



TESIS DOCTORAL



CONTROL PREDICTIVO GENERALIZADO
DE PROCESOS INDUSTRIALES:
FORMULACIONES APROXIMADAS

Carlos Bordóns Alba

Sevilla, octubre de 1994



TESIS DOCTORAL



CONTROL PREDICTIVO GENERALIZADO
DE PROCESOS INDUSTRIALES:
FORMULACIONES APROXIMADAS

C. Bordóns

E.F. Camacho

Autor: Carlos Bordóns Alba

Director: Eduardo Fernández Camacho

T. 123

Resumen



TESIS DOCTORAL



**CONTROL PREDICTIVO GENERALIZADO
DE PROCESOS INDUSTRIALES:
FORMULACIONES APROXIMADAS**

por

Carlos Bordóns Alba

Ingeniero Industrial por la E.T.S. de I.I.

de la Universidad de Sevilla

Presentada en la

Escuela Técnica Superior de Ingenieros Industriales

de la

Universidad de Sevilla

para la obtención del

Grado de Doctor Ingeniero Industrial

Sevilla, octubre de 1994

ESCUELA TÉCNICA SUPERIOR DE INGENIEROS
INDUSTRIALES

UNIVERSIDAD DE SEVILLA
SECRETARÍA GENERAL

Queda registrada esta Tesis Doctoral
al folio 25 número 99 del libro
correspondiente.

Sevilla, 18 OCT. 1994

El Jefe del Negociado de Tesis,

F. Lero Raffill

A Ana

Agradecimientos

Deseo expresar mi más sincero agradecimiento a todas aquellas personas que han contribuido a la realización de esta tesis, mencionando especialmente al profesor Eduardo Fernández Camacho, bajo cuya dirección ha sido realizada, por lo que he aprendido de él y por su continuo apoyo, y a todos los compañeros y amigos del Departamento de Ingeniería de Sistemas y Automática por la ayuda tanto técnica como personal que me han prestado.

Resumen

Esta tesis se centra en los problemas asociados a la realización práctica de los Controladores Predictivos. La complejidad matemática necesaria para su implantación así como los problemas asociados a la robustez son aspectos determinantes para su aplicación en la práctica industrial. Se abordan estos problemas formulando uno de los controladores más populares (Control Predictivo Generalizado, GPC) de una forma sencilla de realizar y de ajustar, válida para la mayoría de los procesos industriales, estudiando su robustez desde distintos enfoques.

Esta tesis quiere contribuir a llenar el vacío existente entre la forma empírica en que se tiende a usar el Control Predictivo en la práctica industrial y las potentes (aunque a veces enrevesadas) técnicas desarrolladas por los investigadores. Se pretende mostrar cómo el empleo de una técnica de control tan poderosa no requiere necesariamente complejos algoritmos de control.

En primer lugar se lleva a cabo una revisión histórica del desarrollo del Control Predictivo, presentando la problemática asociada a esta disciplina. Posteriormente se muestran las principales características del Control Predictivo Basado en Modelo así como los elementos comunes a todas las técnicas de control que hacen uso de esta idea, poniendo especial énfasis en el Control Predictivo Generalizado y los métodos afines que presentan resultados sobre estabilidad.

A continuación se formula el GPC para una amplia gama de procesos industriales, mostrando cómo esta técnica puede ser programada con un conjunto limitado de instrucciones, disponibles en cualquier sistema de control distribuido, y que el tiempo de cálculo empleado, incluso para la sintonización, es realmente pequeño. El método para realizar el GPC se basa en el hecho de que una amplia clase de procesos industriales pueden describirse mediante unos pocos parámetros, y que se puede obtener un conjunto de funciones simples que relacionen los coeficientes del GPC con los del proceso a controlar. Usando estas funciones, la realización de un GPC resulta casi tan simple como la de un PID. El método está desarrollado para la gran mayoría de procesos industriales, incluyendo sistemas con efecto integral, donde algunos de los controladores predictivos más empleados en la industria no pueden ser aplicados. Se presentan algunos ejemplos que muestran la facilidad de implantación y los buenos resultados obtenidos.

El concepto de robustez resulta primordial para la implantación con éxito de cualquier controlador en la práctica. Se analiza pues la robustez que el método presentado posee, estudiando el comportamiento cuando se tienen en cuenta las discrepancias entre el proceso real y el modelo utilizado, es decir, en presencia de incertidumbres. Los modelos simplificados de los procesos que se han utilizado hacen si cabe más necesario

este estudio, de forma que se garanticen resultados válidos no sólo para el modelo aproximado, sino para una zona alrededor de éste en la que presumiblemente se encuentra el proceso real. El estudio se aborda considerando por una lado las incertidumbres estructuradas y por otro las no estructuradas. En el primer caso se obtienen unos límites en las variaciones de los parámetros que describen la incertidumbres, presentando una zona segura en las proximidades del proceso nominal. En el segundo caso se lleva a cabo un estudio en frecuencia, obteniendo una familia de posibles procesos en un entorno del nominal para la cual el controlador es robusto.

Por último se presentan aplicaciones del método que abarcan un amplio campo dentro de los procesos industriales y se pueden considerar como un banco de pruebas del controlador. Se incluye aquí la aplicación al control adaptativo de una Planta Solar, obteniendo resultados reales sobre un proceso difícil de controlar. También se aplica a un proceso típico en la industria como es un evaporador donde se corroboran algunos de los resultados obtenidos sobre robustez. Se termina con la aplicación a una Planta Piloto de laboratorio, con el controlador implantado sobre un sistema de control distribuido comercial.

Se finaliza la Tesis exponiendo la conclusiones más importantes y proponiendo futuras líneas de investigación en el campo del Control Predictivo.

Índice general

Resumen de la Tesis	vii
Lista de figuras	xix
Lista de tablas	xxv
1 Introducción	1
1.1 Desarrollo del control predictivo	2
1.2 Objetivos de la tesis	4
2 Control predictivo basado en modelo	7
2.1 Características del control predictivo basado en modelo	7
2.2 Estrategia del MPC	9
2.3 Elementos del MPC	11
2.3.1 Modelo de predicción	12
2.3.2 Función objetivo	16
2.3.3 Obtención de la ley de control	19

2.4	Revisión de algunos algoritmos	20
2.5	Control Predictivo basado en la Respuesta Impulsional	24
2.5.1	Modelo del proceso y predicción	25
2.5.2	Ley de control	26
2.6	Control Predictivo Generalizado	27
2.6.1	Introducción	27
2.6.2	Formulación del Control Predictivo Generalizado	28
2.6.3	Consideración de ruidos coloreados	33
2.7	Control Predictivo Restringido de Horizonte Deslizante	34
2.7.1	Introducción	34
2.7.2	Cálculo de la ley de control	34
2.7.3	Propiedades	37
2.8	GPC estable	38
2.8.1	Introducción	38
2.8.2	Desarrollo de la ley de control	39
2.9	Control predictivo multivariable	42
2.10	GPC con restricciones	44
3	Realización simple del GPC para procesos industriales	47
3.1	Modelo del proceso	48
3.1.1	Identificación del proceso: el método de la curva de reacción . . .	49

3.2	Procesos con retardo múltiplo del tiempo de muestreo	51
3.2.1	Modelo discreto del proceso	51
3.2.2	Formulación del problema	52
3.2.3	Cálculo de los coeficientes del controlador	55
3.2.4	Ponderación del esfuerzo de control	57
3.2.5	Algoritmo de Control	58
3.2.6	Un ejemplo	59
3.3	Procesos con retardo no múltiplo del tiempo de muestreo	62
3.3.1	Modelo discreto del proceso	62
3.3.2	Coefficientes del controlador	65
3.3.3	Ejemplo	68
3.4	Procesos con efecto integral	70
3.4.1	Obtención de la ley de control	71
3.4.2	Coefficientes del controlador	72
3.4.3	Ejemplo de aplicación	74
3.5	Consideración de referencias en rampa	77
3.5.1	Aplicación	79
4	Robustez frente a incertidumbres estructuradas	83
4.1	Incetidumbres	83
4.1.1	Incetidumbres paramétricas	85

4.1.2	Dinámica inmodelada	88
4.1.3	Ambos tipos de incertidumbres	92
4.2	Estabilidad en presencia de incertidumbres estructuradas	92
4.2.1	Desarrollo del predictor	93
4.2.2	Ecuación característica	95
4.2.3	Resolución de la ecuación característica	98
4.3	Influencia de incertidumbres paramétricas	99
4.3.1	Incertidumbre en el polo	100
4.3.2	Incertidumbre en la ganancia	101
4.3.3	Incertidumbre en m	101
4.3.4	Consideración de todas las incertidumbres paramétricas	102
4.4	Influencia de dinámica inmodelada	103
4.4.1	Polo inmodelado	103
4.4.2	Error en estimación del retardo	107
4.4.3	Ambos factores	109
4.5	Efecto conjunto	111
4.6	Influencia del parámetro λ	112
5	Robustez frente a incertidumbres no estructuradas	115
5.1	Incertidumbres no estructuradas	115
5.1.1	Descripción de la incertidumbres	115

5.1.2	Descripción del proceso	117
5.2	Criterios de robustez	119
5.2.1	Herramientas de diseño robusto	119
5.2.2	Medida de la robustez del GPC	121
5.3	Límites de robustez	124
5.3.1	Estructura de control	125
5.3.2	Obtención de los límites de robustez	126
5.3.3	Influencia del parámetro λ	132
5.4	Relación con incertidumbres estructuradas	133
5.4.1	Incertidumbre en la ganancia	134
5.4.2	Incertidumbre en el retardo	136
5.4.3	Comparación de resultados	139
5.5	Mejora de la robustez mediante polinomios de ponderación	140
5.5.1	Elección del polinomio T	141
5.5.2	Relación con otras formulaciones	142
6	Aplicaciones	143
6.1	Control de una planta solar	143
6.1.1	Descripción de la planta	143
6.1.2	Estrategia de control	145
6.1.3	Resultados	146

6.2	Control de la composición en un evaporador	146
6.2.1	Descripción del proceso	146
6.2.2	Obtención del modelo lineal	151
6.2.3	Diseño del controlador	153
6.2.4	Pruebas	154
6.3	Planta Piloto	158
6.3.1	Descripción de la planta	159
6.4	Control de la planta	162
6.4.1	Control de caudal	162
6.4.2	Control de la temperatura a la salida del intercambiador	165
6.4.3	Control de la temperatura en el depósito	167
6.4.4	Control del nivel del depósito	168
6.4.5	Comentarios	170
7	Conclusiones	173
A	Sistema de control distribuido Integral 2000	175
A.1	Generalidades	175
A.2	Descripción del PMC	177
A.2.1	Funcionamiento	177
A.2.2	Configuración de la memoria	178

A.2.3 Programación de las secuencias	178
A.3 El supervisor	180

Lista de figuras

2.1	Estrategia del Control Predictivo	10
2.2	Estructura básica del MPC	11
2.3	Respuesta impulsional y ante escalón	13
2.4	Trayectoria de referencia	18
2.5	Constrained Receding-Horizon Predictive Control	35
2.6	Estabilización del bucle	39
2.7	Configuración del controlador	42
3.1	Respuesta del sistema	49
3.2	Curva de reacción	50
3.3	Esquema de control	54
3.4	Coefficientes del controlador	55
3.5	Influencia de la ponderación del esfuerzo de control	58
3.6	Evolución de la temperatura de salida ante un escalón	60
3.7	Código en lenguaje C del programa de control	62

3.8	Respuesta del proceso ante cambios en la referencia y en λ	63
3.9	Esquema de control	65
3.10	Parámetros l_{y1} , l_{y2} y l_{u1} aen función de a y m	66
3.11	Funcionamiento con cambios en las condiciones de operación	70
3.12	Coefficientes l_{y1} , l_{y2} , l_{y3} y l_{u1}	74
3.13	Evolución del motor ante cambios en la referencia	76
3.14	Evolucion del motor con freno electromagnético	77
3.15	Evolucion de la salida ante referencia variable	81
4.1	Incertidumbres en los parámetros del sistema	87
4.2	Respuesta ante escalón para sistemas con distintos polos	89
4.3	Respuesta en frecuencia con polo inmodelado	90
4.4	Respuesta en frecuencia para diferentes retardos	91
4.5	Predictor	93
4.6	Diagrama de bloques de la ley de control	96
4.7	Ley de control con predictor	97
4.8	Diagrama de bloques simplificado	97
4.9	Límites superior e inferior de incertidumbre en el polo	101
4.10	Límites superiores de incertidumbre en la ganancia para $d=5$ y $d=10$	102
4.11	Zona estable con incertidumbres en polo y ganancia para retardo 10	103
4.12	Límites de α para distintos valores de γ y del retardo	104

4.13	Influencia del parámetro m en la estabilidad para distintos valores de k y del retardo	106
4.14	Polo inmodelado para distintos retardos	107
4.15	Comportamiento con $k=1$	108
4.16	Influencia en la estabilidad de los errores en el retardo	109
4.17	Efecto conjunto de polo inmodelado y error en el retardo	110
4.18	Efecto conjunto de incertidumbres paramétricas y dinámica inmodelada	112
4.19	Influencia del parámetro λ en los límites de robustez	113
5.1	Representación de las incertidumbres en diagrama de bloques	117
5.2	Discrepancia entre el proceso nominal y el real	117
5.3	Familia \mathcal{F} para distintas frecuencias	118
5.4	Tratamiento de la incertidumbres paramétricas como no estructuradas .	120
5.5	Proceso y controlador con realimentación unitaria	122
5.6	Estabilidad robusta	123
5.7	Equivalente discreto	124
5.8	Diagrama de bloques del GPC	125
5.9	Forma característica de los límites \bar{l}_a y \bar{l}_m	127
5.10	Diagrama polar del proceso \tilde{G} y de $\tilde{G}C$ mostrando los límites para una determinada frecuencia	128
5.11	Límite \bar{l}_a para retardo 1 y diversos valores del polo	129
5.12	Límite \bar{l}_a para $m=0$ y diversos retardos	130

5.13	Límite \bar{l}_a para $m=0.5$ y diversos retardos	131
5.14	Límite \bar{l}_a para $m=1$ y diversos retardos	132
5.15	Influencia del retardo en \bar{l}_a	133
5.16	Influencia del parámetro λ en \bar{l}_a	134
5.17	Tratamiento de la variación de la ganancia como incertidumbre no estructurada	135
5.18	Límite de estabilidad para $a=0.8$, $m=1$ y $d=5$	136
5.19	Límite de estabilidad para $a=0.9$, $m=1$ y $d=10$	137
5.20	Incertidumbre en el retardo	137
5.21	Zona estable a determinada frecuencia	138
5.22	Procesos con retardos mayores que el nominal	139
5.23	Procesos con retardos mayores y menores que el nominal	140
6.1	Esquema de la planta solar	144
6.2	GPC adaptativo: temperatura de salida, caudal y radiación solar	147
6.3	Esquema del evaporador	148
6.4	Respuesta del sistema ante escalón a la entrada	153
6.5	Funcionamiento del GPC en el evaporador	155
6.6	Comparación de GPC y PI para un cambio en la referencia	156
6.7	Comparación de GPC y PI para cambios en la alimentación	157
6.8	Influencia de errores en la estimación de la ganancia y el retardo	158

6.9	Influencia del error en la estimación del retardo	159
6.10	Planta Piloto	160
6.11	Esquema general de la planta	161
6.12	Control de caudal	164
6.13	Código del GPC en lenguaje ITER	166
6.14	Comportamiento del intercambiador de calor	167
6.15	Evolución de la temperatura en el depósito	169
6.16	Evolución del nivel en el depósito	170
A.1	Sistema de control INTEGRAL 2000	176

Lista de Tablas

3.1	Parámetros del sistema para distintas condiciones de operación	68
3.2	Coefficientes del controlador	69
3.3	Coefficientes que aparecen en la ley de control en cada caso. El \times indica que el coeficiente existe	80
6.1	Variables del proceso y valores en el punto de operación	149

Capítulo 1

Introducción

Se puede considerar que en la actualidad el objetivo de un sistema de control consiste en actuar sobre las variables manipuladas de forma que puedan satisfacerse múltiples y cambiantes criterios de funcionamiento (económicos, de seguridad, medioambientales o de calidad) en presencia de cambios en las características del proceso.

El amplio abanico de metodologías actuales de control de procesos se enfrenta al cumplimiento de este objetivo. La diferencia entre las diversas técnicas radica básicamente en los compromisos hechos en la formulación matemática de los criterios de funcionamiento y en la elección de la forma de representar el proceso. La representación matemática de muchos de estos criterios se lleva a cabo en la forma de funciones objetivo dinámicas y de restricciones mientras que el proceso se representa como un modelo dinámico con incertidumbres asociadas. La importancia de las incertidumbres está siendo cada vez más reconocida y por tanto incluida explícitamente en la formulación de los controladores.

Las técnicas de Control Predictivo Basado en Modelo (Model Based Predictive Control, MPC) parecen constituir unas poderosas herramientas para afrontar estos retos. MPC, en su forma más general, acepta cualquier tipo de modelos, funciones objetivo o restricciones, siendo la metodología que actualmente puede reflejar más directamente los múltiples criterios de funcionamiento relevantes en la industria. Quizás sea ésta la principal razón del éxito de estas técnicas en numerosas aplicaciones de la industria de procesos, unida a que es la forma más general de formular el problema de control en el dominio del tiempo, de manera que puede resultar fácil de aceptar por el personal de la industria.

El Control Predictivo constituye un campo muy amplio de métodos de control de-

sarrollados en torno a ciertas ideas comunes e integra diversas disciplinas como control óptimo, control estocástico, control de procesos con tiempos muertos, control multivariable o control con restricciones.

1.1 Desarrollo del control predictivo

A partir de finales de los años 1970s aparecieron distintos artículos que mostraban el incipiente interés de la industria por el MPC, principalmente las publicaciones de Richalet *et al.* [77], [76] presentando el Control Predictivo Heurístico Basado en Modelo (Model Predictive Heuristic Control, MPHc), posteriormente conocido también como Control Algorítmico Basado en Modelo (Model Algorithmic Control MAC) y de Cutler y Ramaker [26] con Control con Matriz Dinámica (Dynamic Matrix Control, DMC). En ambos algoritmos, se usa explícitamente un modelo dinámico del proceso (respuesta impulsional en el primero de ellos y ante escalón en el segundo) para predecir el efecto de las acciones de control futuras en la salida, las cuales son determinadas minimizando el error predicho sujeto a restricciones de operación. La optimización se repite en cada instante de muestreo con información actualizada del proceso. Estas formulaciones eran de naturaleza heurística y algorítmica e intentaban aprovechar el creciente potencial de los computadores digitales por aquella época.

Estos procedimientos están íntimamente relacionados con el problema de control óptimo en tiempo mínimo y con programación lineal [93]. Una de las ideas centrales del MPC, el horizonte deslizante, fue propuesta ya por Propoi en 1963 [71], enmarcada en el *open-loop optimal feedback* (realimentación óptima en bucle abierto), problema sobre el que se trabajó extensamente en los años setenta.

Rápidamente el MPC adquirió gran popularidad en las industrias de procesos químicos principalmente debido a la simplicidad del algoritmo y al uso del modelo de respuesta impulsional o en escalón, que aunque posea muchos más parámetros que las formulaciones en el espacio de estados o entrada/salida suele ser preferido por ser intuitivo y necesitar menos información *a priori* para identificar. En [33] se puede encontrar una completa relación de aplicaciones en el sector petroquímico durante la década de los ochenta. La mayoría de estas aplicaciones fueron llevadas a cabo sobre sistemas multivariados incluyendo restricciones. A pesar de su éxito, estas formulaciones carecían de teorías formales que proporcionaran resultados de estabilidad y robustez; de hecho el caso de horizonte finito parecía demasiado complicado de analizar salvo en casos muy específicos.

Independientemente fue surgiendo otra línea de trabajo en torno a las ideas del

control adaptativo, desarrollando estrategias esencialmente para procesos monovariantes formuladas con modelos entrada/salida (que aunque conceptualmente sencillas de extrapolar al caso multivariable no lo son tanto en la práctica). Se pueden incluir aquí el Control Autoajutable Basado en Predictor (Predictor-based Self-tuning Control) de Peterka [68], diseñado para minimizar, para los valores predichos más recientes, el valor esperado de un criterio cuadrático en un horizonte de control dado (finito o asintóticamente infinito) o el Control Adaptativo de Horizonte Extendido (Extended Horizon Adaptive Control EHAC) de Ydstie [91]. Este último método intenta mantener la salida futura (calculada mediante una ecuación diofántica) próxima a la referencia en un instante de tiempo posterior al retardo del proceso y permite distintas estrategias según la señal de control se considere constante en un determinado intervalo o se quiera minimizar su valor o su incremento. El Control Adaptativo de Predicción Extendida (Extended Prediction Self Adaptive Control, EPSAC) debido a De Keyser *et al.* [42] propone una señal de control constante a partir del instante actual¹ mientras que usa un predictor subóptimo en lugar de resolver una ecuación diofántica. En este contexto también aparece el Control Predictivo Generalizado (Generalized Predictive Control, GPC) presentado por Clarke *et al.* en 1987 [23], el cual tomando ideas del Controlador de Mínima Varianza Generalizado (Generalized Minimum Variance, GMV) [20] es quizás el método más popular en la actualidad y será objeto de un estudio detallado en los capítulos siguientes.

Existen numerosas formulaciones de controladores predictivos basadas en las mismas ideas comunes, entre las que se pueden incluir: "Multistep Multivariable Adaptive Control (MUSMAR) [38]", "Multipredictor Receding Horizon Predictive Control (MURHAC) [54]", "Predictive Functional Controller (PFC) [75]" o "Unified Predictive Control (UPC) [85]".

La formulación del MPC también puede llevarse a cabo en el espacio en estados [59]. Este hecho permite no sólo el uso de teoremas bien conocidos de la teoría de espacios de estados, sino que también facilita la generalización a casos más complejos como sistemas con perturbaciones estocásticas y ruido en las medidas. Extendiendo el modelo de respuesta ante escalón y usando las técnicas conocidas de estimación de estados, se pueden tratar también los procesos con integradores. Las técnicas de estimación del estado provenientes del control óptimo estocástico pueden ser usadas para la predicción sin añadir complicación adicional [51]. Esta perspectiva conduce a reglas simples de sintonización para estabilidad y robustez: el controlador MPC puede ser interpretado como un compensador basado en un observador del estado y su estabilidad, prestaciones y robustez son determinadas por los polos del observador (que pueden ser determinados directamente por un nuevo parámetro ajustable) y del regulador (determinados por los horizontes, ponderaciones, etc). Un análisis de de las características inherentes a todos

¹Nótese que debido al horizonte deslizante la señal real no tiene por qué mantenerse constante.

los algoritmos de MPC (y especialmente del GPC) desde el punto de vista de la teoría lineal óptimo cuadrática gaussiana puede ser encontrado en [6].

Aunque los primeros trabajos sobre el GPC (Mohtadi [58]) probaron algunos teoremas específicos sobre estabilidad usando relaciones con el espacio de estados y estudiaron la influencia de los polinomios de filtrado sobre la mejora de la robustez, había que admitir que la falta de resultados generales sobre estabilidad en los controladores de horizonte finito era un serio inconveniente. A causa de esto, en los noventa apareció una nueva línea de trabajo en nuevos métodos de control predictivo con estabilidad garantizada. Dos métodos: Control Predictivo Restringido de Horizonte Deslizante (Constrained Receding Horizon Predictive Control, CRHPC) (Clarke y Scattolini [21]) y Control Estabilizador de Entrada/Salida y Horizonte Deslizante (Stabilizing Input/Output Receding Horizon Control, SIORHC) (Mosca *et al.* [61]) se desarrollaron independientemente y demostraron la estabilidad imponiendo restricciones de igualdad en la salida tras un horizonte finito. Con el mismo objetivo, Kouvaritakis *et al.* [46] presentaron el GPC *estable* (SGPC), una formulación que garantiza la estabilidad del bucle cerrado gracias a la estabilización del proceso previamente a la minimización de la función objetivo.

También se han obtenido importantes resultados en un problema complicado de antemano como el de la estabilidad de controladores de horizonte deslizante con restricciones [73], [95],[78]. Campo y Morari [17] y Allwright [2] han obtenido nuevos resultados usando diseño robusto. La idea central es tener en cuenta las incertidumbres en el proceso de forma explícita y diseñar el MPC optimizando la función objetivo para la situación más desfavorable de las incertidumbres. Estos resultados permiten aventurar que el MPC experimentará aún un mayor auge tanto en el mundo académico como en la práctica industrial.

1.2 Objetivos de la tesis

A pesar de las diversas aplicaciones llevadas a cabo tanto en el mundo industrial como en centros de investigación, el Control Predictivo no ha alcanzado aún en la industria la popularidad que su potencial permite augurar.

Aunque puede decirse que en determinados campos (industria aeroespacial por ejemplo) se han ido incorporando nuevas estrategias de control avanzado (sobre todo los llamados *métodos modernos* basados en la descripción interna de los sistemas), no ha ocurrido lo mismo en los procesos industriales, donde el impacto de las nuevas técnicas de control ha sido limitado. De hecho, resulta paradójico que la mayoría de

los modernos sistemas de control por computador combinen los más recientes avances en campos como Electrónica, Informática o Telecomunicaciones, con algoritmos de control desarrollados en los años cuarenta y cincuenta. De hecho, no ha existido una transición suave de las técnicas ampliamente aceptadas y probadas (PID, predictor de Smith, desacoplo...) a las técnicas avanzadas, quizás debido a la falta de comunicación entre la comunidad académica y el personal de la industria.

Por otro lado, la implantación de determinadas técnicas requiere en ocasiones complejos cálculos matemáticos que, aunque no constituyen problema alguno en la comunidad científica, representan un serio impedimento para su puesta en práctica por personal de la industria.

Otro tema que ha impedido una amplia extensión a la industria de las últimas técnicas de control es la robustez de los controladores. Este es un tema fundamental en cualquier aplicación práctica y que ha estado prácticamente ausente de la literatura de control desde los años sesenta hasta casi los ochenta. Ha sido necesario esperar hasta casi finales de los setenta para considerar la robustez como un objetivo de diseño; es necesario contar con los errores de modelado o incertidumbres del proceso de forma que el sistema real cumpla los requerimientos de control a pesar de diferir en cierto grado del modelo.

Teniendo en cuenta las consideraciones anteriores, esta tesis pretende contribuir a llenar el vacío entre los desarrollos teóricos y las aplicaciones prácticas en Control Predictivo. Para ello se pretende:

- Estudiar los diferentes desarrollos de controladores predictivos existentes para encontrar una formulación de los algoritmos de forma simplificada que permita una fácil aplicación a la mayoría de los procesos industriales. Esta formulación debe ser intuitiva de forma que el controlador pueda ser implantado y usado por personal sin excesivos conocimientos. Debe requerir poca carga de cálculo, de manera que pueda ser programado en cualquier sistema de control por simple que éste sea con la misma facilidad que un PID digital. Su puesta en marcha y sintonización deben ser además sencillas, similares a la de los métodos de control más aceptados en la industria.

En definitiva se busca un desarrollo que combine la potencia de los controladores predictivos (concretamente el Control Predictivo Generalizado) con la simplicidad y facilidad de sintonización de los tradicionales PIDs. El desarrollo se llevará a cabo para los procesos más comunes de la industria proporcionando unas reglas de ajuste similares a las ampliamente empleadas de Ziegler-Nichols [96].

- Estudiar la robustez de la formulación propuesta. Se pretende estudiar la robustez

del controlador desarrollado para modelos sencillos cuando actúa en presencia de incertidumbres estructuradas y no estructuradas. Se buscan resultados que garanticen cierta confianza en el controlador aunque existan diferencias entre el modelo nominal y el proceso real. No se busca un diseño robusto, sino cuantificar la robustez que el sistema posee.

- Realizar diversas aplicaciones prácticas de las formulaciones realizadas, que permitan comprobar cómo una herramienta tan poderosa como es el Control Predictivo puede ser aplicada en cualquier sistema de control comercial sin la necesidad de complejos algoritmos de control.

Capítulo 2

Control predictivo basado en modelo

En este capítulo se van a describir las principales características del Control Predictivo Basado en Modelo así como los elementos comunes a todos los controladores que utilizan esta idea, mostrando las diversas alternativas que se emplean en las distintas formulaciones. Se revisarán algunos de los métodos más conocidos, haciendo hincapié en el método de la respuesta impulsional y en el Control Predictivo Generalizado. Este último, así como los métodos derivados de él que presentan resultados sobre estabilidad serán estudiados con mayor detalle. Por último, se tratan brevemente dos casos de gran trascendencia en la industria: procesos multivariables y existencia de restricciones.

2.1 Características del control predictivo basado en modelo

El Control Predictivo Basado en Modelo, Model (Based) Predictive Control (MBPC ó MPC), no es una estrategia de control específica, sino que se trata más bien de un campo muy amplio de métodos de control desarrollados en torno a ciertas ideas comunes. Estos métodos de diseño conducen a controladores lineales que poseen prácticamente la misma estructura y presentan suficientes grados de libertad. Las ideas que aparecen en mayor o menor medida en toda la familia de controladores predictivos son básicamente:

- Uso explícito de un modelo para predecir la salida del proceso en futuros instantes de tiempo (horizonte).
- Cálculo de la señales de control minimizando una cierta función objetivo.
- Estrategia deslizante, de forma que en cada instante el horizonte se va desplazando hacia el futuro, lo que implica aplicar la primera señal de control en cada instante y desechar el resto, repitiendo el cálculo en cada instante de muestreo.

Los distintos algoritmos de MPC (también llamado Long-Range Predictive Control o LRPC) difieren entre sí casi exclusivamente en el modelo usado para representar el proceso y los ruidos y en la función de coste a minimizar. Aunque las diferencias puedan parecer pequeñas a priori, pueden provocar distintos comportamientos en bucle cerrado, siendo críticas para el éxito de un determinado algoritmo en una determinada aplicación.

Es un tipo de control de naturaleza abierta dentro del cual se han desarrollado muchas realizaciones, encontrando gran aceptación tanto en aplicaciones industriales como en el mundo académico. En la actualidad existen numerosas aplicaciones de controladores predictivos funcionando con éxito, tanto en la industria de procesos como en control de motores o Robótica. Por ejemplo en [18] se presentan aplicaciones en la industria cementera, en torres de secado o en brazos de robot, mientras que en [74], [77] se presentan desarrollos para columnas de destilación, plantas de PVC, generadores de vapor o servos. El buen funcionamiento de estas aplicaciones muestra la capacidad del MPC de conseguir sistemas de control de elevadas prestaciones capaces de operar sin apenas intervención durante largos períodos de tiempo.

El MPC presenta una serie de ventajas sobre otros métodos, entre las que destacan:

- resulta particularmente atractivo para personal sin un conocimiento profundo de control, puesto que los conceptos resultan muy intuitivos, a la vez que la sintonización es relativamente fácil.
- puede ser usado para controlar una gran variedad de procesos, desde aquéllos con dinámica relativamente simple hasta otros más complejos incluyendo sistemas con grandes retardos, de fase no mínima o inestables.
- puede tratar con facilidad el caso multivariable.
- posee intrínsecamente compensación del retardo.
- introduce de forma natural control *feedforward* para compensación de perturbaciones medibles.

- resulta conceptualmente simple la extensión al tratamiento de restricciones, que pueden ser incluídas de forma sistemática durante el proceso de diseño.
- es muy útil cuando se conocen las futuras referencias (robótica o procesos en *batch*).
- es una metodología completamente abierta basada en algunos principios básicos que permite futuras extensiones.

Pero, lógicamente, también presenta inconvenientes. Unos de ellos es la carga de cálculo necesaria para la resolución de algunos algoritmos. Aunque hoy día esto no es un problema esencial, hay que considerar que muchos computadores industriales no están a la última en cuanto a potencia de cálculo y, sobre todo, que la mayor parte del tiempo disponible en el computador de proceso debe ser empleada para otras funciones distintas del propio algoritmo de control (comunicación con el operario, alarmas, registros, etc). Aún así el mayor inconveniente viene marcado por la necesidad de disponer de un modelo apropiado del proceso. El algoritmo de diseño está basado en el conocimiento previo del modelo y es independiente de éste, pero está claro que las prestaciones obtenidas dependerán de las discrepancias existentes entre el proceso real y el modelo usado.

En la práctica, el MPC ha demostrado ser una estrategia razonable para control industrial, a pesar de la escasez de resultados teóricos en algunos puntos como la estabilidad o robustez. Quizás lo que proporciona mayor potencia al método sea la introducción de restricciones, lo cual augura la aparición de nuevas formulaciones en un futuro próximo.

2.2 Estrategia del MPC

La metodología de todos los controladores pertenecientes a la familia del MPC se caracteriza por la estrategia siguiente, representada en la figura 2.1:

1. En cada instante t y haciendo uso del modelo del proceso se predicen las futuras salidas para un determinado horizonte N , llamado horizonte de predicción. Estas salidas predichas, $\hat{y}(t+k | t)$ ¹ para $k = 1 \dots N$ dependen de los valores conocidos hasta el instante t (entradas y salidas pasadas) y de las señales de control futuras $u(t+k | t)$, $k = 0 \dots N - 1$ que se pretenden mandar al sistema y que son las que se quieren calcular.

¹la notación indica el valor de la variable en el instante $t+k$ calculado en el instante t .

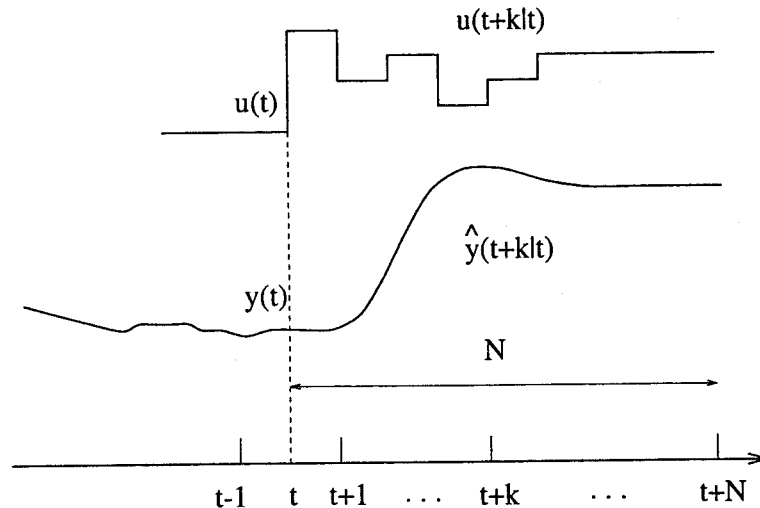


Figura 2.1: Estrategia del Control Predictivo

2. El conjunto de señales de control futuras se calcula optimizando un determinado criterio en el que se pretende mantener el proceso lo más próximo posible a la trayectoria de referencia $w(t+k)$ (que puede ser directamente el set-point o una suave aproximación a éste). Este criterio suele tomar la forma de una función cuadrática de los errores entre la salida predicha y la trayectoria de referencia también predicha, incluyendo en muchos casos el esfuerzo de control. Si el criterio es cuadrático, el modelo lineal y no existen restricciones se puede obtener una solución analítica, en otro caso se debe usar un método iterativo de optimización. Adicionalmente se hace alguna suposición sobre la estructura de la ley de control futura, como por ejemplo que va a ser constante a partir de cierto instante.
3. La señal de control $u(t | t)$ es enviada al proceso mientras que las siguientes señales de control calculadas son desechadas, puesto que en el siguiente instante de muestreo ya se conoce $y(t+1)$ y se repite el paso 1 con este nuevo valor y todas las secuencias son actualizadas. Se calcula por tanto el $u(t+1 | t+1)$ (que en principio será diferente al $u(t+1 | t)$ al disponer de nueva información), haciendo uso del concepto de horizonte deslizante.

Para llevar a cabo esta estrategia, se usa una estructura como la mostrada en la figura 2.2. Se hace uso de un modelo para predecir las salidas futuras del proceso, basándose en las futuras señales de control propuestas. Estas señales son calculadas por el optimizador teniendo en cuenta la función de coste (donde aparece el futuro error de seguimiento) así como las restricciones. Por tanto el modelo juega un papel decisivo

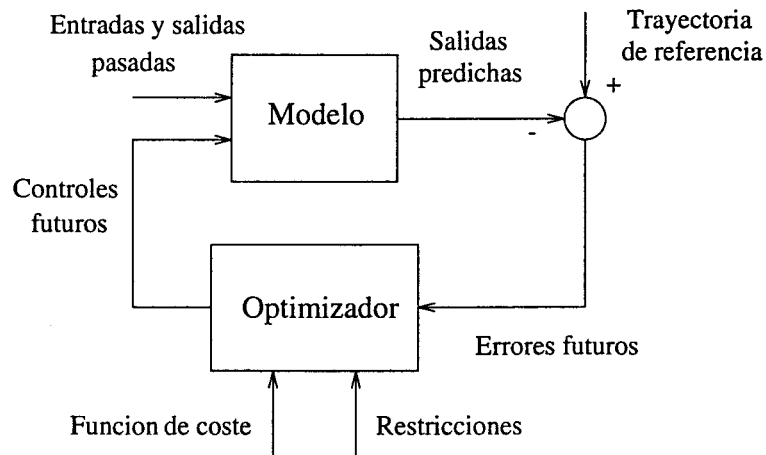


Figura 2.2: Estructura básica del MPC

en el controlador. El modelo elegido debe ser capaz de capturar la dinámica del proceso para poder predecir las salidas futuras al mismo tiempo que debe ser sencillo de usar y de comprender.

El optimizador es otra parte fundamental de la estrategia pues proporciona las acciones de control. Si la función de coste es cuadrática, el mínimo puede ser obtenido como una función de las entradas y salidas pasadas y de la trayectoria de referencia. Sin embargo, cuando existen restricciones de desigualdad la solución debe ser calculada por métodos numéricos con más carga de cálculo.

2.3 Elementos del MPC

Todos los algoritmos presentados poseen elementos comunes y para cada uno de estos elementos se pueden elegir diversas opciones, dando lugar pues a distintos algoritmos. Estos elementos son:

- Modelo de predicción
- Función objetivo
- Obtención de la ley de control

2.3.1 Modelo de predicción

La piedra angular del MPC es el modelo; un diseño completo debe incluir los mecanismos necesarios para la obtención del mejor modelo posible, el cual debe ser lo suficientemente rico para capturar al máximo la dinámica del proceso y debe ser capaz de permitir el cálculo de las predicciones a la vez que sea intuitivo y permita un análisis teórico. El uso del modelo del proceso viene determinado por la necesidad del cálculo de la salida predicha en instantes futuros $\hat{y}(t+k | t)$. Las diferentes estrategias de MPC pueden usar distintos modelos para representar la relación de las salidas con las entradas medibles, algunas de las cuales serán variables manipuladas y otras se pueden considerar como perturbaciones medibles, que pueden ser compensadas por acción *feedforward*. Además se tendrá en cuenta un modelo de las perturbaciones, para intentar describir el comportamiento que no aparece reflejado en el modelo del proceso, englobándose aquí el efecto de las entradas no medibles, el ruido y los errores de modelado.

Para el estudio se puede separar el modelo en dos partes: el modelo del proceso propiamente dicho y el modelo de las perturbaciones. Cualquier método usará ambas partes para la predicción.

Modelo del Proceso

Casi todas las formas posibles de modelar un proceso aparecen en alguna formulación de MPC siendo las más usadas las siguientes:

- Respuesta impulsional. También conocido por secuencia de ponderación o modelo de convolución, aparece en el MAC y como caso especial también en el GPC y el EPSAC. La salida viene relacionada con la entrada por la ecuación

$$y(t) = \sum_{i=1}^{\infty} h_i u(t-i)$$

donde h_i son los valores muestreados obtenidos al someter al proceso a un impulso unitario de amplitud igual al período de muestreo (ver figura 2.3a). Esta suma es truncada y sólo se consideran N valores (por tanto sólo permite representar procesos estables y sin integradores), teniendo

$$y(t) = \sum_{i=1}^N h_i u(t-i) = H(z^{-1})u(t) \quad (2.1)$$

donde $H(z^{-1}) = h_1 z^{-1} + h_2 z^{-2} + \dots + h_N z^{-N}$. Un inconveniente de este método es el gran número de parámetros que necesita, ya que N suele ser un valor elevado

(del orden de 40-50). La predicción vendrá dada por:

$$\hat{y}(t+k | t) = \sum_{i=1}^N h_i u(t+k-i | t) = H(z^{-1})u(t+k | t)$$

Este método es ampliamente aceptado en la práctica industrial debido a que es muy intuitivo pues refleja claramente la influencia de un determinado fenómeno en una determinada salida. Nótese que si el proceso es multivariable, las diferentes salidas reflejarán el efecto de las M entradas de la forma:

$$y_j(t) = \sum_{k=1}^M \sum_{i=1}^N h_i^{kj} u^k(t-i)$$

Una gran ventaja de este método es que no requiere información previa sobre el proceso, con lo que el procedimiento de identificación se simplifica, a la vez que permite describir fácilmente dinámicas complejas como fase no mínima o retardos.

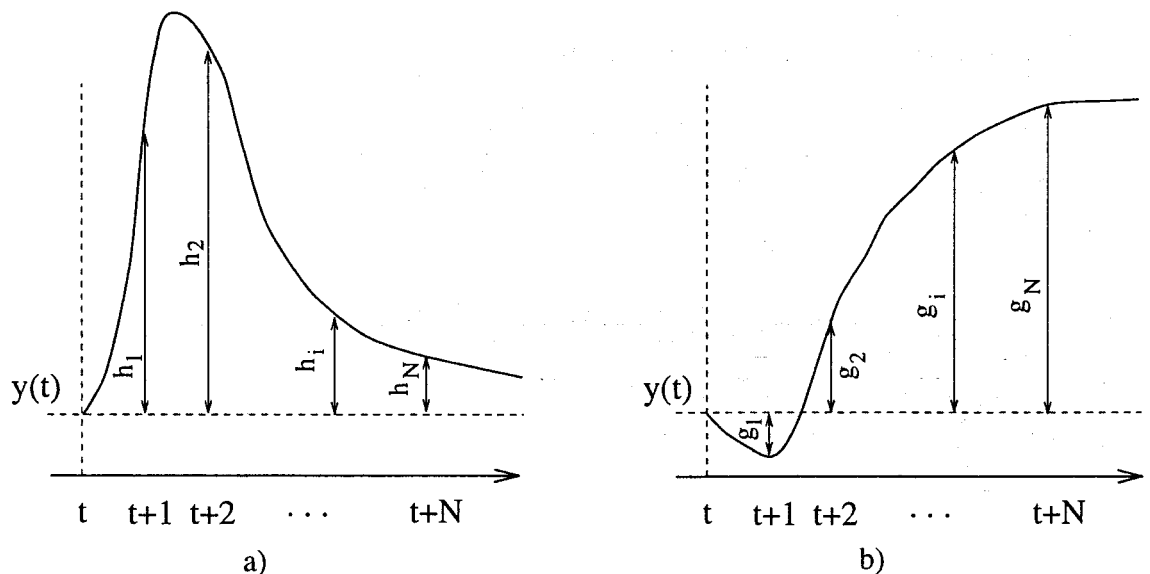


Figura 2.3: Respuesta impulsional y ante escalón

- Respuesta ante escalón. Usado por el DMC y sus variantes, es muy similar al anterior sólo que ahora la señal de entrada es un escalón. Para sistemas estables se tiene la respuesta truncada que será

$$y(t) = y_0 + \sum_{i=1}^N g_i \Delta u(t-i) = y_0 + G(z^{-1})(1-z^{-1})u(t) \quad (2.2)$$

donde las g_i son los valores muestreados ante la entrada en escalón y $\Delta u(t) = u(t) - u(t-1)$, según se muestra en la figura 2.3b. El valor de y_0 puede tomarse 0 sin pérdida de generalidad, con lo cual el predictor será:

$$\hat{y}(t+k | t) = \sum_{i=1}^N g_i \Delta u(t+k-i | t)$$

Este método presenta las mismas ventajas e inconvenientes que el anterior.

- **Función de transferencia.** Usado por GPC, UPC, EPSAC, EHAC, MUSMAR o MUR-HAC entre otros, utiliza el concepto de función de transferencia $G = B/A$ con lo que la salida viene dada por:

$$A(z^{-1})y(t) = B(z^{-1})u(t)$$

$$\begin{aligned} A(z^{-1}) &= 1 + a_1 z^{-1} + a_2 z^{-2} + \dots + a_{na} z^{-na} \\ B(z^{-1}) &= b_1 z^{-1} + b_2 z^{-2} + \dots + b_{nb} z^{-nb} \end{aligned}$$

Por tanto la predicción vendrá dada por

$$\hat{y}(t+k | t) = \frac{B(z^{-1})}{A(z^{-1})} u(t+k | k)$$

Esta representación es válida también para procesos inestables y posee la ventaja de necesitar pocos parámetros, aunque es fundamental un conocimiento *a priori* del proceso sobre todo en cuanto al orden de los polinomios A y B .

- **Espacio de estados.** Usado por ejemplo en el PFC, tiene la siguiente representación:

$$\begin{aligned} x(t) &= Ax(t-1) + Bu(t-1) \\ y(t) &= Cx(t) \end{aligned}$$

siendo x el estado y A , B y C las matrices del sistema, de entrada y de salida respectivamente. Para este modelo la predicción viene dada por [4]

$$\hat{y}(t+k | t) = C\hat{x}(t+k | t) = C[A^k x(t) + \sum_{i=1}^k A^{i-1} B u(t+k-i | t)]$$

Posee la ventaja de que sirve para multivariable a la vez que permite analizar la estructura interna del proceso (aunque a veces los estados obtenidos al discretizar no tienen ningún significado físico). Los cálculos pueden ser complicados, con la necesidad adicional de incluir un observador si los estados no son accesibles.

- Otros. También existen modelos no lineales para representar el proceso aunque el problema de su uso radica en que dan lugar a que el problema de optimización sea más complejo. Otras formas de representación son las redes neuronales [87] y la lógica borrosa [84], que han sido usadas en diferentes aplicaciones.

Modelo de las perturbaciones

De tanta importancia como la elección de un determinado modelo del proceso es la elección del modelo utilizado para representar la perturbaciones. Un modelo bastante extendido es el Autorregresivo Integrado de Media Nula (Auto-Regressive and Integrated Moving Average, ARIMA), en el que las perturbaciones, es decir las diferencias entre la salida medida y la calculada por el modelo vienen dadas por

$$n(t) = \frac{C(z^{-1})e(t)}{D(z^{-1})}$$

donde el polinomio $D(z^{-1})$ incluye explícitamente el integrador $\Delta = 1 - z^{-1}$, $e(t)$ es un ruido de media cero y normalmente el polinomio C se considera igual a uno. Este modelo se considera apropiado para dos tipos de perturbaciones: cambios aleatorios ocurridos en instantes aleatorios (por ejemplo cambio en la calidad del material) y "Brownian motion" (en procesos con balance de energía) y es usado directamente en el GPC, EPSAC, EHAC o UPC y con ligeras variaciones en otros métodos. Nótese que al incluir un integrador se consigue un control con error nulo en régimen permanente (*offset-free*).

Haciendo uso de la ecuación diofántica

$$1 = E_k(z^{-1})D(z^{-1}) + z^{-k}F_k(z^{-1}) \quad (2.3)$$

se tiene

$$n(t) = E_k(z^{-1})e(t) + z^{-k}\frac{F_k(z^{-1})}{D(z^{-1})}e(t) \quad n(t+k) = E_k(z^{-1})e(t+k) + F_k(z^{-1})n(t)$$

y la predicción será

$$\hat{n}(t+k|t) = F_k(z^{-1})n(t) \quad (2.4)$$

Si se combina la ecuación (2.4) con un modelo de función de transferencia (como el usado en el GPC), haciendo $D(z^{-1}) = A(z^{-1})(1 - z^{-1})$ se puede obtener la predicción de la salida:

$$\begin{aligned} \hat{y}(t+k|t) &= \frac{B(z^{-1})}{A(z^{-1})}u(t+k|t) + F_k(z^{-1})(y(t) - \frac{B(z^{-1})}{A(z^{-1})}u(t)) \\ \hat{y}(t+k|t) &= F_k(z^{-1})y(t) + \frac{B(z^{-1})}{A(z^{-1})}(1 - z^{-k}F_k(z^{-1}))u(t+k|t) \end{aligned}$$

y haciendo uso de (2.3) se obtiene la siguiente expresión para el predictor de k pasos

$$\hat{y}(t+k | t) = F_k(z^{-1})y(t) + E_k(z^{-1})B(z^{-1}) \Delta u(t+k | t)$$

Como caso particular del ARIMA se puede incluir la perturbación constante

$$n(t) = \frac{e(t)}{1 - z^{-1}}$$

cuya mejor predicción será $\hat{n}(t+k | t) = n(t)$. Este modelo de perturbación, combinado con el modelo de respuesta ante escalón es el usado en el DMC y métodos afines.

Una extensión de la anterior sería la "drift disturbance" usada en el PFC

$$n(t) = \frac{e(t)}{(1 - z^{-1})^2}$$

siendo la predicción óptima $\hat{n}(t+k | t) = n(t) + c(t)k$. Análogamente otros modelos polinómicos de orden mayor pueden ser usados.

2.3.2 Función objetivo

Los diversos algoritmos de MPC proponen distintas funciones de coste para la obtención de la ley de control. En general se persigue que la salida futura en el horizonte considerado siga a una determinada señal de referencia al mismo tiempo que se puede penalizar el esfuerzo de control requerido para hacerlo. La expresión general de tal función objetivo será:

$$J(N_1, N_2, Nu) = E \left\{ \sum_{j=N_1}^{N_2} \delta(j) [y(t+j | t) - w(t+j)]^2 + \sum_{j=1}^{Nu} \lambda(j) [\Delta u(t+j-1)]^2 \right\} \quad (2.5)$$

En algunos métodos (por ejemplo MAC ó DMC) el segundo sumando, que considera el esfuerzo de control, no es tenido en cuenta, mientras que en otros (UPC) también aparecen directamente los valores de la señal de control (no sus incrementos). En la función de coste se pueden considerar:

- Parámetros: N_1 y N_2 son los horizonte mínimo y máximo de coste (o de predicción) y Nu es el horizonte de control, que no tiene por qué coincidir con el horizonte máximo, como se verá posteriormente. El significado de N_1 y N_2 resulta bastante intuitivo: marcan los límites de los instantes en que se desea que

la salida siga a la referencia. Así, si se toma un valor grande de N_1 es porque no importa que haya errores en los primeros instantes, lo cual provocará una respuesta suave del proceso. Nótese que para procesos con tiempo muerto d no tiene sentido que N_1 sea menor que dicho valor puesto que la salida no empezará a evolucionar hasta el instante $t + d$. Además, si el proceso es de fase no mínima, este parámetro permite eliminar de la función objetivo los primeros instantes de respuesta inversa.

Los coeficientes $\delta(j)$ y $\lambda(j)$ son secuencias que ponderan el comportamiento futuro. Usualmente se consideran valores constantes o secuencias exponenciales. Por ejemplo se puede conseguir un peso exponencial de $\delta(j)$ a lo largo del horizonte usando:

$$\delta(j) = \alpha^{N_2-j}$$

Si α está comprendido entre 0 y 1 indica que se penaliza más a los errores más alejados del instante t que a los más próximos, dando lugar a un control más suave y con menor esfuerzo. Si, por el contrario, $\alpha > 1$ es que se penalizan más los primeros errores, provocando un control más brusco. En el PFC sólo se contabiliza el error en ciertos puntos (puntos de coincidencia); esto se consigue fácilmente en la función objetivo dando a los elementos de la secuencia $\delta(j)$ valor uno en dichos puntos.

Todos estos valores pueden ser usados como parámetros de sintonización, obteniendo un abanico muy amplio de posibilidades con las que se puede cubrir un rango muy amplio de opciones, desde un control *standard* hasta una estrategia diseñada a medida para un proceso en particular.

- Trayectoria de referencia: Una de las ventajas del control predictivo es que si se conoce *a priori* la evolución futura de la referencia, el sistema puede empezar a reaccionar antes de que el cambio se haya efectivamente realizado, evitando los efectos del retardo en la respuesta del proceso. En muchas aplicaciones la evolución futura de la referencia $r(t + k)$ es conocida de antemano, como en Robótica, servos o procesos en *batch*; en otras aplicaciones aunque la referencia sea constante, se puede conseguir una sensible mejora de prestaciones simplemente conociendo el instante de cambio de valor y adelantándose a esa circunstancia.

En el criterio de minimización (2.5), la mayoría de los métodos suelen usar una trayectoria de referencia $w(t + k)$ que no tiene por qué coincidir con la referencia real. Normalmente será una suave aproximación desde el valor actual de la salida $y(t)$ a la referencia conocida mediante un sistema de primer orden:

$$w(t) = y(t) \quad w(t + k) = \alpha w(t + k - 1) + (1 - \alpha)r(t + k) \quad k = 1 \dots N \quad (2.6)$$

α es un parámetro comprendido entre 0 y 1 (mientras más próximo a 1 más suave será la aproximación) que constituye un valor ajustable que influirá en la respuesta dinámica del sistema. En la figura 2.4 se muestra la forma de la

trayectoria cuando la referencia $r(t+k)$ es constante y para dos valores distintos de α ; para valores pequeños de este parámetro se tiene un seguimiento rápido (w_1) mientras que si aumenta la trayectoria de referencia será w_2 dando lugar a una respuesta más suave.

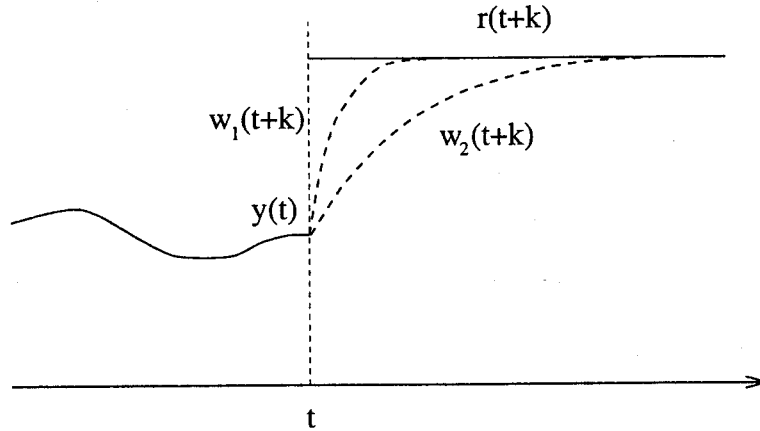


Figura 2.4: Trayectoria de referencia

Otra estrategia es la usada en el PFC, útil para setpoints variables:

$$w(t+k) = r(t+k)\alpha^k(y(t) - r(t))$$

La trayectoria de referencia puede servir para especificar el comportamiento en bucle cerrado; esta idea es usada en el GPC o el EPSAC definiendo una salida auxiliar

$$\psi(t) = P(z^{-1})y(t)$$

El error en la función objetivo vendrá dado ahora por $\psi(t+k) - w(t+k)$. $P(z^{-1})$ tiene ganancia estática unidad y lo que se consigue es la generación de una trayectoria de referencia con la dinámica definida por $1/P(z^{-1})$ y valor inicial el de la salida medida. En [22] se muestra que si se puede conseguir un control *dead-beat* en $\psi(t)$, tal que

$$\psi(t) = B(z^{-1})w(t)$$

siendo $B(z^{-1})$ un determinado polinomio con ganancia unidad, la respuesta en bucle cerrado del proceso será claramente:

$$y(t) = B(z^{-1})/P(z^{-1})w(t)$$

que en definitiva equivale a colocar los polos del bucle cerrado en los ceros del polinomio de diseño $P(z^{-1})$.

- Restricciones: En la práctica, todos los procesos están sujetos a restricciones. Los actuadores tienen un campo limitado de acción así como una determinada velocidad de cambio (*slew rate*), como es el caso de las válvulas, limitadas por las posiciones de totalmente abierta o cerrada y por la velocidad de respuesta. Razones constructivas, de seguridad o medioambientales o bien los propios rangos de los sensores pueden causar límites en las variables de proceso, tales como niveles en depósitos, caudales en tuberías o temperaturas y presiones máximas. Además, normalmente las condiciones de operación vienen definidas por la intersección de ciertas restricciones por motivos fundamentalmente económicos, con lo que el sistema de control operará cerca de los límites. Todo lo expuesto anteriormente hace necesaria la introducción de restricciones en la función a minimizar.

Muchos algoritmos predictivos tienen intrínsecamente en cuenta el tema de las restricciones (MAC, DMC) por lo cual han tenido gran éxito en la industria, mientras que otros las pueden incorporar *a posteriori* (GPC)[9]. Normalmente se considerarán límites en la amplitud y el *slew rate* de la señal de control y límites en las salidas:

$$\begin{aligned} u_{min} &\leq u(t) \leq u_{max} && \forall t \\ du_{min} &\leq u(t) - u(t-1) \leq du_{max} && \forall t \\ y_{min} &\leq y(t) \leq y_{max} && \forall t \end{aligned}$$

con la adición de estas restricciones a la función objetivo, la minimización resulta más compleja, no pudiendo obtenerse la solución analíticamente como en el caso sin restringir.

2.3.3 Obtención de la ley de control

Para obtener los valores $u(t+k | t)$ será necesario minimizar la funcional J de la ecuación (2.5). Para ello se calculan los valores de las salidas predichas $\hat{y}(t+k | t)$ en función de valores pasados de entradas y salidas y de señales de control futuras, haciendo uso del modelo que se haya elegido y se sustituyen en la función de coste, obteniendo una expresión cuya minimización conduce a los valores buscados. Para el criterio cuadrático si el modelo es lineal y no existen restricciones se puede obtener una solución analítica, en otro caso se debe usar un método iterativo de optimización.

De cualquiera de las maneras la obtención de la solución no resulta trivial pues existirán $N_2 - N_1 + 1$ variables independientes, valor que puede ser elevado (del orden de 10 a 30). Con la idea de reducir estos grados de libertad se puede proponer cierta estructura a la ley de control. Además se ha encontrado [43] que esta estructuración de la ley de control produce una mejora en la robustez y en el comportamiento general

del sistema, debido fundamentalmente a que el hecho de permitir la libre evolución de las variables manipuladas (sin estructurar) puede conducir a señales de control de alta frecuencia no deseables y que en el peor de los casos podrían conducir a inestabilidad.

Esta estructura de la ley de control se plasma en el uso del concepto de horizonte de control (Nu), usado en DMC, GPC, EPSAC ó EHAC, que consiste en considerar que tras un cierto intervalo $Nu < N_2$ no hay variación en las señales de control propuestas, es decir:

$$\Delta u(t + j - 1) = 0 \quad j > Nu$$

lo cual es equivalente a dar pesos infinitos a los cambios en el control a partir de cierto instante. El caso límite sería considerar Nu igual a 1 con lo que todas las acciones futuras serían iguales a $u(t)$ ².

Otra forma de estructurar la ley de control es mediante el uso de funciones base, procedimiento usado en el PFC y que consiste en representar la señal de control como combinación lineal de unas funciones base predeterminadas:

$$u(t + k) = \sum_{i=1}^n \alpha_i(t) B_i(k)$$

las B_i se eligen según la naturaleza del proceso y de la referencia, siendo normalmente de tipo polinómico

$$B_0 = 1 \quad B_1 = k \quad B_2 = k^2 \dots$$

Como se indicó previamente, en presencia de restricciones no existe una solución analítica, teniendo que recurrir a métodos de programación cuadrática que conllevan gran carga de cálculo. Sin embargo sí existe una solución analítica cuando se impone la condición de que la salida alcanza el valor de referencia en un determinado instante, método usado en el "Constrained Receding Horizon Predictive Control", (CRHPC) [21], muy cercano al GPC y que garantiza resultados de estabilidad, el cual será tratado más adelante.

2.4 Revisión de algunos algoritmos

A continuación se repasan brevemente algunos de los métodos más populares en orden a mostrar sus características más sobresalientes. Estudios comparativos pueden

²Recuérdese que debido al horizonte deslizante, la señal de control se recalcula en el siguiente muestreo.

encontrarse en [44], [33] y [47]. Se van a tratar aquí brevemente los métodos considerados más representativos: DMC, MAC, EPSAC y EHAC dejándose el GPC para la sección siguiente donde será estudiado con más detalle.

Control con Matriz Dinámica (Dynamic Matrix Control)

El DMC utiliza la respuesta ante escalón (2.2) para modelar el proceso, considerando sólo los N primeros términos, asumiendo por tanto que el proceso es estable y sin integradores. Este valor de N es a su vez el horizonte de predicción, haciendo uso además de un horizonte de control $Nu \leq N$. En cuanto a las perturbaciones se considera que su valor a lo largo de todo el horizonte será el mismo que en el instante t , es decir, será igual al valor medido de la salida menos el postulado por el modelo:

$$\hat{n}(t+k|t) = \hat{n}(t|t) = y_m(t) - \hat{y}(t|t)$$

y por tanto el valor predicho de la salida será

$$\hat{y}(t+k|t) = \sum_{i=1}^k g_i \Delta u(t+k-i) + \sum_{i=k+1}^N g_i \Delta u(t+k-i) + \hat{n}(t+k|t)$$

donde el primer sumando contiene las futuras acciones de control que deben ser calculadas, el segundo contiene los valores pasados de las acciones de control y es por tanto conocido, y el último representa las perturbaciones.

La función de coste puede ser sólo la minimización de los errores futuros o puede incluir también el esfuerzo de control, presentando en tal caso la forma genérica (2.5). Una de las características de este método que le ha proporcionado gran popularidad en la industria es la adición de restricciones, de forma que al criterio de minimización habrá que añadirle ecuaciones de la forma:

$$\sum_{i=1}^N C_{y_i}^j \hat{y}(t+k|t) + C_{u_i}^j u(t+k-i) + c^j \leq 0 \quad j = 1 \dots N_c$$

Siguiendo el comportamiento de los MPCs, la optimización (numérica por la presencia de restricciones) es llevada a cabo en cada instante de muestreo y el valor de $u(t)$ es enviado al proceso.

Los inconvenientes de este método son por un lado las gran dimensión del problema (lo que conlleva gran carga de cálculo) y por otro la incapacidad de poder trabajar con procesos inestables.

Control Algorítmico Basado en Modelo (Model Algorithmic Control)

También conocido por Control Heurístico Predictivo Basado en Modelo (Model Predictive Heuristic Control), cuyo software se llama IDCOM (Identification-Command), es muy similar al método anterior con algunas diferencias.

En primer lugar utiliza un modelo de respuesta impulsional (2.1) válido sólo para procesos estables, en el cual aparece el valor de $u(t)$ en lugar de $\Delta u(t)$. Por otro lado, no hace uso del concepto de horizonte de control con lo que en los cálculos aparecen tantas señales de control como salidas futuras.

Introduce una trayectoria de referencia como sistema de primer orden que evoluciona desde la salida actual hasta el setpoint según una determinada constante de tiempo, siguiendo la expresión (2.6). La varianza del error entre esta trayectoria y la salida es lo que se pretende minimizar en la función objetivo. La estimación de las perturbaciones se lleva a cabo por la siguiente expresión recursiva:

$$\hat{n}(t+k | t) = \alpha \hat{n}(t+k-1 | t) + (1-\alpha)(y_m(t) - \hat{y}(t | t))$$

con $\hat{n}(t | t) = 0$ y $0 \leq \alpha < 1$. α es un parámetro ajustable muy relacionado con el tiempo de respuesta, el ancho de banda y la robustez del sistema en bucle cerrado [33].

También considera restricciones en los actuadores así como en variables internas o salidas secundarias. Diversos algoritmos pueden ser usados para optimizar en presencia de las restricciones, desde el presentado inicialmente por Richalet *et al.* que puede ser usado también para la identificación de la respuesta impulsional, hasta otros que se muestran en [80].

Control Adaptativo de Predicción Extendida (Extended Prediction Self Adaptive Control)

El planteamiento del EPSAC es diferente al de los métodos anteriores. Para predicción el proceso es modelado mediante la función de transferencia

$$A(z^{-1})y(t) = B(z^{-1})u(t-d) + v(t)$$

donde d es el retardo y $v(t)$ la perturbación. El modelo puede ser extendido con un término $D(z^{-1})d(t)$, con $d(t)$ una perturbación medible para incluir efecto *feedforward*. Usando este modelo se obtiene la predicción según se indica en [44].

Una característica del método es que la estructuración de la ley de control es extremadamente simple, ya que se reduce a considerar que la señal de control va a permanecer constante a partir del instante t , es decir, $\Delta u(t+k) = 0$ para $k > 0$. En definitiva:

el horizonte de control se reduce a 1 y por tanto el cálculo se reduce a un solo valor: $u(t)$. Para obtener este valor se usa una función de coste de la forma:

$$\sum_{k=d}^N \gamma(k)[w(t+k) - P(z^{-1})\hat{y}(t+k|t)]^2$$

siendo $P(z^{-1})$ un polinomio de diseño con ganancia estática unidad y el factor $\gamma(k)$ una secuencia de ponderación, análoga a las aparecidas en (2.5).

La señal de control puede ser calculada analíticamente (lo cual representa una ventaja sobre los métodos anteriores) de la forma:

$$u(t) = \frac{\sum_{k=d}^N \alpha_k [w(t+k) - P(z^{-1})\hat{y}(t+k|t)]}{\sum_{k=d}^N \gamma(k)\alpha_k^2}$$

siendo α_k la respuesta impulsional discreta del sistema.

Control Adaptativo de Horizonte Extendido (Extended Horizon Adaptive Control)

El EHAC considera el proceso modelado por su función de transferencia sin tener presente un modelo de la perturbaciones:

$$A(z^{-1})y(t) = B(z^{-1})u(t-d)$$

y pretende minimizar la discrepancia entre el modelo y la referencia en el instante $t+N$: $\hat{y}(t+N|t) - w(t+N)$, siendo $N \geq d$. La solución a este problema no es única (a menos que $N = d$); una posible estrategia es considerar que el horizonte de control es 1, es decir,

$$\Delta u(t+k-1) = 0 \quad 1 < k \leq N-d$$

o minimizar el esfuerzo de control:

$$J = \sum_{k=0}^{N-d} u^2(t+k)$$

Existe una versión incremental del EHAC que permite tratar fácilmente las perturbaciones en la carga, que consiste en considerar

$$J = \sum_{k=0}^{N-d} \Delta u^2(t+k)$$

En esta formulación se usa un predictor de N pasos de la forma

$$\hat{y}(t + N | t) = y(t) + F(z^{-1}) \Delta y(t) + E(z^{-1})B(z^{-1}) \Delta u(t + N - d)$$

siendo $E(z^{-1})$ y $F(z^{-1})$ polinomios que satisfacen la ecuación

$$(1 - z^{-1}) = A(z^{-1})E(z^{-1})(1 - z^{-1}) + z^{-N}F(z^{-1})(1 - z^{-1})$$

siendo el grado de E igual a $N - 1$. Una ventaja de este método es que proporciona una sencilla solución analítica, resultando

$$u(t) = u(t - 1) + \frac{\alpha_0(w(t + N) - \hat{y}(t + N | t))}{\sum_{k=0}^{N-d} \alpha_k^2}$$

con α_k el coeficiente correspondiente a $\Delta u(t + k)$ en la ecuación del predictor.

La ley de control depende pues sólo de los parámetros del proceso y por tanto puede hacerse fácilmente autoajutable si se dispone de un identificador en línea. Como se ve el único parámetro de ajuste es el horizonte de predicción N , lo cual simplifica el uso pero proporciona poca libertad para el diseño.

Se observa que no se puede usar una trayectoria de referencia puesto que sólo se considera el error en un instante $(t + N)$, así como tampoco se pueden ponderar los esfuerzos de control en cada punto, con lo que no se pueden eliminar ciertas frecuencias en la actuación.

2.5 Control Predictivo basado en la Respuesta Impulsional

Quizás la formulación más simple e intuitiva de Control Predictivo sea la basada en las ideas de Richalet *et al.* [77] [76]. Hace uso de la respuesta impulsional truncada del proceso y proporciona una solución analítica simple en ausencia de restricciones. Este método ha sido claramente aceptado en la práctica como lo demuestra la gran cantidad de aplicaciones funcionado actualmente en diversos campos, donde gran parte de su éxito se debe al modelo del proceso utilizado. Como se sabe, los modelos basados en la función de transferencia pueden provocar resultados con grandes errores cuando el orden del modelo no es correcto. En el entorno industrial la estructura del proceso es afectada más frecuentemente por las perturbaciones que por las variables medidas, lo cual exige una actualización continua de los parámetros del modelo. Por otro lado,

la representación mediante la respuesta impulsional es una buena elección pues su identificación es relativamente sencilla.

2.5.1 Modelo del proceso y predicción

La salida del sistema en el instante t está relacionada con las entradas por medio de los coeficientes de la respuesta impulsional de la forma:

$$y(t) = \sum_{j=1}^N h_j u(t-j) = H(z^{-1})u(t)$$

Al considerar la respuesta truncada a N elementos, se asume que el sistema es estable y causal. Usando este modelo, el predictor de k pasos se puede escribir como:

$$\hat{y}(t+k | t) = \sum_{j=1}^N h_j u(t+k-j) + \hat{n}(t+k | t)$$

donde el sumatorio se puede dividir en dos términos:

$$f_r(t+k) = \sum_{j=k+1}^N h_j u(t+k-j) \quad f_o(t+k) = \sum_{j=1}^k h_j u(t+k-j)$$

tales que f_r representa la respuesta libre, que es el valor esperado de $y(t+j)$ considerando acciones futuras de control nulas, y f_o es la respuesta forzada, es decir, el componente adicional de respuesta debido al conjunto propuesto de futuras acciones de control.

Se asume que las perturbaciones se mantendrán constantes en el futuro con el mismo valor que en el instante t , es decir, $\hat{n}(t+k | t) = \hat{n}(t | t)$ que es la salida medida menos la predicha por el modelo nominal:

$$\hat{n}(t+k | t) = \hat{n}(t | t) = y(t) - \sum_{j=1}^N h_j u(t-j)$$

Entonces la predicción viene dada por:

$$\hat{y}(t+k | t) = f_r + f_o + \hat{n}(t | t)$$

Siendo M el horizonte y considerando \mathbf{u}_+ como el vector de las acciones de control propuestas, \mathbf{u}_- el de acciones de control pasadas, \mathbf{y} las salidas predichas, \mathbf{n} las perturbaciones y la trayectoria de referencia \mathbf{w} como una suave aproximación al setpoint actual (ecuación (2.4)):

$$\mathbf{u}_+ = \begin{bmatrix} u(t) \\ u(t+1) \\ \vdots \\ u(t+M+1) \end{bmatrix} \quad \mathbf{u}_- = \begin{bmatrix} u(t-N+1) \\ u(t-N+2) \\ \vdots \\ u(t-1) \end{bmatrix} \quad \mathbf{y} = \begin{bmatrix} \hat{y}(t+1) \\ \hat{y}(t+2) \\ \vdots \\ \hat{y}(t+M) \end{bmatrix}$$

$$\mathbf{n} = \begin{bmatrix} \hat{n}(t+1) \\ \hat{n}(t+2) \\ \vdots \\ \hat{n}(t+M) \end{bmatrix} \quad \mathbf{w} = \begin{bmatrix} w(t+1) \\ w(t+2) \\ \vdots \\ w(t+M) \end{bmatrix}$$

y definiendo las matrices

$$\mathbf{H}_1 = \begin{bmatrix} h_1 & 0 & \cdots & 0 \\ h_2 & h_1 & \cdots & 0 \\ \cdots & \cdots & \ddots & \cdots \\ h_M & h_{M-1} & \cdots & h_1 \end{bmatrix} \quad \mathbf{H}_2 = \begin{bmatrix} h_N & h_{N-1} & \cdots & h_i & \cdots & h_2 \\ 0 & h_N & \cdots & h_j & \cdots & h_3 \\ \cdots & \cdots & \ddots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & h_N & \cdots & h_{M+1} \end{bmatrix}$$

se puede escribir el predictor como:

$$\mathbf{y} = \mathbf{H}_1 \mathbf{u}_+ + \mathbf{H}_2 \mathbf{u}_- + \mathbf{n}$$

2.5.2 Ley de control

La función objetivo minimiza el error así como el esfuerzo de control. Si se expresan los errores futuros como:

$$\mathbf{e} = \mathbf{w} - \mathbf{y} = \mathbf{w} - \mathbf{H}_2 \mathbf{u}_- - \mathbf{n} - \mathbf{H}_1 \mathbf{u}_+ = \mathbf{w} - \mathbf{f} - \mathbf{H}_1 \mathbf{u}_+$$

donde el vector \mathbf{f} contiene los términos que dependen de valores conocidos (entradas pasadas, salida actual y referencias), entonces la función de coste se puede escribir como:

$$J = \mathbf{e}^T \mathbf{e} + \lambda \Delta \mathbf{u}_+^T$$

siendo λ el factor de penalización de la variación de las variables de control. Si no existen restricciones la solución puede ser obtenida analíticamente, proporcionando:

$$\Delta \mathbf{u}_+ = (\mathbf{H}_1^T \mathbf{H}_1 + \lambda \mathbf{I})^{-1} \mathbf{H}_1^T (\mathbf{w} - \mathbf{f}) \quad (2.7)$$

Al ser una estrategia de horizonte deslizante, solamente se usa el primer elemento de este vector, $\Delta u(t)$, desechando el resto y repitiendo los cálculos en el siguiente período de muestreo.

El cálculo de la ley de control (2.7) es relativamente simple comparado con otras formulaciones aunque requiere la inversión de una matriz de dimensión $M \times M$. Nótese que si el número de señales futuras a calcular se elige como un valor $P < M$, entonces la matriz es de dimensión $P \times P$, ya que \mathbf{H}_1 es de dimensión $M \times P$, lo cual reduce los cálculos necesarios.

La simplicidad de este algoritmo así como la posibilidad de incluir restricciones ha convertido esta formulación en una de las usadas con mayor frecuencia en la industria actualmente.

2.6 Control Predictivo Generalizado

2.6.1 Introducción

El Control Predictivo Generalizado GPC fué propuesto por Clarke *et al.* [23] y se ha convertido en uno de los métodos más populares en el ámbito del Control Predictivo tanto en el mundo industrial como en el académico. Se ha empleado con éxito en numerosas aplicaciones industriales, [18], mostrando buenas prestaciones a la vez que un cierto grado de robustez respecto a sobreparametrización o retardos mal conocidos. Puede resolver muchos problemas de control diferentes para un amplio campo de procesos con un número razonable de variables de diseño, que son especificadas por el operario dependiendo del conocimiento previo del proceso y de los objetivos de control. Pero a pesar de este éxito en la práctica, este método adolece de la ausencia de un análisis teórico completo que estudie la influencia de los parámetros de diseño (horizontes, secuencias de ponderación) sobre la estabilidad del bucle cerrado así como de resultados de robustez.

La idea básica del GPC es calcular una secuencia de futuras acciones de control de tal forma que minimice una función de coste multipaso. El índice a minimizar es la esperanza matemática de una función cuadrática que mide por un lado la distancia entre la salida predicha del sistema y una cierta trayectoria de referencia hasta el horizonte de predicción, y por otro el esfuerzo de control necesario para obtener dicha salida. Esta idea ha sido usada por Lelic y Zarrop [53] y Lelic y Wellstead [52] para la obtención de un controlador por asignación de polos generalizado que pertenece a

la clase de los controladores de horizonte extendido y es una extensión de los bien conocidos controladores por asignación de polos.

El Control Predictivo Generalizado tiene muchas ideas en común con otros controladores predictivos previamente mencionados ya que está basado en las mismas ideas pero posee a su vez algunas diferencias. Como se verá más adelante, es capaz de proporcionar una solución analítica (en ausencia de restricciones), puede trabajar con procesos inestables o de fase no mínima e incorpora el concepto de horizonte de control así como la consideración en la función de coste de ponderación de los incrementos en las acciones de control. Las diversas posibilidades disponibles para el GPC conducen a una gran variedad de objetivos de control comparado con otras realizaciones, algunas de las cuales pueden ser consideradas como subconjuntos o casos límites del GPC.

2.6.2 Formulación del Control Predictivo Generalizado

La mayoría de los procesos de una sola entrada y una sola salida (single-input single-output, SISO), al ser considerados en torno a un determinado punto de trabajo y tras ser linealizados, pueden ser descritos de la siguiente forma:

$$A(z^{-1})y(t) = z^{-d}B(z^{-1})u(t-1) + C(z^{-1})e(t) \quad (2.8)$$

donde $u(t)$ y $y(t)$ son respectivamente la señal de control y la salida del proceso y $e(t)$ es un ruido blanco de media cero. A , B y C son los siguientes polinomios en el operador de desplazamiento hacia atrás z^{-1} :

$$\begin{aligned} A(z^{-1}) &= 1 + a_1z^{-1} + a_2z^{-2} + \dots + a_naz^{-na} \\ B(z^{-1}) &= b_0 + b_1z^{-1} + b_2z^{-2} + \dots + b_nbz^{-nb} \\ C(z^{-1}) &= 1 + c_1z^{-1} + a_2z^{-2} + \dots + c_ncz^{-nc} \end{aligned}$$

donde d es el tiempo muerto del sistema.

Este modelo es conocido como Autorregresivo de Media Móvil (Controller Autoregressive Moving-Average CARMA). En muchas aplicaciones industriales en las que las perturbaciones son no-estacionarias resulta más conveniente [23] el uso de un modelo CARMA integrado, dando lugar al CARIMA, que viene descrito por:

$$A(z^{-1})y(t) = B(z^{-1})z^{-d}u(t) + C(z^{-1})\frac{e(t)}{\Delta} \quad \text{con } \Delta = 1 - z^{-1}$$

Por simplicidad, a partir de ahora el polinomio C se va a tomar igual a 1. Nótese que en el caso de que C^{-1} pueda ser truncado se puede absorber en A y B . El caso general de ruido coloreado es tratado más adelante.

El algoritmo del Control Predictivo Generalizado consiste en aplicar una secuencia de señales de control que minimice una función de coste de la forma:

$$J(N_1, N_2, Nu) = E\left\{ \sum_{j=N_1}^{N_2} \delta(j)[\hat{y}(t+j|t) - w(t+j)]^2 + \sum_{j=1}^{Nu} \lambda(j)[\Delta u(t+j-1)]^2 \right\} \quad (2.9)$$

donde $E\{\cdot\}$ es la esperanza matemática e $\hat{y}(t+j|t)$ es la predicción óptima j pasos hacia delante de la salida del proceso con datos conocidos hasta el instante t , N_1 y N_2 son los horizontes mínimo y máximo de coste, Nu es el horizonte de control y $\delta(j)$ y $\lambda(j)$ son las secuencias de ponderación mientras que $w(t+j)$ es la futura trayectoria de referencia, que se puede calcular según se muestra en la figura 2.4. En [23] se considera $\delta(j)$ igual a 1 y $\lambda(j)$ constante.

El objetivo es pues el cálculo de la futura secuencia de control $u(t)$, $u(t+1)$,... de tal manera que la salida futura del proceso $y(t+j)$ permanezca próxima a $w(t+j)$. Esto se logra minimizando $J(N_1, N_2, Nu)$.

Con la intención de minimizar la función de coste, se obtendrá previamente la predicción óptima de $y(t+j)$ para $j \geq N_1$ y $j \leq N_2$. Considérese la siguiente ecuación diofántica:

$$\begin{aligned} 1 &= E_j(z^{-1}) \Delta A + z^{-j} F_j(z^{-1}) \\ 1 &= E_j(z^{-1}) \tilde{A} + z^{-j} F_j(z^{-1}) \end{aligned} \quad (2.10)$$

Los polinomios E_j and F_j están únicamente definidos con grados $j-1$ y na respectivamente. Se pueden obtener dividiendo 1 entre $\tilde{A}(z^{-1})$ hasta que el resto pueda ser factorizado como $z^{-j} F_j(z^{-1})$. El cociente de la división es entonces el polinomio $E_j(z^{-1})$.

Si se multiplica la ecuación (2.8) por $E_j(z^{-1}) z^j$

$$\tilde{A}(z^{-1}) E_j(z^{-1}) y(t+j) = E_j(z^{-1}) B(z^{-1}) \Delta u(t+j-d-1) + E_j(z^{-1}) e(t+j) \quad (2.11)$$

Teniendo en cuenta (2.10), la ecuación (2.11) queda:

$$(1 - z^{-j} F(z^{-1})) y(t+j) = E_j(z^{-1}) B(z^{-1}) \Delta u(t+j-d-1) + E_j(z^{-1}) e(t+j)$$

La cual puede ser escrita como

$$y(t+j) = F(z^{-1}) y(t) + E_j(z^{-1}) B(z^{-1}) \Delta u(t+j-d-1) + E_j(z^{-1}) e(t+j) \quad (2.12)$$

Al ser el grado del polinomio $E_j(z^{-1}) = j - 1$ los términos del ruido en la ecuación (2.12) están todos en el futuro. La mejor predicción de $y(t + j)$ será por consiguiente:

$$\hat{y}(t + j | t) = G_j(z^{-1}) \Delta u(t + j - d - 1) + F(z^{-1})y(t)$$

donde $G_j(z^{-1}) = E_j(z^{-1})B(z^{-1})$

Resulta simple demostrar que los polinomios E_j y F_j se pueden obtener recursivamente, de forma que los nuevos valores en el paso $j + 1$ (E_{j+1} y F_{j+1}) sean función de los del paso j . La recursividad de la ecuación diofántica ha sido demostrada en [23]. A continuación se muestra una demostración más simple. Existen otras formulaciones del GPC que no están basadas en la recursividad de esta ecuación [1].

Considérese que los polinomios E_j y F_j se han obtenido dividiendo 1 entre $\tilde{A}(z^{-1})$ hasta que el resto haya sido factorizado como $z^{-j}F_j(z^{-1})$.

Con:

$$\begin{aligned} F_j(z-1) &= f_{j,0} + f_{j,1}z^{-1} + \dots + f_{j,na}z^{-na} \\ E_j(z-1) &= e_{j,0} + e_{j,1}z^{-1} + \dots + e_{j,j-1}z^{-(j-1)} \end{aligned}$$

Supóngase que se utiliza el mismo procedimiento para obtener E_{j+1} y F_{j+1} , es decir, dividir 1 entre $\tilde{A}(z^{-1})$ hasta que el resto se pueda factorizar como $z^{-(j+1)}F_{j+1}(z^{-1})$ con

$$F_{j+1}(z-1) = f_{j+1,0} + f_{j+1,1}z^{-1} + \dots + f_{j+1,na}z^{-na}$$

Está claro que solamente es necesario dar un paso más en la división para obtener los polinomios E_{j+1} y F_{j+1} . E_{j+1} vendrá dado por:

$$E_{j+1}(z-1) = E_j(z-1) + e_{j+1,j}z^{-j} \quad \text{con } e_{j+1,j} = f_{j,0}$$

Los coeficientes del polinomio F_{j+1} se pueden expresar como:

$$f_{j+1,i} = f_{j,i+1} - f_{j,0} \tilde{a}_{i+1} \quad i = 0 \dots na - 1$$

El polinomio G_{j+1} puede ser obtenido recursivamente como sigue:

$$\begin{aligned} G_{j+1} &= E_{j+1}B = (E_j + f_{j,0}z^{-j})B \\ G_{j+1} &= G_j + f_{j,0}z^{-j}B \end{aligned}$$

Es decir, los primeros j coeficientes de G_{j+1} serán idénticos a los de G_j mientras que el resto viene dado por:

$$g_{j+1,j+i} = g_{j,j+i} + f_{j,0} b_i \quad \text{para } i = 0 \dots nb$$

Para resolver el GPC es necesario obtener el conjunto de señales de control $u(t)$, $u(t+1)$, ..., $u(t+N)$ que minimizan la ecuación (2.9). Al tener el proceso un retardo de d períodos de muestreo, la salida sólo se verá influenciada por la señal $u(t)$ después del instante $d+1$. Los valores N_1 , N_2 y Nu que marcan el horizonte pueden ser definidos como $N_1 = d+1$, $N_2 = d+N$ y $Nu = N$. No tiene sentido hacer $N_1 < d+1$ ya que los términos de (2.9) sólo dependerán de las señales de control pasadas. Por otro lado, haciendo $N_1 > d+1$ los primeros puntos de la secuencia de salida, que serán los mejor estimados, no se tendrán en cuenta.

El conjunto de las j predicciones óptimas:

$$\begin{aligned}\hat{y}(t+d+1|t) &= G_{d+1} \Delta u(t) + F_{d+1}y(t) \\ \hat{y}(t+d+2|t) &= G_{d+2} \Delta u(t+1) + F_{d+2}y(t) \\ &\vdots \\ \hat{y}(t+d+N|t) &= G_{d+N} \Delta u(t+N-1) + F_{d+N}y(t)\end{aligned}$$

puede ser escrito en forma matricial como:

$$\mathbf{y} = \mathbf{G}\mathbf{u} + \mathbf{F}(z^{-1})y(t) + \mathbf{G}'(z^{-1}) \Delta u(t-1) \quad (2.13)$$

Donde

$$\mathbf{y} = \begin{bmatrix} \hat{y}(t+d+1|t) \\ \hat{y}(t+d+2|t) \\ \vdots \\ \hat{y}(t+d+N|t) \end{bmatrix} \quad \mathbf{u} = \begin{bmatrix} \Delta u(t) \\ \Delta u(t+1) \\ \vdots \\ \Delta u(t+N-1) \end{bmatrix}$$

$$\mathbf{G} = \begin{bmatrix} g_0 & 0 & \dots & 0 \\ g_1 & g_0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ g_{N-1} & g_{N-2} & \dots & g_0 \end{bmatrix}$$

$$\mathbf{G}'(z^{-1}) = \begin{bmatrix} G_{d+1}(z^{-1}) - g_0 \\ G_{d+2}(z^{-1}) - g_0 - g_1 z^{-1} \\ \vdots \\ G_{d+N}(z^{-1}) - g_0 - g_1 z^{-1} - \dots - g_{N-1} z^{-(N-1)} \end{bmatrix}$$

$$\mathbf{F}(z^{-1}) = \begin{bmatrix} F_{d+1}(z^{-1}) \\ F_{d+2}(z^{-1}) \\ \vdots \\ F_{d+N}(z^{-1}) \end{bmatrix}$$

Al depender los últimos términos de la ecuación (2.13) sólo del pasado, pueden agruparse en \mathbf{f} , dando lugar a:

$$\mathbf{y} = \mathbf{G}\mathbf{u} + \mathbf{f} \quad (2.14)$$

Entonces la ecuación (2.9) puede escribirse como:

$$J = (\mathbf{G}\mathbf{u} + \mathbf{f} - \mathbf{w})^T (\mathbf{G}\mathbf{u} + \mathbf{f} - \mathbf{w}) + \lambda \mathbf{u}^T \mathbf{u} \quad (2.15)$$

donde:

$$\mathbf{w} = \left[w(t+d+1) \quad w(t+d+2) \quad \dots \quad w(t+d+N) \right]^T \quad (2.16)$$

La ecuación (2.15) se puede poner como:

$$J = \frac{1}{2} \mathbf{u}^T \mathbf{H} \mathbf{u} + \mathbf{b} \mathbf{u} + \mathbf{f}_0 \quad (2.17)$$

donde:

$$\begin{aligned} \mathbf{H} &= 2(\mathbf{G}^T \mathbf{G} + \lambda \mathbf{I}) \\ \mathbf{b} &= 2(\mathbf{f} - \mathbf{w})^T \mathbf{G} \\ \mathbf{f}_0 &= (\mathbf{f} - \mathbf{w})^T (\mathbf{f} - \mathbf{w}) \end{aligned}$$

El mínimo de J , siempre que no existan restricciones en la señal de control, puede ser calculado igualando a cero el gradiente de J , lo cual conduce a:

$$\mathbf{u} = -\mathbf{H}^{-1} \mathbf{b}^T \quad (2.18)$$

Debido al uso de la estrategia deslizante, sólo el primer elemento del vector \mathbf{u} es aplicado realmente, repitiendo de nuevo el mismo procedimiento al siguiente instante de muestreo. La solución propuesta involucra la inversión (o al menos la triangularización) de una matriz de dimensión $N \times N$, lo cual conlleva una gran carga de cálculo. El concepto ya usado en otros métodos de horizonte de control es empleado en [23] con la finalidad de reducir la cantidad de cálculo, asumiendo que las señales de control permanecerán en un valor constante a partir del intervalo $Nu < N$. Por tanto la dimensión de la matriz que hay que invertir queda reducida a $Nu \times Nu$, quedando la carga de cálculo reducida (en el caso límite de $Nu = 1$, se reduce al caso escalar, como en el EPSAC) aunque restringiendo la optimalidad. Para aumentar la rapidez del algoritmo, fundamental sobre todo en el caso adaptativo, se presenta en esta tesis un método desarrollado para la gran mayoría de procesos industriales y con una carga de cálculo mínima. También existen realizaciones usando redes neuronales de Hopfield para obtener algoritmos rápidos [72].

2.6.3 Consideración de ruidos coloreados

Cuando el polinomio del ruido $C(z^{-1})$ de la ecuación (2.8) no es igual a 1 la predicción cambia ligeramente. Para calcular el predictor en esta situación se debe resolver la siguiente ecuación diofántica:

$$C(z^{-1}) = E_j(z^{-1})\tilde{A}(z^{-1}) + z^{-j}F_j(z^{-1}) \quad (2.19)$$

con $\delta(E_j(z^{-1})) = j - 1$ y $\delta(F_j(z^{-1})) = \delta(\tilde{A}(z^{-1})) - 1$

Multiplicando la ecuación (2.8) por $\Delta E_j(z^{-1})$ y usando (2.19)

$$C(z^{-1})(y(t+j) - E_j(z^{-1})e(t+j)) = E_j(z^{-1})B(z^{-1})\Delta u(t+j-1) + F_j(z^{-1})y(t)$$

Como los términos de ruido están todos en el futuro, el valor esperado para el término de la izquierda de la ecuación anterior es:

$$E[C(z^{-1})(y(t+j) - E_j(z^{-1})e(t+j))] = C(z^{-1})\hat{y}(t+j|t)$$

El valor esperado de la salida puede ser generado por la ecuación:

$$C(z^{-1})\hat{y}(t+j|t) = E_j(z^{-1})B(z^{-1})\Delta u(t+j-1) + F_j(z^{-1})y(t) \quad (2.20)$$

Nótese que esta ecuación de predicción podría ser usada para generar las predicciones de forma recursiva. Se puede obtener una expresión explícita para la predicción óptima de j pasos resolviendo la siguiente ecuación diofántica:

$$1 = C(z^{-1})M_j(z^{-1}) + z^{-k}N_j(z^{-1}) \quad (2.21)$$

con $\delta(M_j(z^{-1})) = j - 1$ y $\delta(N_j(z^{-1})) = \delta(C(z^{-1})) - 1$.

Multiplicando la ecuación (2.20) por $M_j(z^{-1})$ y usando (2.21),

$$\hat{y}(t+j|t) = M_j E_j(z^{-1})B(z^{-1})\Delta u(t+j-1) + M_j(z^{-1})F_j(z^{-1})y(t) + N_j(z^{-1})y(t)$$

que se puede expresar como:

$$\hat{y}(t+j|t) = G(z^{-1})\Delta u(t+j-1) + Gp(z^{-1})\Delta u(t+j-1) + (M_j(z^{-1})F_j(z^{-1}) + N_j(z^{-1}))y(t)$$

con $\delta(G(z^{-1})) < j$. Estas predicciones se pueden usar en la función de coste que se minimiza de la misma forma que el caso de ruido blanco.

2.7 Control Predictivo Restringido de Horizonte Deslizante

2.7.1 Introducción

A pesar de la gran aceptación del GPC en la industria, aún no había sido desarrollada una teoría formal del control predictivo en general, existiendo un vacío en cuestiones fundamentales como estabilidad y robustez. De hecho, la mayoría de los resultados sobre estabilidad se limitaban al caso límite de horizonte infinito (no utilizado en la práctica), y no existió una teoría clara que relacione el comportamiento del bucle cerrado con los parámetros de diseño como horizontes y secuencias de ponderación.

Con la idea de resolver alguno de los inconvenientes anteriores, aparece una variación sobre la formulación básica del GPC, llamada Control Predictivo Restringido de Horizonte Deslizante (Constrained Receding-Horizon Predictive Control, CRHPC), desarrollado por Clarke y Scattolini [21], [64], [62], que permite obtener resultados sobre estabilidad y robustez para horizontes pequeños. La idea consiste básicamente en obligar a que la salida coincida con la referencia durante un cierto intervalo al final del horizonte de predicción, según se muestra en la figura 2.5. Se emplean algunos de los grados de libertad de las señales de control futuras para forzar la salida mientras que el resto queda disponible para minimizar la función de coste en un cierto intervalo.

2.7.2 Cálculo de la ley de control

La obtención de la ley de control se realiza de manera similar al GPC, aunque los cálculos resultan algo más complejos. Las señales de control deben minimizar la función objetivo

$$J(N_y, N_u) = E \left\{ \sum_{i=1}^{N_y} \mu(i) [y(t+i|t) - y^o(t+i)]^2 + \sum_{j=0}^{N_u} \lambda(j) [\Delta u(t+j)]^2 \right\} \quad (2.22)$$

donde $y^o(t+i)$ es la referencia, sujeta a las restricciones

$$\begin{aligned} y(t+N_y+i) &= y^o(t+N_y+1) & i &= 1 \dots m \\ \Delta u(t+N_u+j) &= 0 & j &> 0 \end{aligned} \quad (2.23)$$

Siendo los parámetros de diseño los valores de los horizontes, de las secuencias de ponderación $\mu(i)$ y $\lambda(i)$ así como el valor de m , que marca los instantes de coincidencia entre la salida y la referencia.

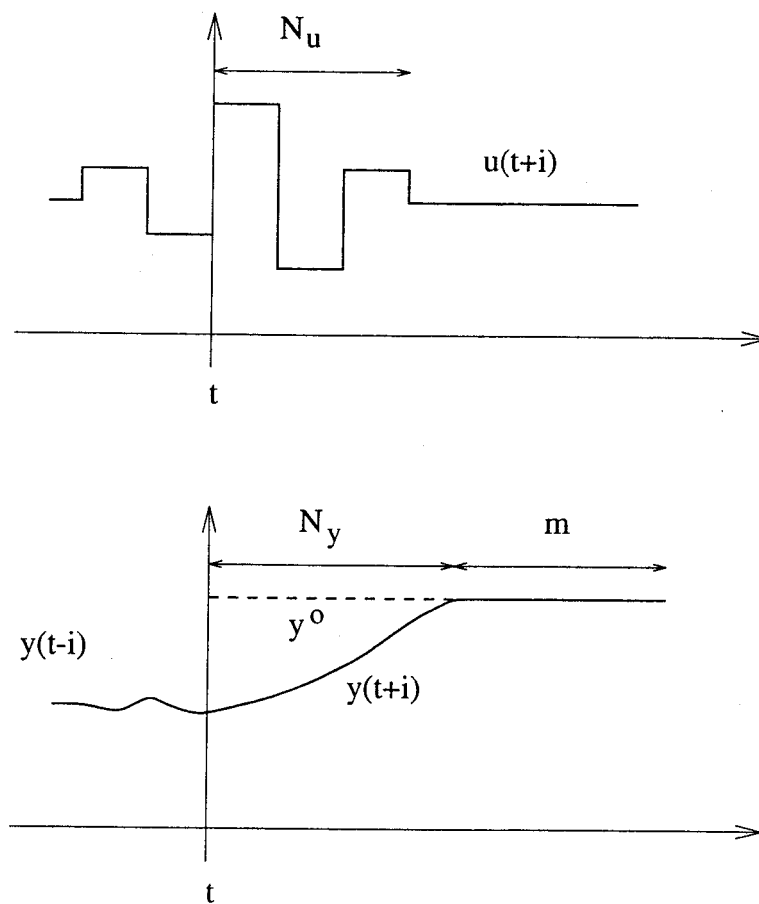


Figura 2.5: Constrained Receding-Horizon Predictive Control

Para resolver el problema de la optimización es necesario conocer la predicción de la salida, lo que se consigue definiendo los polinomios $E_i(z^{-1})$, $F_i(z^{-1})$ y $G_i(z^{-1})$ siendo E_i de grado $i - 1$ y tales que:

$$\begin{aligned} 1 &= E_i(z^{-1})A \Delta(z^{-1}) + z^{-i}F_i(z^{-1}) \\ z^{-d}E_i(z^{-1})B(z^{-1}) &= \sum_{h=1}^i s_h z^{-h} + z^{-i+1}G_i(z^{-1}) \end{aligned} \quad (2.24)$$

siendo s_h los coeficientes de la respuesta ante escalón del proceso, el cual puede ser modelado por

$$\Delta(z^{-1})A(z^{-1})y(t) = B(z^{-1}) \Delta u(t - d)$$

Multiplicando la ecuación anterior por $E_i(z^{-1})$ y teniendo en cuenta (2.24), se llega fácilmente a

$$y(t+i) = \sum_{h=1}^i s_h \Delta u(t+i-h) + f(t+i)$$

$$f(t+i) = F_i(z^{-1})y(t) + G_i(z^{-1})\Delta u(t-1)$$

Si se definen ahora las siguientes secuencias de futuras variables:

$$\begin{aligned} Y(t) &= [y(t+1) y(t+2) \dots y(t+N_y)]^T \\ Y^o(t) &= [y^o(t+1) y^o(t+2) \dots y^o(t+N_y)]^T \\ \Delta U(t) &= [\Delta u(t) \Delta u(t+1) \dots \Delta u(t+N_u)]^T \\ F(t) &= [f(t+1) f(t+2) \dots f(t+N_y)]^T \\ \bar{Y}(t) &= [y(t+N_y+1) y(t+N_y+2) \dots y(t+N_y+m)]^T \\ \bar{Y}^o(t) &= [y^o(t+N_y+1) y^o(t+N_y+2) \dots y^o(t+N_y+m)]^T \\ \bar{F}^o(t) &= [f(t+N_y+1) f(t+N_y+2) \dots f(t+N_y+m)]^T \end{aligned}$$

las siguientes matrices, de dimensión $N_y \times (N_u + 1)$ y $m \times N_u + 1$ respectivamente:

$$G = \begin{bmatrix} s_1 & 0 & 0 & \dots & 0 \\ s_2 & s_1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ s_{N_y} & s_{N_y-1} & s_{N_y-2} & \dots & s_{N_y-N_u} \end{bmatrix}$$

$$\bar{G} = \begin{bmatrix} s_{N_y+1} & s_{N_y} & \dots & s_{N_y-N_u+1} \\ s_{N_y+2} & s_{N_y+1} & \dots & s_{N_y-N_u+2} \\ \vdots & \vdots & \vdots & \vdots \\ s_{N_y+m} & s_{N_y+m-1} & \dots & s_{N_y-N_u+m} \end{bmatrix}$$

y las secuencias de ponderación

$$\begin{aligned} M(t) &= \text{diag} \{ \mu(1), \mu(2), \dots, \mu(N_y) \} \\ \Lambda(t) &= \text{diag} \{ \lambda(0), \lambda(1), \dots, \lambda(N_u) \} \end{aligned}$$

se puede escribir:

$$\begin{aligned} Y(t) &= G \Delta U(t) + F(t) \\ \bar{Y}(t) &= \bar{G} \Delta U(t) + \bar{F}(t) \end{aligned}$$

Entonces la función de coste (2.22) y las restricciones (2.23) se pueden escribir como:

$$\begin{aligned} J &= [Y(t) - Y^o(t)]^T M(t) [Y(t) - Y^o(t)] + \Delta U^T(t) \Lambda(t) \Delta U(t) \\ \bar{G} \Delta U(t) + \bar{F}(t) &= \bar{Y}^o(t) \end{aligned}$$

La solución se obtiene haciendo uso de los multiplicadores de Lagrange. Si se considera el caso más común de secuencia de peso constante, es decir, $M(t) = \mu I$, $\Lambda(t) = \lambda I$, la solución puede escribirse como:

$$\begin{aligned} \Delta U(t) = & (\mu G^T G + \lambda I)^{-1} [\mu G^T (Y^o(t) - F(t)) + \bar{G}^T (\bar{G} (\mu G^T G + \lambda I)^{-1} \bar{G}^T)^{-1} \\ & \times (\bar{Y}^o(t) - \bar{F}(t) - \mu \bar{G} (\mu G^T G + \lambda I)^{-1} G^T (Y^o(t) - F(t)))] \end{aligned} \quad (2.25)$$

Al ser una estrategia deslizante, sólo el primer elemento del vector $\Delta U(t)$ es utilizado, desechándose los restantes y procediendo a un nuevo cálculo en el siguiente instante de muestreo. Este método proporciona pues una solución analítica que, como se observa en (2.25) resulta más compleja que la proporcionada por el GPC. La carga de cálculo puede resultar considerable pues se realizan diversas operaciones con matrices incluyendo inversión, aunque los cálculos se puedan optimizar sabiendo que G es triangular y las matrices a invertir son simétricas. Este factor puede ser determinante en el caso adaptativo donde todos los vectores y matrices pueden cambiar constantemente.

La obtención de la señal de control precisa la inversión de la matriz $\bar{G}(\mu G^T G + \lambda I)^{-1} \bar{G}^T$. A partir de la definición de la matriz \bar{G} se deduce que debe cumplirse la condición $m \leq N_u + 1$, lo cual puede ser interpretado como que el número m de restricciones a la salida no puede ser mayor que el de variaciones en la señal de control $N_u + 1$. Otra condición para la invertibilidad es que $m \leq n + 1$, ya que el coeficiente s_i de la respuesta ante escalón es una combinación lineal de los previos $n + 1$ valores (siendo n el orden del sistema). Nótese que esta última condición va a impedir que el orden de la matriz a invertir (que es $m \times m$) sea muy grande, ya que en la mayoría de casos prácticos el valor de n no pasará de dos o tres.

2.7.3 Propiedades

Como se ha indicado, una de las ventajas de este método es que permite obtener resultados de estabilidad para horizontes finitos, compensando en cierta manera la carga de cálculo que acarrea. Se presentan a continuación las propiedades más sobresalientes del CRHPC, cuya demostración puede encontrarse en [64] y se basan en formular la ley de control (2.25) en el espacio de estados. El hecho de que la salida coincida con la referencia en un cierto intervalo garantiza la monotonicidad de la ecuación de Riccati asociada y por tanto la estabilidad según los resultados de Kwon y Pearson [50].

Propiedad 1. Si $N_y = N_u > n + d + 1$ y $m = n + 1$ entonces el sistema en bucle cerrado es asintóticamente estable.

Propiedad 2. Si $N_u = n + d$ y $m = n + 1$ la ley de control resulta ser del tipo *dead-beat*.

Propiedad 3. Si el sistema es asintóticamente estable, $\mu = 0$, $m = 1$ y existe un valor ν tal que

$$s_\nu \leq s_{\nu+1} \leq \dots \leq s_\infty, \quad s_\nu > s_\infty/2, \quad s_\infty > 0$$

o bien

$$s_\nu \geq s_{\nu+1} \geq \dots \geq s_\infty, \quad s_\nu < s_\infty/2, \quad s_\infty < 0$$

entonces, para cualquier $N_y = N_u \leq \nu - 1$ el sistema en bucle cerrado es asintóticamente estable.

Propiedad 4. Bajo las condiciones de la propiedad anterior, el sistema en bucle cerrado es asintóticamente estable para $N_u = 0$ y $N_y \geq \nu - 1$.

Propiedad 5. Si el sistema en bucle abierto es asintóticamente estable, $\mu = 0$, $m = 1$, $N_u = 0$ y existen ciertas constantes $K > 0$ y $0 < \eta < 1$ tales que

$$|s_i - s_\infty| \leq K\eta^i, \quad i \geq 0$$

entonces, para cualquier N_y tal que

$$|s_{N_y+1}| > K \frac{1+\eta}{1-\eta} \eta^{N_y+1}$$

el sistema en bucle cerrado es asintóticamente estable.

Se ve pues que el método es capaz de estabilizar cualquier tipo de proceso como sistemas inestables o de fase no mínima. Al igual que el GPC, el uso de entradas y salidas filtradas puede resultar en un control por asignación de polos. El uso del polinomio $P(z^{-1})$ introducido en [24] puede servir para mostrar la íntima relación entre el control predictivo y la asignación de polos [21]. Este polinomio aparece al definir una salida auxiliar $\psi(t) = P(z^{-1})y(t)$, la cual sustituye a $y(t)$ en (2.22) y (2.23) de forma que los polos deseados del bucle cerrado son las raíces de $z^{n+1}P(z^{-1})$.

2.8 GPC estable

2.8.1 Introducción

Siguiendo con la intención de resolver la falta de resultados sobre estabilidad que posee el GPC, Rossiter y Kouvaritakis han desarrollado una formulación que asegura que

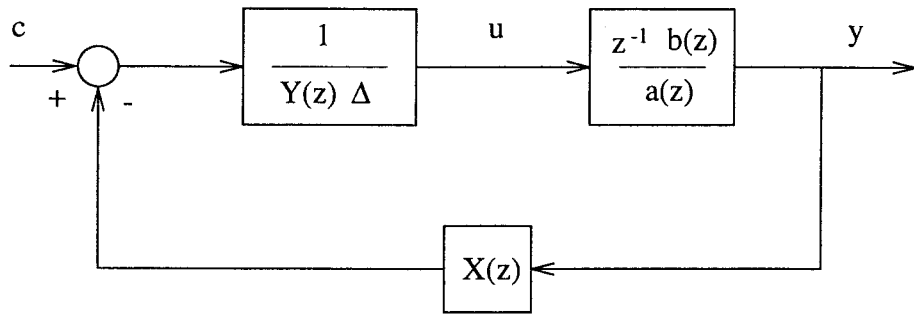


Figura 2.6: Estabilización del bucle

la función de coste asociada es monótona decreciente, garantizando de esta forma la estabilidad del bucle cerrado. Esta estrategia es llamada por ello GPC Estable (Stable Generalized Predictive Control, SGPC) [46], [79] y se basa en estabilizar el bucle cerrado antes de la aplicación de la estrategia de control. Ahora las acciones de control que hay que calcular son los valores de la referencia que se envían al bucle cerrado, y no las entradas al sistema, que serán función de éstas. Para estabilizar el bucle se elige un controlador que conduzca a respuestas finitas ante impulso.

2.8.2 Desarrollo de la ley de control

Para la obtención de la ley de control se considera el sistema modelado por

$$G(z) = \frac{z^{-1}b(z)}{a(z)} = \frac{b_1z^{-1} + b_2z^{-2} + \dots + b_nz^{-n}}{1 + a_1z^{-1} + a_2z^{-2} + \dots + a_nz^{-n}} \quad (2.26)$$

donde por simplicidad y sin pérdida de generalidad se ha considerado que el orden del numerador y del denominador es el mismo; si el retardo es mayor que la unidad, basta con hacer los primeros términos de $b(z)$ iguales a cero. Como se ha indicado, antes de optimizar la función de coste, se estabiliza el bucle con los polinomios $X(z)$ e $Y(z)$ según se muestra en la figura 2.6. La señal c es la referencia para el bucle cerrado, que es el valor que aparecerá en la función de coste.

Los polinomios $X(z)$ e $Y(z)$ satisfacen las siguientes relaciones:

$$a(z)Y(z)\Delta(z) + z^{-1}b(z)X(z) = 1$$

$$K(z) = \frac{X(z)}{Y(z)\Delta(z)}$$

con $\Delta(z) = 1 - z^{-1}$ y $K(z)$ el controlador total. Con estas relaciones se deduce que:

$$y(z) = z^{-1}b(z)c(z)$$

$$\Delta u(z) = A(z)c(z)$$

siendo $A(z) = a(z)\Delta$.

Con las relaciones anteriores se puede obtener la función de coste que, al igual que el GPC estándar, mide la discrepancia de las salidas futuras con la referencia así como el esfuerzo de control necesario. El coste debe ser expresado como función de los futuros valores de c , de forma que con su minimización se obtengan estos valores que, como se ha dicho, proporcionan la referencia al bucle ya estabilizado. Para ello se consideran los siguientes vectores:

$$\begin{aligned} y^+ &= [y(t+1) \ y(t+2) \ \dots \ y(t+n_y)]^T \\ c^+ &= [c(t+1) \ c(t+2) \ \dots \ c(t+n_c)]^T \\ \Delta u^+ &= [\Delta u(t) \ \Delta u(t+1) \ \dots \ \Delta u(t+n_y-1)]^T \\ c^- &= [c(t) \ c(t-1) \ \dots \ c(t-n)]^T \end{aligned}$$

donde n_y y n_c son los horizontes de salida y de referencia.

Simulando la ecuación (2.26) hacia delante en el tiempo se puede escribir:

$$\begin{aligned} y^+ &= \Gamma_b c^+ + H_b c^- + M_b c^\infty \\ \Delta u^+ &= \Gamma_a c^+ + H_A c^- + M_A c^\infty \end{aligned} \quad (2.27)$$

donde los dos últimos términos de cada ecuación representan la respuesta libre y_f y Δu_f de y e Δu respectivamente. Las matrices Γ_A , Γ_b , H_A , H_b , M_A y M_b son matrices que pueden ser fácilmente obtenidas según se muestra en [46] y c^∞ es el vector de $n_y + n_c$ filas que contiene los futuros valores deseados de la señal de mando c . Los valores de este vector se eligen para asegurar que no haya *offset* en régimen permanente. Normalmente, si $r(t+i)$, $i = 1, \dots, n_y$ es el setpoint a seguir por la salida, se elige c^∞ como $[1, 1, \dots, 1]^T \times r(t+n_y)/b(1)$. Si se consideran cambios en escalón en la referencia, se puede elegir c^∞ como el vector de referencias futuras premultiplicado por una matriz E formada por las últimas $n_y - n_c$ filas de I_{n_y} .

Con todo lo anterior, la función de coste se puede escribir como:

$$J = \|r^+ - y^+\|_2^2 + \lambda \|\Delta u^+\|_2^2 = [c^+ - c_o]^T S^2 [c^+ - c_o] + \gamma$$

donde

$$\begin{aligned} S^2 &= \Gamma_b^T \Gamma_b + \lambda \Gamma_A^T \Gamma_A \\ c_o &= S^{-2} [\Gamma_b^T (r^+ - y_f) - \lambda \Gamma_A^T \Delta u_f] \\ \gamma &= \|r^+ - y_f\|^2 + \lambda \|\Delta u_f\|^2 - \|c_o\|^2 \end{aligned}$$

Al ser γ una constante conocida, la ley de control del SGPC vendrá dada por la minimización del coste $J = \|S(c^+ - c_o)\|$, y la ley de control del SGPC viene definida

por

$$\Delta u(t) = A(z)c(t) \quad c(t) = \frac{p_r(z)}{p_c(z)}r(t + ny)$$

donde $p_r(z)$ y $p_c(z)$ son polinomios definidos por

$$\begin{aligned} p_r &= e^T T [\Gamma_b^T (I - M_b E) - \Gamma_b^T \Gamma_A E] \\ p_c &= e^T T [\Gamma_b^T H_b + \lambda \Gamma_b H_A] \end{aligned} \quad (2.28)$$

y $T = (\Gamma_b^T \Gamma_b + \lambda \Gamma_A \Gamma_A)^{-1}$. e es el primer vector de la base estándar y los polinomios vienen dados en grado descendiente de z .

La estabilidad del algoritmo es una consecuencia del hecho de que y y Δu están relacionadas con c a través de operadores de respuesta impulsional finita ($z^{-1}b(z)$ y $A(z)$), con lo que se puede establecer que el coste es monótonamente decreciente.

La predicción obtenida en 2.27 es en bucle abierto, la cual se puede mejorar como se muestra en [79] simulando las ecuaciones del bucle cerrado $Y(z) \Delta u = c - X(z)y$ y $A(z)y = b(z) \Delta u$, obteniendo:

$$\begin{aligned} y^+ &= \Gamma_b c^+ + M_b c^\infty - P_1 y^- - P_2 \Delta u^- \\ \Delta u^+ &= \Gamma_A c^+ + M_A c^\infty - P_3 y^- - P_4 \Delta u^- \end{aligned} \quad (2.29)$$

donde

$$\begin{aligned} P_1 &= C_b H_{z^{-1}X} + C_Y H_A & P_2 &= C_b H_Y - C_Y H_b \\ P_3 &= C_A H_{z^{-1}X} + C_{z^{-1}X} H_A & P_4 &= C_A H_Y - C_{z^{-1}X} H_b \end{aligned}$$

usando la ecuación (2.29) en la función de coste, resulta una solución óptima que depende de los valores pasados y futuros de r y los valores actuales y pasados de y y Δu . Teniendo en cuenta el controlador que estabiliza el sistema, la configuración completa es la que se muestra en la figura 2.7 y resulta ser:

$$D(z) \Delta(z)u(z) = p_r(z)r(z) + N(z)y(z) \quad (2.30)$$

El prefiltro $p_r(z)$ viene dado por (2.28) mientras que el numerador y denominador del controlador puede calcularse directamente de la solución de la siguiente identidad de Bezout:

$$A(z)D(z) + z^{-1}b(z)N(z) = p_c(z)$$

con p_c obtenido según (2.28). Por tanto la ecuación (2.30) resulta ser una expresión simple de la ley de control óptima del SGPC.

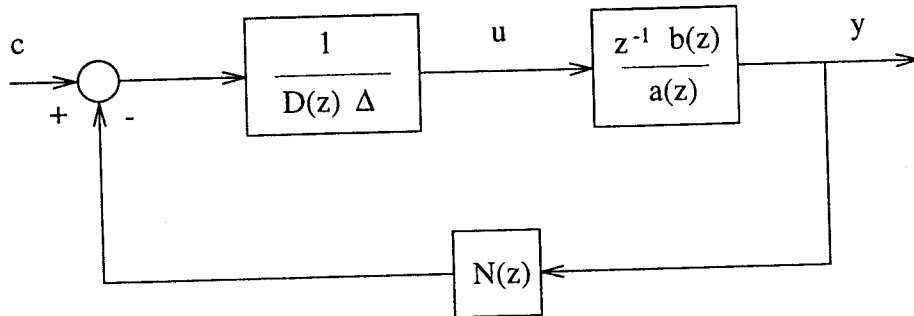


Figura 2.7: Configuración del controlador

2.9 Control predictivo multivariable

La mayoría de los procesos industriales poseen múltiples entradas y salidas. En ciertos casos un cambio en una de las variables manipuladas afecta principalmente a su correspondiente variable controlada y se puede considerar el proceso como un conjunto de sistemas monovariabes desacoplados. En muchas ocasiones, cuando se cambia una de las variables manipuladas, no sólo se ve afectada su correspondiente variable de proceso, sino que también influye sobre el resto. Estas interacciones pueden derivar en pérdida de prestaciones o conducir incluso a la inestabilidad. Cuando las interacciones no son despreciables, el proceso debe ser considerado como multivariable.

La forma más utilizada de controlar sistemas multivariables es mediante el uso de desacopladores para disminuir las interacciones, para posteriormente controlar cada sistema considerado independiente con un controlador monovariabes. Ello requiere emparejar correctamente entre sí las entradas y las salidas y además es necesario que ambas coincidan en número. Un desacoplo total es en general muy difícil de conseguir, sobre todo en procesos con dinámicas complejas o grandes tiempos muertos.

Una de las ventajas del MPC, que le ha hecho cosechar gran éxito, es su capacidad para manejar los procesos multivariables de forma directa [82],[89].

La obtención de la ley de control de un GPC multivariable sigue los mismos pasos que en el caso monovariabes aunque los cálculos asociados son más complejos. La diferencia estriba en que el proceso debe ser descrito ahora mediante matrices polinomiales en la variable z .

El modelo CARIMA para un proceso multivariable de n entradas y m salidas es:

$$\mathbf{A}(z^{-1})\mathbf{y}(t) = \mathbf{B}(z^{-1})\mathbf{u}(t-1) + \frac{1}{\Delta}\mathbf{C}(z^{-1})\mathbf{e}(t)$$

donde $\mathbf{A}(z^{-1})$ y $\mathbf{C}(z^{-1})$ son matrices polinomiales mónicas de dimensión $n \times n$ y $\mathbf{B}(z^{-1})$ es otra matriz de $n \times m$, definidas como:

$$\begin{aligned}\mathbf{A}(z^{-1}) &= I_{n \times n} + A_1 z^{-1} + A_2 z^{-2} + \cdots + A_{n_a} z^{-n_a} \\ \mathbf{B}(z^{-1}) &= B_0 + B_1 z^{-1} + B_2 z^{-2} + \cdots + B_{n_b} z^{-n_b} \\ \mathbf{C}(z^{-1}) &= I_{n \times n} + C_1 z^{-1} + C_2 z^{-2} + \cdots + C_{n_c} z^{-n_c}\end{aligned}$$

Las variables $y(t)$, $u(t)$ y $e(t)$ son el vector de n salidas, el vector de m entradas y el de n ruidos en el instante t respectivamente y el criterio a minimizar es:

$$J(N_1, N_2, Nu) = \sum_{j=N_1}^{N_2} \|\hat{y}(t+j|t) - w(t+j)\|_R^2 + \sum_{j=1}^{Nu} \|\Delta u(t+j-1)\|_Q^2 \quad (2.31)$$

donde $\hat{y}(t+j|t)$ es el predictor óptimo de j pasos y R y Q son matrices de peso definidas positivas.

La predicción también se puede calcular con una ecuación diofántica de la forma:

$$\mathbf{C}(z^{-1}) = \mathbf{E}_j(z^{-1})\tilde{\mathbf{A}}(z^{-1}) + z^{-j}\mathbf{F}_j(z^{-1})$$

considerando $\mathbf{C}(z^{-1}) = I_{n \times n}$ en el caso de ruido blanco.

La obtención de las m señales de control que deben ser aplicadas en el instante t es análoga al caso monovariable y se encuentra desarrollada en [15].

Aparte de la mayor carga de cálculo impuesta por el hecho de trabajar con matrices polinomiales, también aparece el problema de la obtención de los polinomios $\mathbf{A}(z^{-1})$ y $\mathbf{B}(z^{-1})$ a partir de la representación en forma de matriz de transferencia $\mathbf{T}(z^{-1})$. La matriz de transferencia es la forma más usada de representar los procesos multivariables, ya que puede ser obtenida mediante análisis frecuencial o aplicando el método de la curva de reacción a cada entrada, midiendo la ganancia estática, constante de tiempo y tiempo muerto para cada una de las salidas. Conocida la matriz $\mathbf{T}(z^{-1})$ se buscan las matrices polinomiales $\mathbf{A}(z^{-1})$ y $\mathbf{B}(z^{-1})$ tales que:

$$\mathbf{T}(z^{-1}) = \mathbf{A}(z^{-1})^{-1}\mathbf{B}(z^{-1})z^{-1}$$

La forma más sencilla de encontrar $\mathbf{A}(z^{-1})$ y $\mathbf{B}(z^{-1})$ es hacer $\mathbf{A}(z^{-1})$ diagonal con sus elementos iguales a los mínimos comunes múltiplos de los denominadores de la correspondiente fila de $\mathbf{T}(z^{-1})$. Entonces $\mathbf{B}(z^{-1}) = \mathbf{A}(z^{-1})\mathbf{T}(z^{-1})z$. Las matrices que se obtienen no tienen por qué ser coprimas por la izquierda; si se desea que sí lo sean, se puede lograr según se muestra en [35].

Otra problemática que aparece en la práctica es la asociada a la existencia de tiempos muertos. Aunque por la propia naturaleza del Control Predictivo los tiempos muertos aparecen implícitamente en el controlador, la existencia de valores dispares para distintas parejas de entrada-salida puede dar lugar a problemas numéricos, debido a la gran cantidad de ceros que aparecen en la matriz $\mathbf{B}(z^{-1})$. Este problema es tratado en detalle en [15]

2.10 GPC con restricciones

Como se indicó en la sección 2.3.2 en la práctica todos los procesos están sujetos a restricciones de uno u otro tipo. Sin embargo, éstas son ignoradas por la mayoría de las técnicas de control. Una de las características del Control Predictivo Basado en Modelo, y que ha tenido gran influencia en su aceptación en la práctica industrial es el hecho de ser la única metodología capaz de manejar las restricciones de forma sistemática durante la fase de diseño del controlador.

En la práctica es habitual usar el GPC estándar (tal como se ha descrito en la sección 2.6) para calcular la señal $u(t)$. Si ésta viola las restricciones entonces se satura a sus límites. El caso de que las señales de control futuras sobrepasen las restricciones (bien de amplitud o de velocidad de cambio) no es casi nunca considerado pues en la mayoría de los casos ni siquiera se calculan. Lógicamente esta forma de operar no garantiza el óptimo, ya que se puede trabajar en un punto de operación con mayor valor de la función objetivo y por tanto se pueden reducir las prestaciones. En cualquier caso, las variables manipuladas siempre pueden mantenerse dentro de ciertos límites por medio del programa de control o por los actuadores.

No ocurre así con las variables controladas, cuya violación de las restricciones puede ocasionar daños físicos en los equipos y pérdidas en la producción. Estas violaciones no pueden resolverse saturando las acciones de control, sino que deben ser tratadas aprovechando la capacidad de predicción del MPC. El sistema de control debe anticiparse al incumplimiento de las restricciones y corregirlas de forma adecuada.

Aparte de las restricciones en la amplitud y velocidad de cambio de las señales de control y de amplitud en las salidas, normalmente impuestas por razones de seguridad, también se puede forzar a que la respuesta del proceso posea ciertas características [48] como mantener la salida dentro de una banda, evitar sobreoscilaciones o comportamiento de fase no mínima, obligar a una respuesta monótona, tener en cuenta las no linealidades de los actuadores o imponer restricciones de igualdad al estado final (usado en el CRHPC).

Todas las restricciones anteriores pueden ser expresadas en función del vector \mathbf{u} de incrementos de las señales de control como $\mathbf{R} \mathbf{u} \leq \mathbf{c}$ y $\mathbf{A} \mathbf{u} = \mathbf{a}$, siendo \mathbf{R} y \mathbf{A} matrices y \mathbf{c} y \mathbf{a} vectores que tomarán distinta forma según el tipo de restricciones.

El problema ahora consiste en minimizar la función objetivo del GPC:

$$J = \frac{1}{2} \mathbf{u}^T \mathbf{H} \mathbf{u} + \mathbf{b} \mathbf{u} + \mathbf{f}_0$$

sujeta al cumplimiento del anterior conjunto de restricciones. Se tiene por tanto el problema de la optimización de una función cuadrática con restricciones lineales, conocido como Programación Cuadrática. La solución analítica del GPC sin restricciones ya no es válida y hay que recurrir a los métodos numéricos propios de la programación cuadrática (para una revisión de los principales, ver [15]). También se puede emplear la Programación Dinámica para resolver la optimización con restricciones, método más general que permite manejar modelos no lineales y funciones objetivo no necesariamente cuadráticas aunque lleva asociada una gran carga de cálculo; ésta puede ser reducida usando tablas precalculadas, como se muestra en [27].

Uno de los problemas que aparecen ahora, aparte de la mayor carga de cálculo, es el de encontrar una solución posible. Si las restricciones son únicamente en la señal de control, siempre se puede tomar la solución del problema sin restringir y limitar las actuaciones, pero no puede hacerse lo mismo para las restricciones en las salidas. Existen diversas sugerencias propuestas para tratar este tema; ver por ejemplo [73].

Otro de los problemas asociados es el de la estabilidad. Hasta hace poco tiempo no se habían obtenido resultados de estabilidad en presencia de restricciones. Sin embargo, han aparecido nuevas aproximaciones que se basan en la idea de que se puede demostrar que las funciones de costo son monótonamente decrecientes y, si hay una solución posible, pueden ser interpretadas como funciones de Lyapunov garantizando de este modo la estabilidad. Por un lado se tienen los resultados ya expuestos de Clarke y Scattolini con el CRHPC [21] imponiendo restricciones en el estado final y por otro están los obtenidos por Rawlings y Muske [73] considerando la función objetivo compuesta de dos partes: una con horizonte finito y restringida, y otra con horizonte infinito y sin restricciones. La parte infinita y sin restricciones se puede resolver por una ecuación de Riccati con lo que se puede obtener una función de costo dependiente del estado, que es posteriormente introducida en el problema finito de optimización el cual es resuelto por métodos numéricos.

Capítulo 3

Realización simple del GPC para procesos industriales

Una de las razones del éxito de los tradicionales controladores PID en la industria es que son realmente fáciles de realizar y su sintonización resulta muy sencilla usando ciertas reglas heurísticas, siendo las de Ziegler-Nichols [96] unas de las más usadas en la práctica industrial. Como se vió en el capítulo anterior, un Controlador Predictivo Generalizado da lugar a una ley de control lineal muy fácil de realizar siempre que los parámetros del controlador sean conocidos. Sin embargo el cálculo de los parámetros del GPC requiere ciertos cálculos matemáticos relativamente complejos, tales como la resolución de la ecuación diofántica, la formación de las matrices \mathbf{G} , \mathbf{G}' y el vector \mathbf{f} y luego la resolución de un sistema de ecuaciones lineales. Este hecho, que no constituye ningún problema en la comunidad académica, donde normalmente se dispone de paquetes de cálculo matemático, puede resultar un factor determinante a la hora de ser usado en la práctica industrial, donde se está acostumbrado a formas más simples de realizar y sintonizar controladores.

Los cálculos mencionados anteriormente sólo deben ser realizados una vez en procesos con parámetros fijos, pero si éstos cambian, los parámetros del GPC deben ser calculados de nuevo en tiempo real cuando se usa una estrategia de control adaptativo. En la práctica esto puede constituir una importante dificultad pues por un lado, muchos sistemas de control industrial no están diseñados para realizar cálculos complejos, y por otro, hay que considerar que la mayor parte del tiempo del computador de proceso se utiliza para otros cometidos diferentes del algoritmo de control en sí mismo (comunicación con el operador, tratamiento de alarmas, visualización y registro, etc), lo que provocará que el cálculo de los parámetros del GPC sea irrealizable para el tiempo de muestreo requerido y el número de lazos a controlar por el sistema.

Este capítulo pretende mostrar cómo el GPC puede ser fácilmente realizado y sintonizado para una amplia gama de procesos industriales. Se mostrará que el GPC puede ser programado con un conjunto limitado de instrucciones, disponibles en cualquier sistema de control distribuido, y que el tiempo de cálculo empleado, incluso para la sintonización, es realmente pequeño. El método para realizar el GPC se basa en el hecho de que una amplia clase de procesos industriales pueden describirse por unos pocos parámetros, y que se puede obtener un conjunto de funciones simples del tipo Ziegler-Nichols que relacionen los parámetros del GPC con los del proceso a controlar. Usando estas funciones, la realización de un GPC resulta casi tan simple como la de un PID.

3.1 Modelo del proceso

Cuando se consideran pequeños cambios en torno a un punto de funcionamiento la mayoría de los procesos industriales pueden ser descritos mediante un modelo lineal, normalmente de alto orden. La razón de esto estriba en que la mayoría de los procesos están compuestos de muchos elementos dinámicos, usualmente de primer orden, de forma tal que el modelo completo es de orden igual al número de elementos. De hecho, cada elemento de almacenamiento de masa o energía da lugar a un elemento de primer orden en el modelo. Por ejemplo un intercambiador de calor puede ser modelado dividiéndolo en fracciones de tubo cada una de las cuales puede ser considerada como un sistema de primer orden. El modelo resultante tendrá un orden igual al número de tramos en que se ha dividido el tubo, que será un número elevado. Estos modelos de elevado orden son muy difíciles de manejar en control. Pero, afortunadamente, es posible aproximar el comportamiento de tales procesos de elevado orden por un sistema con una constante de tiempo y un retardo.

Se puede considerar un proceso con N elementos de primer orden en serie, cada uno teniendo una constante de tiempo τ/N , tal como se indica en [28]. La función de transferencia del sistema viene dada por

$$G(s) = \frac{1}{(1 + \frac{\tau}{N}s)^N}$$

Haciendo variar el valor de N desde 1 hasta ∞ , la respuesta cambia desde una de primer orden hasta un retardo puro (igual a τ). Si, como ocurre en muchos procesos, una constante de tiempo es mucho mayor que el resto, las constantes de tiempo más pequeñas se unen para producir un retraso que actúa como un retardo puro. En esta situación la dinámica es debida casi exclusivamente a la constante de tiempo mayor, tal como se puede ver en la figura 3.1. Por tanto es posible aproximar el modelo de

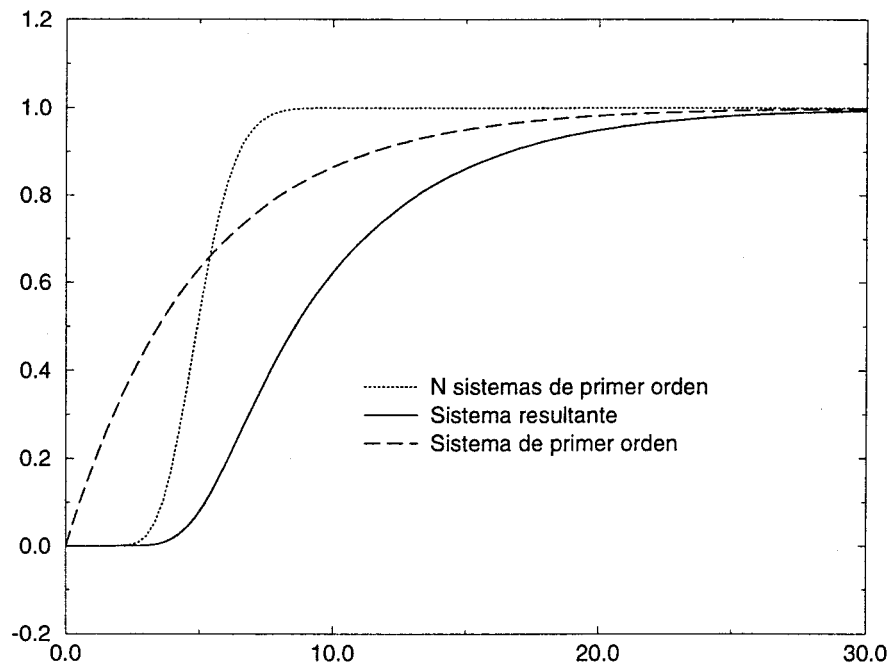


Figura 3.1: Respuesta del sistema

un sistema dinámico complejo de orden elevado mediante un proceso de primer orden combinado con un elemento de retardo. Este tipo de sistemas puede ser descrito por la siguiente función de transferencia:

$$G(s) = \frac{K}{1 + \tau s} e^{-s\tau_d} \quad (3.1)$$

donde K es la ganancia estática del proceso, τ es la constante de tiempo y τ_d es el retardo.

3.1.1 Identificación del proceso: el método de la curva de reacción

Una vez que se ha definido la estructura del modelo, el siguiente paso es la correcta elección de los valores de los parámetros. Para su identificación se le debe aplicar al sistema un estímulo apropiado a la entrada. En el método de la curva de reacción se aplica una entrada en escalón, es decir, una entrada con amplio espectro de frecuencias, y se registra la salida del proceso para ajustar el modelo a los valores medidos.

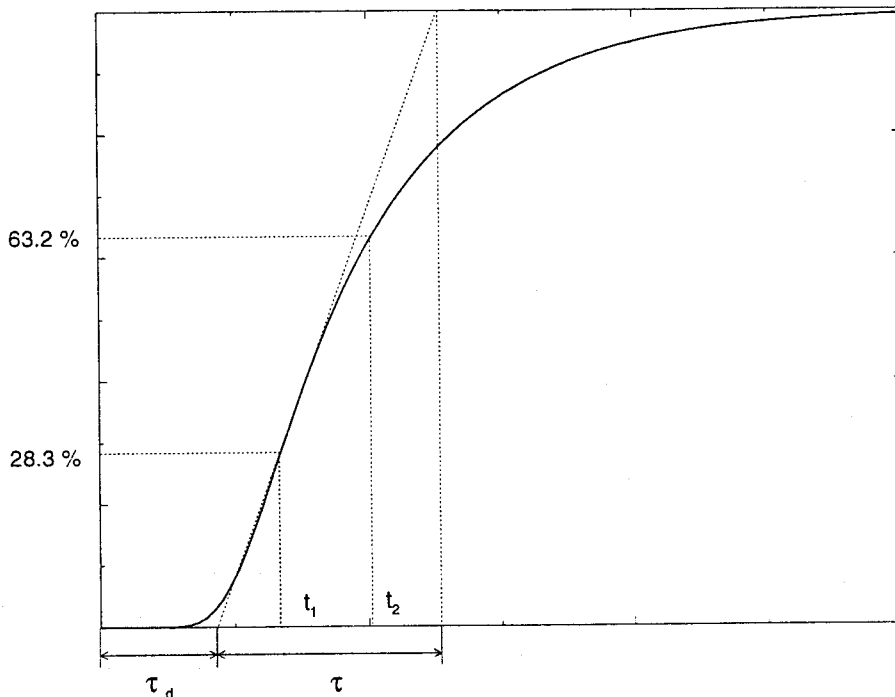


Figura 3.2: Curva de reacción

La respuesta ante escalón o curva de reacción tiene la forma que se muestra en la figura 3.2. Se provoca un escalón de magnitud Δu en la variable manipulada y se obtiene la evolución temporal de la variable de proceso $y(t)$. Los parámetros del proceso de la ecuación (3.1) se pueden obtener midiendo el incremento producido en la salida tras el transitorio, Δy , y dos tiempos: t_1 que es el instante en que la salida alcanza el 28.3 por ciento del valor estacionario, y t_2 cuando la salida alcanza el 63.2 por ciento. Con el uso de estos valores, los parámetros del proceso vienen dados por:

$$\begin{aligned}
 K &= \frac{\Delta y}{\Delta u} \\
 \tau &= 1.5(t_2 - t_1) \\
 \tau_d &= 1.5(t_1 - \frac{1}{3}t_2)
 \end{aligned}
 \tag{3.2}$$

Otra forma similar y quizás más intuitiva para la obtención de los parámetros del proceso consiste en encontrar el punto de inflexión de la respuesta temporal y dibujar la tangente a la curva por dicho punto [39]. La ganancia estática es obtenida por la expresión anterior mientras que τ y τ_d vienen dados directamente de la curva, como se ve en la figura. Los valores obtenidos son similares en ambos casos.

El método de la curva de reacción es probablemente uno de los más usados en la industria para la sintonización de controladores, como es el caso del método de Ziegler-Nichols para sintonizar PIDs, y es usado en la etapa de presintonización de algunos

controladores autoajustables comerciales.

En las dos próximas secciones se va a realizar el desarrollo del GPC para los procesos que pueden identificarse usando este método; se tratará primero el caso más simple en el que el retardo es un múltiplo entero del tiempo de muestreo y posteriormente su ampliación al caso de retardos no enteros. Más adelante también se tratarán los procesos con integradores, que no pueden ser identificados por la curva de reacción al presentar una salida que aumenta continuamente con el tiempo ante una entrada en escalón.

3.2 Procesos con retardo múltiplo del tiempo de muestreo

3.2.1 Modelo discreto del proceso

Cuando el retardo τ_d es un múltiplo entero del tiempo de muestreo T ($\tau_d = dT$), la correspondiente función de transferencia discreta del sistema continuo (3.1) toma la forma:

$$G(z^{-1}) = \frac{bz^{-1}}{1 - az^{-1}} z^{-d} \quad (3.3)$$

donde los parámetros discretos a , b y d pueden obtenerse fácilmente de los continuos discretizando el sistema. Con el empleo de un mantenedor de orden cero, se obtienen las siguientes expresiones:

$$a = e^{-\frac{T}{\tau}} \quad b = K(1 - a) \quad d = \frac{\tau_d}{T}$$

Si para modelar las perturbaciones se emplea un modelo CARIMA (Controlled Auto-Regressive and Integrated Moving-Average), es decir, autorregresivo y con media móvil integrado, con el polinomio del ruido igual a la unidad, se obtiene la siguiente ecuación

$$(1 - az^{-1})y(t) = bz^{-d}u(t - 1) + \frac{\varepsilon(t)}{\Delta}$$

donde $u(t)$ e $y(t)$ son las secuencias de control y salida del proceso respectivamente, $\varepsilon(t)$ es un ruido blanco de media nula y $\Delta = 1 - z^{-1}$. Esta ecuación se puede transformar en:

$$y(t + 1) = (1 + a)y(t) - ay(t - 1) + b \Delta u(t - d) + \varepsilon(t + 1) \quad (3.4)$$

3.2.2 Formulación del problema

Como se ha mostrado en el capítulo anterior, el algoritmo del Control Predictivo Generalizado consiste en aplicar la secuencia de control que minimice una función de costo multipaso de la forma

$$J(N_1, N_2) = E \left\{ \sum_{j=N_1}^{N_2} \delta(j) [\hat{y}(t+j|t) - w(t+j)]^2 + \sum_{j=1}^{N_2-d} \lambda(j) [\Delta u(t+j-1)]^2 \right\} \quad (3.5)$$

Nótese que el mínimo horizonte de salida N_1 debe tomar un valor mayor que el retardo d ya que las salidas para instantes anteriores a $d+1$ no pueden ser afectadas por la primera acción de control $u(t)$. A partir de ahora N_1 y N_2 se considerarán como $N_1 = d+1$ y $N_2 = d+N$, siendo N el horizonte de control.

Si en el instante t se dispone de las predicciones de la salida $\hat{y}(t+d+j-1|t)$ y $\hat{y}(t+d+j-2|t)$, entonces, teniendo en cuenta la ecuación (3.4), el mejor valor futuro para $\hat{y}(t+d+j|t)$ lo proporciona:

$$\begin{aligned} \hat{y}(t+d+j|t) &= (1+a)\hat{y}(t+d+j-1|t) - a\hat{y}(t+d+j-2|t) + \\ &+ b \Delta u(t+j-1) \end{aligned} \quad (3.6)$$

Aplicando esta ecuación sucesivamente para $j = 1, 2, \dots, i$ se tiene

$$\hat{y}(t+d+i|t) = G_i(z^{-1})\hat{y}(t+d|t) + D_i(z^{-1}) \Delta u(t+i-1) \quad (3.7)$$

donde $G_i(z^{-1})$ es de grado 1 y $D_i(z^{-1})$ es de grado $i-1$. Nótese que cuando $\delta(i) = 1$ y $\lambda(i) = \lambda$, los polinomios $D_i(z^{-1})$ son iguales a los $G_i(z^{-1})$ dados en [23] para el caso de $d = 0$ y que los términos $f(t+i)$ de esta referencia son iguales a $G_i(z^{-1})y(t)$ de la ecuación (3.7).

Si se introduce $\hat{y}(t+d+i|t)$ en la ecuación (3.5), $J(N)$ es función exclusivamente de $\hat{y}(t+d|t)$, $\hat{y}(t+d-1|t)$, $\Delta u(t+N_2-d-1)$, $\Delta u(t+N_2-d-2) \dots \Delta u(t)$ y de la trayectoria de referencia.

La minimización de $J(N)$ respecto a $\Delta u(t)$, $\Delta u(t+1) \dots \Delta u(t+N-1)$ conduce a

$$\mathbf{M} \mathbf{u} = \mathbf{P} \mathbf{y} + \mathbf{R} \mathbf{w} \quad (3.8)$$

donde

$$\mathbf{u} = [\Delta u(t) \quad \Delta u(t+1) \quad \cdots \quad \Delta u(t+N-1)]^T$$

$$\mathbf{y} = [\hat{y}(t+d|t) \quad \hat{y}(t+d-1|t)]^T$$

$$\mathbf{w} = [w(t+d+1) \quad w(t+d+2) \quad \cdots \quad w(t+d+N)]^T$$

\mathbf{M} y \mathbf{R} son matrices de dimensión $N \times N$ y \mathbf{P} de dimensión $N \times 2$.

Llamando q a la primera fila de la matriz \mathbf{M}^{-1} , $\Delta u(t)$ viene dada por

$$\Delta u(t) = q \mathbf{P} \mathbf{y} + q \mathbf{R} \mathbf{w}$$

Si se considera que el setpoint no va a cambiar (o no se conoce su evolución futura) $w(t+d+i)$ será igual al valor actual de la referencia $r(t)$. Por tanto la trayectoria de referencia se puede escribir como:

$$\mathbf{w} = [1 \cdots 1] r(t)$$

Y por consiguiente el incremento de la señal de control $\Delta u(t)$ resulta:

$$\Delta u(t) = l_{y1} \hat{y}(t+d|t) + l_{y2} \hat{y}(t+d-1|t) + l_{r1} r(t) \quad (3.9)$$

Donde $q \mathbf{P} = [l_{y1} \quad l_{y2}]$ y $l_3 = \sum_{i=1}^N q_i \sum_{j=1}^N r_{ij}$. Los coeficientes l_{y1} , l_{y2} , l_{r1} , son funciones de a , b , $\delta(i)$ y $\lambda(i)$. Si se considera que el proceso posee ganancia estática unidad (es decir, $b = 1 - a$), los coeficientes que aparecen en (3.9) sólo dependerán de $\delta(i)$ y $\lambda(i)$ (que se consideran fijos) y del polo del sistema que cambiará en el caso de una estrategia de control adaptativo. Al hacer esto se está trabajando realmente con un esfuerzo de control λ normalizado. Esta consideración obliga a dividir la señal de control por la ganancia estática (en el caso general de que ésta sea distinta de la unidad).

La esquema resultante de control se muestra en la figura 3.3. Los parámetros estimados del proceso se utilizan para calcular los coeficientes del controlador (l_{y1} , l_{y2} , l_{r1}). El predictor (3.6) proporciona los valores $\hat{y}(t+d|t)$ y $\hat{y}(t+d-1|t)$ y la señal de control se divide por la ganancia estática del proceso para obtener un sistema con ganancia estática unidad.

Los coeficientes del controlador no dependen del retardo d y para valores fijos de $\delta(i)$ y $\lambda(i)$, son función exclusivamente del valor estimado del polo (\hat{a}). La forma estándar de calcular los coeficientes del controlador sería mediante el cálculo de las matrices \mathbf{M} , \mathbf{P} y \mathbf{R} y la resolución de la ecuación (3.8) seguida por la generación de la ley de control proporcionada por (3.9). Esto implica la triangularización de una matriz de dimensión $N \times N$, lo cual puede resultar prohibitivo para algunas aplicaciones en tiempo real.

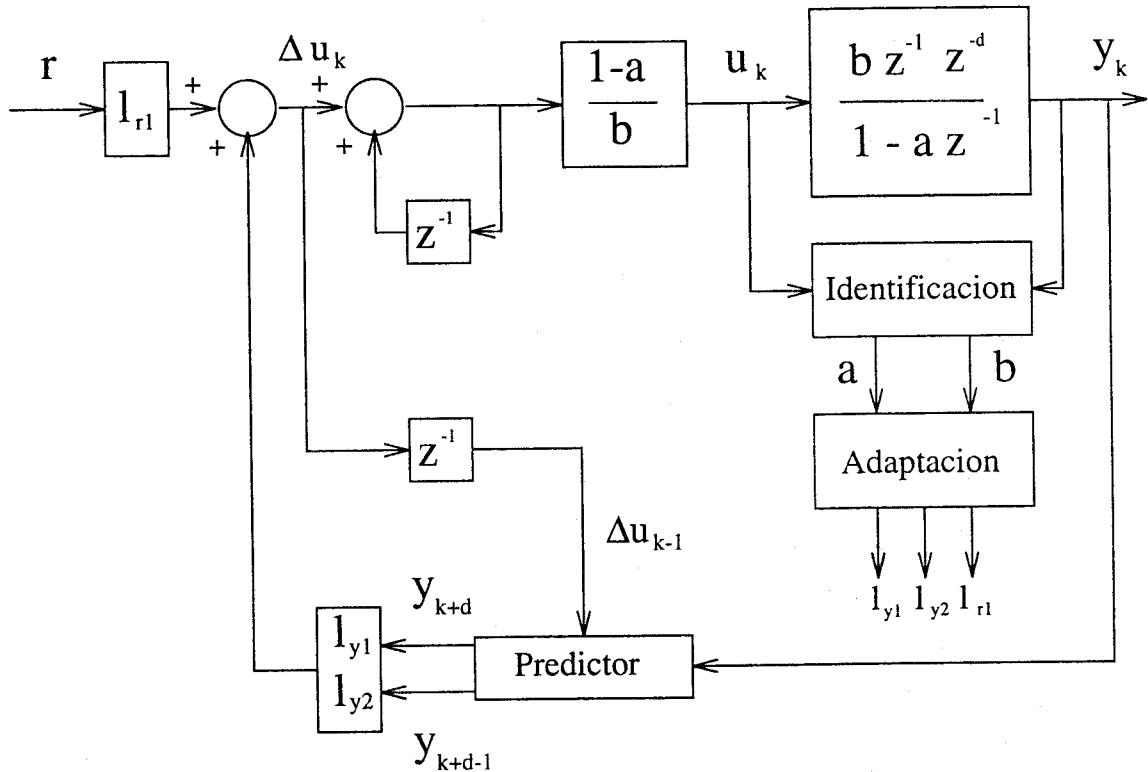


Figura 3.3: Esquema de control

Como se muestra en [14] se pueden obtener los coeficientes del controlador interpolando en un conjunto de valores previamente calculados, como se ve en la figura 3.4. Nótese que esto se puede llevar a cabo en este caso debido a que los coeficientes sólo dependen de un parámetro. El número de puntos precalculados depende de la variabilidad de los parámetros del proceso y de la exactitud que se desee. El conjunto no tiene que ser uniforme y se pueden calcular más puntos en aquellas zonas donde los coeficientes varíen sustancialmente para obtener una mejor aproximación o reducir el tamaño de almacenamiento requerido.

El predictor que se necesita en el algoritmo para calcular $\hat{y}(t+d | t)$, $\hat{y}(t+d-1 | t)$ se puede obtener aplicando la ecuación (3.6) secuencialmente para $j = 1 - d \dots 0$. En realidad consiste en un modelo del proceso que es proyectado hacia el futuro con los valores conocidos de entradas y salidas pasadas y sólo requiere un cálculo sencillo y directo.

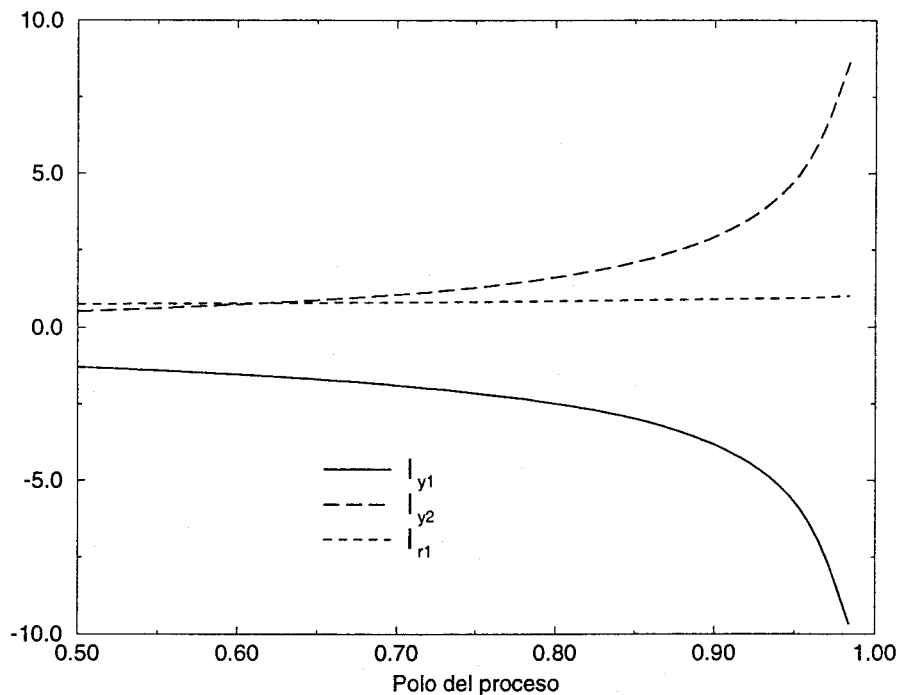


Figura 3.4: Coeficientes del controlador

3.2.3 Cálculo de los coeficientes del controlador

El algoritmo descrito anteriormente puede ser usado para el cálculo de los coeficientes del GPC para procesos que pueden ser descritos mediante la ecuación (3.4) (que, como se ha indicado permite describir a una gran cantidad de procesos industriales) en la zona de variación de los parámetros del proceso.

En la práctica el tiempo de muestreo usado en el computador de control dependerá de la respuesta temporal del proceso. En [39] se recomienda un tiempo de muestreo entre $1/15$ y $1/4$ de T_{95} (tiempo que emplea la salida del sistema en llegar al 95 por ciento de su valor final). Por consiguiente en la mayoría de los procesos cuando son muestreados con la frecuencia adecuada el valor del polo de la función de transferencia discreta se encontrará entre 0.5 y 0.95.

En la figura 3.4 aparecen los coeficientes (l_{y1} , l_{y2} y l_{r1}) calculados según se ha explicado previamente para unos pesos $\delta(i) = \delta^i$ y $\lambda(i) = \lambda$ con $\delta = 1$, $\lambda = 0.8$ y $N = 15$ en función del polo del proceso, el cual varía desde 0.5 hasta 0.99 con incrementos de 0.0056.

Una característica interesante de estos coeficientes es que debido a que la ganancia

del bucle cerrado debe ser la unidad, la suma de los tres valores debe ser cero para cualquier zona de trabajo. Por tanto sólo es necesario el cálculo de dos de ellos, por ejemplo l_{y1} y l_{y2} , viniendo l_{r1} dado como $-l_{y1} - l_{y2}$.

Los valores de los coeficientes en función del polo pueden ser almacenados en una tabla para distintos valores del polo del proceso. De esta forma su empleo requiere una búsqueda en la correspondiente tabla con interpolación si es necesario, lo que aparte de necesitar cierta cantidad de memoria para el almacenamiento, necesita cierto tiempo que puede ser crucial en muchas aplicaciones. Resulta pues más conveniente disponer de unas sencillas expresiones analíticas que relacionen los coeficientes del controlador con el polo del proceso. Observando la figura 3.4 se puede deducir que funciones de la siguiente forma pueden servir para relacionar los coeficientes con el polo:

$$l_{yi} = k_{1i} + k_{2i} \frac{a}{k_{3i} - a} \quad i = 1, 2 \quad (3.10)$$

Los coeficientes k_{ji} se pueden calcular mediante un ajuste por mínimos cuadrados haciendo uso de los valores calculados para diferentes valores del polo a . Se puede escribir la ecuación (3.10) como:

$$al_{yi} = l_{yi}k_{3i} - k_{1i}k_{3i} + a(k_{1i}a - k_{2i})$$

Si se repite esta ecuación para los N_p puntos usados para obtener la aproximación se tiene:

$$\begin{bmatrix} a^1 l_{yi}^1 \\ a^2 l_{yi}^2 \\ \vdots \\ a^{N_p} l_{yi}^{N_p} \end{bmatrix} = \begin{bmatrix} l_{yi}^1 & 1 & a^1 \\ l_{yi}^2 & 1 & a^2 \\ \vdots & \vdots & \vdots \\ l_{yi}^{N_p} & 1 & a^{N_p} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} e^1 \\ e^2 \\ \vdots \\ e^{N_p} \end{bmatrix} \quad (3.11)$$

donde l_{yi}^j , a^j , e^j , $j = 1 \dots N_p$ son los N_p valores del polo del sistema, los coeficientes precalculados y los errores de la aproximación respectivamente. con $x_1 = k_{3i}$, $x_2 = -k_{1i}k_{3i}$ y $x_3 = k_{1i}a - k_{2i}$.

En forma matricial la ecuación (3.11) resulta:

$$\mathbf{Y} = \mathbf{M} \mathbf{X} + \mathbf{E}$$

Los valores óptimos de \mathbf{X} serán los resultantes de la minimización de $\mathbf{E}^T \mathbf{E}$, dando lugar a:

$$\mathbf{X} = (\mathbf{M}^T \mathbf{M})^{-1} \mathbf{M}^T \mathbf{Y}$$

con lo que los coeficientes buscados son:

$$\begin{aligned}k_{3i} &= x_1 \\k_{1i} &= -x_2/k_{3i} \\k_{2i} &= k_{1i} - x_3\end{aligned}$$

En el caso de considerar un valor de $\lambda = 0.8$ y trabajando con un horizonte de control de 15, los coeficientes vienen dados por:

$$\begin{aligned}l_{y1} &= -0.845 - 0.564 \frac{a}{1.05-a} \\l_{y2} &= 0.128 + 0.459 \frac{a}{1.045-a} \\l_{r1} &= -l_{y1} - l_{y2}\end{aligned}$$

Estas expresiones proporcionan una buena aproximación a los verdaderos coeficientes, ajustándose al conjunto de valores precalculados con un error máximo menor que el uno por ciento de los valores nominales para el campo de interés de variación del polo.

3.2.4 Ponderación del esfuerzo de control

El esfuerzo de control aparece en la función de costo (3.5) ponderado por el factor λ y por tanto la señal de control de la ecuación (3.9) verá sus coeficientes influenciados por este valor. Mientras mayor sea λ menores variaciones en la señal de control se van a permitir. Si se le asigna un valor pequeño la respuesta del sistema será rápida ya que el controlador tiende a minimizar el error de seguimiento sin considerar el esfuerzo requerido para ello. Los coeficientes l_{y1} , l_{y2} y l_{r1} y por tanto los polos del bucle cerrado dependerán del valor de λ .

La figura 3.5 muestra el módulo del mayor polo en bucle cerrado cuando λ cambia de 0 a 0.8 con incrementos de 0.2. Como se observa, el módulo del mayor polo en bucle cerrado disminuye con λ , indicando que el sistema evoluciona más rápidamente, llegando a comportamientos del tipo *dead-beat* para $\lambda = 0$,

También se pueden obtener funciones analíticas que expresen la influencia de λ en los coeficientes del controlador. Para ello se considera que los k_{ij} de la ecuación (3.10) varían con λ , y se va a obtener una aproximación con funciones de la forma: $sign(k_{ij})e^{c_0 + c_1\lambda + c_2\lambda^2}$.

Aplicando logaritmos los coeficientes c_1, c_2 y c_3 pueden ser calculados por un ajuste

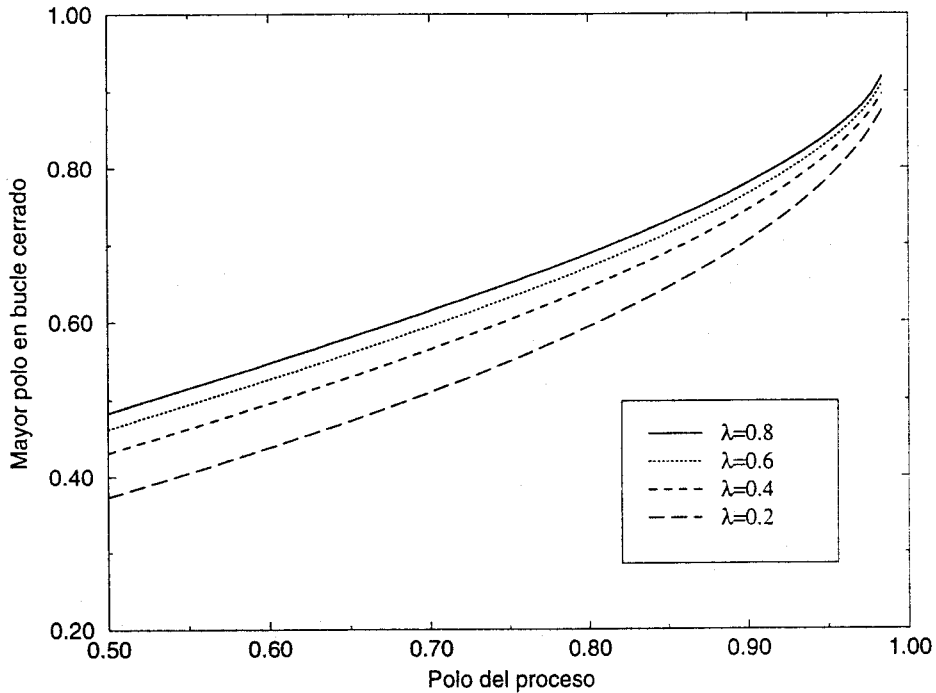


Figura 3.5: Influencia de la ponderación del esfuerzo de control

polinómico para valores de λ entre 0.2 y 1.2, obteniendo las siguientes expresiones:

$$\begin{aligned}
 k_{11} &= -e^{0.3598-0.9127\lambda+0.3165\lambda^2} \\
 k_{21} &= -e^{0.0875-1.2309\lambda+0.5086\lambda^2} \\
 k_{31} &= 1.05 \\
 k_{12} &= e^{-1.7383-0.40403\lambda} \\
 k_{22} &= e^{-0.32157-0.81926\lambda+0.3109\lambda^2} \\
 k_{32} &= 1.045
 \end{aligned} \tag{3.12}$$

Los valores de los coeficientes del controlador l_{y1} y l_{y2} obtenidos al introducir las k_{ij} anteriores en la ecuación (3.10) resultan ser buenas aproximaciones a los valores reales. El máximo error relativo para $0.55 < a < 0.95$ y $0.2 \leq \lambda \leq 1.2$ es menor que el 3 por ciento.

3.2.5 Algoritmo de Control

Una vez que se ha elegido el valor de λ , los términos k_{ij} se pueden calcular fácilmente usando las expresiones (3.12) para, usando (3.10) obtener los coeficientes del controlador en función del polo del proceso. Mientras que en el caso más simple de un

controlador fijo este proceso sólo debe realizarse una vez, en el caso de un control adaptativo, los pasos a seguir se reflejan en el siguiente algoritmo:

0. Calcular k_{ij} con las expresiones (3.12).
1. Llevar a cabo la identificación.
2. Hacer $l_{yi} = k_{1i} + k_{2i} \frac{\hat{a}}{k_{3i} - \hat{a}}$ para $i = 1, 2$ y $l_{r1} = -l_{y1} - l_{y2}$
3. Calcular $\hat{y}(t + d | t)$ e $\hat{y}(t + d - 1 | t)$ usando la ecuación (3.6) recursivamente.
4. Calcular la señal de control $u(t)$ con:

$$\Delta u(t) = l_{y1} \hat{y}(t + d | t) + l_{y2} \hat{y}(t + d - 1 | t) + l_{r1} r(t)$$
5. Dividir la señal de control por la ganancia estática.
6. Volver al paso 1.

En el caso de un controlador con coeficientes fijos el algoritmo se simplifica pues los coeficientes sólo se tienen que calcular una vez (salvo que se quiera cambiar λ durante su funcionamiento) y en cada instante de muestreo sólo es necesario realizar los pasos 3 y 4.

Obsérvese la simplicidad del algoritmo, que puede ser implantado en cualquier sistema de control industrial por pocas facilidades de programación que éste posea, y por personal sin profundos conocimientos de programación.

3.2.6 Un ejemplo

Con la intención de mostrar lo fácilmente que puede ser usado el método anterior, se presenta la aplicación al control de un proceso típico como puede ser un calentador de agua. Primero se lleva a cabo la identificación por el método de la curva de reacción y luego se aplica el GPC precalculado con las aproximaciones hechas anteriormente.

El proceso consiste básicamente el calentamiento de un determinado caudal de agua fría por la energía aportada por el combustible, cuyo caudal es regulado mediante una válvula de control. La variable de proceso es la temperatura de salida mientras que la variable manipulada es el caudal de combustible.

Los valores estacionarios de las variables son: temperatura de entrada 20°C , temperatura de salida 50°C y la válvula de combustible al 18.21 por ciento. Para la

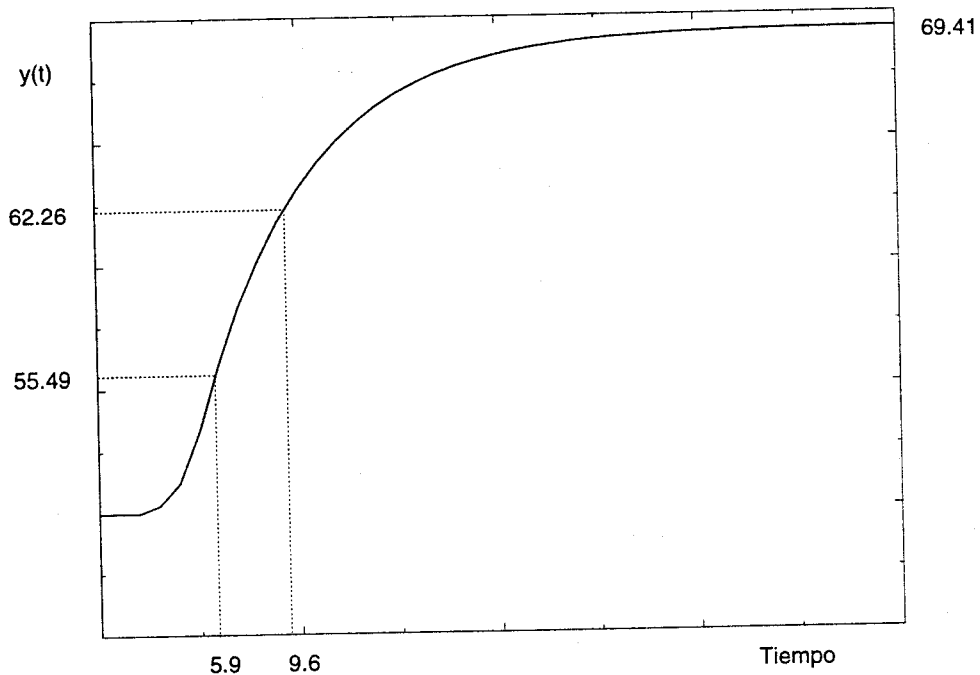


Figura 3.6: Evolución de la temperatura de salida ante un escalón

identificación de los parámetros del proceso en estas condiciones se cambia la apertura de la válvula a un 30 por ciento y se registra la evolución temporal de la salida, que obtiene un valor estacionario de $y = 69.411 \text{ } ^\circ\text{C}$ según se observa en la figura 3.6. Los parámetros del proceso pueden ser obtenidos directamente de esta respuesta tal como se explicó en la sección 3.1.1. Los tiempos medidos son $t_1 = 5.9$ y $t_2 = 9.6$ segundos, con lo que los parámetros del proceso resultan ser:

$$\begin{aligned} K &= \frac{\Delta y}{\Delta u} = \frac{69.411 - 50}{30 - 18.21} = 1.646 \\ \tau &= 1.5(t_2 - t_1) = 1.5(9.6 - 5.9) = 5.55 \\ \tau_d &= 1.5(t_1 - \frac{1}{3}t_2) = 1.5(5.9 - \frac{9.6}{3}) = 4.05 \end{aligned}$$

con estos valores, y usando un tiempo de muestreo de un segundo, el retardo se puede considerar como entero, obteniendo la siguiente función de transferencia discreta:

$$G(z^{-1}) = \frac{0.2713z^{-1}}{1 - 0.8351z^{-1}}z^{-4}$$

Como se está suponiendo que los parámetros del proceso no cambian, los coeficientes del controlador l_{y1} , l_{y2} y l_{r1} se calculan una sola vez. En este caso, eligiendo un valor de $\lambda = 0.8$, los k_{ij} se obtienen de la ecuación (3.12):

$$\begin{aligned}
 k_{11} &= -0.845 \\
 k_{21} &= -0.564 \\
 k_{31} &= 1.05 \\
 k_{12} &= 0.128 \\
 k_{22} &= 0.459 \\
 k_{32} &= 1.045
 \end{aligned}$$

que dan lugar a:

$$l_{y1} = -3.0367 \quad l_{y2} = 1.9541 \quad l_{r1} = 1.0826$$

En cada instante de muestreo sólo será necesario calcular las salidas predichas y la ley de control. Se necesitan las predicciones $\hat{y}(t+4 | t)$ e $\hat{y}(t+3 | t)$, que se calculan a partir de la siguiente ecuación con $i = 1 \dots 4$

$$\hat{y}(t+i | t) = (1+a)\hat{y}(t+i-1 | t) - a\hat{y}(t+i-2 | t) + b \Delta u(t+i-5)$$

y la ley de control, siendo G la ganancia estática (que no es la unidad) viene dada por:

$$u(t) = u(t-1) + (l_{y1}\hat{y}(t+4 | t) + l_{y2}\hat{y}(t+3 | t) + l_{r1}r)/G$$

La programación de este algoritmo en cualquier computador se traduce en un programa realmente simple. Para demostrar este hecho se incluye el trozo de código necesario en lenguaje C (figura 3.7). Se necesitan dos vectores \mathbf{u} e \mathbf{y} , el primero se usa para almacenar los valores pasados de la señal de control y el segundo para las salidas. Al ser el retardo de 4, los vectores serán:

$$\begin{aligned}
 \mathbf{y} &= [y(t-1), y(t), \hat{y}(t+1), \hat{y}(t+2), \hat{y}(t+3), \hat{y}(t+4)] \\
 \mathbf{u} &= [u(t), u(t-1), u(t-2), u(t-3), u(t-4), u(t-5)]
 \end{aligned}$$

Obsérvese que \mathbf{y} contiene las salidas predichas así como las salidas en los instantes t y $t-1$, ya que estos dos valores son necesarios en las dos primeras predicciones $\hat{y}(t+1)$, $\hat{y}(t+2)$.

La respuesta del bucle cerrado ante cambios en la referencia de $+10^\circ C$ se muestra en la figura 3.8, donde también se observa la evolución de la señal de control. Adicionalmente, la ponderación del esfuerzo de control λ se cambia en $t = 60$ del valor inicial de 0.8 a un valor más pequeño de 0.3; nótese que el esfuerzo de control aumenta


```

/* Predictor */
for (i=2; i<=5; i++)
y[i] = 1.8351*y[i-1] - 0.8351*y[i-2] + 0.2713*(u[6-i] - u[5-i]);

/* Ley de Control */
u[0] = u[1] + (- 3.0367*y[5] + 1.9541*y[4] + 1.08254*r) / 1.646 ;

/* Actualizacion */
for (i=5; i>0; i--) u[i] = u[i-1];
y[0] = y[1];

```

Figura 3.7: Código en lenguaje C del programa de control

y la salida tiende a ser más rápida. Por otro lado, si λ toma un valor mayor como 1.3 el comportamiento tiende a ser más lento; este cambio se lleva a cabo en el instante $t = 100$. Nótese que el cambio de λ se realiza sin excesivos cálculos, pues sólo hay que cambiar las k_{ij} ; sin embargo en el algoritmo estándar cualquier cambio de λ obligaría a invertir una nueva matriz (aunque los parámetros del proceso sean fijos).

3.3 Procesos con retardo no múltiplo del tiempo de muestreo

3.3.1 Modelo discreto del proceso

La ecuación (3.3) no se puede emplear cuando el retardo τ_d del proceso no es un múltiplo entero del tiempo de muestreo T ($dT \leq \tau_d \leq (d+1)T$) En esta situación el retardo fraccionario puede ser aproximado [31] por los dos primeros términos de la expansión de Padé y la función de transferencia discreta se convierte en:

$$G(z^{-1}) = \frac{b_0 z^{-1} + b_1 z^{-2}}{1 - a z^{-1}} z^{-d} \quad (3.13)$$

Como se observa esta función de transferencia es ligeramente diferente de la del modelo previo (equation(3.3)), pues presenta un cero adicional apareciendo en consecuencia un nuevo parámetro en el denominador. Aunque en muchas situaciones se

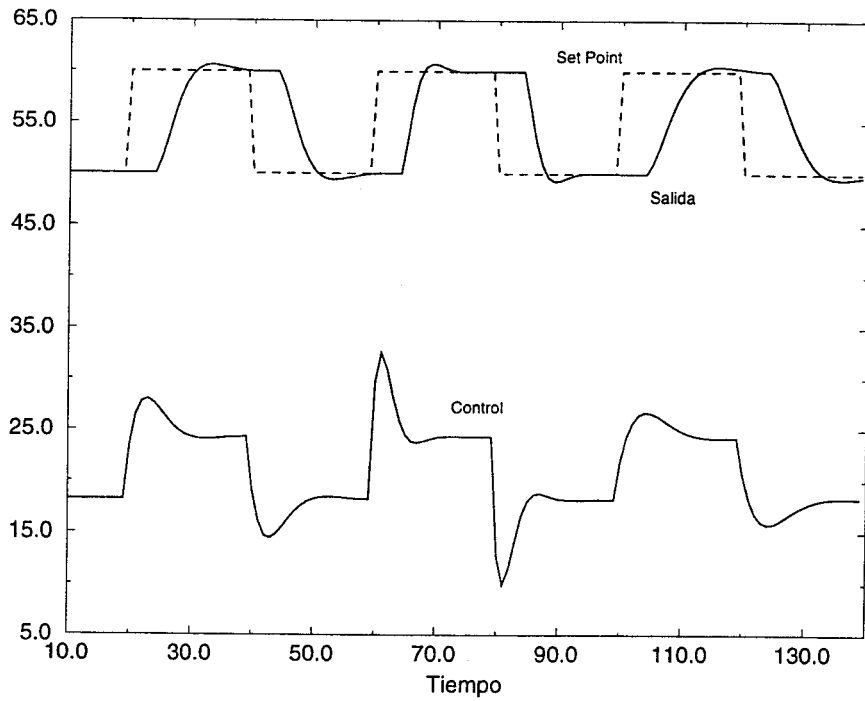


Figura 3.8: Respuesta del proceso ante cambios en la referencia y en λ .

puede jugar con el tiempo de muestreo de forma que el retardo sea entero, en otros casos éste será sin remedio un valor fraccionario. Para estos casos se va a utilizar un procedimiento análogo al presentado previamente para obtener un desarrollo fácil de implantar del GPC. Algunos pasos se omitirán por simplicidad, pues el desarrollo es análogo al de la sección previa.

El siguiente procedimiento se puede emplear para obtener los parámetros discretos a , b_0 y b_1 [31]: primero se descompone el retardo en $\tau_d = dT + \epsilon T$ con $0 < \epsilon < 1$. Entonces los parámetros son:

$$a = e^{-\frac{T}{\tau}} \quad b_0 = K(1-a)(1-\alpha) \quad b_1 = K(1-a)\alpha \quad \text{con } \alpha = \frac{a(a^{-\epsilon} - 1)}{1-a}$$

La función J que va a ser minimizada es también la de la ecuación (3.5). Usando el modelo CARIMA con el polinomio del ruido igual a 1, se puede escribir el sistema como:

$$(1 - az^{-1})y(t) = (b_0 + b_1z^{-1})z^{-d}u(t-1) + \frac{\varepsilon(t)}{\Delta}$$

que se puede transformar en:

$$y(t+1) = (1+a)y(t) - ay(t-1) + b_0 \Delta u(t-d) + b_1 \Delta u(t-d-1) + \varepsilon(t+1) \quad (3.14)$$

Si se conocen $\hat{y}(t+d+i-1|t)$ e $\hat{y}(t+d+i-2|t)$ entonces se ve claramente a partir de la ecuación (3.14), que el mejor valor esperado para $\hat{y}(t+d+i|t)$ viene dado por:

$$\begin{aligned} \hat{y}(t+d+i|t) &= (1+a)\hat{y}(t+d+i-1|t) - a\hat{y}(t+d+i-2|t) + \\ &+ b_0 \Delta u(t+i-1) + b_1 \Delta u(t+i-2) \end{aligned} \quad (3.15)$$

Si se introduce $\hat{y}(t+d+i|t)$ en la función a minimizar, $J(N)$ es función de $\hat{y}(t+d|t)$, $\hat{y}(t+d-1|t)$, $\Delta u(t+N_2-d-1)$, $\Delta u(t+N_2-d-2) \dots \Delta u(t)$, $\Delta u(t-1)$ y la trayectoria de referencia.

La minimización de $J(N)$ respecto a $\Delta u(t)$, $\Delta u(t+1) \dots \Delta u(t+N-1)$ conduce a

$$\mathbf{M} \mathbf{u} = \mathbf{P} \mathbf{y} + \mathbf{R} \mathbf{w} + \mathbf{Q} \Delta u(t-1) \quad (3.16)$$

donde

$$\begin{aligned} \mathbf{u} &= [\Delta u(t) \Delta u(t+1) \dots \Delta u(t+N-1)]^T \\ \mathbf{y} &= [\hat{y}(t+d|t) \hat{y}(t+d-1|t)]^T \\ \mathbf{w} &= [w(t+d+1) w(t+d+2) \dots w(t+d+N)]^T \end{aligned}$$

\mathbf{M} y \mathbf{R} son matrices de dimensión $N \times N$, \mathbf{P} de dimensión $N \times 2$ y \mathbf{Q} de $N \times 1$.

Obsérvese que el término $\mathbf{Q} \Delta u(t-1)$ no aparecía en el caso de retardo entero debido a los diferentes parámetros del proceso y por tanto ahora la ley de control no será la misma. Llamando q a la primera fila de la matriz \mathbf{M}^{-1} , $\Delta u(t)$ viene dado por:

$$\Delta u(t) = q \mathbf{P} \mathbf{y} + q \mathbf{R} \mathbf{w} + q \mathbf{Q} \Delta u(t-1)$$

Si se considera la trayectoria de referencia como

$$\mathbf{w} = [1 \dots 1] r(t)$$

el incremento en la señal de control $\Delta u(t)$ resulta:

$$\Delta u(t) = l_{y1} \hat{y}(t+d|t) + l_{y2} \hat{y}(t+d-1|t) + l_{r1} r(t) + l_{u1} \Delta u(t-1) \quad (3.17)$$

Donde $q \mathbf{P} = [l_{y1} \ l_{y2}]$, $l_{r1} = \sum_{i=1}^N q_i \sum_{j=1}^N r_{ij}$ y $l_{u1} = q \mathbf{Q}$. En la figura 3.9, se muestra el esquema de control resultante, donde los valores $\hat{y}(t+d|t)$, $\hat{y}(t+d-1|t)$ se obtienen usando el predictor descrito previamente, que es simplemente el modelo del proceso proyectado hacia el futuro y resulta muy simple de calcular.

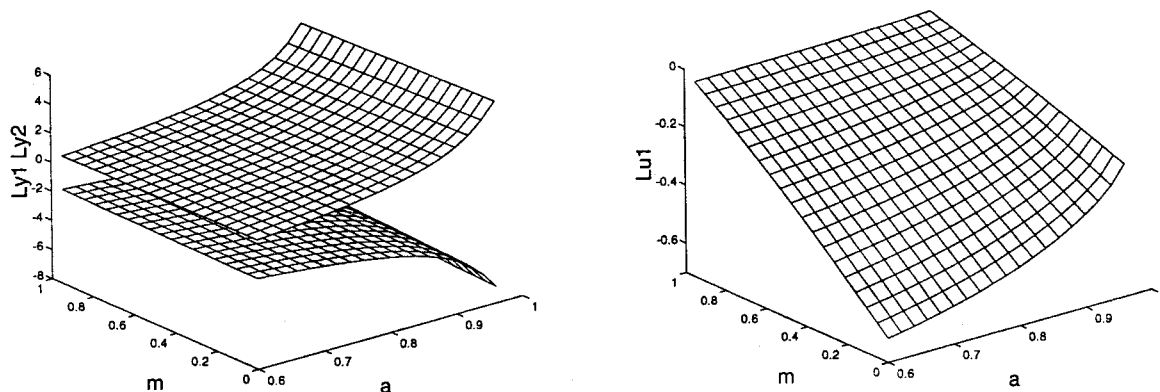


Figura 3.10: Parámetros l_{y1} , l_{y2} y l_{u1} en función de a y m

de m . La forma estándar de calcular los coeficientes sería a través de las matrices \mathbf{M} , \mathbf{P} , \mathbf{R} \mathbf{Q} y la resolución de la ecuación (3.16) para generar la ley de control (3.17). Esto requiere la triangularización de una matriz de $N \times N$ que involucra una gran carga de cálculo, determinante en su aplicación en tiempo real.

Tal como se indicó en el caso de procesos con retardo entero, los coeficientes se pueden obtener interpolando en un conjunto de valores previamente calculados, como se muestran en la figura 3.10 donde aparecen los valores de l_{y1} , l_{y2} y l_{u1} en función de a y m . Ahora hay que realizar la búsqueda en una tabla bidimensional lo cual puede necesitar más tiempo; interesará por tanto obtener unas aproximaciones analíticas de los coeficientes en función de los valores usuales de variación de a y m .

Se usará el algoritmo descrito anteriormente para la obtención de los coeficientes para la familia de procesos que puede ser descrita por (3.13) en el campo de interés de variación de los parámetros que la describen. Esta región va a estar delimitada por los valores del polo en el intervalo $[0.5, 0.95]$ y por el otro parámetro m variando entre 0 y 1.

Las curvas que se muestran en la figura 3.10 son las correspondientes a los coeficientes l_{y1} y l_{y2} para $\delta(i) = \delta^i$ y $\lambda(i) = \lambda$ con $\delta = 1$, $\lambda = 0.8$ y $N = 15$. Los coeficientes l_{y1} , l_{y2} y l_{r1} no son independientes entre sí, pues para mantener la ganancia estática del bucle cerrado en la unidad su suma debe ser cero. Este hecho implica que sólo hay que calcular tres de los cuatro coeficientes, resultando $l_{r1} = -l_{y1} - l_{y2}$.

Se van a aproximar los coeficientes por funciones de la forma:

$$k_{1i}(m) + k_{2i}(m) \frac{a}{k_{3i}(m) - a} \quad (3.18)$$

Los coeficientes $k_{ji}(m)$ dependen del valor de m y se pueden calcular por un ajuste de mínimos cuadrados usando los valores previamente calculados de l_i para diferentes valores de a y m , obteniendo polinomios de bajo orden que proporcionan una buena aproximación para las $k_{ji}(m)$.

En el caso de $m = 0.5$, $\lambda = 0.8$ y para un horizonte de control de 15 los coeficientes vienen dados por:

$$\begin{aligned} l_{y1} &= -0.9427 - 0.5486 \frac{a}{1.055-a} \\ l_{y2} &= 0.1846 + 0.5082 \frac{a}{1.0513-a} \\ l_{u1} &= -0.3385 + 0.0602 \frac{a}{1.2318-a} \\ l_{r1} &= -l_{y1} - l_{y2} \end{aligned}$$

Estas expresiones proporcionan una buena aproximación a los verdaderos coeficientes del controlador y se ajustan a los datos precalculados con un error máximo menor que el dos por ciento para el campo de interés de variación de los parámetros del proceso.

La influencia de la ponderación del esfuerzo de control λ en los coeficientes también debe ser tomada en cuenta. Los coeficientes son mayores para valores pequeños de λ para producir un esfuerzo de control mayor, por lo que este valor debe ser incluido en las funciones de aproximación. Con un procedimiento análogo al de la sección anterior, los valores de $k_{ij}(m)$ de las expresiones (3.18) se pueden aproximar por funciones de λ , obteniendo un error final cercano al tres por ciento.

El algoritmo en el caso adaptativo considera los cambios en los parámetros del proceso y se muestra a continuación:

1. Llevar a cabo la identificación.
2. Calcular $k_{ij}(m, \lambda)$.
3. Hacer $l_{yi}, l_{u1} = k_{1i} + k_{2i} \frac{a}{k_{3i}-a}$ y $l_{r1} = -l_{y1} - l_{y2}$
4. Calcular $\hat{y}(t+d|t)$ e $\hat{y}(t+d-1|t)$ con la ecuación (3.15).
5. Hallar el valor de $u(t)$ con:

$$\Delta u(t) = (l_{y1}\hat{y}(t+d|t) + l_{y2}\hat{y}(t+d-1|t) + l_{r1}r(t))/G + l_{u1} \Delta u(t-1)$$
6. Volver al paso 1.

Caudal	K	τ	τ_d	a	$b_0(\times 10^{-3})$	$b_1(\times 10^{-3})$	d
4.5	0.107	62.8	15.6	0.9234	7.2402	0.9485	3
4	0.112	46.56	15.65	0.8981	9.9901	1.4142	3
3.5	0.112	40.49	16.37	0.8838	9.6041	3.4067	3

Tabla 3.1: Parámetros del sistema para distintas condiciones de operación

3.3.3 Ejemplo

Se presenta a continuación un ejemplo de diseño y aplicación del controlador. Se trata de controlar la composición del destilado en una columna de destilación binaria y está tomado de [86]. La variable manipulada es el caudal de reflujo mientras que la variable controlada es la composición del destilado. Aunque es un proceso no lineal, puede ser descrito por un modelo de primer orden más retardo en diferentes puntos de trabajo. Esta hipótesis puede considerarse razonable pues de hecho una columna de destilación está compuesta de un cierto número de platos, cada uno de los cuales es un elemento de primer orden.

Como el proceso es no lineal, la respuesta varía para diferentes condiciones de operación, de tal manera que se obtienen distintos valores de los parámetros K , τ y τ_d al cambiar el caudal de reflujo de 3.5 a 4.5 mol/min (ver tabla 3.1). En base a estas condiciones, variaciones en los parámetros del proceso de $0.107 \leq K \leq 0.112$, $15.6 \leq \tau_d \leq 16.37$ y $40.49 \leq \tau \leq 62.8$ deberían ser consideradas (τ y τ_d en minutos).

Si se considera un tiempo de muestreo de cinco minutos, el retardo no es entero, con lo que se debe usar la función de transferencia (3.13). Los parámetros discretos se muestran en la misma tabla. El polo puede cambiar entre 0.8838 y 0.9234, mientras que el parámetro m lo hará entre 0.7381 y 0.8841.

Para hacer frente a las variaciones de la dinámica del sistema dependiendo del punto de operación, se puede diseñar un GPC con sus coeficientes variables según el punto de trabajo, es decir, un *gain-scheduling*. Para ello se calcula el conjunto de coeficientes y, dependiendo del caudal de reflujo, los coeficientes del controlador se obtendrán interpolando en dicho conjunto.

Para un valor fijo de λ , se usan los coeficientes $k_{ij}(m)$ para calcular los valores de $l_i(a, m)$. Para $\lambda = 0.8$, se pueden usar las siguientes aproximaciones:

$$\begin{aligned}
k_{11} &= 0.141 * m^2 - 0.125 * m - 0.920 \\
k_{21} &= -0.061 * m^2 + 0.202 * m - 0.625 \\
k_{31} &= -0.015 * m + 1.061 \\
k_{12} &= -0.071 * m^2 + 0.054 * m + 0.180 \\
k_{22} &= -0.138 * m + 0.575 \\
k_{32} &= -0.015 * m + 1.058 \\
k_{14} &= -0.115 * m^2 + 0.847 * m - 0.729 \\
k_{24} &= -0.113 * m + 0.112 \\
k_{34} &= -0.071 * m^2 + 0.091 * m + 1.181
\end{aligned}$$

Los coeficientes del GPC para diferentes condiciones de operación se incluyen en la tabla 3.2.

Caudal	l_{y1}	l_{y2}	l_{r1}	l_{u1}
4.5	-4.577	3.629	0.948	-0.035
4	-3.881	2.954	0.927	-0.039
3.5	-3.637	2.745	0.892	-0.091

Tabla 3.2: Coeficientes del controlador

Como d es de tres períodos de muestreo, el predictor es:

$$\hat{y}(t+i | t) = (1+a)\hat{y}(t+i-1 | t) - a\hat{y}(t+i-2 | t) + b_0 \Delta u(t+i-3) + b_1 \Delta u(t+i-4)$$

y la ley de control,

$$u(t) = u(t-1) + (l_{y1}\hat{y}(t+3 | t) + l_{y2}\hat{y}(t+2 | t) + l_{r1}r)/G + l_{u1} \Delta u(t-1)$$

donde G es la ganancia estática y los coeficientes del controlador y del predictor (que también varían) se calculan interpolando con el caudal, una vez filtrado por un filtro paso bajo, en la tablas anteriores.

Para mostrar el funcionamiento del sistema, se provocan cambios en el setpoint que cubren la región tratada, y se puede observar (figura 3.11) que la salida sigue fielmente a la referencia a pesar de los cambios en la dinámica del sistema debidos al cambio de punto de trabajo.

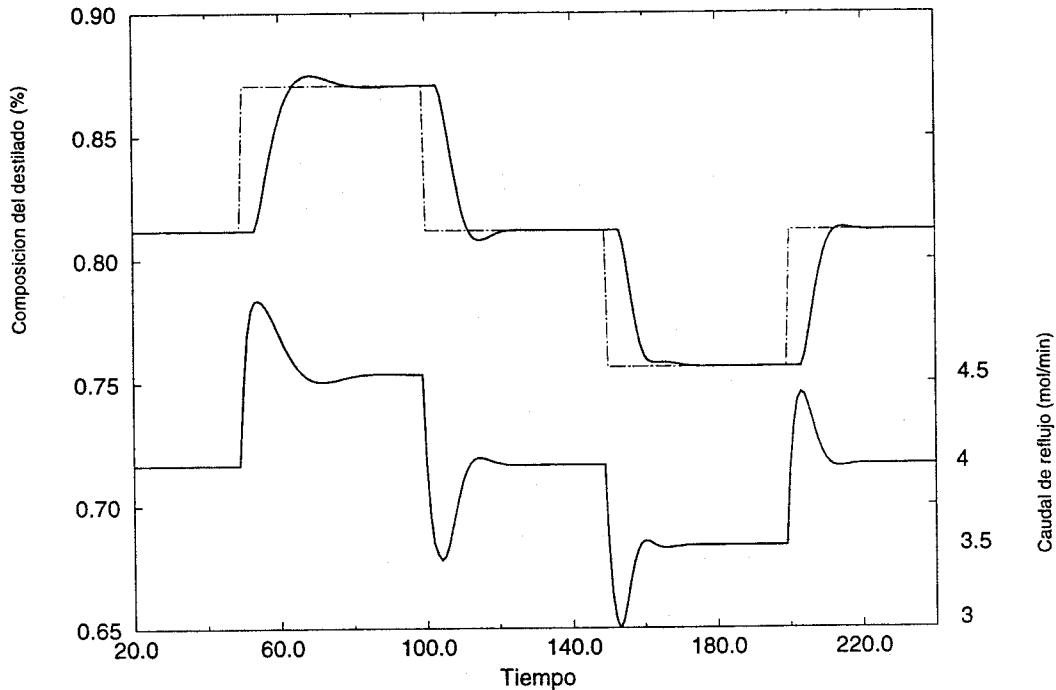


Figura 3.11: Funcionamiento con cambios en las condiciones de operación

3.4 Procesos con efecto integral

Es corriente encontrar en la práctica industrial cierta clase de procesos que incluyen un efecto integral. La salida de estos procesos crece indefinidamente cuando son excitados por una entrada en escalón. Este puede ser el caso de un depósito, donde el nivel crece continuamente siempre que el caudal de entrada sea mayor que el de salida. También el ángulo girado por un motor eléctrico crece al ser alimentado mientras la carga no iguale al par motor. El comportamiento de estos procesos difiere drásticamente del de los considerados hasta el momento en el capítulo. Estos procesos no pueden ser modelados ya por un sistema de primer orden más retardo puro, sino que se necesita la adición de un término $1/s$ para modelar el efecto integral. Por tanto la función de transferencia para este tipo de procesos será:

$$G(s) = \frac{K}{s(1 + \tau s)} e^{-\tau_d s} \quad (3.19)$$

En el caso general de retardo no múltiplo del tiempo de muestreo la función de transferencia discreta cuando se emplea un mantenedor de orden cero viene dada por:

$$G(z) = \frac{b_0 z^{-1} + b_1 z^{-2} + b_2 z^{-3}}{(1 - z^{-1})(1 - a z^{-1})} z^{-d} \quad (3.20)$$

en el caso más simple de que el retardo sea múltiplo entero b_2 desaparece.

En esta sección se va a obtener la ley de control del GPC para procesos que puedan ser descritos mediante la ecuación (3.19). Es importante destacar que otras formulaciones de control predictivo no son capaces de controlar este tipo de procesos pues emplean la respuesta truncada ante impulso o escalón, que no es válida para procesos inestables. Como el GPC hace uso de la función de transferencia, no aparecen problemas con sistemas inestables.

3.4.1 Obtención de la ley de control

El procedimiento para la obtención de la ley de control es análogo al usado en las secciones anteriores, aunque lógicamente el predictor será distinto y la expresión final variará ligeramente.

Usando el modelo CARIMA con el polinomio del ruido igual a 1, el sistema se puede escribir como:

$$(1 - z^{-1})(1 - az^{-1})y(t) = (b_0 + b_1z^{-1} + b_2z^{-2})z^{-d}u(t - 1) + \frac{\varepsilon(t)}{\Delta}$$

la cual se puede transformar en:

$$\begin{aligned} y(t + 1) &= (2 + a)y(t) - (1 + 2a)y(t - 1) + ay(t - 2) + \\ &+ b_0 \Delta u(t - d) + b_1 \Delta u(t - d - 1) + b_2 \Delta u(t - d - 2) + \varepsilon(t + 1) \end{aligned}$$

Si se conocen los valores de $\hat{y}(t + d + i - 1 | t)$, $\hat{y}(t + d + i - 2 | t)$ e $\hat{y}(t + d + i - 3 | t)$ la mejor salida predicha en el instante $t + d + i$ será:

$$\begin{aligned} \hat{y}(t + d + i | t) &= (2 + a)\hat{y}(t + d + i - 1 | t) - (1 + 2a)\hat{y}(t + d + i - 2 | t) + \\ &+ a\hat{y}(t + d + i - 3 | t) + b_0 \Delta u(t + i - 1) + b_1 \Delta u(t + i - 2) + b_2 \Delta u(t + i - 3) \end{aligned}$$

Ahora la función de costo es función de las salidas predichas $\hat{y}(t + d | t)$, $\hat{y}(t + d - 1 | t)$ e $\hat{y}(t + d - 2 | t)$, así como de las futuras acciones de control $\Delta u(t + N - 1)$, $\Delta u(t + N - 2)$... $\Delta u(t)$, y de las entradas pasadas $\Delta u(t - 1)$ y $\Delta u(t - 2)$ y, por supuesto, de la trayectoria de referencia.

La minimización de $J(N_1, N_2, Nu)$ da lugar a la siguiente ecuación matricial para el cálculo de \mathbf{u} :

$$\mathbf{M} \mathbf{u} = \mathbf{P} \mathbf{y} + \mathbf{R} \mathbf{w} + \mathbf{Q}_1 \Delta u(t - 1) + \mathbf{Q}_2 \Delta u(t - 2)$$

donde \mathbf{M} y \mathbf{R} son matrices de dimensión $N \times N$, \mathbf{P} de dimensión $N \times 2$ and \mathbf{Q}_1 y \mathbf{Q}_2 de $N \times 1$. Siguiendo la misma nomenclatura que en las secciones anteriores \mathbf{u} es el vector de las futuras señales de control e \mathbf{y} el de las salidas predichas.

El primer elemento del vector \mathbf{u} se puede obtener llamando q a la primera fila de la matriz \mathbf{M}^{-1} :

$$\Delta u(t) = q \mathbf{P} \mathbf{y} + q \mathbf{R} \mathbf{w} + q \mathbf{Q}_1 \Delta u(t-1) + q \mathbf{Q}_2 \Delta u(t-2)$$

Si se asume que la referencia va a ser constante sobre el horizonte de predicción e igual al setpoint actual:

$$\mathbf{w} = [1 \dots 1]r(t+d)$$

resulta que la ley de control es:

$$\begin{aligned} \Delta u(t) &= l_{y1} \hat{y}(t+d|t) + l_{y2} \hat{y}(t+d-1|t) + l_{y3} \hat{y}(t+d-2|t) + l_r r(t+d) + \\ &+ l_u \Delta u(t-1) + l_{u2} \Delta u(t-2) \end{aligned} \quad (3.21)$$

siendo $q \mathbf{P} = [l_{y1} \ l_{y2} \ l_{y3}]$, $l_{r1} = \sum_{i=1}^N (q_i \sum_{j=1}^N r_{ij})$, $l_{u1} = q \mathbf{Q}_1$ y $l_{u2} = q \mathbf{Q}_2$.

Por consiguiente la ley de control resulta ser una expresión lineal que depende de seis coeficientes que a su vez dependen de los parámetros del proceso (excepto del retardo) y de la ponderación que se le asigne al esfuerzo de control en la función de costo, λ . Además, uno de estos coeficientes es una combinación lineal de los otros, puesto que se debe cumplir la siguiente relación para obtener un bucle cerrado con ganancia estática unidad:

$$l_{y1} + l_{y2} + l_{y3} + l_r = 0$$

3.4.2 Coeficientes del controlador

La ley de control (3.21) resulta muy fácil de aplicar siempre que se conozcan los coeficientes l_{y1} , l_{y2} , l_{y3} , l_{r1} , l_{u1} y l_{u2} . Resulta pues crucial disponer de expresiones que relacionen estos coeficientes con los parámetros del sistema. De forma similar a lo visto anteriormente, se van a obtener simples relaciones analíticas.

Como el proceso es modelado por (3.20), se necesitan cuatro parámetros (a , b_0 , b_1 y b_2) para describirlo. Se van a obtener ahora relaciones simples entre los coeficientes del controlador y los parámetros del sistema para el caso en que el tiempo muerto es entero, es decir, $b_2 = 0$.

De forma similar a los procesos sin integrador, se puede considerar que el sistema cumple $(b_0 + b_1)/(1 - a) = 1$, para trabajar con procesos normalizados. Posteriormente los valores calculados deberán ser divididos por este valor que en general será distinto de la unidad.

Se van a obtener los coeficientes del controlador como función del polo a y de un parámetro:

$$n = \frac{b_0}{b_0 + b_1}$$

Este parámetro se mueve en un margen muy estrecho para cualquier proceso. Como b_0 y b_1 están relacionados con los parámetros continuos por (ver[4]):

$$b_0 = K(T + \tau(-1 + e^{-\frac{T}{\tau}})) \quad b_1 = K(\tau - e^{-\frac{T}{\tau}}(T + \tau)) \quad (3.22)$$

entonces

$$n = \frac{a - 1 - \log a}{(a - 1) \log a}$$

que para los valores más usuales del polo del sistema va a variar entre $n = 0.5$ y $n = 0.56$. Por consiguiente se pueden expresar los coeficientes del controlador como funciones del polo y n para un valor fijo de λ .

La figura 3.12 muestra la forma de los coeficientes para un valor fijo de $\lambda = 1$. Se puede ver que los coeficientes dependen principalmente del polo a , siendo prácticamente independientes de n salvo en el caso de l_{u1} . Funciones de la forma

$$f(a, n, \lambda) = k_1(n, \lambda) + k_2(n, \lambda) \frac{a}{k_3(n, \lambda) - a}$$

donde k_i pueden ser aproximadas por:

$$\begin{aligned} k_{y1,1} &= -e^{0.955-0.559\lambda+0.135\lambda^2} \\ k_{y1,2} &= -e^{0.5703-0.513\lambda+0.138\lambda^2} \\ k_{y1,3} &= 1.0343 \\ k_{y2,1} &= e^{0.597-0.420\lambda+0.0953\lambda^2} \\ k_{y2,2} &= e^{1.016-0.4251\lambda+0.109\lambda^2} \\ k_{y2,3} &= 1.0289 \\ k_{y3,1} &= -e^{-1.761-0.422\lambda+0.071\lambda^2} \\ k_{y3,2} &= -e^{0.103-0.353\lambda+0.089\lambda^2} \\ k_{y3,3} &= 1.0258 \\ k_{u1,1} &= 1.631n - 1.468 + 0.215\lambda - 0.056\lambda^2 \\ k_{u1,2} &= -0.124n + 0.158 - 0.026\lambda + 0.006\lambda^2 \\ k_{u1,3} &= 1.173 - 0.019\lambda \end{aligned} \quad (3.23)$$

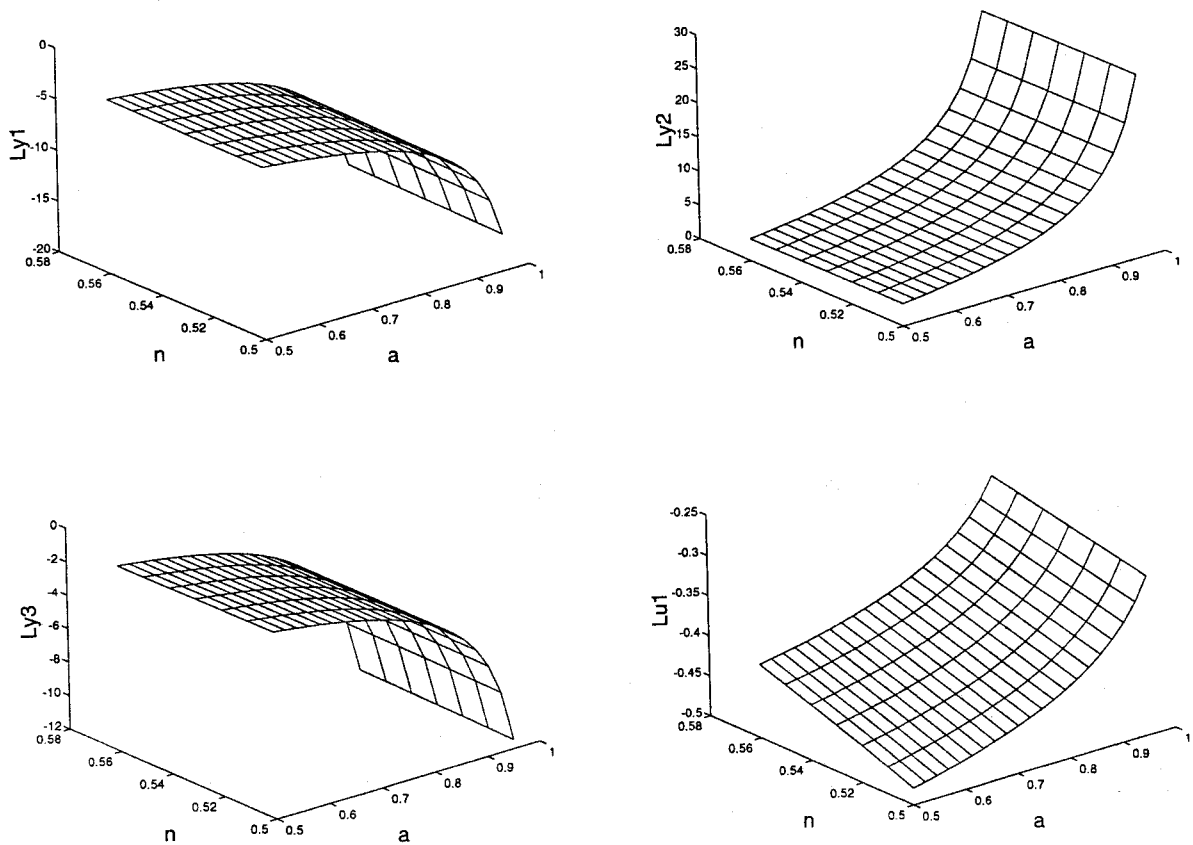


Figura 3.12: Coeficientes l_{y1} , l_{y2} , l_{y3} y l_{u1}

proporcionan buenas aproximaciones para l_{y1} , l_{y2} , l_{y3} , l_{r1} y l_{u1} en la zona normal de variación de los parámetros del proceso. Obsérvese que no se suministra una función aproximada para l_{r1} , ya que es linealmente dependiente de los otros coeficientes. Estas funciones se adaptan a los datos calculados con un error máximo menor del 1.5 por ciento de los valores nominales. Se pueden obtener aproximaciones aún más precisas si el campo de variabilidad de los parámetros del proceso es menor.

3.4.3 Ejemplo de aplicación

La ley de control (3.21) va a ser aplicada a un proceso muy común como es un motor de corriente continua. Si la entrada del proceso es la tensión de alimentación (U) y la salida es el ángulo girado por el eje (θ) está claro que el proceso posee un efecto integral, ya que la posición crece indefinidamente mientras se alimenta con cierta tensión. Para

la obtención de un modelo que describa el comportamiento del motor, se tienen en cuenta el par de inercia (proporcional a la aceleración angular) y el par de fricción dinámica (proporcional a la velocidad angular) cuya suma se iguala al par desarrollado por el motor, que depende de la tensión que se le aplique. Se trata de un sistema de primer orden respecto de la velocidad, pero si se considera la posición como la salida del proceso, el sistema es de segundo orden:

$$J \frac{d^2\theta}{dt^2} + f \frac{d\theta}{dt} = M_m$$

y la función de transferencia será:

$$\frac{\theta(s)}{U(s)} = \frac{K}{s(1 + \tau s)}$$

donde K y τ son parámetros constructivos del motor.

El controlador se va a implantar sobre un motor real con tensión de alimentación de 24 Voltios e intensidad nominal de 1.3 Amperios, sometido a una carga constante. Para la obtención experimental de los parámetros del motor se utiliza el método de la curva de reacción, provocando un escalón en la tensión de alimentación y midiendo la evolución de la velocidad angular (que es un sistema de primer orden). Los parámetros obtenidos son:

$$K = 2.5 \quad \tau = 0.9 \text{ segundos}$$

y retardo cero. Tomando un tiempo de muestreo de $T = 0.06$ segundos se obtiene la función de transferencia discreta:

$$G(z) = \frac{0.004891z^{-1} + 0.004783z^{-2}}{(1 - z^{-1})(1 - 0.935507z^{-1})}$$

Para un valor de $\lambda = 2$ los parámetros de la ley de control, calculados según (3.23) son:

$$\begin{aligned} l_{y1} &= -11.537 \\ l_{y2} &= 19.242 \\ l_{y3} &= -8.207 \\ l_{u1} &= -0.118 \\ l_{r1} &= 0.502 \end{aligned}$$

Los resultados del seguimiento de la posición del eje a cambios en escalón de la referencia se muestran en la figura 3.13, donde se observa que apenas existe sobreoscilación. El sistema posee una zona muerta de forma que es insensible ante señales de control menores que 0.7 V; para evitar esto se añade una no-linealidad que obliga a la señal de control a mantenerse siempre por encima de este umbral.

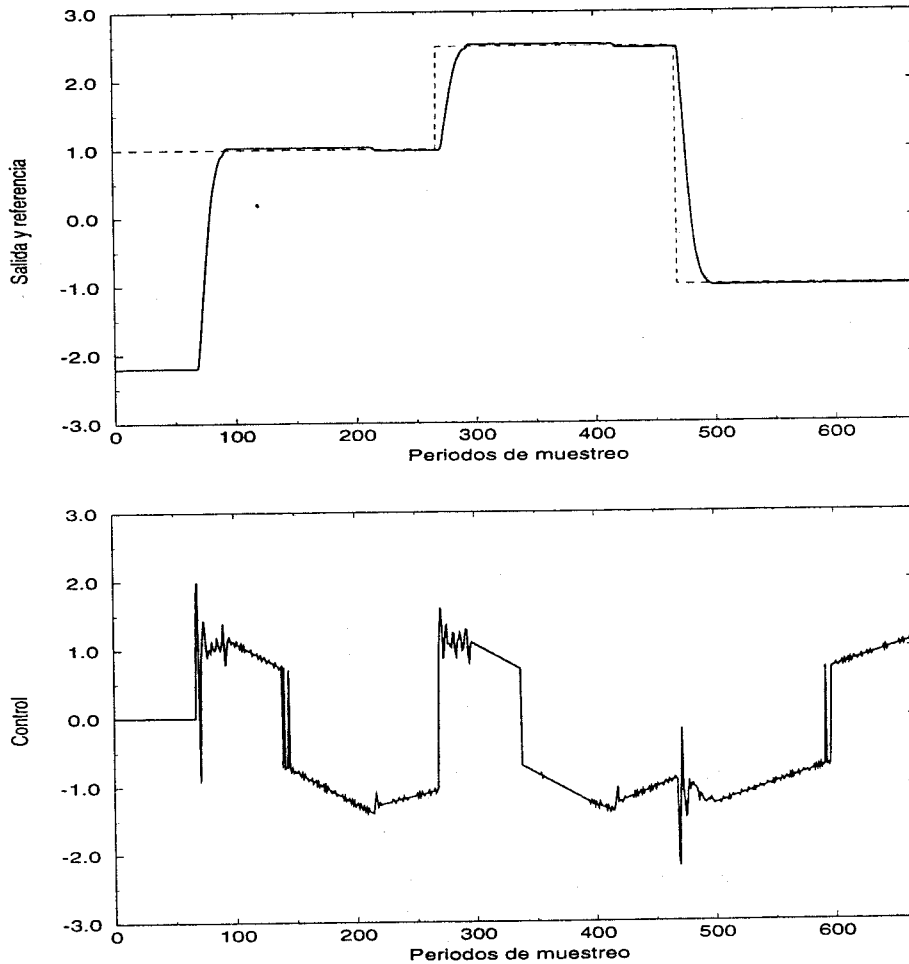


Figura 3.13: Evolución del motor ante cambios en la referencia

Es importante considerar que el tiempo de muestreo es muy pequeño (0.06 segundos) lo que podría hacer imposible la implantación del algoritmo GPC estándar. Sin embargo, debido a la sencilla formulación aquí empleada, la implantación se reduce al cálculo de la expresión (3.21) y apenas consume tiempo en cualquier computador por poco potente que este sea.

El proceso puede ser perturbado por la adición de un freno electromagnético que además del propio efecto de aumento de carga hace cambiar la constante de fricción del motor y por consiguiente los parámetros usados en el modelo no van a ser buenos. Aún así, se ve en la figura 3.14 que el GPC controla el sistema razonablemente bien a pesar de la aparición de alguna sobreoscilación.

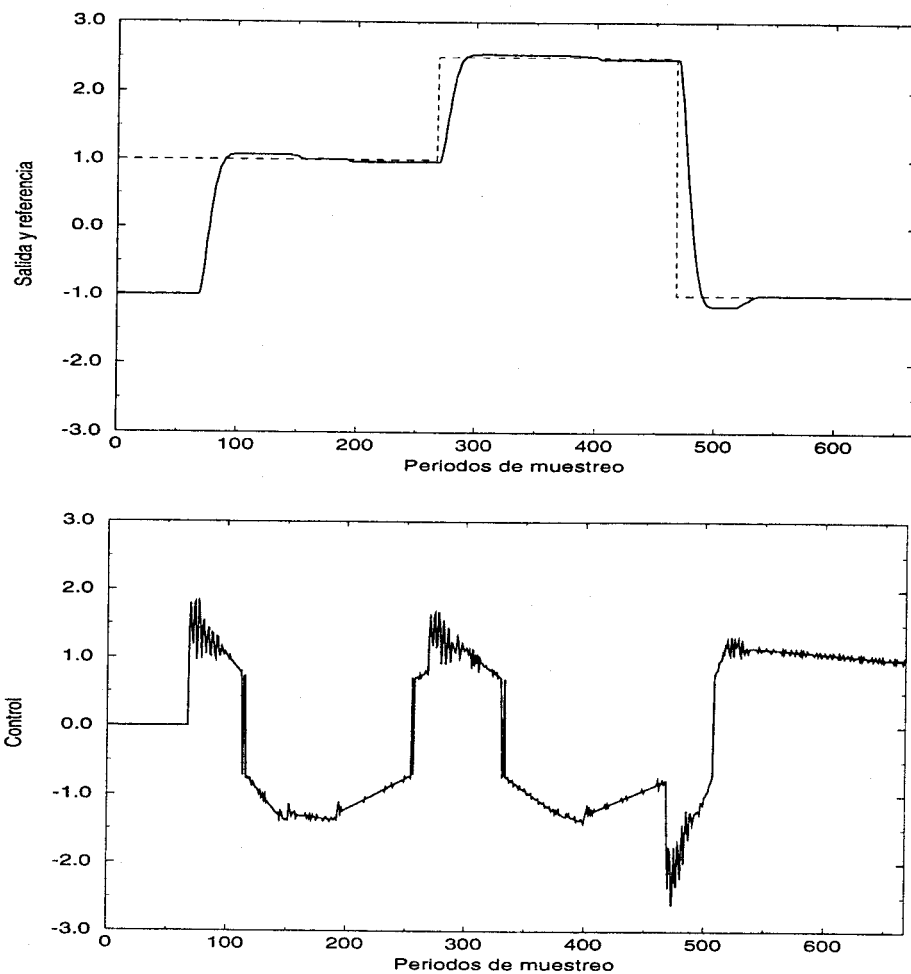


Figura 3.14: Evolucion del motor con freno electromagnético

3.5 Consideración de referencias en rampa

Es usual que la señal de referencia de un proceso se mantenga en un valor constante y se mueva a otros valores también constantes durante la operación normal de la planta. Esto es lo que se ha considerado hasta el momento, es decir, $w(t + d + 1) = w(t + d + 2) \dots = r(t)$, siendo $r(t)$ el setpoint en el instante t que se mantendrá en un valor fijo.

Pero la referencia no evolucionará así en todas las situaciones. En muchas ocasiones puede evolucionar como una rampa, cambiando suavemente a otro valor constante. En general será deseable que la salida siga una trayectoria combinada de rampas y escalones.

Esta situación se da con frecuencia en diversos procesos industriales. De hecho es corriente en la industria alimentaria o farmacéutica la necesidad de mantener un perfil de temperaturas dado por escalones y rampas. También, por ejemplo, en el control de motores o en aplicaciones de robótica interesa que la posición o la velocidad sigan evoluciones de ese tipo.

Por tanto se va a reformular el cálculo del GPC cuando la referencia es una rampa, definida por un parámetro α que indica el incremento (o decremento si es negativo) en cada período de muestreo. Por tanto ahora la trayectoria de referencia es:

$$\begin{aligned} w(t+d+1) &= r(t+d) + \alpha \\ w(t+d+2) &= r(t+d) + 2\alpha \\ &\dots \dots \\ w(t+d+N) &= r(t+d) + N\alpha \end{aligned}$$

Con el procedimiento empleado a lo largo de este capítulo y para sistemas de primer orden con retardo, se llega a que

$$\mathbf{M} \mathbf{u} = \mathbf{P} \mathbf{y} + \mathbf{R} \mathbf{w} + \mathbf{Q} \Delta u(t-1)$$

donde \mathbf{M} y \mathbf{R} son matrices de dimensión $N \times N$, \mathbf{P} de dimensión $N \times 2$ y \mathbf{Q} de $N \times 1$:

$$\begin{aligned} \mathbf{u} &= [\Delta u(t) \ \Delta u(t+1) \ \dots \ \Delta u(t+N-1)]^T \\ \mathbf{y} &= [\hat{y}(t+d|t) \ \hat{y}(t+d-1|t)]^T \\ \mathbf{w} &= [w(t+d+1) \ w(t+d+2) \ \dots \ w(t+d+N)]^T \end{aligned}$$

Si q es la primera fila de la matriz \mathbf{M}^{-1} , entonces $\Delta u(t)$ viene dada por

$$\Delta u(t) = q \mathbf{P} \mathbf{y} + q \mathbf{R} \mathbf{w} + q \mathbf{Q} \Delta u(t-1) \quad (3.24)$$

Si al producto de la fila q por la matriz \mathbf{R} se le denomina \mathbf{h} , que es un vector de N elementos, el término de la ecuación anterior que incluye a la referencia es $\mathbf{h} \mathbf{w}$, que toma la forma:

$$\mathbf{h} \mathbf{w} = \sum_{i=1}^N h_i r(t+d+i) = \sum_{i=1}^N h_i (r(t+d) + \alpha i) = \sum_{i=1}^N h_i r(t+d) + \alpha \sum_{i=1}^N h_i i$$

por tanto se puede escribir

$$\mathbf{h} \mathbf{w} = l_{r1} r(t+d) + \alpha l_{r2}$$

Con lo que la ley de control queda:

$$\Delta u(t) = l_{y1}\hat{y}(t+d|t) + l_{y2}\hat{y}(t+d-1|t) + l_{r1}r(t+d) + \alpha l_{r2} + l_{u1} \Delta u(t-1)$$

Siendo $q \mathbf{P} = [l_{y1} \ l_{y2}]$, $l_{u1} = q \mathbf{Q}$, $l_{r1} = \sum_{i=1}^N (q_i \sum_{j=1}^N r_{ij})$, $l_{r2} = \alpha \sum_{i=1}^N h_i$. Por tanto la ley de control es lineal, dependiendo ahora de un coeficiente más: l_{r2} debido exclusivamente a la referencia en rampa. Resulta evidente que si la rampa degenera en un escalón ($\alpha = 0$) se obtiene la ley de control ya conocida.

Por consiguiente la única consecuencia de la consideración de referencias en rampa es la adición de un nuevo parámetro en la ley de control, mientras que los otros permanecen inalterados. El predictor sigue siendo el mismo y el algoritmo de resolución no difiere del visto previamente. El nuevo parámetro l_{r2} es independiente del retardo y sólo depende como los otros de los parámetros del proceso (a y m) y del esfuerzo de control λ , pudiendo obtenerse fácilmente una aproximación analítica de la misma forma que los demás parámetros. Se sigue cumpliendo la relación $l_{y1} + l_{y2} + l_{r1} = 0$, por lo que sólo hay que calcular realmente cuatro de los cinco valores.

Se han considerado hasta ahora las situaciones de proceso sin integrador, proceso con integrador, referencia constante y referencia en rampa, apareciendo en cada uno de ellos un coeficiente nuevo en la ley de control.

Para englobar todas estas situaciones, la ley de control tendrá la siguiente expresión general:

$$\begin{aligned} \Delta u(t) = & l_{y1}\hat{y}(t+d|t) + l_{y2}\hat{y}(t+d-1|t) + l_{y3}\hat{y}(t+d-2|t) + \\ & + l_{r1}r(t+d) + \alpha l_{r2} + l_{u1} \Delta u(t-1) + l_{u2} \Delta u(t-2) \end{aligned}$$

Algunos de los coeficientes que aparecen en la expresión anterior serán nulos según el caso considerado. Los parámetros del controlador para cada situación se muestran en la tabla 3.3.

3.5.1 Aplicación

Como aplicación del caso de referencias en rampa, se va a considerar el motor del ejemplo anterior con una señal de referencia que es una combinación de escalones y rampas, definida por un valor nulo de α en las zonas constantes y un valor distinto de cero en las rampas.

Proceso	Referencia	l_{y1}	l_{y2}	l_{y3}	l_{u1}	l_{u2}	l_{r1}	l_{r2}
$\frac{k}{1+\tau s} e^{-\tau_d s}$ τ_d entero	Constante	×	×	0	0	0	×	0
	Rampa	×	×	0	0	0	×	×
$\frac{k}{1+\tau s} e^{-\tau_d s}$ τ_d no entero	Constante	×	×	0	×	0	×	0
	Rampa	×	×	0	×	0	×	×
$\frac{k}{s(1+\tau s)} e^{-\tau_d s}$ τ_d entero	Constante	×	×	×	×	0	×	0
	Rampa	×	×	×	×	0	×	×
$\frac{k}{s(1+\tau s)} e^{-\tau_d s}$ τ_d no entero	Constante	×	×	×	×	×	×	0
	Rampa	×	×	×	×	×	×	×

Tabla 3.3: Coeficientes que aparecen en la ley de control en cada caso. El × indica que el coeficiente existe

Se usan los mismos valores de los parámetros del controlador que en el ejemplo anterior con la adición de $l_{r2} = 2.674$. Como $(b_0 + b_1)/(1 - a) = 0.15$, la ley de control viene dada por:

$$\Delta u(t) = -76.92y(t) + 128.29y(t-1) - 54.72y(t-2) + 3.34r(t) + 17.82\alpha - 0.12\Delta u(t-1)$$

Como el retardo del sistema puede considerarse nulo, las salidas predichas son conocidas en el instante t . Los resultados obtenidos se muestran en la figura 3.15, donde se observa cómo la posición del motor sigue a la referencia en rampa sin problemas.

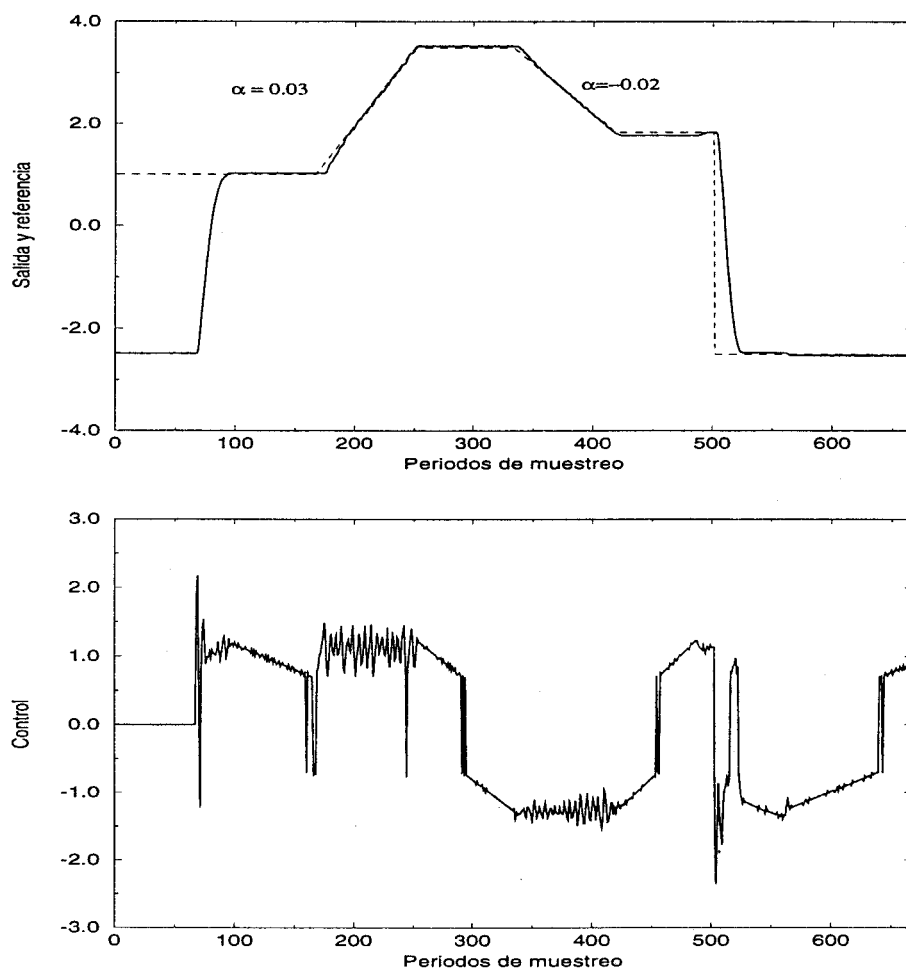


Figura 3.15: Evolucion de la salida ante referencia variable

Capítulo 4

Robustez frente a incertidumbres estructuradas

4.1 Incertidumbres

La elaboración de modelos matemáticos de procesos del mundo real requiere la adopción de simplificaciones. No existe en la práctica ningún modelo matemático que pueda describir exactamente un proceso físico. Independientemente de la técnica de diseño utilizada, los controladores son siempre diseñados basados en la información (necesariamente incompleta) sobre el comportamiento dinámico del proceso. Sea el proceso descrito mediante un sistema de ecuaciones diferenciales parciales o simplemente mediante una ganancia y una constante de tiempo, siempre hay que tener presente cómo los errores de modelado pueden afectar adversamente al comportamiento del sistema controlado. Se desea que el controlador sea insensible a estas incertidumbres en el modelo, es decir, se pretende que sea *robusto*.

Aunque el concepto de robustez parece de gran utilidad práctica, ha estado prácticamente ausente de la literatura de control desde los años sesenta a los ochenta. Como indica Morari [60] quizás ésta sea una de las razones por las que las técnicas de diseño desarrolladas en ese tiempo no hayan tenido una mayor aceptación en la práctica industrial. Se pretende pues tratar el tema de la robustez del controlador presentado en el capítulo previo. No se trata de diseñar el controlador utilizando criterios de robustez, sino que se pretende determinar la que éste posee. En cualquier caso, existen desarrollos de controladores predictivos diseñados usando criterios robustos, por ejemplo Campo [17] o Papavasiliou [67] y variaciones del algoritmo GPC estándar para

mejorar la robustez [22]. También existen desarrollos de controladores predictivos de horizonte deslizante usando un prefiltrado para asegurar altos márgenes de robustez a altas frecuencias [92].

Aunque ningún modelo pueda representar el proceso perfectamente, resulta evidente que algunos lo harán con mayor fidelidad que otros. De todas formas, ningún modelo se podrá considerar completo sin tener en cuenta los posibles errores de modelado; a estos errores se le llamarán incertidumbres y existen diversas formas de representarlas. Las representaciones de las incertidumbres varían según la estructuralidad que contengan, reflejando de algún modo el conocimiento de los mecanismos físicos que causan la discrepancia entre el modelo y el proceso así como la capacidad de formalizar estos mecanismos de forma que puedan ser manipulados.

En muchos casos las incertidumbres se pueden expresar en forma estructurada, como expresiones en función de unos determinados parámetros que puedan ser considerados en la función de transferencia [32]. Sin embargo, siempre suele haber errores remanentes a altas frecuencias que no pueden modelarse así, constituyendo lo que serían incertidumbres no estructuradas [30]. En este capítulo se van a considerar sólo las primeras, dejándose para el siguiente el estudio de las incertidumbres no estructuradas.

Como se vió en el Capítulo 3, el modelo de sistema con que se trabaja (primer orden con retardo puro) a pesar de su simplicidad es válido para una gran mayoría de procesos. Pero resulta fundamental considerar el caso en que el modelo no sea capaz de describir completamente toda la dinámica del proceso real. Es decir, el proceso se acerca *razonablemente* al modelo pero hay ocasiones en las que éste no lo describe correctamente.

Al considerar el modelo con ganancia, constante de tiempo y retardo puro de la forma

$$G(s) = \frac{K}{1 + \tau s} e^{-s\tau_d} \quad (4.1)$$

se puede tener en cuenta la influencia de incertidumbres en los tres parámetros que definen el sistema. Como ya se indicó, en tiempo discreto y para el caso más general de retardo no múltiplo del tiempo de muestreo, la función de transferencia del modelo es

$$G(z^{-1}) = \frac{b_0 z^{-1} + b_1 z^{-2}}{1 - a z^{-1}} z^{-d} \quad (4.2)$$

donde la ganancia estática es $\frac{b_0+b_1}{1-a}$, el polo es a y el retardo es d .

Se van a considerar por un lado incertidumbres en el valor de los parámetros del modelo sin cambiar la estructura de éste (es decir, cada parámetro podrá encontrarse dentro de un determinado intervalo) y por otro lado la influencia de dinámica inmodelada, que alterará el orden del modelo empleado.

4.1.1 Incertidumbres paramétricas

Se va a abordar en esta sección el estudio de las posibles incertidumbres en los parámetros que definen el proceso pero sin alterar el orden de éste. Es decir, la estructura del proceso va a ser la misma que la del modelo pero con parámetros distintos de los "reales". Habrá pues un conjunto de parámetros estimados, que estarán relacionados con los reales mediante una serie de constantes. En el caso de que no existieran incertidumbres, estos dos conjuntos serían idénticos. Para ver los parámetros que se van a tener en cuenta, la función de transferencia del sistema, ecuación (4.2), se puede poner de la forma:

$$G(z^{-1}) = \frac{b_0 + b_1}{1 - a} \frac{(1 - a)(mz^{-1} + (1 - m)z^{-2})}{1 - az^{-1}} z^{-d} \quad (4.3)$$

donde $G_e = \frac{b_0+b_1}{1-a}$, es la ganancia estática, y el resto es un sistema de ganancia estática unidad. El parámetro m , como se vió en el capítulo anterior, sirve para modelar el retardo fraccionario entre d y $d + 1$ y por consiguiente variará entre 0 y 1. Por tanto se van a considerar en este apartado errores en la estimación de a , G_e y m . El error en la estimación del retardo se va a abordar en el apartado siguiente, considerándolo como una dinámica inmodelada ya que hace cambiar el orden del sistema, aunque también se podría haber considerado como una incertidumbre paramétrica. Antes de estudiar la influencia de los errores en la estimación sobre el proceso con el controlador, se verá cuál es la discrepancia entre proceso y modelo considerados en bucle abierto para tener una idea de qué representa realmente cada incertidumbre.

Incetidumbre en el polo

Como se sabe, el polo está relacionado con la constante de tiempo ($a = e^{-\frac{T}{\tau}}$), por tanto aquí se tratan los errores en la estimación de la constante de tiempo del proceso. Para modelar esta incertidumbre, se considera que el proceso tiene el polo en $\alpha \times a$

mientras que el modelo lo tiene en a . En este caso y manteniendo la ganancia estática del proceso, éste puede ser modelado por:

$$G(z^{-1}) = \frac{(1 - \alpha a)(b_0 z^{-1} + b_1 z^{-2})}{(1 - a)(1 - \alpha a z^{-1})} z^{-d}$$

en consecuencia, para valores de α mayores que la unidad, el modelo trabaja con constantes de tiempo mayores que las reales, ocurriendo lo contrario para valores menores que la unidad. Gráficamente se puede ver la influencia del error en la estimación del polo en la figura 4.1a, donde se muestran las respuestas ante una entrada en escalón del modelo y del proceso con α mayor y menor que 1.

Incertidumbre en la ganancia

Otro factor clave que caracteriza el comportamiento del proceso es la ganancia estática. Errores en su estimación pueden conducir a comportamientos del proceso distintos de los del modelo. Se va suponer que el proceso puede ser modelado por una ganancia γG_e , siendo $G_e = \frac{b_0 + b_1}{1 - a}$. La ganancia suele tener un considerable efecto sobre la estabilidad al cerrar el bucle, debido a su influencia sobre el margen de fase. Será pues interesante ver cómo altera el funcionamiento del proceso una vez éste está siendo controlado. Haciendo variar el valor de γ alrededor de la unidad, obtenemos distintos comportamientos, como se puede observar en la figura 4.1b. La función de transferencia que modela el proceso será pues:

$$G(z^{-1}) = \frac{\gamma(b_0 z^{-1} + b_1 z^{-2})}{(1 - a z^{-1})} z^{-d}$$

Incertidumbre en el valor de m

El parámetro m aparece cuando se quieren considerar retardos que no son múltiplos del período de muestreo. En este caso, haciendo una aproximación de Padé, se puede aproximar el retardo entre d y $d + 1$ por el polinomio $(b_0 z^{-1} + b_1 z^{-2})z^{-d}$. Poniendo la función de transferencia según (4.3), está claro que haciendo variar m entre 0 y 1, se consiguen representar los retardos fraccionarios. La figura 4.1c refleja este hecho. En la estimación de este parámetro se va a considerar también un posible error, por lo que se trabajará con βm . El valor de β se moverá en el entorno de la unidad, aunque imponiendo siempre que el valor \hat{m} esté por debajo de 1, ya que en caso contrario se

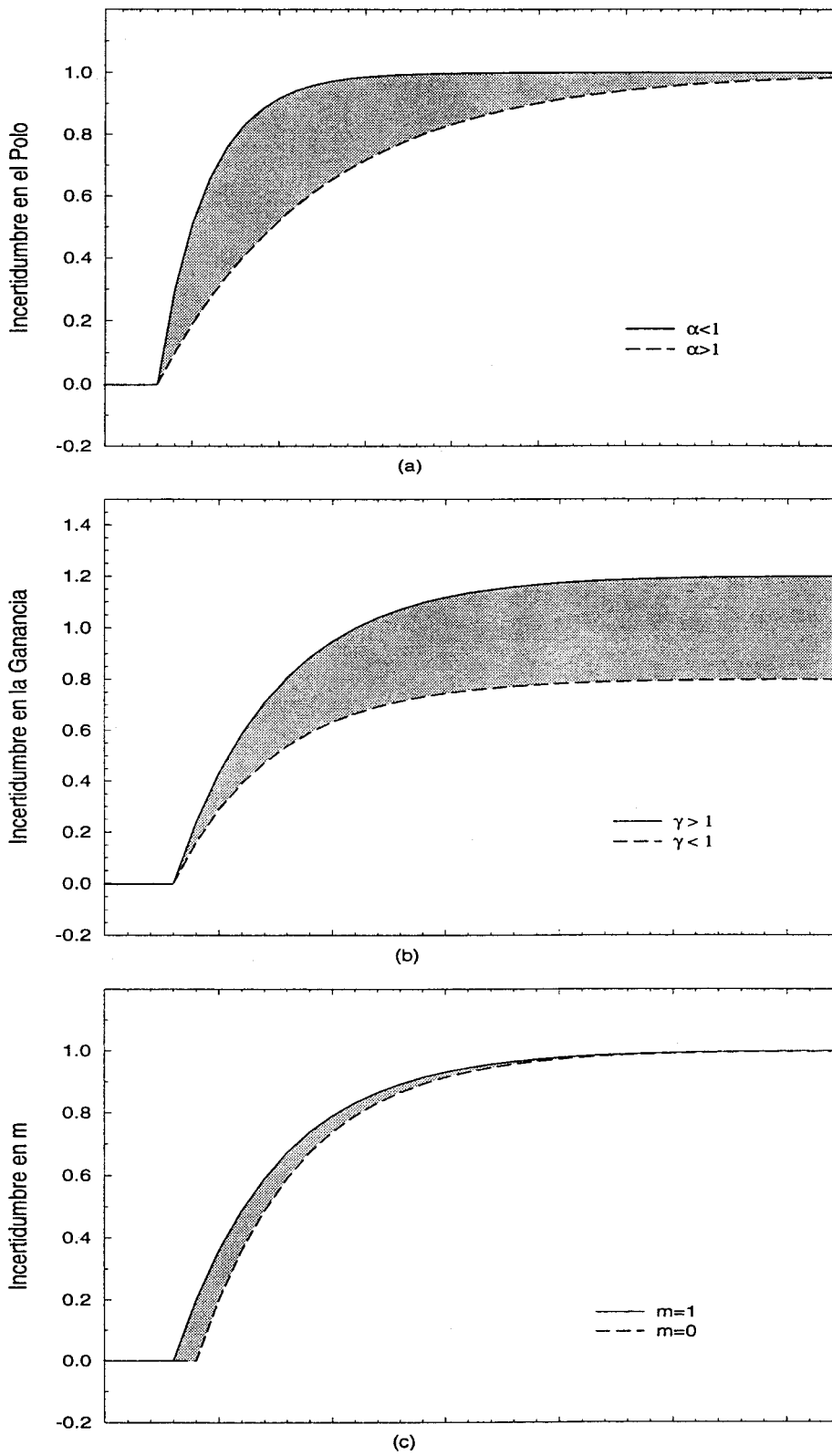


Figura 4.1: Incertidumbres en los parámetros del sistema

estaría trabajando fuera de la zona entre d y $d + 1$. Si esto ocurriera, correspondería al caso de error en la estimación del retardo, que se estudiará más adelante. Nótese además que la ganancia estática no se ve alterada por la estimación de este parámetro, ya que se debe cumplir que $\hat{b}_0 + \hat{b}_1 = b_0 + b_1$.

4.1.2 Dinámica inmodelada

Como se ha indicado previamente, existe otro tipo de incertidumbres estructuradas distintas a la puramente paramétricas; son aquellas que hacen variar el orden del modelo que describe el proceso. Son incertidumbres que poseen cierta *estructura*, ya que pueden modelarse mediante la adición de nuevos términos a la función de transferencia del modelo, pero estos términos no son simples parámetros, sino que incluyen dinámica. A continuación el estudio se va a centrar en los factores que se consideran más relevantes, como son la existencia de un nuevo polo sin modelar y los cambios en el retardo.

Sistema con un polo inmodelado

Se va a considerar aquí el caso de que el proceso real tenga otro polo aparte del que aparece en el modelo. Es decir, si el modelo tiene el polo en a , el proceso tendrá además otro situado en $k \times a$, siendo k una constante entre 0 y 1. Por consiguiente el proceso puede ser descrito mediante la siguiente función de transferencia:

$$G(z^{-1}) = \frac{(b_0 z^{-1} + b_1 z^{-2})(1 - ka)}{(1 - az^{-1})(1 - kaz^{-1})} z^{-d} \quad (4.4)$$

Se puede observar que, aparte del nuevo término $1 - kaz^{-1}$, aparece en el denominador el valor $1 - ka$. Este término se añade para que no haya discrepancia entre la ganancia estática del proceso y la del modelo.

La constante k tomará siempre valores inferiores a la unidad, es decir, el polo inmodelado será siempre menos dominante que el conocido (que es el que aparece en el modelo). Esto será así derivado de la circunstancia de que al identificar un sistema se pierde con más facilidad la información de alta frecuencia (polos más rápidos). La figura 4.2 puede servir de ejemplo; en ella se muestra la respuesta ante una entrada en escalón de un sistema con un solo polo (que correspondería al modelo usado), otro con un polo diez veces menor que el anterior y el resultante con los dos polos. Se

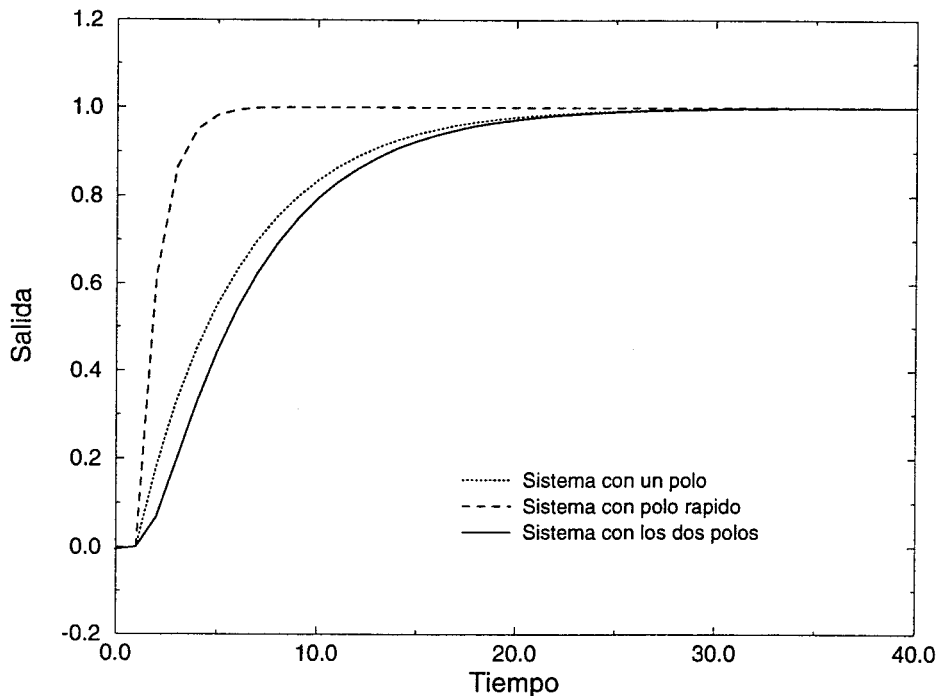


Figura 4.2: Respuesta ante escalón para sistemas con distintos polos

observa que el comportamiento del sistema resultante se parece más al del polo mayor, quedando la dinámica del polo rápido enmascarada.

La discrepancia de comportamiento entre el modelo y el sistema real con el polo añadido en $k \times a$ se aprecia con más claridad observando la respuesta en frecuencia mediante el diagrama de Bode. En la Figura 4.3 se muestra la respuesta del modelo con polo en $a = 0.8$ (línea continua) y el sistema con otro polo en $k \times a = 0.1$ (línea discontinua). Se puede apreciar que así como el comportamiento es prácticamente idéntico para frecuencias bajas, la discrepancia tanto en módulo como en fase se hacen notar para frecuencias mayores. Se pone de manifiesto un denominador común de la dinámica inmodelada cual es el aumento de la incertidumbre con la frecuencia.

Sistemas con estimación errónea del retardo

La otra circunstancia que se plantea en este capítulo es la incertidumbre en el valor del retardo, es decir, el modelo trabaja con un valor de retardo mientras que el proceso real puede tener otro ligeramente diferente. De este modo, el modelo viene dado por la ecuación (4.5) mientras que el proceso estará modelado mediante:

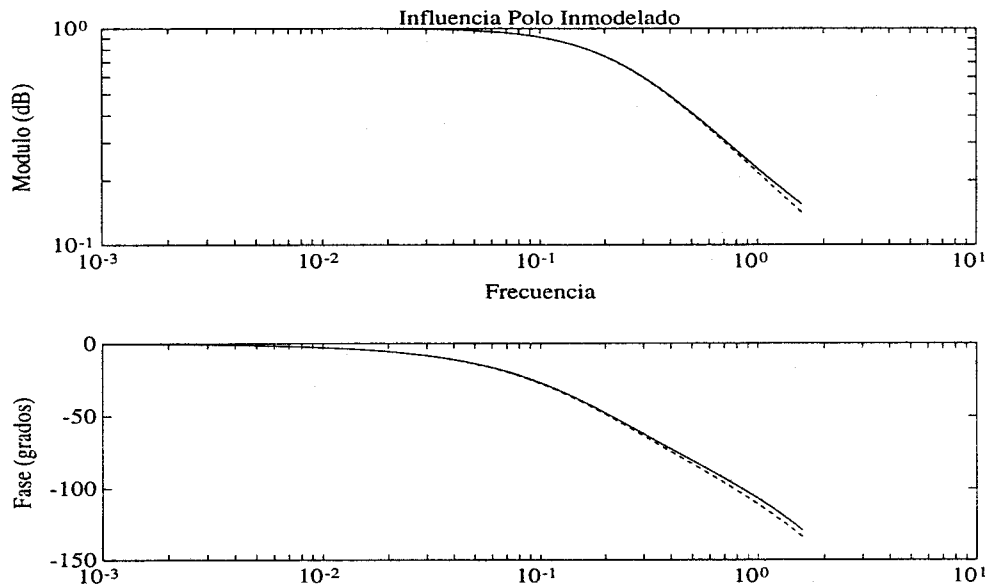


Figura 4.3: Respuesta en frecuencia con polo inmodelado

$$G(z^{-1}) = \frac{b_0 z^{-1} + b_1 z^{-2}}{1 - a z^{-1}} z^{-D} \quad (4.5)$$

siendo D el retardo real que puede ser mayor o menor que d , en general $D = d + i$, con $i = \pm 1, \pm 2, \dots, \pm N_d$. Como se mencionó en la sección anterior, este hecho también se podría considerar como incertidumbre paramétrica, pero se ha considerado más aconsejable su inclusión aquí puesto que los cambios en el retardo hacen cambiar el orden del sistema debido al término z^i .

Esta situación puede darse por diversos motivos, por ejemplo:

- Conocimiento erróneo del retardo desde un principio.
- Variación del retardo a lo largo de la vida del proceso. Por ejemplo el retardo en un intercambiador de calor puede aumentar por acumulación de suciedad en los tubos.
- Cambios en el régimen de funcionamiento. Los retardos debidos a fenómenos de transporte varían cuando cambia el caudal.

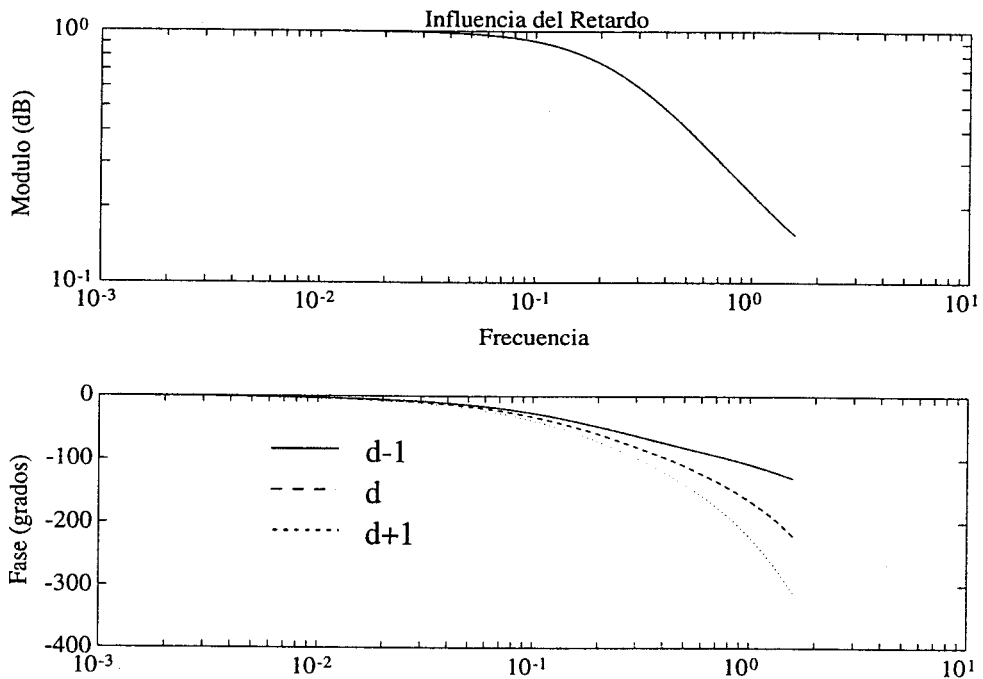


Figura 4.4: Respuesta en frecuencia para diferentes retardos

El retardo es un factor que no es fácil de conocer exactamente (sobre todo si es variable) y cuya influencia en el comportamiento de los controladores es importante; de hecho, muchas técnicas de control ampliamente aceptadas como el regulador de Mínima Varianza, muestran un comportamiento pobre o incluso inestable si el retardo no se conoce exactamente [31] [22]. Aún así en la mayoría de los casos se supone perfectamente conocido. De hecho, los algoritmos más ampliamente utilizados de estimación de parámetros suelen considerar el retardo conocido y constante [81].

Como es bien sabido, el retardo es un factor que afecta mucho a la estabilidad. Su variación afecta exclusivamente a la fase de la respuesta en frecuencia, dejando el módulo inalterado. Como puede desprenderse de la figura 4.4, donde se muestra la respuesta en frecuencia de un sistema cuando se incrementa y decreta el retardo en una unidad, la fase se muestra claramente alterada, afectando este hecho al margen de fase y por tanto a la estabilidad.

4.1.3 Ambos tipos de incertidumbres

Si se quiere tener en cuenta el caso más completo en que el proceso presente todas las incertidumbres descritas anteriormente, entonces puede ser modelado por la siguiente función de transferencia:

$$G(z^{-1}) = \frac{\gamma(1 - \alpha a)(1 - k\alpha a)(b_0 + b_1)(\beta m z^{-1} + (1 - \beta m)z^{-2})}{(1 - a)(1 - \alpha a z^{-1})(1 - k\alpha a z^{-1})} z^{-D} \quad (4.6)$$

A partir de aquí se va a suponer el caso más genérico (α, β y γ distintos de la unidad, k distinto de cero y D distinto de d), anulando en los distintos casos las incertidumbres que no aparezcan.

4.2 Estabilidad en presencia de incertidumbres estructuradas

Una vez descritos los tipos de incertidumbres que se van a estudiar, se presenta en este apartado un estudio de su influencia sobre el comportamiento del proceso (con sus incertidumbres) operando con un controlador que ha sido desarrollado para el modelo ideal. Es decir, tanto el predictor como los parámetros l_i son calculados para un modelo que no coincide exactamente con el proceso real que se va a controlar.

Se plantea la siguiente pregunta: ¿qué discrepancias se permiten entre el proceso y el modelo para que el sistema controlado sea estable? Para resolverla se presenta aquí una herramienta de cálculo que permite conocer la estabilidad al trabajar con un controlador desarrollado para un modelo definido por los parámetros (a, b_0, b_1, d) cuando el proceso presenta incertidumbres. En definitiva se va a aplicar en este apartado el método que permitirá ver si el proceso real (relacionado con el modelo por los parámetros α, β, γ, k y D) en bucle cerrado es estable o no.

Para llevar a cabo el cálculo de la estabilidad del sistema en bucle cerrado será necesario estudiar los polos del conjunto formado por el proceso y el controlador desarrollado para el modelo. El control queda definido por la ecuación:

$$\Delta u(t) = l_{y1} \hat{y}(t + d | t) + l_{y2} \hat{y}(t + d - 1 | t) + l_{r1} r(t) + l_{u1} \Delta u(t - 1) \quad (4.7)$$

los parámetros del controlador l_{y1}, l_{y2}, l_{r1} y l_{u1} han sido precalculados para el modelo (no para el proceso pues lógicamente éste no se conoce). Análogamente el predictor

trabaja con los parámetros del modelo, aunque se va actualizando con los valores leídos de la salida producida por el proceso real. Ahora pues para el cálculo de la función de transferencia en bucle cerrado no se puede considerar $\hat{y}(t+d | t) = z^d y(t)$, sino que hay que desarrollar el predictor como función de los valores $u(t)$ e $y(t)$ reales del proceso. Este desarrollo se llevará a cabo a continuación.

4.2.1 Desarrollo del predictor

Se necesita pues un predictor que proporcione la salida predicha en los instantes $t+d$ y $t+d-1$ con los datos disponibles en el instante t , es decir, con los valores de la señal de control anteriores a t y con los valores leídos de la salida hasta el instante t inclusive, ver figura 4.5.

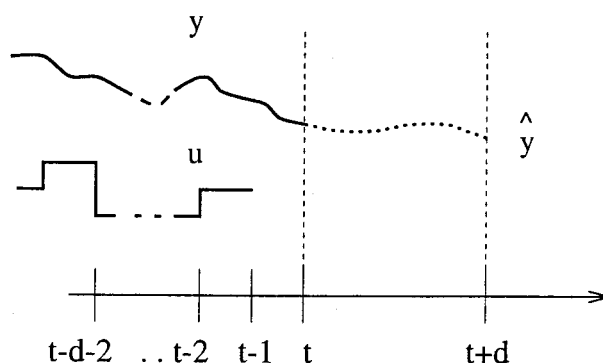


Figura 4.5: Predictor

Por tanto se pretende llegar a:

$$\hat{y}(t+d | t) = E^d(z^{-1})y(t) + F^d(z^{-1})u(t)$$

siendo E^d y F^d dos polinomios en z^{-1} que dependerán del valor del retardo. Para encontrar estos polinomios se tomará la ecuación en diferencias que define el modelo del sistema y se aplicará secuencialmente hasta obtener el valor predicho en $t+d$ en función de los conocidos en t . En [23] el predictor es calculado recursivamente, de forma que los polinomios que definen $\hat{y}(t+j)$ son obtenidos en función de los del instante $j-1$. Aquí se va a formalizar el predictor de otra forma diferente, de manera que los polinomios anteriores puedan ser calculados directamente para d y $d-1$, sin tener que calcular todos los valores previos. De este modo, partimos de que el modelo puede ser escrito de la forma:

$$(1 - az^{-1})y(t) = (b_0 + b_1z^{-1})z^{-d}u(t-1) + \frac{\varepsilon(t)}{\Delta}$$

pudiendo transformarse en

$$y(t+1) = (1+a)y(t) - ay(t-1) + b_0 \Delta u(t-d) + b_1 \Delta u(t-d-1) + \varepsilon(t+1) \quad (4.8)$$

como la componente de ruido está en el futuro, el predictor óptimo será:

$$\hat{y}(t+1) = (1+a)y(t) - ay(t-1) + b_0 \Delta u(t-d) + b_1 \Delta u(t-d-1) \quad (4.9)$$

análogamente para el instante $t+2$

$$\hat{y}(t+2) = (1+a)\hat{y}(t+1) - ay(t) + b_0 \Delta u(t-d+1) + b_1 \Delta u(t-d)$$

Si se sustituye $\hat{y}(t+1)$ por su valor dado en la ecuación (4.9), y avanzando en el tiempo hasta $t+d$ se llega a

$$\hat{y}(t+d) = e^d_0 y(t) + e^d_1 y(t-1) + f^d_1 u(t-1) + f^d_2 u(t-2) + \dots + f^d_{d+2} u(t-d-2)$$

Por consiguiente, usando la transformada en z , se puede escribir

$$\hat{y}(t+d) = E^d(z^{-1})y(t) + F^d(z^{-1})u(t) \quad (4.10)$$

El polinomio E^d es siempre de grado 1, ya que en el cálculo de la salida predicha en $t+d$ sólo intervienen las salidas en t y $t-1$, mientras que el polinomio F^d es de grado $d+2$; por tanto se necesitan las señales $u(t-d-2)$, $u(t-d-1)$, ..., $u(t-1)$. Los términos de estos polinomios son los siguientes:

$$\begin{aligned} e_0 &= \sum_{i=0}^d a^i \\ e_1 &= -\sum_{i=1}^d a^i \\ f_0 &= 0 \\ f_1 &= b_0 \\ &\vdots \\ f_i &= a^{i-2}(ab_0 + b_1) \\ &\vdots \\ f_{d+1} &= a^{d-1}b_1 - b_0 \sum_{i=0}^{d-1} a^i \\ f_{d+2} &= -b_1 \sum_{i=0}^{d-1} a^i \end{aligned} \quad (4.11)$$

Igualmente se obtiene que

$$\hat{y}(t + d - 1) = E^{d-1}(z^{-1})y(t) + F^{d-1}u(t)$$

siendo E^{d-1} de grado 1 y F^{d-1} de grado $d-1$ y con análogas expresiones a las ecuaciones (4.11) simplemente sustituyendo d por $d-1$. Siempre se considera $d > 0$, ya que para $d = 0$ se tiene que $\hat{y}(t + d) = \hat{y}(t) = y(t)$, $\hat{y}(t + d - 1) = \hat{y}(t - 1) = y(t - 1)$.

Se dispone ya pues de la ecuación (4.10) que proporciona la relación buscada entre salida predichas y valores conocidos de entradas y salidas en el instante t ; con ella y la ley de control dada por la ecuación (4.7) ya se puede abordar el estudio de la función de transferencia en bucle cerrado.

4.2.2 Ecuación característica

Se pretende pues calcular la función de transferencia que relaciona la salida del sistema controlado con la referencia, sabiendo que la ley de control y las predicciones vienen dadas por

$$\begin{aligned} \Delta u(t) &= l_{y1}\hat{y}(t + d | t) + l_{y2}\hat{y}(t + d - 1 | t) + l_{r1}r(t) + l_{u1} \Delta u(t - 1) \\ \hat{y}(t + d | t) &= E^d(z^{-1})y(t) + F^d(z^{-1})u(t) \\ \hat{y}(t + d - 1 | t) &= E^{d-1}(z^{-1})y(t) + F^{d-1}(z^{-1})u(t) \end{aligned}$$

cuando el proceso real involucra incertidumbres estructuradas, viniendo representado por:

$$G(z^{-1}) = \frac{B(z^{-1})}{A(z^{-1})} = \frac{\gamma(1 - \alpha a)(1 - k\alpha a)(b_0 + b_1)(\beta m z^{-1} + (1 - \beta m)z^{-2})}{(1 - a)(1 - \alpha a z^{-1})(1 - k\alpha a z^{-1})} z^{-D} \quad (4.12)$$

El sistema resultante se puede ver representado en diagrama de bloques en la figura 4.6, y se va a proceder al cálculo de la función de transferencia resultante.

Para ello, el bloque de realimentación $l_{y1}\hat{y}(t + d | t) + l_{y2}\hat{y}(t + d - 1 | t)$ debe ser expresado en función de $y(t)$. Teniendo en cuenta que $G(z^{-1})$ relaciona $y(t)$ con $u(t)$ se tiene que:

$$\begin{aligned} l_{y1} \hat{y}(t + d | t) + l_{y2}\hat{y}(t + d - 1 | t) &= \\ &= l_{y1}(E^d(z^{-1})y(t) + F^d(z^{-1})u(t)) + l_{y2}(E^{d-1}y(t) + F^{d-1}(z^{-1})u(t)) = \end{aligned}$$

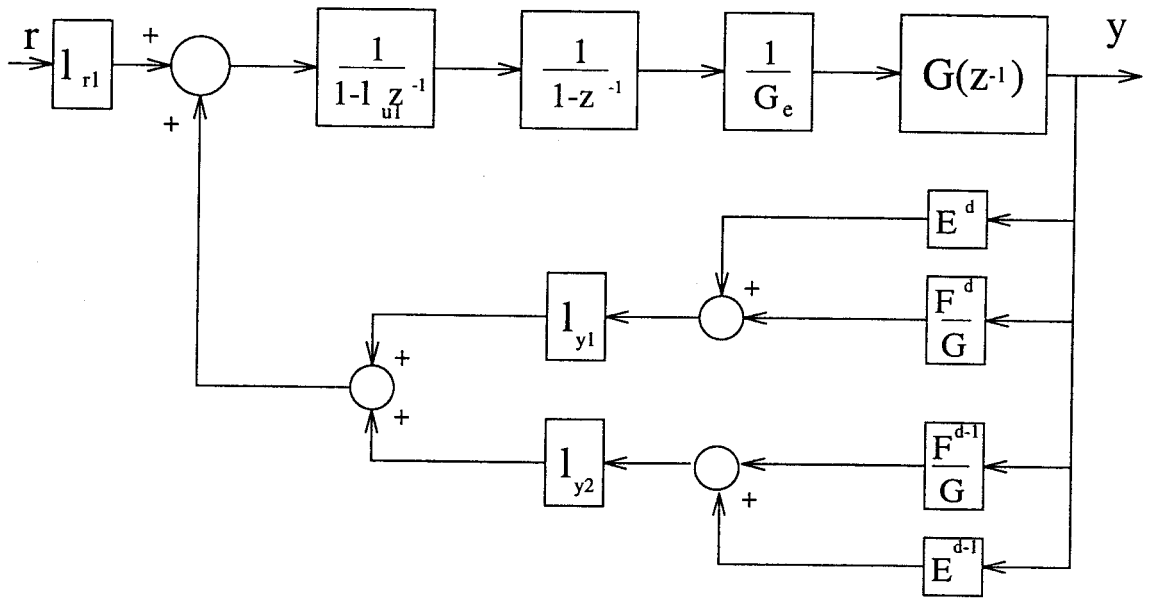


Figura 4.7: Ley de control con predictor

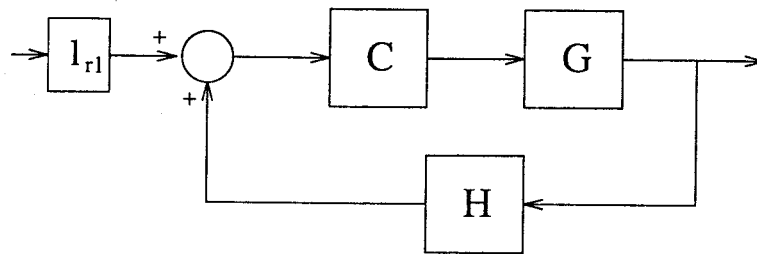


Figura 4.8: Diagrama de bloques simplificado

$$y(t) = C G l_{r1} r(t) + C G H y(t) \Rightarrow y(t)(1 - C G H) = C G l_{r1} r(t)$$

y por consiguiente la función de transferencia en bucle cerrado $G_{bc}(z^{-1})$ será:

$$G_{bc}(z^{-1}) = \frac{C(z^{-1})G(z^{-1})l_{r1}}{1 - C(z^{-1})G(z^{-1})H(z^{-1})}$$

Haciendo $G = \frac{B}{A}$ se obtiene que

$$G_{bc} = \frac{\frac{G l_{r1}}{(1-l_{u1}z^{-1})(1-z^{-1})G_e}}{1 - \frac{G\{l_{y1}(E^d + \frac{F^d}{G}) + l_{y2}(E^{d-1} + \frac{F^{d-1}}{G})\}}{(1-l_{u1}z^{-1})(1-z^{-1})G_e}}$$

$$G_{bc} = \frac{Gl_{r1}}{(1 - l_{u1}z^{-1})(1 - z^{-1})G_e - \{l_{y1}(E^dG + F^d) + l_{y2}(E^{d-1}G + F^{d-1})\}} = \frac{Bl_{r1}}{A(1 - l_{u1}z^{-1})(1 - z^{-1})G_e - \{l_{y1}(E^dB + F^dA) + l_{y2}(E^{d-1}B + F^{d-1}A)\}}$$

En consecuencia, la ecuación característica será:

$$A(z^{-1})(1 - l_{u1}z^{-1})(1 - z^{-1})G_e - B(z^{-1})(l_{y1}E^d + l_{y2}E^{d-1}) - A(z^{-1})(l_{y1}F^d + l_{y2}F^{d-1}) = 0 \quad (4.13)$$

la cual, sustituyendo A y B por sus valores, según ecuación (4.12), puede escribirse como:

$$\{(1 - a)(1 - \alpha az^{-1})(1 - k\alpha az^{-1})\} \{(1 - l_{u1}z^{-1})(1 - z^{-1})G_e - (l_{y1}F^d + l_{y2}F^{d-1})\} - \{\gamma(1 - \alpha a)(1 - k\alpha a)(b_0 + b_1)(\beta m z^{-1} + (1 - \beta m)z^{-2})z^{-D}\} (l_{y1}E^d + l_{y2}E^{d-1}) = 0$$

Obsérvese la diferencia entre esta ecuación característica y la que se obtendría para el caso de ausencia de incertidumbres. Aparte de los nuevos parámetros que aparecen, ahora el grado se ve influenciado por la dinámica inmodelada (polo inmodelado y error en el retardo). Esta circunstancia hace que varíe también la estructura de la ecuación característica, aumentando el polo inmodelado el grado de la ecuación en una unidad, mientras que el efecto del retardo incrementará o decrementará el grado según la relación entre d y D . Se observa pues el diferente efecto que incertidumbres paramétricas y dinámica inmodelada producen sobre la ecuación característica. En cualquier caso se comprueba que para $\alpha = \beta = \gamma = 1$, $k = 0$ y $D = d$ el resultado es el mismo que el obtenido anteriormente, como era de esperar.

4.2.3 Resolución de la ecuación característica

Se pretende resolver la ecuación característica para obtener los polos del sistema en bucle cerrado. La solución se va a obtener numéricamente dada la dificultad de obtener una solución analítica debido a la complejidad de la ecuación, no sólo por el elevado número de parámetros que aparecen, sino también porque su grado varía por el efecto de la dinámica inmodelada. Se obtendrán pues las soluciones para todos los valores de los parámetros del sistema (a , m y d) y de las incertidumbres (α , β , γ , k y D). Para el límite de estabilidad será necesario calcular sólo el módulo del mayor de los polos y ver si se encuentra dentro del círculo unidad.

Para calcular las raíces se escribe previamente la ecuación en la variable z . Para ello habrá que eliminar los términos en potencias negativas de z de la ecuación (4.13). Si se trabaja con los polinomios recíprocos \tilde{A} , \tilde{B} , \tilde{E} y \tilde{F} , siendo:

$$\begin{aligned}\tilde{A}(z) &= z^2 A(z^{-1}) \\ \tilde{B}(z) &= z^{D+2} B(z^{-1}) \\ \tilde{E}(z) &= z E(z^{-1}) \\ \tilde{F}(z) &= z^{d+2} F(z^{-1})\end{aligned}$$

se tiene que

$$\begin{aligned}\tilde{A}z^{-2}\{(z - l_{u1})(z - 1)z^{-2}G_e - (l_{y1}z^{-(d+2)}\tilde{F}^d + l_{y2}z^{-(d+1)}\tilde{F}^{d-1})\} - \\ - z^{-(D+2)}\tilde{B}(l_{y1}z^{-1}\tilde{E}^d + l_{y2}z^{-1}\tilde{E}^{d-1}) = 0\end{aligned}$$

A continuación se pueden distinguir dos casos según sea el valor del retardo del proceso D :

- $D > d$, es decir, $D = d + i$ con $i = 1, 2, \dots, N_d$. Se multiplicará la ecuación anterior por z^{D+3} , obteniendo

$$\tilde{A}\{(z - l_{u1})(z - 1)z^{d+i-1}G_e - z^{i-1}(l_{y1}\tilde{F}^d + l_{y2}\tilde{F}^{d-1})\} - \tilde{B}(l_{y1}\tilde{E}^d + l_{y2}\tilde{E}^{d-1}) = 0 \quad (4.14)$$

- $D < d$, es decir, $i = -1, -2, \dots, -N_d$. Multiplicando (4.14) por z^{d+4} :

$$\tilde{A}\{(z - l_{u1})(z - 1)z^d G_e - (l_{y1}\tilde{F}^d + l_{y2}\tilde{F}^{d-1})\} - \tilde{B}z^{i+1}(l_{y1}\tilde{E}^d + l_{y2}\tilde{E}^{d-1}) = 0 \quad (4.15)$$

Por consiguiente, estas expresiones serán las utilizadas para calcular los polos del bucle cerrado. Se puede observar que el grado de las ecuaciones varía con D y también con la relación entre D y d . En los siguientes apartados se estudiarán los polos para los distintos tipos de incertidumbre, primero por separado y luego globalmente.

4.3 Influencia de incertidumbres paramétricas

Se pretende encontrar los límites entre los que pueden hallarse los parámetros descritos anteriormente para preservar la estabilidad en bucle cerrado. El problema podría ser abordado haciendo uso de los teoremas de Kharitonov sobre estabilidad de sistemas, [5]. Pero no se va enfocar desde ese punto de vista por varias razones:

- Los teoremas de Kharitonov presentan condiciones necesarias y suficientes muy sencillas para estabilidad robusta, pero sólo para sistemas continuos.
- Las extensiones de estos teoremas para sistemas en tiempo discreto no presentan unos resultados tan sencillos y son más restrictivas [40].
- La obtención de condiciones se complica cuando aumenta el orden de la ecuación característica.
- Al considerar el bucle cerrado, los coeficientes de la ecuación característica pueden ser función de varios parámetros del proceso, obteniendo pues limitaciones a estos coeficientes cruzados, no a los del proceso.

De cualquier forma, la obtención de resultados para el proceso y el controlador que se están considerando se va a realizar numéricamente dentro de unos determinados límites. Bastará con resolver para un cierto conjunto de valores la ecuación (4.14). En [7] se presenta una primera aproximación a la robustez para unos determinados valores de variación de los parámetros. Se pretende ahora realizar un estudio exhaustivo sobre la influencia de las incertidumbres en los diferentes parámetros analizando su peso relativo. Primero se considerarán los caso por separado y luego conjuntamente.

4.3.1 Incertidumbre en el polo

A continuación se buscan los límites entre los que puede moverse el polo del proceso alrededor del considerado en el modelo preservando la estabilidad (una primera aproximación al problema se presenta en [13]). La cuestión es pues: ¿entre qué valores puede variar α ? Lógicamente estos límites dependerán de los otros parámetros del proceso. Se puede comprobar que m apenas influye y por tanto basta con calcular α_{min} y α_{max} para los posibles valores del polo y el retardo. Limitando los posibles valores de α al intervalo $[0.8, 1.2]$ (errores del 20 %) el alcance de α viene reflejado en la figura 4.9. En ella se representan las líneas que delimitan las regiones estables en función del polo del modelo para diferentes retardos, es decir, la zona "segura" es aquella comprendida entre la línea superior (α_{max}) y la inferior (α_{min}) para cada valor de d .

De esta figura se desprende que para una amplia zona de trabajo ($a < 0.75$) y para valores normales del retardo el sistema continúa siendo estable. Para polos mayores el límite superior va disminuyendo debido casi exclusivamente al hecho de que el bucle abierto ya sería inestable, es decir $\alpha \times a \geq 1$. Sólo para sistemas muy lentos y con grandes retardos la zona de movimiento de α se estrecha.

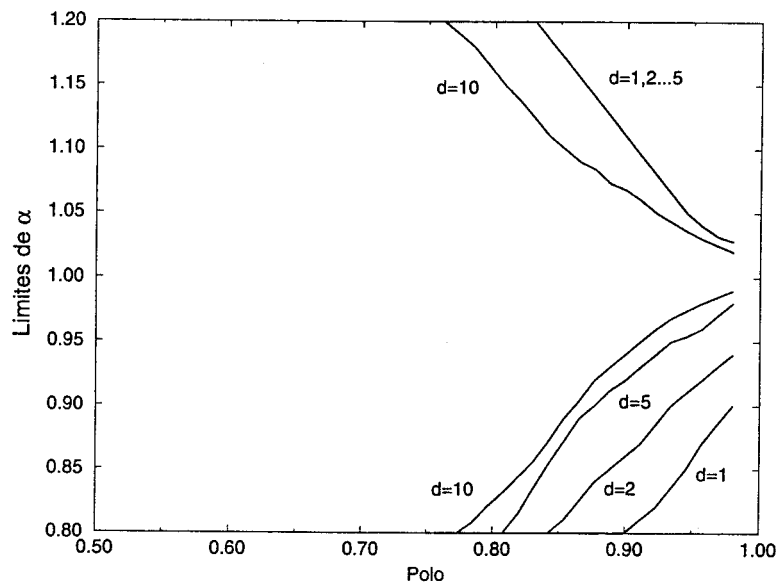


Figura 4.9: Límites superior e inferior de incertidumbre en el polo

4.3.2 Incertidumbre en la ganancia

Análogamente a la sección anterior, se van a obtener los límites de incertidumbre en la ganancia cuando la del modelo es G_e y la del proceso es $\gamma \times G_e$. En realidad la ganancia es un valor relativamente fácil de calcular, bien en bucle abierto usando la curva de reacción o también en bucle cerrado simplemente midiendo el cociente entre la salida y la señal de control en estado estacionario. Aun así, se va a permitir que γ se mueva entre 0.5 y 2, lo cual es un error considerable. Sin embargo, para la zona de trabajo considerada, el límite inferior γ_{min} está siempre por debajo de 0.5 y γ_{max} siempre sobre 1.5, es decir, se permiten incertidumbres en el valor de la ganancia del orden del 50%. Para retardos pequeños (1,2) el límite superior sobrepasa siempre el valor $\gamma = 2$ y sólo para retardos del orden de 10, se acerca al valor 1.5. En la figura 4.10 se muestran los límites superiores para retardos 5 y 10; los límites inferiores no se muestran pues son siempre menores que 0.5. Se puede concluir pues que el controlador es muy robusto frente a este tipo de errores.

4.3.3 Incertidumbre en m

Aunque este parámetro debe ser tenido en cuenta para la realización de un estudio completo, se comprueba fácilmente que su influencia es despreciable. Así, para todos los valores posibles del conjunto (a, m, d) , se ha comprobado que para el proceso modelado

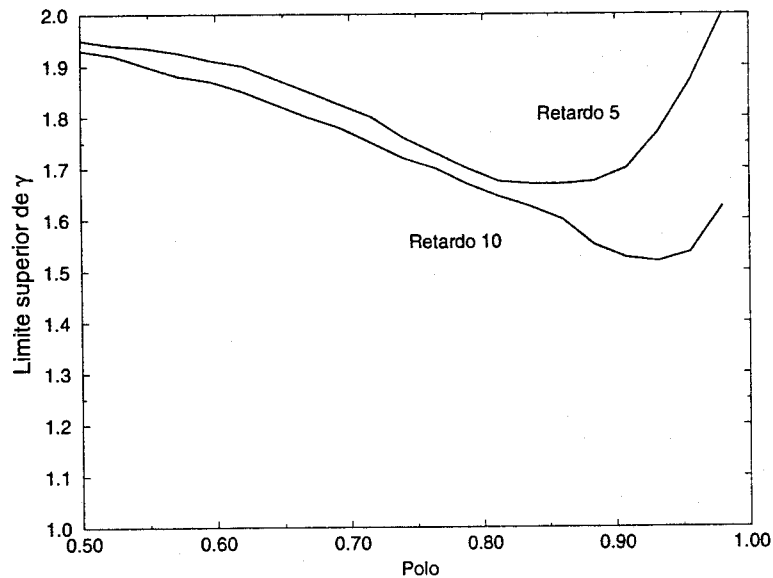


Figura 4.10: Límites superiores de incertidumbre en la ganancia para $d=5$ y $d=10$

con $\beta \times m$, el valor de β se puede mover libremente por el intervalo $[0, 3]$ sin alcanzar la inestabilidad en ningún caso. Nótese que estos límites son exagerados, pero de esta forma permiten despreciar el efecto de este parámetro.

4.3.4 Consideración de todas las incertidumbres paramétricas

Viendo las distintas incertidumbres por separado, se ha podido concluir que la del polo es la más importante, seguida de la de la ganancia, quedando en último lugar la de m que se puede despreciar. Pero, ¿qué ocurre si aparecen las tres al mismo tiempo? Lógicamente el efecto conjunto puede ser diferente y en principio producirá efectos más perniciosos. Pues bien, se han obtenido resultados para la zona de trabajo de los parámetros del proceso y de las posibles incertidumbres, algunos de los cuales se muestran a continuación. La influencia de m y β se vuelve a comprobar que resulta despreciable, por lo que se va a eliminar de los resultados. Por tanto el sistema vendrá definido por a y d y la zona estable por la superficie acotada por los valores máximos y mínimos de α y γ . Para poder representarlo gráficamente se dibujan distintas superficies para los valores discretos del retardo. Por ejemplo, en la figura 4.11 se muestra la zona segura para retardo $d = 10$: en el eje vertical se encuentran los valores límites de α para los valores usuales del polo y γ entre 0.5 y 1.5. Es decir, en la figura de la izquierda la zona "segura" es el volumen comprendido entre los ejes y la superficie mientras que en la figura de la derecha es el volumen limitado inferiormente por la superficie dibujada y superiormente por el plano horizontal de cota $\alpha = 1$.

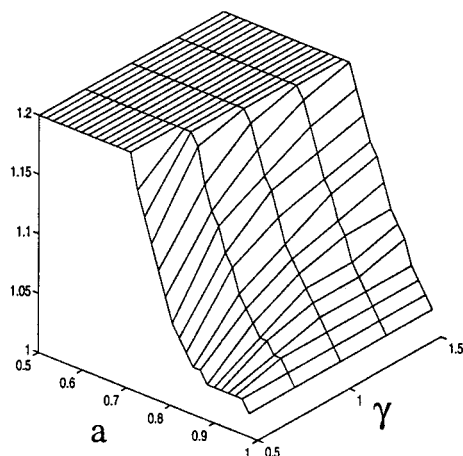
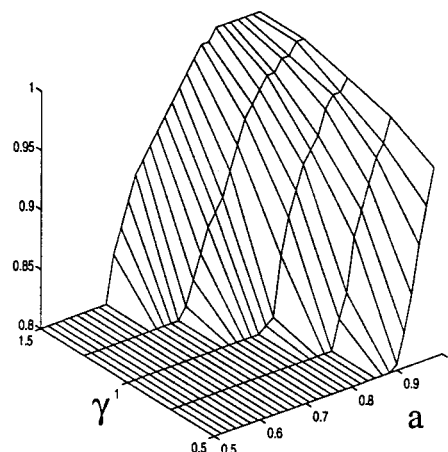
Limite superior de α Limite inferior de α 

Figura 4.11: Zona estable con incertidumbres en polo y ganancia para retardo 10

Se deduce que la zona segura disminuye ligeramente con las dos incertidumbres, pero no de forma alarmante, teniendo en cuenta que además se trata de un retardo considerable. Para una visión más cuantitativa, se presentan, para una serie de retardos, gráficas en dos dimensiones, que aunque menos vistosas quizás presenten más utilidad. Son cortes de las superficies para valores de γ iguales a 0.5, 1 y 1.5 (ver figura 4.12). donde puede verse que si bien el efecto conjunto es peor, aún así se puede trabajar con un cierto margen de confianza.

4.4 Influencia de dinámica inmodelada

Se va a estudiar la influencia de la dinámica inmodelada (polo y retardo) sobre la estabilidad, considerando que no aparecen incertidumbres paramétricas; en definitiva, se tendrá que $\alpha = \beta = \gamma = 1$.

4.4.1 Polo inmodelado

Se va estudiar en esta sección la influencia que sobre la estabilidad en bucle cerrado tiene el polo inmodelado, considerando que el retardo del modelo coincide con el del proceso,

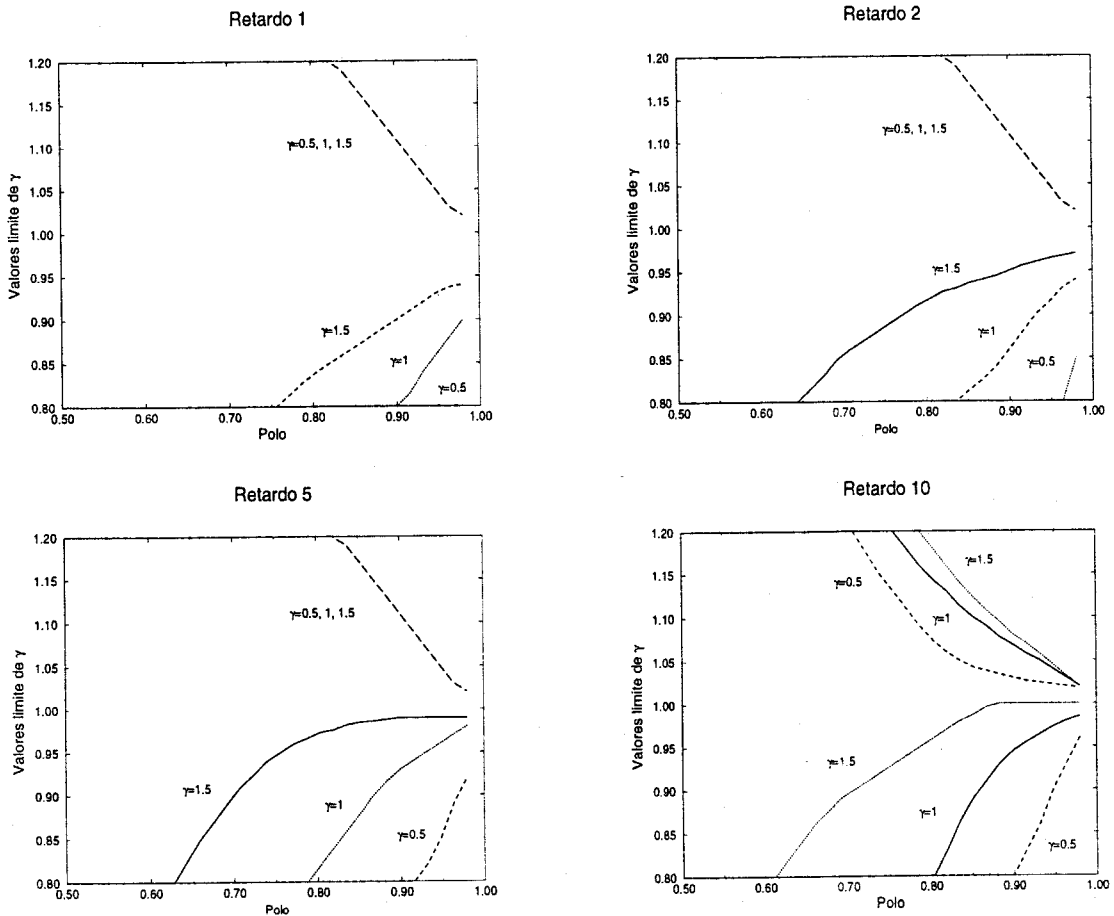


Figura 4.12: Límites de α para distintos valores de γ y del retardo

dejándose para más adelante la consideración de los dos factores conjuntamente. Por tanto el proceso vendrá descrito por la ecuación (4.4).

Se buscan pues los polos en bucle cerrado para diversos valores de a, b_0, b_1, d y k . Habrá pues que encontrar las soluciones de la ecuación característica, obtenida haciendo $d = D$ en (4.13).

En una primera aproximación a la ecuación se observa que en ella aparece el valor del retardo d , y como se verá a continuación, los polos van a depender de este valor, que por ahora se supone conocido. Nótese que para el modelo sin polo inmodelado, el bucle cerrado era independiente del valor del retardo, al incluir el GPC un predictor analítico. Aunque dicho valor sea conocido por el predictor, el hecho de la incertidumbre del polo dará lugar a que el retardo tenga influencia en la estabilidad del bucle cerrado.

La ecuación característica posee $d + 4$ raíces; para ver la estabilidad se va a tomar el mayor de los correspondientes $d + 4$ polos del sistema en bucle cerrado y se verá si se encuentra en el interior del círculo unidad. Realizando estos cálculos para todas las posibles variaciones de los parámetros del sistema y para los valores de k comprendidos entre 0 y 1, se pueden estudiar exhaustivamente todas las posibilidades existentes. Aunque esto es perfectamente factible, resulta interesante considerar cuáles de los parámetros del sistema influyen sensiblemente en la estabilidad, por si se puede descartar alguno.

El sistema se puede parametrizar usando:

- El valor del polo a , que variará desde valores del orden de 0.5 hasta valores próximos, pero menores que la unidad (sólo se consideran pues sistemas estables en bucle abierto)
- El parámetro $m = b_0/(b_0 + b_1)$, que indica la cercanía del retardo fraccionario al valor d o al $d + 1$. Como ya se indicó, para $m = 1$ el retardo es d y para $m = 0$ es $d + 1$.
- El retardo puro d , que se podrá variar desde cero hacia adelante.

Empezando por el parámetro m , se pueden obtener los mayores polos para diversos valores de retardo y de k haciendo variar los valores de m y de a , deduciendo según se desprende de la figura 4.13 que la variación del mencionado parámetro influye mínimamente en los resultados.

En la figura se plasman algunos casos para los valores extremos de m , quedando las curvas para otros valores de este parámetro encerradas entre las dos anteriores. Si acaso se observa que menores valores de m (la línea continua corresponde a $m = 0$ y la discontinua a $m = 1$) hacen aumentar la inestabilidad, debido a que se estaría trabajando con retardos mayores. Por tanto para seguir analizando el problema, y mientras no se diga lo contrario, se va a trabajar con un valor de m constante, que puede ser el caso más desfavorable $m = 0$.

En cambio los valores del polo y retardo del sistema sí van a influir considerablemente en los límites de estabilidad. Tomando como muestra algunos valores concretos de retardo ($d = 1, 2, 5, 10$) y haciendo variar la constante k entre 0 y 1, se pueden ver los resultados de estabilidad en la figura 4.14 para todos los valores del polo en bucle abierto. En ella se aprecia que el módulo del mayor polo en bucle cerrado permanece casi siempre dentro del círculo unidad incluso para valores muy grandes y poco usuales de k . Sólo para retardos grandes se pierde robustez.

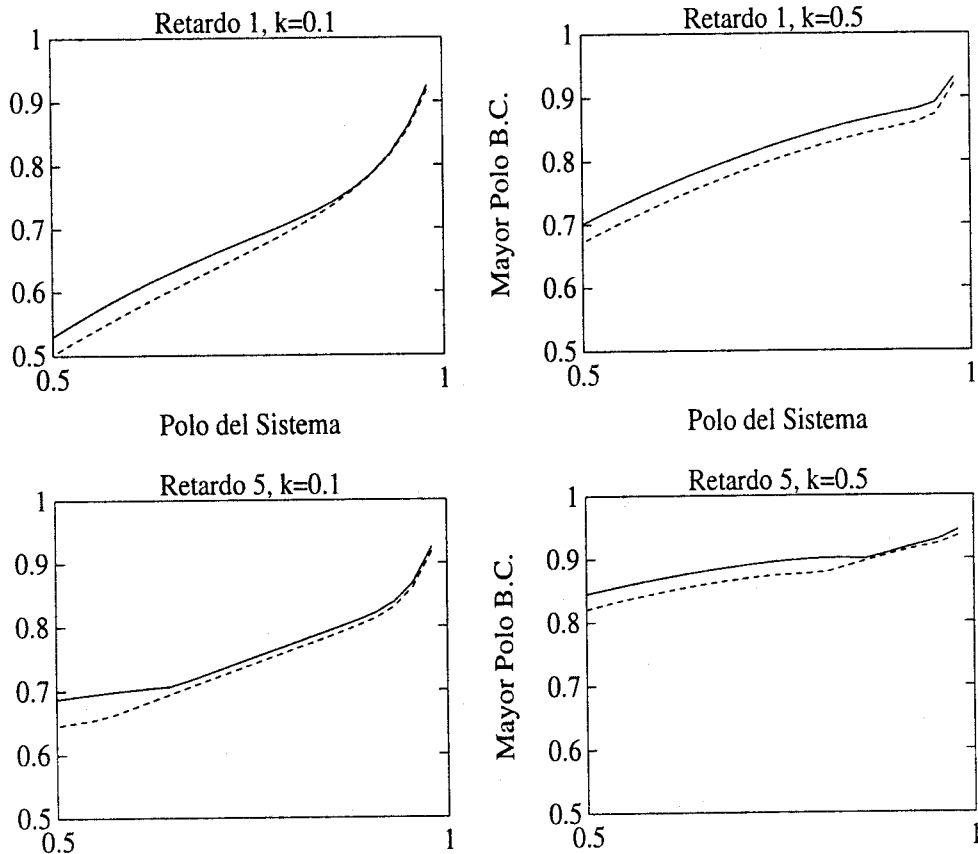


Figura 4.13: Influencia del parámetro m en la estabilidad para distintos valores de k y del retardo

Como ejemplo de comportamiento ante un caso límite, se muestra la respuesta de un sistema con el polo inmodelado de la misma magnitud que el conocido, es decir, $k = 1$. Se puede ver a partir de la figura 4.15 que en bucle abierto el proceso (línea continua) y el modelo (discontinua) difieren bastante, siendo pues un caso muy desfavorable. Aún así, se observa que el GPC lleva el proceso hasta la referencia sin error permanente, aunque lógicamente con una gran sobreoscilación.

En cualquier caso, se puede concluir que el GPC es muy robusto frente a la dinámica inmodelada de este tipo, pues para valores *razonables* del parámetro k e incluso para retardos grandes, se mantiene dentro de la zona de estabilidad.

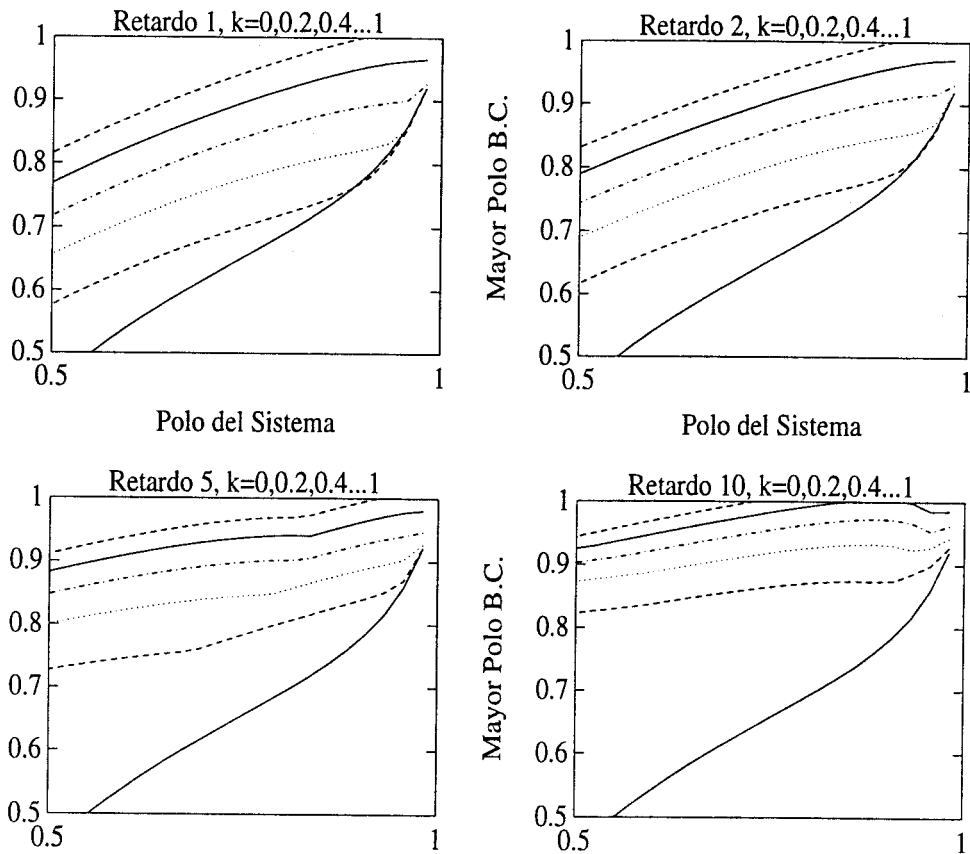


Figura 4.14: Polo inmodelado para distintos retardos

4.4.2 Error en estimación del retardo

El otro factor que se considera aquí es, como ya se ha indicado, la estimación errónea del retardo del proceso. El modelo tiene un retardo d , mientras que el proceso puede describirse por $D = d + i$, con $i = \pm 1, \pm 2, \dots, \pm N_d$, es decir, puede haber discrepancias en cantidades enteras; el caso de discrepancias fraccionarias queda englobado en la incertidumbre en el parámetro m . Análogamente a la sección anterior, se va a estudiar la influencia de este factor individualmente, dejando para más adelante su efecto con el conjunto de las incertidumbres.

Por tanto, será necesario resolver la ecuación característica cuando el proceso viene descrito por la ecuación (4.5). Se resuelven numéricamente las ecuaciones (4.14) y (4.15) y a continuación se presentan algunos de los resultados obtenidos.

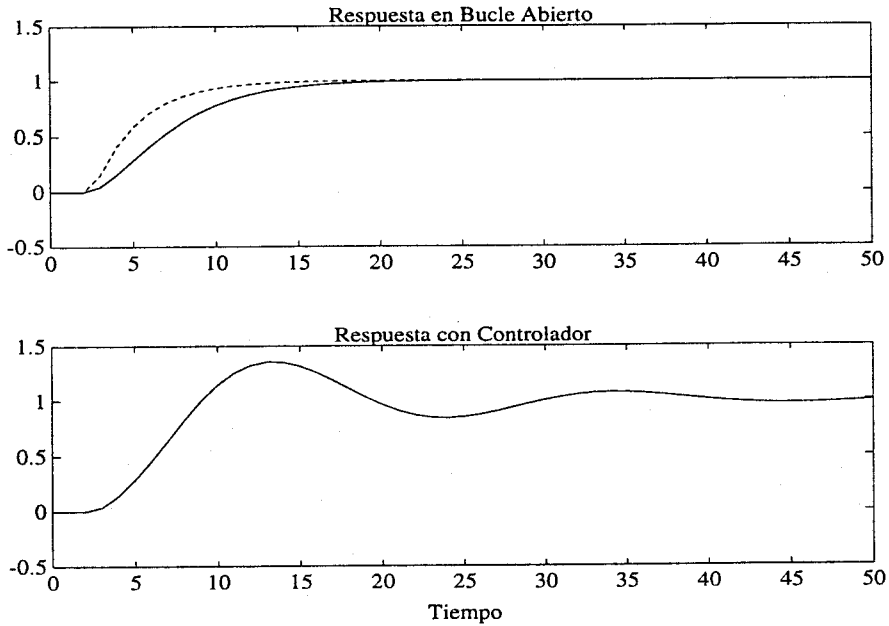


Figura 4.15: Comportamiento con $k=1$

Análogamente al caso anterior, el valor de m no influye excesivamente en la estabilidad, por lo que se representan los resultados obtenidos para algunos valores concretos de d y para todo el rango de variación del polo a .

Así, en la figura 4.16 se muestra el módulo del mayor polo para retardos del modelo de $d = 1, 2, 5, 10$ con discrepancias de $i = \pm 1, \pm 2$. Como se ve, para retardos pequeños, la estabilidad queda garantizada para errores de hasta dos unidades en todo el rango del polo, pero cuando se trabaja con retardos mayores esto sólo ocurre para escasos valores de a , e incluso para retardo 10 sólo se permite un error de una unidad.

Se puede concluir pues que una buena estimación del retardo es fundamental para el correcto funcionamiento del GPC, ya que para errores de más de una unidad, el sistema puede hacerse inestable si el retardo del proceso es elevado. Aunque en general se considera al GPC bastante robusto en cuanto a cambios en el retardo (ver por ejemplo [24]), esto es debido principalmente a que el modelo usado para el proceso emplea un polinomio B extendido capaz de absorber estos cambios. Resulta pues conveniente, en los casos en que se sospeche de estimaciones erróneas, utilizar algún algoritmo de cálculo de retardos, como por ejemplo el TIDEA [31], que además es particularmente sencillo para modelos de primer orden.

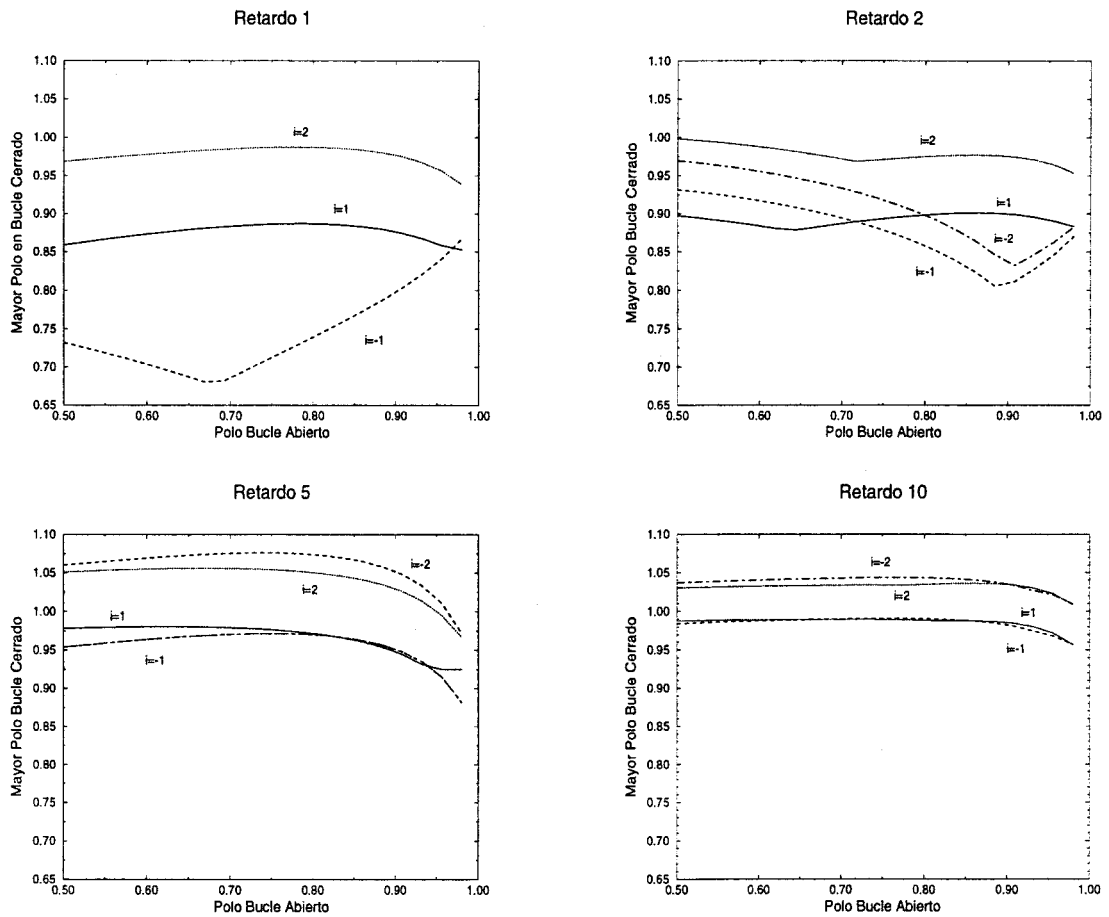


Figura 4.16: Influencia en la estabilidad de los errores en el retardo

4.4.3 Ambos factores

Aunque ya se han visto los dos factores que provocan dinámica inmodelada por separado, se pretende acto seguido estudiar su efecto conjunto. Lógicamente la robustez variará por la concurrencia de ambos factores, los cuales como ya se ha visto tienen distinto efecto sobre la estabilidad. La forma de calcular la influencia de ambos se reduce a resolver la ecuación característica para distintos valores de k y D .

En la figura 4.17 se muestran los valores de k_{max} para cuatro valores del retardo y toda la zona de variación del polo. Como se podría esperar se ve que para valores de i distintos de 0, es decir, con errores en la estimación del retardo, la k_{max} varía; en general este valor disminuye al aumentar el de i , de forma que para $i = 2$ es inestable para todos los valores del polo si el retardo es mayor o igual a 5 (curvas que no aparecen

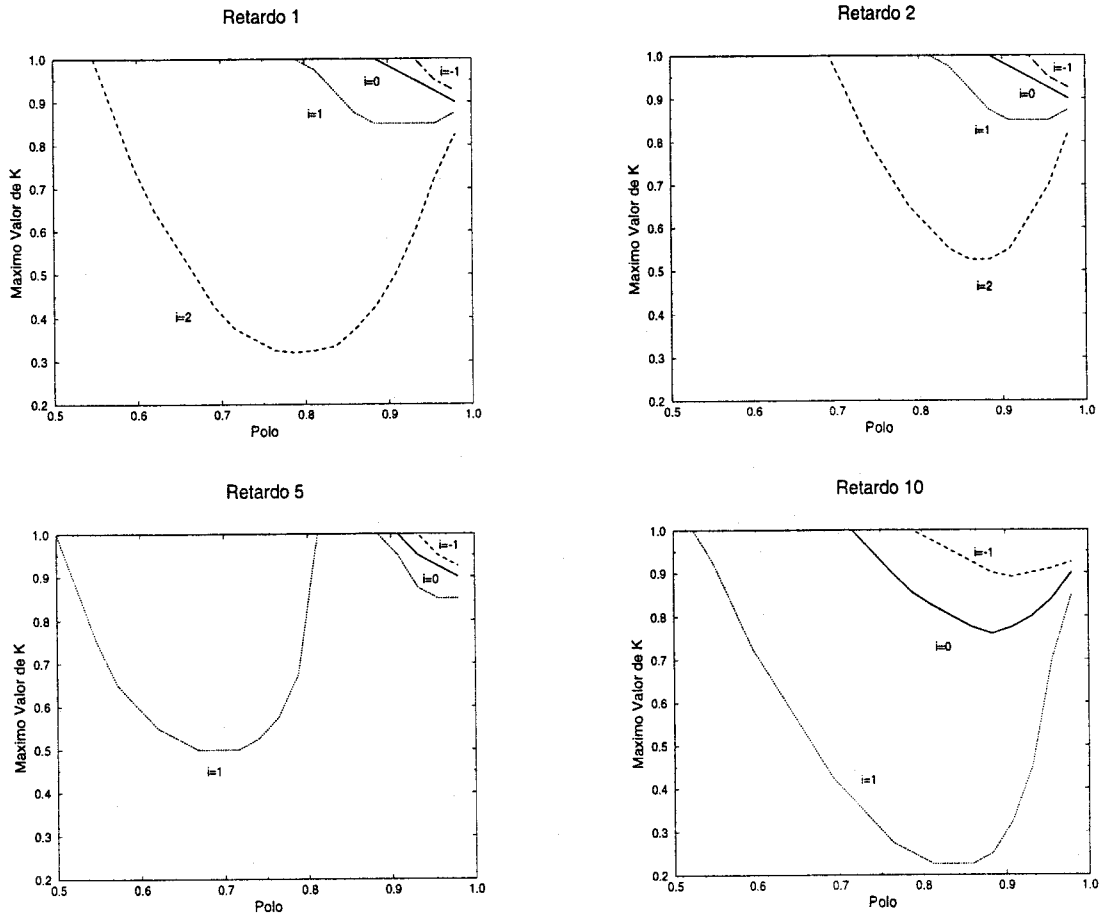


Figura 4.17: Efecto conjunto de polo inmodelado y error en el retardo

en la figura). Para $i = -1$ se permiten valores de k_{max} superiores a los del caso en que $i = 0$, aunque al hacerse más negativo se pierde estabilidad (de hecho, para $i = -2$ y $d > 5$ la k_{max} es 0).

Se puede deducir en general que aunque la concurrencia de ambos factores disminuye la robustez, siempre existe una banda de ± 1 unidad en la estimación del retardo en la que se permite la existencia de un polo inmodelado considerable (de hecho siempre con un 20% del valor de a como mínimo para toda la zona considerada). Se aprecia también que si el retardo del proceso es una unidad menor que el del modelo el comportamiento es más robusto.

4.5 Efecto conjunto

Aunque se han estudiado los dos tipos de incertidumbres estructuradas por separado, también cabe considerar el efecto conjunto de ambas sobre la robustez. Un estudio detallado de todos los fenómenos conllevaría la consideración de todos los posibles valores de los parámetros que definen el proceso (a , m y d) y las incertidumbres (α , β , γ , k y D) con lo cual sería necesario trabajar en un espacio de dimensión ocho. Esto no tiene mucho sentido, teniendo en cuenta que además en el caso de concurrencia de todas las incertidumbres el parecido entre el modelo y el proceso podría ser prácticamente nulo. Se ha preferido presentar unos resultados genéricos que muestren la zona estable para zona limitada de incertidumbres y comentar cualitativamente la influencia de cada factor. De cualquier manera, el caso de incertidumbre simultánea de todos los parámetros puede ser tratado como incertidumbre no estructurada desde el punto de vista que se tratará en el capítulo próximo.

La zona de incertidumbres está formada por los siguientes valores:

- α entre 0.9 y 1.1
- β entre 0.8 y 1.2
- γ entre 0.8 y 1.2
- k entre 0 y 0.2
- D entre $d + 1$ y $d - 1$

para esta zona "razonable" de incertidumbres se calcula el límite de estabilidad para el conjunto de parámetros que define el modelo. Los resultados para retardos 1 y 5 se presentan en la figura 4.18, donde se muestran los límites de la zona estable, quedando ésta encerrada dentro de las curvas dibujadas.

De la figura se desprende que la zona estable se ve claramente limitada por el valor máximo del polo; así para cualquier retardo se observa que no se permiten polos mayores de 0.9. La razón de este hecho es evidente: al tolerarse una incertidumbre en el polo de un 10%, el propio proceso en bucle abierto tendría el polo en la frontera del círculo unidad, encontrándose en el límite de estabilidad. Si no se juega con la incertidumbre en el polo la zona estable es considerablemente mayor. También se observa que el valor de m apenas influye mientras que el incremento del retardo va disminuyendo la zona segura; de hecho para retardos mayores ésta resulta sensiblemente inferior.

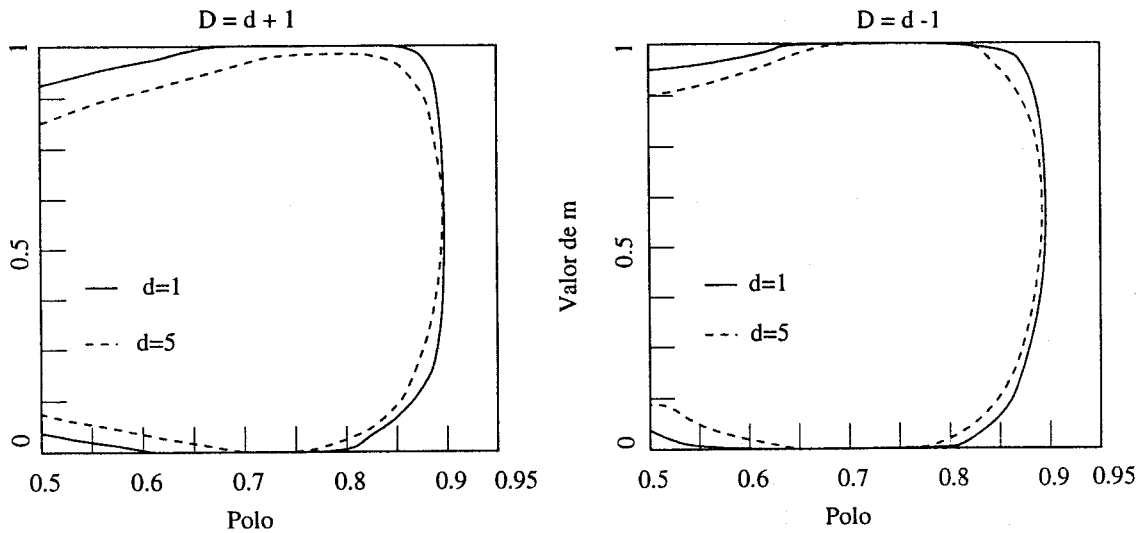


Figura 4.18: Efecto conjunto de incertidumbres paramétricas y dinámica inmodelada

En general se puede deducir que el factor más desfavorable es la incertidumbre en el retardo, ya que para diferencias de dos o más unidades se pierde robustez sensiblemente. En cambio, el controlador se comporta de forma muy robusta frente al resto de incertidumbres. De hecho las paramétricas tienen un amplio rango de variación siempre que la del polo no haga al sistema en bucle abierto inestable; el polo inmodelado apenas influye para un rango normal de valores.

En cualquier caso, el algoritmo GPC presentado en el capítulo anterior ha demostrado ser muy robusto respecto al tipo de incertidumbres que se han considerado. Para pequeños retardos el sistema en bucle cerrado es estable para grandes incertidumbres en la ganancia y constantes de tiempo con errores de más de 100 %.

4.6 Influencia del parámetro λ

Los resultados numéricos presentados hasta el momento han sido calculados para un determinado valor del parámetro λ que, como se sabe es el que pondera el esfuerzo de control en la función a minimizar. Valores grandes de λ darán lugar a un control suave mientras que valores muy pequeños proporcionarán acciones de control más bruscas. Como valor apropiado para unas prestaciones "normales" del GPC se puede escoger el valor de $\lambda = 0.8$. Este valor, para el campo de variación del proceso nominal que se está considerando proporciona una respuesta en bucle cerrado con un polo claramente dominante que da lugar a una constante de tiempo del orden del 60 – 70% de la de

bucle abierto. Las zonas estables presentadas a lo largo del capítulo se pueden obtener igualmente para distintos valores de λ , ya que lo único que varía con este parámetro son los valores de l_{y1} , l_{y2} , l_{r1} y l_{u1} , según la ley ya presentada.

Los límites resultarán afectados por la variación de la λ en la medida que al disminuir este parámetro se pierde robustez. Aunque este hecho se puede comprobar para todos los casos presentados en el capítulo, se presentan sólo algunos resultados para el último caso tratado (un conjunto de todas las incertidumbres estructuradas). En la figura 4.19 se muestran los límites para tres valores diferentes del peso en el esfuerzo de control, observándose la pérdida de robustez con la disminución de λ para un valor fijo del retardo ($d = 5$). Para valores pequeños ($\lambda = 0.3$) la disminución resulta apreciable. Según se deduce de las gráficas, se observa que se puede conseguir un diseño más robusto trabajando con valores grandes de este parámetro aunque a costa de perder prestaciones, ya que con λ demasiado grande la respuesta del sistema se puede hacer excesivamente lenta.

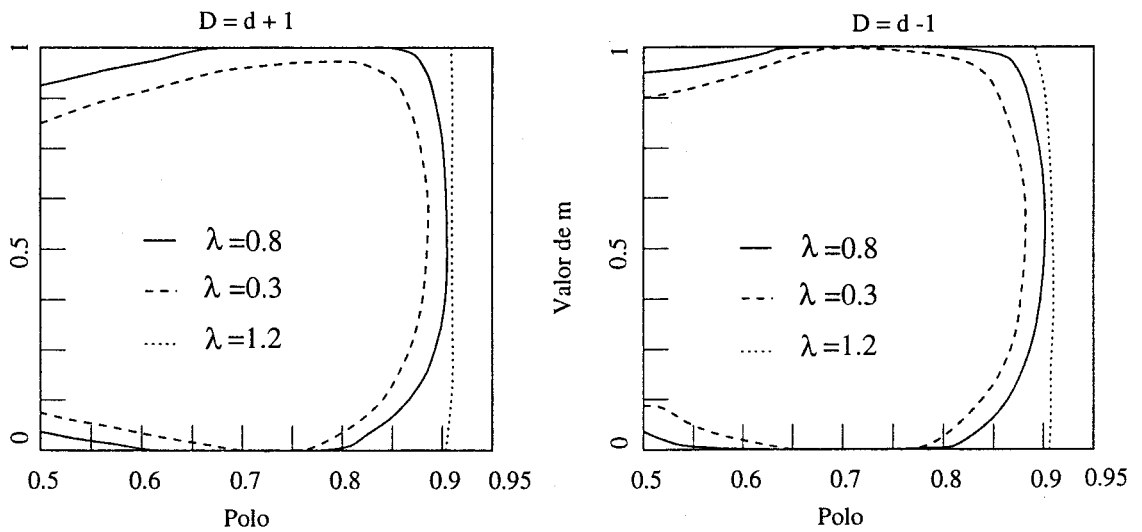


Figura 4.19: Influencia del parámetro λ en los límites de robustez

Capítulo 5

Robustez frente a incertidumbres no estructuradas

5.1 Incertidumbres no estructuradas

5.1.1 Descripción de la incertidumbres

En el capítulo anterior se hizo la distinción entre dos tipos de incertidumbres: las estructuradas y las que no lo son. Abordado pues el estudio de las primeras se procede ahora a considerar la influencia que las incertidumbres no estructuradas tienen en el comportamiento del proceso en bucle cerrado.

Muchas incertidumbres en el proceso pueden ser parametrizadas de alguna forma, tal como se vió en el capítulo anterior. Así un proceso no lineal puede ser representado por un modelo lineal en torno a un punto de equilibrio, debiendo cambiar el valor de los parámetros al cambiar éste. En otros casos en que el modelo lineal pueda representar con suficiente exactitud el proceso, las propias condiciones de trabajo pueden provocar cambios en los parámetros. Por ejemplo, el funcionamiento a mayores caudales reducirá los tiempos muertos y las constantes de tiempo. Estas situaciones se pueden representar mediante límites paramétricos y hacer el tratamiento mostrado en el capítulo anterior. Estas incertidumbres se dan en general a bajas frecuencias; sin embargo siempre existirán errores a frecuencias más altas que difícilmente pueden ser tratados de la forma anterior. Estos errores pueden ser debidos a diversos efectos tales como resonancias electromagnéticas, retardos, dinámicas de válvulas y actuadores o

procesos de difusión. Debido a ello a altas frecuencias incluso el orden del modelo es desconocido.

También hay que hacer notar que las incertidumbres pueden ser usadas para representar efectos que de hecho no son del todo inciertos. Tal puede ser el caso de un elemento no lineal, que aunque puede ser bien modelado, no puede aparecer como tal en un método de diseño que sólo contemple modelos lineales. En otros casos, se pueden ignorar deliberadamente efectos dinámicos conocidos y modelables para así conseguir un modelo nominal más sencillo con la idea de aplicar un procedimiento de diseño determinado (tal puede ser el caso de aplicar el GPC con un modelo de primer orden a sabiendas de que el proceso es de orden superior).

Haciendo uso de las ideas de Doyle y Stein [30], el proceso real se encontrará en un entorno del proceso nominal que será representado por un modelo conocido. Esta proximidad entre proceso real (G) y nominal (\tilde{G}) puede ser representada en el dominio de la frecuencia en forma aditiva:

$$G(i\omega) = \tilde{G}(i\omega) + \Delta G(i\omega)$$

con

$$|\Delta G(i\omega)| < l_a(\omega) \quad \forall \omega$$

Nótese que esta representación no impone ninguna forma o estructura particular a la incertidumbre $\Delta G(i\omega)$, sino simplemente un límite. También la incertidumbre se puede expresar en forma multiplicativa

$$G(i\omega) = [I + L(i\omega)]\tilde{G}(i\omega)$$

con

$$|L(i\omega)| < l_m(\omega) \quad \forall \omega$$

Con esta representación G está en un entorno normalizado de \tilde{G} , de forma que el sistema compensado con el controlador C va a tener los mismos límites de incertidumbre que el proceso.

A falta de un nombre mejor, las incertidumbres representadas mediante las ecuaciones anteriores reciben el nombre de *incertidumbres no estructuradas*. En la figura 5.1 se pueden ver ambas representaciones de las incertidumbres en diagrama de bloques.

Aunque los efectos de incertidumbres a frecuencias bajas pueden ser en muchos casos representados mediante incertidumbres estructuradas, a frecuencias altas la discrepancia entre el proceso real y el nominal puede resultar considerable por la aparición de efectos realmente difíciles de modelar, poniendo de manifiesto un comportamiento como el descrito en la figura 5.2a. De ella se desprende que la respuesta es muy similar

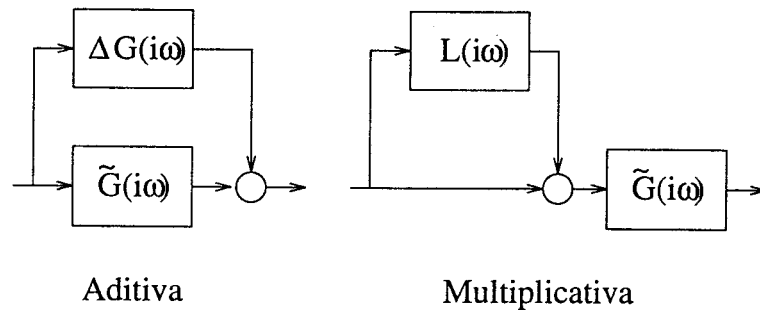


Figura 5.1: Representación de las incertidumbres en diagrama de bloques

para frecuencias bajas pero la discrepancia puede aumentar considerablemente para frecuencias altas, con lo cual la franja en torno al proceso nominal es estrecha al principio y luego debe expandirse sensiblemente. Esta franja viene definida por el valor del límite l_m , que irá acorde con este comportamiento siendo mucho menor que la unidad para ω pequeñas y superando la unidad conforme ésta se hace mayor (figura 5.2b).

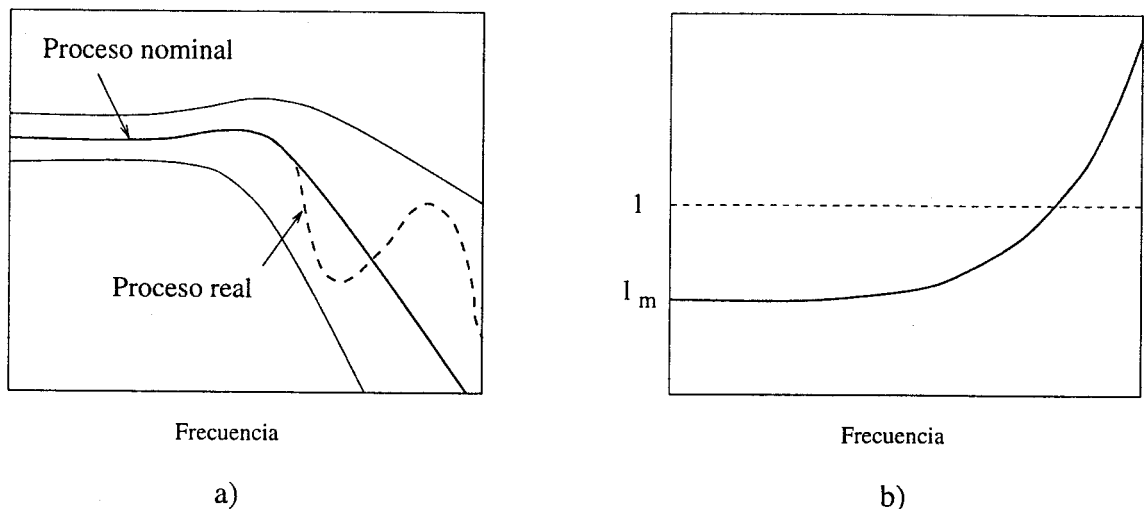


Figura 5.2: Discrepancia entre el proceso nominal y el real

5.1.2 Descripción del proceso

A partir de ahora se asumirá que el comportamiento dinámico de un determinado proceso es descrito no por un modelo lineal invariante en el tiempo, sino por una familia de modelos lineales invariantes. Entonces los posibles procesos reales (G) se

encontrarán en una zona próxima al proceso nominal (\tilde{G}), el cual será modelado por un sistema de primer orden con retardo.

Se definirá por tanto una familia \mathcal{F} de procesos en el dominio de la frecuencia, que en el plano de Nyquist vendrá representada por una región en el entorno de la planta nominal para cada frecuencia ω . Si se define esta familia como

$$\mathcal{F} = \{G : |G(i\omega) - \tilde{G}(i\omega)| \leq \bar{l}_a(\omega)\}$$

la región consiste en un disco con centro en $G(i\omega)$ y radio $\bar{l}_a(\omega)$. Por tanto cualquier miembro de la familia cumple la condición

$$G(i\omega) = \tilde{G}(i\omega) + l_a(i\omega) \quad |l_a(i\omega)| \leq \bar{l}_a(\omega)$$

Esta región cambiará con ω puesto que l_a lo hace y por tanto para describir la familia \mathcal{F} tendremos una zona formada por los discos a distintas frecuencias, como se muestra en la figura 5.3.

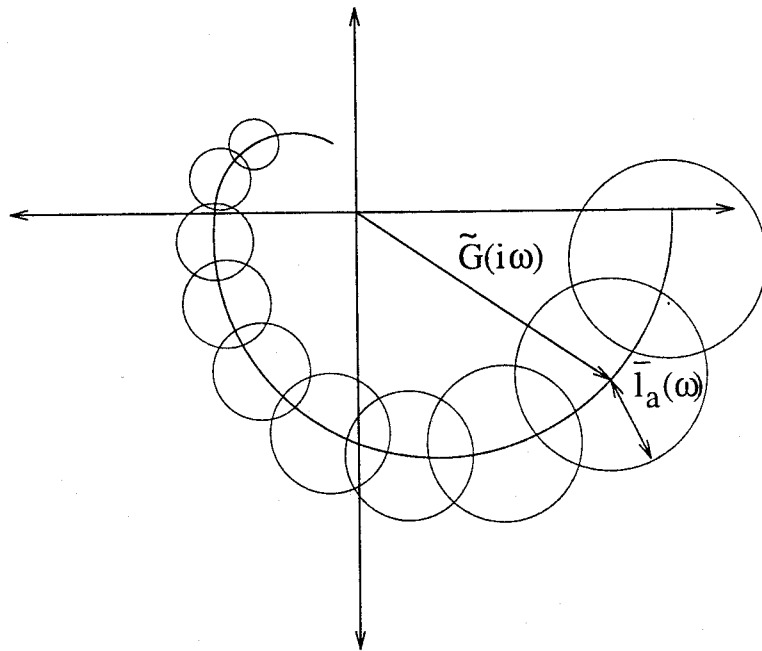


Figura 5.3: Familia \mathcal{F} para distintas frecuencias

Si se desea trabajar con incertidumbres multiplicativas, la familia de procesos puede ser descrita mediante:

$$\mathcal{F} = \left\{G : \left| \frac{G(i\omega) - \tilde{G}(i\omega)}{\tilde{G}(i\omega)} \right| \leq \bar{l}_m(\omega) \right\} \quad (5.1)$$

simplemente considerando

$$l_m(i\omega) = \frac{l_a(i\omega)}{\tilde{G}(i\omega)} \quad \bar{l}_m(i\omega) = \frac{\bar{l}_a(\omega)}{|\tilde{G}(i\omega)|}$$

Por consiguiente cualquier miembro de la familia \mathcal{F} cumple que:

$$G(i\omega) = \tilde{G}(i\omega)(1 + l_m(i\omega)) \quad |l_m(i\omega)| \leq \bar{l}_m(i\omega)$$

Esta representación de las incertidumbres en el plano de Nyquist como un disco en torno al proceso nominal puede englobar cualquier conjunto de incertidumbres estructuradas. Por ejemplo, para un sistema de primer orden con retardo puro de la forma $G(s) = ke^{-\tau_d s}/(1 + \tau s)$ las incertidumbres en los parámetros k , τ y τ_d dan lugar a una zona tal como la descrita en la figura 5.4a (para un determinado valor de ω). Siempre se podrá encontrar un disco que circunscriba dicha zona, con lo cual existirá un límite $\bar{l}_a(\omega)$ que permita tratar estas incertidumbres como no estructuradas y usar los criterios de robustez que se expondrán más adelante .

De cualquier manera, se observa que es una postura un tanto conservadora, como se puede deducir de la figura 5.4b, pero es la mejor si no se conoce a priori la estructura de las incertidumbres. Si éstas tienen una estructura y pueden ser descritas por una serie de parámetros, entonces conviene abordar el problema desde otro punto de vista, haciendo uso de los resultados del capítulo precedente.

En cualquier caso hay que hacer notar que esta representación de las incertidumbres no es válida para describir familias de procesos cuyos polos puedan cruzar el eje imaginario (o el círculo unidad en sistemas discretos), ya que si un proceso G tiene un polo en el eje para una frecuencia ω_o entonces $|G(i\omega_o)| = \infty$ y por tanto $\bar{l}_a(\omega_o) = \bar{l}_m(\omega_o) = \infty$. Por consiguiente los elementos de la familia no pueden tener distinto número de polos inestables.

5.2 Criterios de robustez

5.2.1 Herramientas de diseño robusto

Tradicionalmente y para sistemas monovariantes se han usado métodos gráficos para el estudio en frecuencia de la robustez del sistema, empleando los conceptos de margen de fase y margen de ganancia. Los métodos de diseño robusto surgidos en los últimos años

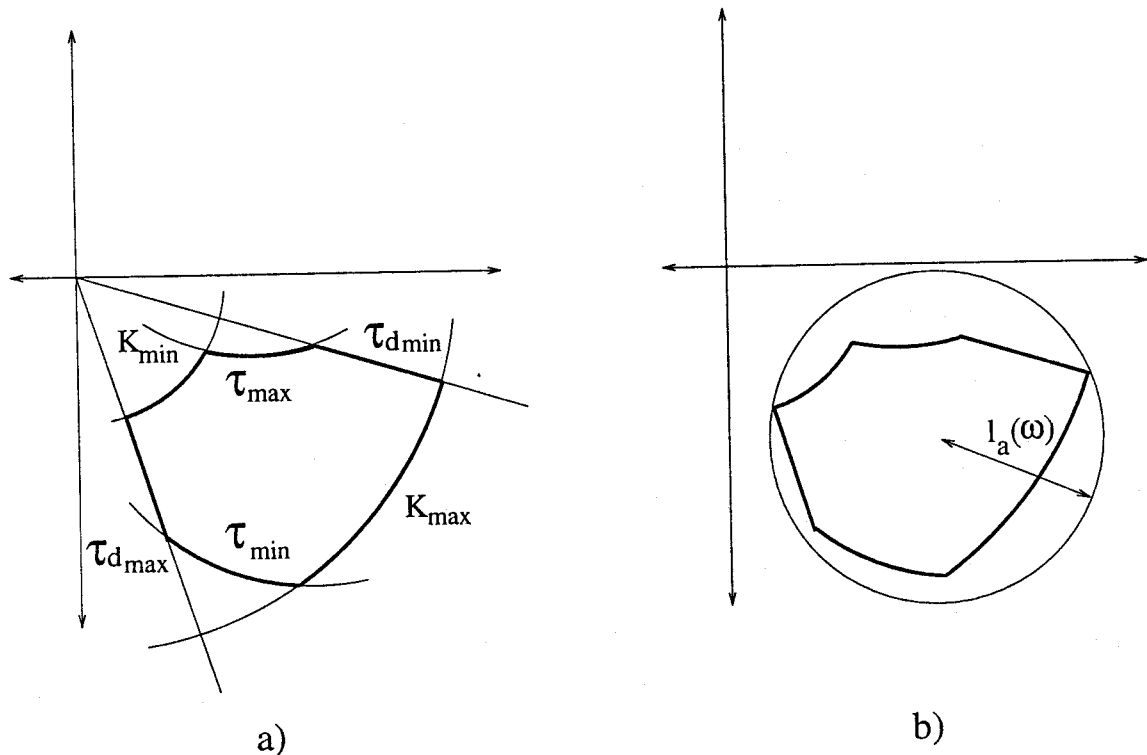


Figura 5.4: Tratamiento de la incertidumbres paramétricas como no estructuradas

se basan la extrapolación de estos conceptos a sistemas multivariables. Pero ello no es ni mucho menos algo inmediato, ya que es necesario generalizar la idea de ganancia y desfase, obteniendo alguna magnitud equivalente que permita un tratamiento análogo.

Doyle y Stein [30] proponen la extensión de las ideas clásicas de monovariable a sistemas con varias entradas y salidas, permitiendo no sólo la obtención de márgenes de robustez sino el diseño de controladores robustos. Para ello hacen uso del valor singular de una matriz de transferencia¹, valor que proporciona una idea del "tamaño" de la matriz [57]. Los valores singulares máximo y mínimo de una matriz de transferencia dependerán de ω y pueden representarse gráficamente como diagrama magnitud-frecuencia análogamente al diagrama de Bode de un sistema monovariable.

Estas ideas dieron lugar al uso del método de optimización \mathcal{H}_∞ de Zames [94]² para el diseño de controladores [49]. La norma ∞ de un sistema con matriz de transferencia F se obtiene calculando el mayor valor singular para cada frecuencia de $F(i\omega)$ y

¹Se definen los valores singulares de una matriz A como $\sqrt{\lambda_i(A^*A)}$ siendo λ_i el i -ésimo autovalor y A^* la matriz conjugada traspuesta.

²La norma \mathcal{H}_∞ para un sistema multivariable descrito por una matriz de transferencia F se define como $\|F\|_\infty = \sup_\omega \|F(i\omega)\|_2$ con $\|F\|_2 = \max_i \sigma_i(F)$ siendo σ_i el valor singular i -ésimo de la matriz F .

tomando el valor máximo de ellos. El uso de esta norma permite el diseño robusto, de forma que el controlador \mathcal{H}_∞ es aquél que minimiza la norma ∞ de la función de sensibilidad ponderada por una cierta función de la frecuencia.

Un aspecto fundamental es el de traducir el conocimiento práctico sobre las incertidumbres del proceso a términos cuantitativos que permitan la aplicación de las técnicas de \mathcal{H}_∞ . El método del valor singular estructurado (μ) desarrollado Doyle [29] constituye una poderosa herramienta para el estudio de la estabilidad y prestaciones en presencia de incertidumbres y existen varias realizaciones en este contexto como por ejemplo [3]. El problema del análisis del comportamiento mediante el valor singular estructurado es una técnica relativamente asequible [69], mientras que la síntesis del controlador requiere cálculos considerables (algunas técnicas de resolución se muestran en [55]). Aún así existen implantaciones de controladores sintetizados con la técnica μ funcionando con éxito en diversas aplicaciones en la industria de procesos y en el sector aeroespacial [90].

Estos métodos todavía no han llegado al límite y constituyen un campo abierto en el que aparecen continuamente nuevos resultados, también para sistemas muestreados, como por ejemplo [45], [83], o para sistemas no lineales o de parámetros distribuidos. Las herramientas de diseño de controladores robustos encuentran su campo de aplicación preferentemente en sistemas multivariables conllevando una considerable carga de cálculo por lo que su utilización como herramienta de diseño, sobre todo en el caso adaptativo dista mucho de ser sencilla. Lógicamente estos métodos pueden ser usados para sistemas monovariantes restringiendo la dimensión a 1, pero en muchos casos se pueden obtener idénticos resultados haciendo uso de las mismas ideas pero con formulaciones más simples como se verá a continuación.

5.2.2 Medida de la robustez del GPC

Al igual que en el capítulo anterior, el objetivo no es hacer un diseño robusto. Se busca cuantificar en lo posible la robustez del Control Predictivo Generalizado para sistemas monovariantes de primer orden con retardo. Para ello se hará uso de los conceptos en que se basan los métodos de control robusto nombrados anteriormente pero aplicados al caso escalar.

Sea el proceso G perteneciente a la familia \mathcal{F} , definida por la expresión (5.1) y en la que todos sus miembros tienen un número N de polos inestables, con un controlador C en serie y realimentación unitaria tal como muestra la figura 5.5 y tal que dicho controlador con el proceso nominal \tilde{G} sea estable. Entonces, haciendo uso del criterio

de estabilidad de Nyquist³ todo proceso G será estable si el número neto de vueltas alrededor del punto $(-1,0)$ en el plano de Nyquist para CG coincide con el valor de N . Esto es lo mismo que decir que la zona donde se encuentran todos los procesos G de la familia no incluye al punto -1 .

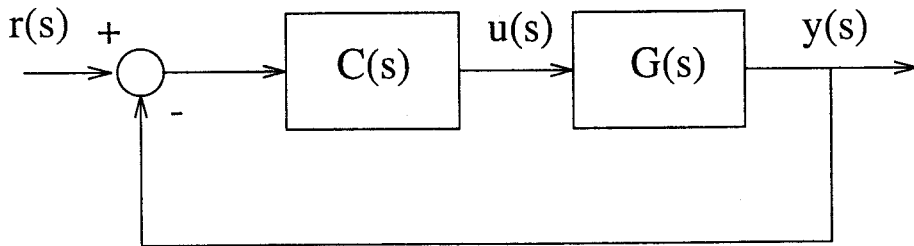


Figura 5.5: Proceso y controlador con realimentación unitaria

Haciendo uso de la función de sensibilidad complementaria T definida como

$$T(i\omega) = \frac{G(i\omega)C(i\omega)}{1 + G(i\omega)C(i\omega)}$$

se tiene el siguiente teorema:

Teorema de Estabilidad Robusta (ver [60]). *Supóngase que todos los procesos G de una familia \mathcal{F}*

$$\mathcal{F} = \left\{ G : \left| \frac{G(i\omega) - \tilde{G}(i\omega)}{\tilde{G}(i\omega)} \right| \leq \bar{l}_m(\omega) \right\}$$

tienen el mismo número de polos inestables y que un controlador C estabiliza el proceso nominal \tilde{G} . Entonces el sistema posee estabilidad robusta con el controlador C si y sólo si la función de sensibilidad complementaria para el proceso nominal satisface la siguiente relación:

$$|\tilde{T}(i\omega)| \bar{l}_m(\omega) < 1 \quad \forall \omega \quad (5.2)$$

La demostración es sencilla y se deduce de la figura 5.6. La condición de que la zona incierta no incluya al punto -1 equivale a que la distancia $|1 + \tilde{G}C|$ sea mayor al radio del disco, $|\tilde{G}C| \bar{l}_m$, es decir:

$$|1 + \tilde{G}C(i\omega)| > |\tilde{G}C(i\omega)| \bar{l}_m(\omega) \quad \forall \omega$$

o, lo que es lo mismo

$$\left| \frac{\tilde{G}C(i\omega)}{1 + \tilde{G}C(i\omega)} \right| \bar{l}_m(\omega) = |\tilde{T}(i\omega)| \bar{l}_m(\omega) < 1 \quad \forall \omega$$

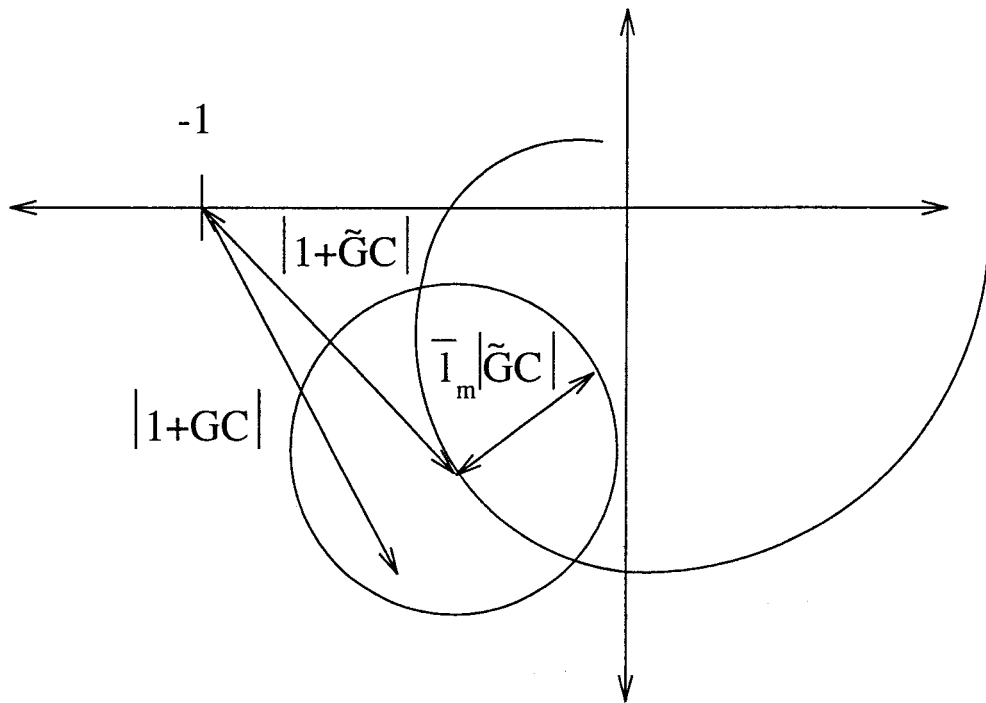


Figura 5.6: Estabilidad robusta

Nótese que la condición (5.2) puede escribirse como

$$\|\tilde{T}\bar{l}_m\|_\infty \equiv \sup_{\omega} |\tilde{T}\bar{l}_m(\omega)| < 1$$

lo cual indica que la condición de estabilidad robusta impone un límite a la norma ∞ de la función de sensibilidad complementaria ponderada por \bar{l}_m .

Este criterio puede ser fácilmente generalizado para sistemas discretos. Para ello se tiene el proceso con un mantenedor de orden cero y un prefiltro según se indica en la figura 5.7, donde también se muestra el equivalente discreto, englobando $G(z)$ el mantenedor y el posible prefiltro. Entonces el teorema anterior quedaría como:

Teorema de Estabilidad Robusta (ver [60]). *Supóngase que todos los procesos $G(s)$ de una familia \mathcal{F}*

$$\mathcal{F} = \left\{ G(s) : \left| \frac{G(e^{i\omega T}) - \tilde{G}(e^{i\omega T})}{\tilde{G}(e^{i\omega T})} \right| \leq \bar{l}_m(\omega) \right\}$$

tienen el mismo número de polos inestables, los cuales no se vuelven inobservables por el muestreo, y que un controlador $C(z)$ estabiliza el proceso nominal $\tilde{G}(s)$. Entonces

³Este criterio puede ser encontrado en cualquier libro clásico de control, ver por ejemplo [65].

el sistema posee estabilidad robusta con el controlador C si y sólo si la función de sensibilidad complementaria para el proceso nominal satisface la siguiente relación:

$$|\tilde{T}(e^{i\omega T})| \bar{l}_m(\omega) < 1 \quad 0 \leq \omega \leq \pi/T \quad (5.3)$$

Como se ve la condición es análoga a la de continuo. Nótese que si se cumple para el intervalo de ω dado también lo hará para todo ω .

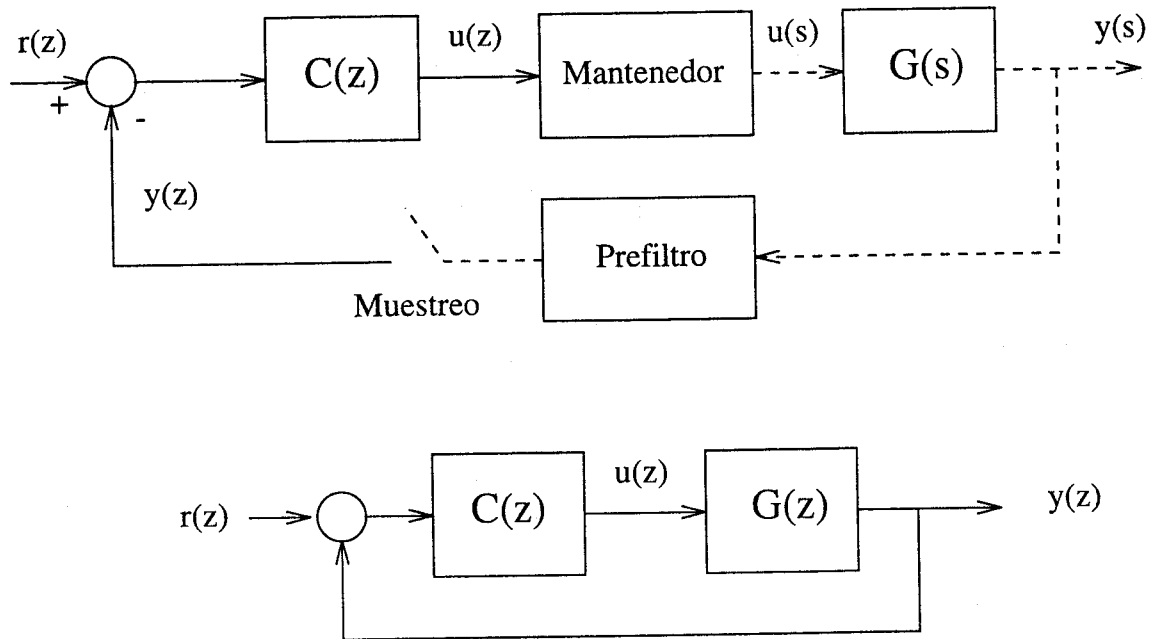


Figura 5.7: Equivalente discreto

Este criterio de robustez se va a emplear en la próxima sección para obtener los límites entre los que se podrá encontrar la familia de procesos alrededor del proceso nominal manteniendo la estabilidad del bucle cerrado.

5.3 Límites de robustez

En esta sección se va a proceder a la obtención de los límites de robustez del GPC precalculado en presencia de incertidumbres no estructuradas, usando los fundamentos expuestos previamente. Ahora se puede abordar el problema desde otro ángulo. Si se dispone del controlador desarrollado para el proceso nominal ¿cuál será la familia alrededor de dicho proceso para la cual el controlador es robusto? Es decir, se buscará

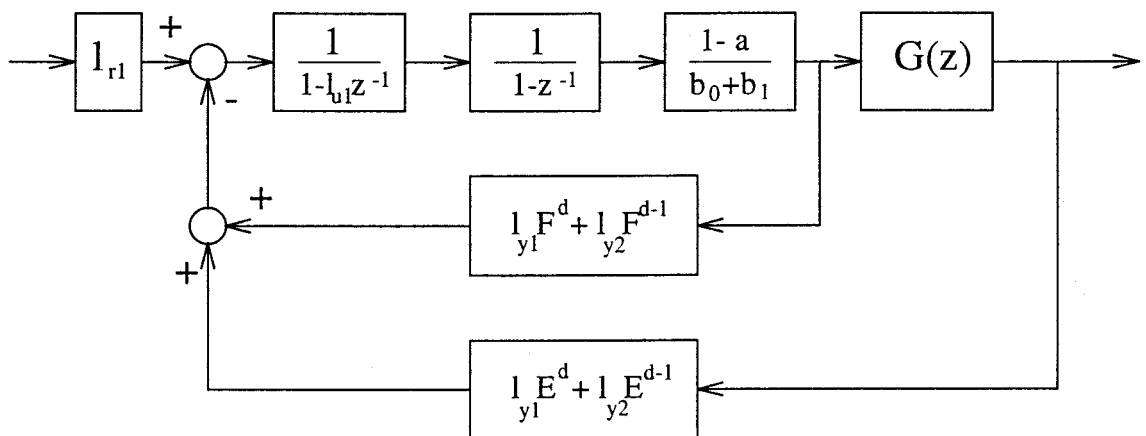
el límite de robustez $\bar{l}_a(\omega)$ que proporciona una zona alrededor de la planta nominal en la cual el controlador mantiene la estabilidad del bucle cerrado.

5.3.1 Estructura de control

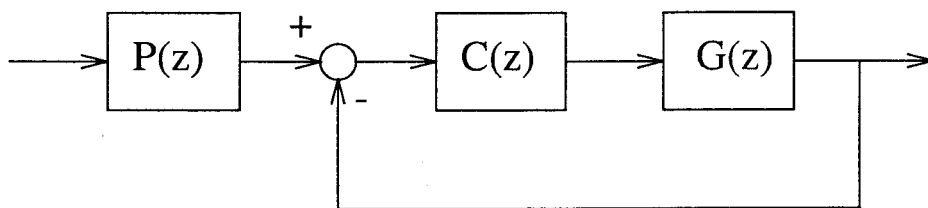
Para obtener este límite hay que poner el GPC en la forma descrita en la figura 5.7. Usando el predictor cuya expresión se formuló en el capítulo anterior

$$\hat{y}(t+d|t) = E^d(z^{-1})y(t) + F^d(z^{-1})u(t)$$

el diagrama de bloques del sistema con controlador queda tal como se muestra en la figura 5.8a.



a)



b)

Figura 5.8: Diagrama de bloques del GPC

Operando se obtiene el diagrama de bloques de 5.8b, resultando el controlador:

$$C(z^{-1}) = -\frac{l_{y1}E^d(z^{-1}) + l_{y2}E^{d-1}(z^{-1})}{(1 - l_{u1}z^{-1})(1 - z^{-1})G_e - (l_{y1}F^d(z^{-1}) + l_{y2}F^{d-1}(z^{-1}))}$$

Por tanto, con esta configuración se puede abordar el estudio de la estabilidad robusta del proceso G con el controlador C . Los resultados sobre estabilidad serán ciertos siempre que el bloque P sea estable. Por consiguiente antes que nada hay que estudiar la estabilidad de P , definido como:

$$P(z^{-1}) = -\frac{l_{r1}}{l_{y1}E^d(z^{-1}) + l_{y2}E^{d-1}(z^{-1})}$$

sabiendo que (ver capítulo anterior)

$$\begin{aligned} E^d(z^{-1}) &= e_0^d + e_1^d z^{-1} = \sum_{i=0}^d a^i - \sum_{i=1}^d a^i z^{-1} \\ E^{d-1}(z^{-1}) &= e_0^{d-1} + e_1^{d-1} z^{-1} = \sum_{i=0}^{d-1} a^i - \sum_{i=1}^{d-1} a^i z^{-1} \end{aligned}$$

Se buscan pues las raíces del polinomio $l_{y1}E^d + l_{y2}E^{d-1}$. Este polinomio es de grado 1 y su raíz será:

$$z = \frac{l_{y1}e_1^d + l_{y2}e_1^{d-1}}{l_{y1}e_0^d + l_{y2}e_0^{d-1}} = \frac{(l_{y1} + l_{y2}) \sum_{i=1}^{d-1} a^i + l_{y1}a^d}{(l_{y1} + l_{y2}) \sum_{i=0}^{d-1} a^i + l_{y1}a^d}$$

como $l_{y1} + l_{y2} = -l_{r1}$ con $l_{r1} > 0$ y l_{y1} es siempre negativo, la raíz será tal que

$$|z| = \left| \frac{-l_{r1} \sum_{i=1}^{d-1} a^i + l_{y1}a^d}{-l_{r1} \sum_{i=0}^{d-1} a^i + l_{y1}a^d} \right| < 1$$

La última desigualdad es debida a que el sumatorio del denominador tiene un término más que el del numerador y todos los términos son negativos, por lo cual el cociente siempre es menor que uno y positivo, encontrándose pues la raíz en el círculo unidad. Por tanto queda demostrado que P es estable y no va a influir en los resultados sobre estabilidad del bucle cerrado.

5.3.2 Obtención de los límites de robustez

Como ya se ha indicado varias veces a lo largo de esta tesis, el tipo de proceso con el que se trabaja es aquél que puede ser modelado por un sistema de primer orden

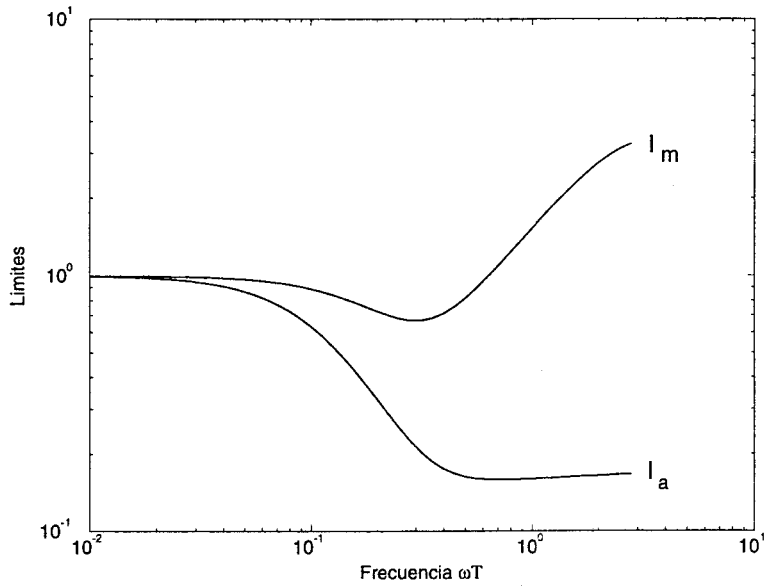


Figura 5.9: Forma característica de los límites \bar{l}_a y \bar{l}_m

más un retardo puro. Por tanto, trabajando ya en discreto se tendrá que la función de transferencia del proceso nominal será

$$\tilde{G}(z^{-1}) = \frac{b_0 z^{-1} + b_1 z^{-2}}{1 - a z^{-1}} z^{-d} \quad (5.4)$$

y la del controlador

$$C(z^{-1}) = \frac{l_{y1} E^d(z^{-1}) + l_{y2} E^{d-1}(z^{-1})}{(1 - l_{u1} z^{-1})(1 - z^{-1}) G_e - (l_{y1} F^d(z^{-1}) + l_{y2} F^{d-1}(z^{-1}))}$$

Con estas expresiones de G y C y haciendo uso de la condición (5.3) se van a obtener los límites de robustez para los sistemas que puedan ser descritos por (5.4), para todos los valores de los parámetros que describen el sistema (a , m y d). Así, para cada valor de la frecuencia ω los límites se pueden calcular como:

$$\bar{l}_m = \left| \frac{1 + \tilde{G}C}{\tilde{G}C} \right| \quad \bar{l}_a = \bar{l}_m | \tilde{G} |$$

En la figura 5.9 se muestra la forma genérica que toman los límites en función de ωT para determinados valores de a , m y d . Se observa que ambos límites son prácticamente constantes e iguales a la unidad para frecuencias bajas y a partir de cierto valor evolucionan (el límite aditivo \bar{l}_a disminuye y el multiplicativo \bar{l}_m aumenta). Nótese que estas curvas demuestran el alto grado de robustez del proceso con GPC ya que el límite

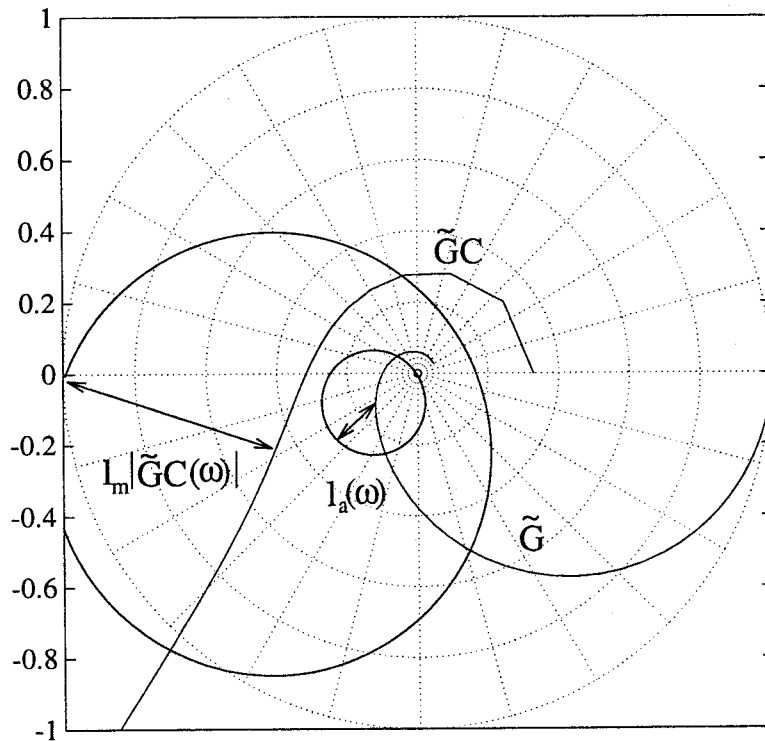


Figura 5.10: Diagrama polar del proceso \tilde{G} y de $\tilde{G}C$ mostrando los límites para una determinada frecuencia

\bar{l}_m es relativamente grande (igual a 1) a bajas frecuencias (donde normalmente las incertidumbres multiplicativas son siempre menores que la unidad) y aumenta considerablemente con la frecuencia, tal como lo suelen hacer las incertidumbres según se vió en la figura 5.2b. El pequeño valor de \bar{l}_a a altas frecuencias no debe resultar preocupante, ya que esto es debido a que el proceso en sí tiene ganancia muy pequeña. Recuérdese que ambos límites son dependientes y están relacionados por $\bar{l}_a = \bar{l}_m | \tilde{G} |$.

En la figura 5.10 se muestran las respuestas en frecuencia del proceso nominal solo y con el controlador así como los discos de radio \bar{l}_a y $\bar{l}_m | \tilde{G}C |$ para una frecuencia determinada. Dentro del disco de radio \bar{l}_a se pueden encontrar todos los procesos G pertenecientes a la familia \mathcal{F} en torno al proceso nominal definido por ($a = 0.9, m = 0, d = 0$), manteniendo la estabilidad del bucle cerrado.

Se puede observar que las formas de dichas respuestas conducen a las formas de \bar{l}_a y \bar{l}_m . Así, para frecuencias bajas, $\tilde{G}C(\omega)$ tiene un módulo grande (debido al término

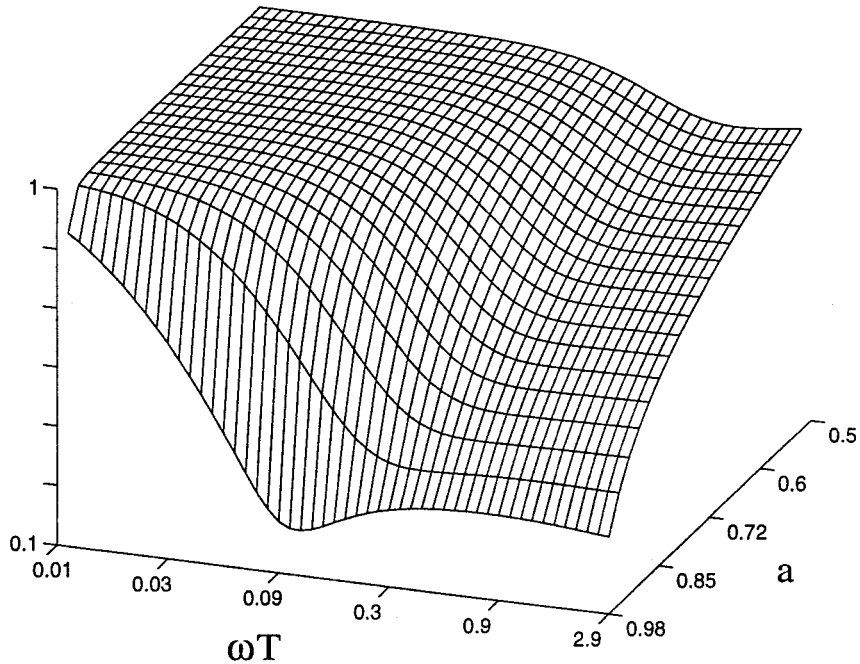


Figura 5.11: Límite \bar{l}_a para retardo 1 y diversos valores del polo

integral) con lo que su distancia al punto -1 coincidirá prácticamente con dicho módulo, conduciendo pues a un valor de \bar{l}_m igual a la unidad. Posteriormente, al aumentar ω llega un momento en que $\tilde{G}C$ se va alejando del -1 (sin disminuir prácticamente el valor de su módulo) y por consiguiente el límite puede aumentar sin peligro. Se pretende ahora ver cómo son estas zonas de estabilidad para todos los valores de los parámetros del modelo.

Para un valor concreto del retardo $d = 1$ y para un valor medio de m , ($m = 0.5$) se presenta en la figura 5.11 la evolución del límite \bar{l}_a con la frecuencia (ωT) para un campo de evolución de a entre 0.5 y 0.98. Como cabría esperar se ve que el límite es menor para polos mayores, ya que con polos del bucle abierto próximos al límite del círculo unidad la incertidumbre permitida es menor ya que sería más fácil entrar en la zona inestable ya en bucle abierto.

Se presentan ahora los resultados obtenidos en forma de gráficas para valores representativos del proceso nominal. Se van a tomar tres valores significativos de m : 0, que equivale a retardo $d + 1$, 1, que equivale a retardo d y 0.5 que representaría el retardo fraccionario en la mitad del intervalo $[d, d+1]$. Se presentan resultados para unos retardos determinados: 1, 2, 5 y 10 y para todos los valores del polo a entre 0.5 y

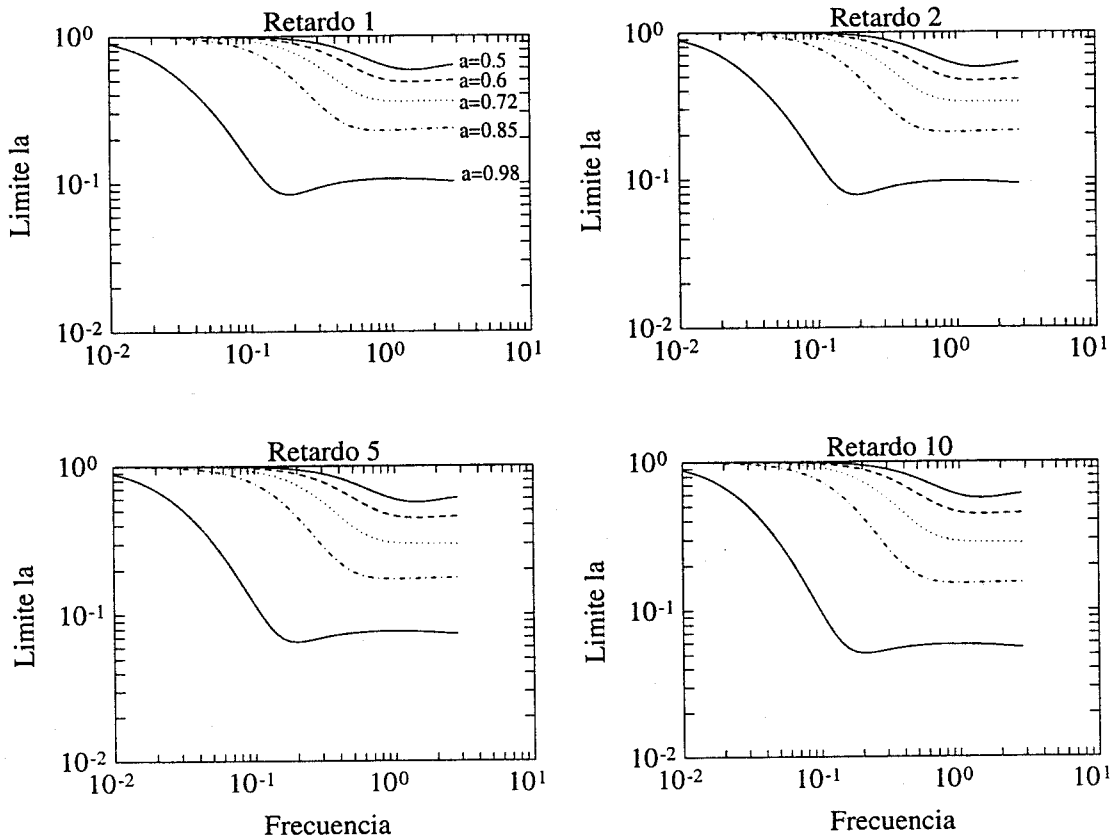


Figura 5.12: Límite \bar{l}_a para $m=0$ y diversos retardos

0.98.

En la figura 5.12 se tienen los límites de robustez para ciertos valores del polo y del retardo en bucle abierto en el caso de $m = 0$. Se observa claramente que en todos los casos el límite disminuye al aumentar el polo y al aumentar el valor del retardo; además empieza a disminuir a frecuencias más bajas, por lo que el ancho de banda también disminuye. Estos resultados concuerdan en líneas generales con los obtenidos para incertidumbres estructuradas. Para $m = 0.5$ se tienen análogos resultados en la figura 5.13 así como para $m = 1$ en la figura 5.14.

De dichas figuras se puede deducir que el valor de m no influye excesivamente, aunque valores pequeños de este parámetro reducen ligeramente los límites puesto que corresponden a retardos fraccionarios cercanos al $d = 1$. Sin embargo el tiempo muerto sí es un factor determinante. Su influencia queda más patente en la figura 5.15, donde para valores fijos de a y m se muestra el límite \bar{l}_a para diferentes retardos. Como era de esperar el aumento del retardo reduce claramente los límites de robustez. Es de

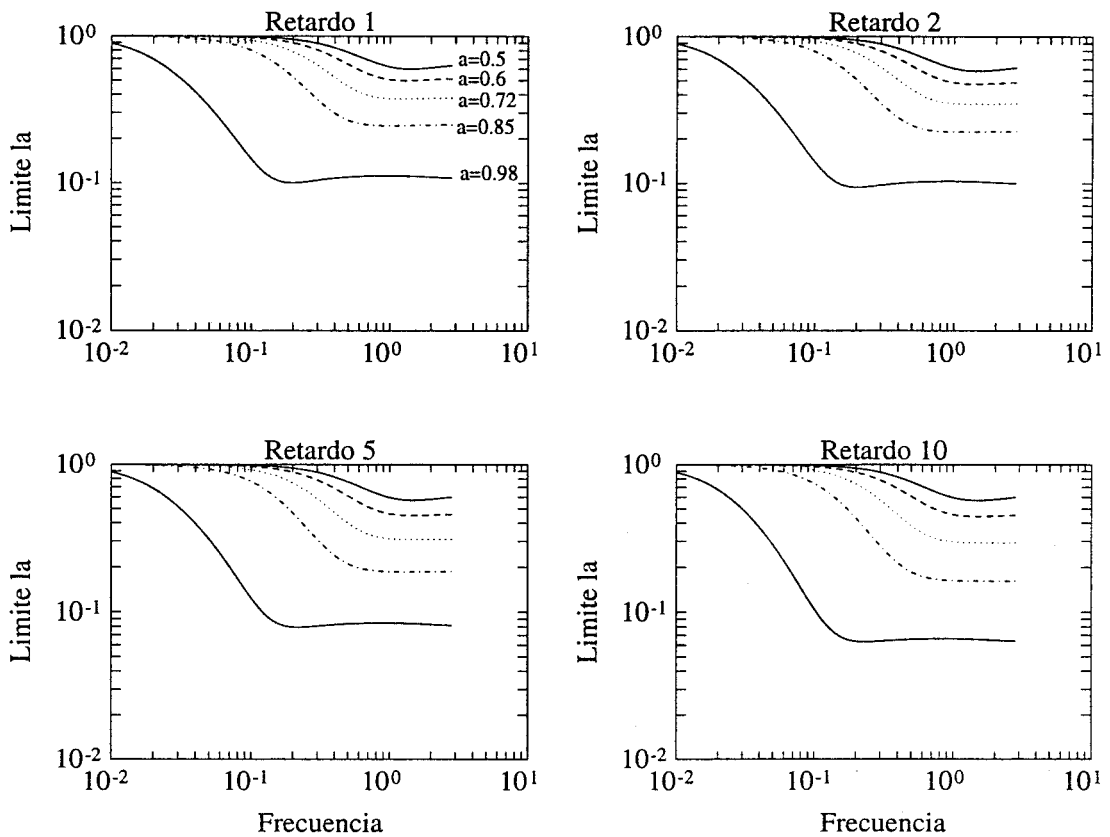


Figura 5.13: Límite \bar{l}_a para $m=0.5$ y diversos retardos

destacar que los valores obtenidos confirman al GPC como un método de gran robustez, lo cual corrobora los resultados presentados en el capítulo anterior.

Forma de uso de las curvas

Las curvas obtenidas no sólo son una muestra de la robustez del GPC, sino que constituyen una herramienta para saber lo cerca que se encuentra cada proceso de la zona peligrosa en que el sistema en bucle cerrado se puede hacer inestable.

El uso de las curvas que proporcionan los valores límite de \bar{l}_a resulta muy sencillo (siempre que se conozca algo de las incertidumbres). Una vez conocido el proceso nominal que constituye el "núcleo" de la familia \mathcal{F} de procesos, se dibujan en torno a él y para cada valor de ω los discos de radio $\bar{l}_a(\omega)$ obtenidos de las gráficas mostradas (ver figura 5.10). Si se conocen algunos puntos de la respuesta en frecuencia del proceso real, se puede comprobar fácilmente si éste se mantiene en el interior de los discos.

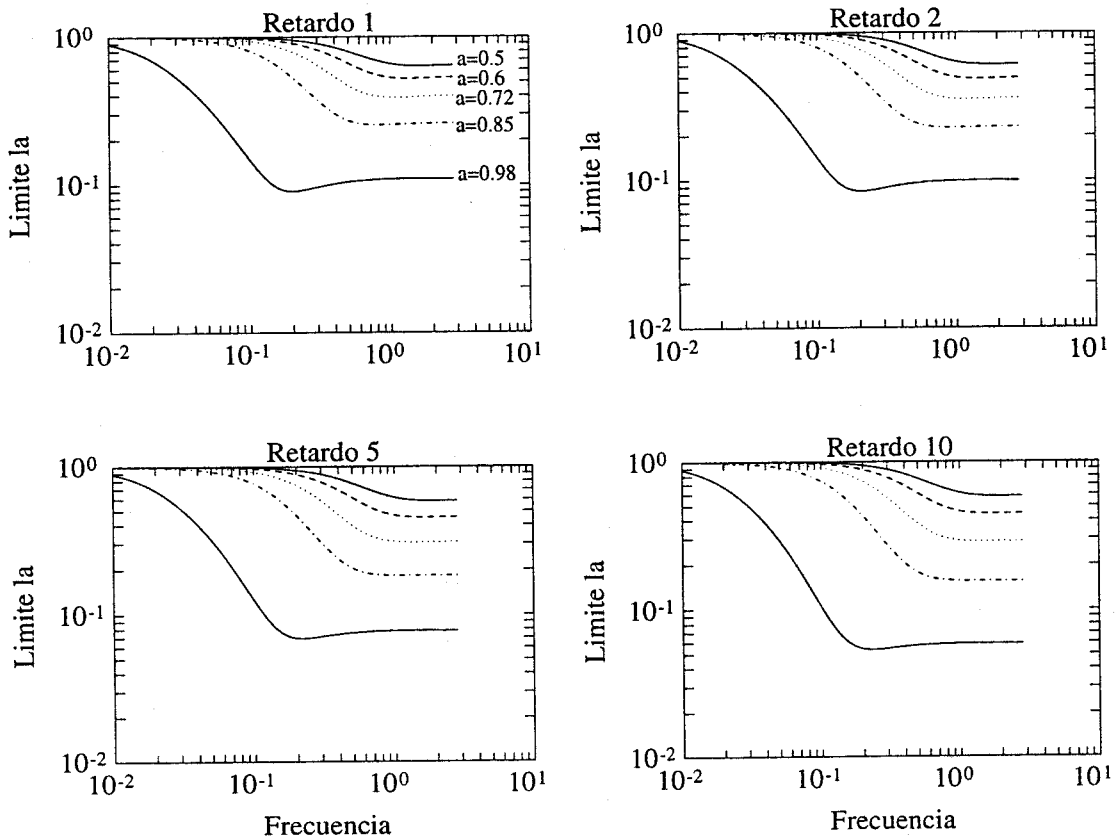


Figura 5.14: Límite \bar{l}_a para $m=1$ y diversos retardos

5.3.3 Influencia del parámetro λ

Los resultados numéricos presentados hasta el momento han sido calculados para un determinado valor del parámetro λ que, como se sabe, es el que pondera el esfuerzo de control en la función a minimizar. Valores muy pequeños de λ proporcionarán acciones de control bruscas mientras que valores grandes darán lugar a un control más suave. Como valor apropiado para unas prestaciones "normales" del GPC se puede tomar $\lambda = 0.8$. Este valor, para el campo de variación del proceso nominal que se está considerando proporciona una respuesta en bucle cerrado con un polo claramente dominante que da lugar a una constante de tiempo del orden del 60 – 70% de la de bucle abierto. Para obtener los límites usando un valor distinto de λ el procedimiento a seguir es el mismo.

El límite $\bar{l}_a(\omega)$ resultará afectado por la variación de λ en la manera que al disminuir este parámetro se pierde robustez, tal como se vió en el capítulo anterior. En la figura 5.16 se muestran los límites para tres valores diferentes del peso en el esfuerzo de

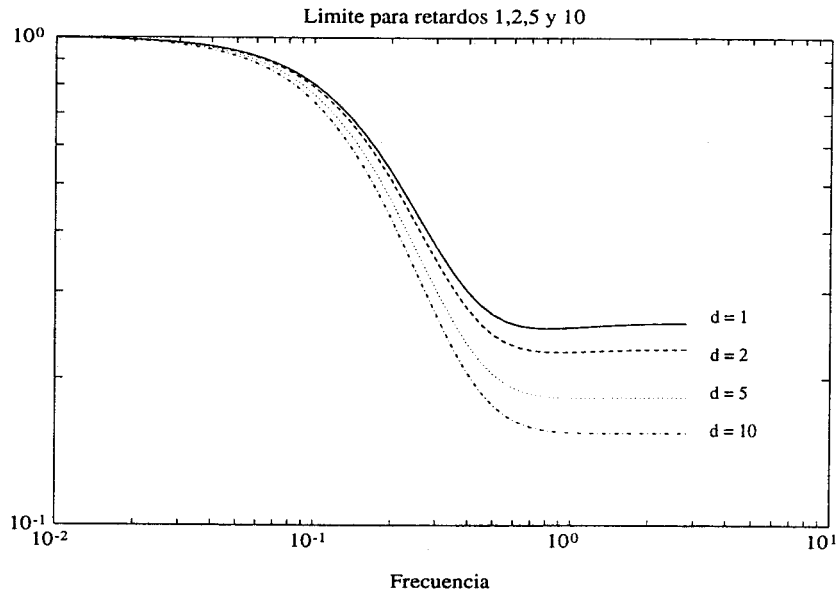


Figura 5.15: Influencia del retardo en \bar{l}_a

control, observándose la pérdida de robustez con la disminución de λ . Para valores pequeños ($\lambda = 0.3$) la disminución de \bar{l}_a resulta apreciable. Obsérvese que se puede conseguir un diseño más robusto trabajando con valores grandes de este parámetro aunque a costa de perder prestaciones, ya que con λ demasiado grande la respuesta del sistema se puede hacer excesivamente lenta.

5.4 Relación con incertidumbres estructuradas

Como se indicó previamente, la consideración de la incertidumbres según (5.1) puede englobar a las incertidumbres estructuradas, de tal manera que los límites obtenidos tratando a las incertidumbres estructuradas como si no lo fueran pueden ser más conservadores que si realmente se conoce la estructura de las incertidumbres y se aplican los resultados del capítulo precedente. Ambos límites serán más similares cuanto más se pueda asemejar la zona incierta a un disco.

Se pretende ver en esta sección que ambos límites son congruentes. Para ello se compararán los resultados obtenidos para dos casos muy claros: incertidumbre en la ganancia y en el retardo.

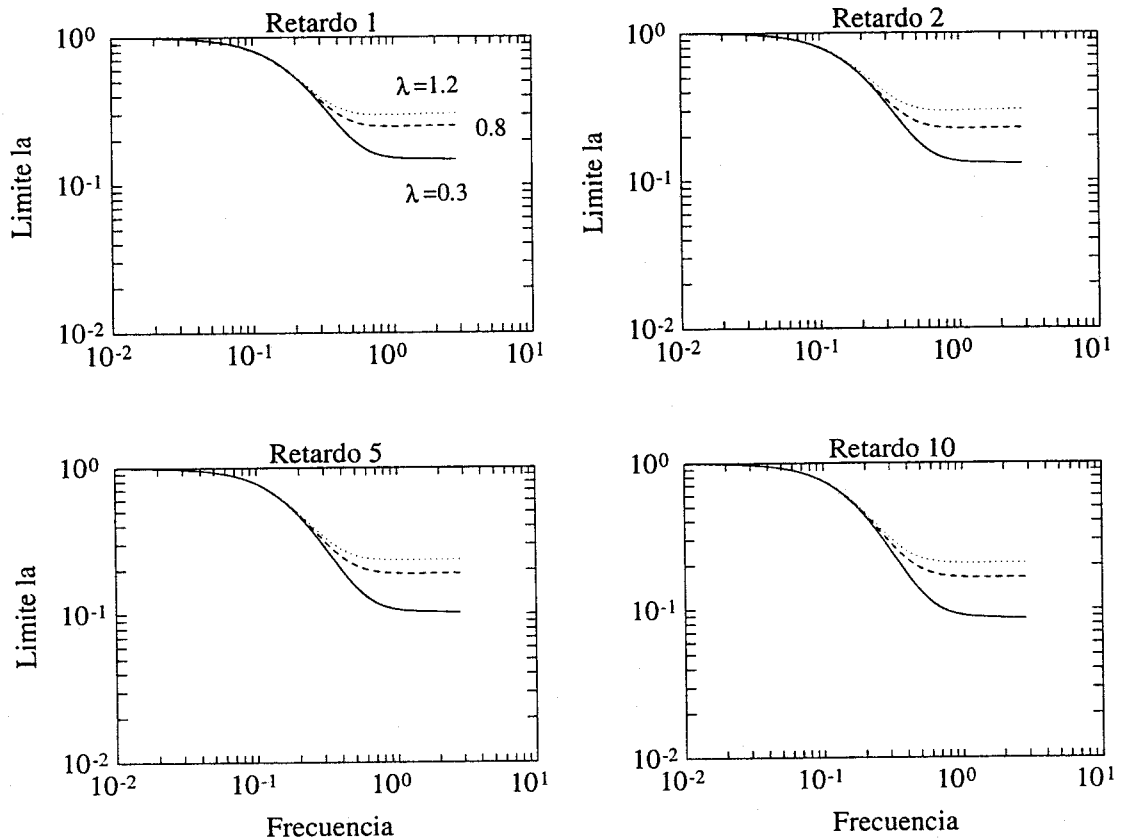


Figura 5.16: Influencia del parámetro λ en \bar{l}_a

5.4.1 Incertidumbre en la ganancia

En esta situación el proceso nominal \tilde{G} tiene una ganancia G_e mientras que la del proceso real será γG_e pudiendo ser γ mayor o menor que la unidad. Este factor provocará que ambos procesos tengan idéntica fase pero distinto módulo. Usando la descripción de incertidumbres no estructuradas utilizada hasta ahora, el radio del disco para cada ω será la diferencia entre ganancias del proceso y del nominal (figura 5.17). Es decir, la familia \mathcal{F} será:

$$\mathcal{F} = \{G : |G(i\omega) - \tilde{G}(i\omega)| \leq |\gamma - 1| |\tilde{G}(i\omega)| = \bar{l}_a(\omega)\}$$

Se busca pues el límite superior del radio de los discos de forma que se mantenga la estabilidad para todos los procesos de la familia \mathcal{F} . El límite de γ vendrá definido pues por

$$|\gamma - 1| = \min_{\omega} \frac{\bar{l}_a(\omega)}{|\tilde{G}(i\omega)|} \quad (5.5)$$

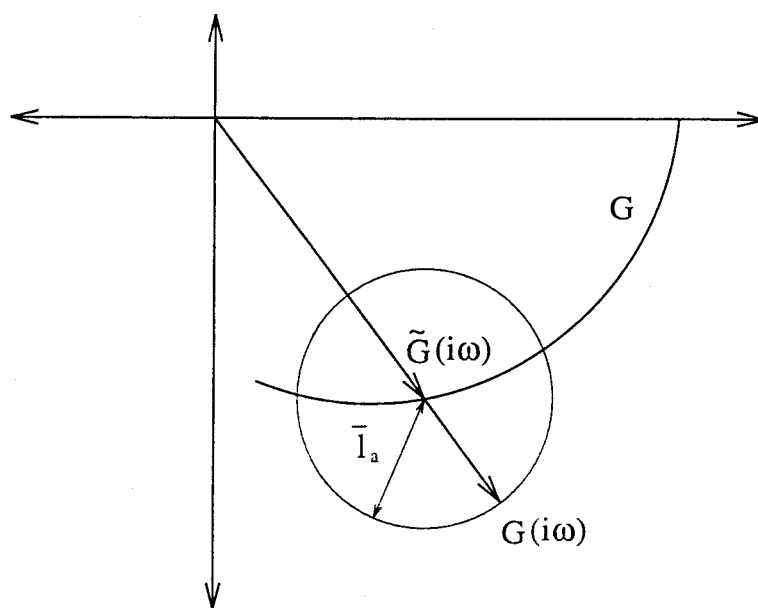


Figura 5.17: Tratamiento de la variación de la ganancia como incertidumbre no estructurada

Para un proceso nominal con los parámetros $a = 0.8$, $m = 1$ y $d = 5$ se muestra en la figura 5.18 el límite de estabilidad para ganancias de los procesos de la familia mayores que la del nominal. Todo proceso que esté en el interior de la zona sombreada será estable con el controlador diseñado para el proceso nominal. Se observa que un proceso con $\gamma = 1.6$ se encuentra siempre dentro de la zona segura, mientras que si se toma $\gamma = 1.8$ el proceso se sale de la franja estable para ciertos valores de omega, con lo cual el bucle cerrado será inestable. De hecho el valor máximo de γ se obtiene de la expresión (5.5) y para este caso en concreto toma el valor $\gamma = 1.68$.

Si se toma otro caso de estudio como por ejemplo $a = 0.9$, $m = 1$ y $d = 10$ se tienen los resultados de la figura 5.19. En ella se ve que el proceso con $\gamma = 1.5$ permanece en la zona estable mientras que si se aumenta γ hasta el valor 1.6 ya se sale de ella. Numéricamente se obtiene de la expresión 5.5 que el límite es $\gamma = 1.52$.

Estos resultados obtenidos para unos casos determinados se pueden comparar con los presentados en el capítulo anterior. Así, observando la figura 4.10 se tiene que para $a = 0.8$ y retardo 5 el proceso es estable para $\gamma = 1.6$ pero no lo es para $\gamma = 1.8$. En la misma figura se puede comprobar que para $a = 0.9$ y retardo 10 la estabilidad se mantiene para $\gamma = 1.5$ pero no para $\gamma = 1.6$. Se pueden tomar otros valores cualesquiera y comprobar que los resultados son similares, como era de esperar. La similitud entre ambas interpretaciones es debido a que en la zona más "peligrosa", es decir, cuando el proceso tiene fase cercana a -180° , el disco de radio \bar{I}_a que engloba la discrepan-

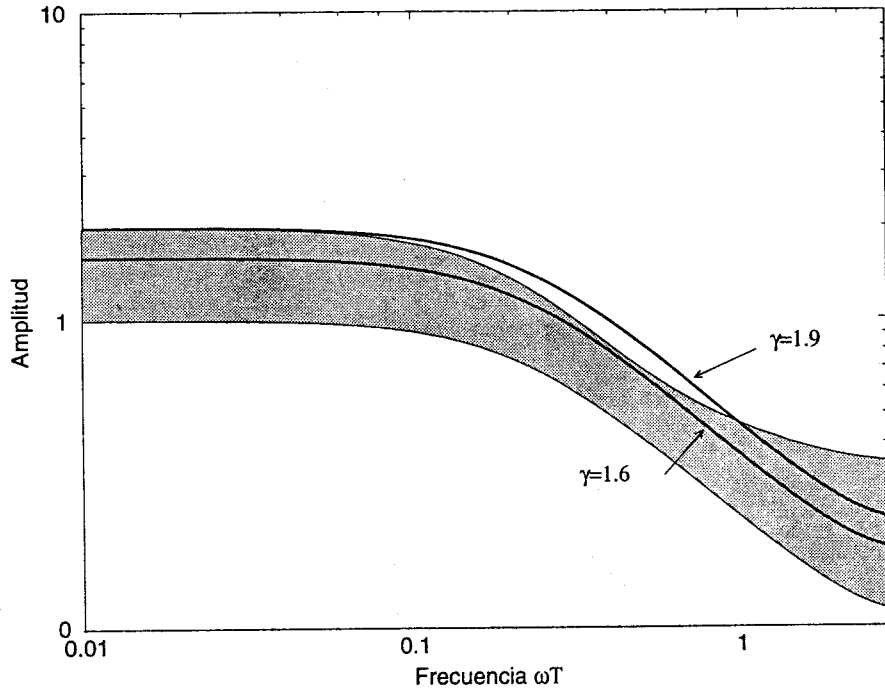


Figura 5.18: Límite de estabilidad para $a=0.8$, $m=1$ y $d=5$

cia de módulo (ver figura 5.17) es una buena representación de dicha incertidumbre estructurada.

5.4.2 Incertidumbre en el retardo

Se va a tratar ahora la incertidumbre en el retardo, que realmente puede ser tratada como estructurada, ya que se puede modelar cambiando en la $G(z)$ el término z^{-d} por $z^{-d \pm i}$. Análogamente a la ganancia, se va a considerar como no estructurada para aplicar los criterios de robustez anteriormente expuestos.

Como se sabe, un cambio en el retardo de cualquier proceso afecta exclusivamente al comportamiento de la fase de éste, dejando el módulo inalterado. Este hecho queda reflejado para una determinada frecuencia en la figura 5.20a. El radio \bar{l}_a del disco determina el desfase máximo que se permite entre los procesos de una misma familia.

La relación entre \bar{l}_a y α se puede obtener gráficamente de la figura 5.20b.

$$\operatorname{sen} \frac{\alpha(\omega)}{2} = \frac{\bar{l}_a(\omega)/2}{|\tilde{G}(i\omega)|}$$

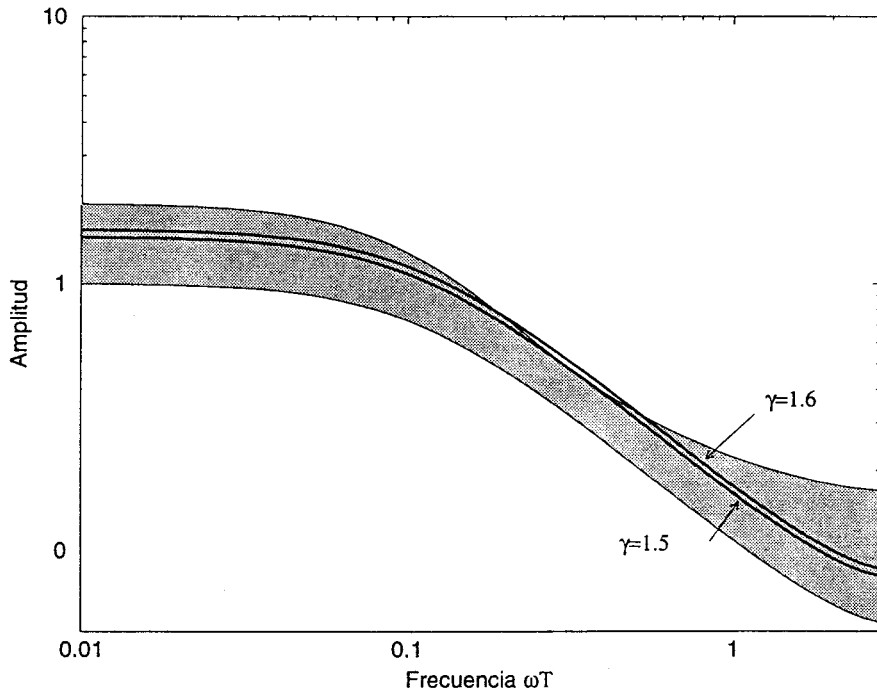


Figura 5.19: Límite de estabilidad para $a=0.9$, $m=1$ y $d=10$

y por tanto

$$\alpha(\omega) = 2 \arcsen \frac{\bar{l}_a(\omega)/2}{|\tilde{G}(i\omega)|} = 2 \arcsen \frac{\bar{l}_m(\omega)}{2} \tag{5.6}$$

Nótese que para aquellas frecuencias en que el límite sea tal que $\bar{l}_a > 2 |\tilde{G}|$ se permite cualquier desfase (ver figura 5.21) con lo cual cualquier variación en el retardo exclusivamente mantendría el proceso dentro de la zona de estabilidad (siempre que se trabajara en dichas frecuencias).

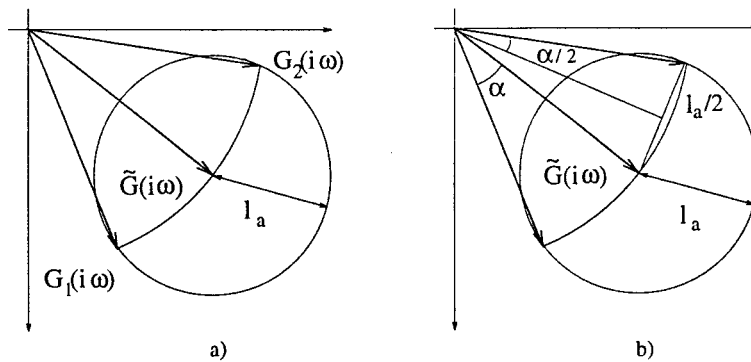


Figura 5.20: Incertidumbre en el retardo

Las diferencias en la fase al cambiar el retardo se acentúan al aumentar la frecuencia, siendo muy considerables para discrepancias de dos unidades o más con el proceso nominal (ver por ejemplo figura 4.4). Según esto se puede aventurar que los límites permitidos al retardo serán pequeños. A continuación se muestran algunos resultados que confirman esta hipótesis.

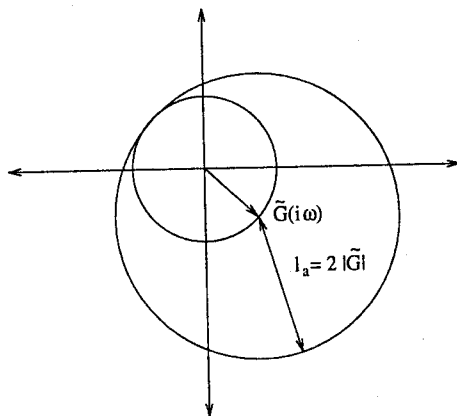


Figura 5.21: Zona estable a determinada frecuencia

Considérese por ejemplo un sistema con polo en $a = 0.8$, $m = 0$ y retardo $\bar{5}$. Se obtiene el límite superior \bar{l}_a tal como se ha indicado previamente y a partir de éste se obtiene la zona robusta para variaciones exclusivamente en el retardo usando la expresión (5.6). En la figura 5.22 se muestra esta zona para retardos mayores, obtenida restando al valor de la fase del proceso el valor de α para cada frecuencia. Se muestran también las fases de los procesos con retardos una y dos unidades mayores que, como se puede observar, van divergiendo conforme se aumenta la frecuencia. Por otro lado, la zona estable se va estrechando, provocando que la fase para retardo igual a 7 se salga de dicha zona, indicando que el comportamiento es inestable. Estos resultados se pueden comparar con los de la figura 4.16; en ella se muestra que para una discrepancia de una unidad en el retardo el sistema es estable y para dos ya deja de serlo. Los resultados de ambas formulaciones parecen por tanto similares. Análogo resultado se obtiene para retardos inferiores al nominal.

Para otros procesos nominales se pueden comparar resultados con los obtenidos para las incertidumbres estructuradas. Por ejemplo para $a = 0.9$, $m = 0$ y $d = 2$, se observa (figura 5.23) que la zona segura engloba en su interior los procesos con una unidad de discrepancia en el retardo (desde retardo 1 a retardo 3 inclusive). Este resultado puede ser comparando con la figura 4.16; en ella se indica que la discrepancia permitida es de dos unidades, es decir, desde retardo nulo hasta retardo 4 inclusive. Ambos resultados no concuerdan, entonces ¿cuál es el resultado correcto? Para una variación exclusivamente en el retardo el resultado correcto lo proporciona la figura 4.16, debido a que el resultado obtenido en esta sección se refiere a cualquier incertidumbre que

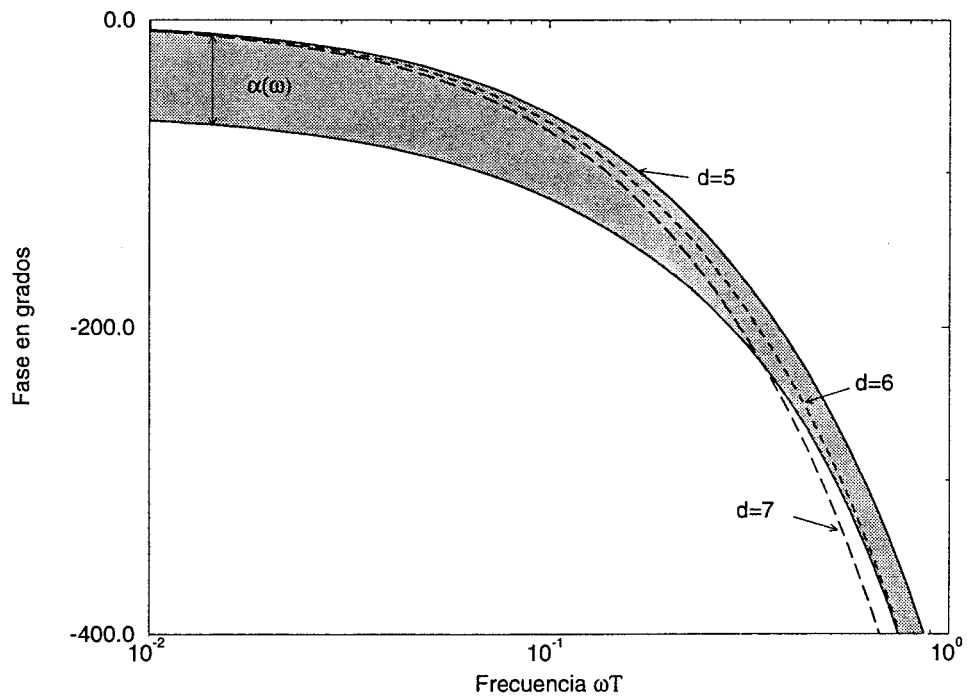


Figura 5.22: Procesos con retardos mayores que el nominal

pueda ser englobada en el disco elegido (ver figura 5.20), el cual ciertamente incluye otros efectos. El tema es pues que al intentar tratar el retardo como incertidumbre no estructurada se están incluyendo otros efectos adicionales.

Los resultados obtenidos con este método para el retardo son claramente conservadores debido a la zona elegida para modelar la incertidumbre. Ahora en la zona "peligrosa" el disco no proporciona una buena representación de la incertidumbre en el retardo, a diferencia del disco que representaba la incertidumbre en la ganancia del apartado anterior. Para fases cercanas -180° los posibles procesos $G(\omega)$ con igual módulo pero retardo distinto a $\tilde{G}(\omega)$ no están mucho más próximos al punto -1 que $G(\omega)$ aunque el disco incluirá procesos que sí lo estarán y por tanto harán disminuir el límite.

5.4.3 Comparación de resultados

Los resultados obtenidos en este capítulo son cualitativamente iguales a los obtenidos en el capítulo previo. Es decir, el factor que más afecta a la robustez es el retardo, por sus efectos a frecuencias altas y el aumento del polo provoca pérdida de robustez mientras que el parámetro m apenas influye.

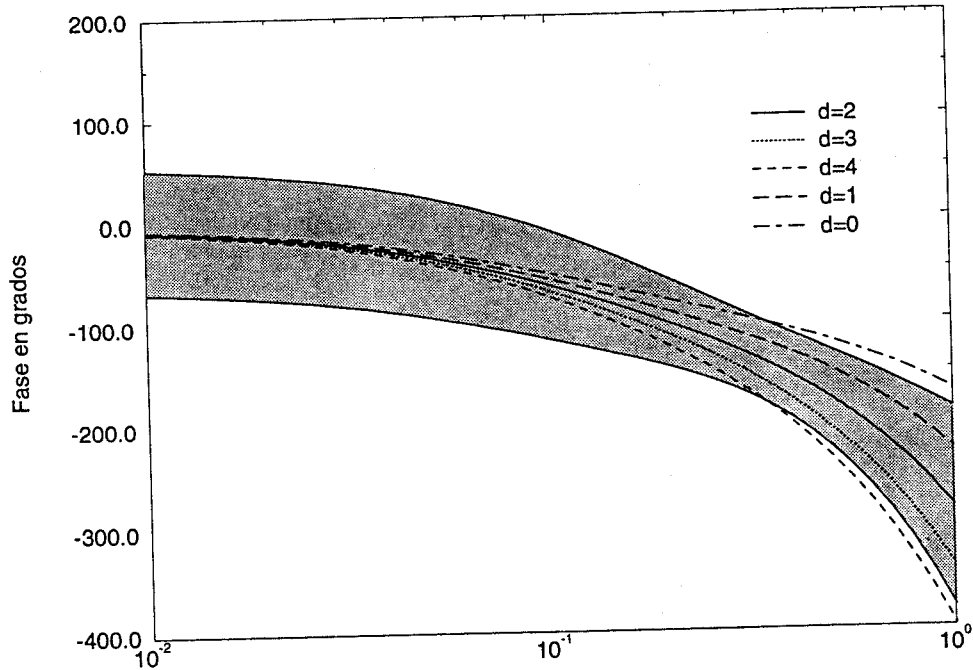


Figura 5.23: Procesos con retardos mayores y menores que el nominal

El estudio en frecuencia proporciona resultados más conservadores, de manera que se asegura una zona estable. Si se dispone de alguna información sobre las incertidumbres de forma que puedan ser parametrizadas, se puede afinar un poco más en la definición de esta zona estable, obteniendo resultados más concretos. Aunque, como se ha mostrado, con el estudio en frecuencia se pueden tratar también las incertidumbres estructuradas obteniendo resultados similares a los del capítulo previo, parece más congruente usar una metodología u otra según la naturaleza de las incertidumbres. De este modo, si se conoce el parámetro responsable de la incertidumbre se usarán los resultados del capítulo anterior mientras que si ésta puede ser debida a la variación conjunta de muchos efectos es recomendable usar los resultados de éste.

5.5 Mejora de la robustez mediante polinomios de ponderación

Como se ha indicado, no se ha pretendido diseñar un controlador robusto, sino cuantificar de alguna manera la robustez que esta formulación del GPC posee (se ha resultado pues el problema de análisis pero no el de síntesis). Pero, de cualquier forma, no se pueden pasar por alto las técnicas existentes para mejorar la robustez del algoritmo

estándar.

Ya Clarke en su presentación del Control Predictivo Generalizado [23] apunta la posibilidad del uso de polinomios adicionales $P(z^{-1})$ y $T(z^{-1})$ como elementos de diseño análogos a los empleados en el Controlador de Mínima Varianza Generalizada [19] como polinomios de ponderación. El primero de ellos sería empleado para seguimiento del modelo mientras que el segundo podría usarse para la mejora de la robustez.

En [22] se muestra la reformulación de algoritmo del GPC incluyendo estos polinomios. Para ello se parte del modelo CARIMA expresado de la forma:

$$A(z^{-1})y(t) = B(z^{-1})u(t-1) + \frac{T(z^{-1})}{\Delta}\xi(t)$$

El polinomio $T(z^{-1})$ ha sido considerado hasta el momento igual a 1, describiendo de esta forma las perturbaciones más usuales. Ahora se considerará como un parámetro de diseño. Por tanto las predicciones no serán óptimas pero como contraprestación se podrá ganar robustez frente a incertidumbres, interpretación similar a la usada por Ljung [56]. Este polinomio pues puede ser considerado como un prefiltro o como un observador. Una selección apropiada de este polinomio puede reducir los errores de predicción y proporcionar mayor robustez, como se muestra en [70].

El polinomio $P(z^{-1})$ puede aparecer al definir una salida auxiliar

$$\psi(t) = P(z^{-1})y(t)$$

de forma que pondera la influencia de la salida. Este polinomio permite ampliar los objetivos de control, usando sus raíces como parámetros de diseño. De esta forma se puede conseguir un control *dead-beat*, por asignación de polos o LQ.

Introduciendo estos polinomios el algoritmo del GPC es replanteado en [22] obteniendo una nueva formulación. Para ello la ecuación diofántica que proporciona la predicción cambia de $1 = E_j(z^{-1})A\Delta + z^{-j}F_j(z^{-1})$ a $T(z^{-1}) = E_j(z^{-1})A\Delta + z^{-j}F_j(z^{-1})$ y se sigue un procedimiento análogo al mostrado en [23]. El problema consiste ahora en la elección de los valores apropiados de T y P . Al estar interesados en la robustez, se busca el valor óptimo para T .

5.5.1 Elección del polinomio T

La elección del polinomio de filtrado T no es trivial. Normalmente se asume que el filtrado *más fuerte* (considerando más fuerte al filtrado con ancho de banda más pequeño

o a igualdad de ancho de banda con mayor pendiente) posee mejores propiedades de robustez para incertidumbres a altas frecuencias. Pero esto no es siempre cierto, como se muestra con algunos contraejemplos en [92]. En dicho artículo se presentan las líneas maestras para la correcta elección del polinomio T basadas en el aumento del margen de robustez a altas frecuencias y se llega a la conclusión de que para procesos estables en bucle abierto la mejor elección es

$$T(z^{-1}) = A(1 - \beta z^{-1})^{N_1 - \delta(P)}$$

donde β estará cerca de la raíz dominante de A , N_1 es el mínimo horizonte de predicción y $\delta(P)$ es el grado del polinomio P .

Nótese que la idea de filtrado para mejorar la robustez también subyace en el Control por Modelo Interno (IMC) [60], donde una vez obtenido el controlador que proporciona unas determinadas características, es desintonizado con un filtro para así aumentar la robustez.

5.5.2 Relación con otras formulaciones

En el artículo previamente citado también se estudia la relación con \mathcal{H}_∞ y se comparan ambos métodos, obteniéndose resultados equivalentes en cuanto a robustez. Esto implica que el prefiltrado con T puede ser una alternativa asequible, reduciendo considerablemente los cálculos necesarios para el diseño en el entorno \mathcal{H}_∞ , cálculos apreciables sobre todo en una estrategia adaptativa.

Con la introducción del polinomio T se puede aumentar la robustez pero no hay que olvidar que ya no se tiene la predicción óptima. De alguna forma esta idea es similar a la empleada con el Regulador Lineal Cuadrático Gaussiano (LQG) para recuperar las buenas propiedades de robustez que pierde el Regulador Lineal Cuadrático (LQR) con la inclusión del observador (ver por ejemplo [36]). Esto se consigue con el método LQG/LTR y consiste en recuperar la función de transferencia en bucle abierto (Loop Transfer Recovery, LTR), de forma que se aproxime a la función de transferencia en bucle abierto del método LQR. Esto se puede lograr actuando sobre los parámetros del filtro de Kalman, trabajando con unas covarianzas *ficticias* (ver [37]). De esta forma se gana robustez a costa de empeorar la estimación.

Tanto en un caso como en otro no hay porqué considerar preocupante la pérdida de optimalidad en la predicción o en la estimación. Siendo pragmáticos, lo que interesa es que el controlador cumpla sus objetivos y sea robusto, siendo la predicción un medio y no un fin, por lo que el método empleado para conseguir los fines puede resultar irrelevante.

Capítulo 6

Aplicaciones

En este capítulo se van a presentar tres aplicaciones de Control Predictivo Generalizado usando el método presentado y analizado a lo largo de la tesis. Estas aplicaciones abarcan un amplio campo dentro de los procesos industriales y se pueden considerar como un banco de pruebas suficientemente amplio del controlador.

En primer lugar se presenta la aplicación al control adaptativo de una planta solar, obteniendo resultados reales sobre un proceso difícil de controlar. Posteriormente se aplica a un proceso típico en la industria como es un evaporador (simulado mediante su modelo no lineal) donde se corroboran algunos de los resultados obtenidos sobre robustez. Se termina con la aplicación real a una planta piloto de laboratorio, con el controlador implantado sobre un sistema de control distribuido comercial.

6.1 Control de una planta solar

6.1.1 Descripción de la planta

El proceso a controlar se trata del campo de colectores solares distribuidos de la Plataforma Solar de Almería (Tabernas, Almería). Este campo consiste básicamente (ver figura 6.1) en una tubería por la que circula aceite y sobre la cual se concentra la radiación solar por medio de espejos cilindro-parabólicos, los cuales siguen el movimiento del sol en un eje. Se dispone en total de 480 módulos agrupados en diez lazos paralelos, cada uno de los cuales mide 172 metros (de los cuales 30 son pasivos, es decir, no

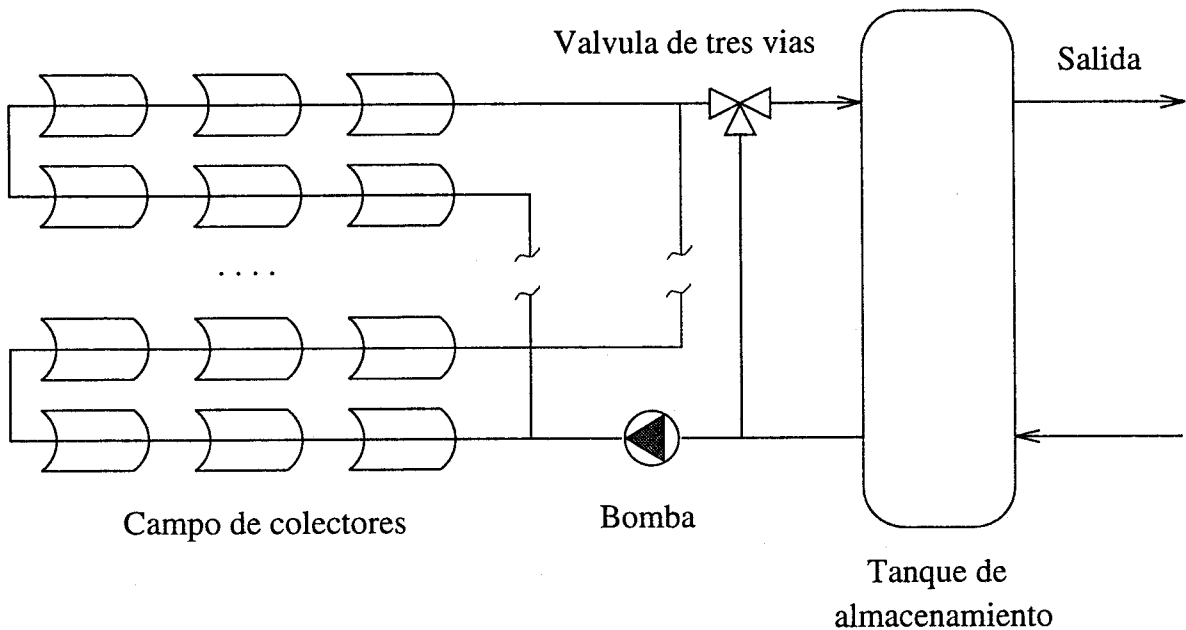


Figura 6.1: Esquema de la planta solar

reciben la radiación concentrada). El aceite se calienta tras atravesar el campo y se introduce entonces en un tanque de almacenamiento para ser usado en la producción de electricidad o bien en una planta desalinizadora. El aceite extraído de la parte inferior del tanque (donde la temperatura es menor) vuelve a recircular por el campo. Además se dispone de una válvula de tres vías que permite la recirculación del aceite hasta que éste alcanza una temperatura adecuada para entrar en el tanque de almacenamiento. Una descripción más completa del campo se puede encontrar en [41].

Una característica determinante de una planta solar es que la fuente de energía no puede ser manipulada. Además, la intensidad de la radiación solar depende de las variaciones cíclicas diarias y estacionales así como de las nubes, humedad ambiental o transparencia del aire. El objetivo del sistema de control es mantener una temperatura de salida del aceite constante a pesar de las cambiantes condiciones de operación. Este objetivo se logra variando el caudal de fluido a través del campo.

El mantenimiento de una temperatura constante a lo largo del día mientras las condiciones solares cambian requiere amplias variaciones en el caudal, lo que da lugar a una variación considerable de la dinámica del proceso. Los parámetros del controlador deben ajustarse a las condiciones de operación, por lo que se va a emplear un controlador autoajustable. Debido a la dinámica tan variable que posee y a las fuertes perturbaciones, esta planta ha sido usada para probar distintos tipos de controladores [12], [11].

6.1.2 Estrategia de control

Para llevar cabo el control se emplea un modelo lineal de primer orden con retardo que relaciona el caudal de aceite con la temperatura de salida. Este modelo se obtiene realizando diversas pruebas en bucle abierto examinando la curva de reacción. De estas pruebas se deduce que el retardo τ_d es relativamente pequeño comparado con la constante de tiempo τ , lo que justifica la elección de un tiempo de muestreo T igual al valor más pequeño de τ_d (que corresponde a las condiciones de operación de caudal mayor). Se tiene por tanto

$$G(z) = \frac{bz^{-1}}{1 - az^{-1}} z^{-k}$$

con $k = 1$ en las condiciones de caudal alto.

El sistema de control va a incluir, aparte del propio GPC un controlador *feedforward* en serie [10]. La inclusión de este elemento queda justificada por la necesidad de rechazar las perturbaciones de variables medibles como son la radiación solar y la temperatura de entrada al campo del fluido. Los cambios de estas variables provocan cambios en la salida que no son debidos a la acción de control, lo que daría lugar a un ajuste innecesario de los parámetros estimados por el identificador. Para más detalles sobre el cálculo y la utilidad del *feedforward* ver [16]. En el esquema de control se considera este término como incluido en el proceso, con lo que la señal de control es la temperatura de referencia del *feedforward*.

El diseño del controlador sigue los pasos del capítulo 3. Al trabajar en la región de caudal alto, la ley de control viene dada por

$$\Delta u(t) = l_{y1} \hat{y}(t + 1 | t) + l_{y2} y(t) + l_{r1} r(t)$$

y el predictor es

$$\hat{y}(t + j + 1 | t) = (1 + a)\hat{y}(t + j | t) - a\hat{y}(t + j - 1 | t) + b \Delta u(t + j - 1)$$

donde el polo a varía entre 0.85 y 0.95.

Los parámetros del sistema se calculan en línea mediante un estimador de mínimos cuadrados recursivo, con factor de olvido variable y una supervisión que lo desconecta cuando las señales no contienen suficiente dinámica.

6.1.3 Resultados

El GPC ha sido probado con éxito en la planta solar por Berenguel [11]. En la figura 6.2 se muestran los resultados de operación durante un día típico (caracterizado por la evolución de la radiación solar). El caudal debe evolucionar entre 4.5 y 7 litros por segundo para mantener la temperatura a pesar de los cambios en la dinámica del proceso. Estas pruebas han sido realizadas con un valor pequeño del peso del esfuerzo de control λ y proporciona una respuesta rápida a los cambios de setpoint; si se desea menor sobreoscilación este factor debe ser incrementado.

Aunque pueda parecer que los resultados no son muy buenos (por la existencia de sobreoscilaciones), hay que considerar que se trata de un proceso realmente difícil de controlar. En cualquier caso los resultados obtenidos por este método superar a los obtenidos con controladores PID fijos o adaptativos, como se muestra también en [11].

6.2 Control de la composición en un evaporador

Como banco de pruebas del GPC se ha elegido un proceso muy común en la industria como es un evaporador. Este proceso involucra un buen número de variables interrelacionadas entre sí y aunque puede parecer algo simple si se compara con otros procesos de mayor dimensión, permite comprobar el funcionamiento de cualquier técnica de control. Los resultados que se presentan en esta sección han sido obtenidos vía simulación sobre un modelo no lineal del proceso.

6.2.1 Descripción del proceso

El proceso en cuestión es un evaporador de circulación forzada en el cual la materia prima es mezclada con una extracción del producto y es bombeada a través de un intercambiador de calor vertical por el que circula vapor de agua, que se condensa en los tubos. La mezcla se evapora y pasa por una vasija de separación, donde se separan el líquido y el vapor. El primero se hace recircular extrayendo parte como producto final mientras que el vapor es condensado con agua de refrigeración. El esquema de este proceso, que es usado en muchos sectores productivos, como por ejemplo en industrias azucareras, se muestra en la figura 6.3.

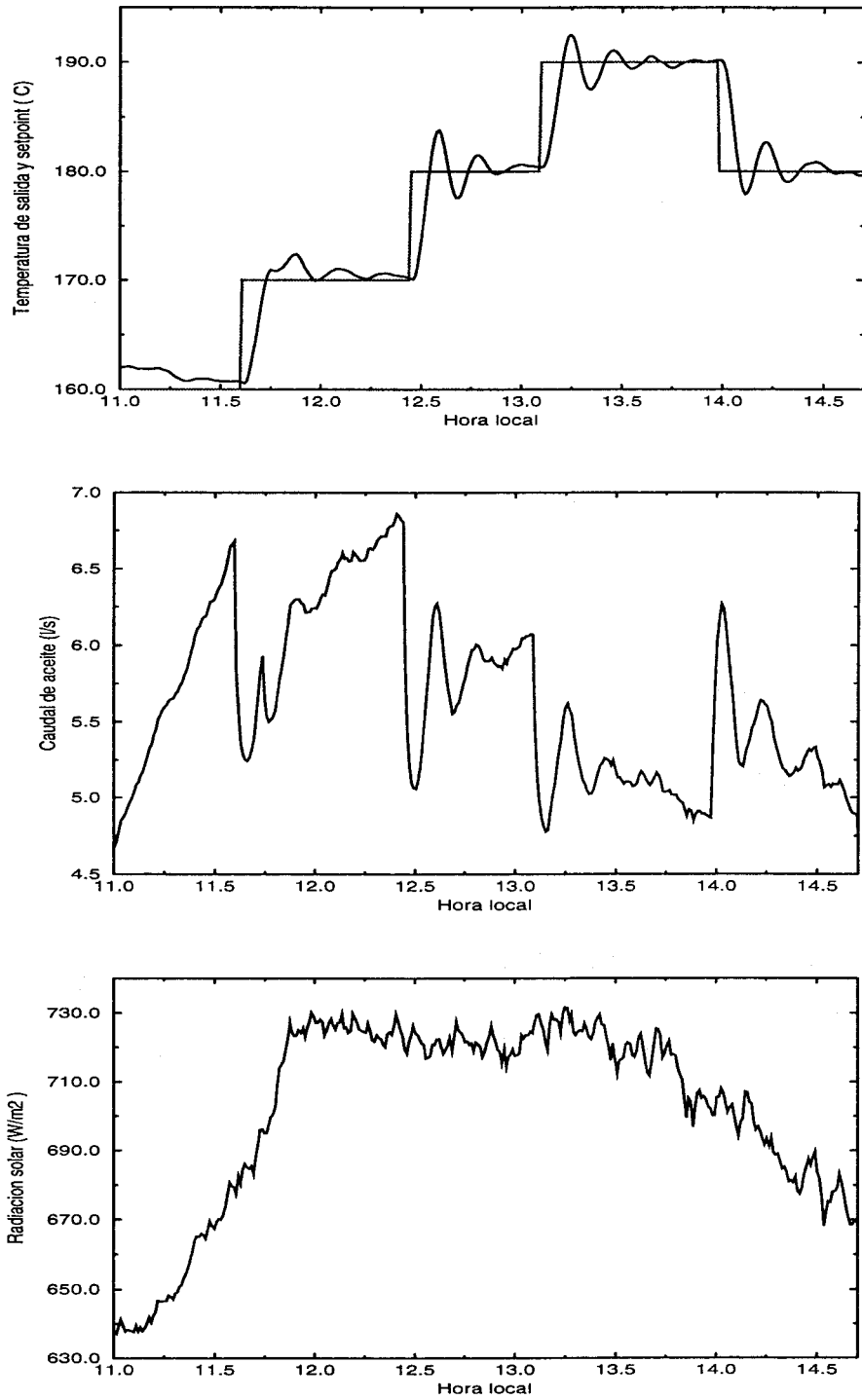


Figura 6.2: GPC adaptativo: temperatura de salida, caudal y radiación solar

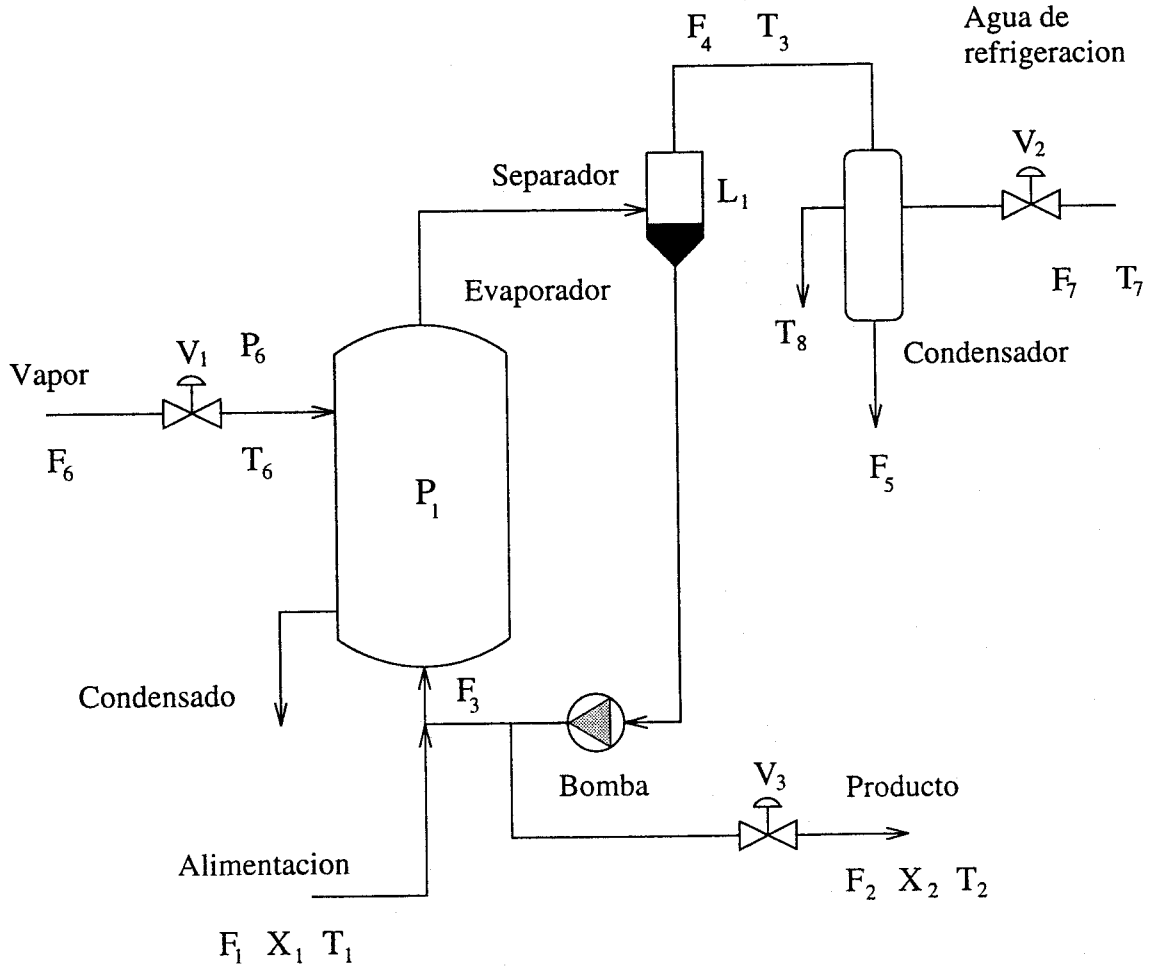


Figura 6.3: Esquema del evaporador

El comportamiento del proceso puede ser modelado mediante una serie de ecuaciones obtenidas al realizar un balance de materia y energía, además de hacer ciertas suposiciones más o menos realistas. Las ecuaciones que describen el comportamiento del proceso se han extraído de [63]. En la tabla 6.1 quedan recogidas las principales variables, así como sus valores en el punto de operación.

Balance de materia en el líquido. Un balance de materia en el líquido (solvente y soluto) proporciona:

$$\rho A \frac{dL_1}{dt} = F_1 - F_4 - F_2 \quad (6.1)$$

siendo ρ la densidad del líquido y A la sección del separador. El producto ρA se puede considerar constante e igual a 20 kg/m.

Balance de materia en el soluto. El balance de materia en el soluto en la fase

Variable	Descripción	Valor	Unidades
F_1	Caudal de alimentación	10.0	kg/min
F_2	Caudal de producto	2.0	kg/min
F_3	Caudal de recirculación	50.0	kg/min
F_4	Caudal de vapor	8.0	kg/min
F_5	Caudal de condensado	8.0	kg/min
X_1	Composición de la alimentación	5.0	%
X_2	Composición del producto	25.0	%
T_1	Temperatura de alimentación	40.0	grados
T_2	Temperatura del producto	84.6	grados
T_3	Temperatura de la fase vapor	80.6	grados
L_1	Nivel del separador	1.0	m
P_1	Presión de operación	50.5	kPa
F_6	Caudal de vapor	9.3	kg/min
T_6	Temperatura del vapor	119.9	grados
P_6	Presión del vapor	194.7	kPa
Q_6	Aporte de calor	339.0	kW
F_7	Caudal de agua de refrigeración	208.0	kg/min
T_7	T. entrada agua refrigeración	25.0	grados
T_8	T. salida agua refrigeración	46.1	grados
Q_7	Calor cedido en el condensador	307.9	kW

Tabla 6.1: Variables del proceso y valores en el punto de operación

líquida conduce a:

$$M \frac{dX_2}{dt} = F_1 X_1 - F_2 X_2 \quad (6.2)$$

M es la cantidad total de líquido en el evaporador, considerada constante e igual a 20 kg.

Balance de materia en el vapor de proceso. La cantidad total de vapor de agua puede ser expresada en función de la presión que existe en el sistema según:

$$C \frac{dP_1}{dt} = F_4 - F_5 \quad (6.3)$$

siendo C una constante que convierte la masa de vapor en una presión equivalente y toma un valor de 4 kg/kPa.

Balance de energía en el líquido. Se supone que el líquido se encuentra siempre en su punto de ebullición y que la mezcla es perfecta. La temperatura del líquido

se puede obtener linealizando la curva de saturación en torno al punto de trabajo y añadiendo un término para incluir la elevación del punto de ebullición debido a la presencia del soluto:

$$T_2 = 0.5616P_1 + 0.3126X_2 + 48.43 \quad (6.4)$$

Análogamente la temperatura del vapor es:

$$T_3 = 0.507P_1 + 55.0 \quad (6.5)$$

Se puede considerar que la dinámica del balance de energía es muy rápida, de manera que

$$F_4 = (Q_6 - F_1 C_p (T_2 - T_1)) / \lambda \quad (6.6)$$

donde el calor específico $C_p = 0.07 \text{ kW/K(kg/min)}$ y el calor latente de vaporización $\lambda = 38.5 \text{ kW/(kg/min)}$. El cambio del calor sensible entre T_2 y T_3 para F_4 se considera pequeño comparado con el calor latente. Además, se asume que no existen pérdidas de calor hacia el ambiente ni incrementos debidos al bombeo.

Intercambiador La presión del vapor P_6 determina la temperatura del vapor bajo ciertas condiciones de saturación. Aproximando la relación entre presión y temperatura del vapor saturado por una linealización en torno al punto de trabajo se tiene:

$$T_6 = 0.1538P_6 + 90.0 \quad (6.7)$$

El aporte de calor viene dado por

$$Q_6 = UA_1(T_6 - T_2) \quad (6.8)$$

donde UA_1 es el coeficiente global de transferencia de calor multiplicado por el área, que es una función del caudal a través del intercambiador:

$$UA_1 = 0.16(F_1 + F_3)$$

El caudal de vapor se puede calcular usando el calor latente del vapor en las condiciones de saturación, $\lambda_s = 36.6 \text{ kW/(kg/min)}$:

$$F_6 = Q_6 / \lambda_s \quad (6.9)$$

En relación con la dinámica general, la del intercambiador se considera muy rápida.

Condensador La dinámica de condensador también puede considerarse muy rápida. El balance energético en el agua de refrigeración conduce a la siguiente relación:

$$Q_7 = F_7 C_p (T_8 - T_7)$$

Si UA_2 , el coeficiente global de transferencia de calor multiplicado por el área, se considera constante e igual a 6.84 kW/K la ecuación de transferencia de calor puede aproximarse por:

$$Q_7 = UA_2(T_3 - 0.5(T_7 + T_8))$$

Si se elimina T_8 de las dos últimas ecuaciones:

$$Q_7 = \frac{UA_2(T_3 - T_7)}{1 + UA_2/(2C_p F_7)} \quad (6.10)$$

y por tanto

$$T_8 = T_7 + Q_7/(F_7 C_p) \quad (6.11)$$

El caudal de condensado viene dado por

$$F_5 = Q_7/\lambda \quad (6.12)$$

con $\lambda = 38.5 \text{ kW/K(kg/min)}$ (calor latente de vaporización del agua en las condiciones de trabajo).

Grados de libertad Como se ve existen doce ecuaciones para veinte variables, existiendo por tanto ocho grados de libertad. Para cerrar el problema es necesario pues considerar ocho ecuaciones más; éstas serán las que proporcionan los valores de las variables manipuladas y de las perturbaciones:

- Tres variables manipuladas: La presión del vapor P_6 , que depende de la apertura de la válvula V_1 , el caudal de agua de refrigeración F_7 , gobernado por la válvula V_2 y el caudal de producto F_2 , con V_3 .
- Cinco perturbaciones: Caudal de alimentación F_1 , caudal de recirculación F_3 , composición y temperatura de la alimentación X_1 y T_1 , y temperatura del agua de refrigeración T_7 .

Con estas consideraciones se puede obtener una solución única que permite calcular el valor de las restantes variables.

6.2.2 Obtención del modelo lineal

Como se deduce de las ecuaciones (6.1)-(6.12) el proceso es un sistema no lineal y con una fuerte interrelación entre las variables. Aún así, para el diseño del controlador se va a trabajar con un modelo lineal y con varios lazos independientes entre sí. Está

claro que estas hipótesis de trabajo no son correctas tal como se pondrá de manifiesto al poner en funcionamiento el control.

El control del evaporador incluye el mantenimiento de unas condiciones de trabajo estables así como la obtención de producto de una determinada calidad. Para obtener el primer objetivo es necesario controlar la masa y la energía del sistema, lo cual se lleva a cabo manteniendo constantes el nivel en el separador L_1 y la presión de operación P_1 . Para ello se utilizarán dos controladores locales tipo PI, de forma que actuando sobre el caudal de producto F_2 se controla el nivel L_1 y con el agua de refrigeración F_7 se controla la presión de operación P_1 . La justificación de la elección de los emparejamientos y la sintonización de estos lazos se puede encontrar en [63] y [25].

El otro objetivo es la obtención de una determinada composición de producto. Esto se consigue actuando sobre la variable manipulada restante que es la presión del vapor que aporta la energía al evaporador, P_6 . La interacción entre las variables es muy fuerte, como se puede comprobar por el método de Bristol [8], e incluso se podría habería emparejado X_2 con F_7 y P_1 con P_6 .

Para la obtención del modelo lineal del lazo de composición se provoca un escalón en la entrada (P_6) y se estudia el efecto sobre la salida (X_2). Como era de esperar, y según se desprende de la figura 6.4, donde se muestra la evolución de las variables más significativas a un escalón del 10% en la presión del vapor, la interacción entre las variables es considerable.

La evolución de la composición no sigue pues el patrón de un sistema de primer orden, debido principalmente a que la experiencia no se ha realizado en bucle abierto, ya que están actuando los reguladores de nivel y presión que, como se ha indicado, deben estar activados para un funcionamiento estable del evaporador y afectan indirectamente a la composición.

Se va a intentar aproximar el comportamiento del lazo $X_2 - P_6$ al de un modelo de primer orden con retardo de la forma

$$G(s) = \frac{K}{1 + \tau s} e^{-\tau_d s}$$

que, como se sabe, a pesar de su simplicidad es muy usado en la práctica. Para obtener los parámetros del modelo se va a emplear el método de la curva de reacción, que proporciona a partir de la gráfica de la respuesta del sistema a una entrada en escalón los valores de K , τ y τ_d .

Debido a la no linealidad, el comportamiento del sistema va a ser distinto para entradas de distinto valor y de distinto signo. Realizando diversas experiencias para

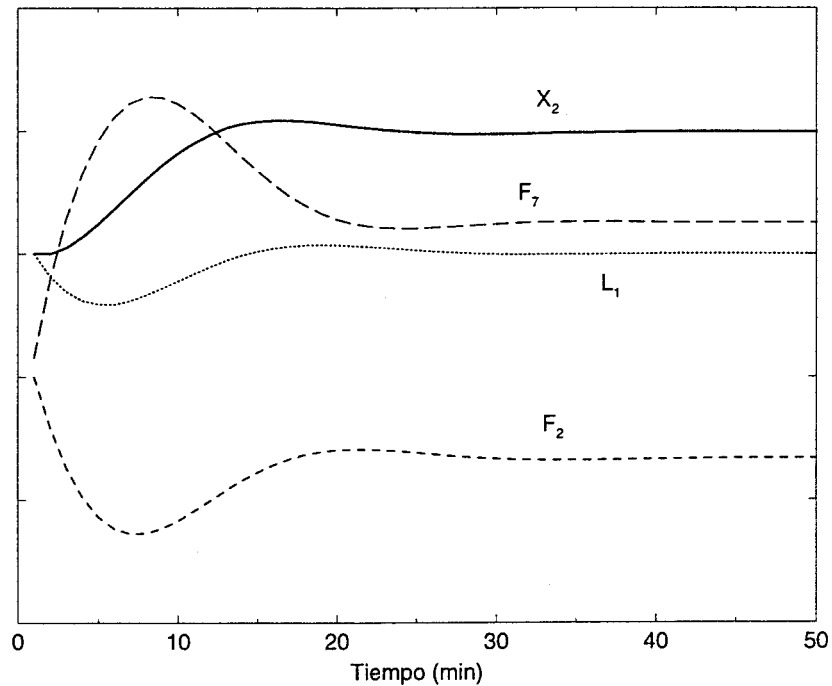


Figura 6.4: Respuesta del sistema ante escalón a la entrada

escalones de distinto signo y magnitud, se pueden considerar apropiados los siguientes parámetros:

$$k = 0.234 \text{ \%/kPa} \quad \tau = 4.5 \text{ min} \quad \tau_d = 3.5 \text{ min}$$

Al tomar un tiempo de muestreo de un minuto el retardo no es entero, por lo que la función de transferencia en discreto en el entorno del punto de trabajo será (expresiones de conversión en el capítulo 3)

$$G(z^{-1}) = \frac{0.02461z^{-1} + 0.02202z^{-2}}{1 - 0.8007374z^{-1}} z^{-3}$$

Esta función de transferencia se empleará para el diseño del controlador a pesar de las limitaciones que posee por la existencia de los fenómenos anteriormente expuestos.

6.2.3 Diseño del controlador

Una vez obtenido un modelo lineal del proceso, el diseño del controlador es directo si se usa el GPC precalculado. Basta con calcular los parámetros que aparecen en la ley de control:

$$\Delta u(t) = l_{y1}\hat{y}(t+d|t) + l_{y2}\hat{y}(t+d-1|t) + l_{r1}r(t) + l_{u1}\Delta u(t-1) \quad (6.13)$$

Si se quiere diseñar un regulador fijo (no adaptativo), basta con calcular estos parámetros una sola vez. En el caso del evaporador con $a = 0.8007374$, $m = b_0/(b_0 + b_1) = 0.528$ y para un valor de λ de 1.2 se tiene:

$$\begin{aligned} l_{y1} &= -2.2748 \\ l_{y2} &= 1.5868 \\ l_{r1} &= 0.6879 \\ l_{u1} &= -0.1862 \end{aligned}$$

siendo por tanto la señal de control en cada instante

$$u(t) = 0.8138u(t-1) + 0.1862u(t-2) - 2.2748\hat{y}(t+d | t) + 1.5868\hat{y}(t+d-1 | t) + 0.6879r(t)$$

Para completar el cálculo de la ley de control (6.13) se necesitan los valores predichos de la salida en los instantes $t+d$ y $t+d-1$. Este cálculo se realiza con facilidad dada la simplicidad del modelo. Basta con lanzar la ecuación del modelo hacia el futuro:

$$\begin{aligned} \hat{y}(t+i) &= (1+a)\hat{y}(t+i-1) - a\hat{y}(t+i-2) + b_0(u(t-d+i-1) - \\ &\quad - u(t-d+i-2)) + b_1(u(t-d+i-2) - u(t-d+i-3)) \quad i = 1 \dots d \end{aligned}$$

siendo los elementos $\hat{y}(t) = y(t)$ e $\hat{y}(t-1) = y(t-1)$ valores conocidos en el instante t . Nótese que si se quisiera hacer el controlador adaptativo, bastaría únicamente con calcular el nuevo valor de las l_i cuando cambiasen los parámetros del sistema. La simplicidad de la ley de control obtenida resulta manifiesta, siendo comparable a la un PID digital, y resulta pues fácil de implantar en cualquier sistema de control por simple que éste sea

6.2.4 Pruebas

Se presentan a continuación algunos resultados de aplicar la ley de control anterior sobre el evaporador (simulado mediante su modelo no lineal). A pesar de que las simplificaciones que se han empleado en la fase de diseño (sistema monovariable, modelo lineal de primer orden) no eran muy realistas, se obtiene un comportamiento razonablemente bueno del sistema en bucle cerrado.

Así en la figura 6.5 se muestra el comportamiento del proceso ante cambios en la referencia de la composición. Se observa que la salida sigue claramente a la referencia aunque con una alguna sobreoscilación inicial. Hay que tener en cuenta que los lazos considerados independientes presentan gran interrelación entre sí, y particularmente la

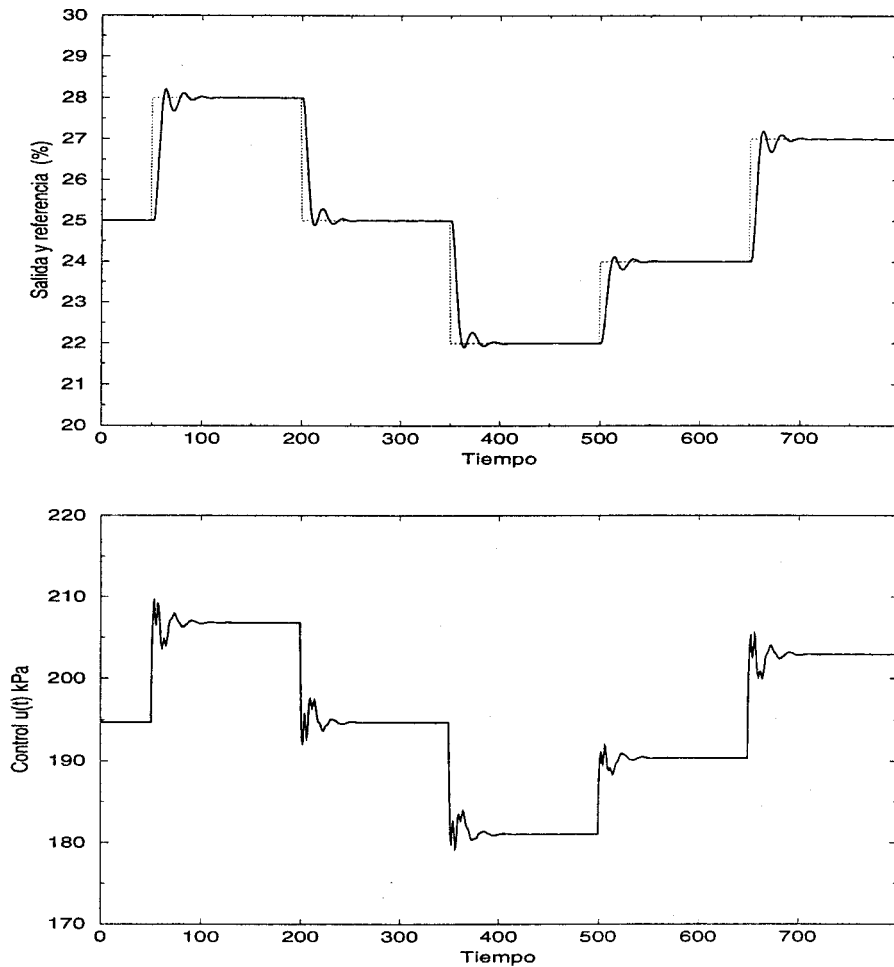


Figura 6.5: Funcionamiento del GPC en el evaporador

composición se ve muy alterada por las variaciones en el caudal de agua de refrigeración F_7 , el cual varía continuamente para mantener constante la presión de operación.

A pesar de la sobreoscilación el comportamiento puede considerarse bueno. Se puede comparar con el que se obtiene con un controlador clásico como es un PI. Unos buenos valores para el ajuste de este controlador son los calculados en [63], modificando ligeramente los obtenidos por el método de Ziegler-Nichols:

$$K = 1.64 \text{ kPa/\%} \quad T_I = 3.125 \text{ min}$$

En la figura 6.6 se comparan ambos reguladores para un cambio en la referencia del 28 al 25 por ciento. Se nota que el GPC es más rápido y sobreoscila menos que el PI a la vez que no introduce gran complejidad en el diseño, tal como se ha visto en la sección anterior. Las respuestas de ambos controladores ante cambios en la alimentación apa-

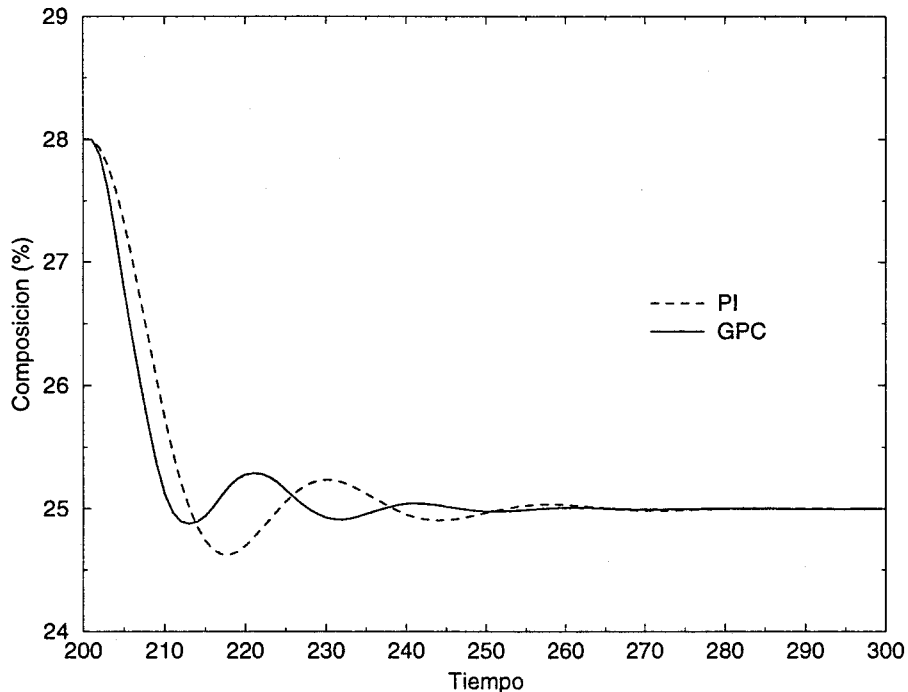


Figura 6.6: Comparación de GPC y PI para un cambio en la referencia

recen reflejadas en la figura 6.7, donde en el instante $t = 50$ se cambia el caudal de alimentación de 10 a 11 kg/min y en $t = 200$ la composición a la entrada cambia bruscamente del 5 al 6 por ciento. Se ve que estos cambios afectan considerablemente a la composición del producto y aunque ambos controladores devuelven la salida al valor de referencia, el GPC lo hace antes y con menor sobreoscilación, reduciendo los picos en torno a un 30%. La diferencia con un PI sintonizado directamente por Ziegler-Nichols es aún más notable.

También se pueden hacer pruebas en cuanto al estudio de la robustez. Realmente ya se sabe que el modelo que se ha usado no corresponde con el real (que no es de primer orden ni lineal) y por tanto queda probado que el controlador posee ya cierta robustez. Pero para corroborar los resultados de robustez presentados previamente se va a suponer que el modelo que se ha usado es perfecto y se verá qué ocurre con las incertidumbres en el modelo.

Si se trabaja por ejemplo con un modelo que posee un error en la estimación del polo, de forma que $\hat{a} = \alpha a$ con $\alpha = 0.9$, es decir, un error del 10%, habrá que calcular los nuevos parámetros l_i usando $\hat{a} = 0.9 \times a = 0.9 \times 0.8007374 = 0.7206$ (supuesto que $a = 0.8007374$ es el valor "real"). Recalculando pues la ley de control (incluido el predictor) para este nuevo valor, y dejando la ganancia inalterada, se obtiene la respuesta que se muestra en la figura 6.8. Como se ve, la composición apenas se ve

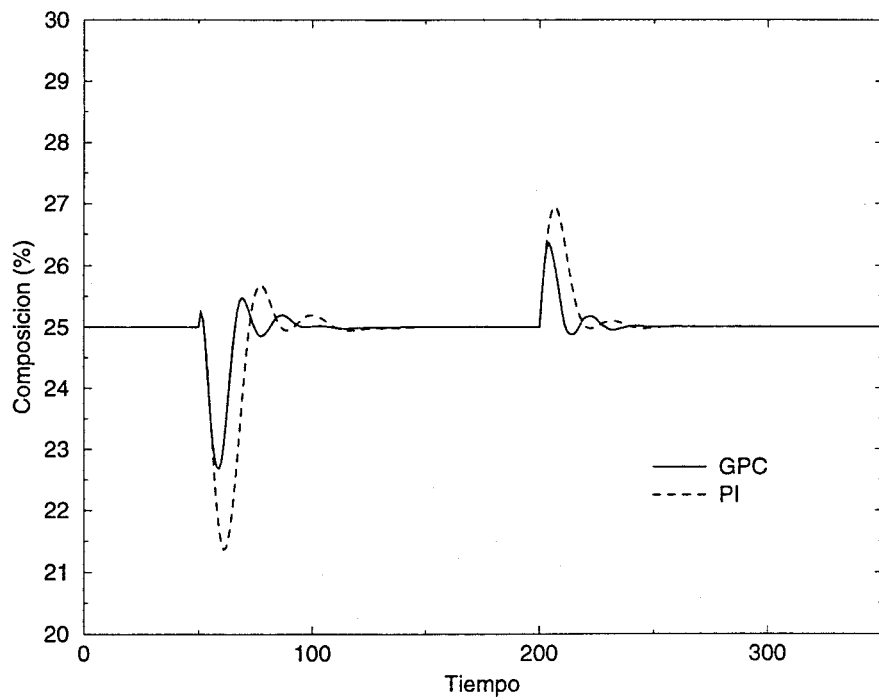


Figura 6.7: Comparación de GPC y PI para cambios en la alimentación

alterada por este error de modelado, obteniéndose una respuesta similar a la del modelo inicial.

Lo mismo se puede hacer cambiando la ganancia del sistema. Para los valores del modelo considerado como "bueno" se puede ver que incertidumbres de hasta el 100% en la ganancia se pueden permitir sin ningún problema. Aumentando pues hasta el doble la ganancia del modelo usado y calculando la nueva ley de control, se obtiene un comportamiento más lento pero que no deja de ser satisfactorio, tal como se muestra en la misma figura.

El conocimiento del retardo es un factor fundamental en los métodos predictivos basados en modelo, del tal forma que grandes errores en su estimación pueden dar lugar a inestabilidad. Mientras que para una diferencia de una unidad la respuesta apenas varía, no ocurre lo mismo si la discrepancia es de dos unidades o más. En la figura 6.9 se muestra el efecto de usar un modelo con retardo 2 siendo el "real" igual a 3. La respuesta es defectuosa aunque sin llegar a la inestabilidad. Nótese que una vez detectado un fallo en el valor del retardo la corrección es muy sencilla puesto que basta con calcular más o menos valores de $\hat{y}(t+d)$ usando la ecuación del modelo cuyos demás parámetros no han cambiado. Además los valores de los coeficientes l_i no hay que cambiarlos al ser independientes del retardo.

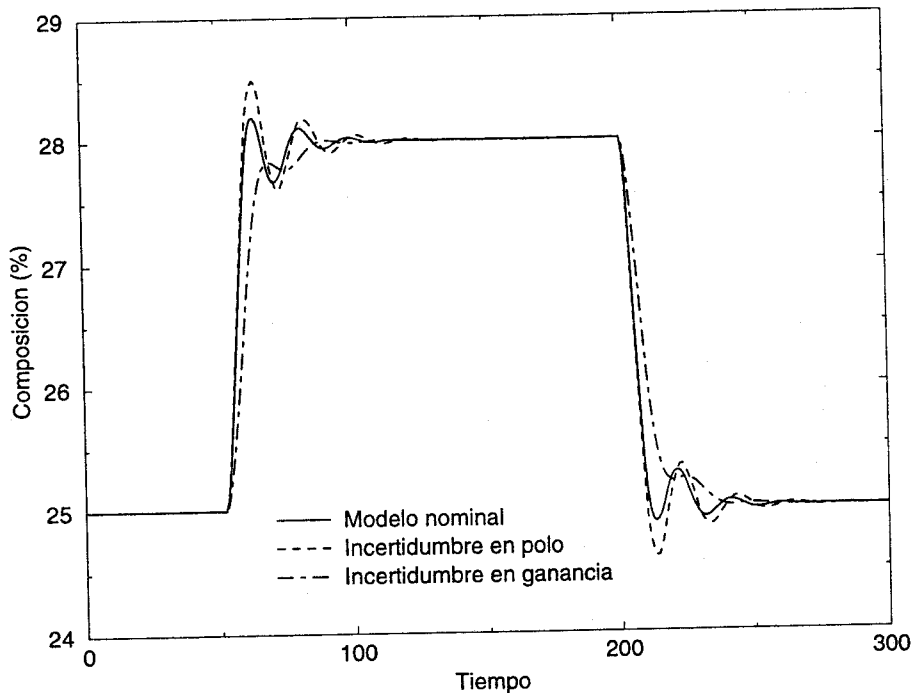


Figura 6.8: Influencia de errores en la estimación de la ganancia y el retardo

6.3 Planta Piloto

Con la intención de mostrar cómo se puede implantar el GPC en un sistema de control distribuido industrial, se presentan en esta sección aplicaciones al control de las magnitudes más usuales que se pueden encontrar en la industria de procesos, tales como caudales, niveles o temperaturas.

Como banco de pruebas de nuevas estrategias de control se dispone en el laboratorio del Departamento de Ingeniería de Sistemas y Automática de una Planta Piloto para control e instrumentación industrial. Esta Planta es básicamente un sistema con agua como fluido de trabajo en el que pueden tener lugar diversos procesos termodinámicos con intercambio de masa y energía. En esencia consiste en un depósito con calentamiento interno con una serie de tuberías de entrada y salida y un circuito de recirculación con un intercambiador de calor. En la figura 6.10 se muestra una imagen de la Planta.

El diseño de la Planta permite el ensayo de diversas estrategias de control en un gran número de lazos. Según la configuración elegida se puede controlar temperatura, caudal, presión o nivel. Para ello se dispone de cuatro actuadores: tres válvulas automáticas y una resistencia eléctrica que calienta el interior del depósito. Posterior-

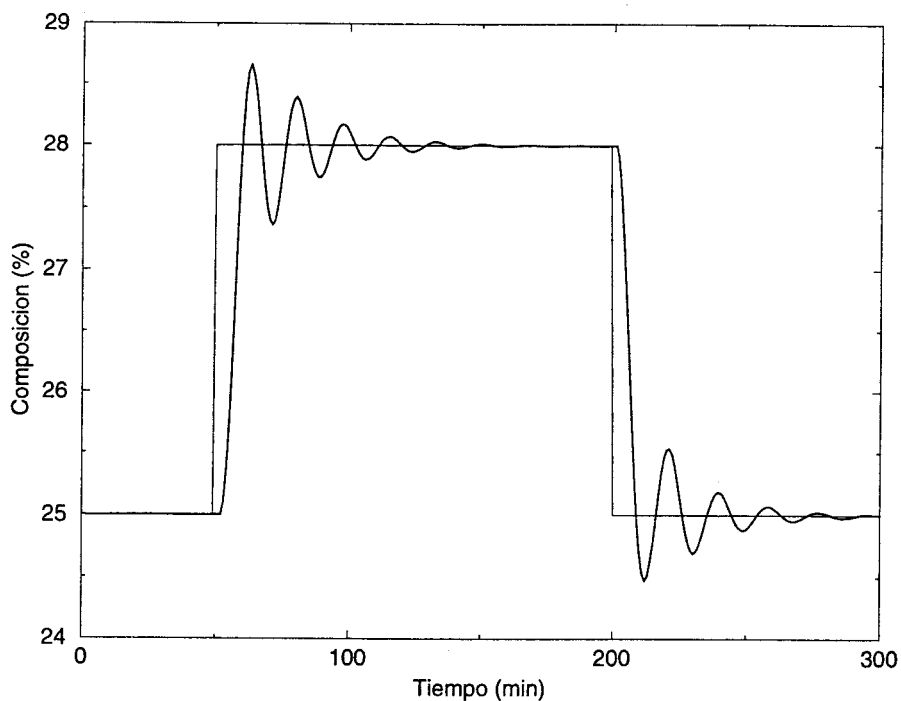


Figura 6.9: Influencia del error en la estimación del retardo

mente se elegirán algunos de los posibles lazos (considerados independientes) para la implantación de controladores GPC.

6.3.1 Descripción de la planta

El esquema de la Planta se muestra en la figura 6.11 donde se señalan los principales elementos así como la localización de los diversos instrumentos.

Los principales elementos que la forman son:

- **Circuito de alimentación.** La Planta dispone de dos tuberías de entrada, una de agua fría (a temperatura ambiente) y otra de agua caliente (a unos 70°C) con unas condiciones nominales de caudal y presión de 10 l/min y 2 bar para el agua fría y 5 l/min y 1 bar para la caliente. Las temperaturas y caudales de entradas son medidos con termopares y placas-orificio respectivamente, disponiendo de válvulas motorizadas para la regulación de los caudales de entrada.
- **Depósito.** Es un tanque de 1 m de altura y 20 cm de diámetro interior, térmicamente aislado, con un volumen aproximado de 31 l. Puede trabajar presurizado (hasta

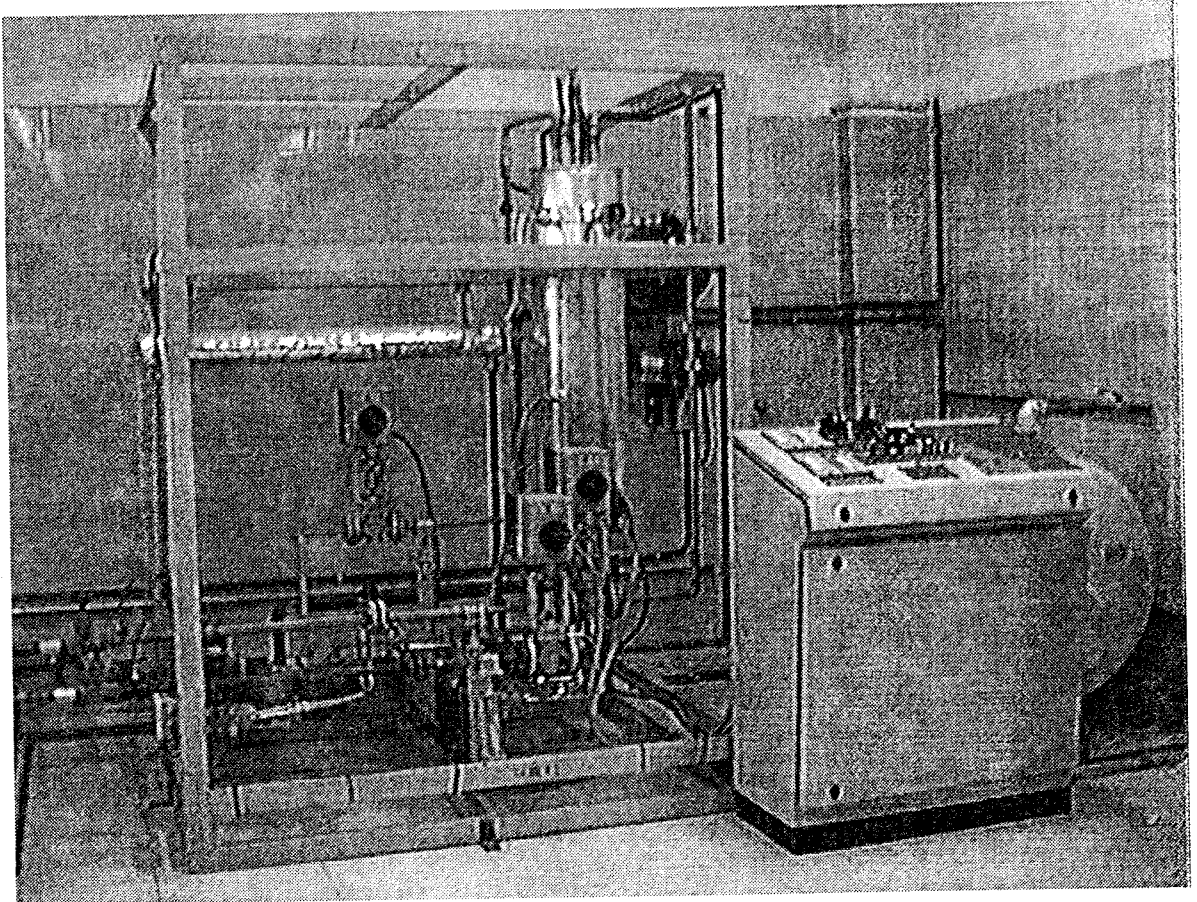


Figura 6.10: Planta Piloto

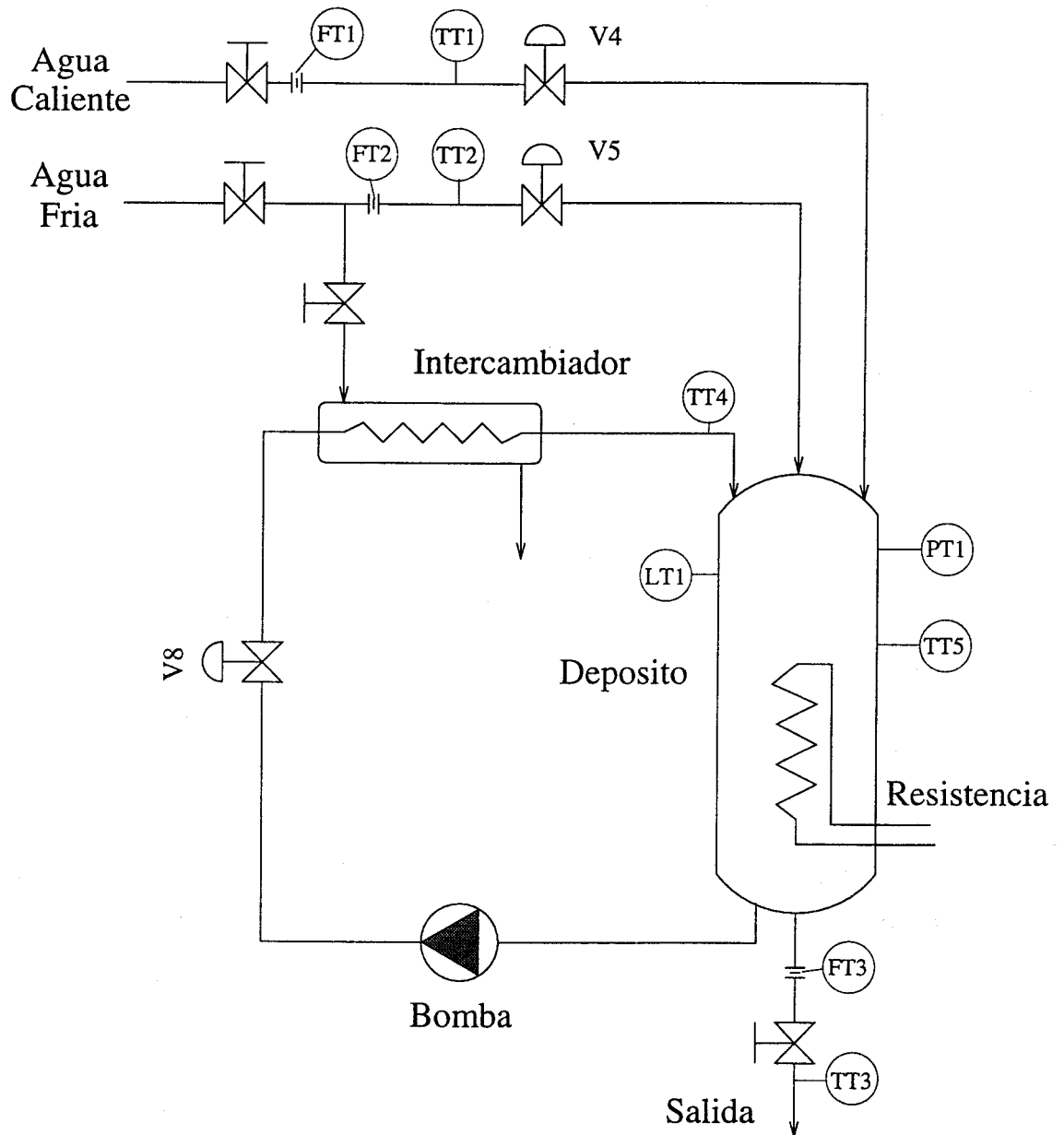


Figura 6.11: Esquema general de la planta

un límite de 4 bar) o a presión atmosférica, según la posición de la válvula de venteo. Dispone en su interior de una resistencia eléctrica de 15 kW para calentamiento, así como de un rebosadero, una tubería de salida y otra para la recirculación del agua a través del intercambiador.

- Circuito de recirculación. El agua caliente del depósito puede ser enfriada con la entrada de agua fría mediante el circuito de refrigeración. Este circuito consta de una bomba centrífuga que hace circular el agua caliente de la parte inferior del depósito a través de un intercambiador de calor de haz tubular, para retornar a menor temperatura a la parte superior del mismo.

6.4 Control de la planta

Para el control de la instalación se dispone de un sistema de control distribuido INTEGRAL 2000 de ORSI AUTOMAZIONE, formado por un controlador y un supervisor conectados por una red local de alta velocidad. El primero de ellos se encarga de ejecutar las rutinas de control digital y analógico mientras que el segundo actúa como plataforma de programación y comunicación con el operador. (Ver el apéndice para una descripción más detallada del sistema de control). En este sistema se van a implantar los algoritmos del GPC vistos anteriormente. Este sistema de control constituye un ejemplo típico de controlador industrial, recogiendo las características más usuales de sistemas de tamaño medio que pueden encontrarse hoy día en el mercado. Como en cualquier computador de control las facilidades de cálculo son limitadas y el tiempo disponible para la ejecución del algoritmo de control es escaso por la atención que precisan otras operaciones. Por tanto constituye una excelente plataforma para la implantación del GPC precalculado en entornos industriales.

De todos los posibles lazos que se pueden controlar, se van a mostrar los resultados obtenidos en algunas situaciones: control del caudal de agua fría FT_2 con la válvula V_5 , control de la temperatura de salida del intercambiador de calor TT_4 con la válvula V_8 , control del nivel en el depósito LT_1 con el caudal de agua fría mediante la válvula V_5 y control de la temperatura del depósito TT_5 con la resistencia eléctrica.

6.4.1 Control de caudal

Como ejemplo de regulación de un lazo sencillo se ha elegido el control del caudal de agua fría mediante la electroválvula V_5 . Al provenir todo el agua que se suministra a

la planta de un único grupo de presión, las variaciones que se impongan al caudal de agua caliente o al de agua fría del intercambiador de calor afectarán al caudal de agua fría como perturbaciones. La regulación del caudal de agua fría no sólo es importante como lazo de control en sí mismo, sino que también puede ser necesaria como variable auxiliar para controlar temperatura o nivel en el depósito.

La dinámica de este lazo viene marcada principalmente por la válvula de regulación. Se trata de una válvula motorizada con posicionador con un tiempo de apertura total de 110 segundos, lo que provocará que la evolución del caudal no sea muy rápida. El comportamiento del caudal se va a aproximar al de un sistema de primer orden más retardo.

Primeramente se obtendrán los parámetros que identifican el sistema haciendo uso de la curva de reacción, para luego obtener los coeficientes del GPC con el método descrito en el capítulo 3. Para ello, trabajando con la válvula abierta a un 70% (caudal de 3.98 l/min), se produce un escalón hasta el 80%, obteniendo tras el transitorio un valor estacionario de 6.33 l/min. Del registro de la evolución del caudal se puede calcular:

$$K = 0.25 \text{ (l/min)/ \%} \quad \tau = 10.5 \text{ segundos} \quad \tau_d = 10 \text{ segundos}$$

que en discreto proporciona la función de transferencia:

$$G(z) = \frac{0.043z^{-1}}{1 - 0.8265z^{-1}} z^{-5}$$

si se elige un tiempo de muestreo de $T = 2$ segundos. La señal de control es fácilmente calculada usando la expresión conocida:

$$u(t) = u(t-1) + (l_{y1}\hat{y}(t+5) + l_{y2}\hat{y}(t+4) + l_{r1}r(t))/K$$

siendo $u(t)$ la posición de V_5 e $y(t)$ el valor de FT_2 . Usando las fórmulas aproximadas (3.10) con $\lambda = 0.8$, las ganancias del controlador resultan:

$$\begin{aligned} l_{y1} &= -2.931 \\ l_{y2} &= 1.864 \\ l_{r1} &= 1.067 \end{aligned}$$

El comportamiento del control se muestra en la figura 6.12 cuando se producen cambios en el setpoint (nótese que aunque el tiempo de muestreo es de 2 segundos, los valores se representan cada 10 segundos por limitaciones del sistema de control). Obsérvese la aparición de perturbaciones en la medida del caudal; algunas de ellas son producidas por alteraciones en los otros caudales de la instalación y otras son

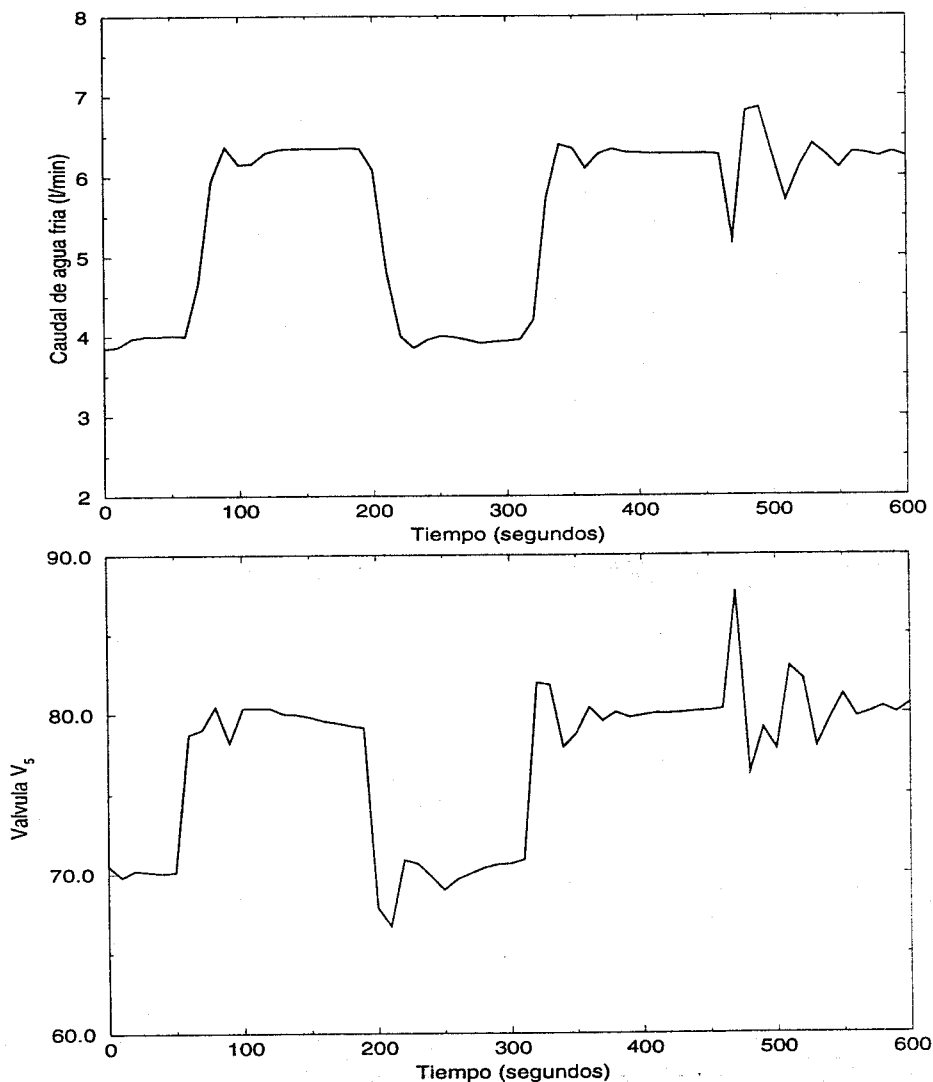


Figura 6.12: Control de caudal

ruidos eléctricos provocados por la conexión de otros aparatos cercanos a la instalación (principalmente robots). Se han producido cambios en el setpoint entre 4 y 6.3 litros por minuto. El valor medido del flujo sigue con rapidez a los cambios de la referencia (teniendo en cuenta la dinámica lenta de la válvula) con una sobreoscilación máxima del 13%. Para estudiar la respuesta del sistema ante perturbaciones externas se cierra parcialmente la válvula manual de agua fría. Se puede ver que el caudal disminuye bruscamente pero el controlador actúa con rapidez, rechazando la perturbación.

Para dejar patente la facilidad con la que se puede implantar un GPC en un sistema de control industrial usando las fórmulas de aproximación, se muestra en la figura 6.13 el código para el lenguaje de programación del INTEGRAL 2000, llamado ITER. El

cálculo de los parámetros del controlador se basa en los parámetros del proceso, los cuales pueden ser cambiados en línea por el operador o por un identificador si se emplea control adaptativo. Si se usara la formulación *standard* del controlador se necesitarían cálculos mucho más complejos y sería prácticamente imposible programar el GPC en este sistema de control.

6.4.2 Control de la temperatura a la salida del intercambiador

Dentro de la Planta, se puede considerar el intercambiador de calor como proceso independiente. El intercambiador hace disminuir la temperatura del agua de recirculación, impulsada por la bomba, usando para ello un caudal constante de agua fría. La forma de controlar la temperatura de salida es haciendo variar el caudal de agua de recirculación con la electroválvula V_8 ; se consigue pues la temperatura deseada a costa de variaciones en el caudal. En definitiva el intercambiador no es más que un haz de tubos con agua caliente en su interior que intercambia calor con el agua fría exterior. Se puede considerar pues formado por un gran número de elementos de primer orden que unidos se comportan como un sistema de primer orden con retardo puro (ver capítulo 3). Por ello se va a aproximar el sistema $TT_4 - V_5$ por una función de transferencia de ese tipo.

Siguiendo el procedimiento empleado hasta ahora, se calculan los parámetros del sistema, que se usarán para la ley de control. Algunos de los resultados obtenidos se muestran en la figura 6.14. Debe tenerse en cuenta que, por no ser el intercambiador independiente del resto de la planta, su temperatura de salida influye, a través del depósito, en la de entrada, produciendo cambios en las condiciones de trabajo. A pesar de ello el comportamiento es razonablemente bueno.

Se producen cambios en el setpoint de $38\text{ }^\circ\text{C}$ a $34\text{ }^\circ\text{C}$. Como se puede observar en la figura, la temperatura de salida del intercambiador evoluciona hacia el nuevo valor del setpoint suavemente sin presentar oscilaciones. Se introducen ahora dos tipos de perturbaciones externas. Primero se cierra la válvula manual de agua fría de refrigeración durante unos instantes. Como es de esperar la temperatura a la salida del intercambiador se incrementa rápidamente debido a esta fuerte perturbación y luego es llevada al valor deseado por el GPC. La segunda perturbación es causada disminuyendo el ciclo de trabajo de la resistencia calentadora, decrementando de esta forma la temperatura de entrada del agua caliente y cambiando por tanto el punto de operación. Como se puede ver, el GPC rechaza casi completamente esta perturbación, causada por el cambio en la dinámica.

```

;----- ALGORITMO DE CONTROL GPC -----
STARTPROG
;-----CALCULO DE LOS PARAMETROS DEL GPC (L) -----
DEFSUB CALCULA_L(VAR FLOAT LY1,LY2,LR1;FLOAT LAMBDA,A)
  LOCAL FLOAT K11,K21,K31,K12,K22,K32
  K11=-EXP(0.3598-0.9127*LAMBDA+0.3165*SQR(LAMBDA))
  K21=-EXP(0.0875-1.2309*LAMBDA+0.5086*SQR(LAMBDA))
  K31=1.05
  K12=EXP(-1.7383-0.40403*LAMBDA)
  K22=EXP(-0.32157-0.81926*LAMBDA+0.3109*SQR(LAMBDA))
  K32=1.045
  LY1=K11+K21*(A/(K31-A))
  LY2=K12+K22*(A/(K32-A))
  LR1=-LY1-LY2
  RETURN
ENDSUB
;----- CONTROLADOR -----
DEFSUB FLOAT CONTROL_GPC(FLOAT Y,Yt1;VAR FLOAT CFP,YP,U)
  ;Variables locales
  LOCAL INTEGER I, D
  LOCAL FLOAT Ugpc,G, LAMBDA,LY1,LY2,LR1,R,A,B
  ;Set Point
  R=CFP[10]
  ;Parametro LAMBDA
  LAMBDA=CFP[0]
  ;Span de la salida
  SPAN_GPC=CFP[13]
  ;Parametros de la Funcion de Transferencia
  A=CFP[6]
  B=CFP[7]
  D=FTOI(CFP[9])
  ;Ganancia del proceso
  G=B/(1.0-A)
  ;Predictor
  YP[0]=Y
  YP[1]=(1.0+A)*YP[0]-A*Yt1+B*(U[D]-U[D+1])
  FOR I=2 TO D
  YP[I]=(1.0+A)*YP[I-1]-A*YP[I-2]+B*(U[D-I+1]-U[D-I+2])
  ENDFOR
  ;Parametros del controlador
  CALCULA_L(LY1,LY2,LR1,LAMBDA,A)
  ;Accion de control
  IF (D=0) THEN
    Ugpc=U[1]+(LY1*YP[D]+LY2*Yt1+LR1*R)/G
  ELSE
    Ugpc=U[1]+(LY1*YP[D]+LY2*YP[D-1]+LR1*R)/G
  ENDIF
  RETURN(Ugpc)
ENDSUB
ENDPROG

```

Figura 6.13: Código del GPC en lenguaje ITER

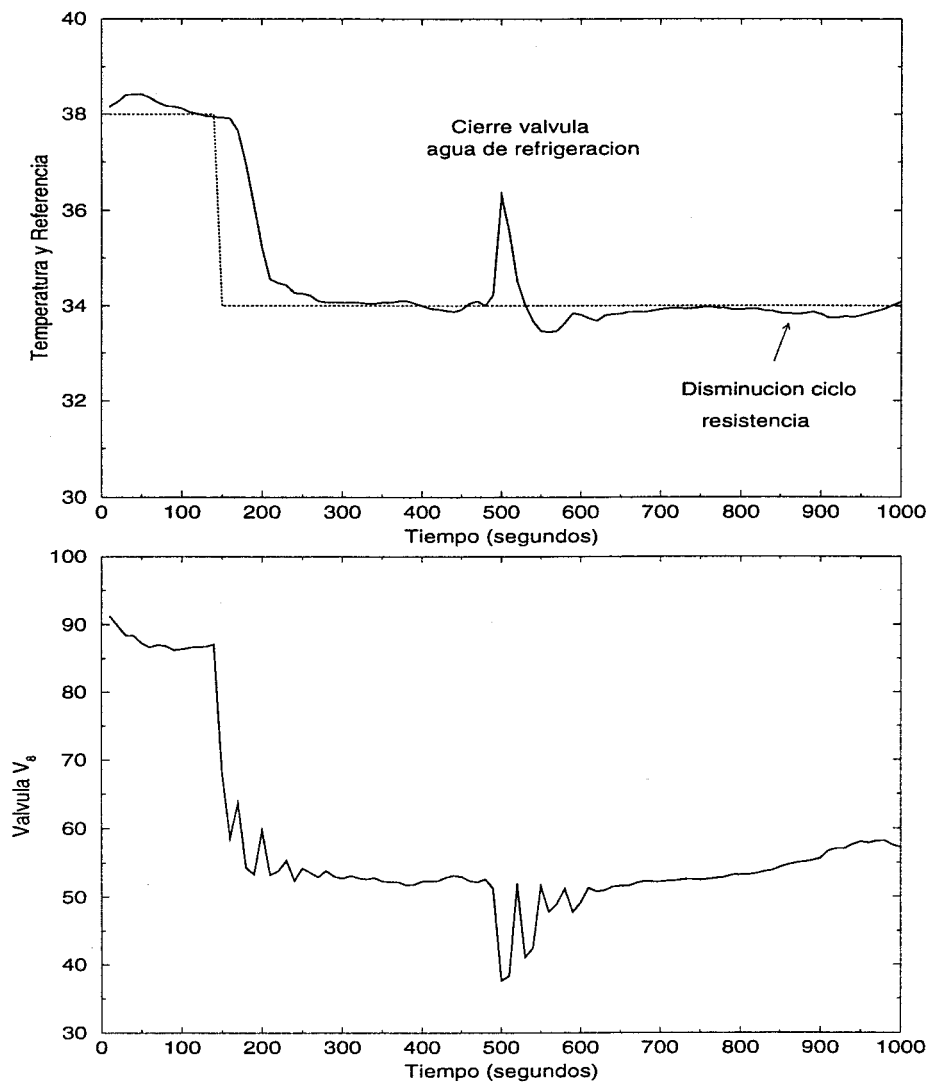


Figura 6.14: Comportamiento del intercambiador de calor

6.4.3 Control de la temperatura en el depósito

La siguiente aplicación también corresponde a un caso típico en la industria de procesos: el control de la temperatura de un líquido en el interior de un depósito. La variable de proceso es la temperatura interior TT_5 mientras que la variable manipulada es el ciclo de trabajo de la resistencia eléctrica (la resistencia sólo puede estar encendida o apagada, viniendo dada la acción de control por el tiempo dentro de cada ciclo de trabajo que se encuentra a ON).

Se observa que el proceso posee efecto integral, pues la temperatura crecería inde-

finidamente para un escalón en la resistencia, al estar el tanque aislado térmicamente. Por tanto, la función de transferencia que modela este lazo debe ser de la forma:

$$G(s) = \frac{K}{s(1 + \tau s)} e^{-\tau_d s}$$

Se identifica el sistema provocando escalones pseudoaleatorios en la entrada entre 20 y 70%, obteniendo los siguientes valores:

$$K = 0.41 \quad \tau = 50 \text{ segundos} \quad \tau_d = 50 \text{ segundos}$$

Si se elige un tiempo de muestreo de $T = 10$ segundos, usando las expresiones (3.22) la función de transferencia discreta es:

$$G(z) = \frac{0.3746z^{-1} + 0.35z^{-2}}{(1 - z^{-1})(1 - 0.8187z^{-1})} z^{-5}$$

Se han realizado diversas pruebas para mostrar el funcionamiento del bucle cerrado. Estas pruebas se han llevado a cabo con la bomba de recirculación funcionando de forma que el intercambiador de calor extrae una cantidad de calor prácticamente constante del interior del tanque. La figura 6.15 muestra los resultados obtenidos. En primer lugar se simula un fallo en el actuador, de forma que la resistencia se apaga durante un cierto tiempo, con lo que la temperatura empieza a decaer. Cuando el actuador vuelve a funcionar, el sistema de control es capaz de llevar la temperatura al valor deseado con un suave transitorio. También se muestra la evolución ante un cambio en el setpoint de 51°C a 60°C , observando que el sistema evoluciona bien.

6.4.4 Control del nivel del depósito

El nivel es otra de las variables que más aparecen en el control de procesos. En la Planta Piloto, el nivel del depósito se puede controlar por los caudales de agua de alimentación (fría o caliente). En este ejemplo se usa el de agua fría gobernado por la válvula V_5 . Identificando el sistema alrededor del punto de funcionamiento nominal (70% de apertura) se obtiene la función de transferencia:

$$G(s) = \frac{1.12}{1 + 87s} e^{-45s}$$

Con un tiempo de muestreo de $T = 10$ segundos el retardo es fraccionario, con lo que los coeficientes del controlador se calculan según se muestra en la sección 3.3.2.

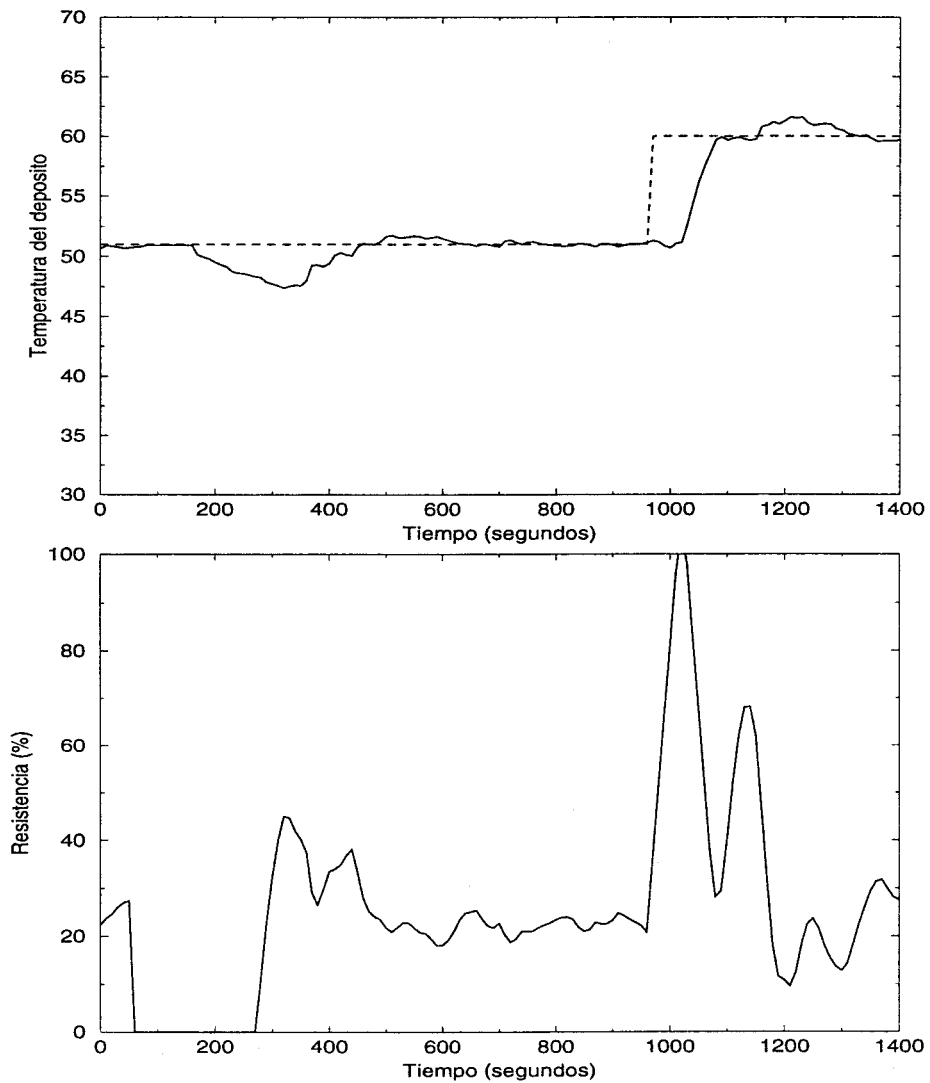


Figura 6.15: Evolución de la temperatura en el depósito

Los resultados obtenidos al trabajar con un peso del esfuerzo de control de $\lambda = 1$ y horizonte 15 se muestran en la figura 6.16. El setpoint se cambia de un 70 a un 75% y, como puede observarse, el nivel evoluciona suavemente entre ambas referencias. Para observar las cualidades de rechazo a las perturbaciones se provoca una perturbación consistente en una apertura repentina de la válvula de agua caliente, dando lugar a un aumento de nivel. Como puede observarse, la perturbación es rechazada tras un suave transitorio.

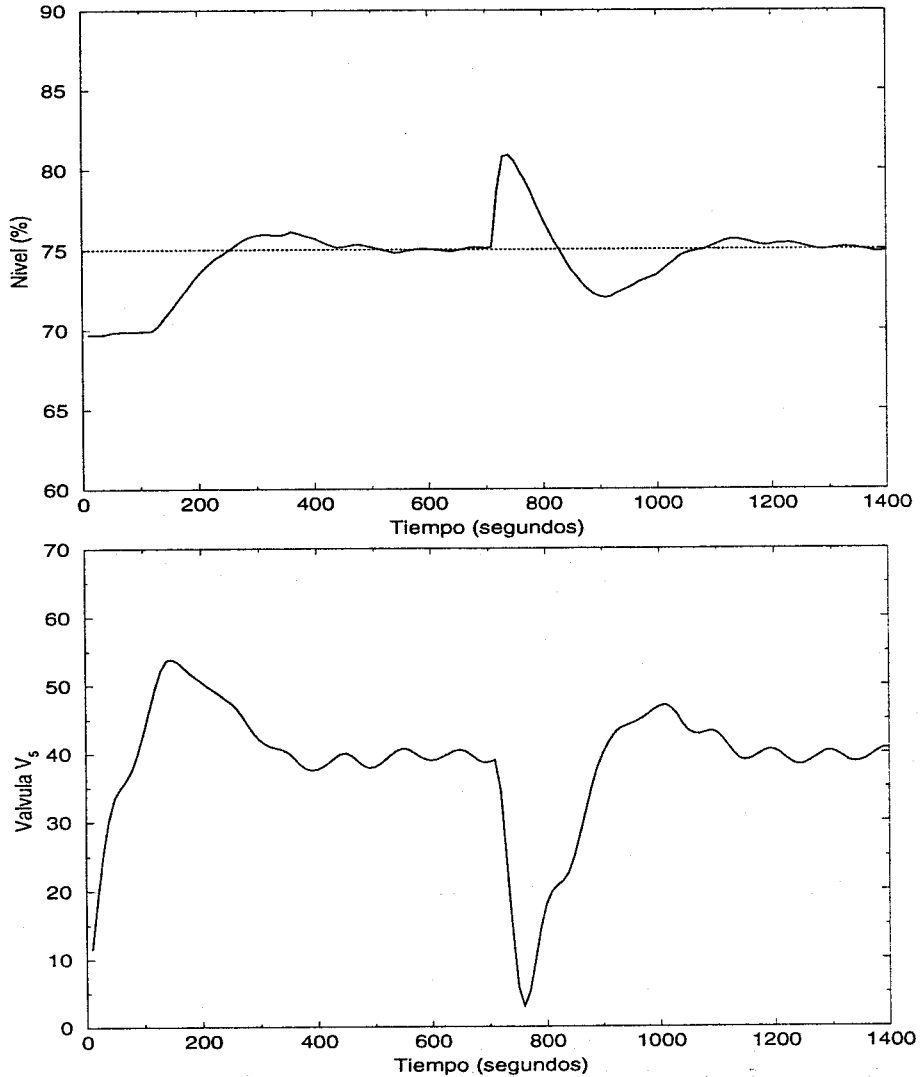


Figura 6.16: Evolución del nivel en el depósito

6.4.5 Comentarios

El principal objetivo de los ejemplos presentados en esta sección ha sido mostrar lo fácilmente que se puede implantar un Controlador Predictivo Generalizado en un sistema de control distribuido industrial usando la técnica presentada en el capítulo 3. Los GPCs fueron programados con facilidad usando el lenguaje ITER del sistema de control INTEGRAL 2000.

Aunque la comparación de los resultados por el GPC con los obtenidos mediante el uso de otra técnica de control no era uno de los objetivos, GPC ha producido mejores

resultados que los tradicionales PIDs en los casos que se han tratado, con una carga de cálculo comparable. En todos los lazos, los resultados obtenidos por PIDs ajustados por el método de Ziegler-Nichols resultaron pobres y sólo tras un largo período de ajuste manual se obtuvieron mejores resultados, aunque siempre inferiores al GPC (ver [88]). En cambio, el cálculo de los coeficientes del GPC se llevó a cabo casi al momento, sin la necesidad de un largo período de sintonización.

Capítulo 7

Conclusiones

Se ha propuesto una formulación del Control Predictivo Generalizado válida para la mayoría de procesos industriales que reduce considerablemente las necesidades de cálculo asociadas a este tipo de controladores. Esta formulación proporciona unas expresiones muy simples, análogas a las empleadas en la práctica industrial para la sintonización de PIDs, que proporcionan los coeficientes de una sencilla ley de control lineal en función de los parámetros del proceso. La simplicidad del controlador permite un ajuste de parámetros en tiempo real, bien por el uso de una estrategia adaptativa o bien por decisión del operador. Esta técnica puede ser implantada en cualquier sistema de control por pocas facilidades de cálculo que posea, pues la dificultad de los cálculos es análoga a la de un PID digital.

Se ha estudiado la robustez de la técnica propuesta como paso previo a la aplicación en procesos reales, donde las incertidumbres pueden afectar a la estabilidad del bucle cerrado. El estudio ha sido realizado desde dos puntos de vista: incertidumbres estructuradas y no estructuradas. Ambas aproximaciones muestran la robustez del controlador, asegurando amplios límites de robustez en situaciones bastante desfavorables. La comparación de ambos resultados indica que los obtenidos considerando las incertidumbres como no estructuradas son más conservadores, reflejando que un conocimiento más específico de las causas de las incertidumbres puede proporcionar unos límites de robustez más realistas.

Los resultados anteriores se han comprobado experimentalmente en diversos procesos reales (Planta Solar, intercambiador, depósito, motor) o simulados de forma realista (calentador, columna de destilación, evaporador), mostrando la facilidad con que puede implantarse así como las buenas prestaciones y la robustez. La programación del método en un sistema de control distribuido comercial, con escasas facilidades de

cálculo y ciertas limitaciones temporales, muestra cómo este método que posee todas las ventajas del Control Predictivo puede implantarse con facilidad eliminando el inconveniente de la complejidad.

Como futuras líneas de investigación se proponen:

- Control predictivo multivariable. Aunque teóricamente es un tema resuelto, su aplicación práctica conlleva muchos problemas. La obtención de algoritmos con pocos requerimientos de cálculo y que resuelvan ciertos problemas prácticos como los retardos parece ser un campo abierto.
- Control con restricciones. Es un campo fundamental en la práctica industrial. Escasean resultados de estabilidad y robustez, así como nuevos métodos de optimización (redes neuronales, por ejemplo) que permitan una solución rápida y simple del problema.

Apéndice A

Sistema de control distribuido Integral 2000

A.1 Generalidades

El control y supervisión de la Planta Piloto descrita en el capítulo 6 se ha llevado a cabo mediante el sistema de control distribuido INTEGRAL 2000, fabricado por la empresa italiana ORSI AUTOMAZIONE [66]. Es un sistema abierto que permite llevar a cabo el control y supervisión de la planta así como una gestión integrada. Está basado en la interconexión de diversos módulos (controladores locales, autómatas programables, computadores de control, supervisión o gestión, etc) a través de una red local de alta velocidad ITERNET determinística con protocolo "Token passing" modificado y velocidad de 2,5 Mbaudios. Esta modularidad permite la interconexión de los distintos niveles de la estructura jerárquica de control: desde el control local de un simple lazo con un PID hasta una supervisión mediante un sistema experto como el G2 [34].

En el caso que nos ocupa, el sistema de control está formado únicamente por un controlador, llamado PMC (Programmable Multifunctional Controller) y un supervisor. El primero es el encargado de resolver de una manera coordinada todos los problemas integrados en la regulación y manipulación de la instalación, adquisición y regulación de las variables analógicas, secuenciamiento de las tareas de control, etc. Está constituido básicamente por un microprocesador y una serie de tarjetas analógicas y digitales de entrada/salida, pudiendo ser todos sus elementos redundantes para aumentar la fiabilidad. El supervisor es un computador personal con dos monitores desde el cual

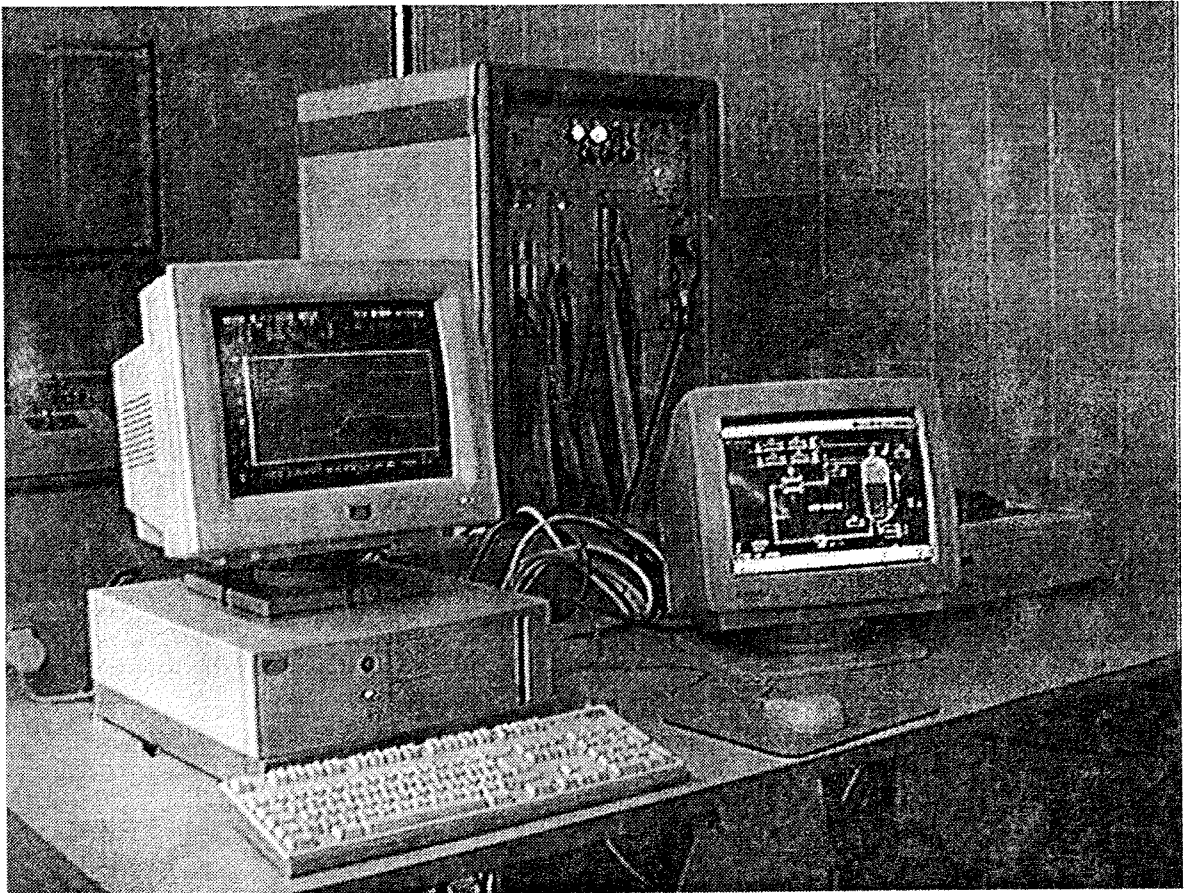


Figura A.1: Sistema de control INTEGRAL 2000

se lleva a cabo la programación del PMC y la comunicación con el operario. Permite la visualización de sinópticos de los diversos procesos que se estén controlando, el cambio de consignas y parámetros y es el encargado de llevar a cabo la gestión de las alarmas así como de los archivos históricos. Desde el supervisor se pueden detener determinadas secuencias de control e incluso se pueden modificar éstas sin interrumpir el control normal de la instalación. En la figura A.1 se ven ambos módulos, conectados mediante la red local con cable coaxial.

A.2 Descripción del PMC

A.2.1 Funcionamiento

El PMC de la serie 10 posee una CPU Intel 8086 con coprocesador matemático 8087. Funciona a una frecuencia de 8 MHz y está dotado de un sistema operativo residente de tipo multitarea (actualmente está en el mercado una versión más potente basada en el microprocesador 30386 y a mayor frecuencia de trabajo). Cada tarea o secuencia puede ser clasificada de manera independiente o coordinada por otras secuencias, capaces de gestionar las adecuadas funciones de control; además cada función puede ser escrita en el lenguaje de programación más adecuado de los tres que permite el PMC (que se verán más adelante). El sistema operativo permite la ejecución "simultánea" de hasta 256 secuencias por tiempo compartido, tomando como base de temporización una unidad de tiempo denominada *tick* que puede estar comprendida entre 55 y 200 milisegundos, usándose típicamente el valor de 100 ms, que se puede considerar más que suficiente para el tipo de procesos que aparece en la Planta Piloto.

Las secuencias se programan para que se ejecuten cada cierto número de *ticks* o bien pueden entrar en funcionamiento por una determinada situación, como en el caso de tratamiento de alarmas. Cada vez que tiene lugar el *tick* se realizan las siguientes operaciones:

1. Comunicación con el exterior a través de la red local. En nuestro caso los únicos módulos son el PMC y el supervisor y por tanto en esta fase lo que se hace es transmitir los valores actuales de las variables de la memoria del PMC al supervisor (para representación, alarmas e históricos) mientras que éste manda los valores (parámetros, consignas ...) que deban ser cambiados.
2. Actualización de todas las entradas y salidas, tanto analógicas como digitales.
3. Ejecución de todas las secuencias que estén activas en ese momento, con el orden que se haya establecido.
4. Actualización de las salidas analógicas y digitales.
5. Espera hasta el siguiente *tick*.

Cada secuencia está definida por un número de orden y por un número de temporización expresado en *ticks*. Además cada secuencia puede estar activa o inactiva, pudiendo una secuencia activar o desactivar a otra o desactivarse a sí misma. Existen cuatro secuencias especiales que siempre deben estar presentes en una aplicación:

- Secuencia 0: se ejecuta cuando se dispone de un PMC redundante y se produce la conmutación de memoria entre el "Master" y el "Slave".
- Secuencia 1: se ejecuta al iniciar una aplicación, cada vez que se pasa el sistema de STOP a RUN. En ella se coloca toda la fase de iniciación de variables y todo aquello que interese ejecutar una sola vez al arrancar la aplicación.
- Secuencia 2: se ejecuta cuando se restablece la tensión tras una caída estando el sistema en RUN.
- Secuencia 3: se ejecuta siempre que el sistema sufre una pérdida de tensión estando en RUN. Dispone de una batería tampón para realizar una parada ordenada.

A.2.2 Configuración de la memoria

El PMC dispone de una memoria global de 64 K, que es compartida por todas las secuencias. Los registros de la memoria pueden ser tratados como tipo carácter, entero o flotante. También se puede acceder a un determinado bit dentro de un registro así como formar vectores de registros¹. Las variables pueden a su vez ser permanentes (no pierden su valor en ausencia de tensión) o volátiles (sí lo pierden). El reparto de la memoria entre los distintos tipos de variables es modificable según la aplicación.

Existen dos métodos de trabajo: real y simbólico. En el primero de ellos toda la programación debe realizarse con el nombre de los registros, lo cual conlleva más trabajo en la depuración y puesta a punto de los programas. Por ello normalmente se usará el modo simbólico, que facilita el desarrollo de la aplicación. Para realizar esto es necesario construir una base de datos en la que se especifique la relación de cada nombre simbólico con el registro real.

A.2.3 Programación de las secuencias

Existen tres lenguajes de programación que se pueden usar indistintamente para crear las diferentes secuencias de que conste la aplicación, estando cada uno de ellos orientado a resolver problemas específicos. Cada secuencia puede estar escrita en un lenguaje y utilizar rutinas escritas en otro distinto. Estos lenguajes son:

¹Sin embargo no se permite la agrupación en matrices, por lo que un algoritmo que requiera cálculo matricial no puede ser programado directamente.

- lenguaje LADDER: es el típico lenguaje de contactos, usado en la mayoría de los automatismos. Es sencillo y muy intuitivo, a la vez que dispone de un editor gráfico con funciones predefinidas (contadores, temporizadores, biestables, puertas lógicas, histéresis...). Este lenguaje está especialmente indicado para la programación de la parte lógica de la planta, siendo fácil de aprender por personal no cualificado a la vez que resulta rápido en ejecución.
- lenguaje de Bloques: es un lenguaje semigráfico de alto nivel preparado para trabajar en entorno Windows. Un programa consiste en una serie de bloques conectados entre sí, de forma análoga a los diagramas de bloques usados en control. Los bloques pueden ser funciones de librería (controladores PID, filtros, entradas/salidas...) o definidas por el usuario.
- lenguaje ITER: es un lenguaje de alto nivel de tipo procedural, capaz de realizar los cálculos matemáticos más usuales y con estructuras de control como bucles o sentencias condicionales. De gran utilidad para personal con conocimientos de programación en otro lenguaje. Posee ciertas restricciones respecto a otros lenguajes de alto nivel, ya que, por ejemplo, no permite las operaciones con matrices.

Aparte de estos tres lenguajes propios del sistema de control, se puede añadir a la aplicación cualquier rutina escrita en otro lenguaje de alto nivel, como *C* o FORTRAN (siempre que hayan sido escritas teniendo en cuenta las limitaciones del PMC). Al final todas las secuencias se compilan y se montan juntas, dando lugar a un único ejecutable que es el que se envía al PMC. En cualquier caso, es imprescindible especificar para cada secuencia su número (para establecer la prioridad) y su temporización expresada en *ticks*. Como ya se indicó, la programación se lleva a cabo en el supervisor, y luego el programa ejecutable se envía por la red al PMC, donde se puede sustituir por otro sin necesidad de detener el sistema.

Las prestaciones del PMC en cuanto a velocidad y potencia de cálculo no son muy elevadas, quedando muy lejos de las que pueden ofrecer hoy día los computadores personales más potentes. Este hecho puede considerarse normal en muchos sistemas de control y corrobora en cierta forma la necesidad de utilizar algoritmos simplificados de los controladores para poder ser implantados en muchos sistemas comerciales empleados en la práctica.

A.3 El supervisor

El supervisor es un computador personal con microprocesador 80386 conectado al PMC por medio de la red. Puede llevar la supervisión de hasta 16 PMCs, encargándose de la programación previa y de la comunicación con el operario durante el funcionamiento normal de la aplicación. Posee un monitor principal para la programación de las secuencias, supervisión de alarmas, control de históricos y comunicaciones y un monitor gráfico para la presentación de las pantallas gráficas interactivas.

El supervisor posee las características habituales de otros sistemas de control comerciales, con pantallas gráficas fáciles de crear y de comprender (ver por ejemplo [88]), cuyos valores se refrescan continuamente en tiempo real con los registros extraídos de la memoria de PMC a través de la red. Las órdenes del operador, como cambios en los parámetros de los controladores o de las consignas de los lazos son inmediatamente transmitidas al PMC. Las principales funciones de las cuales se encarga el supervisor son:

- Gestión de las pantallas gráficas interactivas. Incluye el tratamiento de las acciones del operador, representación de sinópticos y representación numérica o gráfica de la evolución de las diversas variables.
- Tratamiento de alarmas. Avisa al operador de la aparición de alarmas a la vez que lleva un registro de los instantes de activación, reconocimiento y desaparición.
- Archivos históricos. Permite la generación *on-line* de archivos históricos tanto en disco duro como diskette así como su visualización en tiempo real.
- Detección de fallos del hardware. Chequea continuamente el estado de las tarjetas de entrada/salida, detectando los posibles errores que puedan aparecer.
- Depuración de programas. Permite la depuración de los programas mientras éstos se están ejecutando, monitorizando el valor que toman los registros de la memoria del PMC. También se pueden programar otras aplicaciones mientras el PMC sigue controlando el sistema.

Bibliografía

- [1] P. Albertos y R. Ortega. On Generalized Predictive Control: Two Alternative Formulations. *Automatica*, 25(5):753–755, 1989.
- [2] J.C. Allwright. *Advances in Model-Based Predictive Control*, capítulo On min-max Model-Based Predictive Control. Oxford University Press, 1994.
- [3] N. Amann y F. Allgöwer. Design of Robustly Performing Controllers for a Class of Practical Control Problems. En *Proceedings of the 2nd European Control Conference, Delft, Holanda.*, páginas 721–726, 1993.
- [4] K.J. Aström y B. Wittenmark. *Computer Controlled Systems. Theory and Design*. Prentice-Hall. Englewood Cliffs, NJ, 1984.
- [5] S. Basu. New Results in Multidimensional Robust Stability. En *IEEE International Symposium on Circuits and Systems*, 1990.
- [6] R.R. Bitmead, M. Gevers, y V. Wertz. *Adaptive Optimal Control. The Thinking Man's GPC*. Prentice-Hall, 1990.
- [7] C. Bordons y E.F. Camacho. Realización de un GPC para Procesos Industriales. En *Jornadas de Control Inteligente. Sevilla*, 1991.
- [8] E.H. Bristol. On a new measure of interaction for multivariable process control. *IEEE Trans. on Automatic Control*, 11(1):133–4, 1966.
- [9] E.F. Camacho. Constrained Generalized Predictive Control. *IEEE Trans. on Automatic Control*, 38(2):327–332, 1993.
- [10] E.F. Camacho y M. Berenguel. *Advances in Model-Based Predictive Control*, capítulo Application of Generalized Predictive Control to a Solar Power Plant. Oxford University Press, 1994.
- [11] E.F. Camacho, M. Berenguel, y C. Bordons. Adaptive Generalized Predictive Control of a Distributed Collector Field. *IEEE Trans. on Control Systems Technology*, diciembre 1994.

- [12] E.F. Camacho, M. Berenguel, y F.R. Rubio. Application of a Gain Scheduling Generalized Predictive Controller to a Solar Power Plant. *Control Engineering Practice*, 2(2):227–238, 1994.
- [13] E.F. Camacho y C. Bordons. Simple Implementation of Generalized Predictive Self Tuning Controllers for Industrial Processes. En *Proceedings 1st European Control Conference, Grenoble*, páginas 344–349, julio 1991.
- [14] E.F. Camacho y C. Bordons. Implementation of Self Tuning Generalized Predictive Controllers for the Process Industry. *Int. Journal of Adaptive Control and Signal Processing*, 7:63–73, 1993.
- [15] E.F. Camacho y C. Bordons. *Model Predictive Control in the Process Industry*. Springer-Verlag (en prensa), 1995.
- [16] E.F. Camacho, F.R. Rubio, y F.M. Hughes. Self-tuning Control of a Solar Power Plant with a Distributed Collector Field. *IEEE Control Systems Magazine*, páginas 72–78, 1992.
- [17] P.J. Campo y M. Morari. Robust Model Predictive Control. En *American Control Conference, Minneapolis, Minnesota*, 1987.
- [18] D. W. Clarke. Application of Generalized Predictive Control to Industrial Processes. *IEEE Control Systems Magazine*, 122:49–55, 1988.
- [19] D. W. Clarke y P. J. Gawthrop. Self-tuning Controller. *Proceedings IEE*, 122:929–934, 1975.
- [20] D. W. Clarke y P. J. Gawthrop. Self-tuning Control. *Proceedings IEEE*, 123:633–640, 1979.
- [21] D. W. Clarke y R. Scattolini. Constrained Receding-horizon Predictive Control. *Proceedings IEE*, 138(4):347–354, July 1991.
- [22] D.W. Clarke y C. Mohtadi. Properties of Generalized Predictive Control. *Automatica*, 25(6):859–875, 1989.
- [23] D.W. Clarke, C. Mohtadi, y P.S. Tuffs. Generalized Predictive Control - Part I. The Basic Algorithm. *Automatica*, 23(2):137–148, 1987.
- [24] D.W. Clarke, C. Mohtadi, y P.S. Tuffs. Generalized Predictive Control - Part II. Extensions and Interpretations. *Automatica*, 23(2):149–160, 1987.
- [25] A. Creus. *Instrumentación Industrial*. Marcombo, 1989.
- [26] C.R. Cutler y B.C. Ramaker. Dynamic Matrix Control- A Computer Control Algorithm. En *Automatic Control Conference, San Francisco*, 1980.

- [27] A.P. de Madrid, M. Santos, S. Dormido, y F. Morilla. *Advances in Model-Based Predictive Control*, capítulo Constrained Generalized Predictive Control with Dynamic Programming. Oxford University Press, 1994.
- [28] P.B. Deshpande y R.H. Ash. *Elements of Computer Process Control*. ISA, 1981.
- [29] J. C. Doyle. Analysis of Control Systems with Structured Uncertainty. *Proceedings IEE, Part D*, 129:242–250, 1982.
- [30] J. C. Doyle y G. Stein. Multivariable Feedback Design: Concepts for a Classical/Modern Synthesis. *IEEE Trans. on Automatic Control*, 36(1):4–16, 1981.
- [31] G. Ferretti, C. Manfezzoni, y R. Scattolini. Recursive Estimation of Time Delay in Sampled Systems. *Automatica*, 27(4):653–661, 1991.
- [32] Y.K. Foo y Y.C. Soh. Robust Stability Bounds for Systems with Structured and Unstructured Perturbations. *IEEE Trans. on Automatic Control*, 38(7), 1993.
- [33] C.E. García, D.M. Prett, y M. Morari. Model Predictive Control: Theory and Practice—a Survey. *Automatica*, 25(3):335–348, 1989.
- [34] Genysim. *Manuales Sistema Experto G2*, 1992.
- [35] G. Goodwin y K. Sin. *Adaptive Filtering, Prediction and Control*. Prentice-Hall, 1984.
- [36] F. Gordillo. *Contribuciones al problema del control óptimo*. Tesis doctoral, Universidad de Sevilla, 1994.
- [37] F. Gordillo y F.R. Rubio. Self-tuning Controller with LQG/LTR Structure. En *Proceedings 1st European Control Conference, Grenoble*, páginas 2159–2163, julio 1991.
- [38] C. Greco, G. Menga, E. Mosca, y G. Zappa. Performance Improvement of Self Tuning Controllers by Multistep Horizons: the MUSMAR Approach. *Automatica*, 20:681–700, 1984.
- [39] R. Isermann. *Digital Control Systems*. Springer-Verlag, 1981.
- [40] E.I. Jury. Robustness of Discrete Systems: a Review. En *11th IFAC World Congress, Vol 5*, páginas 184–189, Tallin, URSS, 1990.
- [41] A. Kalt. Distributed Collector System Plant Construction Report. *IEA/SSPS Operating agent DFVLR, Colonia*, 1982.
- [42] R.M.C. De Keyser y A.R. Van Cuawenberghe. Extended Prediction Self-adaptive Control. En *IFAC Symp. on Identification and System Parameter Estimation, York, Reino Unido*, páginas 1317–1322, 1985.

- [43] R.M.C. De Keyser. Basic Principles of Model Based Predictive Control. En 1st *European Control Conference, Grenoble*, páginas 1753–1758, julio 1991.
- [44] R.M.C. De Keyser, Ph.G.A. Van de Velde, y F.G.A. Dumortier. A Comparative Study of Self-adaptive Long-range Predictive Control Methods. *Automatica*, 24(2):149–163, 1988.
- [45] M. Khammash. Necessary and Sufficient Conditions for the Robustness of Time-Varying Systems with Application to Sampled-Data Systems. *IEEE Trans. on Automatic Control*, 38(1):49–57, 1993.
- [46] B. Kouvaritakis, J.A. Rossiter, y A.O.T Chang. Stable Generalized Predictive Control: An Algorithm with Guaranteed Stability. *Proceedings IEE, Part D*, 139(4):349–362, 1992.
- [47] K. Krämer y H. Ubehauen. Predictive Adaptive Control. Comparison of Main Algorithms. En *Proceedings 1st European Control Conference, Grenoble*, páginas 327–332, julio 1991.
- [48] A.G. Kutnetsov y D.W. Clarke. *Advances in Model-Based Predictive Control*, capítulo Application of constrained GPC for improving performance of controlled plants. Oxford University Press, 1994.
- [49] H. Kwakernaak. Robust Control and \mathcal{H}_∞ -Optimization. Tutorial Paper. *Automatica*, 29(2):255–273, 1993.
- [50] W.H. Kwon y A.E. Pearson. On Feedback Stabilization of Time-Varying Discrete Linear Systems. *IEEE Trans. on Automatic Control*, 23:479–481, 1979.
- [51] J.H. Lee, M. Morari, y C.E. García. State-space Interpretation of Model Predictive Control. *Automatica*, 30(4):707–717, 1994.
- [52] M.A. Lelic y P.E. Wellstead. Generalized Pole Placement Self Tuning Controller. Part 1. Basic Algorithm. *Int. J. Control*, 46(2):547–568, 1987.
- [53] M.A. Lelic y M.B. Zarrop. Generalized Pole Placement Self Tuning Controller. Part 2. Application to Robot Manipulator Control. *Int. J. Control*, 46(2):569–601, 1987.
- [54] J.M. Lemos y E. Mosca. A Multipredictor-Based LQ Self-tuning Controller. En *IFAC Symp. on Identification and System Parameter Estimation, York, Reino Unido*, páginas 137–141, 1985.
- [55] J.L. Lin, I. Postlethwaite, y D.W. Gu. μ -K-iteration: A new algorithmy for μ -synthesis. *Automatica*, 29(1):219–224, 1993.
- [56] L. Ljung. *System Identification. Theory for the user*. Prentice-Hall, 1987.

- [57] A. G. MacFarlane y B. Kouvaritakis. A Design Technique for Linear Multivariable Feedback Systems. *Int. J. Control*, 25:837–879, 1977.
- [58] C. Mohtadi. *Advanced Self-tuning Algorithms*. Tesis doctoral, Oxford University, U.K., 1986.
- [59] M. Morari. *Advances in Model-Based Predictive Control*, capítulo Model Predictive Control: Multivariable Control Technique of Choice in the 1990s? Oxford University Press, 1994.
- [60] M. Morari y E. Zafriou. *Robust Process Control*. Prentice-Hall, 1989.
- [61] E. Mosca, J.M. Lemos, y J. Zhang. Stabilizing I/O Receding Horizon Control. En *IEEE Conference on Decision and Control*, 1990.
- [62] E. Mosca y J. Zhang. Stable Redesign of Predictive Control. *Automatica*, 28:1229–1233, 1992.
- [63] R.B. Newell y P.L. Lee. *Applied Process Control. A Case Study*. Prentice-Hall, 1989.
- [64] G. De Nicolao y R. Scattolini. *Advances in Model-Based Predictive Control*, capítulo Stability and Output Terminal Constraints in Predictive Control. Oxford University Press, 1994.
- [65] K. Ogata. *Modern Control Engineering*. Prentice-Hall, 1989.
- [66] Orsi. *Manuales Sistema INTEGRAL 2000*, 1991.
- [67] G.C. Papavasiliou y J.C. Allwright. A Descendent Algorithm for a Min-Max Problem in Model Predictive Control. En *Proceedings of the 30th Conference on Decision and Control*, Brighton, Inglaterra, 1991.
- [68] V. Peterka. Predictor-based Self-tuning Control. *Automatica*, 20(1):39–50, 1984.
- [69] I. Postlethwaite y S. Skogestad. *Essays on Control: Perspectives in the Theory and its Applications*, capítulo Robust Multivariable Control Using \mathcal{H}_∞ Methods: Analysis, Design and Industrial Applications. Birkhäuser, 1993.
- [70] C. Prada, J. Serrano, P. Vega, y M.A. Vega. *Advances in Model-Based Predictive Control*, capítulo A comparative study of DMC and GPC controllers. Oxford University Press, 1994.
- [71] A.I. Propoi. Use of LP Methods for Synthesizing Sampled-Data Automatic Systems. *Automn Remote Control*, 24, 1963.

- [72] J.M. Quero y E.F. Camacho. Neural Generalized Predictive Controllers. En *Proc. of the IEEE International Conference on System Engineering, Pittsburg*, 1990.
- [73] J. Rawlings y K. Muske. The Stability of Constrained Receding Horizon Control. *IEEE Trans. on Automatic Control*, 38:1512–1516, 1993.
- [74] J. Richalet. Industrial Applications of Model Based Predictive Control. *Automatica*, 29(5):1251–1274, 1993.
- [75] J. Richalet, S. Abu el Ata-Doss, C. Arber, H.B. Kuntze, A. Jacubash, y W. Schill. Predictive Functional Control. Application to Fast and Accurate Robots. En *Proc. 10th IFAC Congress, Munich*, 1987.
- [76] J. Richalet, A. Rault, J.L. Testud, y J. Papon. Algorithmic Control of Industrial Processes. En *4th IFAC Symposium on Identification and System Parameter Estimation. Tbilisi URSS*, 1976.
- [77] J. Richalet, A. Rault, J.L. Testud, y J. Papon. Model Predictive Heuristic Control: Application to Industrial Processes. *Automatica*, 14(2):413–428, 1978.
- [78] J.A. Rossiter y B. Kouvaritakis. Constrained Stable Generalized Predictive Control. *Proceedings IEE, Part D*, 140(4), 1993.
- [79] J.A. Rossiter y B. Kouvaritakis. *Advances in Model-Based Predictive Control*, capítulo Advances in Generalized and Constrained Predictive Control. Oxford University Press, 1994.
- [80] R. Rouhani y R.K. Mehra. Model Algorithmic Control; Basic Theoretical Properties. *Automatica*, 18(4):401–414, 1982.
- [81] F.R. Rubio. *Control Adaptativo de Procesos Industriales. Aplicación a una Central Solar*. Tesis doctoral, Universidad de Sevilla, 1985.
- [82] S.L. Shah, C. Mohtadi, y D.W. Clarke. Multivariable Adaptive Control without a Prior Knowledge of the Delay Matrix. *Systems and Control Letters*, 9:295–306, 1987.
- [83] N. Sivashankar y P.P. Khargonekar. Robust Stability and Performance Analysis of Sampled-Data Systems. *IEEE Trans. on Automatic Control*, 38(1):58–69, 1993.
- [84] I. Skrjanc y D. Matko. *Advances in Model-Based Predictive Control*, capítulo Fuzzy Predictive Controller with Adaptive Gain. Oxford University Press, 1994.
- [85] R. Söeterboek. *Predictive Control. A unified approach*. Prentice-Hall, 1992.
- [86] E. Srinivasa y M. Chidambaram. Robust Control of a distillation column by the method of inequalities. *Journal of Process Control*, 1(3):171–176, 1993.

- [87] Y. Tan y R. De Keyser. *Advances in Model-Based Predictive Control*, capítulo Neural Network Based Predictive Control. Oxford University Press, 1994.
- [88] E.R. Velasco. *Control y Supervisión de Planta Piloto mediante Sistema de Control Distribuido*. P.F.C. Univesidad de Sevilla, 1994.
- [89] R.A.J. De Vries y H.B. Verbruggen. *Advances in Model-Based Predictive Control*, capítulo Multivariable Unified Predictive Control. Oxford University Press, 1994.
- [90] S. J. Williams. Introduction to μ . En *Robust Control System Design using H_∞ and Related Methods*. Cambridge, Reino Unido, 1991.
- [91] B.E. Ydstie. Extended Horizon Adaptive Control. En *Proc. 9th IFAC World Congress, Budapest, Hungría*, 1984.
- [92] T.W. Yoon y D.W. Clarke. Prefiltering in Receding-Horizon Predictive Control. *Internal Report 1995/93, University of Oxford, Department of Engineering Science*, 1993.
- [93] L.A. Zadeh y B.H. Whalen. On Optimal Control and Linear Programming. *IRE Trans. Aut. Control*, 7(4), 1962.
- [94] G. Zames. Feedback and Optimal Sensitivity: Model Reference Transformations, Weighted Seminorms, and Approximate Inverses. En *Proc. 17th Allerton Conf.*, páginas 744–752, 1979.
- [95] A. Zheng y M. Morari. Stability of Model Predictive Control with Soft Constraints. *Internal Report. California Institute of Technology*, 1994.
- [96] J.G. Ziegler y N.B. Nichols. Optimum Settings for Automatic Controllers. *Trans. ASME*, 64:759–768, 1942.