como

perturbación de carga en la entrada del proceso. Como el sistema en lazo cerrado es estable y tiene acción integral el error de control irá a cero. Encontramos así que

$$u(\infty) - u(0) = k_i \int_0^{\infty} e(t) dt.$$

Como la perturbación se aplica en la entrada del proceso, el cambio en la señal de control es igual al cambio de la perturbación. Por lo tanto, $u(\infty) - u(0) = 1$ y obtenemos

$$IE = \int_0^{\infty} e(t) dt = \frac{1}{k_i} = \frac{T_i}{K}. \quad (4.54)$$

□

La ganancia integral k_i es así inversamente proporcional a la integral del error originado por un escalón unitario como perturbación de carga aplicada a la entrada del proceso.

Respuesta del punto de consigna

Las siguientes especificaciones sobre el punto de consigna se expresan generalmente en el dominio del tiempo. Pueden incluir requisitos sobre el tiempo de subida, tiempo de asentamiento, razón de decaimiento, sobrelongación, y error en estado estacionario frente a cambios en escalón en el punto de consigna. Estas cantidades se definen como sigue, ver la Figura 4.21.

- El *tiempo de subida* T_r se define o como la inversa de la máxima pendiente de la respuesta a un escalón o como el tiempo que tarda la respuesta a un escalón en cambiar del 10 por ciento al 90 por ciento de su valor en estado estacionario.
- El *tiempo de asentamiento* T_s es el tiempo que le lleva a la respuesta en escalón permanecer dentro del p por ciento de su valor en estado estacionario. Se suelen utilizar normalmente los valores $p = 1, 2, y 5$ por ciento del valor en estado estacionario.
- La *razón de decaimiento* d es la razón entre dos máximos consecutivos del error para un cambio en escalón en el punto de consigna o en la carga; ver Figura 2.35. El valor $d = 1/4$, que se llama razón de decaimiento de un cuarto, ha sido usado tradicionalmente. Este valor es, sin embargo, normalmente demasiado alto, como se demostrará posteriormente.
- La *sobrelongación* o es la razón entre la diferencia entre el primer pico y el valor en estado estacionario de la respuesta a un escalón. Con frecuencia se da en tanto por ciento. En aplicaciones de control industrial es común especificar una sobrelongación máxima de 8 a 10 por ciento. Sin embargo en muchas situaciones es deseable, tener una respuesta sobreamortiguada con ninguna sobrelongación.
- El *error en estado estacionario* $e_{ss} = y_{sp} - y_0$ es el error de control en estado estacionario e . Esto es siempre cero para un controlador con acción integral.

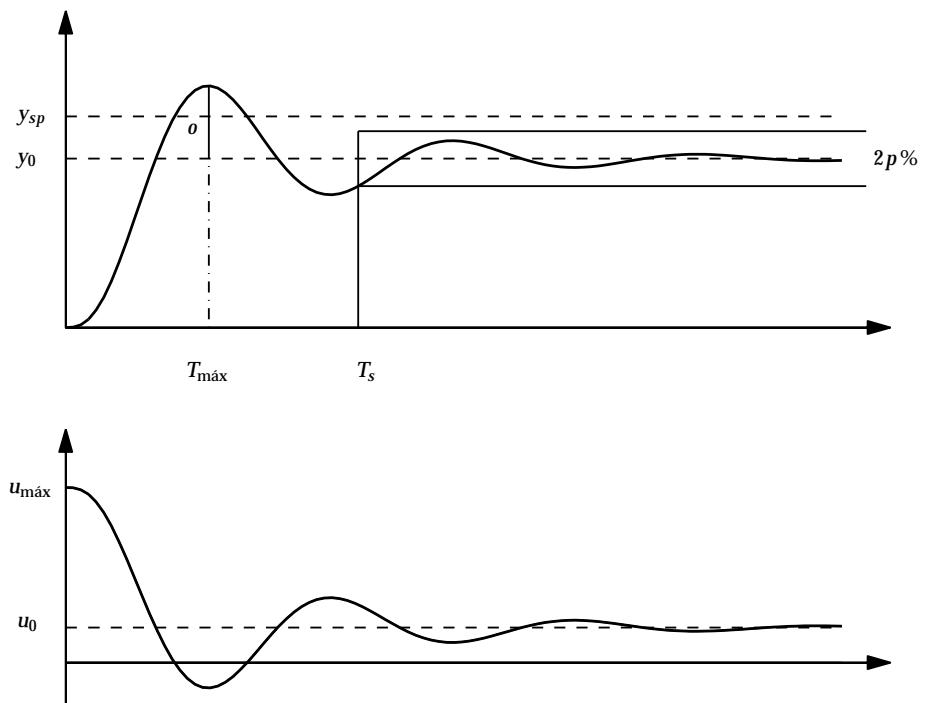


Figura 4.21 Especificaciones sobre el seguimiento del punto de consigna basadas en la respuesta temporal a un escalón unitario en el punto de consigna. La curva superior muestra la respuesta de la salida, y la curva inferior muestra la señal de control correspondiente.

Los actuadores pueden tener limitaciones de velocidad, lo que significa que los cambios a un escalón en la señal de control no aparecerán instantáneamente. En sistemas de control de movimiento es a menudo más relevante considerar respuestas a señales en rampa en vez de señales en escalón.

Respuesta a perturbaciones de carga

La respuesta a las perturbaciones de carga es de importancia fundamental en control de procesos. La Figura 4.22 muestra la salida para una perturbación de carga en escalón aplicada en la entrada del proceso y algunos valores que se usan para caracterizar la respuesta. La figura muestra el *error máximo* e_{\max} , el tiempo que lleva alcanzar el máximo T_{\max} , y el tiempo de asentamiento T_s . Además de estos números, también se suelen usar la integral del error (IE) o la integral del error absoluto (IAE) para caracterizar la respuesta de la perturbación de carga. El error máximo para un escalón unitario y el tiempo donde se alcanza se puede aproximar por

$$e_{\max} = \frac{1}{1 + kP(0)} \quad (4.55)$$

$$T_{\max} \approx \frac{3}{\omega_{ms}}.$$

Ilustramos estas estimaciones con un ejemplo.

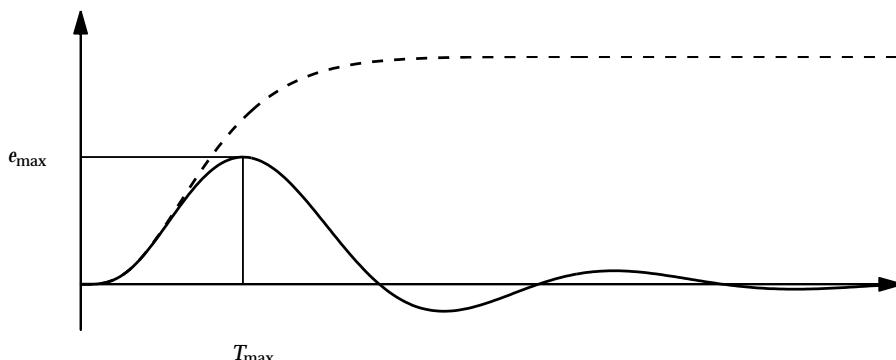


Figura 4.22 El error debido a una perturbación en la carga en escalón unitario en la entrada del proceso y algunos valores usados para caracterizar la atenuación de las perturbaciones de carga. La curva a tramos muestra el error en lazo abierto.

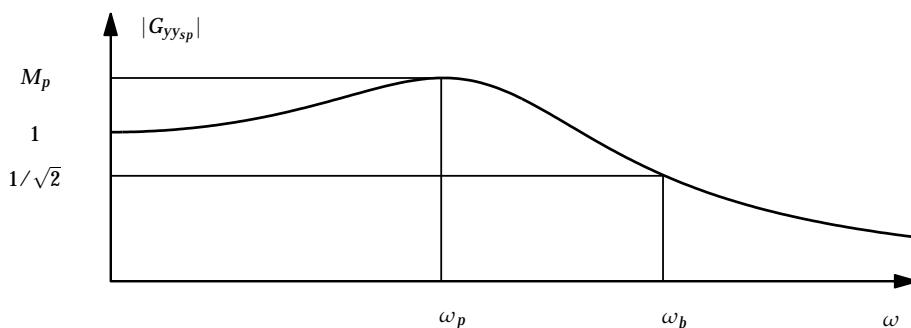


Figura 4.23 Curva de ganancia para la función de transferencia del punto de consigna a la salida.

EJEMPLO 4.7—ESTIMACIÓN DEL ERROR MÁXIMO

Cuando un proceso con la función de transferencia $P(s) = (s + 1)^{-4}$ se controla por un controlador PI con parámetros $k = 0,78$ y $k_i = 0,38$, tenemos $\omega_{ms} = 0,559$, $e_{\max} = 0,59$, y $T_{\max} = 5,15$. Las estimas anteriores dan $e_{\max} \approx 0,56$, $T_{\max} = 5,6$.

□

Criterios basados en la respuesta en frecuencia

Las especificaciones pueden también relacionarse con la respuesta en frecuencia. Como las especificaciones se enfocaron originalmente sobre la respuesta al punto de consigna fue natural considerar la función de transferencia del punto de consigna a la salida. En la Figura 4.23 se muestra una curva de ganancia típica para esta respuesta. Es natural requerir que la ganancia en estado estacionario sea uno. Especificaciones típicas son entonces las siguientes:

- El *pico de resonancia* M_p es el valor máximo de la respuesta en frecuencia.
- El *pico de frecuencia* ω_p es la frecuencia donde ocurre el máximo.

- El *ancho de banda* ω_b es la frecuencia donde la ganancia ha disminuido a $1/\sqrt{2}$.

Para un sistema con realimentación del error la función de transferencia del punto de consigna a la salida es igual a la función de transferencia de la sensibilidad complementaria, y tenemos $M_p = M_t$.

Las especificaciones pueden también relacionarse con la función de transferencia del lazo. Parámetros útiles previamente analizados son

- Frecuencia de cruce de ganancia ω_{gc} .
- Margen de ganancia g_m .
- Margen de fase φ_m .
- Máxima sensibilidad M_s .
- Frecuencia donde la función de sensibilidad tiene su máximo ω_{ms} .
- Frecuencia de cruce de sensibilidad ω_{sc} .
- Sensibilidad complementaria máxima M_t .
- Frecuencia donde la función de sensibilidad complementaria tiene su máximo ω_{mt} .

Relaciones entre las especificaciones en el dominio del tiempo y el dominio de la frecuencia

Hay relaciones aproximadas entre las especificaciones en el dominio del tiempo y de la frecuencia. Sea $G(s)$ la función de transferencia del punto de consigna a la salida. En el dominio del tiempo la velocidad de respuesta se caracteriza por el tiempo de subida T_r , el tiempo de residencia medio T_{ar} , o el tiempo de asentamiento T_s . En el dominio de la frecuencia el tiempo de respuesta se puede caracterizar por el ancho de banda en lazo cerrado ω_b , la frecuencia de cruce de ganancia ω_{gc} , y la frecuencia de sensibilidad ω_{ms} . El producto del ancho de banda y el tiempo de subida es aproximadamente constante, y tenemos

$$T_r \omega_b \approx 2. \quad (4.56)$$

Ha sido previamente mostrado que

$$T_{ar} = -\frac{G'(0)}{G(0)};$$

ver (2.16).

La sobrelongación de la respuesta a un escalón o está relacionado con el pico M_p de la respuesta en frecuencia en el sentido de que un pico mayor normalmente implica una sobrelongación mayor. Desgraciadamente, no hay relaciones sencillas porque la sobrelongación también depende de cuán rápidamente decrece la respuesta en frecuencia. Para $M_p < 1,2$ la sobrelongación o en la respuesta a un escalón está a menudo próxima a $M_p - 1$. Para valores mayores de M_p la sobrelongación es normalmente menor que $M_p - 1$. Estas relaciones no se

mantienen para todos los sistemas; hay sistemas con $M_p = 1$ que tienen una sobreelongación positiva. Estos sistemas tienen funciones de transferencias que decaen rápidamente alrededor del ancho de banda.

Para evitar sobreelongaciones en sistemas con realimentación del error es por lo tanto aconsejable requerir que el máximo de la función de sensibilidad complementaria sea pequeña, del orden de, $M_t = 1,1 - 1,2$ con el fin de evitar la gran sobreelongación en la respuesta a un escalón para las señales de referencia.

Evaluación del comportamiento

Antes de diseñar un controlador es útil hacer una evaluación preliminar del comportamiento que se puede alcanzar. Es interesante conocer si un controlador PID es suficiente o si el comportamiento puede mejorarse sustancialmente utilizando un controlador más complejo. Es también interesante conocer si un controlador PI es suficiente o si la acción derivativa da mejoras significativas. Para hacer la evaluación necesitamos alguna medida del comportamiento. En esta sección emplearemos la frecuencia de cruce de ganancia ω_{gc} como una medida. Cuando el margen de fase es 60° esta frecuencia es igual a la frecuencia de cruce de sensibilidad ω_{sc} . Recuerde de la Sección 4.6 que perturbaciones con frecuencias más bajas que ω_{sc} se reducen por realimentación. Para márgenes de fase menores que 60° tenemos $\omega_{sc} < \omega_{gc}$, y para márgenes de fase mayores $\omega_{sc} > \omega_{gc}$. La atenuación de las perturbaciones de carga se mejora así con frecuencias de cruce de ganancia creciente.

Dinámicas de procesos con propiedades de fase no mínima como un retardo de tiempo imponen limitaciones fundamentales sobre el comportamiento alcanzable que se puede expresar por la inigualdad

$$\omega_{gc}L < a, \quad (4.57)$$

donde a es un número menor que 1. Como el retardo de tiempo verdadero L rara vez se conoce se puede aproximar por el retardo de tiempo aparente L_a . La Figura 4.24 muestra el producto $\omega_{gc}L_a$ para un gran conjunto de sistemas bajo control PID robusto. Los círculos que representan sistemas FOTD muestran que el producto es 0,5 para sistemas FOTD. Para sistemas de orden superior con dinámicas dominadas por la constante de tiempo el producto $\omega_{gc}L_a$ es mayor que 0,5 porque el retardo de tiempo aparente del modelo FOTD aproximante es mayor que el retardo de tiempo verdadero del sistema.

Considérese un sistema en lazo cerrado con una función de transferencia del proceso $P(s)$ y un controlador con función de transferencia $C(s)$. La frecuencia de cruce de ganancias se define por

$$\arg P(i\omega_{gc}) + \arg C(i\omega_{gc}) = -\pi + \varphi_m. \quad (4.58)$$

Obsérvese que se utilizan unidades de radianes en esta ecuación. Un controlador PD con filtrado tiene un adelanto de fase máximo del orden de 60° , un controlador proporcional tiene retardo de fase cero, un controlador PI tiene un retraso de fase de 45° , y un controlador PID puede tener un adelanto de fase de 45° . Si se desea un margen de fase de 45° se sigue de la Ecuación 4.58 que las frecuencias de cruce para control PI, PID, y PD son las frecuencias donde el proceso tiene

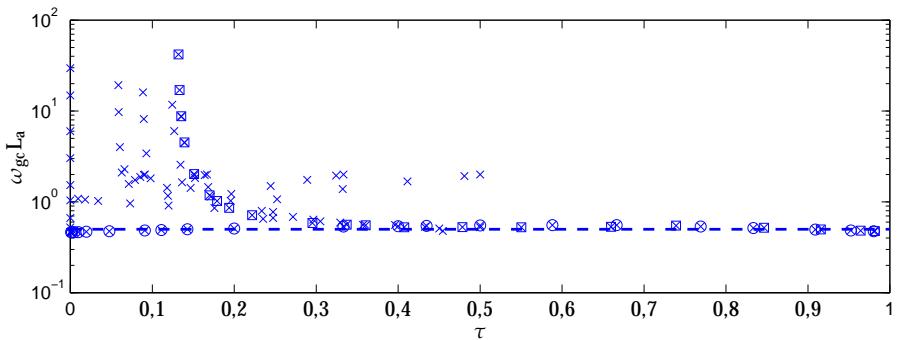


Figura 4.24 El producto de la frecuencia de cruce de ganancia ω_{gc} y el retardo de tiempo aparente L_a como función del retardo de tiempo normalizado, para un gran conjunto de sistemas bajo control PID. Los círculos muestran resultados para sistemas FOTD y los cuadrados para sistemas SOTD. Los controladores PID se diseñan para dar una sensibilidad combinada $M = 1.4$. Todos los sistemas se describen en la Sección 7.2. La línea a tramos da la relación $\omega_{gc}L_a = 0.5$.

un retardo de fase de 90° , 135° , y 195° , respectivamente. Estas frecuencias se denotan como ω_{90} , ω_{135} , y ω_{195} . Se puede obtener una estimación de las ganancias del controlador requeridas calculando las ganancias del proceso en las frecuencias correspondientes. Nótese que esta evaluación sólo requiere la función de transferencia del proceso. Lo ilustramos con dos ejemplos.

EJEMPLO 4.8—PROCESOS CON MÚLTIPLES CONSTANTES DE TIEMPO

Sea un proceso con la función de transferencia

$$P(s) = \frac{1}{(s+1)^4}.$$

Tenemos $\omega_{90} = 0,41$ y $K_{90} = 0,73$, donde K_{90} denota la ganancia del proceso en ω_{90} . Más aún, tenemos $\omega_{135} = 0,67$, $K_{135} = 0,48$, y $\omega_{195} = 1,14$, $K_{195} = 0,19$. Podemos así esperar que perturbaciones con frecuencias más bajas que $0,4$ rad/s pueden reducirse mediante control PI. Como ω_{135} es moderadamente mayor que ω_{90} podemos esperar que un controlador PI puede mejorarse algo introduciendo la acción derivativa. La ganancia de un controlador PID se puede esperar que sea dos veces más grande que para control PI. También obsérvese que el retardo de tiempo aparente es $L = 2,14$ y que $\omega_{gc} = 0,47$ que está en buen acuerdo con (4.24). \square

EJEMPLO 4.9—UN PROCESO CON CONSTANTE DE TIEMPO DOMINANTE

Sea un proceso con la función de transferencia

$$P(s) = \frac{1}{(s+1)(0,1s+1)(0,01s+1)(0,001s+1)}.$$

Tenemos $\omega_{90} = 3,0$, $K_{90} = 0,3$, $\omega_{135} = 9,9$, $K_{135} = 0,07$, y $\omega_{195} = 47,5$, $K_{195} = 0,004$. Podemos así esperar que perturbaciones con frecuencias más bajas que 3

Cuadro 4.2 Parámetros de controladores PI para el proceso $P(s) = (s + 1)^{-3}$ diseñado con diferentes M_s .

M_s	k	k_i	M_{un}	b	ω_{ms}	IAE	T_s	M_t
1,2	0,355	0,171	0,426	1,00	0,671			1,00
1,4	0,633	0,325	0,866	1,00	0,74	3,07	10,3	1,00
1,6	0,862	0,461	1,379	0,93	0,79	2,28	7,87	1,05
1,8	1,056	0,580	1,901	0,70	0,83	2,00	6,77	1,24
2,0	1,222	0,685	2,444	0,50	0,86	1,89	6,27	1,45

rad/s se puedan reducir mediante control PI. Con un controlador PID pueden reducirse perturbaciones con frecuencias hasta 9,9 rad/s. En este caso, hay beneficios de comportamiento significativos por usar la acción derivativa. Como ω_{195} es mucho mayor que ω_{135} puede haber beneficios sustanciales si se utilizan controladores más complejos. Como la ganancia del proceso K_{135} es tan baja la mejora en beneficios requieren controladores con alta ganancia, y los beneficios pueden no ser realizables a menos que el ruido del sensor sea muy bajo. \square

Parámetros de diseño

En el diseño e implementación del sistema de control es conveniente tener un parámetro que se pueda cambiar para influir en los compromisos claves en el problema de diseño. Se puede obtener un comportamiento expresado por tiempo de respuesta rápido y buena atenuación de las perturbaciones de carga, pero se pueden necesitar grandes señales de control. Requisitos más estrictos sobre robustez pueden conducir a un comportamiento más pobre.

El compromiso entre comportamiento y robustez varía entre los diferentes problemas de control. Por lo tanto, es deseable tener un parámetro de diseño para cambiar las propiedades del sistema en lazo cerrado. Idealmente, el parámetro debería estar directamente relacionado con el comportamiento o la robustez del sistema; no debería estar orientado al proceso. Debería tener buenos valores por defecto de forma que un usuario no se vea forzado a seleccionar algún valor. Esto es de especial importancia cuando se usa el procedimiento de diseño para la sintonía automática. El parámetro de diseño debería también tener una buena interpretación física y límites naturales para simplificar su ajuste.

La conducta de un sistema puede a menudo caracterizarse por unos pocos de polos dominantes que están próximos al origen. Cuando hay un polo dominante real, este polo se puede usar como un parámetro de diseño. Esto se emplea, por ejemplo, en el método de diseño de la sintonía lambda, que se analizará en Sección 6.5. Cuando los polos dominantes son complejos la distancia del origen de los polos ω_0 y su amortiguamiento relativo ζ son buenos parámetros de diseño. Esto se aplica a controladores basados en el diseño por asignación de polos, que se presentará en la Sección 6.4. La sensibilidad máxima M_s o la sensibilidad combinada M son buenas variables de diseño para problemas de regulación. Esto se

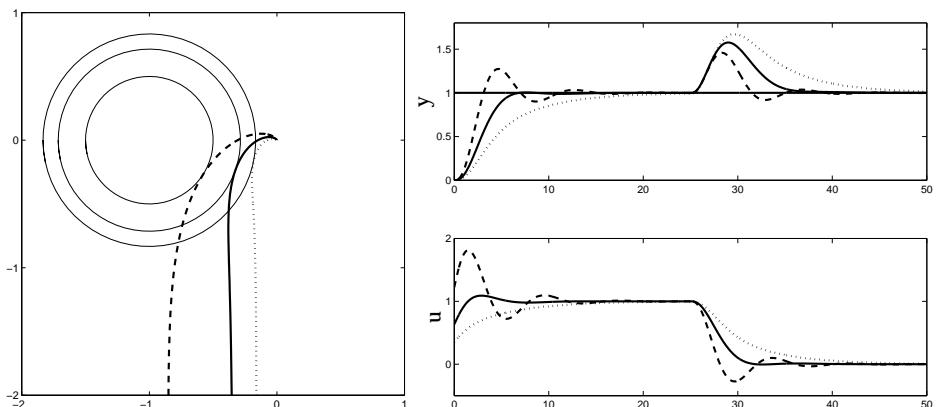


Figura 4.25 Ilustra los efectos de utilizar M_s como un parámetro de diseño. Las curvas de la izquierda muestran los diagramas de Nyquist de las funciones de transferencias de lazo junto con los círculos M constante $M_s=1.2$ (punteado), 1.4, 2.0 (a tramos). Las curvas de la derecha muestran las salidas del proceso y las señales de control para los diferentes valores de los parámetros de diseño.

ilustra en Figura 4.25, que muestra el efecto de M_s sobre las respuestas del tiempo y de frecuencia para un controlador PI, y en la Tabla 4.2, que da los valores numéricos de los parámetros del controlador y algunos criterios. La respuesta a un escalón $M_s = 1.2$ no tiene sobreelongación y tiempos de asentamiento relativamente grandes. El tiempo de asentamiento disminuye y la sobreelongación aumenta cuando el valor de M_s se aumenta. Nótese que se emplea la ponderación del punto de consigna para valores más grandes de M_s para reducir la sobreelongación. El comportamiento también mejora con M_s crecientes. Los valores del IAE disminuyen por un factor del orden de 2. Aparte de su uso en diseño es también posible implementar sistemas donde el usuario puede ajustar en línea los parámetros de diseño.

4.10 Resumen

En esta sección hemos resumido algunas cuestiones importantes para el diseño de sistemas de control, con una atención particular para el control PID. Un análisis del lazo de realimentación básico mostró que es necesario considerar seis funciones de transferencia (la banda de los seis) para determinar las propiedades de un lazo de realimentación. Esto se desprecia severamente en la mayoría de los textos elementales sobre control y en la literatura de controladores PID. A continuación se presentó la noción de estabilidad. Esto es importante porque el riesgo de inestabilidad es la principal desventaja de la realimentación. También se introdujeron criterios de estabilidad y márgenes de estabilidad. Los criterios de estabilidad también hicieron posible obtener las regiones de parámetros que dan un sistema en lazo cerrado estable bajo control PID. La caracterización de un sistema en lazo cerrado por sus polos y ceros proporciona un conocimiento muy valioso, y está también estrechamente relacionada con muchos métodos

de diseño. También se introdujeron la función de sensibilidad y la función de sensibilidad complementaria, que son útiles para expresar la robustez frente a variaciones de los parámetros. A continuación se estudió el problema del diseño del controlador, y una serie de criterios utilizados para dar especificaciones sobre un sistema de control. Los factores claves son las perturbaciones de carga, el ruido de medida, la robustez, y la respuesta al punto de consigna. Un resultado agradable es que para sistemas que tienen dos grados de libertad es posible diseñar para perturbaciones y robustez. La respuesta al punto de consigna deseada puede entonces obtenerse utilizando una acción feedforward. Para control PID la ponderación del punto de consigna es una forma especial de controlador con dos grados de libertad que a menudo es suficiente. Se demuestra también que los requisitos claves se pueden parametrizar de una forma sencilla. La respuesta frente a perturbaciones de carga se deducen de la ganancia integral del controlador k_I . Los efectos del ruido de medida se tienen en cuenta por la ganancia del ruido M_{un} , que mantiene una relación sencilla con los parámetros del controlador. La robustez se considera por las sensibilidades máximas M_s y M_t .

4.11 Notas y referencias

El diseño de sistemas de control es complicado porque hay que considerar muchos factores y se tienen que hacer muchos compromisos. Es por lo tanto natural que llevase tiempo antes de que se desarrollase una buena comprensión de este tema. Los primeros trabajos sobre diseño de control se basaron en las ecuaciones diferenciales que describen al sistema en lazo cerrado. Un enfoque típico fue ajustar los parámetros del controlador de forma que los polos en lazo cerrado tenían propiedades deseadas. Métodos sistemáticos para el diseño de sistemas de control aparecieron en los 1940 cuando emergió el campo del control. Los métodos de diseño se basaron en la respuesta en frecuencia, los cálculos se daban en gráficas, y el modelado con frecuencia se hacía experimentalmente perturbando al sistema con señales sinusoidales; ver [55, 187, 66, 78]. Es notable que se le diese un interés particular a la robustez frente a variaciones del proceso. El libro [176] da una presentación del tema. Este libro también da relevancia al concepto importante de controladores que tienen dos grados de libertad. Tales controladores admiten un desacople de las respuestas a puntos de consigna y perturbaciones de carga.

Hubo un desplazamiento de paradigma en los 1960 cuando volvieron a aparecer las ecuaciones diferenciales con el nombre de sistemas en el espacio de estado; ver [372]. Esto coincidió con la aparición de los computadores digitales, que permitían hacer cálculos numéricos eficientes. Las ideas importantes de control óptimo y filtrado de Kalman son contribuciones claves; ver [44, 190, 189, 191, 282, 41, 68].

Hubo un desarrollo muy dinámico de la teoría, muchos métodos de diseño y se desarrollaron también técnicas computacionales eficientes ver [57].

El tema de la robustez fue desgraciadamente despreciado durante un largo período de tiempo. Esto se remedió con la aparición del método denominado teoría \mathcal{H}_∞ , que condujo a una reconciliación con los métodos clásicos de respuesta en frecuencia. Los libros [99, 376, 317] dan una perspectiva equilibrada. Los criterios de robustez M_s , M_t , y M_σ son resultados de la teoría de control robusto.

Un nuevo criterio de robustez interesante que se enfoca sobre variaciones en los parámetros del proceso han sido sugeridos en [153] y [154]. La cuestión de las limitaciones fundamentales está estrechamente relacionada con la robustez tal como se analiza en [38]. Para control de procesos el retardo de tiempo verdadero es un factor limitante clave. Nótese que el retardo de tiempo verdadero puede ser diferente del retardo de tiempo aparente obtenido cuando se ajustan modelos FOTD.

Muchos practicantes de control han sido completamente conscientes de la importancia del compromiso entre comportamiento y robustez; ver [310], y es ahora gratificante ver que la teoría de control robusto ha hecho posible fusionar la teoría y la práctica; ver [268].

En la literatura sobre control PID ha habido un largo debate, sobre si los métodos de sintonía deberían basarse sobre la respuesta a cambios en el punto de consigna o de perturbaciones de carga. Es sorprendente que una gran cantidad de trabajos sólo muestren la respuesta de la salida del proceso a un cambio en escalón del punto de consigna. Como la regulación en estado estacionario es el problema esencial en control de procesos, las respuestas a perturbaciones en la carga son más importantes que las respuestas a cambios en el punto de consigna como ha sido puesto de manifiesto muchas veces por Shinskey; ver por ejemplo [314]. Una de las conclusiones útiles de la teoría de control robusto es que se requieren seis respuestas para tener una comprensión completa de un sistema en lazo cerrado,

Otra lección de la teoría de control robusto es que la atenuación de las altas frecuencias mejora la robustez. Esto es una buena razón para utilizar un filtraje efectivo en el control PID.

5

Diseño feedforward

5.1 Introducción

Feedforward es una técnica sencilla y potente que complementa a la realimentación. La acción feedforward se puede emplear tanto para mejorar las respuestas a cambios en el punto de consigna como para reducir el efecto de las perturbaciones medibles. El uso del feedforward para mejorar la respuesta a cambios en el punto de consigna ha sido ya analizado en conexión con la ponderación del punto de consigna en la Sección 3.4. Ahora damos un tratamiento sistemático del diseño del control feedforward y también presentamos el diseño de sistemas con seguimiento de modelo. El caso especial de ponderación del punto de consigna se estudiará con cierto detalle, y presentaremos métodos para determinar los pesos del punto de consigna. También mostraremos cómo se puede usar la acción feedforward para reducir el efecto de las perturbaciones que se pueden medir.

5.2 Respuesta mejorada del punto de consigna

Feedforward se puede usar muy efectivamente para mejorar la respuesta del punto de consigna del sistema. Utilizando feedforward es también posible separar el problema de diseño en dos partes. Se diseña primero el controlador por realimentación para lograr robustez y buen rechazo de las perturbaciones y el feedforward se trata a continuación para dar una buena respuesta a cambios en el punto de consigna.

El uso efectivo del feedforward requiere una estructura del sistema que tiene dos grados de libertad. Un ejemplo de tal sistema se muestra en la Figura 3.10. Primero se asume que el sistema tiene la estructura mostrada en la Figura 5.1. Sea el proceso que tiene la función de transferencia $P(s)$. Supongamos que ha sido diseñado un controlador por realimentación $C(s)$, que da buen rechazo de perturbaciones y buena robustez, y consideraremos el problema de diseñar un compensador feedforward que da una buena respuesta a cambios en el punto de consigna.

El compensador feedforward se caracteriza por la función de transferencia $M_u(s)$ y $M_y(s)$, donde $M_y(s)$ da la respuesta del punto de consigna deseado. El

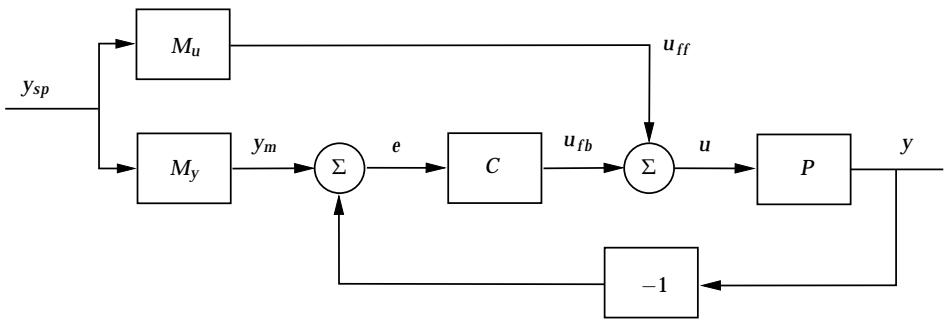


Figura 5.1 Diagrama de bloques de un sistema con dos grados de libertad.

sistema trabaja como sigue. Cuando se cambia el punto de consigna la función de transferencia $M_u(s)$ genera la señal u_{ff} , que da la salida deseada cuando se aplica como entrada al proceso. La salida deseada y_m se genera por $M_y(s)$. Bajo condiciones ideales esta señal es igual a la salida del proceso y . El error de control e es cero, y la señal de realimentación u_{fb} permanece constante. Si hay perturbaciones o errores de modelado las señales y_m y y diferirán. La realimentación intenta traer el error a cero. La función de transferencia del punto de consigna a la salida del proceso es

$$G_{yy_{sp}}(s) = \frac{P(CM_y + M_u)}{1 + PC} = M_y + \frac{PM_u - M_y}{1 + PC}. \quad (5.1)$$

El primer término representa la función de transferencia deseada. El segundo término se puede hacer pequeño de dos maneras. Se puede usar compensación feedforward para hacer $PM_u - M_y$ pequeño, o compensación por realimentación para que el error sea pequeño haciendo la ganancia del lazo PC grande. La condición para feedforward ideal es

$$M_y = PM_u. \quad (5.2)$$

Obsérvese el carácter diferente de la realimentación y del feedforward. Con feedforward se intenta que dos funciones de transferencias coincidan, y con realimentación se pretende hacer el error pequeño dividiendo por un número grande. Con un controlador que tiene acción integral la ganancia del lazo es muy grande para frecuencias pequeñas. Es así suficiente asegurarse de que la condición para feedforward ideal se mantiene en frecuencias más altas. Esto es más fácil que satisfacer la condición (5.2) para todas las frecuencias.

Sistemas inversos

De la ecuación (5.2) se sigue que el compensador feedforward M_u es

$$M_u = P^{-1} M_y, \quad (5.3)$$

lo que significa que contiene una inversa del modelo del proceso P . Un tema clave en el diseño de compensadores feedforward es así encontrar dinámicas inversas. Es fácil calcular formalmente la inversa. Hay, sin embargo, problemas fundamentales graves en la inversión de un sistema, que se ilustra mediante el siguiente ejemplo.

EJEMPLO 5.1—INVERSA DE UN SISTEMA FOTD

El sistema

$$P(s) = \frac{1}{1+sT} e^{-sL}$$

tiene la inversa formal

$$P^{-1}(s) = (1+sT)e^{sL}.$$

Este sistema no es un sistema dinámico causal porque el término e^{sL} representa una predicción. El término $(1+sT)$ requiere una derivada ideal, que también es problemático tal como se analizó en la Sección 3.3. La implementación de la acción feedforward requiere así de aproximaciones. \square

EJEMPLO 5.2—INVERSA DE UN SISTEMA CON CERO EN EL SEMIPLANO DERECHO

El sistema

$$P(s) = \frac{s-1}{s+2}$$

tiene la inversa

$$P^{-1}(s) = \frac{s+2}{s-1}.$$

Obsérvese que esta inversa es un sistema inestable. \square

Se sigue de (5.2) que habrá cancelaciones polos-ceros cuando se diseñe un compensador feedforward. Los polos y ceros cancelados deben ser estables y suficientemente rápidos; en caso contrario, habrá señales en el sistema que crecerán exponencialmente o decaerán muy lentamente.

Las dificultades en el cálculo de inversas pueden evitarse restringiendo la elección de M_y . Como $M_u = P^{-1}M_y$ podemos requerir que la función de transferencia M_y tenga un retardo de tiempo que sea al menos tan grande como el retardo de tiempo de P . Mas aún, M_y y P deben tener los mismos ceros en el semiplano derecho. Para evitar diferenciación, el exceso de polos en M_y debe ser al menos tan grande como el exceso de polos en P . Una posibilidad es aproximar la dinámica del proceso por un modelo sencillo y escoger M_y como un modelo que tiene la misma estructura. Para diseñar un compensador feedforward tenemos así que calcular inversas del sistema aproximadas con propiedades apropiadas.

Inversas aproximadas

En la Sección 2.8 se analizaron diferentes formas de encontrar modelos de procesos aproximados. Ahora veremos un método adicional que se adapta al diseño de control feedforward.

Denotemos por P^\dagger la inversa aproximada de la función de transferencia P . Una aproximación común en control de procesos es despreciar todas las dinámicas y simplemente tomar la inversa de la ganancia estática, i.e.:

$$P^\dagger(s) = P(0)^{-1}.$$

Una serie de resultados sobre inversas de sistemas más precisas ha sido derivada en teoría de sistemas. Algunos de estos serán mostrados aquí. Nótese que la función de transferencia inversa sólo tiene que ser pequeña para aquellas frecuencias donde la función de sensibilidad es grande.

EJEMPLO 5.3—INVERSA APROXIMADA DEL SISTEMA FOTD
El sistema

$$P(s) = \frac{1}{1+sT} e^{-sL}$$

tiene la inversa aproximada

$$P^\dagger(s) = \frac{1+sT}{1+sT/N},$$

donde N da el rango de frecuencias donde la inversión es válida. \square

EJEMPLO 5.4—INVERSA APROXIMADA DE UN SISTEMA CON CERO EN EL SEMIPLANO DERECHO-RHP

El sistema

$$P(s) = \frac{s-1}{s+2}$$

tiene la inversa

$$P^\dagger(s) = \frac{s+2}{s+1}.$$

Obsérvese que el cero inestable en P da lugar a un polo en P^\dagger que es la imagen especular del cero inestable. \square

Un modelo sencillo para sistemas con respuesta monótona a un escalón tiene la función de transferencia

$$P(s) = \frac{K}{(1+sT)^n} e^{-sL}. \quad (5.4)$$

Llamamos a esto el modelo NOTD porque tiene un retardo de tiempo y n constantes de tiempo iguales. Se puede hacer la aproximación ajustando la función de transferencia en algunas frecuencias relevantes. Suponiendo que necesitamos un ajuste perfecto en $\omega = 0$ y $\omega = \omega_0$ encontramos que

$$\begin{aligned} P(0) &= K \\ |P(i\omega_0)| &= \frac{1}{(1+(\omega_0 T)^2)^{n/2}} \\ \arg P(i\omega_0) &= -n \arctan \omega_0 T - \omega_0 L. \end{aligned}$$

Resolviendo las ecuaciones encontramos

$$\begin{aligned} K &= P(0) \\ T &= \frac{\sqrt{|P(i\omega_0)|^{-n/2} - 1}}{\omega_0} \\ L &= -\frac{\arg P(i\omega_0) + n \arctan \omega_0 T}{\omega_0}. \end{aligned} \quad (5.5)$$

Se requiere un buen ajuste en la frecuencia ω_{ms} de sensibilidad máxima. Como esta frecuencia se conoce cuando el controlador de realimentación C ha sido diseñado es natural escoger $\omega_0 = \omega_{ms}$.

Daremos un ejemplo para ilustrar la precisión de la aproximación.

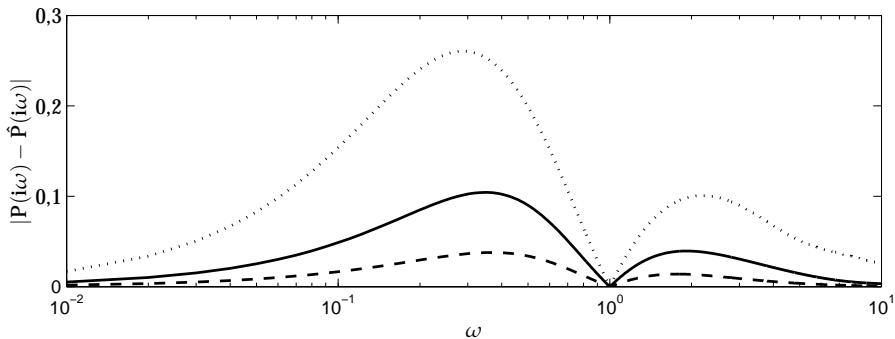


Figura 5.2 Error cuando se ajustan modelos NOTD de diferentes órdenes a la función de transferencia $P(s) = 1/(s+1)^4$ para $n = 1$ (punteada), $n = 2$ (continua), and $n = 3$ (a tramos).

Cuadro 5.1 Parámetros y errores máximos cuando se ajustan modelos NOTD de diferentes órdenes a la función de transferencia $P(s) = 1/(s+1)^4$.

n	ω	K	L	T	e_{\max}	ω_{\max}
1	0,5	1	1,9566	2,4012	0,1828	1,7400
2	0,5	1	1,1352	1,5000	0,0710	1,4500
3	0,5	1	0,5169	1,1773	0,0255	1,3300
1	1,0	1	1,8235	3,8730	0,2603	0,2800
2	1,0	1	1,0472	1,7321	0,1043	0,3400
3	1,0	1	0,4737	1,2328	0,0378	0,3600

EJEMPLO 5.5—CUATRO CONSTANTES DE TIEMPO IGUALES
Considérese un proceso con la función de transferencia

$$P(s) = \frac{1}{(s+1)^4}.$$

En la Figura 5.2 mostramos el error $|P(i\omega) - \hat{P}(i\omega)|$ para modelos NOTD con diferentes n , y en la Tabla 5.1 damos los parámetros y el error máximo para diferentes ajustes. Obsérvese que se obtienen errores relativamente grandes, del 20 al 30 por ciento, para modelos de primer orden, y se consiguen reducciones significativas aumentando el orden del modelo. \square

Para un proceso dado por (5.4) es razonable escoger el modelo de la respuesta como

$$M_y = \frac{1}{(1+sT_m)^n} e^{-sL}.$$

Entonces se sigue de (5.2) que el compensador feedforward está dado por

$$M_u = \frac{1}{K} \left(\frac{1+sT}{1+sT_m} \right)^n. \quad (5.6)$$

En este caso particular el compensador feedforward consiste así de un modelo del proceso y una red de adelanto-retardo o retardo-adelanto.

Hay situaciones donde se desea que un lazo de realimentación debería tener una respuesta al punto de consigna con un tiempo de respuesta especificado. Un caso típico es cuando se mezclan algunas sustancias que vienen de fuentes diferentes. Cuando se hacen cambios de producción es muy deseable que todos los sistemas reaccionen a los cambios de producción de la misma forma. Es muy fácil conseguir esto cuando la dinámica del proceso requerida es lenta en comparación con el ancho de banda de la realimentación, porque se sigue de (5.1) que la respuesta del punto de consigna no es muy sensible al modelo del proceso. Ilustramos esto con un ejemplo.

EJEMPLO 5.6—RESPUESTA LENTA DEL PUNTO DE CONSIGNA

Sea un proceso con la función de transferencia

$$P(s) = \frac{1}{(s+1)^4},$$

controlada con un controlador PI con $K = 0,775$ y $T_i = 2,05$. Esto da $M_s = 2$ y $\omega_{ms} = 0,559$. Aproximando el modelo del proceso con un modelo de primer orden FOTD da los parámetros $K_p = 1$, $T = 2,51$, y $L = 1,94$, ver (5.5). Supóngase que la respuesta deseada del punto de consigna está dada por

$$M_y(s) = \frac{1}{1 + sT_r}.$$

La Figura 5.3 muestra las respuestas del punto de consigna para diferentes valores de T_r . La figura muestra que la respuesta a las perturbaciones de carga es la misma en ambos casos y que la respuesta del punto de consigna tiene la conducta esperada. Obsérvese las distorsiones de las curvas para $T_r = 2$; se deben al hecho de que el modelo no se ajusta tan bien para altas frecuencias. Una regla heurística es que el modelo de primer orden es razonable para $\omega_{ms} T_r > 2$. En este caso esto da $T_r > 3,6$. Se requieren modelos más precisos para obtener la conducta deseada para $T_r = 2$. \square

La ventaja de utilizar un controlador con dos grados de libertad es que la buena atenuación de la perturbación se puede mantener mientras que se hace más lenta la respuesta del punto de consigna.

5.3 Ponderación del punto de consigna

Para controladores PID sencillos puede no ser necesario usar un sistema completo con dos grados de libertad. La respuesta deseada del punto de consigna se puede a menudo mantener simplemente ajustando los pesos del punto de consigna; ver Sección 3.4. Para determinar los pesos del punto de consigna consideramos la función de transferencia desde el punto de consigna a la salida del proceso, y escogemos los parámetros del punto de consigna de forma que la ganancia más grande de esta función de transferencia sea uno o próximo a uno. Esto da una

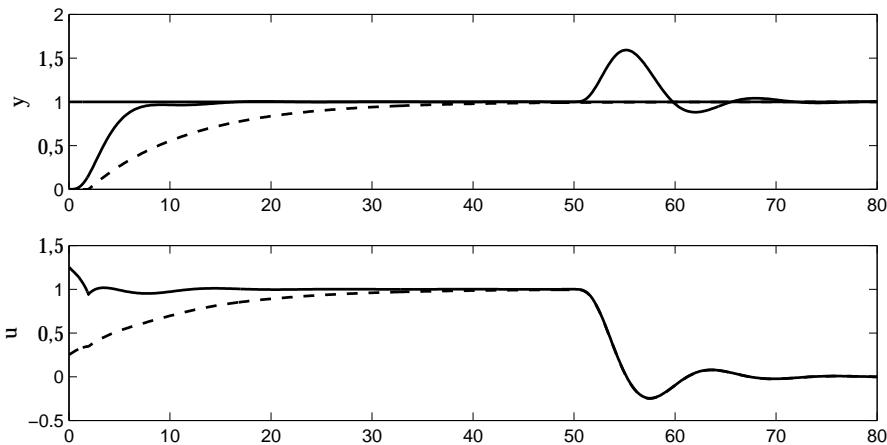


Figura 5.3 Respuestas del punto de consigna y y de las perturbaciones de carga del proceso $P(s) = 1/(s+1)^4$ con un controlador PI y un compensador feedforward basado en el modelo FOTD para los tiempos de respuestas deseadas $T_r = 10$ (línea a tramos) y $T_r = 2$ (línea continua).

respuesta del punto de consigna sin sobreelongación para la mayoría de los sistemas.

Se sigue de la Figura 3.10 y de la Ecuación 3.20 que la función de transferencia del punto de consigna a la salida del proceso es

$$G_{yy_{sp}}(s) = \frac{k_i + bks + ck_d s^2}{k_i + ks + k_d s^2} \frac{P(s)C(s)}{1 + P(s)C(s)} = \frac{k_i + bks + ck_d s^2}{k_i + ks + k_d s^2} T(s). \quad (5.7)$$

Una posibilidad para conseguir la ganancia más grande próxima a uno es especificar que la sensibilidad máxima M_t está próxima a uno. En tal caso puede no ser necesario usar la ponderación del punto de consigna. Para diseños con valores más grandes de M_t podemos simplemente calcular el máximo de $|G_{yy_{sp}}(i\omega)|$ y ajustar los valores de b y c que da un valor próximo a uno. El peso c se fija con frecuencia a cero. En ese caso, sólo hay un parámetro para escoger. Si $|G_{yy_{sp}}(i\omega)|$ es mayor que uno para $b = 0$, se puede usar un filtrado paso baja del punto de consigna para reducir más aún la magnitud de $|G_{yy_{sp}}(i\omega)|$. El filtro del punto de consigna $F_{sp}(s)$ se puede determinar de la forma siguiente: sea m_s el máximo de la función de transferencia (5.7) con $b = c = 0$, y sea ω_{sp} la frecuencia donde ocurre el máximo. Un filtro de primer orden

$$F_{sp} = \frac{1}{1 + sT_{sp}},$$

tiene la magnitud $1/m_s$ en la frecuencia ω_{sp} si la constante de tiempo es

$$T_{sp} = \frac{1}{\omega_{sp}} \sqrt{m_s^2 - 1}.$$

Alimentando el punto de consigna a través de un filtro paso baja diseñado de esta forma se reducirá la magnitud en la frecuencia ω_{sp} a uno.

Una desventaja con la ponderación del punto de consigna y el filtrado es que la respuesta del punto de consigna puede ser innecesariamente lenta.

5.4 Feedforward neutro

Una elección muy simple de control feedforward para sistemas con respuestas a un escalón monótonas que satisfacen (5.2) están dadas por

$$\begin{aligned} M_y &= \frac{P}{P(0)} = \frac{P}{K_p} \\ M_u &= \frac{1}{K_p}. \end{aligned} \tag{5.8}$$

Esto significa que la respuesta deseada del punto de consigna es la respuesta normalizada del sistema en lazo abierto. Como $M_u = 1/K_p$ la señal de control es proporcional al punto de consigna. A un cambio en escalón en el punto de consigna la señal de control cambia también en escalón al valor constante que da el estado estacionario deseado, y permanece en ese valor. El diseño de una acción feedforward neutra es así muy simple.

Se puede sustituir un modelo de proceso complicado por un modelo aproximado. Para control PID es natural basar el diseño de la compensación feedforward en el modelo NOTD. Una forma de determinar los parámetros apropiados es ajustar el modelo en la frecuencia ω_{ms} donde la función de sensibilidad tiene su valor más grande. Ilustramos el procedimiento de diseño con un ejemplo.

EJEMPLO 5.7—CUATRO CONSTANTES DE TIEMPO IGUALES

Sea un proceso con la función de transferencia

$$P(s) = \frac{1}{(s+1)^4}.$$

Un controlador PI con una especificación de $M_s = 2$ para este sistema da los parámetros $K = 0,775$, $T_i = 2,05$, y $\omega_{ms} = 0,56$ de un modelo aproximado. La Ecuación 5.5 da los parámetros $K = 1$, $L = 1,94$, y $T = 2,50$. La Figura 5.4 muestra la respuesta del sistema a un escalón y a perturbaciones en la carga. Obsérvese que hay una depresión en la señal de control alrededor del instante de tiempo $t = 2$. La razón está en la diferencia entre el proceso y el modelo usado para diseñar el compensador feedforward. Esto se ilustra en la Figura 5.5, que muestra las respuestas del proceso y el modelo. Nótese que el proceso responde más rápido que el modelo inicialmente. Hay entonces un error, que se compensa por la realimentación.

Se puede mejorar la respuesta del punto de consigna utilizando una aproximación mejor del modelo del proceso. Una posibilidad es ajustar un modelo NOTD de segundo orden. Tal modelo tiene los parámetros $K = 1$, $T = 1,52$, y $L = 1,13$. La Figura 5.6 muestra las respuestas del sistema a un escalón y a

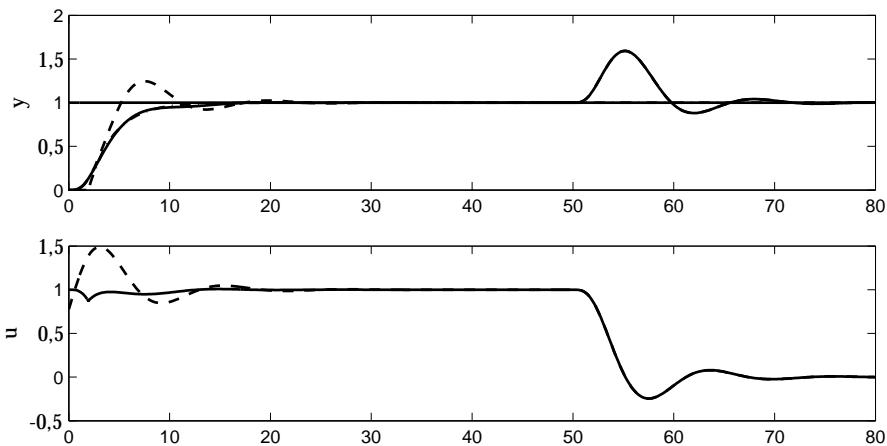


Figura 5.4 Respuestas del punto de consigna y y perturbaciones de carga del proceso $P(s) = 1/(s+1)^4$ con un controlador PI (línea a tramos) y compensación feedforward basada en el modelo FOTD (línea continua).

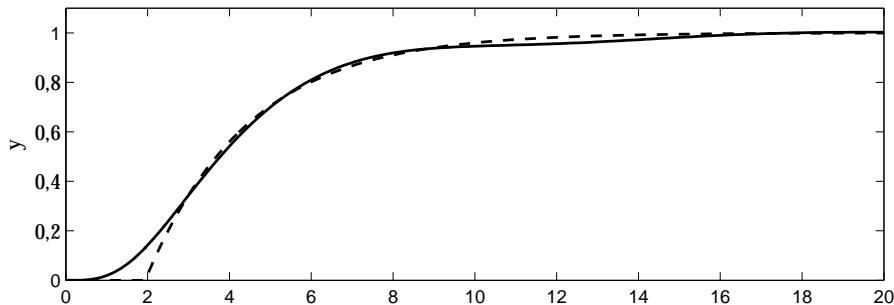


Figura 5.5 Respuestas a un escalón del proceso P (línea continua) y el modelo usado para diseñar el compensador feedforward (línea a tramos).

perturbaciones en la carga. Comparada con la Figura 5.4 la señal de control está más próxima al valor ideal $u = 1$ y la respuesta del punto de consigna es un poco mejor. La Figura 5.7 muestra la comparación de la salida del modelo y_m y la salida del proceso. Una comparación con la Figura 5.5 muestra que el modelo de segundo orden da un mejor ajuste. Una comparación de la Figura 5.4 con la Figura 5.6 también ilustra que la compensación feedforward requiere de un buen modelado. \square

En el control de temperatura es a menudo deseable tener un controlador sin sobreelongación frente a respuestas en escalón. El próximo ejemplo ilustra como se puede emplear la compensación feedforward para conseguirlo.

EJEMPLO 5.8—CONSTANTES DE TIEMPO DISTRIBUIDAS

Sea un proceso con la función de transferencia

$$P(s) = \frac{1}{\cosh \sqrt{s}}.$$

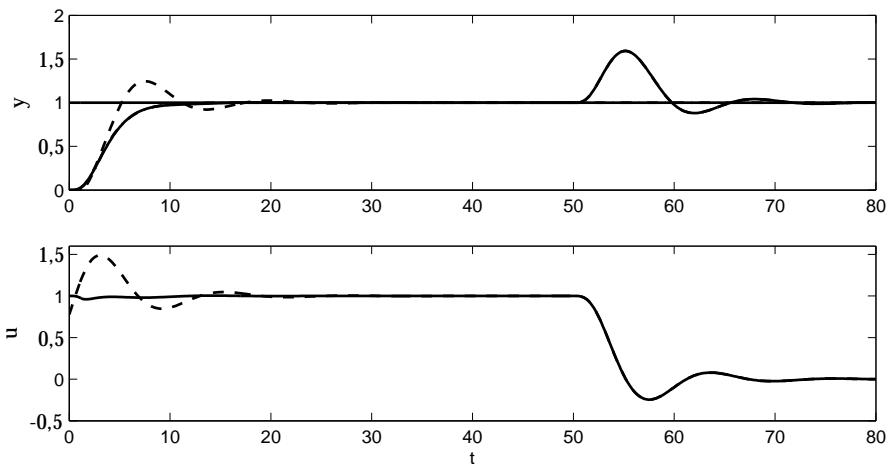


Figura 5.6 Respuestas a punto de consigna y perturbaciones de carga del proceso $P(s) = 1/(s+1)^4$ con un controlador PI (línea a tramos) y compensación feedforward basada en un modelo SOTD (línea continua).

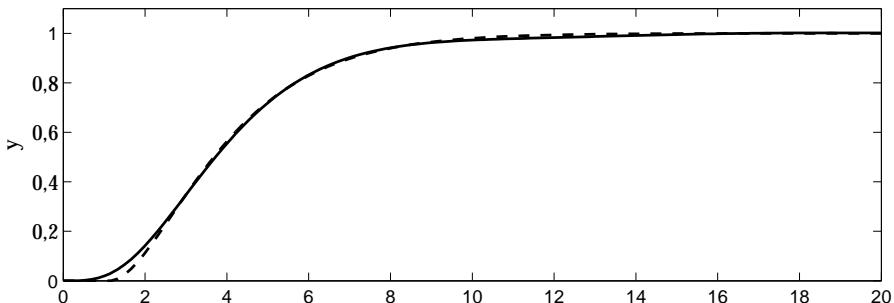


Figura 5.7 Respuestas a un escalón del proceso P (línea continua) y el modelo usado para diseñar la compensación feedforward (dilínea a tramos).

Un controlador PI agresivo con $M_s = 2$ tiene $K = 2,66$, $T_i = 0,197$, y $\omega_{ms} = 9,68$. Incluso con $b = 0$ este controlador da una sobreelongación tal como se muestra por la curva a tramos en la Figura 5.8. Ajustando un modelo FOTD en las frecuencias 0 y ω_{ms} da $K = 1$, $T = 0,408$, y $L = 0,0917$. El error en la función de transferencia es menor que el 5 por ciento. La Figura 5.8 muestra una simulación del sistema con compensación feedforward neutra basada en ese modelo. La figura muestra que la compensación feedforward neutra consigue la respuesta deseada. \square

Sistema oscilatorio

El control PID no es la mejor estrategia para sistemas oscilatorios ya que se puede obtener mucho mejor comportamiento con controladores más complejos. Sin embargo, el control PID se emplea algunas veces para tales sistemas, y el compor-

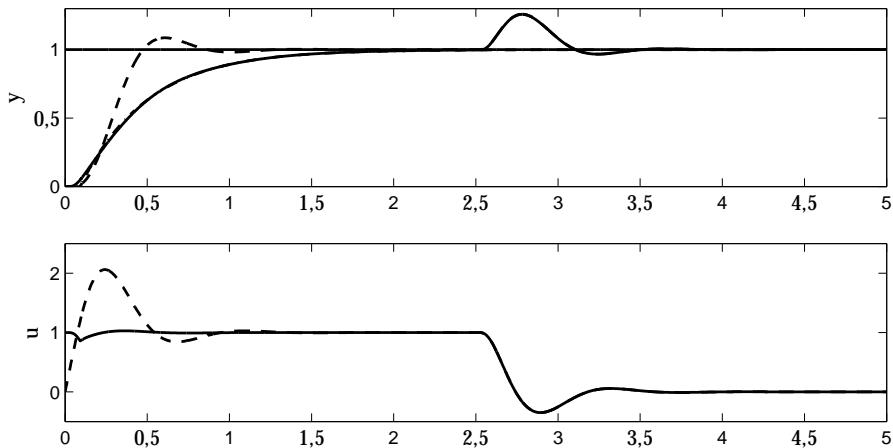


Figura 5.8 Respuestas a puntos de consigna y perturbaciones de carga del proceso $P(s) = 1/\cosh \sqrt{s}$ con un controlador PI (líneas a tramos) y un compensador feedforward neutro basado en un modelo FOTD de primer orden (líneas continuas).

tamiento de un controlador PID convencional puede a menudo mejorarse mediante compensación feedforward. Sin embargo, la compensación feedforward neutra, que da una respuesta similar al sistema no controlado no se puede usar porque dará una respuesta que es demasiado oscilatoria. Ilustraremos como se puede emplear la compensación feedforward mediante un ejemplo.

EJEMPLO 5.9—SISTEMA OSCILATORIO

Sea un sistema con la función de transferencia

$$P(s) = \frac{9}{(s+1)(s^2 + 0,1s + 9)}.$$

El modo oscilatorio tiene un amortiguamiento relativo $\zeta = 0,03$, que es bastante bajo.

Parámetros razonables para un controlador PI para el sistema son $K = -0,167$ y $T_i = -0,210$. Como el controlador tiene ganancia negativa, se debe utilizar ponderación del punto de consigna con $b = 0$ para obtener una respuesta razonable. La sobreelongación es, no obstante, todavía sustancial como se ve por la curva a tramos de la Figura 5.9. Para diseñar la compensación feedforward escogemos una respuesta deseada dada por la función de transferencia

$$M_y = \frac{9}{(s+1)(s^2 + 6s + 9)}.$$

La dinámica de este sistema es esencialmente la misma que para el proceso, pero los polos complejos ahora tienen amortiguamiento crítico. Se sigue de (5.2) que

$$M_u = \frac{s^2 + 0,1s + 9}{s^2 + 6s + 9}.$$

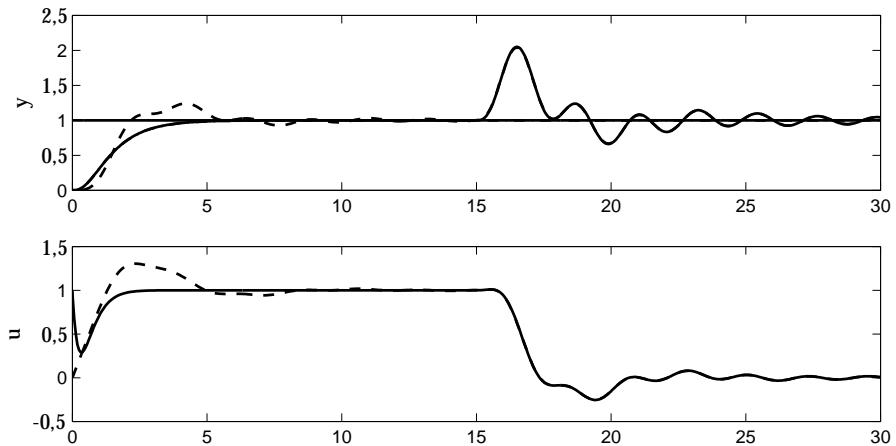


Figura 5.9 Respuestas a punto de consigna y perturbaciones de carga del proceso $P(s) = 9/(s+1)(s^2 + 0,2s + 9)$ con un controlador PI (líneas a tramos) y una compensación feedforward (líneas continuas).

Esta función de transferencia es próxima a uno para todas las frecuencias excepto aquellas que corresponden a los modos oscilatorios donde tiene baja ganancia. La función de transferencia así bloquea las señales que pueden excitar a los modos oscilatorios. La Figura 5.9 muestra la respuesta del sistema a puntos de consigna y perturbaciones de carga. Es claro que se mejora sustancialmente la respuesta del punto de consigna mediante el uso de la compensación feedforward. La respuesta de la perturbación de carga es todavía bastante pobre, lo cual refleja el hecho de que el control PI no es apropiado para sistemas altamente oscilatorios. \square

5.5 Respuesta rápida del punto de consigna

Con compensación feedforward neutra no hay sobreelongación en la señal de control. Es posible obtener repuestas más agresivas si permitimos que la señal de control tenga sobreelongación. Esto se logra simplemente requiriendo una respuesta más rápida. Para hacer esto el modelo debe ser también preciso sobre un rango más amplio de frecuencias. Sin embargo la sobreelongación en la señal de control puede aumentar muy rápidamente lo que aumenta el tiempo de respuesta tal como se ilustra en el ejemplo siguiente.

EJEMPLO 5.10—RESPUESTA RÁPIDA DEL PUNTO DE CONSIGNA

Sea el sistema

$$P(s) = \frac{1}{(s+1)^4}.$$

Suponga que se desea tener una respuesta del punto de consigna dada por

$$M_y = \frac{1}{(sT_m + 1)^4}.$$

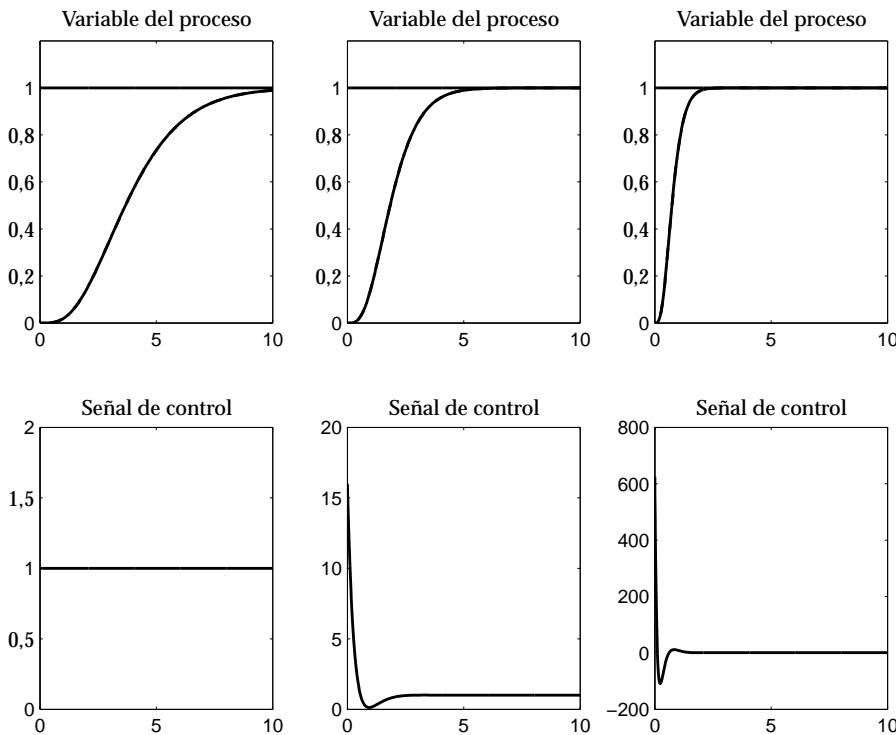


Figura 5.10 Respuestas del punto de consigna del proceso $P(s) = (s + 1)^{-4}$ con compensadores feedforward diseñados para dar $M_u(s) = (sT_m + 1)^{-4}$ para $T_m = 1$ (izquierda), $0,5$ y $0,2$ (derecha).

Se sigue de (5.3) que

$$M_u = \frac{(s + 1)^4}{(sT_m + 1)^4}.$$

Para feedforward neutro tenemos $T_m = 1$, que da $M_u = 1$. En general, tenemos

$$M_u(\infty) = T_m^{-4}.$$

La ganancia del controlador así aumenta muy rápidamente con valores decrecientes de T_m . Esto se ilustra en la simulación que se muestra en Figura 5.10, que muestra la respuesta para $T_m = 1$ (feedforward neutro), $T_m = 0,5$, y $T_m = 0,2$. Los valores iniciales de la señal de control son 1 , 16 , y 625 , respectivamente.

Nótese que la potencia 4 en las expresiones se debe al hecho de que el proceso tiene un exceso de polos de 4. En la práctica, la saturación del actuador determina lo que se puede lograr. \square

Control en tiempo óptimo

El ejemplo claramente ilustra que se puede usar la compensación feedforward para obtener respuestas rápidas del punto de consigna pero que ello requiere de

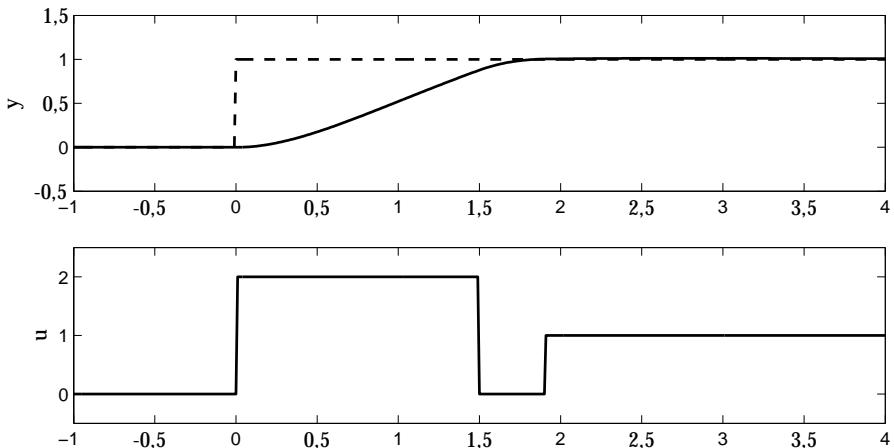


Figura 5.11 Cambio del punto de consigna en tiempo óptimo para el proceso $P = (s + 1)^{-2}$.

modelos que sean válidos sobre un rango amplio de frecuencias y que se puedan necesitar señales de control muy grandes. El tamaño de la señal de control depende críticamente del exceso de polos del proceso. En la práctica, es también necesario tener en cuenta que las señales de control tienen rango limitado. Es, por lo tanto, muy natural buscar estrategias que lleven a la salida del proceso de un punto de consigna al otro en tiempo mínimo. Este problema se resuelve por teoría de control óptimo. Se sabe que para sistemas lineales la solución es un control bang-bang lo que significa que la señal de control comuta entre sus valores extremos. En la Figura 5.11 se da un ejemplo, que muestra la solución de tiempo mínimo para el proceso $P = (s + 1)^{-2}$ cuando la señal de control está limitada a valores entre 0 y 2. En este caso el control es muy simple. Sin embargo, puede haber un gran número de comutaciones para sistemas de orden superior o para sistemas oscilatorios. A causa de su complejidad no es viable emplear control óptimo excepto en situaciones muy especiales. Por lo tanto se desarrollarán métodos aproximados.

Control con pulsos en escalón

Para sistemas estables con respuestas monótonas a un escalón a menudo se pueden obtener respuestas rápidas del punto de consigna con señales de control que tienen la forma mostrada en la Figura 5.11. Esto significa que inicialmente se usa la máxima señal de control. La señal de control se comuta entonces a su valor más bajo, y finalmente se da la señal de control que da el estado estacionario deseado. Si el pulso inicial se aproxima mediante un impulso obtenemos la situación que se muestra en la Figura 5.12. Suponiendo que el sistema está inicialmente en reposo su salida está entonces dada por

$$y(t) = ag(t) + bh(t - L),$$

donde h es la respuesta en escalón y g la respuesta impulsional del sistema. Los parámetros a , b , y L deberían escogerse de forma que la respuesta coincida con la

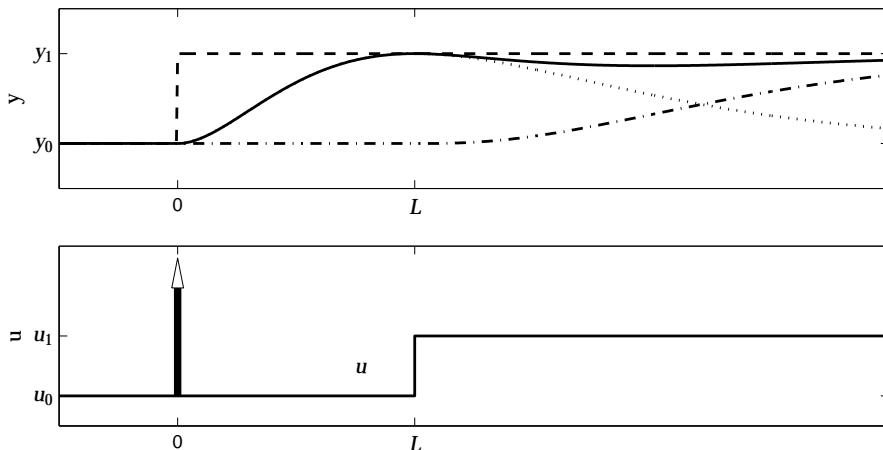


Figura 5.12 Respuesta a un impulso (línea punteada) y a un escalón retardado (línea de punto y raya) para un sistema con respuesta monótona a un escalón. La línea a tramos es el punto de consigna, y la línea continua es la salida del proceso, compuesta por la suma de las respuestas a un escalón y un impulso.

respuesta deseada tanto como sea posible. Para hacer esto el parámetro a debería escogerse como $y_{sp}/g_{máx}$, donde $g_{máx}$ es el máximo de la respuesta impulsional $g(t)$. El parámetro b debería seleccionarse de forma que se obtenga el estado estacionario deseado. Por tanto, $b = y_{sp}/K_p$ donde K_p es la ganancia del proceso en estado estacionario. El parámetro L debería ajustarse para mantener la salida tan próxima como sea posible al punto de consigna. Estas elecciones implican que el tiempo de asentamiento del sistema es igual al tiempo donde la respuesta impulsional alcanza su máximo. El tiempo de asentamiento en lazo cerrado es así coincidente con el tiempo de respuesta natural del sistema. No es por supuesto posible tener un impulso como una entrada. El impulso se aproxima por lo tanto por un pulso con una amplitud que corresponde al valor máximo de la señal de control. La duración se escoge de forma que el área bajo el pulso es igual a a . Se puede hacer una sintonía final de los parámetros dados anteriormente por optimización. Ilustramos el procedimiento por un ejemplo.

EJEMPLO 5.11—RESPUESTA RÁPIDA A CAMBIOS EN EL PUNTO DE CONSIGNA

Sea un sistema con la función de transferencia

$$P(s) = \frac{1}{(s+1)^4}.$$

La Figura 5.13 compara la respuesta rápida del punto de consigna con control PI regular con dos parámetros de ajuste. La respuesta rápida del punto de consigna se ha calculado con $u_{máx} = 4$ y $u_{mín} = -4$, y el tiempo de subida resultante y el tiempo de asentamiento son aproximadamente 4 unidades de tiempo. Los controladores han sido diseñados con moldeo del lazo para sensibilidades máximas $M_s = 1,4$ y $M_s = 2,0$. Los parámetros del controlador correspondiente son $K = 0,43$, $T_i = 2,25$, y $b = 1$ para $M_s = 1,4$, y $K = 0,78$, $T_i = 2,05$, y $b = 0,23$

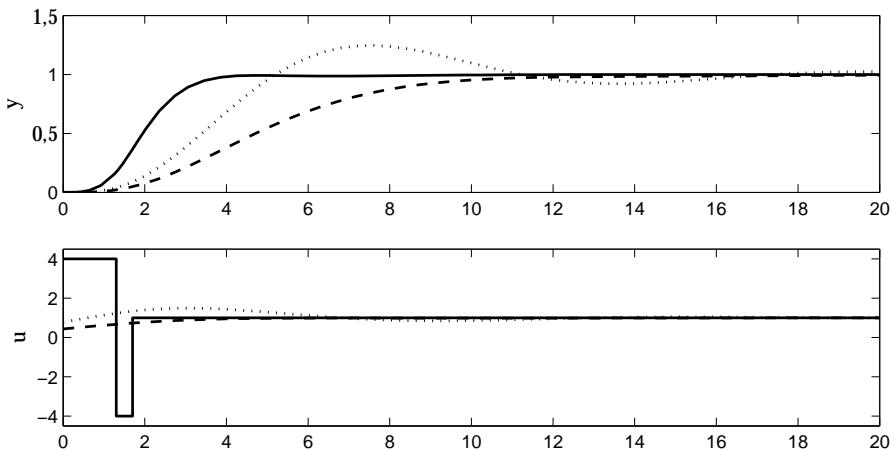


Figura 5.13 Comparación entre la estrategia de la respuesta del punto de consigna rápido (continua) y control PI con $M_s = 1,4$ (a tramos) y $M_s = 2,0$ (punteada) para $P(s) = 1/(s + 1)^4$.

para $M_s = 2,0$. Ambos diseños PI son claramente superados por el método del pulso-escalón. El tiempo de subida es un factor 2–3 más grande, y el tiempo de asentamiento aproximadamente 3 veces mayor. La razón es, por supuesto, que se utiliza mucho menos la autoridad de control que se tiene disponible. Si se aumenta el peso del punto de consigna b y/o M_s , aumentará el tamaño de la señal de control. Esto conduce a un tiempo de subida más rápido, pero a expensas de una mayor sobreelongación, de forma que el tiempo de asentamiento puede realmente ser incluso más alto. \square

Limitaciones de velocidad

La idea de la respuesta rápida del punto de consigna puede también aplicarse al caso cuando hay limitaciones de velocidad. Esto se ilustra en la Figura 5.14, que muestra una simulación del proceso con la función de transferencia $P = 1/(s + 1)^4$ cuando hay limitaciones de velocidad $|du/dt| < 2$. Es también posible combinar limitaciones de velocidad y nivel.

5.6 Atenuación de la perturbación

Las perturbaciones se pueden eliminar por realimentación. Con un sistema realimentado es, sin embargo, necesario que haya un error antes de que el controlador pueda tomar acciones para eliminar las perturbaciones. En algunas situaciones, es posible medir perturbaciones antes de que hayan influido en el proceso. Es entonces natural intentar eliminar los efectos de las perturbaciones antes de que se hayan creado errores de control. Este paradigma de control se llama *feedforward*. El principio se ilustra en la Figura 5.15.

En la Figura 5.15 la función de transferencia del proceso P se compone de dos factores, $P = P_1 P_2$. Una perturbación medida d entra en la sección del proceso P_2 .

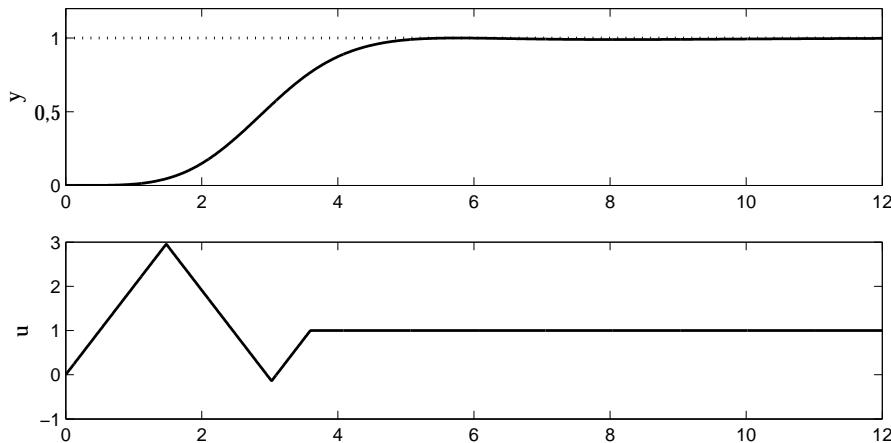


Figura 5.14 La salida del proceso y y la señal de control para cambios rápidos del punto de consigna con limitaciones de velocidad $|du/dt| < 2$ para el proceso $P(s) = 1/(s + 1)^4$.

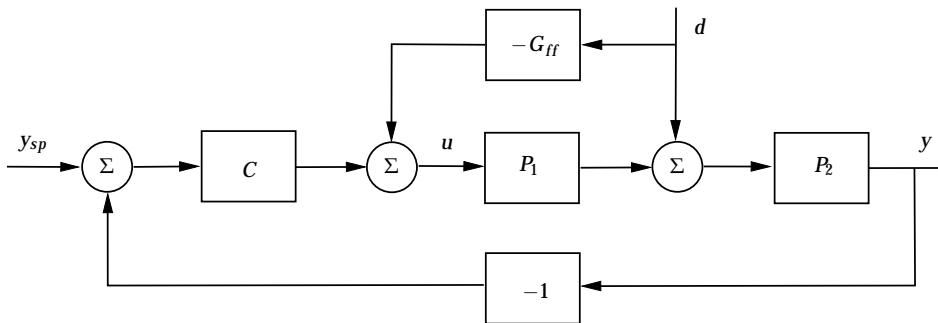


Figura 5.15 Diagrama de bloques de un sistema donde una perturbación medida d se reduce por una combinación de realimentación y feedforward.

La perturbación medida se alimenta en la entrada del proceso vía la función de transferencia feedforward G_{ff} . La función de transferencia de la perturbación de carga a la salida del proceso es

$$G_{yd}(s) = \frac{P_2(1 - P_1 G_{ff})}{1 + PC} = P_2(1 - P_1 G_{ff})S, \quad (5.9)$$

donde $S = 1/(1 + PC)$ es la función de sensibilidad. Esta ecuación muestra que hay dos formas de reducir la perturbación. Podemos intentar hacer $1 - P_1 G_{ff}$ pequeño por una elección adecuada de la función de transferencia feedforward G_{ff} , o podemos hacer la función de transferencia del lazo PC grande por realimentación. Feedforward y realimentación se pueden también combinar.

Obsérvese que con la acción feedforward estamos intentando hacer la diferencia entre dos términos pequeños, pero con realimentación simplemente multiplicamos por un pequeño número. Una consecuencia inmediata es que la acción feedforward es más sensible que la realimentación. Con realimentación hay

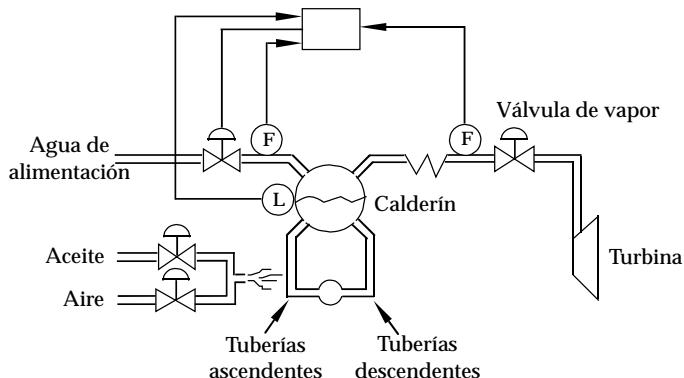


Figura 5.16 Diagrama esquemático de un calderín con control de nivel.

un riesgo de inestabilidad; no hay tal riesgo con la acción feedforward. Realimentación y feedforward son por lo tanto complementarias, y es útil combinarlas.

Un compensador feedforward ideal está dado por

$$G_{ff} = P_1^{-1} = \frac{P_{yd}}{P_{yu}}, \quad (5.10)$$

donde P_{yd} es la función de transferencia de d a y y $P_{yu} = P$ es la función de transferencia de u a y . El compensador feedforward ideal se forma al tomar la inversa de la dinámica del proceso P_1 . Esta inversa con frecuencia no es realizable, por lo que tienen que usarse aproximaciones.

Feedforward es más efectivo cuando la perturbación d entra pronto en el proceso. Esto ocurre cuando la mayoría de la dinámica está en la sección del proceso P_2 . Cuando $P_1 = 1$, y por lo tanto $P_2 = P$, el compensador feedforward ideal es por lo tanto realizable, y los efectos de la perturbación se pueden eliminar de la salida del proceso y . Por otra parte, cuando la dinámica entra tarde en el proceso, de forma que $P_1 \approx P$, los efectos de la perturbación se ven en la salida del proceso y en el mismo tiempo que se ven en la señal feedforward. En este caso, no hay ventaja en utilizar feedforward comparado con la realimentación.

Aplicaciones

En muchas aplicaciones de control hay algunos procesos en serie. En tales casos, es a menudo fácil medir perturbaciones y usar feedforward. Aplicaciones típicas del control feedforward son el control del nivel del calderín en calderas de vapor, el control de columnas de destilación y los trenes de laminación. A continuación se da una aplicación de realimentación y feedforward combinado.

EJEMPLO 5.12—CONTROL DEL NIVEL DEL CALDERÍN

En la Figura 5.16 se muestra un diagrama simplificado de una caldera de vapor. El agua en los tubos ascendentes se calienta por los quemadores. El vapor generado en los tubos ascendentes, que es más ligero que el agua, se eleva hacia el calderín. Esto origina una circulación alrededor del lazo consistente en los tubos

ascendentes, el calderín, y los tubos descendentes. El vapor se separa del agua del calderín. El caudal de vapor que va a la turbina se controla por la válvula de admisión de vapor.

Es importante mantener constante el nivel del agua en el calderín. Un nivel de agua demasiado bajo da un enfriamiento insuficiente de los tubos ascendentes, y hay un riesgo de que se quemen. Con un nivel de agua demasiado alto, se puede introducir agua en las turbinas, lo cual puede originar un daño. Hay un sistema de control para mantener constante el nivel. El problema de control es difícil a causa del denominado *efecto de contracción y expansión*. Se puede explicar como sigue. Suponga que el sistema está en equilibrio con un nivel constante del calderín. Si se aumenta el caudal de vapor abriendo la válvula de la turbina, la presión en el calderín decaerá. La disminución de presión origina la generación de burbujas extra en el calderín y en los tubos ascendentes. Como resultado de ello, el nivel del calderín inicialmente aumentará. Como se extrae más vapor del calderín, su nivel finalmente disminuirá. Este fenómeno, que se llama el *efecto de contracción y expansión*, origina graves dificultades en el control del nivel del calderín. Matemáticamente, también da lugar en la función de transferencia a un cero en el semiplano derecho.

El problema se puede resolver introduciendo la estrategia de control que se muestra en la Figura 5.16. Consiste en una combinación de realimentación y feed-forward. Hay una realimentación del nivel del calderín al controlador, pero hay también una acción feedforward de la diferencia entre el caudal del vapor y el caudal de alimentación de agua de forma que se consigue que ambos se equilibren rápidamente. □

5.7 Resumen

En este capítulo se ha presentado el diseño de la compensación feedforward. La acción feedforward se puede usar para reducir el efecto de las perturbaciones medibles. El diseño del feedforward es esencialmente una cuestión de encontrar modelos inversos del proceso. Para hacer esto han sido analizadas diferentes técnicas. La mayor parte del capítulo se ha dedicado a la respuesta del punto de consigna. Se ha utilizado una estructura con dos grados de libertad. Esto proporciona una separación clara de la regulación y la respuesta del punto de consigna y de la realimentación y de la acción feedforward. Se ha supuesto que previamente ha sido diseñado el controlador por realimentación. Una forma sencilla de modificar la respuesta del punto de consigna es usar una ponderación del mismo. Si no se pueden obtener los resultados deseados por ponderar a cero el punto de consigna se puede emplear un conjunto completo de esquemas con dos grados de libertad. Esto posibilita el hacer una separación completa entre la respuesta de la perturbación de carga y la respuesta del punto de consigna. El tema crucial de diseño es decidir la velocidad de respuesta que es posible alcanzar. Para sistemas con respuestas monótonas al punto de consigna ha sido propuesta la noción de compensación feedforward neutro. Han sido también analizadas otras variantes. Finalmente ha sido demostrado que se pueden obtener respuestas rápidas del

punto de consigna utilizando métodos no lineales.

Se deben tomar cuidados especiales cuando se implementa control feedforward, en caso contrario puede ocurrir la saturación del término integral (windup). La implementación de control feedforward se analiza en la Sección 13.4.

5.8 Notas y referencias

La compensación feedforward es un complemento útil de la realimentación. Se utilizó en amplificadores electrónicos incluso antes de que emergiera el amplificador realimentado tal como se analiza en [51]. La utilización de la acción feedforward en control de procesos fue introducida en [307]. La efectividad del feedforward para mejorar la respuesta del punto de consigna utilizando una estructura del sistema con dos grados de libertad (2DOF) se introdujo en [176]. La ponderación del punto de consigna, que es una forma sencilla de 2DOF, ha sido utilizada en una extensión limitada en los primeros controladores PID donde los pesos han sido 0 o 1. La utilización de pesos continuamente ajustables apareció en los 1980. El empleo de la acción feedforward para reducir el efecto de las perturbaciones medidas es complicada de aplicar en los sistemas de control de procesos construidos sobre componentes separadas pero muy fácil en los modernos sistemas de control distribuido; ver [49] y [1]. Las aplicaciones de la compensación feedforward están ganando en popularidad. Están también emergiendo métodos para evaluar las mejoras potenciales que se consiguen al utilizar feedforward; ver [278, 280, 279].

6

Diseño PID

6.1 Introducción

Este capítulo describe métodos para encontrar parámetros de un controlador PID, que es un caso especial del problema de diseño de sistemas de control que fue analizado en el Capítulo 4. El diseño de controladores PID difiere del problema de diseño general porque se restringe la complejidad del controlador. Los métodos de diseño generales dan un controlador con una complejidad que coincide con el modelo del proceso. Para obtener un controlador con una complejidad restringida podemos o simplificar el modelo del proceso de forma que el diseño da un controlador PID, o podemos diseñar un controlador para un modelo complejo y aproximar con un controlador PID. Otra razón de por qué emergieron métodos de diseño especiales para controladores PID es el deseo de tener métodos de diseño sencillos que se puedan emplear por personas con pocos conocimientos sobre control. La situación ha cambiado sustancialmente con la aparición de herramientas de sintonía y sintonizadores automáticos, que han hecho posible mejorar el conocimiento del proceso y permitido el uso de cálculos más extensivos. Esto ha llevado el diseño de los controladores PID más próximo a la corriente principal del diseño de sistemas de control.

En este capítulo se ha intentado mantener un equilibrio al proporcionar una perspectiva histórica y presentar métodos potentes de diseño. La Sección 6.2 describe los métodos desarrollados por Ziegler y Nichols, que han tenido un gran impacto sobre la práctica del control PID incluso aunque no den como resultado buenas sintonías. También se analizan algunas extensiones de los métodos de Ziegler-Nichols.

Con frecuencia es necesario complementar los métodos de diseño con una sintonía fina manual para obtener los objetivos deseados de la dinámica en lazo cerrado. Estas reglas de sintonía manual se analizan en la Sección 6.3.

La Sección 6.4 presenta el método de asignación de polos, que es una de las corrientes principales de métodos en el diseño de sistemas de control. Para aplicar este método es necesario aproximar la dinámica del proceso por un modelo de primer orden para control PI y un modelo de segundo orden para control PID. En lugar de intentar posicionar todos los polos en lazo cerrado, se puede intentar asignar solamente unos pocos polos dominantes. Tales métodos se presentan

en la Sección 6.4. El método de diseño de asignación de polos dominantes es el método de la sintonía lambda, presentado en la Sección 6.5.

En la Sección 6.6, se presentan métodos de sintonía algebraicos. En estos métodos, los parámetros del controlador se obtienen de las especificaciones mediante un cálculo algebraico directo. En estas técnicas es también necesario aproximar la dinámica del proceso por modelos de bajo orden.

Muchas técnicas para diseño de sistemas de control se basan en optimización. Esto da una forma muy flexible de equilibrar criterios de diseños que están en conflicto. Es también posible aplicarlo directamente a controladores que tienen complejidad restringida. En la Sección 6.7 se exponen una serie de usos de la optimización para control PID.

El moldeo del lazo es otra técnica bien conocida para el diseño de sistemas de control. En la Sección 6.8 se muestra como se puede utilizar esto para control PID. Esto da un método de diseño muy flexible, que permite un compromiso agradable entre comportamiento y robustez. Un análisis del método también da una comprensión de las dificultades con la acción derivativa.

En las Secciones 6.9 y 6.10 se dan las conclusiones y las referencias.

6.2 Método de Ziegler-Nichols y procedimientos relacionados

Dos métodos clásicos para determinar los parámetros de controladores PID fueron presentados por Ziegler y Nichols en 1942. estos métodos son todavía ampliamente utilizados en su forma original o en algunas de sus modificaciones. A menudo constituyen la base de los procedimientos de sintonía utilizados por los fabricantes de controladores y la industria de procesos. Los métodos se basan en la determinación de algunas características de la dinámica del proceso. Los parámetros del controlador se expresan entonces en términos de estas características mediante fórmulas sencillas. Es sorprendente que los métodos sean tan ampliamente referenciados porque dan moderadamente buena sintonía sólo en situaciones restringidas. Explicaciones plausibles pueden ser la simplicidad de los métodos y el hecho de que se puedan usar como ejercicios sencillos para los estudiantes en cursos básicos de control.

El método de la respuesta en escalón

El primer método de diseño presentado por Ziegler y Nichols se basa en obtener información del proceso en la forma de la respuesta a un escalón en lazo abierto. Este método se puede ver como un método tradicional basado en modelado y control donde se utiliza un modelo muy simple del proceso. La respuesta a un escalón se caracteriza por sólo dos parámetros a y L , tal como se muestra en la Figura 6.1. Compare también con la Figura 2.32.

En primer lugar se determina el punto donde la pendiente de la respuesta a un escalón tiene su máximo y se dibuja la tangente en este punto. Las intersecciones entre la tangente y los ejes de coordenadas dan los parámetros a y L . En el Capítulo 2, se obtuvo con estos parámetros un modelo del proceso que se va a controlar. Esto corresponde a modelar un proceso por un integrador y un retardo de tiempo. Ziegler y Nichols han dado directamente los parámetros del PID como

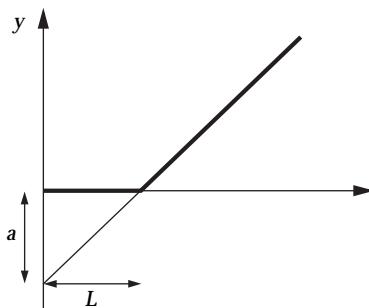


Figura 6.1 Caracterización de la respuesta en escalón en el método de la respuesta a un escalón de Ziegler-Nichols.

Cuadro 6.1 Parámetros del controlador para el método de la respuesta a un escalón de Ziegler-Nichols.

Controlador	aK	T_i/L	T_d/L	T_p/L
P	1			4
PI	0,9	3		5,7
PID	1,2	2	$L/2$	3,4

funciones de a y L . En la Tabla 6.1 se muestran estos valores. Se da también en la tabla una estima del período T_p del sistema en lazo cerrado.

EJEMPLO 6.1—MÉTODO DE LA RESPUESTA A UN ESCALÓN DE ZIEGLER-NICHOLS

El método de Ziegler-Nichols se aplicará a un proceso con la función de transferencia

$$P(s) = \frac{1}{(s+1)^3}. \quad (6.1)$$

Medidas de la respuesta a un escalón dan los parámetros $a = 0,218$ y $L = 0,806$. Los parámetros del controlador se pueden ahora determinar de la Tabla 6.1. Los parámetros de un controlador PID son $K = 5,50$, $T_i = 1,61$, y $T_d = 0,403$. En la Figura 6.2 se muestra la respuesta del sistema en lazo cerrado a un cambio en escalón en el punto de consigna seguido por un cambio en escalón en la carga. La conducta del controlador es tal como cabría esperar. La razón de decaimiento para la respuesta en escalón está próxima a un cuarto. Es más pequeña para la perturbación en la carga. La sobreelongación en la respuesta del punto de consigna es demasiado grande. Esto se puede mejorar por la ponderación del punto de consigna b . Compárese con la Sección 3.4. \square

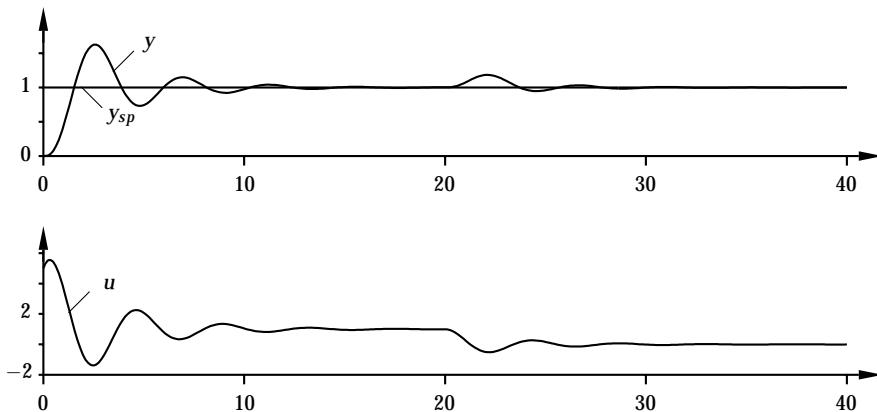


Figura 6.2 Respuesta del punto de consigna y y de la perturbación en la carga de un proceso con función de transferencia $1/(s + 1)^3$ controlado por un controlador PID sintonizado con el método de la respuesta a un escalón de Ziegler-Nichols. Los diagramas muestran el punto de consigna y_{sp} , la salida del proceso y , y la señal de control u .

El método de la respuesta en frecuencia

Este método se basa también en una caracterización simple de la dinámica del proceso. El diseño se basa en el conocimiento del punto del diagrama de Nyquist de la función de transferencia del proceso $P(s)$ donde dicho diagrama intersecta al eje real negativo. En la Sección 2.4 este punto se caracteriza por K_{180} y ω_{180} . Por razones históricas el punto se ha referido como el punto último y se define por los parámetros $K_u = 1/K_{180}$ y $T_u = 2\pi/\omega_{180}$, que se llaman la *ganancia última* y el *período último*. Estos parámetros se pueden determinar de la forma siguiente. Se conecta un controlador al proceso, y se establecen los parámetros de forma que la acción de control es proporcional, i.e., $T_i = \infty$ y $T_d = 0$. Se aumenta lentamente la ganancia hasta que el proceso comienza a oscilar. La ganancia cuando esto ocurre es K_u , y el período de la oscilación es T_u . Tenemos $K_u = 1/K_{180}$ y $T_u = 2\pi/\omega_u$. Los parámetros se pueden también determinar aproximadamente mediante la realimentación de un relé tal como se analiza en la Sección 2.7.

Ziegler-Nichols han dado fórmulas sencillas para los parámetros del controlador en términos de la ganancia y del período último que se muestra en la Tabla 6.2. En la tabla se da también una estimación del período T_p de la dinámica dominante del sistema en lazo cerrado.

El método de la respuesta en frecuencia se puede ver también como un procedimiento de sintonía empírico donde los parámetros del controlador se obtienen por experimentos directos sobre el proceso combinado con algunas reglas sencillas. Para un controlador proporcional la regla es simplemente aumentar la ganancia hasta que el proceso oscila y entonces reducir la ganancia en un 50 por ciento.

Ilustramos el procedimiento de diseño con un ejemplo.

6.2 Método de Ziegler-Nichols y procedimientos relacionados

Cuadro 6.2 Parámetros del controlador para el método de la respuesta en frecuencia de Ziegler-Nichols.

Controlador	K/K_u	T_i/T_u	T_d/T_u	T_p/T_u
P	0,5			1,0
PI	0,4	0,8		1,4
PID	0,6	0,5	0,125	0,85

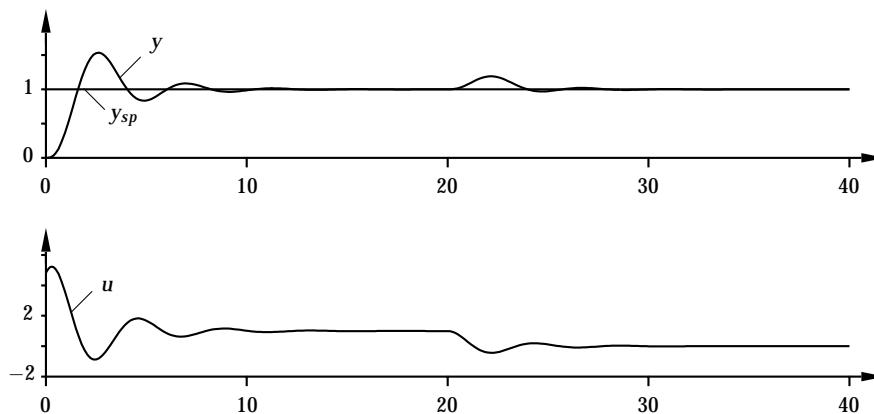


Figura 6.3 Respuesta del punto de consigna y_{sp} y de la perturbación en la carga de un proceso con la función de transferencia $1/(s+1)^3$ controlado por un PID que se sintoniza con el método de la respuesta en frecuencia de Ziegler-Nichols. Los diagramas muestran el punto de consigna y_{sp} , la salida del proceso y , y la señal de control u .

EJEMPLO 6.2—EL MÉTODO DE LA RESPUESTA EN FRECUENCIA DE ZIEGLER-NICHOLS

Considérese el mismo proceso como en el Ejemplo 6.1. El proceso dado por (6.1) tiene la ganancia última $K_u = 8$ y el período último $T_u = 2\pi/\sqrt{3} = 3,63$. La Tabla 6.2 da los parámetros $K = 4,8$, $T_i = 1,81$, y $T_d = 0,44$ para un controlador PID. En la Figura 6.3 se muestran las respuestas del punto de consigna en lazo cerrado y de las perturbaciones de carga cuando se aplica el controlador al proceso que está dado por (6.1).

Los parámetros y el comportamiento de los controladores obtenidos con el método de la respuesta en frecuencia están próximos a aquellas obtenidas en el método de la respuesta a un escalón. Las respuestas están ligeramente mejor amortiguadas. \square

Las reglas de sintonía de Ziegler-Nichols fueron originalmente diseñadas para dar sistemas con buenas respuestas a perturbaciones en la carga. Se obtuvieron mediante una gran cantidad de simulaciones de muchos sistemas diferentes con evaluación manual de los resultados. El criterio de diseño fue la razón de decaimiento de un cuarto, que es a menudo demasiado grande, como se observa en

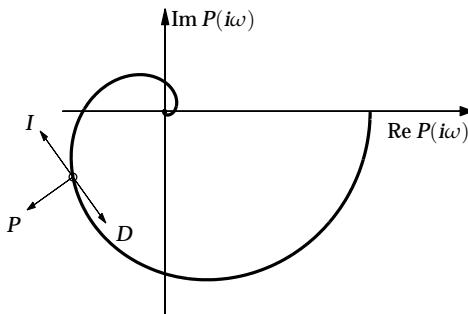


Figura 6.4 Ilustra que un punto del diagrama de Nyquist de la función de transferencia del proceso se puede mover a otra posición mediante control PID. El punto marcado con un círculo se puede mover en las direcciones $P(i\omega)$, $-iP(i\omega)$, y $iP(i\omega)$ cambiando respectivamente la ganancia proporcional, integral, y derivativa.

los ejemplos. Por esta razón el método de Ziegler-Nichols requiere a menudo una modificación o re-sintonía. Como el objetivo de diseño primario era reducir las perturbaciones de carga, con frecuencia es necesario escoger cuidadosamente la ponderación del punto de consigna con el fin de obtener una respuesta satisfactoria del punto de consigna.

Una interpretación del método de la respuesta en frecuencia

El método de respuesta en frecuencia se puede interpretar como un método donde se posiciona un punto del diagrama de Nyquist. Con control PI o PID, es posible mover un punto dado de la curva de Nyquist de la función de transferencia del proceso a una posición arbitraria en el plano complejo, tal como se indica en la Figura 6.4. Cambiando la ganancia, se mueve un punto del diagrama de Nyquist radialmente desde el origen. El punto se puede mover en la dirección ortogonal cambiando la ganancia integral o derivativa. Obsérvese que con parámetros positivos del controlador el punto se puede mover a un cuarto del plano con control PI o PD y a la mitad del plano con control PID. Desde este punto de vista el método de Ziegler-Nichols se puede interpretar como un método primitivo de moldeo del lazo donde un punto de la función de transferencia del lazo se mueve a un punto deseado.

El método de la respuesta en frecuencia comienza con la determinación del punto $(-1/K_u, 0)$ donde el diagrama de Nyquist de la función de transferencia en lazo abierto intersecta al eje real negativo.

Investigaremos ahora como se cambia el punto último por el controlador. Para un controlador PI con sintonía de Ziegler-Nichols tenemos $K = 0,4K_u$ y $\omega_u T_i = (2\pi/T_u)0,8T_u = 5,02$. Por lo tanto, la función de transferencia del controlador PI en la frecuencia última es

$$C(i\omega_u) = K \left(1 + \frac{1}{i\omega_u T_i} \right) = 0,4K_u(1 - i/5,02) = K_u(0,4 - 0,08i).$$

El punto último se mueve así a $-0,4 + 0,08i$. Esto significa que se introduce un retraso de $11,2^\circ$ en la frecuencia última.

6.2 Método de Ziegler-Nichols y procedimientos relacionados

Para un controlador PID tenemos $K = 0,6K_u$, $\omega_u T_i = \pi$, y $\omega_u T_d = \pi/4$. La respuesta en frecuencia del controlador ω_u es

$$C(i\omega_u) = K \left(1 + i \left(\omega_u T_d - \frac{1}{\omega_u T_i} \right) \right) \approx 0,6K_u(1 + 0,467i).$$

Este controlador da un avance de fase de 25° en la frecuencia última. La función de transferencia del lazo es

$$G_\ell(i\omega_u) = P(i\omega_u)C(i\omega_u) = -0,6(1 + 0,467i) = -0,6 - 0,28i.$$

El método de respuesta en frecuencia de Ziegler-Nichols para un controlador PID así mueve el punto último $(-1/K_u, 0)$ al punto $-0,6 - 0,28i$. La distancia de este punto al punto crítico es 0,5. Esto significa que el método da una sensibilidad que es siempre mayor que 2.

Ha sido sugerido por Pessen mover el punto último a $-0,2 - 0,36i$ o $-0,2 - 0,21i$. Suda empleó aproximaciones para obtener $M_t = 1,3$ moviendo el punto crítico a $-0,628 - 0,483i$.

Diseño de un controlador PI con un margen de fase dado

Usando la idea de que el controlador PI se puede interpretar como mover un punto de la función de transferencia del lazo es fácil desarrollar un método de diseño que da un sistema en lazo cerrado con un margen de fase dado. Sea la función de transferencia del proceso

$$P(i\omega) = \alpha(\omega) + i\beta(\omega) = \rho(\omega)e^{i\psi(\omega)}.$$

Con control PI la función de transferencia del lazo es

$$G_l(i\omega) = \left(k - i \frac{k_i}{\omega} \right) (\alpha(\omega) + i\beta(\omega)) = \alpha(\omega)k + \frac{\beta(\omega)k_i}{\omega} + i \left(\beta(\omega)k - \frac{\alpha(\omega)k_i}{\omega} \right).$$

Sea ω_{gc} la frecuencia de cruce de ganancia; requiriendo que el sistema tenga un margen de fase φ_m se sigue que

$$G_l(i\omega_{gc}) = -\cos(\varphi_m) - i\sin(\varphi_m),$$

que implica que

$$\begin{aligned} \alpha(\omega_{gc})k + \frac{\beta(\omega_{gc})k_i}{\omega_{gc}} &= -\cos(\varphi_m) \\ \beta(\omega_{gc})k - \frac{\alpha(\omega_{gc})k_i}{\omega_{gc}} &= -\sin(\varphi_m). \end{aligned}$$

Resolviendo esta ecuación para k y k_i da

$$\begin{aligned} k &= -\frac{\alpha(\omega_{gc}) \cos \varphi_m + \beta(\omega_{gc}) \sin \varphi_m}{\alpha^2(\omega_{gc}) + \beta^2(\omega_{gc})} = -\frac{1}{\rho(\omega_{gc})} \cos(\varphi_m - \psi(\omega_{gc})) \\ k_i &= \omega_{gc} \frac{\alpha(\omega_{gc}) \sin \varphi_m - \beta(\omega_{gc}) \cos \varphi_m}{\alpha^2(\omega_{gc}) + \beta^2(\omega_{gc})} = \frac{\omega_{gc}}{\rho(\omega_{gc})} \sin(\varphi_m - \psi(\omega_{gc})). \end{aligned} \tag{6.2}$$

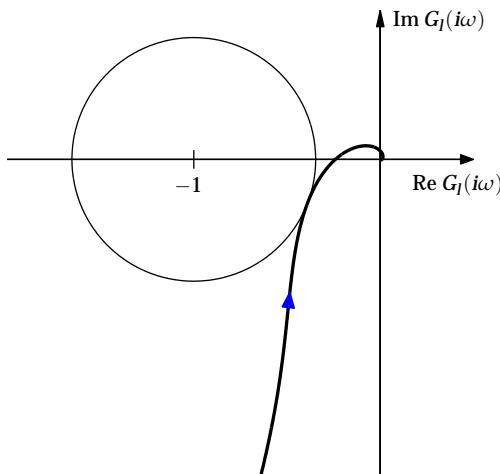


Figura 6.5 El diagrama de Nyquist para la función de transferencia del lazo G_l para control PI del proceso $P(s) = e^{-\sqrt{s}}$. El controlador se diseñó para dar el margen de fase de 60° .

Es así directo calcular las ganancias del controlador cuando se da la frecuencia de cruce de ganancia. Valores razonables de la frecuencia de cruce de ganancia están en el rango $\omega_{90} \leq \omega_{gc} \leq \omega_{180-\varphi_m}$. El método se puede mejorar barriendo sobre ω_{gc} para maximizar la ganancia integral. Aplicando el método para diseñar un controlador PI para el proceso $P(s) = e^{-\sqrt{s}}$ con un margen de fase de 60° da $\omega_{gc} = 5,527$, $K = 4,79$ y $T_i = 0,392$ y $M_s = 1,53$. En la Figura 6.5 se muestra el diagrama de Nyquist de la función de transferencia del lazo .

Relaciones entre los métodos de sintonía de Ziegler-Nichols

El método de respuesta a un escalón y el método de respuesta en frecuencia no dan los mismos valores de los parámetros del controlador. Comparando los Ejemplos 6.1 y 6.2 encontramos que las ganancias del controlador son 5,5 y 4,8 y que los tiempos integrales son 1,61 y 1,81. El método de la respuesta a un escalón dará en general ganancias más grandes y tiempos integrales más pequeños. Esto se ilustra más aun en el ejemplo siguiente.

EJEMPLO 6.3—PROCESO CON INTEGRACIÓN Y RETARDO

Considérese un proceso con la función de transferencia

$$P(s) = \frac{K_v}{s} e^{-sL},$$

que es el modelo originalmente usado por Ziegler y Nichols para derivar sus reglas de sintonía para el método de la respuesta en escalón. Para este proceso tenemos que $a = K_v L$. La frecuencia última es $\omega_u = \pi/2L$, que da el período último $T_u = 4L$, y la ganancia última es $K_u = \pi/2K_v L$.

Con control PI el método de la respuesta a un escalón da los siguientes parámetros:

$$K = \frac{0,9}{K_v L}, \quad T_i = 3L.$$

Esto se puede comparar con los parámetros

$$K = \frac{0,63}{K_v L}, \quad T_i = 3,2L$$

obtenidos con el método de la respuesta en frecuencia. Obsérvese que los tiempos integrales están dentro del 10 por ciento, pero que el método de la respuesta en escalón da una ganancia que es cerca de un 40 por ciento más elevada.

Los parámetros PID obtenidos del método de la respuesta en escalón son

$$K = \frac{1,2}{bL}, \quad T_i = 2L \text{ and } T_d = \frac{L}{2},$$

y aquellos dados por los métodos de la respuesta en frecuencia son

$$K = \frac{0,94}{bL}, \quad T_i = 2L \text{ and } T_d = \frac{L}{2}.$$

Ambos métodos dan los mismos valores de los tiempos integral y derivativo, pero el método de la respuesta a un escalón da una ganancia que es cerca del 25 por ciento más grande que la del método de respuesta en frecuencia. \square

EJEMPLO 6.4—PROCESOS CON RETARDO PURO

Sea un proceso con la función de transferencia

$$P(s) = K_p e^{-sL}.$$

En este caso encontramos que $a = \infty!$ El método de la respuesta en frecuencia da así una ganancia nula del controlador para control PI y PID.

El período último es $T_u = 2L$, y la ganancia última es $K_u = 1/K_p$. Utilizando el método de la respuesta en frecuencia se sigue de la Tabla 6.2 que para control PI $KK_p = 0,4$ y $T_i = 1,6L$. El controlador PI da un resultado muy pobre tal como se ilustra en la Figura 6.6. La acción integral es demasiado pequeña, lo que implica que se emplea un tiempo muy grande para que el error tienda a cero. Como comparación se muestra también la respuesta con un controlador PI que tiene $KK_p = 0,25$ y $T_i = 0,35$. Este controlador tiene una respuesta mucho mejor frente a perturbaciones en la carga.

Para control PID el método de la respuesta en frecuencia da $KK_p = 0,6$, $T_i = L$ y $T_d = 0,25$, lo que da como resultado un sistema inestable en lazo cerrado. \square

Estos ejemplos muestran que pueden existir considerables diferencias entre los parámetros del controlador obtenidos por la respuesta a un escalón y los métodos de respuesta en frecuencia.

El método de Chien, Hrones, y Reswick

Ha habido muchas sugerencias para modificar los métodos de Ziegler-Nichols. Hay métodos que usan la misma información acerca del proceso que los métodos de Ziegler-Nichols, pero se modifican los coeficientes en las Tablas 6.1 y 6.2. Los

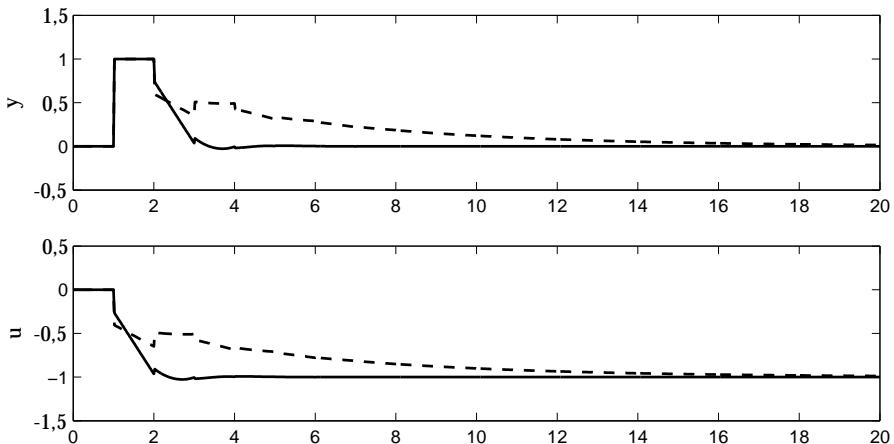


Figura 6.6 Respuestas frente a una perturbación en la carga para un proceso con retardo ($L = 1$) con controladores PI sintonizados por el método de Ziegler-Nichols de la respuesta en frecuencia (a tramos) y un método propio (continua).

fabricantes de controladores emplean muchos métodos de este tipo. Hay también otros métodos que usan más datos del proceso. Muchos métodos se basan en la idea de que el proceso se approxima por el modelo FOTD

$$P(s) = \frac{K_p}{1 + sT} e^{-sL}.$$

Como una ilustración describiremos un método desarrollado por Chien, Hrones, y Reswick (CHR). Su método da un sistema en lazo cerrado con una robustez ligeramente mejor que el método de Ziegler-Nichols. Los criterios de diseño utilizados fueron «respuesta más rápida sin sobrelongación» o «respuesta más rápida con 20 por ciento de sobrelongación». Propusieron diferentes reglas de sintonía para perturbaciones en la carga y respuesta a cambios en el punto de consigna.

Para sintonizar el controlador de acuerdo con el método CHR, los parámetros a y L del modelo del proceso se determinan primero de la misma forma que para el método de la respuesta en escalón de Ziegler-Nichols. Los parámetros del controlador se dan entonces como funciones de estos dos parámetros. En la Tabla 6.3 se muestran las reglas de sintonía para la respuesta frente a perturbaciones en la carga. Las reglas de sintonía en la Tabla 6.3 tienen en general ganancias más bajas que las correspondientes a la regla de Ziegler-Nichols en la Tabla 6.1.

Chien, Hrones, y Reswick encontraron que la sintonía para la respuesta a cambios en el punto de consigna era diferente de la sintonía frente a perturbaciones en la carga. En ese tiempo las ventajas de la ponderación del punto de consigna y de los sistemas con dos grados de libertad no se conocían. Se requería un parámetro adicional, la constante de tiempo T , y las ganancias del controlador eran en general más bajas; véase la Tabla 6.4.

Cuadro 6.3 Parámetros del controlador obtenidos de la respuesta a perturbaciones en la carga del método de Chien, Hrones y Reswick.

Controlador	No sobreelongación			20 % sobreelongación		
	aK	T_i/L	T_d/L	aK	T_i/L	T_d/L
P	0,3			0,7		
PI	0,6	4,0		0,7	2,3	
PID	0,95	2,4	0,42	1,2	2,0	0,42

Cuadro 6.4 Parámetros del controlador obtenidos de la respuesta a cambios en el punto de consigna del método de Chien, Hrones y Reswick.

Controlador	No sobreelongación			20 % sobreelongación		
	aK	T_i/L	T_d/L	aK	T_i/L	T_d/L
P	0,3			0,7		
PI	0,35	1,2		0,6	1,0	
PID	0,6	1,0	0,5	0,95	1,4	0,47

El método de Cohen-Coon

El método de Cohen-Coon se basa también en un modelo FOTD del proceso

$$P(s) = \frac{K_p}{1+sT} e^{-sL}.$$

El criterio principal de diseño es el rechazo de las perturbaciones de carga. Intenta posicionar los polos dominantes que dan una razón de decaimiento de la amplitud de un cuarto. Para controladores P y PD los polos se ajustan para dar una ganancia máxima del controlador, sujeta a la restricción de la razón de decaimiento. Esto minimiza el error en estado estacionario debido a las perturbaciones en la carga. Para control PI y PID se maximiza la ganancia integral $k_i = K/T_i$. Esto corresponde a la minimización de la integral del error IE, debido a una perturbación de carga de un escalón unitario. Para controladores PID se asignan tres polos en lazo cerrado; dos polos son complejos, y el tercer polo real se posiciona a la misma distancia del origen que los otros polos. La configuración de polos se ajusta para dar una razón de decaimiento en amplitud de un cuarto, y la distancia de los polos al origen se ajusta para minimizar IE.

Como el proceso se caracteriza por tres parámetros (K_p , L , y T), es posible dar fórmulas de sintonía donde los parámetros del controlador se expresan en términos de estos parámetros. Tales fórmulas fueron derivadas por Cohen y Coon

Cuadro 6.5 Parámetros del controlador del método de Cohen-Coon.

Controlador	aK	T_i/L	T_d/L
P	$1 + \frac{0,35\tau}{1 - \tau}$		
PI	$0,9 \left(1 + \frac{0,092\tau}{1 - \tau} \right)$	$\frac{3,3 - 3,0\tau}{1 + 1,2\tau}$	
PD	$1,24 \left(1 + \frac{0,13\tau}{1 - \tau} \right)$		$\frac{0,27 - 0,36\tau}{1 - 0,87\tau}$
PID	$1,35 \left(1 + \frac{0,18\tau}{1 - \tau} \right)$	$\frac{2,5 - 2,0\tau}{1 - 0,39\tau}$	$\frac{0,37 - 0,37\tau}{1 - 0,81\tau}$

basándose en cálculos analíticos y numéricos. En la Tabla 6.5 se dan las fórmulas. Los parámetros $a = K_p L/T$ y $\tau = L/(L+T)$ se usan en la tabla para facilitar la comparación con la sintonía de Ziegler-Nichols. Una comparación con la Tabla 6.1 muestra que los parámetros del controlador están próximos a los obtenidos por el método de la respuesta a un escalón de Ziegler-Nichols para τ pequeños. También se observa que el tiempo integral disminuye para τ crecientes, lo que es deseable tal como se encontró en Sección 6.2. Una peculiaridad es que las ganancias van a infinito cuando τ va a 1, lo que no es correcto. El método también sufre de tener una razón de decaimiento demasiado grande, lo que significa que el sistema en lazo cerrado obtenido tiene un pobre amortiguamiento y una alta sensibilidad.

Comentario

Las reglas de sintonía de Ziegler-Nichols son simples e intuitivas, requieren poco conocimiento del proceso, y se pueden aplicar con un esfuerzo modesto. El proceso se caracteriza por dos parámetros que se pueden determinar mediante experimentos sencillos. El método de la respuesta en frecuencia tiene la ventaja de que los parámetros K_u y T_u son más fáciles de determinar de forma precisa que los parámetros a y L , que se usan por el método de la respuesta a un escalón.

Los métodos son todavía ampliamente utilizados aunque dan sistemas en lazo cerrado que no son robustos. Las reglas se combinan a menudo con una sintonía manual, que se presentará en Sección 6.3. La desventaja principal con los métodos es que utilizan poca información del proceso y que el criterio de diseño de razón de decaimiento de la amplitud de un cuarto da sistemas en lazo cerrado con pobre robustez. No está claro por qué se utilizó este criterio de diseño. Las respuestas de las perturbaciones de carga parecen bastante razonables, pero sin análisis y estudios de sensibilidad no es obvio que los sistemas en lazo cerrado no son robustos. Las simulaciones mostradas en las Figura 6.2 y Figura 6.3 indican que los métodos dan un control razonable. Simulaciones repetidas con perturbaciones en los parámetros del controlador revelan muy claramente que el sistema en lazo cerrado no es robusto. Sistemas como los que se muestran en los Ejemplos 6.3 y 6.4 también ilustran que no es suficiente caracterizar el proceso solamente

por dos parámetros.

Ha sido propuestos un número muy grande de variaciones de los métodos de Ziegler-Nichols. Hemos seleccionado para presentar dos métodos. Las modificaciones del método de Chien-Hrones-Reswick dan sistemas con una robustez algo mejor, pero todavía utilizan demasiada poca información del proceso. El método de Cohen-Coon usa tres parámetros para caracterizar el proceso, pero aún emplea el amortiguamiento de un cuarto de la amplitud como un criterio de diseño.

En el Capítulo 7 desarrollaremos nuevos métodos que abordan las grandes desventajas de los métodos de Ziegler-Nichols aunque reteniendo su simplicidad.

6.3 Sintonía empírica basada en reglas

Como los métodos de Ziegler-Nichols solo dan valores «aproximados», es necesario complementar los métodos mediante sintonía manual para obtener propiedades razonables en lazo cerrado. La sintonía manual se realiza típicamente con experimentos sobre el proceso en lazo cerrado. Se introduce una perturbación como un cambio en el punto de consigna o bien como un cambio en la variable de control. Se observa la respuesta en lazo cerrado, y se ajustan los parámetros del controlador. Los ajustes se basan en reglas simples, que dan pautas para cambiar los parámetros. Las reglas se desarrollaron mediante la realización de una gran cantidad de experimentos. Lo que sigue es un conjunto simple de reglas:

- Aumentando la ganancia proporcional disminuye la estabilidad
- El error decae más rápidamente si se disminuye el tiempo de integración
- Disminuyendo el tiempo de integración disminuye la estabilidad
- Aumentando el tiempo derivativo mejora la estabilidad

Finalmente, las reglas de sintonía han sido también formalizadas en diferentes tipos de sistemas formales basados en reglas tales como sistemas expertos o lógica borrosa.

Los mapas de sintonía son una forma de expresar las reglas de sintonía. El objetivo de estos mapas es proporcionar intuición acerca de cómo influyen los cambios en los parámetros del controlador sobre la conducta del sistema en lazo cerrado. Los mapas de sintonía son simples tablas de respuestas transitorias o en frecuencia que corresponden a variaciones sistemáticas en los parámetros del controlador. En la Figura 6.7 se da un ejemplo de un mapa de sintonía.

La figura ilustra cómo la respuesta de la perturbación de carga está influenciada por cambios en la ganancia y en el tiempo integral. El modelo del proceso

$$P(s) = \frac{1}{(s+1)^8}$$

ha sido utilizado en el ejemplo. El método de respuesta en frecuencia de Ziegler-Nichols da los parámetros del controlador $K = 1,13$, $T_i = 7,58$, y $T_d = 1,9$. La figura muestra claramente los beneficios de obtener un valor más pequeño de T_i . Del análisis de la figura parece que los valores $K = 1$ y $T_i = 5,0$ son razonables.

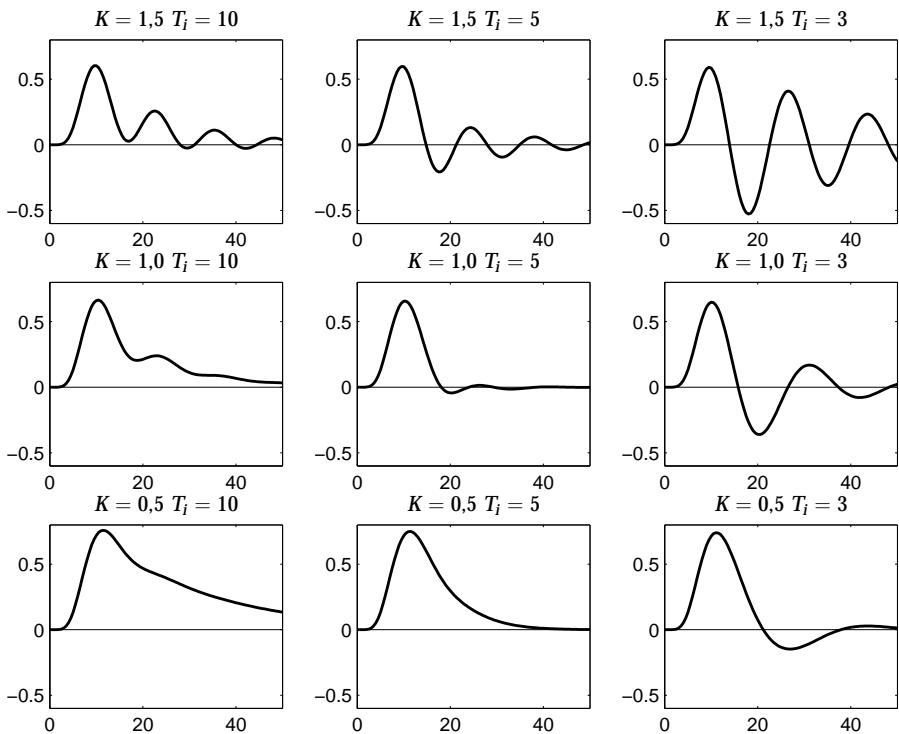


Figura 6.7 Mapa de sintonía para control PID de un proceso con la función de transferencia $P(s) = (s + 1)^{-8}$. La figura muestra las respuestas a una perturbación en escalón unitaria en la entrada del proceso. El parámetro T_d tiene el valor 1.9.

La figura también muestra que la elección de T_i es relativamente crítica. También se observa que no se pueden implementar los controladores en forma serie con $T_i < 7.6$ (comparar con la Sección 3.4).

En la Figura 6.8 se muestra un tipo diferente de mapa de sintonía, que representa los diagramas de Nyquist de la función de transferencia del lazo. La figura muestra que algunos de los diagramas de Nyquist se doblan demasiado hacia la derecha en las bajas frecuencias; ver las figuras en la posiciones izquierda con $T_i = 10$. Esto significa que el controlador introduce demasiado adelanto de fase. Esto se reduce al disminuir el parámetro T_i .

Un estudio comparativo de diagramas como los de la Figura 6.7 y Figura 6.8 es una buena forma de desarrollar la intuición para las relaciones entre las respuestas temporal y frecuencial. Una manera incluso mejor es usar el software interactivo que está ahora emergiendo.

Conducta contra intuitiva

Las reglas comunes para sintonía manual dicen que el sistema se hace menos oscilatorio si se reduce la ganancia, si se aumenta el tiempo integral, y si se aumenta el tiempo derivativo. Comparar con la Figura 6.4. Estas reglas se mantienen para el sistema que se muestra en la Figura 6.7 y Figura 6.8. Hay, sin embargo,

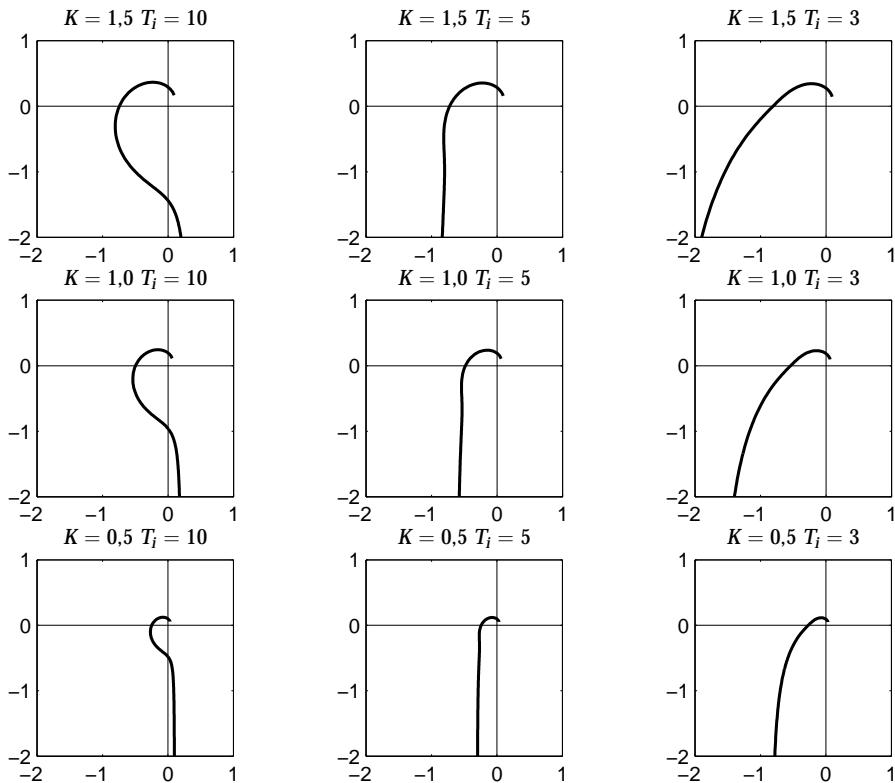


Figura 6.8 Mapa de sintonía para control PID de un proceso con la función de transferencia $P(s) = (s + 1)^{-8}$. La figura muestra los diagramas de Nyquist de las funciones de transferencia del lazo. El parámetro T_d tiene el valor 1.9.

situaciones donde estas reglas no se cumplen. El siguiente es un ejemplo común sencillo.

EJEMPLO 6.5—CONTROL PI DE UN INTEGRADOR
Sea un proceso con la función de transferencia

$$P(s) = \frac{1}{s},$$

y un controlador PI con la función de transferencia

$$C(s) = K\left(1 + \frac{1}{sT_i}\right).$$

La función de transferencia del lazo es

$$G_l(s) = P(s)C(s) = K \frac{1 + sT_i}{s^2 T_i}.$$

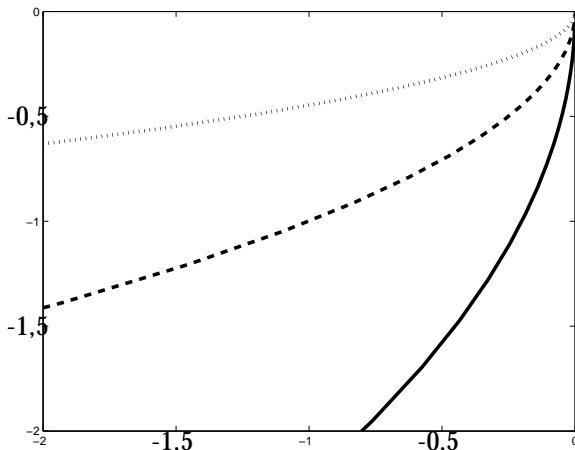


Figura 6.9 Diagramas de Nyquist para la función de transferencia del lazo para un integrador con control PI. El tiempo de integración T_i es constante, y la ganancia tiene los valores $K = 0,2$ (punteado), 1 (a tramos), y 5 (continua). Obsérvese la conducta contra intuitiva de que el margen de fase aumenta cuando aumenta la ganancia del controlador.

y la ecuación característica es

$$s^2 + Ks + \frac{K}{T_i} = 0.$$

Identificando esto con un sistema de segundo orden estándar $s^2 + 2\zeta\omega_0 + \omega_0^2$ encontramos que

$$\zeta = \sqrt{\frac{K}{2T_i}}.$$

Se sigue de esta ecuación que el amortiguamiento aumenta cuando se aumenta la ganancia del controlador lo que es contrario a la intuición desarrollada para los sistemas simples. Esto se ilustra también por los diagramas de Nyquist en la Figura 6.9. Se ve como el diagrama de Nyquist se mueve fuera del punto crítico -1 cuando la ganancia aumenta. La razón para esto es que el diagrama de Nyquist está muy próximo al eje imaginario negativo para grandes ω . Observe que un pequeño retardo de tiempo o una pequeña constante de tiempo destruirá esta propiedad. \square

Situaciones como esta hace difícil el establecer reglas eficientes que cubran un amplio rango de condiciones.

Una inigualdad para el tiempo de integración

Es útil tener una forma sencilla de juzgar si la acción integral de un controlador es demasiado débil, como en los tres ejemplos de la izquierda y de la parte inferior en la Figura 6.7 y la Figura 6.8. Tal criterio se puede basar en un cálculo de la conducta asintótica de la función de transferencia del lazo. Para un proceso con

función de transferencia P y un controlador PI con función de transferencia C tenemos

$$\begin{aligned} G_\ell(s) &= P(s)C(s) \approx (P(0) + sP'(0)) K \left(1 + \frac{1}{sT_i} \right) \\ &= \frac{KP(0)}{sT_i} + KP(0) + \frac{KP'(0)}{T_i} + KP'(0)s. \end{aligned}$$

Así, para bajas frecuencias la asymptota del diagrama de Nyquist es paralela al eje imaginario con la parte real igual a

$$KP(0) + \frac{KP'(0)}{T_i} = KK_p \left(1 - \frac{T_{ar}}{T_i} \right),$$

donde $K_p = G(0)$ es la ganancia estática del proceso, y T_{ar} es el tiempo de residencia medio. Es razonable requerir que la parte real de la asymptota sea menor que $-0,5$. Esto da

$$T_i < T_{ar} \frac{2KK_p}{1 + 2KK_p} < T_{ar}. \quad (6.3)$$

Para el sistema en la Figura 6.7 y la Figura 6.8, obtenemos el requisito $T_i < 6,0$ para los sistemas de la fila superior, $T_i < 5,3$ para los sistemas de la fila del medio, y $T_i < 4,0$ para los sistemas de la fila inferior. Esto significa que la condición (6.3) excluye a los tres ejemplos de la izquierda y de las filas del medio e inferior en la Figura 6.7 y la Figura 6.8.

Se puede usar la inigualdad para el tiempo de integración dada por (6.3) para tener una idea de las limitaciones de las reglas de Ziegler-Nichols para sistemas con grandes retardos de tiempo. Sea un proceso con la función de transferencia

$$P(s) = K_p \frac{e^{-sL}}{1 + sT}.$$

Para este sistema tenemos $T_{ar} = L + T$. Considérese un controlador PI sintonizado por el método de la respuesta a un escalón de Ziegler-Nichols. Se sigue de la Tabla 6.1 que $KK_p = 0,9T/L$ y $T_i = 3L$. La ecuación 6.3 entonces da

$$3L < (L + T) \frac{1,8T}{L + 1,8T},$$

que implica que $L < 0,38T$. Esto significa que el método de la respuesta a un escalón de Ziegler-Nichols para control PI no dará buen control a menos que el retardo de tiempo es suficientemente pequeño. Compárese con el Ejemplo 6.4.

Comentario

La sintonía manual se utilizó antes de que estuviera disponible cualquier método sistemático de sintonía. Es un complemento necesario para el método de Ziegler-Nichols. Es esencial para todos los practicantes de control ganar experiencia para juzgar las propiedades de los sistemas en lazo cerrado y cambiar los parámetros del controlador para modificar la conducta. La evaluación se puede basar en un simple test de saltos donde se perturba el punto de consigna o la salida

del controlador o mediante medidas de respuesta en frecuencia más elaboradas. Es necesario conocer la conducta contra intuitiva de procesos con acción integral que se ilustraron en el Ejemplo 6.5. Los sistemas basados en reglas han sido formalizados cuando se desarrollaron los autosintonizadores basados en sistemas expertos y lógica borrosa. En la Sección 6.7 presentaremos métodos sistemáticos para mejorar la sintonía basados en optimización.

6.4 Asignación de polos

Muchas propiedades de un sistema en lazo cerrado se expresan por sus polos. La idea con la asignación de polos es diseñar un controlador que da un sistema en lazo cerrado con polos en lazo cerrado deseados. El método requiere un modelo completo del proceso. Sujeto a algunas condiciones técnicas es posible encontrar un controlador que da los polos en lazo cerrado deseados, a condición de que el controlador sea suficientemente complejo. Para usar el método para control PID es necesario restringir la complejidad del modelo mediante varios métodos de aproximación. Los polos seleccionados deben entonces escogerse con cuidado en orden a asegurar que el modelo aproximado es válido para frecuencias que correspondan a los polos seleccionados.

Un refinamiento del procedimiento es considerar también los ceros de las funciones de transferencias. Esto es particularmente relevante para la respuesta del punto de consigna. Los ceros de la función de transferencia que se originan por el controlador pueden verse influenciados por la ponderación del punto de consigna.

EJEMPLO 6.6—CONTROL PI DE UN SISTEMA DE PRIMER ORDEN

Supóngase que el proceso se puede describir por el siguiente modelo de primer orden

$$P(s) = \frac{K_p}{1 + sT},$$

que tiene sólo dos parámetros, la ganancia del proceso K_p y la constante de tiempo T . Sea el proceso controlado por un controlador PI estándar con ponderación del punto de consigna,

$$\begin{aligned} C(s) &= K\left(1 + \frac{1}{sT_i}\right) \\ C_{ff}(s) &= K\left(b + \frac{1}{sT_i}\right). \end{aligned}$$

El sistema en lazo cerrado es de segundo orden. La función de transferencia del lazo es

$$G_\ell(s) = P(s)C(s) = \frac{K_p K (1 + sT_i)}{sT_i(1 + sT)} = \frac{K_p K (s + 1/T_i)}{T(s + 1/T)},$$

y el polinomio característico

$$s^2 + \frac{1 + K_p K}{T}s + \frac{K_p K}{TT_i}. \quad (6.4)$$

El sistema en lazo cerrado tiene dos polos a los que se les puede dar valores arbitrarios por una elección apropiada de la ganancia K y del tiempo integral T_i del controlador. Ahora supóngase que los polos en lazo cerrado se caracterizan por su amortiguamiento relativo ζ y su frecuencia ω_0 . El polinomio característico en lazo cerrado entonces se hace

$$s^2 + 2\zeta\omega_0 s + \omega_0^2. \quad (6.5)$$

Identificando los coeficientes de iguales potencias de s en (6.4) y (6.5) obtenemos

$$\begin{aligned} K &= \frac{2\zeta\omega_0 T - 1}{K_p} \\ T_i &= \frac{2\zeta\omega_0 T - 1}{\omega_0^2 T} \\ k_i &= \frac{K}{T_i} = \frac{\omega_0^2 T}{K_p}. \end{aligned} \quad (6.6)$$

Es conveniente usar los parámetros ω_0 y ζ como parámetros de diseño; ω_0 determina la velocidad de respuesta y ζ determina la forma de la respuesta.

Con los parámetros del controlador dados por (6.6) el sistema en lazo cerrado se caracteriza por *la banda de los seis*, ver Ecuación (4.2).

$$\begin{aligned} \frac{PC}{1+PC} &= \frac{(2\zeta\omega_0 - 1/T)s + \omega_0^2}{s^2 + 2\zeta\omega_0 s + \omega_0^2} & \frac{C}{1+PC} &= \frac{K(s+1/T_i)(s+1/T)}{s^2 + 2\zeta\omega_0 s + \omega_0^2} \\ \frac{P}{1+PC} &= \frac{K_p s / T}{s^2 + 2\zeta\omega_0 s + \omega_0^2} & \frac{1}{1+PC} &= \frac{s(s+1/T)}{s^2 + 2\zeta\omega_0 s + \omega_0^2} \\ \frac{PC_{ff}}{1+PC} &= \frac{b(2\zeta\omega_0 - 1/T)s + \omega_0^2}{s^2 + 2\zeta\omega_0 s + \omega_0^2} & \frac{C_{ff}}{1+PC} &= \frac{K(b s + 1/T_i)(s+1/T)}{s^2 + 2\zeta\omega_0 s + \omega_0^2}. \end{aligned} \quad (6.7)$$

El valor máximo de la función de transferencia desde una perturbación en la carga a la entrada del proceso hasta la salida del proceso es

$$\max_{\omega} |G_{xd}(i\omega)| = \max_{\omega} \left| \frac{P(i\omega)}{1 + P(i\omega)C(i\omega)} \right| = \frac{K_p}{\omega_0 T \min(1, \zeta)}.$$

Para tener buen rechazo de las perturbaciones de carga es así deseable escoger ω_0 tan grande como sea posible. El valor máximo de ω_0 está limitado por la magnitud de las señales de control y la validez del modelo del proceso. La función de transferencia desde el ruido de medida a la señal de control tiene la magnitud K para altas frecuencias. Si K_{\max} es el valor permisible más grande de la ganancia del controlador se sigue de (6.6) que

$$\omega_0 T < \frac{1 + K_p K_{\max}}{2\zeta}.$$

Sea T_e la suma de las constantes de tiempo despreciadas o de los retardos de tiempo y utilizando la regla heurística de que el error de fase debería ser menor que

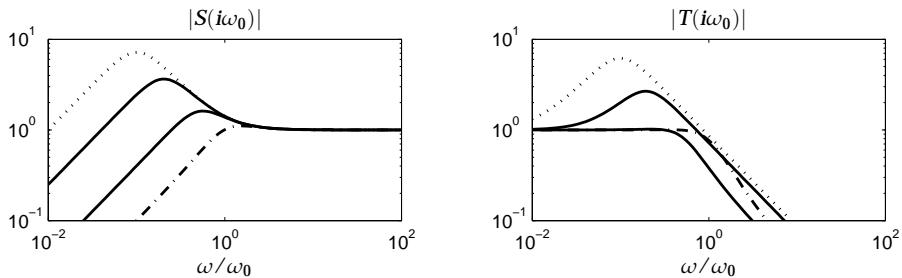


Figura 6.10 Curvas de ganancia de las funciones de sensibilidad para $\zeta = 0,7$ y $\omega_0 T = 0,1, 0,2, 0,5$, y 1 . La curva punteada corresponde a $\omega_0 L = 0,1$ y la curva de punto y raya a $\omega_0 L = 1$.

$\pm 15^\circ$ encontramos que ω_0 debe escogerse de forma que $\omega_0 T_e < 0,25$. Comparar con la Sección 2.8.

La frecuencia ω_0 seleccionada no debería ser demasiado pequeña. Una indicación de esto viene dada por la Ecuación 6.6, que muestra que la ganancia proporcional es negativa si $2\zeta\omega_0 T < 1$. en la Figura 6.10 se observa una evidencia más, que muestra los diagramas de Bode de las curvas de ganancia de las funciones de sensibilidad para diferentes valores de $\omega_0 T$. La figura muestra que las sensibilidades son grandes cuando $\omega_0 T$ es pequeña. El máximo de la función de sensibilidad es aproximadamente $M_s = 1/(2\zeta\omega_0 T)$. Una elección razonable del parámetro ω_0 es así

$$\frac{1}{2\zeta} \leq \omega_0 T < \min\left(\frac{0,25}{T_e}, \frac{1 + K_p K_{\max}}{2\zeta}\right). \quad (6.8)$$

El límite inferior corresponde a control integral puro; ver (6.6).

Se sigue de (6.7) que la función de transferencia desde el punto de consigna hasta la salida del proceso tiene un cero en $s = -1/(bT_i)$. Para evitar una sobreelongación excesiva en la respuesta del punto de consigna, el parámetro b debería escogerse de forma que el cero está a la izquierda de los polos dominantes en lazo cerrado. Un valor razonable es $b = 1/(\omega_0 T_i)$, que coloca el cero en $s = -\omega_0$. Esto da

$$b = \frac{1}{2\zeta - 1/(\omega_0 T)}.$$

Es particularmente importante usar un pequeño valor de b cuando $\omega_0 T$ es pequeña y para sistemas inestables donde T es negativo. Se obtiene una respuesta a cambios en el punto de consigna que no tiene sobreelongación escogiendo $b = 0$ y $\zeta \geq 1$.

La razón de por qué las sensibilidades son grandes para pequeños valores de $\omega_0 T$ es que el polinomio característico (6.5) es una pobre elección para diseños donde el sistema en lazo cerrado es más lento que el sistema en lazo abierto. En tales casos es mejor hacer un diseño que cancele el polo del proceso y produzca un sistema en lazo cerrado con una constante de tiempo T_0 . Tal controlador tiene

los parámetros

$$\begin{aligned} K &= \frac{T}{K_p T_0} \\ T_i &= T, \end{aligned} \tag{6.9}$$

y da un sistema en lazo cerrado con $M_s = M_t = 1$. El controlador no es apropiado cuando $\omega_0 T_0 > 1$ porque se sigue de (6.7) que la función de transferencia desde las perturbaciones de carga a la salida del proceso es

$$\frac{P}{1 + PC} = \frac{sK_p T_0}{(1 + sT)(1 + sT_0)}.$$

La atenuación de las perturbaciones de carga es así pobre para grandes valores de T_0/T . \square

EJEMPLO 6.7—CONTROL PI DE UN PROCESO CON DOS POLOS REALES

Se supone que el proceso está caracterizado por el modelo de segundo orden

$$P(s) = \frac{K_p}{(1 + sT_1)(1 + sT_2)},$$

y que se utiliza el controlador PI. La función de transferencia del lazo se hace

$$G_\ell(s) = P(s)C(s) = \frac{K_p K(1 + sT_1)}{sT_1(1 + sT_1)(1 + sT_2)} = \frac{K_p K(s + 1/T)}{T_1 T_2 (s + 1/T_1)(s + 1/T_2)},$$

y el polinomio característico es

$$s^3 + \left(\frac{1}{T_1} + \frac{1}{T_2} \right) s^2 + \frac{1 + K_p K}{T_1 T_2} s + \frac{K_p K}{T_1 T_2 T_1}. \tag{6.10}$$

A los ceros de este polinomio de tercer orden no se le pueden asignar valores arbitrarios ya que el controlador sólo tiene dos parámetros. En particular, encontramos que el coeficiente de s^2 está dado por las constantes de tiempo del proceso. Sin embargo, si también consideramos la frecuencia ω_0 como un parámetro es posible hacer coincidir este polinomio con (6.10) to

$$(s + \alpha\omega_0)(s^2 + 2\zeta\omega_0 s + \omega_0^2).$$

Igualando coeficientes de iguales potencias de s obtenemos

$$\begin{aligned} \omega_0 &= \frac{T_1 + T_2}{(\alpha + 2\zeta)T_1 T_2} \\ K &= \frac{(1 + 2\alpha\zeta)\omega_0^2 T_1 T_2 - 1}{K_p} \\ T_i &= \frac{K_p K}{\alpha\omega_0^3 T_1 T_2} \\ k_i &= \frac{\alpha\omega_0^3 T_1 T_2}{K_p}. \end{aligned}$$

Es posible obtener así un diseño que da una configuración preestablecida de los polos con control PI, i.e., especificados α y ζ . El parámetro ω_0 es un factor de escala que se determina por la dinámica del proceso. \square

EJEMPLO 6.8—CONTROL PID DE UN PROCESO CON DOS POLOS REALES
Suponga que el proceso se caracteriza por el modelo de segundo orden

$$P(s) = \frac{K_p}{(1+sT_1)(1+sT_2)}.$$

Este modelo tiene tres parámetros. Utilizando un controlador PID, que también tiene tres parámetros, es posible colocar arbitrariamente los tres polos del sistema en lazo cerrado. La función de transferencia del controlador PID se puede escribir como

$$C(s) = \frac{K(1+sT_i+s^2T_iT_d)}{sT_i}.$$

El polinomio característico del sistema en lazo cerrado es

$$s^3 + s^2 \left(\frac{1}{T_i} + \frac{1}{T_2} + \frac{K_p K T_d}{T_1 T_2} \right) + s \left(\frac{1}{T_1 T_2} + \frac{K_p K}{T_1 T_2} \right) + \frac{K_p K}{T_1 T_2 T_i}. \quad (6.11)$$

Un polinomio característico en lazo cerrado apropiado para un sistema de tercer orden es

$$(s + \alpha\omega_0)(s^2 + 2\zeta\omega_0 s + \omega_0^2), \quad (6.12)$$

que contiene dos polos dominantes con amortiguamiento relativo ζ y frecuencia ω_0 , y un polo real localizado en $-\alpha\omega_0$. Identificando los coeficientes de iguales potencias de s en las Ecuaciones 6.11 y 6.12 da

$$\begin{aligned} \frac{1}{T_i} + \frac{1}{T_2} + \frac{K_p K T_d}{T_1 T_2} &= \omega_0(\alpha + 2\zeta) \\ \frac{1}{T_1 T_2} + \frac{K_p K}{T_1 T_2} &= \omega_0^2(1 + 2\zeta\omega_0) \\ \frac{K_p K}{T_1 T_2 T_i} &= \alpha\omega_0^3. \end{aligned}$$

Resolviendo estas ecuaciones se obtienen los siguientes parámetros para el controlador:

$$K = \frac{T_1 T_2 \omega_0^2 (1 + 2\alpha\zeta) - 1}{K_p}$$

$$T_i = \frac{T_1 T_2 \omega_0^2 (1 + 2\alpha\zeta) - 1}{T_1 T_2 \alpha \omega_0^3}$$

$$T_d = \frac{T_1 T_2 \omega_0 (\alpha + 2\zeta) - T_1 - T_2}{T_1 T_2 \omega_0^2 (1 + 2\alpha\zeta) - 1}$$

$$k_i = \frac{\alpha \omega_0^3 T_1 T_2}{K_p}.$$

Si se cumple que $c = 0$, la función de transferencia desde el punto de consigna a la salida del proceso tiene un cero en $s = -1/(bT_i)$. Para evitar una sobreelongación excesiva en la respuesta del punto de consigna, se puede escoger el parámetro b de forma que este cero cancela el polo en $s = -\alpha\omega_0$. Esto da

$$b = \frac{1}{\alpha\omega_0 T_i} = \frac{\omega_0^2 T_1 T_2}{\omega_0^2 T_1 T_2 (1 + 2\alpha\zeta) - 1}.$$

También, se observa que se obtiene control PI para

$$\omega_0 = \omega_c = \frac{T_1 + T_2}{(\alpha + 2\zeta)T_1 T_2}.$$

La elección de ω_0 puede ser crítica. El tiempo derivativo es negativo para $\omega_0 < \omega_c$. Así, la frecuencia ω_c proporciona una cota inferior del ancho de banda. La ganancia aumenta rápidamente con ω_0 . La cota superior para el ancho de banda viene dada por la validez del modelo. \square

El caso general

Como hay una relación entre la complejidad del modelo y la complejidad del controlador es natural preguntarse cuál es el modelo más general que dará controladores PI y PID. Un controlador PI tiene dos parámetros que son suficientes para caracterizar una ecuación de segundo orden; esto permite un modelo del proceso de primer orden. El sistema en el Ejemplo 6.6 es así el sistema más general donde la asignación de polos dará un controlador PI.

Como un controlador PID tiene tres parámetros, es posible determinar todos los parámetros de una ecuación de tercer orden. Con control PID es así posible usar asignación de polos para un sistema de segundo orden. El sistema de segundo orden más general no es el del Ejemplo 6.8, sino el del próximo ejemplo.

Si sólo se especifica una estructura de los polos es suficiente un controlador PI para un sistema de segundo orden y un controlador PID para un sistema de tercer orden.

EJEMPLO 6.9—SISTEMA GENERAL DE SEGUNDO ORDEN

Suponga que el proceso se caracteriza por el modelo de segundo orden

$$P(s) = \frac{b_1 s + b_2}{s^2 + a_1 s + a_2}. \quad (6.13)$$

Este modelo tiene cuatro parámetros. Tiene dos polos que pueden ser reales o complejos, y tiene un cero. Este modelo representa a muchos procesos, sistemas oscilatorios, y sistemas con ceros en el semiplano derecho. El cero en el semiplano derecho se puede usar también como una aproximación de un retardo de tiempo. Supongamos que el proceso se controla por un controlador PID parametrizado

como

$$C(s) = k + \frac{k_i}{s} + k_d s$$

$$C_{ff}(s) = bk + \frac{k_i}{s} + ck_d s.$$

El sistema en lazo cerrado es de tercer orden, y el polinomio característico es

$$s(s^2 + a_1 s + a_2) + (b_1 s + b_2)(k_d s^2 + ks + k_i).$$

Una ecuación característica en lazo cerrado apropiada de un sistema de tercer orden es

$$(s + \alpha\omega_0)(s^2 + 2\zeta\omega_0 s + \omega_0^2).$$

Igualando coeficientes de iguales potencias en s en estas ecuaciones da las siguientes ecuaciones:

$$\begin{aligned} a_1 + b_2 k_d + b_1 k &= (\alpha\omega_0 + 2\zeta\omega_0)(1 + b_1 k_d) \\ a_2 + b_2 k + b_1 k_i &= (1 + 2\alpha\zeta)\omega_0^2(1 + b_1 k_d) \\ b_2 k_i &= \alpha\omega_0^3(1 + b_1 k_d). \end{aligned}$$

Éste es un conjunto de ecuaciones lineales en los parámetros del controlador. La solución es directa pero tediosa y está dada por

$$\begin{aligned} k &= \frac{a_2 b_2^2 - a_2 b_1 b_2 (\alpha + 2\zeta) \omega_0 - (b_2 - a_1 b_1)(b_2(1 + 2\alpha\zeta)\omega_0^2 + \alpha b_1 \omega_0^3)}{b_2^3 - b_1 b_2^2 (\alpha + 2\zeta) \omega_0 + b_1^2 b_2 (1 + 2\alpha\zeta)\omega_0^2 - \alpha b_1^3 \omega_0^3} \\ k_i &= \frac{(-a_1 b_1 b_2 + a_2 b_1^2 + b_2^2) \alpha \omega_0^3}{b_2^3 - b_1 b_2^2 (\alpha + 2\zeta) \omega_0 + b_1^2 b_2 (1 + 2\alpha\zeta)\omega_0^2 - \alpha b_1^3 \omega_0^3} \\ k_d &= \frac{-a_1 b_2^2 + a_2 b_1 b_2 + b_2^2 (\alpha + 2\zeta) \omega_0 - b_1 b_2 \omega_0^2 (1 + 2\alpha\zeta) + b_1^2 \alpha \omega_0^3}{b_2^3 - b_1 b_2^2 (\alpha + 2\zeta) \omega_0 + b_1^2 b_2 (1 + 2\alpha\zeta)\omega_0^2 - \alpha b_1^3 \omega_0^3}. \end{aligned} \quad (6.14)$$

Estas fórmulas son bastante útiles porque muchos procesos pueden describirse aproximadamente por la función de transferencia dada por (6.13).

La función de transferencia desde el punto de consigna a la salida del proceso es

$$G_{yy_{sp}}(s) = \frac{(b_1 s + b_2)(c k_d s^2 + b k s + k_i)}{(s + \alpha \omega_0)(s^2 + 2\zeta\omega_0 s + \omega_0^2)}.$$

Los parámetros b y c tienen una fuerte influencia sobre la respuesta de esta función de transferencia. \square

Las fórmulas dadas en el Ejemplo 6.9 son particularmente útiles en los casos en los que estamos «llevando» el controlador PID a situaciones extremas. Las reglas de sintonía estándar típicamente no funcionarán en estos casos. Ejemplos comunes son sistemas con ceros en el semiplano derecho y sistemas con modos oscilatorios pobremente amortiguados. Para ilustrar esto consideraremos un ejemplo.

EJEMPLO 6.10—SISTEMA OSCILATORIO CON CEROS EN EL SEMIPLANO DERECHO
 Sea un sistema con la función de transferencia

$$P(s) = \frac{1-s}{s^2+1}.$$

Este sistema tiene un cero en el semiplano derecho y dos polos complejos no amortiguados. El proceso es difícil de controlar. Para proporcionar amortiguamiento para los polos no amortiguados en $s = \pm i$ es necesario tener una ganancia de control razonable en $\omega = 1$. Esto es difícil porque el cero en el semiplano derecho en $s = 1$ implica que la frecuencia de cruce de ganancia debería ser menor que 0,5 en orden a tener un sistema en lazo cerrado razonablemente robusto. Ninguno de los métodos estándar para la sintonía de controladores PID funciona bien para este sistema. Para aplicar el método de asignación de polos especificamos que el sistema en lazo cerrado tiene el polinomio característico

$$s^3 + 2s^2 + 2s + 1.$$

Las fórmulas en el Ejemplo 6.9 dan un controlador con los parámetros $k = 0$, $k_i = 1/3$, y $k_d = 2/3$. Esto se puede verificar también con un cálculo sencillo. Observe que la ganancia proporcional es cero y que el controlador tiene dos ceros complejos en $\pm i\sqrt{2}$. Tal controlador se puede implementar con un controlador PID que tiene la forma no interactuante. Comparar con la Sección 3.2. \square

Usar modelos aproximados

Como la asignación de polos sólo dará controladores PID si el modelo del proceso es de segundo orden o menor es necesario desarrollar modelos aproximados para poder utilizar asignación de polos. En la Sección 2.8 se presentaron diferentes métodos de aproximación. En esta sección ilustraremos el método con algunos ejemplos.

Sea un proceso descrito por la función de transferencia

$$P(s) = \frac{1}{(1+s)(1+0,2s)(1+0,05s)(1+0,01s)}. \quad (6.15)$$

Este proceso tiene cuatro constantes de tiempo de valores 1, 0,2, 0,05, y 0,01. Las aproximaciones se pueden hacer de diferentes maneras.

EJEMPLO 6.11—APROXIMACIÓN POR UN SISTEMA DE PRIMER ORDEN
 Si los requisitos de control no son demasiado estrictos, podemos intentar aproximar la función de transferencia por

$$P(s) = \frac{1}{1+1,26s},$$

donde la constante de tiempo es el tiempo de residencia medio. Tal como se analizó en la Sección 2.8, esta aproximación es buena en bajas frecuencias. La suma de

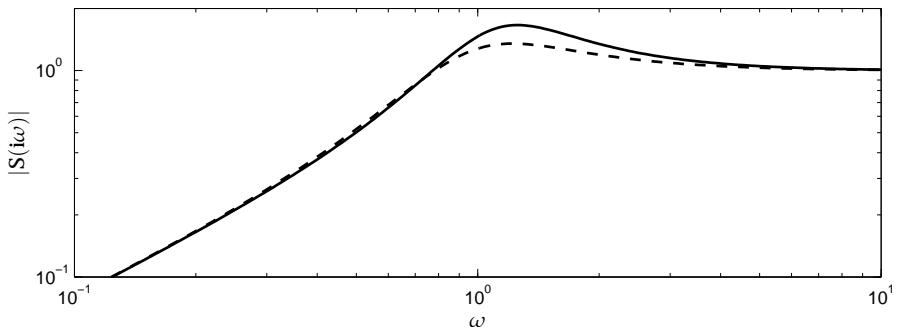


Figura 6.11 Funciones de sensibilidad para el sistema aproximado (a tramos) y el sistema verdadero en el Ejemplo 6.11.

las constantes de tiempo despreciadas es $T_e = 0,26$. El error de fase es menor que 15° para frecuencias por debajo de 1 rad/s. Si se diseña un controlador PI por el método de asignación de polos con $\zeta = 0,5$, se obtienen los siguientes parámetros del controlador,

$$K = 1,26\omega_0 - 1$$

$$T_i = \frac{1,26\omega_0 - 1}{1,26\omega_0^2}$$

$$b = \frac{1,26\omega_0}{1,26\omega_0 - 1}.$$

donde se escoge b de forma que el cero se hace $s = -\omega_0$. Si el modelo del proceso fuera correcto, el margen de fase con $\zeta = 0,5$ sería 50° . Debido a las aproximaciones realizadas, el margen de fase será menor. Disminuirá con ω_0 . Para $\omega_0 = 1$ el margen de fase es $\varphi_m = 42^\circ$. Los polos en lazo cerrado para el sistema son $-100, -20, -4,99, -0,46 \pm 1,02i$. Los polos en lazo cerrado obtenidos cuando se aplica el controlador al modelo simplificado son $-0,5 \pm 0,87i$. A causa de la aproximación los polos dominantes difieren de los valores de diseño. La diferencia aumenta con ω_0 crecientes. El sistema se hace inestable para $\omega_0 = 3,8255$.

La Figura 6.11 muestra las funciones de sensibilidad para el sistema exacto y el aproximado. Las sensibilidades máximas son $M_t = 1,35$ y $M_s = 1,66$, respectivamente. Esto indica que los polos en lazo cerrado deben escogerse con cuidado cuando se emplea el método de asignación de polos. \square

El siguiente ejemplo muestra lo que sucede cuando se aproxima el sistema por un modelo de segundo orden.

EJEMPLO 6.12—APROXIMACIÓN POR UN MODELO DE SEGUNDO ORDEN
Sea el sistema dado por (6.15). Aproximar la función de transferencia por

$$P(s) = \frac{1}{(1+s)(1+0,26s)}.$$

Se obtiene manteniendo la constante de tiempo más grande y aproximando las tres constantes de tiempo más cortas por su suma. La suma de las constantes

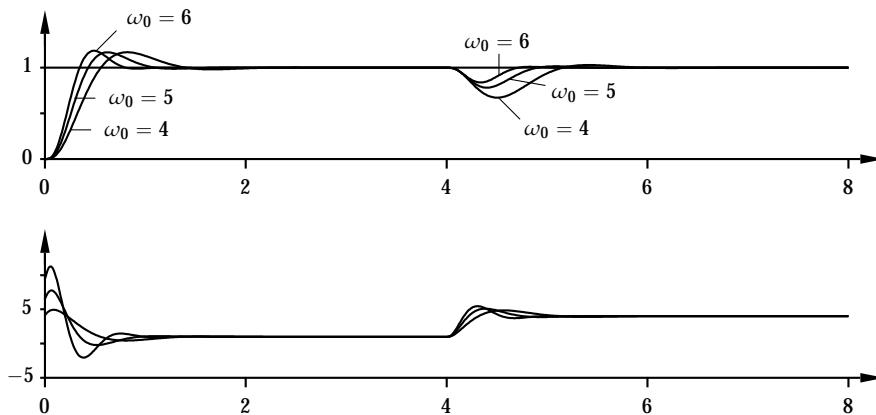


Figura 6.12 Respuestas a cambios en el punto de consigna y perturbaciones en la carga del proceso con dos polos controlados por un controlador PID sintonizado de acuerdo con el Ejemplo 6.12. Se muestran las respuestas para $\omega_0 = 4, 5$, y 6 . El diagrama superior muestra el punto de consigna $y_{sp} = 1$ y la salida del proceso y , y el diagrama inferior muestra la señal de control u .

de tiempo despreciadas es $T_e = 0,06$. El error de fase es menor que 15° para frecuencias por debajo de $4,4$ rad/s. Haciendo una aproximación del modelo del proceso que sea válida para frecuencias más elevadas que en el ejemplo previo, podemos diseñar un controlador más rápido. Si se escogen $\zeta = 0,5$ y $\alpha = 1$ en (6.16), los cálculos de diseño en el Ejemplo 6.8 dan los siguientes parámetros para el controlador PID:

$$\begin{aligned} K &= 0,52\omega_0^2 - 1 \\ T_i &= \frac{0,52\omega_0^2 - 1}{0,26\omega_0^3} \\ T_d &= \frac{0,52\omega_0 - 1,26}{0,52\omega_0^2 - 1} \\ b &= \frac{0,26\omega_0^2}{0,52\omega_0^2 - 1}. \end{aligned} \tag{6.16}$$

En este caso, se obtiene control PI puro para $\omega_0 = 2,4$. La ganancia derivativa se hace negativa para anchos de bandas más cortos. La aproximación desprecia la constante de tiempo 0,05. Si se requieren las dinámicas despreciadas para dar un error de fase, de a lo más, $0,3$ rad (17 grad) se puede obtener con un ancho de banda de, $\omega_0 < 6$ rad/s. En la Figura 6.12, se demuestra la conducta del control para $\omega_0 = 4, 5$, y 6 .

La especificación del ancho de banda en lazo cerrado deseado es clave, ya que la ganancia del controlador aumenta rápidamente con el ancho de banda especificado. Es también crucial para conocer el rango de frecuencia donde el modelo es válido. Alternativamente, se puede emplear una cota superior para la ganancia del controlador para limitar el ancho de banda. Observe el efecto de cambiar la frecuencia de diseño ω_0 . El sistema con $\omega_0 = 6$ responde más rápido

y tiene un error más pequeño cuando está sujeto a perturbaciones en la carga. El diseño no funcionará bien cuando se aumenta ω_0 por encima de 8. \square

Diseño de polos dominantes

En el diseño por asignación de polos se intenta asignar todos los polos en lazo cerrado. Una desventaja del método es que es difícil especificar muchos polos en lazo cerrado. En la Sección 4.5 se mencionó que la conducta de un sistema puede con frecuencia caracterizarse por unos pocos polos dominantes. Se puede por lo tanto intentar colocar algunos polos dominantes. Ilustraremos esto con algunos ejemplos.

EJEMPLO 6.13—UN CONTROLADOR TIPO INTEGRADOR PURO

Sea un proceso con la función de transferencia $P(s)$ y un controlador que es un integrador

$$C(s) = \frac{k_i}{s}.$$

Los polos en lazo cerrado son

$$1 + k_i \frac{P(s)}{s} = 0.$$

Como el controlador tiene un parámetro ajustable, es posible asignar un polo. Para obtener un polo en $s = -a$ el parámetro del controlador debería escogerse como

$$k_i = \frac{a}{P(-a)}. \quad (6.17)$$

Para obtener una buena atenuación de las perturbaciones de carga seleccionaremos el polo en lazo cerrado de forma que la ganancia integral k_i sea lo más grande posible. Por ejemplo, si

$$P(s) = \frac{1}{(s+1)^2},$$

obtenemos

$$k_i = a(-a+1)^2 = a^3 - 2a^2 + a,$$

que tiene su máximo $4/27$ para $a = 1/3$. \square

EJEMPLO 6.14—CONTROL PI

Un controlador PI tiene dos parámetros. En consecuencia, es necesario asignar dos polos. Sea un proceso con función de transferencia $P(s)$, y el controlador parametrizado como

$$C(s) = k + \frac{k_i}{s}.$$

La ecuación característica en lazo cerrado es

$$1 + \left(k + \frac{k_i}{s}\right) P(s) = 0.$$

Requiere que esta ecuación tenga raíces en

$$p_{1,2} = \omega_0 \left(-\zeta_0 \pm i\sqrt{1 - \zeta_0^2} \right) = \omega_0 e^{i(\pi \pm \gamma)} = \omega_0 (-\cos \gamma \pm i \sin \gamma),$$

donde $\gamma = \arccos \zeta_0$. La condición de que el sistema en lazo cerrado tiene un polo p_1 es así

$$1 + \left(k + \frac{k_i}{p_1} \right) P(p_1) = 0. \quad (6.18)$$

Ésta es una ecuación lineal con variable compleja con dos incógnitas. Para resolverla introducimos $a(\omega_0)$ y $\phi(\omega_0)$, definido como

$$P\left(\omega_0 e^{i(\pi-\gamma)}\right) = a(\omega_0) e^{i\phi(\omega_0)}.$$

Observe que $P\left(\omega_0 e^{i(\pi-\gamma)}\right)$ representa los valores de la función de transferencia sobre el rayo $e^{i(\pi-\gamma)}$. Cuando $\gamma = \pi/2$, entonces $P\left(\omega_0 e^{i(\pi-\gamma)}\right) = P(i\omega_0)$, que es la respuesta en frecuencia normal.

La Ecuación 6.18 se puede escribir como

$$1 + \left(k + \frac{k_i}{\omega_0 e^{i(\pi-\gamma)}} \right) a(\omega_0) e^{i\phi(\omega_0)} = 0.$$

esta ecuación que es lineal en k y k_i , tiene la solución

$$\begin{aligned} k &= -\frac{\sin(\phi(\omega_0) + \gamma)}{a(\omega_0) \sin \gamma} \\ k_i &= -\frac{\omega_0 \sin \phi(\omega_0)}{a(\omega_0) \sin \gamma}. \end{aligned} \quad (6.19)$$

Observe que $\phi(\omega_0)$ es cero para $\omega_0 = 0$ y típicamente negativo cuando ω_0 aumenta. Esto implica que la ganancia proporcional es negativa y la ganancia integral positiva pero pequeña para ω_0 pequeña. Cuando ω_0 aumenta ambos k y k_i aumentarán inicialmente. Para valores más grandes de ω_0 ambos parámetros disminuirán. Requeriendo que ambos parámetros sean positivos, encontramos que ω_0 se debe seleccionar de forma que

$$\gamma < -\phi(\omega_0) < \pi.$$

El tiempo integral del controlador es

$$T_i = \frac{k}{k_i} = \frac{\sin(\phi(\omega_0) + \gamma)}{\omega_0 \sin \phi(\omega_0)}.$$

Observe que T_i es independiente de $a(\omega_0)$. □

Cuadro 6.6 Parámetros del controlador para el diseño de polos dominantes de un controlador PI para un proceso que es un retardo de tiempo puro.

ζ	k	$k_i L$	T_i/L	$\omega_0 L$	M_s	IAE/L
0,1	0,388	1,50	0,258	1,97	6,34	4,03
0,2	0,343	1,27	0,270	1,93	3,60	2,42
0,5	0,244	0,847	0,288	1,86	1,99	1,56
0,707	0,195	0,688	0,284	1,88	1,69	1,54
1,0	0,135	0,541	0,250	2,00	1,49	1,85

EJEMPLO 6.15—UN PROCESO RETARDO PURO

Sea un proceso con la función de transferencia

$$P(s) = e^{-sL}.$$

Utilizando control integral puro, se sigue de la Ecuación 6.17 que $k_i = ae^{-aL}$. La ganancia tiene su valor más grande $k_i = e^{-1}/L$ para $a = 1/L$. La función de transferencia del lazo es entonces

$$G_I(s) = P(s)C(s) = \frac{1}{esL} e^{-sL}.$$

La sensibilidad del sistema es $M_s = 1,39$, que es un valor razonable.

Con control PI se sigue de la Ecuación 6.19 que

$$k = \frac{\sin(\omega_0 L \sin \gamma - \gamma)}{\sin \gamma} e^{-\omega_0 L \cos \gamma}$$

$$k_i = \omega_0 \frac{\sin(\omega_0 L \sin \gamma)}{\sin \gamma} e^{-\omega_0 L \cos \gamma}.$$

Para minimizar IE , determinamos el valor de ω_0 que maximiza k_i . En la Tabla 6.6 se muestran los resultados. Esta tabla también da los valores M_s y el IAE . El IAE tiene un mínimo en $\zeta \approx 0,6$. Observe que hay variaciones significativas en la ganancia pero que los valores del tiempo de integración son relativamente constante para todos los valores del parámetro de diseño ζ . El valor de IAE debería ser pequeño para dar buen rechazo de la perturbación de carga, y M_s debería ser pequeño para dar buena robustez. La tabla ilustra el compromiso entre estos objetivos. Para obtener una robustez razonable de $M_s < 2$, el amortiguamiento relativo debería ser mayor que 0,5.

Observe que para $\zeta = 1$ obtenemos $k = e^{-2}$ y $k_i = 4e^{-2}/L$. Esto se puede comparar con $k_i = e^{-1}L$ para control I puro. Con control PI la ganancia integral puede así aumentarse por un factor de 1.5 comparado con un controlador I. Para un sistema bien amortiguado ($\zeta = 0,707$) la ganancia es aproximadamente 0,2 y el tiempo integral es $T_i = 0,28L$. Esto se puede comparar con los valores 0,45 y 2L obtenidos con el método de respuesta en frecuencia de Ziegler-Nichols. El diseño

de polos dominantes da así un controlador con una acción integral mucho más fuerte que el método de Ziegler-Nichols. En el Ejemplo 6.4 encontramos que esto era altamente deseable.

En resumen, encontramos que un proceso con una dinámica de retardo puro se puede controlar bastante bien con un controlador PI. \square

El diseño de polos dominantes es un caso especial de asignación de polos donde sólo se intenta colocar algunos polos dominantes. Para controladores P, I, o D puros se puede colocar un polo. Para controladores PI y PD hay dos polos dominantes, que pueden ser convenientemente parametrizados con el amortiguamiento relativo ζ . El método se hace más complicado para control PID. Despues del diseño es necesario comprobar que los polos en lazo cerrado obtenidos son realmente dominantes. Es también necesario evaluar la robustez del sistema en lazo cerrado.

Comentario

La asignación de polos es un método estándar para el diseño de sistemas de control. Las especificaciones se dan en términos de todos los polos del sistema en lazo cerrado o posiblemente solamente la estructura de los polos. Se requiere un buen juicio para escoger adecuadamente los polos. Cuando se emplea asignación de polos la complejidad del controlador se determina a partir de la complejidad del modelo del proceso. Para obtener un controlador PID se requiere que el modelo sea de bajo orden o que el modelo se aproxime por un modelo de bajo orden. Los retardos de tiempo son a menudo aproximados cuando se usa asignación de polos. No hay forma natural de introducir una restricción de robustez en la asignación de polos. El sistema en lazo cerrado resultante debe analizarse para asegurar que es suficientemente robusto.

6.5 Sintonía lambda

La sintonía lambda es un caso especial de asignación de polos que se utiliza con frecuencia en la industria de procesos. El proceso se modela por el modelo FOTD

$$P(s) = \frac{K_p}{1 + sT} e^{-sL}.$$

Diferentes aproximaciones del retardo de tiempo L resulta en ambos controladores PI y PID.

Control PI

Si se usa un controlador PI con la función de transferencia

$$C(s) = K \frac{1 + sT_i}{sT_i}$$

con el tiempo integral T_i escogido igual a la constante de tiempo T del proceso, la función de transferencia del lazo se hace

$$G_I(s) = P(s)C(s) = \frac{K_p K}{sT} e^{-sL} \approx \frac{K_p K(1 - sL)}{sT},$$

donde la función exponencial ha sido aproximada usando un desarrollo en serie de Taylor. La ecuación característica del sistema en lazo cerrado es

$$s(T - K_p K L) + K_p K = 0.$$

Requeriendo que el polo en lazo cerrado es $s = -1/T_{cl}$, donde T_{cl} es la constante de tiempo en lazo cerrado deseada, encontramos

$$K_p K = \frac{T}{L + T_{cl}},$$

que da la siguiente sencilla regla de sintonía

$$\begin{aligned} K &= \frac{1}{K_p} \frac{T}{L + T_{cl}} \\ T_i &= T. \end{aligned} \tag{6.20}$$

El tiempo de repuesta en lazo cerrado T_{cl} es el parámetro de diseño. En el trabajo original de Dahlin [92] fue denotado como $T_{cl} = \lambda$, lo que explica el nombre de sintonía lambda.

La elección de T_{cl} es crítica. Una regla heurística muy común es escoger $T_{cl} = 3T$ para un controlador robusto y $T_{cl} = T$ para sintonía agresiva cuando los parámetros del proceso están bien determinados. Ambas elecciones conducen a controladores con ganancia cero y tiempo integral cero para sistemas con retardo de tiempo puro. Para procesos dominados por el retardo es por lo tanto algunas veces recomendado escoger T_i como el valor mas grande de T y $3L$.

Una desventaja con la sintonía lambda es que se cancela el polo del proceso. Esto no es serio para procesos dominados por el retardo. La ganancia integral es

$$k_i = \frac{K}{T_i} = \frac{1}{K_p(L + T_{cl})}.$$

Cuando T_{cl} es proporcional a T la ganancia integral es pequeña para T grandes. La respuesta a perturbaciones de carga es así muy pobre para proceso dominados por la constante de tiempo.

En este último caso es útil hacer un diseño que no cancele el polo del proceso. Cuando el proceso FOTD se controla con un controlador PI la función de transferencia del lazo es

$$G_I(s) = P(s)C(s) = \frac{K_p K (1 + sT_i)e^{-sL}}{sT_i(1 + sT)} \approx \frac{K_p K (1 + sT_i)(1 - sL)}{sT_i(1 + sT)},$$

donde la función exponencial ha sido aproximada por un desarrollo en serie de Taylor. La ecuación característica es de segundo orden:

$$s^2 \left(\frac{T_i T}{K_p K} - T_i L \right) + s \left(T_i + \frac{T_i}{K_p K} - L \right) + 1 = 0.$$

Comparando esto con la ecuación característica deseada,

$$s^2 T_{cl}^2 + 2\zeta T_{cl}s + 1 = 0,$$

da los parámetros del controlador

$$\begin{aligned} K &= \frac{L + 2\zeta T_{cl}}{T_{cl}^2 + T_{cl}^2/(K_p K) + 2\zeta T_{cl}L + L^2} \\ T_i &= \frac{K_p K(L + 2\zeta T_{cl})}{1 + K_p K}. \end{aligned} \quad (6.21)$$

Estas reglas de sintonía se pueden aplicar también a procesos integrantes a condición de que T_{cl} se escoja adecuadamente. Para procesos dominados por la constante de tiempo es razonable escoger T_{cl} proporcional a L .

Control PID

Para la obtención del diseño PID, se usa la forma interactuante del controlador PID (3.8):

$$C'(s) = K' \frac{(1 + sT'_i)(1 + sT'_d)}{sT'_i}.$$

Se aproxima el retardo de tiempo utilizando (2.59), lo que da la función de transferencia del proceso

$$P(s) = \frac{K_p}{1 + sT} e^{-sL} \approx \frac{K_p(1 - sL/2)}{(1 + sT)(1 + sL/2)}.$$

El tiempo integral se escoge como $T'_i = T$ y el tiempo derivativo como $T'_d = L/2$. Los ceros del controlador entonces cancelarán los polos del proceso, y la función de transferencia del lazo se hace

$$G_l(s) = P(s)C'(s) \approx \frac{K_p K' (1 - sL/2)}{sT}.$$

La ecuación característica es

$$s(T - K_p K' L/2) + K_p K' = 0.$$

Requeriendo que el polo en lazo cerrado sea $s = -1/T_{cl}$ encontramos

$$K_p K' = \frac{T}{L/2 + T_{cl}},$$

que da las siguientes sencillas reglas de sintonía:

$$\begin{aligned} K' &= \frac{1}{K_p} \frac{T}{L/2 + T_{cl}} \\ T'_i &= T \\ T'_d &= \frac{L}{2}. \end{aligned}$$

Utilizando (3.9), los parámetros correspondientes para el controlador PID no interactuante son

$$\begin{aligned} K &= \frac{1}{K_p} \frac{L/2 + T}{L/2 + T_{cl}} \\ T_i &= T + L/2 \\ T_d &= \frac{TL}{L + 2T}. \end{aligned} \quad (6.22)$$

Observe que no hay acción derivativa para procesos que sean retardo puro ($T = 0$).

Comentario

La sintonía lambda es un caso especial de asignación de polos. Es un método sencillo que puede dar buenos resultados en ciertas circunstancias a condición de que el parámetro de diseño se escoge adecuadamente. El método básico cancela un polo del proceso lo que se traducirá en una pobre respuesta de las perturbaciones de carga en procesos dominados por constante de tiempo. Se pueden hacer diversas modificaciones ad hoc, pero esto requiere un cierto conocimiento del proceso.

6.6 Diseño algebraico

Hay algunos métodos de sintonía algebraicos donde la función de transferencia del controlador se obtiene de las especificaciones por un cálculo algebraico directo. Los métodos están estrechamente relacionados con la asignación de polos.

Formas estándar

Una cuestión fundamental es determinar las funciones de transferencia que dan respuestas apropiadas frente a cambios en el punto de consigna. Esto se puede hacer comenzando con una función de transferencia de una forma dada y determinando los parámetros de manera que se minimice algún criterio del error tales como IAE, ISE, o ITAE.

Ejemplos típicos son

$$\begin{aligned} G_1 &= \frac{\omega_0^2}{s^2 + 2\zeta\omega_0 + \omega_0^2} \\ G_2 &= \frac{\alpha\omega_0^3}{(s^2 + 2\zeta\omega_0 s + \omega_0^2)(s + \alpha\omega_0)} \\ G_3 &= \frac{\omega_0(s + \beta\omega_0)}{\beta(s^2 + 2\zeta\omega_0 s + \omega_0^2)} \\ G_4 &= \frac{\alpha\omega_0^2(s + \beta\omega_0)}{\beta(s^2 + 2\zeta\omega_0 s + \omega_0^2)(s + \alpha\omega_0)}. \end{aligned} \quad (6.23)$$

El parámetro ω_0 es un factor de escala que determina la velocidad de respuesta. Los parámetros α , β , y ζ determinan la forma de las funciones de transferencia. El

amortiguamiento relativo ζ está típicamente en el rango de 0,5 a 1. Los parámetros α y β tienen una influencia significativa si son menores que uno. Disminuyendo α hace la respuesta más lenta y reduce la sobreelongación. Disminuyendo β hace la respuesta más rápida y aumenta la sobreelongación. Han habido muchos esfuerzos para encontrar parámetros que optimizan diversos criterios. Sea un sistema donde el proceso tiene la función de transferencia $P(s)$ y las funciones de transferencia del controlador son

$$C(s) = K \left(1 + \frac{1}{sT_i} + sT_d \right)$$

$$C_{ff}(s) = K \left(b + \frac{1}{sT_i} + scT_d \right).$$

La función de transferencia en lazo cerrado desde el punto de consigna hasta la salida del proceso es entonces

$$G_{yy_{sp}} = \frac{PC_{ff}}{1 + PC}.$$

En primer lugar se escogen los parámetros del controlador K , T_i , y T_d de forma que coincidan con el denominador de la función de transferencia especificada, y se seleccionan los pesos del punto de consigna b y c para que concuerde el numerador de la función de transferencia especificada. Como los controladores que son sencillos tienen sólo unos pocos parámetros es necesario que las funciones de transferencia escogidas sean suficientemente simples.

Para sistemas con realimentación del error donde $C(s) = C_{ff}(s)$ es posible dar una expresión explícita para la función de transferencia del controlador:

$$C = \frac{1}{P} \cdot \frac{G_{yy_{sp}}}{1 - G_{yy_{sp}}}. \quad (6.24)$$

Para asegurarse de que el controlador obtenido es un controlador PID es necesario hacer aproximaciones o cancelaciones tal como se analizó en la Sección 2.8.

Se sigue de (6.24) que todos los polos y ceros del proceso se cancelan por el controlador a menos que $G_{yy_{sp}}$ tenga los correspondientes polos y ceros. Esto significa que la realimentación del error no puede aplicarse cuando el proceso tiene polos y ceros mal amortiguados. El método también dará una pobre respuesta a perturbaciones en la carga cuando se cancelan los polos lentos del proceso.

Hay muchas versiones diferentes de los métodos de diseño algebraico. Es suficiente presentar algunos casos.

Método de Haalman

Para sistemas con un retardo de tiempo L , Haalman ha sugerido seleccionar la función de transferencia del lazo

$$G_l(s) = P(s)C(s) = \frac{2}{3Ls} e^{-sL}.$$

El valor $2/3$ se encontró minimizando el error cuadrático medio para un cambio en escalón en el punto de consigna. Esta elección da una sensibilidad $M_s = 1,9$,

que es un valor razonable. Observe que es sólo el retardo de tiempo del proceso lo que influye sobre la función de transferencia del lazo. Todos los otros polos y ceros del proceso se cancelan, lo que puede conducir a dificultades.

Aplicando el método de Haalman a un proceso con la función de transferencia

$$P(s) = \frac{K_p}{1 + sT} e^{-sL}$$

da el controlador

$$C(s) = \frac{2(1+sT)}{3K_p L s} = \frac{2T}{3K_p L} \left(1 + \frac{1}{sT}\right),$$

que es un controlador PI con $K = 2T/3K_p L$ y $T_i = T$. Estos parámetros se pueden comparar con los valores $K = 0,9T/L$ y $T_i = 3L$ obtenidos por el método de la respuesta a un escalón de Ziegler-Nichols.

Comparando el método de Haalman con la sintonía lambda encontramos que los tiempos integrales son los mismos que las ganancias son las mismas si seleccionamos $T_{cl} = L/2$. Como la sintonía lambda se basa en aproximaciones del retardo de tiempo parece más razonable usar el método de Haalman cuando L es grande. .

Aplicando el método de Haalman a un proceso con la función de transferencia

$$P(s) = \frac{K_p}{(1 + sT_1)(1 + sT_2)} e^{-sL}$$

da un controlador PID con parámetros $K = 2(T_1 + T_2)/3K_p L$, $T_i = T_1 + T_2$, y $T_d = T_1 T_2 / (T_1 + T_2)$. Para procesos más complejos es necesario aproximarlos para obtener una función de transferencia de la forma deseada tal como se analizó en Sección 2.8.

La Figura 6.13 muestra una simulación del método de Haalman para un sistema con retardo de tiempo normalizado $\tau = 0,5$. La figura muestra que las respuestas son buenas.

Peligros por la cancelación de los polos del proceso que son lentos

Una característica del método de Haalman es que los polos y ceros del proceso se cancelan por polos y ceros en el controlador. Cuando se cancelan los polos y ceros, habrá modos no controlables en el sistema en lazo cerrado. Esto puede llevar a un pobre comportamiento si se excitan dichos modos. El problema es particularmente grave si los modos cancelados son lentos o inestables. Utilizamos un ejemplo para ilustrar lo que puede acontecer.

EJEMPLO 6.16—PÉRDIDA DE CONTROLABILIDAD DEBIDA A LA CANCELACIÓN
Sea un sistema en lazo cerrado donde un proceso con la función de transferencia

$$P(s) = \frac{1}{1 + sT} e^{-sL}$$

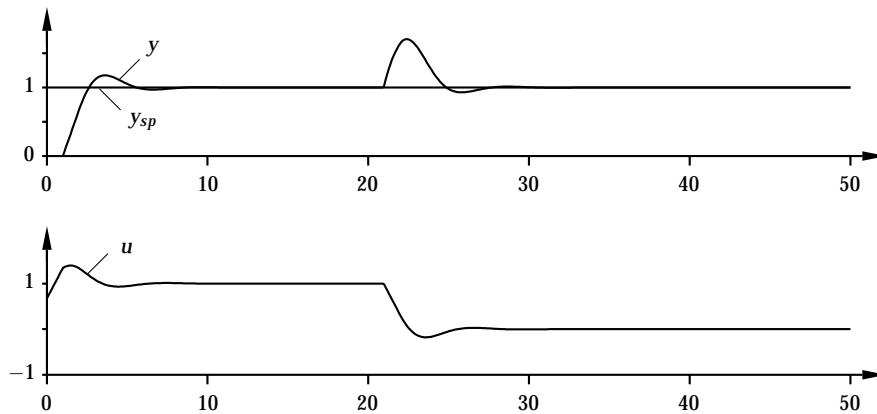


Figura 6.13 Simulación de un sistema en lazo cerrado obtenido por el método de Haalman. La función de transferencia de la planta es $P(s) = e^{-s}/(s+1)$. Los diagramas muestran el punto de consigna y_{sp} , la salida del proceso y , y la señal de control u .

se controla con un controlador PI cuyos parámetros se escogen de forma que se cancela el polo del proceso. La función de transferencia del controlador es entonces

$$C(s) = K \left(1 + \frac{1}{sT} \right) = K \frac{1 + sT}{sT}.$$

El proceso se puede representar por la ecuación

$$\frac{dy(t)}{dt} = \frac{1}{T} (u(t-L) - y(t)), \quad (6.25)$$

y el controlador se puede describir por

$$\frac{du(t)}{dt} = -K \left(\frac{dy(t)}{dt} + \frac{y(t)}{T} \right). \quad (6.26)$$

Considérese la conducta del sistema en lazo cerrado cuando se escogen las condiciones iniciales como $y(0) = 1$ y $u(t) = 0$ para $-L < t < 0$. Sin realimentación la salida está dada por

$$y_{ol}(t) = e^{-t/T}.$$

Para calcular la salida para el sistema en lazo cerrado primero eliminamos $y(t)$ entre (6.25) y (6.26). Esto da

$$\frac{du(t)}{dt} = -\frac{K}{T} u(t-L).$$

Se sigue así que $u(t) = 0$, y (6.25) entonces implica que

$$y_{cl}(t) = e^{-t/T} = y_{ol}(t).$$

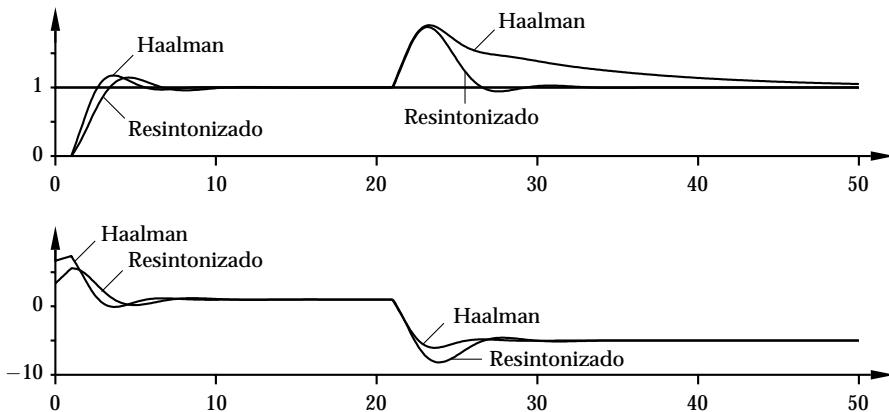


Figura 6.14 Simulación de un sistema en lazo cerrado obtenido por el método de Haalman. La función de transferencia del proceso es $P(s) = e^{-s}/(10s + 1)$, y los parámetros del controlador son $K = 6,67$ y $T_i = 10$. El diagrama superior muestra el punto de consigna $y_{sp} = 1$ y la salida del proceso y , y el diagrama inferior muestra la señal de control u . La figura también muestra las respuestas con un controlador resintonizado con $K = 6,67$, $T_i = 3$, y $b = 0,5$.

Las trayectorias del sistema en lazo cerrado y del sistema en lazo abierto son así las mismas. La señal de control es cero, lo que significa que el controlador no intenta reducir el error de control. \square

El ejemplo claramente indica que hay desventajas con la cancelación de los polos del proceso. En la Figura 6.14 se muestra otra ilustración del fenómeno, que es una simulación de un sistema en lazo cerrado donde se diseña el controlador por el método de Haalman. Esta simulación es idéntica a la simulación de la Figura 6.13, pero la constante de tiempo del proceso es ahora 10 en lugar de 1 para la simulación de la Figura 6.14.

En este caso encontramos que la respuesta del punto de consigna es excelente pero la respuesta a perturbaciones en la carga es muy pobre. La razón para esto es que el controlador cancela el polo $s = -0,1$ al tener un cero del controlador en $s = -0,1$. Observe que la salida del proceso después de una perturbación en la carga decrece con la constante de tiempo $T = 10$ pero que la señal de control es prácticamente constante debido a la cancelación. La atenuación de las perturbaciones de carga se mejora considerablemente al reducir el tiempo integral del controlador tal como se muestra en la Figura 6.14.

Hemos así mostrado que la cancelación de los polos del proceso pueden dar sistemas con pobre rechazo de las perturbaciones de carga. Observe que esto no se refleja en las simulaciones a menos que se excite al proceso. Por ejemplo, no se observará en una simulación de una respuesta en escalón en el punto de consigna. Podemos también preguntarnos por qué no hay una gran diferencia en las simulaciones de las Figuras 6.13 y 6.14. La razón es que el polo cancelado en la Figura 6.14 es lento en comparación con los polos en lazo cerrado, pero es de la misma magnitud que los polos en lazo cerrado en la Figura 6.13.

Podemos así concluir que la cancelación de polos se puede hacer para sistemas de retardo dominante pero no para sistemas en los que dominan las constantes

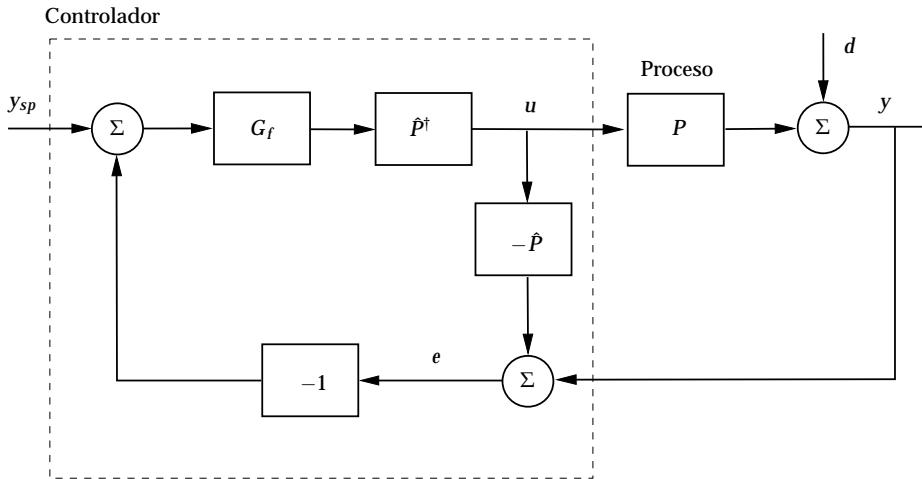


Figura 6.15 Diagrama de bloques de un sistema en lazo cerrado con un controlador basado en el principio del modelo interno.

de tiempo.

Control con modelo interno (IMC)

El principio del modelo interno es un método general para el diseño de sistemas de control que se puede aplicar al control PID. En la Figura 6.15 se muestra un diagrama de bloques de tales sistemas. En el diagrama se supone que todas las perturbaciones que actúan sobre el proceso se reducen a una perturbación equivalente d en la salida del proceso. En la figura \hat{P} denota un modelo del proceso, \hat{P}^\dagger es una inversa aproximada de \hat{P} , y G_f es un filtro paso baja. El nombre *controlador con modelo interno* deriva del hecho de que el controlador contiene un modelo del proceso internamente. Este modelo se conecta en paralelo con el proceso.

Si el modelo coincide con el proceso, i.e., $\hat{P} = P$, la señal e es igual a la perturbación d para todas las señales de control u . Si $G_f = 1$ y \hat{P}^\dagger es una inversa exacta del proceso, entonces la perturbación d se cancelará perfectamente. El filtro G_f se introduce para obtener un sistema que es menos sensible a los errores de modelado. Una elección común es $G_f(s) = 1/(1 + sT_f)$, donde T_f es un parámetro de diseño.

El controlador obtenido por el principio del modelo interno se puede representar como un controlador serie ordinario con la función de transferencia

$$C = \frac{G_f \hat{P}^\dagger}{1 - G_f \hat{P}^\dagger \hat{P}}. \quad (6.27)$$

De esta expresión se sigue que los controladores de este tipo cancelan los polos y ceros del proceso.

El principio del modelo interno dará típicamente controladores de orden más alto. Sin embargo, es posible obtener desde el principio, haciendo hipótesis especiales, controladores PI o PID. Para ver esto considere un proceso con la función

de transferencia

$$P(s) = \frac{K_p}{1+sT} e^{-sL}. \quad (6.28)$$

Una inversa aproximada está dada por

$$\hat{P}^\dagger(s) = \frac{1+sT}{K_p}.$$

Observe que no se intenta encontrar una inversa del retardo de tiempo. Escogiendo el filtro

$$G_f(s) = \frac{1}{1+sT_f},$$

y aproximan el retardo de tiempo por

$$e^{-sL} \approx 1 - sL,$$

La ecuación 6.27 da ahora

$$C(s) = \frac{1+sT}{K_p s(L+T_f)},$$

que es un controlador PI. Observe que este controlador es idéntico al obtenido con la sintonía lambda si $T_f = T_{cl}$; ver Ecuación 6.20.

Si en lugar de esto el retardo de tiempo se approxima por una aproximación de Padé de primer orden,

$$e^{-sL} \approx \frac{1 - sL/2}{1 + sL/2},$$

La Ecuación 6.27 da en su lugar el controlador PID

$$C(s) = \frac{(1 + sL/2)(1 + sT)}{K_p s(L + T_f + sT_f L/2)} \approx \frac{(1 + sL/2)(1 + sT)}{K_p s(L + T_f)}.$$

Para procesos descritos por la Ecuación 6.28, encontramos así que el principio del modelo interno producirá controladores PI o PID. Se pueden utilizar de la manera usual aproximaciones como las analizadas en la Sección 2.8 para obtener controladores PI y PID para procesos más complejos.

Una característica interesante del controlador con modelo interno es que la robustez se considera explícitamente en el diseño. La robustez se puede ajustar seleccionando adecuadamente el filtro G_f . Se puede encontrar un compromiso entre comportamiento y robustez empleando la constante del filtro como un parámetro de diseño. El controlador IMC se puede diseñar para dar una excelente respuesta frente a cambios en el punto de consigna. Como el método de diseño inherentemente implica que se cancelan los polos y los ceros de la planta, la respuesta a las perturbaciones de la carga pueden ser pobres si los polos cancelados son lentos en comparación con los polos dominantes. Comparar con las respuestas en la Figura 6.14. El controlador IMC pueden también verse como una extensión del predictor de Smith; ver Capítulo 8.

Controlador con modelo interno de Skogestad (SIMC)

Skogestad ha desarrollado una versión del método de sintonía del control con modelo interno para control PID que evita algunas de las desventajas anteriormente mencionadas. El método comienza con un modelo FOTD para control PI o un modelo SOTD para control PID. Se requiere que el sistema en lazo cerrado debería tener la función de transferencia

$$G_{yysp} = \frac{1}{1 + sT_{cl}} e^{-sL}$$

Para un sistema FOTD se sigue entonces de la Ecuación 6.24 que la función de transferencia del controlador es

$$C(s) = \frac{1 + sT}{K_p(1 + sT_{cl} - e^{-sL})} \approx \frac{1 + sT}{sK_p(T_{cl} + L)}$$

donde la función exponencial se aproxima utilizando un desarrollo en serie de Taylor. En contraste con la recomendación para control IMC la constante de tiempo en lazo cerrado es proporcional al retardo de tiempo L . Se recomienda la elección $T_{cl} = L$. Se modifica también el término integral para procesos dominados por la constante de tiempo. La regla de sintonía para control PI es

$$\begin{aligned} K &= \frac{T}{2K_pL} \\ T_i &= \min(T, 8L) \end{aligned} \tag{6.29}$$

Se utilizan los mismos parámetros para un controlador PID en forma serie, y el tiempo derivativo se escoge como la constante de tiempo más pequeña.

Comentario

Los métodos de diseño analíticos están muy estrechamente relacionados con la asignación de polos. La principal diferencia es que especifica la función de transferencia completa en lugar de los polos en lazo cerrado. Una característica agradable es que los cálculos requeridos son muy simples. Una desventaja es que se cancelan los polos del proceso. Esto es particularmente serio para sistemas dominados por una constante de tiempo.

6.7 Métodos de optimización

La optimización es una herramienta poderosa para el diseño de controladores. El método es conceptualmente simple. Se especifica una estructura del controlador con unos pocos parámetros. Las especificaciones se expresan como inigualdades de funciones de los parámetros. La especificación que es más importante se escoge como la función a optimizar. El método se adapta bien a los controladores PID donde se dan la estructura del controlador y su parametrización. Hay algunas trampas cuando se utiliza la optimización. Hay que tener cuidado cuando se formula el criterio y las restricciones; en algunos casos, un criterio será realmente

Cuadro 6.7 Parámetros del controlador obtenidos de la minimización de la integral del error absoluto, IAE. K_{hf} es la ganancia de alta frecuencia del controlador.

L	IAE	M_s	K_{hf}	aK	T_i/L
0,0	0		∞		
0,2	0,14	3,3	4,7	0,94	2,9
0,5	0,60	3,0	2,0	1,0	2,2
1,0	1,5	2,4	1,0	1,0	1,4
2,0	3,2	2,1	0,60	1,2	1,0
5,0	7,7	2,0	0,42	2,1	0,6
10,0	15	1,9	0,37	3,7	0,53

óptimo, pero el controlador puede todavía no ser apropiado ya que no se ha considerado una restricción. Otra dificultad es que la función de coste puede tener muchos mínimos locales. Una tercera es que los cálculos requeridos pueden ser con facilidad excesivos. Pueden también surgir problemas numéricos. Sin embargo, la optimización es una buena herramienta que ha sido utilizada satisfactoriamente para el diseño de controladores PID. En esta sección analizamos algunos de estos métodos.

Un aviso

Los algoritmos de optimización son muy potentes. Resolverán cualquier criterio que se formule. Es por lo tanto muy importante formular los problemas correctamente e introducir todas las restricciones que son relevantes. En el control PID es particularmente importante introducir restricciones de robustez. Esto con frecuencia no se ha considerado en muchos trabajos sobre el uso de la optimización en control PID. El ejemplo que sigue ilustra lo que puede acontecer.

EJEMPLO 6.17—UN CONTROLADOR PI OPTIMIZADO PARA IAE

Sea un proceso con la función de transferencia

$$P(s) = \frac{1}{s+1} e^{-sL}.$$

La Tabla 6.7 da los parámetros del controlador que minimizan IAE para perturbaciones en la carga. En la tabla se muestran también algunos otros criterios. La tabla muestra que la integral del error absoluto aumenta con L , tal como cabría esperar. La tabla muestra que la sensibilidad máxima es grande para prácticamente todos los sistemas, particularmente aquellos con pequeños L . La tabla también indica que la ganancia del controlador en alta frecuencia es elevada para valores pequeños de L . Por ejemplo, si requerimos $M_s < 1,8$, que es una exigencia de robustez relativamente modesta, ninguno de los sistemas es aceptable. \square

El ejemplo ilustra la necesidad de considerar todos los aspectos de un problema cuando lo formulamos. Desgraciadamente, esto no se observa en muchos de los primeros trabajos sobre sintonía de controladores.

Fórmulas de sintonía basadas en optimización

Se han dedicado muchos estudios al desarrollo de reglas de sintonía basadas en optimización. Muy a menudo se ha considerado un proceso descrito por

$$P = \frac{K_p}{1 + sT} e^{-sL}$$

Se han calculado las funciones de coste obtenidas para cambios en escalón unitario en el punto de consigna y se han ajustado a los valores numéricos obtenidos fórmulas del tipo

$$p = a \left(\frac{L}{T} \right)^b,$$

donde p es un parámetro del controlador y a y b son constantes. En muchos casos, el criterio es IAE para perturbaciones en la carga, que con frecuencia da sistemas con bajo amortiguamiento y pobre sensibilidad. Las fórmulas dadas a menudo sólo se cumplen para un pequeño rango de retardos de tiempo normalizados, e.g., $0,2 < \tau < 0,6$. Debería también observarse que los criterios basados en cambios en el punto de consigna pueden a veces ser erróneos porque con frecuencia no se observa que los cambios en el punto de consigna están drásticamente influenciados por diferentes ponderaciones en dicho punto de consigna.

Módulo óptimo y simétrico óptimo

El módulo óptimo (BO) y el simétrico óptimo (SO) son dos métodos para seleccionar y sintonizar controladores que también pueden ser vistos como diseños analíticos donde las funciones de transferencia deseadas dadas por las Ecuaciones 6.23 se obtienen por optimización. Los acrónimos BO y SO se derivan de las palabras alemanas Beträgs Optimum y Symmetrische Optimum. Los métodos se desarrollaron para servomotores donde la respuesta a cambios en el punto de consigna es particularmente importante. La idea básica es encontrar un controlador que haga que la respuesta en frecuencia desde el punto de consigna hasta la salida de la planta sea tan próxima a uno como sea posible para bajas frecuencias. Si $G(s)$ es la función de transferencia desde el punto de consigna a la salida, el controlador se determina de tal forma que $G(0) = 1$ y que $d^n|G(i\omega)|/d\omega^n = 0$ en $\omega = 0$ para tantos n como sea posible. Una propiedad interesante es que el método de diseño toma en cuenta explícitamente las dinámicas no modeladas. Ilustramos la idea con unos pocos ejemplos.

EJEMPLO 6.18—SISTEMA DE SEGUNDO ORDEN

Sea la función de transferencia

$$G(s) = \frac{a_2}{s^2 + a_1 s + a_2},$$

que se ha escogido de forma que $G(0) = 1$. Consideremos en primer lugar como deberían escogerse los parámetros de forma que se obtenga una respuesta en frecuencia máximamente plana. Tenemos

$$|G(i\omega)|^2 = \frac{a_2^2}{a_1^2 \omega^2 + (a_2 - \omega^2)^2} = \frac{a_2^2}{a_2^2 + \omega^2(a_1^2 - 2a_2) + \omega^4}.$$

escogiendo $a_1 = \sqrt{2a_2}$ encontramos

$$|G(i\omega)|^2 = \frac{a_2^2}{a_2^2 + \omega^4}.$$

Las primeras tres derivadas de $|G(i\omega)|$ desaparecerán en el origen. La función de transferencia tiene entonces la forma

$$G(s) = \frac{\omega_0^2}{s^2 + \omega_0 s \sqrt{2} + \omega_0^2}.$$

La respuesta en escalón de un sistema con esta función de transferencia tiene una sobreelongación $o = 4\%$. El tiempo de asentamiento al 2% del valor en estado estacionario es $T_s = 6/\omega_0$. \square

Si la función de transferencia G en el ejemplo se obtiene por realimentación del error de un sistema con la función de transferencia del lazo G_{BO} , la función de transferencia del lazo es

$$G_{BO}(s) = \frac{G(s)}{1 - G(s)} = \frac{\omega_0^2}{s(s + \sqrt{2}\omega_0)}, \quad (6.30)$$

que es la función de transferencia del lazo deseada para el método llamado módulo óptimo.

El cálculo en el Ejemplo 6.18 se puede realizar con más esfuerzo para sistemas de orden más elevado. Lo ilustramos con otro ejemplo.

EJEMPLO 6.19—SISTEMA DE TERCER ORDEN SIN CEROS

Sea la función de transferencia

$$G(s) = \frac{a_3}{s^3 + a_1 s^2 + a_2 s + a_3}.$$

Despues de algunos cálculos obtenemos

$$|G(i\omega)| = \frac{a_3}{\sqrt{a_3^2 + (a_2^2 - 2a_1 a_3)\omega^2 + (a_1^2 - 2a_2)\omega^4 + \omega^6}}.$$

Cinco derivadas de $|G(i\omega)|$ desaparecerán en $\omega = 0$, si los parámetros son tales que $a_1^2 = 2a_2$ y $a_2^2 = 2a_1 a_3$. La función de transferencia entonces se hace

$$G(s) = \frac{\omega_0^3}{s^3 + 2\omega_0 s^2 + 2\omega_0^2 s + \omega_0^3} = \frac{\omega_0^3}{(s + \omega_0)(s^2 + \omega_0 s + \omega_0^2)}. \quad (6.31)$$

La respuesta en escalón de un sistema con esta función de transferencia tiene una sobreelongación $o = 8,1\%$. El tiempo de asentamiento al 2% del valor en estado estacionario es $9,4/\omega_0$. Se puede obtener un sistema con esta función de

transferencia en lazo cerrado con un sistema que tenga realimentación del error y la función de transferencia del lazo

$$G_l(s) = P(s)C(s) = \frac{\omega_0^3}{s(s^2 + 2\omega_0 s + 2\omega_0^2)}.$$

La función de transferencia del lazo (6.31) se puede obtener también de otras funciones de transferencia de lazo si el controlador tiene ponderación del punto de consigna. Por ejemplo, si un proceso con la función de transferencia

$$P(s) = \frac{\omega_0^2}{s(s + 2\omega_0)}$$

se controla por un controlador PI que tiene parámetros $K = 2$, $T_i = 2/\omega_0$, y $b = 0$, la función de transferencia del lazo se hace

$$G_{SO} = \frac{\omega_0^2(2s + \omega_0)}{s^2(s + 2\omega_0)}. \quad (6.32)$$

El óptimo simétrico trata de obtener la función de transferencia del lazo dada por la Ecuación 6.32. Observe que el diagrama de Bode de esta función de transferencia es simétrica alrededor de la frecuencia $\omega = \omega_0$. Esta es la motivación para el nombre de óptimo simétrico.

Si se utiliza un controlador PI con $b = 1$, la función de transferencia desde el punto de consigna hasta la salida del proceso se hace

$$G(s) = \frac{G_{SO}(s)}{1 + G_{SO}(s)} = \frac{(2s + \omega_0)\omega_0^2}{(s + \omega_0)(s^2 + \omega_0 s + \omega_0^2)}.$$

Esta función de transferencia no es máximamente plana a causa del cero en el numerador. Este cero también dará una respuesta del punto de consigna con una gran sobreelongación, del 43 por ciento. \square

Los métodos BO y SO se pueden llamar métodos de deformación del lazo ya que ambos métodos intentan obtener una función de transferencia del lazo específica. Los métodos de diseño se pueden describir como sigue. Se establece en primer lugar cuál de las funciones de transferencia, G_{BO} o G_{SO} , es más apropiada. La función de transferencia del controlador $C(s)$ se escoge entonces tal que la función de transferencia del lazo $G_l(s) = P(s)C(s)$ cumple las especificaciones. Ilustramos los métodos con los ejemplos siguientes.

EJEMPLO 6.20—CONTROL BO

Sea un proceso con la función de transferencia

$$P(s) = \frac{K_p}{s(1 + sT)}. \quad (6.33)$$

Con un controlador proporcional la función de transferencia del lazo se hace

$$G_\ell(s) = \frac{KK_p}{s(1+sT)}.$$

Para hacer esta función de transferencia igual a G_{BO} dada por la Ecuación 6.30 se debe requerir que

$$\omega_0 = \frac{\sqrt{2}}{2T}.$$

La ganancia del controlador debería escogerse como

$$K = \frac{\omega_0\sqrt{2}}{2K_p} = \frac{1}{2K_p T}.$$

□

EJEMPLO 6.21—CONTROL SO

Sea un proceso con la misma función de transferencia como en el ejemplo previo (Ecuación 6.33). Con un controlador PI que tiene la función de transferencia

$$C(s) = \frac{K(1+sT_i)}{sT_i},$$

obtenemos la función de transferencia del lazo

$$G_l(s) = P(s)C(s) = \frac{K_p K(1+sT_i)}{s^2 T_i (1+sT)}.$$

Esto es idéntico a G_{SO} si escogemos

$$K = \frac{1}{2K_p T}$$

$$T_i = 4T.$$

Para obtener la función de transferencia dada por la Ecuación 6.31 el peso del punto de consigna b debería ser cero. □

Un procedimiento de diseño Un procedimiento de diseño sistemático se puede basar en los métodos BO y SO. El método de diseño consiste en dos pasos. En el primer paso la función de transferencia del proceso se simplifica a una de las

siguientes formas:

$$\begin{aligned} P_1(s) &= \frac{K_p}{1+sT} \\ P_2(s) &= \frac{K_p}{(1+sT_1)(1+sT_2)}, \quad T_1 > T_2 \\ P_3(s) &= \frac{K_p}{(1+sT_1)(1+sT_2)(1+sT_3)}, \quad T_1 > T_2 > T_3 \\ P_4(s) &= \frac{K_p}{s(1+sT)} \\ P_5(s) &= \frac{K_p}{s(1+sT_1)(1+sT_2)}, \quad T_1 > T_2. \end{aligned}$$

Los polos del proceso se pueden cancelar mediante ceros del controlador para obtener la función de transferencia del lazo deseada. Un polo lento se puede aproximar por un integrador; los polos rápidos se pueden concentrar juntos tal como se analizó en Sección 2.8. La regla heurística dada en los trabajos originales del método es que las constantes de tiempo T tales que $\omega_0 T < 0,25$ se pueden considerar como integradores.

El controlador se deriva de la misma forma que en los Ejemplos 6.20 y 6.21 seleccionando los parámetros de forma que la función de transferencia del lazo concuerda con G_{BO} o G_{SO} . Haciendo esto obtenemos los resultados que se resumen en la Tabla 6.8. Observe, por ejemplo, que los Ejemplos 6.20 y 6.21 corresponden a las entradas Proceso G_4 en la tabla. Es natural considerar la constante de tiempo más pequeña como una aproximación de la dinámica despreciada en el proceso. Es interesante observar que es esta constante de tiempo la que determina el ancho de banda del sistema en lazo cerrado.

La respuesta del punto de consigna para el método BO es excelente. Observe que es necesario usar un controlador con una estructura de dos grados de libertad o un prefiltro para evitar una gran sobreelongación para el método SO. Nótese también que los polos del proceso se cancelan en los casos marcados C1 o C2 en la Tabla 6.8. La respuesta a las perturbaciones de carga será pobre si el polo cancelado es lento comparado con la dinámica en lazo cerrado, que se caracteriza por ω_0 en la Tabla 6.8.

Estos principios de diseño se pueden extender a otros procesos además de aquellos listados en la tabla.

EJEMPLO 6.22—APLICACIÓN DE BO Y SO

Sea un proceso con la función de transferencia

$$P(s) = \frac{1}{(1+s)(1+0,2s)(1+0,05s)(1+0,01s)}. \quad (6.34)$$

Como esta función de transferencia es de cuarto orden, el procedimiento de diseño no se puede aplicar directamente. Mostramos como se obtienen diferentes controladores dependiendo de las aproximaciones hechas. El comportamiento del sistema en lazo cerrado depende de la aproximación. Utilizamos el parámetro ω_0 como una medida aproximada del comportamiento.

Cuadro 6.8 Los parámetros del controlador obtenidos con los métodos BO y SO. La entrada P da la función de transferencia del proceso, la entrada C da la estructura del controlador, y la entrada M indica si se utiliza el método BO o SO. En la entrada Notas, A1 significa que $1 + sT_1$ se aproxima por sT_1 , y C1 significa que se cancela la constante de tiempo T_i .

P	C	M	Notas	KK_p	T_1	T_d	ω_0	b	c
P_1	I	BO		0.5	T		$0.7/\bar{T}$		
P_2	P	BO	A1	$\frac{T_1}{2T_2}$			$0.7/\bar{T}_2$	1	
P_2	PI	BO	C1	$\frac{T_1}{2T_2}$	T_1		$0.7/\bar{T}_2$	1	
P_2	PI	SO	A1	$\frac{T_1}{2T_2}$	$4T_2$		0.5	0	
P_3	PD	BO	A1, C2	$\frac{T_1}{2T_3}$		T_2	$0.7/\bar{T}_3$	1	1
P_3	PID	BO	C1, C2	$\frac{T_1 + T_2}{2T_3}$	$T_1 + T_2$	$\frac{T_1 + T_2}{\bar{T}_1 + \bar{T}_2}$	$0.7/\bar{T}_3$	1	1
P_3	PID	SO	A1, C2	$\frac{T_1(T_2 + 4T_3)}{8T_3^2}$	$T_2 + 4T_3$	$\frac{4T_2 T_3}{T_2 + 4T_3}$	$0.5/\bar{T}_3$	0	
P_4	P	BO		1	\bar{T}		$0.7/\bar{T}$	1	
P_4	PI	SO		\bar{T}	$4T$		$0.5/\bar{T}$	0	
P_5	PD	BO	C1	$\frac{1}{2T_2}$		T_1	$0.7/\bar{T}_2$	1	1
P_5	PD	SO	A1	$\frac{T_1}{8T_2^2}$		$4T_2$	$0.5/\bar{T}_2$	1	0
P_5	PID	SO	C1	$\frac{T_1 + 4T_2}{8T_2^2}$	$T_1 + 4T_2$	$\frac{4T_1 T_2}{T_1 + 4T_2}$	$0.5/\bar{T}_2$	$\frac{T_1}{T_1 + 4T_2}$	0

Si es aceptable un controlador con bajo comportamiento, el proceso (6.34) se puede aproximar por

$$P(s) = \frac{1}{1 + 1,26s}. \quad (6.35)$$

La aproximación tiene un error de fase menor que 10° para $\omega \leq 1,12$. Se sigue de la Tabla 6.8 que el sistema (6.35) se puede controlar con un controlador tipo integrador con

$$k_i = \frac{K}{T_i} = \frac{0,5}{1,26} = 0,4.$$

esto da un sistema en lazo cerrado con $\omega_0 = 0,55$.

Se obtiene un sistema en lazo cerrado con mejor comportamiento si la función de transferencia (6.34) se aproxima por

$$P(s) = \frac{1}{(1 + s)(1 + 0,26s)}. \quad (6.36)$$

Se mantiene así la constante de tiempo más pequeña, y las constantes de tiempo restantes se aproximan concentrándolas todas juntas. La aproximación tiene un error de fase menor que 10° para $\omega \leq 5,15$. Se puede diseñar un controlador PI usando el método BO. De la Tabla 6.8 se obtienen los parámetros $K = 1,92$ y $T_i = 1$. El sistema en lazo cerrado tiene $\omega_0 = 2,7$.

Si la función de transferencia se aproxima por

$$P(s) = \frac{1}{(1 + s)(1 + 0,2s)(1 + 0,06s)}, \quad (6.37)$$

la aproximación tiene un error de fase menor que 10° para $\omega \leq 26,6$. El método BO se puede usar también en este caso. La Tabla 6.8 da los parámetros del controlador $K = 10$, $T_i = 1,2$, y $T_d = 0,17$. La estructura del controlador se define por los parámetros $b = 1$ y $c = 1$. Este controlador da un sistema en lazo cerrado con $\omega_0 = 11,7$.

El método puede también aplicarse al sistema (6.37). La Tabla 6.8 da los parámetros del controlador $K = 15,3$, $T_i = 0,44$, $T_d = 0,11$, y $b = 0,45$. Para estos parámetros obtenemos $\omega_0 = 8,3$.

Se pueden obtener controladores con diferentes propiedades al aproximar la función de transferencia de diferentes formas. En la Tabla 6.9 se da un resumen de las propiedades obtenidas del sistema en lazo cerrado, donde IAE se refiere a la repuesta a las perturbaciones de carga. Observe que el Controlador 2 cancela un polo del proceso con constante de tiempo de 1 s y que el Controlador 3 cancela los polos del proceso con constantes de tiempo 1 s y 0,25 s. Esto explica por qué el IAE cae drásticamente para el Controlador 4, que no cancela ningún polo del proceso. El Controlador 4 realmente tiene un ancho de banda ω_0 menor que el Controlador 3.

En la Figura 6.16 se muestra la simulación de los diferentes controladores. La simulación y los datos mostrados en la Tabla 6.9 claramente ilustran los beneficios del modelo mejorado y controladores más complicados. \square

Cuadro 6.9 Resultados obtenidos con diferentes controladores diseñados por los métodos BO y SO en el Ejemplo 6.22. La frecuencia ω_m define el límite superior cuando el error de fase es menor que el 10 %.

Controlador	K	T_i	T_d	k_i	b	c	ω_0	ω_m	IAE
1					0,4		0,55	1,12	2,7
2	1,92	1		0,52	1		2,7	5,15	0,52
3	10	1,2	0,17	8,3	1	1	11,7	26,6	0,12
4	15,3	0,44	0,11	35	0,45	0	8,3	26,6	0,029

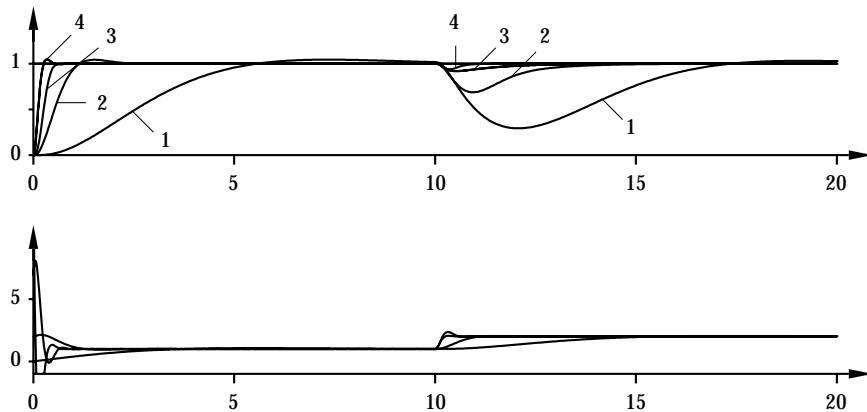


Figura 6.16 Simulación del sistema en lazo cerrado con diferentes controladores diseñados por los métodos BO y SO dados en la Tabla 6.9. El diagrama superior muestra el punto de consigna y_{sp} y la salida del proceso y , y el diagrama inferior muestra la señal de control u .

Diseño para rechazo de perturbaciones

Los métodos de diseño analizados hasta aquí se han basado en una caracterización de la dinámica del proceso. Las propiedades de las perturbaciones sólo han influido indirectamente en el diseño. Se utilizó una perturbación en la carga en la forma de un escalón, y en algunos casos se minimizó una función de coste basada en el error debido a una perturbación en la carga. El ruido de medida se incorporó también al limitar la ganancia de alta frecuencia del controlador.

En esta sección, analizamos brevemente métodos de diseño que directamente intentan conseguir un compromiso entre atenuación de las perturbaciones de carga y amplificación del ruido de medida debido a la realimentación.

Sea el sistema mostrado en la Figura 6.17. Observe que se filtra la señal de medida antes de que se alimente al controlador. Sean D y N las transformadas de Laplace de la perturbación de carga y del ruido de medida, respectivamente. La

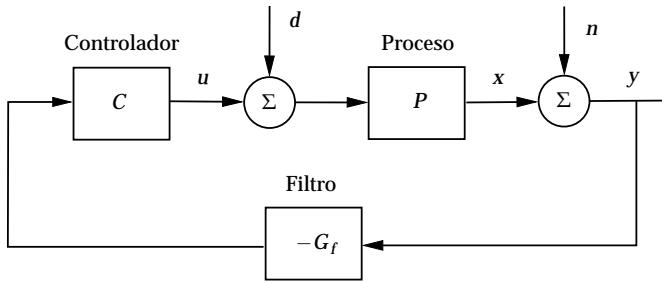


Figura 6.17 Diagrama de bloques de un sistema en lazo cerrado.

salida del proceso y la señal de control están entonces dada por

$$\begin{aligned} X &= \frac{P}{1 + G_\ell} D - \frac{G_\ell}{1 + G_\ell} N \\ U &= -\frac{G_\ell}{1 + G_\ell} D - \frac{CG_f}{1 + G_\ell} N, \end{aligned} \quad (6.38)$$

donde $G_\ell = PCG_f$ es la función de transferencia del lazo. Pueden ahora darse diferentes hipótesis acerca de las perturbaciones y de los diferentes criterios de diseño. Lo ilustramos por un ejemplo.

EJEMPLO 6.23—DISEÑO PARA RECHAZO DE PERTURBACIONES

Suponga que las funciones de transferencia en la Figura 6.17 están dadas por

$$P = \frac{1}{s}, \quad G_f = 1, \quad C = k + \frac{k_i}{s}.$$

Más aún, suponga que n es un ruido estacionario con densidad espectral ϕ_n y que d se obtiene enviando ruido estacionario con el espectro ϕ_d a través de un integrador. Esto es una forma de modelar la situación de que la perturbación de carga tiene una deriva y que el ruido de medida tiene frecuencias altas.

Con la hipótesis dada, la Ecuación 6.38 entonces se hace

$$\begin{aligned} X &= \frac{s}{s^2 + ks + k_i} \frac{1}{s} D_1 - \frac{sk + k_i}{s^2 + ks + k_i} N \\ U &= -\frac{sk + k_i}{s^2 + ks + k_i} \frac{1}{s} D_1 - \frac{s^2k + k_is}{s^2 + ks + k_i} N, \end{aligned}$$

donde hemos supuesto

$$D(s) = \frac{1}{s} D_1(s).$$

Si n y d_1 son ruidos blancos se sigue que la varianza de x está dada por

$$J = Ex^2 = \frac{1}{2kk_i} \phi_d + \frac{1}{2} \left(k + \frac{k_i}{k} \right) \phi_n.$$

Esta ecuación claramente indica el compromiso al diseñar el controlador. El primer término del lado derecho es la contribución a la varianza debido a la perturbación de carga. El segundo término representa la contribución debida al ruido de medida. Observe que la atenuación de las perturbaciones de carga aumenta cuando k y k_i aumentan, pero los valores más grandes de k y k_i también aumenta la contribución del ruido de medida.

Podemos intentar encontrar valores de k y k_i que minimizan J . Un cálculo directo da

$$k = \sqrt{2} \left(\frac{\phi_d}{\phi_n} \right)^{1/4}$$

$$k_i = \sqrt{\frac{\phi_d}{\phi_n}}.$$

esto significa que los parámetros del controlador están determinados únicamente por la razón de las intensidades del ruido del proceso y el ruido de medida. También observe que con estos parámetros el polinomio característico en lazo cerrado es

$$s^2 + \omega_0 s \sqrt{2} + \omega_0^2,$$

con $\omega_0 = \sqrt{\phi_v/\phi_e}$. El sistema óptimo tiene así un amortiguamiento relativo $\zeta = 0,707$ y un ancho de banda que está dado por la razón de las intensidades de la perturbación de carga y el ruido de medida. \square

Comentario

Las técnicas de optimización son muy potentes. Cuando se utilizan es esencial incluir en la formulación todos los aspectos relevantes del problema; si no se hace así, el denominado controlador óptimo puede tener muy malas propiedades. En esta sección hemos cubierto algunos métodos de optimización que han sido utilizados para el control PID.

Los métodos BO y SO se usan ampliamente en servosistemas. La optimización es encontrar una función de transferencia desde el punto de consigna hasta la salida del proceso que sea máximamente plana. Los métodos están principalmente adaptados para sistemas sin retardo de tiempo. Se pueden tratar pequeños retardos de tiempo mediante aproximación.

Una característica interesante del procedimiento es el empleo de aproximaciones; se desprecian los polos rápidos y las constantes de tiempo lentas, y las dinámicas lentas se aproximan por integradores. La incertidumbre del modelo también aparece explícitamente en el diseño porque el ancho de banda que se puede lograr se determina por las constantes de tiempo más lentas despreciadas.

Los métodos se pueden interpretar como asignación de polos donde el polinomio característico en lazo cerrado deseado para el módulo óptimo es

$$A_{BO}(s) = s^2 + \omega_0 s \sqrt{2} + \omega_0^2$$

y para el óptimo simétrico es

$$A_{SO}(s) = (s + \omega_0)(s^2 + \omega_0 s + \omega_0^2)$$

Hay posibilidades para combinar ambos esquemas. Una desventaja con todos los métodos de diseño de este tipo es que los polos del proceso se pueden cancelar. Esto puede llevar a una pobre atenuación de las perturbaciones de carga si los polos cancelados se excitan por perturbaciones y si son lentes comparadas con los polos en lazo cerrado dominantes.

6.8 Deformación robusta del lazo

Los métodos de diseño analizados hasta aquí tienen todos la propiedad de que la robustez a variaciones del proceso tiene que ser comprobada después del diseño. Uno de los grandes avances en la teoría de control en el final del último siglo fue la emergencia de los métodos de diseño con robustez garantizada (la denominada teoría \mathcal{H}^∞). En esta sección presentaremos un método para el diseño de controladores PID con este mismo espíritu. En la Sección 4.6 se demostró que las condiciones de robustez se pueden expresar en términos de discos circulares que son regiones prohibidas para el diagrama de Nyquist de la función de transferencia del lazo. Para control PID estas condiciones dan un conjunto de valores admisibles de los parámetros del controlador, llamadas la región de robustez. La atenuación de las perturbaciones de carga de baja frecuencia es inversamente proporcional a la ganancia integral k_i . La inyección del ruido de medida se captura por la ganancia del controlador k para control P y PI o la ganancia derivativa k_d para control PD y PID. El método de diseño es maximizar la ganancia integral k_i sujeto a restricciones sobre la robustez y la inyección de ruido. Se obtiene entonces buena respuesta al punto de consigna mediante ponderación del punto de consigna o acción feedforward tal como se analizó en la Sección 5.3. Este método de diseño trae el diseño de controladores PID dentro de la corriente principal del diseño de sistemas de control.

La región de robustez

En la Sección 4.6 se demostró que la robustez a variaciones del proceso se puede expresar por la sensibilidad máxima M_s , la sensibilidad máxima complementaria M_t , o por la sensibilidad conjunta M . Todas estas condiciones dicen que el diagrama de Nyquist de la función de transferencia del lazo debería evitar círculos que encierran al punto crítico. Para control PID de un proceso con función de transferencia dada la restricción de robustez se traduce en restricciones sobre los parámetros del controlador, llamada la región de robustez. Para determinar la región de robustez consideraremos un proceso con la función de transferencia $P(s)$ y un controlador PID ideal con la función de transferencia $C(s)$. La función de transferencia del lazo es $G_l(s)$, y el cuadrado de la distancia desde un punto del diagrama de Nyquist de la función de transferencia del lazo al punto $-c$ es

$$f(k, k_i, k_d, \omega) = |c + G_l(i\omega)|^2 = |c + (k + i(k_d\omega - k_i/\omega))P(i\omega)|^2;$$

y la restricción de robustez se hace

$$f(k, k_i, k_d, \omega) \geq r^2. \quad (6.39)$$

Introducimos

$$P(i\omega) = \alpha(\omega) + i\beta(\omega) = \rho(\omega)e^{i\varphi(\omega)}, \quad (6.40)$$

donde

$$\begin{aligned}\alpha(\omega) &= \rho(\omega) \cos \varphi(\omega), \\ \beta(\omega) &= \rho(\omega) \sin \varphi(\omega).\end{aligned}$$

El siguiente cálculo directo pero tedioso muestra que la función f se puede escribir como

$$\begin{aligned}f(k, k_i, k_d, \omega) &= \left| c + (k + i(k_d\omega - k_i/\omega))(\alpha(\omega) + i\beta(\omega)) \right|^2 \\ &= |c + \alpha k + \beta(k_d\omega - k_i/\omega) + i(\beta k + \alpha(k_d\omega - k_i/\omega))|^2 \\ &= c^2 + \rho^2 k^2 + 2c\alpha k + \rho^2(k_d\omega - k_i/\omega)^2 - 2\beta c(k_d\omega - k_i/\omega) \\ &= \rho^2 \left(k + \frac{\alpha c}{\rho^2} \right)^2 + \frac{\rho^2}{\omega^2} \left(k_i + \frac{\omega \beta c}{\rho^2} - k_d \omega^2 \right)^2 \geq r^2,\end{aligned}$$

donde el argumento ω en las funciones α y β ha sido eliminado para simplificar la escritura. Insertando los argumentos la condición de robustez se puede escribir como

$$\left(\frac{\rho(\omega)}{r} \right)^2 \left(k + \frac{\alpha(\omega)c}{\rho(\omega)^2} \right)^2 + \left(\frac{\rho(\omega)}{\omega r} \right)^2 \left(k_i + \frac{\omega \beta(\omega)c}{\rho(\omega)^2} - \omega^2 k_d \right)^2 \geq 1. \quad (6.41)$$

Para tener un sistema en lazo cerrado estable hay también una condición de enlazamiento requerida por el teorema de estabilidad de Nyquist. La restricción de robustez implica así que los parámetros del controlador deben pertenecer a una región llamada la *región de robustez*, ver Figura 6.18. El diseño de controladores PID puede así formularse como el siguiente problema de programación semi-infinito: maximizar k_i sujeto a la restricción de robustez (6.41) y restricciones sobre k y k_d .

La Figura 6.18 da una buena comprensión del problema de diseño. El controlador PI, que maximiza la ganancia integral, puede encontrarse de la intersección de la región de robustez con el plano $k_d = 0$. El mejor controlador PI tiene $k = 0,4$ y $k_i = 0,2$. Se pueden obtener valores de la ganancia integral que son cinco veces más grande cuando se usa la acción derivativa.

El problema de optimización no es directo porque la restricción (6.41) debe satisfacerse para toda ω , y el conjunto de parámetros que satisfacen la restricción no es necesariamente convexo. Antes de resolver el problema de optimización investigaremos por lo tanto el conjunto de restricción.

Una interpretación geométrica

La restricción de robustez (6.41) tiene una interpretación agradable. Para ω y k_d fijas representa el exterior de una elipse en el plano $k-k_i$; ver Figura 6.19. La elipse tiene su centro en $k = \alpha c / \rho^2$ y $k_i = \omega \beta c / \rho^2$, y sus ejes son paralelos a los ejes de coordenada. La mitad del eje horizontal tiene longitud r/ρ , y la mitad del

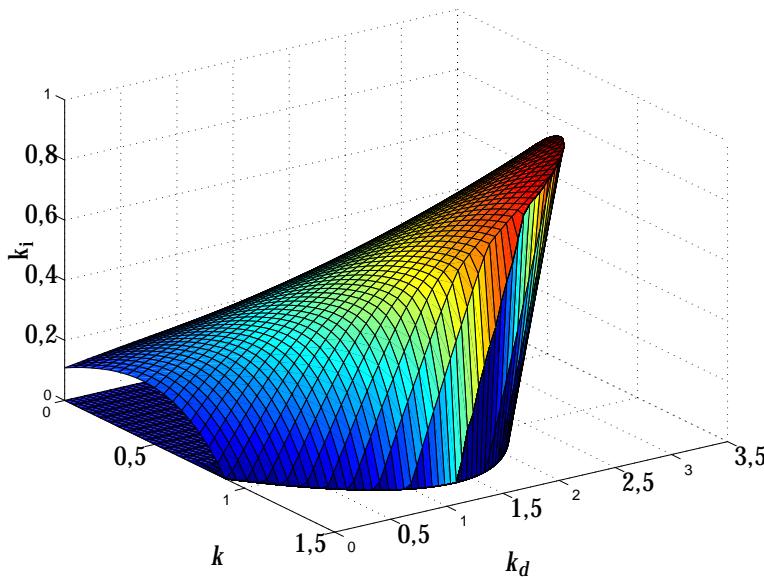
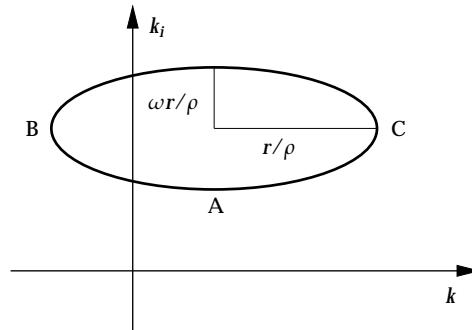


Figura 6.18 Región de robustez para un proceso con la función de transferencia $P(s) = 1/(s+1)^4$ y el criterio de robustez $M_s \leq 1,4$.



eje vertical tiene longitud $\omega r/\rho$. El centro de la elipse está sobre la frontera de estabilidad.

Cuando ω va desde 0 a ∞ las elipses tienen una envolvente

$$\begin{aligned} f(k, k_i, k_d, \omega) &= r^2, \\ \frac{\partial f}{\partial \omega}(k, k_i, k_d, \omega) &= 0, \end{aligned} \tag{6.42}$$

que define una frontera de la restricción de sensibilidad. Suponiendo que el proceso tiene ganancia positiva la otra frontera está dada por el plano $k - k_d$. Como la función f es cuadrática en k_i la envolvente tiene dos ramas; solo una rama corresponde a sistemas en lazo cerrado estables.

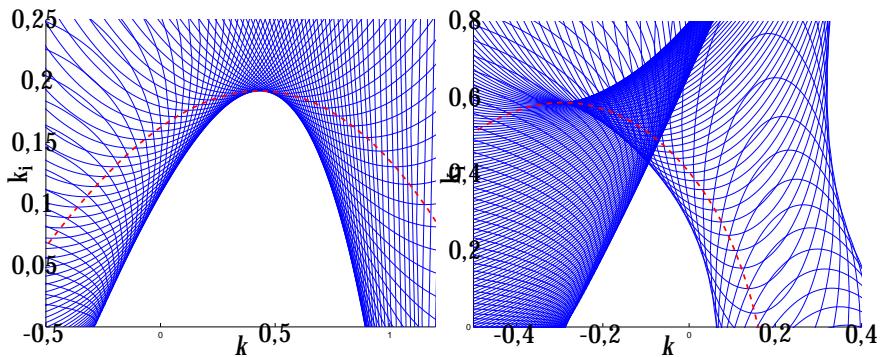


Figura 6.20 Ilustración geométrica de las elipses generadas por la restricción de sensibilidad (6.39) y la envolvente generada por ella. Las curvas de la izquierda se generan para un sistema con la función de transferencia $P(s) = (s+1)^{-4}$ con la restricción $M_s = 1,4$. Las curvas de la derecha se generan para un sistema con la función de transferencia $P(s) = 1/(s+1)(s^2 + 0,2s + 9)$ con la restricción $M_s = 1,4$.

Habiendo comprendido la naturaleza de las restricciones es conceptualmente fácil resolver el problema de optimización encontrando el valor más grande de k_i en la envolvente. Puede haber máximos locales y la envolvente puede tener aristas. Esto se ilustra en la Figura 6.20, que muestra las envolventes y el lugar del vértice inferior de la elipse en dos casos. La figura de la izquierda tiene una envolvente suave, y el lugar del vértice inferior coincide con la envolvente en el máximo. La figura de la derecha tiene una envolvente con una arista en el valor máximo de k_i . Como la generación de la envolvente consume bastante tiempo, es deseable encontrar algoritmos que puedan dar una solución más efectiva. Es también de interés caracterizar los casos donde hay solamente un máximo local.

Envolvente regular

Consideraremos primero el caso donde la envolvente es regular y no tiene puntos angulosos cerca del máximo. El valor máximo de k_i para k_d fijo entonces ocurre en una tangencia con el vértice inferior de la elipse; ver Figura 6.19. El lugar del vértice vertical inferior está dada por

$$\begin{aligned} k(\omega) &= -\frac{\alpha c}{\rho^2} = -\frac{c}{\rho(\omega)} \cos \varphi(\omega), \\ k_i(\omega) &= -\frac{\omega \beta c}{\rho^2} + \omega^2 k_d = -\frac{\omega}{\rho(\omega)}(r + c \operatorname{sen} \varphi(\omega)) + \omega^2 k_d. \end{aligned} \quad (6.43)$$

Se muestra con una línea a tramos en la Figura 6.20. El valor máximo de k_i puede así encontrarse maximizando k_i en el lugar del vértice inferior. Diferenciando la expresión para k_i en (6.43) da

$$\begin{aligned} \frac{dk_i}{d\omega} &= -\frac{d}{d\omega}\left(\frac{\omega(r + c \operatorname{sen} \varphi)}{\rho}\right) + 2\omega k_d \\ &= (r + c \operatorname{sen} \varphi)\left(\frac{\omega \rho'}{\rho^2} - \frac{1}{\rho}\right) - \frac{\omega \varphi' c \cos \varphi}{\rho} + 2\omega k_d = 0. \end{aligned}$$

Para simplificar la escritura hemos eliminado el argumento ω de las funciones α , β , y φ . Dividiendo la ecuación anterior por ω y multiplicándola por ρ , la condición para el extremo se hace

$$h_{PID}(\omega) = (r + c \sin \varphi) \left(\frac{\rho'}{\rho} - \frac{1}{\omega} \right) - c \varphi' \cos \varphi + 2\rho k_d = 0. \quad (6.44)$$

Para determinar el óptimo tenemos así que encontrar la solución ω_{PID}^* de esta ecuación; los parámetros del controlador se obtienen entonces de la Ecuación 6.43.

La Ecuación 6.44 se satisface para un mínimo, un máximo, o un punto silla. Para asegurar que hay un máximo debe requerirse que

$$\frac{d^2 f}{d\omega^2}(\omega^*) > 0. \quad (6.45)$$

Para garantizar que se satisface globalmente la restricción (6.41) tenemos que evaluarlo para toda ω . Esto puede hacerse mediante el diagrama de Nyquist de la función de transferencia del lazo.

La Ecuación 6.44 se puede resolver iterativamente por bisección o por el método de Newton-Raphson. Ambos métodos convergen muy rápido, pero requieren condiciones iniciales apropiadas. Observe, sin embargo, que en general, pueden haber algunas soluciones que pueden encontrarse por comenzar la iteración desde diferentes condiciones iniciales.

Para clases especiales de sistemas es posible dar buenas condiciones iniciales. Sean sistemas donde la función de transferencia $P(s)$ tiene ganancia en baja frecuencia positiva y

$$\begin{aligned} \frac{d \arg P(i\omega)}{d\omega} &< 0, \\ \frac{d \log_{10} |P(i\omega)|}{d \log_{10} \omega} &< 1. \end{aligned} \quad (6.46)$$

Estas condiciones implican que la cantidad $\rho'/\rho - 1/\omega$ es negativa. Para control PI, cuando $k_d = 0$, se sigue de (6.44) y (6.46) que $h_{PI}(\omega_{90}) > 0$ y además que $h_{PI}(\omega_{180-\arcsin r/c}) < 0$. La Ecuación 6.44 entonces tiene una raíz en el intervalo

$$\omega_{90} < \omega_{PI}^* \leq \omega_{180-\arcsin(r/c)}. \quad (6.47)$$

La condición de monotonicidad (6.46) así sólo tiene que ser válida en el intervalo (6.47). Si la condición (6.46) se mantiene se sigue de la Ecuación 6.43 y 6.47 que ambos k y k_i son positivos. Muchos procesos satisfacen esta condición.

Control PD

Para control PD es natural maximizar la ganancia proporcional sujeta a la restricción de robustez. Trabajando los detalles para el caso de envolventes regulares encontramos que el problema se pueden resolver como sigue: encontrar un valor de ω tal que

$$h_{PD}(\omega) = (r + c \cos \varphi) \frac{\rho'}{\rho} + c \varphi' \sin \varphi = 0. \quad (6.48)$$

Entonces calcular las ganancias del controlador de las ecuaciones

$$\begin{aligned} k(\omega) &= -\frac{\alpha c}{\rho^2} - \frac{r}{\rho} = -\frac{r + c \cos \varphi}{\rho}, \\ k_d(\omega) &= \frac{\beta c}{\omega \rho^2} = \frac{c \sin \varphi}{\omega \rho}. \end{aligned} \quad (6.49)$$

Si ρ'/ρ es negativa (6.48) siempre tiene una solución ω_{PD}^* en el intervalo

$$\omega_{180} < \omega_{PD}^* < \omega_{270-\arcsin(r/c)} = \omega_{180+\arccos(r/c)}. \quad (6.50)$$

La fórmula y el código para el diseño de controladores PD puede también usarse simplemente haciendo la observación que diseñar un controlador PD para el sistema $P(s)$ es lo mismo que diseñar un controlador PI para el sistema $sP(s)$.

Un algoritmo de diseño

Obtuvimos el siguiente algoritmo para resolver el problema de diseño en el caso de envolventes regulares.

ALGORITMO 6.1—DISEÑO DEL CONTROLADOR PARA ENVOLVENTES REGULARES

1. Diseñar un controlador PD resolviendo (6.48) por bisección comenzando con el intervalo $(\omega_{180}, \omega_{180+\arccos(r/c)})$. La solución da la frecuencia ω_{PD}^* .
2. Diseñar un controlador PI resolviendo (6.44) con $k_d = 0$ por bisección comenzando con el intervalo $(\omega_{90}, \omega_{180-\arcsin(r/c)})$. La solución da la frecuencia ω_{PI}^* .
3. Diseñar un controlador PID para k_d fija resolviendo (6.44) por bisección comenzando con el intervalo $(\omega_{PI}^*, \omega_{PD}^*)$. Aumentar k_d al valor más grande para el cual se satisface la restricción de robustez.
4. Verificar por cálculo que hay una envolvente regular (6.45) o por el diagrama de Nyquist de la función de transferencia del lazo.

□

Si la envolvente no es regular la solución obtenida por iteración corresponde a un máximo local de la distancia desde el punto crítico al diagrama de Nyquist. El diagrama de Nyquist entonces entra en la región de restricción para puntos alrededor del máximo.

Envolvente con puntos angulosos

El valor máximo de k_i también puede ocurrir en un punto donde la envolvente tiene una arista. Esto se ilustra en la Figura 6.21. Los vértices B y C de la elipse en la Figura 6.19 están dados por

$$\begin{aligned} k(\omega) &= -\frac{\alpha c}{\rho^2} \pm \frac{r}{\rho} = -\frac{\alpha(\omega)c \cos \phi(\omega)}{\rho(\omega)} \pm \frac{r}{\rho(\omega)}, \\ k_i(\omega) &= -\frac{\omega \beta c}{\rho^2} + \omega^2 k_d = -\frac{\omega c \sin \phi(\omega)}{\rho(\omega)} + \omega^2 k_d, \end{aligned} \quad (6.51)$$

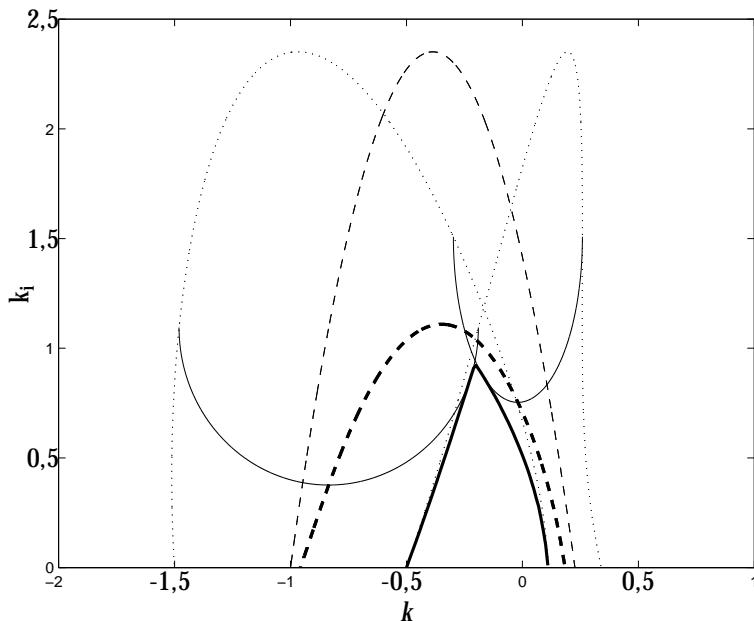


Figura 6.21 Ilustración geométrica de la restricción de sensibilidad (6.39) y la envolvente generada por ella. La envolvente se representa por la línea continua gruesa; el lugar del vértice inferior por la línea a tramos gruesa. La mitad de la elipse se muestra con una línea continua fina. El lugar del centro de las elipses se muestra como una línea a tramos fina, y el lugar de los vértices verticales mediante líneas punteadas. Las curvas se generan para un sistema con la función de transferencia $P(s) = 1/(s+1)(s^2 + 0.2s + 9)$ con la restricción $M_s = 2.0$.

donde el vértice de la izquierda corresponde a un signo menos y el vértice de la derecha a un signo más. El lugar de estos vértices se muestra en líneas punteadas finas, y el lugar del centro de las elipses en líneas a tramos finas. La envolvente se representa como una línea continua gruesa, y el lugar del vértice inferior de la elipse por líneas a tramos gruesas. Obsérvese que el máximo ocurre en la intersección de elipses correspondientes a dos frecuencias diferentes, ω_1 y ω_2 ; ver Figura 6.21. La condición de envolvente (6.42) se satisface entonces para ambas frecuencias. Esto da la condición

$$\begin{aligned} f(k, k_i, k_d, \omega_1) &= R^2, \\ \frac{\partial f}{\partial \omega}(k, k_i, k_d, \omega_1) &= 0, \\ f(k, k_i, k_d, \omega_2) &= R^2, \\ \frac{\partial f}{\partial \omega}(k, k_i, k_d, \omega_2) &= 0. \end{aligned} \tag{6.52}$$

En el diagrama de Nyquist esto corresponde al caso cuando la función de transferencia del lazo es tangente al círculo M en dos puntos.

Es así posible caracterizar el punto donde k_i tiene su valor máximo por ecuaciones algebraicas. Esto significa que el problema de diseño se reduce a resolver

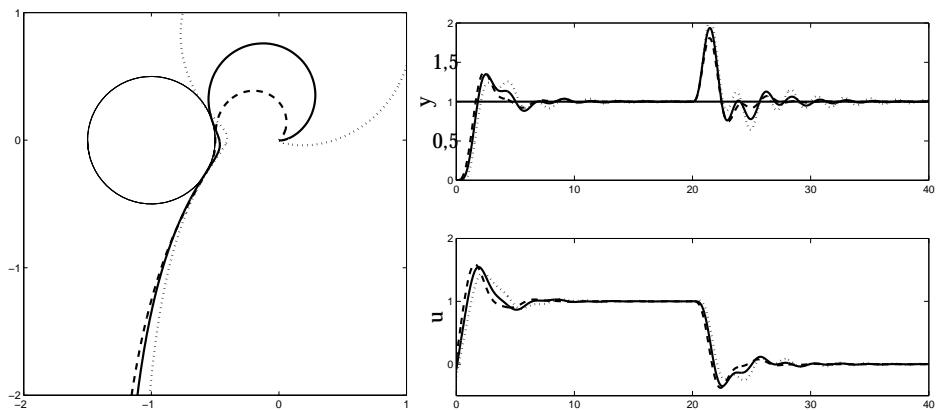


Figura 6.22 Diagramas de Nyquist de la función de transferencia del lazo y respuestas temporales para el Ejemplo 6.24 con $a = 0,2$ (punteada), $0,5$, y $1,0$ (a tramos), cuando se diseña para $M_s = 2,0$.

ecuaciones algebraicas, (6.52), y que se evitan procedimientos de búsqueda elaborados. La ecuación se puede resolver utilizando el método de Newton-Raphson.

Se pueden obtener buenos valores iniciales que son esenciales para la iteración de Newton-Raphson approximando la envolvente por los lugares del lugar horizontal derecho y el lugar del vértice inferior de la elipse; ver Figura 6.21. Ilustraremos el caso de las envolventes con puntos angulosos con un ejemplo.

EJEMPLO 6.24—UN SISTEMA OSCILATORIO

Sea el proceso con la función de transferencia

$$P(s) = \frac{9}{(s+1)(s^2 + as + 9)}.$$

Este es un proceso interesante desde dos puntos de vista. En primer lugar, el sistema tiene dos polos oscilatorios con amortiguamiento relativo $\zeta = a/6$. Cuando se disminuye el parámetro a se hace más y más difícil controlar el proceso. En segundo lugar, dependiendo del valor del parámetro a la envolvente puede tener una derivada continua, $a \geq 1,0653$, o un punto anguloso, $a < 1,0653$.

Para el caso cuando la envolvente tiene un punto anguloso, se diseñó un controlador PI para $M_s = 2,0$. En la Figure 6.22 se muestran los diagramas de Nyquist y las respuestas temporales para los casos $a = 0,2$, $0,5$, y $1,0$. El controlador se comporta razonablemente bien a pesar de que los polos están pobremente amortiguados. En la Tabla 6.10 se muestran los parámetros del controlador y las frecuencias en las cuales la función de transferencia del lazo es tangente al círculo M_s . Obsérvese en la Tabla 6.10 como la ganancia proporcional es negativa para pequeños valores de a . Ésta es la única forma de aumentar el amortiguamiento de los polos oscilatorios con un controlador PI.

Finalmente, ilustramos cómo nuestro método de diseño proporcionará un controlador PI razonable para el caso extremo de $a = 0$. Con el parámetro de diseño $M_s = 1,4$ obtenemos los parámetros del controlador $K = -0,183$, $k_i = 0,251$, y

Cuadro 6.10 Parámetros interesantes cuando se diseña un controlador para $M_s = 2,0$ y diferentes valores de a en el Ejemplo 6.24.

a	K	k_i	ω_1	ω_2
0,0	-0,29	0,68	0,97	2,75
0,1	-0,25	0,82	1,08	2,71
0,2	-0,20	0,93	1,16	2,67
0,5	-0,09	1,17	1,37	2,55
1,0	0,09	1,38	1,65	2,30
2,0	0,48	1,54	2,79	2,79

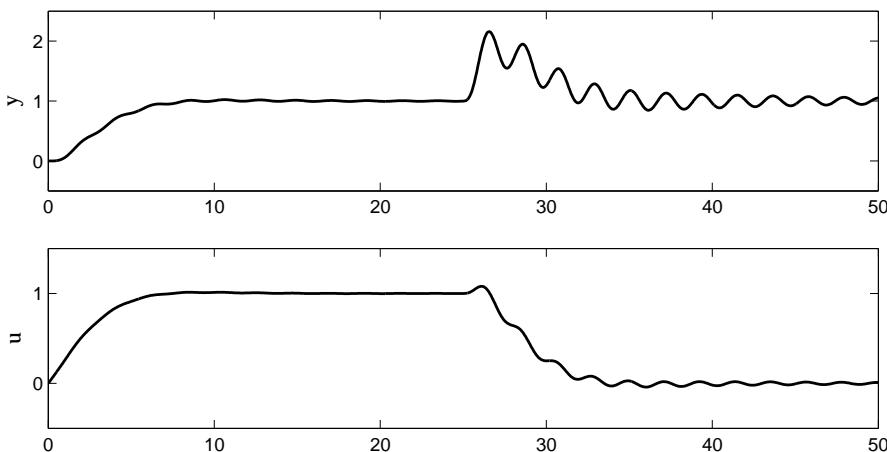


Figura 6.23 Respuesta temporal del sistema en lazo cerrado del Ejemplo 6.24 obtenido para $a = 0$, cuando se diseña el controlador PI para $M_s = 1,4$.

$b = 0$. En la Figura 6.23 se muestran las respuestas temporales. Observamos que la respuesta a cambios en el punto de consigna es bastante razonable, incluso si hay una traza de modos pobemente amortiguados. Sin embargo la perturbación en la carga excitará los modos oscilatorios. El hecho de que el controlador PI es incapaz de proporcionar amortiguamiento de estos modos es claramente observable en la figura. \square

El acantilado de la derivada

Envoltorios regulares se encuentran frecuentemente para control PI de sistemas con respuesta en frecuencia esencialmente monótonas, y para control PID con valores moderados de k_d . Sin embargo, la optimización de k_i en la región de robustez a menudo da controladores con propiedades no deseables. Esto se puede comprender a partir del diagrama de la región de robustez en la Figura 6.18, que

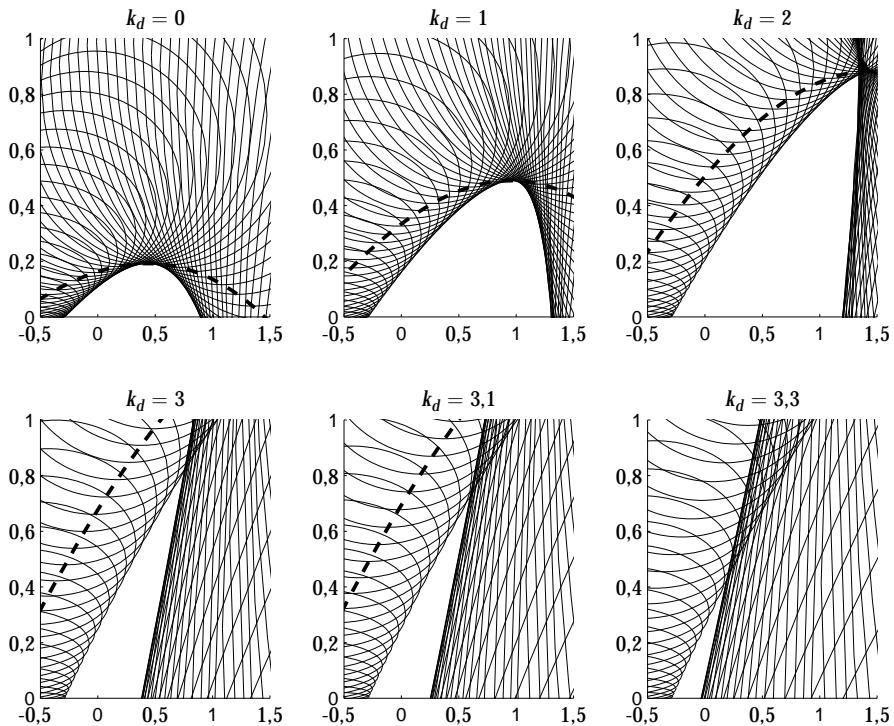


Figura 6.24 Cortes de la región de robustez para ganancia derivativa constante k_d . Las curvas se calculan para el control PID del proceso $P(s) = 1/(s+1)^4$. Observese las aristas cortantes de la región para grandes valores de k_d (el acantilado de la derivada).

muestra que el máximo valor de k_i ocurre en una arista. Tal solución es poco deseable porque pequeños cambios en los parámetros del controlador darán cambios drásticos en k_i . Esto se ilustra también en la Figura 6.24, que muestra las intersecciones de la superficie de robustez para valores fijos de k_d . La figura muestra que para control PI ($k_d = 0$) la envolvente es regular en el máximo $k_i = 0,2$ que ocurre para $k = 0,4$. La ganancia integral k_i puede aumentarse sustancialmente al introducir la acción derivativa. Sin embargo, con valores más grandes de k_d el máximo de k_i ocurre en una arista. La ganancia integral tiene su máximo $k_i = 0,9$ para $k = 0,925$ y $k_d = 2,86$. El comportamiento es muy sensible a las variaciones en los parámetros del controlador en el máximo. Un cambio marginal de los parámetros del controlador pueden hacer al sistema inestable. El controlador que maximiza k_i también tiene otras desventajas, que se ilustran por el ejemplo siguiente.

EJEMPLO 6.25—EL ACANTILADO DE LA DERIVADA

Sea un proceso con la función de transferencia

$$P(s) = \frac{1}{(s+1)^4}.$$

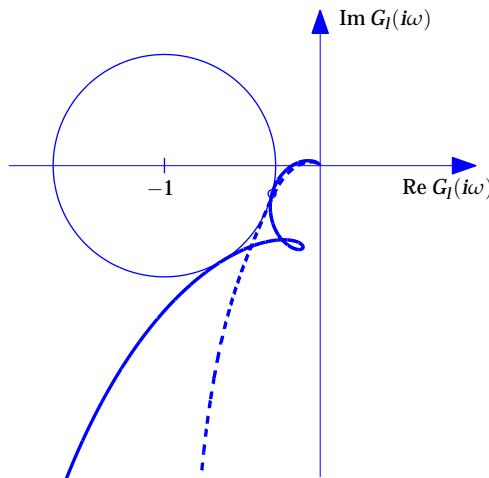


Figura 6.25 El diagrama de Nyquist de la función de transferencia del lazo para control PID del proceso $P(s) = 1/(s+1)^4$, con un controlador que tiene los parámetros $k = 0,925$, $k_i = 0,9$, and $k_d = 2,86$.

Maximizando la ganancia integral k_i sujeta a la restricción de robustez $M_s \leq 1,4$, da los parámetros del controlador $k = 0,925$, $k_i = 0,9$, y $k_d = 2,86$.

En la Figura 6.25 se muestra el diagrama de Nyquist para la función de transferencia del lazo. Observe que el diagrama de Nyquist tiene un bucle. Esto ocurrirá siempre que el máximo se produce donde la envolvente tiene una arista. El controlador obtenido tiene un avance da fase excesivo, que se produce al tener un controlador PID con ceros complejos, $T_i < 4T_d$. En el caso particular tenemos $T_i = 0,33 T_d$. En la Figura 6.26 se muestran los diagrama temporales de la respuesta del sistema a cambios en escalón en el punto de consigna y perturbaciones en la carga. Las respuestas son oscilatorias.

Para comparación hemos mostrado los diagramas de Nyquist y las respuestas temporales para un controlador PID donde $T_i = 4T_d$. Los parámetros del controlador son $k = 1,1$, $k_i = 0,36$, y $k_d = 0,9$. Las respuestas de este controlador son incluso mejor, si el pico en la respuesta a las perturbaciones de carga es mas grande. \square

Evitando el acantilado de la derivada

Hay algunas formas de modificar el problema de diseño para evitar las dificultades asociadas con el acantilado de la derivada. Una manera es introducir condiciones que no permiten al diagrama de Nyquist tener bucles. Otra alternativa es requerir que $T_i > \alpha T_d$. También se ha intentado fijar la ganancia derivativa al valor obtenido por un controlador PD. Esto no elimina los bucles del diagrama de Nyquist en todos los casos. La maximización de k_i puede también sustituirse por maximizar la integral del error absoluto debido a las perturbaciones de carga.

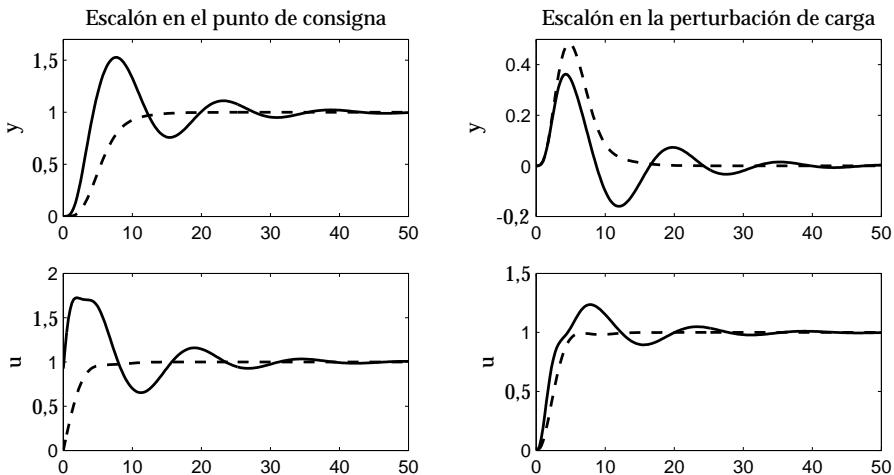


Figura 6.26 Resuestas temporales para control PID del proceso $P(s) = 1/(s+1)^4$, con el controlador teniendo parámetros $k = 0,925$, $k_i = 0,9$, y $k_d = 2,86$ (líneas continuas) y $k = 1,1$, $k_i = 0,36$, y $k_d = 0,9$ (líneas a tramos).

El método MIGO

Después de muchos intentos se ha encontrado que una solución simple es restringir la ganancia derivativa de forma que el máximo ocurre en un punto donde $\partial k_i / \partial k = 0$. Esto evita tener un máximo en una arista. El algoritmo es directo.

ALGORITMO 6.2—DISEÑO MIGO DE UN CONTROLADOR PID

1. Fijar la ganancia derivativa k_d . Encontrar parámetros del controlador resolviendo (6.44); entonces calcular las ganancias del controlador de (6.43).
2. Calcular el valor de M para un rango de frecuencias alrededor de ω^* , y comprobar la restricción de robustez $M \geq M_{crit}$.
3. Aumentar k_d hasta que se obtiene el máximo valor que satisface la restricción de robustez.

□

Este método de diseño particular se llama MIGO (M constrained Integral Gain Optimization).

Un algoritmo para un controlador en forma serie

Frecuentemente sucede que el método de diseño MIGO da parámetros del controlador tales que $T_i < 4T_d$. En la Sección 3.2 se demostró que tales controladores no se pueden implementar en forma serie. Es por lo tanto de interés tener controladores donde los parámetros se restringen a $T_i \geq 4T_d$. Cuando la razón $n = T_i/T_d \geq 4$, el controlador se puede escribir como

$$C(s) = k \left(1 + \frac{1}{sT_i} + sT_d \right) = k' \frac{(T'_i s + 1)(T'_d s + 1)}{T'_i s}, \quad (6.53)$$

donde

$$\begin{aligned} k &= k' \frac{T'_i + T'_d}{T'_i} \\ T_i &= T'_i + T'_d \\ T_d &= \frac{T'_i T'_d}{T'_i + T'_d}. \end{aligned} \quad (6.54)$$

Introduciendo $n' = T'_i / T'_d$, también se sigue que

$$n = \frac{(1+n')^2}{n'}. \quad (6.55)$$

Observe que $n' = 1$ corresponde a $n = 4$.

Se sigue de la Ecuación 6.54 que $T_i = nT_d$ da la siguiente relación entre los parámetros del controlador

$$k_i = \frac{k^2}{nk_d}.$$

A continuación se desarrollará un algoritmo sencillo para maximizar la ganancia integral de un controlador PID con $T_i = nT_d$ sujeta a una restricción de robustez. En primer lugar hacemos la observación que el control PID del proceso $P(s)$ da la función de transferencia del lazo

$$G_l(s) = P(s)C(s) = k' \frac{(1+sT'_i)(1+sT'_d)}{sT'_i} P(s) = k' \frac{(1+sT'_i)}{sT'_i} \left((1+sT'_d)P(s) \right).$$

Esto es idéntico a la función de transferencia del lazo para control PI del proceso

$$P'(s) = (1+sT'_d)P(s).$$

Como hay algoritmos eficientes para control PI obtenemos el siguiente algoritmo iterativo.

ALGORITMO 6.3—DISEÑO DE UN CONTROLADOR PID CON $T_i = 4T_d$

1. Comenzar diseñando un controlador PI para el proceso $P(s)$. Esto da un controlador con el tiempo integral $T_i = k/k_i$. Fijar $T'_1 = T_i/2$ y $j = 1$.
2. Diseñar un controlador PI para el proceso $P'(s) = (1+sT'_1)P(s)$. Sea el tiempo integral del controlador T'_j . Fijar $T'_{j+1} = (T'_j + T'_i)/2$ y repetir hasta que T'_j converge a T' . Sea k' la ganancia del controlador.
3. Los parámetros del controlador son $k = 2k'$, $T_i = 2T'$ y $T_d = T'/2$.

□

Ejemplos

El método de diseño se ilustrará mediante dos ejemplos.

Cuadro 6.11 Los parámetros del controlador obtenidos mediante diseño por deformación del lazo con $M_s = 1,4$ para un proceso con la función de transferencia $P(s) = (s + 1)^{-4}$.

Controlador	K	k_i	k_d	b	T_i	T_d	IAE
PD	1,333	0	1,333	1	0	1	∞
PI	0,433	0,192	0	0,14	2,25	0	5,20
PID MIGO	1,305	0,758	1,705	0*	1,72	1,31	2,25
PID $T_i = 4T_d$	1,132	0,356	0,900	0,9	3,18	0,80	2,51

EJEMPLO 6.26—CUATRO CONSTANTES DE TIEMPO IGUALES

Sea un sistema con la función de transferencia

$$P(s) = \frac{1}{(s + 1)^4}.$$

La Tabla 6.11 resume las propiedades de los controladores PD, PI, y PID diseñados para $M_s = 1,4$. El controlador PD se diseñó maximizando la ganancia proporcional; los controladores PI y PID maximizando la ganancia integral. También se diseñó un controlador PID con la restricción adicional $T_i = 4T_d$. En la Figura 6.27 se muestran las respuestas a cambios en el punto de consigna y a perturbaciones en la carga.

Los controladores PID tienen mejor comportamiento que el controlador PI. La ganancia integral es 2 a 3 veces más grande y el IAE un factor de 2 más pequeño. El controlador PID MIGO tiene $T_i = 1,3T_d$. La tabla muestra que el comportamiento se reduce cuando los parámetros del controlador se restringen a $T_i = 4T_d$. Observe que muchos controladores comerciales PID tiene la restricción $T_i \geq 4T_d$ incorporada, ya que se basan en la forma serie; ver (3.10).

El parámetro b se calcula tal como se describe en la Section 5.3. El cálculo para el controlador PID MIGO muestra que la sobreelongación no se puede reducir suficientemente si no se emplea ponderación del punto de consigna, lo que se indica por la entrada 0* en la tabla. En este caso se recomienda usar un diseño feedforward propio para un sistema con dos grados de libertad. Tal diseño puede mejorar significativamente la respuesta del punto de consigna. \square

EJEMPLO 6.27—CUATRO CONSTANTES DE TIEMPO AMPLIAMENTE DISTRIBUIDAS

Sea un sistema con la función de transferencia

$$P(s) = \frac{1}{(s + 1)(0,1s + 1)(0,01s + 1)(0,001s + 1)}.$$

La Tabla 6.12 resume las propiedades de los controladores PD, PI y PID. Todos los controladores fueron diseñados con la restricción de que la sensibilidad máxima no sea mayor que $M_s = 1,4$. El controlador PD se diseñó maximizando la ganancia proporcional, y los controladores PI y PID maximizando la ganancia integral.

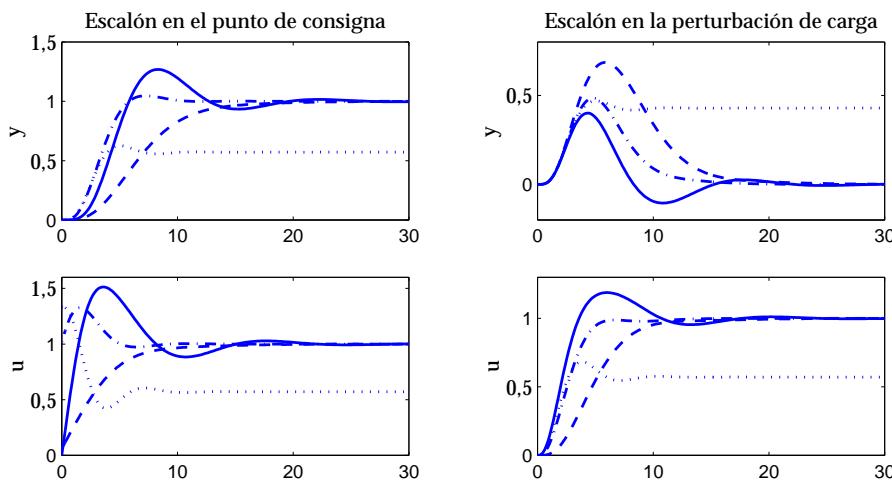


Figura 6.27 Respuestas para el sistema $P(s) = (s + 1)^{-4}$ con los controladores en la Tabla 6.11 a cambios en escalón unitario en el punto de consigna (izquierda) y perturbaciones en la carga (derecha). Las líneas punteadas muestran las respuestas con el controlador PD, a tramos con el controlador PI, punto y raya con el controlador PID con parámetros restringidos a $T_i = 4T_d$, y líneas continuas con el controlador PID diseñado usando el método MIGO.

También fue diseñado un controlador PID con la restricción adicional $T_i = 4T_d$. En la Figura 6.28 se muestran las respuestas a cambios en el punto de consigna y perturbaciones en la carga para los diferentes controladores.

La Tabla 6.12 y la Figura 6.28 muestran que la acción derivativa mejora drásticamente el comportamiento. Las ganancias proporcionales de los controladores con acción derivativa son alrededor de 90, mientras el controlador PI tiene la ganancia 4.2. Se sigue de (4.40) que el pico máximo de la respuesta a la perturbación en la carga es alrededor de 0,01 para controladores con acción derivativa y cerca de 20 veces más grande para control PI. El pico es tan grande que la respuesta a la perturbación en la carga para control PI está bastante fuera de la gráfica. El tiempo de respuesta se aumenta también drásticamente cuando se usa la acción derivativa. Las ganancias integrales de los controladores con acción derivativa

Cuadro 6.12 Parámetros del controlador obtenidos por diseño de deformación del lazo con $M_s = 1,4$ para un proceso con la función de transferencia $P(s) = 1/(s + 1)(0,1s + 1)(0,01s + 1)(0,001s + 1)$.

Controlador	K	k_i	k_d	b	T_i	T_d	IAE
PD	91,7	0	4,4	1	0	0,048	∞
PI	4,21	8,53	0	1	0,494	0	0,1044
PID MIGO	85,5	1488	3,87	0	0,057	0,045	0,00143
PID $T_i = 4T_d$	86,7	518	3,63	0,6	0,168	0,042	0,00143

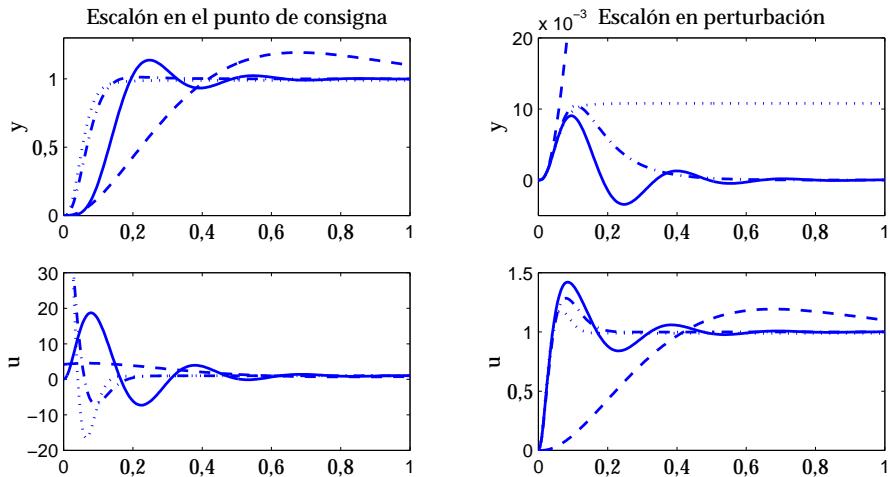


Figura 6.28 Respuestas para el sistema $P(s) = 1/(s+1)(0.1s+1)(0.01s+1)(0.001s+1)$ con los controladores en la Tabla 6.12. Las líneas punteadas muestran las respuestas con el controlador PD controller, a tramos con el controlador PI, punto y raya con el controlador PID con parámetros restringidos a $T_i = 4T_d$, y líneas continuas con el controlador PID diseñado utilizando el método MIGO. La respuesta a las perturbaciones de carga para el controlador PI está fuera de escala.

son también mucho mas grandes que para el caso del control PI. □

Es interesante comparar los Ejemplos 6.26 y 6.27. Para el sistema con cuatro constantes de tiempo iguales en el Ejemplo 6.26 la ganancia integral se puede aumentar por un factor de 3 al introducir la acción derivativa, mientras que para el sistema con cuatro constantes de tiempo distribuidas, en el Ejemplo 6.27, se puede aumentar por un factor de 200. La principal diferencia es que el sistema en el Ejemplo 6.27 está dominado por la constante de tiempo; tiene un retardo de tiempo normalizado $\tau = 0,07$. El sistema en el Ejemplo 6.26 tiene $\tau = 0,33$. El retardo de tiempo normalizado es un buen indicador de los beneficios de la acción derivativa. En el Capítulo 7 se mostrará que es posible una gran mejora en comportamiento con acción derivativa para procesos con pequeños retardos de tiempo normalizados (procesos dominados por constantes de tiempo).

6.9 Resumen

En este capítulo han sido presentadas una serie de técnicas para diseñar controladores PID, comenzando con métodos del tipo de Ziegler-Nichols, donde la dinámica del proceso se caracterizaba mediante unos pocos parámetros que se podían obtener de experimentos sencillos. Estos métodos han tenido mucha influencia y han sido utilizados ampliamente por los fabricantes. A pesar de su popularidad hay dos desventajas con el método de Ziegler-Nichols, la hipótesis fundamental de amortiguamiento de un cuarto en la amplitud, que resulta en sistemas con muy mala robustez, y el conocimiento limitado del proceso que se utiliza. En el Capítulo 7 se desarrollarán métodos que evitan ambas dificultades.

Los métodos estándar para el diseño de sistemas de control también se pueden adaptar para el diseño de controladores PID. Cuando se utilizan técnicas analíticas hay una correspondencia entre la complejidad del modelo y del controlador, y es necesario aproximar la dinámica del proceso por sistemas de primer y segundo orden. Las técnicas de reducción de modelos son por lo tanto necesarias para aplicar asignación de polos al control PID. En particular es necesario aproximar el retardo de tiempo. Las dinámicas no modeladas limitan el comportamiento que se puede alcanzar, y los polos en lazo cerrado que se pueden escoger.

Otra forma de usar la asignación de polos es fijar una configuración de polos y determinar ambos los parámetros del controlador y la magnitud de los polos. De esta forma es posible usar modelos de segundo orden para diseñar controladores PI y modelos de tercer orden para diseñar controladores PID. Otra forma de aplicar asignación de polos al control PID es colocar solamente los polos dominantes. La ventaja de este enfoque es que se puede aplicar a modelos de orden arbitrario. Se ha dado una especial atención a una técnica particular de asignación de polos llamada sintonía lambda, que ha sido utilizada ampliamente en la industria de procesos.

Se han analizado también una serie de métodos denominados de diseño algebraico. En estos métodos se da la función de transferencia en lazo cerrado y se obtienen los parámetros del controlador mediante cálculos algebraicos. Los parámetros del controlador se pueden determinar también por métodos de optimización, donde se intenta optimizar criterios que especifican el comportamiento sujeto a varias restricciones. Hay muchos métodos que reflejan la riqueza del problema de control. Se le ha prestado una particular atención a dos métodos, BO y SO, que se usan con frecuencia en el control de movimiento.

También se presenta un nuevo método de diseño desarrollado por los autores y sus colaboradores. En este método se intenta optimizar la atenuación de la perturbación sujeta a restricciones sobre la robustez. El método da una forma sencilla de equilibrar la atenuación de las perturbaciones en la carga con la inyección de ruido de medida que es inevitable cuando se utiliza realimentación. Combinando este método con la ponderación del punto de consigna, o una acción feedforward más elaborada, se tiene una forma agradable de conseguir también una buena respuesta a cambios en el punto de consigna. El método se puede ver como una adaptación del diseño robusto al control PID.

6.10 Notas y referencias

Hay una literatura muy amplia sobre sintonía de controladores PID. Buenas fuentes generales son los libros [321, 95, 309, 239, 90, 329, 261, 229, 354, 356, 287, 87, 302, 257, 82, 242]. Los libros claramente muestran la necesidad de una variedad de técnicas, reglas de sintonía sencillas, así como procedimientos más elaborados que se basan en el modelado de procesos, formulación de especificaciones, y diseño de control. Incluso si se utilizan reglas heurísticas sencillas, es importante darse cuenta que no hay un sustituto para el conocimiento y la comprensión. La sintonía satisfactoria de un controlador no se puede hacer sin un conocimiento acerca del modelado del proceso y de la teoría del control. Es también necesario saber que hay muchos tipos diferentes de problemas de control y en consecuencia

muchos métodos de diseño diferentes. Emplear sólo un método es tan peligroso como solamente creer en las reglas de sintonía empíricas. Los problemas de control se pueden especificar de muchas formas diferentes. Una buena revisión de diversas formas de especificar requerimientos sobre un sistema de control se da en [342, 224, 57]. Para formular especificaciones es necesario conocer los factores que limitan fundamentalmente el comportamiento de un sistema de control.

Los trabajos seminales de [378, 379] son los primeros intentos de desarrollar métodos sistemáticos para la sintonía de controladores PID. Una perspectiva interesante sobre estos trabajos se recoge en una entrevista con Ziegler; ver [54]. El método CHR, descrito en [81], es una modificación del método de Ziegler-Nichols. Este es uno de los primeros trabajos donde se menciona que se requieren diferentes métodos de sintonía para respuesta al punto de consigna y para respuesta a perturbaciones en la carga. Buena respuesta a perturbaciones en la carga es a menudo el criterio relevante en aplicaciones de control de procesos. Observe que las respuestas se pueden sintonizar independientemente teniendo un controlador que admite una estructura de dos grados de libertad. La utilidad de un parámetro de diseño se menciona también en el trabajo CHR. A pesar de sus desventajas, el método de Ziegler-Nichols ha sido el fundamento de muchos métodos de sintonía; ver [336, 228, 149]. En [362, 361, 111] se presentaron cartas de sintonía.

Los métodos de deformación del lazo fueron inspirados por los métodos de diseño de control clásicos basados en la respuesta en frecuencia; ver [342]. Aplicaciones al control PID se encuentran en [277, 135, 77, 370].

La idea del diseño algebraico se presentó en [342] y [251] como un método de diseño sistemático para especificaciones dadas; una presentación más reciente se encuentra en [57]. El diseño algebraico se aplicó al control de procesos en [323, 43, 153]. Los trabajos originales sobre el método de sintonía λ son [92] y [168]. El método se llama algunas veces el método de Dahlin; ver [95]. El método es muy popular en la industria de las papeleras donde ha sido utilizado para desarrollar procedimientos de sintonía estandarizados; ver [304] y [9]. La sintonía lambda está estrechamente relacionada con el predictor de Smith y el controlador con modelo interno; ver [323, 80, 79, 296]. Las técnicas de sintonía desarrolladas en [320, 272, 273, 319, 181] son otros ejemplos del enfoque analítico al diseño. En [296] se demostró que el control con modelo interno se reduce a un control PI y PID cuando se hacen aproximaciones adecuadas del retardo de tiempo. Un nuevo método de diseño algebraico, descrito en [153, 154] se utiliza en un controlador PID desarrollado por Foxboro. Una característica interesante es que la respuesta deseada está dada como un sistema de orden elevado.

El método de sintonía analítico da controladores que cancelan polos y ceros de la función de transferencia del proceso. Esto conduce a la falta de observabilidad o controlabilidad. Hay severas desventajas en esto como ha sido apuntado muchas veces, e.g., en [79, 312, 246]. La respuesta a perturbaciones en la carga serán muy lentas para procesos con dinámica dominada por la constante de tiempo. En [79] se da una modificación que no cancela el polo del proceso. El controlador con modelo interno de Skogestad se presenta en [318]. Este controlador evita la cancelación mediante una modificación del tiempo integral en dinámicas dominadas por la constante de tiempo.

Muchos métodos para el diseño del control se basan en técnicas de optimi-

zación. Este enfoque tiene la ventaja de que captura muchos aspectos diferentes del problema de diseño. Hay también un software potente que se puede utilizar. En [57] se encuentra un análisis general sobre el uso de la optimización para el diseño de sistemas de control. Los trabajos [298, 219] dan controladores que se optimizan con respecto a los criterios ISE, IAE, y ITAE. Otras aplicaciones para el control PID se dan en [164, 363, 259, 346, 218, 231, 370, 270, 364, 281, 377]. Los métodos BO y SO fueron introducidos en [195, 196]. Una presentación de estos métodos con muchos ejemplos se encuentra en [113, 112].

La asignación de polos es un método de diseño algebraico directo que se utiliza mucho en ingeniería de control; ver [342]. Tiene la ventaja de que se especifican directamente los polos en lazo cerrado. Muchos otros métodos de diseño se pueden interpretar también como asignación de polos. Los trabajos [105, 129] muestran cómo muchas propiedades del sistema en lazo cerrado se pueden deducir de los polos en lazo cerrado. Esto da una buena guía para seleccionar los polos en lazo cerrado que son apropiados. Un ejemplo temprano de la asignación de polos es [86, 88, 89]. Puede ser difícil escoger los polos deseados en lazo cerrado para sistemas de orden elevado. Esto se evita especificando sólo unos pocos de polos, como en el método de diseño de polos dominantes descrito en [276, 274, 275].

El desarrollo del control robusto fue un gran avance de la teoría de control que hizo posible tomar en cuenta explícitamente la robustez en el diseño del control; ver [99, 175, 131], [317, 376, 348]. Estas ideas fueron aplicadas al control PID en los trabajos [267, 25, 268, 266, 206]. Los métodos presentados en la Sección 6.8 se basan en estos trabajos.

Hay comparativamente pocos trabajos sobre controladores PID que consideran la naturaleza aleatoria de las perturbaciones. Los trabajos [132, 128, 109] son excepciones.

Hay muchos trabajos sobre comparaciones de algoritmos de control y métodos de sintonía. El trabajo [240] da muchos consejos útiles; otros trabajos interesantes son [243, 118, 119].

7

Una sustitución a Ziegler-Nichols

7.1 Introducción

Debido a la gran aceptación de los controladores PID, es muy útil disponer de reglas de sintonía simples que puedan ser aplicadas a un amplio rango de procesos. Este hecho justifica la longevidad de las reglas de Ziegler-Nichols. Estas reglas han sido utilizadas durante más de medio siglo a pesar incluso de tener graves inconvenientes. En este capítulo presentamos nuevos métodos de sintonía basados en el espíritu de los de Ziegler y Nichols.

Tal y como se comentó en la Sección 4.2, el diseño de sistemas de control es un gran problema. Cualquier problema de diseño debería tomar en cuenta las perturbaciones de carga, el ruido de medida y el seguimiento de referencias. Para el desarrollo de las reglas simples seguiremos las principales ideas utilizadas por Ziegler y Nichols. Nos centraremos por tanto en las perturbaciones de carga mediante la maximización de la ganancia integral pero nos desviaremos de Ziegler y Nichols con la incorporación de una restricción de robustez. En este capítulo hemos escogido el requisito de que la sensibilidad combinada sea menor que $M = 1.4$. Por otro lado, el ruido de medida será gestionado desintonizando los controladores si las ganancias son demasiado grandes y el seguimiento de referencias se aborda utilizando la ponderación del punto de consigna.

El procedimiento que utilizamos es esencialmente el mismo empleado por Ziegler y Nichols, donde hemos seleccionado un gran conjunto de procesos representativos incluyendo una amplia variedad de sistemas, fundamentalmente con respuestas a escalón monótonas y típicos en control de procesos. Los controladores para cada proceso del conjunto son obtenidos aplicando el diseño MIGO que fue descrito en la Sección 6.8, el cual está basado en el criterio dado anteriormente. Una vez obtenidos los parámetros del controlador, tendremos que encontrar las correlaciones con los parámetros del procesos normalizado. Las reglas simples de sintonía obtenidos son llamadas AMIGO, que quiere decir diseño Approximate MIGO (diseño MIGO aproximado).

El procedimiento muestra que es en realidad posible obtener fórmulas simples de sintonía. Un resultado muy importante es que es necesario utilizar más

información sobre el proceso que la utilizada por Ziegler y Nichols. Reglas de sintonía basadas en respuestas en escalón se pueden utilizar para modelos FOTD, donde es necesario utilizar los tres parámetros K_p , L y T y no dos parámetros, $a = K_p L/T$ y L , tal y como fue sugerido por Ziegler y Nichols. Para control PI es posible obtener reglas de sintonía cercanas a reglas óptimas para el conjunto de prueba al completo. En el caso de reglas para control PID, se pueden obtener reglas óptimas para procesos balanceados y con retardo dominante. Para procesos con constante de tiempo dominante es necesario disponer de mejor información sobre el proceso.

Para el método de respuesta en frecuencia donde Ziegler y Nichols caracterizaron el proceso por dos parámetros, K_{180} y T_{180} , hemos mostrado que es necesario considerar un tercer parámetro, por ejemplo, la ganancia estática K_p . Incluso con tres parámetros no es posible obtener reglas cercanas al óptimo para todos los procesos en el conjunto de ensayo. Sin embargo, es posible obtener reglas de sintonía conservadoras para ambos tipo de controladores, PI y PID.

El método de diseño utilizado puede resultar en controladores con altas ganancias para procesos con constante de tiempo dominante. Esto podría llevar a grandes variaciones en la señal de control en presencia de ruido. En algunos casos, puede ser por tanto necesario buscar un compromiso entre la atenuación de perturbaciones de carga y la inyección de ruido de medida. Esto podría ser llevado a cabo mediante la desintonización de controladores. En este capítulo también se mostrarán algunos métodos sobre estas ideas.

El análisis de controladores sobre el conjunto de ensayo ha permito también aumentar la compresión sobre los controles PI y PID. Se muestra que la acción derivativa sólo aporta mejoras moderadas para procesos balanceados y con retardo dominante pero que proporciona grandes mejoras para procesos con constante de tiempo dominante. También se muestra que existe un amplio rango de procesos donde es ventajoso tener $T_i < 4T_d$. Hay que tener en cuenta que los controladores implementados con la estructura en serie no permiten esta sintonía. Se puede obtener también interesantes fórmulas que dan la razón de tiempo de residencia medio para sistemas en lazo abierto y lazo cerrado. Esto hace posible estimar los tiempos de respuesta en lazo cerrado esperados.

En la siguiente sección se presenta el conjunto de procesos de prueba. Haciendo uso de este conjunto, en la Sección 7.3 se derivan las reglas de sintonía de AMIGO basadas en experimentos de respuesta a escalón para controladores PI y en la Sección 7.4 para controladores PID. Las reglas de sintonía de AMIGO basadas en experimentos de respuesta en frecuencia se presentan en la Sección 7.5. Tal y como se comenta en la Sección 7.6, se pueden obtener reglas de sintonía más eficientes para el control PID de procesos con constante de tiempo dominante si se hace uso de un modelo de segundo orden. En la Sección 7.7, las reglas MIGO y AMIGO son comparadas para tres procesos diferentes, uno con constante de tiempo dominante, uno balanceado y otro con retardo dominante respectivamente. La Sección 7.8 y la Sección 7.9 tratan los problemas de filtrado de ruido y reducción de ganancias a altas frecuencias mediante la desintonía de controladores.

7.2 El conjunto de prueba

El control PID no es apropiado para todo tipo de procesos. En [138] se sugiere utilizar controladores PID para aquellos procesos que vienen caracterizados esencialmente por respuestas a escalón monótonas. Una forma de poder caracterizar tales procesos es utilizar el índice de monotonía:

$$\alpha = \frac{\int_0^\infty g(t)dt}{\int_0^\infty |g(t)|dt}, \quad (7.1)$$

donde g es la respuesta a un impulso del sistema. Sistemas con $\alpha = 1$ tienen respuestas a escalón monótonas, y sistemas con $\alpha > 0,8$ se consideran esencialmente monótonos. Las reglas de sintonía presentes en este capítulo son derivadas utilizando un conjunto de ensayo de procesos esencialmente monótonos.

Los siguientes 134 procesos se utilizaron para derivar las reglas de sintonía:

$$P_1(s) = \frac{e^{-s}}{1+sT},$$

$$T = 0,02, 0,05, 0,1, 0,2, 0,3, 0,5, 0,7, 1, \\ 1,3, 1,5, 2, 4, 6, 8, 10, 20, 50, 100, 200, 500, 1000$$

$$P_2(s) = \frac{e^{-s}}{(1+sT)^2},$$

$$T = 0,01, 0,02, 0,05, 0,1, 0,2, 0,3, 0,5, 0,7, 1, \\ 1,3, 1,5, 2, 4, 6, 8, 10, 20, 50, 100, 200, 500$$

$$P_3(s) = \frac{1}{(s+1)(1+sT)^2},$$

$$T = 0,005, 0,01, 0,02, 0,05, 0,1, 0,2, 0,5, 2, 5, 10$$

$$P_4(s) = \frac{1}{(s+1)^n},$$

$$n = 3, 4, 5, 6, 7, 8$$

$$P_5(s) = \frac{1}{(1+s)(1+\alpha s)(1+\alpha^2 s)(1+\alpha^3 s)},$$

$$\alpha = 0,1, 0,2, 0,3, 0,4, 0,5, 0,6, 0,7, 0,8, 0,9$$

$$P_6(s) = \frac{1}{s(1+sT_1)} e^{-sL_1},$$

$$L_1 = 0,01, 0,02, 0,05, 0,1, 0,2, 0,3, 0,5, 0,7, 0,9, 1,0, \quad T_1 + L_1 = 1$$

$$P_7(s) = \frac{T}{(1+sT)(1+sT_1)} e^{-sL_1}, \quad T_1 + L_1 = 1,$$

$$T = 1, 2, 5, 10 \quad L_1 = 0,01, 0,02, 0,05, 0,1, 0,3, 0,5, 0,7, 0,9, 1,0$$

$$P_8(s) = \frac{1 - \alpha s}{(s+1)^3}, \quad \alpha = 0,1, 0,2, 0,3, 0,4, 0,5, 0,6, 0,7, 0,8, 0,9, 1,0, 1,1$$

$$P_9(s) = \frac{1}{(s+1)((sT)^2 + 1,4sT + 1)}, \quad T = 0,1, 0,2, 0,3, 0,4, 0,5, 0,6, 0,7, 0,8, 0,9, 1,0. \quad (7.2)$$

Estos procesos son representativos para muchos de los sistemas encontrados en el control de procesos. El conjunto de ensayo incluye sistemas con retardo dominante, con constante de tiempo dominante y procesos integradores. Todos los procesos tienen respuesta a escalón monótona excepto P_8 y P_9 , donde el rango de parámetro para estos procesos se escogen de tal forma que los sistemas son esencialmente monótonos con $\alpha \geq 0,8$. El tiempo de retardo normalizado varía entre 0 y 1 para el proceso P_1 y sólo de 0,14 a 1 para P_2 . El proceso P_6 es un sistema con integrador y por tanto $\tau = 0$. El resto de procesos tiene valores de τ en el rango $0 < \tau < 0,5$.

7.3 Control PI

Los procesos en el conjunto de ensayo (7.2) son primero aproximados por simples modelos FOTD

$$P(s) = \frac{K_p}{1+sT} e^{-sL}, \quad (7.3)$$

donde K_p es la ganancia estática, T es la constante de tiempo (también llamado retraso), y L es el tiempo de retardo. Los procesos con componente integrativa son aproximados por el modelo

$$P(s) = \frac{K_v}{s} e^{-sL}, \quad (7.4)$$

donde K_v es la ganancia de velocidad y L es el tiempo de retardo. El modelo (7.4) se puede considerar como el límite de (7.3) haciendo que K_p y T tiendan a infinito de tal manera que $K_p/T = K_v$ sea constante. Los parámetros de (7.3) y (7.4) pueden determinarse a partir de un experimento de respuesta a escalón utilizando los métodos presentados en la Sección 2.7.

Las reglas de sintonía fueron obtenidas de la siguiente manera. Se aplicó el método de diseño MIGO (ver Sección 6.8) con $M = 1,4$ a todos los procesos del conjunto de ensayo (7.2) obteniendo los parámetros del controlador PI, K y T_i . Entonces, se obtuvieron las reglas AMIGO encontrando las relaciones entre los parámetros del controlador y los parámetros del proceso.

La Figura 7.1 muestra las relaciones entre los parámetros del controlador y los del proceso para todos los procesos del conjunto de ensayo.

La ganancia del controlador se normaliza multiplicándola por la ganancia estática del proceso K_p o con el parámetro $a = K_p L / T = K_v L$. El tiempo integral se normaliza dividiéndolo por T o por L . Los parámetros para el proceso integrativo P_6 son sólo normalizados con a y L ya que K_p y T son infinitos para estos procesos. En la Figura 7.1 los parámetros del controlador se dibujan frente al tiempo de retardo normalizado $\tau = L/(L+T)$.

La figura muestra que existe una buena correlación entre los parámetros del controlador normalizado y el tiempo de retardo normalizado. Esto indica que es posible desarrollar buenas reglas de sintonía basadas en modelos FOTD. Sin embargo, hay que tener en cuenta que existen variaciones significativas en los parámetros con el tiempo de retardo normalizado τ .

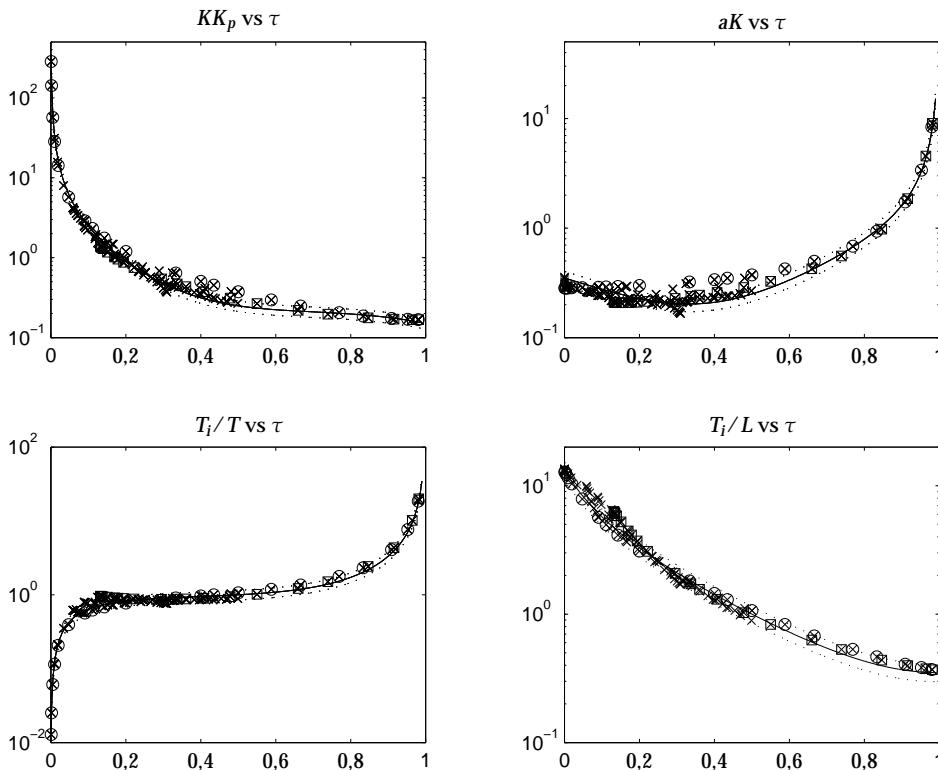


Figura 7.1 Representación de los parámetros del controlador PI normalizado frente al tiempo de retardo normalizado τ . Las líneas continuas corresponden a la regla de diseño AMIGO (7.5), y las líneas de puntos suspensivos indican variaciones en los parámetros del 15 por ciento. Los círculos representan los parámetros obtenidos para el proceso P_1 y los cuadrados los obtenidos para el proceso P_2 .

Ziegler y Nichols intentaron encontrar reglas que no dependiesen de τ . La figura 7.1 muestra que los parámetros normalizados KK_p , aK , T_i/T , y T_i/L varían tanto como dos órdenes de magnitud con τ , de lo que por tanto se deduce que no es posible encontrar reglas eficientes que no dependan de τ .

La líneas sólidas de la Figura 7.1 corresponden a fórmula de sintonía AMIGO,

$$K = \frac{0,15}{K_p} + \left(0,35 - \frac{LT}{(L+T)^2} \right) \frac{T}{K_p L} \quad (7.5)$$

$$T_i = 0,35L + \frac{13LT^2}{T^2 + 12LT + 7L^2},$$

y las líneas con puntos suspensivos muestran los límites para variaciones del 15 por ciento en los parámetros del controlador. Casi todos los procesos incluidos en el conjunto de ensayo caen dentro de estos límites.

Para procesos integrativos, K_p y T tienden a infinito y $K_p/T = K_v$. Por con-

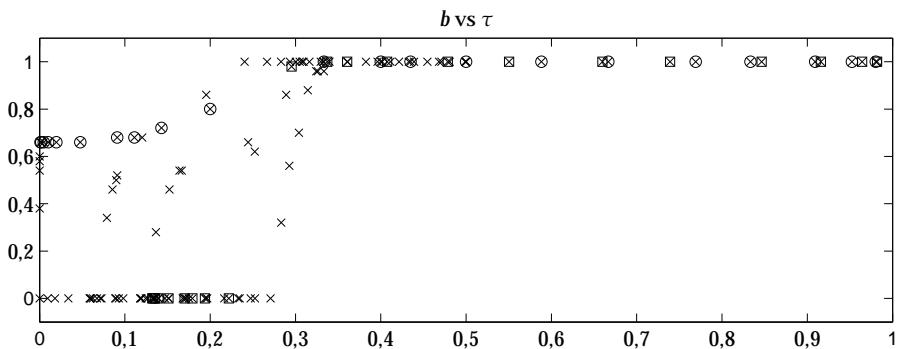


Figura 7.2 Ponderación del punto de consigna como una función de τ para el conjunto de ensayo (7.2). Los círculos representan los parámetros obtenidos para el proceso P_1 y los cuadrados los parámetros obtenidos desde el proceso P_2 .

siguiente, las reglas de sintonía AMIGO (7.5) pueden ser simplificadas a

$$K = \frac{0,35}{K_v L} \quad (7.6)$$

$$T_i = 13,4L.$$

para procesos integrativos.

La regla de sintonía (7.5) se puede ver como una sustitución al método de Ziegler-Nichols basado en la respuesta a escalón para control PI. Obsérvese que la regla fue diseñada para una sensibilidad de $M = 1,4$. Se pueden obtener reglas similares para otros valores del parámetro de diseño.

Ponderación del punto de consigna

El método de diseño MIGO también da valores apropiados de b , los cuales se determinan tal que el pico de resonancia de la función de transferencia entre el punto de consigna y la salida del proceso llegue a estar cercano a uno, tal como se comentó en la Sección 5.3. La Figura 7.2 muestra los valores del parámetro b para el conjunto de ensayo (7.2).

La Figura 7.2 muestra que la correlación con el parámetro τ no es tan buena como para los parámetros de realimentación y que existe una diferencia más grande entre los modelos FOTD puros P_1 y los otros procesos. La ponderación del punto de consigna debería ser $b = 1$ para procesos con $\tau > 0,3$.

7.4 Control PID

La Figura 7.3 ilustra las relaciones entre los parámetros del controlador PID obtenidos utilizando el diseño MIGO y los parámetros de proceso para todos los procesos en el conjunto de ensayo.

La figura indica que las variaciones de los parámetros del controlador normalizado son de varios órdenes de magnitud. Podemos por tanto concluir que no

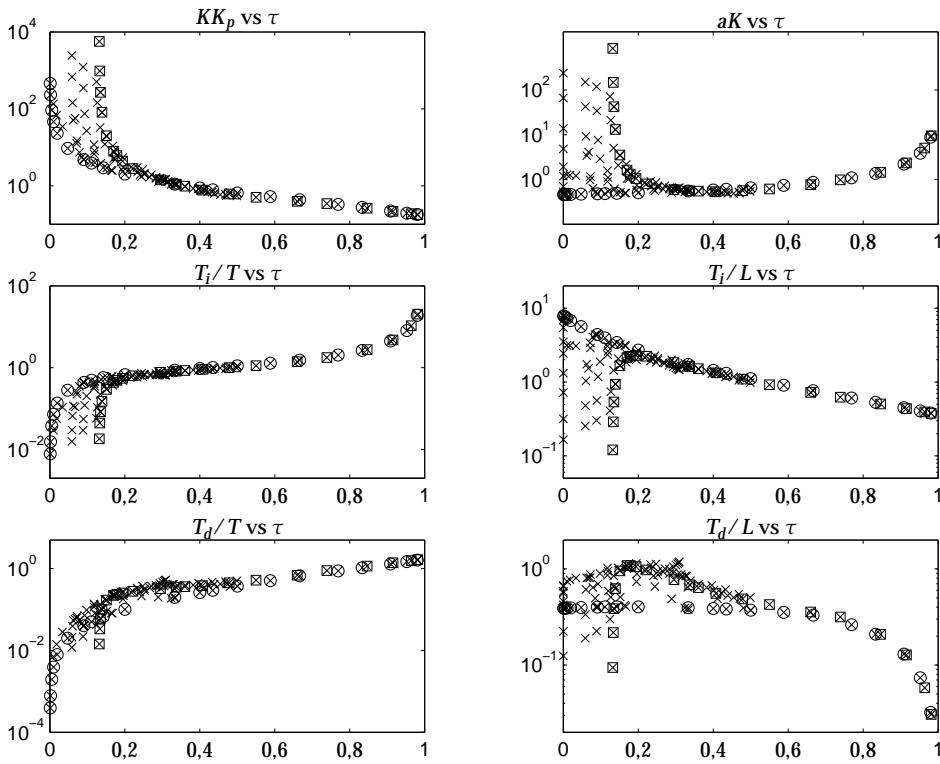


Figura 7.3 Parámetros del controlador PID normalizado como una función del tiempo de retardo normalizado τ . Los controladores para procesos P_1 son marcados con \circ y los controladores para P_2 con \square .

es posible encontrar buenas reglas de sintonía universales que no dependan del tiempo de retardo normalizado τ . Recuérdese que Ziegler y Nichols sugirieron las reglas $aK = 1,2$, $T_i = 2L$ y $T_d = 0,5L$. La Figura 7.3 muestra que esos parámetros son sólo apropiados para muy pocos procesos del conjunto de ensayo.

En la Figura 7.3, los parámetros del controlador para procesos P_1 son marcados con círculos mientras que para P_2 con cuadrados. Para $\tau < 0,5$, las ganancias para P_1 son típicamente más pequeñas que para los otros procesos y el tiempo integral más grande. Este es lo contrario que lo que ocurrió para el control PI tal y como se puede ver en la Figura 7.1. El proceso P_2 tiene una ganancia que es más grande y un tiempo integral más bajo que para los otros procesos.

Para el control PI, fue posible obtener reglas de sintonía simples donde los parámetros del controlador obtenidos con las reglas AMIGO difirieron menos del 15 por ciento que aquellas obtenidas con las reglas MIGO para la mayoría de los procesos en el conjunto de ensayo. La Figura 7.3 indica que sólo se pueden obtener reglas de sintonía universales para el control PID con $\tau \geq 0,5$.

Para $\tau < 0,5$ existe una propagación significante de los parámetros normalizados, lo que implica que no es posible encontrar reglas de sintonía universales para procesos con constante de tiempo dominante. Obsérvese que la ganancia

y el tiempo integral están bien definidos para $0,3 < \tau < 0,5$ pero que hay una variación considerable del tiempo derivativo normalizado en ese intervalo.

Debido a la gran dispersión en los valores de los parámetros para $\tau < 0,5$, en estos casos vale la pena modelar el proceso más exactamente para obtener una buena sintonía de controladores PID. Los modelos de proceso (7.3) y (7.4) modelan procesos estables con tres parámetros y procesos integrativos con dos parámetros. En la práctica, es muy difícil obtener más parámetros de proceso más allá del experimento de respuesta a escalón. Un experimento de respuesta a escalón no es por tanto siempre suficiente para sintonizar con exactitud controladores PID para $\tau < 0,5$.

Sin embargo, es posible encontrar reglas de sintonía *conservadoras* para $\tau < 0,5$ escogiendo aquellos controladores de la Figura 7.3 con ganancias más pequeñas y tiempos integrales más grandes. Antes de desarrollar tales reglas, discutiremos la razón de por qué no se pueden encontrar reglas de sintonía universales para procesos con constante de tiempo dominante.

Problemas con la estructura FOTD

El criterio utilizado es maximizar la ganancia integral k_i . Las limitaciones fundamentales vienen dadas por el tiempo de retardo *verdadero* del proceso, el cual hemos denotado por L_0 . La ganancia integral es proporcional a la frecuencia de cruce de ganancia ω_{gc} del sistema en lazo cerrado. La frecuencia de cruce de ganancia ω_{gc} está típicamente limitada a

$$\omega_{gc}L_0 < 0,5.$$

Cuando un proceso se aproxima por un modelo FOTD, el tiempo de retardo aparente L es más grande que el tiempo de retardo verdadero L_0 porque las constantes de tiempo se aproximan por retardos adicionales. Esto implica que la ganancia integral obtenida para el modelo FOTD será más baja que para un diseño basado en el modelo real. La situación es particularmente pronunciada para sistemas con τ pequeña.

Considere un control PI de sistemas de primer orden, por ejemplo, procesos con las funciones de transferencia

$$P(s) = \frac{K_p}{1+sT} \quad \text{o} \quad P(s) = \frac{K_v}{s}.$$

Debido a que estos sistemas no tienen tiempos de retardo, no existe limitación en las dinámicas y por tanto se pueden obtener arbitrariamente altas ganancias integrables. Ya que estos procesos pueden ser ajustados perfectamente por los modelos (7.3) y (7.4), las reglas de diseño reflejan esta propiedad. Los parámetros del procesos son $L = 0$, $a = 0$ y $\tau = 0$, y ambos, los métodos de diseño MIGO y AMIGO (7.5) dan ganancias integrales infinitas.

Considérese un control PID de un sistema de segundo orden con las funciones de transferencia

$$P(s) = \frac{K_v}{s(1+sT_1)} \quad \text{y} \quad P(s) = \frac{K_p}{(1+sT_1)(1+sT_2)}.$$

Debido a que los sistemas no tienen tiempos de retardo, es posible tener controladores con ganancias integrales arbitrariamente altas. La primera función de transferencia tiene $\tau = 0$ y el proceso de segundo orden tiene valores de τ en el rango $0 \leq \tau < 0,13$, donde $\tau = 0,13$ corresponde a $T_1 = T_2$. Cuando estas funciones de transferencia se aproximan por un modelo FOTD, una de las constantes de tiempo será aproximada por un tiempo de retardo. Ya que el modelo aproximado tiene un tiempo de retardo, existirán limitaciones en la ganancia integral.

Podemos por tanto concluir que para $\tau < 0,13$ hay procesos en el conjunto de ensayo que permiten ganancias integrales infinitamente grandes. Esto explica la amplia dispersión de los parámetros del controlador para pequeños valores de τ . La difusión es infinitamente grande para $\tau < 0,13$ y decrece para valores más grandes de τ . Por consiguiente, para pequeños τ el modelado mejorado aporta un beneficio significante.

Una vía para evitar la dificultad consiste en utilizar un modelo más complicado, tal como

$$P(s) = \frac{b_1 s + b_2 s}{s^2 + a_1 s + a_2} e^{-sL}.$$

Sin embargo, es muy difícil estimar los parámetros de este modelo de manera precisa desde un simple experimento de respuesta a escalón. Reglas de diseño para modelos con cinco parámetros puede ser también engoroso. Debido a que el problema ocurre para valores pequeños de tau, puede ser posible aproximar el proceso por

$$P(s) = \frac{K_v}{s(1+sT)} e^{-sL},$$

que sólo tiene tres parámetros. En lugar de desarrollar reglas de sintonía para modelos más complicados, puede ser mejor computar simplemente parámetros del controlador basados en el modelo estimado.

Reglas de sintonía conservadoras (AMIGO)

La Figura 7.3 muestra que no es posible encontrar reglas de sintonía óptimas para controladores PID que estén basadas en los modelos simples (7.3) o (7.4). Sin embargo, es posible encontrar reglas de sintonía conservadoras robustas con un rendimiento más bajo. Las reglas son cercanas al diseño MIGO para el proceso P_1 , por ejemplo, el proceso que da el controlador con ganancia más pequeña y el tiempo integral más grande; ver Figura 7.3.

Las reglas de sintonía AMIGO sugeridas para controladores PID son

$$\begin{aligned} K &= \frac{1}{K_p} \left(0,2 + 0,45 \frac{T}{L} \right) \\ T_i &= \frac{0,4L + 0,8T}{L + 0,1T} L \\ T_d &= \frac{0,5LT}{0,3L + T}. \end{aligned} \tag{7.7}$$

Para procesos con integrador, la ecuación 7.7 puede escribirse como

$$\begin{aligned} K &= \frac{0,45}{K_v L} \\ T_i &= 8L \\ T_d &= 0,5L. \end{aligned} \tag{7.8}$$

La Figura 7.4 compara la regla de sintonía (7.7) con los parámetros del controlador dado en la Figura 7.3.

La regla de sintonía (7.7) describe bien la ganancia del controlador K para un proceso con $\tau > 0,3$. Para pequeños τ , la ganancia del controlador se ajusta bien para procesos P_1 , pero la regla AMIGO subestima la ganancia para otros procesos.

El tiempo integral T_i está bien descrito por la regla de sintonía (7.7) para $\tau > 0,2$. Para pequeños τ , el tiempo integral se ajusta bien para procesos P_1 , pero la regla AMIGO lo sobreestima para otros procesos.

La regla de sintonía (7.7) describe bien el tiempo derivativo T_d para procesos con $\tau > 0,5$. En el rango $0,3 < \tau < 0,5$, el tiempo derivativo puede ser hasta un factor de 2 más grande que el valor obtenido por la regla AMIGO. Si los valores del tiempo derivativo para la regla AMIGO se utilizan en este rango, se disminuye la robustez; el valor de M se puede incrementar en un 15 por ciento. Para $\tau < 0,3$, la regla de sintonía AMIGO da un tiempo derivativo que algunas veces es más pequeño y otras vez más grande que el obtenido con MIGO. A pesar de esto, se observa que el método AMIGO aporta una sintonía más conservadora para todos los procesos en el conjunto de ensayo, principalmente a causa de que la ganancia del controlador disminuye y el tiempo integral se incrementa.

La regla de sintonía (7.7) tiene la misma estructura que el método Cohen-Coon, pero los parámetros difieren significativamente.

Robustez

La Figura 7.5 muestra las curvas de Nyquist de las funciones de transferencia del lazo obtenidas cuando los procesos en el conjunto de ensayo (7.2) se controlan con los controladores PID sintonizados con la regla conservadora AMIGO (7.7).

Cuando se utiliza MIGO, todas las curvas de Nyquist están fuera del círculo M . Con AMIGO, existen algunos procesos donde las curvas de Nyquist están dentro del círculo. Una investigación muestra que la acción derivativa es demasiado pequeña en esos casos comparada con las curvas de T_d/L frente τ en la Figura 7.4. El incremento de M es a lo sumo del 15 por ciento con la regla AMIGO. Si el incremento no es aceptable, se puede aumentar la acción derivativa o disminuir la ganancia en un 15 por ciento.

Ponderación del punto de consigna

La Figura 7.6 muestra los valores del parámetro b para el conjunto de ensayo (7.2).

La correlación entre b y τ no es tan buena, pero una regla conservadora y simple consiste en escoger b como

$$b = \begin{cases} 0, & \text{para } \tau \leq 0,5 \\ 1, & \text{para } \tau > 0,5. \end{cases} \tag{7.9}$$

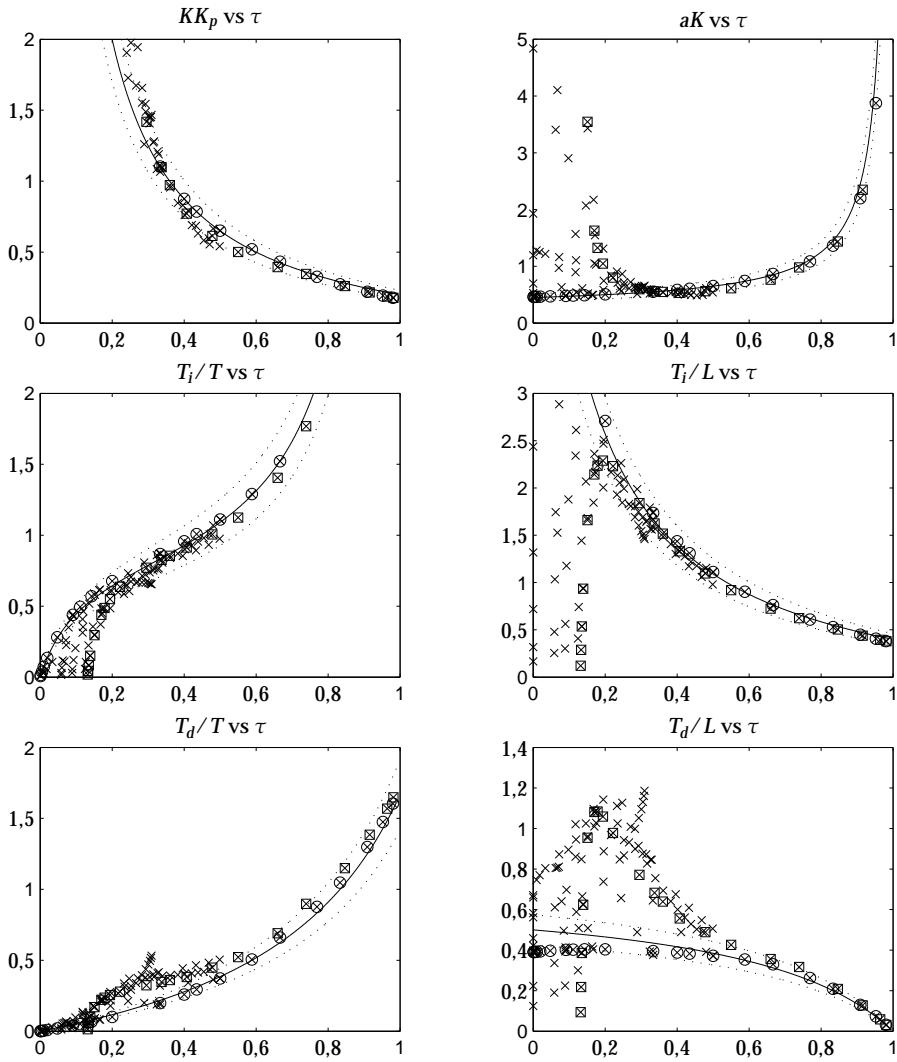


Figura 7.4 Parámetros del controlador normalizado como función del tiempo de retardo normalizado τ . La línea continua corresponde a la regla de sintonía (7.7) y las líneas en puntos suspensivos indican variaciones del 15 por ciento en los parámetros. Los círculos representan los parámetros obtenidos para el proceso P_1 y los cuadrados los obtenidos con el proceso P_2 .

Los beneficios de la acción derivativa

Debido a que se escogió la maximización de la ganancia integral como un criterio de diseño, podemos juzgar los beneficios de la acción derivativa por la razón de la ganancia integral para controles PI y PID. La Figura 7.7 muestra esta razón para el conjunto de ensayo, excepto para unos pocos procesos con una razón alta para pequeños valores de τ .

La figura muestra que los beneficios de la acción derivativa son marginales para procesos con retardo dominante pero que los beneficios se incrementan con

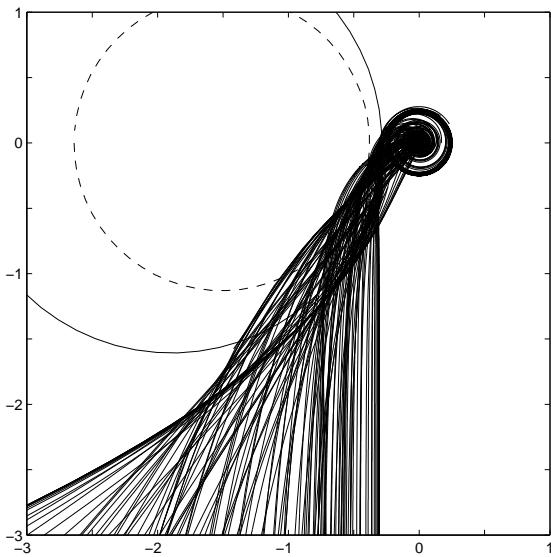


Figura 7.5 Curvas de Nyquist de las funciones de transferencia del lazo obtenidas cuando controladores PID sintonizados acorde a (7.7) se aplican al conjunto de ensayo (7.2). El círculo de línea continua corresponde a $M = 1,4$ y el de líneas a tramos a un círculo donde M se incrementa en un 15 por ciento.

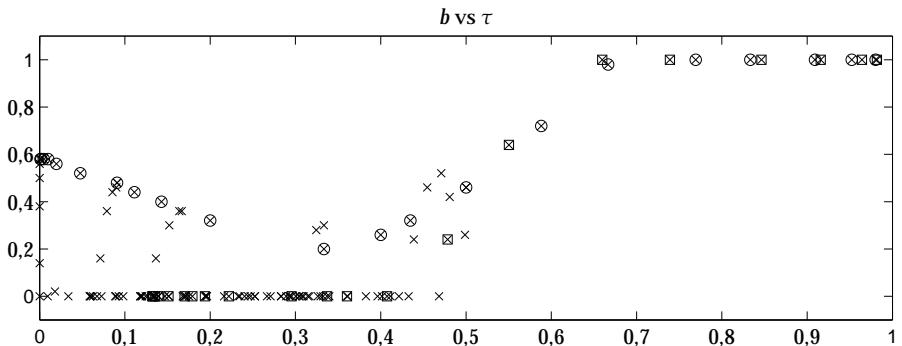


Figura 7.6 Ponderación del punto de consigna como una función del τ para el conjunto de ensayo (7.2). Los círculos representan los parámetros obtenidos con el proceso P_1 y los cuadrados los obtenidos con el proceso P_2 .

la disminución de τ . Para $\tau = 0,5$, se puede doblar la ganancia integral, y para valores de $\tau < 0,15$, la ganancia integral se puede incrementar arbitrariamente para algunos procesos.

La razón T_i/T_d

La razón T_i/T_d es de interés por varias razones. Es una medida de importancia relativa de las acciones integral y derivada. Muchos controladores PID se implementan en serie, lo que requiere que la razón sea mayor que 4. Muchas reglas clásicas de sintonía fijan la razón a 4. La Figura 7.8 muestra dicha razón para el

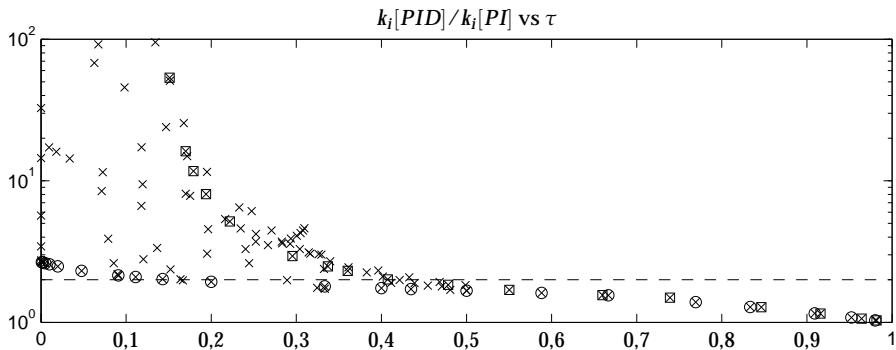


Figura 7.7 La razón de la ganancia integral con controles PI y PID como función del tiempo de retardo normalizado τ . La línea discontinua corresponde a la razón $k_i[\text{PID}]/k_i[\text{PI}] = 2$. Los controladores para los procesos P_1 son marcados con círculos y los controladores para P_2 con cuadrados.

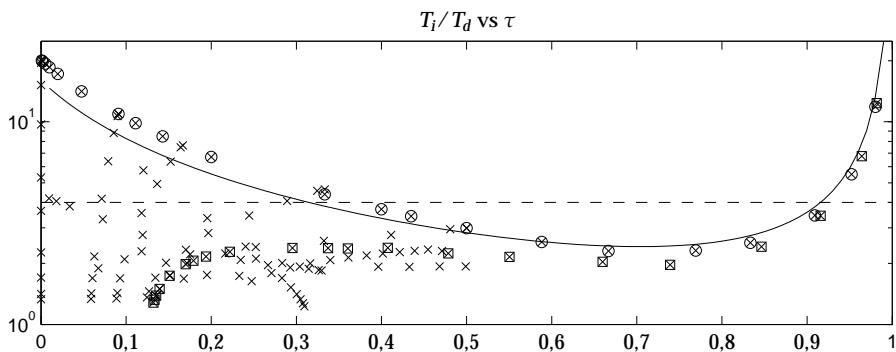


Figura 7.8 La razón entre T_i y T_d como función del tiempo de retardo normalizado τ . El proceso P_1 está marcado con círculos y el proceso P_2 con cuadrados. La línea discontinua corresponde a la razón $T_i/T_d = 4$ y la línea continua a la razón dada por la regla de sintonía AMIGO (7.7).

conjunto completo del ensayo.

La figura muestra que existe una variación significante en la razón T_i/T_d , particularmente para pequeños valores de τ . La razón es cercana a 2 para $0,5 < \tau < 0,9$ e incrementa a infinito cuando τ se aproxima a 1 porque la acción derivativa es cero para procesos con tiempo de retardo puro.

La Figura 7.8 también muestra la razón obtenida por la regla de sintonía AMIGO (7.7). La razón es menor que cuatro para procesos con $0,3 < \tau < 0,9$, lo que quiere decir que la regla de sintonía no puede ser utilizada para controladores implementados en la forma en serie para estos procesos. Sin embargo, se observa que los cambios de rendimiento y robustez son mínimos si la regla de sintonía (7.7) se modifica tal que $T_d = T_i/4$ para estos procesos. La Figura 7.9 muestra los diagramas de Nyquist de las funciones de transferencia del lazo obtenidas cuando los procesos en el conjunto de ensayo con $0,3 < \tau < 0,9$ se sintonizan tal que la ganancia K y el tiempo integral T_i se obtienen de (7.7), y el tiempo derivativo

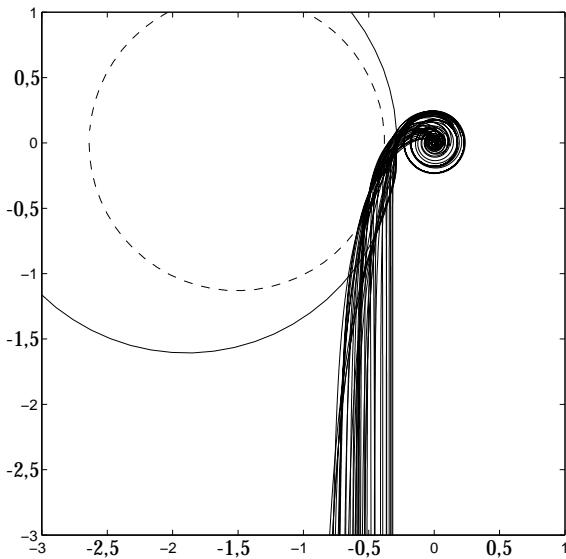


Figura 7.9 Curvas de Nyquist de la función de transferencia del lazo obtenidas desde los procesos del conjunto de ensayo con $0,3 < \tau < 0,9$ cuando el controlador es sintonizado con $T_d = T_i/4$.

se calcula como $T_d = T_i/4$. La figura muestra que la robustez es más o menos la misma que para (7.7) (comparar con la Figura 7.5).

El tiempo de residencia medio

El parámetro T_{63} , que es el tiempo cuando la respuesta a un escalón ha alcanzado el 63 por ciento (un factor de $(1 - 1/e)$) de su valor en régimen permanente, es una medida razonable del tiempo de respuesta de sistemas estables. Es sencillo determinar el parámetro mediante simulación, pero no mediante cálculos analíticos. Para procesos FOTD, tenemos $T_{ar} = T_{63}$. El tiempo de residencia medio T_{ar} es de hecho una buena estimación del T_{63} para sistemas con respuesta a un escalón esencialmente monótona. Para todos los procesos estables en el conjunto de ensayo tenemos $0,99 < T_{63}/T_{ar} < 1,08$.

El tiempo de residencia medio es fácil de calcular analíticamente. Considérese un sistema en lazo cerrado obtenido cuando un proceso con función de transferencia $P(s)$ se controla con un controlador PID con ponderación del punto de consigna. La función de transferencia desde el punto de consigna a la salida es

$$G_{yy_{sp}}(s) = \frac{P(s)C_{ff}(s)}{1 + P(s)C(s)},$$

donde

$$C_{ff}(s) = bK + \frac{k_i}{s}.$$

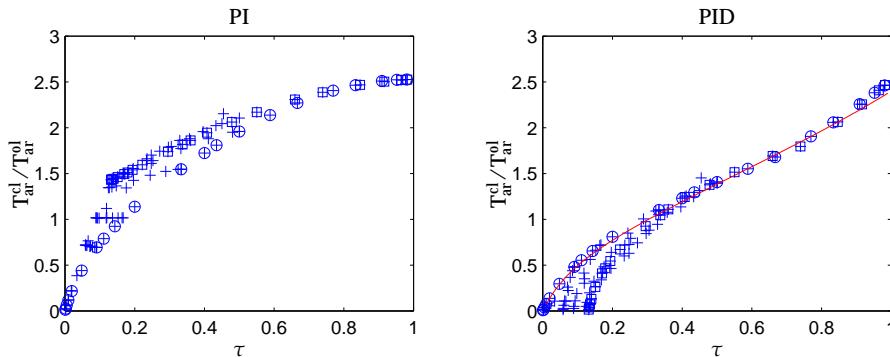


Figura 7.10 La razón del tiempo de residencia medio del sistema en lazo cerrado y en lazo abierto para control PI a la izquierda y para control PID a la derecha.

Tras una serie de cálculos sencillos, aunque tediosos, se obtiene

$$T_{ar} = -\frac{G'_{yy_{sp}}(0)}{G_{yy_{sp}}(0)} = T_i \left(1 - b + \frac{1}{KK_p} \right). \quad (7.10)$$

La Figura 7.10 muestra los tiempos de residencia medios del sistema en lazo cerrado divididos por el tiempo de residencia medio del sistema en lazo abierto. La Figura 7.10 muestra que para el control PID el sistema en lazo cerrado es más rápido que el sistema en lazo abierto cuando $\tau < 0,3$ y más lento para $\tau > 0,3$.

7.5 Métodos de respuesta en frecuencia

En esta sección investigaremos si es posible obtener reglas de sintonía simples similares a las del método de respuesta en frecuencia de Ziegler-Nichols.

Parametrización

Ziegler-Nichols caracterizaron los procesos por dos parámetros K_{180} y T_{180} cuando desarrollaron su método de respuesta en frecuencia para la sintonía del controlador. Las reglas de sintonía de Ziegler-Nichols no usaban suficiente información y resultan en una sintonía demasiada agresiva, dando lugar a sistemas en lazo cerrado no robustos. Cuando se investigó el método de respuesta a un escalón se encontró que se podría obtener una mejora significativa mediante la inclusión de un tercer parámetro adicional del proceso, la ganancia estática. En esta sección, se investigará si para el método basado en el dominio de la frecuencia se obtienen mejoras similares.

Para el método de respuesta a un escalón utilizamos el tiempo de retardo normalizado τ como un parámetro que caracteriza al proceso. En este caso, el parámetro correspondiente en el dominio de la frecuencia es la razón de ganancia $\kappa = K_{180}/K_p$.

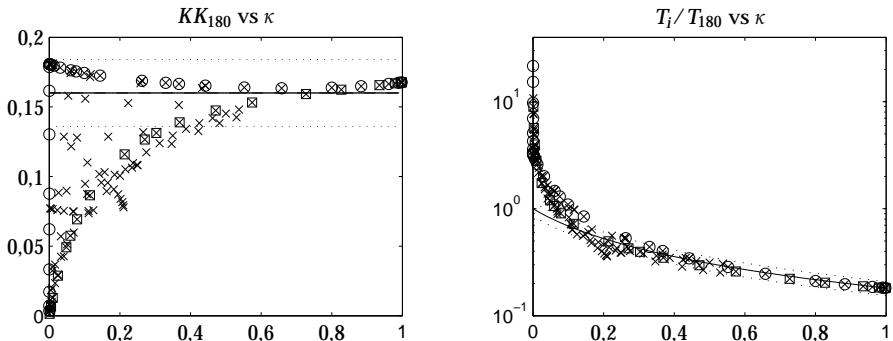


Figura 7.11 Parámetros del controlador normalizado frente a la razón de ganancia κ para procesos estables con $M = 1.4$. Las líneas continuas corresponden a la regla de sintonía (7.11), y las líneas punteadas indican variaciones del 15 por ciento en dicha regla. Los círculos representan datos obtenidos para el proceso P_1 y los cuadrados datos obtenidos para P_2 .

Reglas de sintonía de control PI para procesos balanceados y con retardo dominante

El método de diseño MIGO ha sido aplicado a todos los procesos en el conjunto de ensayo (7.2). La Figura 7.11 muestra los parámetros del controlador obtenidos para $M = 1.4$. La figura muestra que existe una propagación significativa de los parámetros del controlador para procesos con constante de tiempo dominante.

Las reglas de sintonía de Ziegler-Nichols tiene los valores constantes $KK_{180} = 0,4$ y $T_i/T_{180} = 0,8$, para todos los valores de κ . La Figura 7.11 muestra que puede ser razonable tener un valor constante KK_{180} para $\kappa > 0,5$, pero no para pequeños valores de κ . La ganancia $KK_{180} = 0,4$ sugerida por Ziegler y Nichols es claramente demasiado alta, lo que explica la pobre robustez de su método. El tiempo integral sugerido por Ziegler y Nichols, $T_i = 0,8T_{180}$, es demasiado alto excepto para procesos con valores muy pequeños de κ .

La Figura 7.11 muestra que no es posible capturar todos los datos mediante una regla de sintonía. Sin embargo, puede ser posible obtener un regla para procesos balanceados y de retardo dominante. La Figura 7.11 muestra las gráficas correspondientes a la siguiente regla de sintonía.

$$KK_{180} = 0,16 \quad (7.11)$$

$$\frac{T_i}{T_{180}} = \frac{1}{1 + 4,5\kappa}.$$

La regla de sintonía (7.11) no es apropiada para procesos con constante de tiempo dominante, pero genera parámetros del controlador que están bastante cerca del óptimo para procesos con $\kappa > 0,2$. En particular, obsérvese que la razón T_i/T_{180} se reduce en un factor de tres cuando κ crece de 0,2 a 1.

La Figura 7.12 muestra los diagramas de Nyquist obtenidos para todos los procesos en el conjunto de ensayo con $\kappa > 0,1$ cuando se utiliza la regla de sintonía (7.11).

La figura muestra que todas las funciones de transferencia de lazo están cercanas al círculo M .

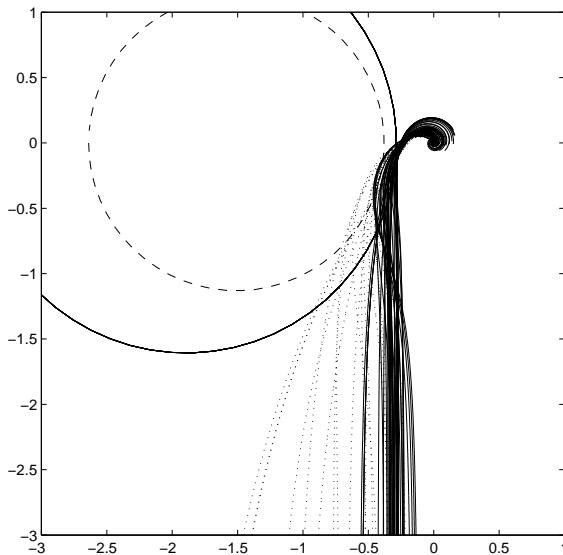


Figura 7.12 Los diagramas de Nyquist de las funciones de transferencia del lazo obtenidos con controladores PI sintonizados de acuerdo con (7.11) se aplican a los procesos en el conjunto de ensayo con $\kappa > 0,1$. Las funciones de transferencia correspondientes a los procesos con $0,1 < \kappa < 0,2$ se muestran con líneas punteadas. El círculo continuo corresponde a $M = 1,4$ y el de líneas discontinuas a un círculo donde M se incrementa en un 15 por ciento.

Reglas de sintonía de control PID para procesos balanceados y con retardo dominante

Los parámetros de los controladores PID para todos los procesos del conjunto de ensayo (7.2) fueron calculados utilizando el diseño MIGO con las restricciones descritas en la sección anterior. El parámetro de diseño se seleccionó como $M = 1,4$.

La Figura 7.13 ilustra las relaciones entre los parámetros del controlador obtenidos con el diseño MIGO y los parámetros de procesos para todos los procesos en el conjunto de ensayo.

En la Figura 7.13, los parámetros del controlador de los procesos P_1 están marcados con círculos y los correspondientes a P_2 están representados con cuadrados. Para $\kappa < 0,3$, la ganancia para P_1 es típicamente más pequeña y el tiempo integral más grande que para los otros procesos. Esto es lo contrario que lo que ocurrió para el control PI. El proceso P_2 tiene una ganancia más grande y un tiempo integral más pequeño que para la mayoría de los procesos.

La figura indica que las variaciones de los parámetros del controlador normalizado son más de un orden de magnitud. Por consiguiente, no es posible encontrar buenas reglas de sintonía universales que no dependan de la razón de ganancia κ . Ziegler y Nichols sugirieron la regla $KK_{180} = 0,6$, $T_i/T_{180} = 0,5$ y $T_d/T_{180} = 0,125$, la cual está representada en la Figura 7.13 por las líneas discontinuas. La regla de Ziegler-Nichols es sólo apropiada para muy pocos procesos en el conjunto de ensayo. La ganancia del controlador es demasiado alta excepto para algunos procesos con valores muy pequeños de κ .

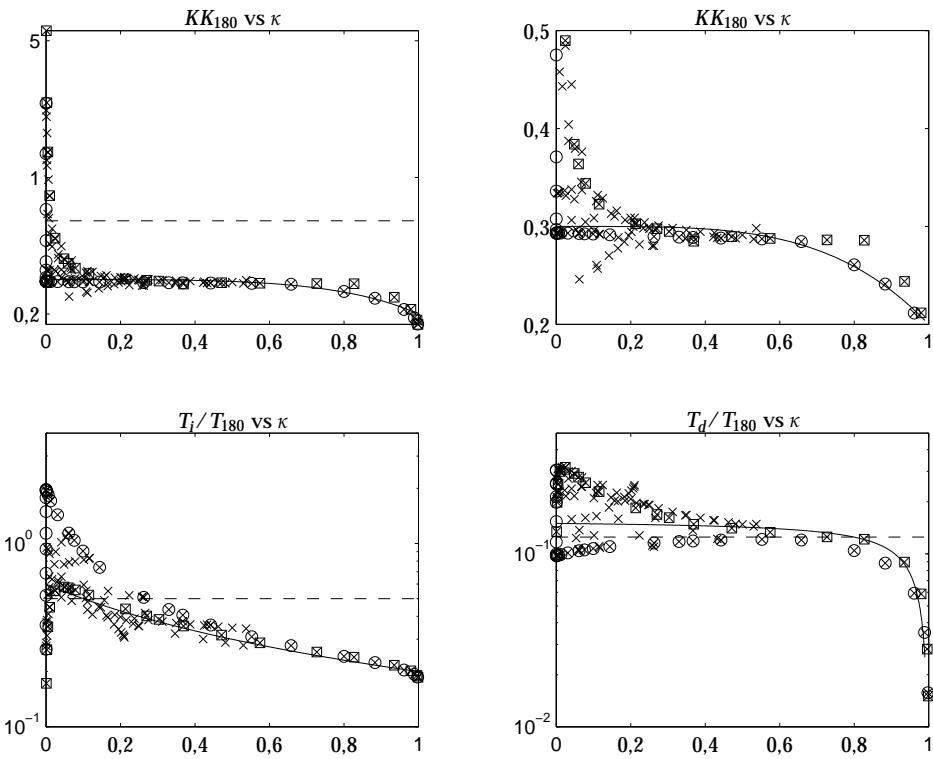


Figura 7.13 Parámetros del controlador PID normalizado como función de la razón de ganancia κ . Los parámetros obtenidos para el proceso P_1 están marcados por círculos y los obtenidos para el proceso P_2 por cuadrados. Las líneas discontinuas indican la regla de sintonía de Ziegler-Nichols y las líneas continuas corresponden a la regla (7.12).

Aunque la Figura 7.13 muestra que no es posible capturar todos los datos por una regla de sintonía, queda claro que se puede encontrar una regla de sintonía para procesos balanceados y de retardo dominante. La Figura 7.13 muestra en líneas continuas las gráficas correspondientes a la siguiente regla de sintonía.

$$\begin{aligned} K &= (0,3 - 0,1\kappa^4)/K_{180} \\ T_i &= \frac{0,6}{1 + 2\kappa} T_{180} \\ T_d &= \frac{0,15(1 - \kappa)}{1 - 0,95\kappa} T_{180}. \end{aligned} \quad (7.12)$$

La regla de sintonía (7.12) no es apropiada para procesos con constante de tiempo dominante, pero aporta parámetros del controlador que están muy cerca del óptimo para procesos con $\kappa > 0,2$. La Figura 7.14 ilustra los diagramas de Nyquist obtenidos para todos los procesos en el conjunto de ensayo con $\kappa > 0,1$ cuando se utiliza la regla de sintonía (7.12). La figura muestra que todas las funciones de transferencia del lazo permanecen muy cerca del círculo M .

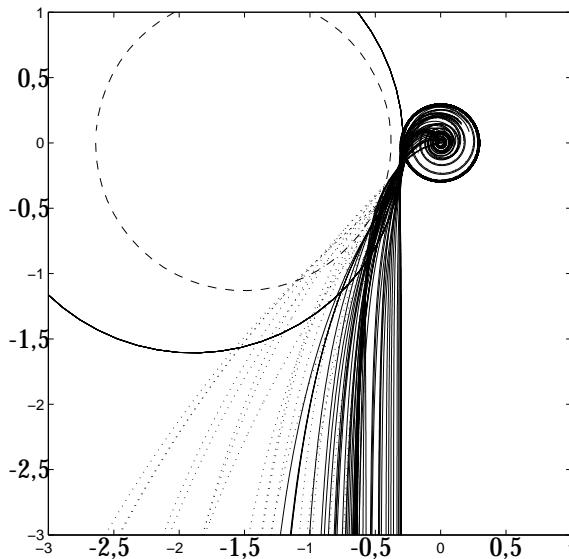


Figura 7.14 Curvas de Nyquist de las funciones de transferencia del lazo obtenidas cuando controladores PID sintonizados de acuerdo con (7.12) son aplicados a procesos del conjunto de ensayo con $\kappa > 0.1$. Las funciones de transferencia correspondientes a procesos con $0.1 < \kappa < 0.2$ se muestran con líneas punteadas. El círculo sólido corresponde a $M = 1.4$ y el punteado a un círculo donde M está incrementado en un 15 por ciento.

7.6 Control PID basado en un modelo de segundo orden

En esta sección se presentan las reglas de sintonía basadas en el modelo SOTD (2.47). El modelo SOTD se puede obtener utilizando el método combinado de la respuesta a un escalón y la respuesta en frecuencia presentadas en la Sección 2.7.

La Figura 7.15 muestra los parámetros del controlador K , $k_i = K/T_i$ y $k_d = KT_d$ para todos los procesos en el conjunto de ensayo, excepto para procesos integrativos, dibujados respecto del tiempo de retardo normalizado $\tau_1 = L_1/T_{63}$. La Figura 7.16 muestra los parámetros del controlador T_i y T_d con diferentes normalizaciones. Obsérvese que las escalas son también diferentes. Comparando los resultados con la Figura 7.3, donde se ha utilizado un modelo FOTD más simple, se observa una mejora significativa, particularmente para valores pequeños del tiempo de retardo normalizado. Esto no es sorprendente ya que el rendimiento alcanzable viene fundamentalmente dado por el tiempo de retardo y la mejora se debe principalmente a estimaciones mejoradas del tiempo de retardo verdadero.

Las figuras muestran que existe una considerable propagación de los valores de los parámetros. En la Figura 7.15, los parámetros KL_1/T_{63} , $k_i T_{63}$ y $k_i L_1$ se extienden sobre dos décadas. El rango de variación es más grande para otras normalizaciones; por ejemplo, el parámetro KK_p varía sobre cinco décadas. Obsérvese también que existe una dispersión en los valores, particularmente en k_d y T_d . Esto significa que no podemos pretender encontrar buenas fórmulas donde los parámetros normalizados sean funciones sólo de τ_1 .

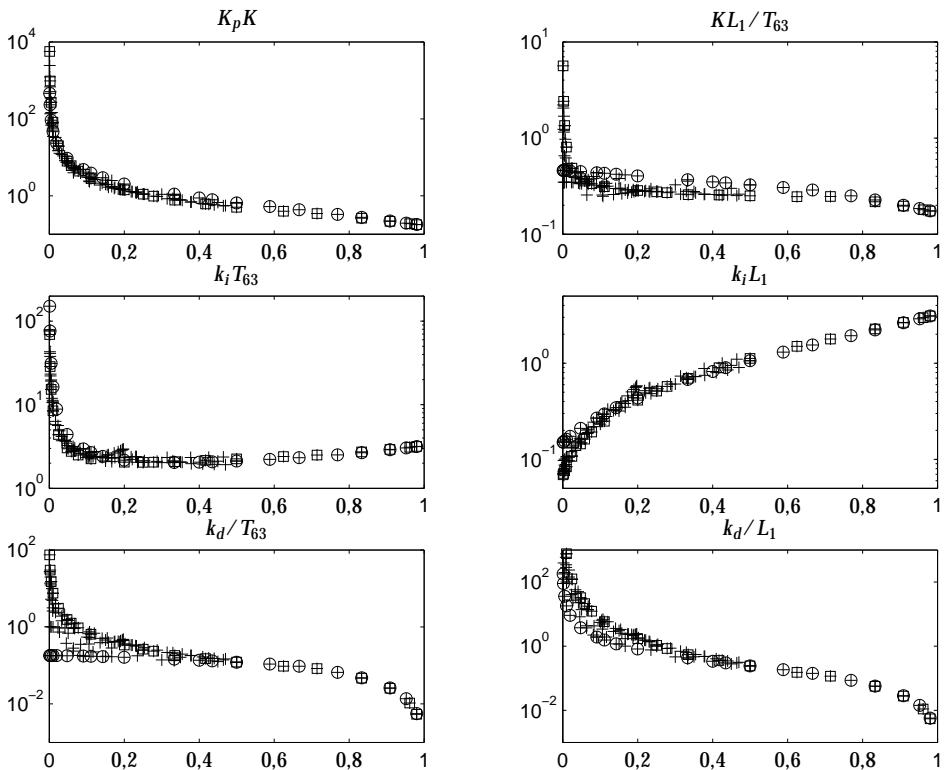


Figura 7.15 Parámetros K , k_i y k_d del controlador normalizado para los procesos en el conjunto de ensayo dibujados respecto de τ_1 . Los datos para procesos P_1 están marcados por círculos y los de P_2 por cuadrados.

Estructura de las fórmulas de sintonía

Para obtener destreza con las parametrizaciones prácticas, veremos algunos casos especiales.

Para procesos de retardo dominante donde $L_1 \gg T_1 > T_2$, el modelo (2.47) puede aproximarse por

$$P(s) = K_p e^{-sL_1}.$$

La acción derivativa no puede utilizarse para este proceso. Diseñando un controlador PI llegamos a

$$C(s) = K + \frac{k_i}{s} = \frac{0,1677}{K_p} + \frac{0,4618}{sL_1 K_p},$$

donde los valores numéricos se dan para un diseño con $M = 1,4$. Despreciando el tiempo de retardo y utilizando los valores numéricos de los parámetros del controlador, encontramos que el sistema en lazo cerrado es de primer orden con el polo $sL_1 = -0,4$.

Si la dinámica del proceso es un tiempo de retardo con una constante de tiem-

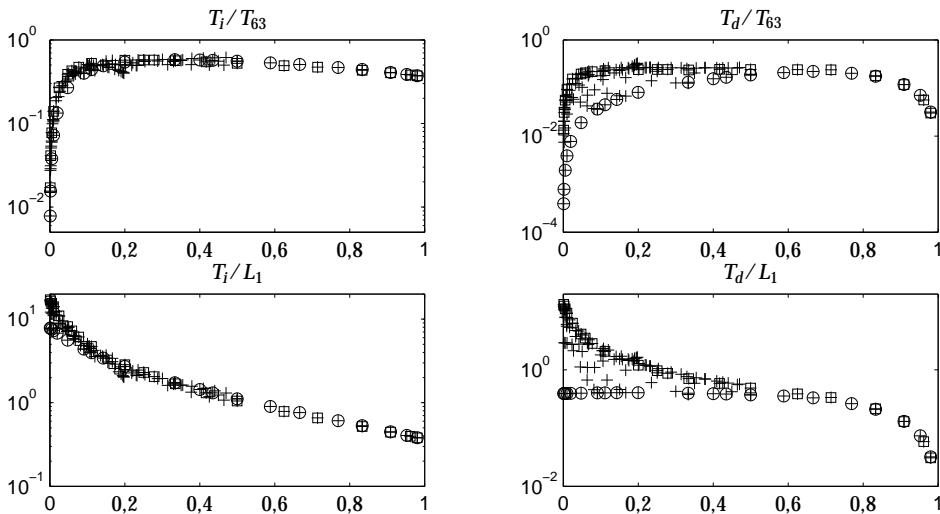


Figura 7.16 Parámetros T_i y T_d del controlador normalizado para los procesos del conjunto de ensayo respecto de τ_1 . Los datos para los procesos P_1 están marcados por círculos y los de P_2 por cuadrados.

po pequeña,

$$P(s) = \frac{K_p}{1+sT} e^{-sL},$$

encontramos que las funciones de transferencia de lazo bajo control PID con ganancia derivativa k_d como

$$G_L(s) \approx \frac{k_d K_p}{T} e^{-sL}.$$

El diagrama de Nyquist de esta función de transferencia es un círculo alrededor del origen con radio $k_d K_p / T$. El criterio de diseño utilizado la sensibilidad combinada requiere que el radio sea menor que $(M-1)/M$. La mayor ganancia derivativa es por lo tanto

$$k_d = \frac{T}{K_p} \frac{M-1}{M}.$$

Para procesos con retardo dominante, el tiempo derivativo es por consiguiente proporcional a la constante de tiempo.

El controlador PID para el proceso

$$P(s) = \frac{K_p}{sT} e^{-sL_1}$$

es

$$C(s) = K + \frac{k_i}{s} + k_d s = \frac{0,4603T}{K_p L_1} + \frac{0,05841T}{s K_p L_1^2} + \frac{0,1796sT}{K_p},$$

donde los valores numéricos se dan para un diseño con $M = 1,4$. Despreciando el tiempo de retardo y utilizando los valores numéricos de los parámetros del

controlador, encontramos que el sistema en lazo cerrado es de segundo orden con los polos $sL_1 = -0,2 \pm 0,11i$; donde el polo dominante es por tanto $\omega_d = 0,2$.

El controlador PID para el proceso

$$P(s) = \frac{K_p}{s^2 T_1 T_2} e^{-sL_1}$$

es

$$C(s) = K + \frac{k_i}{s} + k_d s = \frac{0,02140 T_1 T_2}{K_p L_1^2} + \frac{0,001218 T_1 T_2}{K_p L_1^3 s} + \frac{0,3 T_1 T_2 s}{K_p L_1},$$

donde los valores numéricos se dan para un diseño con $M = 1,4$. Despreciando el tiempo de retardo y utilizando los valores numéricos de los parámetros del controlador, encontramos que el sistema en lazo cerrado es de tercer orden con los polos $sL_1 = -0,23$ y $sL_1 = 0,035 \pm 0,064i$; donde el polo dominante es por tanto $\omega_d = 0,07$.

Parametrización

Considerando los casos especiales tratados anteriormente, es razonable intentar sintonizar fórmulas que tienen la siguiente forma

$$\begin{aligned} K_p K &= \alpha_1 + \alpha_2 \frac{T_1}{L_1} + \alpha_3 \frac{T_2}{L_1} + \alpha_4 \frac{T_1 T_2}{L_1^2} \\ K_p k_i &= \beta_1 \frac{1}{L_1} + \beta_2 \frac{T_1}{L_1^2} + \beta_3 \frac{T_2}{L_1^2} + \beta_4 \frac{T_1 T_2}{L_1^3} \\ K_p k_d &= \left(\gamma_1 L_1 + \gamma_2 T_1 + \gamma_3 T_2 + \gamma_4 \frac{T_1 T_2}{L_1} \right) \frac{T_1 + T_2}{T_1 + T_2 + L_1}. \end{aligned} \quad (7.13)$$

Esto ajustará los controladores para los casos especiales. Los coeficientes de las ganancias proporcional e integral se obtienen simplemente añadiendo los coeficientes para los procesos prototípico. Debido a la estructura de la fórmula, esto dará automáticamente una interpolación entre procesos con retardo puro y doble integrador con retardo. Este procedimiento no funcionará para la acción derivativa. En este caso, simplemente hemos tomado una media ponderada con pesos L_1 y $T_1 + T_2$.

Haciendo un ajuste por mínimos cuadrados de los parámetros en (7.13) utilizando los parámetros del conjunto de ensayo, se obtienen los resultados mostrados en la Tabla 7.1.

Parámetros finales

Parece razonable hacer las siguientes aproximaciones.

$$\begin{aligned} \alpha_1 &= 0,19 & \alpha_2 &= 0,37 & \alpha_3 &= 0,18 & \alpha_4 &= 0,02 \\ \beta_1 &= 0,48 & \beta_2 &= 0,03 & \beta_3 &= -0,0007 & \beta_4 &= 0,0012 \\ \gamma_1 &= 0,29 & \gamma_2 &= 0,16 & \gamma_3 &= 0,20 & \gamma_4 &= 0,28. \end{aligned} \quad (7.14)$$

Los parámetros β_3 y β_4 son demasiado pequeños. Esto significa que la ganancia integral está esencialmente determinada por los parámetros K_p , T_1 y L_1 . Esto

Cuadro 7.1 Parámetros ajustados a la fórmula de sintonía para conjuntos de datos diferentes; P^\dagger denota todos los procesos excepto el proceso integrativo P_6 .

Par	P_1	P_2	P_1, P_2	P^\dagger	e^{-s}	e^{-s}/s	e^{-s}/s^2
α_1	0,1755	0,1815	0,1823	0,1903	0,1677	-	-
α_2	0,4649	-0,0215	0,4607	0,3698	-	0,4603	-
α_3	0	0,6816	0,0930	0,1777	-	-	-
α_4	0	0,0210	0,0211	0,0196	-	-	0,02140
β_1	0,5062	0,4613	0,4800	0,4767	0,4618	-	-
β_2	0,0587	-0,2028	0,0596	0,0310	-	0,05841	-
β_3	0	0,2877	-0,0367	0,0017	-	-	-
β_4	0	0,0013	0,0013	0,0012	-	-	0,001218
γ_1	0,3026	0,2864	0,2971	0,2918	-	-	-
γ_2	0,1805	0,0590	0,1814	0,1654	-	0,1796	-
γ_3	0	0,2464	0,0814	0,2033	-	-	-
γ_4	0	0,3090	0,3096	0,2772	-	-	0,3

explica por qué existe una buena correlación en los datos para la ganancia integral en la Figura 7.15. La correlación para la ganancia proporcional en la Figura 7.15 es buena pero no tanto como para k_i , ya que los parámetros α_3 y α_4 son más grandes. La correlación es pobre para k_d porque los parámetros γ_3 y γ_4 son grandes.

Procesos integrativos

Investigaremos el procesos P_6 con el fin de investigar que las fórmulas también trabajan para procesos integrativos. Un modelo para procesos integrativos puede ser obtenido tomando el límite de

$$P(s) = \frac{K_p}{(1+sT_1)(1+sT_2)} e^{-sL_1}$$

ya que K_p y T_1 tienden a infinito de tal manera que $K_p/T_1 = K_v$. El modelo entonces llega a ser

$$P(s) = \frac{K_v}{s(1+sT_2)} e^{-sL_1}.$$

La fórmula de sintonía (7.13) es por tanto

$$\begin{aligned} K_v K &= \alpha_2 \frac{1}{L_1} + \alpha_4 \frac{T_2}{L_1^2} \\ K_v k_i &= \beta_2 \frac{1}{L_1^2} + \beta_4 \frac{T_2}{L_1^3} \\ K_v k_d &= \gamma_2 + \gamma_4 \frac{T_2}{L_1}. \end{aligned} \tag{7.15}$$

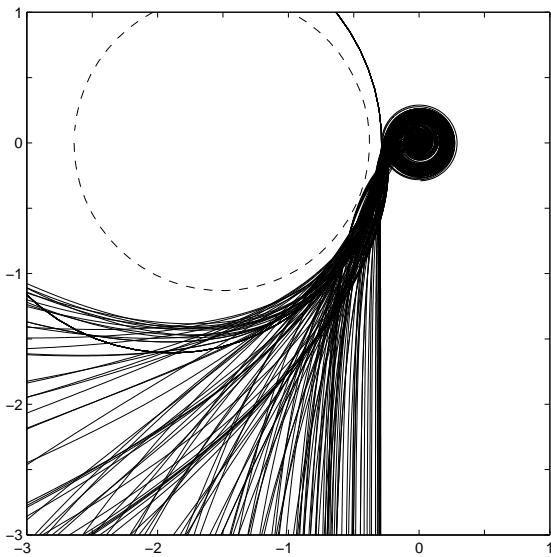


Figura 7.17 Diagramas de Nyquist de las funciones de transferencia del lazo obtenidas cuando controladores PID se sintonizan de acuerdo con (7.13), (7.15) y (7.14) se aplican al conjunto de ensayo (7.2). El círculo continuo corresponde a $M = 1,4$ y el punteado a uno donde M se incrementa en un 15 por ciento.

Validación

La Figura 7.17 muestra los diagramas de Nyquist de las funciones de transferencia del lazo obtenidas cuando los procesos en el conjunto de ensayo (7.2) se controlan con los controladores PID sintonizados con las reglas (7.13), (7.15) y (7.14).

Cuando se usa MIGO, todos los diagramas de Nyquist están fuera del círculo unidad en la figura. Con la regla aproximada, existen algunos procesos donde los diagramas de Nyquist están dentro del círculo. El incremento de M es, sin embargo, menos del 15 por ciento para todos los procesos en el conjunto de ensayo.

7.7 Comparación de los métodos

Esta sección presenta algunos ejemplos para ilustrar el método AMIGO y para compararlo con los diseños MIGO para controladores PI y PID. Se tratan tres ejemplos, un proceso con constante de tiempo dominante, un proceso con retardo dominante y un procesos con retardo y constante de tiempo balanceados.

EJEMPLO 7.1—DINÁMICA DE CONSTANTE DE TIEMPO DOMINANTE
Considérese un proceso con la función de transferencia

$$P(s) = \frac{1}{(1+s)(1+0,1s)(1+0,01s)(1+0,001s)}.$$

Cuadro 7.2 Parámetros del controlador obtenidos con las reglas de sintonía MIGO y AMIGO para el proceso de constante de tiempo dominante del Ejemplo 7.1.

Controlador	Diseño	K	T_i	T_d	b	k_i
PI	MIGO	3,56	0,660		0	5,39
	AMIGO-escalón	4,13	0,539		0	7,66
PID	MIGO	56,9	0,115	0,0605	0	495
	AMIGO-escalón	6,44	0,361	0,0367	0	17,8
	AMIGO-escalón+frecuencia	59,6	0,127	0,0523	0	468

Ajustando el modelo (7.3) al proceso, encontramos que el tiempo de retardo y las constantes de tiempo aparentes son $L = 0,075$ y $T = 1,04$, lo cual lleva a $\tau = 0,067$. La dinámica es por tanto de constante de tiempo dominante. Los datos de la respuesta en frecuencia correspondiente que son necesarios para el diseño AMIGO son $K_{180} = 0,0091$ y $T_{180} = 0,199$. Como la ganancia estática es $K_p = 1$, la razón de ganancia viene dada por $\kappa = K_{180}/K_p = 0,0091$. Debido a que el proceso es de constante de tiempo dominante con $\kappa < 0,1$, no se pueden utilizar las reglas AMIGO basadas en los datos de respuesta en frecuencia para este proceso. Ajustando ahora por un modelo de segundo orden (2.47), se obtienen los parámetros $T_1 = 0,980$, $T_2 = 0,108$ y $L_1 = 0,010$.

En la Tabla 7.2 se presentan los parámetros del controlador con las reglas de sintonía MIGO y AMIGO.

La Figura 7.18 muestra las respuestas del sistema a cambios en el punto de consigna y en las perturbaciones de carga cuando los controladores se sintonizan con los diseños MIGO y AMIGO.

La figura muestra que la regla AMIGO aporta respuestas que están cerca de la regla MIGO para control PI. Sin embargo, ya que este es un proceso de constante de tiempo dominante, la regla de sintonía AMIGO para control PID es conservativa comparada a la regla MIGO. Esto es obvio en la figura.

Las respuestas obtenidas utilizando el modelo SOTD no están presentadas en la figura, sin embargo la Tabla 7.2 muestra que los parámetros del controlador están cerca del diseño MIGO.

Obsérvese que las magnitudes de las señales de control son similares en las perturbaciones de carga, pero existe una mayor diferencia en el tiempo de respuesta. Las diferencias en las respuestas muestran claramente la importancia de reaccionar rápidamente.

El ejemplo muestra que la acción derivativa puede aportar mejoras drásticas en rendimiento para procesos con constante de tiempo dominante. También demuestra que el rendimiento de control puede aumentarse obteniendo mejores modelos del proceso que (7.3). \square

En lo siguiente consideraremos un proceso donde la constante de tiempo y el retardo están balanceados.

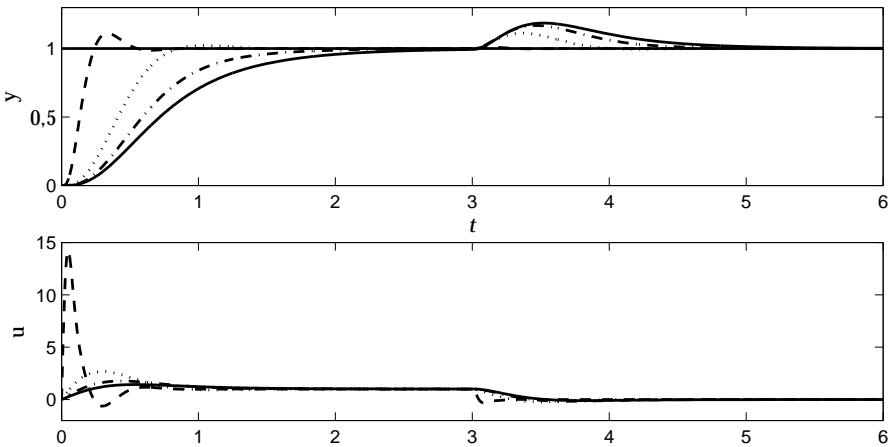


Figura 7.18 Respuestas a un escalón unitario en el instante de tiempo 0 en el punto de consigna y a un escalón unitario en la perturbación de carga en el instante de tiempo 3 para controladores PI diseñados con MIGO (línea continua) y AMIGO (línea de punto y raya), y además para controladores PID diseñados con MIGO (línea a tramos) y AMIGO-escalón (línea punteada) para el proceso de constante de tiempo dominante del Ejemplo 7.1.

EJEMPLO 7.2—CONSTANTE DE TIEMPO Y RETARDO BALANCEADOS

Considérese un proceso con la función de transferencia

$$P(s) = \frac{1}{(s+1)^4}.$$

Ajustando el modelo (7.3) al proceso encontramos que el retardo y las constantes de tiempo aparentes son $L = 1,42$ y $T = 2,90$, y por tanto $L/T = 0,5$ y $\tau = 0,33$. Los datos de respuesta en frecuencia necesarios para el diseño AMIGO son $K_{180} = 0,250$ y $T_{180} = 6,28$. La razón de ganancia llega a ser $\kappa = K_{180}/K_p = 0,25$. Ajustando por un modelo de segundo orden (2.47), se obtienen los parámetros $T_1 = 1,73$, $T_2 = 1,73$ y $L_1 = 1,05$.

En la Tabla 7.3 se presentan los parámetros del controlador obtenidos con las reglas de sintonía MIGO y AMIGO.

La Figura 7.19 muestra las respuestas del sistema a cambios en el punto de consigna y a perturbaciones de carga utilizando los diseños MIGO y AMIGO-step.

La figura muestra que las respuestas a la perturbación de carga obtenidas para MIGO y AMIGO son bastante similares, lo que era de esperar debido a la similitud en los parámetros del controlador. La diferencia en la respuesta a cambios en el punto de consigna entre los diseños MIGO y AMIGO se debe a la diferencia en las ponderaciones b del punto de consigna de los dos diseños.

La ganancia integral k_i es alrededor de tres veces más grande para el control PID que para el control PI. Este resultado está en concordancia con la Figura 7.7. □

Cuadro 7.3 Parámetros del controlador con las reglas de sintonía MIGO y AMIGO para el proceso con constante de tiempo y retardo balanceados en Ejemplo 7.2.

Controlador	Diseño	K	T_i	T_d	b	k_i
PI	MIGO	0,432	2,43		1	0,178
	AMIGO-escalón	0,414	2,66		0	0,156
	AMIGO-frecuencia	0,640	2,96			0,216
PID	MIGO	1,19	2,22	1,21	0	0,536
	AMIGO-escalón	1,12	2,40	0,619	0	0,467
	AMIGO-frecuencia	1,20	2,51	0,927		0,478
	AMIGO-escalón+frecuencia	1,15	2,17	1,32	0	0,506

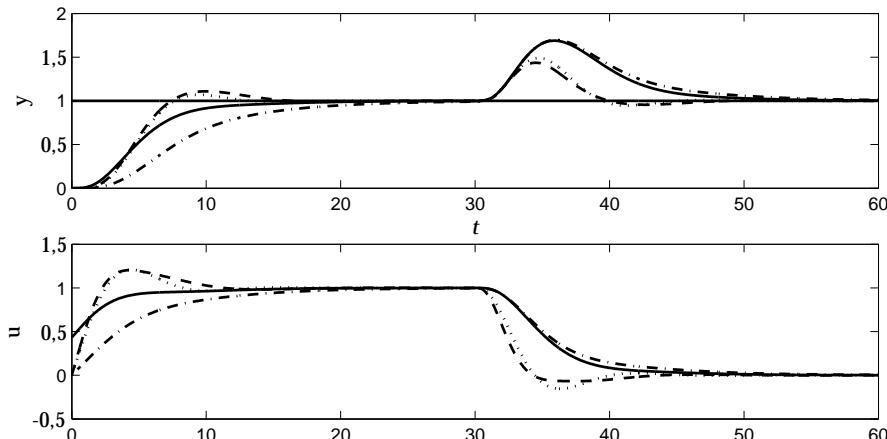


Figura 7.19 Respuesta a un escalón unitario en el instante de tiempo 0 en el punto de consigna y y a un escalón unitario en la perturbación de carga en el instante de tiempo 30 para controladores PI diseñados con MIGO (línea continua) y AMIGO-escalón (línea de punto y raya), y controladores PID diseñados con MIGO (línea a tramos) y AMIGO-escalón (línea punteada) para el proceso con constante de tiempo y retardo balanceados del Ejemplo 7.2.

Finalmente, consideraremos un ejemplo donde la dinámica está dominada por el tiempo de retardo.

EJEMPLO 7.3—DINÁMICA CON TIEMPO DE RETARDO DOMINANTE

Considérese un proceso con la función de transferencia

$$P(s) = \frac{1}{(1 + 0,05s)^2} e^{-s}.$$

Aproximando el proceso con el modelo (7.3) se obtienen los parámetros de proceso $L = 1,01$, $T = 0,0932$ y $\tau = 0,92$. El gran valor de τ muestra que el proceso es de retardo dominante. Los datos de respuesta en frecuencia necesarios para el

Cuadro 7.4 Parámetros del controlador obtenidos con las reglas de sintonía MIGO y AMIGO para el proceso de retardo dominante del Ejemplo 7.3.

Controlador	Diseño	K	T_i	T_d	b	k_i
PI	MIGO	0,170	0,404		1	0,421
	AMIGO-escalón	0,175	0,360		1	0,486
	AMIGO-frecuencia	0,163	0,407			0,400
PID	MIGO	0,216	0,444	0,129	1	0,486
	AMIGO-escalón	0,242	0,474	0,119	1	0,511
	AMIGO-frecuencia	0,212	0,446	0,0957		0,475
	AMIGO-escalón+frecuencia	0,218	0,453	0,129	1	0,481

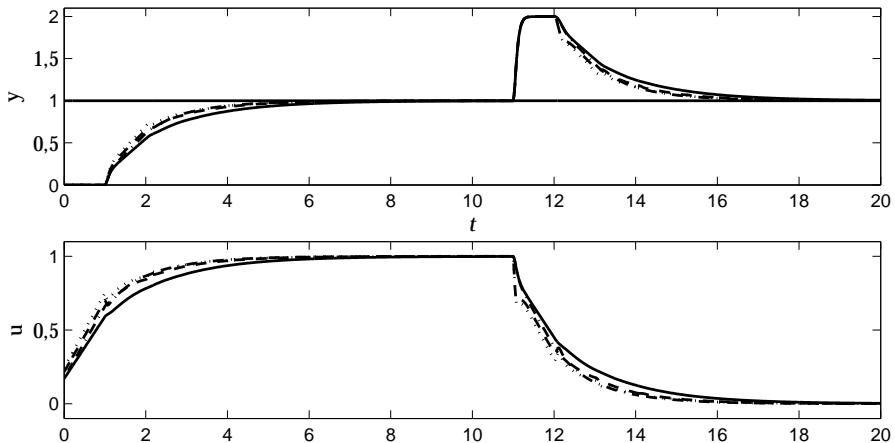


Figura 7.20 Respuestas a un cambio en escalón unitario en el instante de tiempo 0 en el punto de consigna y a un escalón unitario en el instante de tiempo 10 en las perturbaciones de carga para controladores PI diseñados con MIGO (línea continua) y AMIGO-step(línea de punto y raya), y controladores PID diseñados con MIGO (línea a tramos) y AMIGO-escalón (línea punteada) para el proceso de retardo dominante del Ejemplo 7.3.

diseño AMIGO son $K_{180} = 0,980$ y $T_{180} = 2,20$, siendo la razón de ganancia $\kappa = K_{180}/K_p = 0,98$. El proceso tiene la misma estructura que (2.47) donde los parámetros de este modelo vienen dados por $T_1 = T_2 = 0,05$ y $L_1 = 1$.

En la Tabla 7.4 se muestran los parámetros del controlador obtenidos con las reglas de sintonía MIGO y AMIGO.

La Figura 7.20 muestra las respuestas del sistema a cambios en el punto de consigna y perturbaciones de carga. Las respuestas obtenidas para los diseños MIGO y AMIGO son similares. También se observa que las diferencias son mejores entre el control PI y PID, lo que era de esperar ya que el proceso es de retardo dominante. \square

7.8 Ruido de medida y filtrado

Hasta el momento hemos centrado la atención en las perturbaciones de carga en la robustez frente a variaciones en el proceso. En muchos casos, también es necesario considerar el ruido de medida. Este es particularmente el caso para procesos con constante de tiempo dominante donde la maximización de la ganancia integral resulta en controladores con alta ganancia. El ruido de medida puede entonces dar lugar a grandes señales de control. En casos extremos, la señales de control pueden ser demasiado grandes y el actuador puede llegar a saturarse. El efecto del ruido de medida se puede estimar de la función de transferencia que relaciona el ruido de medida con la señal de control:

$$G_{un} = -\frac{C}{1 + PC}. \quad (7.16)$$

Debido a que el ruido de medida tiene típicamente altas frecuencias, las propiedades a altas frecuencias de la función de transferencia son particularmente importantes.

El efecto del ruido de medida se puede paliar filtrando la señal de medida tal y como se muestra en la Figura 4.3. La función de transferencia que relaciona el ruido de medida y la salida del controlador tiene la forma

$$G_{un} = -\frac{CG_f}{1 + PCG_f}. \quad (7.17)$$

Una función de transferencia típica del filtro viene dada por

$$G_f(s) = \frac{1}{1 + sT_f + (sT_f)^2/2}; \quad (7.18)$$

Véase (3.16). Añadiendo un filtro se reducirá la robustez del controlador, pero es sencillo recuperar la robustez rediseñando el controlador para un proceso con la función de transferencia $P(s)G_f(s)$. El procedimiento de diseño comienza diseñando un controlador PID ideal para el proceso $P(s)$. El diseño da una guía para escoger la constante de tiempo del filtro T_f ; típicamente un fracción del tiempo integral para el control PI o del tiempo derivativo para el control PID. Se diseña entonces un controlador PID ideal para el proceso $P(s)G_f(s)$, y el controlador para el proceso $P(s)$ es $C(s)G_f(s)$. Si es necesario, el procedimiento se puede iterar varias veces. El añadir un filtro mejora la atenuación del ruido de medida a costa de empeorar la atenuación de perturbaciones de carga, por lo que la elección de diseño final es un compromiso. Se muestra el procedimiento a continuación mediante un ejemplo.

EJEMPLO 7.4—EFECTO DEL FILTRADO.

Considérese el sistema de constante de tiempo dominante del Ejemplo 7.1. La Tabla 7.2 muestra que el diseño MIGO da un controlador con alta ganancia, $k =$

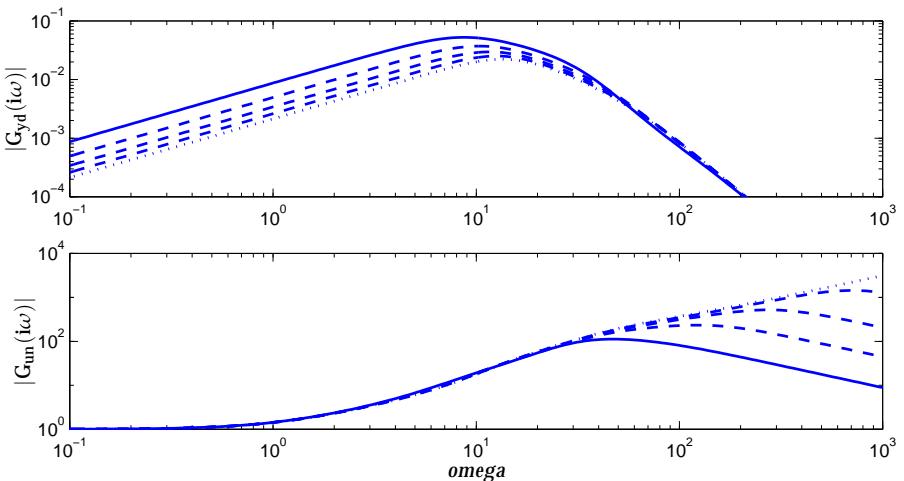


Figura 7.21 Curvas de ganancia de las funciones de transferencia que relaciona las perturbaciones de carga y la salida del proceso (superior), y el ruido de medida y la salida del controlador (inferior). Las curvas para el control PID ideal se muestran con líneas punteadas, para control PID con filtrado y constantes de tiempo $T_f = 0,002, 0,005$ y $0,010$ con líneas discontinuas, y para $T_f = 0,02$ con líneas continuas.

Cuadro 7.5 Parámetros del controlador obtenidos en el Ejemplo 7.4. Compárese con el Ejemplo 7.1.

T_f	K	T_i	T_d	k_i	k_d/T_f	M_{un}
0,000	59,6	0,127	0,0523	468	∞	∞
0,002	52,6	0,138	0,0546	382	1436	1436
0,005	44,7	0,153	0,0578	293	516	520
0,010	35,6	0,176	0,0624	203	222	234
0,020	25,1	0,220	0,0705	115	88,6	112

56,9, lo que resulta en una buena atenuación de las perturbaciones de carga con ganancia integral $k_i = 495$. La función de transferencia que relaciona el ruido de medida con la salida del controlador tiene alta ganancia a altas frecuencias, como se muestra en el diagrama de Bode de la Figura 7.21. El tiempo derivativo es $T_f = 0,06$ y constantes de tiempo razonables del filtro están en el rango $T_f = 0,002 - 0,020$.

Para diseñar controladores para el sistema $P(s)G_f(s)$, hemos aproximado la función de transferencia utilizando la regla de la mitad de Skogestad. Comenzando con el modelo SOTD utilizado en el Ejemplo 7.1, se aborda el filtrado añadiendo $T_f/2$ a la constante de tiempo T_2 y al tiempo de retardo L_1 . La combinación del proceso $P(s)$ y el filtro $G_f(s)$ se representa entonces por el modelo SOTD (2.47) con $T_1 = 0,980$, $T_2 = 0,108 + T_f/2$ y $L_1 = 0,010 + T_f/2$. La ecuación (7.13) entonces da los parámetros del controlador mostrados en la Tabla 7.5. La ganancia

del controlador decrece en un factor de 3 a medida que incrementa la constante del filtro, el tiempo integral aumenta por un factor de 2 y el tiempo derivativo aumenta en un 40 por ciento. La ganancia integral k_i decrece por una factor de 4 y la ganancia máxima a altas frecuencias de G_{un} decrece en varios órdenes de magnitud.

La Tabla 7.5 también muestra la máxima ganancia, M_{un} , de la función de transferencia $G_{un}(s)$ y su estimación k_d/T_f dada por (4.44). Esta simple estimación es excepcionalmente exacta para pequeños valores de las constantes de tiempo del filtro.

En la Figura 7.22 también se ilustran las propiedades de los diferentes controladores, la cual muestra las respuestas del sistema frente a perturbaciones de carga y ruido de medida para diferentes controladores diseñados con diferentes valores de la constante de tiempo del filtro T_f .

Obsérvese que existen grandes variaciones en la señal de control para $T_f = 0,002$ incluso si el ruido en la salida del proceso no es demasiado grande. El motivo de esto es que la ganancia del controlador es demasiado elevada.

Las Figuras 7.21 y 7.22 muestran una buena ilustración del compromiso entre la atenuación de las perturbaciones de carga y la inyección del ruido de medida. El compromiso final es siempre subjetivo, pero una moderada cantidad de filtrado es siempre útil debido a que el efecto del ruido de medida puede disminuirse significativamente con sólo un moderado incremento de la ganancia integral. En este caso particular, un valor de T_f alrededor de 0,01 es una elección razonable.

□

7.9 Desintonía

La regla de sintonía AMIGO lleva de manera natural a la desintonía. Para control PI, el rechazo a perturbaciones de carga se puede caracterizar por la ganancia integral $k_i = K/T_i$. La amplificación del ruido de medida se puede caracterizar por la ganancia del controlador K . Debido a que el ruido de medida tiene típicamente altas frecuencias, la variación de la señal de control generada por el ruido de medida es aproximadamente $Kn(t)$, donde $n(t)$ es el ruido de medida.

La Figura 7.23 muestra el dominio de robustez de un controlador PI para procesos típicos de primer orden. Todas las ganancias en el área blanca satisfacen la condición de robustez de que las sensibilidades combinadas son menores que $M = 1,4$. Alguna combinación de los parámetros del controlador en este rango es por tanto admisible desde un punto de vista de robustez.

La atenuación de perturbaciones se deduce de la ganancia integral k_i . Asumiendo que las perturbaciones de carga entran a la entrada del proceso, la función de transferencia que relaciona las perturbaciones de carga con la salida del proceso está aproximadamente dada por

$$G(s) = \frac{s}{k_i}.$$

La atenuación de la perturbación de carga es por tanto inversamente proporcional

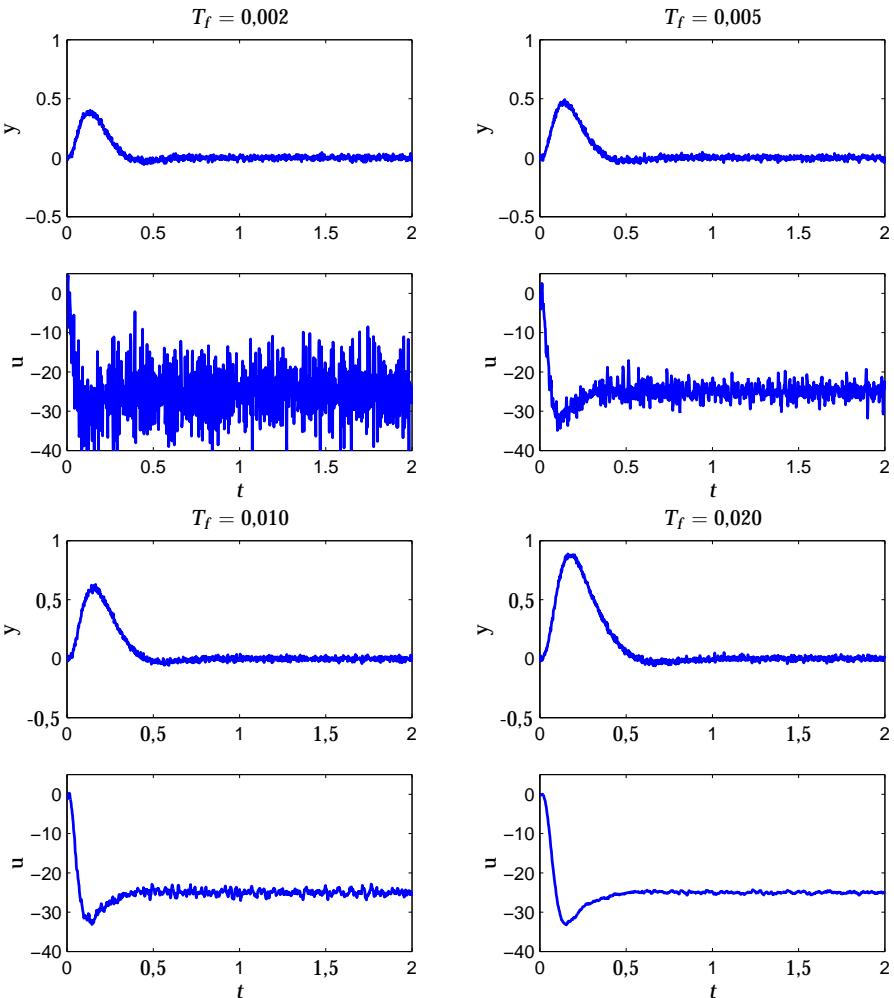


Figura 7.22 Simulación de control PID del sistema del Ejemplo 7.4. Las medidas se filtran con un filtro de segundo orden (7.18) con las constantes de tiempo $T_f = 0.002, 0.005, 0.010$ y 0.020 . Para cada constante del filtro, se escogen los parámetros del controlador para maximizar la ganancia integral sujeta a la restricción de robustez $M = 1.4$. El sistema está sometido a una perturbación de carga de 25 aplicada en el instante de tiempo 0 y a ruido de medida.

a k_i . La medida de ruido posee típicamente altas frecuencias, y para altas frecuencias la función de transferencia que relaciona el ruido de medida con la señal de control está aproximadamente dada por

$$G(s) = K.$$

La inyección del ruido de medida es por consiguiente proporcional a la ganancia del controlador K . Debido a que todos los valores de K y k_i que satisfacen el requerimiento de robustez están dados en la Figura 7.23, es sencillo conseguir un compromiso entre la atenuación de las perturbaciones de carga y la inyección de

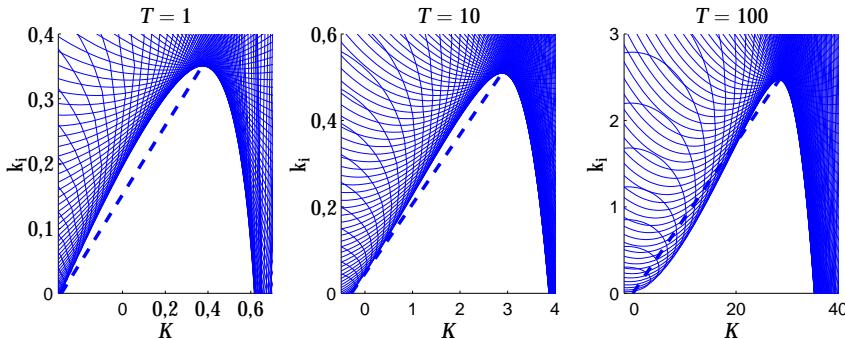


Figura 7.23 La restricción de sensibilidad para un sistema con $M = 1,4$ y la función de transferencia $P(s) = e^{-s}/(1+sT)$, con $T = 1, 10$ y 100 .

ruido de medida.

La Figura 7.23 indica qué variaciones de la variable de control debido al ruido de medida pueden reducirse simplemente disminuyendo la ganancia proporcional. La penalización por esto es una pobre atenuación de las perturbaciones de carga. Un adecuado compromiso cuantitativo se realiza sencillamente basándose en la Figura 7.23. En lugar de escoger los valores máximos de la ganancia integral, deberíamos simplemente escoger una combinación de la ganancia proporcional e integral sobre el borde izquierdo de la región de robustez de la figura. Debido a que la figura no está disponible cuando se utilizan las fórmulas de sintonía simples, en ese caso desarrollaremos una expresión aproximada para el límite izquierdo de la región de robustez.

Una primera tentativa

Una posibilidad es reducir las ganancias tal y como se indica en la línea recta de la Figura 7.23. Esta línea pasa a través del pico con parámetros K^0 y k_i^0 obtenidos por el diseño nominal. Cuando la ganancia integral es cero, el límite de robustez va a través del punto

$$\begin{aligned} KK_p &= -1 + \frac{1}{M_s} = -\frac{M_s - 1}{M_s} = -\alpha \\ k_i &= 0. \end{aligned} \tag{7.19}$$

Una línea a través de este punto y el extremo es

$$k_i = k_i^0 \frac{KK_p + \alpha}{K^0 K_p + \alpha}. \tag{7.20}$$

Obsérvese que no es útil reducir la ganancia proporcional por debajo del valor $K = 0$ cuando el controlador se reduce a un controlador integral puro. La Figura 7.23 muestra que la fórmula (7.20) es conservativa para $T = 1$ y $T = 10$ pero no para $T = 100$ ya que la línea estará parcialmente fuera del límite de robustez para este proceso. Nosotros utilizaremos la fórmula de desintonía (7.20) para procesos

con $\tau > 0,1$. Para procesos de constante de tiempo dominante con $\tau < 0,1$, se necesitan mejores aproximaciones del límite de robustez.

La Figura 7.23 indica que para $T = 100$, el lado más bajo en la parte inferior izquierda del límite de robustez tiene la forma de una parábola. Para obtener una mejor aproximación del límite de la parte inferior izquierda de la región de robustez, consideraremos primero un ejemplo sencillo donde el límite de robustez se puede calcular explícitamente. Un integrador con retardo es un caso extremo de un proceso de constante de tiempo dominante, pero empezaremos por determinar el límite de robustez para incluso el caso más sencillo.

Un integrador puro

Considérese un integrador puro

$$P(s) = \frac{1}{s}.$$

La función de transferencia del lazo con control PI es

$$G_I(s) = \frac{Ks + k_i}{s^2} = \frac{k_i}{s^2} + \frac{K}{s}.$$

Requeriendo que la función de transferencia del lazo esté fuera de un círculo con radio r y centro $-c$ se obtiene

$$\left| c + G_I(i\omega) \right|^2 \geq r^2. \quad (7.21)$$

Sin embargo

$$\begin{aligned} \left| c + G_I(i\omega) \right|^2 &= \left| c - \frac{k_i}{\omega^2} - i\frac{K}{\omega} \right|^2 = \left(c - \frac{k_i}{\omega^2} \right)^2 + \left(\frac{K}{\omega} \right)^2 \\ &= \frac{k_i^2}{\omega^4} + \frac{K^2 - 2ck_i}{\omega^2} + c^2 = \left(\frac{k_i}{\omega^2} + \frac{K^2 - 2ck_i}{2k_i} \right)^2 + c^2 - \left(\frac{K^2 - 2ck_i}{2k_i} \right)^2. \end{aligned}$$

La condición de robustez puede por tanto escribirse como

$$\left| c + G_I(i\omega) \right|^2 = \left(\frac{k_i}{\omega^2} + \frac{K^2 - 2ck_i}{2k_i} \right)^2 + c^2 - \left(\frac{K^2 - 2ck_i}{2k_i} \right)^2 \geq r^2.$$

El lazo izquierdo tiene su valor mínimo para

$$\omega^2 = \frac{2k_i^2}{2ck_i - K^2},$$

donde requerimos que $2ck_i \geq K^2$. La condición de robustez impone por consiguiente la siguiente restricción entre las ganancias integral y proporcional:

$$\left(\frac{2ck_i - K^2}{2k_i} \right)^2 \leq c^2 - r^2.$$

Igualmente se alcanza para

$$\frac{2ck_i - K^2}{2k_i} = \sqrt{c^2 - r^2},$$

o

$$k_i = \frac{K^2}{2(c - \sqrt{c^2 - r^2})} = \frac{K^2(c + \sqrt{c^2 - r^2})}{2r^2}, \quad (7.22)$$

que es una parábola en el plazo K , k_i .

Variables no-normalizadas

Hasta ahora hemos utilizado variables escaladas. Si consideramos un proceso con la función de transferencia

$$P(s) = \frac{K_V}{s} = \frac{K_p}{sT},$$

la ecuación llega a ser

$$k_i = \frac{K_p K^2 (c + \sqrt{c^2 - r^2})}{2Tr^2}.$$

Para un diseño basado en una restricción sobre M_s tenemos $c = 1$ y $r = 1/M_s$. Por tanto,

$$\frac{c + \sqrt{c^2 - r^2}}{2r^2} = \frac{M_s(M_s + \sqrt{M_s^2 - 1})}{2}.$$

Para un diseño con restricciones iguales en ambas, sensibilidad y sensibilidad complementaria tenemos

$$r = \frac{2M - 1}{2M(M - 1)}$$

$$c = \frac{2M^2 - 2M + 1}{2M(M - 1)}.$$

Esto implica $c^2 - r^2 = 1$, y obtenemos

$$\frac{c + \sqrt{c^2 - r^2}}{2r^2} = \frac{c + 1}{2r^2} = M(M - 1).$$

Resumiendo, encontramos que la restricción de robustez para un integrador puro viene dada por

$$k_i = \beta \frac{K_p K^2}{T}, \quad (7.23)$$

donde

$$\beta = \begin{cases} M_s \left(M_s + \sqrt{M_s^2 - 1} \right) / 2 & \text{para diseño basado sobre } M_s \\ M(M - 1) & \text{para diseño basado sobre } M. \end{cases} \quad (7.24)$$

La ecuación 7.23 implica que la ganancia integral se reduce por un factor n^2 cuando la ganancia se reduce por un factor de n . Debido a que $T_i = K/k_i$, encontramos que el tiempo integral aumenta con un factor de n .

La regla de desintonía (7.23) se ha derivado para un integrador sin tiempo de retardo. Para tratar con el proceso (7.3), primero observamos en la Figura 7.23 y en la ecuación 7.19 que la parábola pasa a través del punto $KK_p = -\alpha$ para $k_i = 0$. Para el proceso (7.3), la regla de desintonía (7.23) debería ser por tanto reemplazada por

$$k_i = \beta \frac{(\alpha + KK_p)^2}{K_p(L + T)}, \quad (7.25)$$

donde la constante de tiempo T ha sido sustituida por la constante de tiempo efectiva $T + L$.

Combinando los resultados

Hemos obtenido dos fórmulas de desintonía. La fórmula (7.20) basada en la extrapolación lineal genera buenos resultados para procesos con $\tau > 0.1$, y procesos con $\tau < 0.1$ siempre que la reducción de ganancia sea moderada. La fórmula (7.25) aporta buenos resultados para procesos con constantes de tiempo fuertemente dominantes con gran reducción de ganancia. Por tanto, es natural combinar ambas fórmulas, lo que dará un buen ajuste a la parte izquierda de la restricción de robustez en la Figura 7.23.

Las fórmulas (7.20) y (7.25) dan el mismo resultado para

$$\begin{aligned} k_i^0 \frac{\alpha + KK_p}{\alpha + K^0 K_p} &= \beta \frac{(\alpha + KK_p)^2}{K_p(L + T)} \\ & \text{o} \\ KK_p &= \frac{k_i^0 K_p(L + T)}{\beta(\alpha + K^0 K_p)} - \alpha. \end{aligned} \quad (7.26)$$

Resumiendo, hemos obtenido la siguiente fórmula para desintonía del controlador PI. Primero escogemos una ganancia $K < K_0$, y entonces determinamos la ganancia integral de la siguiente manera. Para procesos con $\tau > 0.1$, determinar k_i de (7.20). Para procesos con $\tau < 0.1$, calcular la ganancia integral de

$$k_i = \begin{cases} k_i^0 \frac{\alpha + KK_p}{\alpha + K^0 K_p} & \text{for } KK_p \geq \frac{k_i^0 K_p(L + T)}{\beta(\alpha + K^0 K_p)} - \alpha \\ \beta \frac{(\alpha + KK_p)^2}{K_p(L + T)} & \text{for } KK_p < \frac{k_i^0 K_p(L + T)}{\beta(\alpha + K^0 K_p)} - \alpha. \end{cases} \quad (7.27)$$

Obsérvese que esta ecuación es una aproximación de la parte izquierda de la restricción de robustez en la Figura 7.23.

Ejemplos

Se mostrará la regla de desintonía (7.27) mediante algunos ejemplos. Primero, tratamos un sólo proceso con la estructura (7.3).

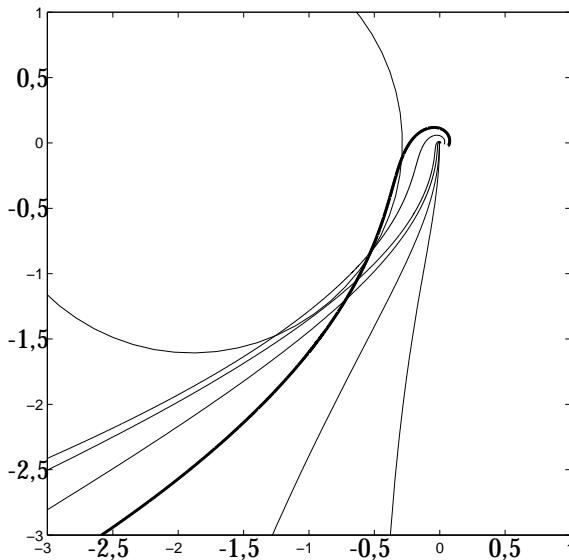


Figura 7.24 Diagramas de Nyquist de las funciones de transferencia del lazo para control PI del proceso (7.28). La línea gruesa corresponde al controlador óptimo y las líneas delgadas a controladores donde la ganancia se reduce por factores 0,5, 0,1, 0,05, 0,01 y 0,005, respectivamente. El círculo muestra la restricción de robustez $M = 1,4$.

EJEMPLO 7.5—DESINTONÍA

Un controlador PI diseñado para el proceso

$$P(s) = \frac{1}{1 + 1000s} e^{-s} \quad (7.28)$$

utilizando el diseño AMIGO (7.5) con la restricción de robustez $M = 1,4$, tiene los parámetros de controlador $K = 349$ y $T_i = 13,2$, que da una ganancia integral $k_i = 26,4$. El proceso es casi un integrador con retardo, $P(s) \approx 0,001e^{-s}/s$, con un tiempo de retardo normalizado $\tau \approx 0,001$, lo que explica la alta ganancia del controlador. La Figura 7.24 muestra los diagramas de Nyquist de la función de transferencia del lazo, así como las curvas obtenidas cuando se reduce la ganancia por los factores 0,5, 0,1, 0,05, 0,01 y 0,005, respectivamente, utilizando la regla de desintonía (7.27). La figura muestra que las funciones de transferencia del lazo de los sistemas desintonizados permanecen cerca de la región de robustez.

La Figura 7.25 muestra cómo la ganancia integral k_i y el tiempo integral T_i cambian cuando se reduce la ganancia. Obsérvese que el tiempo integral permanece casi constante siempre que la reducción de ganancia sea realizada de acuerdo con la parte lineal de (7.27). La relación lineal se reemplaza por la relación cuadrática cuando la ganancia es menor que $K \approx 135$. La reducción de ganancia en ese punto es $K/K^0 \approx 135/349 \approx 0,4$.

□

En el siguiente ejemplo, la regla de desintonía (7.27) se aplica a un gran conjunto de ensayo de procesos.

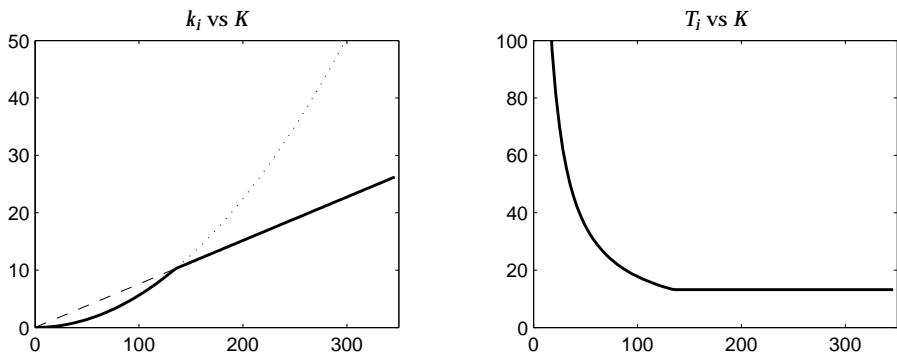


Figura 7.25 Relaciones entre la ganancia reducida K y la ganancia integral k_i (izquierda) y el tiempo integral T_i (derecha). La línea discontinua corresponde a la regla de desintonía (7.20) y la línea punteada a la regla (7.25).

EJEMPLO 7.6—DESINTONÍA APLICADA AL CONJUNTO DE ENSAYO

La regla de desintonía (7.27) ha sido aplicada a todos los procesos del conjunto de ensayo (7.2). La Figura 7.26 muestra los diagramas de Nyquist de las funciones de transferencia del lazo obtenidas cuando los controladores PI son desintonizados utilizando (7.27). La figura muestra cuatro casos; el lazo original y lazos donde la ganancia del controlador se reduce por factores 0,5, 0,1 y 0,01. Sólo se muestran aquellos sistemas donde la ganancia del controlador es mayor que 0,5. Esta es la razón por la que sólo tres casos están a la izquierda cuando la ganancia del controlador se reduce por un factor de 0,01.

El ejemplo muestra que las funciones de transferencia del lazo permanecen cerca de la región de robustez y que la regla de desintonía trabaja bien para todos los proceso del conjunto de ensayo. \square

Una interpretación de asignación de polos

Existen situaciones donde el tiempo de respuesta y el ancho de banda son de gran importancia. En tales casos, el problema de desintonía puede ser resuelto utilizando un simple enfoque de asignación de polos. Despreciando el tiempo de retardo de la función de transferencia del lazo cuando se combina un controlador PI con el modelo del proceso (7.3) lleva a

$$G_I(s) = \frac{K_p K s + K_p k_i}{s(1 + sT)}.$$

El polinomio característico es

$$s^2 + s \frac{1 + K_p K}{T} + \frac{k_i K_p}{T}.$$

Comparando esto con el polinomio estándar $s^2 + 2\zeta\omega s + \omega^2$ encontramos

$$\begin{aligned} 1 + K_p K &= 2\zeta\omega T \\ K_p k_i &= \omega^2 T. \end{aligned} \tag{7.29}$$

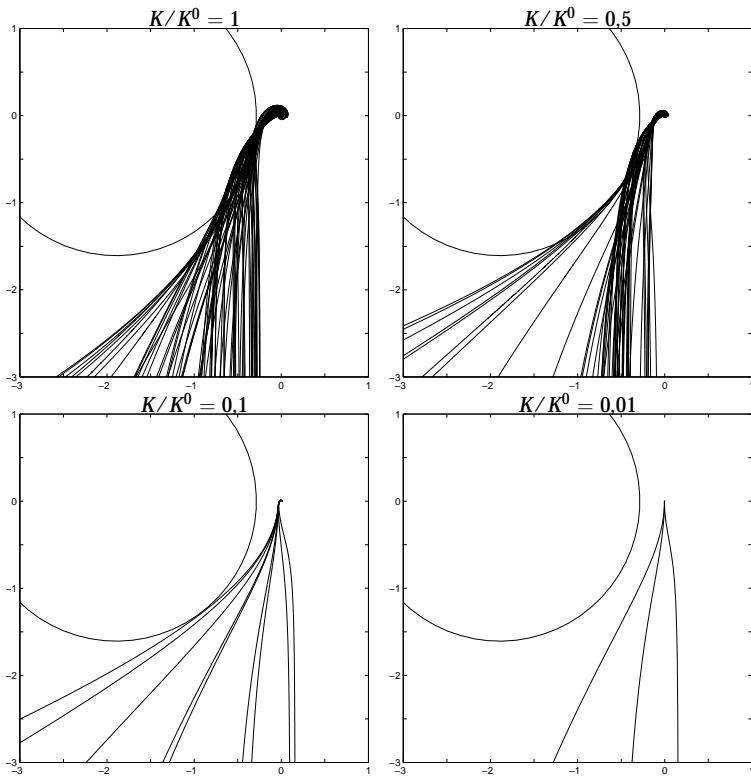


Figura 7.26 Diagramas de Nyquist de las funciones de transferencia del lazo donde el controlador es desintonizado utilizando la regla (7.27). Los círculos muestran el margen de robustez $M = 1,4$.

Cuadro 7.6 Parámetros del controlador, frecuencia y amortiguamiento.

T	K^0	k_i^0	ω	ζ	ω^e	ζ^e
5	1,21	0,296	0,343	0,910	0,222	0,830
10	2,82	0,513	0,226	0,845	0,216	0,804
20	6,24	0,99	0,222	0,815	0,217	0,794
100	34	4,94	0,222	0,788	0,211	0,784

Utilizando los valores numéricos para el proceso (7.3) con $K_p = 1$ y $L = 1$, obtenemos los valores de la Tabla 7.6 para diferentes valores de la constante de tiempo T . Los parámetros del controlador óptimo K^0 y k_i^0 se determinan del diseño MIGO con $M = 1,4$.

Las dos últimas columnas son la frecuencia y el amortiguamiento cuando la constante de tiempo T se reemplaza por la constante de tiempo efectiva $T_e = T + L$. Obsérvese que la frecuencia y el amortiguamiento son particularmente constantes para el rango completo de parámetros.

Otra forma de desintonizar el controlador consiste en utilizar la ecuación 7.29 y reducir la frecuencia natural. Esto da lugar a

$$\omega = \frac{1 + KK_p}{2\zeta T_e}$$

y

$$k_i = \frac{(1 + KK_p)^2}{4\zeta^2 K_p T_e}. \quad (7.30)$$

Esto es similar, pero un tanto diferente, que la expresión parabólica en (7.27).

Control PID

Para el control PID, es natural comenzar una reducción de ganancia a altas frecuencias disminuyendo la ganancia derivativa. Un vía para realizar esto es la siguiente. Sean K^{PID} , k_i^{PID} y k_d^{PID} las ganancias de los controladores PID obtenidos por la fórmula de sintonía del método AMIGO, y K^{PI} y k_i^{PI} las ganancias correspondientes para el controlador PI. Siguiendo las ideas utilizadas en el control PI, obtendremos el controlador desintonizado mediante interpolación lineal. Esto da lugar a

$$\begin{aligned} K &= K^{PI} + \frac{k_d}{k_d^{PID}}(K^{PID} - K^{PI}) \\ k_i &= k_i^{PI} + \frac{k_d}{k_d^{PID}}(k_i^{PID} - k_i^{PI}). \end{aligned} \quad (7.31)$$

Esto resulta en una vía natural para desintonizar el controlador PID hasta llegar a ser un controlador PI. Si se requieren más reducciones de ganancia, podemos proceder como con los controladores PI.

EJEMPLO 7.7—DESINTONÍA APLICADA AL CONJUNTO DE ENSAYO

La regla de desintonía (7.31) ha sido aplicada al conjunto de ensayo (7.2). La Figura 7.27 muestra los diagramas de Nyquist de las funciones de transferencia del lazo obtenidos cuando los controladores PID son desintonizados utilizando (7.31). La figura muestra cuatro casos, el lazo original sintonizado con las reglas de sintonía AMIGO (7.7) y los bucles donde la ganancia derivativa se reduce por factores 0,1, 0,01 y 0. El último caso da un controlador PI puro.

El ejemplo muestra que las funciones de transferencia del lazo permanecen cerca de la región de robustez y que la regla de desintonía trabaja bien para procesos en el conjunto de ensayo. \square

7.10 Resumen

En esta sección se ha tratado de desarrollar reglas simples de sintonía basadas en el espíritu del trabajo realizado por Ziegler y Nichols en 1940. El objetivo ha sido realizar reglas que puedan ser utilizadas mediante sintonía manual y en autosintonizadores para un amplio rango de procesos. Los métodos fueron desarrollados

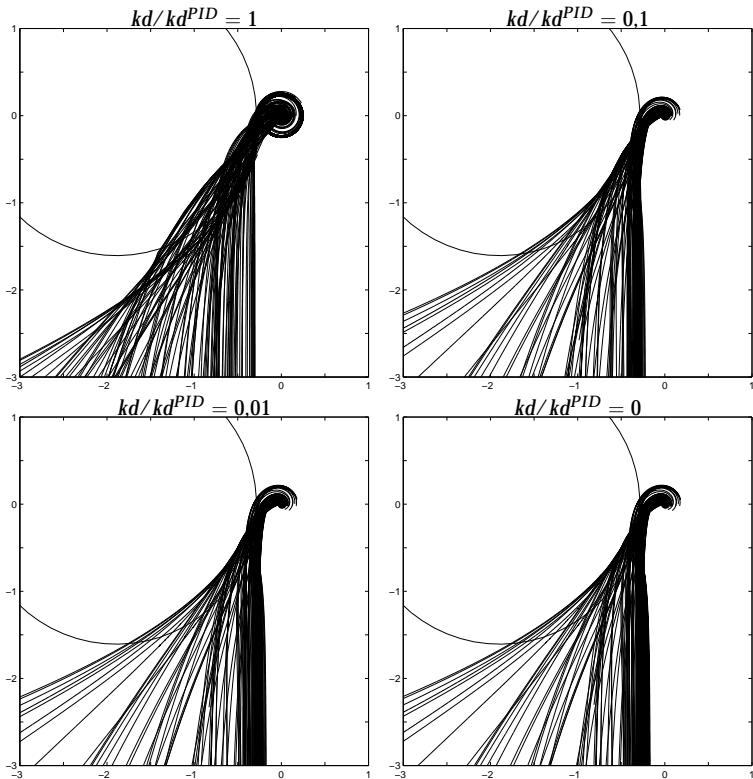


Figura 7.27 Diagramas de Nyquist de las funciones de transferencia del lazo donde el controlador es desintonizado utilizando la regla (7.31). Los círculos muestran la restricción de robustez $M = 1,4$.

aplicando las técnicas de deformación del lazo presentadas en la Sección 6.8 para un gran conjunto de ensayo de procesos representativos. Los parámetros del controlador obtenidos fueron entonces correlacionados con características sencillas de la dinámica del proceso.

Una observación interesante fue que existen significantivas diferencias entre procesos con dinámicas de retardo dominante y con constante de tiempo dominante. Para capturar dicha diferencia, la dinámica del proceso debe caracterizarse para al menos tres parámetros. Obsérvese que Ziegler y Nichols utilizaron sólo dos parámetros. Una posible elección es: ganancia del proceso K_p , constante de tiempo aparente T y tiempo de retardo aparente L . Estos parámetros pueden obtenerse a través de un experimento de respuesta a un escalón. Ver Sección 2.7. El tiempo de retardo relativo $\tau = L/(L + T)$, que varía de 0 a 1, se utiliza para la caracterización rudimentaria de dinámicas. Los procesos con pequeños τ se llaman de constante de tiempo dominante, aquellos con τ cercanos a uno se conocen como de retardo dominante y los procesos con τ alrededor de 0,5 se denominan balanceados.

Se obtuvieron resultados muy satisfactorios para control PI, donde los pará-

metros de la sintonía MIGO pueden ajustarse por

$$K = \frac{0,15}{K_p} + \left(0,35 - \frac{LT}{(L+T)^2} \right) \frac{T}{K_p L} \quad (7.32)$$

$$T_i = 0,35L + \frac{13LT^2}{T^2 + 12LT + 7L^2},$$

para todo el conjunto de ensayo. La regla de sintonía, que hemos llamado AMIGO (Approximate MIGO), aportó buenos resultados para todos los procesos en el conjunto de ensayo pasando desde procesos integrativos a procesos con tiempo de retardo puro.

Los valores numéricos en (7.32) se basan en una sensibilidad combinada $M = 1,4$. La forma de las reglas de sintonía son la misma para los otros valores de M pero siendo los diferentes valores numéricos de los coeficientes.

Para el control PID de procesos con $\tau > 0,3$, fue también posible encontrar la regla de sintonía simple

$$K = \frac{1}{K_p} \left(0,2 + 0,45 \frac{T}{L} \right) \quad (7.33)$$

$$T_i = \frac{0,4L + 0,8T}{L + 0,1T} L$$

$$T_d = \frac{0,5LT}{0,3L + T}.$$

Esta regla de sintonía dió una regla de sintonía conservadora para procesos con constante de tiempo dominante. Puede, por tanto, utilizadarse para el completo rango de procesos probando que una regla de sintonía conservadora es aceptable. La acción derivativa puede dar beneficios apropiados para procesos con constante de tiempo dominante. Una estimación cuantitativa puede obtenerse comparando las ganancias integrales k_i de (7.32) y (7.33).

Para algunos procesos con constante de tiempo dominante es posible proporcionar reglas de sintonía con mucho mejor rendimiento que (7.33). Cuando la dinámica del proceso se caracteriza por los parámetros K_p , L y T , ambos, el tiempo de retardo y las constantes de tiempo pequeñas se capturan en L . Para procesos con constante de tiempo dominante, se puede mejorar el rendimiento si la constante de tiempo y el tiempo de retardo se separan utilizando un mejor modelado. Para un proceso caracterizado por un modelo SOTD con cuatro parámetros, la regla de sintonía AMIGO es

$$K_p K = \alpha_1 + \alpha_2 \frac{T_1}{L_1} + \alpha_3 \frac{T_2}{L_1} + \alpha_4 \frac{T_1 T_2}{L_1^2}$$

$$K_p k_i = \beta_1 \frac{1}{L_1} + \beta_2 \frac{T_1}{L_1^2} + \beta_3 \frac{T_2}{L_1^2} + \beta_4 \frac{T_1 T_2}{L_1^3} \quad (7.34)$$

$$K_p k_d = \left(\gamma_1 L_1 + \gamma_2 T_1 + \gamma_3 T_2 + \gamma_4 \frac{T_1 T_2}{L_1} \right) \frac{T_1 + T_2}{T_1 + T_2 + L_1},$$

donde los parámetros vienen dados por

$$\begin{aligned}\alpha_1 &= 0,19 & \alpha_2 &= 0,37 & \alpha_3 &= 0,18 & \alpha_4 &= 0,02 \\ \beta_1 &= 0,48 & \beta_2 &= 0,03 & \beta_3 &= -0,0007 & \beta_4 &= 0,0012 \\ \gamma_1 &= 0,29 & \gamma_2 &= 0,16 & \gamma_3 &= 0,20 & \gamma_4 &= 0,28.\end{aligned}\quad (7.35)$$

Esta regla de sintonía es similar a (7.33) para procesos con dinámicas balanceadas y retardo dominante, pero sin embargo típicamente genera una mayor ganancia para procesos con constante de tiempo dominante. Esta regla requiere modelos del proceso mejores. Es difícil obtener dos constantes de tiempos de un experimento de respuesta a un escalón. En este caso, se puede hacer uso de la técnica de identificación de sistemas o de los métodos de respuesta a un escalón y frecuencia combinados descritos en la Sección 2.7.

También se han desarrollado reglas de sintonía basadas en la respuesta en frecuencia. En este caso, los parámetros se escogieron como: ganancia del proceso K_p , ganancia última K_{180} y periodo último T_{180} . El parámetro $\kappa = K_{180}/K_p$ se utilizó para clasificar los procesos. Esta elección ajusta lo que se utiliza en los autosintonizadores basados en realimentación con relé. La regla de sintonía AMIGO para controladores PI basada en los datos de la respuesta en frecuencia es

$$\begin{aligned}KK_{180} &= 0,16 \\ \frac{T_i}{T_{180}} &= \frac{1}{1 + 4,5\kappa}\end{aligned}\quad (7.36)$$

y las reglas de sintonía para controladores PID son

$$\begin{aligned}K &= (0,3 - 0,1\kappa^4)/K_{180} \\ T_i &= \frac{0,6}{1 + 2\kappa} T_{180} \\ T_d &= \frac{0,15(1 - \kappa)}{1 - 0,95\kappa} T_{180}.\end{aligned}\quad (7.37)$$

Estas reglas de sintonía dan buenas sintonías para procesos balanceados y de retardo dominante con $\kappa > 0,2$, pero no son apropiadas para procesos con constante de tiempo dominante.

Las reglas de sintonía AMIGO optimizan la atenuación de las perturbaciones de carga con una robustez especificada. El ruido de medida se puede tratar mediante filtrado de la salida del proceso, existiendo ventajas significativas de utilizar un filtro de segundo orden. La dinámica del filtro puede determinarse de una manera sencilla aplicando las reglas AMIGO con $T_f/2$ añadidos a T y L para el modelo FOTD y con $T_f/2$ añadido a T_2 y L para el modelo SOTD. De esta forma es posible conseguir un compromiso entre la atenuación de las perturbaciones de carga y la inyección del ruido de medida.

Se ha desarrollado un método sistemático de desintonía de controladores para determinar una ganancia específica del controlador.

7.11 Notas y referencias

Este capítulo se basa en el trabajo de los autores y sus estudiantes. La motivación fue obtener una mejor comprensión de la información requerida para el desarrollo de buenas reglas de sintonía y encontrar reglas de sintonía que puedan utilizarse para sintonía manual y automática. La base del trabajo es el método de diseño robusto MIGO, el cual se desarrolla en [25] para control PI y en [21, 265] para control PID. En ciertas circunstancias, es ventajoso tener $T_i < 4T_d$ tal y como ha sido indicado en [208]. Reglas de sintonía para este caso concreto se pueden encontrar en [352]. El método MIGO requiere conocimiento de la función de transferencia del proceso. Las reglas de sintonía AMIGO para control PI y PID descritas en [138, 140, 139] se pueden aplicar sólo cuando se conocen las características de las respuestas a un escalón y en frecuencia del proceso. Gran parte del material de este capítulo no ha sido publicado anteriormente.

8

Control predictivo

8.1 Introducción

Un controlador PI sólo considera datos presentes y pasados, y un controlador PID además predice el comportamiento futuro del proceso mediante extrapolación lineal. Han existido muchos intentos para encontrar otras formas de predecir el comportamiento futuro del proceso y tener esta predicción para generar las señales de control. Buenas predicciones pueden mejorar el rendimiento del controlador, particularmente cuando el sistema tiene tiempos de retardo, muy comunes en control de procesos. Los tiempos de retardo pueden surgir de un mecanismo de un retardo puro causado por el transporte o por el tiempo consumido debido a computación o comunicación. Los retardos pueden ser también causados por medidas obtenidas en análisis fuera de línea. Pueden además aparecer cuando un sistema de alto orden o una ecuación en diferencias parciales se aproxima con un modelo de bajo orden, como en la conducción de calor. Los retardos de tiempo aparecen en muchos de los modelos comentados en este libro. Un controlador no comentado hasta ahora que podría tratar los procesos con grandes retardos de tiempo fue propuesto por Smith en 1957. El controlador se conoce ahora como predictor de Smith y se puede ver como un nuevo tipo de controlador pero también se puede interpretar como una ampliación de un controlador PID. Existen muchos otros controladores con capacidades predictivas. El controlador predictivo basado en modelo es una gran clase de controladores que está llegando a ser cada vez más popular.

En este capítulo empezamos presentando el predictor Smith en la Sección 8.2. Este controlador puede dar mejoras significativas en la respuesta a cambios en el punto de consigna, pero puede ser también muy sensible a incertidumbres de modelado. Esto se muestra en la Sección 8.3 donde analizamos el sistema en lazo cerrado utilizando el predictor de Smith. El análisis además muestra que los conceptos de márgenes de ganancia y fase no son suficientes para caracterizar la robustez del sistema. La razón es que el diagrama de Nyquist de la función de transferencia del lazo puede tener grandes bucles a frecuencias más altas que la frecuencia de cruce de ganancia. La robustez se captura bien por las propiedades de la *banda de los cuatro*, y además existe otra medida de robustez clásica, el margen de retardo, que aporta una buena comprensión. En la Sección 8.4 se describe un

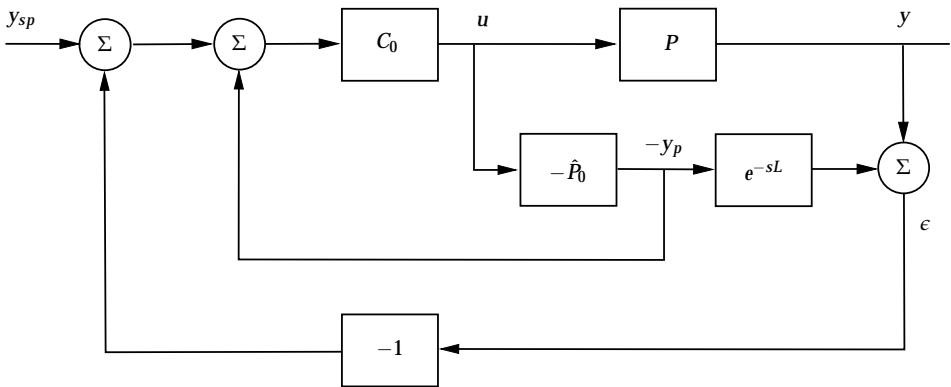


Figura 8.1 Diagrama de bloques de un sistema con predictor de Smith.

tipo especial del predictor de Smith llamado controlador PPI, que es un controlador más simple y más robusto. En la Sección 8.6 está dedicada al Control predictivo basado en modelo, una forma de predicción más general que está ganando en popularidad.

8.2 El predictor de Smith

Para describir la idea del predictor de Smith consideramos un proceso con un retardo de tiempo L y factorizamos la función de transferencia del proceso como

$$P(s) = P_0(s)e^{-sL}, \quad (8.1)$$

donde la función de transferencia P_0 no tiene retardo de tiempo alguno. La Figura 8.1 muestra un diagrama de bloques del sistema en lazo cerrado con un predictor de Smith. El controlador consiste en un controlador PI o PID ordinario C_0 y un modelo del proceso \hat{P} , conectado en paralelo con el proceso. Si el modelo es idéntico al proceso, la señal y_p presenta una salida sin retardo o, equivalentemente, una predicción de lo que sería la salida si no hubiese retardo. Utilizando el modelo, es por lo tanto posible generar una predicción de la salida. La señal y_p se realimenta al controlador, existiendo además una realimentación adicional desde la salida del proceso y para contemplar las perturbaciones de carga. Si el modelo \hat{P} es idéntico al proceso P y no existen perturbaciones actuando sobre el proceso, la señal ϵ es cero. Esto quiere decir que la realimentación exterior al lazo no aporta ninguna contribución, y la relación entrada-salida del sistema viene dada por

$$G_{yy_{sp}} = \frac{PC_0}{1 + P_0C_0} = \frac{P_0C_0}{1 + P_0C_0}e^{-sL}. \quad (8.2)$$

El controlador C_0 puede consiguiente diseñarse como un sistema sin retardo y la respuesta del sistema en lazo cerrado tendrá simplemente un retardo de tiempo adicional.

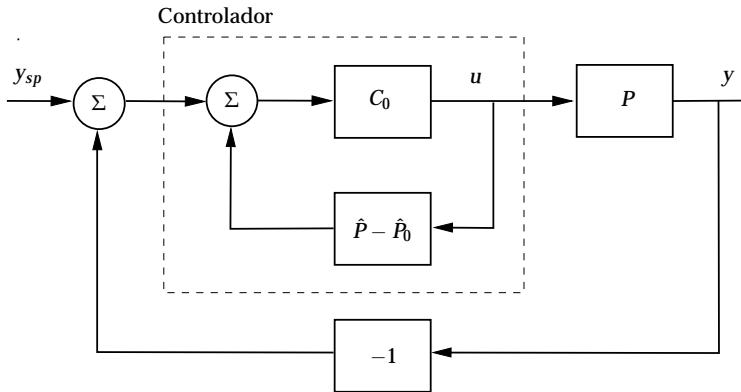


Figura 8.2 Otra representación de un sistema con predictor de Smith.

El sistema mostrado en la Figura 8.1 puede también representarse por el diagrama de bloques de la Figura 8.2, que es un lazo de realimentación convencional con un proceso P y un controlador C , donde el controlador posee la función de transferencia

$$C = \frac{C_0}{1 + C_0(\hat{P}_0 - \hat{P})} = \frac{C_0}{1 + C_0\hat{P}_0(1 - e^{-sL})}. \quad (8.3)$$

La función de transferencia $\hat{P}_0 e^{-sL}$ es la función de transferencia del modelo del proceso utilizado para el diseño del controlador. El controlador C es por tanto obtenido envolviendo una realimentación alrededor del controlador C_0 . La relación entrada-salida del controlador C puede escribirse como

$$U(s) = C_0(s)(E(s) - \hat{P}_0(s)(1 - e^{-sL})U(s)), \quad (8.4)$$

donde $U(s)$ y $E(s)$ son las transformadas de Laplace de la señal de control y del error respectivamente. El término $\hat{P}_0(s)(1 - e^{-sL})U(s)$ puede interpretarse físicamente como el efecto de predicción de las señales de control sobre la salida en el intervalo $(t - L, t)$. El predictor de Smith puede por tanto interpretarse como un controlador PI convencional donde los efectos de las acciones de control pasadas se sustraen del error. El controlador puede ser comparado con un controlador PID, que predice la salida actual del proceso mediante interpolación lineal, tal y como se muestra en la Figura 3.5. Este tipo de predicción es menos efectiva para sistemas con retardos de tiempos porque las salidas futuras del proceso están fuertemente influenciadas por las acciones de control pasadas en lugar de entradas actuales.

Se ilustrarán las propiedades del predictor de Smith mediante un ejemplo.

EJEMPLO 8.1—SISTEMA DE PRIMER ORDEN CON RETARDO

Considérese un proceso con la función de transferencia

$$P(s) = \frac{K_p}{1 + sT} e^{-sL}. \quad (8.5)$$

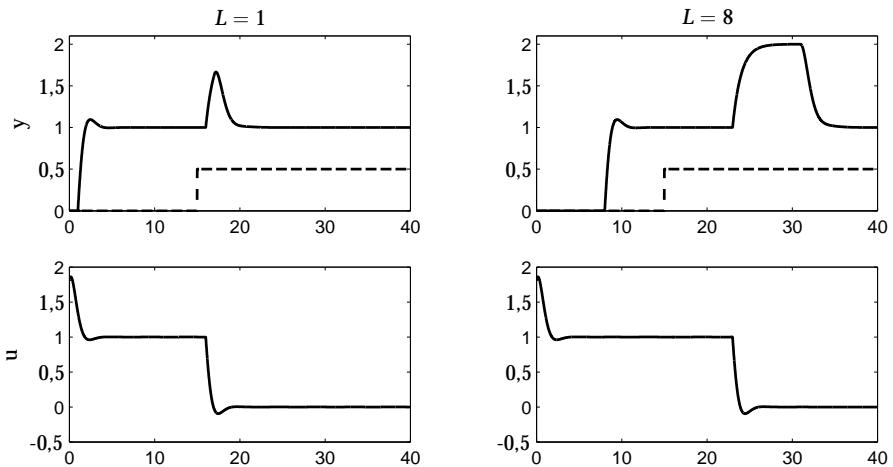


Figura 8.3 Respuestas del sistema en lazo cerrado con predictor de Smith. El proceso tiene la función de transferencia $P(s) = e^{-sL}/(s + 1)$ y la figura muestra la respuesta para $L = 1$ y 8 . La línea discontinua corresponde a la perturbación de carga.

Un controlador PI que da el polinomio característico

$$s^2 + 2\zeta\omega_0 s + \omega_0^2$$

para el proceso sin retardo se diseña tal y como se describe en la Sección 6.4. El controlador es

$$C_0(s) = K \left(1 + \frac{1}{sT_i} \right),$$

donde

$$\begin{aligned} K &= \frac{2\zeta\omega_0 T - 1}{K_p} \\ T_i &= \frac{K_p K}{\omega_0^2 T}. \end{aligned} \tag{8.6}$$

La Figura 8.3 muestra las respuestas del sistema frente a un cambio en escalón unitario en el punto de consigna y en las perturbaciones de carga a la entrada del proceso. La perturbación de carga se aplica en el tiempo $t = 15$ en todos los casos. La constante de tiempo es igual a uno en todos los casos y se modifica el retardo de tiempo L . El controlador PI es diseñado para dar un sistema en lazo cerrado con $\omega_0 = 2$ y $\zeta = 0,7$ sin considerar el retardo de tiempo.

La figura muestra que las respuestas al punto de consigna tienen la misma forma pero con un desplazamiento que cambia con el retardo de tiempo del proceso. La forma es la misma que se obtendría para un sistema sin retardo de tiempo, siendo ésta una propiedad del sistema bastante notable.

Las formas de las respuestas a los cambios en las perturbaciones de carga cambian con el retardo de tiempo L , de tal manera que a medida que éste aumenta el sistema toma más tiempo para reaccionar. La parte inicial de las respuestas es similar pero con diferentes retardos. Debido a las variaciones del retardo, el

tiempo para que el sistema se recupere del efecto de la perturbación de carga varía con dicho retardo. \square

Analizando los resultados, puede parecer sorprendente que es posible obtener buenas respuestas incluso cuando el retardo de tiempo es superior a $L = 8$. A continuación, analizaremos los sistemas obtenidos cuando se utiliza el predictor de Smith para comprender mejor su comportamiento.

El predictor

De (8.3) se puede ver que el predictor de Smith se puede interpretar como una conexión en cascada de un controlador convencional C_0 y un bloque con la función de transferencia

$$C_{pred} = \frac{1}{1 + C_0(\hat{P}_0 - \hat{P})} = \frac{1}{1 + C_0\hat{P}_0(1 - e^{-sL})}. \quad (8.7)$$

Para obtener las respuestas mostradas en la Figura 8.3, la función de transferencia C_{pred} compensa el retardo de tiempo del proceso. Intuitivamente, esto se puede entender de la siguiente manera. Si se asume que $C_0\hat{P}_0 \approx -1$; entonces de (8.7) se tiene que

$$C_{pred} \approx e^{sL}.$$

Esto quiere decir que la función de transferencia $C_{pred}(s)$ actúa como un predictor ideal. Por consiguiente podemos esperar que la función de transferencia $C_{pred}(s)$ se comporte como un predictor ideal para aquellas frecuencias donde $C_0(i\omega)\hat{P}_0(i\omega)$ es cercano a -1 . Obsérvese que no es posible tener $C_0(i\omega)\hat{P}_0(i\omega) = -1$ para cualquier frecuencia debido a que la función de transferencia (8.2) sería entonces inestable. Las propiedades de la función de transferencia (8.7) se mostrarán mediante un ejemplo.

EJEMPLO 8.2—PREDICTOR PARA SISTEMA DE PRIMER ORDEN CON RETARDO DE TIEMPO

Considérese el mismo sistema que en el Ejemplo 8.1. Asumiendo que no existen errores de modelado se deduce que $\hat{P} = P = P_0 e^{-sL}$, que combinado con un controlador PI da lugar al predictor

$$C_{pred} = \frac{1}{1 + C_0(P_0 - P)} = \frac{1}{1 + \frac{K_p K(1 + sT_i)}{sT_i(1 + sT)}(1 - e^{-sL})}. \quad (8.8)$$

De (8.8) se deduce que $C_{pred}(i\omega) = 1$ para $\omega L = 2\pi, 4\pi, 6\pi, \dots$ y que $C_{pred}(s)$ tiende a 1 para grandes valores de s . La función de transferencia C_{pred} posee el desarrollo en serie

$$C_{pred}(s) = \frac{T_i}{T_i + K_p K L} \left(1 + \frac{K_p K L}{T_i + K_p K L} \left(T + \frac{L}{2} - T_i \right) s + \dots \right).$$

La ganancia estática de C_{pred} decrece a medida que aumenta L y es siempre menor que uno. La Figura 8.4 muestra el diagrama de Bode de la función de transferencia para $L = 8$. La figura muestra que la función de transferencia da un avance

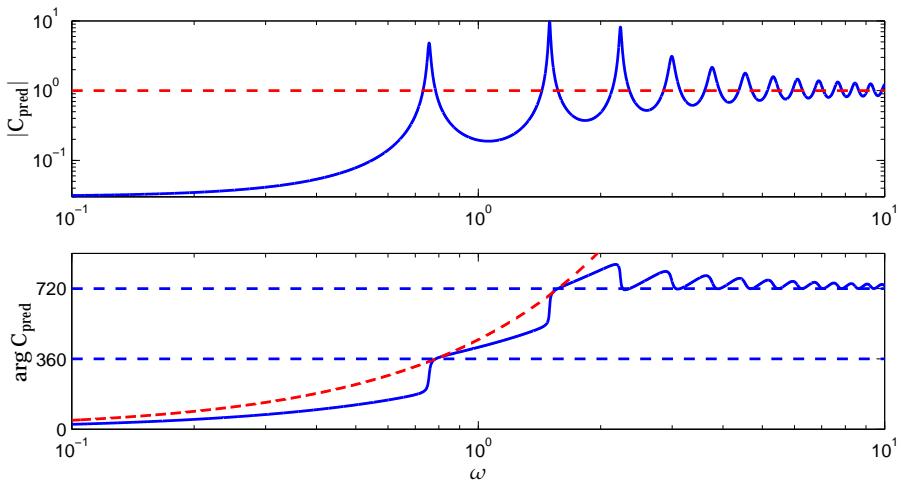


Figura 8.4 Diagrama de Bode de las funciones de transferencia del lazo $C_{\text{pred}}(s)$ dado por (8.8) para $L = 8$ (línea continua) y para el predictor ideal e^{sL} (línea discontinua)

de fase muy grande, más de 800° . La comparación con la curva de fase de un predictor ideal muestra que el sistema se approxima bien al predictor ideal para ciertas frecuencias. Las curvas continuas y discontinuas están muy próximas para aquellas frecuencias donde la curva de ganancia posee picos. Obsérvese, sin embargo, que las curvas de ganancia son distintas. El predictor ideal tiene ganancia constante, pero la ganancia de la función de transferencia C_{pred} cambia en varios órdenes de magnitud.

Ahora investigaremos cómo se genera el amplio avance de fase. La Figura 8.5 muestra los diagramas de Nyquist de la función de transferencia C_{pred} para $K_p = 1$, $T = 1$, $K = 1,8$, $T_i = 0,45$, y $L = 1, 2,5, 4$ y 8 . Para $L = 1$, el avance de fase más grande está próximo a 90° . El avance de fase aumenta a medida que aumenta L , tal como se indica en la curva para $L = 2,5$ donde aumenta la parte circular de la curva de Nyquist. El diagrama de Nyquist tiende a infinito para $L = 2,99$, lo que indica que la función de transferencia tiene polos sobre el eje imaginario. Para L más grandes, el diagrama de Nyquist rodea al origen, lo que quiere decir que el avance de fase es más de 360° . La curva para $L = 4$ muestra que el avance de fase más grande es más de 450° . Cuando se incrementa L más allá, el diagrama de Nyquist va otra vez a infinito para $L = 6,40$, y para un L mayor existen dos rodeos al origen, indicando que el avance de fase es más de 720° . La curva para $L = 8$ muestra que el mayor avance de fase es más de 800° .

Para deformar la curva con $L = 2,5$ continuamente hasta la curva con $L = 4$ en la Figura 8.5, la curva debe ir a infinito para algún valor intermedio de L . En este caso particular, el diagrama de Nyquist de C_{pred} va a infinito para $L = 2,99, 6,40, 9,80, 13,40, 17,00, 20,6, \dots$. Esto quiere decir que la función de transferencia C_{pred} es inestable para algunos valores de L . Tiene dos polos en el semiplano derecho para $2,99 < L < 6,40$, cuatro polos en el semiplano derecho para $6,40 < L < 9,80$, etc. Para la simulación con $L = 10$ en la Figura 8.3, la función de transferencia del predictor tiene seis polos en el semiplano derecho. Por lo tanto, el predictor (8.3)

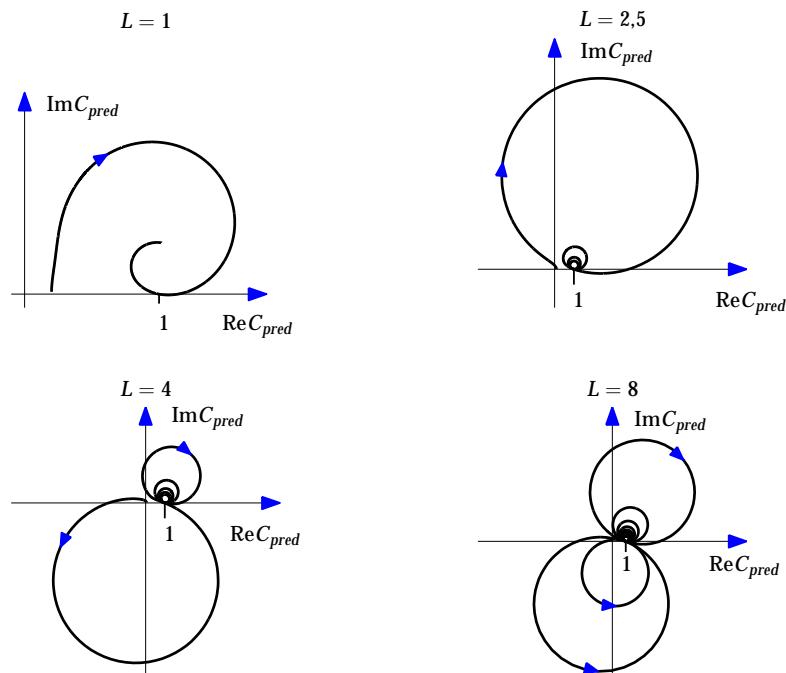


Figura 8.5 Diagramas de Nyquist de la función de transferencia C_{pred} para el sistema del Ejemplo 8.2 con $L = 1, 2,5, 4$, y 8 . Las curvas para $L = 1$ y $2,5$ poseen los giros en el sentido de las agujas del reloj, la curva para $L = 4$ primero presenta un giro en el sentido contrario a las agujas del reloj antes de presentar el resto de giros en el sentido de las agujas del reloj, la curva para $L = 8$ primero presenta dos vueltas en el sentido contrario a las agujas del reloj y el resto de vueltas en sentido contrario.

consigue un gran avance de fase mediante los polos en el semiplano derecho. □

Con controladores inestables existen graves inconvenientes. De la integral de Bode (4.28) se deduce que los polos en el semiplano derecho incrementan la sensibilidad. Es decir, la respuesta notable frente a cambios en el punto de consigna de la Figura 8.3 tiene un precio. En la siguiente sección se analizarán algunas de estas cuestiones.

8.3 Análisis del control del predictor de Smith

Ahora se investigará el sistema en lazo cerrado obtenido cuando un proceso se controla haciendo uso del predictor de Smith. Sea P la función de transferencia del proceso y (8.3) la función de transferencia del predictor de Smith, encon-

tramos

$$\begin{aligned}
 G_{yy_{sp}} &= \frac{PC}{1+PC} = \frac{PC_0}{1+\hat{P}_0C_0+(P-\hat{P})C_0} = \frac{P_0C_0}{1+P_0C_0}e^{-sL} \\
 G_{yd} &= \frac{P}{1+PC} = \frac{P(1+(\hat{P}_0-\hat{P})C_0)}{1+\hat{P}_0C_0+(P-\hat{P})C_0} = P\left(1-\frac{P_0C_0}{1+P_0C_0}e^{-sL}\right) \\
 -G_{un} &= \frac{C}{1+PC} = \frac{C_0}{1+\hat{P}_0C_0+(P-\hat{P})C_0} = \frac{C_0}{1+P_0C_0} \\
 -G_{yn} &= \frac{1}{1+PC} = \frac{1+(\hat{P}_0-P)C_0}{1+\hat{P}_0C_0+(P-\hat{P})C_0} = 1-\frac{P_0C_0}{1+P_0C_0}e^{-sL},
 \end{aligned} \tag{8.9}$$

donde la última igualdad se obtiene asumiendo que el modelo es perfecto, por ejemplo, $\hat{P} = P$. La forma de la función de transferencia desde la salida del proceso $G_{yy_{sp}}$ muestra que aparte del retardo de tiempo, las respuestas frente al punto de consigna son las mismas que para el sistema sin retardo de tiempo. La función de transferencia G_{un} desde el ruido de medida a la señal de control es la misma que la función de transferencia desde el punto de consigna a la salida del controlador. Esta función de transferencia es la misma que para un sistema sin retardo.

Estabilidad

De (8.9) se deduce que el sistema en lazo cerrado tiene los polos del proceso en lazo abierto y los ceros de

$$1 + \hat{P}_0C_0 + (P - \hat{P})C_0 \approx 1 + \hat{P}_0C_0,$$

donde la aproximación es válida cuando $\hat{P} \approx P$. Los ceros de esta función pueden escogerse para ser estables mediante un controlador apropiado C_0 . Para tener un sistema en lazo cerrado estable, se debe requerir también que el proceso sea estable. Esto quiere decir que el predictor de Smith no funciona para procesos inestables en lazo abierto. En la Sección 8.5 se mostrarán algunas modificaciones para eliminar esta dificultad.

Respuesta a perturbaciones de carga

Cuando se desprecian los errores de modelado, la respuesta a perturbaciones de carga a la entrada del proceso viene dada por la función de transferencia

$$G_{yd} = P\left(1 - \frac{P_0C_0}{1+P_0C_0}e^{-sL}\right);$$

ver (8.9). El segundo término tiene un retardo de tiempo L . Si una perturbación ocurre en el instante de tiempo 0 se deduce que la respuesta en el intervalo $0 \leq t < L$ es la misma que la respuesta del sistema en lazo abierto. En la Figura 8.3 se da una ilustración.

Supóngase que el proceso P es estable con ganancia estática K_p y que el controlador C_0 posee acción integral con ganancia integral k_i . Un desarrollo en serie

de G_{yd} para pequeños valores de s resulta en

$$G_{yd}(s) \approx K_p \left(1 - \frac{K_p k_i}{s + K_p k_i} (1 - Ls) \right) = K_p \frac{s + K_p k_i L s}{s + K_p k_i} \approx \left(K_p L + \frac{1}{k_i} \right) s. \quad (8.10)$$

Como $G_{yd}(0) = 0$, no existe error en régimen permanente para cambios en escalón en las perturbaciones de carga. Además, la integral del error para la perturbación de carga en la forma de un escalón unitario es

$$IE = K_p L + \frac{1}{k_i}. \quad (8.11)$$

Obsérvese que el primer término $K_p L$ depende sólo del proceso y que el segundo término $1/k_i$ depende solamente del controlador.

La función de transferencia P tiene un polo en el origen para procesos que tiene acción integral. Para tales procesos y un controlador con acción integral, tenemos $P(s) \approx K_v/s$ y $C(s) \approx k_i/s$ para pequeños valores de s . Esto implica que

$$G_{yd}(s) \approx \frac{K_v}{s} \left(1 - \frac{K_v k_i}{s^2 + K_v k_i} (1 - Ls) \right) = \frac{K_v s^2 + K_v k_i L s}{s^2 + K_v k_i} \approx K_v L. \quad (8.12)$$

Esto quiere decir que existirá error en régimen permanente para procesos con integración incluso si el controlador posee acción integral. Por consiguiente, la recuperación frente a perturbaciones de carga será muy lenta para procesos con dinámicas lentas. Obsérvese que el sistema en lazo cerrado es estable incluso aunque P contenga un integrador. La razón es que el integrador de P se cancela con un cero de la función de transferencia $1 - PC_0/(1 + P_0 C_0)$. Tal y como se comentará en la Sección 8.5, se han propuesto varias modificaciones del predictor de Smith para procesos con integración.

Las funciones de sensibilidad

En el caso ideal $\hat{P} = P$, de (8.9) se deduce que las funciones de sensibilidad y sensibilidad complementaria son

$$\begin{aligned} S &= 1 - \frac{PC_0}{1 + P_0 C_0} = 1 - \frac{P_0 C_0}{1 + P_0 C_0} e^{-sL} = 1 - T_0 e^{-sL} \\ T &= \frac{PC_0}{1 + P_0 C_0} = \frac{P_0 C_0}{1 + P_0 C_0} e^{-sL} = T_0 e^{-sL}, \end{aligned} \quad (8.13)$$

donde T_0 es la función de sensibilidad complementaria para el sistema nominal sin retardo. Obsérvese que las curvas de ganancia de T y T_0 son idénticas. La curva de ganancia de la función de sensibilidad complementaria es independiente de L .

EJEMPLO 8.3—SISTEMA DE PRIMERO ORDEN CON RETARDO

Para el sistema de primer orden del ejemplo 8.1 donde el controlador C_0 se diseñó para $\omega_0 = 2$, las funciones de sensibilidad son

$$\begin{aligned} T(s) &= \frac{K_p K (1 + sT_i)}{sT_i(1 + sT) + K_p K (1 + sT_i)} e^{-sL} = \frac{s\omega_0^2 T / (2\zeta\omega_0 T - 1) + \omega_0^2}{s^2 + 2\zeta\omega_0 s + \omega_0^2} e^{-sL} \\ S(s) &= 1 - T(s). \end{aligned}$$

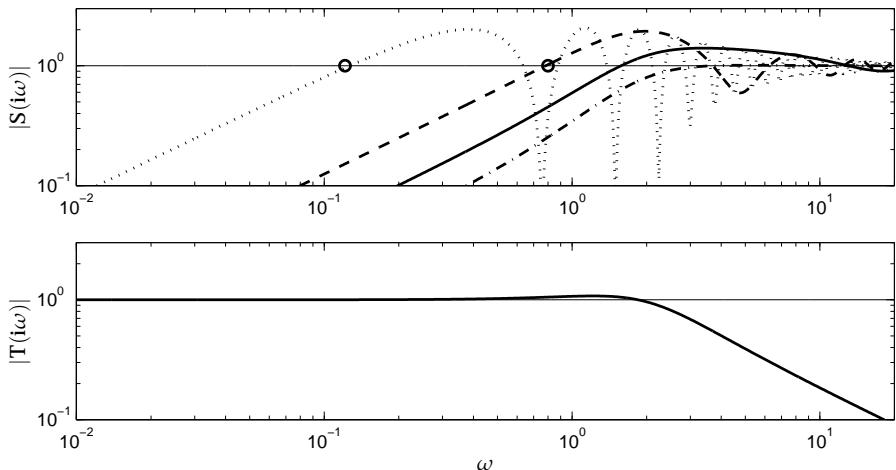


Figura 8.6 Curvas de ganancia para las funciones de sensibilidad del sistema en el Ejemplo 8.3 con $L = 0$ (línea de punto y raya), 0,25 (línea continua), 1 (línea discontinua), and 8 (línea punteada).

La Figura 8.6 muestra que las curvas de ganancia de las funciones de sensibilidad para $L = 0, 0,25, 1$ y 8 , corresponden a $\omega_0 L = 0, 0,5, 2$ y 16 . La máxima sensibilidad aumenta rápidamente con L , donde tenemos $M_s = 1,1, 1,4, 1,6$ y 2 para $L = 0, 0,25, 0,4$ y $1,2$, respectivamente. Para $L > 1,2$, la sensibilidad máxima permanece próxima a $M_s = 2$. Obsérvese además que la sensibilidad para bajas frecuencias aumenta rápidamente cuando L aumenta.

Las diferencias entre las propiedades de baja frecuencia de las funciones de sensibilidad en la Figura 8.6 son sencillas de explicar a partir de (8.10). La asíntota a bajas frecuencias de las curvas de ganancia de la función de sensibilidad intersecta con la línea de magnitud unitaria para $\omega = k_i/(1 + k_i K_p L)$. Para el sistema de la figura, tenemos $K_p = 1$ y $k_i = 4$, y las intersecciones se denotan por círculos. \square

Las funciones de sensibilidad en la Figura 8.6 son típicas para sistemas con predictores de Smith. La función de sensibilidad complementaria está cercana a uno para las frecuencias hasta el ancho de banda ω_b del sistema nominal sin retardo. La función de sensibilidad tiene el comportamiento típico oscilatorio mostrado en la figura e intersecta con la línea $|S| = 1$ varias veces. Para grandes retardos, la frecuencia de cruce de sensibilidad es aproximadamente $\omega_{sc} = k_i/(1 + k_i K_p L)$, reflejando el hecho de que la atenuación de las perturbaciones de carga es pobre para grandes L . Obsérvese también que los picos más pronunciados de la función de sensibilidad están próximos a $M_s = 2$ en la frecuencia donde $|T(i\omega)| \approx 1$.

Robustez

Para controladores con acción integral tenemos que $T(0) = 1$. Sea ω_b una frecuencia tal que $|T(i\omega)|$ es próximo a 1 para $0 \leq \omega \leq \omega_b$. Si $\omega_b L \geq \pi$, entonces de (8.13) se deduce que la sensibilidad máxima es alrededor de $M_s = 2$. Para tener sensibilidades más pequeñas es por lo tanto necesario requerir que $\omega_b L$ no sea

tan grande. De (4.32) se deduce que es posible tener perturbaciones en el proceso tal que

$$\frac{|\Delta P(i\omega)|}{|P(i\omega)|} < \frac{1}{|T(i\omega)|}$$

sin hacer al sistema inestable. Para frecuencias menores que ω_b , la parte derecha es igual a uno. Esta desigualdad implica que la región de incertidumbre es un círculo con centro en $P(i\omega)$ que pasa a través del origen. Si consideramos sólo variaciones en la fase, las variaciones admisibles son por lo tanto 60° o $\pi/3$ rad. Como el cambio de fase es $\omega_b L$, encontramos que

$$|\omega_b \Delta L| < \frac{\pi}{3},$$

lo que da la siguiente estimación de las variaciones permisibles en el retardo de tiempo

$$\frac{|\Delta L|}{L} < \frac{\pi}{3\omega_b L} \approx \frac{1}{\omega_b L}. \quad (8.14)$$

Por lo tanto, para controladores con grandes valores de $\omega_b L$ se requiere un conocimiento exacto del retardo de tiempo. Considérese, por ejemplo, el sistema de la Figura 8.3 con $L = 8$. En este caso, tenemos $\omega_b L = 16$, lo que implica que el error permisible en el retardo de tiempo es como máximo un 6 por ciento.

La función de transferencia del lazo

El análisis de las funciones de sensibilidad indica que la robustez del sistema en lazo cerrado con el predictor de Smith puede ser pobre cuando $\omega_b L$ es grande. El análisis de la función de transferencia del lazo da una compresión adicional.

Cuando no existen errores de modelado, la función de transferencia del lazo obtenido utilizando el predictor de Smith es

$$PC = \frac{PC_0}{1 + C_0(P_0 - P)} = \frac{P_0 C_0 e^{-sL}}{1 + P_0 C_0 (1 - e^{-sL})}. \quad (8.15)$$

La Figura 8.7 muestra los diagramas de Nyquist de la función de transferencia del lazo para diferentes valores de L .

Para $L = 1$, el diagrama de Nyquist tiene un lazo de tamaño moderado. El lazo incrementa a medida que aumenta L , tal y como se puede observar comparando los casos para $L = 1$ y $L = 2$ en la Figura 8.7. El lazo es casi circular para valores de L superiores a 2. Para $L = 2,99$, el lazo es infinitamente grande, y para $2,99 < L < 6,40$ la función da dos vueltas alrededor del punto crítico, una para ω positivo y otra para ω negativo. Obsérvese que sólo hemos mostrado la rama del diagrama de Nyquist correspondiente a $0 \leq \omega < \infty$. Los polos inestables son los polos de la función de transferencia del predictor (8.7) y el número de vueltas alrededor del punto crítico incrementa a medida que L aumenta. Para $L = 8$, se obtienen cuatro vueltas alrededor del punto crítico.

La Figura 8.8 muestra los diagramas de Bode de la función de transferencia del lazo para los casos $L = 1$ y $L = 8$.

Las funciones de transferencia del lazo cambian drásticamente con L . La frecuencia de cruce de ganancia es 0,82 para $L = 1$ y decrece a 0,13 para $L = 8$.

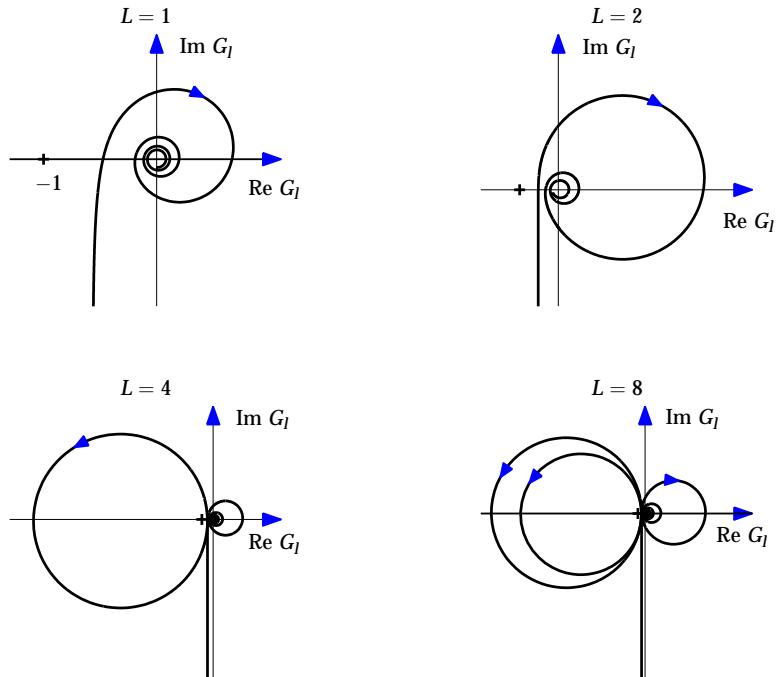


Figura 8.7 Diagramas de Nyquist de la función de transferencia del lazo para un sistema FOTD, (8.5), con el controlador del predictor de Smith. El punto crítico -1 está marcado con un +.

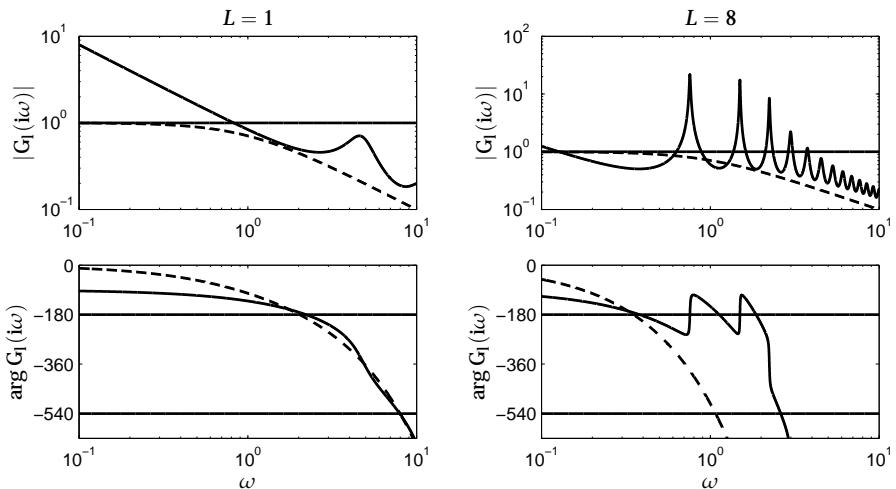


Figura 8.8 Diagramas de Bode de la función de transferencia del lazo (línea continua) y la función de transferencia del proceso (línea discontinua) para un sistema FOTD (8.5) con un predictor de Smith. Las curvas de la izquierda son para $L = 1$ y las de la derecha son para $L = 8$.

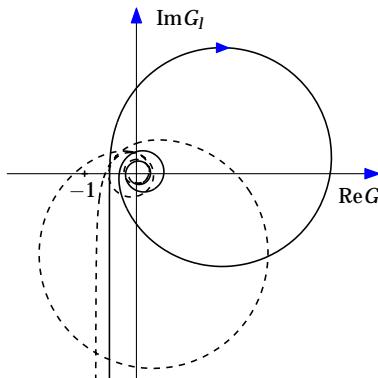


Figura 8.9 Diagramas de Nyquist de las funciones de transferencia del lazo para el sistema del Ejemplo 8.1 con $L = 2$ en el caso nominal (línea continua) y cuando el retardo de tiempo se aumenta en un 30 por ciento (línea discontinua).

Estos valores se ajustan bastante bien con el límite de rendimiento $\omega_{gc}L \approx 1$ dado por (4.57). Obsérvese que la curva de ganancia para $L = 8$ tiene varios cruces a altas frecuencias. La frecuencia de cruce de ganancia es más pequeña para $L = 8$ incluso si el tiempo de subida para cambios en el punto de consigna es el mismo para ambos sistemas. Los picos altos de la curva corresponden a los lazos del diagrama de Nyquist en la Figura 8.7.

El diagrama de Bode del sistema en lazo abierto se muestra en las líneas discontinuas de la Figura 8.8. Obsérvese que el controlador aporta un gran avance de fase a las frecuencias correspondientes a los dos primeros picos, que representan los polos inestables del controlador.

El margen de retardo

Las medidas de robustez clásicas, margen de ganancia y fase, no capturan las propiedades de los diagramas de Nyquist del tipo mostrado en la Figura 8.7, donde la curva de Nyquist posee grandes lazos. Esto se ilustra en la Figura 8.9, que muestra el diagrama de Nyquist de la función de transferencia del lazo para el caso $L = 2$ y para un sistema donde el retardo de tiempo del proceso ha sido aumentado en un 30 por ciento. La figura muestra que el sistema llega a ser inestable cuando el retardo de tiempo es incrementado en un 30 por ciento. Nótese que el lazo grande es el que cruza el punto crítico -1 y no la parte de la curva de Nyquist cercana al margen de ganancia. Para capturar este efecto, se introduce la medida de robustez denominada *el margen de retardo*. El margen de retardo se define como el cambio en el tiempo de retardo requerido para hacer un sistema inestable. Para sistemas con $L = 2$ y $L = 8$ en la Figura 8.7, los márgenes de retardo son 27 y 7 por ciento, respectivamente.

Obsérvese que las funciones de sensibilidad también capturan la robustez en los casos de las funciones de transferencia del lazo como las mostradas en la Figura 8.7. La sensibilidad a variaciones en el retardo de tiempo se puede estimar por (8.14), que da márgenes de retardo de 25 y 6 por ciento para los sistemas de la Figura 8.7 con $L = 2$ y $L = 8$. Estos números están próximos a los valores obtenidos utilizando el margen de retardo.

Otra forma de cuantificar la robustez consiste en explorar la sensibilidad del lazo cerrado a variaciones en los parámetros del proceso. Para procesos FOTD, tenemos

$$P(s) = \frac{K_p}{1+sT} e^{-sL}$$

Por tanto,

$$\log P = \log K_p - \log(1+sT) - sL$$

Diferenciando se obtiene

$$\frac{dP}{P} = \frac{dK_p}{K_p} - \frac{sdT_p}{1+sT_p} - sdL = \frac{dK_p}{K_p} - \frac{sdT_p}{1+sT_p} - sL \frac{dL}{L}$$

Para sistemas con grandes retardos de tiempo, el último término es dominante, lo que quiere decir que la sensibilidad a variaciones en el retardo de tiempo es la restricción crítica.

$$\frac{|dL|}{L} < \frac{1}{\omega_b L |T(i\omega)|},$$

y obtenemos la siguiente estimación del margen de retardo:

$$d_m = \max \frac{|dL|}{L} < \max \frac{1}{\omega_b L |T(i\omega)|}.$$

Resumen

El predictor de Smith hace posible obtener mejoras sustanciales en la respuesta a cambio en el punto de consigna tal y como se muestra en la Figura 8.3. El controlador se obtiene de una manera sencilla primero diseñando el controlador C_0 para el sistema nominal P_0 sin tiempo de retardo. Entonces, el predictor de Smith se obtiene colocando en cascada C_0 con un predictor C_{pred} , que eficazmente elimina el retardo de tiempo. Una característica interesante del predictor de Smith es que utiliza acciones de control pasadas para la predicción, lo que hace posible en principio compensar el retardo. Sin embargo, el controlador puede tener polos inestables. El producto $\omega_b L$, donde ω_b es el ancho de banda del sistema nominal en lazo cerrado $T_0 = P_0 C_0 / (1 + P_0 C_0)$ y L es el retardo de tiempo, es un parámetro crucial. El número de polos inestables del controlador crece con $\omega_b L$, y controladores con polos en el semiplano derecho presentan una pobre robustez. Además, variaciones permisibles en el tiempo de retardo son inversamente proporcionales a $\omega_b L$. Para disponer de un sistema en lazo cerrado robusto, es por lo tanto necesario restringir $\omega_b L$. En el Ejemplo 8.3 encontramos que, por ejemplo, para tener $M_s = 1,4$ fue necesario conseguir $\omega_b L < 0,5$.

8.4 El controlador PPI

En esta sección describiremos casos especiales del predictor de Smith que dan lugar a controladores de una forma particularmente simple. El predictor de Smith tratado en el Ejemplo 8.1 se basó en un modelo FOTD. El criterio de diseño fue encontrar un controlador que diese un sistema de segundo orden con polos que

tengan un determinado amortiguamiento ζ y una frecuencia ω_0 para el sistema sin retardo de tiempo. Otro posible diseño sería escoger un controlador que cancelase el polo del proceso y hacer el otro polo del sistema en lazo cerrado igual a $s = -1/T_{cl}$, donde T_{cl} es el tiempo de respuesta deseado del sistema en lazo cerrado. El método de diseño resulta en los siguientes parámetros del controlador;

$$K = \frac{T}{T_{cl}K_p}, \quad T_i = T.$$

La función de transferencia del lazo del sistema nominal sin retardo es $P_0C_0 = 1/(sT_{cl})$, y el controlador tiene la función de transferencia

$$C(s) = \frac{1+sT}{K_p s T_{cl}} \frac{1}{1 + \frac{1}{sT_{cl}}(1 - e^{-sL})}. \quad (8.16)$$

La función de transferencia del lazo es

$$P(s)C(s) = \frac{1}{sT_{cl}} \frac{1}{1 + \frac{1}{sT_{cl}}(1 - e^{-sL})}. \quad (8.17)$$

Como el polo del proceso se cancela, se debería requerir que el polo del proceso sea rápido en comparación con la dinámica dominante en lazo cerrado; ver la Sección 6.6. El parámetro de sintonía es: el tiempo de respuesta en lazo cerrado T_{cl} .

La relación entrada-salida del controlador (8.16) se puede escribir como

$$\begin{aligned} U(s) &= \frac{1+sT}{K_p s T_{cl}} E(s) - \frac{1}{sT_{cl}}(1 - e^{-sL}) U(s) \\ &= \frac{1+sT}{K_p s T_{cl}} \left(E(s) - \frac{K_p}{1+sT}(1 - e^{-sL}) U(s) \right) = \frac{1+sT}{K_p s T_{cl}} E_p(s), \end{aligned} \quad (8.18)$$

donde $E_p(s)$ es la transformada de Laplace del error predicho

$$e_p(t) = y_{sp}(t) - y(t) - \tilde{y}(t),$$

y

$$\tilde{Y}(s) = \frac{K_p}{1+sT}(1 - e^{-sL}) U(s).$$

El término $\tilde{y}(t)$ representa el efecto de las acciones de control tomadas en el intervalo $(t - L, t)$ sobre la salida del sistema. El controlador puede por lo tanto interpretarse como un controlador PI que actúa sobre un error predicho, que es el error actual compensado por las acciones de control pasadas que no han aparecido todavía a la salida. El controlador se denomina *controlador PI de predicción o controlador PPI*.

El controlador es particularmente simple si $T_{cl} = T$. La relación entrada-salida del controlador llega a ser entonces

$$U(s) = K E(s) + \frac{e^{-sL}}{1+sT_i} U(s).$$

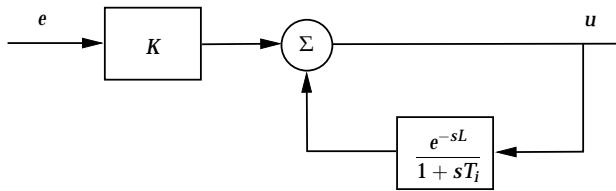


Figura 8.10 Diagrama de bloques de una implementación del controlador PPI con $T_{cl} = T_i$.

La Figura 8.10 muestra un diagrama de bloques que describe esta ecuación. Obsérvese la gran similitud con el controlador PI mostrado en la Figura 3.3. Existen varias versiones del controlador donde la ganancia se reemplaza por un controlador PD.

El predictor

El controlador PPI (8.16) es una combinación en cascada de un controlador PI y un predictor con la función de transferencia

$$C_{pred}(s) = \frac{1}{1 + \frac{1}{sT_{cl}}(1 - e^{-sL})}. \quad (8.19)$$

Aparte del escalado en frecuencia, el predictor de Smith se caracteriza completamente por la razón T_{cl}/L . Se puede observar que el predictor no tiene polos en el semiplano derecho para algunos valores de T_{cl} . La razón de esto es que la función de transferencia del lazo del sistema nominal sin retardo posee fase constante.

Un desarrollo en serie de la función de transferencia (8.19) para valores pequeños de s resulta en

$$\begin{aligned} C_{pred}(s) &\approx \frac{1}{1 + L/T_{cl} - sT_{cl}(L/T_{cl})^2/2 + \dots} \\ &\approx \frac{1}{1 + L/T_{cl}} \left(1 + \frac{1}{2} \frac{(L/T_{cl})^2}{1 + L/T_{cl}} T_{cl}s + \dots \right). \end{aligned} \quad (8.20)$$

La ganancia estática es $C_{pred}(0) = 1/(1 + L/T_{cl})$, y también se deduce que C_{pred} tiende a 1 cuando s tiende a infinito. La Figura 8.11 muestra el diagrama de Bode del predictor (8.19). También se muestra el diagrama de Bode del predictor ideal $e^{sT_{pred}}$ por cuestiones comparativas, donde

$$T_{pred} = \frac{1}{2} \frac{(L/T_{cl})^2}{1 + L/T_{cl}}, \quad (8.21)$$

y un predictor basado en diferenciación. El predictor basado en diferenciación se ha ajustado para dar la máxima ganancia como el predictor (8.19). Se pueden observar diferencias entre los predictores. El predictor ideal tiene ganancia unidad para todas las frecuencias mientras que los otros predictores poseen ganancias más altas a altas frecuencias y más bajas a bajas frecuencias. El predictor (8.19) aporta un mayor avance de fase que el predictor basado en diferenciación, pero el avance de fase cae rápidamente a altas frecuencias.

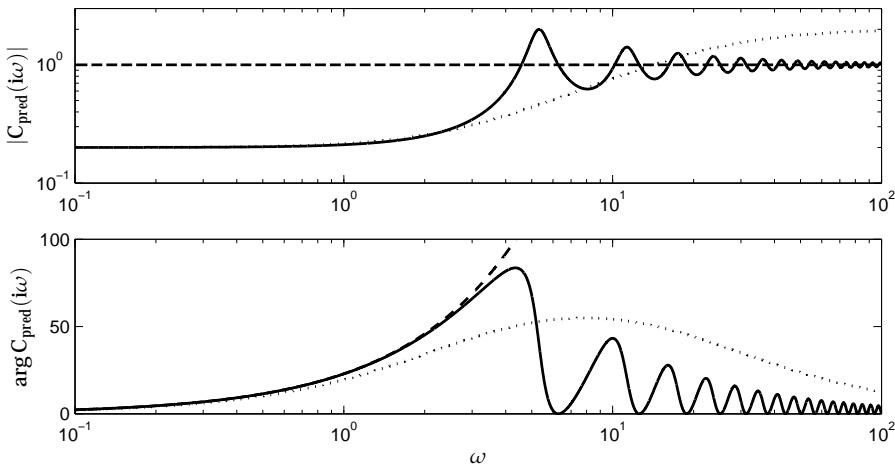


Figura 8.11 Diagramas de Bode para el predictor (8.19) (línea continua), un predictor basado en diferenciación (línea punteada) y el predictor ideal e^{sT} (línea a tramos).

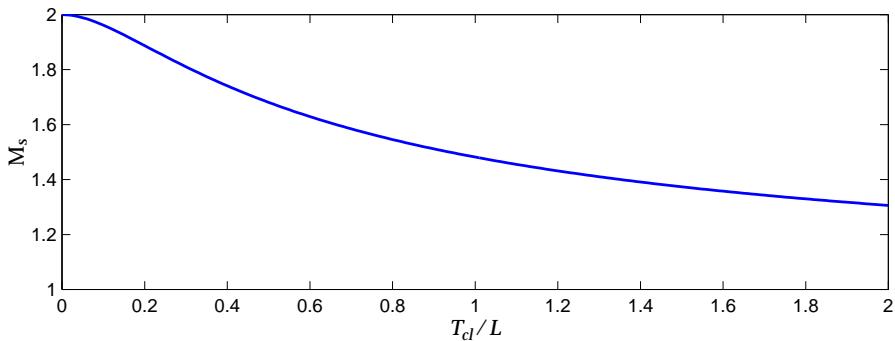


Figura 8.12 Sensibilidad máxima M_s del sistema en lazo cerrado con el controlador PPI (8.16) como función de T_{cl}/L .

Elecciones de diseño

La elección del parámetro de diseño T_{cl} es un compromiso entre robustez y rendimiento. El tiempo de respuesta viene dado directamente por T_{cl} , donde tiempos rápidos de respuesta requieren pequeños valores de T_{cl} . La robustez está gobernada por la razón T_{cl}/L . La función de sensibilidad viene dada por

$$S = 1 - \frac{e^{-sL}}{1 + sT_{cl}}.$$

La Figura 8.12 muestra la sensibilidad máxima como una función de T_{cl}/L . Obsérvese que la sensibilidad máxima tiene la propiedad $M_s \leq 2$. Para tener $M_s \leq 1,6$ se requiere que $T_{cl} > 0,66L$, y para $M_s \leq 1,4$ se necesita que $T_{cl} > 1,4L$. Para disponer de una robustez razonable, el tiempo de respuesta no puede seleccionarse mucho más pequeño que L . De (8.14) se deduce que el máximo error

relativo en el retardo de tiempo viene dado por

$$\frac{|\Delta L|}{L} \leq \frac{T_{cl}}{L}.$$

Si se permiten sensibilidades máximas como $M_s = 2$ y el retardo de tiempo se conoce de manera precisa, es posible permitir valores más pequeños en la razón T_{cl}/L .

Para sistemas sobreamortiguados, la integral del error IE es una medida que es sencilla de calcular. De (8.18), el controlador PPI en el dominio del tiempo es

$$u(t) = \frac{T}{K_p T_{cl}} e(t) + \frac{1}{K_p T_{cl}} \int_0^t e(t) dt - \frac{1}{T_{cl}} \int_0^t (u(t) - u(t-L)) dt. \quad (8.22)$$

Para calcular la integral del error para el controlador PPI, se asumirá que el sistema está inicialmente en reposo y que se aplica un escalón unitario a la perturbación de carga en la entrada del proceso. Debido a que el control posee acción integral, tenemos que $u(\infty) = 1$. Por consiguiente,

$$\int_0^\infty (u(t) - u(t-L)) dt = L.$$

Después de una perturbación de carga unitaria, de (8.22) se deduce que

$$u(\infty) - u(0) = 1 = \frac{1}{K_p T_{cl}} \int_0^\infty e(t) dt - \frac{L}{T_{cl}}.$$

Por tanto, la integral del error resulta ser

$$IE_{\text{PPI}} = K_p(L + T_{cl}).$$

La integral del error consiste de dos términos. El primer término, $K_p L$, se debe al retardo de tiempo y no puede ser influenciado por el controlador. El segundo término, $K_p T_{cl}$, se puede hacer pequeño especificando una constante de tiempo en lazo cerrado T_{cl} pequeño. Sin embargo, un valor pequeño de T_{cl} resultará en una pobre robustez.

Es interesante comparar el rendimiento del controlador PPI con el del controlador PID. En la Sección 4.9, se mostró que la integral del error para un controlador PID es

$$IE_{\text{PID}} = \frac{T_i}{K} = \frac{1}{k_i}.$$

De (7.7) se tiene que un controlador PID para procesos con retardo dominante y sintonizado para $M_s = 1,4$ tiene $k_i K_p L = 0,5$. Esto da $IE = 2K_p L$, que está próximo al valor $IE = 2,4K_p L$ obtenido para el controlador PPI. Por lo tanto, obtenemos la conclusión de que el controlador PPI no aporta significativamente mejor rendimiento frente a perturbaciones de carga que un controlador PI si ambos controladores tienen la misma robustez. La principal ventaja del controlador PPI es su capacidad para mejorar el seguimiento de puntos de consigna; ver Figura 8.3.

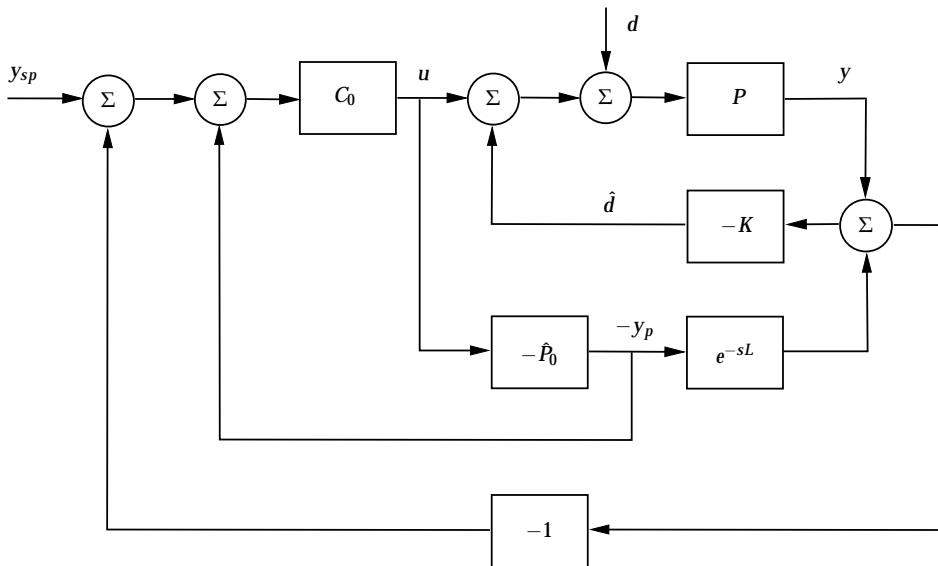


Figura 8.13 Predictor de Smith modificado para procesos integradores.

8.5 Predictores para procesos integradores

El predictor de Smith básico tiene propiedades útiles, pero también posee algunos inconvenientes graves. No puede utilizarse con procesos inestables, y posee error en régimen permanente frente a perturbaciones de carga para procesos con integración. Como veremos, se han propuesto algunas modificaciones.

Para procesos con integración, se ha sugerido modificar el predictor de Smith, como se muestra en la Figura 8.13, para obtener error cero en régimen permanente frente a perturbaciones de carga. La razón de esta modificación se puede entender del principio del control basado en modelo interno. La señal \hat{d} que se realimenta es una estimación de la perturbación de carga.

En la Figura 8.13, las funciones de transferencia desde el punto de consigna y_{sp} y la perturbación de carga d a la salida y vienen dadas por

$$Y = \frac{PC_0(1 + K\hat{P})}{1 + C_0(\hat{P}_0 - \hat{P}) + P(K + KC_0\hat{P}_0 + C_0)} Y_{sp} + \frac{P(1 + C_0(\hat{P}_0 - \hat{P}))}{1 + C_0(\hat{P}_0 - \hat{P}) + P(K + KC_0\hat{P}_0 + C_0)} D. \quad (8.23)$$

Cuando $s \rightarrow 0$, resultan las siguientes aproximaciones:

$$C_0 \approx \frac{k_i}{s} \quad \hat{P}_0 - \hat{P} \approx \frac{K_v}{s}(1 - e^{-sL}) \approx K_v L.$$

Si asumimos también que $\hat{P} = P$, se puede mostrar que la función de transferencia entre y_{sp} e y llega a ser uno, y la función de transferencia entre d e y llega a ser cero cuando $s \rightarrow 0$.

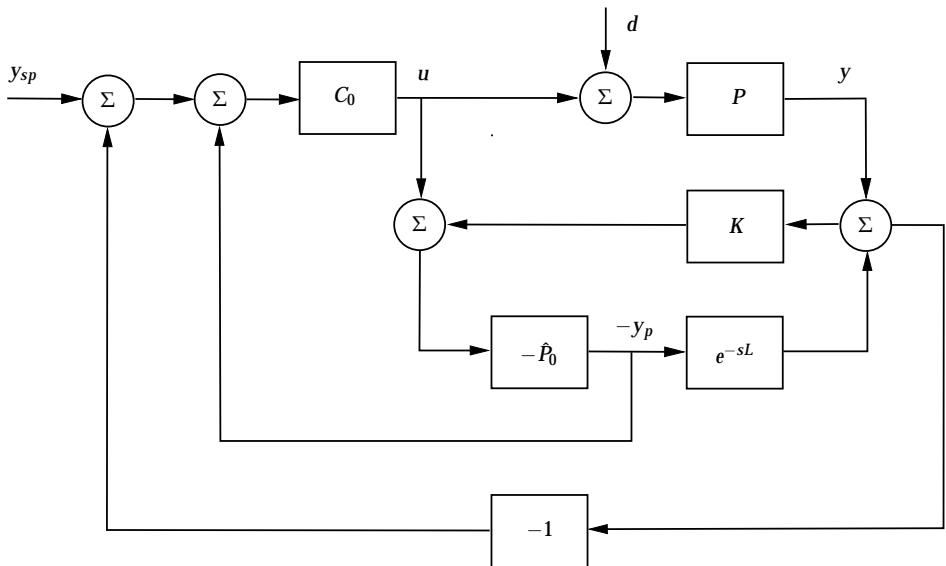


Figura 8.14 Predictor de Smith modificado para procesos integradores.

Otra modificación para procesos integradores se da en la Figura 8.14. La variable y_p es una estimación de la señal de medida no retardada

$$Y = P_0(U + D).$$

La estimación viene dada por

$$Y_p = \hat{P}_0(U + K(Y - \hat{Y})).$$

Cuando \hat{P}_0 es estable, se puede utilizar el valor $K = 0$, correspondiente al predictor de Smith original. Para procesos integradores, es, sin embargo, necesario tener $K \neq 0$.

De la Figura 8.14, las funciones de transferencia desde el punto de consigna y_{sp} y la perturbación de carga d a la salida y vienen dadas por

$$\begin{aligned} Y &= \frac{PC_0(1 + KP)}{1 + K\hat{P} + \hat{P}_0 - \hat{P} + PC_0(1 + K\hat{P}_0)} Y_{sp} \\ &\quad + \frac{P(1 + K\hat{P} + \hat{P}_0 - \hat{P})}{1 + K\hat{P} + \hat{P}_0 - \hat{P} + PC_0(1 + K\hat{P}_0)} D. \end{aligned} \tag{8.24}$$

Bajo la hipótesis de que $\hat{P} = P$, se puede mostrar que la función de transferencia entre y_{sp} e y es uno, y la función de transferencia entre d e y es cero cuando $s \rightarrow 0$.

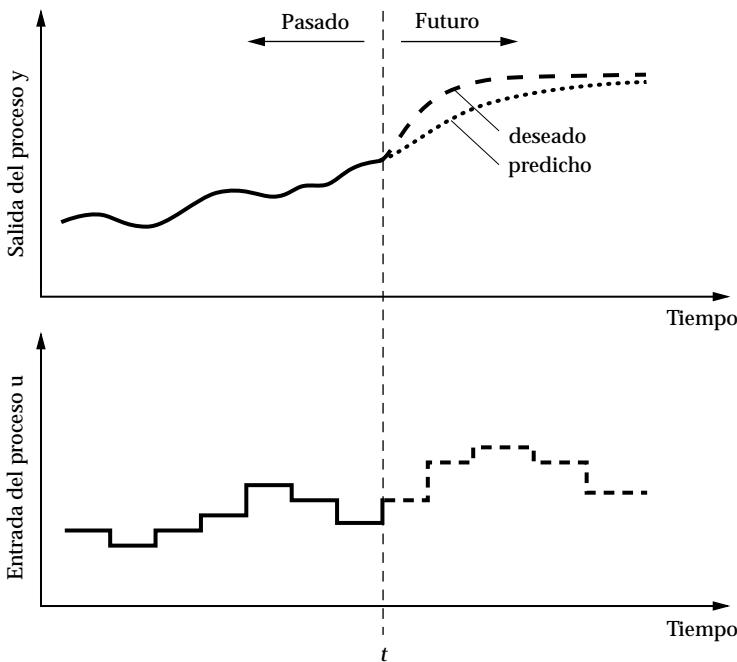


Figura 8.15 Ilustración del control predictivo basado en modelo.

8.6 Control predictivo basado en modelo

El control predictivo basado en modelo se basa en la predicción del comportamiento futuro del proceso utilizando un modelo del proceso y la optimización del comportamiento del sistema sobre un horizonte finito de tiempo. La implementación se obtiene aplicando la parte inicial de la señal de control y repitiendo el proceso sobre un horizonte de tiempo desplazado. Este procedimiento se denomina *control de horizonte deslizante* o *control de horizonte móvil*. Haciendo referencia a la Figura 8.15, el algoritmo se puede describir como sigue:

- 1: Obtener un modelo del proceso.
- 2: Considérese la situación en el instante t . Se observan las entradas y las salidas pasadas del proceso u e y ; ver Figura 8.15. El comportamiento futuro del proceso se predice bajo la suposición de que el modelo del proceso y las señales de control futuras $u_f = u(\tau)$, $t \leq \tau < t + t_h$ son conocidas.
- 3: Se calcula la señal de control u_f para obtener el comportamiento futuro deseado.
- 4: Se aplica la parte inicial de la señal de control u_f en el intervalo $[t, t + h]$.
- 5: Modificar el tiempo a $t + h$, y repetir el procedimiento desde el paso 2.

Los pasos se pueden llevar a cabo de diferentes maneras y existen un gran número de algoritmos. Se pueden utilizar diferentes modelos; modelos físicos,

modelos entrada-salida y modelos en el espacio de estados. Además, el método se puede utilizar tanto para sistemas de una entrada y una salida como para sistemas de múltiples entradas y salidas.

El comportamiento deseado se puede especificar de muchas formas. Un procedimiento común consiste en especificar el comportamiento futuro deseado mediante un modelo matemático, por ejemplo, uno que indica cómo aproximarse al punto de consigna. La desviación del comportamiento deseado se puede formular como un problema de optimización para minimizar la desviación entre el comportamiento actual y el deseado, posiblemente con una penalización sobre las señales de control. El paso 2 es un problema de optimización en lazo abierto donde la optimización se lleva a cabo sobre un horizonte de tiempo finito. La realimentación se obtiene aplicando sólo la parte inicial de la señal de control, desplazando posteriormente el horizonte hacia adelante y repitiendo entonces el problema de optimización.

El control predictivo basado en modelo es particularmente simple para sistemas muestreados donde la señal de control es constante entre intervalos de muestreo. El parámetro h se puede escoger como el intervalo de muestreo, y el horizonte de predicción t_h se selecciona típicamente como un número pequeño de intervalos de muestreo. La mayoría de los controladores predictivos se han desarrollado para sistemas muestreados.

Una propiedad muy útil del control predictivo basado en modelo es que se pueden tener en cuenta de manera sistemática restricciones sobre la señal de control y la salida del proceso. Una elección común es formular el problema tal que se pueden utilizar algoritmos eficientes de programación cuadrática. Una dificultad clave en el control predictivo basado en modelo consiste en asegurar estabilidad cuando el horizonte es finito. Existen numerosos trabajos de investigación dedicados a este problema.

Un ejemplo simple

Para ilustrar las ideas daremos algunos detalles mediante un caso sencillo. Considerese el modelo de proceso muestreado

$$y(t) + a_1y(t-h) + \cdots + a_ny(t-nh) = b_1u(t-h) + b_2u(t-2h) + \cdots + b_nu(t-nh), \quad (8.25)$$

donde y es la salida y u la entrada del proceso. Considérese la situación en el instante t . El comportamiento pasado viene completamente caracterizado por

$$\mathcal{Y}_t = (y(t), y(t-h), \dots, u(t-h), u(t-2h), \dots). \quad (8.26)$$

Utilizando el modelo, es sencillo predecir los valores futuros de la salida del proceso como función de las señales de control actuales y futuras:

$$\mathcal{U}_t = (u(t), u(t+h), \dots, u(t+Nh)). \quad (8.27)$$

El comportamiento futuro deseado se puede caracterizar especificando la trayectoria de referencia para las salidas futuras del proceso, tal y como se indica en la Figura 8.15, y dando una función de coste que penaliza las desviaciones $e(t) =$

$y(t) - y_d(t)$ desde la salida deseada $y_d(t)$ y los incrementos de la señal de control $\Delta u(t) = u(t) - u(t-h)$

$$J(u(t), u(t+h), \dots, u(t+Nh)) = \sum_{k=1}^{t+N} e(t+kh)^2 + \rho(\Delta u(t+(k-1)h))^2. \quad (8.28)$$

Además pueden existir restricciones sobre las entradas y las salidas del proceso y sobre el incremento de la señal de control.

Las señales de control futuras u_t son entonces calculadas minimizando J sujeta a las restricciones. La señal de control $u(t)$ se aplica, y el procedimiento al completo es repetido. La señal de control es una función de las entradas y las salidas pasadas

$$u(t) = F(y(t), y(t-h), \dots, y(t-nh), u(t-h), y(t-2h), \dots, u(t-nh)),$$

donde la función F se obtiene implícitamente resolviendo un problema de optimización.

Un caso particular simple es cuando el modelo del proceso es de primer orden en los incrementos de las entradas y salidas del mismo, lo cual se ilustra mediante el siguiente ejemplo.

EJEMPLO 8.4—MPC PARA SISTEMA DE PRIMER ORDEN

Sea el modelo del proceso sea

$$\Delta y(t+h) = -a\Delta y(t) + b\Delta u(t),$$

donde $\Delta y(t) = y(t) - y(t-h)$ y $\Delta u(t) = u(t) - u(t-h)$. Sea la trayectoria deseada una señal $y_d(t)$ que empieza en $y(t)$ y se aproxima al punto de consigna y_{sp} exponencialmente con una constante T_{cl} . La salida del proceso deseada en el instante de tiempo $t+h$ es entonces

$$y_d(t+h) = y(t) + (1 - e^{-h/T_{cl}})(y_{sp} - y(t)).$$

Asumiendo que no existen penalizaciones sobre las señales de control, la salida deseada del proceso puede obtenerse en el siguiente periodo de muestreo. Equiparando $y(t+h)$ con $y_d(t+h)$ se obtiene

$$y(t+h) = y(t) + \Delta y(t+h) = y(t) - a\Delta y(t) + b\Delta u(t) = y(t) + (1 - e^{-h/T_{cl}})(y_{sp} - y(t)).$$

Resolviendo esta ecuación para $\Delta u(t)$ se tiene que

$$\Delta u(t) = \frac{1 - e^{-h/T_{cl}}}{b}(y_{sp} - y(t)) + \frac{a}{b}\Delta y(t),$$

que es un controlador PI con ganancias

$$k = \frac{a}{b}$$

$$k_i = \frac{1 - e^{-h/T_{cl}}}{b}.$$

Obsérvese que la ganancia proporcional sólo depende del modelo del proceso y que la ganancia integral dependen de la velocidad de respuesta deseada T_{cl} . \square

Es sencillo tratar con problemas que tienen múltiples entradas y salidas, siendo posible además incluir restricciones. Existen muchos casos especiales y variantes del control predictivo basados en modelo. Algunos serán descritos brevemente y para más detalles nos remitimos a las referencias.

El algoritmo de Dahlin-Higham

Uno de los primeros algoritmos de control predictivo basado en modelo fue desarrollado para el control de las máquinas de papel. El algoritmo se basa en un proceso del modelo en términos de un modelo FOTD

$$P(s) = \frac{K_p}{1 + sT} e^{-sL},$$

y la respuesta deseada frente a cambios en el punto de consigna viene dada por

$$G_{yy_{sp}} = \frac{1}{1 + sT_{cl}} e^{-sL}.$$

Asumiendo que la señal de control es constante sobre los intervalos de muestreo de longitud $h = L/n$, donde n es un entero, se obtiene el modelo de proceso muestreado

$$y(t+h) = ay(t) + K_p(1-a)u(t-nh).$$

La respuesta deseada frente a cambios en el punto de consigna viene dada por la ecuación en diferencias

$$y_d(t+h) = a_d y_d(t) + (1-a_d)y_{sp}(t-nh).$$

Introduciendo el operador de desplazamiento hacia atrás q^{-1} definido por

$$q^{-1}y(t) = y(t-h), \quad (8.29)$$

el modelo del proceso se puede escribir como

$$y(t) = \frac{K_p(1-a)q^{-(n+1)}}{1 - aq^{-1}} u(t) = P(q^{-1})u(t).$$

Sea el controlador caracterizado por

$$u(t) = C(q^{-1})(y_{sp}(t) - y(t)).$$

La relación entrada-salida del sistema en lazo cerrado es entonces

$$y(t) = \frac{P(q^{-1})C(q^{-1})}{1 + P(q^{-1})C(q^{-1})} y_{sp}(t).$$

Utilizando el operador de desplazamiento hacia atrás, la respuesta deseada viene dada por

$$y_d(t) = \frac{(1 - a_d)q^{-(n+1)}}{1 - a_d q^{-1}} y_{sp}(t) = G_d(q^{-1}) y_{sp}(t),$$

donde $a_d = e^{-h/T_{cl}}$. Equiparando esto con la salida del proceso se obtiene

$$\frac{P(q^{-1})C(q^{-1})}{1 + P(q^{-1})C(q^{-1})} = G_d(q^{-1}) = \frac{(1 - a_d)q^{-(n+1)}}{1 - a_d q^{-1}}.$$

Resolviendo esta ecuación con respecto a $C(q^{-1})$ se tiene

$$C(q^{-1}) = \frac{G_d(q^{-1})}{P(q^{-1})(1 - G_d(q^{-1}))} = \frac{(1 - a_d)(1 - aq^{-1})}{K_p(1 - a_d q^{-1} - (1 - a_d)q^{-(n+1)})}.$$

El controlador puede entonces describirse por

$$u(t) = \frac{1 - a_d}{K_p} (e(t) - ae(t - h)) + a_d u(t - h) + (1 - a_d) u(t - (n + 1)h).$$

Este controlador posee acción integral y las salidas pasadas se utilizan para predicción.

Control de matriz dinámica (DMC)

En el control de matriz dinámica (dynamic matrix control, DMC), el proceso se modela por un modelo de respuesta impulsional finita

$$y(t) = b_1 u(t - h) + b_2 u(t - 2h) + \dots + b_n u(t - nh), \quad (8.30)$$

y el criterio consisten en minimizar la función de coste

$$J(u(t), u(t + h), \dots, u(t + (n - 1)h)) = \sum_{k=1}^n e^2(t + kh),$$

donde

$$e(t + kh) = y_d(t + kh) - b_1 u(t + kh - h) - b_2 u(t + kh - 2h) - \dots - b_n u(t + kh - nh).$$

El problema de optimización es sencillo de resolver debido a que e es una función lineal de las variables de control futuras y la función de coste es cuadrática. Obsérvese que el modelo (8.30) también es útil si existen múltiples entradas y salidas, donde en ese caso los coeficientes b_i serían matrices. Estos coeficientes se denominan *matrices dinámicas* ya que reflejan la dinámica de la respuesta, lo que motivó el nombre de DMC. En terminología estándar de control, los parámetros son simples coeficientes de la respuesta impulsional. En los primeros usos del control de matriz dinámica, era práctica común determinar las matrices b_i de una simple medida de la respuesta a impulso o a escalón.

Un inconveniente con DMC es que puede ser necesario un gran número de parámetros si la dinámica del proceso es muy lenta. El algoritmo DMC fue más tarde generalizado a QDMC (control cuadrático de matriz dinámica), que puede tener en cuenta restricciones sobre la señal de control.

Control de mínima varianza

El controlador de mínima varianza es un controlador predictivo para sistemas con perturbaciones aleatorias donde el criterio consiste en minimizar la varianza de las fluctuaciones en la salida del proceso. El algoritmo se desarrolló originalmente para el control de máquinas de papel donde la naturaleza estocástica de las perturbaciones es tan importante como la dinámica del proceso. Empezaremos con un ejemplo sencillo.

EJEMPLO 8.5—CONTROL DE MÍNIMA VARIANZA

Sea un modelo

$$y(t+h) = -ay(t) + bu(t) + e(t+h) + ce(t),$$

donde u es la variable de control, y es la salida del proceso y e es una secuencia de variables aleatorias independientes con media cero y desviación estándar σ . El periodo de muestreo es h .

Considérese la situación en el instante t . La salida del proceso $y(t)$ es conocida y la salida en el instante $t+h$ puede tomar valores arbitrarios escogiendo la señal de control $u(t)$. La señal aleatoria $e(t+h)$ es independiente de las entradas y salidas pasadas \mathcal{Y}_t dadas por (8.26). Además, $e(t)$ se puede calcular desde las entradas y salidas pasadas \mathcal{Y}_t . La ley de control que minimiza la desviación desde el punto de consigna y_{sp} viene dada por

$$u(t) = \frac{ay(t) - ce(t)}{b},$$

Si se utiliza esta ley de control, encontramos que $y(t) = e(t)$, que quiere decir que la salida es ruido blanco. El cálculo de $e(t)$ desde las entradas pasadas es por tanto trivial y la ley de control es

$$u(t) = \frac{a-c}{b}y(t).$$

□

En el caso general, el modelo del proceso es

$$a(q^{-1})y(t) = b(q^{-1})u(t) + c(q^{-1})e(t). \quad (8.31)$$

donde u es la entrada del proceso, y es la salida del proceso y e es una secuencia de variables Gausianas aleatorias independientes con media cero y varianza σ . $a(q^{-1})$, $b(q^{-1})$ y $c(q^{-1})$ son polinomios en el operador de desplazamiento hacia atrás

$$\begin{aligned} a(q^{-1}) &= 1 + a_1q^{-1} + a_2q^{-2} + \dots + a_nq^{-n} \\ b(q^{-1}) &= b_\ell q^{-\ell} + b_{\ell+1}q^{-\ell-1} + \dots + b_nq^{-n} \\ c(q^{-1}) &= 1 + c_1q^{-1} + c_2q^{-2} + \dots + c_nq^{-n}. \end{aligned}$$

Por temas de simplicidad se han escogido todos los polinomios del mismo grado. Esto no presenta pérdida de generalidad ya que podemos permitir el hacer coeficientes iguales a cero. El coeficiente b_ℓ es el primer coeficiente que no desaparece en el polinomio $b(q^{-1})$. El número ℓ es un parámetro importante denominado retraso de entrada-salida y además introducimos el polinomio $b'(q^{-1}) = q^\ell b(q^{-1})$.

Es natural asumir que no existen factores comunes para los tres polinomios $a(q^{-1})$, $b(q^{-1})$ y $c(q^{-1})$. Se supone que el polinomio $c(q^{-1})$ tiene todos sus ceros fuera del círculo unidad. El modelo (8.31) captura la dinámica tanto del proceso como de sus perturbaciones.

El control de mínima varianza está muy relacionado con la predicción, y por tanto determinaremos primero un predictor para la salida del proceso cuando la entrada u es cero. La predicción de ℓ pasos hacia adelante viene dada por

$$c(q^{-1})\hat{y}(t + \ell) = g(q^{-1})y(t),$$

donde el polinomio $g(q^{-1})$ es

$$a(q^{-1})f(q^{-1}) + q^{-1}g(q^{-1}) = c(q^{-1}).$$

Obsérvese que la dinámica del predictor viene dada por el polinomio $c(q^{-1})$ en el modelo (8.31). La predicción del error

$$\epsilon(t) = f(q^{-1})e(t)$$

tiene la varianza

$$E\epsilon^2 = \sigma^2 \sum_0^{\ell-1} f_k^2. \quad (8.32)$$

La estrategia de control de mínima varianza simple viene dada por

$$u(t) = -\frac{s(q^{-1})}{r(q^{-1})}y(t) = -\frac{g(q^{-1})}{b'(q^{-1})f(q^{-1})}y(t), \quad (8.33)$$

y el error de control es

$$y(t) = f(q^{-1})e(t). \quad (8.34)$$

El error bajo el control de mínima varianza es por lo tanto igual al error en la predicción de la salida ℓ pasos hacia adelante. El error de control es una media móvil de orden $\ell - 1$. Es así sencillo determinar si un proceso está bajo control de mínima varianza simplemente calculando la función de correlación de la salida. Como el error de control es de media móvil de orden $\ell - 1$, su función de covarianza es cero para todos los retrasos mayores que ℓ .

La robustez del control de mínima varianza está fuertemente influenciada por la elección del tiempo de muestreo. En la práctica, es bueno escoger h mayor que $L/2$.

8.7 Resumen

El rendimiento de un controlador PI se puede mejorar incorporando capacidades predictivas. La acción derivativa es una posibilidad, pero existen otras alternativas. El predictor de Smith y el control predictivo basado en modelo son útiles para sistemas con retardo de tiempo cuando se dispone de modelos apropiados. En tales casos se puede obtener grandes mejoras en la respuesta al punto de consigna. El controlador predictivo PI es una versión simple del predictor de Smith teniendo la ventaja sobre un controlador PID de que permite alcanzar un mayor avance de fase. Una paradoja es que el controlador predictivo sólo aporta mejoras modestas en comparación con los controladores PI para proceso con dinámicas de retardo dominante, siendo tales mejoras significativas para procesos con constantes de tiempo dominantes. Los controladores predictivos basados en modelo son más generales que los predictores de Smith y se pueden utilizar también para tratar con sistemas de múltiples entradas y salidas. Además, pueden tener en cuenta restricciones de manera sistemática.

Como los controladores predictivos se basan en modelos matemáticos, es importante que los modelos sean precisos, siendo particularmente importante tener una buena estimación del retardo de tiempo. Se realizó un análisis exhaustivo de la robustez del predictor de Smith, estando disponibles resultados similares para otros controladores predictivos. El resultado clave es que la sensibilidad a errores de modelado está muy relacionada con el parámetro $\omega_b L$, donde ω_b es el ancho de banda en lazo cerrado o L/T_{cl} donde T_{cl} es el tiempo de respuesta en lazo cerrado cuando se desprecia el retardo L . La robustez requerida es que ambos parámetros no sean demasiado pequeños. Una regla razonable es que los parámetros deberían ser menores que 0.5.

8.8 Notas y referencias

En [323] se propuso un controlador para sistemas con retardo de tiempo, también tratado en el libro [322]. Una explicación del mecanismo que genera grandes avances de fase se da en [33]. Se han presentado muchas modificaciones del predictor de Smith; ver [22, 233, 234, 194, 207]. El controlador de [134], el controlador PPI en [143] y el controlador PID τ de [315] son casos especiales del predictor de Smith. Los trabajos [297, 241, 182] comparan predictores de Smith con controladores PID.

El control de mínima varianza fue desarrollado en la fase inicial del control por computador de máquinas de papel con el fin de encontrar una estrategia de control que minimizase las fluctuaciones en variables cualitativas. Un resultado clave es que la varianza más pequeña que se puede alcanzar es la varianza del error en la predicción de la salida sobre el retardo de tiempo del proceso. El control de mínima varianza fue primero publicado por [31] y en [39] se dio una perspectiva de su uso. El control de mínima varianza requiere un modelo de las perturbaciones y de la dinámica del proceso. En [12] se desarrolló un método para obtener esta información de experimentos sobre el proceso y su aplicación al modelado y control de máquinas de papel se realizó en [32]. El controlador autosintonizado [28] se puede ser como una automatización de identificación de

sistemas y control de mínima varianza.

El controlador presentado en [92] y [168] se puede ver como una versión en tiempo discreto del predictor de Smith. Ambos, el predictor de Smith y el controlador de Dahlin-Higham, que son versiones tempranas del control predictivo basado en modelo [312], fueron desarrollados para aplicaciones de control de procesos. Existen numerosas versiones de control predictivo basado en modelo; ver [293], [91] y [115]. Hay varios libros recientes sobre control predictivo basado en modelo [4, 200, 225]. Los trabajos resumen [292, 285] contienen muchas referencias. Los trabajos [209], [98] y [369], y el libro [52] dan una perspectiva práctica. Aunque el control predictivo basado en modelo fue originalmente pensado para sistemas multivariados, y su uso ha sido sugerido como un reemplazo al control PID; ver [222].

9

Sintonía automática y adaptación

9.1 Introducción

La sintonía automática, o autosintonía, es un método donde el controlador se sintoniza automáticamente bajo demanda de un usuario. Tradicionalmente, el usuario pulsaba un botón o enviaba una orden al controlador. La sintonía automática de controladores PID se puede llevar a cabo mediante la combinación de los métodos de determinadas dinámicas de procesos, descritas en el Capítulo 2, con los métodos para el cálculo de parámetros de controladores PID, descritos en los Capítulo 4, 6 y 7. Un procedimiento de sintonía automática consiste en tres pasos:

- Generación de la perturbación del proceso.
- Evaluación de la respuesta de la perturbación.
- Cálculo de los parámetros del controlador.

Éste es el mismo procedimiento que utiliza un ingeniero experimentado cuando sintoniza un controlador manualmente. El proceso debe ser perturbado de alguna manera para determinar su dinámica. Esto se puede hacer de múltiples maneras, por ejemplo, aplicando escalones, pulsos o sinusoides a la entrada del proceso. La evaluación de la respuesta a la perturbación puede incluir una determinación del modelo del proceso o una simple caracterización de la respuesta.

La experiencia industrial ha mostrado claramente que la sintonía automática es algo altamente útil y deseable. La sintonía automática se denomina algunas veces sintonía bajo demanda o sintonía de una vez. Los controladores PID comerciales con capacidades de sintonía automática están disponibles desde principios de los ochenta.

La sintonía automática se puede incorporar en un controlador. Puede realizarse utilizando dispositivos externos conectados al lazo de control sólo durante la fase de sintonía, de tal manera que se muestren los parámetros del controlador cuando el experimento de sintonía ha finalizado. Debido a que los dispositivos de

sintonía se supone que trabajan junto con los controladores de distintos fabricantes, éstos deben aportar suficiente información sobre el controlador para dar una sugerencia adecuada de cada parámetro.

Incluso cuando se utilizan dispositivos de sintonía automática es importante poseer cierto conocimiento del proceso. Esto se analiza en la Sección 9.2. La sintonía automática es sólo una manera de utilizar la técnica adaptativa. La Sección 9.3 muestra un resumen de varias técnicas adaptativas, así como una discusión sobre su uso. Los enfoques de sintonía automática se pueden dividir en dos categorías, los enfoques basados en modelo y los enfoques basados en reglas. En los enfoques basados en modelo, se obtiene un modelo del proceso de manera explícita y la sintonía se basa en dicho modelo. La Sección 9.4 trata enfoques donde el modelo se obtiene a partir de experimentos de la respuesta transitoria, experimentos de respuesta en frecuencia y estimación de parámetros. En los enfoques basados en reglas, no se obtiene un modelo explícito del proceso. En lugar de eso, la sintonía se basa en reglas similares a aquellas que utiliza un operador experimentado cuando sintoniza el controlador de manera manual. En la Sección 9.5 se describe el enfoque basado en reglas. La Sección 9.7 trata con la sintonización reaumentada iterativa, que es un método iterativo para la sintonía de controladores.

En la Sección 9.8 se presentan unos cuantos productos con capacidades adaptativas. Esta sección ilustra cómo se utilizan algunas de las ideas de adaptación en productos, pero no se tiene la intención de realizar una presentación exhaustiva de este tema. El capítulo finaliza con algunas conclusiones y referencias en las Secciones 9.9 y 9.10.

9.2 Conocimiento del proceso

En este capítulo describiremos varios métodos de sintonía automática. Antes de entrar en detalles, debemos enfatizar que el mal comportamiento en un lazo de control no puede ser siempre corregido mediante la sintonía del controlador. Es absolutamente necesario comprender la razón de dicho mal comportamiento.

El proceso puede ser diseñado de manera incorrecta de modo que existan grandes retardos de tiempo, grandes constantes de tiempo, no linealidades y respuestas inversas. Los sensores y actuadores pueden estar colocados incorrectamente o mal montados, de tal manera que pueden incorporar malas dinámicas. Ejemplos típicos son los termopares con cubiertas pesadas que hacen que su respuesta sea lenta o motores con válvulas todo-nada con largos tiempos de recorrido. Las válvulas pueden estar sobreestimadas de manera que sólo pueden actuar en una pequeña región. El recorrido del sensor puede ser demasiado amplio de forma que se obtiene una pobre resolución en la medida, o el sensor puede aportar también excesivo ruido de medida.

Además pueden existir fallos y desgastes en el equipamiento del proceso. Las válvulas pueden tener excesiva fricción estática o tener histéresis debido al desgaste. Los sensores pueden tener derivas y modificar sus medidas debido a la contaminación.

Si el lazo de control se está comportando de manera no satisfactoria, es esencial que primero determinemos la razón del problema antes de realizar la sintonía. Por supuesto, sería altamente deseable tener ayudas de los ingenieros de

procesos para realizar la diagnosis. La sintonía automática puede realmente llevar a resultados erróneos si no se aplica con cuidado. Por ejemplo, considérese un lazo de control que oscila a causa de la fricción en el actuador. Particularmente, todos los dispositivos de sintonía intentarán estabilizar la oscilación reduciendo la ganancia del controlador. ¡Esto sólo incrementará el período de oscilación! Estas importantes cuestiones se tratarán en un capítulo separado de este libro, el Capítulo 10. Recuérdese que ninguna cantidad de la bien llamada «inteligencia» de un elemento puede reemplazar el conocimiento del proceso real.

9.3 Técnicas adaptativas

Las técnicas de sintonía automática surgieron de la investigación control adaptativo. La adaptación fue originalmente desarrollada para tratar con procesos con características que eran variables en el tiempo o con distintas condiciones de operación. Particularmente, todas las técnicas adaptativas se pueden utilizar para sintonía automática. El controlador adaptativo es simplemente puesto en marcha hasta que los parámetros convergen y entonces los parámetros permanecen constantes. El inconveniente con este enfoque es que los controladores adaptativos pueden requerir información previa. Existen numerosas técnicas especiales que se pueden utilizar. La experiencia industrial ha mostrado que la sintonía automática es probablemente la aplicación más útil de las técnicas adaptativas. La planificación de ganancia es también una técnica muy efectiva para trabajar con procesos que cambian sus características con el punto de operación. En esta sección se dará un resumen de estas técnicas. En este libro, la frase *técnicas adaptativas* incluirán auto sintonía, planificación de ganancia y adaptación.

Control adaptativo

Un controlador adaptativo ajusta sus parámetros continuamente para tener en cuenta los cambios en la dinámica del proceso y las perturbaciones. La adaptación se puede aplicar a los parámetros de control por realimentación y por prealimentación.

Existen dos tipos de controladores adaptativos basados en métodos directos e indirectos. En un método directo, los parámetros del controlador se ajustan directamente a partir de los datos de operación en lazo cerrado. En los métodos indirectos, los parámetros de controlador se obtienen indirectamente primero mediante la actualización del modelo del proceso en línea y después determinando los parámetros del controlador mediante algún método de diseño de control. El sistema de modelo de referencia es un controlador adaptativo indirecto. El regulador auto sintonizado se puede implementar de ambas maneras, con control directo e indirecto. Existe un gran número de métodos disponibles para ambas formas directa e indirecta. Tales métodos se pueden describir en términos de los métodos utilizados para modelado y diseño de control.

En la Figura 9.1 se muestra el diagrama de bloques de un controlador adaptativo indirecto. Hay un estimador de parámetros que determina los parámetros del modelo basándose en las observaciones de las entradas y salidas del proceso. También existe un bloque de diseño que calcula los parámetros del controlador a

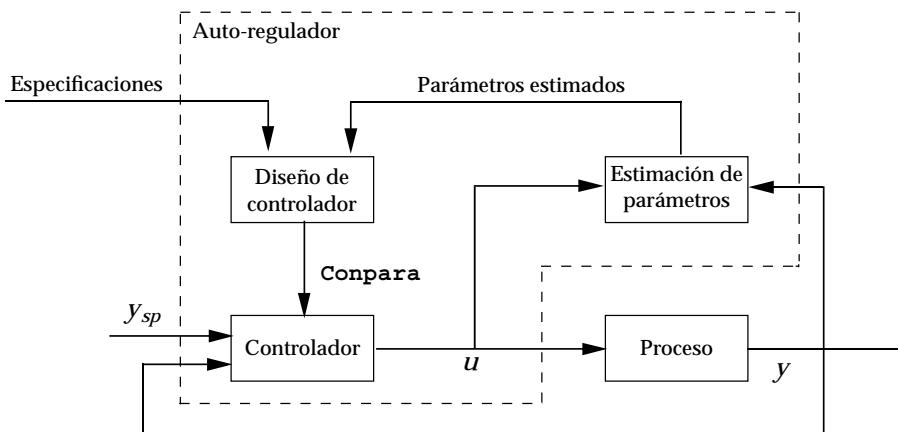


Figura 9.1 Diagrama de bloques de un controlador adaptativo.

partir de los parámetros del modelo. Si el sistema se opera como un sintonizador, el proceso se excita mediante una señal de entrada. Los parámetros se pueden estimar recursivamente o en modo por lotes. Entonces, se calculan los parámetros del controlador y se pone en marcha. Si el sistema se opera como un controlador adaptativo, se calculan los parámetros recursivamente y los parámetros del controlador se actualizan cuando se obtienen los nuevos valores.

Sintonía automática

Por sintonía automática (o auto sintonía) entendemos un método donde un controlador se sintoniza automáticamente o bajo demanda por el usuario. Típicamente, el usuario pulsará un botón o enviará un comando al controlador. La experiencia industrial ha mostrado claramente que ésta es una característica altamente útil y deseable. La sintonía automática es algunas veces denominada sintonía bajo demanda o sintonía de una vez. La capacidad de auto-sintonía se puede incorporar en los controladores. Prácticamente, todos los controladores se pueden beneficiar de herramientas para sintonía automática, lo que simplificaría drásticamente su uso. Los controladores de lazo único y los sistemas distribuidos para control de procesos son importantes áreas de aplicación. La mayoría de estos controladores son del tipo PID. Actualmente, la sintonía automática es ampliamente utilizada en los controladores PID.

La auto sintonía se puede llevar a cabo con dispositivos externos que se conectan a un proceso. Como estos sistemas tienen que trabajar con controladores de diferentes fabricantes, se les deben proporcionar con información sobre la estructura del controlador para dar un sugerencia apropiada sobre cada parámetro. Tal información incluye la estructura del controlador (estándar, en serie, o en paralelo), período de muestreo, constantes de tiempo del filtro, y unidades de los distintos parámetros del controlador (ganancia o banda proporcional, minutos o segundos, tiempo o repetición/tiempo).

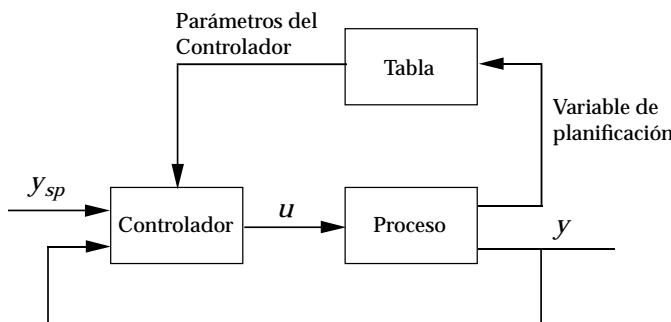


Figura 9.2 Diagrama de bloques de un sistema de planificación de ganancia.

Planificación de ganancia

La planificación de ganancia es una técnica que trata con procesos no lineales, procesos con variaciones en el tiempo o situaciones donde los requisitos sobre el control cambian con las condiciones de operación. Para utilizar la técnica, es necesario encontrar variables medibles, denominadas variables de planificación, que se correlacionan bien con los cambios en la dinámica del proceso. La variable de planificación puede ser, por ejemplo, la señal medible, la señal de control o una señal externa. Por razones históricas, la frase *planificación de ganancia* se utiliza incluso si se cambian otros parámetros que no sean la ganancia, por ejemplo, el tiempo derivativo o el tiempo integral. La planificación de ganancia es una manera muy efectiva de controlar sistemas cuya dinámica cambia con las condiciones de operación. La planificación de ganancia no ha sido utilizada mucho a causa del esfuerzo que requiere su implementación. Sin embargo, cuando se combina con la auto-sintonía, la planificación de ganancia es muy sencilla de utilizar.

La Figura 9.2 muestra el diagrama de bloques de un sistema con planificación de ganancia. El sistema se puede interpretar como un sistema con dos lazos. Existe un lazo interno, compuesto por el proceso y el controlador, y un lazo externo que ajusta los parámetros del controlador basándose en condiciones de operación. Además, existen situaciones donde no hay lazo externo y la variable de planificación no está afectada por la salida del controlador.

La noción de planificación de ganancia fue utilizada originalmente para sistemas de control de vuelo, pero se está utilizando de forma creciente en el área de control de procesos. De hecho, es un elemento estándar en cualquier lazo simple de controladores PID. Para aplicaciones de control de procesos, se pueden obtener mejoras significativas haciendo uso de un pequeño conjunto de parámetros del controlador.

La planificación de ganancia es generalmente una alternativa a la adaptación. Tiene la ventaja de que puede seguir cambios rápidos en las condiciones de operación. El problema clave consiste en encontrar las variables de planificación. Algunas elecciones posibles son la señal de control, la variable del proceso o una señal externa. La tasa de producción es generalmente una buena elección en aplicaciones de control de procesos, ya que las constantes de tiempo y los retardos de tiempo son por lo general inversamente proporcionales a la tasa de producción.

El desarrollo de una planificación de ganancia puede tomar un esfuerzo de in-

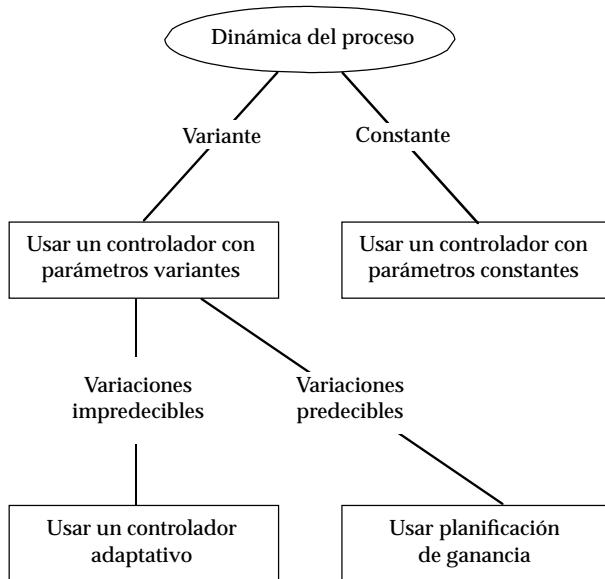


Figura 9.3 Cuándo utilizar las diferentes técnicas adaptativas.

geniería considerable. La disponibilidad de un mecanismo de sintonía automática puede reducir significativamente el esfuerzo ya que el ajuste puede entonces determinarse experimentalmente. Primero se determina una variable de planificación. Su rango se cuantifica en un número de condiciones de operación discretas. Entonces, se determinan los parámetros del controlador mediante sintonía automática cuando el sistema está funcionando en un punto de operación. Los parámetros se almacenan en una tabla. El procedimiento se repite hasta que se cubren todos los puntos de operación. De esta forma, es sencillo instalar un sistema de planificación de ganancia en un sistema de control por computador mediante la programación de una tabla que almacena los parámetros del controlador así como los mecanismos de actualización.

Uso de técnicas adaptativas

Hemos descrito tres técnicas que son útiles para tratar con procesos que tienen propiedades cambiantes con el tiempo o que presentan distintas condiciones de operación. En la Figura 9.3 se muestra un diagrama para guiar en la elección sobre las distintas técnicas adaptativas.

El rendimiento de los controladores es la primera cosa a considerar. Si los requisitos son modestos, se puede utilizar un controlador con parámetros constantes y sintonía conservadora. Para demandas de rendimiento más altas, se deberían considerar otras soluciones. Si la dinámica del proceso es constante, se debería utilizar un controlador con parámetros constantes. Los parámetros del controlador se pueden obtener utilizando auto-sintonía.

Si la dinámica del proceso o la naturaleza de las perturbaciones son cambiantes, es sencillo compensar estos cambios cambiando el controlador. Si las variaciones se pueden predecir a partir de la señales medidas, se debería utilizar

una planificación de ganancia por su simplicidad y por dar un rendimiento superior y más robusto que las técnicas de adaptación continua. Ejemplos típicos son las variaciones causadas debidas a no-linealidades en el lazo de control. Se puede utilizar la auto sintonía para construir la tabla de planificación de ganancia.

Existen también casos donde las variaciones en la dinámica del proceso no son predecibles. Ejemplos típicos son cambios debidos a variaciones no medibles en las materias primas, desgastes, contaminación, etc. Estas variaciones no pueden manejarse por planificación de ganancia ya que no se dispone de la variable de planificación, pero se pueden tratar mediante adaptación. Generalmente se utiliza un procedimiento de auto sintonía para inicializar el controlador adaptativo. Algunas veces se denomina pre-sintonía o sintonía inicial.

El control feedforward merece una mención especial. Es un método muy potente para tratar con perturbaciones medibles. Sin embargo, el uso del control feedforward requiere buenos modelos de la dinámica del proceso. Es difícil sintonizar automáticamente los lazos del control feedforward bajo demanda debido a que el operador generalmente no puede manipular la perturbación utilizada para el controlador feedforward. Para sintonizar el controlador feedforward es necesario esperar a una perturbación apropiada. Por consiguiente, la adaptación es particularmente útil para el control feedforward.

9.4 Métodos basados en modelo

Esta sección presenta un resumen de los enfoques de sintonía automática que se basan en una derivación explícita de un modelo del proceso. Los modelos se pueden obtener de varias formas tal y como se mostró en el Capítulo 2. En esta sección, describimos enfoques basados en respuestas transitorias, respuesta en frecuencia y estimación de parámetros. Los métodos se pueden caracterizar con términos de métodos en lazo abierto o lazo cerrado.

Métodos de respuesta transitoria

Los auto sintonizadores se pueden basar en el análisis de la respuesta transitoria en lazo abierto o lazo cerrado. En la Sección 2.7 se analizaron los métodos para determinar la respuesta transitoria. La mayoría de los métodos más usados se basan en respuestas a escalón o a pulso, pero existen métodos que pueden utilizar muchos otros tipos de perturbaciones.

Sintonía en lazo abierto Se puede obtener un modelo de proceso sencillo a partir de un experimento de la respuesta transitoria en lazo abierto. Se aplica un escalón o un pulso a la entrada del proceso y se mide la respuesta. Para llevar a cabo tal experimento, el proceso debe ser estable. Si se utiliza un ensayo basado en pulso, el proceso puede incluir un integrador. Es importante que el proceso esté en equilibrio cuando comience el experimento.

Sólo existen uno o dos parámetros que deben establecerse *a priori*, concretamente, la amplitud y la duración de la señal. La amplitud debería escogerse suficientemente grande de forma que la respuesta sea fácilmente visible por encima del nivel de ruido. Por otro lado, debería ser tan pequeña como sea posible

para no perturbar el proceso más de lo necesario y mantener la dinámica lineal. El nivel de ruido se puede determinar automáticamente al inicio del experimento de sintonía. Por consiguiente, debe ser posible decidir la magnitud por parte del operador.

La duración del experimento es el segundo parámetro que normalmente se estable *a priori*. Si la respuesta no es conocida, es muy difícil determinar si una respuesta a escalón se ha asentado o no. Un enfoque intuitivo consiste en decir que la señal de medida ha alcanzado su nuevo estado estacionario si la velocidad de cambio es suficientemente pequeña. Sin embargo, esta velocidad de cambio está relacionada con las constantes del tiempo del proceso que son desconocidas. Si se utilizar un ensayo de pulso, su duración debería estar también relacionada con las constantes de tiempo del proceso.

Se pueden utilizar muchos métodos para extraer las características del proceso a partir de un experimento basado en la respuesta transitoria. La mayoría de los auto sintonizadores determinan la ganancia estática, la constante de tiempo aparente y el retardo de tiempo aparente. La ganancia estática es sencilla de encontrar de manera exacta a partir de un experimento de respuesta a escalón comparando los valores estacionarios de la señal de control y la señal de medida antes y después del cambio en escalón. La constante de tiempo y el tiempo de retardo se puede obtener de varias formas, véase la Sección 2.7.

Los métodos de respuesta transitoria se utilizan generalmente en modo de pre-sintonía en dispositivos de sintonía más complicados. Se explota la principal ventaja de estos métodos, se requiere poco conocimiento previo del proceso. También es sencillo explicar los métodos al personal de planta. Sin embargo, el principal inconveniente es que son sensibles a perturbaciones. Este problema es menos importante si se utiliza sólo para la fase de pre-sintonía.

Sintonía en lazo cerrado La sintonía automática basada en la identificación de la respuesta transitoria se puede llevar a cabo en lazo cerrado. Los escalones o los pulsos se añaden entonces en el punto de consigna o en la señal de control. También existen auto sintonizadores que no introducen ninguna perturbación transitoria. En lugar de esto se utilizan las perturbaciones causadas por los cambios en el punto de consigna o las perturbaciones de carga. En estos casos es necesario detectar que las perturbaciones son suficientemente grandes comparadas con el nivel de ruido.

Los métodos de sintonía en lazo cerrado no se pueden utilizar sobre procesos de los que no se posee información. Alguna clase de pre sintonía se debe llevar a cabo siempre para cerrar el lazo de una manera satisfactoria. Por otro lado, usualmente no se suele requerir ninguna información adicional *a priori*. La magnitud de los cambios en escalón en el punto de consigna se determinan de manera sencilla a partir del cambio deseado o aceptado en la señal de medida.

Como el objetivo del diseño es una respuesta transitoria adecuada en lazo cerrado, es razonable basar la sintonía sobre las respuestas en lazo cerrado. Es sencillo dar especificaciones de diseño en términos de las respuesta transitoria en lazo cerrado, por ejemplo, amortiguamiento, sobreoscilación, constantes de tiempo en lazo cerrado, etc. El inconveniente es que la relación entre estas especificaciones y los parámetros del PID es normalmente bastante complicada. Por lo tanto, se requiere cierta lógica o heurística.

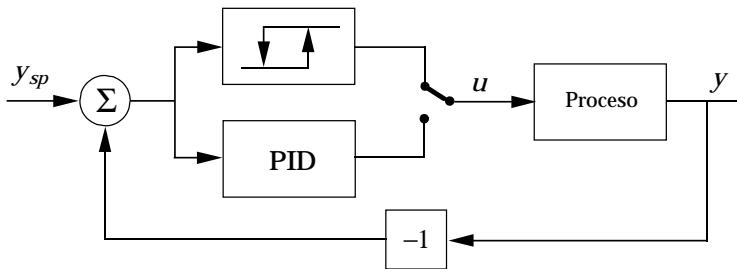


Figura 9.4 El auto-sintonizador de relé. En el modo sintonía, el proceso se conecta a la realimentación de relé.

Métodos de respuesta en frecuencia

Existen también auto-sintonizadores basados en métodos de respuesta en frecuencia. En la Sección 2.7, se mostró cómo las técnicas de respuesta en frecuencia podrían utilizarse para determinar la dinámica del proceso.

Uso del método del relé En los métodos tradicionales de respuesta en frecuencia, la función de transferencia de un proceso se determina midiendo las respuestas en estado estacionario a entradas sinusoidales. Una dificultad de este enfoque es que las frecuencias apropiadas de la señal de entrada se deben escoger *a priori*. En la Sección 2.7, se describió un método especial donde se generaba automáticamente la frecuencia apropiada de la señal de entrada. La idea fue simplemente introducir una realimentación no lineal de tipo relé para generar una oscilación de ciclo límite. Con un relé ideal, el método genera una señal de entrada al proceso con un período cercano a la frecuencia última del sistema en lazo abierto.

En la Figura 9.4 se muestra un diagrama de bloques de un auto-sintonizador basado en el método del relé. Obsérvese que existe una comutación que selecciona entre la realimentación del relé o la realimentación ordinaria con el PID. Cuando se desea sintonizar el sistema, la función PID se desconecta y el sistema se conecta al control realimentado con relé. Entonces, el sistema comienza a oscilar. El período y la amplitud de la oscilación se determinan cuando se alcanza una oscilación en estado estacionario, obteniéndose el período último y la ganancia última. Se pueden determinar entonces los parámetros del controlador PID a partir de estos valores. El controlador PID entonces se comuta automáticamente otra vez y el control se ejecuta con los nuevos parámetros del PID.

Se debe especificar por adelantado la amplitud inicial del relé. Se puede utilizar un lazo de realimentación desde la medida de la amplitud de la oscilación a la amplitud del relé para asegurar que la salida está dentro de unos límites razonables durante la oscilación. En el relé es sencillo introducir histéresis, lo que reduce el efecto del ruido de medida e incrementa también el período de la oscilación. Con la histéresis existe un parámetro adicional que se puede ajustar automáticamente basado en la determinación del nivel de ruido de medida. Obsérvese que no es necesario conocer *a priori* las escalas de tiempo ya que la frecuencia última se determina automáticamente a través del experimento.

En el método del relé, se genera una oscilación con la frecuencia apropiada mediante una no linealidad estática. Incluso, el orden de magnitud de la cons-

tante de tiempo del proceso no tiene por qué ser conocida. Por consiguiente, este método no es sólo apropiado como un dispositivo de sintonía, sino que se puede utilizar también en pre sintonía. También puede ser apropiado para la determinación de los tiempos de muestreo en controladores digitales.

El método de sintonía del relé puede además modificarse para identificar varios puntos en la curva de Nyquist. Esto se puede llevar a cabo realizando varios experimentos con distintos valores de amplitud y de histéresis en el relé. También se puede incluir un filtro con características conocidas en el lazo para identificar otros puntos del diagrama de Nyquist.

Métodos en Línea El análisis de la respuesta en frecuencia también se puede utilizar para la sintonía en línea de controladores PID. Introduciendo un filtro paso-banda, se puede investigar el contenido de la señal a distintas frecuencias. Utilizando este conocimiento, se puede identificar el modelo de un proceso en términos de puntos del diagrama de Nyquist y seguir en línea. En este auto sintonizador, es crucial la elección de las frecuencias en los filtros paso-banda. Esta elección se puede simplificar utilizando el procedimiento de sintonía descrito anteriormente en la fase de pre sintonía.

Métodos de estimación de parámetros

Un procedimiento de sintonía común consiste en utilizar estimación de parámetros recursiva para determinar un modelo de bajo orden en tiempo discreto del proceso. Los parámetros del modelo de bajo orden resultante se utilizan entonces en un esquema de diseño para calcular los parámetros del controlador. Un auto sintonizador de este tipo puede además operar como un controlador adaptativo que cambia los parámetros del controlador continuamente. Por lo tanto, los auto sintonizadores basados en esta idea poseen generalmente una opción para adaptación continua.

La principal ventaja de los auto sintonizadores de este tipo es que no requieren ningún tipo específico de señal de excitación. La señal de control puede ser, por ejemplo, una secuencia de cambios manuales de la señal de control o las señales obtenidas durante una operación normal. Un inconveniente con este tipo de auto sintonizadores es que requieren información previa. Se debe especificar el tiempo de muestreo para el procedimiento de identificación, que debería estar relacionado con las constantes de tiempo del sistema en lazo cerrado. Como la identificación se realiza en línea, se requiere un controlador que al menos sea capaz de estabilizar el sistema. Los sistemas basados en este procedimiento de identificación necesitan una fase de pre sintonía, que puede basarse en los métodos presentados anteriormente en esta sección.

9.5 Métodos basados en reglas

Esta sección trata métodos de sintonía automática que no utilizan un modelo explícito del modelo. En este caso la sintonía se basa en la idea de imitar la sintonía manual mediante la experiencia del ingeniero del proceso.

La sintonía del controlador es un compromiso entre rendimiento y robustez. La Tabla 9.1 muestra cómo la estabilidad y la velocidad de respuesta cambian en

Cuadro 9.1 Reglas empíricas de los efectos de los parámetros del controlador sobre la velocidad y la estabilidad del lazo cerrado.

	Velocidad	Estabilidad
K aumenta	aumenta	decrece
T_i aumenta	decrece	aumenta
T_d aumenta	aumenta	aumenta

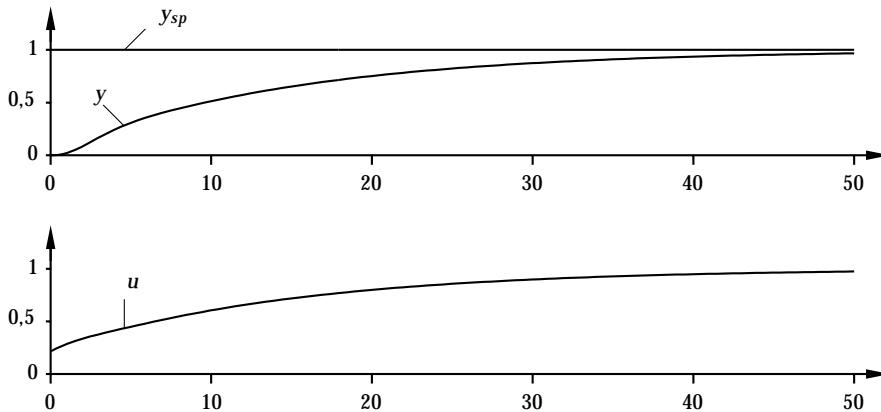


Figura 9.5 Respuesta al punto de consigna donde una regla correcta consiste en incrementar la ganancia y decrementar el tiempo integral. El diagrama superior muestra el punto de consigna y_{sp} y la salida del proceso y y el inferior la señal de control u .

la medida que los parámetros del controlador PID varían. Obsérvese que la tabla sólo contiene reglas heurísticas. Sin embargo, existen excepciones. Por ejemplo, un incremento de la ganancia generalmente resulta en un control más estable cuando el proceso contiene un integrador. Estas mismas reglas se pueden mostrar en los mapas de sintonía. Ver, por ejemplo, el mapa de sintonía para control PI de la Figura 6.7.

Los procedimientos de sintonía automática basados en reglas esperan a respuestas transitorias, cambios en escalón o perturbaciones de carga de la misma forma que los métodos basados en modelo. Cuando se produce tal perturbación, se observa el comportamiento del proceso controlado. Si el control se desvía de las especificaciones, los parámetros del controlador se ajustan en base a algunas reglas.

Las Figuras 9.5 y 9.6 muestran cambios en el punto de consigna de los lazos de control con un controlador PI mal sintonizado. La respuesta mostrada en la Figura 9.5 es muy lenta. En este caso, una regla correcta es incrementar la ganancia y decrementar el tiempo integral. La Figura 9.6 también muestra una respuesta lenta a causa de un tiempo integral demasiado grande. La respuesta es además oscilatoria debido a una excesiva ganancia. Por tanto, una regla correcta consiste en decrementar ambas, la ganancia y el tiempo integral.

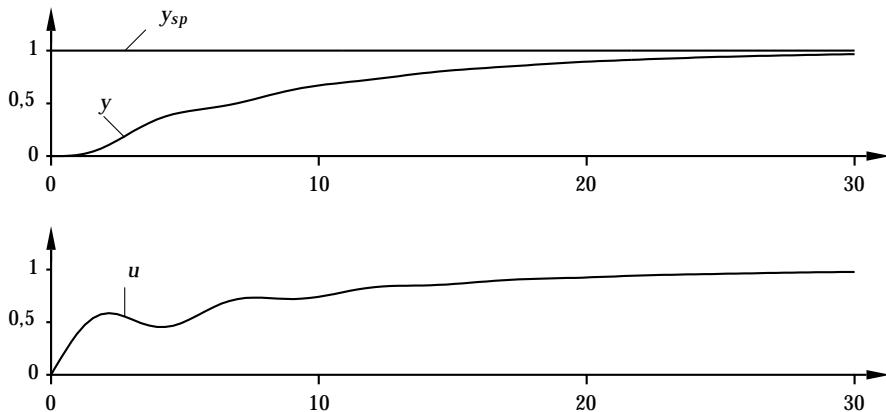


Figura 9.6 Respuesta en el punto de consigna donde una regla correcta es disminuir la ganancia y el tiempo integral. El diagrama superior muestra el punto de consigna y_{sp} y la salida del proceso y y el inferior la señal de control u .

Si se obtienen gráficas como las de las Figuras 9.5 y 9.6, para un operador experimentado es sencillo aplicar reglas correctas de sintonía. Para obtener un procedimiento de sintonía automática basado en reglas, las gráficas deben ser sustituidas por cantidades que caractericen las respuestas. Comúnmente, se utilizan cantidades de sobreoscilación y razón de decaimiento para caracterizar la estabilidad del lazo de control y la constante de tiempo y la frecuencia de oscilación para caracterizar la velocidad del lazo.

Es relativamente fácil obtener reglas relevantes para decir si los diferentes parámetros del controlador deberían ser aumentados o disminuidos. Sin embargo, es muy complicado determinar *cuánto* se deberían aumentar o disminuir. Por tanto, los métodos basados en reglas son más apropiados para adaptación continua donde se realizan pequeños cambios sucesivos en los parámetros del controlador después de cada transitorio.

Los métodos basados en reglas tienen una gran ventaja en comparación con los enfoques basados en modelos que utilizan adaptación continua. Los basados en reglas manejan perturbaciones de carga eficientemente y a la misma vez cambios en el punto de consigna. Los enfoques basados en modelo son más apropiados para cambios en el punto de consigna. Sin embargo, cuando se produce una perturbación de carga, la respuesta transitoria se produce debido a una señal de entrada desconocida. En estos casos no es sencillo obtener un modelo de entrada-salida bajo tales circunstancias.

Un inconveniente con los enfoques basados en reglas es que normalmente asumen que los cambios en el punto de consigna o las perturbaciones de carga son escalones o impulsos aislados. Si se producen dos cambios en el punto de consigna o en la perturbación de carga en un espacio corto de tiempo puede resultar en una salida del proceso que da lugar a una regla de sintonía del controlador errónea.

9.6 Supervisión de controladores adaptativos

La sintonía automática y la planificación de ganancia han sido bien aceptadas por la industria de procesos y en la actualidad es común encontrar ambas en los controladores individuales y en los sistemas de control distribuido. Existen numerosos auto sintonizadores bien diseñados que son sencillos de utilizar. Sin embargo, el uso industrial del controlador adaptativo «verdadero» está limitado, existiendo varias razones para ello. Una es que muchos de los controladores que han sido probados industrialmente no han sido suficientemente robustos. Esto ha empañado la técnica con un tanto de mala reputación. Los algoritmos adaptativos se deben proporcionar con un sistema supervisor que se encarga de tener cuidado de aquellas condiciones de operación para las cuales el algoritmo no fue diseñado.

El problema no es sólo para los controladores adaptativos. *Cada* controlador necesita una capa de supervisión. El controlador PID simple, por ejemplo, tiene protección de antiwindup para tratar con la situación de saturación en la señal de control, funciones para la transferencia sin salto para la conmutación de modos entre control manual y automático, funciones para la transferencia sin salto frente al cambio de parámetros, y algunas veces limitaciones de zonas muerta y la velocidad en la señal de control. Esta sección analiza algunas funciones de supervisión para controladores adaptativos.

Inicialización

El primer tópico a tener en cuenta es la inicialización del controlador adaptativo. La inicialización debería asegurar que cuando se inicie la adaptación se utilicen parámetros adecuados del controlador. Un controlador adaptativo requiere también parámetros adicionales que se deberían obtener en la fase de inicialización. Por ejemplo, los controladores adaptativos, tanto los basados en modelo como los basados en reglas, necesitan conocer la escala de tiempo del proceso. Esto se utiliza para establecer los períodos de muestreo y las constantes de tiempo.

En controladores adaptativos de propósito especial, la inicialización se puede llevar a cabo manualmente mediante un usuario experimentado. Sin embargo, en controladores adaptativos multipropósito, esta fase no se debería dejar para usuarios sin experiencia. Se debería llevar a cabo automáticamente. Sin embargo, casi todos los controladores adaptativos industriales multipropósito tienen alguna función de sintonía automática o pre sintonía que inicializa el controlador adaptativo. Estos procedimientos pueden estar basados en experimentos de respuesta a escalón, que proporcionan la escala de tiempo del proceso en términos de retardo de tiempo aparente y la constante de tiempo aparente. También se pueden basar en experimentos de realimentación del relé, donde en tales casos la escala de tiempo del proceso se obtiene en términos de la frecuencia última ω_u .

En lo que sigue, se asume que la escala de tiempo del proceso ha sido obtenida y está disponible en el controlador adaptativo, siendo denotada por T_p . También se asume que los cálculos de diseño se realizarán de tal manera que T_p sea además proporcional a la constante de tiempo en lazo cerrado. El procedimiento de inicialización no es sólo invocado una vez que se instala el controlador adaptativo. Parte del procedimiento de inicialización tiene que utilizarse durante las transiciones de modo y también en los cambios de parámetros.

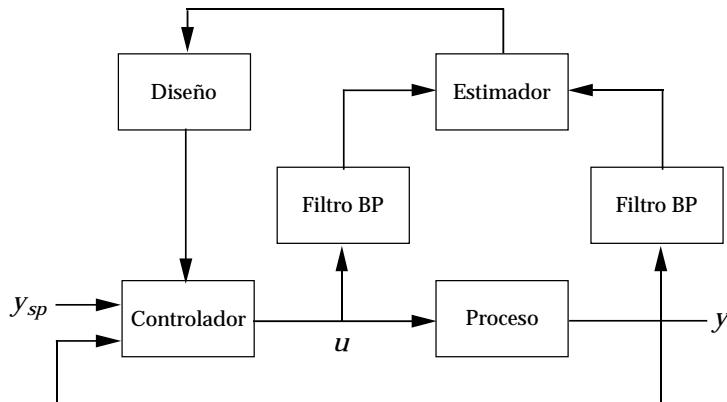


Figura 9.7 Filtrado de la señal de control y la señal medida.

Una característica importante del procedimiento de inicialización es obtener constantes de tiempo apropiadas para los filtros. Las componentes de bajas frecuencias de las señales deberían reducirse para eliminar los términos de desviación. Los componentes de altas frecuencias están normalmente contaminados por ruido de medida que perturba al estimador de parámetros. Por tanto, la señal de control y la señal de medida deberían filtrarse por un filtro paso-banda antes de llegar al estimador de parámetros. Véase la Figura 9.7.

Es importante que la estimación de parámetros se base en datos relevantes. Si el orden del modelo es bajo, es particularmente importante que el modelo se ajuste a los datos en la región de frecuencia apropiada para el diseño del controlador, a saber, el rango de frecuencia alrededor de la frecuencia última. Este rango de frecuencia se determina mediante la elección de las constantes de tiempo de los filtros paso-banda. El rango de frecuencia se puede establecer estrecho o ancho, dependiendo del objetivo de control y del orden del modelo estimado. En el controlador ECA600, véase Sección 9.8, se utiliza un controlador paso-banda de rango estrecho y se identifica un modelo de sólo dos parámetros. Los modelos que tienen más parámetros requieren filtros con rango de frecuencia más amplio.

Detección de la excitación

El estimador de parámetros es la parte central del controlador adaptativo. Generalmente se utiliza un estimador basado en mínimos cuadrados recursivos. Esto se puede describir por

$$\begin{aligned}
 \hat{\theta}(t) &= \hat{\theta}(t-1) + P(t)\varphi(t)\epsilon(t) \\
 \epsilon(t) &= y(t) - \varphi(t)^T\hat{\theta}(t-1) \\
 P(t) &= P(t-1) - \frac{P(t-1)\varphi(t)\varphi(t)^TP(t-1)}{1 + \varphi(t)^TP(t-1)\varphi(t)},
 \end{aligned} \tag{9.1}$$

donde $\hat{\theta}$ es el parámetro estimado, P es la matriz de covarianza y φ es el vector de regresión, que normalmente contiene las medidas y las señales de control

retrasadas. Para conseguir el seguimiento de variaciones en la dinámica del proceso, es necesario tener más en cuenta los datos recientes que los antiguos. Esto generalmente se asegura introduciendo un factor de olvido λ y modificando la matriz de covarianza acorde a

$$P(t) = \frac{1}{\lambda} \left(P(t-1) - \frac{P(t-1)\varphi(t)\varphi(t)^T P(t-1)}{\lambda + \varphi(t)^T P(t-1) \varphi(t)} \right). \quad (9.2)$$

Un factor del olvido en el rango $0 < \lambda < 1$ impide que la matriz de covarianza converja a cero. La elección de λ es un compromiso entre la velocidad de adaptación y la velocidad de robustez. Por ejemplo, disminuyendo λ resultará un incremento en la frecuencia de adaptación pero disminuirá además la robustez. La introducción de un factor de olvido puede causar problemas si la excitación no es suficientemente buena. Supóngase por ejemplo que φ es cero durante un cierto período. De la ecuación 9.2 se deduce entonces que la matriz de covarianza aumentará exponencialmente. Existen varias formas de superar este problema, como por ejemplo, utilizando un factor de olvido variable o utilizando un olvido direccional. Se ha propuesto también reinicializar la matriz de covarianza periódicamente para asegurar que P permanece dentro de ciertos límites. Esto resolverá seguramente el problema numérico pero de tal manera que la incertidumbre de la estimación variará periódicamente, algo que no es conveniente.

El problema de la excitación no es sólo un problema numérico. El problema consiste además en asegurar que al estimador de parámetros se le proporciona con suficiente información relevante para producir un modelo del proceso fiable. En principio hay dos soluciones a este problema:

1. Asegurar que la excitación está siempre presente mediante la inclusión de señales de excitación a la entrada del proceso.
2. Asegurar que la estimación se lleva a cabo sólo cuando hay suficiente excitación natural del proceso.

El primer enfoque puede parecer atractivo. Una señal de excitación que es tan pequeña que es difícilmente perceptible comparada con el ruido de medida normal no hará mucho daño. Desafortunadamente, tal excitación no es de mucha ayuda para la estimación de parámetros. La señal de excitación debe tener una amplitud significativa para que sea de utilidad. La fricción u otras no linealidades pueden, de lo contrario, distorsionar o incluso eliminar la respuesta de la salida del proceso perdiéndose la excitación. Una señal de excitación con una amplitud significativa causa degradación del control y puede por lo tanto ser solamente aceptada durante períodos cortos tal y como durante un experimento de sintonía automática. Por estas razones, el primer enfoque es rara vez utilizado en controladores industriales. En lugar de ello, se utiliza el segundo enfoque.

Para asegurar que la estimación se realiza sólo después de cambios significativos en el punto de consigna o en la perturbación de carga cuando existe suficiente excitación, se requiere un procedimiento que mida la excitación.

A continuación se describe un enfoque conveniente para la detección de la excitación similar al utilizado en el controlador ECA600. La idea básica es realizar un filtrado paso-alto de la señal medida. Cuando la magnitud de la variable filtrada excede un cierto límite, se concluye que la excitación es suficientemente alta

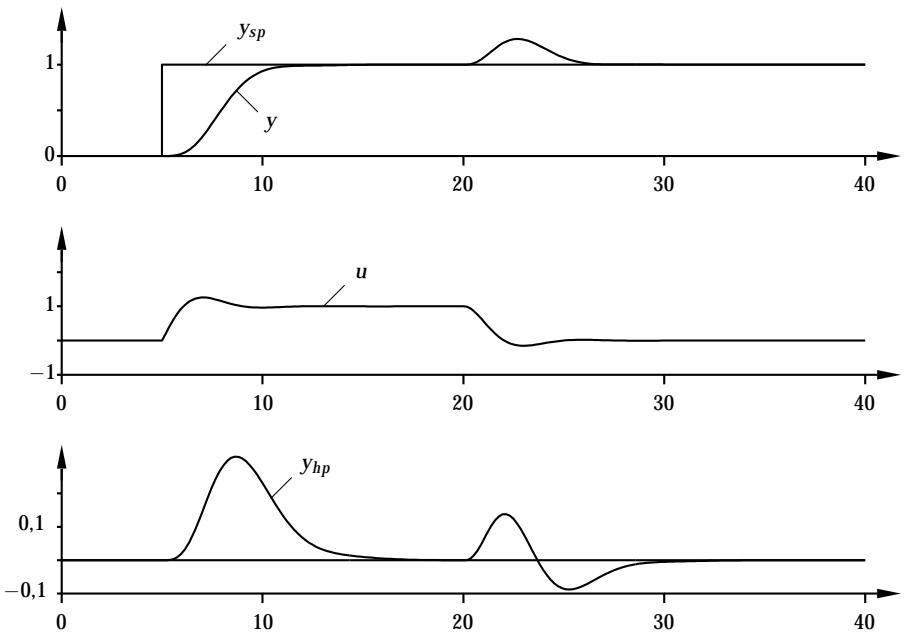


Figura 9.8 Detección de la excitación utilizando un filtrado paso-alto de la señal medida. La figura muestra las respuestas a cambios en el punto de consigna en $t = 5$ y en la perturbación de carga en $t = 20$.

para adaptación. El filtro paso-alto viene dado por

$$Y_{hp} = \frac{s}{s + \omega_{hp}} Y, \quad (9.3)$$

donde Y es la transformada de Laplace de la salida del proceso y y Y_{hp} es la correspondiente señal filtrada. El filtro posee una ganancia unitaria a altas frecuencias. La frecuencia ω_{hp} se escoge para ser inversamente proporcional a la escala de tiempo del proceso T_p .

La Figura 9.8 muestra una simulación donde la señal de medida es introducida a través del filtro paso-alto (9.3). De la figura, es obvio que la salida del filtro paso-alto es apropiada para la detección de excitación. La excitación es alta y la adaptación puede ser iniciada cuando la magnitud de $|y_{hp}|$ sea grande.

El próximo problema consiste en decidir cuándo la excitación es tan baja que la estimación debería ser interrumpida otra vez. Un enfoque consiste en permitir la adaptación mientras $|y_{hp}|$ permanezca en valores grandes. Un inconveniente con este enfoque es que existen retardos en el estimador. Esto quiere decir que incluso si $|y_{hp}|$ es pequeño, puede existir todavía excitación en las señales filtradas en el estimador de parámetros. Una solución al problema es simplemente permitir la adaptación durante un tiempo fijo después que se haya detectado la excitación.

Detección de las perturbaciones de carga

Los controladores adaptativos basados en modelo tienen problemas con las per-

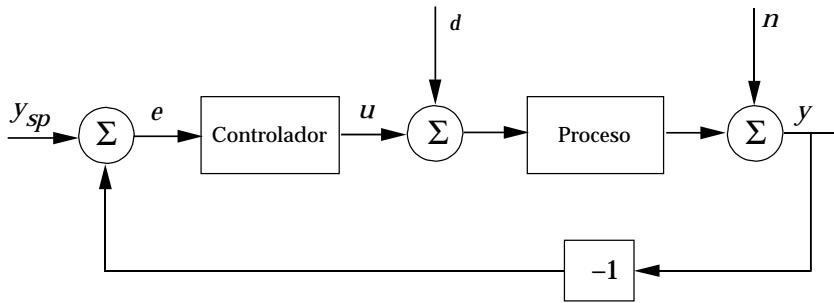


Figura 9.9 Diagrama de bloques de un lazo de realimentación simple.

turbaciones de carga. Para observar esto, considérese el diagrama de bloques de la Figura 9.9. La salida del proceso y viene dada por

$$Y(s) = P(s)(U(s) + D(s)) + N(s),$$

donde $P(s)$ es la función de transferencia del proceso, $U(s)$ es la transformada de Laplace de la señal de control u , $D(s)$ es la transformada de Laplace de la perturbación de carga d y $N(s)$ es la transformada de Laplace del ruido de medida n . Se asume que el ruido de medida sólo contiene altas frecuencias y que éstas son filtradas por filtros en el controlador. El término del ruido no es por tanto considerado a continuación. La salida del proceso se puede descomponer en dos términos,

$$y(t) = y_u(t) + y_d(t), \quad (9.4)$$

donde y_u proviene de la señal de control e y_d proviene de las perturbaciones de carga.

En (9.1), el error de predicción en el estimador basado en mínimos cuadrados viene dado por

$$\epsilon(t) = y(t) - \hat{y}(t) = y(t) - \varphi(t)^T \hat{\theta}(t-1). \quad (9.5)$$

El estimador basado en mínimos cuadrados trata de minimizar $\epsilon(t)$, por ejemplo, para hacer que la salida del proceso predicha $\hat{y}(t)$ sea igual a la salida del proceso real $y(t)$. Se asume implícitamente que

$$y(t) = y_u(t) = \varphi(t)^T \theta(t-1),$$

donde $\theta(t)$ son parámetros del proceso real. Si esta hipótesis es válida, por ejemplo, si $y(t) = y_u(t)$, los parámetros estimados $\hat{\theta}(t)$ convergerán a los valores reales $\theta(t)$, considerando que la excitación es suficiente. Sin embargo, si la salida del proceso viene dada por la Ecuación 9.4, y si $y_d(t)$ tiene componentes de frecuencia en la región de estimación, los parámetros estimados no convergerán a sus valores reales.

Este es un problema muy serio en las aplicaciones de control de procesos. En control de procesos, los cambios en el punto de consigna se realizan generalmente sólo durante los cambios de producción. (Excepciones pueden ser controladores secundarios en configuraciones en cascada). Esto quiere decir que las perturbaciones de carga generalmente son las únicas señales de excitación. Tanto para

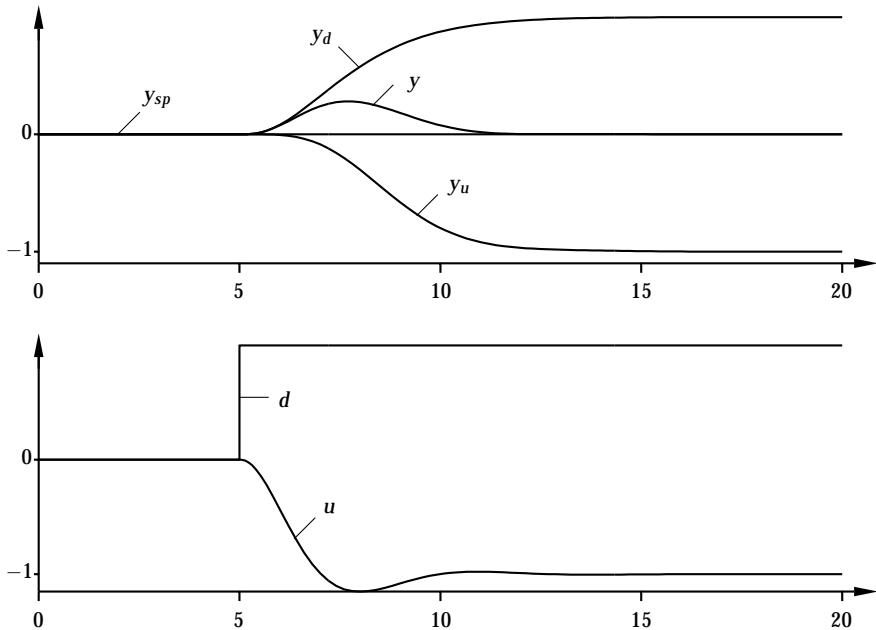


Figura 9.10 El diagrama superior muestra el punto de consigna $y_{sp} = 0$, la señal medida y , su componente de carga y_d y su componente de la señal de control y_u . El diagrama inferior muestra la perturbación de carga d y la señal de control u .

controladores basados en reglas como basados en modelo, existen posibilidades de obtener información útil considerando que las perturbaciones de carga vienen en la forma de transitorios aislados. Dicha solución se presentará a continuación.

La Figura 9.10 muestra los distintos componentes de la salida del proceso después de un cambio en escalón en la perturbación de carga. Rápidamente después del cambio de carga, $y(t) \approx y_d(t)$, por ejemplo, los cambios en el proceso están causados sólo a la perturbación de carga d . Transcurrido un rato, la contribución de la señal de control u es la componente dominante.

Una solución para el problema de identificación es evitar la adaptación durante la primera fase de la respuesta, donde y_d domina sobre y_u . La adaptación se debería iniciar en la segunda fase donde la excitación fundamental en $y(t)$ está originada por la señal de control.

Para utilizar esta solución, se necesita un procedimiento que detecta las perturbaciones de carga. Esta detección debe ser rápida, de forma que la adaptación se interrumpe tan rápido como sea posible. La detección se puede realizar de la siguiente manera. Primero, se filtra la señal de control por un filtro paso-alto de la misma forma que la señal medida en la ecuación 9.3:

$$U_{hp} = \frac{s}{s + \omega_{hp}} U. \quad (9.6)$$

La Figura 9.11 ilustra el mismo experimento que en la Figura 9.8, pero también se presenta el valor del filtrado paso-alto de la señal de control. En lo que sigue se asume que el proceso tiene una ganancia estática positiva, por ejemplo, $P(0) > 0$,

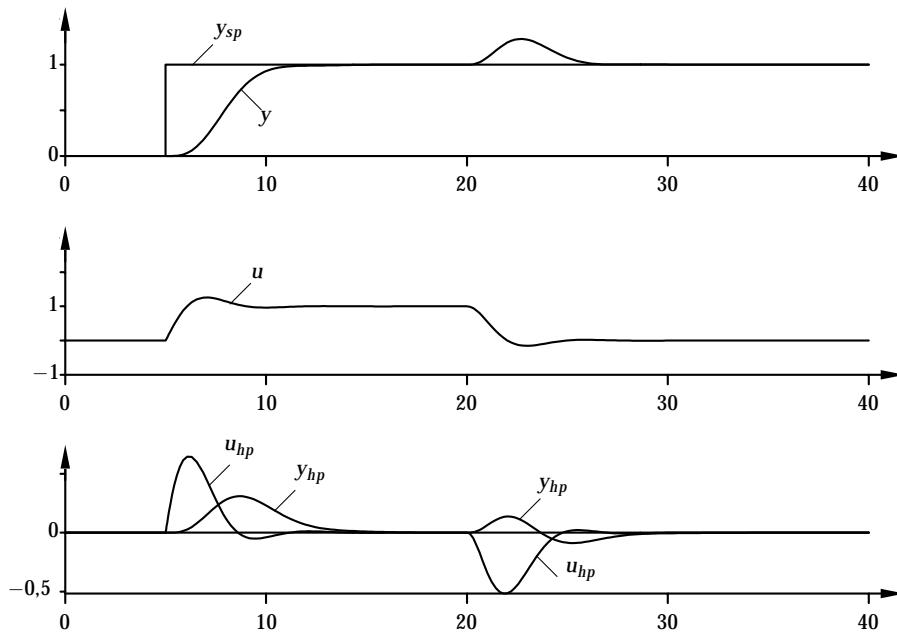


Figura 9.11 Detección de la excitación utilizando filtros paso-alto. La figura muestra las respuestas a cambios en el punto de consigna en $t = 5$ y a perturbaciones de cargar en $t = 20$.

y que todos los ceros están en el semiplano izquierdo. Después de un cambio en el punto de consigna, ambos, y_{hp} y u_{hp} , van en la misma dirección, mientras que cuando se produce una perturbación de carga van en la dirección contraria. Esta diferencia se puede utilizar para distinguir entre cambios en el punto de consigna y perturbaciones de carga. De esta manera, es posible retrasar la adaptación y evitarla durante la primera fase de la respuesta a perturbaciones de carga y llevarla a cabo sólo durante la segunda fase.

Otra manera más sencilla de evitar la adaptación durante la primera fase de una respuesta a perturbación de carga se puede obtener del error de predicción $\epsilon(t)$; véase ecuación 9.5. El error de predicción se puede escribir como

$$\begin{aligned}\epsilon(t) &= y(t) - \varphi(t)^T \hat{\theta}(t-1) \\ &= (y_u(t) - \varphi(t)^T \hat{\theta}(t-1)) + y_d(t).\end{aligned}\quad (9.7)$$

Por consiguiente, $\epsilon(t)$ está formado por dos términos, el error de predicción real que queremos minimizar y la componente de perturbación de carga de la salida del proceso. Si asumimos que la dinámica del proceso cambia lentamente, el primer término permanecerá limitado y se pueden detectar las grandes perturbaciones de carga a través de la magnitud de $|\epsilon(t)|$. Si restringimos la adaptación a aquellos períodos donde $|\epsilon(t)|$ es pequeño, seremos capaces de seguir variaciones lentas en el proceso y evitaremos también la adaptación cuando $|y_d(t)|$ llega a ser muy grande.

Detección de oscilación

Las oscilaciones con contenido de alta frecuencia cerca de la frecuencia última forman una excitación ideal para control adaptativo si son causadas por variaciones en el punto de consigna o controladores de altas ganancias que resultan en pequeños márgenes de estabilidad. En estos casos, $y(t) = y_u(t)$.

Desafortunadamente, las oscilaciones en los lazos de control son normalmente generadas por otras fuentes. Una causa común de oscilaciones es un movimiento de deslizamiento a causa de la fricción de las válvulas. Véase Sección 10.2. Si no se toman precauciones, el controlador adaptativo interpretará estas oscilaciones como causadas por una ganancia del lazo demasiado alta. Esto quiere decir que el controlador se desintonizará. Esto ocurre para ambos controladores adaptativos, los basados en reglas y los basados en modelo. El movimiento de deslizamiento se puede modelar como una perturbación de carga.

Otra razón de las oscilaciones en los lazos de control puede ser que una oscilación de perturbación de carga externa perturbe el proceso. Esta perturbación puede, por ejemplo, estar causada por un lazo de control vecino con movimiento de deslizamiento.

En estos casos, ya no es cierto que $y(t) = y_u(t)$, sino que la componente de carga $y_d(t)$ es la dominante. Esto quiere decir que estas perturbaciones proporcionarán el estimador del proceso con desinformación de la misma forma que las perturbaciones de carga comentadas en la sección anterior. Para evitar el problema, las oscilaciones tienen que ser detectadas de la misma forma que las perturbaciones de carga, de forma que la adaptación se pueda inhibir cuando estas perturbaciones están presentes. En la Sección 10.4 se presentan estos procedimientos de detección.

Saturación de la señal

Cuando la salida del proceso se satura, ya no es cierto que $y(t) = y_u(t)$. Supóngase que la salida del proceso entra en saturación en el límite y_{limit} . La salida del proceso se puede describir entonces de la siguiente manera:

$$y(t) = y_u(t) + (y_{\text{limit}} - y_u(t)).$$

El segundo término de la parte derecha se puede interpretar como una componente de la perturbación de carga. Por lo tanto, tenemos el mismo problema que se analizó en las secciones anteriores. Por consiguiente, la estimación debería interrumpirse cuando y se satura. Otra vez, es útil disponer de un temporizador conectado a esta interrupción, tal que la estimación se desactiva durante un rato después del periodo de saturación. Esto es para evitar estimaciones erróneas durante los transitorios.

También puede ser deseable interrumpir la adaptación cuando la señal de control u satura. Esto puede parecer confuso ya que en este caso $y(t) = y_u(t)$. Sin embargo, en este caso si están presentes las perturbaciones de carga puede no llegar a ser cierto que $y_u(t)$ domine sobre $y_d(t)$ en la segunda fase de la respuesta a perturbaciones de carga.

Transiciones de modo

Un controlador de parámetros constantes opera principalmente en dos modos:

- Modo manual
- Modo automático

La transferencia sin salto entre los diferentes modos de operación se lleva a cabo asegurando que a todos los estados se les asigna valores apropiados en las transiciones. Si esto no se realiza adecuadamente, ocurrirán «saltos» en la señal de control durante las transiciones de modo. Véase la Sección 13.5.

Un controlador adaptativo tiene tres modos de operación:

- Modo manual
- Modo automático
- Modo adaptativo

Aquí, es importante asegurar transferencia sin salto también entre los dos primeros modos y el tercer modo adaptativo. La estimación de parámetros se desconecta normalmente cuando el controlador está en modo manual y generalmente también cuando está en modo automático sin adaptación. Es por consiguiente importante inicializar todos los estados adicionales que vienen dados en el estimador de parámetros cuando se inicia el controlador adaptativo.

Una inicialización errónea del estimador de parámetros resultará en «saltos» en los parámetros estimados. Estos saltos no son siempre inmediatamente visibles como un «salto» en la señal de control, pero pueden deteriorar el control de otras maneras ya que proporcionan un modelo del proceso erróneo.

Los estados más importantes del estimador de parámetros están dados por las ecuaciones 9.1 y 9.2. Se debería asignar un valor grande a la matriz de covarianza $P(t)$ cuando los parámetros estimados son inciertos. Sin embargo, cuando los parámetros del controlador se inicializan mediante un procedimiento de sintonía automática o cuando por otras razones se consideran que son precisos, $P(t)$ no se debería reinicializar con un valor grande sino mantenerse cercana a sus valor en estado estacionario.

El vector de residuos $\varphi(t)$ contiene normalmente señales de control y medida retrasadas. Este vector se debería inicializar con los valores actuales de estas señales.

También es importante que se proporcionen con estados correctos a los filtros y a las funciones de supervisión. Esto puede algunas veces llevarse a cabo introduciendo un retardo en el estimador. Supóngase, por ejemplo, que el controlador se comuta de modo manual a modo adaptativo y que la salida del proceso no está cerca del punto de consigna. Esto quiere decir que inmediatamente obtenemos un transitorio en la commutación de modo. Si el procedimiento de detección de la excitación está activo, el mecanismo de adaptación puede empezar entonces antes de que los estados hayan obtenido sus valores apropiados. Este problema se puede evitar retrasando el procedimiento de detección de la excitación en la transición de modo.

Se debe realizar también una re inicialización del controlador adaptativo si se modifican los parámetros relacionados con la adaptación. Supóngase, por ejemplo, que se modifica el período de muestro por el usuario o que el procedimiento de sintonía automática está en operación, lo que resulta en nuevos valores del período de muestreo. Esto quiere decir que se debe realizar una re inicialización

total del controlador adaptativo, con las nuevas constantes de tiempo del filtro, etc.

Otra transición de modo sucede si se combina el controlador adaptativo con una planificación de ganancia. En este caso, se debería realizar una re inicialización si existe una comutación en la planificación de ganancia.

También es generalmente posible reiniciar el controlador adaptativo, de forma que los parámetros estimados $\hat{\theta}(t)$ se reinicializan a algunos valores prefijados, normalmente aquellos obtenidos durante la fase de inicialización.

Límites de los parámetros estimados

Existe una región del espacio de parámetros donde la información suministrada durante la fase de inicialización es relevante. Dentro de esta región, está la información a priori sobre si es correcto el tiempo del proceso T_p , y si son adecuados los períodos de muestreo y las constantes de tiempo de los filtros. Si la dinámica del proceso cambia tanto que los parámetros estimados tienen a salirse de esta región, el comportamiento del controlador podría ser pobre.

Es, por consiguiente, ventajoso limitar los parámetros estimados a una región permisible. La adaptación puede continuar fuera de la región, pero el algoritmo debería ser reinicializado de forma que se obtengan nuevos parámetros apropiados para la nueva región. Utilizando la planificación de ganancia es posible, por ejemplo, tener algunas regiones con diferentes períodos de muestreo y constantes de tiempo del filtro.

Puede ser difícil encontrar tales regiones si el modelo estimado es de orden elevado. Es más sencillo cuando el orden del modelo es más bajo y quizás es posible encontrar interpretaciones físicas de los parámetros. En los controladores PID adaptativos, hay generalmente límites sobre la ganancia, el tiempo integral y el tiempo derivativo.

Existe otra razón para acotar los parámetros estimados, que se relaciona con la excitación necesaria para la estimación de parámetros. Supóngase, por ejemplo, que los parámetros estimados cambian tanto que se obtiene un ancho de banda en lazo cerrado muy pequeño. La excitación en la banda de frecuencia de interés será entonces baja y obtendremos una adaptación muy lenta.

Puede ser ventajoso tener límites sobre la velocidad de los cambios de las estimaciones. Esto se hace para disminuir los efectos de señales imprevistas u otros errores. Esta característica se puede comparar con los limitadores de velocidad que generalmente se utilizan en controladores estándar.

9.7 Sintonía de realimentación iterativa

La sintonía de realimentación iterativa, IFT, es un método iterativo en línea para el ajuste de los parámetros del controlador. La idea clave es una manera inteligente de calcular el gradiente del error del controlador con respecto a los parámetros del controlador.

Sea un sistema estándar con realimentación del error. Supóngase también que

se desea minimizar la función de coste

$$J = \int_0^T f(y(t), u(t)) dt$$

para un controlador PID con la parametrización

$$C(s) = k + \frac{k_i}{s} + k_d s.$$

Para minimizar el criterio, es útil conocer el gradiente de la función de coste con respecto a los parámetros del controlador. La derivada parcial de J con respecto a la ganancia del controlador k viene dada por

$$\frac{\partial J}{\partial k} = \int_0^T \left(\frac{\partial f(y(t), u(t))}{\partial y} \frac{\partial y}{\partial k} + \frac{\partial f(y(t), u(t))}{\partial u} \frac{\partial u}{\partial k} \right) dt. \quad (9.8)$$

Para evaluar la parte derecha necesitamos las derivadas parciales

$$y_k = \frac{\partial y}{\partial k}, \quad u_k = \frac{\partial u}{\partial k}.$$

Éstas se pueden calcular convenientemente a partir de las transformadas de Laplace. Tenemos que

$$\begin{aligned} Y_k &= \frac{\partial Y}{\partial C} \frac{\partial C}{\partial k} = \frac{\partial Y}{\partial C} \\ Y_{k_i} &= \frac{\partial Y}{\partial C} \frac{\partial C}{\partial k_i} = \frac{1}{s} \frac{\partial Y}{\partial C} \\ Y_{k_d} &= \frac{\partial Y}{\partial C} \frac{\partial C}{\partial k_d} = s \frac{\partial Y}{\partial C}. \end{aligned} \quad (9.9)$$

La salida del proceso viene dada por

$$\begin{aligned} Y &= \frac{PC}{1+PC} Y_{sp} + \frac{P}{1+PC} D + \frac{1}{1+PC} N \\ &= \left(1 - \frac{1}{1+PC}\right) Y_{sp} + \frac{P}{1+PC} D + \frac{1}{1+PC} N, \end{aligned}$$

y el error de control por

$$E = Y_{sp} - Y = \frac{1}{1+PC} Y_{sp} - \frac{P}{1+PC} D - \frac{1}{1+PC} N.$$

Utilizando esta expresión para el error encontramos

$$\begin{aligned} \frac{\partial Y}{\partial C} &= \frac{P}{(1+PC)^2} Y_{sp} - \frac{P^2}{(1+PC)^2} D - \frac{P}{(1+PC)^2} N \\ &= \frac{P}{1+PC} \left(\frac{1}{1+PC} Y_{sp} - \frac{P}{1+PC} D - \frac{1}{1+PC} N \right). \end{aligned}$$

Por lo tanto,

$$\frac{\partial Y}{\partial C} = \frac{P}{1+PC}E = \frac{1}{C}\frac{PC}{1+PC}E. \quad (9.10)$$

Las derivadas parciales de la salida con respecto a los parámetros del controlador se pueden calcular de una manera similar. Tenemos

$$\begin{aligned} U_k &= \frac{\partial U}{\partial C} \frac{\partial C}{\partial k} = \frac{\partial U}{\partial C} \\ U_{k_i} &= \frac{\partial U}{\partial C} \frac{\partial C}{\partial k_i} = \frac{1}{s} \frac{\partial U}{\partial C} \\ U_{k_d} &= \frac{\partial U}{\partial C} \frac{\partial C}{\partial k_d} = s \frac{\partial U}{\partial C}. \end{aligned} \quad (9.11)$$

Cálculos directos muestran que la derivada de la sensibilidad de la salida viene dada por

$$\frac{\partial U}{\partial C} = \frac{1}{1+PC}E. \quad (9.12)$$

Las Ecuaciones 9.10 y 9.12 se pueden utilizar para calcular las derivadas de la sensibilidad necesarias para la optimización. El error E es conocido, pero existe una dificultad porque no se conoce la función de transferencia del proceso P . Esta dificultad se puede sortear de la siguiente manera:

- Realizar un experimento y almacenar la salida y_1 y la señal del error de control e_1 .
- Realizar un segundo experimento de la misma duración donde se escoge el punto de consigna como el error de control e_1 del primer experimento. Almacenar la salida y_2 y el error de control e_2 de este experimento.

La salida y el error de control del segundo experimento vienen dados por

$$\begin{aligned} Y_2 &= \frac{PC}{1+PC}E_1 + \frac{P}{1+PC}D_2 + \frac{1}{1+PC}N_2 \\ &= \frac{1}{C}\frac{\partial Y_1}{\partial C} + \frac{P}{1+PC}D_2 + \frac{1}{1+PC}N_2 \\ E_2 &= \frac{1}{1+PC}E_1 - \frac{P}{1+PC}D_2 - \frac{1}{1+PC}N_2 \\ &= \frac{\partial U_1}{\partial C} + \frac{P}{1+PC}D_2 + \frac{1}{1+PC}N_2. \end{aligned}$$

Los términos D_2 y N_2 no están correlacionados con E_1 si los experimentos están bien separados en el tiempo. Sus efectos se pueden hacer arbitrariamente pequeños seleccionando grandes secuencias de datos. Por lo tanto,

$$\begin{aligned} \frac{\partial Y_1}{\partial C} &\approx CY_2 \\ \frac{\partial U_1}{\partial C} &\approx E_2. \end{aligned} \quad (9.13)$$

El segundo experimento, por consiguiente, da una estimación de las derivadas de la sensibilidad de la entrada y la salida con respecto a los parámetros del controlador. Combinando esto con la entrada y la salida del primer experimento podemos calcular ahora de (9.8) el gradiente de la función de coste con respecto a los parámetros del controlador desde la ecuación. Los parámetros del controlador pueden entonces ajustarse recursivamente. Resumiendo, obtenemos el siguiente algoritmo.

ALGORITMO 9.1—SINTONÍA DE REALIMENTACIÓN ITERATIVA

1. Hacer un experimento de duración fija y almacenar la salida y_1 y la señal del error de control e_1 .
2. Hacer un segundo experimento de la misma duración donde se escoje el punto de consigna como el error de control e_1 del primer experimento. Almacenar la salida y_2 y el error de control e_2 de este experimento.
3. Calcular el gradiente de la función de coste a partir de las ecuaciones 9.8, 9.9, 9.11 y 9.13.
4. Modificar los parámetros del controlador utilizando el gradiente.
5. Repetir desde 1 hasta que el gradiente sea suficientemente pequeño.

□

La misma idea se puede aplicar a un controlador con dos grados de libertad pero entonces se requiere incluir un tercer experimento. Una propiedad agradable de la sintonía de realimentación iterativa es que se puede utilizar para muchos controladores y criterios diferentes. Además, está particularmente bien adaptada para la optimización con respecto a las perturbaciones estocásticas estacionarias.

9.8 Productos comerciales

Para ilustrar cómo las técnicas adaptativas se utilizan industrialmente, presentamos algunas características de los controladores industriales. En lugar de dar una presentación exhaustiva, hemos seleccionado unos pocos productos para mostrar el amplio rango de técnicas y hemos escogido productos que poseen un buen registro de seguimiento. También hemos seleccionado productos para los cuales se han publicado descripciones razonablemente detalladas; más productos se describen en el libro [347] y en las revisiones en revistas comerciales.

Foxboro EXACTTM(760/761)

Foxboro fue una de las primeras compañías en anunciar productos que utilizaban técnicas adaptativas. El controlador monolazo Foxboro EXACTTM(760/761) que utilizaba adaptación basada en reconocimiento de patrones fue anunciado por Foxboro en octubre de 1984. El controlador fue más tarde mejorado con más características y Foxboro ha continuado ampliando los usos de la adaptación en un rango de productos incluyendo su sistema DCS Foxboro I/ATM. Las ideas se describen en [58] y [62] y los detalles sobre el sistema se encuentran en [59] y [64].

Foxboro continuó el desarrollo de la adaptación, y auto sintonía y adaptación están ahora disponibles en sus sistemas de control distribuido bajo el nombre comercial Exact MV™. En [154] se encuentra una presentación de los detalles del sistema. Se utilizan tres bloques de función, PIDA, FBTUNE y FFTUNE para implementar el controlador. PIDA es un controlador PID avanzado, FBTUNE, que maneja la sintonía de las ganancias de realimentación, tiene funciones para la presintonía y la adaptación, y FFTUNE tiene funciones para la sintonía de ganancias de la acción feedforward y de la planificación de ganancia.

Estructura del controlador Foxboro utiliza una estructura del controlador donde la acción integral se implementa con realimentación positiva alrededor de una constante de tiempo como se muestra en la Figura 3.3. Esta implementación da un controlador en forma serie, véase (3.8). También está disponible en el sistema un controlador con una estructura especial denominado $\text{PID}\tau$. Este controlador es un controlador PID donde la acción integral se implementa mediante una realimentación positiva alrededor de una constante de tiempo con un retardo de tiempo tal como se muestra en la Figura 8.10. Esta disposición da un controlador con más adelanto de fase que un controlador PID normal. Como el adelanto de fase se asocia también con una alta ganancia, es necesario proporcionar un buen filtrado si existe ruido de medida. El controlador se puede interpretar como un controlador donde se predice la salida futura con una combinación de las entradas y salidas pasadas del controlador, véase en la Sección 8.4 el análisis del controlador PPI. El controlador $\text{PID}\tau$ puede también considerarse como una forma especial de un predictor de Smith. El controlador $\text{PID}\tau$ aporta una mejora de rendimiento significativa para procesos con constante de tiempo dominante, pero requiere una sintonía cuidadosa.

Reconocimiento de patrones La adaptación basada en reconocimiento de patrones se puede ver como una automatización del procedimiento utilizado por un ingeniero de procesos experto cuando está sintonizando un controlador. La siguiente descripción sigue la presentación descrita en [59]. El error de control después de las perturbaciones del proceso se analizan y se modifican los parámetros del controlador. Si los parámetros del controlador son razonables, se obtiene una respuesta transitoria del error del tipo de la que se muestra en la Figura 9.12. Se utilizan las reglas heurísticas para detectar que una perturbación adecuada ha ocurrido y para detectar los picos e_1 , e_2 , e_3 y el período de la oscilación T_p . Se utilizan también heurísticas para cambiar los parámetros del controlador si la respuesta es sobreamortiguada. El transitorio se caracteriza cuantitativamente en términos de dos parámetros, sobreelongación (o) y amortiguamiento (D), que se definen como

$$o = \left| \frac{e_2}{e_1} \right|, \quad D = \frac{e_3 - e_2}{e_1 - e_2}, \quad (9.14)$$

donde e_1 , e_2 y e_3 son los picos de las transiciones mostradas en la Figura 9.12. Obsérvese que la definición utilizada de amortiguamiento es igual que la raíz cuadrada de la razón de decaimiento (2.49). Por lo tanto, el amortiguamiento de un cuarto de la amplitud corresponde a $D = 0,5$.

Los parámetros del controlador se ajustan utilizando reglas heurísticas para obtener los amortiguamientos y la sobreelongación deseadas. En la Sección 6.3 se

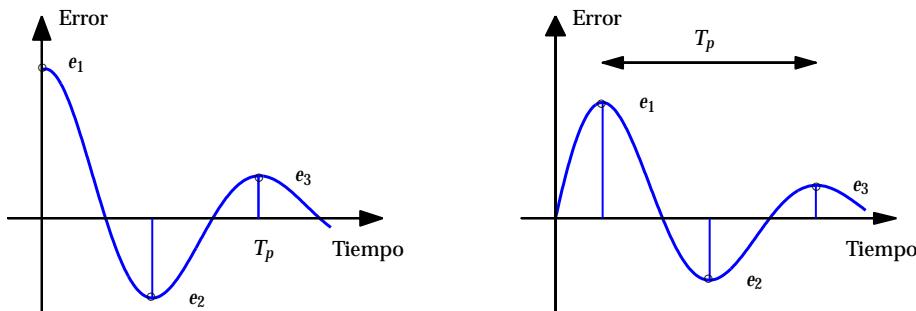


Figura 9.12 Respuesta en el error del control frente a un cambio en escalón en el punto de consigna (curva izquierda) y en las perturbaciones de carga (curva derecha).

analizan algunas reglas y se muestra el efecto de los parámetros del controlador en el mapa de sintonía de la Figura 6.7.

Pre sintonía es la denominación de Foxboro para la auto sintonía. El controlador tiene un conjunto de parámetros que deben ser dados o por el usuario a partir del conocimiento previo del lazo o estimados utilizando la función de pre sintonía.

- Valores iniciales de PB , T'_i y T'_d .
- Banda de ruido (NB). El controlador comienza la adaptación cuando la señal de error excede dos veces a NB .
- Tiempo de espera máximo (W_{\max}). El controlador espera durante un tiempo de W_{\max} para la ocurrencia de un segundo pico.

Si el usuario es incapaz de proporcionar los parámetros requeridos, se puede activar una función de pre sintonía que estima estas cantidades. Para activar la función de pre-sintonía, el controlador debe ponerse primero en modo manual. Cuando se activa la función de pre sintonía, se genera una entrada en escalón. Los parámetros del proceso, ganancia estática K_p , retardo de tiempo L y constante de tiempo T , se obtienen entonces de un simple análisis de la curva de reacción del proceso. Los parámetros del controlador se calculan utilizando un fórmula del tipo de Ziegler-Nichols:

$$PB = 120K_p L / T, \quad T'_i = 1,5L, \quad T'_d = T_i / 6. \quad (9.15)$$

El tiempo de respuesta máximo, W_{\max} , se determina también de la respuesta a un escalón durante $W_{\max} = 5L$.

La banda de ruido se determina durante la última fase del modo de pre-sintonía. La señal de control se retorna primero al nivel antes del cambio en escalón. Con el controlador todavía en modo manual y la señal de control mantenida constante, la salida se pasa a través de un filtro paso-alto. La banda de ruido se calcula como una estimación de la amplitud pico a pico de la salida del filtro paso-alto. La banda de ruido estimada se utiliza también para ajustar la acción derivativa.

Existen una serie de parámetros opcionales. Si estos no se suministran por el usuario, entonces se utilizarán los valores por defecto. Los parámetros opcionales son los que se muestran a continuación (entre paréntesis se dan los valores por defecto):

- Máximo amortiguamiento permitido (0,3)
- Máxima sobreelongación permitida (0,5)
- Factor derivativo (1). El término derivativo se multiplica por el factor derivativo. Este permite que la influencia derivativa sea ajustada por el usuario. Establecer el factor derivativo a cero resulta en un control PI.
- Límite de cambio (10). Este factor limita los parámetros del controlador a un cierto rango. Por lo tanto, el controlador no establecerá los valores de PB , T'_i y T'_d más grandes que diez veces o más pequeño que una décima que sus valores iniciales si se utiliza un valor por defecto de 10 para el límite de cambio.

La pre sintonía ha sido mejorada en los últimos productos de Foxboro. El proceso se excita con un pulso doble, véase la Figura 2.33, en lugar de un escalón como se utilizaba en el sistema original. Se ajusta un modelo FOTD o SOTD a los datos a partir del experimento, tal y como se describe en [313]. Los parámetros del controlador se calculan a partir del modelo mediante un método analítico robusto novedoso descrito en [154]. La adaptación de los parámetros de ganancia de la realimentación se realiza todavía utilizando el método de reconocimiento de patrones.

Los controladores en el sistema Foxboro DCS tienen filtros de adelanto-retraso para el control feedforward de las perturbaciones medibles. Las ganancias del controlador feedforward se sintonizan ajustando un modelo en tiempo continuo de bajo orden utilizando el método de los momentos, véase [154].

ABB

La adaptación en los sistemas de ABB tiene sus orígenes en un auto sintonizador basado en la realimentación del relé que se desarrolló primero por la compañía NAF a principios de 1980. El primer sistema fue parte de un pequeño sistema DCS (de unos 50 lazos) denominado SDM-20TM introducido en 1982 y un controlador monolazo ECA-40TM introducido en 1986. Estos sistemas utilizaban también auto sintonía para construir la planificación de ganancias. La compañía NAF sufrió una serie de traspasos, SattControl, Alfa Laval Automation, y es ahora parte de ABB. Las técnicas adaptativas se desarrollaron añadiendo adaptación continua, adaptación de las ganancias de realimentación y feedforward, y diagnosis. Todas estas características fueron introducidas en el ECA600TM, el cual fue anunciado en 1988. La tecnología es ahora una parte integral del sistema ABB DCS Industrial IT System 800xATM, que incorpora también control borroso y control predictivo basado en modelo. Existen varios tipos de controladores PID, donde las versiones avanzadas dan acceso a más parámetros. Hay también un controlador PPI para sistemas con retardo de tiempo.

Las partes esenciales de esta tecnología se describen en [40, 15, 17, 18, 19, 20, 144, 137, 1].

Flujo de datos bidireccional Los sistemas de control distribuido se programaban tradicionalmente de forma gráfica utilizando un lenguaje orientado a bloques. Un inconveniente de estos sistemas tradicionales es que el flujo de datos es unidireccional. Esto lleva a una latencia impredecible del sistema, lo que es particularmente notable en sistemas grandes cuando los cálculos pasados para evitar windup se propagan a través de algunos lazos. Una característica novedosa e interesante del sistema ABB 800xA es un estructura de datos denominada *conexión de control* que permite flujo de datos bidireccional entre los módulos de control. Esta característica hace posible implementar la protección windup de una manera elegante que evita la latencia incluso en sistemas complejos con muchos lazos en cascada.

Auto sintonía del relé La auto sintonía se realiza utilizando el método del relé estudiado en la Sección 2.7. El sintonizador se opera típicamente como sigue. El proceso se lleva a un punto de operación deseado, mediante el operador en modo manual o bien por un controlador previamente sintonizado en modo automático. Cuando el lazo está en estacionario, el operador pulsa un botón de sintonía. Despues de un corto período, donde el nivel de ruido se mide automáticamente, se introduce un relé con histéresis en el lazo, y el controlador PID se desconecta temporalmente (ver Figura 9.4). El ancho de la histéresis se establece automáticamente, basándose en el nivel de ruido que hay en el proceso. Cuanto más bajo es el nivel de ruido, más pequeña es la amplitud requerida de la señal de medida. La amplitud del relé se controla de forma que la oscilación se mantiene en un nivel mínimo por encima del nivel de ruido. Cuando se obtiene una oscilación de amplitud y período constante, se interrumpe el experimento de relé y se calcula $P(i\omega_0)$, que es el valor de la función de transferencia P a la frecuencia de oscilación ω_0 , utilizando el análisis de la función descriptiva.

Estructuras de control y diseño del controlador Algunos controladores PID y PPI están disponibles en los sistemas ABB. Las versiones avanzadas dan acceso al usuario a muchos parámetros. El algoritmo PID en el controlador ECA600TM es de la forma serie y los controladores en el sistema DCS utilizan la forma paralelo.

El procedimiento de identificación proporciona un modelo del proceso en términos de un punto $P(i\omega_0)$ sobre el diagrama de Nyquist. Hay también un ensayo para determinar si la dinámica del proceso es de constante de tiempo dominante. La frecuencia ω_0 depende de la histéresis del relé. Es típicamente menor que ω_{180} , lo que es ventajoso; véase la Sección 7.5. Al introducir el controlador PID $C(i\omega)$ en el lazo de control, es posible mover el punto correspondiente a ω_0 sobre el diagrama de Nyquist de la función de transferencia del lazo a una localización deseada. En el caso normal, el punto deseado es

$$P(i\omega_0)C(i\omega_0) = 0,5e^{-i135\pi/180}. \quad (9.16)$$

Como hay tres parámetros K , T_i y T_d , y el criterio de diseño (9.16) sólo especifica dos parámetros, se introduce la restricción adicional

$$T'_i = 4T'_d. \quad (9.17)$$

El procedimiento normal puede dar ganancias muy altas para sistemas con constante de tiempo dominante. Si esto se detecta, se utiliza un controlador PI con sintonía conservadora

$$K = 0,5 / |P(i\omega_0)|, \quad T_i = 4 / \omega_0 \quad (9.18)$$

Se puede también utilizar una sintonía diferente para procesos con retardo dominante. Entonces se puede utilizar un controlador PI con la regla de sintonía

$$K = 0,25 / |P(i\omega_0)|, \quad T_i = 1,6 / \omega_0 \quad (9.19)$$

En las versiones iniciales del controlador (ECA-40TM y ECA-600TM), el usuario puede influir sobre la sintonía seleccionando normal/PI/retardo de tiempo. Las últimas versiones del controlador obtienen más información sobre el proceso haciendo un cambio en escalón después del experimento del relé. Esto permite obtener la ganancia estática del proceso y la razón de ganancia κ , pudiendo mejorarse la sintonía sin intervención del usuario. En el sistema ABB 800xATM, esto se puede llevar a cabo utilizando las reglas de sintonía presentadas en [19]. Se pueden obtener más mejoras utilizando los resultados del Capítulo 7.

Planificación de ganancia La planificación ganancia se introdujo en los primeros controladores SDM-20TM y ECA-40TM. Fue muy sencillo construir las planificaciones utilizando auto sintonía y esta característica fue bien recibida por los usuarios. La planificación de ganancia se hizo por lo tanto una característica estándar de casi todos los controladores. Los usuarios pueden seleccionar la variable de planificación como la señal de control, la salida del proceso medida o una señal externa. Es importante que las variables de planificación no cambien demasiado rápido, por lo que se utilizan el filtrado y las histéresis para señales como la salida del proceso que pueden cambiar rápidamente. En los primeros sistemas había tres conjuntos de parámetros disponibles, pero en las últimas versiones se pueden utilizar tablas de parámetros más grandes. Los parámetros se obtienen utilizando el auto sintonizador una vez en cada condición de operación. Los rangos de la variable de planificación donde se utilizan distintos parámetros se pueden ajustar también por el usuario.

Realimentación adaptativa La información del experimento de realimentación del relé se utiliza para inicializar el controlador adaptativo. La Figura 9.7 muestra el principio del controlador adaptativo. La idea clave es seguir el punto del diagrama de Nyquist obtenido por el auto sintonizador del relé. Se realiza de la siguiente forma. La señal de control u y la señal medida y se filtran a través de un filtro paso-banda estrecho centrado a la frecuencia ω_0 . Esta frecuencia se obtiene a partir del experimento del relé. Las señales se analizan entonces en un estimador de mínimos cuadrados, que proporciona una estima del punto $P(i\omega_0)$.

Feedforward adaptativo La acción feedforward de las perturbaciones medibles puede mejorar frecuentemente el rendimiento de manera significativa. Una acción feedforward adaptativa ha sido una característica de todos los controladores comenzando con el ECA400TM. La diagnosis para la evaluación en línea del valor

potencial de la acción feedforward es un tema activo de investigación; ver [278], [280] y [279].

El control feedforward adaptativo se basa en un modelo simple

$$y(t) = au(t - 4h) + bv(t - 4h), \quad (9.20)$$

donde y es la señal medida, u es la señal de control y v es la señal de perturbación que debería ser pre alimentada. El tiempo de muestreo h se determina del experimento del relé como $h = T_0/8$, donde T_0 es el periodo de oscilación. Los parámetros a y b se estiman recursivamente por un algoritmo de mínimos cuadrados. El compensador feedforward tiene la estructura simple

$$\Delta u_{ff}(t) = k_{ff}(t)\Delta v(t), \quad (9.21)$$

donde la ganancia feedforward k_{ff} se calcula a partir los parámetros estimados del proceso

$$k_{ff}(t) = -0,8 \frac{\hat{b}(t)}{\hat{a}(t)}. \quad (9.22)$$

La interfaz hombre/máquina Los auto-sintonizadores basados en la realimentación del relé se pueden implementar con interfaces hombre-máquina muy sencillas. En muchos casos es suficiente proporcionar a los controladores con simplemente un botón para iniciar la sintonía. La planificación de ganancia se puede implementar también de una manera muy amigable para el usuario. Muchos de los problemas normalmente asociados con la implementación de los controladores adaptativos se pueden evitar porque el auto sintonizador da buenos valores iniciales.

La experiencia industrial ha indicado también que existe una ventaja significativa combinando la adaptación con diagnosis y supervisión. Por ejemplo, carece de sentido sintonizar un controlador si se tiene un mal actuador en el lazo.

Gestión del proceso Emerson

Las técnicas adaptativas utilizadas en los sistemas Emerson vuelven a los sistemas DCS de Provox™ y RS3™, donde se introdujeron el Sintonizador Inteligente de Fisher-Rosemount y el Planificador de Ganancia. El uso de la adaptación se ha expandido en el sistema Delta V. En [52] se pueden encontrar información relativamente detallada sobre las técnicas utilizadas, que además posee numerosas referencias. El sistema tiene capacidades para auto sintonía, planificación de ganancia y adaptación. Hay también un programa para control borroso y control predictivo basado en modelo.

La sintonía automática se basa en la realimentación del relé. El rango de la oscilación del relé es típicamente un pequeño porcentaje del rango de la señal completa. Analizando la parte inicial del primer escalón se obtiene una estimación del retardo de tiempo aparente. Cuando se tiene disponible una estimación del retardo de tiempo, es posible también obtener un modelo FOTD. Se pueden mostrar los parámetros K_p , T y L del modelo FOTD. La sintonía es típicamente llevada a cabo en unos pocos periodos de la oscilación. Como se dispone de un modelo FOTD, es posible utilizar algunas técnicas de sintonía. Las opciones disponibles

incluyen sintonía Ziegler-Nichols, sintonía IMC y sintonía Lambda por nombrar unas pocas. El sistema se estructura de forma que un usuario sin experiencia tiene pocas elecciones, pero un usuario experto tiene disponible a su vez múltiples opciones. Hay también un simulador incorporado de manera que la sintonía se puede probar sobre el modelo del proceso antes de ponerla en funcionamiento sobre el proceso real.

El control adaptativo se basa en los datos del proceso durante una operación normal. La excitación también se puede proporcionar. El sistema consiste en un supervisor, un generador de excitación, adaptadores para la ganancia, el tiempo de integración y el tiempo derivativo, y una red de seguridad. El objetivo de la adaptación es obtener una respuesta oscilatoria amortiguada ligeramente. El enfoque es similar al utilizado en el Exact de Foxboro.

La planificación de ganancia se realiza mediante la estimación de las características estáticas del proceso. La interpolación se realiza utilizando técnicas beras.

Honeywell

Los productos Honeywell que utilizan control adaptativo comenzaron con el controlador monolazo UDC 6000TM que tenía una función adaptativa denominada Accutune. Las técnicas adaptativas fueron desarrolladas posteriormente y son componentes esenciales del sistema DCS de Honeywell TDC 3000TM.

UDC 6000TM La adaptación en el UDC 6000TM combina procedimientos basados en modelo y procedimientos basados en reglas. El modelado se basa en un experimento de respuesta a un escalón. El usuario lleva en modo manual la variable del proceso a un punto lejos del punto de consigna deseado y espera a que se alcance el estado estacionario. Comutando al modo de sintonía se inicia un experimento de respuesta a un escalón en lazo abierto, donde el tamaño del escalón se calcula para hacer que la variable del proceso llegue al punto de consigna.

Durante el experimento, la variable del proceso y su derivada se monitorizan continuamente. El retardo de tiempo L se calcula como el intervalo de tiempo entre el cambio en escalón y el momento en el que la variable de proceso cruza un cierto límite pequeño.

Si la derivada de la variable de proceso decrece continuamente desde el punto de inicio, se concluye que el proceso es de primer orden y se determina un modelo FOTD a partir de unos pocos puntos de la respuesta a escalón. Los cálculos se pueden llevar a cabo antes de alcanzar el estado estacionario, y es de notar que el proceso se identificará en un tiempo menor a una tercera parte de la constante de tiempo.

Si la derivada de la variable de proceso aumenta a una máxima y entonces decrece, el proceso se identifica como un proceso de segundo orden y se determina un modelo SOTD a partir de la respuesta en escalón. El controlador se comuta entonces a modo automático y se controla al punto de consigna utilizando los parámetros del controlador preliminares cuando se alcanza la pendiente máxima de la salida del proceso, pero es necesario esperar al régimen permanente para obtener un modelo completo. En [20] se pueden encontrar más detalles sobre el procedimiento de modelado.

Cuando se ha obtenido un modelo, se calculan los parámetros del controlador a partir del modelo y el controlador se pasa a modo de control automático.

El controlador utilizado tiene la forma serie con la función de transferencia

$$C(s) = K \frac{(1 + sT'_i)(1 + sT'_d)}{sT'_i(1 + 0,125sT'_d)}.$$

Obsérvese que la constante de tiempo del filtro es 1/8 del tiempo derivativo. El diseño del controlador se basa en la asignación de polos del procedimiento tipo Dahlin-Higham donde se cancelan los polos del proceso. Existen algunas versiones diferentes dependiendo del orden del proceso y del retardo de tiempo. Para sistemas con retardo de tiempo, la constante de tiempo en lazo cerrado se escoge como $T_{cl} = L + T/3$. El controlador UDC 6000 tiene también adaptación continua, que se activa cuando la variable de proceso cambia más de 0,3 por ciento desde el punto de consigna o si el punto de consigna cambia más de un valor establecido.

Honeywell LOOPTUNE™ El sistema TDC3000 tiene un amplio rango de controladores; controladores básicos, controladores extendidos, controladores multi-función, gestores de procesos y módulos de gestión. LOOPTUNE es un paquete de programas en el sistema que sintoniza lazos con controladores PID.

El algoritmo de sintonía no cuenta con ningún modelo particular del sistema. El rendimiento se evalúa utilizando la función de coste cuadrática

$$J = \frac{1}{N} \sum_{t=1}^N ((1 - \rho)(y(t) - y_{sp}(t))^2 + \rho(u(t) - u(t-1))^2), \quad (9.23)$$

donde N es el horizonte de evaluación y ρ un factor de ponderación que balanza el error de control frente a los cambios en el actuador. Los parámetros del controlador se modifican uno cada vez y la función de coste se evalúa sobre un horizonte de tiempo dado N . Para obtener una estimación fiable se requiere un valor grande de N , pero entonces la evaluación emplea un tiempo mayor. El conocimiento del proceso se puede utilizar para mejorar la búsqueda de buenos parámetros del controlador dirigiendo la búsqueda hacia las ganancias del controlador mayores y los tiempos integrales más pequeños.

Yokogawa SLPC-181, 281

Los Yokogawa SLPC-181 y 281 utilizan un modelo del proceso como un sistema de primer orden con retardo de tiempo para calcular los parámetros del PID. Para obtener el modelo se utiliza una técnica de programación no lineal. Los parámetros del PID se calculan de las ecuaciones desarrolladas a partir de exhaustivas simulaciones. Sin embargo, las ecuaciones exactas no han sido publicadas.

Se utilizan dos estructuras de controlador distintas.

$$\begin{aligned} 1: \quad u &= K \left(-y + \frac{1}{T_i} \int edt - T_d \frac{dy_f}{dt} \right) \\ 2: \quad u &= K \left(e + \frac{1}{T_i} \int edt - T_d \frac{dy_f}{dt} \right) \end{aligned} \quad (9.24)$$

Cuadro 9.2 Especificaciones de respuesta al punto de consigna utilizadas en el Yokogawa SLPC-181 y en el 281.

Tipo	Características	Criterio
1	no sobreelongación	no sobreelongación
2	5 % sobreelongación	ITAE mínimo
3	10 % sobreelongación	IAE mínimo
4	15 % sobreelongación	ISE mínimo

donde y_f se genera filtrando y con un filtro de primer orden que posee una constante de tiempo T_d/N . La primera estructura se recomienda si el rechazo de la perturbación de carga más importante, y la estructura 2 si lo más importante son las respuestas a cambios en el punto de consigna. El punto de consigna se puede además pasar a través de dos filtros en serie:

$$\text{Filtro 1: } \frac{1 + \alpha_i s T_i}{1 + s T_i} \quad \text{Filtro 2: } \frac{1 + \alpha_d s T_d}{1 + s T_d} \quad (9.25)$$

donde α_i y α_d son parámetros establecidos por el usuario, principalmente para ajustar la sobreelongación de la respuesta al punto de consigna. Los efectos de estos dos filtros son esencialmente equivalentes a la ponderación del punto de consigna. Se puede mostrar que $\alpha_i = b$, donde b es el factor de ponderación del punto de consigna.

El usuario especifica el tipo de rendimiento de la respuesta del punto de consigna de acuerdo con la Tabla 9.2. Una sobreelongación alta dará, por supuesto, una respuesta más rápida. El controlador tiene cuatro modos adaptativos:

Modo auto. El controlado adaptativo está activo. Los controladores PID se actualizan automáticamente.

Modo monitorización. En este modo, sólo se visualizan el modelo calculado y los parámetros del PID. Este modo es útil para validar la función adaptativa o comprobar las variaciones de la dinámica del proceso durante la operación.

Modo auto arranque. Este modo se utiliza para calcular los parámetros PID iniciales. Se utiliza una respuesta en lazo abierto a un escalón para estimar el modelo.

Modo bajo demanda. Este modo se utiliza para realizar un cambio en el punto de consigna. Cuando se solicita la sintonía bajo demanda, se aplica un cambio en escalón a la entrada del proceso en lazo cerrado. El controlador estima el modelo del proceso utilizando la respuesta en lazo cerrado resultante.

El controlador monitoriza constantemente el rendimiento del sistema computando la razón de las varianzas de la salida del proceso y del modelo. Se espera obtener una razón alrededor de 1. Si es más grande que 2 o menor que 0,5, se da un mensaje de advertencia para la re sintonía del controlador. El retardo de tiempo y la compensación feedforward están disponibles para el controlador de ganancia constante, pero el fabricante no recomienda que se utilice en conjunto con la adaptación.

Techmation Protuner

El Protuner es un analizador del proceso de Techmation Inc. Consiste en un paquete de programas para computadores personales y un módulo de interfaz con cables para conectarse a la salida del proceso y a la señal de control del lazo que se va a analizar. Protuner monitoriza un experimento de respuesta a un escalón, calcula la respuesta en frecuencia del proceso basado en datos experimentales y sugiere los parámetros del controlador basado en varios métodos de sintonía.

Información previa Antes de realizar el análisis del proceso, el usuario debe proporcionar alguna información sobre el proceso y el controlador. Esto se realiza utilizando un par de «Menús de configuración». Se debe dar la siguiente información del proceso:

- Los rangos de las señales de control y de medida.
- Se debe determinar si el proceso es estable o si tiene acción integral.

Para ser capaces de establecer parámetros del controlador relevantes, se deben proporcionar los siguientes datos del controlador:

- Tipo P (ganancia o banda proporcional).
- Tipo I (segundos, segundos/repetición, minutos, or minutos/repetición).
- Estructura del controlador (ideal, serie o paralelo).
- Tiempo de muestreo.
- Constante de tiempo del filtro (si existe un filtro paso-bajo conectado a la señal de medida).

Antes de que se pueda llevar a cabo el experimento de sintonía, el usuario debe además especificar un tiempo de muestra. Éste es el tiempo durante el cual los datos se almacenarán durante el experimento. Es importante escoger el tiempo de muestra lo suficientemente grande, de forma que la respuesta a un escalón se asiente antes de que el tiempo de muestra finalice. En el caso de un experimento en lazo abierto para un proceso con integración, la respuesta debe alcanzar una velocidad constante de cambio cuando el experimento finalice.

Determinando el modelo del proceso El procedimiento de sintonía se basa en un experimento de respuesta a un escalón. Se puede realizar bien en lazo abierto o bien en lazo cerrado. Se recomienda el experimento en lazo abierto. Cuando el usuario envía la orden de inicio, se muestran en la pantalla la salida del proceso y la señal de control donde el eje de abscisas representa el tiempo de muestra definido por el usuario. El usuario realiza entonces un cambio en escalón en la señal de control. Si el experimento se realiza en lazo cerrado, se introduce un escalón en el punto de consigna.

Existen varias capacidades para editar los datos obtenidos del experimento de respuesta a un escalón. Se pueden eliminar valores fuera de rango y los datos se pueden filtrar. Estas características son de gran utilidad porque hacen posible superar problemas que generalmente ocurren cuando se realizan experimentos en procesos industriales.

Cuando los datos se han editado, Protuner calcula la respuesta en frecuencia del proceso. El resultado se puede visualizar también en un diagrama de Bode, de Nyquist o de Nichols. Entonces, se muestran la ganancia estática, la constante de tiempo dominante y el retardo de tiempo aparente, así como la ganancia última y el período último.

Cálculos de diseño Los parámetros del controlador se calculan a partir de la respuesta en frecuencia utilizando una técnica especial. Ésta se basa en la cancelación de los polos del proceso por los ceros del controlador. El tiempo integral y el tiempo derivativo se determinan primero para llevar a cabo esta cancelación. La ganancia es entonces determinada para cumplir los márgenes de fase y de ganancia predeterminados.

Protuner proporciona varios opciones de diseño. Los parámetros del controlador se dan para respuestas en lazo cerrado lentas (críticamente amortiguadas), medianas (levemente sobreamortiguadas) y rápidas (razón de decaimiento de 0,38). Las diferentes opciones de diseño se obtienen especificando diferentes valores de los márgenes de ganancia y fase. El Protuner proporciona diferentes parámetros del controlador dependiendo si se consideran el punto de consigna o las perturbaciones de carga. Se proporcionan parámetros del controlador P, PI y PID. El usuario debe suministrar las ponderaciones del punto de consigna para la acción proporcional y derivativa y la ganancia de altas frecuencias de la parte derivativa.

Evaluación Es posible evaluar el rendimiento del sistema en lazo cerrado de varias maneras. Se puede dibujar la respuesta en frecuencia combinada, por ejemplo, de la función de transferencia en lazo cerrado $G_l(i\omega) = P(i\omega)C(i\omega)$ en un diagrama de Bode, de Nyquist o de Nichols. De esta manera, se pueden comprobar los márgenes de ganancia y de fase o el valor M_s .

Protuner también tiene capacidad de simulación. Es posible simular la respuesta en lazo cerrado del proceso utilizando el controlador sugerido. Para hacer esto, es necesario proporcionar algunos parámetros adicionales del controlador, concretamente, las ponderaciones del punto de consigna b y c , y el factor de limitación de la ganancia derivativa N . Utilizando la capacidad de simulación, es además posible investigar los efectos del ruido y diseñar filtros para reducir tales efectos.

Algunas reflexiones personales

Las técnicas adaptativas se han utilizado extensivamente en la industria desde la mitad de los 80. Las técnicas se han mostrado útiles y se continúan desarrollando productos, pero existe un claro potencial para mejorar los productos existentes.

De los resultados en el Capítulo 7 se pueden aprender varias lecciones. Una observación es que es útil caracterizar la dinámica del proceso con tres parámetros. La dinámica puede entonces clasificarse como de retardo dominante, balanceada o de constante de tiempo dominante. Se puede obtener un control equilibrado utilizando un modelo FOTD para sistemas con dinámicas balanceadas o de retardo dominante, pero el rendimiento de control se puede mejorar para sistemas de constante de tiempo dominante utilizando un mejor modelo. En la Sección 2.7, se mostró que es difícil obtener un modelo SOTD a partir de un experimento de respuesta a un escalón. Por lo tanto, no es posible diseñar auto-

sintonizadores para un control equilibrado basado en una respuesta a un escalón o del conocimiento de la ganancia y la frecuencia última. Una indicación de esto es que Foxboro comutó a utilizar un doble pulso en lugar de un escalón. El doble pulso se puede realmente interpretar como una versión corta de un experimento con realimentación del relé.

Es altamente deseable llevar a cabo la sintonía en un corto espacio de tiempo, tal y como se mostró para el Honeywell UDC 6000TM. Una ventaja del auto-sintonizador del relé es que generalmente la sintonía se efectúa en un tiempo mucho más corto que el tiempo de asentamiento medio del sistema. En particular, el tiempo puede ser mucho más corto para sistemas con dinámica de constante de tiempo dominante. Una cuestión interesante es por lo tanto qué información se puede deducir de un experimento de realimentación del relé. La experiencia de Emerson indica que al menos se puede determinar un modelo FOTD a partir de un experimento de realimentación del relé. Explorar en detalle la información que se puede deducir de un experimento con relé, es por lo tanto una tarea de investigación interesante y útil. Si se obtienen mejores modelos, es posible utilizar los algoritmos presentados en el Capítulo 7 para dar un control más equilibrado. Las ganancias potenciales son particularmente grandes para procesos con constante de tiempo dominante.

Los enfoques libres de modelo para control adaptativo tienen muchas características atractivas pero su principal desventaja es que la sintonía requiere mucho tiempo. La sintonía se puede hacer más efectiva utilizando sintonía con realimentación iterativa, que también calcula las estimaciones del gradiente de la función de coste; ver Sección 9.7. Los controladores adaptativos basados en modelos más elaborados como el controlador autosintonizado presentado en la Sección 9.7, es una alternativa al control basado en modelo. Tales controladores trabajan muy bien pero hasta ahora han requerido usuarios con grandes conocimientos. Puede existir una posibilidad para hacerlos más simples explotando la información obtenida de la sintonía automática. Esto puede ser además el camino para introducir la adaptación en los controladores predictivos basados en modelo.

Herramientas como Techmations Protuner, que permiten la simulación del proceso con diferentes configuraciones del controlador, son muy útiles para usuarios avanzados. Hay muchos componentes disponibles para construir tales sistemas en los sistemas actuales DCS. Es, por lo tanto, natural proporcionar tales herramientas como una parte integral de los sistemas.

9.9 Resumen

Una característica esencial de la realimentación es que se puede utilizar para diseñar sistemas que son insensibles a las variaciones del proceso. Cuando existen grandes variaciones, el rendimiento se puede mejorar ajustando los parámetros del controlador. Las técnicas adaptativas están, por lo tanto, siendo incrementalmente utilizadas en los controladores PID para adaptar los parámetros del controlador a los cambios en la dinámica del proceso o en las perturbaciones. En este capítulo, hemos dado una amplia presentación de una variedad de métodos adaptativos cubriendo la sintonía automática, la planificación de ganancia y la adaptación continua. Estas técnicas se utilizan de varias formas.

En la sintonía automática, los parámetros del controlador se ajustan bajo demanda del usuario. La planificación de ganancia se puede utilizar cuando existe una variable planificación medible que se correlaciona bien con los cambios del proceso. Los parámetros del controlador se obtienen de una tabla, que aporta los parámetros del controlador como función de la variable de planificación. La auto sintonía se puede utilizar para construir dicha tabla. Cuando está disponible una variable de planificación, se puede utilizar el control adaptativo.

Se presentaron métodos basados en modelo y basados en características, donde se prestó atención particular a la realimentación del relé para la auto-sintonía, la estimación de parámetros y la sintonía con realimentación iterativa.

El controlador adaptativo obtiene el conocimiento requerido de la entrada y la salida del proceso. El control adaptativo es menos robusto que la planificación de ganancia y requiere funciones de supervisión. La supervisión de controladores adaptativos fue por lo tanto descrita.

También se efectuó una corta presentación de algunos controladores adaptativos industriales donde se han utilizado los métodos adaptativos de manera satisfactoria.

9.10 Notas y referencias

Los controladores con sintonía automática surgieron de la investigación sobre control adaptativo. En [103, 34, 61, 36] se pueden encontrar artículos de revisión sobre técnicas adaptativas. En los libros [158, 148, 29], se puede encontrar un tratamiento más detallado. En [185, 116, 193, 248, 366, 14] se encuentran algunas revisiones de diferentes enfoques y productos.

En los sintonizadores automáticos se utilizan muchos enfoques diferentes. Los sistemas descritos en [255, 204, 335] se basan en técnicas de respuestas transitórias. El artículo [151] se basa en la correlación cruzada. En [374, 180] se propone el uso de la representación en series ortonormales de la respuesta a un escalón del sistema. El reconocimiento de patrones, que fue la base para el controlador EXACTTM de Foxboro, se describe en [60, 61, 58, 62, 59, 64, 283, 5, 199, 263, 332]. La auto sintonía basada en la realimentación del relé se trató en [16, 17, 136, 300, 150, 212, 301, 349]. La sintonía con realimentación iterativa se presentó en [170]. Es más efectiva que la búsqueda directa porque se utiliza información del gradiente.

Las técnicas adaptativas tradicionales basadas en la identificación de sistemas y el diseño de control han sido aplicadas al control PID. La identificación se basa generalmente en la estimación de parámetros de un modelo de función de transferencia. En [163, 173, 367, 368, 73] se muestran algunos ejemplos de este enfoque. Existen también algunos sistemas donde el controlador se actualiza directamente, como los descritos en [288, 230, 289]. La supervisión de controladores adaptativos se estudia en [183, 331, 83, 213, 137].

En [347] se presenta una revisión de algunos productos que utilizan adaptación. Algunas ayudas para la sintonía se implementan en computadores de mano o como programas de computador donde el usuario introduce la información del proceso a través de un teclado; ver [53, 345, 366].

Los artículos [238, 237] describen los productos Fisher Rosemount para sintonía y planificación de ganancia. Las implementaciones en el sistema Delta VTM

DCS se describen en el libro [52]. Los sistemas Yokogawa se analizan en [335] y [366].

Existen comparaciones de diferentes auto sintonizadores y controladores adaptativos, pero pocos resultados de estos estudios han llegado a dominio público. Algunos artículos que tratan de este tema son [249, 250, 103, 102]. En [167, 71] se describen algunas experiencias operacionales.

10

Evaluación del lazo y del rendimiento

10.1 Introducción

El diseño, la sintonía y la implementación de estrategias de control y controladores es sólo la primera fase en la solución de un problema de control. La segunda fase incluye la operación, la supervisión y el mantenimiento. Esta fase ha sido tradicionalmente realizada de manera manual, pero el interés sobre funciones de supervisión automáticas se ha incrementado significativamente en los últimos años a causa de la reducción de personal en la industria de procesos.

Este capítulo trata sobre los métodos para la puesta en servicio, la supervisión y la diagnosis de lazos de control. Los métodos de adaptación presentados en el Capítulo 9 se dividieron en dos categorías, sintonía bajo demanda y adaptación continua. Los procedimientos para la supervisión y la diagnosis se pueden clasificar de la misma manera. Los llamamos evaluación del lazo y del rendimiento. Los procedimientos de evaluación del lazo se utilizan para investigar las propiedades del lazo de control, por ejemplo, niveles de señales, niveles de ruido, no linealidades y condiciones del equipo. Los procedimientos de evaluación del rendimiento se utilizan para supervisar los lazos de control durante la operación y asegurar que cumplen las especificaciones. Los fallos en cumplir las especificaciones pueden deberse a problemas del equipo, no linealidades u otras variaciones en la dinámica del proceso o en su entorno.

El capítulo comienza con una presentación de los problemas que ocurren en las válvulas. Estos problemas se identifican como una de las principales razones del mal rendimiento del lazo de control. Las Secciones 10.3 y 10.4 tratan de la evaluación del lazo y del rendimiento, respectivamente. La sintonía y la diagnosis poseen muchos aspectos en común. Tales aspectos se analizan en la Sección 10.5.

10.2 Válvulas

Las válvulas de control están sujetas a desgaste. Tras algún tiempo de operación, este desgaste resulta en fricción e histéresis que deteriora el rendimiento de con-

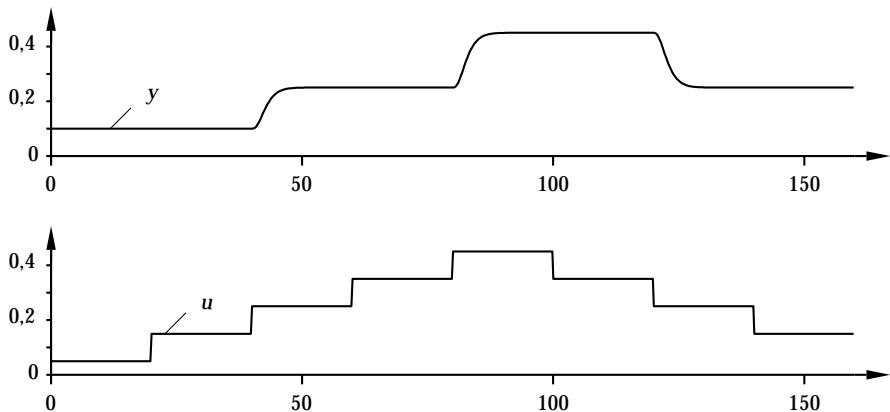


Figura 10.1 Procedimiento para comprobar la cantidad de fricción en una válvula. El diagrama superior muestra la salida del proceso y y el diagrama inferior muestra la señal de control u .

trol. Además, las válvulas son generalmente no lineales y están sobredimensionadas. Por tanto, las válvulas han sido identificadas como la principal fuente de problemas a nivel de lazo en el control de procesos. Por consiguiente, dedicaremos esta sección a estos problemas.

Fricción en la válvula

La alta fricción en la válvula es una causa común de problemas. Por supuesto, en la válvula siempre existe una fricción estática (stiction), pero si el mantenimiento de la válvula es insuficiente, la fricción puede ser tan grande que se degrada el rendimiento de control. La cantidad de fricción puede medirse fácilmente haciendo pequeños cambios en la señal de control y observando cómo reaccionan las salidas del proceso. El procedimiento se muestra en la Figura 10.1. En la figura, la salida del proceso sólo responde a la señal de control cuando los cambios en la señal de control son suficientemente grandes para vencer a la fricción estática.

La fricción en la válvula resulta en un movimiento de deslizamiento. Este fenómeno se muestra en la Figura 10.2. Supongamos que la válvula está clavada (pegada) en una cierta posición debido a la fricción. Si existe error de control, la acción integral del controlador hará que la salida del controlador aumente hasta que la presión en el actuador sea suficientemente alta para vencer a la fricción estática. En este momento, la válvula se mueve (despegue) a una nueva posición donde se clava (pega) otra vez. Esta posición de la válvula es normalmente tal que la salida del proceso se mueve al otro lado del punto de consigna, lo que quiere decir que el procedimiento se repite. La salida del proceso oscilará por lo tanto alrededor del punto de consigna. El patrón de la Figura 10.2, donde la señal de medida está cercana a una onda cuadrada y la señal de control a una onda triangular, es el movimiento típico de pegue y despegue o deslizamiento.

Muchos operadores desintonizan el controlador cuando ven oscilaciones como la mostrada en la Figura 10.2, ya que creen que las oscilaciones se deben a una mala sintonía del controlador. Desafortunadamente, la mayoría de controladores

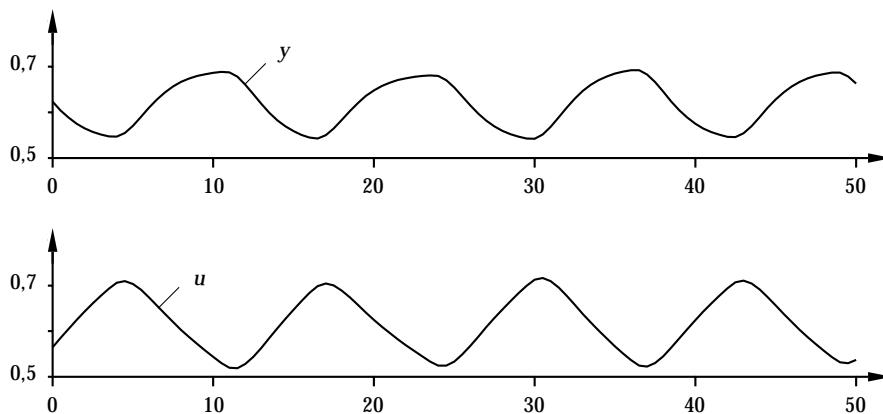


Figura 10.2 Movimiento de deslizamiento por fricción en la válvula y la acción integral. El diagrama superior muestra la salida del proceso y y el diagrama inferior muestra la señal de control u .

adaptativos hacen lo mismo. Lo que se debería hacer cuando un lazo de control empieza a oscilar es, primero, determinar la causa de la oscilación. En la Figura 10.3 se presenta una buena forma de hacer esto.

El primer problema a determinar es si las oscilaciones son generadas dentro o fuera del lazo de control. Esto se puede hacer desconectando la realimentación, por ejemplo, conmutando el controlador a modo manual. Si la oscilación todavía persiste, las perturbaciones se deben generar fuera del lazo; en caso contrario se producirían dentro del lazo. Podría existir una situación cuando el lazo de control oscila a causa de la fricción de la válvula, incluso cuando el controlador está en modo manual, concretamente, si la fricción ocurre en la válvula piloto de un posicionador en lugar de en la propia válvula.

Si las perturbaciones se generan dentro del lazo, la causa puede ser la fricción en la válvula o un controlador mal sintonizado. Si la fricción está presente o no, se puede determinar haciendo pequeños cambios en la señal de control y comprobando si la señal de medida sigue, tal y como se muestra en la Figura 10.1. Si la fricción está causando oscilaciones, la solución al problema es el mantenimiento de la válvula.

Si las perturbaciones se generan fuera del lazo de control, se podría intentar, por supuesto, encontrar el origen de las perturbaciones para eliminarlas. Esto no es siempre posible, incluso si se localiza el origen. Uno puede entonces intentar prealimentar la perturbación al controlador y de esta manera reducir sus efectos sobre el lazo de control actual. Véase la Sección 5.6.

Histéresis en la Válvula

Debido al desgaste, generalmente existe histéresis (huelgo) en la válvula o el actuador. La cantidad de histéresis se puede medir tal como se muestra en la Figura 10.4. El experimento comienza con dos cambios en escalón en la señal de control en la misma dirección. La anchura de la histéresis se cerrará si el primer escalón es suficientemente grande. Esto quiere decir que el segundo escalón se realiza sin

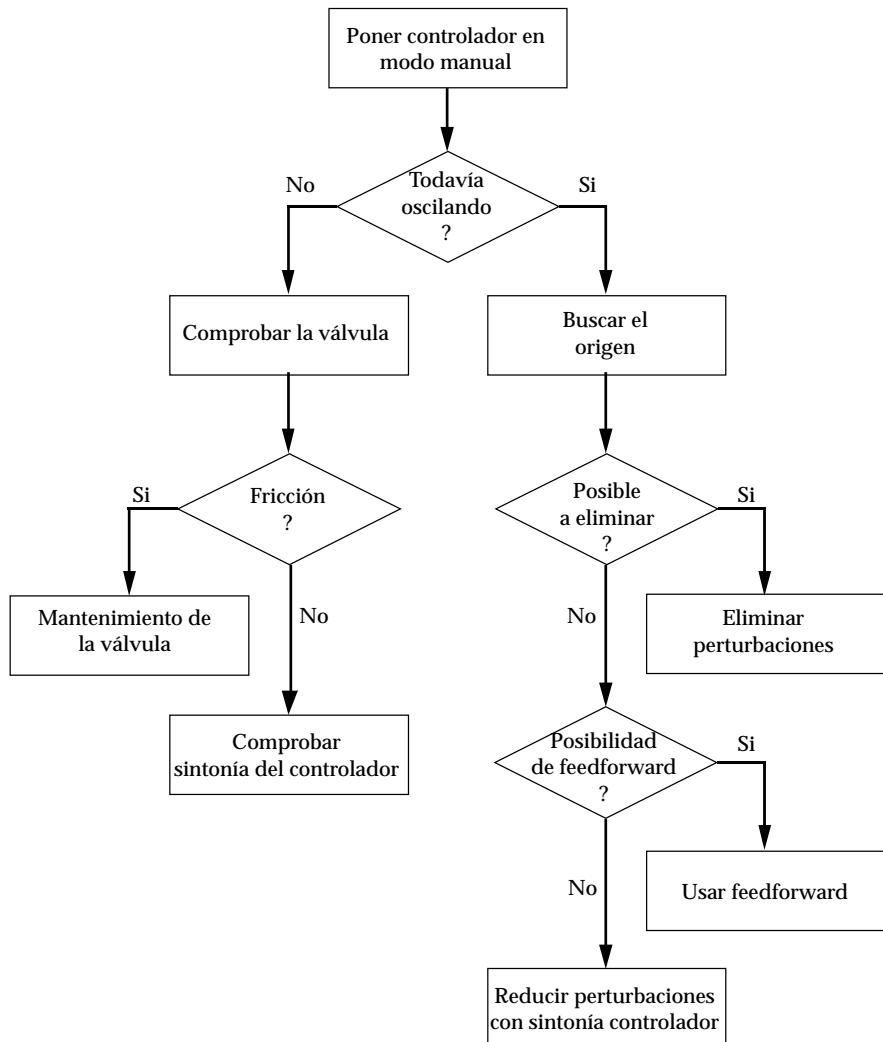


Figura 10.3 Procedimiento de diagnosis para descubrir la causa de las oscilaciones y acciones recomendadas para eliminarlas.

histéresis. El tercer escalón se realiza entonces en la dirección contraria. La señal de control tiene entonces que recorrer la anchura completa de la histéresis antes de que la válvula se mueva. Si los últimos dos escalones son del mismo tamaño, la histéresis es $\Delta y / K_p$, donde Δy es la diferencia entre las salidas del proceso después del segundo y tercer escalón (ver Figura 10.4), y K_p es la ganancia estática del proceso (fácilmente obtenida desde la Figura 10.4).

La histéresis se puede determinar también a partir de un barrido continuo sobre partes del rango de operación. La Figura 10.5 muestra las salidas del proceso de un proceso con fricción y otro con histéresis, respectivamente, cuando la entrada del proceso es en forma de rampa desde cero a uno y después vuelta a

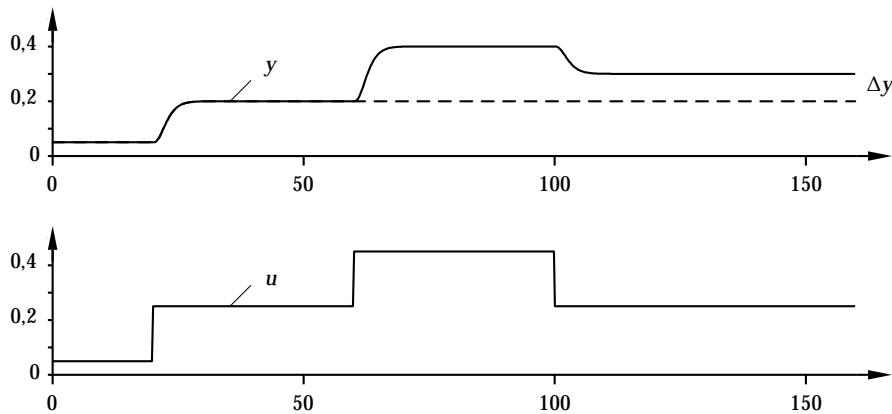


Figura 10.4 Procedimiento para comprobar la histéresis de la válvula. El diagrama superior muestra la salida del proceso y y el inferior muestra la señal de control u .

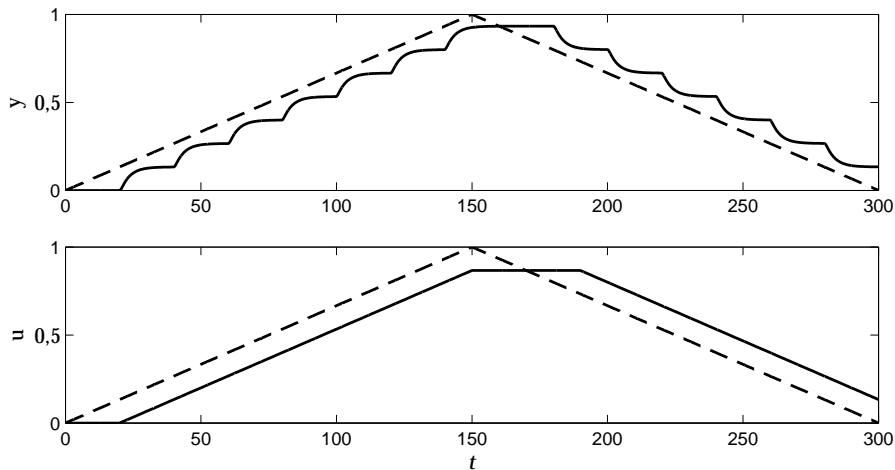


Figura 10.5 Salidas del proceso (líneas continuas) y señales de control (líneas a tramos) para proceso con fricción (gráfica superior) e histéresis (gráfica inferior).

cero otra vez. En la Figura 10.6 se presentan los correspondientes diagramas de fase. Uno puede fácilmente medir la cantidad de histéresis a partir del diagrama de fase. Barridos de este tipo se hacen convenientemente durante la puesta en servicio.

La Figura 10.7 el control en lazo cerrado de un proceso con 10 por ciento de histéresis en la válvula. El proceso es

$$P(s) = \frac{1}{(1 + 0.05s)^2} e^{-0.3s},$$

y el controlador es un controlador PI con parámetros $K = 0,35$ y $Ti = 0,15$. La señal de control tiene que recorrer la anchura de la histéresis para mover la válvula

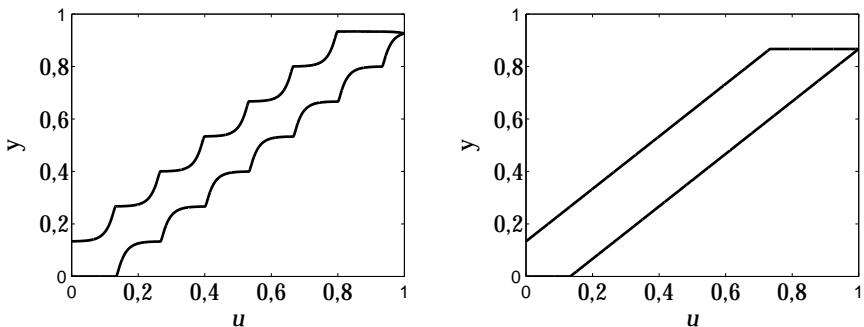


Figura 10.6 Diagramas de fase de las señales de la Figura 10.5 para el proceso con fricción (izquierda) e histéresis (derecha).

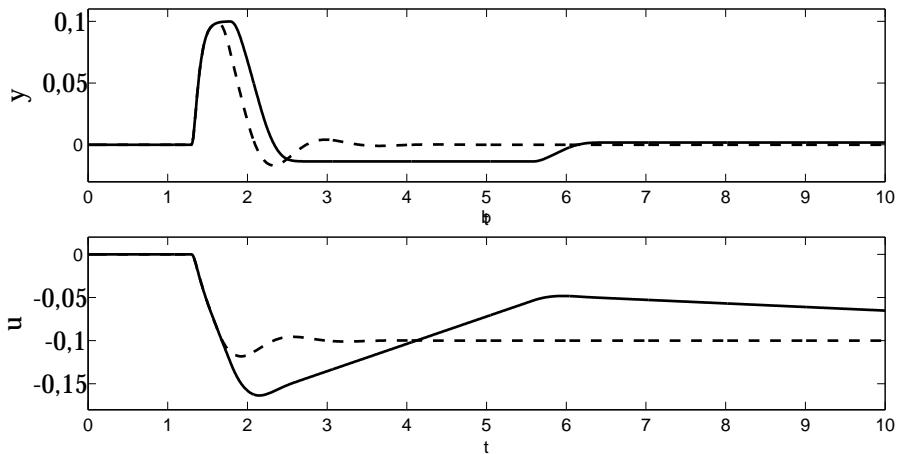


Figura 10.7 Control en lazo cerrado con histéresis en la válvula. El diagrama superior muestra la salida del proceso y y el inferior la señal de control u . Las líneas punteadas muestran el control sin histéresis. Las líneas continuas muestran el control con una histéresis del 10 por ciento (0.1).

la. Por lo tanto, obtenemos las derivas lineales típicas en la señal de control tal y como se muestra en la Figura 10.7.

Si se aplica un auto sintonizador de relé a un proceso con histéresis, la ganancia del proceso estimada será más pequeña que el valor real. Esto conlleva una ganancia del controlador demasiado elevada. Un auto sintonizador basado en un experimento de respuesta a un escalón trabajará adecuadamente si la anchura está cerrada antes de que se realice el experimento de respuesta a un escalón. (Comparar con el segundo escalón de la Figura 10.4).

10.3 Evaluación del lazo

Esta sección sugiere ensayos que son útiles para realizar sobre el lazo de control. Estos ensayos se deberían llevar a cabo de manera regular y especialmente en conexión con la sintonía del controlador. Los ensayos para fricción e histéresis, presentados en la Sección 10.2, son dos procedimientos importantes de la evaluación del lazo. Los experimentos sugeridos en la Sección 2.7 para obtener la dinámica del proceso son también procedimientos de asignación del lazo para la sintonía de los controladores. Las comprobaciones y ensayos añadidos en esta sección son básicas pero a menudo se olvidan o desprecian.

Rangos de la señal

El rango de señal de la señal de medida se relaciona con la resolución del sensor. Un rango de señal grande significa que la resolución es baja. Para obtener una alta resolución, es por lo tanto importante restringir el rango de la señal a aquellos valores que son relevantes para control.

Si el elemento de control final es una válvula, el rango de salida se determina por el tamaño de la válvula. Las válvulas suelen estar sobredimensionadas. Las principales razones son inseguridad entre los ingenieros combinado con un temor de instalar una válvula que sea demasiado pequeña para proporcionar los caudales máximos posibles.

Una válvula grande no tiene la misma exactitud que una más pequeña. Los problemas de fricción y huelgo discutidos en la sección anterior son más severos si la válvula está sobredimensionada.

Si los rangos de la señal se escogen adecuadamente y el proceso es lineal, la ganancia estática ideal del proceso es $P(0) = 1$. Si la ganancia estática es uno, la señal medida alcanza su valor máximo cuando la señal de control está en su máximo valor. A causa de las válvulas sobredimensionadas, la ganancia estática del proceso es generalmente mayor que uno en aplicaciones de control de procesos.

Relaciones estáticas de entrada-salida

Desde un punto de vista de control, se desea tener una relación entrada-salida lineal estática. Esta relación es, sin embargo, generalmente no lineal, debido principalmente a una válvula de característica no lineal. Las no linealidades pueden también suceder en sensores o en el propio proceso.

Si el proceso es no lineal, el control se puede mejorar utilizando una planificación de ganancia u otras formas de linealización. Como se indicó en la Sección 9.3, es importante comprender la causa de la no linealidad para determinar una referencia adecuada de planificación de ganancia.

La característica estática del proceso se puede obtener determinando la relación estática entre la señal de control y la señal medida. Esto se puede hacer realizando cambios en escalón en la señal de control y midiendo los correspondientes cambios en la salida del proceso; ver Figura 10.8.

La característica mostrada en la Figura 10.8 es obviamente no lineal. Tiene una ganancia más elevada cuando las señales de control son más grandes. Si los valores estacionarios de la señal de medida se dibujan frente a la señal de control, obtenemos la característica estática del proceso. Véase la Figura 10.8. Un gráfico como esta revela si la planificación de ganancia es adecuada o no.

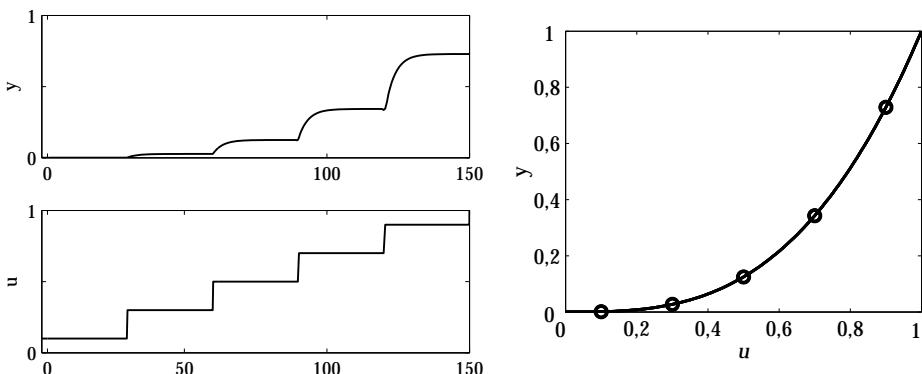


Figura 10.8 Los diagramas de la izquierda muestran un procedimiento para determinar la característica estática del proceso. La señal de control u se cambia en modo de escalones, y se determinan los cambios correspondientes en la salida del proceso y . El diagrama de la derecha muestra la característica estática del proceso, por ejemplo, la salida del proceso y como función de la señal de control u .

Perturbaciones

Otro tema importante a considerar antes de la sintonía del controlador es la actuación de las perturbaciones sobre el lazo de control. Hemos indicado que es importante conocer si las principales perturbaciones son cambios en el punto de consigna (el problema del servo) o las perturbaciones de carga (el problema del regulador).

También es importante investigar el nivel del ruido de medida y su contenido en frecuencia. Compárese con la Sección 2.6. Si el nivel de ruido es alto, puede ser necesario filtrar la señal de medida antes de introducir el algoritmo de control. Ésta es una manera sencilla de eliminar el ruido de alta frecuencia. Si existen perturbaciones con gran contenido de frecuencia cercano a la frecuencia última, no es posible utilizar un filtrado paso-bajo para eliminarlas. La acción feedforward es una posibilidad, siempre que sea posible medir el origen de las perturbaciones. Se pueden utilizar filtros notch si el ruido se concentra en un rango de frecuencia estrecho. Véase la Sección 2.6 donde se analiza el modelado y las medidas del ruido.

10.4 Evaluación del rendimiento

La evaluación del lazo, seguida por acciones apropiadas como mantenimiento de la válvula, selección de los rangos de señal, linealización de las no linealidades y sintonía del controlador, debería mantener el lazo de control en buena forma.

Tras algún tiempo de operación, el rendimiento puede, sin embargo, deteriorarse debido a variaciones en el proceso y la operación. Por lo tanto, es importante supervisar los lazos de control y detectar estas degradaciones. Esta supervisión ha sido tradicionalmente realizada por humanos, pero la reducción de personal en la industria de procesos combinada con el incremento en la demanda de calidad ha sido una fuerza motriz para el desarrollo de procedimientos de moni-

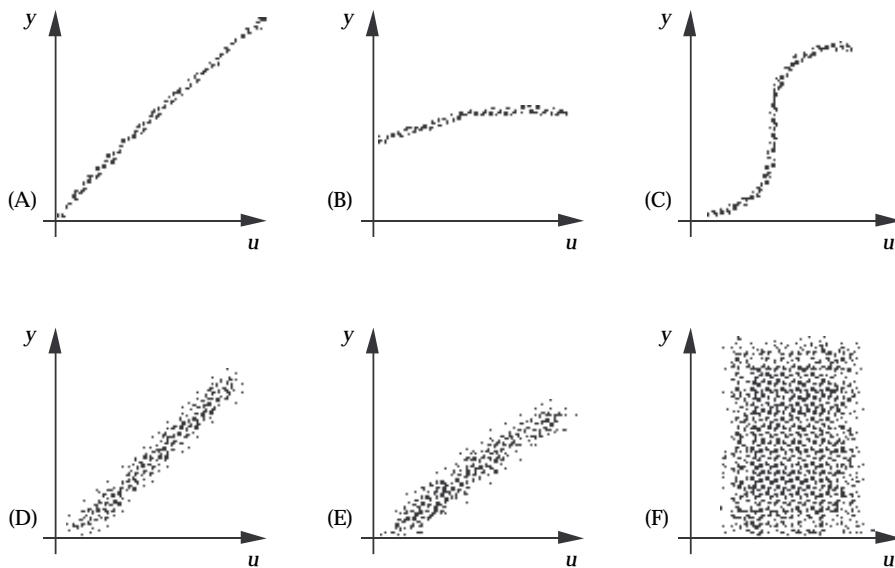


Figura 10.9 Ejemplos de datos de entrada-salida estática registrados durante una operación normal. Los resultados que se muestran en A, B y C indican un problema de servo puro. Los resultados en F indican un problema de regulación puro. Los D y E son una mezcla de casos. El caso B indica una pobre resolución del sensor y el caso E indica una mala elección del tamaño del actuador.

torización y evaluación automáticos del rendimiento. Esta sección proporciona algunos procedimientos para la supervisión automática del rendimiento del lazo de control.

La relación estática entrada-salida

Si se dispone de un detector para la estacionariedad, es sencillo guardar una estadística de la fracción de tiempo en la que el sistema está en estacionario. La relación estática de entrada-salida se puede obtener entonces simplemente registrando la entrada y la salida del proceso durante las condiciones del estacionario. Para obtener buenos datos se deberían filtrar con respecto a la escala de tiempo del lazo cerrado. Entonces se obtienen gráficas como las mostradas en la Figura 10.9. De estas curvas se pueden determinar si las principales variaciones en la salida se deben a cambios en el punto de consigna o a perturbaciones de carga, por ejemplo, si tratamos con un problema tipo servo o con un problema de regulación. Tendremos un problema tipo servo si los datos experimentales dan una curva bien definida y un problema de regulación si no hay una relación definida entre las entradas y las salidas. Un indicador que puede ser también útil es una simple estadística de la fracción del tiempo total donde existen cambios en escalón o transitorios debidos a cambios en el punto de consigna.

Para un problema tipo servo, también se pueden determinar las variaciones en la ganancia estática del sistema. Esto da una indicación valiosa de si se requiere una planificación de ganancia. La curva de ganancia estática se puede utilizar también para propósitos de diagnóstico. Cambios en la curva indican cambios en

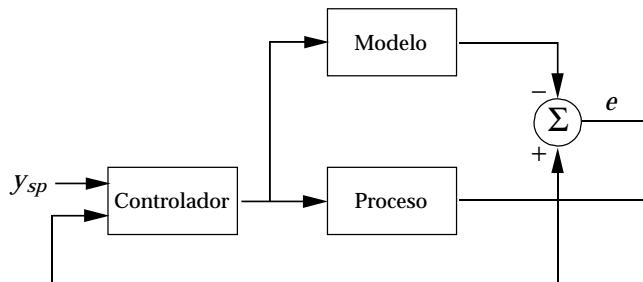


Figura 10.10 Detección de fallos basada en modelo.

el proceso. Comparando la pendiente de la curva de la ganancia estática con la ganancia incremental del proceso medida durante la sintonía o adaptación, podemos también obtener indicaciones de si existe o no histéresis en el lazo. También se pueden deducir si los actuadores están adecuadamente dimensionados.

Diagnóstico basado en modelo

La mayoría de los procedimientos de supervisión automática están, en principio, basados en la idea mostrada en la Figura 10.10. Si se dispone de un modelo del proceso, la señal de control se puede alimentar a la entrada del modelo. Comparando la salida del modelo con la salida del proceso real, se puede detectar cuando cambia la dinámica del proceso. Si el modelo es bueno, la diferencia entre la salida del modelo y la salida del proceso (e) es pequeña. Si cambia la dinámica del proceso e no será ya pequeña, porque las dos respuestas a la señal de control serán diferentes.

Índice de Harris

Una de las funciones de supervisión más ampliamente aplicadas se basa en el índice de Harris. La idea consiste en calcular la varianza de la salida del proceso, en línea o fuera de línea, y entonces compararla con la varianza mínima que se pueda conseguir. El problema se analizó en la Sección 2.6. El índice de Harris se define como

$$I_H = 1 - \frac{\sigma_{MV}^2}{\sigma_y^2},$$

donde σ_{MV}^2 es la varianza mínima de la salida del proceso y σ_y^2 es la varianza actual de la salida del proceso. El índice de Harris, I_H , toma valores entre cero y uno. Si el índice está próximo a cero, la varianza actual está cercana a la varianza mínima, lo que quiere decir que el lazo de control se comporta de manera satisfactoria. Si la varianza actual es grande, el índice de Harris está próximo a uno.

El método requiere que sea conocida la varianza mínima σ_{MV}^2 . Una característica agradable del método es que se puede determinar la varianza mínima sólo del retardo de tiempo, lo que significa que el modelado se puede realizar relativamente simple. Un inconveniente es que la varianza mínima normalmente no se puede alcanzar con un controlador tan sencillo como un controlador PID, lo que indica que es difícil determinar valores razonables del índice de Harris. Además,

incluso si es posible obtener control de mínima varianza, este control no es generalmente deseable ya que puede ser muy agresivo.

Por estas razones, se han presentado muchas variaciones del índice de Harris donde la varianza mínima σ_{MV}^2 se sustituye por la varianza obtenida utilizando otros objetivos de diseño y donde se tienen en cuenta las limitaciones de la estructura de control PID. El principal inconveniente de estos enfoques es que requieren un modelo más preciso del proceso.

Las herramientas de monitorización de rendimiento basadas en el enfoque del índice de Harris proporcionan información sobre algún rendimiento del lazo comparado con el rendimiento ideal. No existe intención de detectar alguna causa posible de un mal rendimiento. Hay otras herramientas de monitorización del rendimiento que no miran al rendimiento global, sino que tratan de detectar ciertos tipos de problemas. Algunos de estos se presentan en las próximas subsecciones.

Lazos de control oscilatorios

El problema más serio a nivel de lazo es que muchos lazos de control oscilan. Hay algunas causas posibles para estas oscilaciones; véase la Sección 10.2. Una razón podría ser que una carga oscilatoria está perturbando el lazo. Las perturbaciones de carga de baja frecuencia se eliminan eficientemente por el controlador, ya que un controlador con acción integral proporciona una alta ganancia del lazo a bajas frecuencias. Como el proceso normalmente tiene un carácter paso-bajo, se filtran las perturbaciones de carga de altas frecuencias por el proceso. Por lo tanto, los componentes de altas frecuencias en la señal de medida no se introducen normalmente en el proceso sino en el sensor o en las conexiones entre el sensor y el controlador. Debido a que no contienen información relevante alguna acerca del estado del proceso, se deberían filtrar por el controlador. También es importante no transferir estas señales a la salida del controlador, ya que pueden causar deterioro en el equipamiento de actuación.

Las perturbaciones con mucha energía cerca de la frecuencia última ω_u son demasiado rápidas para tratarse eficientemente por el controlador y son demasiado lentas para filtrarse. Estas perturbaciones podrían incluso amplificarse a causa de la realimentación.

Un controlador mal sintonizado puede ser otra razón de las oscilaciones, en particular en plantas no lineales donde un cambio en el punto de operación podría resultar en una ganancia del lazo demasiado alta. Sin embargo, los controladores en las plantas de control de procesos se sintonizan generalmente de manera conservadora y la mala sintonía del controlador no suele ser la causa más probable de las oscilaciones.

La razón más común de las oscilaciones en los lazos de control es, sin embargo, la fricción en la válvula, resultando en un movimiento de deslizamiento tal y como se analizó en la Sección 10.2.

Detección Las oscilaciones en los lazos de control se pueden detectar de varias formas. Una forma es hacer un análisis espectral de la señal de medida y buscar los picos del espectro. Una dificultad es que las oscilaciones generalmente están lejos de ser ondas sinusoidales, lo que quiere decir que no aparecen picos distintos en el espectro.

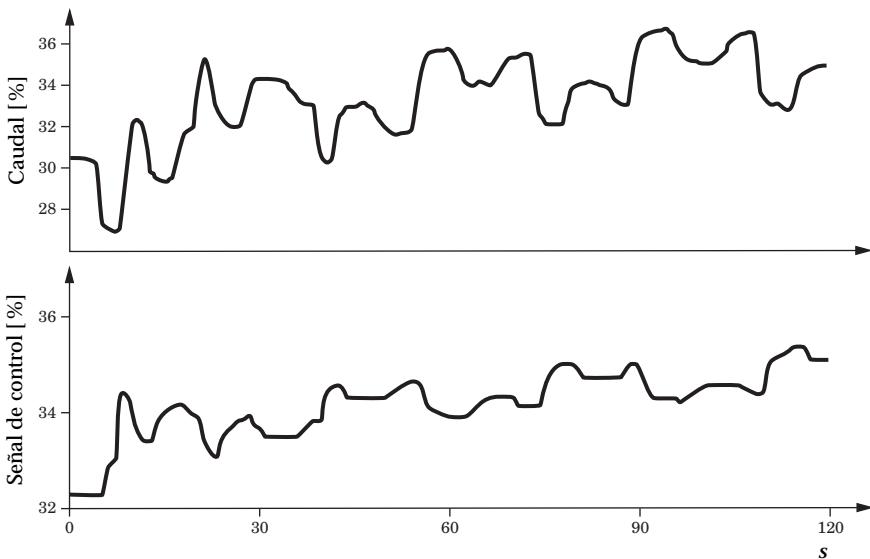


Figura 10.11 Movimiento de pegue y despegue en un lazo de control de flujo.

La Figura 10.11 muestra un registro del lazo de control de caudal de una fábrica de papel con alta fricción estática en la válvula. La figura muestra el resultado de un cambio en escalón en el punto de consigna. El control utilizado fue un controlador PI con ganancia $K = 0,30$ y tiempo integral $T_i = 34$ s. Obsérvese que las oscilaciones están lejos de una onda sinusoidal pura. Una resintonía del controlador dio los parámetros $K = 0,19$ y $T_i = 2$ s. ¡Se ve que el tiempo integral se ha disminuido de 34 s a 2 s! En la Figura 10.12 se muestra un experimento de respuesta a escalón utilizando los nuevos parámetros del controlador. El tiempo de un asentamiento es significativamente más corto que en la Figura 10.11. También es más obvio que las oscilaciones son realmente causadas por la fricción, ya que el patrón típico de la señal de medida está cercano a una señal cuadrada y la señal de control a una señal triangular.

Otro enfoque para detectar oscilaciones es investigar las características del error de control. La idea que hay detrás de este procedimiento de detección es estudiar la magnitud de la integral absoluta del error (*IAE*) entre cruces sucesivos por cero del error de control, por ejemplo,

$$IAE = \int_{t_{i-1}}^{t_i} |e(t)| dt, \quad (10.1)$$

donde t_{i-1} y t_i son dos instancias consecutivas de cruces por cero. Se asume que el controlador tiene acción integral, de forma que el error medio es cero.

Durante los períodos de un buen control, la magnitud del error de control es pequeña y los tiempos entre los cruces por cero son relativamente cortos. Esto quiere decir que los valores de *IAE* calculados con (10.1) son pequeños cuando el control es bueno.

Cuando ocurre un perturbación de carga, la magnitud de $e(t)$ aumenta, y exis-

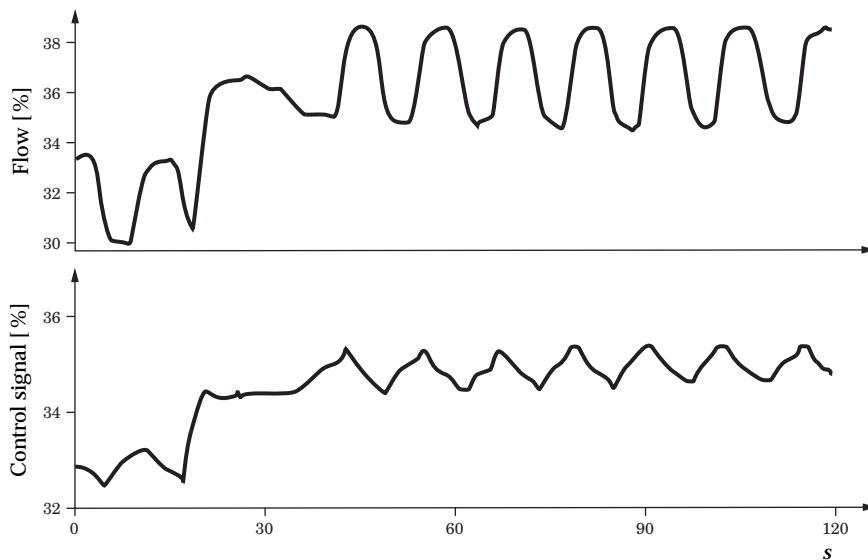


Figura 10.12 Movimiento de deslizamiento en un lazo de control de caudal con el controlador resintonizado.

te un período relativamente largo sin cruces por cero. Esto significa que el correspondiente valor IAE se hace grande.

Cuando el lazo de control comienza a oscilar, habrá una alta frecuencia de valores grandes de IAE . Esta observación se utiliza para detectar oscilaciones en el lazo de control.

EJEMPLO 10.1—CONTROL DE LA CONCENTRACIÓN DE PULPA

El ejemplo siguiente se ha tomado de una sección del control de concentración de pulpa en una fábrica de papel, donde la pulpa es diluida con agua para obtener una concentración deseada. La válvula de agua tenía una fricción elevada, y se conectó un procedimiento de detección de oscilación al controlador. El controlador era un PI con ganancia $K = 0,33$ y tiempo integral $T_i = 24$ s.

La Figura 10.13 muestra 10 minutos de datos del lazo de control de concentración. La primera gráfica muestra la salida del proceso, la concentración de pulpa en tanto por ciento. Debido a la alta fricción en la válvula de agua, el proceso está oscilando con una amplitud de un pequeño porcentaje. La primera gráfica muestra también una estimación del punto de consigna, ya que esta variable no fue registrada. La estimación se obtiene simplemente mediante un filtrado pasabajo de la salida del proceso.

La segunda gráfica muestra la señal de control en tanto por ciento. Es obvio que el controlador intenta eliminar sin éxito la oscilación.

La tercera gráfica muestra el IAE calculado entre sucesivos cruces por cero del error de control. La gráfica también muestra el IAE_{lim} , que es el límite de lo que se considera valores grandes de IAE . En esta implementación, el valor de IAE_{lim} se determina automáticamente de los parámetros del controlador en cada lazo. Los valores de IAE son significativamente más grandes que IAE_{lim} , lo que indica

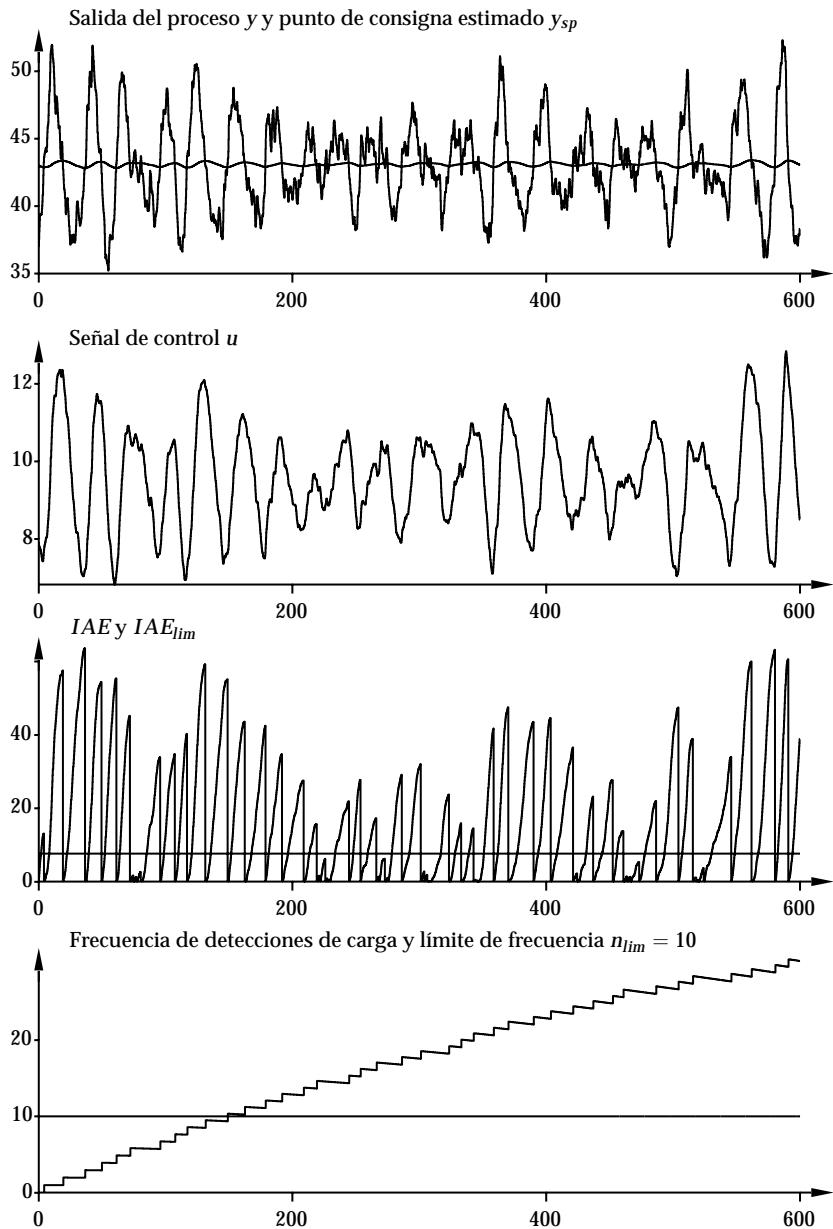


Figura 10.13 El procedimiento de detección de oscilación aplicado a un lazo de control de concentración de pulpa.

que el lazo está oscilando.

Finalmente, la cuarta gráfica muestra la velocidad de las detecciones de carga y el límite de velocidad $n_{lim} = 10$. La velocidad excede el límite después de unos tres minutos, y el procedimiento de detección genera una alarma.

Este ejemplo muestra cómo el procedimiento de detección de oscilación gestiona la detección de oscilaciones en los lazos de control. Las oscilaciones actuales

son fácilmente apreciables en la Figura 10.13. Sin embargo, los operadores de procesos rara vez tiene acceso a estas clases de gráficas, sino que generalmente poseen gráficas de barras de baja resolución. La presente oscilación ha estado presente durante un largo tiempo sin ser descubierta por los operadores de proceso.

□

Diagnosis Como el lazo de control puede oscilar por varias razones, es importante no sólo detectar la oscilación, sino también encontrar sus causas. Esto se puede realizar manualmente como se describió en la Sección 10.2.

Se han realizado intentos de desarrollar procedimientos para la diagnosis automática. Aquí se puede utilizar la diferencia en el espectro. Cuando un lazo de control oscila a causa de una ganancia del lazo demasiado alta, el error de control está generalmente cercano a una onda sinusoidal, lo que resulta en un único pico en el espectro. Lo mismo sucede en la mayoría de los casos cuando el lazo está oscilando a causa de perturbaciones externas. Sin embargo, cuando el lazo de control está oscilando debido a la fricción en la válvula, se pueden encontrar varios picos en el espectro.

Lazos de control lentos

Las oscilaciones en los lazos de control son comunes, pero la situación contraria también suele ser habitual, concretamente, que los lazos de control sean lentos debido a una sintonía conservadora. Esto causa grandes y largas desviaciones innecesarias del punto de consigna durante las perturbaciones de carga.

La razón principal para los controladores con sintonía conservadora es la falta de tiempo. Los ingenieros sintonizan los controladores hasta que ellos consideran que es «suficiente». No tienen tiempo para optimizar el control. Muchos controladores se sintonizan sólo una vez cuando se instalan y no se vuelven a sintonizar nunca más. Para mantener la estabilidad cuando cambian las condiciones de operación, los controladores se sintonizan para el «peor caso». Una mejor solución sería, por supuesto, utilizar una planificación de ganancia y quizás adaptación. Cuando se resintoniza un controlador, es principalmente porque las condiciones de proceso causan control oscilatorio. En otras palabras, cuando los controladores se resintonizan, se desintonizan. Cuando las condiciones del proceso cambian a control lento, el controlador no se suele resintonizar otra vez.

Detección La Figura 10.14 muestra dos respuestas en la entrada perturbaciones de carga en forma de cambios en escalón en la entrada del proceso. Una respuesta es buena, con una rápida recuperación sin ninguna sobreelongación. La segunda respuesta es muy lenta. Un hecho que caracteriza la segunda respuesta es que existe un largo período donde ambos, la salida del proceso y la señal de control u , van a la deriva lentamente en la misma dirección. Esta característica se utiliza para la detección.

Ambas respuestas tienen una fase inicial donde las dos señales van en direcciones contrarias, por ejemplo, $\Delta u \Delta y < 0$, donde Δu y Δy son los incrementos de las dos señales. Lo que caracteriza la respuesta lenta es que después de esta fase inicial existe un largo período de tiempo donde la correlación entre los dos

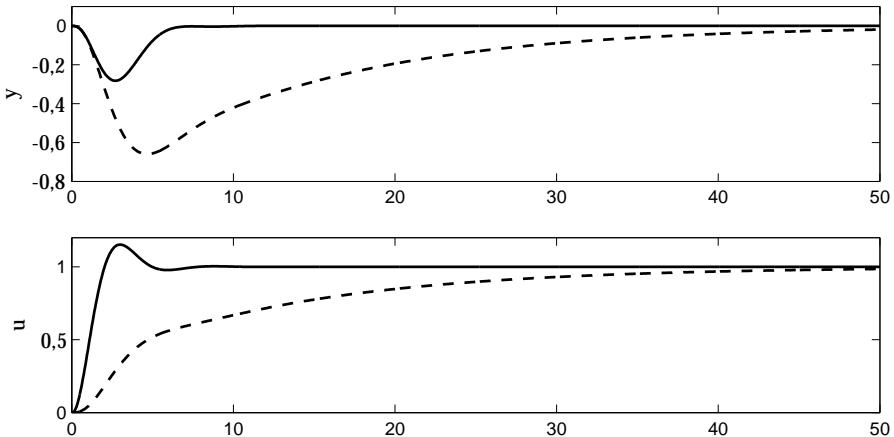


Figura 10.14 Una respuesta buena (líneas continuas) y una respuesta lenta (líneas a tramos) a un cambio en escalón en la perturbación de carga a la entrada del proceso.

incrementos de señal es positiva. Esta observación forma la base del índice Idle, que expresa la relación entre los tiempos de correlación positiva y negativa entre los incrementos de señal.

Para formar el índice Idle, se calculan primero los períodos de tiempo donde las correlaciones entre los incrementos de señal son positivos y negativos. El siguiente procedimiento se actualiza cada período de muestreo

$$t_{\text{pos}} = \begin{cases} t_{\text{pos}} + h & \text{si } \Delta u \Delta y > 0 \\ t_{\text{pos}} & \text{si } \Delta u \Delta y \leq 0 \end{cases}$$

$$t_{\text{neg}} = \begin{cases} t_{\text{neg}} + h & \text{si } \Delta u \Delta y < 0 \\ t_{\text{neg}} & \text{si } \Delta u \Delta y \geq 0, \end{cases}$$

donde h es el período de muestreo. El índice Idle I_I se define entonces por

$$I_I = \frac{t_{\text{pos}} - t_{\text{neg}}}{t_{\text{pos}} + t_{\text{neg}}}. \quad (10.2)$$

Obsérvese que I_I está limitado al intervalo $[-1, 1]$. Un valor positivo de I_I cercano a uno significa que el control es lento. El índice Idle para la respuesta lenta en la Figura 10.14 es $I_I = 0,82$. Un valor negativo de I_I cercano a -1 se puede obtener en un lazo de control bien sintonizado. El índice Idle para la respuesta buena en la Figura 10.14 es $I_I = -0,63$. Sin embargo, valores negativos de Idle cercanos a -1 se obtienen también en lazos de control oscilatorios. Por lo tanto, es deseable combinar el cálculo del índice Idle con un procedimiento de detección de oscilación como el descrito anteriormente.

El cálculo del índice Idle se puede realizar fuera de línea o en línea utilizando una versión recursiva. Como el método se basa en las características de los incrementos de señal, es sensible al ruido. Por consiguiente, es importante filtrar las señales adecuadamente antes de que sean diferenciadas.

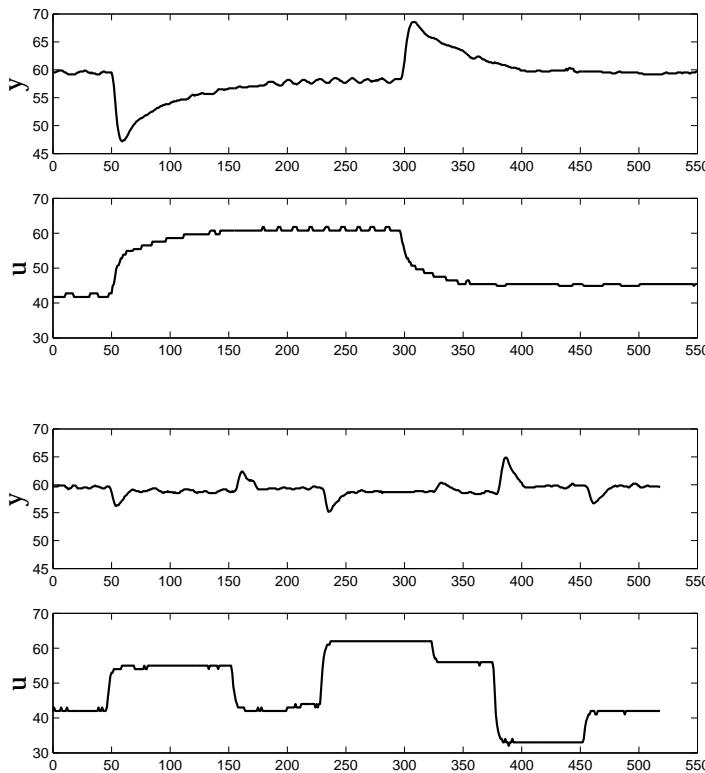


Figura 10.15 Control de un intercambiador de calor. Las gráficas muestran las respuestas a perturbaciones de carga para un lazo de control lento con índice Idle $I_I = 0,8$ (parte superior) y un lazo sintonizado adecuadamente con índice Idle $I_I = 0,3$ (parte inferior).

EJEMPLO 10.2—CONTROL DE UN INTERCAMBIADOR DE CALOR

Este ejemplo se ha tomado de un intercambiador de calor industrial. El objetivo de control es controlar la temperatura del agua del lado secundario controlando el caudal de vapor de agua en el lado primario.

Las gráficas superiores de la Figura 10.15 muestran las respuestas a perturbaciones de carga obtenidas con un controlador PI sintonizado de manera conservadora. Los parámetros del controlador fueron $K = 0,01$ y $T_i = 30\text{s}$. Las señales son relativamente ruidosas a causa de la baja resolución, 1 por ciento, de la salida del controlador. El control es lento, algo que también se refleja en el índice Idle que da un valor de $I_I = 0,8$.

La estructura del controlador se cambió a un controlador PID y se sintonizó adecuadamente, resultando en los parámetros $K = 0,025$, $T_i = 8\text{s}$ y $T_d = 2\text{s}$. El comportamiento del control mejorado se muestra en las gráficas inferiores de la Figura 10.15. La recuperación tras las perturbaciones de carga es significativamente más rápida, aunque sin una sobreelongación apreciable. La ganancia integral k_i se incrementa casi en un factor de diez. Las mejoras se observan también en el índice Idle, que se redujo a $I_I = 0,3$. \square

10.5 Sintonía y diagnosis integradas

Los procedimientos de diagnosis están relacionados con las técnicas adaptativas de varias formas. Hemos señalado la importancia de comprobar las válvulas antes de aplicar un procedimiento de sintonía automática. Si esto no se hace, el procedimiento de sintonía automática no proporcionará los parámetros del controlador adecuados. Por esta razón, sería deseable tener estas comprobaciones incorporadas en los procedimientos de sintonía automática. Tales dispositivos no están todavía disponibles, y las comprobaciones adecuadas, por lo tanto, las debe realizar el operador.

Los métodos de detección en línea están relacionados con el controlador adaptativo continuo. El control adaptativo monitoriza el rendimiento del lazo de control y los cambios de los parámetros del controlador, si cambia la dinámica del proceso. Los procedimientos de evaluación del rendimiento también monitorizan el rendimiento del lazo de control. Si cambia la dinámica del proceso, estos métodos proporcionan una alarma en lugar de modificar los parámetros del controlador. Como un ejemplo, en la Figura 10.3 hemos visto que es importante determinar *por qué* el rendimiento se ha modificado antes de realizar cualquier acción. La mayoría de los controladores adaptativos aplicados a un proceso con fricción estática desintonizarán el controlador, debido que interpretan las oscilaciones como una mala sintonía del controlador. Consecuentemente, es deseable dotar a los controladores adaptativos con métodos de detección en línea, de forma que se puedan detectar otras razones de un mal rendimiento en el lazo de control que no sean sólo por mala sintonía. La carencia de estas clases de procedimientos de detección en controladores adaptativos es quizás la principal razón por la que hay relativamente pocas aplicaciones de control adaptativo continuo disponibles hoy día.

10.6 Resumen

Es importante realizar una evaluación del lazo de control antes de sintonizar el controlador. Esta evaluación incluye comprobaciones del equipo tales como sensores y válvulas, rangos de señal, no linealidades y perturbaciones.

Cuando se llevan a cabo la evaluación del lazo y la sintonía del controlador, el lazo de control se debería comportar bien. Debido a los cambios en el proceso y su operación, el lazo de control se puede degradar después de algún tiempo de operación. Es, por lo tanto, importante supervisar los lazos de control. Esto se realiza tradicionalmente por operarios, pero cada vez más se utilizan métodos de supervisión automática en el control de proceso.

En esta sección, se han presentado algunos ejemplos de herramientas de monitorización del lazo y del rendimiento. La sección sólo ha aportado un breve resumen de esta área, centrándose sólo en métodos de lazos simples. Recientemente, se han realizado muchos intentos para desarrollar métodos de monitorización del rendimiento de secciones del proceso incluyendo varios lazos de control. Sin embargo, estos procedimientos son rara vez de carácter general, sino que se desarrollan para plantas específicas.

10.7 Notas y referencias

El primer trabajo sobre detección de fallos fue desarrollado por [169]. Los problemas asociados con las válvulas de control se trasladaron a una audiencia más amplia a principios de los noventa; ver [107, 50]. En ese tiempo, hubo además una coincidencia de que era beneficioso evaluar el rendimiento de lazos de control; véase [310, 311, 37]. El índice de Harris [96], [161] se basa en la comparación con el rendimiento obtenido por el control de mínima varianza [32]. El concepto ha sido extendido y aplicado en varias aplicaciones de control de procesos; ver por ejemplo [93, 325, 159, 201, 202, 159, 262, 223, 160, 338]. El procedimiento de detección de oscilación se describe en [142] y [337], y el índice Idle se presentó en [144]. En [286, 178, 174] se presentan buenas revisiones sobre esta área. Un método para reducir el efecto de la fricción en válvulas se desarrolló por [146].

11

Interacción

11.1 Introducción

Hasta el momento, nos hemos centrado en el control de lazos simples con un sensor, un actuador y un controlador. En aplicaciones prácticas, un sistema de control puede tener muchos lazos, algunas veces miles. A pesar de esto, un sistema de control grande puede tratarse como un solo lazo si la interacción entre los lazos es despreciable. Sin embargo, existen situaciones cuando puede haber una interacción considerable entre los diferentes lazos de control. Un caso típico es cuando se combinan varias corrientes para obtener una mezcla deseada. En tal caso, está claro que los lazos interactúan. Otros casos son el control de calderas, máquinas de papel, columnas de destilación, reactores químicos, intercambiadores de calor, redes de distribución de vapor, sistemas de transmisión, y sistemas de aire acondicionado. Los procesos que tienen muchas variables de control y muchas variables de medida se denominan sistemas de múltiples entradas y múltiples salidas (MIMO). Debido a las interacciones, puede ser difícil controlar tales sistemas como un solo lazo.

Un tratamiento razonablemente completo de sistemas multivariados está fuera del objetivo de este libro. En este capítulo, trataremos brevemente algunas cuestiones sobre interacción de lazos que son de relevancia particular para el control PID. La Sección 11.2 presenta ejemplos sencillos que ilustran lo que puede ocurrir en la interacción de lazos. En particular, se muestra que los parámetros del controlador en un lazo pueden tener influencia significativa sobre la dinámica de otros lazos. También se introduce la matriz de ganancias relativas de Bristol, que es una manera sencilla de caracterizar las interacciones. Se trata el problema del emparejamiento de entradas y salidas, y se muestra que las interacciones pueden generar ceros en un sistema multivariado. En la Sección 11.3, presentamos un método de diseño basado en desacoplo, que es una extensión natural de los métodos de sintonía para sistemas de una entrada y una salida. La Sección 11.4 presenta los problemas que ocurren en los sistemas de transmisión con motores paralelos. El capítulo finaliza con un resumen y referencias.

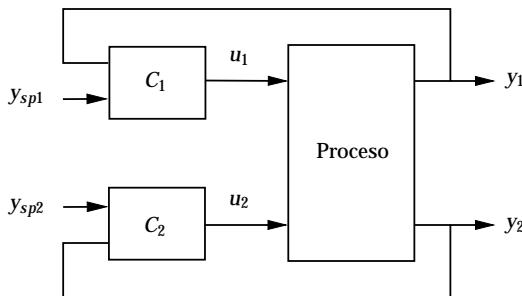


Figura 11.1 Diagrama de bloques de un sistema con dos entradas y dos salidas (TITO).

11.2 Interacción de lazos sencillos

En esta sección, ilustraremos algunos efectos de la interacción en el caso más simple de un sistema con dos entradas y dos salidas. Tal sistema se denomina sistema TITO. El sistema se puede representar por las ecuaciones

$$\begin{aligned} Y_1(s) &= p_{11}(s)U_1(s) + p_{12}U_2(s) \\ Y_2(s) &= p_{21}(s)U_1(s) + p_{22}U_2(s), \end{aligned} \quad (11.1)$$

donde $p_{ij}(s)$ es la función de transferencia desde la entrada j -ésima a la salida i -ésima. Las funciones de transferencia p_{11} , $p_{12}(s)$, $p_{21}(s)$ y p_{22} se pueden combinar en la matriz

$$P(s) = \begin{pmatrix} p_{11}(s) & p_{12}(s) \\ p_{21}(s) & p_{22}(s) \end{pmatrix}, \quad (11.2)$$

que se denomina función de transferencia o matriz de función de transferencia del sistema. Algunos efectos sobre interacción se ilustrarán con un ejemplo.

EJEMPLO 11.1—EFECTOS DE LA INTERACCIÓN

Considérese el sistema descrito por el diagrama de bloques de la Figura 11.1. El sistema tiene dos entradas y dos salidas. Existen dos controladores, el controlador C_1 controla la salida y_1 con la entrada u_1 y el controlador C_2 controla la salida y_2 con la entrada u_2 . Un efecto de la interacción es que la sintonía de un lazo puede influir en el otro lazo. Esto se muestra en la Figura 11.2, que muestra una simulación del primer lazo cuando C_1 es un controlador PI y $C_2 = k$ es un controlador proporcional.

El ejemplo muestra que la ganancia del segundo lazo tiene una influencia significativa sobre el comportamiento del primer lazo. La respuesta del primer lazo es buena cuando el segundo lazo se desconecta, $k_2 = 0$, pero el sistema llega a ser más lento cuando se incrementa la ganancia del segundo lazo. El sistema es inestable para $k_2 = 0,8$.

Un simple análisis nos permite entender lo que ocurre. En este caso particular,

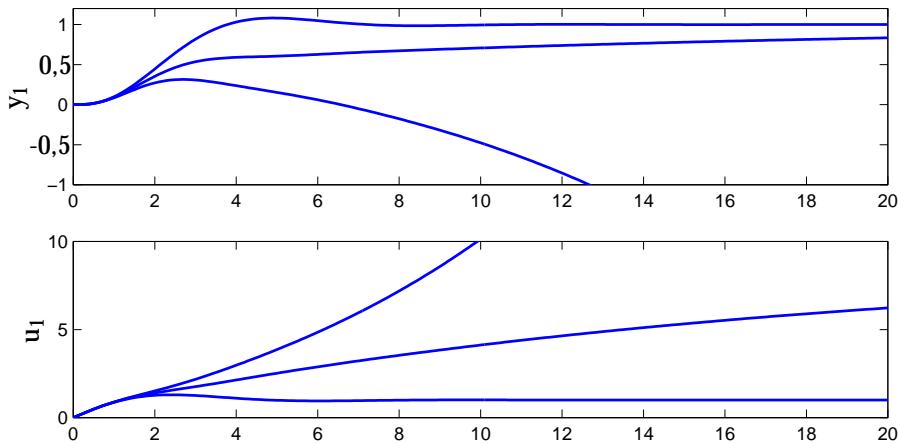


Figura 11.2 Simulación de las respuestas a escalones en los puntos de consigna para el lazo 1 del sistema en la Figura 11.1. El controlador C_1 es un controlador PI con ganancias $k_1 = 1$, $k_i = 1$, y C_2 es un controlador proporcional con ganancias $k_2 = 0, 0,8$ y $1,6$.

el sistema se describe por

$$\begin{aligned} Y_1(s) &= \frac{1}{(s+1)^2} U_1(s) + \frac{2}{(s+1)^2} U_2(s) \\ Y_2(s) &= \frac{1}{(s+1)^2} U_1(s) + \frac{1}{(s+1)^2} U_2(s). \end{aligned}$$

La realimentación en el segundo lazo es $U_2(s) = -k_2 Y_2(s)$. Introduciendo esto en la segunda ecuación da

$$U_2(s) = -\frac{k_2}{s^2 + 2s + k_2 + 1} U_1(s),$$

y la inserción de esta expresión para $U_2(s)$ en la primera ecuación resulta en

$$Y_1(s) = g_{11}^{cl}(s) U_1(s) = \frac{s^2 + 2s + 1 - k_2}{(s+1)^2(s^2 + 2s + 1 + k_2)} U_1(s).$$

Esta ecuación muestra claramente que la ganancia k_2 en el segundo lazo tiene un efecto significativo sobre la dinámica que relaciona u_1 e y_1 . La ganancia estática es

$$g_{11}^{cl}(0) = \frac{1 - k_2}{1 + k_2}.$$

Obsérvese que la ganancia decrece cuando k_2 aumenta y que la ganancia se vuelve negativa para $k_2 > 1$. \square

El ejemplo indica que existe una necesidad de tener alguna manera de determinar si las interacciones pueden causar dificultades. Una medida simple de interacción se comentará a continuación.

Matriz de ganancias relativas de Bristol

Una manera sencilla de investigar el efecto de la interacción consiste en analizar cómo se ve influenciada la ganancia estática del proceso de un lazo por las ganancias en los otros lazos. Considérese primero el sistema con dos entradas y dos salidas mostrado en la Figura 11.1. Investigaremos cómo la ganancia estática en el primer lazo es influenciada por el controlador en el segundo lazo. Para evitar hacer hipótesis específicas sobre el controlador, Bristol asumió que el segundo lazo estaba en control perfecto, lo que quiere decir que la salida del segundo lazo es cero. Entonces, de (11.1) se deduce que

$$\begin{aligned} Y_1(s) &= p_{11}(s)U_1(s) + p_{12}U_2(s) \\ 0 &= p_{21}(s)U_1(s) + p_{22}U_2(s). \end{aligned}$$

Eliminando $U_2(s)$ de la primera ecuación se obtiene

$$Y_1(s) = \frac{p_{11}(s)p_{22}(s) - p_{12}(s)p_{21}(s)}{p_{22}(s)} U_1(s).$$

La razón de las ganancias estáticas del lazo 1 cuando el segundo lazo está abierto y cuando el segundo lazo está cerrado es por tanto

$$\lambda = \frac{p_{11}(0)p_{22}(0)}{p_{11}(0)p_{22}(0) - p_{12}(0)p_{21}(0)}. \quad (11.3)$$

El parámetro λ se denomina *índice de interacción de Bristol* para sistemas TITO. Obsérvese que el índice se refiere a condiciones estáticas. En la práctica, esto se puede interpretar también como la interacción de señales a bajas frecuencias. No existe interacción si $p_{12}(0)p_{21}(0) = 0$, lo que implica que $\lambda = 1$. Valores pequeños o negativos de λ indican que existen interacciones. Considérese, por ejemplo el sistema del Ejemplo 11.1 donde el índice de interacción es $\lambda = -1$, que indica que las interacciones plantean dificultades severas.

El índice de interacción se puede generalizar para sistemas con muchas entradas y muchas salidas. La idea es comparar las ganancias estáticas de una salida cuando los otros lazos están abiertos con las ganancias cuando todas las otras salidas son cero. El resultado se puede resumir en la *matriz de ganancias relativas de Bristol* (RGA) que se define como

$$R = P(0) .* P^{-T}(0), \quad (11.4)$$

donde $P(0)$ es la ganancia estática del sistema, $P^{-T}(0)$ la traspuesta de la inversa de $P(0)$, y $.*$ denota la multiplicación de matrices elemento a elemento. El elemento r_{ij} es la razón entre las ganancias estáticas en lazo abierto y en lazo cerrado de la señal de entrada u_j a la salida y_i . Se puede observar que la matriz R es simétrica y que todas las filas y columnas suman uno. Obsérvese que la matriz de ganancias relativas de Bristol sólo captura el comportamiento del proceso a bajas frecuencias.

La matriz de ganancias relativas para el sistema (11.1) viene dada por

$$R = \begin{pmatrix} \lambda & 1-\lambda \\ 1-\lambda & \lambda \end{pmatrix}, \quad (11.5)$$

donde λ es el índice de interacción (11.3). No existe interacción si $\lambda = 1$. Esto quiere decir que el segundo lazo no tiene impacto sobre el primer lazo y viceversa. Si λ está entre 0 y 1, el lazo cerrado tiene una ganancia más alta que el lazo abierto. El efecto es más severo para $\lambda = 0,5$. Si λ es mayor que 1, el lazo cerrado tiene una ganancia inferior que la del lazo abierto. Cuando λ es negativa, la ganancia del primer lazo cambia de signo cuando se cierra el segundo lazo. Por lo tanto, el efecto de las interacciones es así severo.

Emparejamiento

Para controlar un sistema lazo a lazo, primero debemos decidir cómo se deberían conectar los controladores, por ejemplo, si y_1 en la Figura 11.1 debería ser controlarse por u_1 o u_2 . Esto se denomina el *problema del emparejamiento*.

La matriz de ganancia relativa se puede utilizar como una guía para el emparejamiento. No existe interacción si $\lambda = 1$. Si $\lambda = 0$, tampoco existe interacción, pero los lazos se deberían intercambiar. Los lazos deberían intercambiarse cuando $\lambda < 0,5$. Si $0 < \lambda < 1$, la ganancia del primer lazo aumenta cuando se cierra el segundo lazo, y si $\lambda > 1$ la ganancia del lazo cerrado es menor que la ganancia del lazo abierto. Bristol recomendó que el emparejamiento se debería realizar de forma que las ganancias relativas correspondientes sean positivas y tan cercanas a uno como fuese posible. El emparejamiento de señales con ganancias relativas negativas debería evitarse. Si las ganancias están fuera del intervalo $0,67 < \lambda < 1,5$, el desacoplamiento puede mejorar el control significativamente. Ilustraremos el emparejamiento con un ejemplo.

EJEMPLO 11.2—EMPAREJAMIENTO DE SEÑALES

Considérese el sistema del Ejemplo 11.1. La matriz de ganancias estáticas es

$$P(0) = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}.$$

Su inversa es

$$P^{-1}(0) = \begin{pmatrix} -1 & 2 \\ 1 & -1 \end{pmatrix},$$

y la matriz de ganancias relativas viene dada por

$$R = P(0) * P^{-T}(0) = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix} * \begin{pmatrix} -1 & 1 \\ 2 & -1 \end{pmatrix} = \begin{pmatrix} -1 & 2 \\ 2 & -1 \end{pmatrix},$$

que quiere decir que $\lambda = -1$. La regla de emparejamiento dice que y_1 se debería emparejar con u_2 .

Cuando $u_1 = -k_2 y_2$, la relación entre u_2 y y_1 viene dada por

$$Y_1(s) = g_{12}^{cl}(s) U_2(s) = \frac{2s^2 + 4s + 2 + k_2}{(s+1)^2(s^2 + 2s + 1 + k_2)} U_2(s),$$

y la ganancia estática es

$$g_{12}^{cl}(0) = \frac{2 + k_2}{1 + k_2}.$$

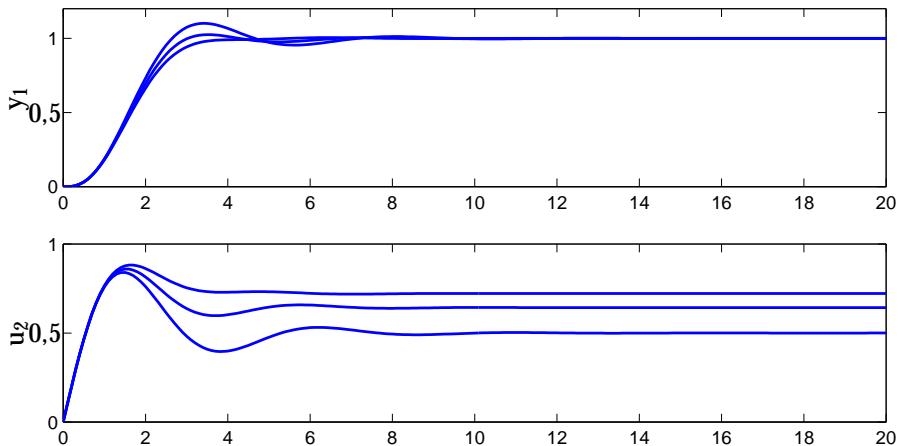


Figura 11.3 Simulación de las respuestas a escalón en el punto de consigna para y_1 del sistema en la Figura 11.1 cuando los lazos son commutados tal que el controlador para y_1 es $U_2 = C_1(s)(Y_{sp1} - Y_1)$ y el controlador para y_2 es $u_1 = -k_2 y_2$ with $k_2 = 0, 0,8$ y $1,6$. El controlador C_1 es un controlador PI con ganancias $k_1 = 1$, $k_i = 1$.

La ganancia decrece a medida que k_2 aumenta, pero nunca es negativa para $k_2 > 0$. Existe interacción pero no tan severa como para el emparejamiento de y_1 con u_1 . En la Figura 11.3 se muestran las propiedades del sistema en lazo cerrado. Una comparación con la Figura 11.2 muestra que existe una reducción drástica en la interacción cuando se comutan las entradas. \square

Ceros multivariados

En la Sección 4.3 encontramos que los ceros en el semiplano derecho imponían restricciones severas sobre el rendimiento alcanzable. Para sistemas de una entrada y una salida, los ceros se pueden encontrar por simple inspección. Sin embargo, los ceros de sistemas multivariados se pueden crear también por la interacción. Una definición de ceros que también funciona para sistemas multivariados es que los ceros son los polos del sistema inverso. Los ceros del sistema (11.1) vienen dados por

$$\det P(s) = p_{11}(s)p_{22}(s) - p_{12}(s)p_{21}(s) = 0. \quad (11.6)$$

Los ceros en el semiplano derecho son de particular interés porque imponen limitaciones sobre el rendimiento alcanzable. Mostraremos esto mediante un ejemplo.

EJEMPLO 11.3—SISTEMA DE ROSEN BROCK

Considérese un sistema con la función de transferencia

$$P(s) = \begin{pmatrix} p_{11}(s) & p_{12}(s) \\ p_{21}(s) & p_{22}(s) \end{pmatrix} = \begin{pmatrix} \frac{1}{s+1} & \frac{2}{s+3} \\ \frac{1}{s+1} & \frac{1}{s+1} \end{pmatrix}. \quad (11.7)$$

Las dinámicas de los subsistemas son muy benévolas. No hay limitaciones de dinámicas en el control de ningún lazo individual. La matriz de ganancias relativas es

$$R = \begin{pmatrix} 1 & 2/3 \\ 1 & 1 \end{pmatrix} \cdot * \begin{pmatrix} 3 & -3 \\ -2 & 3 \end{pmatrix} = \begin{pmatrix} 3 & -2 \\ -2 & 3 \end{pmatrix},$$

que muestra que existen interacciones significativas. Utilizando las reglas para emparejamiento, encontramos que es razonable emparejar u_1 con y_1 y u_2 con y_2 . Como $\lambda > 1,5$, podemos esperar dificultades a causa de la interacción. De (11.6) se deduce que los ceros del sistema vienen dados por

$$\det P(s) = \frac{1}{s+1} \left(\frac{1}{s+1} - \frac{2}{s+3} \right) = \frac{1-s}{(s+1)^2(s+3)} = 0.$$

Hay un cero en $s = 1$ en el semiplano derecho, y por lo tanto podemos esperar dificultades cuando se diseñan los lazos de control para tener un ancho de banda mayor que $\omega_0 = 1$.

Considérese, por ejemplo, el problema de controlador la variable y_1 . Si el segundo lazo está abierto, podemos alcanzar una respuesta muy rápida con un controlador PI. Cuando el segundo lazo está cerrado existen, sin embargo, limitaciones severas en el rendimiento a causa de las interacciones, y el lazo de control tiene que ser desintonizado. La Figura 11.4 muestra las respuestas obtenidas con controladores que tienen ganancias $k = 2$ y $k_i = 2$ en ambos lazos. En la figura, hemos realizado primero un escalón unitario en el punto de consigna del primer controlador y después un cambio en escalón en el segundo controlador. La figura muestra que las interacciones son considerables. El sistema llega a ser inestable si la ganancia se aumenta en un factor de 3. \square

El Ejemplo 11.3 ilustra que un sistema multivariable que parece inocente puede tener ceros en el semiplano derecho. El caso contrario también es posible como se muestra en el siguiente ejemplo.

EJEMPLO 11.4—INTERACCIÓN BENÉFICA

Considérese el sistema

$$P(s) = \begin{pmatrix} p_{11}(s) & p_{12}(s) \\ p_{21}(s) & p_{22}(s) \end{pmatrix} = \begin{pmatrix} \frac{s-1}{(s+1)(s+2)} & \frac{s}{(s+1)(s+2)} \\ \frac{-6}{(s+1)(s+2)} & \frac{s-2}{(s+1)(s+2)} \end{pmatrix}. \quad (11.8)$$

El sistema tiene la matriz de ganancias relativas

$$R = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

que indica que y_1 se debería emparejar con u_1 y que y_2 se debería emparejar con u_2 . El sistema multivariable no tiene ceros. Por lo tanto, tenemos la situación interesante de que existen limitaciones severas para controlar tanto el primer como el segundo lazo individualmente a causa de los ceros en el semiplano derecho en

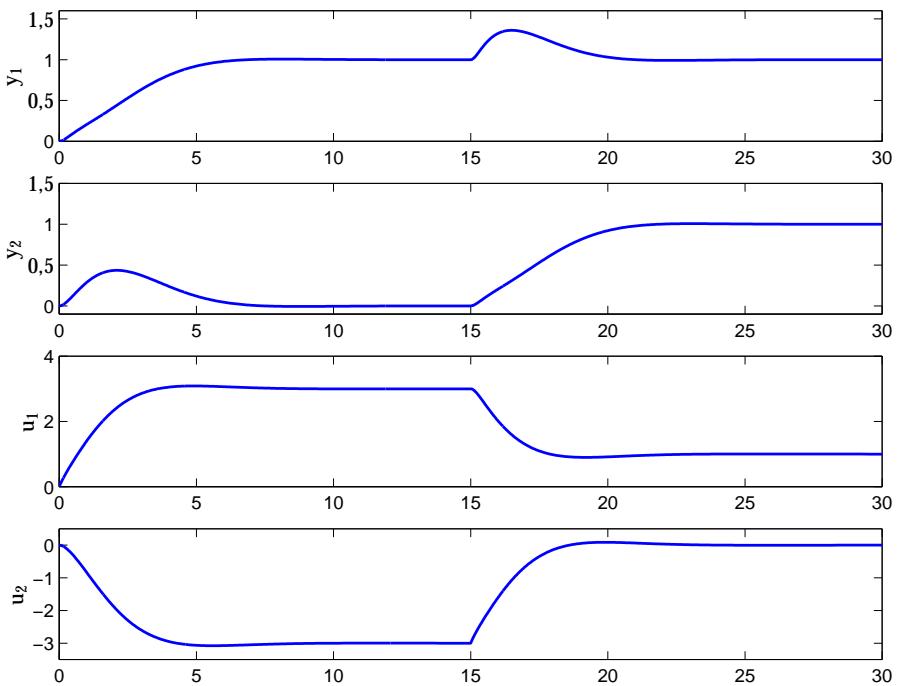


Figura 11.4 Respuestas a un escalón del proceso (11.7) con un control PI en ambos lazos. Ambos controladores tienen las ganancias $k = 2$ y $k_i = 2$. Primero se aplica un escalón en y_{sp1} en el instante de tiempo 0, y un escalón y_{sp2} se aplica a continuación en el instante de tiempo 15.

los elementos p_{11} y p_{22} . Como el sistema multivariable no tiene ningún cero en el semiplano derecho, es posible controlar el sistema multivariable con un ancho de banda alto. Esto se muestra en la Figura 11.4, donde ambos lazos son controlados con controladores PI con ganancias $k = 100$ y $k_i = 2000$. Nótese la rápida respuesta del sistema. Sin embargo, una dificultad es que el sistema llega a ser inestable si uno de los lazos se rompe. \square

11.3 Desacoplo

El desacoplo es una manera sencilla de tratar con la dificultad creada por las interacciones en los lazos. La idea consiste en diseñar un controlador que reduzca el efecto de la interacción. Idealmente, los cambios en un punto de consigna deberían afectar sólo a la salida del proceso correspondiente. Esto se puede llevar a cabo mediante un precompensador que mezcle las señales enviadas desde el controlador a las entradas del proceso. Daremos detalles para sistemas con dos entradas y dos salidas, pero el método se puede aplicar a sistemas con muchas entradas y muchas salidas.

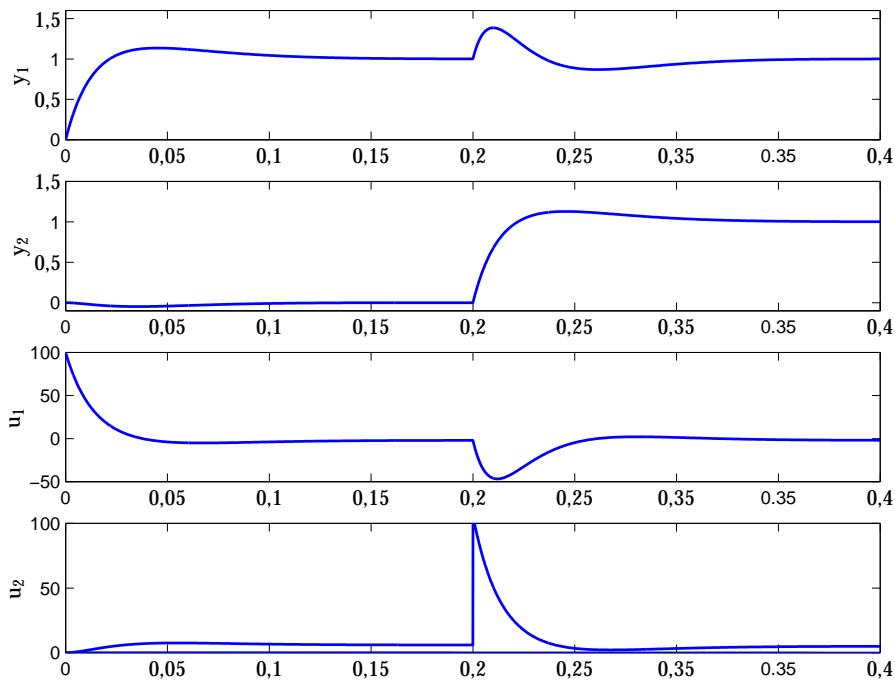


Figura 11.5 Respuestas a un escalón del control PI del proceso (11.8) cuando ambos lazos están cerrados. Se utilizan controladores PI con ganancias $k = 100$ y $k_i = 2000$ en ambos lazos.

Supóngase que el proceso tiene la función de transferencia (11.2) y que $P(0)$ es no singular. Primero introducimos un desacoplador estático $u = D\bar{u}$, donde D es una matriz constante

$$D = \begin{pmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{pmatrix}.$$

La función de transferencia desde \bar{u} a y viene entonces dada por $P(s)D$. La elección

$$D = P^{-1}(0) = \frac{1}{\det P(0)} \begin{pmatrix} p_{22}(0) & -p_{12}(0) \\ -p_{21}(0) & p_{11}(0) \end{pmatrix} \quad (11.9)$$

hace que $P(0)D$ sea la matriz identidad. El sistema $P(s)D$ está, por tanto, estéticamente desacoplado, y el acoplamiento es pequeño a bajas frecuencias. El acoplamiento permanece pequeño si el sistema se controla con controladores desacoplados, a condición de que los anchos de banda de los lazos del control sean suficientemente pequeños.

Suponiendo que los controladores son controladores PID, encontramos que el controlador estéticamente desacoplado se describe por

$$\begin{pmatrix} U_1(s) \\ U_2(s) \end{pmatrix} = \begin{pmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{pmatrix} \begin{pmatrix} \bar{c}_1(s)Y_{sp1}(s) - c_1(s)Y_1(s) \\ \bar{c}_2(s)Y_{sp2}(s) - c_2(s)Y_2(s) \end{pmatrix},$$

donde U es la señal de control, Y la salida del proceso, e Y_{sp} el punto de consigna.

Los controladores son controladores PID con ponderación del punto de consigna, por lo tanto,

$$c_i = k_{Pi} + \frac{k_{Ii}}{s} + k_{Di}s, \quad \bar{c}_i = b_i k_{Pi} + \frac{k_{Ii}}{s},$$

donde b_i es el peso del punto de consigna. Los pesos del punto de consigna influyen en la interacción entre los lazos. Escogiendo $b_i = 0$ se obtiene la interacción más pequeña.

El sistema desacoplado

La función de transferencia del sistema desacoplado es $Q(s) = P(s)D$, donde

$$\begin{aligned} q_{11}(s) &= \frac{p_{11}(s)p_{22}(0) - p_{12}(s)p_{21}(0)}{\det P(0)} \\ q_{12}(s) &= \frac{p_{12}(s)p_{11}(0) - p_{12}(0)p_{11}(s)}{\det P(0)} \\ q_{21}(s) &= \frac{p_{21}(s)p_{22}(0) - p_{21}(0)p_{22}(s)}{\det P(0)} \\ q_{22}(s) &= \frac{p_{22}(s)p_{11}(0) - p_{21}(s)p_{12}(0)}{\det P(0)}. \end{aligned}$$

Del desarrollo se deduce que $Q(0)$ es la matriz identidad. El desarrollo en serie de Taylor de la función de transferencia $Q(s)$ para pequeños $|s|$ da lugar a

$$Q(s) \approx \begin{pmatrix} 1 & \kappa_{12}s \\ \kappa_{21}s & 1 \end{pmatrix}$$

para algunas constantes κ_{12} y κ_{21} . Para bajas frecuencias ω , los elementos diagonales de $Q(s)$ son iguales a 1, y los elementos fuera de la diagonal son proporcionales a s . Si el ancho de banda del controlador PID descentralizado es suficientemente bajo, los términos fuera de la diagonal serán por lo tanto pequeños, y el sistema estará aproximadamente desacoplado. El sistema en lazo cerrado se puede describir por

$$\begin{pmatrix} 1 + q_{11}c_1 & q_{12}c_2 \\ q_{21}c_1 & 1 + q_{22}c_2 \end{pmatrix} Y = \begin{pmatrix} q_{11}\bar{c}_1 & q_{12}\bar{c}_2 \\ q_{21}\bar{c}_1 & q_{22}\bar{c}_2 \end{pmatrix} Y_{sp},$$

donde la dependencia sobre s se suprime para simplificar la notación. Esta ecuación se puede escribir como

$$Y = \tilde{H}Y_{sp},$$

donde

$$\begin{aligned} \tilde{h}_{11} &= \frac{q_{11}\bar{c}_1(1 + q_{22}c_2) - q_{12}q_{21}\bar{c}_1c_2}{(1 + q_{11}c_1)(1 + q_{22}c_2) - q_{12}q_{21}c_1c_2} \\ \tilde{h}_{12} &= \frac{q_{12}\bar{c}_2(1 + q_{22}c_2) - q_{12}q_{22}\bar{c}_2c_2}{(1 + q_{11}c_1)(1 + q_{22}c_2) - q_{12}q_{21}c_1c_2} \\ \tilde{h}_{21} &= \frac{q_{21}\bar{c}_1(1 + c_1q_{11}) - q_{11}q_{21}c_1\bar{c}_1}{(1 + q_{11}c_1)(1 + q_{22}c_2) - q_{12}q_{21}c_1c_2} \\ \tilde{h}_{22} &= \frac{q_{22}\bar{c}_2(1 + q_{11}c_1) - q_{12}q_{21}c_1\bar{c}_2}{(1 + q_{11}c_1)(1 + q_{22}c_2) - q_{12}q_{21}c_1c_2}. \end{aligned}$$

Como diseñamos los controladores de forma que las interacciones son pequeñas, el término $q_{12}q_{21}$ es más pequeño que $q_{11}q_{22}$. La matriz \tilde{H} se puede aproximar entonces por

$$\tilde{H} \approx H = \begin{pmatrix} \frac{q_{11}\bar{c}_1}{1+q_{11}c_1} & \frac{q_{12}\bar{c}_2}{1+q_{11}c_1} \\ \frac{q_{21}\bar{c}_1}{1+q_{22}c_2} & \frac{q_{22}\bar{c}_2}{1+q_{22}c_2} \end{pmatrix}.$$

Los elementos de la diagonal de H son los mismos que para el diseño de control SISO. Se pueden utilizar los métodos estándar para diseño de controladores PI presentados en los Capítulos 6 y 7 para encontrar los controladores c_1 y c_2 . Analizando los elementos fuera de la diagonal, podemos estimar cuán severas son las interacciones. Los controladores pueden tener que desintonizarse para asegurar que las interacciones son tolerables. La interacción se puede reducir arbitrariamente haciendo los lazos de control suficientemente lentos. El análisis de interacción también resulta en una pérdida de rendimiento debido a la interacción. Si se pierde mucho rendimiento, es aconsejable considerar otros métodos de diseño.

Estimando los efectos de la interacción

A continuación desarrollaremos una forma sencilla de estimar los efectos de las interacciones. Los elementos fuera de la diagonal de H vienen dados por

$$h_{12} = \frac{q_{12}\bar{c}_2}{1+q_{11}c_1}$$

$$h_{21} = \frac{q_{21}\bar{c}_1}{1+q_{22}c_2}.$$

Obsérvese que $q_{11}(0) = q_{22}(0) = 1$ y que $q_{12} \approx \kappa_{12}s$ y $q_{21}(s) \approx \kappa_{21}s$ para pequeños valores de s . Como los controladores tienen acción integral, para pequeños valores de s tenemos

$$h_{12}(s) \approx \frac{\kappa_{12}k_{I2}s}{k_{I1}}, \quad h_{21}(s) \approx \frac{\kappa_{21}k_{I1}s}{k_{I2}}.$$

La interacción es así muy pequeña a bajas frecuencias, y por lo tanto podemos garantizar que la interacción es arbitrariamente pequeña para controladores suficientemente lentos. Para estimar el máximo de la interacción, observamos que

$$h_{12} = q_{12}\bar{c}_2S_1, \quad h_{21} = q_{21}\bar{c}_1S_2,$$

donde $S_1 = (1+q_{11}c_1)^{-1}$ y $S_2 = (1+q_{22}c_2)^{-1}$ son las funciones de sensibilidad para los lazos cuando se desprecia la interacción. Una estimación cruda de los términos de interacción es así

$$\max_{\omega} |h_{12}(i\omega)| \approx |\kappa_{12}|k_{I2}M_{s1}$$

$$\max_{\omega} |h_{21}(i\omega)| \approx |\kappa_{21}|k_{I1}M_{s2},$$

donde M_{s1} y M_{s2} son las sensibilidades máximas de los lazos individuales y donde hemos utilizado también la estimación

$$q_{12}(s) \approx \kappa_{12}s, \quad q_{21}(s) \approx \kappa_{21}s$$

y

$$\bar{c}_1 \approx k_{I1}/s, \quad \bar{c}_2 \approx k_{I2}/s.$$

Por consiguiente, la interacción se puede capturar por los índices de interacción

$$\kappa_1 = |\kappa_{12}k_{I2}|M_{s1}, \quad \kappa_2 = |\kappa_{21}k_{I1}|M_{s2}. \quad (11.10)$$

El índice κ_1 describe cómo el segundo lazo influye en el primer lazo, y κ_2 describe cómo el primer lazo influye en el segundo. Obsérvese que el término κ_{12} depende del sistema y de la ganancia integral k_{I2} del segundo lazo. La interacción se puede reducir así haciendo las ganancias integrales más pequeñas. Las estimaciones no son precisas debido a las aproximaciones realizadas, y no son fiables cuando existe una diferencia considerable en los anchos de banda de los lazos.

Ejemplos

El método de diseño se mostrará mediante dos ejemplos. Empezaremos investigando el sistema de Rosenbrock.

EJEMPLO 11.5—SISTEMA DE ROSENROCK

Considérese el sistema en el Ejemplo 11.3 donde el proceso tiene la función de transferencia (11.7). Tenemos

$$D = P^{-1}(0) = \begin{pmatrix} 1 & 2/3 \\ 1 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 3 & -2 \\ -3 & 3 \end{pmatrix}.$$

Si introducimos desacoplamiento estático, la función de transferencia compensada es

$$Q(s) = \begin{pmatrix} \frac{3(1-s)}{(s+1)(s+3)} & \frac{4s}{(s+1)(s+3)} \\ 0 & \frac{1}{s+1} \end{pmatrix} \approx \begin{pmatrix} 1 - 7s/3 & 4s/3 \\ 0 & 1 - s \end{pmatrix}.$$

La interacción viene dada por $\kappa_{12} = 4/3$ y $\kappa_{21} = 0$. Como $\kappa_{21} = 0$, la interacción no aporta limitaciones de rendimiento para el segundo lazo. Sin embargo, existen limitaciones a causa del cero en el semiplano derecho en $s = 1$. El diseño de un controlador PI que minimice la ganancia integral sujeto a las restricciones de que la máxima sensibilidad M_{s1} y la máxima sensibilidad complementaria M_{p1} sean menores que 1,6, da $k_{P1} = 0,2975$ y $k_{I1} = 0,3420$.

Como $\kappa_{12} = 4/3$, hay restricciones sobre el diseño del primer lazo a causa del acoplamiento. Al requerir que el acoplamiento κ_1 sea menor que 0,5 y la sensibilidad máxima M_{s2} sea menor que 1,6, encontramos que la ganancia integral del segundo lazo k_{I2} debe ser menor que $\kappa_1 / (\kappa_{12} M_{s1} M_{s2}) = 0,23$. Para diseñar un controlador PI, utilizamos un procedimiento de asignación donde se cancela el polo rápido del proceso $s = -1$. La ganancia en el segundo lazo es entonces $k_{P2} = 0,23$.

La Figura 11.6 muestra las respuestas en frecuencia de h_{11} , h_{12} y h_{22} . La magnitud máxima del término h_{12} es 0,26, que es la mitad del valor estimado. La

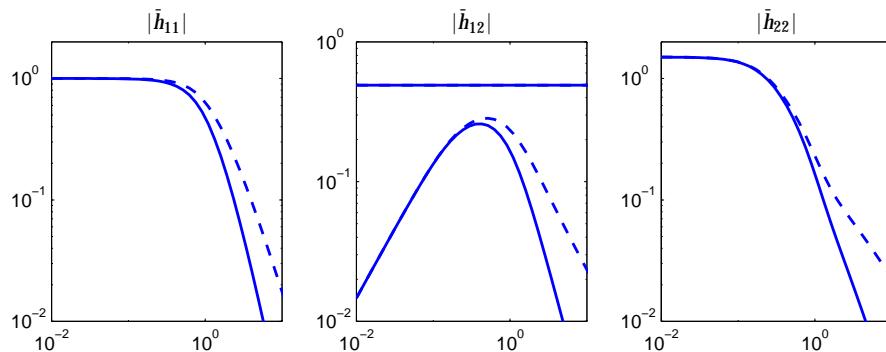


Figura 11.6 Respuestas en frecuencia del sistema en lazo cerrado sin (líneas a tramos) y con (líneas continuas) ponderación del punto de consigna. Obsérvese que sin la ponderación del punto de consigna, la interacción $|\bar{h}_{12}(i\omega)|$ es más grande y se extiende a frecuencias más altas.

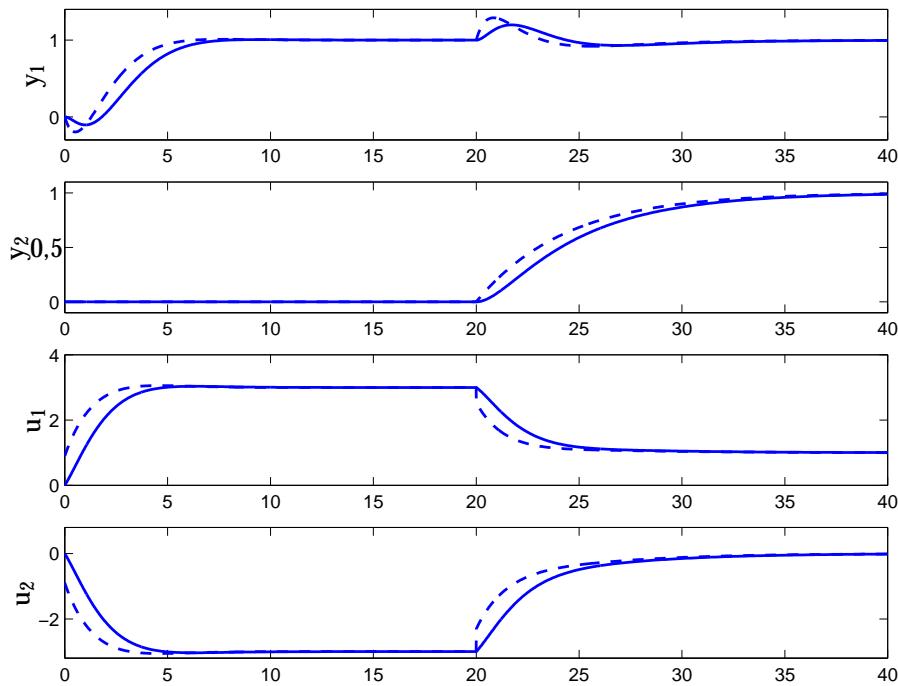


Figura 11.7 Simulación del método de diseño aplicado al sistema de Rosenbrock. La figura muestra la respuesta de las salidas a escalones en las señales de referencia. Los controladores PI tienen ganancias $k_{P1} = 0,30$, $k_{I1} = 0,34$, $k_{P2} = 0,23$, $k_{I2} = 0,23$. Las líneas a tramos muestran los resultados con realimentación del error y las líneas continuas muestran los resultados con pesos del punto de consigna iguales a cero.

razón de esta discrepancia es que la estimación simple $q_{12} \approx \kappa_{12}s$ sobreestima el término.

La Figura 11.7 muestra simulaciones de las respuestas en el punto de consigna para el sistema en lazo cerrado. Las líneas continuas muestran las respuestas para el controlador con ponderación del punto de consigna ($b_1 = b_2 = 0$). En el instante de tiempo $t = 0$ se aplica un escalón unitario en el punto de consigna del primer controlador, y en el instante $t = 20$ para el segundo controlador. La Figura 11.7 muestra las respuestas a un escalón para el controlador sin ponderación del punto de consigna. La figura indica claramente la ventaja de la ponderación del punto de consigna para sistemas multivariados. La razón de tan gran diferencia es que la señal de control es mucho más suave con la ponderación del punto de consigna.

El efecto de la ponderación del punto de consigna se muestra también en la Figura 11.6, que muestra la respuesta en frecuencia del sistema en lazo cerrado con (líneas continuas) y sin (líneas a tramos) ponderación del punto de consigna. La interacción aumenta considerablemente cuando no se aplica la ponderación del punto de consigna. \square

Las columnas de destilación son procesos industriales típicos donde la interacción es significativa. El próximo ejemplo trata con un caso de este tipo.

EJEMPLO 11.6—LA COLUMNA DE DESTILACIÓN BINARIA DE WOOD-BERRY

La columna de destilación binaria de Wood-Berry es un sistema multivariado que ha sido estudiado extensivamente. Un modelo simple del sistema viene dado por la siguiente función de transferencia

$$P(s) = \begin{pmatrix} \frac{12,8e^{-s}}{16,7s+1} & \frac{-18,9e^{-3s}}{21,0s+1} \\ \frac{6,60e^{-7s}}{10,9s+1} & \frac{-19,4e^{-3s}}{14,4s+1} \end{pmatrix}.$$

Al diseñar un desacoplador estático, encontramos que

$$Q(s) = P(s)P^{-1}(0) \approx \begin{pmatrix} 1 - 11,7s & -12,31s \\ -0,5138s & 1 - 17,3s \end{pmatrix}.$$

Por consiguiente, $\kappa_{12} = -12,31$ y $\kappa_{21} = -0,5138$. Diseñando un controlador PI para los elementos de la diagonal maximizando la ganancia integral sujeta a la restricción de robustez $M_s = 1,6$ da lugar a $k_1 = 2,3481$, $k_{11} = 1,5378$, $k_2 = 0,5859$ y $k_{21} = 0,2978$. Las frecuencias de sensibilidad son $\omega_{s1} = 0,30$ y $\omega_{s2} = 0,11$. Obsérvese que el segundo lazo es más lento que el del primer lazo. Tenemos $\kappa_1 = 5,8$ y $\kappa_2 = 1,26$, que indica que la interacción impone restricciones sobre el rendimiento alcanzable y es necesario desintonizar los controladores. Esto se ilustra mediante curvas a tramos en la simulación mostrada en la Figura 11.8. Para reducir las interacciones, desintonizaremos los controladores disminuyendo las ganancias integrales. Como un primer intento, reduciremos ambas ganancias integrales por un factor de cuatro. Esto implica que el error integrado para las perturbaciones de carga es cuatro veces más grande para un lazo desacoplado. Utilizando la regla de reducción de ganancia simple de la Sección 7.9, encontramos

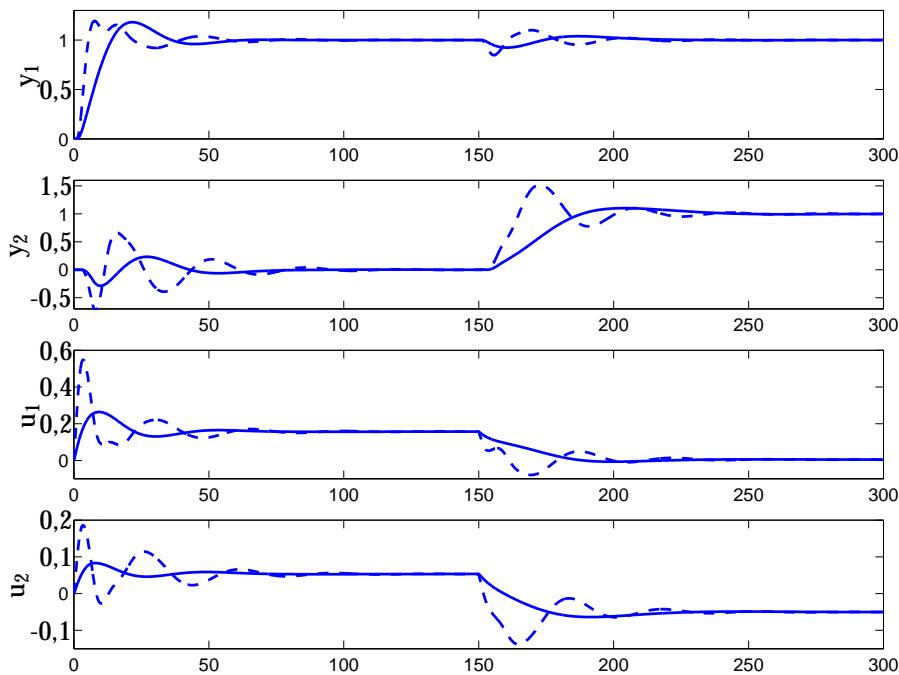


Figura 11.8 Simulación de control desacoplado de la columna de destilación de Wood-Berry. La figura muestra la respuesta de las salidas a escalones en las señales de referencia. Las curvas a tramos muestran las respuestas con controladores PI con ganancias $k_{P1} = 2,348$, $k_{I1} = 1,537$, $k_{P2} = 0,586$ y $k_{I2} = 0,298$. Las líneas continuas muestran las respuestas con controladores PI desintonizados. Las ganancias son $k_{P1} = 1,119$, $k_{I1} = 0,384$, $k_{P2} = 0,293$ y $k_{I2} = 0,0745$. Los pesos del punto de consigna son cero en todos los casos.

que las ganancias proporcionales se deberían reducir por un factor de dos; véase (7.27). Las líneas continuas en la Figura 11.8 muestran que las respuestas dan una reducción significativa de las interacciones. La interacción se puede reducir aún más a costa del precio de un rendimiento más bajo. □

11.4 Sistemas paralelos

Los sistemas que están conectados en paralelo son bastante comunes, particularmente en los sistemas de transmisión. Ejemplos típicos son motores que llevan la misma carga, sistemas de potencia y redes de distribución de vapor. El control de tales sistemas requiere una consideración especial. Para ilustrar las dificultades que pueden surgir, consideraremos la situación con dos motores llevando la misma carga. La Figura 11.9 muestra un diagrama esquemático del sistema.

Sea ω la velocidad angular del eje, J el momento de inercia total y D el coefi-

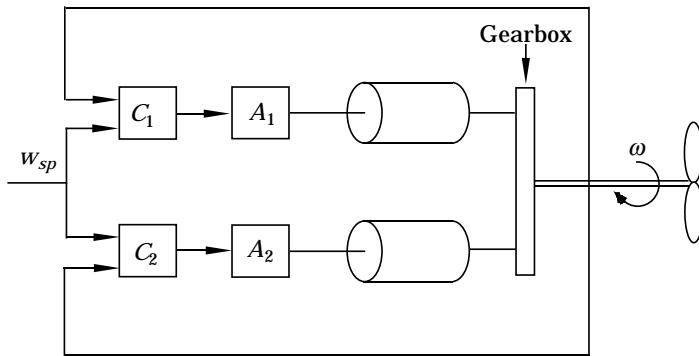


Figura 11.9 Diagrama esquemático de dos motores que llevan la misma carga.

ciente de amortiguamiento. El sistema se puede describir entonces por la ecuación

$$J \frac{d\omega}{dt} + D\omega = M_1 + M_2 - M_L, \quad (11.11)$$

donde M_1 y M_2 son los pares de los motores y M_L es el par de carga.

Control proporcional

Asumir que cada motor se proporciona con un controlador proporcional. Las estrategias de control son entonces

$$\begin{aligned} M_1 &= M_{10} + K_1(\omega_{sp} - \omega) \\ M_2 &= M_{20} + K_2(\omega_{sp} - \omega). \end{aligned} \quad (11.12)$$

En estas ecuaciones los parámetros M_{10} y M_{20} dan los pares proporcionados por cada motor cuando $\omega = \omega_{sp}$ y K_1 y K_2 son las ganancias de los controladores.

De (11.11) y (11.12) se deduce que

$$J \frac{d\omega}{dt} + (D + K_1 + K_2)\omega = M_{10} + M_{20} - M_L + (K_1 + K_2)\omega_{sp}.$$

El sistema en lazo cerrado es así un sistema dinámico de primer orden. Después de las perturbaciones, la velocidad angular alcanza el régimen permanente con una constante de tiempo

$$T = \frac{J}{D + K_1 + K_2}.$$

La velocidad de respuesta viene dada por la suma del amortiguamiento y las ganancias de los controladores. El valor estacionario de la velocidad angular viene dado por

$$\omega = \omega_0 = \frac{K_1 + K_2}{D + K_1 + K_2} \omega_{sp} + \frac{M_{10} + M_{20} - M_L}{D + K_1 + K_2}.$$

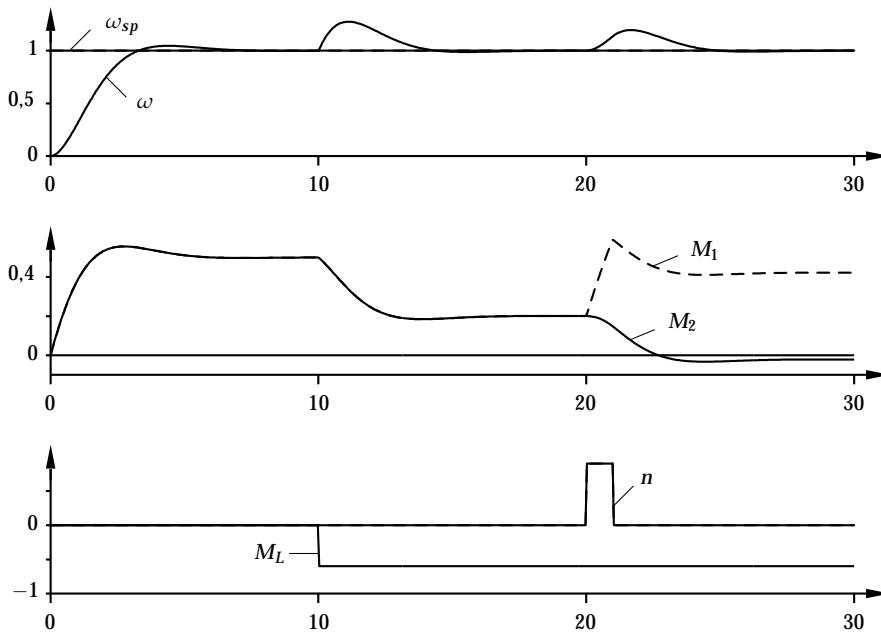


Figura 11.10 Simulación de un sistema con dos motores con controladores PI que mueven la misma carga. La figura muestra el punto de consigna ω_{sp} , la salida del proceso ω , las señales de control M_1 y M_2 , la perturbación de carga M_L , y la perturbación de la medida n .

Esto implica que normalmente existirá un error en régimen permanente. De manera similar, de (11.12) encontramos que

$$\frac{M_1 - M_{10}}{M_2 - M_{20}} = \frac{K_1}{K_2}.$$

La razón de las ganancias del controlador indicará cómo se distribuye la carga entre los motores.

Control proporcional e integral

La manera estándar de eliminar un error en régimen permanente es introducir acción integral. En la Figura 11.10, mostramos una simulación del sistema en la cual los motores tienen controladores PI idénticos. El punto de consigna se modifica en el instante de tiempo 0. Se introduce una perturbación de carga en forma de escalón en el par de carga en el instante de tiempo 10, y en el instante 20 se introduce una perturbación medible en forma de pulso en el segundo controlador. Cuando se produce error de medida, se modifica el balance de los pares de forma que el primer motor toma muchas más carga después de la perturbación. En este caso particular, el segundo motor está realmente frenando. Esto es, por supuesto, altamente no deseable.

Para comprender el fenómeno, mostramos el diagrama de bloques del sistema en la Figura 11.11. La figura muestra que existen dos caminos paralelos en

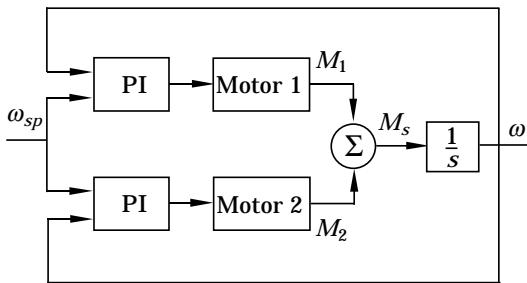


Figura 11.11 Diagrama de bloques para el sistema de la Figura 11.10.

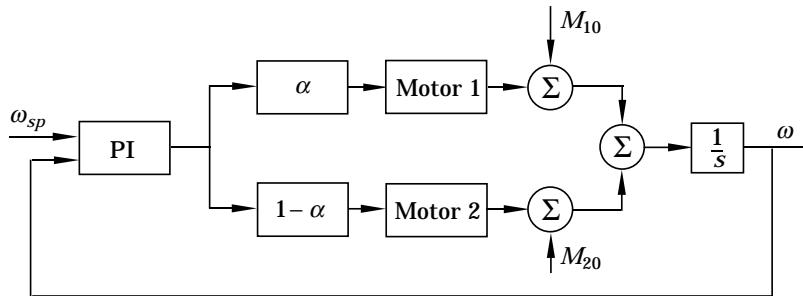


Figura 11.12 Diagrama de bloques del sistema mejorado.

el sistema que contienen integración. Este es un caso estándar donde se pierde observabilidad y controlabilidad. En otras palabras, no es posible cambiar individualmente desde el error las señales M_1 y M_2 . Como el estado no controlable es un integrador, no tiende a cero después de las perturbaciones. Esto quiere decir que los pares pueden tomar valores arbitrarios tras la perturbación. Por ejemplo, puede ocurrir que uno de los motores tome prácticamente toda la carga, siendo claramente una situación indeseable.

Cómo evitar las dificultades

Habiendo comprendido la razón de las dificultades, es sencillo modificar el controlador tal y como se muestra en la Figura 11.12. En este caso, sólo se utiliza un controlador con acción integral. La salida de éste conduce los controladores proporcionales para cada motor. La Figura 11.13 muestra una simulación de este sistema. Las dificultades son claramente eliminadas.

Las dificultades mostradas en los ejemplos con dos motores moviendo la misma carga se acentúan incluso más si hay más motores. En este caso se puede obtener un buen control utilizando un controlador PI y distribuyendo las salidas de este controlador PI a los diferentes motores, cada uno de los cuales tiene un controlador proporcional. Una alternativa es proporcionar un motor con controlador PI y permitir al otro tener un controlador proporcional. Para resumir, hemos encontrado que pueden existir dificultades en sistemas paralelos con acción integral. Las dificultades se originan por la conexión paralela de integradores que pueden producir subsistemas inestables que no son ni controlables ni observa-

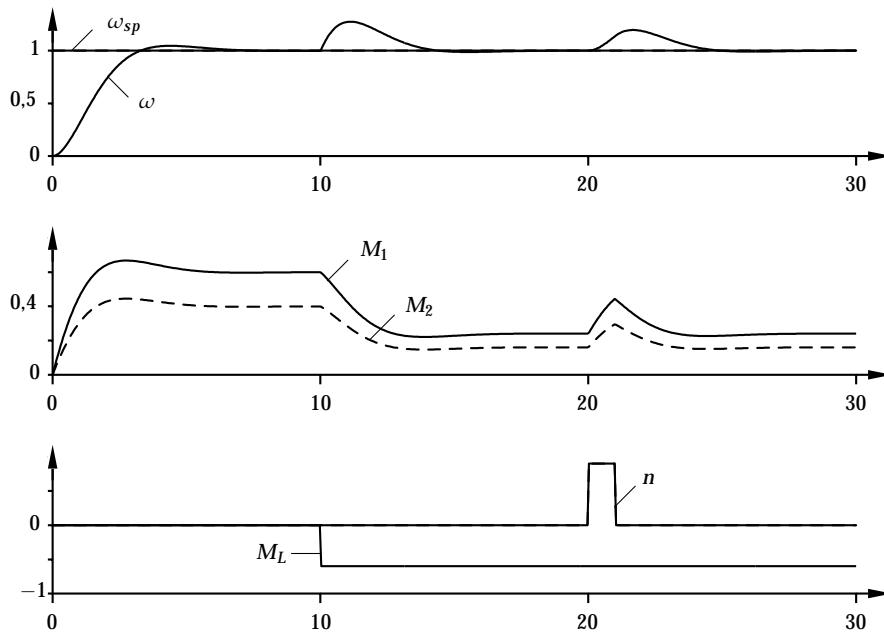


Figura 11.13 Simulación del sistema con el controlador modificado. La figura muestra el punto de consigna ω_{sp} , la salida del proceso ω , las señales de control M_1 y M_2 , la perturbación de carga M_L , y la perturbación de medida n .

bles. En presencia de perturbaciones, estos modos pueden cambiar de manera arbitraria. La solución consiste en cambiar las estrategias de control de forma que exista un sólo integrador.

11.5 Resumen

Aunque un sistema de control grande puede tener muchos sensores y actuadores, generalmente se puede controlar mediante simples controladores de tipo PID. Esto es particularmente sencillo cuando existe poca interacción en el sistema. En este capítulo, hemos presentado medidas simples de interacción, que pueden utilizarse para juzgar si el problema de control se puede resolver utilizando lazos simples. La matriz de ganancias relativas de Bristol se puede utilizar para encontrar emparejamientos de entradas y salidas que son adecuados para control de lazos simples. También se ha presentado un método de diseño simple que se puede aplicar a sistemas con interacción. Este método combina desacoplamiento estático con métodos para el diseño de controladores de lazo simple presentados anteriormente en este libro. Además se ha descrito el control de sistemas de transmisión con motores en paralelo. Para tales sistemas, existen problemas particulares con controladores que tienen acción integral.

11.6 Notas y referencias

En [294, 295, 236] se tratan algunos temas fundamentales relacionados con la interacción en sistemas. La matriz de ganancias relativas se introdujo en [65]. Ha sido utilizada ampliamente y satisfactoriamente en la industria de procesos [308, 236]. Los resultados mejor conocidos del RGA son que una planta con elementos grandes o negativos en su RGA es difícil de controlador y que las variables de entrada y salida se deberían emparejar de forma que los elementos diagonales estén tan cercanos a la unidad como sea posible [133, 316]. El RGA se basa en la ganancia estática del proceso; en [236] se presenta una extensión para tener en cuenta la dinámica. En [75] se desarrolló una medida alternativa denominada el índice interacción en estado estacionario y puede proporcionar una representación más exacta. El desacoplamiento dinámico y estático se trata en muchos libros de texto sobre control de procesos, por ejemplo, [302]. Trabajos recientes para el diseño de control PID desacoplado incluye el trabajo desarrollado por [2]. La desintonía de control PID multivariable se trata de manera heurística en [254]. El método particular presentado en la Sección 11.3 se basa en [23], otros métodos de diseño de sistemas no interactuante se encuentran en [371] y [357]. El control de sistemas con fuerte interacción entre muchos lazos requiere técnicas que son diferentes de aquellas analizadas en este capítulo; véase [91] y [303]. Los sistemas multivariados se tratan en libros de texto estándar en el campo de control de procesos tal como [221, 229, 48, 302]. Además existen libros centrados en sistemas multivariados: ver [308, 317].

12

Paradigmas de control

12.1 Introducción

Los sistemas de control de procesos son normalmente complejos con muchas variables de control y muchas señales de medida. El enfoque ascendente es una de las vías para diseñar tales sistemas. En este procedimiento, el sistema se construye a partir de componentes simples. Los sistemas se pueden implementar de muchas maneras distintas. Originalmente, se realizó por la interconexión de cajas separadas construidas de componentes neumáticos o electrónicos. Hoy día, los sistemas se implementan típicamente en sistemas de control distribuidos que constan de varios computadores conectados jerárquicamente. Los programas para el sistema de control distribuido se construyen típicamente de tal manera que la programación se puede realizar seleccionando e interconectando los componentes. El componente clave, el controlador PID, ha sido ya comentado en detalle. En este capítulo, presentamos algunos de los componentes requeridos para construir sistemas de automatización. También presentamos algunos de los paradigmas básicos que guían la construcción de sistemas complejos.

Se utiliza una colección de paradigmas para control para construir sistemas complejos a partir de componentes sencillos. Los componentes son controladores de tipo PID, filtros lineales y no linealidades estáticas. Las no linealidades típicas son limitadores de amplitud y velocidad de cambio y selectores de señales. La realimentación es un paradigma importante. Los lazos de realimentación simple se utilizan para mantener las variables del proceso constantes o hacerlas que cambien de ciertas maneras específicas. La realimentación se ha analizado extensamente en capítulos previos. Otro paradigma importante es la acción feedforward, que se estudió en el Capítulo 5. El problema clave es determinar las variables de control que se deberían escoger para controlar variables del proceso dadas. Otro problema es que puede existir interacción entre diferentes lazos de control. Esto se trató en el Capítulo 11.

La Sección 12.2 presenta un resumen del problema de diseñar sistemas complejos, y se comparan los enfoques descendente y ascendente. Esta sección también da un resumen del contenido del capítulo. El capítulo finaliza con un ejemplo para ilustrar cómo se pueden utilizar los diferentes componentes y paradigmas. El proceso considerado es un reactor químico, y el diseño se describe en la Sec-

ción 12.9. Algunas de las observaciones importantes realizadas en el capítulo son finalmente resumidas en la Sección 12.10.

12.2 Enfoques ascendente y descendente

Existen dos enfoques generales para diseñar sistemas complejos: ascendente y descendente. En el enfoque *ascendente* o tipo Lego, el sistema se diseña combinando pequeños subsistemas. El enfoque *descendente* comienza con un diseño completo general que se refina sucesivamente. En la práctica, estos esquemas se combinan generalmente. En ambos enfoques, necesitamos conocimiento sobre los bloques o componentes de construcción elementales del sistema. El enfoque ascendente requiere principios para combinar componentes básicos, y el enfoque descendente requiere principios para refinar o descomponer un objetivo de alto nivel tal que se pueda llevar a cabo con los componentes básicos del sistema. Se han descrito previamente en este libro algunos de los componentes y principios de control para la composición y descomposición. En esta sección, daremos un resumen de los enfoques y en secciones posteriores describiremos los componentes y paradigmas que no se han descrito previamente.

El enfoque ascendente

Los sistemas de control complejos se pueden construir a partir de controladores, filtros y elementos no lineales. Los componentes se pueden separar o en piezas físicas o en bloques de funciones implementados en programas que se pueden combinar gráficamente utilizando la idea de copiar y pegar. En los Capítulos 3, 5, 9 y 11 se han descrito los controladores y los filtros. En la Sección 12.6 se analizarán los elementos no lineales.

Principios de control como realimentación, acción feedforward y el seguimiento del modelo se han analizado extensamente en los Capítulos 3 y 5. Otros principios de control importantes tales como *control repetitivo*, *control en cascada*, *control de rango medio*, *control de rango partido*, *control de ratio*, y *control selector* se tratarán en las Secciones 12.3, 12.4, 12.5 y 12.6.

Una ventaja con el enfoque ascendente es que el sistema puede ponerse en marcha y sintonizarse lazo a lazo. Pueden existir dificultades cuando los lazos están interactuando. La desventaja es que no es fácil juzgar si los lazos adicionales aportarán beneficios. El sistema puede ser además difícil de manejar cuando se añaden nuevos lazos.

Soluciones descendentes

Los paradigmas descendentes comienzan generalmente con una formulación del problema en términos de un problema de optimización. Los paradigmas que soportan un enfoque descendente son optimización, realimentación, observadores, control predictivo y linealización. En el enfoque descendente es natural tratar con muchas entradas y muchas salidas de manera simultánea. Como éste no es el tema principal de este libro, daremos sólo una breve introducción. El enfoque descendente conduce generalmente a la estructura de controlador que se muestra en la Figura 12.1. En este sistema, todas las variables de procesos medidas

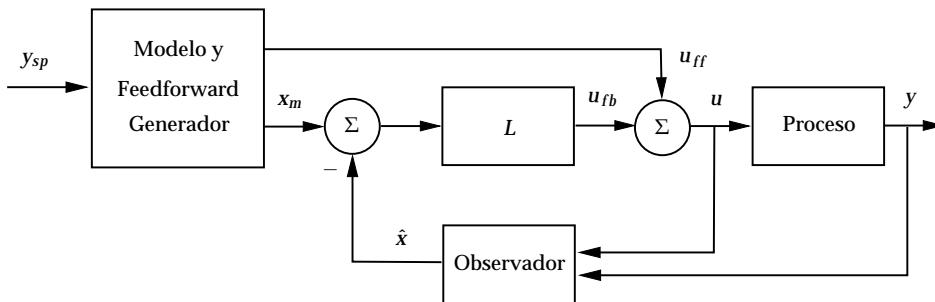


Figura 12.1 Diagrama de bloques de un controlador basado en modelo, realimentación del estado y un observador.

y junto con las variables de control u se envían a un observador, que utiliza la información del sensor y un modelo matemático para generar un vector \hat{x} de buenas estimaciones de las variables internas del proceso y de las perturbaciones importantes. El estado estimado \hat{x} se compara entonces con el estado ideal x_m producido por el generador feedforward, y la diferencia se realimenta al proceso. El generador feedforward también aporta una señal de prealimentación u_{ff} que se envía directamente a las entradas del proceso. El controlador que se muestra en la Figura 12.1 es útil para segmentos del proceso donde existen varias entradas y salidas que interactúan, pero el sistema se complica mucho cuando hay un número grande de entradas y salidas. En tal caso, puede ser mejor descomponer el sistema en varios subsistemas.

Una ventaja con el enfoque descendente es que se tiene en cuenta el comportamiento total del sistema. Un enfoque sistemático basado en modelado matemático y simulación hace sencillo comprender las limitaciones fundamentales. Sin embargo, la puesta en marcha de un sistema es difícil porque se tienen que cerrar de manera simultánea muchos lazos de realimentación. Cuando se utiliza un enfoque descendente es, por lo tanto, una buena práctica sintonizar primero los lazos basándose en simulación, y posiblemente también incluyendo hardware en la simulación del lazo.

Soft Computing

Debido al uso extendido de los computadores en control, existe también una influencia sobre el control por parte de la informática. *Las redes neuronales* y el *control borroso* son dos paradigmas particulares que se originaron y emergieron de la investigación en inteligencia artificial. En las Secciones 12.7 y 12.8 se presentan estos paradigmas. Esta rama de la informática también se denomina *soft computing*.

12.3 Control repetitivo

La atenuación de las perturbaciones ha sido uno de los temas esenciales en este libro. Para el control PID, nos hemos centrado en la eliminación de perturbaciones constantes o lentas. En esta sección, mostraremos que se pueden utilizar ideas

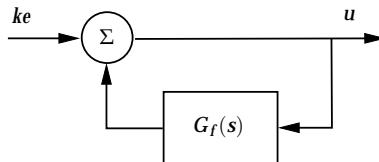


Figura 12.2 Diagrama de bloques de un controlador con realimentación positiva de una señal filtrada.

similares para eliminar otro tipo de perturbaciones, particularmente perturbaciones periódicas. Problemas de este tipo son comunes cuando existen operaciones cíclicas.

En la Sección 4.3 se mostró que la atenuación de perturbaciones se deduce de la función de transferencia que relaciona la perturbación de carga con la salida del proceso

$$G_{yd} = \frac{P}{1 + PC}, \quad (12.1)$$

donde P es la función de transferencia del proceso y C la función de transferencia del controlador, respectivamente. Diseñando un controlador que tenga alta ganancia a una frecuencia particular, las perturbaciones con esas frecuencias se pueden reducir eficientemente. El error de control es cero en régimen permanente si la ganancia es infinita.

Considérese el controlador de la Figura 12.2. Intuitivamente, el sistema funciona como sigue. El filtro G_f filtra la componente de la señal que nos gustaría eliminar y la salida de G_f se realimenta a la entrada con realimentación positiva. El efecto neto consiste en crear una alta ganancia para las frecuencias en la banda de paso del filtro G_f .

Perturbaciones constantes y sinusoidales

Para investigar analíticamente las propiedades del sistema, observamos que el controlador tiene la función de transferencia

$$C(s) = \frac{k}{1 - G_f(s)}. \quad (12.2)$$

Cuando $G_f(s)$ es un filtro paso-bajo con función de transferencia

$$G_f(s) = \frac{1}{1 + sT},$$

encontramos que

$$C(s) = k\left(1 + \frac{1}{sT}\right),$$

que es la función de transferencia de un controlador PI. Obsérvese que la función de transferencia del controlador $C(s)$ tiene ganancia infinita a frecuencia cero, lo que implica que el error en régimen permanente es cero para perturbaciones constantes.

Cuando $G_f(s)$ es el filtro paso-banda

$$G_f(s) = \frac{2\zeta\omega_0 s}{s^2 + 2\zeta\omega_0 s + \omega_0^2},$$

encontramos que

$$C(s) = k \frac{2\zeta\omega_0 s}{s^2 + \omega_0^2}.$$

Obsérvese que esta función de transferencia tiene ganancia infinita para $s = i\omega_0$, lo que implica que el error en régimen permanente es cero para una perturbación sinusoidal de frecuencia ω_0 .

Perturbaciones periódicas

Las perturbaciones periódicas se pueden reducir escogiendo

$$G_f(s) = e^{-sL},$$

donde L es el periodo de la perturbación. Con este filtro encontramos

$$C(s) = \frac{k}{1 - e^{-sL}}. \quad (12.3)$$

La relación entre el error de control y la variable de control es

$$u(t) = ke(t) + u(t - L).$$

La acción de control en el instante de tiempo t es por lo tanto una suma del error de control y la señal de control en el instante de tiempo $t - L$.

El controlador tiene ganancia infinita para $s = 2n\pi i/L$, $n = 0, 1, \dots$. Un controlador de este tipo es particularmente útil cuando las perturbaciones o las variaciones en el punto de consigna son periódicas.

La función de transferencia de la perturbación de carga a la salida (12.1) es

$$G_{yd}(s) = \frac{P(s)}{1 + P(s)C(s)} = \frac{P(s)(1 - e^{-sL})}{1 - e^{-sL} + kP(s)}.$$

La relación entre la perturbación de carga y la salida es entonces

$$(1 - e^{-sL} + kP(s)) Y(s) = P(s)(1 - e^{-sL}) D(s).$$

Obsérvese que la función temporal correspondiente a $(1 - e^{-sL}) D(s)$ es

$$d(t) - d(t - L),$$

que desaparece si D es una perturbación periódica con período L . El error en régimen permanente causado por una perturbación periódica es así cero.

Sin embargo, el rechazo efectivo de las perturbaciones se consigue a un precio que se muestra con el siguiente ejemplo.

EJEMPLO 12.1—UN CASO EXTREMO

Considérese un proceso con la función de transferencia

$$P(s) = e^{-sL},$$

con el controlador

$$C(s) = \frac{1}{1 - e^{-sL}}.$$

que atenúa las perturbaciones periódicas.

La función de transferencia del lazo

$$G_I(s) = \frac{e^{-sL}}{1 - e^{-sL}}$$

es periódica con período $2\pi/L$, y su ganancia es infinita para $\omega = 2n\pi/L$. La respuesta en frecuencia es

$$G_I(i\omega) = -\frac{1}{2} - i\frac{\sin \omega L}{2(1 - \cos \omega L)} = -\frac{1}{2} - i\frac{1}{\tan(\omega L/2)}.$$

El diagrama de Nyquist es una línea vertical a través del punto $G_I = -0,5$ y un medio círculo a la derecha. Esta curva se recorre una vez para $0 \leq \omega \leq 2\pi/L$ e infinitas veces cuando ω tiende a infinito.

El sistema tiene un margen de ganancia de 2 y un margen de fase de 60° . Las funciones de sensibilidad son

$$\begin{aligned} S(s) &= 1 - e^{-sL} \\ T(s) &= e^{-sL}, \end{aligned}$$

y encontramos que $M_s = 2$ y $M_t = 1$.

Una mirada superficial a las medidas de robustez tradicionales como el margen de ganancia $g_m = 2$, el margen de fase $\varphi_m = 60^\circ$, y las sensibilidades máximas $M_s = 2$ y $M_t = 1$ pueden indicar que el sistema es robusto a perturbaciones sobre el proceso.

Sin embargo, el hecho de que $T(i\omega) = 1$ para todas las frecuencias es una indicación de que el sistema tiene propiedades inusuales. Una mayor comprensión se obtiene analizando el efecto de las variaciones de los parámetros. El sistema tiene sólo un parámetro, el retardo de tiempo L , e investigaremos los efectos de las variaciones en este parámetro. El uso de la desigualdad de robustez (4.32), nos convertirá las variaciones del retardo de tiempo a una perturbación aditiva del proceso. Supóngase que el retardo de tiempo cambia de L a $L + \delta L$, entonces

$$e^{-s(L+\delta L)} = e^{-sL}e^{-s\delta L} = e^{-sL} + e^{-sL}(e^{-s\delta L} - 1).$$

Una variación en el retardo de tiempo se puede representar así mediante una perturbación aditiva

$$\Delta P(s) = e^{-sL}(e^{-s\delta L} - 1).$$

Por tanto, $|\Delta P(i\omega)| = |e^{-i\omega\delta L} - 1|$.

Como $|P(i\omega)| = 1$, la desigualdad de robustez (4.32) viene dada por

$$\frac{|\Delta P(i\omega)|}{|P(i\omega)|} = |e^{-i\omega\delta L} - 1| < \frac{1}{|T(i\omega)|} = 1.$$

Esta desigualdad no se satisface para ningún $\delta L > 0$ porque la parte izquierda es 2 y la derecha es 1, y no podemos garantizar estabilidad para un perturbación arbitrariamente pequeña en el retardo de tiempo. \square

El ejemplo muestra que la atenuación efectiva de perturbaciones periódicas se consigue a costa de que el sistema sea extremadamente sensible a variaciones en los parámetros. Un compromiso entre la atenuación de perturbaciones se puede hacer reemplazando $G_f(s)$ en la Figura 12.2 por $\alpha G_f(s)$ con $\alpha < 1$. Entonces, los controladores obtenidos para señales constantes, sinusoidales y periódicas son

$$\begin{aligned} C_{\text{const}}(s) &= \frac{1+sT}{1-\alpha+sT} \\ C_{\text{sine}}(s) &= \frac{s^2+2\zeta\omega_0s+\omega_0^2}{s^2+2(1-\alpha)\zeta\omega_0s+\omega_0^2} \\ C_{\text{per}}(s) &= \frac{1}{1-\alpha e^{-sT}}. \end{aligned}$$

Las ganancias máximas de las funciones transferencia son $1/(1-\alpha)$ en todos los casos. Escogiendo $\alpha < 1$, se disminuye la atenuación de la perturbación pero se mejora la robustez. En la Figura 12.3 se ilustran las propiedades de los controladores C_{const} , C_{sine} y C_{per} , que muestra los diagramas de magnitud de Bode para los controladores. El controlador C_{const} tiene una alta ganancia a bajas frecuencias, el controlador C_{sine} tiene alta ganancia para $\omega = \omega_0$ y el controlador C_{per} tiene alta ganancia a las frecuencias ω_0 , $2\omega_0$, $3\omega_0$, etc.

12.4 Control en cascada

El control en cascada se puede utilizar cuando existen varias señales de medida y una variable de control. Es particularmente útil cuando hay dinámicas significativas, por ejemplo, grandes retardos de tiempo o grandes constantes de tiempo, entre la variable de control y la variable del proceso. Un control más ajustado se puede alcanzar utilizando un señal medida intermedia que responda más rápido a la señal de control. El control en cascada se construye anidando los lazos de control, tal y como se muestra en el diagrama de bloques de la Figura 12.4. El sistema en esta figura tiene dos lazos. El lazo interno se denomina *lazo secundario* y el lazo externo *lazo primario*. La razón de esta terminología es que el lazo externo trata con la señal de medida primaria. Es además posible tener un control en cascada con más lazos de control anidados. El rendimiento de un sistema se puede mejorar con un número de señales de medida hasta un cierto límite. Si se miden todas las variables de estado, no vale generalmente la pena introducir otras variables de medida. En tal caso, el control en cascada es lo mismo que una realimentación del estado. Ilustraremos los beneficios del control en cascada mediante un ejemplo.

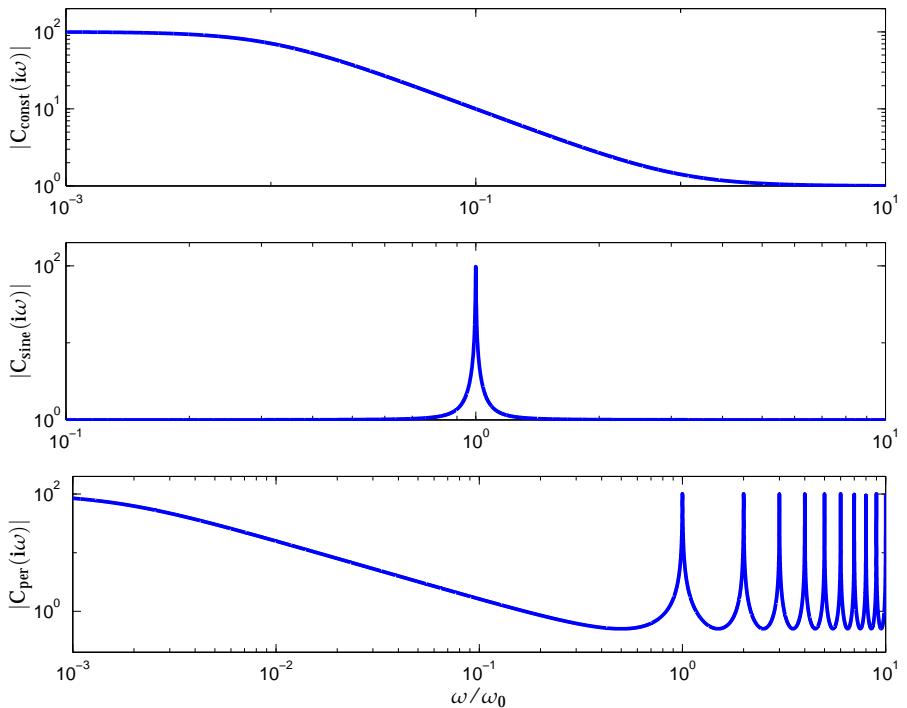


Figura 12.3 Diagramas de magnitud de Bode para los controladores C_{const} (superior), C_{sine} (en el medio) y C_{per} (inferior). El parámetro α es 0.99 en todos los casos, lo que quiere decir que las ganancias más altas de los controladores son 100. Para el filtro paso-banda, tenemos $\zeta = 0.1$, y para el controlador repetitivo tenemos $T = 2\pi/\omega_0$.

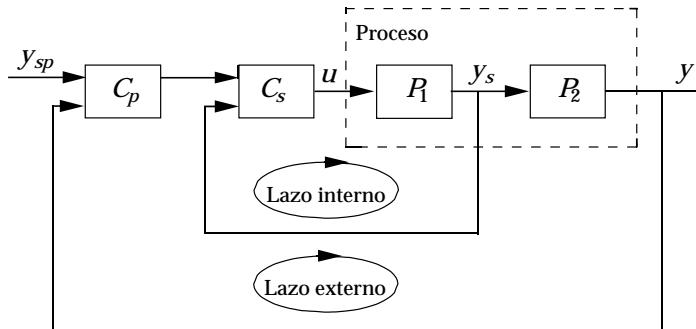


Figura 12.4 Diagrama de bloques de un sistema con control en cascada.

EJEMPLO 12.2—RECHAZO A PERTURBACIONES DE CARGA MEJORADO

Considérese el sistema que se muestra en la Figura 12.4. Sean las funciones de transferencia

$$P_1 = \frac{1}{s+1}$$

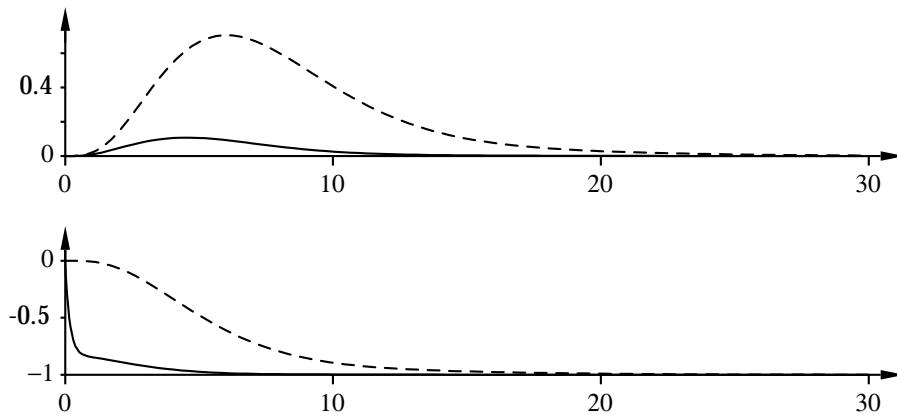


Figura 12.5 Respuestas a perturbación de carga para un sistema con (línea continua) y sin (línea a tramos) control en cascada. El diagrama superior muestra la salida del proceso y y el diagrama inferior la señal de control u .

y

$$P_2 = \frac{1}{(s+1)^3}.$$

Supóngase que una perturbación de carga entra a la entrada del proceso. Existen dinámicas significativas desde la variable de control a la salida primaria. La salida secundaria responde mucho más rápido que la salida primaria. Por tanto, se puede esperar que el control en cascada aporte mejoras.

Las líneas a tramos en la Figura 12.5 muestran la respuesta obtenida con una realimentación convencional utilizando un controlador PI con parámetros $K = 0,37$ y $T_i = 2,2$. Como la respuesta de la variable de medida secundaria a la señal de control es bastante rápida, es posible utilizar ganancias altas del lazo en el lazo secundario. Si el controlador en el lazo interno es de tipo proporcional con ganancia K_s , la dinámica desde el punto de consigna de C_s a la salida del proceso es

$$G(s) = \frac{K_s}{(s+1+K_s)(s+1)^3}.$$

Con $K_s = 5$ en el lazo interno y un control PI con $K = 0,55$ y $T_i = 1,9$ en el lazo externo, se obtienen las respuestas que se muestran en líneas continuas en la Figura 12.5. La figura muestra que la respuesta a la perturbación se mejora sustancialmente utilizando control en cascada. Obsérvese que, en este caso particular, la variable de control cae mucho más rápidamente con control en cascada. La razón principal para esto es el lazo de realimentación interno, que detecta la perturbación mucho más rápidamente que el lazo externo.

El controlador secundario es de tipo proporcional y la ganancia del lazo es 5. Un parte grande de la perturbación se elimina por el lazo interno. El error remanente se elimina a una velocidad más lenta a través de la acción del lazo externo. En este caso, la acción integral en el lazo interno dará siempre una sobreelongación en la respuesta a perturbaciones. \square

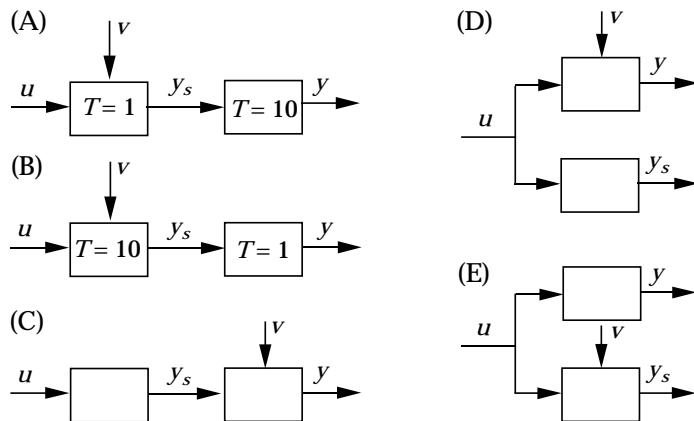


Figura 12.6 Ejemplos de diferentes procesos y configuraciones de medida.

Elección de las variables de medida secundarias

Es importante ser capaz de juzgar si el control en cascada puede aportar mejoras y tener una metodología para escoger la variable de medida secundaria. Esto es sencillo de hacer si recordamos que la idea clave del control en cascada es disponer un lazo de realimentación ajustado alrededor de una perturbación. En el caso ideal, el lazo secundario se puede hacer tan ajustado que el lazo secundario es un servo perfecto donde la variable de medida secundaria responde muy rápidamente a la señal de control. Las reglas básicas para la selección de la variable secundaria son:

- Debería existir una relación bien definida entre las variables de medida primaria y secundaria.
- Las perturbaciones esenciales deberían actuar en el lazo interno.
- El lazo interno debería ser más rápido que el lazo externo. La regla general típica es que los tiempos de residencia medios deberían tener una razón de al menos 5.
- Debería ser posible tener una alta ganancia en el lazo interno.

Una situación común es que el lazo interno es una realimentación alrededor de un actuador. La variable de referencia en el lazo interno puede entonces representar una cantidad física, como caudal, presión, par, velocidad, etc., mientras la variable de control del lazo interno podría ser presión de la válvula, la corriente de control, etc. Esto es también un ejemplo típico donde la realimentación se utiliza para hacer que un sistema se comporte de una manera predictiva simple. También es muy bueno linealizar las características no lineales.

En la Figura 12.6 se muestran una serie de diferentes sistemas de control con una variable de control y dos señales de medida. En la figura, la variable de control se representa por u , la variable de medida primaria por y , la variable de medida secundaria por y_s y la perturbación esencial es v .

Con las reglas dadas, sólo el caso A es apropiado para control en cascada.

Elección de modos de control

Cuando se selecciona la señal de medida secundaria, quedan por escoger los modos de control apropiados para los controladores primario y secundario y sintonizar sus parámetros. La elección se basa en la dinámica del proceso y la naturaleza de las perturbaciones. Es difícil dar reglas generales porque las condiciones pueden variar significativamente. En casos críticos, es necesario analizar y simular. Sin embargo, es útil tener un sentimiento intuitivo de los problemas.

Considérese el sistema de la Figura 12.4. Para tener un control en cascada útil, es necesario que el proceso P_2 sea más lento que P_1 y que las perturbaciones esenciales actúen sobre P_1 . Suponemos que se satisfacen estas condiciones. El controlador secundario se puede escoger generalmente como un controlador proporcional puro o un controlador PD. En algunos casos, la acción integral puede ser útil para mejorar el rechazo a perturbaciones de baja frecuencia. Con controladores que carecen de acción integral, puede existir un error estático en el lazo secundario, aunque esto puede no ser un serio inconveniente. El lazo secundario, como regla general, se utiliza para eliminar perturbaciones rápidas. Las perturbaciones lentas se pueden eliminar fácilmente por el lazo primario, que típicamente tendrá acción integral. Hay además inconvenientes en utilizar la acción integral en el lazo secundario. Con tal sistema siempre existirá sobreelongación en la respuesta del lazo de control primario. La acción integral es necesaria si el proceso P_2 contiene retardos de tiempo esenciales y el proceso P_1 es tal que la ganancia del lazo en el lazo secundario debe ser limitada.

El caso especial donde el proceso P_2 es un integrador puro es bastante común. En este caso, la acción integral en el lazo interno corresponde al control proporcional en el lazo externo. Si se utiliza la acción integral en el lazo interno, se debe reducir la acción proporcional en el lazo externo. Esto es una desventaja significativa para el rendimiento del sistema. Un buen remedio consiste en eliminar el integrador del lazo interno e incrementar la ganancia del lazo externo.

Sintonía y puesta en funcionamiento

Los controladores en cascada se deben sintonizar en una secuencia correcta. El lazo externo se debería poner primero en modo manual cuando se sintoniza el lazo interno. El lazo interno se debería entonces poner en modo automático cuando se sintoniza el lazo externo. El lazo interno se sintoniza generalmente para amortiguamientos críticos o muy críticos, o equivalentemente para una sensibilidad pequeña (M_s). Si esto no se lleva a cabo, hay un pequeño margen para utilizar realimentación en el lazo externo.

La puesta en marcha de los lazos en cascada requiere además algunas consideraciones. Se puede utilizar el siguiente procedimiento, comenzando con ambos controladores en modo manual.

1. Ajustar el punto de consigna del controlador secundario al valor de la variable del proceso secundario.
2. Poner el controlador secundario en modo automático con el punto de consigna interno seleccionado.
3. Ajustar el controlador primario tal que su punto de consigna sea igual a la variable del proceso y sea su señal de control igual al punto de consigna del

controlador secundario.

4. Comutar el controlador secundario al punto de consigna externo.
5. Comutar el controlador primario a modo automático.

Los pasos dados anteriormente se automatizan con diversos grados en los diferentes controladores. Si el procedimiento no se realiza en el modo correcto, se producirán transitorios en las comutaciones.

Saturación del término integral

Si se utiliza la acción integral en ambos lazos de control, secundario y primario, es necesario tener una esquema para evitar el fenómeno de la saturación del término integral ó windup. El lazo interno se puede manejar de la manera típica, pero no es una tarea trivial evitar el windup en el lazo externo. Existen tres situaciones que se deberían considerar:

1. La señal de control en el lazo interno se puede saturar.
2. El lazo de control secundario puede comutarse al punto de consigna interno.
3. El controlador secundario se conmuta de modo automático a manual.

El lazo de realimentación, tal como se ve desde el controlador primario, se rompe en todos estos casos, y es necesario asegurar que su modo integral se trata adecuadamente. Este problema se resuelve de manera automática en una serie de controladores de proceso que incorporan capacidades de control en cascada, pero si construimos el control en cascada utilizando dos controladores independientes, tenemos que resolver el problema por nuestra cuenta. Esto requiere ser capaces de inyectar la señal de seguimiento en el controlador primario.

Si la señal de salida del controlador secundario está limitada, la variable de proceso del controlador secundario se debería escoger como una señal de seguimiento en el controlador primario. Esto también requiere una transferencia digital desde el controlador secundario al primario indicándole cuándo el seguimiento tiene lugar.

En el caso donde el controlador secundario comuta para trabajar de acuerdo con su punto de consigna local en lugar del controlador primario, el punto de consigna local se debería reenviar al controlador primario como una señal de seguimiento. De esta forma, se puede evitar tanto el windup del integrador como los saltos en la transición a control en cascada.

Cuando el controlador secundario comuta a modo manual, la variable de proceso del controlador secundario se debería reenviar al controlador primario como una señal de seguimiento.

Algunas aplicaciones

El control en cascada es una manera conveniente de utilizar medidas extras para mejorar el rendimiento del control. Los siguientes ejemplos muestran algunas aplicaciones.

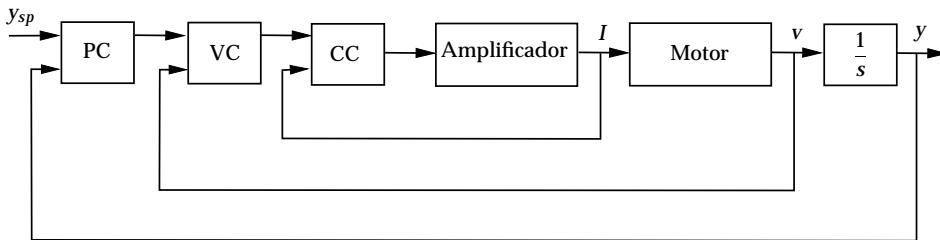


Figura 12.7 Diagrama de bloques de un sistema de control de posición. El sistema tiene tres lazos en cascada con un controlador de corriente (CC) con realimentación de la corriente (I), un controlador de velocidad (VC) con realimentación de la velocidad (v) y un controlador de posición (PC) con realimentación de la posición (y).

EJEMPLO 12.3—POSICIONADORES DE VÁLVULA

Los lazos de control con válvulas neumáticas son una aplicación muy común. En este caso, el lazo interno es una realimentación alrededor de la propia válvula donde se mide la posición de la válvula. El lazo interno reduce las influencias de las variaciones de presión y las distintas no linealidades en el sistema neumático. □

EJEMPLO 12.4—CONTROL DE UN MOTOR

La Figura 12.7 es un diagrama de bloques de un sistema típico de control de un motor. El sistema tiene tres lazos en cascada. El lazo más interno es un lazo de corriente donde se mide dicha variable. El siguiente lazo es el lazo de velocidad, que se basa en la medida de velocidad. El lazo más externo es un lazo de posición. En este caso, la acción integral en el lazo de velocidad es equivalente a una acción proporcional en el lazo de posición. Además, está claro que la acción derivativa del lazo de posición es equivalente a una acción proporcional en el lazo de velocidad. De esto, se deduce directamente que no hay razón para introducir acción integral en el controlador de velocidad o en la acción derivativa en el controlador de posición. □

EJEMPLO 12.5—INTERCAMBIADOR DE CALOR

En la Figura 12.8 se muestra un diagrama esquemático de un intercambiador de calor. El propósito del sistema de control es controlar la temperatura de salida del lado secundario cambiando la válvula del lado primario. El sistema de control utiliza un control en cascada. El lazo secundario es un sistema de control de caudal alrededor de la válvula. La variable de control del lazo primario es el punto de consigna del controlador de caudal. El efecto de las no linealidades en la válvula, así como de las perturbaciones de caudal y presión, se pueden por lo tanto reducir mediante el controlador secundario. □

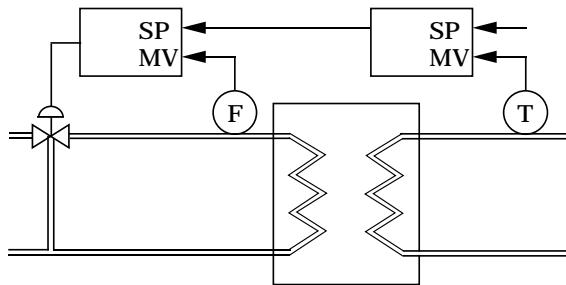


Figura 12.8 Diagrama esquemático de un intercambiador de calor con control en cascada.

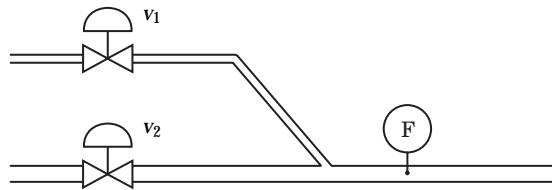


Figura 12.9 Se utilizan dos válvulas para controlar el caudal.

12.5 Control de rango medio y de rango partido

El control en cascada es una estrategia donde se utiliza una señal de control y señales de medida para conseguir dos objetivos de control. La situación dual es cuando se utilizan dos señales de control para controlar una señal de medida. Las dos señales de control se usan algunas veces una cada vez. Este es el caso del control de rango partido. En otras situaciones, es necesario emplear las dos señales de control de manera simultánea. Una situación común es el control de rango medio. En esta sección se describen los controles de rango medio y rango partido.

Control de rango medio

El problema tratado por el control de rango medio se ilustra en la Figura 12.9. La figura ilustra un ejemplo donde se utilizan dos válvulas para controlar un caudal. Una válvula, v_1 es pequeña pero tiene una alta resolución. La otra válvula, v_2 , es grande pero tiene una resolución baja.

Supóngase que la válvula pequeña v_1 está en la mitad de su rango de operación y que sólo perturbaciones pequeñas están actuando sobre el sistema. En este caso, un controlador que manipula la válvula v_1 es capaz de encargarse del problema de control. Sin embargo, cuando se producen grandes perturbaciones, la válvula v_1 se saturará. En este caso, la válvula más grande v_2 debe también manipularse.

En la Figura 12.10 se ilustra la estrategia de control de rango medio. El controlador C_1 toma el punto de consigna y_{sp} y la señal de caudal y como entradas y manipula la válvula pequeña v_1 . Un segundo controlador, C_2 , toma la señal de control de C_1 como entrada e intenta llevarla a un punto de consigna u_{sp} en la mitad de su rango de operación manipulando la válvula grande v_2 . Si ambos

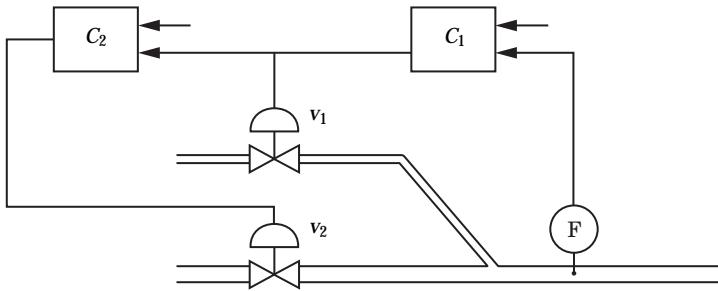


Figura 12.10 Control de rango medio

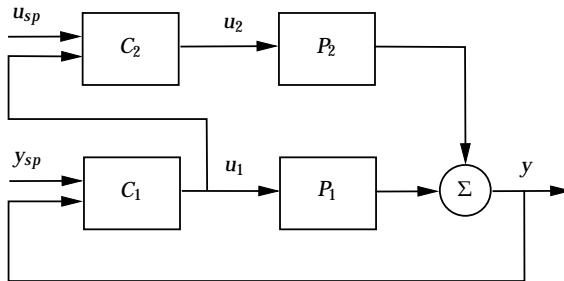


Figura 12.11 Diagrama de bloques de un sistema con control de rango medio.

controladores tiene acción integral, el flujo estará en el punto de consigna y_{sp} y la válvula v_1 estará en el punto de consigna u_{sp} en estado estacionario.

En la Figura 12.11 se muestra un diagrama de bloques de la estrategia de control de rango medio. El proceso P_1 y el controlador C_1 juntos forman un lazo de respuesta rápida. El controlador de rango medio C_2 controla la posición de la válvula del controlador C_1 vía la salida del proceso y . Esto quiere decir que la salida del controlador C_1 se controla retirando la salida del proceso y y del punto de consigna. Si esto se hace lentamente, la desviación del punto de consigna se puede mantener pequeña. Si no, se recomienda utilizar la estructura que se muestra en la Figura 12.12.

En la Figura 12.12 se añade una señal feedforward desde la señal de control u_2 al controlador C_1 . Si el compensador feedforward es

$$C_{ff}(s) = -\frac{P_2(s)}{P_1(s)},$$

el controlador C_2 llevará a cabo el control de rango medio sin ninguna perturbación de la salida del proceso y .

Es probable que la válvula pequeña se sature. A pesar de esto, no es necesario que el controlador C_1 tenga anti windup. Como la señal de control se controla por el controlador C_2 , esto evita el windup del controlador C_1 .

Control de rango partido

En el control de rango partido, el control se comparte por dos controladores que realizan el control de uno en uno. Los sistemas de este tipo son comunes, por

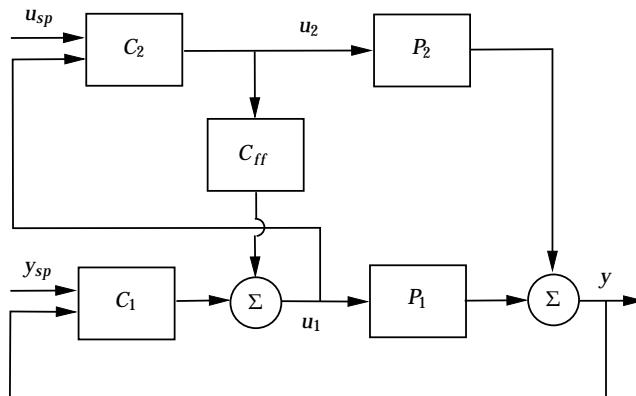


Figura 12.12 Diagrama de bloques de un sistema con control de rango medio.

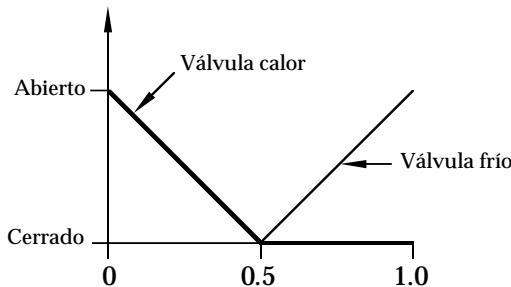


Figura 12.13 Ilustración del concepto de control de rango partido.

ejemplo, en la conexión calentamiento y enfriamiento. Un dispositivo físico se utiliza para calentar y otro para enfriar. Los sistemas de calentamiento y enfriamiento generalmente tienen características estáticas y dinámicas diferentes. En la Figura 12.13 se ilustra el principio del control de rango partido se ilustra, que muestra la relación estática entre las variables de medida y las variables de control. Cuando la temperatura es demasiado lenta, es necesario suministrar calor. Por lo tanto, el calentamiento alcanza su valor máximo cuando la variable de medida es cero. Entonces, decrece linealmente hasta la mitad del rango, donde no se suministran calor. Similarmente, no se aporta enfriamiento cuando la variable de medida está por debajo la mitad del rango. Sin embargo, el enfriamiento se aplica cuando la variable del proceso está por encima la mitad del rango, y entonces aumenta.

Hay una región crítica cuando se comuta de calentamiento a enfriamiento. Para evitar que se produzcan calentamiento y enfriamiento al mismo tiempo, generalmente existe una pequeña zona muerta donde no se aplica ni calor ni frío. Las comutaciones entre los distintos modos de control pueden causar dificultades y oscilaciones.

El control de rango partido se utiliza comúnmente en sistemas para calentamiento y ventilación. También es útil en aplicaciones donde la variable de control varía sobre un rango muy grande. El caudal se separa entonces en caminos

paralelos, cada uno controlado con una válvula.

12.6 Elementos no lineales

Los elementos no lineales se han comentado previamente. En la Sección 3.5, utilizamos un limitador para evitar el fenómeno del windup en un controlador con acción integral. En el Capítulo 9, se mostró que los controladores se podían sintonizar mediante la realimentación de un relé y que el rendimiento se podría mejorar mediante ajuste de ganancia por tabla. En esta sección describimos mas elementos no lineales y también presentamos algunos paradigmas de control que sirven de guía para el uso de estos elementos.

Linealización

La no linealidad en sensores y actuadores se puede compensar de una manera directa. Considérese, por ejemplo, un actuador que tiene las características

$$v = f(u),$$

donde v es la señal de entrada al proceso y u la señal de control. Para compensar la no linealidad, simplemente calculamos la señal de control u_c como si el actuador fuese lineal con ganancia unitaria. La ley de control es

$$u = f^{-1}(u_c),$$

donde f^{-1} es la inversa de la no linealidad del actuador, obteniéndose entonces

$$v = f(u) = f(f^{-1}(u_c)) = u_c.$$

La señal del proceso actuada es entonces idéntica a u_c como se deseaba.

La misma idea se puede aplicar a los sensores. Considérese, por ejemplo, un sensor que tiene la no linealidad $g(x)$. Diseñando un controlador lineal basado en la hipótesis de que el sensor es lineal con ganancia unitaria y alimentando la señal

$$y_c = g^{-1}(y)$$

al controlador, se elimina la no linealidad del sensor.

Se pueden aplicar ideas similares a las no linealidades de los procesos, pero la compensación no es ideal debido a la dinámica. Existe una técnica para compensar las no linealidades denominada linealización por realimentación, pero queda fuera del objetivo de este libro. Existen también situaciones donde las no linealidades son beneficiosas.

Limitadores

Como todos los valores físicos están limitados, es útil también tener dispositivos limitadores en los sistemas de control. Los limitadores se utilizan de distintas maneras. Se pueden emplear para limitar las señales de referencia de forma que

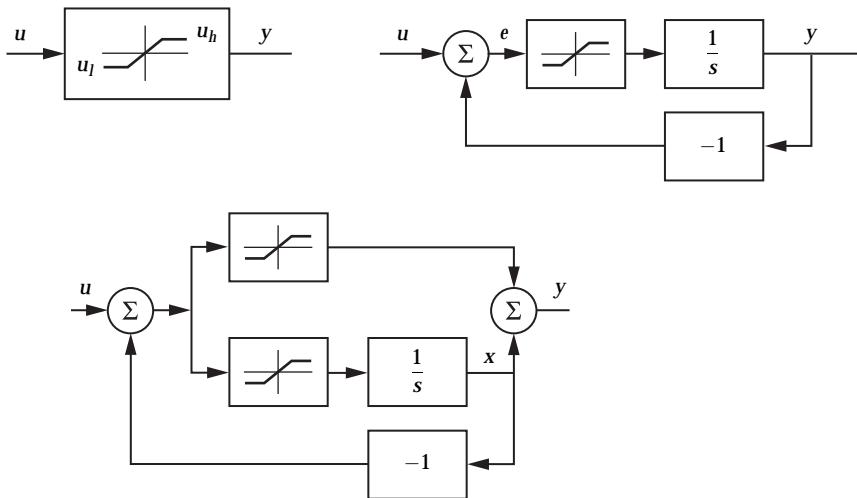


Figura 12.14 Diagrama de bloques de un limitador de amplitud simple (parte superior izquierda), un limitador de velocidad (parte superior derecha) y un limitador de salto y velocidad ó de rampa unitaria (parte inferior).

no estemos generando puntos de consigna que estén demandando cambios mayores o más rápidos de los que puede afrontar un sistema.

En la parte superior izquierda de la Figura 12.14 se muestra un diagrama de bloques de un limitador de amplitud simple. El limitador se puede describir matemáticamente como la no linealidad estática

$$y = \text{sat}(u, u_l, u_h) = \begin{cases} u_l & \text{si } u \leq u_l \\ u & \text{si } u_l < u < u_h \\ u_h & \text{si } u \geq u_h \end{cases}$$

donde u_l y u_h son los límites de saturación.

También es útil limitar la velocidad de cambio de las señales. Esto se puede hacer con el *limitador de velocidad* o *unidad de rampa* que se muestran en la parte superior derecha de la Figura 12.14. La salida sigue la señal de entrada si la velocidad de cambio de la entrada es más pequeña que el límite de velocidad. En estado estacionario, las entradas y las salidas son idénticas porque existe acción integral en el sistema. Como la salida se genera mediante un integrador con señal de entrada limitada, la velocidad de cambio de la salida estará limitada a los límites dados por el limitador. Es posible utilizar límites diferentes aumentando o disminuyendo las velocidades.

En la parte inferior de la Figura 12.14 se muestra un limitador más sofisticado denominado *limitador de salto y velocidad*. La salida seguirá a la entrada para cambios pequeños en la señal de entrada. Para cambios grandes, la salida seguirá a la entrada con una velocidad limitada. El limitador de salto y velocidad se puede

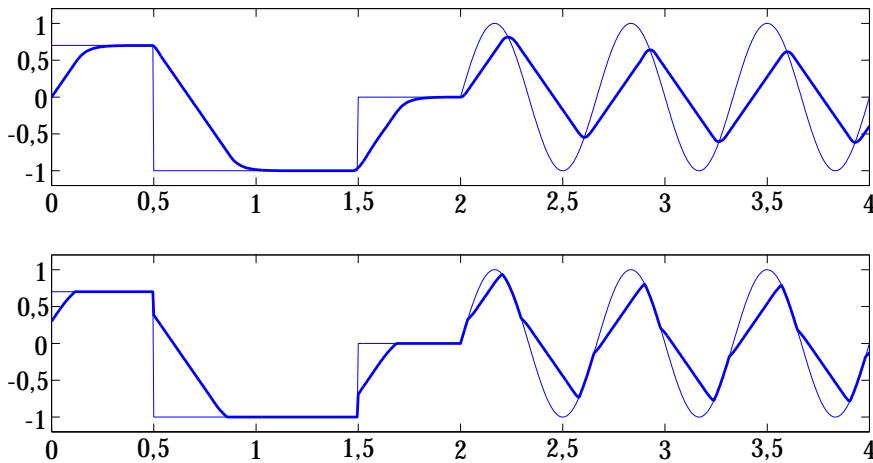


Figura 12.15 Simulación de un limitador de frecuencia (parte superior) y un limitador de salto y frecuencia (parte inferior). La línea delgada muestra la salida del limitador y la línea gruesa muestra la salida del limitador.

describir por

$$\frac{dx}{dt} = \text{sat}(u - x, -a, a)$$

$$y = x + \text{sat}(u - x, -a, a),$$

Si $|u - x| \leq a$, se deduce de las ecuaciones del sistema que $y = u$, y si $u \geq x + a$ se deduce que $dx/dt = a$. Por lo tanto, la señal de salida se aproximará a la señal de entrada a la velocidad a .

Las propiedades de los distintos limitadores se ilustran en la simulación mostrada en la Figura 12.15. La señal de entrada consiste en unos pocos escalones y una sinusode. La curva superior muestra un limitador de velocidad cuando el límite de velocidad es 4. La figura muestra que la velocidad de cambio de la salida está limitada. La respuesta a una entrada sinusoidal muestra claramente que el limitador de velocidad introduce un retraso de fase. La curva inferior muestra la respuesta de un limitador de salto y velocidad. Obsérvese que la salida sigue los cambios rápidos en la entrada siempre que la diferencia entre x y u sea menor que el límite de salto, que es 0,5. La velocidad está limitada a 4.

Control de un tanque de carga

Los problemas de control que se analizaron en el Capítulo 4 fueron todos problemas de regulación donde la tarea consistía en mantener una variable de proceso tan cerca como fuese posible de un punto de consigna dado. Hay muchos otros problemas de control que también son importantes. El control de un tanque de carga es un ejemplo. El propósito de un tanque de carga es actuar como un almacén entre diferentes procesos de producción. El caudal de un proceso se alimenta a otro a través del tanque de carga. Las variaciones en la velocidad de producción se pueden acomodar permitiendo que varíe el nivel del tanque de

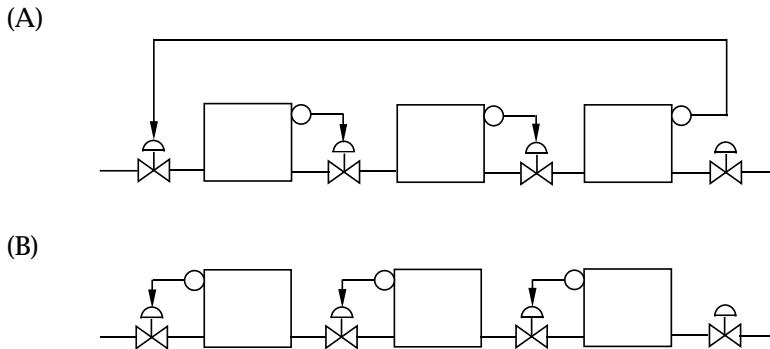


Figura 12.16 Diferentes estructuras para el control del tanque de carga. El caudal de material va de izquierda a derecha. El esquema en A se denomina control en la dirección del caudal. El esquema en B se denomina control en la dirección opuesta al caudal.

carga. Un control de nivel convencional, que trata de mantener el nivel constante, claramente no es apropiado en este caso. Para actuar como un almacén, el nivel debería cambiar realmente.

Existen muchos enfoques para el control del tanque de carga. Una solución común sencilla es utilizar un controlador proporcional con baja ganancia. También se utilizan controladores con zonas muertas o controladores PI no lineales. Un controlador con planificación de ganancia es un método mejor. El nivel del tanque se escoge como la variable de planificación. Se escoge un controlador con baja ganancia cuando el nivel está, por ejemplo, entre el 10 y el 90 por cien, y un controlador con alta ganancia cuando está fuera de estos límites. Hay también esquemas especiales para el control del tanque de carga.

En muchos casos, existen grandes secuencias de tanques de carga y unidades de producción, tal y como se muestra en la Figura 12.16. En la figura se muestran dos estructuras diferentes de control, control en la dirección del caudal y en la dirección opuesta del caudal. Con el control en la dirección del flujo, es sencillo obtener oscilaciones o inestabilidades a causa de la realimentación desde el final de la cadena al inicio.

Control de relación

El control de relación se aplica cuando el objetivo de control consiste en mantener la relación entre dos variables, generalmente caudales, en una cierta razón a . En combustión, por ejemplo, se desea controlar la relación de suministro fuel-aire para que la combustión sea tan eficiente como sea posible. Las mezclas químicas es otro ejemplo donde se desea mantener constante la razón entre diferentes caudales. En los sistemas de mezcla en línea, cuando no hay tanques de mezcla en sentido hacia abajo, esto es de especial importancia.

El control de relación se resuelve normalmente de la forma mostrada en la Figura 12.17. Existen dos lazos de control. El lazo principal consiste en un proceso P_1 y un controlador C_1 . La salida y_1 es el caudal principal y el punto de consigna externo r_1 es el caudal principal deseado. En el segundo lazo, compuesto por el proceso P_2 y el controlador C_2 , se intenta controlar el caudal y_2 tal que la razón y_2/y_1 sea igual a la razón a . En la Figura 12.17, esto se obtiene utilizando una

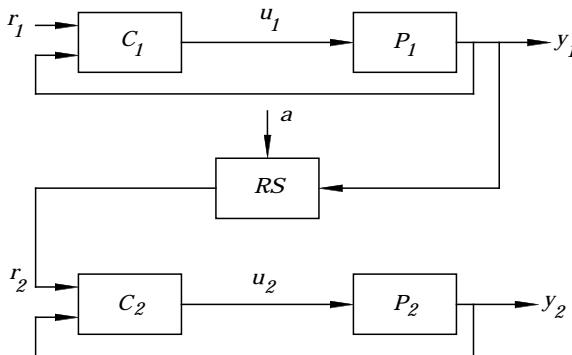


Figura 12.17 Control de relación utilizando una estación de relación (RS) aplicada al caudal principal y_1 .

estación de relación donde el punto de consigna r_2 se determina por

$$r_2(t) = ay_1(t), \quad (12.4)$$

por ejemplo, multiplicando simplemente el caudal principal y_1 por la razón deseada a .

En la ecuación 12.4, el parámetro a se supone constante, aunque esto no es necesario. La razón deseada a es generalmente variable en el tiempo. En combustión, por ejemplo, la razón a se ajusta generalmente basándose en las medidas de O_2 en el escape.

Si los controladores tienen acción integral, la solución mostrada en la Figura 12.17 funcionará en régimen permanente, por ejemplo, $y_1 = r_1$ e $y_2 = ay_1$. Sin embargo, la estación simple de relación no es eficiente durante los transitorios. El segundo caudal y_2 estará siempre retrasado en comparación con el caudal deseado ay_1 . La longitud de este retraso se determina mediante la dinámica del segundo lazo.

Cuando el punto de consigna r_1 está aumentando, el retraso provoca un bajo suministro del producto correspondiente al caudal y_2 , y en cambio cuando r_1 está disminuyendo se produce un exceso. Existen casos donde es importante que nunca se produzca un bajo suministro de alguno de los dos productos. En el caso de la combustión, uno obtiene un suministro bajo de aire durante la parte del transitorio cuando aumenta el punto de consigna externo, pero se produce un exceso cuando el punto de consigna decrece. La solución mostrada en la Figura 12.17 tiene que completarse con algunos selectores lógicos con el fin de evitar que todo el combustible no sea quemado por un suministro bajo de aire. Esto se trata en la siguiente sección.

El principal inconveniente con el enfoque de la estación de simple relación que se muestra en la Figura 12.17 es que el caudal secundario y_2 se retrasa en comparación con el caudal deseado y_1 . Este problema se puede resolver si no se utiliza sólo y_1 para formar el punto de consigna secundario, sino que también se emplea el punto de consigna principal r_1 . En la Figura 12.18 se muestra la estructura, denominada estación de mezcla .

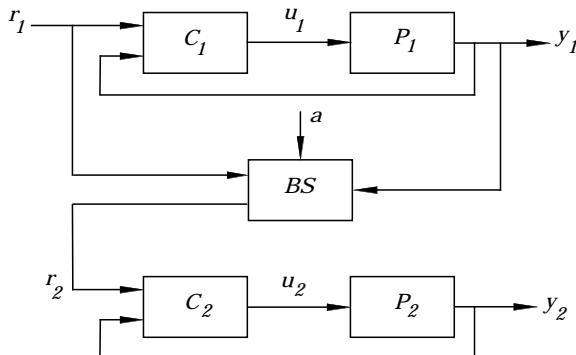


Figura 12.18 Control de relación utilizando la estación de mezcla (BS).

En la estación de mezcla, el punto de consigna secundario se determina como

$$r_2(t) = a(\gamma r_1(t) + (1 - \gamma)y_1(y)). \quad (12.5)$$

La ganancia γ es un factor de ponderación que determina la relación entre el punto de consigna r_1 y el caudal principal y_1 cuando forman el punto de consigna secundario r_s . Cuando $\gamma = 0$, la estación de mezcla es idéntica a la estación de relación.

EJEMPLO 12.6—CONTROL DEL BLANQUEADO DE PULPA

Las estaciones de relación y de mezcla han sido aplicadas en una sección de blanqueado en una fábrica de papel. La pulpa se blanquea añadiendo hidrosulfito al caudal de pulpa. El objeto es mantener constante la razón entre el caudal de pulpa y la de caudal de hidrosulfito.

El diagrama superior de la Figura 12.19 muestra el control utilizando la estación de relación. El controlador del caudal de pulpa, C_1 , es un controlador PI con $K_1 = 0,2$ y $T_{i1} = 4\text{s}$, y el controlador de hidrosulfito, C_2 , es también un controlador PI con $K_2 = 0,078$ and $T_{i2} = 1,07\text{s}$. La figura muestra las respuestas a dos cambios en el punto de consigna en el caudal de pulpa. El caudal de hidrosulfito se escala con la razón deseada, tal que las velocidades de los caudales deseados llegan a ser idénticas. La figura muestra que la estación de relación proporciona la razón correcta en régimen permanente, pero también que hay una desviación entre los dos caudales durante los transitorios. El caudal de hidrosulfito se retrasa en comparación con el caudal de pulpa.

El diagrama inferior de la Figura 12.19 muestra los resultados obtenidos cuando se utiliza la estación de Mezcla con un factor de ganancia $\gamma = 0,75$. En este caso, prácticamente se elimina la diferencia entre los dos caudales.

□

Control selector

El control selector se puede interpretar como la inversa del control de rango partido. En el rango partido existe una señal de medida y varios actuadores. En el

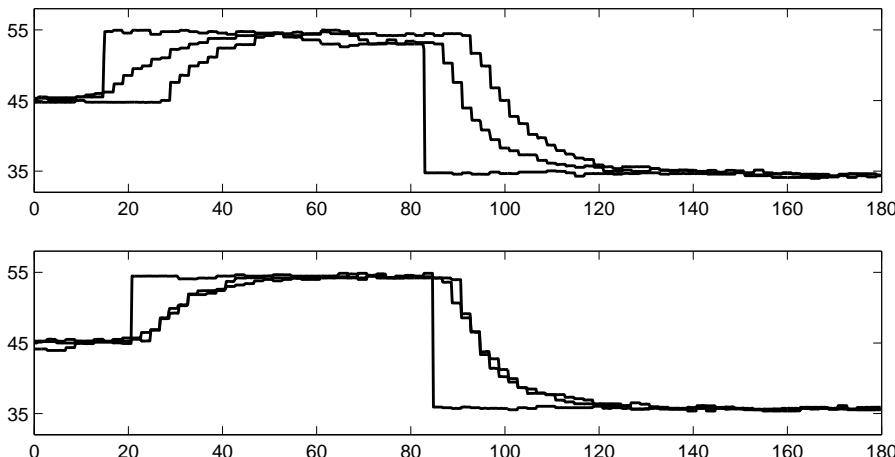


Figura 12.19 Control de relación de un proceso de blanqueado de pulpa utilizando la estación de Relación original (parte superior) y la estación de Mezcla con ganancia $\gamma = 0,75$ (parte inferior). La figura muestra dos cambios en el punto de consigna de la pulpa, el caudal de pulpa (respuesta más rápida) y el caudal de hidrosulfito (respuesta más lenta).

control selector existen muchas señales de medida y un solo actuador. Un selector es un dispositivo estático con muchas entradas y una salida. Existen dos tipos de selectores: *máximo* y *mínimo*. Para un selector de máximo la salida es la más grande de las señales de entrada.

Hay situaciones donde deben tenerse en cuenta algunas variables del proceso controladas. Una variable es la variable controlada primaria, pero se requiere también que otras variables del proceso permanezcan en unos rangos dados. El control selector se puede utilizar para conseguir este objetivo. La idea consiste en utilizar varios controladores y tener un selector que escoja el controlador más apropiado. Un ejemplo de uso es donde la variable controlada primaria es la temperatura y debemos asegurar que la presión no exceda de un cierto rango por razones de seguridad.

En la Figura 12.20 se ilustra el principio del control selector. La variable controlada primaria es la salida del proceso y . Existe una variable de medida auxiliar z que se debería mantener en los límites z_{\min} y z_{\max} . El controlador primario C tiene la variable del proceso y , el punto de consigna y_{sp} y la salida u_n . Existen además controladores secundarios con variables del proceso medidas que son la variable auxiliar z y los puntos de consigna que son límites de la variable z . Las salidas de estos controladores son u_h y u_l . El controlador C es un controlador PI o PID ordinario que proporciona un buen control bajo circunstancias normales. La salida del selector de mínimo es la más pequeña de las señales de entrada y la del selector de máximo es la más grande de las entradas.

Bajo circunstancias normales, la variable auxiliar es mayor que el mínimo valor z_{\min} y más pequeña que el máximo valor z_{\max} . Esto quiere decir que la salida u_h es grande y que la salida u_l es pequeña. Por lo tanto, tanto el selector de máximo como el de mínimo seleccionan u_n . El sistema actúa como si los controladores de máximo y de mínimo no estuvieran presentes. Si la variable z alcanza su límite

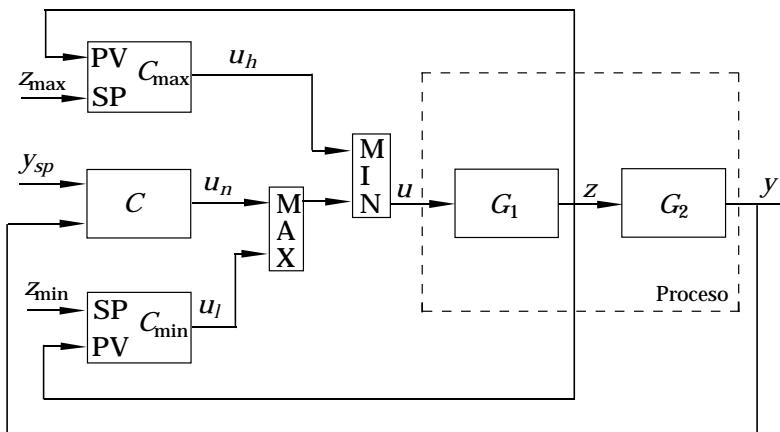


Figura 12.20 Control selector.

superior, la variable u_h llega a ser pequeña y se selecciona por el selector de mínimo. Esto significa que el sistema de control intenta ahora controlar la variable z y conducirla hacia su límite. Una situación similar sucede si la variable z se hace más pequeña que z_{\min} .

En un sistema con selectores, sólo un lazo de control está operando al mismo tiempo. Los controladores se pueden sintonizar de la misma forma que los controladores de lazo único. Cuando los controladores comutantes pueden existir algunas dificultades. Para controladores con acción integral, es además necesario seguir los estados de la integral de aquellos controladores que no están en operación. El control selector es muy común para garantizar que las variables permanezcan dentro de las restricciones. La técnica se utiliza comúnmente en las centrales eléctricas para el control de calderas, sistemas de potencia y reactores nucleares. La ventaja es que se construye a partir de componentes no lineales simples y controladores PI y PID. Una alternativa al control selector es hacer una combinación de controladores ordinarios y lógica. El siguiente ejemplo ilustra el uso del control selector.

EJEMPLO 12.7—CONTROL DE AIRE-COMBUSTIBLE

En la sección anterior comentamos el control de aire-combustible utilizando el control de relación. Cuando se utiliza la estación de relación, puede haber falta de aire a causa de que el punto de consigna del controlador de aire aumenta primero cuando el controlador de combustible ha incrementado el caudal de combustible. Una forma de resolver este problema es utilizar la estación de mezcla. Sin embargo, el sistema no puede compensar las perturbaciones en el canal de aire. Este problema se puede tratar utilizando selectores tal y como se muestra en la Figura 12.21. El sistema utiliza un selector de mínimo y uno de máximo. Hay un controlador PI para el caudal de combustible y otro para el caudal de aire. El punto de consigna para el controlador de aire es el máximo valor entre la señal de referencia y el caudal de combustible. Esto quiere decir que el flujo de aire aumentará tan pronto como se demande más energía. Similarmente, el punto de consigna del combustible es el mínimo valor entre la señal de demanda y el

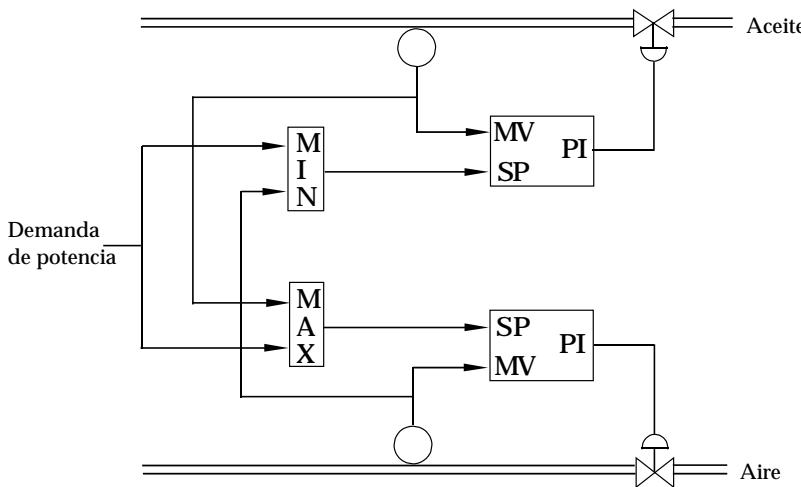


Figura 12.21 Controlador de aire-combustible basado en selectores

caudal de aire. Esto quiere decir que cuando se aumenta la demanda, el punto de consigna del controlador de caudal dual disminuirá inmediatamente, pero el punto de consigna del controlador de aire permanecerá alto hasta que el caudal de combustible haya disminuido realmente. El sistema así asegura que siempre habrá un exceso de aire. □

Selectores de mediana

Un selector de mediana es un dispositivo con muchas entradas y muchas salidas. Su salida selecciona la entrada que representa la mediana actual de las señales de entrada. Un caso especial es un selector de dos salidas de tres, que se utiliza comúnmente para sistemas altamente sensibles. Es posible utilizar sensores y controladores redundantes con el fin de alcanzar una alta fiabilidad. Incorporando selectores de mediana es posible tener un sistema que continuará funcionando incluso si fallan algunos componentes.

12.7 Control de red neuronal

En la sección anterior hemos visto que las no linealidades simples se pueden utilizar de manera muy efectiva en los sistemas de control. En esta sección y en la siguiente, analizaremos algunas técnicas basadas en no linealidades, donde la idea clave consiste en representar funciones de varias variables de una forma compacta. Las ideas se han introducido bajo los nombres de control *neuronal* y *control borroso*. A primera vista, estos métodos pueden parecer bastante complicados, pero una vez se profundiza encontramos que el algoritmo tiene representaciones naturales en forma de implementaciones de funciones no lineales. El encontrar buenas representaciones de funciones lineales no es un problema trivial.

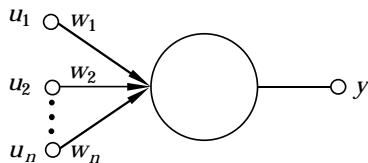


Figura 12.22 Diagrama esquemático de una neurona simple.

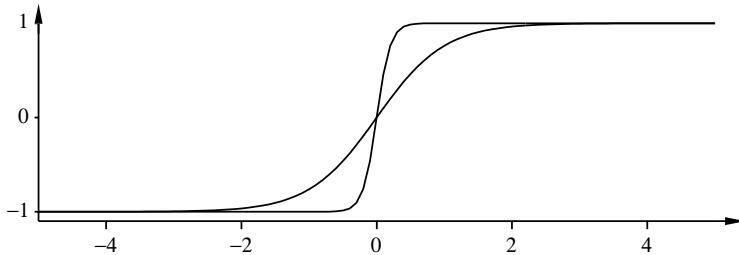


Figura 12.23 Funciones sigmoidales.

Si simplemente intentamos cuantificar las variables y utilizar una interpolación encontramos que el número de entradas en la tabla para representar las funciones crece muy rápidamente con el número de variables. Por ejemplo, si n variables se dividen en cuadrículas de N puntos cada una, encontramos que el número de entradas son N^n . Para una función de cinco variables con $N = 100$ encontramos que se requieren 10^{10} entradas. Otra propiedad útil de las redes neuronales es que hay métodos para ajustar los parámetros de la función a los datos.

Redes neuronales

Las redes neuronales se originaron en intentos de construir modelos simples de la actividad neuronal en el cerebro y hacer dispositivos que pudieran reconocer patrones y llevar a cabo tareas de aprendizaje sencillas. A continuación se muestra una breve descripción que engloba las ideas clave.

Una neurona simple La Figura 12.22 muestra un diagrama esquemático de una neurona simple. El sistema tiene muchas entradas y una salida. Si la salida es y y las entradas son u_1, u_2, \dots, u_n , la relación entrada-salida viene descrita por

$$y = f(w_1 u_1 + w_2 u_2 + \dots + w_n u_n) = f\left(\sum_{k=1}^n w_k u_k\right), \quad (12.6)$$

donde los términos w_i se denominan pesos. La función f se denomina función sigmoidal, y se ilustra en la Figura 12.23. Tal función se puede representar como

$$f(x) = \sinh \alpha x = \frac{e^{\alpha x} - e^{-\alpha x}}{e^{\alpha x} + e^{-\alpha x}} \quad (12.7)$$

donde α es un parámetro. Este modelo de una neurona es así simplemente una función no lineal. Algunas clases especiales de funciones se pueden aproximar por (12.6).

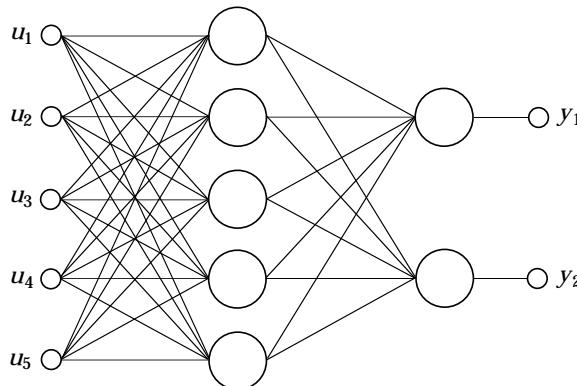


Figura 12.24 Una red neuronal feedforward.

Redes neuronales Se pueden obtener modelos más complicados conectando neuronas entre sí tal y como se muestra en la Figura 12.24. Este sistema se denomina una red neuronal. Generalmente se añade el adjetivo feedforward para indicar que las neuronas están conectadas de una manera feedforward. También existen otros tipos de redes neuronales. En la red feedforward, las neuronas de entrada se conectan a una capa de neuronas, las salidas de las neuronas de la primera capa se conectan a las neuronas de la segunda capa y así sucesivamente hasta que tenemos las salidas. Las capas intermedias en la red se denominan capas ocultas.

Cada neurona se describe por la ecuación (12.6). La relación entrada-salida de una red neuronal es por tanto una función no lineal estática. Consecuentemente, podemos considerar una red neuronal como una forma de construir una función no lineal de varias variables. La representación de la red neuronal implica que una función no lineal de varias variables se construye a partir de dos componentes: una función no lineal simple, la función sigmoidal (12.7) que es una función escalar de una variable, y de operaciones lineales. Es así una manera simple de construir una no linealidad a partir de operaciones simples. Una razón clave por la cual las redes neuronales son interesantes es que prácticamente todas las funciones continuas se pueden aproximar por redes neuronales que contienen una capa oculta. Se ha encontrado que es práctico utilizar más capas ocultas porque entonces se pueden emplear menos pesos. Otra característica práctica de las funciones sigmoidales es que las aproximaciones son locales.

Aprendizaje Obsérvese que una red neuronal tiene muchos parámetros (pesos). Suponiendo que hay n neuronas en una capa, si todas las neuronas se conectan, se requieren n^2 parámetros para describir las conexiones entre dos capas.

Una propiedad interesante de una red neuronal es que existen procedimientos de aprendizaje. Esto es un algoritmo que hace posible encontrar los parámetros (pesos) tal que la función se ajuste a los valores dados de entrada-salida. Los parámetros se obtienen típicamente de manera recursiva dado un valor de entrada a la función y el valor de salida deseado. Entonces, se adaptan los pesos de forma que los datos se ajusten. Posteriormente, se da un nuevo par entrada-salida

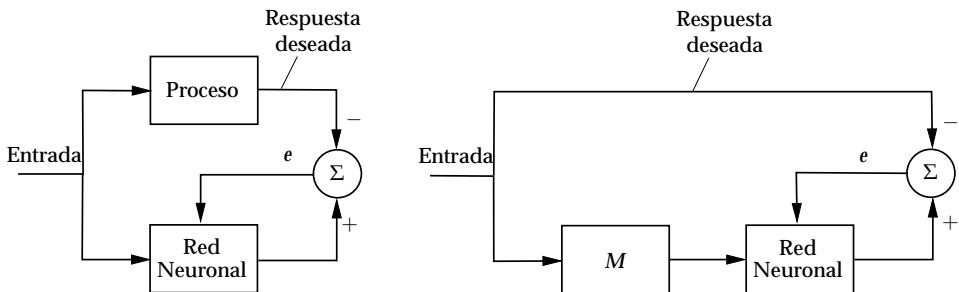


Figura 12.25 Ilustración del entrenamiento de una red feedforward simple. El diagrama de bloques de la izquierda muestra el entrenamiento de una función, y la figura de la derecha el entrenamiento de la función inversa.

y los parámetros se ajustan de nuevo. El procedimiento se repite hasta que se obtenga un buen ajuste para un conjunto de datos razonable. Este procedimiento se denomina entrenamiento de una red. Un método denominado propagación hacia atrás es un método popular para el entrenamiento de redes feedforward. Por esta razón, la red feedforward algunas veces se denomina red de programación hacia atrás. La Figura 12.25 muestra el ajuste de una red neuronal a un conjunto de datos experimentales. Una características interesante es que es posible encontrar tanto la función como su inversa. La función inversa es útil para la compensación de no linealidades en sensores y actuadores.

Aplicaciones de control Una red neuronal feedforward se puede ver como una función no lineal de varias variables con un procedimiento de entrenamiento. La función tiene muchos parámetros (pesos) que se pueden ajustar mediante el procedimiento de entrenamiento de forma que la función se ajusta al conjunto de datos dado. Es un componente del sistema útil incluso aunque es un modelo extremadamente simple de una neurona real. En el control de procesos generalmente podemos hacer buen uso de funciones no lineales, como por ejemplo en la calibración de sensores. Hay muchas situaciones donde un instrumento tiene muchos sensores diferentes, las salidas de las cuales deben combinarse no linealmente para obtener el valor de medida deseado. Las funciones no lineales se pueden utilizar también para reconocimiento de patrones. También, es posible modelar sistemas dinámicos combinando redes neuronales con integradores tal como se ilustra en la Figura 12.26. El sistema de la figura implementa el sistema no lineal

$$\frac{dx}{dt} = f(x, u)$$

$$y = g(x, u),$$

donde las funciones no-lineales se representan por redes neuronales.

12.8 Control borroso

El control borroso es un paradigma de control antiguo que recientemente ha recibido una gran atención . En esta sección daremos una breve descripción de

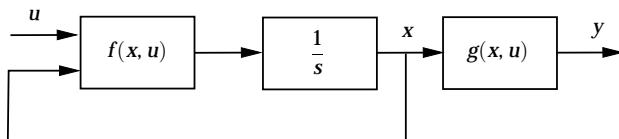


Figura 12.26 Implementación de un sistema dinámico no-lineal utilizando integradores y redes neuronales.

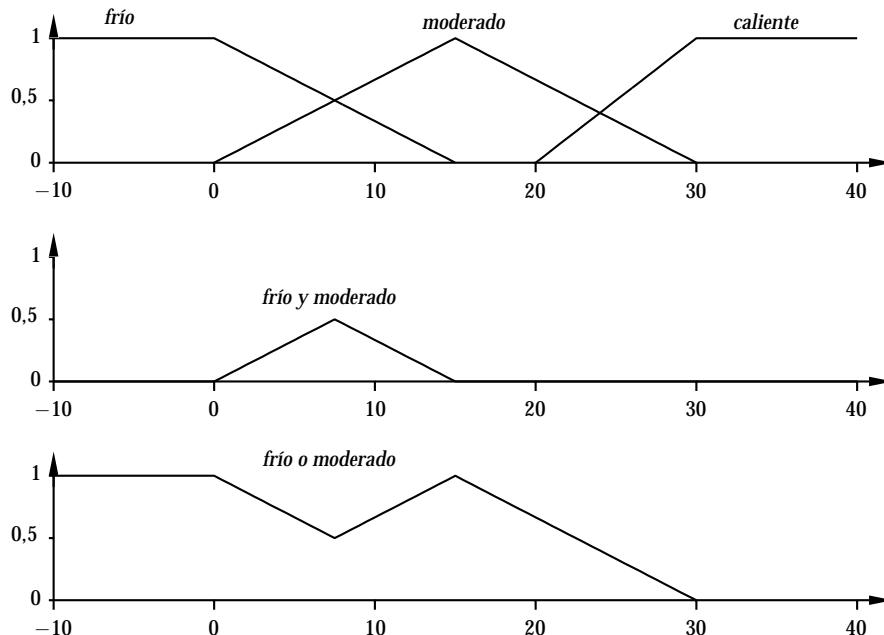


Figura 12.27 Ilustración de la lógica borrosa. El diagrama superior muestra las funciones de pertenencia de *frío*, *moderado* y *caliente*. El diagrama intermedio muestra las funciones de pertenencia para *frío* y *moderado* y el diagrama inferior muestra las funciones de pertenencia para *frío o moderado*.

las ideas clave. Comenzaremos con la lógica borrosa que es la base de su desarrollo.

Lógica borrosa

La lógica booleana ordinaria trata con cantidades que son verdadero o falso. La lógica borrosa es un intento de desarrollar un método para el razonamiento lógico que es menos estricto. Esto se consigue introduciendo variables lingüísticas y asociándolas con *funciones de pertenencia* que toman valores entre 0 y 1. En el control borroso las operaciones lógicas *y*, *o* y *no* son operaciones sobre variables lingüísticas. Estas operaciones se pueden expresar en términos de operaciones sobre las funciones de pertenencia de las variables lingüísticas. Considerense dos variables lingüísticas con las funciones de pertenencia $f_A(x)$ y $f_B(x)$. Las operaciones lógicas se definen por las siguientes operaciones sobre las funciones de

pertenencia:

$$f_{A \text{ y } B} = \min(f_A(x), f_B(x))$$

$$f_{A \text{ o } B} = \max(f_A(x), f_B(x))$$

$$f_{\text{no } A} = 1 - f_A(x).$$

Una variable lingüística donde la función de pertenencia es cero en cualquier lugar excepto para un valor particular se denomina una variable escueta.

Supóngase, por ejemplo que queremos razonar sobre la temperatura. Para este propósito introducimos las variables lingüísticas *frío*, *moderado* y *caliente*, y las asociamos con las funciones de pertenencia mostradas en la Figura 12.27. En la figura también se muestran las variables de pertenencia *frío y moderado* y *frío o moderado*.

Un controlador borroso

En la Figura 12.28 se muestra un diagrama de bloques de un controlador PD borroso. El error de control, que es una variable continua, se alimenta a un sistema lineal que genera la derivada del error. El error y su derivada se convierten a las denominadas variables lingüísticas en un proceso denominada «fusificación». Este procedimiento convierte las variables continuas a una colección de variables lingüísticas. El número de variables lingüísticas es normalmente pequeño, por ejemplo: negativo grande (*NG*), negativo medio (*NM*), negativo pequeño (*NP*), cero (*C*), positivo pequeño (*PP*), positivo medio (*PM*) y positivo grande (*PG*). Como ejemplo, damos reglas para un controlador PD donde el error y su derivada se caracterizan por tres variables lingüísticas (*N*, *C*, *P*) y la variable de control se caracteriza por cinco variables lingüísticas (*NG*, *NM*, *C*, *PM* y *PG*).

- Regla 1: Si e es *C* y de/dt es *P* entonces u es *C*
- Regla 2: Si e es *N* y de/dt es *C* entonces u es *NM*
- Regla 3: Si e es *N* y de/dt es *N* entonces u es *NG*
- Regla 4: Si e es *C* y de/dt es *P* entonces u es *PM*
- Regla 5: Si e es *C* y de/dt es *C* entonces u es *C*
- Regla 6: Si e es *C* y de/dt es *N* entonces u es *NM*
- Regla 7: Si e es *P* y de/dt es *P* entonces u es *PG*
- Regla 8: Si e es *P* y de/dt es *C* entonces u es *PM*
- Regla 9: Si e es *P* y de/dt es *N* entonces u es *C*}

Estas reglas se pueden expresar también en forma de tabla tal y como se muestra en la Tabla 12.1. Las funciones de pertenencia que representan las variables lingüísticas normalmente se solapan (ver Figura 12.27). Debido a esto, existen varias reglas que contribuyen a la señal de control. Las variables lingüísticas que representan la señal de control se calculan como una suma ponderada de las variables lingüísticas de la señal de control. Entonces, la variable lingüística que representa la señal de control se transforma a un número real mediante una operación de «desfusificación». A continuación se dan más detalles.

Inferencia borrosa Se pueden utilizar muchas formas diferentes de funciones de pertenencia. En el control borroso es práctica común utilizar formas triangulares que se solapan tanto para las variables de entradas como de control tal co-

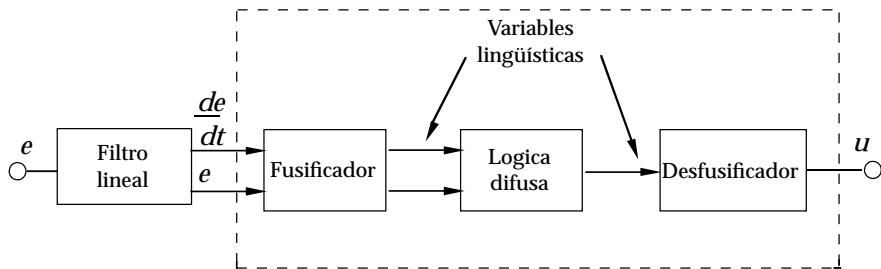


Figura 12.28 Un controlador PD borroso.

Cuadro 12.1 Representación de un control PD difuso en forma de tabla.

		$\frac{de}{dt}$			
		P	C	N	
e		N	C	NM	NG
		C	PM	C	NM
P		PG	PM	C	

mo se muestran en la Figura 12.27. Normalmente, sólo se utilizan unas pocas funciones de pertenencia para las variables de medida.

La lógica borrosa se utiliza sólo en una extensión moderada en el control borroso. Un tema clave es interpretar las expresiones lógicas del tipo que aparece en la descripción del controlador borroso. En control borroso se utilizan algunos métodos especiales. Para describirlos, asumimos que f_A , f_B y f_C son funciones de pertenencia asociadas con las variables lingüísticas A , B y C , y que x e y representan las medidas. Si se miden los valores x_0 e y_0 , se pueden considerar como valores escuetos. La sentencia borrosa

$$\text{Si } x \text{ es } A \text{ y } y \text{ es } B$$

se interpreta entonces como la variable escueta

$$z^0 = \min(f_A(x_0), f_B(y_0))$$

donde y es equivalente a la minimización de funciones de pertenencia. La variable lingüística u definida como

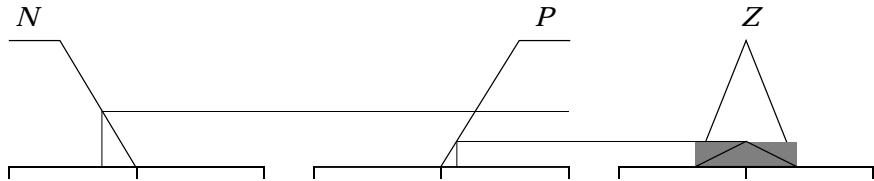
$$\text{Si } x \text{ es } A \text{ o } y \text{ es } B \text{ entonces } u \text{ es } C$$

se interpreta como una variable lingüística con la función de pertenencia

$$f_u(x) = z^0 f_C(x).$$

Si hay algunas reglas, como en la descripción del controlador PD, cada regla se evalúa individualmente. Los resultados obtenidos para cada regla se combinan

Regla 1: Si e es N y de/dt es P entonces u es Z



Regla 2: Si e es N y de/dt es Z entonces u es NM

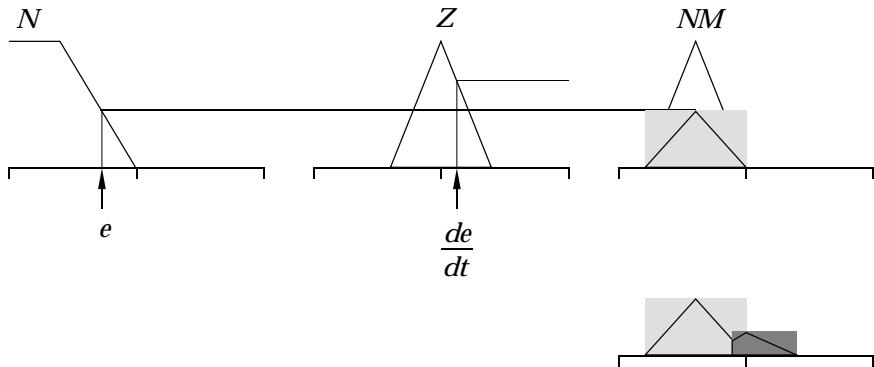


Figura 12.29 Ilustración de la inferencia borrosa con dos reglas utilizando la regla mínimo-máximo.

utilizando el operador \circ . Esto corresponde a tomar el máximo de las funciones de pertenencia obtenidas para cada regla individual.

La Figura 12.29 es una ilustración gráfica para el caso de las dos primeras reglas del controlador PD. La figura muestra cómo se construye la variable lingüística correspondiente a cada regla y cómo se obtiene la señal de control tomando el máximo de las funciones de pertenencia obtenidas a partir de todas las reglas.

El procedimiento de inferencia descrito se denomina «producto-máximo». Esto se refiere a las operaciones sobre las funciones de pertenencia. En control borroso se utilizan también otros procedimientos de inferencia. La operación \circ se representa algunas veces tomando el producto de dos funciones de pertenencia y el operador \circ tomando una suma saturada. También se utilizan combinaciones de los esquemas, siendo posible obtener inferencias «producto-máximo» y «mínimo-suma».

Desfusificación La inferencia borrosa resulta en una variable de control expresada como una variable lingüística y definida por su función de pertenencia. Para aplicar una señal de control debemos tener una variable real. Por tanto, la variable lingüística que define la señal de control se debe convertir a un número real a través de la operación de «desfusificación». Esto se puede hacer de varias formas. Considérese una variable lingüística A con la función de pertenencia $f_A(x)$.

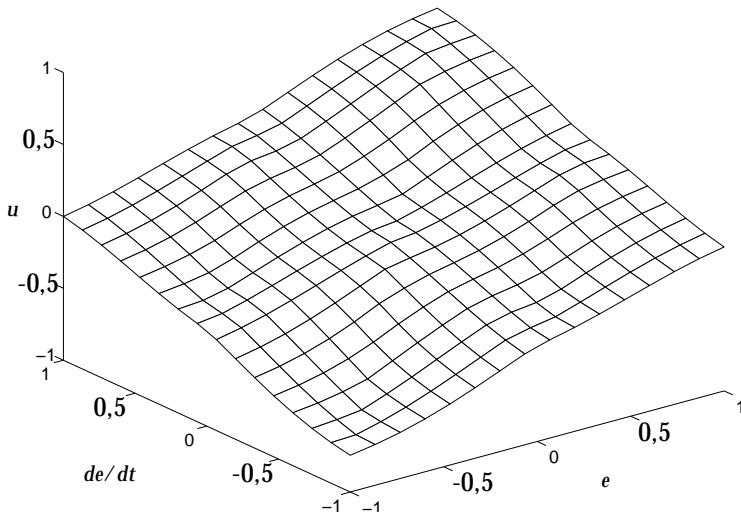


Figura 12.30 Ilustración gráfica de la no linealidad del controlador borroso mostrando la señal de control u como función del error de control e y su derivada.

La desfusificación por valores medios da el valor

$$x_0 = \frac{\int x f_A(x) dx}{\int f_A(x) dx}.$$

La desfusificación por el centroide da una variable real x_0 que satisface

$$\int_{-\infty}^{x_0} f_A(x) dx = \int_{x_0}^{\infty} f_A(x) dx.$$

Control no lineal

Tras haber entrado en detalle, retomamos ahora al controlador PD difuso de la Figura 12.28. Primero observamos que las operaciones de fusificación, lógica borrosa, y desfusificación se pueden describir de una forma muy sencilla. Dejando atrás el vocabulario y considerando el resultado final, un controlador borroso no es nada más que un controlador no lineal. El sistema de la Figura 12.28 se puede expresar de hecho como

$$u = F\left(e, \frac{de}{dt}\right),$$

donde F es una función no lineal de dos variables. Por lo tanto, el controlador borroso PD es un controlador donde la salida es una función no lineal del error e y su derivada de/dt . En la Figura 12.30, damos una ilustración gráfica de la no linealidad definida por reglas dadas para el controlador PD con funciones de pertenencia triangulares estándar y una fusificación del producto. La figura muestra que la función es casi lineal. En este caso particular, el controlador difuso se comportará de forma similar a un controlador lineal PD ordinario.

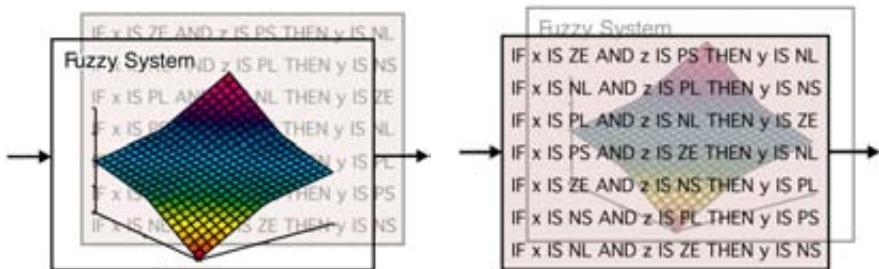


Figura 12.31 Dos vistas de un controlador borroso. La figura de la izquierda muestra que el controlador borroso se puede ver como un controlador no-lineal. Por otro lado, la figura del lado derecho enfatiza las reglas.

El control borroso se puede considerar como una forma de representar una función no lineal, tal y como se muestra en la Figura 12.31. Obsérvese que todavía es necesario tratar con la generación de derivadas o integrales, windup de la integral y todas las otras materias de la misma forma que para un controlador PID ordinario. Podemos también preguntarnos sobre cuándo es útil introducir las no linealidades y la forma que deberían tener.

La representación de una no linealidad por fusificación, lógica borrosa y desfusificación no es muy diferente de la representación de una función no lineal como una tabla con un procedimiento de interpolación. De manera aproximada podemos decir que los valores de la función corresponden a las reglas y las funciones de pertenencia y los procedimientos de fusificación y desfusificación corresponden al mecanismo de interpolación. Para ilustrar esto consideraremos una función de dos variables. Tal función se puede visualizar como una superficie en dos dimensiones. Una función lineal es simplemente un plano inclinado y esta función se puede describir completamente por tres puntos sobre el plano, por ejemplo, con tres reglas. Se pueden obtener superficies o funciones más complejas utilizando más valores de la función. La suavidad de la superficie se expresa mediante los procedimientos de interpolación.

Desde un punto de vista de control, la cuestión clave es comprender cuándo son útiles las no linealidades y qué forma deberían tener. Estas son materias donde queda todavía mucha investigación por hacer. Existen casos donde las no linealidades pueden ser beneficiosas pero también podemos encontrar otros casos donde pueden causar problemas. No es una tarea trivial estudiar lo que sucede y unas pocas simulaciones del comportamiento no son suficientes porque la respuesta de un sistema no lineal es fuertemente dependiente de la amplitud.

También apuntamos que las propiedades del controlador de la Figura 12.28 están fuertemente influenciadas por el filtro lineal utilizado. Por lo tanto, es necesario limitar la ganancia de alta frecuencia de la aproximación de la derivada. También es útil tomar derivadas de la salida del proceso en lugar del error, tal y como se discutió en la Sección 3.3. Se pueden utilizar otros filtros, por ejemplo, añadiendo un integrador a la salida del sistema de la Figura 12.28 obtenemos un controlador PI borroso.

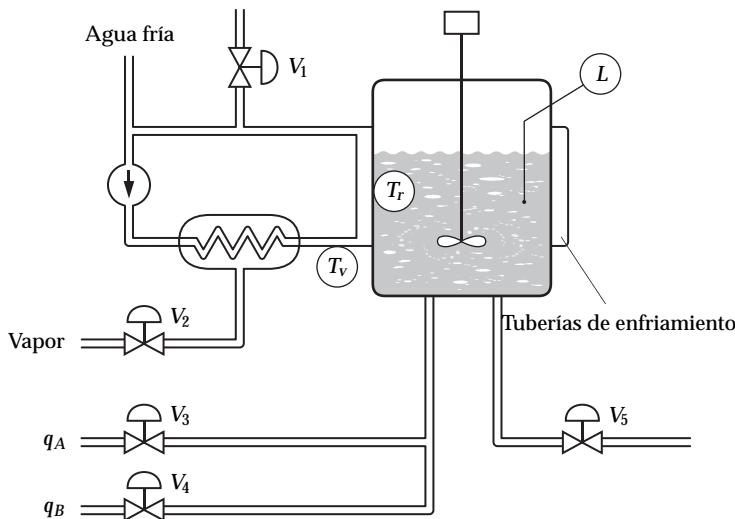


Figura 12.32 Diagrama esquemático de un reactor químico.

Aplicaciones

La representación de la ley de control como una colección de reglas para variables lingüísticas tiene una fuerte atractivo. Es intuitivo explicar heurísticamente cómo funciona el sistema de control. Esto es útil para comunicar estrategias de control a personas con poco entrenamiento formal. Esta es una razón por la que el control borroso es una buena herramienta para la automatización de tareas que se realizan normalmente por los humanos. En este enfoque, se intenta modelar el comportamiento por un operador en términos de reglas lingüísticas. El control borroso se ha utilizado en un número de tareas de control sencillas para electrodomésticos. También se ha utilizado en controladores para procesos que son complicados y mal conocidos. Un ejemplo de este tipo de aplicación es el control de un horno de cemento. Por otro lado, el control borroso se ha utilizado también para la sintonía de controladores.

12.9 Estructurando el sistema

En esta sección ilustramos cómo se pueden construir sistemas complejos a partir de componentes simples utilizando los paradigmas que hemos visto. El problema es bastante complejo. Envuelve la selección de variables de medida, variables de control y requiere una comprensión física significativa del proceso.

El proceso

El proceso a considerar es un reactor químico. La Figura 12.32 muestra un diagrama esquemático. En el reactor se mezclan dos sustancias *A* y *B*, las cuales reaccionan para formar un producto. La reacción es exotérmica, lo que quiere decir que generará calor. El calor se disipa a través de agua que está circulando en las tuberías de refrigeración del reactor. La reacción es muy rápida, de tal forma que el

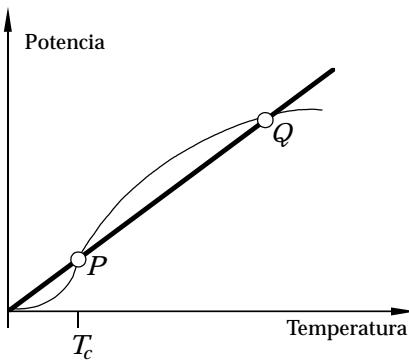


Figura 12.33 Modelo estático del proceso para el reactor exotérmico.

equilibrio se alcanza después de un tiempo que es mucho más corto que el tiempo de residencia del reactor. El caudal q_A de la sustancia A es considerablemente más grande que q_B . La eficiencia de la reacción es esencialmente proporcional al caudal q_B .

Un modelo estático del proceso es útil para comprender el problema de control. La Figura 12.33 muestra la eficiencia y la generación de calor como una función de la temperatura. En la Sección 2.5 se obtuvo un modelo de este tipo. En la figura hemos dibujando una línea recta que corresponde a la potencia de enfriamiento. Existe un equilibrio donde la potencia generada por la reacción es igual a la potencia de enfriamiento representada por los puntos P y Q de la figura. El punto P corresponde a un equilibrio inestable. De la Figura 12.33 se deduce que si la temperatura aumenta por encima de P, la potencia generada por la reacción es mayor que la potencia de enfriamiento, y por lo tanto la temperatura aumentará. El catalizador en el reactor se puede dañar si la temperatura llega a ser demasiado alta. Similarmente, si la temperatura decrece por debajo del punto P, continuará decreciendo y la reacción se parará. Este fenómeno se denomina «congelación». La congelación comienza en la superficie de la tubería de enfriamiento y se expande rápidamente a través del reactor. Si esto ocurre, el reactor debe desconectarse y comenzar de nuevo.

Requisitos de diseño

En la operación de un reactor exotérmico hay riesgos considerables. El reactor puede explotar si la temperatura es demasiado alta. Para reducir el riesgo de explosión, los reactores se sitúan en edificios especiales lejos del operador. Debido al riesgo de explosión, no es factible experimentar con la sintonía del controlador. Consecuentemente, es necesario calcular la sintonía del controlador de antemano y verificar que los ajustes son correctos antes de arrancar el reactor. La seguridad es el requisito primordial del sistema de control. Es importante garantizar que la temperatura de la reacción no será demasiado alta. También es importante asegurar que las perturbaciones del proceso no llevan a una pérdida del caudal de refrigerante y que la agitación no conduce a una explosión. Por otro lado, también se desea operar el reactor eficientemente. Esto quiere decir que se debe evitar la congelación. Además, se desea mantener la eficiencia tan alta como sea posible.

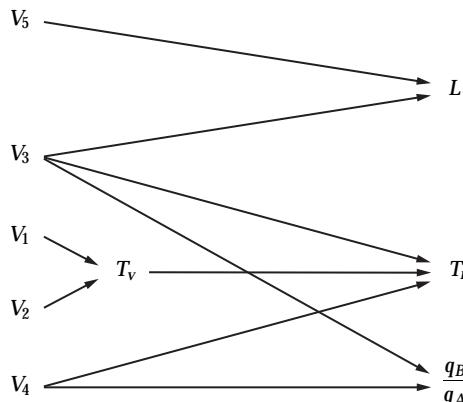


Figura 12.34 Diagrama de causalidad para la variable del proceso.

Debido a los riesgos, es también necesario automatizar el arranque y la parada así como la operación normal. Es deseable evitar el tener que operar el reactor bajo control manual. En este caso particular, el operador puede fijar dos variables: la temperatura del reactor y la razón entre los caudales q_A y q_B . Estas dos variables pueden influir sobre la eficiencia de la reacción y la calidad del producto.

Estructura del controlador

El reactor tiene cinco válvulas. Dos de ellas, V_1 y V_2 , influyen sobre la temperatura del refrigerante. El caudal del reactor se controla con V_3 y V_4 , y el caudal del producto con la válvula V_5 . En esta aplicación particular, la válvula V_5 se controla por escalones corriente abajo del proceso. (Comparar esto con el análisis de los tanques de carga en la Sección 12.6).

Hay cinco señales medidas: la temperatura del reactor T_r , el nivel en el tanque del reactor L , la temperatura de enfriamiento T_v y los caudales q_A y q_B . Las propiedades físicas del proceso dan un estructuración natural del sistema de control. Un balance de masa del material en el reactor se muestra que el nivel se ve influenciado esencialmente por el caudal q_A y la producción demandada. De la estequiometría de la reacción se deduce que la razón de los caudales q_A y q_B se debería mantener constante para obtener una reacción eficiente. La temperatura del reactor está fuertemente afectada por la temperatura del agua, por la temperatura del caudal de refrigerante y los caudales q_A y q_B . La temperatura del refrigerante se ve influenciada por la válvula V_1 que controla la cantidad de caudal y mediante la válvula de vapor V_2 .

Este simple análisis físico lleva al diagrama que se muestra en la Figura 12.34, que ilustra la causalidad de las variables en el proceso. La válvula V_5 se puede considerar como una perturbación ya que se establece por unidades del proceso corriente abajo. La figura 12.34 sugiere que hay tres lazos de control naturales:

- Control de nivel: controlar nivel del tanque con la válvula V_3 .
- Control de temperatura: control de la temperatura del reactor con las válvulas V_1 y V_2 .

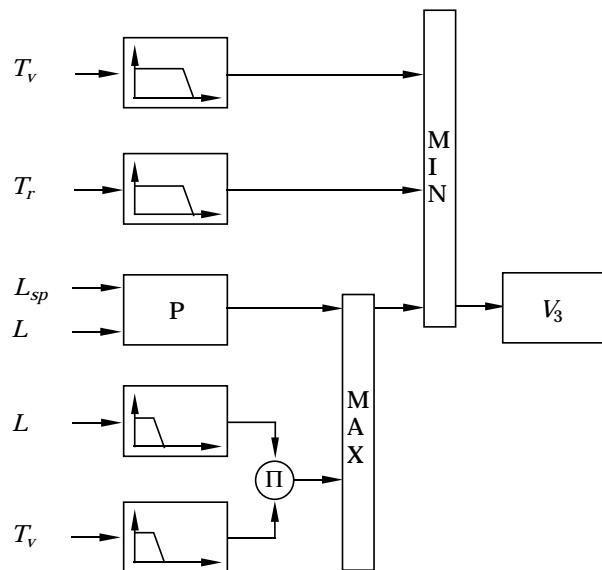


Figura 12.35 Diagrama de bloques para el control de nivel a través de la válvula V_3 .

- Control de la relación de caudal: control de la relación q_B/q_A con la válvula V_4 .

Estos lazos de control se describen en detalle a continuación.

Control de nivel

La Figura 12.35 muestra el diagrama de bloques para el control de nivel. La función primaria es una realimentación proporcional desde el nivel al caudal q_A , que se controla por la válvula V_3 . El reactor se utiliza también como un tanque de carga para aislar la diferencia entre la producción actual y la producción demandada. El nivel en el tanque variará durante las operaciones normales. Unos límites razonables para el nivel del tanque son entre el 50 por ciento y el 100 por ciento. Si se escoge la banda proporcional del controlador como 50 por ciento, la variable de control estará completamente cerrada cuando el tanque esté lleno y medio abierta cuando el tanque esté medio lleno. Es importante que la temperatura del reactor permanezca dentro de los límites dados. Por lo tanto, se restringe el caudal q_A por dos selectores basándose en las medidas de la temperatura en el tanque del reactor (T_r) y la temperatura del refrigerante (T_v). Cuando comienza el reactor, el nivel se mantiene en el límite inferior hasta que la temperatura del refrigerante llega a ser suficientemente alta. Esto se consigue mediante la combinación de limitadores, multiplicadores y selectores, tal y como se muestra en la Figura 12.35.

Control de temperatura

La Figura 12.36 da un diagrama de bloques para el control de la temperatura del reactor. Como la reacción química es rápida en comparación con las dinámicas de temperatura y caudal, el reactor se puede ver desde un punto de vista de

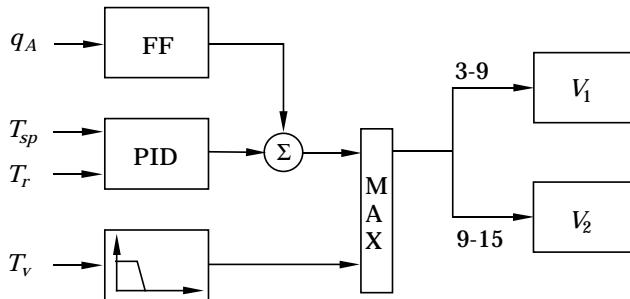


Figura 12.36 Diagrama de bloques del control de temperatura a través de las válvulas V_1 y V_2 .

control como un intercambiador de calor. Durante condiciones normales, la temperatura se controla ajustando el caudal del refrigerante a través de la válvula V_1 . La función de control primaria es una realimentación desde la temperatura a las válvulas V_1 y V_2 . El punto de consigna en este lazo de control se puede ajustar manualmente. Los parámetros del lazo de control se pueden determinar como sigue. La función de transferencia que relaciona el flujo del refrigerante con la temperatura del reactor está aproximadamente dada por

$$G(s) = \frac{K_p}{(1+sT_1)(1+sT_2)}, \quad (12.8)$$

donde las constantes de tiempo tienen los valores $T_1 = 300$ s y $T_2 = 50$ s. El siguiente cálculo proporciona unos valores aproximados del parámetro del controlador. Un controlador proporcional con ganancia K da la función de transferencia del lazo

$$G_0(s) = \frac{KK_p}{(1+sT_1)(1+sT_2)}. \quad (12.9)$$

La ecuación característica del sistema en lazo cerrado es

$$s^2 + s \left(\frac{1}{T_1} + \frac{1}{T_2} \right) + \frac{1+KK_p}{T_1 T_2} = 0.$$

El sistema en lazo cerrado es así sistema de segundo orden. El amortiguamiento relativo ζ y la frecuencia natural no amortiguada ω vienen dadas por

$$2\zeta\omega = \frac{1}{T_1} + \frac{1}{T_2} \approx \frac{1}{T_2} \quad (12.10)$$

y

$$2\zeta\omega^2 = \frac{1+KK_p}{T_1 T_2}. \quad (12.11)$$

La aproximación en la primera expresión viene motivada porque $T_1 \gg T_2$. Con un amortiguamiento relativo de $\zeta = 0,5$, la ecuación (12.10) resulta en $\omega \approx 1/T_2$. Más aún, se sigue que la ecuación (12.11) que

$$1+KK_p = \frac{T_1}{T_2} = \frac{300}{50} = 6.$$

La ganancia del lazo está así esencialmente determinada por la razón de las constantes de tiempo. La ganancia del controlador es

$$K = \frac{5}{K_p},$$

y el sistema en lazo cerrado tiene la frecuencia natural no amortiguada de

$$\omega = 1/T_2 = 0,02 \text{ rad/s.}$$

Si en lugar de ello se escoge un controlador PI, es razonable seleccionar un valor del tiempo integral de

$$T_1 \approx 5 T_2.$$

El control se puede mejorar utilizando acción derivativa. La mejora alcanzable depende de la constante de tiempo del sensor de temperatura. En casos típicos, esta constante de tiempo está entre 10 s y 40 s. Si es tan baja como 10 s, es posible obtener un control mejorado introduciendo una acción derivativa en el controlador. El tiempo derivativo se puede escoger para eliminar la constante de tiempo T_2 . Entonces, obtenemos un sistema con las constantes de tiempo 300 s y 10 s. La ganancia se puede entonces aumentar de forma que

$$1 + KK_p = \frac{300}{10} = 30$$

y la frecuencia natural no amortiguada del sistema viene dada por $\omega \approx 0,1 \text{ rad/s}$. Si la constante de tiempo del sensor de temperatura está entorno a 40 s, la acción derivativa sólo da mejoras marginales.

El calor generado por la reacción química es proporcional al caudal q_A . Para asegurar que la variaciones en q_A se compensan rápidamente, tenemos que introducir una acción feedforward desde el caudal q_A . Este feedforward operará solamente cuando el nivel del tanque sea mayor que el 50 por ciento para evitar la congelación cuando se arranque el reactor.

Para comentar la reacción, el reactor se debe calentar de tal forma que la temperatura en el recipiente de la reacción sea mayor que T_c (comparar con la Figura 12.33). Esto se realiza utilizando la válvula de vapor V_2 . Se utiliza un control de rango partido para las válvulas de vapor y agua (comparar con la Sección 12.6). La válvula de agua se abre para señales bajas (3–9 PSI) y la de vapor para altas presiones (9–15 PSI).

Para evitar que el reactor se congele, es necesario asegurar que la temperatura de la reacción es siempre superior a T_c . Esta es la razón por la que se utiliza una realimentación extra mediante un selector máximo desde la temperatura del agua a T_v . Esta realimentación asegura que la válvula de vapor se abre si la temperatura en el caudal del refrigerante es demasiado baja. Un control en cascada sería una alternativa a esta solución.

Control de relación del caudal

La razón de los caudales q_A y q_B se debe mantener constante. La Figura 12.37 muestra cómo la eficiencia de la reacción depende de q_B cuando q_A se mantiene

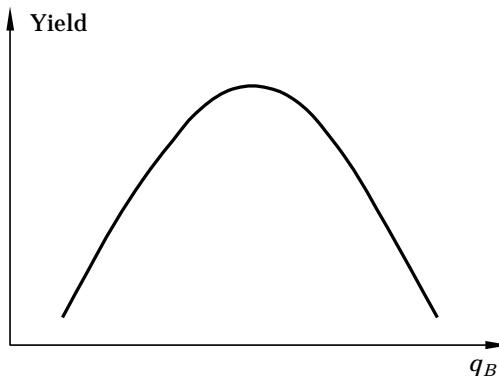


Figura 12.37 Producto de la reacción como función de q_B con q_A constante.

constante. El caudal q_B se controla con un sistema de control de relación (como el que se muestra en la Figura 12.38), que es la función de control primaria. La velocidad de reacción depende fuertemente de q_B . Para disminuir el riesgo de explosión, existe una no linealidad en la realimentación que aumenta la ganancia cuando q_B/q_A es grande. El lazo de caudal tiene varios selectores. En el arranque se desea que la sustancia B no se añada hasta que la temperatura del agua haya alcanzado el valor crítico T_c y el tanque del reactor esté medio lleno. Para alcanzar esto, se ha introducido una realimentación desde la temperatura del agua al nivel del tanque mediante limitadores y un selector de mínimo. Hay también limitadores y un selector que cierra la válvula V_4 si se pierde el caudal q_A . También existe una realimentación directa desde q_A a través de limitadores y selectores y una realimentación desde la temperatura del reactor que cierra la válvula V_4 si la temperatura del reactor es demasiado alta.

Control de inhibición de la válvula de salida

El caudal de salida del reactor se determina por la válvula V_5 . Esta válvula se controla normalmente por escalones del proceso corriente abajo. El control del reactor se puede mejorar introduciendo una orden que depende del estado del reactor. Cuando comienza el reactor, se desea tener la válvula de salida cerrada hasta que el tanque del reactor esté medio lleno y la reacción ha comenzado. Esto se consigue introduciendo el nivel y la temperatura del tanque en el punto de consigna del controlador de la válvula vía limitadores y selectores de mínimos tal y como se muestra en la Figura 12.39. La válvula V_5 se controla normalmente por q_{sp} . El selector de mínimo invalida la orden q_{sp} cuando el nivel L o la temperatura T_r son demasiado bajos.

12.10 Resumen

En este capítulo, hemos ilustrado cómo se pueden construir sistemas de control complejos a partir de componentes simples tales como controladores PID, filtros

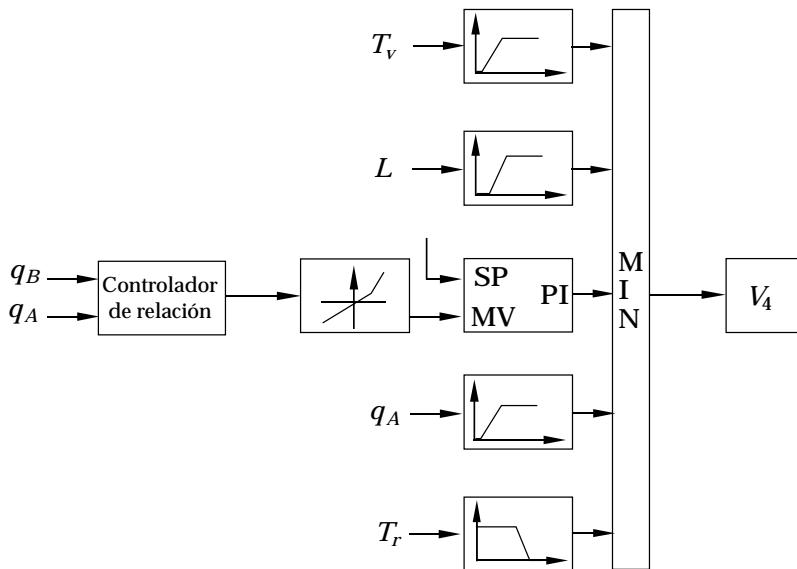


Figura 12.38 Diagrama de bloques para controlar la razón de mezcla q_B/q_A a través de la válvula V_4 .

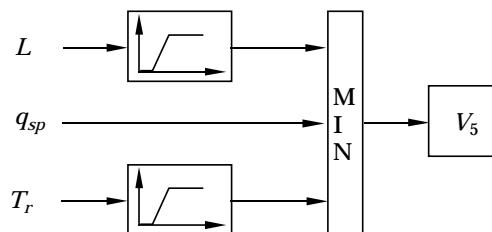


Figura 12.39 Diagrama de bloques para controlar el caudal de salida del reactor a través de la válvula V_5 .

lineales, planificaciones de ganancia y funciones no lineales sencillas. Se han introducido un número de paradigmas de control para guiar el diseño del sistema.

Los paradigmas básicos de control lineal son la realimentación mediante control PID y la acción feedforward. El control en cascada se puede utilizar para aumentar el rendimiento de control a través del uso de medidas adicionales. La realimentación del estado se puede ver como un caso extremo del control en cascada donde todos se miden todos los estados del sistema. Los observadores se pueden utilizar para inferir valores de las variables que no son medidas mediante la combinación de modelos matemáticos con las medidas disponibles. Los controles de rango medio y de rango partido son paradigmas de control donde hay varias señales de control pero sólo una señal medida. Estos paradigmas son el dual de un control en cascada. El control repetitivo es una técnica que es eficiente para casos donde las perturbaciones son periódicas. La idea es crear una ganancia de lazo alta a la frecuencia de la perturbación.

También hemos tratado varios componentes no lineales y paradigmas rela-

cionados incluyendo funciones no lineales, planificadores de ganancia, limitadores y selectores. Recuerde que en la Sección 3.5 se mostró cómo podrían mejorarse los controladores PID mediante funciones no lineales simples para evitar el windup. El control de relación es una estrategia no lineal que admite control de dos variables del proceso de forma que su razón sea constante. En la Sección 9.3, presentamos cómo los planificadores de ganancia se podrían utilizar para tratar con los cambios en la dinámica del proceso. Los planificadores de ganancia y las funciones no lineales son también útiles para el control de tanques de almacenamiento, donde el objetivo no es mantener constantes sus niveles sino permitir que varíen dentro de unos rangos establecidos. El control selector es otro paradigma importante que se utilizó para el control con restricciones donde determinadas variables de proceso tienen que cumplir ciertas restricciones establecidas. También se comentaron brevemente las técnicas neuronales y borrosas, donde se mostró que se podrían interpretar como un control basado en reglas y un control no lineal.

También se dió un ejemplo de cómo se podrían usarlos componentes y los paradigmas para desarrollar un sistema de control de un proceso químico.

12.11 Notas y referencias

Muchos de los aspectos del material de este capítulo se encuentran en libros de texto clásicos sobre control de procesos tales como [69, 309, 48, 302] y en los libros [308, 198] que se centran en sistemas de energía. En [141] se da una presentación más especializada.

Los métodos que se describen en este capítulo se pueden caracterizar como procedimientos ascendentes en el sentido de que un sistema complejo se puede construir mediante la combinación de sistemas simples. En [63] se da una visión interesante sobre esto. Otra posibilidad es un enfoque descendente. En [303] y [247] se puede encontrar un estudio sobre estos temas que queda fuera del objetivo de este libro.

Los controles en cascada y feedforward se tratan en libros de texto estándar sobre control. En [343] se da una presentación con muchos aspectos prácticos. El control selector se utiliza ampliamente en la práctica, y una descripción general se puede encontrar en [35]. Es difícil analizar sistemas no lineales. En [110] se da un análisis de estabilidad de un sistema con selectores. En [145] se presenta la estación de mezcla.

El control borroso ha estado presente durante un largo tiempo; véase [227, 226, 197, 340], y ha recibido mucha atención particularmente en Japón: ver [373, 341, 330, 101, 355]. La técnica se ha utilizado para la automatización de procesos complicados que previamente se controlaban de forma manual. Un ejemplo típico es el control de hornos de cemento; ver [171]. Han existido desarrollos similares utilizando redes neuronales; ver, por ejemplo, [166, 269, 24]. Hacia finales de los 60 hubo mucha actividad con las redes neuronales, pero desapareció rápidamente. Hubo un rápido resurgir sobre su interés en los ochenta. Sin embargo, hay muchas exageraciones en las técnicas borrosas y neuronales y no ha emergido todavía una visión de la relevancia de estos campos en control. El artículo [360] presentó un resumen de los usos posibles de las redes neuronales en el control de

procesos y el trabajo [284] describe una aplicación de control de pH. Los trabajos [211, 179, 332] describen aplicaciones con controladores PID y sobre su sintonía. También existen intentos de fusionar el control neuronal y difuso; ver [67]. La Sección 12.9 se basa en [70].

13

Implementación

13.1 Introducción

Los controladores PID se implementaron originalmente utilizando técnicas analógicas. Los primeros sistemas utilizaron relés neumáticos, fuelles y opresiones de válvulas de aguja. Los motores eléctricos con relé y circuitos de realimentación y los amplificadores operacionales se utilizaron más tarde. Muchas de las características como anti windup y derivación de la salida del proceso en lugar del error del control se incorporaron como «trucos» en estas implementaciones.

Ahora es práctica común implementar controladores PID utilizando microprocesadores y se han redescubierto algunos de los viejos trucos. Algunos controladores PID que se utilizan hoy día tienen características que son heredadas de técnicas antiguas cuando los controladores se implementaban utilizando dispositivos neumáticos. Este es un ejemplo típico del hecho de que las ideas algunas veces cambian a una velocidad mucho más lenta que los dispositivos físicos. Se pueden considerar algunos temas adicionales en conexión con las implementaciones digitales. Las más importantes están relacionadas con el muestreo, la discretización y la cuantificación.

Este capítulo presenta algunas cuestiones de implementación relacionadas con el control PID. La Sección 13.2 muestra un breve resumen de las primeras implementaciones neumáticas, analógicas y electrónicas. La Sección 13.3 trata con aspectos de implementación en computadores tales como muestreo, prefiltrado y discretización del algoritmo PID. En las aplicaciones donde la integración se realiza fuera del controlador, se necesitan algoritmos de velocidad o algoritmos incrementales. Los motores eléctricos son la aplicación más común. En la Sección 13.4 se presentan estos algoritmos, que son útiles incluso cuando la integración se realiza dentro del controlador. En la Sección 13.5 se tratan aspectos operacionales, tales como transferencias sin salto en la conmutación de modos y cambios de parámetros. Un controlador puede tener diferentes salidas dependiendo del dispositivo de actuación que se utilice. En la Sección 13.6 se tratan las salidas del controlador. El capítulo finaliza con un resumen y referencias bibliográficas.

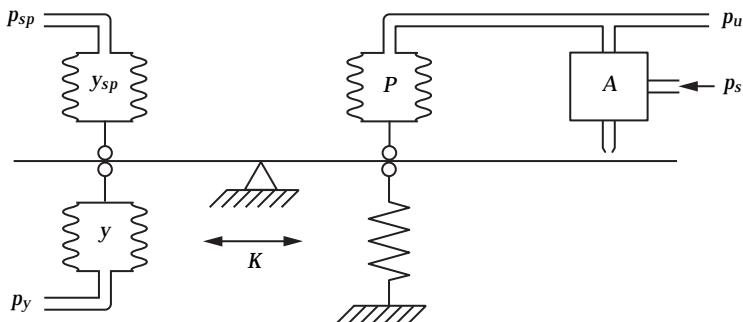


Figura 13.1 Diagrama esquemático de un controlador P neumático basado en el principio de balance de fuerzas.

13.2 Implementaciones analógicas

Las primeras implementaciones de los controladores PID fueron todas analógicas. Esta sección presenta la implementación del controlador neumático y la implementación electrónica-analógica.

El controlador neumático

Esta sección presenta la función básica de los controladores neumáticos. Se han eliminado muchos detalles de los dibujos y de la presentación por temas de claridad. Se pueden encontrar más detalles en las referencias bibliográficas comentadas al final del capítulo.

El controlador P neumático En la Figura 13.1 se muestra un diagrama esquemático de un controlador P neumático basado en el balance de fuerzas. El sistema consiste en una viga que puede rotar alrededor de un punto. La viga se proporciona con tres fuelles, un muelle, un sensor de posición y un amplificador neumático. Los fuelles pueden ejercer fuerzas sobre la viga proporcionales a la presión en los fuelles. El sensor de posición es una válvula de clapeta, que proporciona una señal de presión que es, de manera aproximada, inversamente proporcional a la distancia entre la boquilla y la viga. El amplificador neumático A puede amplificar las señales neumáticas.

Para comprender la operación del sistema se supone que las fuerzas de todos los fuelles son proporcionales a la presión del aire en dichos fuelles. Los dos fuelles de la izquierda reciben las presiones p_{sp} y p_y proporcionales al punto de consigna y la variable medida, respectivamente. El amplificador de presión A recibe la presión de suministro p_s y proporciona la presión de salida p_u , que es la salida del controlador. El fuelle de la derecha, etiquetado como P, es el fuelle de realimentación o fuelle proporcional. En el controlador P, la presión en este fuelle, p_p , es igual a la presión de salida p_u .

Un balance de pares da la siguiente relación entre las presiones:

$$p_u - bias = K(p_{sp} - p_y). \quad (13.1)$$

El término *bias* es la fuerza proporcionada por el muelle. La ganancia K se determina por la posición del punto de balance y puede por lo tanto escogerse ajustando el punto de balance.

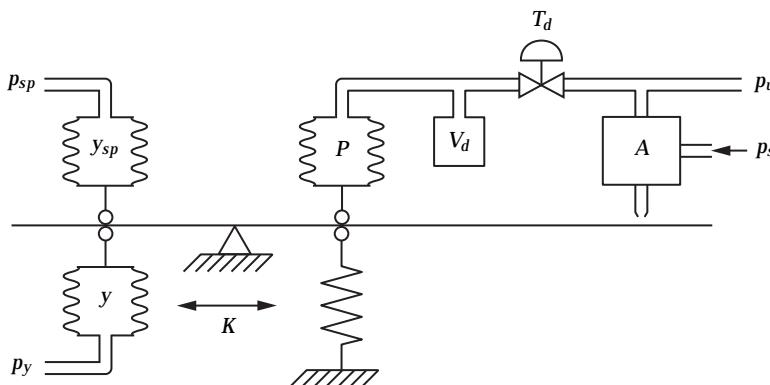


Figura 13.2 Diagrama esquemático de un controlador PD neumático basado en el principio de balance de fuerzas.

tando dicho punto. La ecuación 13.1 es obviamente la ecuación de un controlador P.

Supóngase, por ejemplo, que aumenta la presión del punto de consigna p_{sp} . Entonces, la viga rotará en la dirección positiva conduciendo a un descenso en el caudal de salida de la válvula de la boquilla. Esto causará un incremento de la presión de salida p_u .

El controlador PD neumático En la Figura 13.2 se muestra un controlador PD neumático. En este controlador, se introducen una válvula y un volumen V_d entre el amplificador A y el fuelle de realimentación P. Debido a esta válvula, ya no es cierto que $p_p = p_u$, pero se mantiene la siguiente relación dinámica entre las dos presiones:

$$P_p(s) = \frac{1}{1 + sT_d} P_u(s). \quad (13.2)$$

El valor de la constante de tiempo T_d se puede ajustar por la posición de la válvula.

Como una neutralización originada por el fuelle de realimentación P se retrasa en comparación con el controlador P, un cambio en p_y o p_{sp} resultará inicialmente en un mayor reacción en la presión de salida p_u .

Un balance de pares resulta en las siguientes relaciones entre las presiones:

$$p_p - bias = K(p_{sp} - p_y).$$

De (13.2), esto da en la siguiente presión de salida;

$$P_u(s) = bias + K(1 + sT_d)(P_{sp} - P_y),$$

que es la ecuación de un controlador PD con tiempo derivativo T_d .

El controlador PID neumático En la Figura 13.3 se muestra un controlador PID neumático. En este controlador, el muelle se reemplaza por un fuelle etiquetado

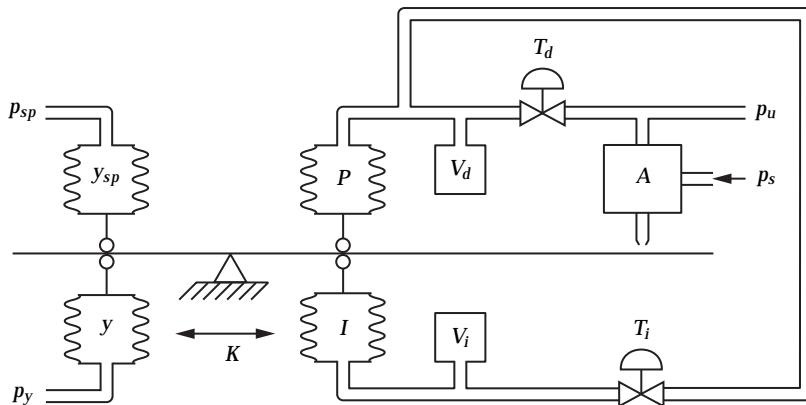


Figura 13.3 Diagrama esquemático de un controlador PID neumático basado en el principio de balance de fuerzas.

como I . Este fuelle se conecta a la presión p_p a través de un volumen V_i y una válvula etiquetada como T_i . La presión en el fuelle I es

$$P_i(s) = \frac{1}{1 + sT_i} P_p(s), \quad (13.3)$$

donde la constante de tiempo T_i se puede ajustar mediante la válvula T_i .

Un balance de pares da las siguientes relaciones entre las presiones:

$$p_p - p_i = K(p_{sp} - p_y).$$

De (13.2) y (13.3), se obtiene la siguiente presión de salida:

$$P_u(s) = K \frac{(1 + sT_i)(1 + sT_d)}{sT_i} (P_{sp}(s) - P_y(s)). \quad (13.4)$$

Esta ecuación muestra que el sistema es un controlador PID en la forma interactuante (ver la Sección 3.2) con ganancia K , tiempo integral T_i y tiempo derivativo T_d .

La idea de utilizar realimentación en el controlador fue una invención muy importante. Tanto la válvula de clapeta como el amplificador neumático son fuertemente no lineales. La disposición con el lazo de realimentación implica que la relación entrada-salida del controlador no cambia mucho incluso si el componente cambia, a condición de que la ganancia sea suficientemente grande. Esta idea, que se denomina realimentación de fuerzas, dió mejoras drásticas en el rendimiento de los controladores. Un ejemplo típico del impacto de la realimentación.

El controlador electrónico analógico

Un controlador PID se puede implementar mediante componentes electrónicos analógicos de muchas formas. Esta sección presenta algunas implementaciones básicas basadas en amplificadores operaciones. Por temas de simplicidad, se han obviado muchos detalles. Como en el caso de los controladores neumáticos, se pueden encontrar más detalles en las referencias bibliográficas.

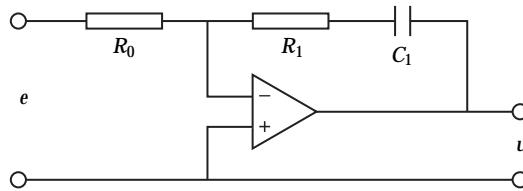


Figura 13.4 Diagrama esquemático de un controlador PI electrónico basado en la realización alrededor de un amplificador operacional.

El controlador PI electrónico En la Figura 13.4 se muestra un controlador PI electrónico.

Una relación aproximada entre las tensiones de entrada e y de salida u viene dada por

$$u = -\frac{Z_1}{Z_0}e,$$

donde Z_0 es la impedancia entre la entrada negativa del amplificador y el voltaje de entrada e y Z_1 es la impedancia entre la entrada cero del amplificador y el voltaje de salida u . Estas impedancias son

$$\begin{aligned} Z_0 &= R_0 \\ Z_1 &= R_1 + \frac{1}{C_1 p}, \end{aligned}$$

donde p es el operador diferencial. Esto da la siguiente relación entre los voltajes de entrada e y de salida u :

$$u = -\frac{Z_1}{Z_0}e = -\frac{R_1}{R_0} \left(1 + \frac{1}{R_1 C_1 p} \right) e.$$

Esto es un controlador PI con parámetros

$$K = \frac{R_1}{R_0} \quad T_i = R_1 C_1.$$

Se obtiene un controlador P eliminando el condensador.

El controlador PD electrónico En la Figura 13.5 se muestra un controlador PD electrónico.

Las impedancias entre la entrada negativa del amplificador y los voltajes de entrada y salida, respectivamente, vienen dadas por

$$\begin{aligned} Z_0 &= \frac{R_0}{1 + R_0 C_0 p} \\ Z_1 &= R_1. \end{aligned}$$

Esto resulta en la siguiente relación entre el voltaje de entrada e y el de salida u :

$$u = -\frac{Z_1}{Z_0}e = -\frac{R_1}{R_0} (1 + R_0 C_0 p) e.$$

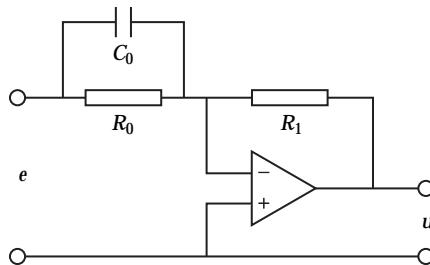


Figura 13.5 Diagrama esquemático de un controlador PD electrónico basado en la realización alrededor de un amplificador operacional.

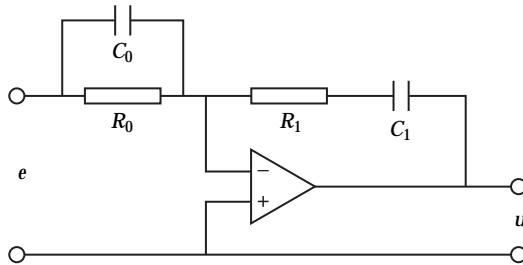


Figura 13.6 Diagrama esquemático de un controlador PID electrónico basado en la realización alrededor de un amplificador operacional.

Esto es un controlador PD con los parámetros

$$K = \frac{R_1}{R_0} \quad T_d = R_0 C_0.$$

Se obtiene un controlador P eliminando el condensador.

El controlador PID electrónico Se puede obtener un controlador PID mediante la combinación de los dos esquemas anteriores tal y como se muestra en la Figura 13.6.

Las impedancias entre la entrada negativa del amplificador y los voltajes de entrada y salida, respectivamente, vienen dadas por

$$Z_0 = \frac{R_0}{1 + R_0 C_0 p}$$

$$Z_1 = R_1 + \frac{1}{C_1 p}.$$

Esto da la siguiente relación entre los voltajes de entrada e y de salida u:

$$u = -\frac{Z_1}{Z_0} e = -\frac{R_1}{R_0} \frac{(1 + R_0 C_0 p)(1 + R_1 C_1 p)}{R_1 C_1 p} e.$$

Esto es un controlador PID en la forma interactuante con parámetros

$$K = \frac{R_1}{R_0} \quad T_i = R_1 C_1 \quad T_d = R_0 C_0.$$

13.3 Implementaciones en computador

Hoy día, la mayoría de los controladores se implementan en computadores. Hay algunos temas que se tienen que tener en cuenta debido al hecho de que las señales se muestran en instantes de tiempo discreto. Estos temas se tratan en esta sección.

Muestreo

Cuando el controlador se implementa en un computador, se leen las entradas analógicas y se establecen las salidas con un cierto período de muestreo. Esto es un inconveniente en comparación con las implementaciones analógicas ya que el muestreo introduce un retardo de tiempo en el lazo de control.

Cuando se utiliza un computador digital para implementar una ley de control, la secuencia ideal de operación es la siguiente:

1. Esperar a la interrupción del reloj.
2. Leer la entrada analógica.
3. Calcular la señal de control.
4. Establecer la salida analógica.
5. Actualizar las variables del controlador.
6. Ir al paso 1.

Con esta implementación se minimiza el retardo. Si la entrada analógica se lee con un período de muestreo h , el retardo medio de la señal medida es $h/2$. El tiempo de cómputo es generalmente corto en comparación con el período de muestreo. Esto quiere decir que el retardo total es alrededor de $h/2$. Sin embargo, la mayoría de los controladores y sistemas de instrumentos no organizan el cálculo de esta manera. Por lo tanto, los retardos introducidos por el muestreo son generalmente algunos períodos de muestreo.

Aliasing

El mecanismo de muestreo introduce algunos fenómenos inesperados que se deben tener en cuenta en una buena implementación digital de un controlador PID. Para explicar éstos, considérense las señales

$$s(t) = \cos(n\omega_s t \pm \omega t)$$

y

$$s_a(t) = \cos(\omega t),$$

donde $\omega_s = 2\pi/h$ [rad/s] es la frecuencia de muestreo. Las fórmulas bien conocidas para la función coseno implican que los valores de la señales en los instantes de muestreo [$kh, k = 0, 1, 2, \dots$] tienen la propiedad

$$s(kh) = \cos(nkh\omega_s \pm \omega kh) = \cos(\omega kh) = s_a(\omega kh).$$

Las señales s y s_a tienen, así, los mismos valores en los instantes de muestreo. Esto quiere decir que no existe forma de separar las señales si sólo se conocen sus

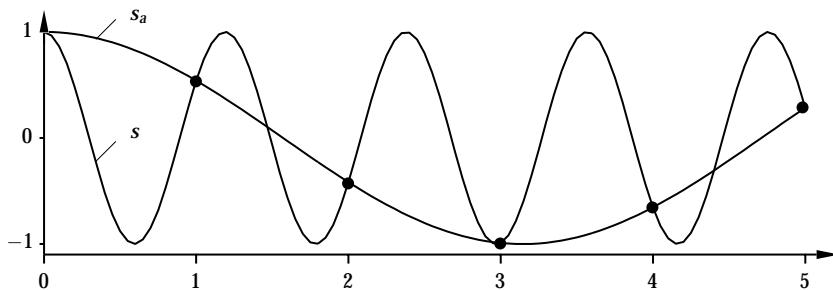


Figura 13.7 Ilustración del efecto de aliasing. El diagrama muestra la señal s y su alias s_a .

valores en los instantes de muestreo. Por tanto, la señal s_a se denomina *alias* de la señal s . Esto se muestra en la Figura 13.7. Una consecuencia del efecto de aliasing es que puede aparecer una perturbación de alta frecuencia después del muestreo como una señal de baja frecuencia. En la Figura 13.7, el período de muestreo es 1 s y la perturbación sinusoidal tiene un período de $6/5$ s. Tras el muestreo, la perturbación aparece como una sinusode con la frecuencia

$$f_a = 1 - \frac{5}{6} = 1/6 \text{ Hz.}$$

En la figura se muestra esta señal de baja frecuencia con un período de tiempo 6 s.

Prefiltrado

Si no se toman las precauciones oportunas, el efecto de aliasing puede crear dificultades significativas. Las altas frecuencias, que en los controladores analógicos normalmente se eliminan efectivamente mediante filtrado paso-bajo, pueden aparecer a causa del aliasing como señales de baja frecuencia en el ancho de banda de un sistema de control muestreado. Para evitar estas dificultades, se debería introducir un prefiltrado analógico (que efectivamente elimina todos los componentes de la señal con frecuencias por encima de la mitad de la frecuencia de muestreo). Estos filtros se denominan filtros anti aliasing. Un filtro de Butterworth de segundo orden es un filtro anti aliasing común. También se utilizan filtros de elevado orden en aplicaciones críticas. En la Figura 13.8 se muestra una implementación de un filtro de esta clase utilizando amplificadores operacionales. La selección del ancho de banda del filtro se ilustra en el siguiente ejemplo.

EJEMPLO 13.1—SELECCIÓN DEL ANCHO DE BANDA DEL FILTRO

Supóngase que se desea que el prefiltrado atenúa las señales por un factor de 16 a la mitad de la frecuencia de muestreo. Si el ancho de banda del filtro es ω_b y la frecuencia de muestreo es ω_s , obtenemos que

$$(\omega_s/2\omega_b)^2 = 16.$$

Por lo tanto,

$$\omega_b = \frac{1}{8} \omega_s.$$

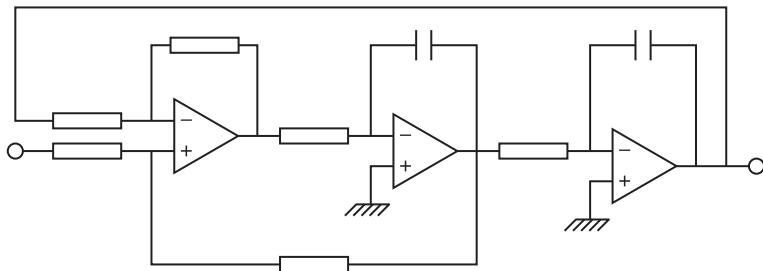


Figura 13.8 Diagrama del circuito de un filtro Butterworth de segundo orden.

□

Obsérvese que la dinámica del prefiltro se combinará con la dinámica del proceso.

Discretización

Para implementar una ley de control en tiempo continuo, tal como un controlador PID en un computador digital, es necesario aproximar las derivadas y la integral que aparecen en la ley de control. A continuación se presentan algunas formas diferentes de hacer esto.

Acción proporcional El término proporcional es

$$P = K(by_{sp} - y).$$

Este término se implementa simplemente reemplazando las variables continuas por sus versiones muestreadas. Por lo tanto,

$$P(t_k) = K(by_{sp}(t_k) - y(t_k)), \quad (13.5)$$

donde $\{t_k\}$ denota los instantes de muestreo, es decir, las veces que el computador lee la entrada analógica.

Acción integral El término integral viene dado por

$$I(t) = \frac{K}{T_i} \int_0^t e(s) ds.$$

Se deduce que

$$\frac{dI}{dt} = \frac{K}{T_i} e. \quad (13.6)$$

Existen varias formas de aproximar esta ecuación. Aproximando la derivada por una diferencia hacia delante resulta en

$$\frac{I(t_{k+1}) - I(t_k)}{h} = \frac{K}{T_i} e(t_k).$$

Esto lleva a la siguiente ecuación recursiva del término integral

$$I(t_{k+1}) = I(t_k) + \frac{Kh}{T_i} e(t_k). \quad (13.7)$$

Si en lugar de esto, la derivada en la ecuación 13.6 se aproxima por la diferencia hacia atrás, se obtiene lo siguiente:

$$\frac{I(t_k) - I(t_k - 1)}{h} = \frac{K}{T_i} e(t_k).$$

Esto lleva a la siguiente ecuación recursiva para el término integral:

$$I(t_{k+1}) = I(t_k) + \frac{Kh}{T_i} e(t_{k+1}). \quad (13.8)$$

Otro método de aproximación simple es el de Tustin. Esta aproximación es

$$I(t_{k+1}) = I(t_k) + \frac{Kh}{T_i} \frac{e(t_{k+1}) + e(t_k)}{2}. \quad (13.9)$$

Otro método es el denominado equivalencia de rampa. Este método proporciona salidas exactas en los instantes de muestreo si la señal de entrada es continua y lineal a tramos entre los instantes de muestreo. El método de equivalencia en rampa proporciona la misma aproximación del término integral que la aproximación de Tustin, es decir, la ecuación 13.9.

Obsérvese que todas las aproximaciones tienen la misma forma, es decir,

$$I(t_{k+1}) = I(t_k) + b_{i1} e(t_{k+1}) + b_{i2} e(t_k), \quad (13.10)$$

pero con diferentes valores de los parámetros b_{i1} y b_{i2} .

Acción derivativa El término derivativo con el filtro clásico de primer orden viene dado por la ecuación 3.14, es decir,

$$\frac{T_d}{N} \frac{dD}{dt} + D = -KT_d \frac{dy}{dt}. \quad (13.11)$$

Esta ecuación se puede aproximar de la misma forma que el término integral.

Aproximando la derivada por una diferencia hacia adelante se obtiene

$$\frac{T_d}{N} \frac{D(t_{k+1}) - D(t_k)}{h} + D(t_k) = -KT_d \frac{y(t_{k+1}) - y(t_k)}{h}.$$

Esto se puede reescribir como

$$D(t_{k+1}) = \left(1 - \frac{Nh}{T_d}\right) D(t_k) - KN(y(t_{k+1}) - y(t_k)). \quad (13.12)$$

Si se aproxima la derivada en la ecuación 13.11 con la diferencia hacia atrás, se obtiene la siguiente ecuación:

$$\frac{T_d}{N} \frac{D(t_k) - D(t_{k-1})}{h} + D(t_k) = -KT_d \frac{y(t_k) - y(t_{k-1})}{h}.$$

Esta ecuación se puede reescribir como

$$D(t_k) = \frac{T_d}{T_d + Nh} D(t_{k-1}) - \frac{KT_d N}{T_d + Nh} (y(t_k) - y(t_{k-1})). \quad (13.13)$$

Utilizando la aproximación de Tustin para aproximar el término derivativo se tiene

$$D(t_k) = \frac{2T_d - Nh}{2T_d + Nh} D(t_{k-1}) - \frac{2KT_d N}{2T_d + Nh} (y(t_k) - y(t_{k-1})). \quad (13.14)$$

Finalmente, la aproximación de equivalencia de rampa da

$$D(t_k) = e^{-Nh/T_d} D(t_{k-1}) - \frac{KT_d(1 - e^{-Nh/T_d})}{h} (y(t_k) - y(t_{k-1})). \quad (13.15)$$

Todas las aproximaciones tienen la misma forma

$$D(t_k) = a_d D(t_{k-1}) - b_d (y(t_k) - y(t_{k-1})), \quad (13.16)$$

pero con valores diferentes de los parámetros a_d y b_d

Las aproximaciones del término derivativo son estables sólo cuando $|a_d| < 1$. Cuando se utiliza la aproximación de la diferencia hacia adelante, se requiere que $T_d > Nh/2$ para asegurar la estabilidad. La aproximación llega a ser inestable para pequeños valores de T_d . Las otras aproximaciones son estables para todos los valores de T_d . Sin embargo, obsérvese que la aproximación de Tustin y la de diferencia hacia adelante dan valores negativos de a_d si T_d es pequeño. Esto no es deseable porque la aproximación exhibirá oscilaciones transitorias. La aproximación hacia atrás da buenos resultados para todos los valores de T_d incluyendo $T_d = 0$.

Para muestreos razonablemente rápidos, sólo hay pequeñas diferencias entre las aproximaciones y además todas son estables. Sin embargo, existen diferencias prácticas. En un controlador de propósito general, es deseable que la acción derivativa se pueda desconectar. Una forma natural de hacer esto es hacer $T_d = 0$. Esto se puede llevar a cabo de manera sencilla cuando la acción derivativa se approxima por la diferencia hacia atrás. Todos los otros métodos resultarán en inestabilidad o desbordamiento para $T_d = 0$. Por lo tanto, la diferencia hacia atrás es una elección razonable para aproximar la derivada.

La Figura 13.9 muestra las curvas de fase para las distintas aproximaciones en tiempo discreto. La aproximación de Tustin y la de equivalencia en rampa dan un mejor ajuste al resultado de tiempo continuo, la aproximación hacia atrás da menos avance de fase y la hacia delante aporta más avance de fase. La aproximación hacia delante se utiliza rara vez debido a los problemas de inestabilidad

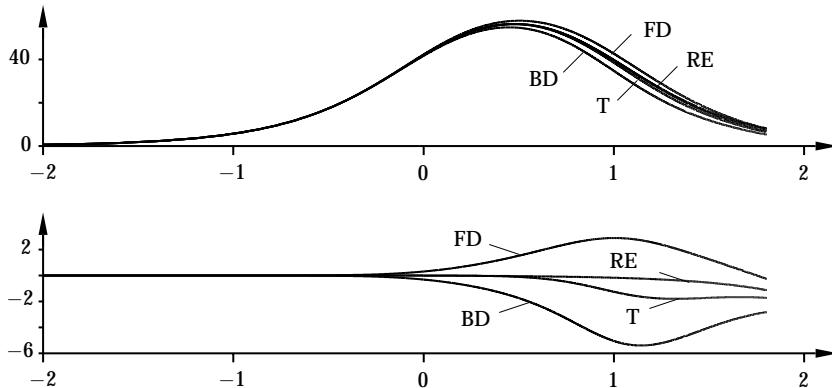


Figura 13.9 Curvas de fase para controladores PD obtenidos mediante diferentes discretizaciones del término derivativo $sT_d/(1+sT_d/N)$ con $T_d = 1$, $N = 10$ y un período de muestreo 0,02. Las discretizaciones son diferencias hacia adelante (FD), diferencias hacia atrás (BD), aproximaciones de Tustin (T) y equivalencia en rampa (RE). El diagrama inferior muestra las diferencias entre las aproximaciones y la curva de fase real.

para pequeños valores del tiempo derivativo T_d . El algoritmo de Tustin se utiliza con bastante frecuencia debido a su simplicidad y su buen ajuste con la función de transferencia de tiempo continuo. La diferencia hacia atrás se utiliza cuando se necesita un algoritmo que se comporte bien para pequeños valores de T_d .

Todas las aproximaciones del controlador PID se pueden representar como

$$R(q)u(kh) = T(q)y_{sp}(kh) - S(q)y(kh), \quad (13.17)$$

donde q es el operador de desplazamiento hacia delante, y los polinomios R , S y T son de segundo orden. Estos polinomios tienen la forma

$$\begin{aligned} R(q) &= (q - 1)(q - a_d) \\ S(q) &= s_0q^2 + s_1q + s_2 \\ T(q) &= t_0q^2 + t_1q + t_2, \end{aligned} \quad (13.18)$$

que significa que la ecuación 13.17 se puede escribir como

$$\begin{aligned} u(kh) &= t_0y_{sp}(kh) + t_1y_{sp}(kh-h) + t_2y_{sp}(kh-2h) \\ &\quad - s_0y(kh) - s_1y(kh-h) - s_2y(kh-2h) \\ &\quad + (1 + a_d)u(kh-h) - a_du(kh-2h). \end{aligned}$$

Cuadro 13.1 Coeficientes de las diferentes aproximaciones del controlador PID en tiempo continuo.

	Hacia adelante	Hacia atrás	Tustin	Equivalencia en rampa
b_{i1}	0	$\frac{Kh}{T_i}$	$\frac{Kh}{2T_i}$	$\frac{Kh}{2T_i}$
b_{i2}	$\frac{Kh}{T_i}$	0	$\frac{Kh}{2T_i}$	$\frac{Kh}{2T_i}$
a_d	$1 - \frac{Nh}{T_d}$	$\frac{T_d}{T_d + Nh}$	$\frac{2T_d - Nh}{2T_d + Nh}$	e^{-Nh/T_d}
b_d	KN	$\frac{KT_d N}{T_d + Nh}$	$\frac{2KT_d N}{2T_d + Nh}$	$\frac{KT_d(1 - e^{-Nh/T_d})}{h}$

Los coeficientes en los polinomios S y T son

$$\begin{aligned}
 s_0 &= K + b_{i1} + b_d \\
 s_1 &= -K(1 + a_d) - b_{i1}a_d + b_{i2} - 2b_d \\
 s_2 &= Ka_d - b_{i2}a_d + b_d \\
 t_0 &= Kb + b_{i1} \\
 t_1 &= -Kb(1 + a_d) - b_{i1}a_d + b_{i2} \\
 t_2 &= Kba_d - b_{i2}a_d.
 \end{aligned} \tag{13.19}$$

En la Tabla 13.1 se dan los coeficientes de los polinomios para los diferentes métodos de aproximación.

Controlador con filtro de segundo orden

Una buena implementación de un controlador PID consiste en combinar un filtrado de segundo orden de la señal medida con un controlador PID ideal; véase la Sección 3.3. Ahora discutiremos cómo se pueden implementar estos controladores. Sea y la señal de medida e y_f la señal filtrada. Tenemos que

$$Y_f(s) = G_f(s)Y(s) = \frac{1}{1 + sT_f + (sT_f)^2/2} Y(s). \tag{13.20}$$

Introduciendo las variables de estado $x_1 = y_f$ y $x_2 = T_f dy_f/dt$, el filtro se puede representar como

$$\begin{aligned}
 T_f \frac{dx_1}{dt} &= x_2 \\
 T_f \frac{dx_2}{dt} &= 2(-x_1 - x_2 + y).
 \end{aligned} \tag{13.21}$$

La derivada filtrada $dy_f/dt = x_2/T_f$ se puede extraer del filtro, y entonces el controlador viene dado por

$$u = k(by_{sp} - y_f) + k_i \int_0^t (y_{sp}(\tau) - y_f(\tau)) d\tau + k_d \frac{dy_f}{dt}. \quad (13.22)$$

Si el controlador PID (13.22) se implementa digitalmente ambas, $x_1 = y_f$ y $x_2 = T_f dy_f/dt$ se tienen que convertir a forma digital. Esta implementación es apropiada para sistemas de propósito especial. Para sistemas de propósito general, el filtro se puede implementar digitalmente. Supóngase que el muestreo tiene un periodo h y los instantes de muestreo vienen representados por t_k . Aproximando la derivada en (13.21) con la diferencia hacia atrás encontramos que

$$\begin{aligned} x_1(t) &= x_1(t-h) + \frac{hT_f}{T_f^2 + 2hT_f + 2h^2} x_2(t-h) \\ &\quad + \frac{2h^2}{T_f^2 + 2hT_f + 2h^2} (y(t) - x_1(t-h)) \\ x_2(t) &= \frac{T_f^2}{T_f^2 + 2hT_f + 2h^2} x_2(t-h) + \frac{2hT_f}{T_f^2 + 2hT_f + 2h^2} (y(t) - x_1(t-h)). \end{aligned}$$

Para obtener una implementación que permita que el parámetro T_f sea cero, introducimos las variables de estado

$$\begin{aligned} y_1 &= x_1 \\ y_2 &= \frac{h}{T_f} x_2. \end{aligned}$$

La ecuación del controlador se puede escribir como

$$\begin{aligned} y_1(t) &= y_1(t-h) + p_1 y_2(t-h) + p_2 (y(t) - y_1(t-h)) = y_1(t-h) + y_2(t) \\ y_2(t) &= p_1 y_2(t-h) + p_2 (y(t) - y_1(t-h)) \\ v(t) &= K(by_{sp} - y_1) - p_4 y_2(t) + I(t-h) \\ &= K(by_{sp} - y_1) - p_2(K + p_4)y(t) \\ &\quad + (p_2(K + p_4) - K)y_1(t-h) + (K - p_1(k + p_4))y_2(t-h) + I(t-h) \\ u(t) &= \text{sat}(v) \\ I(t) &= I(t-h) + p_3(y_{sp}(t) - y_1(t)) + p_5(u(t) - v(t)). \end{aligned} \quad (13.23)$$

donde el término integral se ha aproximado por la diferencia hacia adelante y se ha introducido la protección para windup. Los parámetros del controlador vienen dados por

$$\begin{aligned} p_1 &= \frac{T_f^2}{T_f^2 + 2hT_f + 2h^2} & p_2 &= \frac{2h^2}{T_f^2 + 2hT_f + 2h^2} \\ p_3 &= \frac{Kh}{T_i} & p_4 &= \frac{KT_d}{h} & p_5 &= \frac{h}{T_t} \end{aligned} \quad (13.24)$$

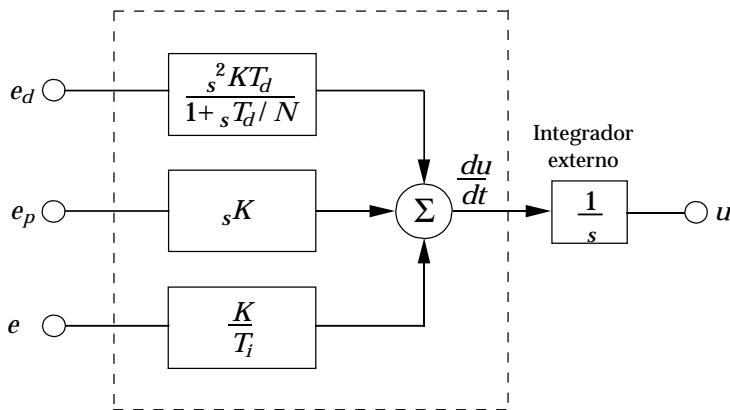


Figura 13.10 Diagrama de bloques de un algoritmo PID en forma de velocidad.

13.4 Algoritmos de velocidad

Los algoritmos descritos hasta ahora se denominan algoritmos posicionales porque la salida de los algoritmos es la variable de control. En ciertos casos, el sistema de control se dispone de tal manera que la señal de control está gobernada directamente por un integrador, por ejemplo, un motor. Entonces, es natural organizar el algoritmo de forma que aporte la velocidad de la variable de control. La variable de control se obtiene entonces integrando su velocidad. Un algoritmo de este tipo se denomina algoritmo de velocidad. En la Figura 13.10 se muestra un diagrama de bloques de un algoritmo de velocidad para un controlador PID.

Los algoritmos de velocidad fueron comúnmente utilizados en muchos de los primeros controladores que se construyeron para motores. En varios casos, la estructura se conservó por parte de los fabricantes cuando se cambió la tecnología con el fin de conservar la compatibilidad funcional con el equipo más antiguo. Otra razón es que muchos temas prácticos, como la protección windup y los cambios de parámetros sin salto, se pueden implementar de manera sencilla utilizando el algoritmo de velocidad. Esto se trata más ampliamente en la Secciónes 3.5 y 13.5. En las implementaciones digitales, los algoritmos de velocidad se denominan también algoritmos incrementales.

Algoritmo incremental

La forma incremental del algoritmo PID se obtiene calculando las diferencias de tiempo de la salida del controlador y sumando los incrementos

$$\Delta u(t_k) = u(t_k) - u(t_{k-1}) = \Delta P(t_k) + \Delta I(t_k) + \Delta D(t_k).$$

En algunos casos la integración se lleva a cabo externamente. Estos es natural cuando se utiliza un motor paso a paso. La salida del controlador debería representar entonces los incrementos de la señal de control, y el motor implementa el integrador. Los incrementos de la parte proporcional, la parte integral y la parte

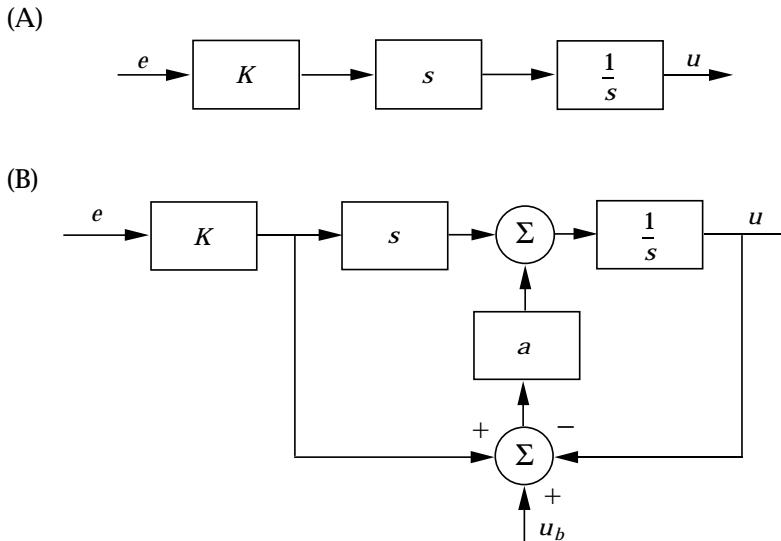


Figura 13.11 Ilustra la dificultad con el controlador proporcional en forma de velocidad (A) y una forma de evitarlo (B).

derivativa se calculan fácilmente a partir de las ecuaciones 13.5, 13.10 y 13.16:

$$\Delta P(t_k) = P(t_k) - P(t_{k-1}) = K(b y_{sp}(t_k) - y(t_k) - b y_{sp}(t_{k-1}) + y(t_{k-1}))$$

$$\Delta I(t_k) = I(t_k) - I(t_{k-1}) = b_{i1} e(t_k) + b_{i2} e(t_{k-1})$$

$$\Delta D(t_k) = D(t_k) - D(t_{k-1}) = a_d \Delta D(t_{k-1}) - b_d (y(t_k) - 2y(t_{k-1}) + y(t_{k-2})).$$

Una ventaja con el algoritmo incremental es que la mayoría de los cálculos se realizan utilizando sólo incrementos. Generalmente se pueden hacer los cálculos utilizando longitud de palabra corta. Sólo en la etapa final cuando se suman los incrementos es cuando se necesita precisión.

Algoritmos de velocidad para controladores sin acción integral

Un algoritmo de velocidad no se puede utilizar directamente para un controlador sin acción integral porque tal controlador no puede mantener el valor estacionario. Esto se puede comprender del diagrama de bloques de la Figura 13.11A, que muestra un controlador proporcional en la forma de velocidad. Se puede obtener estacionariedad para cualquier valor del error de control e , ya que la salida del bloque de derivación es cero para cualquier entrada constante. El problema se puede evitar con la modificación que se muestra en la Figura 13.11B. En este caso, la estacionariedad sólo se obtiene cuando $u = Ke + u_b$, donde u_b es el término de sesgo.

Si se utiliza un controlador PID muestreado, se obtiene una versión simple del método mostrado en la Figura 13.11B implementando un controlador P como

$$\Delta u(t) = u(t) - u(t-h) = Ke(t) + u_b - u(t-h),$$

donde h es el período de muestreo.

Control feedforward

El control feedforward se trató en el Capítulo 5. En el control feedforward, la señal de control se compone de dos términos,

$$u = u_{fb} + u_{ff},$$

donde u_{fb} es el componente de realimentación y u_{ff} es el componente feedforward, bien desde una perturbación medible o desde el punto de consigna.

Para evitar el windup del integrador, es importante que actúe el mecanismo anti windup sobre la señal de control final u y no sólo sobre la componente realimentada u_{fb} .

Desafortunadamente, muchos de los sistemas con instrumentos orientados a bloques disponibles en la actualidad incorporan mecanismos anti windup dentro de los bloques del controlador de realimentación, sin ninguna posibilidad de poder incorporar señales feedforward a estos bloques. Por lo tanto, las señales de feedforward se deben añadir después de los bloques del controlador. Esto puede conducir a windup. Debido a esto, se utilizan algunos trucos para solucionar el problema, tales como alimentar la señal de feedforward a través de filtros paso alto. Sin embargo, estas estrategias llevan a acciones feedforwards menos efectivas.

Los algoritmos incrementales son eficientes para la implementación feedforward. Primero se añaden *los incrementos* de los componentes de realimentación y feedforward,

$$\Delta u = \Delta u_{fb} + \Delta u_{ff}$$

y entonces, componiendo la señal de control como

$$u(t) = u(t - h) + \Delta u(t),$$

se evita el windup. Esto requiere que los bloques de control de realimentación tengan entradas para señales feedforward.

13.5 Aspectos operacionales

La gran mayoría de los controladores se pueden arrancar en dos modos: manual o automático. En modo manual, la salida del controlador se manipula directamente por el operador, que usualmente utiliza botones para aumentar y disminuir la salida del controlador. Un control puede también operar en combinación con otros controladores, tal como en una conexión en cascada o de relación, o con elementos no lineales, tales como multiplicadores y selectores. Esto da lugar a más modos operacionales. Los controladores también tienen parámetros que se pueden ajustar durante la operación. Cuando existen cambios de modos y parámetros, es esencial evitar los transitorios en las commutaciones. La manera en la que se realizan las commutaciones de modo y los cambios de parámetros depende de la estructura seleccionada para el controlador.

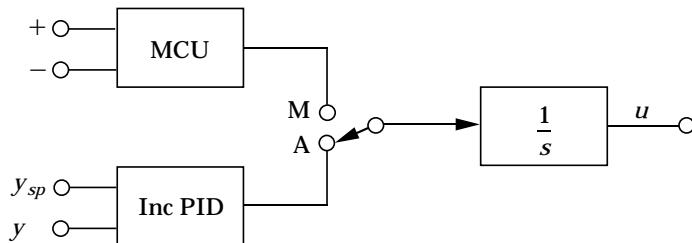


Figura 13.12 Transferencia sin salto en un controlador con salida incremental. MCU representa la unidad de control manual.

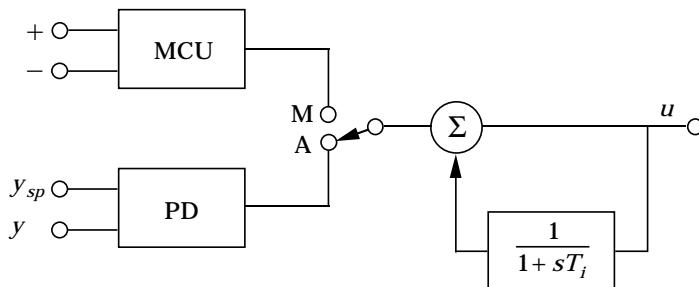


Figura 13.13 Transferencia sin salto en un controlador PID con una implementación en forma serie especial.

Transferencia sin salto entre modo manual y automático

Como el control es un sistema dinámico, es necesario asegurar que el estado del sistema es correcto cuando se realiza la conmutación del controlador entre modo manual y automático. Cuando el sistema está en modo manual, el algoritmo de control produce una señal de control que puede ser diferente de la señal de control generada manualmente. Es necesario asegurar que las dos señales coinciden en el tiempo de la conmutación. Esto se denomina *transferencia sin salto*.

La transferencia sin salto es sencilla de obtener para un controlador en forma incremental. Esto se muestra en la Figura 13.12. El integrador se proporciona con un conmutador tal que las señales se escogen o de la salida manual o de los incrementos automáticos. Como la conmutación sólo influye sobre los incrementos no habrá grandes transitorios.

En la Figura 3.3 se muestra un mecanismo similar que se puede utilizar en la implementación en serie o interactuante de un controlador PID, ver Figura 13.13. En este caso, existirá un transitorio en la conmutación si la salida de la parte del PD no es cero en el instante de conmutación.

Para controladores con implementación paralela, el integrador del controlador PID se puede utilizar para sumar los cambios en modo manual. En la Figura 13.14 se muestra tal controlador. Este sistema da una transición suave entre modo manual y automático a condición de que la conmutación se haga cuando la salida del bloque PD es cero. Si este no es el caso, habrá un transitorio en la conmutación.

También es posible utilizar un integrador separado para los cambios incre-

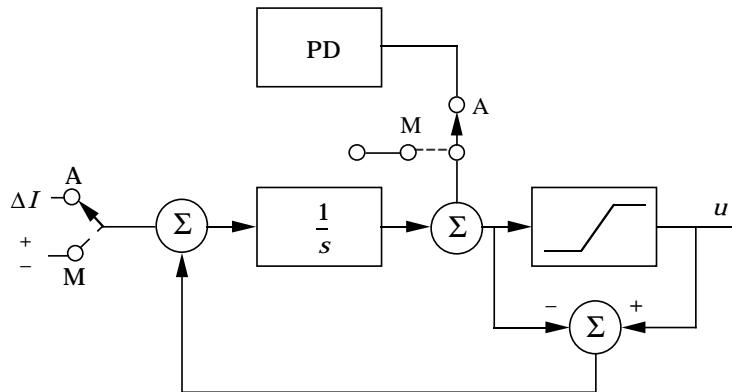


Figura 13.14 Un controlador PID donde se utiliza un integrador tanto para obtener acción integral en modo automático como para sumar las señales incrementales en modo manual.

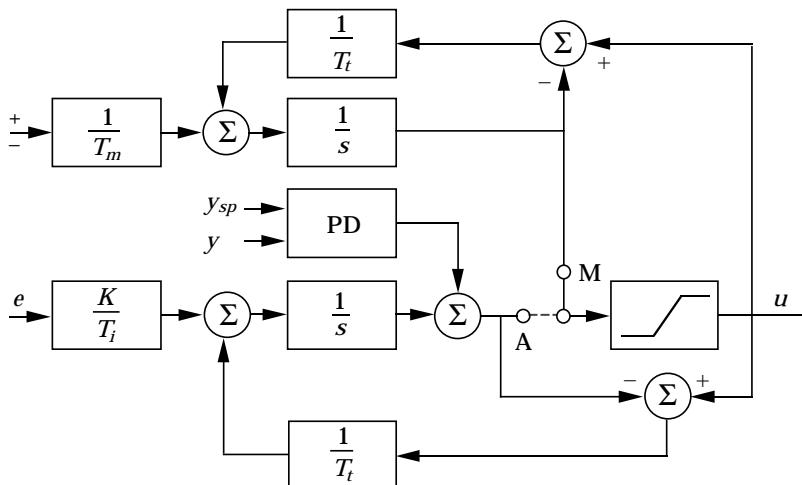


Figura 13.15 Controlador PID con implementación paralela que comunica suavemente entre el control manual y automático.

mentales del dispositivo de control manual. Para evitar transitorios en la comutación de tales sistemas, se necesita asegurar que el integrador del controlador PID se reinicie a un valor adecuado cuando el controlador está en modo manual. De manera similar, el integrador asociado con el control manual se debe reiniciar a un valor adecuado cuando el controlador está en modo automático. Esto se puede realizar con el circuito que se muestra en la Figura 13.15. Con este sistema, el cambio entre manual y automático es suave incluso si el error de control o su derivada es diferente de cero en el instante de comutación. Cuando el controlador opera en modo manual, tal y como se muestra en la Figura 13.15, la realimentación de la salida v del controlador PID sigue a la salida u . Con un seguimiento eficiente, la señal v estará cerca de u todo el tiempo. Hay un mecanismo de seguimiento similar que asegura que el integrador en el circuito de control

manual sigue a la salida del controlador.

Cambios de parámetros en la transferencia sin salto

Un controlador es un sistema dinámico. Un cambio de los parámetros de un sistema dinámico resultará naturalmente en cambios de su salida. Los cambios en la salida se pueden evitar, en algunos casos, mediante cambios simultáneos en el estado del sistema. Los cambios de la salida dependerán también de la realización seleccionada. Con un controlador PID es natural requerir que no existan cambios drásticos en la salida si los parámetros se modifican cuando el error es cero. Esto se cumplirá para todos los algoritmos incrementales porque la salida de un algoritmo incremental es cero cuando la entrada es cero, independientemente de los valores de los parámetros. Sin embargo, para un algoritmo de posición depende de la implementación.

Suponga que el estado se escoge como

$$x_I = \int_0^t e(\tau) d\tau$$

cuando se implementa el algoritmo. El término integral es entonces

$$I = \frac{K}{T_i} x_I.$$

Cualquier cambio de K o T_i resultará en un cambio de I . Para evitar saltos cuando se modifiquen los parámetros, es esencial que el estado se escoja como

$$x_I = \int_0^t \frac{K(\tau)}{T_i(\tau)} e(\tau) d\tau$$

cuando se implementa el término integral.

Con precauciones sensibles, es sencillo asegurar cambios de parámetros sin salto si se cambian cuando el error es cero. Sin embargo, existe un caso donde se tienen que tomar precauciones especiales, concretamente, cuando se utiliza la ponderación del punto de consigna. En tal caso, para tener cambios de parámetros sin salto, es necesario que la cantidad $P + I$ sea invariante a los cambios en los parámetros. Esto quiere decir que cuando se modifican los parámetros, el estado I se debería modificar como sigue:

$$I_{\text{nuevo}} = I_{\text{antiguo}} + K_{\text{antiguo}}(b_{\text{antiguo}} y_{sp} - y) - K_{\text{nuevo}}(b_{\text{nuevo}} y_{sp} - y). \quad (13.25)$$

Para construir los sistemas de automatización, es útil tener módulos apropiados. La Figura 13.16 muestra el diagrama de bloques para un módulo de control manual. Tiene dos entradas: una entrada de seguimiento y una entrada para la señal de control manual. El sistema tiene dos parámetros: la constante de tiempo T_m para la entrada de control manual y la constante de tiempo de reinicio T_t . En las implementaciones digitales es conveniente incluir una característica tal que la señal de referencia se acelera siempre que se pulse uno de los botones de

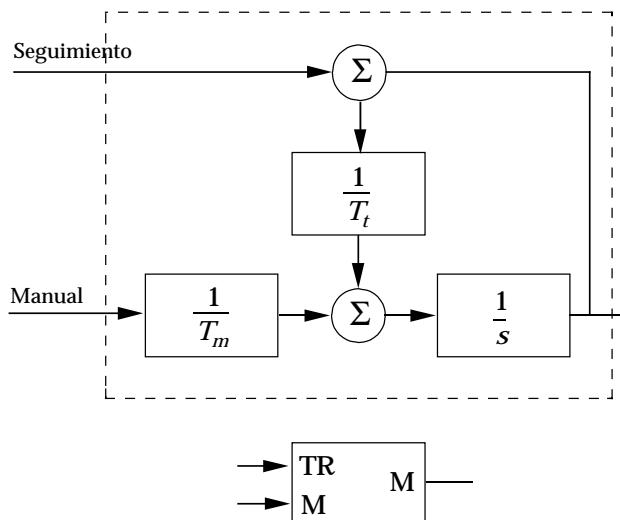


Figura 13.16 Módulo de control manual.

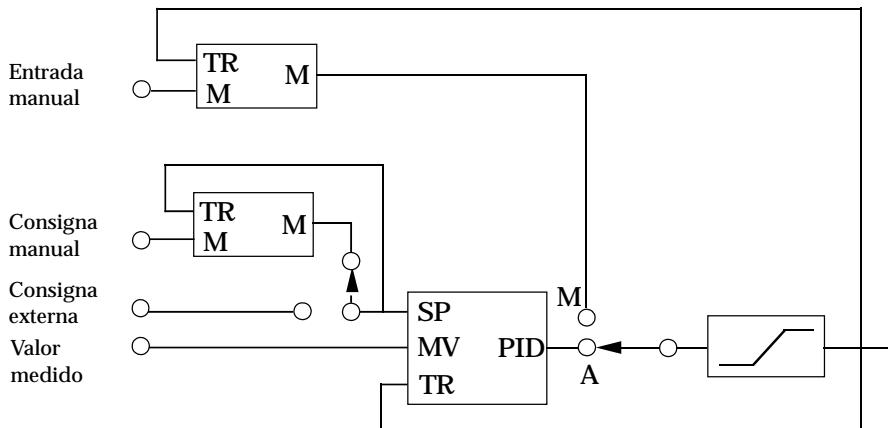


Figura 13.17 Un control PID razonablemente completo con anti windup, modo automático-manual y puntos de consigna manual y externo.

aumentar-disminuir. Utilizando el módulo de control PID y el módulo de control manual descritos en la Figura 13.16, es sencillo construir un controlador completo. La Figura 13.17 muestra un controlador PID con puntos de consigna externo e interno con botones aumentar-disminuir y el modo de control automático. Obsérvese que el sistema sólo tiene dos commutadores.

Código de computador

A modo de ejemplo se darán los códigos de computador para dos controladores PID. Las ecuaciones 13.5, 13.7, 13.9 y 13.13 describen un controlador PID con filtrado de primer orden del término derivativo donde el término derivativo se aproxima con la diferencia hacia atrás. Se proporciona anti windup utilizando el

```

Calcular los coeficientes del controlador
p1=K*b                               "ganancia del punto de consigna
p2=K+K*Td/(Tf+h)                      "ganancia del PD
p3=Tf/(Tf+h)                          "constante del filtro
p4=K*Td/(Tf+h)                        "ganancia derivativa
p5=K*h/Ti                             "ganancia integral
p6=h/Tt                               "ganancia anti-windup

Cambios en los parámetros de la transferencia sin salto
I=I+Kold*(bold*ysp-y)-Knew*(bnew*ysp-y)

Algoritmo de control
adin(ysp)                            "lectura punto de consigna
adin(y)                               "lectura variable del proceso
v=p1*ysp-p2*y+x+I                   "cálculo salida nominal
u=sat(v,ulow,uhigh)                 "saturación salida
daout(u)                             "establecer salida analógica
x=p3*x+p4*y                         "actualizar derivada
I=I+p5*(ysp-y)+p6*(u-v)           "actualizar integral

```

Figura 13.18 Esqueleto del código para implementar un controlador PID con filtrado de primer orden del término derivativo.

esquema descrito en la Sección 3.5. La Figura 13.18 muestra el esqueleto de código del controlador. El lazo principal tiene dos estados, el término integral I , y x que se utiliza para implementar la acción derivativa. Los parámetros p_1, \dots, p_6 se pre calculan para ahorrar tiempo de cómputo en el lazo principal. Estos parámetros tienen que calcularse sólo cuando se modifiquen los parámetros. El término integral se reinicia también tal como viene descrito por (13.25) para evitar transitorios cuando cambian los parámetros. El lazo principal en el algoritmo de control requiere ocho sumas y seis multiplicaciones. Obsérvese que los cálculos se estructuran de forma que existen sólo tres sumas y dos multiplicaciones entre la lectura de las entradas analógicas y el establecimiento de la salida digital. Los estados se actualizan después de fijar la salida digital.

La ecuación 13.23 describe un controlador PID con filtrado de segundo orden de la variable del proceso, donde el filtro se implementa utilizando la diferencia hacia atrás y el término integral se aproxima utilizando la diferencia hacia delante. Se obtiene anti windup mediante el esquema que se muestra en la Figura 3.13. El algoritmo tiene tres estados y_1 , y_2 e I , que representan los estados del filtro y del término integral. El lazo principal en el algoritmo de control requiere diez sumas y siete multiplicaciones. Utilizando un filtro de segundo orden, sólo se necesita un aumento marginal en el tiempo de cómputo. El tiempo entre la lectura de las entradas analógicas y el establecimiento de la salida digital se puede reducir cambiando las coordenadas de la representación del filtro.

```

Calcular los coeficientes del controlador
den=Tf*Tf+2*h*Tf+2*h*h      "denominador
p1=Tf*Tf/den                 "constante del filtro
p2=2*h*h/den                 "constante del filtro
p3=K*h/Ti                     "ganancia integral
p4=K*Td/h                      "ganancia derivativa
p5=h/Tt                         "ganancia anti-windup

Cambios en los parámetros de la transferencia sin salto
I=I+Kold*(bold*ysp-y1)-Knew*(bnew*ysp-y1)

"Algoritmo de control
r=adin(ysp)                   "lectura punto de consigna
y=adin(y)                      "lectura variable del proceso
y2=p1*y2+p2*(y-y1)            "actualización estado filtro y2
y1=y1+y2                      "actualización estado filtro y1
v=K*(b*ysp-y1)-p4*y2+I       "calcular salida nominal
u=sat(v,ulow,uhigh)           "saturación salida
daout(u)                       "establecer salida analógica
I=I+p3*(ysp-y1)+p5*(u-v)     "actualizar integral

```

Figura 13.19 Esqueleto del código para implementar un controlador PID con filtrado de segundo orden de la señal medida.

13.6 Salidas del controlador

Salidas analógicas

Las entradas y salidas de un controlador son normalmente señales analógicas, típicamente 0–20 mA o 4–20 mA. La principal razón de utilizar 4 mA en lugar de 0 mA como límite inferior es que la mayoría de los transmisores se diseñan para la conexión de dos cables. Esto quiere decir que el mismo cable se utiliza para manejar el sensor y transmitir la información desde el sensor. No sería posible manejar el sensor con una corriente de 0 mA. La principal razón de utilizar corriente en lugar de voltaje es para evitar la influencia de las caídas de voltaje a lo largo del cable debido a la resistencia del propio cable (quizás debido a la longitud). En controladores neumáticos, el rango estándar es 3–15 psi.

Tiristores y Triacs

En los controladores de temperatura es práctica común integrar el amplificador de potencia con el controlador. El amplificador de potencia podría ser un tiristor o un triac. Con un tiristor, se conmuta un voltaje de corriente alterna a la carga a un cierto ángulo del voltaje de alterna. Como la relación entre el ángulo y la potencia es no lineal, es crucial utilizar una transformación para mantener una relación lineal. Un triac es también un dispositivo que implementa la conmutación de una

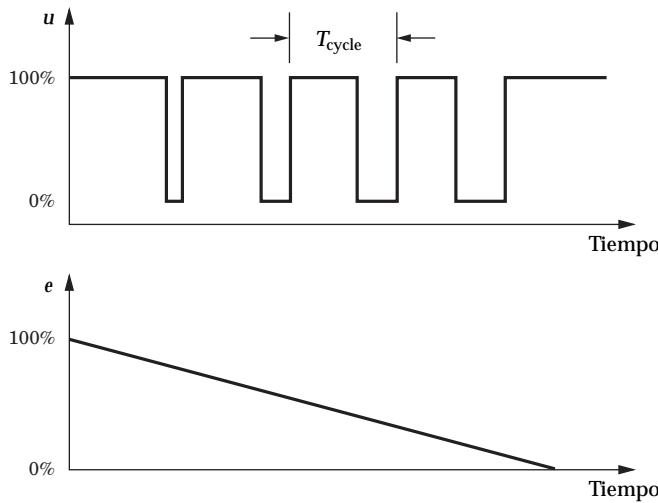


Figura 13.20 Ilustración de la salida del controlador basada en modulación en anchura de pulso.

señal de corriente alterna, pero sólo en el cruce por cero. Tal dispositivo es similar a una salida de pulso.

Modulación de la anchura del pulso

En algunos casos, como con el triac, existe una cuantificación extrema en el sentido de que el actuador sólo acepta dos valores todo/nada. En tal caso, se especifica un tiempo de ciclo T_{ciclo} y el controlador da un pulso con anchura

$$T_{\text{pulso}}(t) = \frac{u(t) - u_{\text{mín}}}{u_{\text{máx}} - u_{\text{mín}}} T_{\text{ciclo}}. \quad (13.26)$$

Una situación similar, aunque un poco distinta, sucede cuando el actuador tiene tres niveles: máximo, mínimo y cero. Un ejemplo típico es una válvula controlada por un motor donde el motor puede permanecer quieto, ir hacia delante o ir hacia atrás.

La Figura 13.20 ilustra la modulación de la anchura del pulso. La figura muestra la salida del controlador P con modulación de anchura del pulso para distintos valores del error de control.

Salida de un pulso con tres posiciones

Si una válvula se mueve mediante un motor eléctrico de velocidad constante, la válvula puede estar en tres estados: «aumentar», «parar» y «disminuir». El control de válvulas con actuadores eléctricos se realiza con una salida del controlador que puede estar en tres estados. La salida de un pulso con tres posiciones se lleva a cabo utilizando dos salidas digitales del controlador. Cuando la primera salida se activa, la posición de la válvula aumentará. Cuando esto ocurre con la segunda salida, la posición de la válvula disminuirá. Si no se activa ninguna de

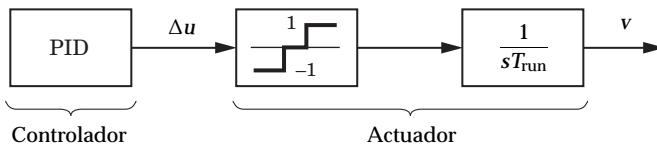


Figura 13.21 Un controlador PID con salida de pulso de tres posiciones combinado con un actuador eléctrico.

las dos salidas, la posición de la válvula es constante. Las dos salidas nunca deben activarse al mismo tiempo.

Normalmente, el controlador tiene una zona muerta y un retardo de tiempo para asegurar que el cambio de dirección del motor no sea ni demasiado frecuente ni demasiado rápido. Esto significa que la salida del controlador es constante siempre que la magnitud del error de control esté en la zona muerta y que la salida se pare durante unos segundos antes de permitir un cambio de dirección.

Un servo motor se caracteriza por su tiempo de arranque T_{run} , que es el tiempo que lleva que el motor vaya desde una posición final a la otra. Como el servo motor tiene una velocidad constante, introduce un integrador en el lazo de control, donde el tiempo de integración se determina de T_{run} . En la Figura 13.21 se muestra un diagrama de bloques de un controlador PID con salida de pulso de tres posiciones combinado con un actuador eléctrico. Supóngase que tenemos una situación en estado estacionario, donde la salida del controlador PID u es igual a la posición v del servo motor. Supóngase además que queremos aumentar rápidamente la salida del controlador en una cantidad Δu . Mientras el incremento de salida está activa, la salida v del servo-motor aumentará de acuerdo con

$$\Delta v = \frac{1}{T_{\text{run}}} \int_0^t 1 dt = \frac{t}{T_{\text{run}}}.$$

Para tener Δv igual a Δu , la integración se debe parar después del tiempo

$$t = \Delta u T_{\text{run}}.$$

En un controlador digital, esto quiere decir que la salida digital correspondiente a una posición de la válvula creciente se durará n períodos de muestreo, donde n viene dado por

$$n = \frac{\Delta u T_{\text{run}}}{h},$$

donde h es el período de muestreo del controlador.

Para ser capaz de llevar a cabo una salida de pulso de tres posiciones correcta, se deben utilizar dos memorias (`Buff_increase` y `Buff_decrease`) para gestionar el número de pulsos que se deberían enviar. En la Figura 13.22 se da un código de computador para la salida de pulso de tres posiciones. Por temas de simplicidad, en el código se han omitido detalles como zona muerta o retardo de tiempo.

De acuerdo con la Figura 13.21, la salida del controlador es Δu en lugar de u en el caso de la salida de pulso de tres posiciones. En el actuador, la parte integral del

```

if delta_u >0 then
    if valve_is_increasing then
        Buff_increase = Buff_increase + n;
    else
        Buff_decrease = Buff_decrease - n;
        if Buff_decrease <0 then
            Buff_increase = - Buff_decrease;
            Buff_decrease = 0;
            valve_is_decreasing = false;
            valve_is_increasing = true;
        end;
    end;
else if delta_u <0 then
    if valve_is_increasing then
        Buff_decrease = Buff_decrease + n;
    else
        Buff_increase = Buff_increase - n;
        if Buff_increase <0 then
            Buff_decrease = - Buff_increase;
            Buff_increase = 0;
            valve_is_increasing = false;
            valve_is_decreasing = true;
        end;
    end;
end;
if Buff_increase >0 then
    Increaseoutput = 1;
    Decreaseoutput = 0;
    Buff_increase = Buff_increase - 1;
else if Buff_decrease >0 then
    Increaseoutput = 0;
    Decreaseoutput = 1;
    Buff_decrease = Buff_decrease - 1;
end;

```

Figura 13.22 Código reducido para una salida de pulso de tres posiciones.

algoritmo de control está fuera del controlador. Esta solución no causa problemas si el algoritmo de control realmente contiene una parte integral. No se puede obtener control P o PD sin información de la posición de la válvula, véase la Figura 13.11.

13.7 Resumen

En este capítulo hemos descrito la implementación de controladores PID. Hemos seguido el desarrollo histórico comenzando con la implementación neumática y electrónica de los controladores analógicos. Se ha presentado con detalle la implementación en computador incluyendo el esqueleto del código. La razón de hacer esto es que muchas características de las implementaciones modernas han heredado algunas características de los antiguos computadores analógicos. Un ejemplo es la preferencia de la implementación de la forma serie.

Es interesante considerar el desarrollo de los controladores. Durante cada fase del desarrollo, la tecnología ha madurado y mejorado, pero el conocimiento se ha ido perdiendo generalmente con los cambios tecnológicos. Por ejemplo, llevó bastante tiempo hasta antes de que se apreciase en las implementaciones en computador la importancia del filtrado de la medida y las protecciones de anti windup. Una razón para ello es que muchos detalles no estuvieron bien documentados y así se olvidaron fácilmente cuando cambió la tecnología. Otra fue que algunas buenas características se obtuvieron automáticamente a causa de características particulares de la tecnología. En este capítulo también se han tratado cuestiones importantes de las operaciones y de las interfaces hombre-máquina.

13.8 Notas y referencias

El libro [172] presenta muchos detalles de las primeras implementaciones de controladores PID utilizando tecnologías neumática, hidráulica y electrónica. En [216, 271] se trata la implementación de controladores neumáticos. Las implementaciones electrónicas se analizan en [6]. Un aspecto interesante es que todos los libros mencionados hasta ahora fueron escritos por los fabricantes de los equipos. El artículo [127] describe los primeros esfuerzos en la implementación digital de controladores PID. Las implementaciones digitales se tratan en detalle en [84, 152, 30]. El artículo [344] da una amplia descripción del desarrollo del controlador de temperatura Eurotherm abarcando un periodo de más de dos décadas y tecnologías desde la electrónica analógica a la digital. El libro [97] describe la implementación de controladores para el control de movimiento. En [27] se puede encontrar código para la implementación de procesadores de señales que admite un muestreo muy rápido.

Bibliografía

- [1] ABB. *Control^{IT}: Analog Process Control — Objects and Design*, version 3.2 edition, September 2002. 3BSE 028 809 R101 Rev B.
- [2] S. Adusumilli, D. E. Rivera, S. Dash, and K. Tsakalis. Integrated MIMO identification and robust PID controller design through loop shaping. In *Proc. American Control Conference (ACC'98)*, volume 2, pages 1230–1234, Philadelphia, PA, June 1998.
- [3] S.G. Akkermans and S.G. Stan. Digital servo IC for optical disc drives. *Control Engineering Practice*, 9(11):1245–1253, 2002.
- [4] Frank Allgower and H Zheng. *Nonlinear model predictive control, progress in systems and control theory*. Birkhauser, Basel, 2000.
- [5] K. L. Anderson, G. L. Blankenship, and L. G. Lebow. A rule-based PID controller. In *Proc. IEEE Conference on Decision and Control*, Austin, Texas, 1988.
- [6] Norman A. Anderson. *Instrumentation for Process Measurement and Control*. Chilton Book Company, Radnor, PA, 1972.
- [7] A. S. Anderssen and E. T. White. Parameter estimation by the transfer function method. *Chemical Engineering Science*, 25:1015–1021, 1970.
- [8] A. S. Anderssen and E. T. White. Parameter estimation by the weighted moments method. *Chemical Engineering Science*, 26:1203–1221, 1971.
- [9] Anonymous. Control loop optimization. Technical Report SSG-5253, Pulp and Paper IndustriesÉngineering Co, Sundsvall, Sweden, 1997.
- [10] M. Araki and H. Taguchi. Two-degree-of-freedom PID controller. *Systems, Control and Information*, 42, 1998.
- [11] M Araki. PID control systems with reference feedforward (PID-FF control systems). In *SICE Annual Conference*, page 31, 1984.
- [12] Karl Johan Åström and Torsten Bohlin. Numerical identification of linear dynamic systems from normal operating records. In *Proc. IFAC Conference on Self-Adaptive Control Systems*, Teddington, UK, 1965.

Bibliografía

- [13] Karl Johan Åström, Hilding Elmquist, and Sven Erik Mattsson. Evolution of continuous-time modeling and simulation. In Richard Zobel and Dietmar Moeller, editors, *Proceedings of the 12th European Simulation Multiconference, ESM'98*, pages 9–18, Manchester, UK, June 1998. Society for Computer Simulation International.
- [14] Karl Johan Åström, Tore Hägglund, C. C. Hang, and W. K. Ho. Automatic tuning and adaptation for PID controllers—A survey. *Control Engineering Practice*, 1(4):699–714, 1993.
- [15] Karl Johan Åström and Tore Hägglund. Automatic tuning of simple regulators. In *Preprints 9th IFAC World Congress*, pages 267–272, Budapest, Hungary, 1984.
- [16] Karl Johan Åström and Tore Hägglund. Automatic tuning of simple regulators with specifications on phase and amplitude margins. *Automatica*, 20:645–651, 1984.
- [17] Karl Johan Åström and Tore Hägglund. *Automatic Tuning of PID Controllers*. ISA, Research Triangle Park, North Carolina, 1988.
- [18] Karl Johan Åström and Tore Hägglund. Practical experiences of adaptive techniques. In *American Control Conference (ACC '90)*, San Diego, California, 1990.
- [19] Karl Johan Åström and Tore Hägglund. New tuning methods for PID controllers. In *European Control Conference*, pages 2456–2462, Rome, Italy, 1995.
- [20] Karl Johan Åström and Tore Hägglund. *PID Controllers: Theory, Design, and Tuning*. ISA, Research Triangle Park, North Carolina, 1995.
- [21] Karl Johan Åström and Tore Hägglund. The future of PID control. *Control Engineering Practice*, 9:1163–1175, 2001.
- [22] Karl Johan Åström, C. C. Hang, and B. C. Lim. A new Smith predictor for controlling a process with an integrator and long dead-time. 39(2), February 1994.
- [23] Karl Johan Åström, Karl Henrik Johansson, and Q.G. Wang. Design of decoupled PI controller for two-by-two systems. *IEE Proceedings on Control Theory and Applications*, 149:74–81, January 2002.
- [24] Karl Johan Åström and Thomas J. McAvoy. Intelligent control. *Journal of Process Control*, 2(2):1–13, 1992.
- [25] Karl Johan Åström, Hélène Panagopoulos, and Tore Hägglund. Design of PI controllers based on non-convex optimization. *Automatica*, 34(5):585–601, May 1998.
- [26] Karl Johan Åström and Lars Rundqvist. Integrator windup and how to avoid it. In *Proceedings of the American Control Conference (ACC '89)*, pages 1693–1698, Pittsburgh, Pennsylvania, 1989.

- [27] Karl Johan Åström and Hermann Steingrímsson. Implementation of a PID controller on a DSP. In Irfan Ahmed, editor, *Digital Control Applications with the TMS 320 Family*, Selected Application Notes, pages 205–238. Texas Instruments, 1991.
- [28] Karl Johan Åström and Björn Wittenmark. On self-tuning regulators. *Automatica*, 9:185–199, 1973.
- [29] Karl Johan Åström and Björn Wittenmark. *Adaptive Control*. Addison-Wesley, Reading, Massachusetts, second edition, 1995.
- [30] Karl Johan Åström and Björn Wittenmark. *Computer-Controlled Systems*. Prentice Hall, third edition, 1997.
- [31] Karl Johan Åström. Computer control of a paper machine — An application of linear stochastic control theory. *IBM Journal of Research and Development*, 11, 1967.
- [32] Karl Johan Åström. *Introduction to Stochastic Control Theory*. Academic Press, New York, 1970.
- [33] Karl Johan Åström. Frequency domain properties of Otto Smith regulators. 26:307–314, 1977.
- [34] Karl Johan Åström. Adaptive feedback control. *Proc. IEEE*, 75:185–217, February 1987. Invited paper.
- [35] Karl Johan Åström. Advanced control methods—Survey and assessment of possibilities. In Morris, Kompass, and Williams, editors, *Advanced Control in Computer Integrated Manufacturing*. Purdue University, West Lafayette, Indiana, 1987.
- [36] Karl Johan Åström. Automatic tuning and adaptive control—Past accomplishments and future directions. In D. M. Prett and C. E. Garcia, editors, *The Second Shell Process Control Workshop*, pages 1–24, Stoneham, Massachusetts, 1990. Butterworths.
- [37] Karl Johan Åström. Assessment of achievable performance of simple feedback loops. 5:3–19, 1991.
- [38] Karl Johan Åström. Limitations on control system performance. *European Journal on Control*, 6(1):2–20, 2000.
- [39] Karl Johan Åström. Control problems in paper making: Revisited. *Pulp & Paper Canada*, 102:6:39–44, 2001.
- [40] K. J. Åström and T. Hägglund. Automatic tuning of simple regulators with specifications on phase and amplitude margins. *Automatica*, 20(5):645–651, 1984.
- [41] M. Athans and P. L. Falb. *Optimal Control*. McGraw-Hill, New York, 1966.
- [42] Derek P. Atherton. *Nonlinear Control Engineering—Describing Function Analysis and Design*. Van Nostrand Reinhold Co., London, UK, 1975.

Bibliografía

- [43] D.P. Atherton. PID controller tuning. *Computing & Control Engineering Journal*, pages 44–50, April 1999.
- [44] R. Bellman. *Dynamic Programming*. Princeton University Press, New Jersey, 1957.
- [45] S. Bennett. *A History of Control Engineering 1800-1930*. Peter Peregrinus, London, 1979.
- [46] S. Bennett. *A History of Control Engineering 1930-1955*. Peter Peregrinus, London, 1993.
- [47] S. Bennett. Past of PID controllers. In *Proc. IFAC Workshop on Digital Control: Past, Present and Future of PID Control*, Terrassa, Spain, 2000.
- [48] Wayne Bequette. *Process Control: Modeling, Design and Simulation*. Prentice Hall, 2003.
- [49] B. Bialkowski. *Process Control Fundamentals for the Pulp & Paper Industry*, chapter Process control sample problems. Tappi Press, 1995.
- [50] W. L. Bialkowski. Dream vs. reality - a view from both sides of the gap. *Pulp & Paper Canada*, 11:19–27, 1994.
- [51] H. S. Black. Inventing the negative feedback amplifier. *IEEE spectrum*, pages 55–60, December 1977.
- [52] T. L. Blevins, G. K. McMillan, W. K. Wojsznis, and Brown M. W. *Advanced Control Unleashed*. ISA, Research Triangle Park, NC, 2003.
- [53] George J. Blickley. PID tuning made easy with hand-held computer. *Control Engineering*, page 99, November 1988.
- [54] George Blickley. Modern control started with Ziegler-Nichols tuning. *Control Engineering*, pages 11–17, October 1990.
- [55] H. W. Bode. *Network Analysis and Feedback Amplifier Design*. Van Nostrand, New York, 1945.
- [56] Torsten Bohlin. *Interactive system identification: Prospects and Pitfalls*. Springer, Berlin, 1991.
- [57] Stephen P. Boyd and Craig H. Barratt. *Linear Controller Design – Limits of Performance*. Prentice Hall Inc., Englewood Cliffs, New Jersey, 1991.
- [58] E. H. Bristol, G. R. Inaloglu, and J. F. Steadman. Adaptive process control by pattern recognition. *Instrum. Control Systems*, pages 101–105, 1970.
- [59] E. H. Bristol and T. W. Kraus. Life with pattern adaptation. In *Proc. 1984 American Control Conference*, San Diego, California, 1984.
- [60] E. H. Bristol. A simple adaptive system for industrial control. *Instrumentation Technology*, June 1967.
- [61] E. H. Bristol. Adaptive control odyssey. In *ISA Silver Jubilee Conference*, Philadelphia, 1970.

- [62] E. H. Bristol. Pattern recognition: An alternative to parameter identification in adaptive control. *Automatica*, 13:197–202, 1977.
- [63] E. H. Bristol. After DDC: Idiomatic (structured) control. In *Proceedings American Institute of Chemical Engineering (AIChE)*, Philadelphia, 1980.
- [64] E. H. Bristol. The EXACT pattern recognition adaptive controller, a user-oriented commercial success. In K. S. Narendra, editor, *Adaptive and Learning Systems*, pages 149–163, New York, 1986. Plenum Press.
- [65] E. Bristol. On a new measure of interaction for multivariable process control. 11:133, 1966.
- [66] G. S. Brown and D. P. Campbell. *Principles of Servomechanisms*. Wiley & Sons, New York, 1948.
- [67] Martin Brown and Chris Harris. *Neurofuzzy Adaptive Modelling and Control*. Prentice Hall, 1994.
- [68] A. E. Bryson and Y. C. Ho. *Applied Optimal Control Optimization, Estimation and Control*. Blaisdell Publishing Company, 1969.
- [69] Page S. Buckley. *Techniques of Process Control*. John Wiley & Sons, Inc., 1964.
- [70] Page S. Buckley. Protective controls for a chemical reactor. *Chemical Engineering*, pages 145–150, April 1970.
- [71] P. J. Callaghan, P. L. Lee, and R. B. Newell. Evaluation of Foxboro controller. *Process Control Engineering*, May:38–40, 1986.
- [72] A. Callender, D. R. Hartree, and A. Porter. Time lag in a control system. *Philos. Trans. A.*, 235:415–444, 1936.
- [73] F. Cameron and D. E. Seborg. A self-tuning controller with a PID structure. *Int. J. Control*, 38(2):401–417, 1983.
- [74] R. H. Cannon. *Dynamics of Physical Systems*. McGraw-Hill, New York, 1967.
- [75] T.-N. Chang and E. J. Davison. Steady-state interaction indices for decentralized unknown systems. In *Preprints 16th IEEE Conf. on Decision and Control*, volume 2, pages 881–887, Athens, Greece, 1987.
- [76] B.-S. Chen and S.-S. Wang. The stability of feedback control with nonlinear saturating actuator: Time domain approach. 33:483–487, 1988.
- [77] C.-L. Chen. A simple method for on-line identification and controller tuning. *AIChE Journal*, 35(12):2037–2039, 1989.
- [78] Harold Chestnut and Robert W. Mayer. *Servomechanisms and Regulating System Design*. Wiley, New York, 1959.
- [79] I.-L. Chien and P. S. Fruehauf. Consider IMC tuning to improve controller performance. *Chemical Engineering Progress*, pages 33–41, October 1990.
- [80] I. L. Chien. IMC-PID controller design—an extension. In *IFAC Symposium, Adaptive Control of Chemical Processes*, pages 155–160, Copenhagen, Denmark, 1988.

Bibliografía

- [81] K. L. Chien, J. A. Hrones, and J. B. Reswick. On the automatic control of generalized passive systems. *Trans. ASME*, 74:175–185, 1952.
- [82] Younqjin Choi and Wan Kyun Chung. *PID Trajectory Tracking Control for Mechanical Systems*. Springer, Berlin, 2004.
- [83] D. W. Clarke and C. E. Hinton. Adaptive control of materials-testing machines. *Automatica*, 33(6):1119–1131, 1997.
- [84] D. W. Clarke. PID algorithms and their computer implementation. *Trans. Inst. Measurement and Control*, 6(6):305–316, 1984.
- [85] Charles M. Close and Dean K. Frederick. *Modeling and Analysis of Dynamic Systems*. Houghton Mifflin, 1993.
- [86] G. H. Cohen and G. A. Coon. Theoretical consideration of retarded control. *Trans. ASME*, 75:827–834, 1953.
- [87] P. Cominos and N. Munro. PID controllers: recent tuning methods and design to specifications. *IEE Proceedings; Control theory and applications*, 149(1):46–53, 2002.
- [88] G. A. Coon. How to find controller settings from process characteristics. *Control Engineering*, 3:66–76, 1956.
- [89] G. A. Coon. How to set three-term controller. *Control Engineering*, 3:71–76, 1956.
- [90] A. B. Corripio. *Tuning of Industrial Control Systems*. ISA, 1990.
- [91] C. R. Cutler and B. C. Ramaker. Dynamic matrix control—A computer control algorithm. In *Proceedings Joint Automatic Control Conference*, San Francisco, California, 1980.
- [92] E. B. Dahlin. Designing and tuning digital controllers. *Instruments and Control Systems*, 42:77–83, June 1968.
- [93] L. Desborough and T. Harris. Performance assessment measures for univariate feedback control. *Canadian Journal of Chemical Engineering*, 70:1186–1197, 1992.
- [94] L. Desbourough and R. Miller. Increasing customer value of industrial control performance monitoring - Honeywell's experience. In *Sixth International Conference on Chemical Process Control*. AIChE Symposium Series Number 326 (Volume 98), 2002.
- [95] Pradeep B. Deshpande and Raymond H. Ash. *Elements of Computer Process Control with Advanced Control Applications*. ISA, Research Triangle Park, North Carolina, 1981.
- [96] W. DeWries and S.M. Wu. Evaluation of process control effectiveness and diagnosis of variation in paper basis weight via multivariate time-series analysis. *IEEE Trans. on Automatic Control*, 23(4):702–708, 1978.
- [97] Yasuhiko Dote. *Servo motor and motion control using digital signal processors*. Chilton Book Company, Radnor, PA, 1972.

- [98] J. J. Downs. Linking control strategy design and model predictive control. pages 411–422, Tucson, AZ, 2001. Chemical Process Control - 6, Assessment and new directions for research (CPC-VI).
- [99] J. C. Doyle, B. A. Francis, and A. R. Tannenbaum. *Feedback Control Theory*. Macmillan, New York, 1992.
- [100] L. H. Dreinhofer. Controller tuning for a slow nonlinear process. *IEEE Control Systems Magazine*, 8(2):56–60, 1988.
- [101] D. Driankov, H. Hellendoorn, and M. Reinfrank. *An Introduction to Fuzzy Control*. Springer-Verlag, 1993.
- [102] G. A. Dumont, J. M. Martin-Sánchez, and C. C. Zervos. Comparison of an auto-tuned PID regulator and an adaptive predictive control system on an industrial bleach plant. *Automatica*, 25:33–40, 1989.
- [103] G. A. Dumont. On the use of adaptive control in the process industries. In Morari and McAvoy, editors, *Proceedings Third International Conference on Chemical Process Control-CPCIII*, Amsterdam, 1986. Elsevier.
- [104] D. P. Eckman. *Principles of industrial process control*. Wiley, New York, 1945.
- [105] O. I. Elgerd and W. C. Stephens. Effect of closed-loop transfer function pole and zero locations on the transient response of linear control systems. *Applications and Industry*, 42:121–127, 1959.
- [106] Hilding Elmqvist, Sven Erik Mattsson, and Martin Otter. Modelica—The new object-oriented modeling language. In Richard Zobel and Dietmar Moeller, editors, *Proceedings of the 12th European Simulation Multiconference, ESM'98*, pages 127–131, Manchester, UK, June 1998. Society for Computer Simulation International.
- [107] David B. Ender. Process control performance: Not as good as you think. *Control Engineering*, 40(10):180–190, 1993.
- [108] H. A. Fertik and C. W. Ross. Direct digital control algorithms with anti-windup feature. *ISA Trans.*, 6(4):317–328, 1967.
- [109] H. A. Fertik. Tuning controllers for noisy processes. *ISA Transactions*, 14:292–304, 1975.
- [110] A. M. Foss. Criterion to assess stability of a ‘lowest wins’ control strategy. *IEEE Proc. Pt. D*, 128(1):1–8, 1981.
- [111] Foxboro, Inc. *Controller Tuning Guide*, 1979.
- [112] F. Fröhr and F. Ortenburger. *Introduction to Electronic Control Engineering*. Siemens Aktiengesellschaft, Heyden & Son Ltd., London, 1982.
- [113] F. Fröhr. Optimierung von Regelkreisen nach dem Betragsoptimum und dem symmetrischen Optimum. *Automatik*, 12:9–14, January 1967.
- [114] S. E. Gallun, C. W. Matthews, C. P. Senyad, and B. Slater. Windup protection and initialization for advanced digital control. *Hydrocarbon Processing*, pages 63–68, June 1985.

Bibliografía

- [115] C. E. Garcia and A. M. Morshedi. Quadratic programming solution of dynamic matrix control (QDMC). *Chemical Engineering Communications*, 46:73–87, 1986.
- [116] P. J. Gawthrop. Self-tuning PID controllers: Algorithms and implementation. 31:201–209, 1986.
- [117] A. Gelb and W. E. Vander Velde. *Multiple-Input Describing Functions and Nonlinear System Design*. McGraw-Hill, New York, 1968.
- [118] J. P. Gerry. A comparison of PID controller algorithms. *Control Engineering*, pages 102–105, March 1987.
- [119] J. Gerry. Tuning process controllers start in manual. *InTech*, May:125–126, 1999.
- [120] L.G. Gibilaro and F.P. Lees. The reduction of complex transfer function models to simple models using the method of moments. *Chem Eng Sci*, 24:85–93, January 1969.
- [121] J. C. Gille, M. J. Pelegrin, and P. Decaulne. *Feedback Control Systems*. McGraw-Hill, New York, 1959.
- [122] A. H. Glattfelder, L. Guzzella, and W. Schaufelberger. Bumpless transfer, anti-reset-windup, saturating and override controls: A status report on self-tuning regulators. In *Proceedings of IMACS-88, Part 2*, pages 66–72, Paris, France, 1988.
- [123] A. H. Glattfelder and Walter Schaufelberger. *Control systems with input and output constraints*. Springer, 2003.
- [124] A. H. Glattfelder and W. Schaufelberger. Start-up performance of different proportional-integral-anti-wind-up regulators. 44:493–505, 1986.
- [125] A. H. Glattfelder and Schaufelberger. Stability analysis of single loop systems with saturation and antireset-windup circuits. 28:1074–1081, 1983.
- [126] Keith Glover. A tutorial on model reduction. In Jan C. Willems, editor, *From data to model*. Springer, Berlin, 1990.
- [127] Kenneth W. Goff. Dynamics in direct digital control II—A systematic approach to DDC design. *ISA Journal*, 13:44–54, December 1966.
- [128] Kenneth W. Goff. Dynamics in direct digital control I—Estimating characteristics and effects of noisy signals. *ISA Journal*, 13:45–49, November 1966.
- [129] D. Graham and R. C. Lathrop. The synthesis of ‘optimum’ transient response: Criteria and standard forms. *Transactions of the AIEE*, 72:273–288, November 1953.
- [130] J. J. Grebe, R. H. Boundy, and R. W. Cermak. The control of chemical processes. *Trans. of American Institute of Chemical Engineers*, 29:211–255, 1933.
- [131] Michael Green and D. J. Limebeer. *Linear Robust Control*. Prentice Hall, Englewood Cliffs, N.J., 1995.

- [132] P. M. E. M. van der Grinten. Finding optimum controller settings. *Control Engineering*, pages 51–56, December 1963.
- [133] P. Grosdidier, M. Morari, and P.R. Holt. Closed-loop properties from steady-state gain information. *Ind. Eng. Chem. Fundamentals*, 24:221–235, 1985.
- [134] A. Haalman. Adjusting controllers for a deadtime process. *Control Engineering*, July 65:71–73, 1965.
- [135] F. Habel. Ein Verfahren zur Bestimmung der Parametern von PI-, PD- und PID-Reglern. *Regelungstechnik*, 28(6):199–205, 1980.
- [136] Tore Hägglund and Karl Johan Åström. Industrial adaptive controllers based on frequency response techniques. *Automatica*, 27:599–609, 1991.
- [137] Tore Hägglund and Karl Johan Åström. Supervision of adaptive control algorithms. *Automatica*, 36:1171–1180, 2000.
- [138] Tore Hägglund and Karl Johan Åström. Revisiting the Ziegler-Nichols tuning rules for PI control. *Asian Journal of Control*, 4(4):364–380, December 2002.
- [139] Tore Hägglund and Karl Johan Åström. Revisiting the Ziegler-Nichols step response method for PID control. *Journal of Process Control*, 14(6):635–650, 2004.
- [140] Tore Hägglund and Karl Johan Åström. Revisiting the Ziegler-Nichols tuning rules for PI control - part II, the frequency response method. *Asian Journal of Control*, 6(4):469–482, 2004.
- [141] Tore Hägglund. *Process Control in Practice*. Chartwell-Bratt Ltd, Bromley, UK, 1991.
- [142] Tore Hägglund. A control-loop performance monitor. *Control Engineering Practice*, 3:1543–1551, 1995.
- [143] Tore Hägglund. An industrial dead-time compensating PI controller. *Control Engineering Practice*, 4:749–756, 1996.
- [144] Tore Hägglund. Automatic detection of sluggish control loops. *Control Engineering Practice*, 7:1505–1511, 1999.
- [145] Tore Hägglund. The Blend station - a new ratio control structure. *Control Engineering Practice*, 9:1215–1220, 2001.
- [146] Tore Hägglund. A friction compensator for pneumatic control valves. *J. of Process Control*, 12:897–904, 2002.
- [147] C.C. Hang and Karl Johan Åström. Relay feedback auto-tuning of process controllers – a tutorial review. *Journal of Process Control*, 12:143–162, 2002.
- [148] Chang C. Hang, Tong H. Lee, and Weng K. Ho. *Adaptive Control*. ISA, Research Triangle Park, North Carolina, 1993.
- [149] C. C. Hang, Karl Johan Åström, and Weng Kuen Ho. Refinements of the Ziegler-Nichols tuning formula. *IEE Proceedings, Part D*, 138(2):111–118, 1991.

Bibliografía

- [150] C. C. Hang, Karl Johan Åström, and W. K. Ho. Relay auto-tuning in the presence of static load disturbance. *Automatica*, 29(2):563–564, 1993.
- [151] C. C. Hang and K. K. Sin. An on-line auto-tuning method based on cross-correlation. *IEEE Transactions on Industrial Electronics*, 38(6):428–437, 1991.
- [152] H. Hanselmann. Implementation of digital controllers—A survey. *Automatica*, 23(1):7–32, 1987. Survey paper.
- [153] D. Hansen, Peter. Robust adaptive PID controller tuning for unmeasured load rejection. In *IFAC Workshop on Digital Control – Past, present, and future of PID Control*, Terrassa, Spain, 2000.
- [154] D. Hansen, Peter. *Adaptive Tuning Methods of the Foxboro I/A System*, chapter Techniques for Adaptive Control, pages 23–54. Elsevier, New York, 2003.
- [155] R. Hanus, M. Kinnaert, and J.-L. Henrotte. Conditioning technique, a general anti-windup and bumpless transfer method. *Automatica*, 23:729–739, 1987.
- [156] R. Hanus. Antiwindup and bumpless transfer: a survey. In *Proceedings of IMACS-88, Part 2*, pages 59–65, Paris, France, 1988.
- [157] P. Harriott. *Process Control*. McGraw-Hill, New York, NY, 1964.
- [158] C. J. Harris and S. A. Billings, editors. *Self-tuning and Adaptive Control: Theory and Applications*. Peter Peregrinus, London, 1981.
- [159] Thomas J. Harris, F. Boudreau, and John F. MacGregor. Performance assessment of multivariable feedback controllers. *Automatica*, 32(11):1505–1518, 1996.
- [160] T.J. Harris, C.T. Seppala, and L.D. Desborough. A review of performance monitoring and assessment techniques for univariate and multivariate control systems. *J. of Process Control*, 9:1–17, 1999.
- [161] T. J. Harris. Assessment of control loop performance. *Canadian Journal of Chemical Engineering*, 67:856–861, 1989.
- [162] D. R. Hartree, A. Porter, A. Callender, and A. B. Stevenson. Time-lag in control systems—II. *Proceedings of the Royal Society of London*, 161:460–476, 1937.
- [163] W. M. Hawk, Jr. A self-tuning, self-contained PID controller. In *Proc. 1983 American Control Conference*, pages 838–842, San Francisco, California, 1983.
- [164] P. Hazebroek and B. L. van der Waerden. Theoretical considerations on the optimum adjustment of regulators. *Trans. ASME*, 72:309–322, 1950.
- [165] H. L. Hazen. Theory of servomechanisms. *JFI*, 218:283–331, 1934.
- [166] R. Hecht-Nielsen. *Neurocomputing*. Addison-Wesley, 1990.
- [167] E. H. Higham. A self-tuning controller based on expert systems and artificial intelligence. In *Proceedings of Control 85*, pages 110–115, England, 1985.

- [168] J. D. Higham. 'Single-term' control of first- and second-order processes with dead time. *Control*, pages 2–6, February 1968.
- [169] D. M. Himmelblau. *Fault Detection and Diagnosis in Chemical and Petrochemical Processes*. Elsevier Scientific, Amsterdam, 1978.
- [170] H Hjalmarsson, M. Gevers, and O Lequin. Iterative feedback tuning: theory and applications. *IEEE Control Systems Magazine*, 18(4):26–41, 1998.
- [171] L. P. Holmblad and J. Østergaard. Control of a cement kiln by fuzzy logic. *F.L. Smidt Review*, 67:3–11, 1981. Copenhagen, Denmark.
- [172] G. Holzbock, Werner. *Automatic Control Theory and Practice*. Reinhold Publishing Company, New York, 1958.
- [173] H. S. Hoopes, W. M. Hawk, Jr., and R. C. Lewis. A self-tuning controller. *ISA Transactions*, 22(3):49–58, 1983.
- [174] A. Horch. *Condition monitoring of control loops*. PhD thesis, Royal Institute of Technology, Stockholm, Sweden, 2000.
- [175] I.M. Horowitz. *Quantitative feedback theory (QFT)*. QFT Publications, Boulder, Colorado, 1993.
- [176] I. M. Horowitz. *Synthesis of Feedback Systems*. Academic Press, New York, 1963.
- [177] G. Howes. Control of overshoot in plastics-extruder barrel zones. In *EI Technology*, No. 3, pages 16–17. Eurotherm International, Brighton, UK, 1986.
- [178] B. Huang and S.L. Shah. *Performance assessment of control loops*. Springer-Verlag, London, 1999.
- [179] Z. Huang. Auto-tuning of PID controllers using neural networks. In *Preprints IFAC International Symposium on Intelligent Tuning and Adaptive Control (ITAC 91)*, Singapore, 1991.
- [180] Mihai Huzmezan, William A. Gough, and Guy Dumont. *Adaptive predictive regulatory control with brainwave*, chapter Techniques for Adaptive Control, pages 99–143. Elsevier, New York, 2003.
- [181] S.-H. Hwang and H.-C. Chang. A theoretical examination of closed-loop properties and tuning methods of single-loop PI controllers. *Chemical Engineering Science*, 42:2395–2415, 1987.
- [182] Ari Ingimundarson and Tore Hägglund. Performance comparison between PID and dead-time compensating controllers. *Journal of Process Control*, 12:887–895, 2002.
- [183] R. Isermann and K. Lachmann. Parameter adaptive control with configuration aids and supervision functions. *Automatica*, 21:623–638, 1985.
- [184] R. Isermann. Practical aspects of process identification. *Automatica*, 16:575–587, 1980.

Bibliografía

- [185] R. Isermann. Parameter adaptive control algorithms—A tutorial. *Automatica*, 18:513–528, 1982.
- [186] A. Ivanoff. Theoretical foundations of the automatic regulation of temperature. *J. Institute of Fuel*, 7:117–138, 1934.
- [187] H. M. James, N. B. Nichols, and R. S. Phillips. *Theory of Servomechanisms*. McGraw-Hill, New York, 1947.
- [188] Rolf Johansson. *System Modeling and Identification*. Prentice Hall, Englewood Cliffs, New Jersey, 1993.
- [189] R. E. Kalman and R. S. Bucy. New results in linear filtering and prediction theory. *Trans ASME (J. Basic Engineering)*, 83 D:95–108, 1961.
- [190] R. E. Kalman. Contributions to the theory of optimal control. *Boletín de la Sociedad Matemática Mexicana*, 5:102–119, 1960.
- [191] R. E. Kalman. New methods and results in linear prediction and filtering theory. Technical Report 61-1, RIAS, February 1961. 135 pp.
- [192] P. Kapasouris and M. Athans. Multivariable control systems with saturating actuators antireset windup strategies. In *Proc. Automatic Control Conference*, pages 1579–1584, Boston, Massachusetts, 1985.
- [193] A. Kaya and S. Titus. A critical performance evaluation of four single loop self-tuning control products. In *Proceedings of the 1988 American Control Conference*, Atlanta, Georgia, 1988.
- [194] I. Kaya and D. P. Atherton. A new PI-PD Smith predictor for control of processes with long time delays. In *Preprints. 14th World Congress of IFAC*, pages 283–288, Beijing, China, 1999.
- [195] C. Kessler. Das symmetrische Optimum, Teil I. *Regelungstechnik*, 6(11):395–400, 1958.
- [196] C. Kessler. Das symmetrische Optimum, Teil II. *Regelungstechnik*, 6(12):432–436, 1958.
- [197] P. J. King and E. H. Mamdani. The application of fuzzy control systems to industrial processes. *Automatica*, 13:235–242, 1977.
- [198] Günter Klefenz. *Automatic Control of Steam Power Plants*. Bibliographisches Institut, third edition, 1986.
- [199] M. Klein, Thomas Marczinkowsky, and M. Pandit. An elementary pattern recognition self-tuning PI-controller. In *Preprints IFAC International Symposium on Intelligent Tuning and Adaptive Control (ITAC 91)*, volume 1, Singapore, 1991.
- [200] B. Kouvaritakis and M. Cannon. *Nonlinear predictive control, theory and practice*. IEE, London, 2001.
- [201] D. Kozub and C.E. Garcia. Monitoring and diagnosis of automated controllers in the chemical process industries. In *AICHE Meeting*, St Louis, MO, 1993.

- [202] D. Kozub and C.E. Garcia. Controller performance monitoring and diagnosis: Experiences and challenges. In *Chemical Process Control V*, pages 83–96, Tahoe City, CA, 1996.
- [203] L. C. Kramer and K. W. Jenkins. A new technique for preventing direct digital control windup. In *Proc. Joint Automatic Control Conference*, pages 571–577, St Louis, Missouri, 1971.
- [204] T. W. Kraus and T. J. Myron. Self-tuning PID controller uses pattern recognition approach. *Control Engineering*, pages 106–111, June 1984.
- [205] N. J. Krikilis. Design of tracking systems subject to actuators and saturation and integrator windup. 39(4):667–682, 1984.
- [206] Birgitta Kristensson. *PID Controllers Design and Evaluation*. PhD thesis, Chalmers, Gothenburg, aug 2003.
- [207] B Kristiansson and B. Lennartsson. Optimal PID controllers including roll off and Smith predictor structure. In *Preprints. 14th World Congress of IFAC*, pages 297–302, Beijing, China, 1999.
- [208] B. Kristiansson and B Lennartsson. Robust and optimal tuning of PI and PID controllers. *IEE Proceedings– Control theory and applications*, 149(1):17–25, 2002.
- [209] R. Kulhavy, J. Lu, and T. Samad. Emerging technologies for enterprise optimization in the process industries. pages 411–422, Tucson, AZ, 2001. Chemical Process Control - 6, Assessment and new directions for research (CPC-VI).
- [210] K. Küpfmüller. Über die Dynamik der selbsttätigen Verstärkungsregler. *ENT*, 5:459–467, 1928.
- [211] C. C. Lee. A self-learning rule-based controller with approximate reasoning and neural nets. In *Preprints 11th IFAC World Congress*, Tallinn, Estonia, 1990.
- [212] A. Leva. PID autotuning algorithm based on relay feedback. *IEE Proceedings D*, 140(5):328–338, 1993.
- [213] Z. Liu. A frequency response based adaptive control for center-driven web winders. In *1998 American Control Conference*, Philadelphia, PA, June 1998.
- [214] L. Ljung and T. Söderström. *Theory and Practice of Recursive Identification*. MIT Press, Cambridge, Massachusetts, 1983.
- [215] L. Ljung. *System Identification—Theory for the User*. Prentice Hall, Englewood Cliffs, New Jersey, 1998. Second Edition.
- [216] Sheldon G. Lloyd and Gerald D. Anderson. *Industrial Process Control*. Fisher Controls Co., Marshalltown, Iowa, 1971.
- [217] L&N. *Leeds & Northrup Technical Journal*. Spring Issue, Number 3, 1968.
- [218] A. M. Lopez, J. A. Miller, C. L. Smith, and P. W. Murrill. Tuning controllers with error-integral criteria. *Instrumentation Technology*, pages 57–62, November 1967.

Bibliografía

- [219] A. M. Lopez, P. W. Murrill, and C. L. Smith. Tuning PI and PID digital controllers. *Instruments and Control Systems*, 42:89–95, February 1969.
- [220] M. P. Lukas. *Distributed Process Control Systems—Their Evaluation and Design*. Van Nostrand Reinhold, New York, 1986.
- [221] William L. Luyben. *Process Modeling, Simulation and Control for Chemical Engineers*. McGraw-Hill, second edition, 1990.
- [222] Joseph Lu. An efficient single-loop MPC algorithm for replacing PID. In *AICHE Annual Conference*, Austin, TX, 2004.
- [223] C. B. Lynch and G. A. Dumont. Control loop performance monitoring. *IEEE Trans. Control Syst. Technol.*, 4:185–192, 1996.
- [224] J. M. Maciejowski. *Multivariable Feedback Design*. Addison-Wesley, Reading, Massachusetts, 1989.
- [225] J. M. Maciejowski. *Predictive control with constraints*. Prentice Hall, Englewood Cliffs, NJ, 2002.
- [226] E. H. Mamdani and S. Assilian. A case study on the application of fuzzy set theory to automatic control. In *Proceedings IFAC Stochastic Control Symposium*, Budapest, Hungary, 1974.
- [227] E. H. Mamdani. Application of fuzzy algorithm for control of simple dynamic plant. *Proc. IEE*, 121:1585–1588, 1974.
- [228] R. J. Mantz and E. J. Tacconi. Complementary rules to Ziegler and Nichols'rules for a regulating and tracking controller. 49:1465–1471, 1989.
- [229] T. E. Marlin. *Process Control*. McGraw-Hill, 2000.
- [230] J. Marsik and V. Strejc. Application of identification-free algorithms for adaptive control. *Automatica*, 25:273–277, 1989.
- [231] S. Marsili-Libelli. Optimal design of PID regulators. 33(4):601–616, 1981.
- [232] C. E. Mason and G. A. Philbrick. Automatic control in the presence of process lags. *Transactions of the ASME*, 62:295–308, 1940.
- [233] M. Matausek and A. Micic. A modified Smith predictor for controlling a process with an integrator and long dead-time. *IEEE Transaction on Automatic Control*, 41:1199–1203, 1996.
- [234] M. Matausek and A. Micic. On the modified Smith predictor for controlling a process with an integrator and long dead-time. *IEEE Transaction on Automatic Control*, 44:1603–1606, 1999.
- [235] J. C. Maxwell. On governors. *Proceedings of the Royal Society of London*, 16:270–283, 1868. Also published in “Mathematical Trends in Control Theory” edited by R. Bellman and R. Kalaba, Dover Publications, New York 1964, pp. 3–17.
- [236] Thomas J. McAvoy. *Interaction Analysis: Principles and Applications*. ISA, Research Triangle Park, North Carolina, 1983.

- [237] Gregory K. McMillan, Willy K. Wojsznis, and Guy T. Borders, Jr. Flexible gain scheduler. In *Advances in Instrumentation and Control*, volume 48 of *ISA Conference*, pages 811–818, 1993.
- [238] Gregory K. McMillan, Willy K. Wojsznis, and Ken Meyer. Easy tuner for DCS. In *Advances in Instrumentation and Control*, volume 48 of *ISA Conference*, pages 703–710, 1993.
- [239] G. K. McMillan. *Tuning and Control Loop Performance*. ISA, Research Triangle Park, North Carolina, second edition, 1983.
- [240] G. K. McMillan. Advanced control algorithms: Beware of false prophecies. *InTech*, pages 55–57, January 1986.
- [241] C. Meyer, D. E. Seborg, and R. K. Wood. A comparison of the Smith predictor and conventional feedback control. *Chemical Engineering Science*, 31:775–778, 1976.
- [242] A. Johnson Michael and Mohammad H. Moradi. *PID Control: New Identification and Design Methods*. Springer, Berlin, 2005.
- [243] J. A. Miller, A. M. Lopez, C. L. Smith, and P. W. Murrill. A comparison of controller tuning techniques. *Control Engineering*, pages 72–75, December 1967.
- [244] N. Minorsky. Directional stability of automatically steered bodies. *J. Amer. Soc. of Naval Engineers*, 34(2):280–309, 1922.
- [245] Charles F. Moore, Cecil L. Smith, and Paul W. Murrill. Improved algorithm for direct digital control. *Instruments & Control Systems*, 43:70–74, January 1970.
- [246] Manfred Morari and Jay H. Lee. Model predictive control: The good, the bad, and the ugly. In *Chemical Process Control, CPCIV*, pages 419–442, Padre Island, TX, 1991.
- [247] Manfred Morari and Evangelos Zafiriou. *Robust Process Control*. Prentice-Hall, Englewood Cliffs, New Jersey, 1989.
- [248] H. M. Morris. How adaptive are adaptive process controllers? *Control Engineering*, 34-3:96–100, 1987.
- [249] C. L. Nachtigal. Adaptive controller performance evaluation: Foxboro EXACT and ASEA Novatune. In *Proceedings ACC-86*, pages 1428–1433, 1986.
- [250] C. L. Nachtigal. Adaptive controller simulated process results: Foxboro EXACT and ASEA Novatune. In *Proceedings ACC-86*, pages 1434–1439, 1986.
- [251] George C. Newton, Jr., Leonard A. Gould, and James F. Kaiser. *Analytical Design of Linear Feedback Controls*. John Wiley & Sons, 1957.
- [252] H. Nicholson, editor. *Modelling of Dynamical Systems, Vol. 1*. Peter Peregrinus, 1980.

Bibliografía

- [253] H. Nicholson, editor. *Modelling of Dynamical Systems, Vol. 2.* Peter Peregrinus, 1981.
- [254] A. Niederlinski. A heuristic approach to the design of linear multivariable interacting control systems. *Automatica*, 7:691–701, 1971.
- [255] Y. Nishikawa, N. Sannomiya, T. Ohta, and H. Tanaka. A method for auto-tuning of PID control parameters. 20:321–332, 1984.
- [256] Harry Nyquist. Regeneration theory. *Bell System Technical Journal*, 11:126–147, 1932. Also published in “Mathematical Trends in Control Theory,” edited by R. Bellman and R. Kalaba, Dover Publications, New York 1964, pp. 83–105.
- [257] Aidan O’Dwyer. *Handbook of PI and PID Controller Tuning Rules.* Imperial College Press, London, 2003.
- [258] Rufus Oldenburg. *Frequency Response.* MacMillan, New York, 1956.
- [259] R. C. Oldenburg and H. Sartorius. A uniform approach to the optimum adjustment of control loops. *Transactions of the ASME*, 76:1265–1279, November 1954.
- [260] W. Oppelt. *Kleines Handbuch technischer Regelvorgänge.* Verlag Chemie, Weinheim, 1964.
- [261] Babatunde A Oquinnaike and W. Harmon Ray. *Process Dynamics, Modeling and Control (Topics in Chemical Engineering).* Oxford University Press, Oxford, 1994.
- [262] J. G. Owen, D Read, H. Blekkenhorst, and A. A. Roche. A mill prototype for automatic monitoring of control loop performance. In *Proc. Control Syst. ’96*, pages 171–178, Halifax, Nova Scotia, Canada, 1996.
- [263] D.J. Pagano. Intelligent tuning of PID controllers based on production rules system. In *Preprints IFAC International Symposium on Intelligent Tuning and Adaptive Control (ITAC 91)*, Singapore, 1991.
- [264] Zalman J. Palmor and Reuel Shinnar. Design of sampled data controllers. *Ind. Eng. Chem. Process Design and Development*, 18(1):8–30, 1979.
- [265] Hélène Panagopoulos, Karl Johan Åström, and Tore Hägglund. Design of PID controllers based on constrained optimisation. *IEE Proc. Control Theory Appl.*, 149(1):32–40, January 2002.
- [266] Hélène Panagopoulos. *PID Control Design, Extension, Application.* PhD thesis, February 2000.
- [267] H. Panagopoulos, K. J. Åström, and T. Hägglund. Design of PI controllers. In *Proc. 1997 IEEE International Conference on Control Applications*, pages 417–422, Hartford, Connecticut, 1997.
- [268] H. Panagopoulos and K. J. Åström. PID control design and H_∞ loop shaping. *Int. J. Robust Nonlinear Control*, 10:1249–1261, 2000.

- [269] H. H. Pao. Use of neural-net technology in control: A survey and a perspective. In *Preprints 11th IFAC World Congress*, Tallinn, Estonia, 1990.
- [270] A. A. Patwardhan, M. N. Karim, and R. Shah. Controller tuning by a least squares method. *AICHE Journal*, 33:1735–1737, October 1987.
- [271] E. Pavlik and B. Machei. *Ein kombiniertes Regelsystem fur die Verfahrensindustrie*. Oldenburg, Munchen, 1960.
- [272] T. J. Pemberton. PID: The logical control algorithm. *Control Engineering*, pages 66–67, May 1972.
- [273] T. J. Pemberton. PID: The logical control algorithm-II. *Control Engineering*, pages 61–63, July 1972.
- [274] Per Persson and Karl Johan Åström. Dominant pole design—A unified view of PID controller tuning. In *Preprints 4th IFAC Symposium on Adaptive Systems in Control and Signal Processing*, pages 127–132, Grenoble, France, July 1992.
- [275] Per Persson and Karl Johan Åström. PID control revisited. In *Preprints IFAC 12th World Congress*, Sydney, Australia, July 1993.
- [276] Per Persson. *Towards Autonomous PID Control*. PhD thesis, Department of Automatic Control, Lund Institute of Technology, Sweden, April 1992.
- [277] B. W. Pessen. How to “tune in” a three mode controller. *Instrumentation*, pages 29–32, Second Quarter 1954.
- [278] Mikael Petersson, Karl-Erik Årzén, and Tore Hägglund. Assessing measurements for feedforward control. In Martins de Carvalho, editor, *European Control Conference - ECC'01*, pages 432–437, Porto, Portugal, September 2001.
- [279] Mikael Petersson, Karl-Erik Årzén, and Tore Hägglund. A comparison of two feedforward control structure assessment methods. *International Journal of Adaptive Control and Signal Processing*, 17(7–9):609–624, August 2003.
- [280] Mikael Petersson, Karl-Erik Årzén, Henrik Sandberg, and Lena de Maré. Implementation of a tool for control structure assessment. In *Proceedings of the 15th IFAC World Congress*, Barcelona, Spain, 2002.
- [281] M. J. G. Polonoyi. PID controller tuning using standard form optimization. *Control Engineering*, pages 102–106, March 1989.
- [282] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mischenko. *The Mathematical Theory of Optimal Processes*. John Wiley, New York, 1962.
- [283] B. Porter, A. H. Jones, and C. B. McKeown. Real-time expert tuners for PI controllers. *IEE Proceedings Part D*, 134(4):260–263, 1987.
- [284] Martin Pottman and Dale E. Seborg. A radial basis function control strategy and its application to a pH neutralization process. In *Proceedings 2nd European Control Conference, ECC '93*, Groningen, The Netherlands, June 1993.

Bibliografía

- [285] S. Joe. Qin and Thomas. A. Badgwell. A survey of industrial model predictive control technology. *Control Engineering Practice*, 11:733–764, 2003.
- [286] S Qin. Control performance monitoring – a review and assessment. *Computers and Chemical Engineering*, 23:173–186, 1998.
- [287] J. Quevedo and T. Escobet. *Digital Control 2000 – Past, present, and future of PID Control*. Pergamon, Oxford, 2000.
- [288] F. Radke and R. Isermann. A parameter-adaptive PID controller with stepwise parameter optimization. 23:449–457, 1987.
- [289] A. B. Rad and P. J. Gawthrop. Explicit PID self-tuning control for systems with unknown time delay. In *IFAC International Symposium ITAC 91 Preprint*, volume 5, Singapore, 1991.
- [290] H. Rake. Step response and frequency response methods. *Automatica*, 16:519–526, 1980.
- [291] James Blake Rawlings and John G Ekerdt. *Chemical Reactor Analysis and Design Fundamentals*. Nob Hill Publishing, WI, 2002.
- [292] J. B. Rawlings. Tutorial overview of model predictive control. *IEEE Control Systems Magazine*, 20, 2000.
- [293] J. Richalet, A. Rault, J. L. Testud, and J. Papon. Model predictive heuristic control: Applications to industrial processes. *Automatica*, 14:413–428, 1976.
- [294] J. Rijnsdorp. Interaction in two-variable control systems for distillation columns – I. *Automatica*, 1:15, 1965.
- [295] J. Rijnsdorp. Interaction in two-variable control systems for distillation columns – II. *Automatica*, 1:29–51, 1965.
- [296] D. E. Rivera, M. Morari, and S. Skogestad. Internal model control—4. PID controller design. *Ind. Eng. Chem. Proc. Dev.*, 25:252–265, 1986.
- [297] C. W. Ross. Evaluation of controllers for deadtime processes. *ISA Transactions*, 16(3):25–34, 1977.
- [298] A. A. Rovira, P. W. Murrill, and C. L. Smith. Tuning controllers for setpoint changes. *Instruments and Control Systems*, pages 67–69, December 1969.
- [299] Lars Rundqvist. Anti-reset windup for PID controllers. In *Preprints 11th IFAC World Congress*, Tallinn, Estonia, 1990.
- [300] Tor Steinar Schei. A method for closed loop automatic tuning of PID controllers. *Automatica*, 28(3):587–591, 1992.
- [301] Tor Steinar Schei. Automatic tuning of PID controllers based on transfer function estimation. *Automatica*, 30(12):1983–1989, 1994.
- [302] Dale E Seborg, Thomas F Edgar, and Duncan A Mellichamp. *Process Dynamics and Control*. Wiley, New York, NY, second edition, 2004.
- [303] D. E. Seborg, T. F. Edgar, and S. L. Shah. Adaptive control strategies for process control: A survey. *AICHE Journal*, 32:881–913, 1986.

- [304] Nancy J. Sell. *Process control fundamentals for the pulp & paper industry*. TAPPI PRESS, 1995.
- [305] J. L. Shearer and Bohdan T. Kulakowski. *Dynamic Modeling and Control of Engineering Systems*. Macmillan, New York, 1990.
- [306] T. Shigemasa, Y. Iino, and M. Kanda. Two degrees of freedom PID auto-tuning controller. In *Proceedings of ISA Annual Conference*, pages 703–711, 1987.
- [307] F. G. Shinskey. Feedforward control applied. *ISA J.*, pages 79–83, March 1963.
- [308] F. G. Shinskey. *Controlling Multivariable Processes*. ISA, Research Triangle Park, North Carolina, 1981.
- [309] F. G. Shinskey. *Process-Control Systems. Application, Design, and Tuning*. McGraw-Hill, New York, third edition, 1988.
- [310] F. G. Shinskey. How good are our controllers in absolute performance and robustness? *Measurement and Control*, 23:114–121, May 1990.
- [311] F. G. Shinskey. Evaluating feedback controllers challenges users and vendors. *Control Engineering*, pages 75–78, September 1991.
- [312] F. G. Shinskey. Model predictors: The first smart controllers. *Instruments and Control Systems*, pages 49–52, September 1991.
- [313] F. G. Shinskey. *Feedback Controllers for the Process Industries*. McGraw-Hill, New York, 1994.
- [314] F. G. Shinskey. *Process-Control Systems. Application, Design, and Tuning*. McGraw-Hill, New York, 4th edition, 1996.
- [315] F. G. Shinskey. PID-deadtime control of distributed processes. *Control Engineering Practice*, 9(11):1177–1183, 2002.
- [316] S. Skogestad and M. Morari. Implications of large RGA-elements on control performance. *Ind. & Eng. Chem. Research*, 26(11):2323–2330, 1987. Also see correction to Eq. 13 in Ind. & Eng. Chem. Research (27:5), 898 (1988).
- [317] S Skogestad and I. Postlethwaite. *Multivariable feedback control: analysis and design*. Wiley, Chichester, UK, 1996.
- [318] S. Skogestad. Simple analytic rules for model reduction and PID controller tuning. *Journal of Process Control*, 13(4):291–309, 2003.
- [319] C. L. Smith, A. B. Corripio, and Jr J. Martin. Controller tuning from simple process models. *Instrumentation Technology*, pages 39–44, December 1975.
- [320] C. L. Smith and P. W. Murrill. A more precise method for tuning controllers. *ISA Journal*, pages 50–58, May 1966.
- [321] C. L. Smith. *Digital Computer Process Control*. Intext Educational Publishers, Scranton, Pennsylvania, 1972.
- [322] Otto J. M. Smith. *Feedback Control Systems*. McGraw-Hill, New York, 1958.

Bibliografía

- [323] O. J. M. Smith. Closed control of loops with dead time. *Chemical Engineering Progress*, 53:217–219, May 1957.
- [324] Torsten Söderström and Petre Stoica. *System Identification*. Prentice-Hall, London, UK, 1989.
- [325] N. Stanfelj, T.E. Marlin, and J.F. MacGregor. Monitoring and diagnosing process control performance: The single-loop case. *Ind. Eng. Chem. Res.*, 32:301–314, 1993.
- [326] Stephanopoulos. *Chemical Process Control. An Introduction to Theory and Practice*. Prentice-Hall, 1984.
- [327] J. T. Stock. Pneumatic process controllers: The ancestry of the proportional-integral-derivative controller. *Trans. of the Newcomen Society*, 59:15–29, 1988.
- [328] V. Strejc. Näherungsverfahren für Aperiodische Übertragscharakteristiken. *Regelungstechnik*, 7(7):124–128, 1959.
- [329] Nobuhide Suda et al. *PID Control*. Asakura Shoten Co., Ltd., Japan, 1992.
- [330] M. Sugeno, editor. *Industrial Applications of Fuzzy Control*. Elsevier Science Publishers BV, The Netherlands, 1985.
- [331] G. A. Sullivan. Adaptive control with expert system based supervisory functions. *Journal of Systems Science*, 27(9):839–850, September 1996.
- [332] R. W. Swiniarski. Neuromorphic self-tuning PID controller uses pattern recognition approach. In *Preprints IFAC International Symposium on Intelligent Tuning and Adaptive Control (ITAC 91)*, Singapore, 1991.
- [333] H. Taguchi and M. Araki. Two-degree-of-freedom PID controllers—their functions and optimal tuning. In *IFAC Workshop on Digital Control – Past, present, and future of PID Control*, Terrassa, Spain, 2000.
- [334] Y. Takahashi, M. J. Rabins, and D. M. Auslander. *Control and Dynamic Systems*. Addison-Wesley, Reading, MA, 1972.
- [335] H. Takatsu, Takafumi Kawano, and K. Kitano. Intelligent self-tuning PID controller. In *Preprints IFAC International Symposium on Intelligent Tuning and Adaptive Control (ITAC 91)*, Singapore, 1991.
- [336] L.-Y. Tan and T. W. Weber. Controller tuning of a third-order process under proportional-integral control. *Industrial & Engineering Chemistry Process Design and Development*, 24:1155–1160, 1985.
- [337] N. F. Thornhill and T. Hägglund. Detection and diagnosis of oscillation in control loops. *Control Engineering Practice*, 5:1343–1354, 1997.
- [338] N. F. Thornhill, M. Oettinger, and P. Fedenczuk. Refinery-wide control loop performance assessment. *Journal of Process Control*, 9:109–124, 1999.
- [339] Michael M. Tiller. *Introduction to Physical Modeling with Modelica*. Kluwer, 2001.
- [340] R. M. Tong. A control engineering review of fuzzy system. *Automatica*, 13:559–569, 1977.

- [341] R. M. Tong. A retrospective view of fuzzy control systems. *Fuzzy Sets and Systems*, 14:199–210, 1984.
- [342] J. Truxal. *Automatic Feedback Control System Synthesis*. McGraw-Hill, New York, 1955.
- [343] G. K. Tucker and D. M. Wills. *A Simplified Technique for Control System Engineering*. Minneapolis-Honeywell Regulator Company. Brown Instruments Division, Philadelphia, PA, 1960.
- [344] George Turnbull. Three-term control in EI. 6:3–7, 1988. Published by Eurotherm International.
- [345] B. Tyreus. TUNEX – an expert system for controller tuning. Technical report, du Pont, 1987.
- [346] P. M. E. M. van der Grinten. Determining plant controllability. *Control Engineering*, pages 87–89, October 1963.
- [347] Vance J. Van Doren. *Techniques for Adaptive Control*. Elsevier, New York, 2003.
- [348] G. Vinnicombe. *Uncertainty and Feedback: \mathcal{H}_∞ loop-shaping and the μ -gap metric*. Imperial College Press, London, 2000.
- [349] A. Voda and I. D. Landau. A method for the auto-calibration of PID controllers. *Automatica*, 31(2), 1995.
- [350] J. Vyshnegradskii. Sur la théorie générale des régulateurs. *Compt. Rend. Acad. Sci. Paris*, 83:318–321, 1876.
- [351] K. S. Walgama and J. Sternby. Inherent observer property in a class of anti-windup compensators. 52(3):705–724, 1990.
- [352] Anders Wallén, Karl Johan Åström, and Tore Hägglund. Loop-shaping design of PID controllers with constant Ti/Td ratio. *Asian Journal of Control*, 4(4):403–409, 2002.
- [353] Anders Wallén. *Tools for Autonomous Process Control*. PhD thesis, Department of Automatic Control, Lund Institute of Technology, Sweden, February 2000.
- [354] Liuping Wang and William R. Cluett. *From Plant Data to Process Control: Ideas for Process Identification and PID Design*. Taylor & Francis, London, 2000.
- [355] Li-Xin Wang. *Adaptive Fuzzy Systems and Control: Design and Stability Analysis*. Prentice Hall, 1994.
- [356] Quing-Guo Wang, Chang Chieh Hang, Tore J. Hägglund, and Kok Kiong Tan. *Advances in PID Control*. Springer, Berlin, 2000.
- [357] Quing-Guo Wang, Ahang Yu, and Chiu Min-Sen. Non-interacting control design for multivariable industrial processes. *Journal of Process Control*, 13:253–265, 2003.

Bibliografía

- [358] Joseph C. Webb. Representative DDC systems. *Instruments & Control Systems*, 40:78–83, October 1967.
- [359] P. E. Wellstead. *Introduction to Physical System Modelling*. Academic Press, 1979.
- [360] M. J. Willis, C. Di Massimo, G. A. Montague, M. T. Tham, and A. J. Morris. Artificial neural networks in process engineering. *IEE Proceedings D*, 138(3):256–266, 1991.
- [361] D. M. Wills. A guide to controller tuning. *Control Engineering*, pages 93–95, August 1962.
- [362] D. M. Wills. Tuning maps for three-mode controllers. *Control Engineering*, pages 104–108, April 1962.
- [363] W. A. Wolfe. Controller settings for optimum control. *Transactions of the ASME*, 64:413–418, 1951.
- [364] S. K. P. Wong and D. E. Seborg. Control strategy for single-input single-output non-linear systems with time delays. 48(6):2303–2327, 1988.
- [365] Shigehiko Yamamoto and Iori Hashimoto. Present status and future needs: The view from Japanese industry. In *Chemical Process Control CPCIV*, pages 1–28, Padre Island, TX, 1991.
- [366] Shigehiko Yamamoto. Industrial developments in intelligent and adaptive control. In *Preprints IFAC International Symposium on Intelligent Tuning and Adaptive Control (ITAC 91)*, Singapore, 1991.
- [367] W. H. Yarber. Electromax V plus, A logical progression. In *Proceedings, Control Expo 84*, 1984.
- [368] W. H. Yarber. Single loop, self-tuning algorithm applied. In *Preprints AIChE Anaheim Symposium*, 1984.
- [369] R. E. Young, R. B. Bartusiak, and R. B. Fontaine. Evolutions of an industrial nonlinear model predictive controller. pages 399–410, Tucson, AZ, 2001. Chemical Process Control - 6, Assessment and new directions for research (CPC-VI).
- [370] Minta Yuwana and Dale E. Seborg. A new method for on-line controller tuning. *AIChE Journal*, 28(3):434–440, 1982.
- [371] Zhang Yuzu, Q. G Wang, and Karl Johan Åström. Dominant pole placement for multi-loop control systems. *Automatica*, 38:1213–1220, 2002.
- [372] L. A. Zadeh and C. A. Desoer. *Series in Systems Science*. McGraw-Hill, Inc., USA, 1963.
- [373] L. A. Zadeh. Fuzzy logic. *IEEE Computer*, pages 83–93, April 1988.
- [374] C. Zervos, P. R. Bélanger, and G. A. Dumont. On PID controller tuning using orthonormal series identification. *Automatica*, 24(2):165–175, 1988.
- [375] C. Zhang and R. J. Evans. Rate constrained adaptive control. 48(6):2179–2187, 1988.

- [376] J.C. Zhou, J.C. Doyle, and K. Glover. *Robust and optimal control*. Prentice Hall, 1996.
- [377] M Zhuang and D. P. Atherton. Optimal PID controller setting using integral performance criteria. *Proc. ACC*, pages 3042–3043, 1991.
- [378] J. G. Ziegler and N. B. Nichols. Optimum settings for automatic controllers. *Trans. ASME*, 64:759–768, 1942.
- [379] J. G. Ziegler and N. B. Nichols. Process lags in automatic-control circuits. *Transactions of the ASME*, 65(5):433–443, 1943.