Los sistemas de primer orden y los Controladores PID

Daniel Chuk
2012

Índice

Los sistemas de primer orden y los Controladores PID

En un lenguaje coloquial, podríamos decir que un controlador es un dispositivo que toma la señal error e, la procesa y genera una acción de control u. En la Figura 1, un proceso $G(s)$ es controlado por un controlador $G_c(s)$ en el marco de un lazo de control.

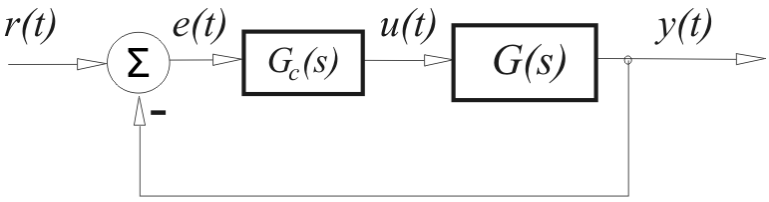


Fig. 1: Representación general de un lazo de control

Los controladores PID (Proporcional – Integral – Derivativo) son, por mucho, los más usados en la industria a pesar de existir algoritmos más eficientes para lograr los objetivos de control (Astrom y Hagglund, 1995) . El motivo es su eficiencia en el logro de los objetivos de control, facilidad de calibración y uso, y robustez.

Antes de avanzar en el estudio de los controladores propiamente, es necesario realizar algunas consideraciones sobre el tipo de modelo de proceso que se toma en cuenta para ello, que es mayormente un proceso de primer orden con retardo puro.

1. Descripción y propiedades de los sistemas de primer orden con retardo puro en el contexto del análisis de controladores PID

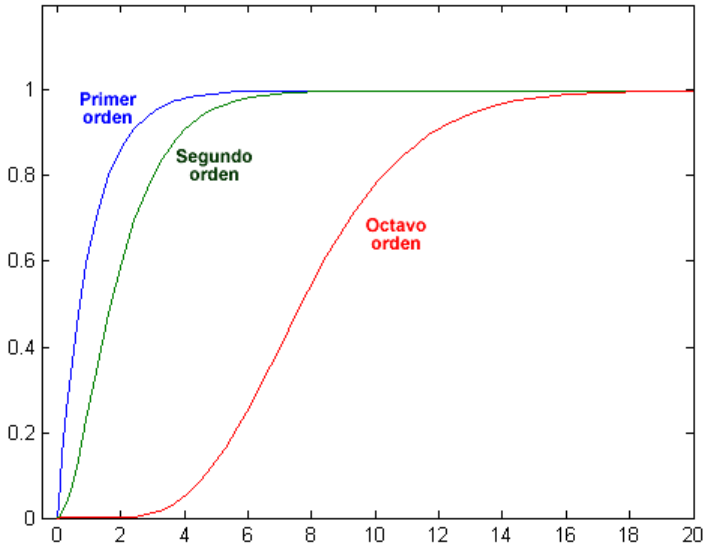


Fig. 2: Aumento del retardo puro aparente con el incremento del orden del sistema

Debe notarse que la gran mayoría de los procesos a controlar, sobre todo en la industria, tienen un alto orden. La Figura 2 muestra la respuesta al escalón de tres sistemas de orden uno, dos y ocho, que no poseen retardos puros. Puede comprobarse que, a medida que el orden del sistema aumenta, aumenta la inercia del sistema y en consecuencia, se incrementa un tiempo en el cual, frente a una excitación, aparentemente no hay modificación alguna. Este tiempo se denomina "retardo puro aparente", el cual sumado a los retardos puros reales del proceso (por ejemplo el generado por una cinta transportadora) definen el retardo puro total del sistema.

Este concepto permite representar los sistemas de alto orden por medio de sistemas de primer orden a los que se les adiciona un retardo puro.

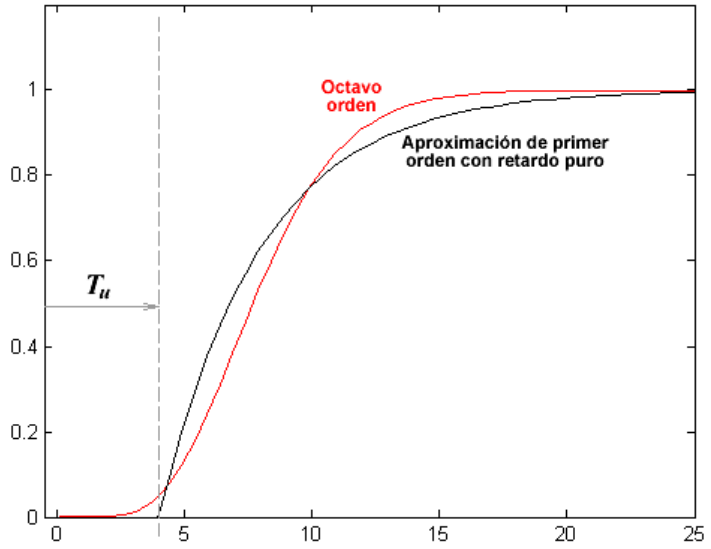


Fig. 3: Aproximación del sistema de octavo orden de la Fig 2 con un Sistema de primer orden más retardo puro.

La expresión matemática de esta aproximación tiene la forma

$$G(s) = \frac{e^{-T_u s} \; K_e}{1 + T_g \, s} \tag{1}$$

Donde K_e: Ganancia estática
T_g: Constante de tiempo
T_u: Retardo puro

Si bien puede parecer una simplificación un tanto burda, los motivos para usarla son los siguientes:

- Describe de una forma simple la gran mayoría de los procesos industriales
- Las reglas empíricas de calibración de parámetros están desarrolladas para este tipo de modelo

1.1 Análisis del sistema de primer orden con retardo puro en los dominios temporal y frecuencial

Sobre este tipo de modelo es apropiado hacer algunas consideraciones que serán útiles a la hora del diseño del controlador correspondiente y del análisis del funcionamiento en lazo cerrado. En particular, interesará la relación entre T_g y T_u.

En el dominio temporal, la respuesta es del tipo

$$y(t) = \begin{cases} 0 & \text{para } t < T_u \\ K_e\left(1 - e^{-\frac{t-T_u}{T_g}}\right) & \text{para } t \geq T_u \end{cases} \qquad (2)$$

De esta expresión se pueden derivar ciertos puntos característicos de la respuesta temporal al escalón unitario que se visualizan en la Figura 4, para un sistema con K_e = 1; T_g = 1 seg ; T_u = 0.5 seg:

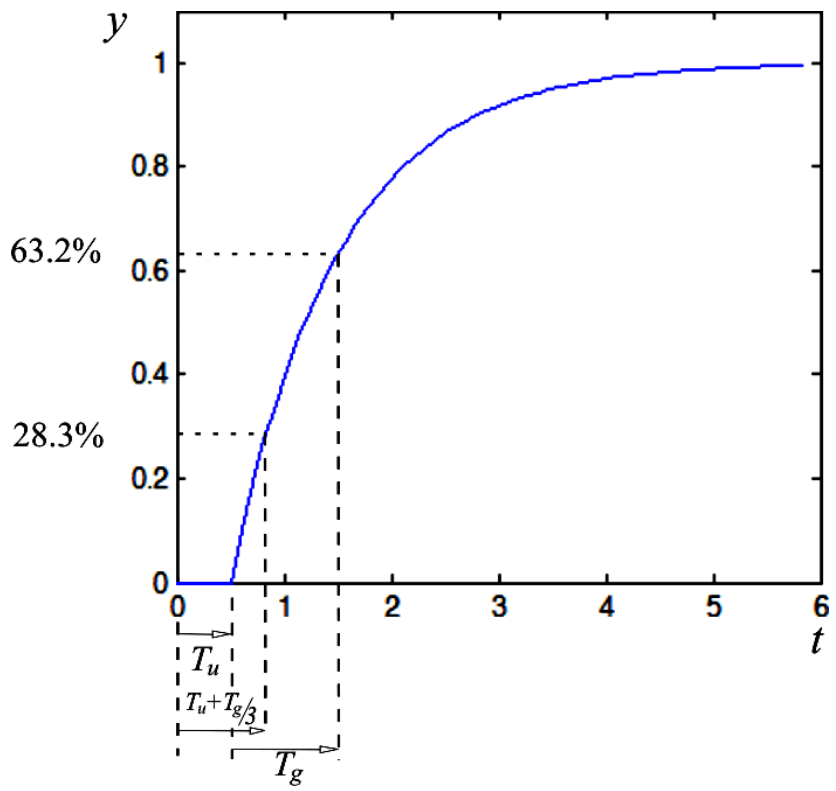


Fig. 4: Puntos característicos de la respuesta temporal de un sistema de primer orden con retardo puro

- El valor de estado estacionario es igual a K_e
- Alcanza el 28.3% de su valor final en el instante $T_u + T_g/3$
- Alcanza el 63.2% de su valor final en el instante $T_u + T_g$

Desde el punto de vista de la ubicación de los polos del sistema, que se relacionan con la respuesta temporal, si el sistema no tuviera un retardo puro, existiría un solo polo real ubicado en s = -1/T_g. Al considerar el retardo puro, el mismo es imposible de graficar ya que equivale a la presencia de infinitos polos. Sin embargo, es posible describir el retardo puro con una aproximación de Padé, por ejemplo de segundo orden, del tipo

$$e^{-T_u s} \cong \frac{1 - \frac{T_u}{2}s + \frac{T_u^2}{12}s^2}{1 + \frac{T_u}{2}s + \frac{T_u^2}{12}s^2} \tag{3}$$

Esta representación agrega en el plano complejo dos polos y dos ceros, de los cuales nos interesan fundamentalmente los dos polos, pues intervienen en la definición de la estabilidad relativa del proceso. La Figura 5 muestra en su sección superior la ubicación en el plano complejo de los polos y ceros del sistema de primer orden con K_e = 1; T_g = 1 seg, con un retardo puro T_u = 0.1 seg., aproximado con la expresión (3). Esto incorpora la presencia de dos polos en s = -30 $\pm j$ 17.32. Como ya se dijo, la estabilidad relativa del sistema está dada por los polos dominantes ubicados más cerca del eje imaginario, por lo cual está claro que en este caso está fijada fundamentalmente por el polo en s = -1/T_g y el retardo puro no tiene una influencia decisiva.

Nótese que la relación entre T_u y T_g es de 1 a 10. Pero si el retardo puro aumenta a T_u = 0.5 seg., la relación es de 5 a 10, ambos valores son comparables y los polos se ubican como se muestra en la parte inferior de la Fig. 5. Los dos polos incorporados por el retardo puro están en s = -6 $\pm j$ 3.46, mucho más cerca del polo en s = -1/T_g y por lo tanto se puede decir que en este caso el retardo puro involucra una inestabilización relativa mayor que en el caso anterior.

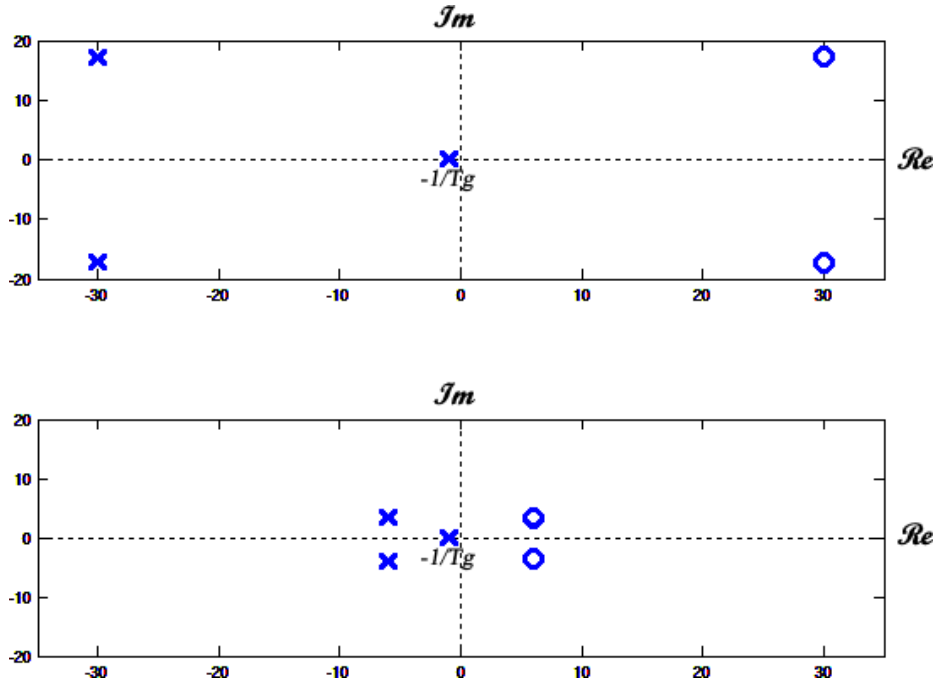


Fig. 5: Ubicación de polos y ceros de un proceso con dos retardos distintos, con aproximación de Padé de 2º orden.

Ya podemos expresar, en forma intuitiva, un resultado importante: **si el retardo puro es comparable con la constante de tiempo del sistema, el sistema será relativamente inestable y, por lo tanto, más difícil de controlar**.

Análisis en el dominio de la frecuencia

Esta idea se reafirma si se comparan, como se muestra en la Fig. 6, los diagramas de Bode del sistema sin retardo (azul), con un retardo de 0.1 seg (negro) y con uno de 0.5 seg. (negro a trazos). Las gráficas de módulo son coincidentes, y se percibe el cambio de comportamiento en la fase. Como es conocido, el módulo del retardo puro es unitario, pero es notable el efecto del retardo puro en el retraso de fase, el cual aumenta directamente con la frecuencia:

$$\angle e^{-T_u s} = \angle e^{-T_u j\omega} = -T_u\omega$$

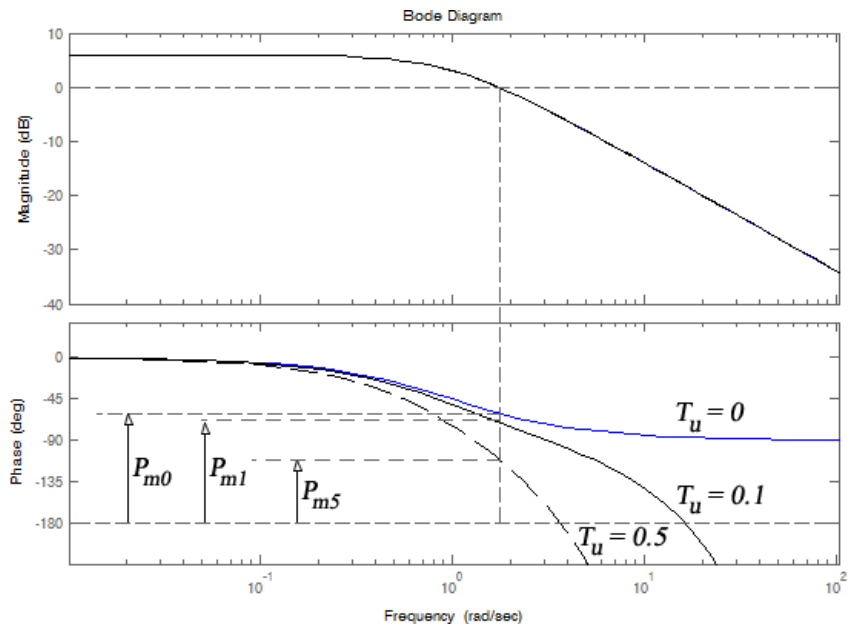


Fig. 6: Comparación de un sistema de primer orden con distintos retardos en el Diagrama de Bode.

El sistema sin retardo (T_u = 0 seg) presenta un margen de fase de P_{m0} = 120°. Al incorporar un retardo de T_u = 0.1 seg el mismo desciende a P_{m1} = 110.07°. Y el sistema con T_u = 0.5 seg tiene tan sólo un margen de fase de P_{m5} = 70.38°. Claramente se observa que el aumento del tiempo muerto involucra directamente un aumento de la inestabilidad relativa del sistema.

Dado que para los tres sistemas el diagrama de módulo, que depende de T_g, permanece sin variación, mientras que el de fase, dependiente mayormente de T_u, modifica sustancialmente su margen de fase, podemos decir algo más fuerte aún en cuanto a la estabilidad del sistema: **la variable más importante en cuanto a la estabilidad de un sistema y la facilidad para controlarlo es el retardo puro T_u.**

1.2 La importancia decisiva del retardo puro T_u

Se puede hacer aún un experimento más: comparar los diagramas de Bode de tres sistemas que tienen igual K_e = 1, con un retardo puro T_u = 0.1 seg., y difieren en las constantes de tiempo, que son T_{g1} = 0.5 seg, T_{g2} = 1 seg y T_{g3} = 4 seg. La Figura 7 muestra dicha comparación.

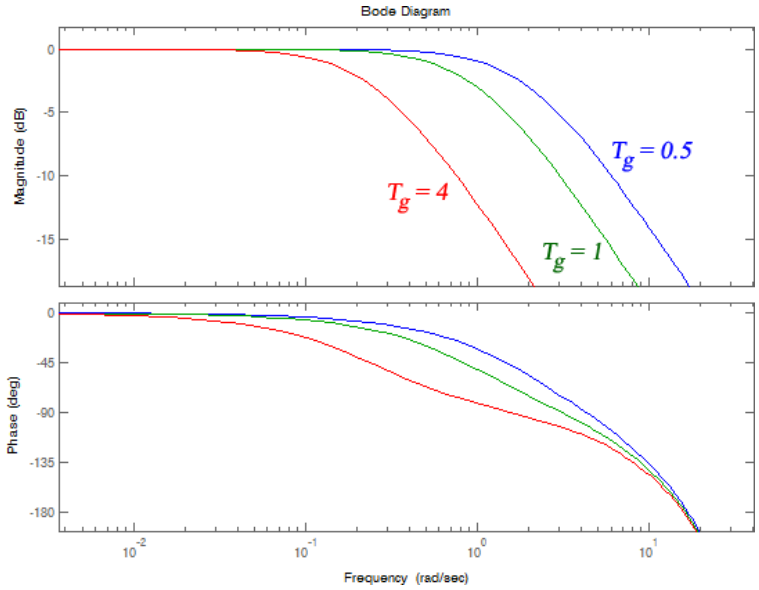


Fig. 7: Comparación de un sistema de primer orden con distintas constantes de tiempo en el Diagrama de Bode.

Ahora bien, como es conocido, multiplicando al sistema por una ganancia variable K y aumentando progresivamente la misma se llega a un punto en que tanto el margen de fase como el de ganancia son nulos y precisamente en esa circunstancia y a esa frecuencia el sistema se vuelve oscilatorio. El resultado de aplicar este procedimiento a los tres procesos señalados en la Fig. 7 se muestra en la Fig.8, y se llega a una importante conclusión: procesos con distintas constantes de tiempo oscilan aproximadamente a la misma frecuencia y por lo tanto

La frecuencia natural de oscilación del proceso depende fundamentalmente de T_u

¿Cuál es esta frecuencia? Para este caso $\omega_n \cong 15.7$ rad/seg. Es interesante comprobar que entonces el período natural de oscilación es $P_n = 2\pi / \omega_n \cong 0.4$, que es 4 veces el tiempo muerto T_u, de lo cual puede generalizarse la expresión

$$P_n \cong 4\, T_u \tag{4}$$

Efectivamente, comparando temporalmente un sistema cualquiera en lazo abierto y luego en lazo cerrado, aumentando en este último caso la ganancia hasta lograr la oscilación, se obtiene la gráfica que se muestra en la Figura 9. Una aproximación a una demostración formal de esta dependencia se encuentra en el Apén-

dice 1. **Este es uno de los resultados más importantes del control**, y como tal es de suponer que se vea reflejado en el diseño de los controladores PID. Más adelante se comprobará que efectivamente es así.

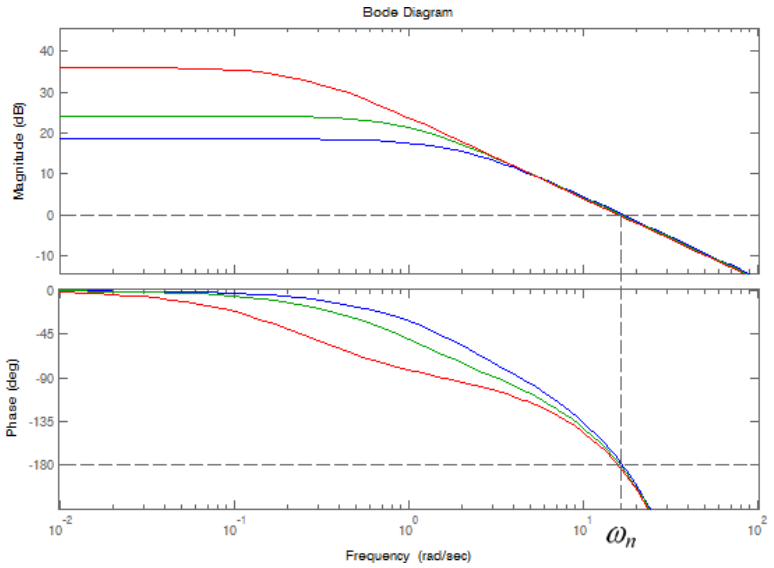


Fig. 8: Frecuencia natural de oscilación de sistemas con igual retardo puro.

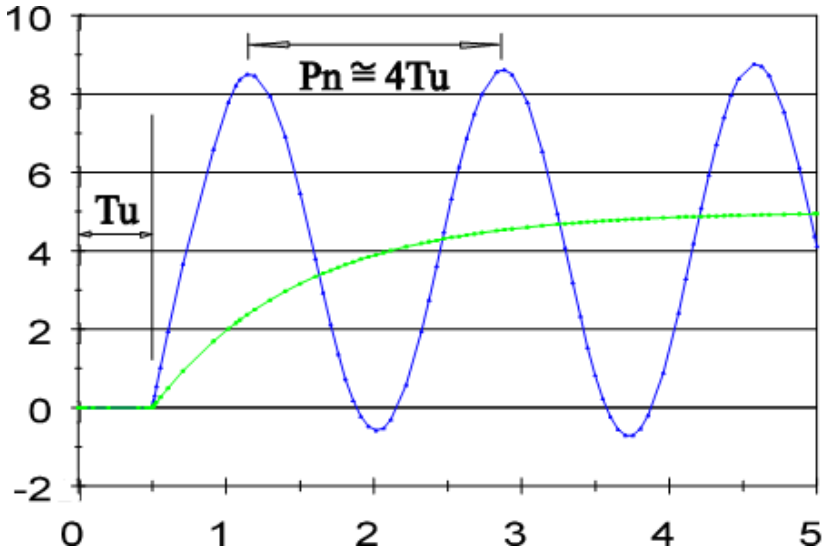


Fig. 9: Relación entre período natural de oscilación y retardo puro.

Por supuesto, este resultado es comprobable también con el método del lugar de las raíces, en el cual es necesario reemplazar el retardo puro por su aproximación de Padé del tipo (3), lo cual se muestra en la Figura 10. Puede verse con claridad que los tres sistemas, a pesar de poseer distintas constantes de tiempo T_g cruzan el eje imaginario –es decir se hacen oscilatorios- a $\omega_n \cong 15.7 \cong 2\pi / T_u$, confirmando la conclusión anterior.

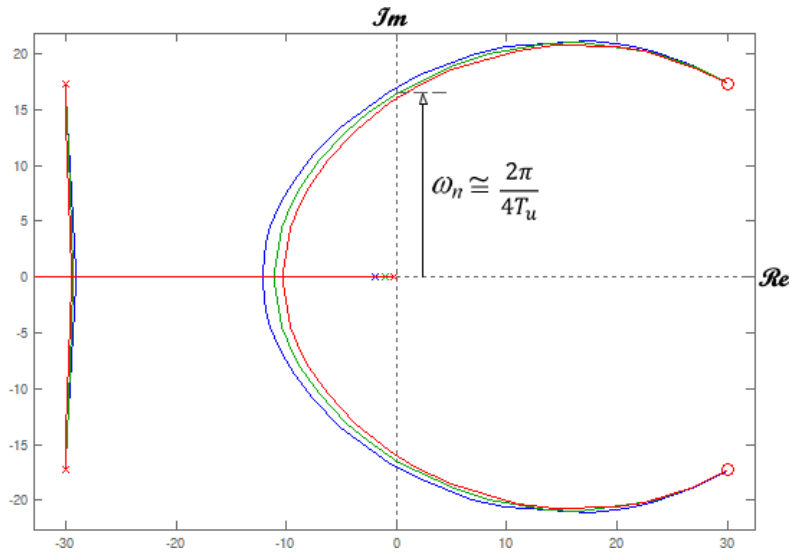


Fig. 10: Comparación de procesos con distinta constante de tiempo en el lugar de las raíces

Es de hacer notar que estos resultados –y, por cierto, el resto de las conclusiones que se derivarán para los controladores PID- son válidos aproximadamente, en términos relativos de T_g y T_u, hasta aproximadamente $T_u = T_g/2$. Cuando T_g y T_u son comparables, o $T_u > T_g$, hay que hacer algunas consideraciones adicionales. Sin embargo, insistimos en que **el sistema de primer orden con retardo puro representa apropiadamente a la gran mayoría de los sistemas industriales**.

2. Especificaciones de los sistemas de control en el dominio temporal

Para poder analizar el desempeño de los controladores PID que se describirán más adelante, es oportuno recordar las especificaciones de diseño que se requieren de un sistema de control. Las mismas se definen en relación a la respuesta $y(t)$ de un sistema con una referencia tipo escalón unitario y son las siguientes (Ver Figura 11):

1. Exactitud. Impuesto un cierto valor de referencia $r(t)$ al cual queremos que el sistema controlado llegue –en este caso este valor es 1 – puede ocurrir que no lo alcance, permaneciendo un error de estado estacionario $e_{ss} = r_{ss} - y_{ss}$ entre la referencia o valor deseado y el valor de estado estacionario, que nos dice cuan exacto es el sistema de control.

2. Velocidad de respuesta o tiempo de crecimiento t_r. Normalmente definida por medio del tiempo que el sistema tarda en llegar del 10% al 90% del valor de estado estacionario y_{ss}. Algunos autores prefieren el tiempo de retardo t_d, ya que este valor incorpora la información del retardo puro T_u. El mismo se define como el tiempo necesario para llegar al 50% del valor de estado estacionario y_{ss}.

3. Sobreelongación o máximo sobreimpulso M_p. Diferencia entre el valor máximo alcanzado y el valor de estado estacionario, la cual se pretende que no sobrepase un cierto porcentaje del valor de estado estacionario.

4. Tiempo de establecimiento t_s. Tiempo que tarda la salida en entrar en una banda ubicada alrededor del valor de estado estacionario y_{ss} y que por lo general se define como un 2% o 5% del mismo.

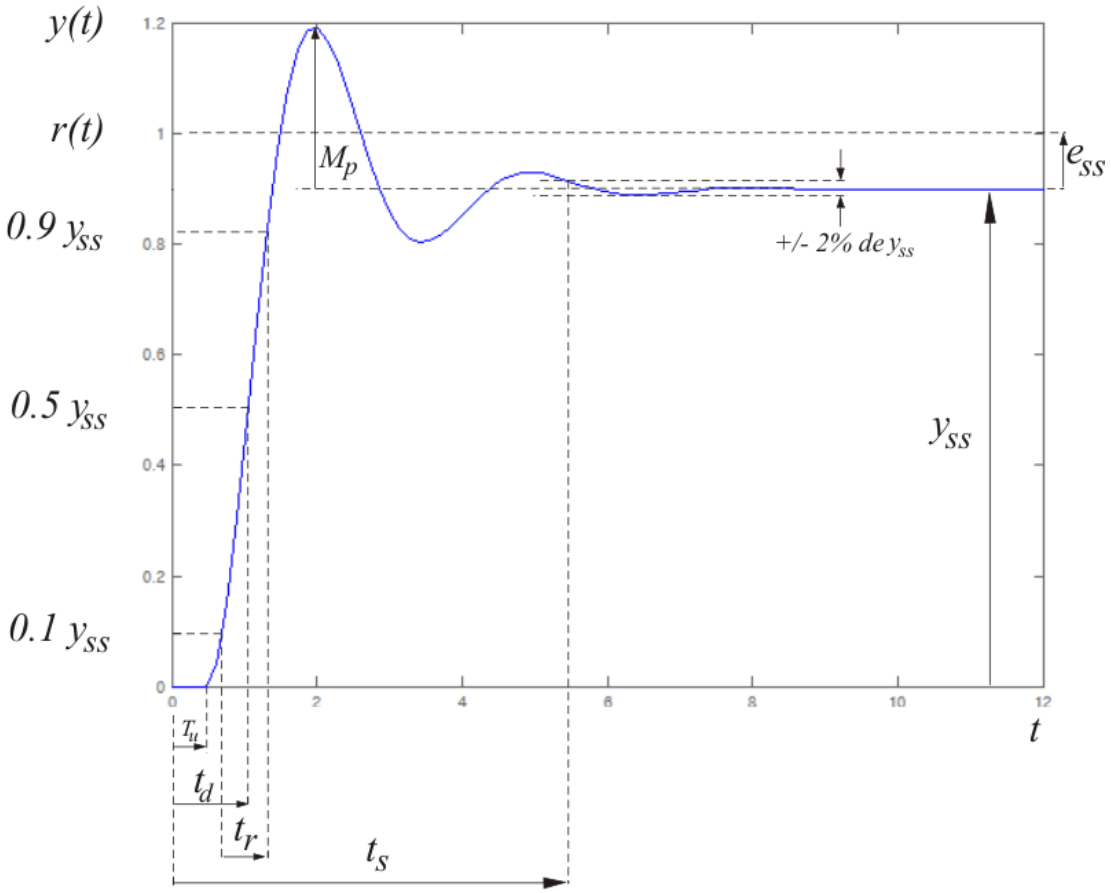


Fig. 11 Especificaciones de un sistema de control en el dominio temporal.

3. El controlador PID

3.1 El Controlador Proporcional P

Retomando la Figura 1, la forma más simple de controlador G_c que puede imaginarse es una constante K_c que multiplique al error:

$$u(t) = K_c\, e(t) \tag{5}$$

Dado que la acción de control resulta proporcional al error, recibe el nombre de *Controlador Proporcional* y se lo identifica con la letra P.

A veces, sobre todo en la industria, se expresa la ganancia proporcional K_c refiriéndose a su inversa porcentual: la *banda proporcional* BP, y se la define como el cambio, expresado en %, que debe sufrir la salida del proceso (variable medida) para que la salida del bloque proporcional pase de su valor mínimo (0%) a su máximo (100%):

$$BP = 100 / K_c \ [\%] \tag{6}$$

Como ya se conoce del análisis del lugar de las raíces de los procesos de primer orden como el de la Fig. 10, los polos resultantes de lazo cerrado, según sea la ganancia K_c, pueden estar en el eje real, con lo que se tiene una respuesta amortiguada, o pueden ser complejos conjugados con parte real negativa, obteniéndose una respuesta subamortiguada, o inclusive puede llegar a la inestabilidad si se sigue haciendo crecer la ganancia K_c. Así, por ejemplo, en la Figura 12 se presenta la respuesta a una referencia tipo escalón unitario de un sistema de primer orden con retardo puro, con parámetros K_e = 2, T_u = 0.1 seg., y T_g = 0.5 seg, con un controlador proporcional con ganancias K_{c1} = 0.5, K_{c2} = 2.25 y K_{c3} = 3. Tal como era de esperarse, la respuesta se vuelve más oscilatoria a medida que aumenta la ganancia. Si se siguiera aumentando la misma hasta K_c = 4.2512, el sistema se volvería marginalmente estable.

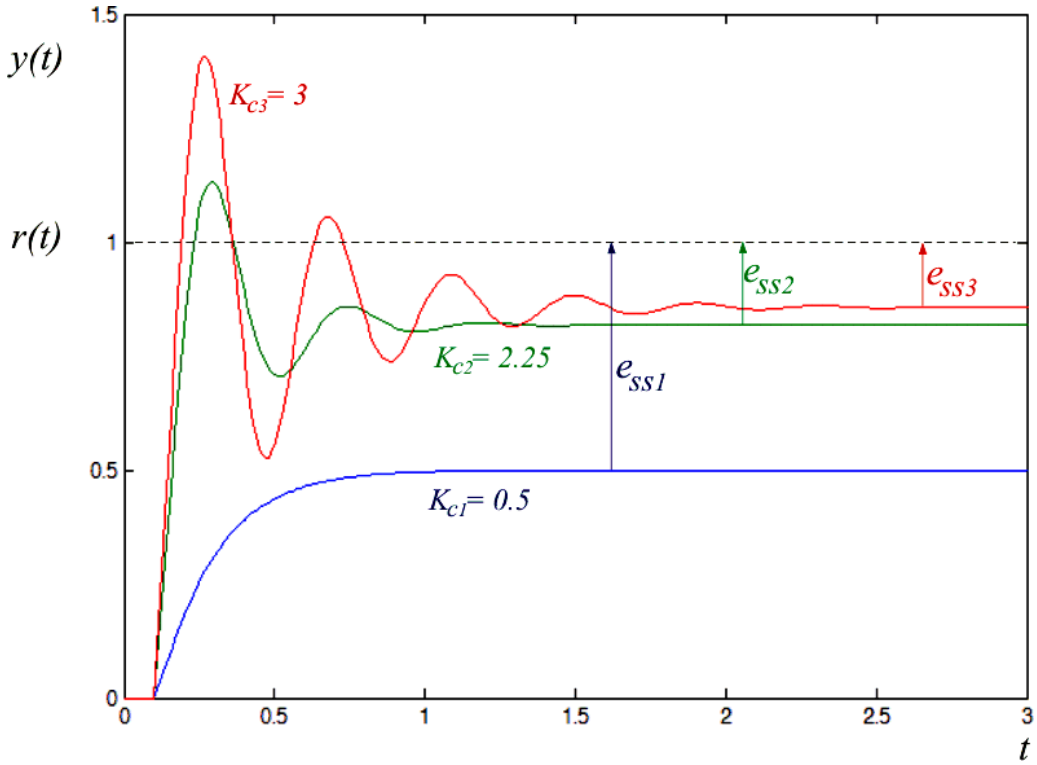


Fig. 12: Respuesta del sistema con un controlador proporcional con distintas ganancias

Una especificación de interés para el diseñador es el error de estado estacionario e_{ss}. Se sabe que la transformada de Laplace del error es

$$E(s) = \frac{R(s)}{1+L(s)}, \tag{7}$$

donde $R(s)$ es la transformada de la entrada de referencia $r(t)$ y $L(s)$ es la función de transferencia de lazo abierto del sistema junto con su controlador. Para este caso, $L(s)$ es

$$L(s) = G_c(s)\, G(s) = K_c \frac{e^{-T_u s}\ K_e}{1+ T_g\, s} \tag{8}$$

También es conocido, por el Teorema del valor final, que se puede calcular el error de estado estacionario como

$$e_{ss} = \lim_{t\to\infty} e(t) = \lim_{s\to 0} s\,E(s) \tag{9}$$

Reemplazando (7) y (8) en (9), y tomando en cuenta que la transformada de un escalón unitario es $R(s) = 1/s$, se obtiene

$$e_{ss} = \lim_{s\to 0} s\frac{1}{s}\frac{1}{1+K_c\,\dfrac{e^{-T_u s}\;K_e}{1+T_g\,s}} = \frac{1}{1+K_c\,K_e}, \tag{10}$$

resultado previsible para cualquier sistema tipo cero, es decir sin un integrador en el origen. Esta expresión confirma lo que se observa en la Fig. 12: aumentando la ganancia del controlador K_c se reduce el error de estado estacionario e_{ss}, **pero nunca se llegará a anularlo**.

Aún a pesar de la imposibilidad de anular el error, cabe preguntarse lo siguiente: si aumentando la ganancia el error disminuye e inclusive el aumento de ganancia nos brinda un menor tiempo de respuesta; ¿porqué no seguir aumentando entonces K_c? Porque la sobreelongación crece también, llegando hasta límites impracticables según el conjunto físico – químico considerado.

Este hallazgo nos alerta en cuanto a una característica fundamental de las especificaciones de los sistemas de control: **son incompatibles entre sí**, de tal manera que al mejorar una, siempre lo hacemos en detrimento de otra. Esto obligará al diseñador a algo muy habitual en la ingeniería: **encontrar una solución de compromiso**.

3.2 El Controlador Proporcional Integral PI

La incapacidad del Controlador Proporcional para anular el error sugiere la adición de un término que integre el error a fin de eliminarlo, con lo cual se llega a un *Controlador Proporcional Integral PI.* La formulación en el tiempo es la siguiente:

$$u(t) = K_c\,[e(t) + K_i \int e(t)dt] \tag{11}$$

Por más chico que sea el error $e(t)$, el término integral incorporado lo acumulará generando una acción de control que terminará por anular el error. El término K_i se denomina *Ganancia integral*, y suele llamársele también *Acción de reset*, dado que su objetivo es "*resetear*" el proceso llevándolo al valor impuesto por la referencia aun a pesar de las perturbaciones.

La función de transferencia del controlador es

$$G_{PI}(s) = \frac{U(s)}{E(s)} = K_c\left(1+\frac{K_i}{s}\right) \tag{12}$$

Un simple análisis dimensional del sumando K_i / s indica que la ganancia integral tiene dimensión de frecuencia. De ahí que también recibe el nombre de *frecuencia de reposición*, y habitualmente se especifica para sistemas industriales en ciclos/minuto ó repeticiones/minuto.

La inversa de la Ganancia integral es la *Constante de tiempo integral* T_i :

$$T_i = 1 / K_i \tag{13}$$

Este tiempo suele definirse como *aquel que debe transcurrir para que la acción integral alcance (iguale o repita) a la acción proporcional*. Esto se ve claramente si proponemos un error constante unitario $e(t) = 1$; la parte proporcional de la Ec. 11 será $u_P(t) = K_c$, mientras la parte integral responderá a $u_I(t) = K_c K_i t$. Obviamente, cuando $t = T_i$ se cumplirá que $u_I(t) = K_c K_i T_i = K_c = u_P(t)$.

Si se coloca la expresión (12) como

$$G_{PI}(s) = K_c \left(\frac{K_i + s}{s} \right) \tag{14}$$

se advierte la dificultad que plantea este controlador: si bien tiene la gran virtud de anular el error de estado estacionario **agrega un polo en el origen** (precisamente el integrador), **lo cual incorpora un elemento de inestabilización**. Esta característica se aprecia en la Figura 13, en la cual se compara la respuesta del ya conocido sistema con parámetros $K_e = 2$, $T_u = 0.1$ seg., y $T_g = 0.5$ seg, en ambos casos con ganancia proporcional $K_c = 2.25$ pero en un caso con ganancia integral $K_i = 0$ –es decir un controlador proporcional- y en otro[1] con $K_i = 3$, en ambos casos frente a una referencia tipo escalón unitario.

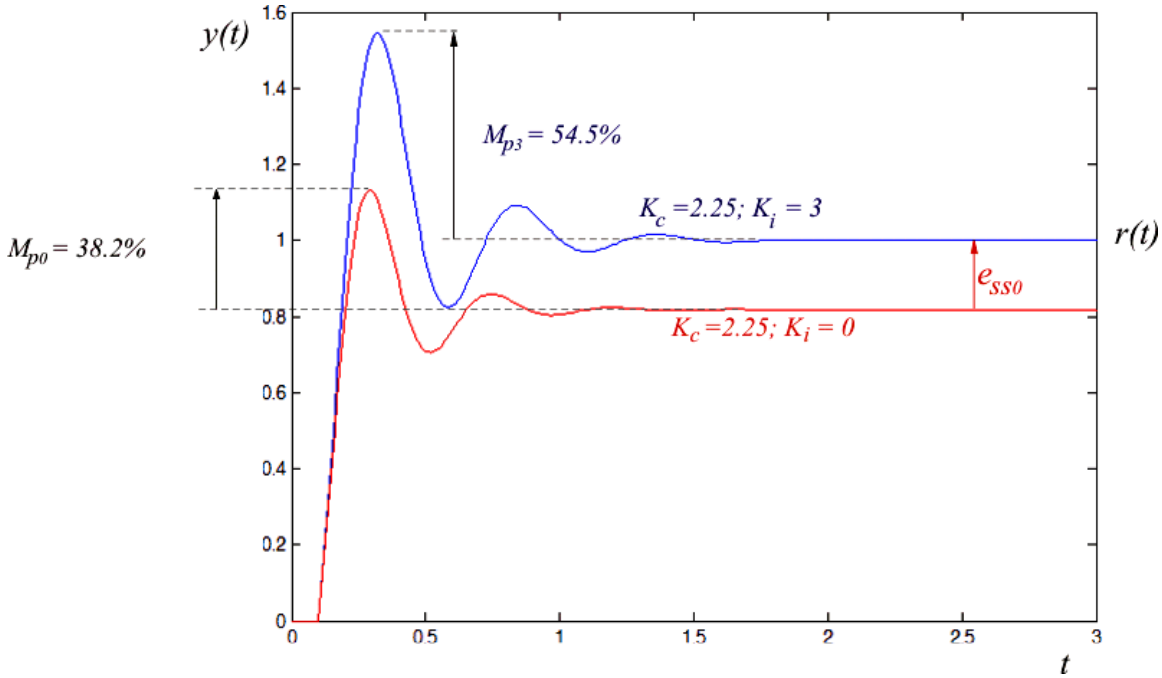


Fig. 13: Comparación temporal de un controlador proporcional con un PI

[1] La selección de los parámetros del controlador puede parecer caprichosa por ahora. Más adelante se explicará el procedimiento para determinar estos valores.

El primero presenta un error de estado estacionario e_{ss0} mientras que el que tiene término integral no nulo consigue eliminar el error. Como contrapartida, se observa la inestabilización relativa en una sobreelongación M_{p3} bastante superior a M_{p0} . También se podrían considerar otros parámetros como el tiempo de establecimiento, que también aumenta.

El efecto que produce el controlador PI en el dominio de la frecuencia puede observarse en la Figura 14. En la misma, la respuesta del sistema con controlador proporcional está graficado en color rojo, la del controlador PI por separado en verde (con frecuencia esquina en K_i = 3) y el sistema controlado con un PI en azul. Observando la fase en baja frecuencia se comprueba que el PI funciona como una red de atraso de fase. Claramente se ve también la reducción en los márgenes de ganancia y fase cuando se incluye el término integral, indicando una menor estabilidad relativa, y confirmando el resultado observado en el dominio temporal.

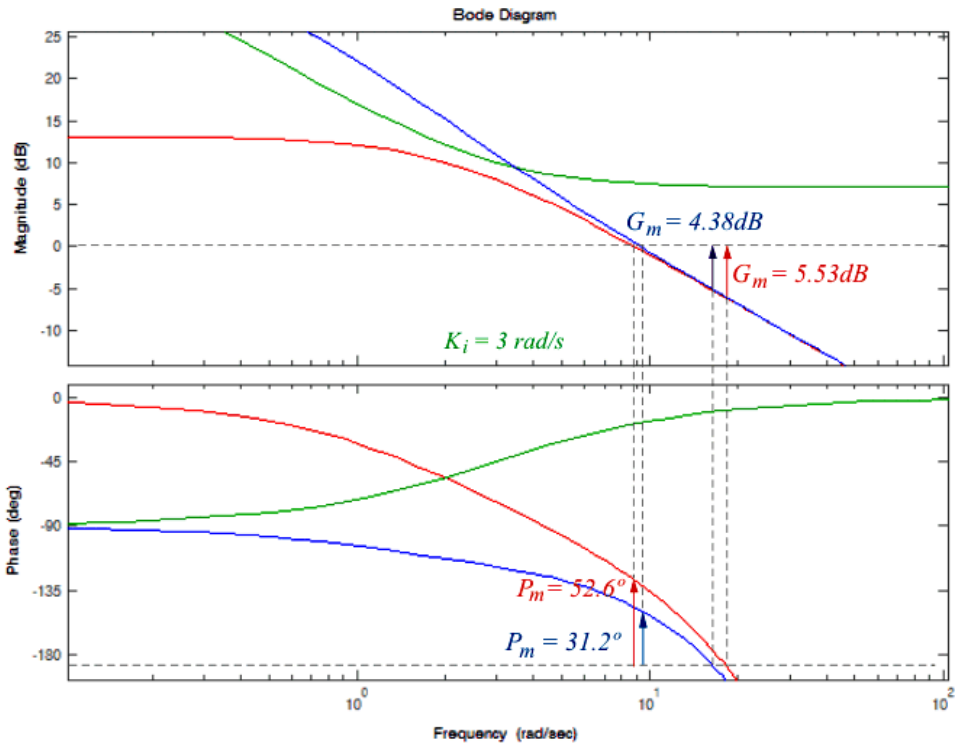


Fig. 14: Comparación en el dominio frecuencial de un controlador proporcional y un PI

3.3 El Controlador Proporcional Integral Derivativo PID

Si la inclusión de un término integral inestabilizó el sistema, la simple intuición indica que la incorporación de un término derivativo contrarrestaría este efecto. En este caso se implementa ponderando la derivada del error con una *Ganancia derivativa* K_d , la cual recibe también el nombre de *Constante de tiempo derivativa* T_d. Con este razonamiento se llega al controlador *Proporcional Integral Derivativo PID*, cuya expresión matemática en el tiempo es

$$u(t) = K_c \left[e(t) + K_i \int e(t)dt + K_d \frac{d(e(t))}{dt} \right] \tag{15}$$

La acción derivativa es la que estabiliza el sistema

La Figura 15 muestra, junto con la ya conocida respuesta del sistema con los controladores P y PI, la que presenta con el controlador PID, en color negro. Se aprecia que este controlador igualmente anula el error de estado estacionario, conserva un buen tiempo de crecimiento pero reduce notablemente el máximo sobreimpulso M_p y el tiempo de establecimiento, señales de una mayor estabilidad relativa.

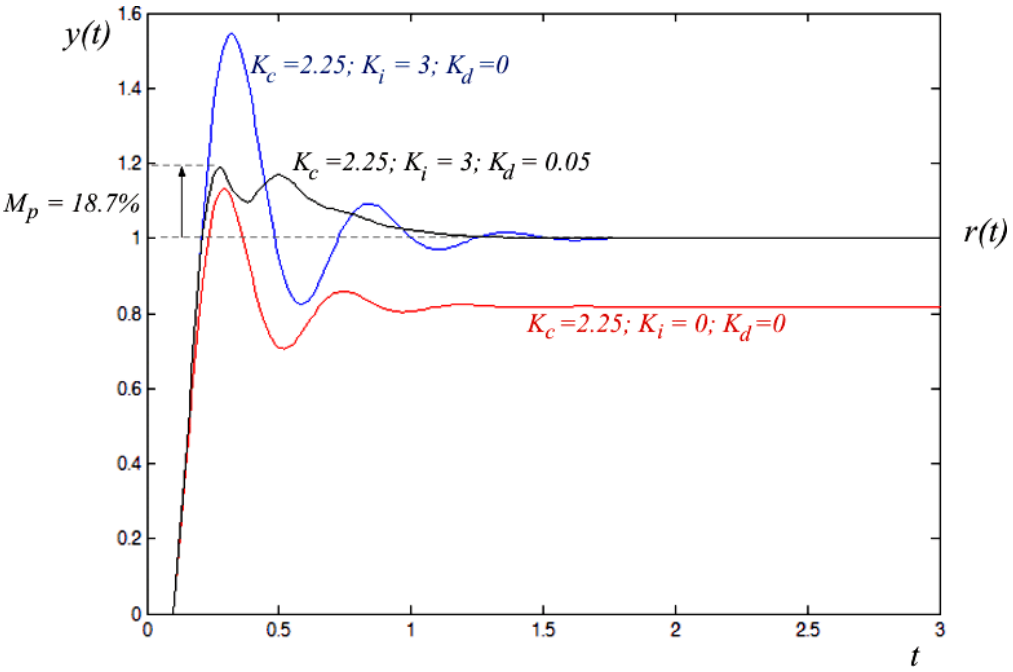


Fig. 15: Comparación temporal de un sistema con los tres tipos de controlador: P, PI y PID.

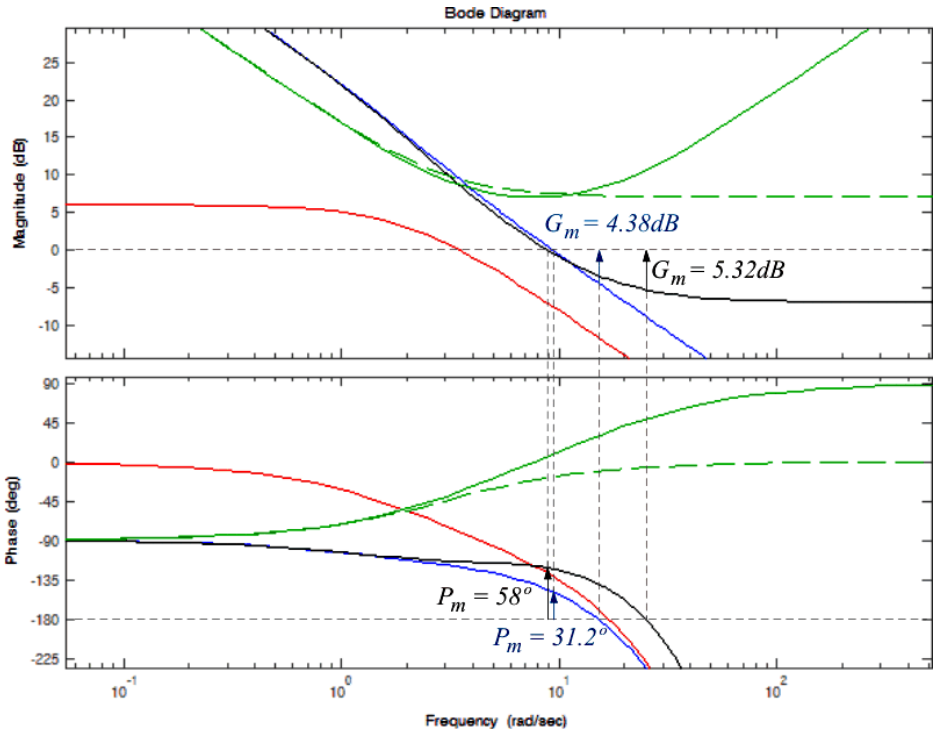


Fig. 16: Comparación en el dominio frecuencial de un controlador PI y un PID

De igual forma, puede realizarse una comparación en el campo de la frecuencia entre el sistema controlado con los distintos controladores, lo cual se grafica en la Figura 16. La respuesta en frecuencia del sistema controlado con un controlador proporcional se muestra en rojo, con un controlador PI en azul y con un PID en negro. En verde a trazos se muestra la respuesta del controlador PI y también en verde con línea llena, la del PID. **El PID puede verse entonces como un filtro rechazabanda o como una red de atraso – adelanto de fase**. Se puede apreciar como el PID amplía tanto el margen de fase como el de ganancia del sistema controlado, lo que concuerda con el análisis anterior en el espacio del tiempo.

Es importante tener en cuenta que **los parámetros seleccionados para el controlador en este ejemplo no son óptimos**, sino que han sido elegidos con un fin didáctico. Como ya se anticipó, en un próximo apartado se discutirán los métodos de selección de los parámetros del controlador.

3.4 Distintas configuraciones de controlador PID

Nótese que en la expresión (15) la ganancia K_c del controlador PID responde a una estructura de diagrama de bloques como la que se muestra en la Figura 17, y recibe la denominación de *No Interactiva*. Tiene como función de transferencia

$$\frac{U(s)}{E(s)} = K_c \left(1 + \frac{K_i}{s} + K_d\, s\right) \tag{16}$$

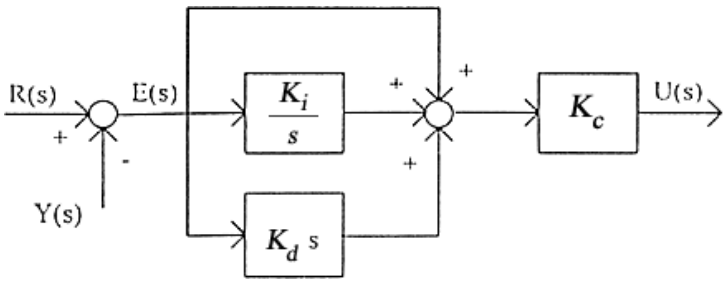


Fig. 17: Estructura de un PID No Interactivo

Esta estructura es la más conocida y la mayoría de los desarrollos de este texto hacen referencia a ella. Existen también otras estructuras como la *Interactiva*, bastante habitual en la industria, cuyo diagrama de bloques se observa en la Figura 18, y responde a la expresión

$$\frac{U(s)}{E(s)} = K'_c \left(1 + \frac{K'_i}{s}\right) \left(1 + K'_d\, s\right) \tag{17}$$

Los efectos de las partes integral y derivativa aparecen multiplicados en lugar de sumados como en el caso No Interactivo.

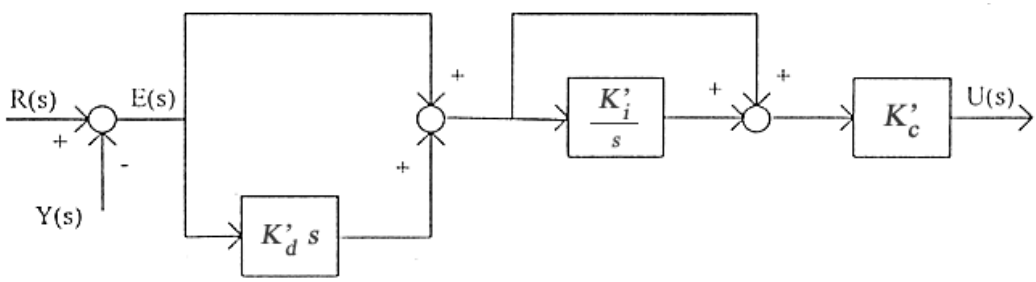


Fig. 18: Estructura de un PID Interactivo

Una tercera configuración es la estructura *Paralelo*, en la cual los tres efectos están completamente independizados:

$$\frac{U(s)}{E(s)} = K''_c + \frac{K''_i}{s} + K''_d\, s \tag{18}$$

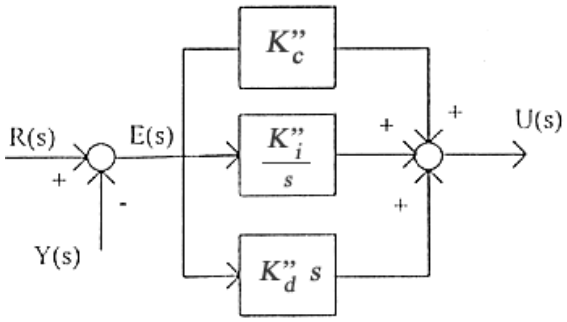


Fig. 19: Estructura de un PID Paralelo

Esta estructura es la usada por el software MATLAB. En ese entorno, Las constantes K''_c, K''_i y K''_d son conocidas como P, I y D.

Nótese que los parámetros están designados con una y dos tildes para los casos Interactivo y Paralelo a fin de señalar que <u>no se trata de los mismos parámetros del controlador *No Interactivo*</u>. La conversión de los parámetros es de fácil resolución y se presenta en la Tabla 1.

4. Sintonía de controladores PID

Obviamente, hay una gran cantidad de combinaciones de K_c, K_i y K_d que funcionarán bien, ¡y más combinaciones aún que lo harán mal! El proceso de asignar valores para los parámetros de un controlador recibe el nombre de *sintonía*, lo cual ya establece una referencia conceptual al campo de la frecuencia.

La sintonía de los controladores PID ha sido encarada por distintos autores a partir de

- las características estimadas del proceso
- el criterio de sintonía elegido

Algunos de los criterios más habituales son el de razón de decaimiento de ¼, MIAE (mínima integral del valor absoluto del error), MISE (del cuadrado del error) y el MITAE (del valor absoluto del error ponderado en el tiempo).

Tabla 1: Conversión de parámetros entre distintos tipos de controlador PID

Tipo de estructura de PID	**Parámetros**		
No Interactivo	K_c	$K'_c\left(1+K'_i K'_d\right)$	K''_c
	K_i	$\frac{K'_i}{1+K'_d}$	$\frac{K''_i}{K''_c}$
	K_d	$\frac{K'_d}{1+K'_d K'_i}$	$\frac{K''_d}{K''_c}$
Interactivo	$\frac{K_c}{2}\left(1+\sqrt{1-4K_iK_d}\right)$	K'_c	$\frac{K''_c}{2}\left(1+\sqrt{1-\frac{4K''_iK''_d}{K''^2_c}}\right)$
	$\frac{1}{2K_i}\left(1+\sqrt{1-4K_iK_d}\right)$	K'_i	$\frac{K''_c}{2K''_i}\left(1+\sqrt{1-\frac{4K''_iK''_d}{K''^2_c}}\right)$
	$\frac{1}{2K_i}\left(1-\sqrt{1-4K_iK_d}\right)$	K'_d	$\frac{K''_c}{2K''_i}\left(1-\sqrt{1-\frac{4K''_iK''_d}{K''^2_c}}\right)$
Paralelo	K_c	$K'_c\left(1+K'_i K'_d\right)$	K''_c
	$\frac{1}{K_i K_c}$	$\frac{1}{K'_i K'_c}$	K''_i
	$K_c K_d$	$K'_c K'_d$	K''_d

Las reglas de sintonía de controladores más comunes que permiten encontrar los parámetros de los mismos fueron desarrolladas por Ziegler y Nichols (Optimun Settings for Automatic Controllers, Transactions of ASME, Nov. 1942). Son reglas empíricas en las cuales el criterio de sintonía es la razón de decaimiento de ¼, es decir se diseña el sistema de tal manera que tenga **respuesta subamortiguada y con una relación de un cuarto entre las amplitudes de las dos primeras oscilaciones** para variaciones en la carga (variable de salida) o en la referencia. La razón de decaimiento debería ser independiente de la variación seleccionada y debería depender solamente de las raíces de la ecuación característica del lazo de control.

El análisis de respuesta en frecuencia fue desarrollado a principios de la Segunda Guerra Mundial. Recién con la difusión de esta herramienta de análisis en el ámbito de la industria alrededor de 1950, se comenzó a apreciar la simplicidad y validez técnica de las reglas de Ziegler y Nichols. Otros autores pulieron estas reglas, pero siguen siendo las más eficaces de que disponemos.

4.1 Métodos de sintonía de Ziegler y Nichols

Ziegler y Nichols proponen dos métodos de sintonía: uno en lazo cerrado y otro en lazo abierto.

4.1.1 Método de lazo cerrado

Como ya se dijo, la sintonía toma en cuenta primeramente las características estimadas del proceso. También se anticipó que los métodos habituales de sintonía consideran un modelo de primer orden con retardo puro del tipo descripto en el apartado 1. Esto significa identificar de alguna manera los valores más importantes del modelo (1). Esto se puede hacer por distintos caminos. Inclusive MATLAB está provisto de algoritmos que permiten realizar tal identificación a partir de las señales entrada – salida del proceso. En el Método de lazo cerrado, Ziegler y Nichols identifican un valor característico del proceso, el Período natural P_n, llevando el proceso a un estado de oscilación.

Los pasos a seguir son los siguientes:

1. Convertir el PID en un simple controlador proporcional, para lo cual:

   1a. Anular la ganancia integral, o lo que es lo mismo definir el tiempo integral (o de reset) en su valor más largo posible.
   1b. Poner la ganancia derivativa (tiempo derivativo) en cero o su valor mínimo.

2. Incrementar la ganancia (disminuyendo la banda proporcional) hasta que el lazo mantenga una oscilación sostenida pequeña. Obtenemos en esta circunstancia dos parámetros: ***la ganancia límite*** (o de oscilación) K_{cu} y ***el período de oscilación*** P_n***, que es el valor más importante.***

3. Ajustar el controlador PID con los siguientes parámetros:

$$K_c = 0.6\, K_{cu}$$
$$K_i = 1/\, T_i = 2/\, P_n$$
$$K_d = T_d = P_n/8$$

Algunas aclaraciones importantes:

a. Ziegler y Nichols definieron estas reglas para una estructura no interactiva de controlador. Si se está trabajando con otra configuración, deberán hacerse las conversiones respectivas.
b. Debe tomarse un especial recaudo en trabajar con las mismas unidades definidas por el controlador usado. Así, por ejemplo, es muy habitual que la Ganancia del controlador esté definida por el fabricante del controlador como:

   K_c = %salida/%entrada = 100/banda proporcional.

   Por lo tanto la ganancia límite deberá estar expresada en las mismas unidades.

   Del mismo modo, es habitual que los tiempos integral y derivativo del controlador estén expresados en minutos, y por lo tanto el período natural de oscilación P_n deberá introducirse en las mismas unidades.

La Tabla 2 resume la asignación de parámetros según Ziegler – Nichols siguiendo el ensayo de lazo cerrado. Se incluye la calibración también para los controladores P y PI. Si bien no es lo más habitual, puede ocurrir por ejemplo que un determinado proceso tenga un polo en el origen, es decir tenga incorporado en sí mis-

mo un integrador puro, lo que haría redundante el integrador del término integral, por lo cual se puede prescindir del mismo y usar tan sólo un controlador proporcional.

Tabla 2. Fórmulas de sintonía de Ziegler – Nichols para lazo cerrado

Tipo	Estructura	K_c	K_i	K_d
P	K_c	$0.5\ K_{cu}$		
PI	$K_c\left(1+\frac{K_i}{s}\right)$	$0.45\ K_{cu}$	$1.2\ /\ P_n$	
PID	$K_c\left(1+\frac{K_i}{s}+K_d\,s\right)$	$0.6\ K_{cu}$	$2\ /\ P_n$	$P_n\ /\ 8$

4.1.2 Método de lazo abierto

Por lo general, la idea de llevar a un sistema a su estado de oscilación, por más que la misma sea pequeña, produce temor en los responsables del proceso. La alternativa es intentar una identificación del modelo del proceso de primer orden con retardo excitándolo con un pequeño escalón en su entrada.

Tal como lo indica el nombre del método, el lazo NO debe estar cerrado, para lo cual se coloca el controlador en modo manual. Este método es conocido como el "método de la curva de reacción", designación debida a Cohen y Coon (1953), quienes también trabajaron sobre la respuesta de un sistema de primer orden con retardo a un escalón de entrada.

Manualmente se debe llevar al proceso a la zona en la que trabajará habitualmente y esperar que desaparezcan las variaciones en la salida. Cuando el proceso está suficientemente estable, se realiza un cambio tipo escalón a la entrada del proceso (salida del controlador) y se observa el comportamiento de la salida. El escalón debe ser del orden del 10% al 25% de la excursión total de la salida del controlador.

Como ya se ha detallado, el sistema responderá con una evolución temporal que puede aproximarse a la que se muestra en la Fig. 4. La graficación de la misma permitiría encontrar los tres parámetros del sistema, a saber K_e, T_g y T_u. Luego las ganancias del controlador se calculan como indica la Tabla 3.

Ahora bien, nótese que K_e y T_g aparecen solamente en el cálculo de la ganancia proporcional, y siempre están asociados en el valor $K_e\ /\ T_g$. Esto simplifica el cálculo ya que este cociente se puede obtener gráficamente como la pendiente del sistema de primer orden justo cuando comienza a reaccionar: se sabe que la respuesta temporal del sistema de primer orden sin retardo frente a un escalón unitario es

$$y(t) = K_e\,(1 - e^{-t/T_g}) \tag{19}$$

Tabla 3. Fórmulas de sintonía de Ziegler – Nichols para lazo abierto

Tipo	Estructura	K_c	K_i	K_d
P	K_c	$\frac{1}{(\frac{K_e}{T_g})T_u}$		
PI	$K_c\left(1+\frac{K_i}{s}\right)$	$\frac{0.9}{(\frac{K_e}{T_g})T_u}$	$\frac{1}{3.3\ T_u}$	
PID	$K_c\left(1+\frac{K_i}{s}+K_d\,s\right)$	$\frac{1.2}{(\frac{K_e}{T_g})T_u}$	$\frac{1}{2\ T_u}$	$0.5\ T_u$

La pendiente buscada, que es la tangente del ángulo α en la Figura 20, se obtiene como la derivada en $t = 0$:

$$\frac{dy}{dx}\bigg|_{t=0} = \frac{K_e}{T_g}e^{-t/T_g}\bigg|_{t=0} = \frac{K_e}{T_g} \tag{20}$$

Tomando en cuenta este dato, se puede concentrar la atención en la primera parte de la evolución del sistema, sin esperar a que llegue a su estado estacionario.

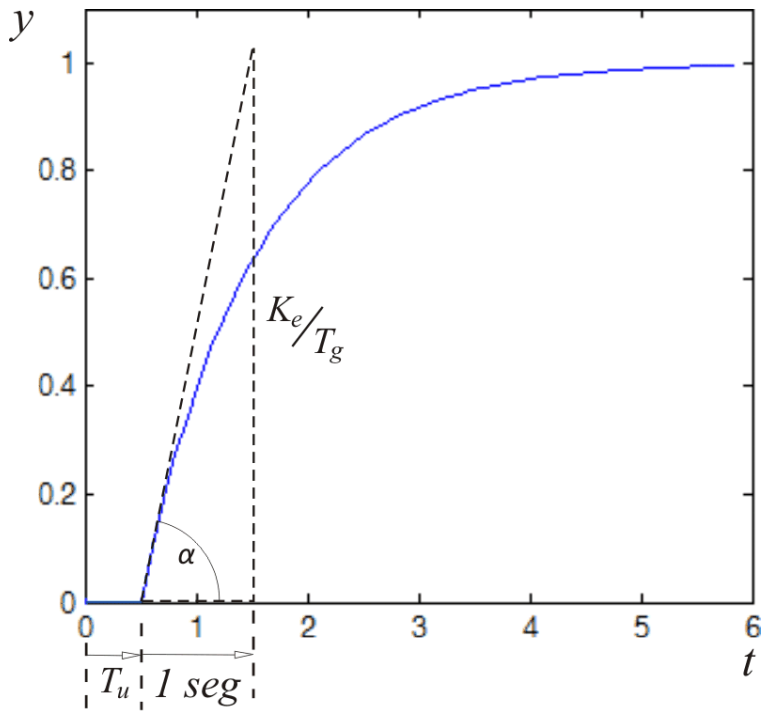


Fig. 20: Determinación de K_e / T_g

Ejemplo de aplicación real

Sea un sistema térmico a controlar con un PID, tal como se muestra en la Fig. 21.

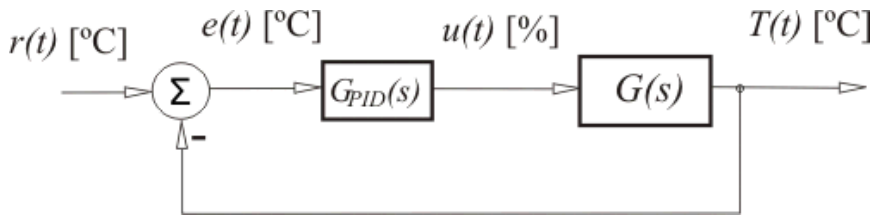


Fig. 21: Diagrama de bloques de un proceso térmico y su control PID

Debe prestarse especial atención a la estructura del controlador y las unidades en las que se manejan el proceso y el controlador. SIEMPRE DEBE PRESTARSE ATENCIÓN A LAS ESPECIFICACIONES DEL MISMO PROVISTAS POR EL FABRICANTE. En este caso se trata de un **PID no interactivo** cuya **salida está expresada en porcentaje** y su **unidad de tiempo es el segundo**.

A fin de calcular los parámetros del controlador, se realiza un ensayo al escalón a fin de relevar la curva de reacción. El ensayo realizado se muestra en la Figura 22. Se coloca el controlador en modo manual, se fija la salida del mismo en el 25%, y se espera a que el proceso se estabilice. Cuando esto ocurre, la temperatura de salida se ubica en 31.3°C. En estas circunstancias , se produce un salto en forma escalón del 25% al 50% disponible del actuador en el t = 54 seg.

La salida, tal como era previsible, se comportó aproximadamente como un sistema constituido por un retardo puro más un sistema de primer orden que tiende a estabilizarse en algo más de 54°C.

Como ya indicamos, nos concentramos en la primera parte de la curva, marcada con un cuadro en la Fig. 22 y que se ve amplificada en la Figura 23. Se obtiene por inspección de la gráfica:

a. ***El retardo puro*** $\boldsymbol{T_u}$, que es el dato más importante. El mismo se estima en T_u = 20 seg.

b. En segundo lugar, obtenemos $\boldsymbol{K_e / T_g}$ como la tangente del ángulo α:

$$\frac{K_e}{T_g} = \frac{\Delta T}{\Delta t}\frac{100}{\Delta u} = \frac{1.2\text{ºC}}{82\text{ seg}}\frac{100}{25\%} = 0.0585\,\frac{\text{ºC}}{\text{seg }\%} \tag{21}$$

Nótese que el valor de $\Delta T / \Delta t$ está multiplicado por $100/\Delta u$. Esto se debe a que el cociente $\Delta T / \Delta t$ es equivalente a K_e / T_g si se hubiese aplicado un escalón unitario o -lo que es equivalente- el 100% de la excursión del actuador. En este caso en particular, sólo se ha aplicado un Δu = 25%. Si se hubiese aplicado el 100%, se hubiese tenido una pendiente 100/Δu = 100/25 = 4 veces más alta.

Con estos valores se procede al cálculo de los parámetros del controlador a partir de las fórmulas empíricas de la Tabla 3:

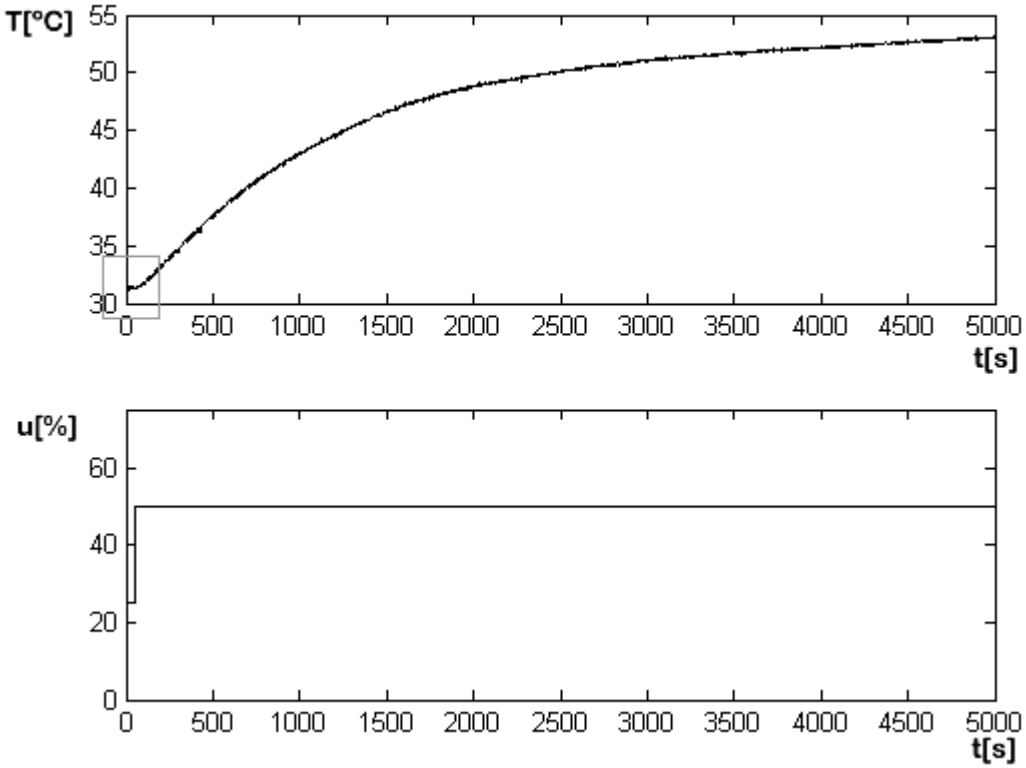


Fig. 22: Comportamiento de un sistema real excitado con un escalón

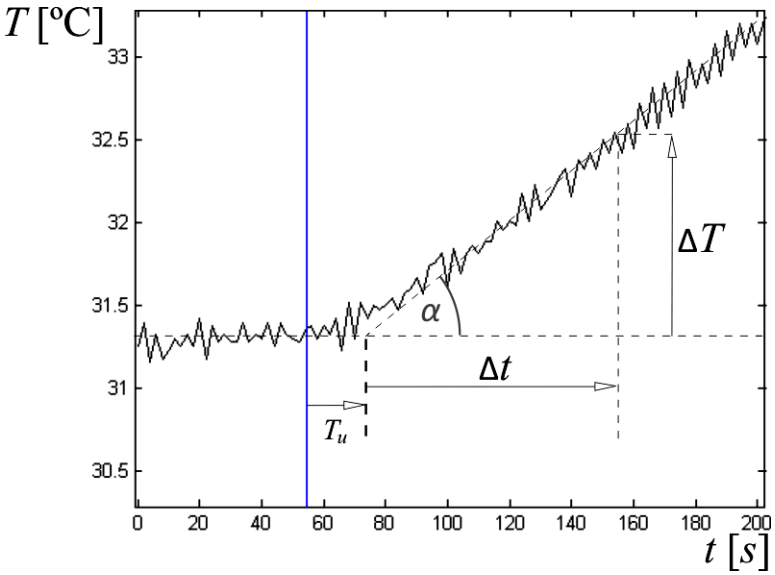


Fig. 23: Amplificación del primer segmento de la evolución temporal de la gráfica anterior.

$$K_c = \frac{1.2}{(\frac{K_e}{T_g})T_u} = \frac{1.2}{\frac{\Delta T100}{\Delta t\ \Delta u}\ T_u} = \frac{1.2}{0.0585\ \frac{ºC}{seg\ \%}\ 20\ seg} = 1.025\ \frac{\%}{ºC}$$

$$K_i = \frac{1}{2\ T_u} = \frac{1}{2\ \ 20seg} = 0.025\frac{1}{seg}$$

$$K_d = 0.5\ T_u = 0.5\ \ 20\ seg = 10\ seg$$

El lector puede realizar un análisis dimensional y comprobar la consistencia con las unidades expresadas en la Fig. 21.

La respuesta del sistema real controlado con estos parámetros frente a un escalón en la referencia de 35ºC a 45ºC se presenta en la Figura 24.

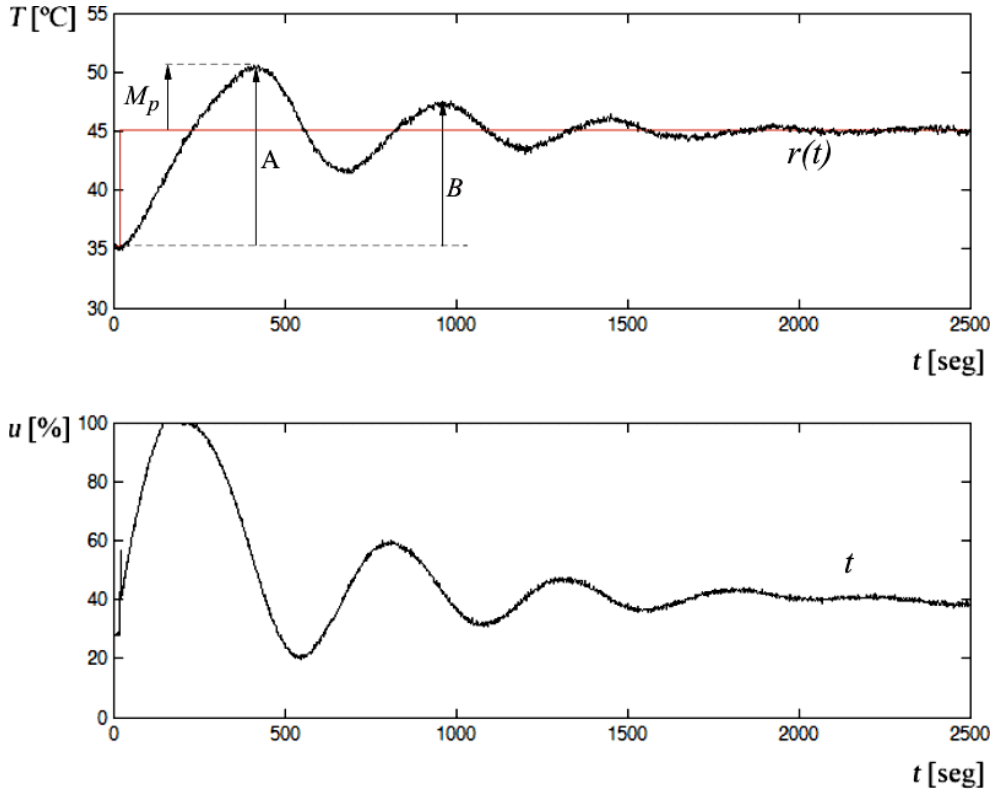


Fig. 24: Respuesta del sistema controlado

Se aprecia que el sistema resuelve correctamente el error de estado estacionario. Una comprobación rápida de que se ha obtenido un buen resultado es calcular la razón de decaimiento, siempre relativa al estado inicial de la variable de 35ºC:

$$r_d = \frac{B - 35ºC}{A - 35ºC} = \frac{47.19ºC - 35ºC}{50.5ºC - 35ºC} = 0.786,$$

con lo que se verifica el decaimiento de aproximadamente ¼.

Otra medida de que el resultado se aproxima a lo propuesto por Ziegler – Nichols es el máximo sobreimpulso, que debe estar alrededor del 60%. Para este caso es

$$M_p = \frac{\text{A - 45ºC}}{\text{45ºC - 35ºC}}\,100\% = 55\%,$$

¿Qué hubiera pasado si la respuesta hubiese sido demasiado oscilatoria? ¿O al contrario, si hubiese sido muy lenta? ¿Cuál es la mejor forma de corregir los parámetros del controlador? Daremos la respuesta en el próximo punto.

5. Características de los procesos controlados con PID con sintonía de Ziegler-Nichols

Salta a la vista que en el método de lazo cerrado las ganancia K_i y K_d dependen <u>sólo</u> del período natural P_n y en el caso del lazo abierto <u>sólo</u> del retardo puro T_u. Todo hace suponer que existe una relación entre el cálculo en lazo abierto y el cálculo en lazo cerrado, lo cual se verifica si consideramos por ejemplo el cálculo de K_i en ambos métodos (se obtendría igual resultado trabajando con K_d):

$K_i = 1 / T_i = 2 / P_n$ Lazo cerrado
$K_i = 1 / T_i = 1 / 2T_u$ Lazo abierto

Igualando estas expresiones se llega a $P_n = 4T_u$, resultado que ya conocíamos de la expresión (4), demostrada en el Apéndice 1. Este resultado confirma, al menos en parte, la validez teórica de los hallazgos empíricos de Ziegler y Nichols.

5.1 Características en el dominio temporal

Es interesante analizar los polos de lazo cerrado del sistema controlado con un PID, ya que la posición de los mismos es indicativa de la respuesta temporal.

Por un lado debe tenerse en cuenta que la función de transferencia del controlador PID, si se reemplazan los valores de K_i y K_d, dados por Ziegler-Nichols es

$$\frac{U(s)}{E(s)} = K_c\left(1 + \frac{K_i}{s} + K_d\, s\right) = \frac{K_c}{s}\left(K_i + s + K_d\, s^2\right) =$$

$$= \frac{1.2}{(\frac{K_e}{T_g})T_u}\,\frac{1}{s}\left(\frac{1}{2T_u} + s + 0.5\, T_u\, s^2\right) \qquad (22)$$

También, dado que el retardo puro $e^{-T_u s}$ no permite expresar la función de transferencia de lazo cerrado en una forma polinomial a fin de encontrar sus raíces, se recurre a la aproximación de Padé de segundo orden (3) del retardo puro. La función de transferencia de lazo abierto queda entonces

$$L_p(s) = \frac{1.2}{(\frac{K_e}{T_g})T_u}\,\frac{1}{s}\,\left(\frac{1}{2T_u} + s + 0.5\,T_u\,s^2\right)\left(\frac{1 - \frac{T_u}{2}s + \frac{T_u^2}{12}s^2}{1 + \frac{T_u}{2}s + \frac{T_u^2}{12}s^2}\right)\frac{K_e}{1 + T_g\,s} =$$

$$= \frac{1.2\,T_g}{T_u}\,\frac{1}{s}\,\left(\frac{1}{2T_u} + s + 0.5\,T_u\,s^2\right)\left(\frac{1 - \frac{T_u}{2}s + \frac{T_u^2}{12}s^2}{1 + \frac{T_u}{2}s + \frac{T_u^2}{12}s^2}\right)\frac{1}{1 + T_g\,s} \qquad (23)$$

Para un proceso caracacterizado por K_e = 3; T_g = 1.5 seg ; T_u = 0.15 seg. y sintonizado según Ziegler-Nichols, el diagrama del lugar de las raíces para una excursión de K_c entre 0 e ∞ es el que se muestra en la Figura 25.

En cuanto al lazo abierto se aprecia:

- Un polo en el origen, incluido por el integrador del PID
- Un polo en -1/ T_g.
- El controlador introduce dos ceros por el término $\frac{1}{2T_u} + s + 0.5\,T_u\,s^2$, que presenta dos raíces coincidentes en -1/ T_u.
- El retardo puro en el formato de Padé, incorpora | dos polos en $-\frac{3}{T_u} \pm \frac{\sqrt{3}}{T_u}$ | dos ceros en $\frac{3}{T_u} \pm \frac{\sqrt{3}}{T_u}$

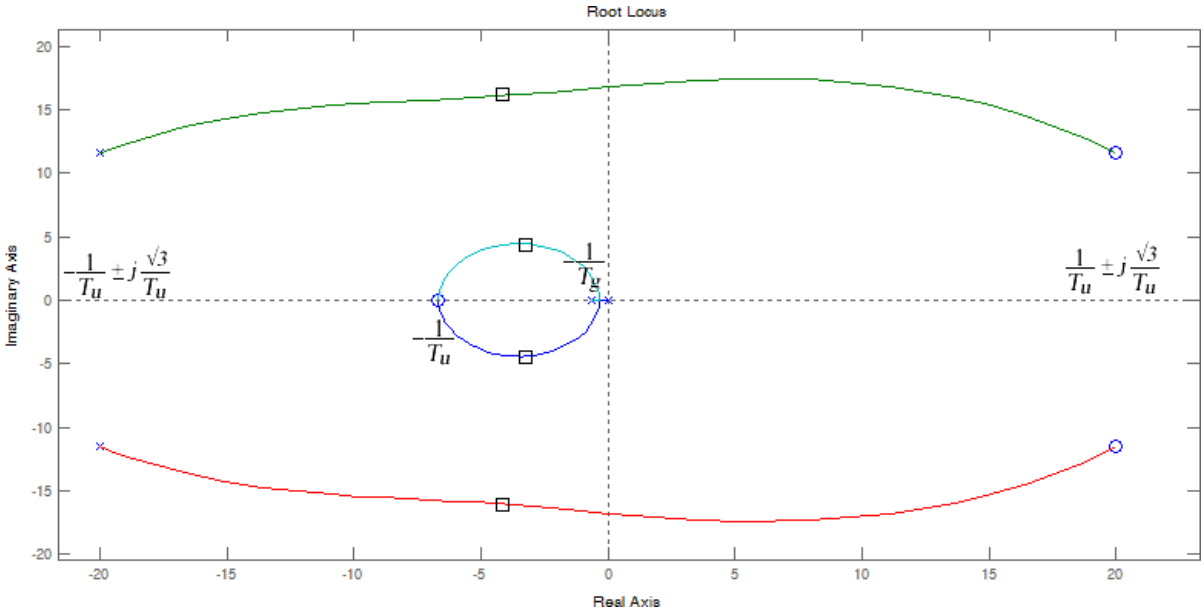


Fig. 25: Lugar de las raíces del sistema de primer orden
con aproximación de Padé del retardo y controlador PID.

Nótese que todos estos valores dependen de T_u que, como venimos anticipando, es el valor más importante a tener en cuenta. Una importantísima conclusión que se consolida a lo largo de todo lo desarrollado es que

Todo depende de los retardos

Para el valor de K_c previsto por Ziegler-Nichols, $K_c = 4$, en lazo cerrado se encuentran dos pares de polos complejos conjugados en

p_1 = -4.1608 + j16.0633 y
p_2 = -3.5476 + j4.4382.

En los mismos centraremos nuestro interés, ya que definen la dinámica de lazo cerrado.

La Figura 26 presenta en color azul la respuesta temporal del sistema con retardo puro (no aproximado por Padé) controlado por un PID sintonizado según Ziegler-Nichols frente a un escalón unitario. En verde y en rojo se observan las respuestas al impulso de los sistemas de segundo orden que responden a los polos p_1 y p_2 respectivamente. Es fácil ver que, tal como era de esperarse, la dinámica del sistema de lazo cerrado responde a una combinación de las dinámicas de estos dos pares de polos.

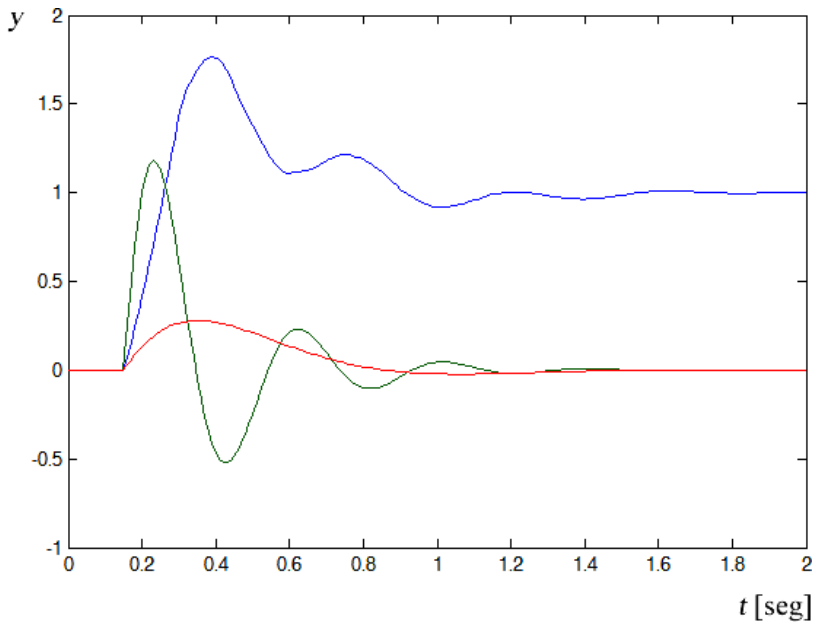


Fig. 26: Dinámica del proceso en lazo cerrado y de los dos pares de polos complejos conjugados representativos.

Ahora bien, si se analiza la ubicación de estos dos pares de polos para una variedad de sistemas, variando en una relación de 1 a 5 los tres parámetros del proceso, se comprueba que el controlador ubica los pares de polos de lazo cerrado p_1 y p_2 cada uno en una misma línea, manteniendo prácticamente sin modificación los ángulos φ_1 y φ_2. Véase la Figura 27. Dado que para un sistema normalizado de segundo orden con función de transferencia

$$C(s) = \frac{\omega_n^2}{s^2 + 2\delta\,\omega_n\, s + \omega_n^2} \tag{24}$$

el valor $\delta = \cos\varphi$ determina la estabilidad relativa del sistema, es decir cuán oscilatoria es la respuesta, luego los valores δ_1 y δ_2 serán prácticamente constantes.

Con esto se concluye que para una variada gama de procesos que se puedan representar por un sistema de primer orden con retardo puro,

Si se usa un controlador PID sintonizado según Ziegler-Nichols, la gran mayoría de los procesos presentarán una respuesta en el tiempo con características dinámicas muy semejantes.

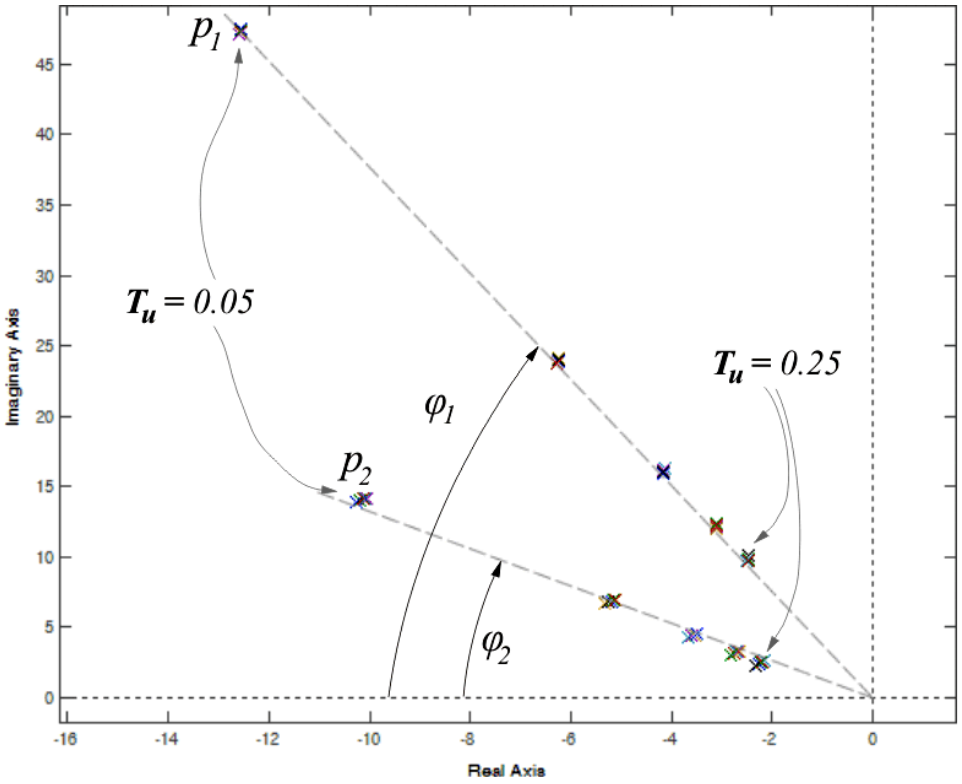


Fig. 27: Ubicación de los polos de lazo cerrado para una variadad de procesos.

Las características aproximadas más importantes de este tipo de respuesta temporal son:

- Sobreelongación de alrededor del 55%.
- Razón de amortiguamiento de ¼ entre la primera y la segunda oscilación.
- Luego de tres oscilaciones el sistema habrá alcanzado su estado estacionario.
- Cuando se presenta una perturbación, durante $P_n/2$ la salida se comporta como si no existiera control, y recién luego se comienza a manifestar el efecto de la acción de control. Véase la Figura 28.

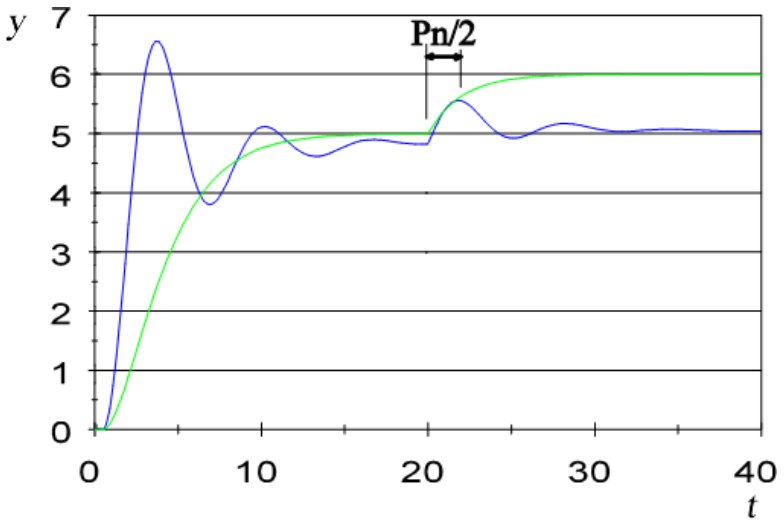


Fig. 28: Retardo frente a una perturbación.

Se insiste en que todas estas características, si bien muy uniformes para una amplia variedad de procesos, son sólo aproximadas, ya que hay procesos que se asemejan más que otros a un sistema de primer orden con retardo puro. Además, deben considerarse las alinealidades propias del proceso, no contempladas en este tipo de modelamiento.

5.2 Características en el dominio frecuencial

La Figura 29 muestra en color verde el diagrama de Bode de un controlador PID sintonizado según Ziegler-Nichols; en azul, el sistema de primer orden con retardo sin controlar, y finalmente en rojo el sistema controlado.

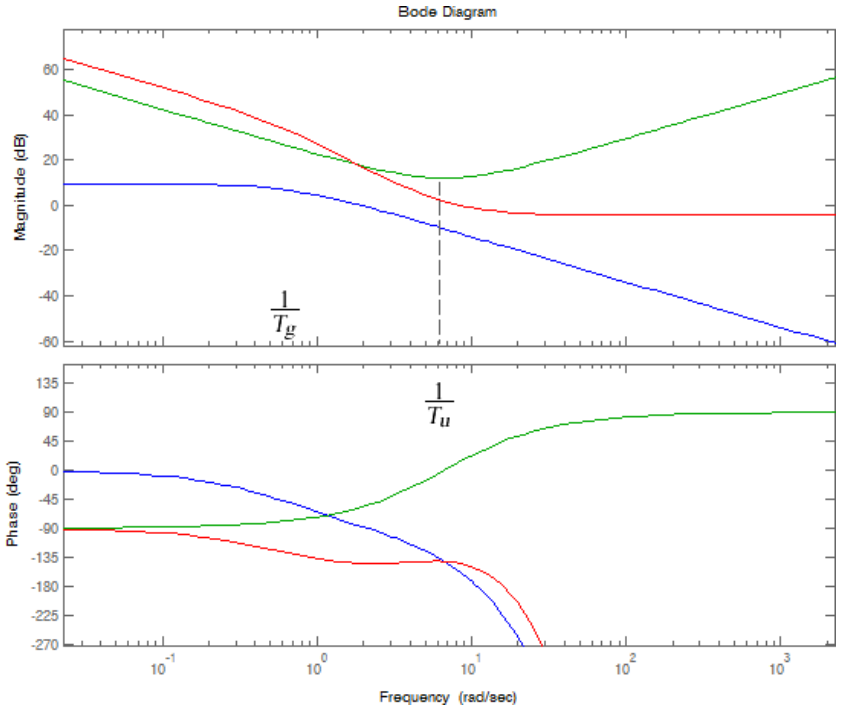


Fig. 29: Diagrama de Bode del sistema y su controlador

Ya se vio en la sección 3.3 (Fig. 16) que el PID puede concebirse como un filtro rechazabanda. También se sabe que la función de transferencia del controlador (22) presenta dos ceros coincidentes en -1/T_u. Esto significa que esa es la posición de ambas frecuencias esquina y por lo tanto el punto de mayor atenuación del filtro. La característica más importante de este controlador es entonces que, más allá de la ganancia K_c, que no hará más que bajar o subir la gráfica de Bode, **la sintonización depende solo de T_u.**

La Figura 30 presenta, en color azul y en línea llena, el diagrama de Bode de un sistema controlado con un PID, mientras que a trazos se grafica el diagrama del PID por separado. En rojo, exactamente lo mismo pero referido a un sistema que presenta un retardo puro 4 veces mayor al del primer sistema, y que en este caso comienza a ser comparable con la constante de tiempo del sistema T_g.

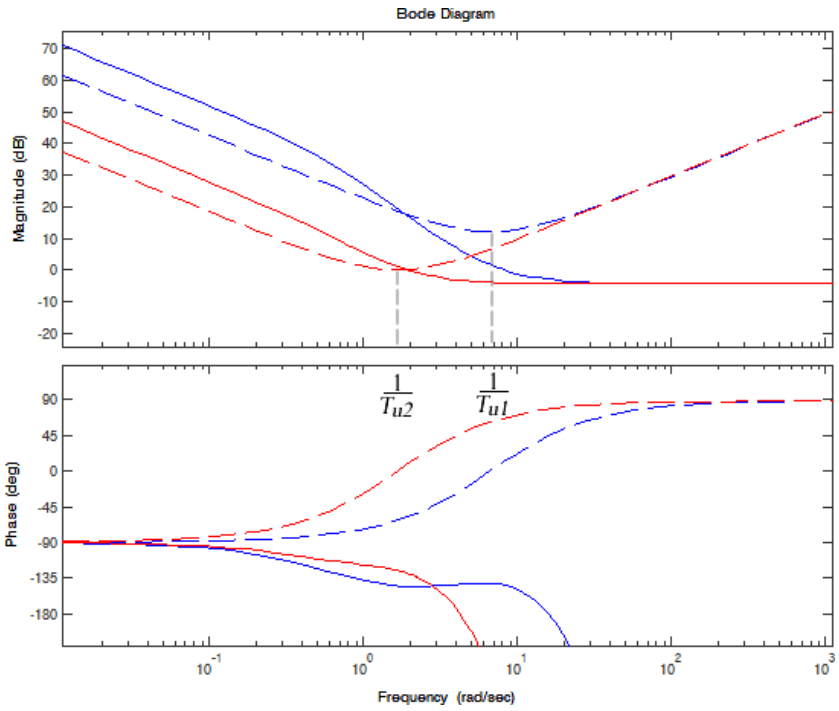


Fig. 30: Comparación de diagramas de Bode para
dos sistemas con diferentes retardos puros

En la Figura 31 se observa una amplificación de la Fig. 30, en la cual se puede verificar que una virtud importante del controlador PID sintonizado con Ziegler-Nichols es mantener márgenes de ganancia y fase relativamente constantes para una variedad de procesos, en particular para procesos con distintos retardos. Las variaciones en los márgenes se acrecentarán, como ya se sabe, a medida que el retardo puro T_u se asemeje a la constante de tiempo del sistema T_g, pero de igual manera sigue representando una muy buena selección de parámetros.

Como puede apreciarse, sintonizar un proceso que tiene un retardo puro mayor significa, en términos prácticos, desplazar la curva del controlador hacia una zona de menor frecuencia, despreciando la modificación efectuada sobre la ganancia, que no es tan importante. A la inversa, si el retardo fuese menor.

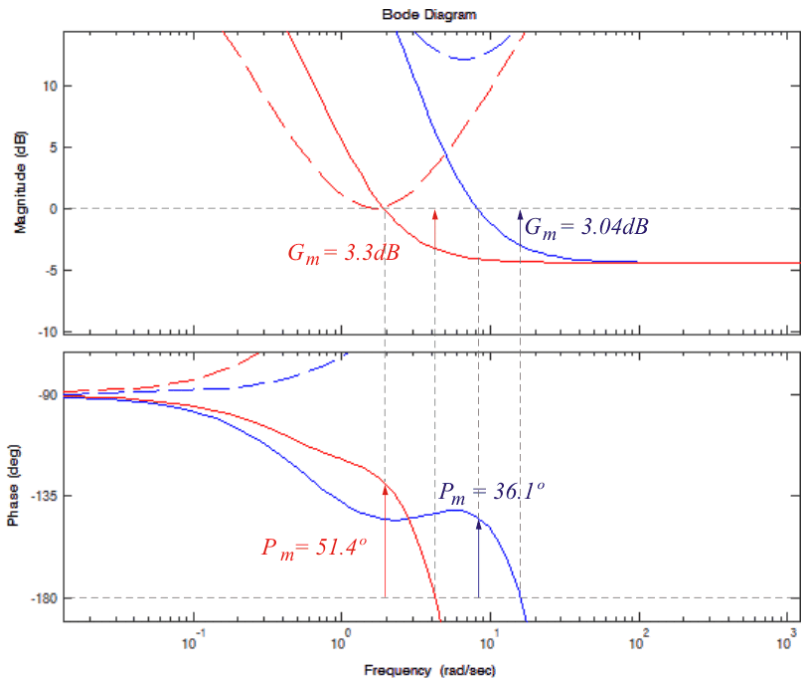


Fig. 31: Márgenes de ganancia y fase para dos procesos controlados con PID sintonizados con Ziegler-NIchols

5.2.1 Una metodología para sintonizar un controlador PID

De lo expuesto puede derivarse una metodología para realizar una puesta en funcionamiento de un controlador PID con seguridad. Debe tenerse en cuenta:

a. *Qué estructura de controlador se está usando*, y eventualmente realizar las conversiones correspondientes.
b. *Tener cuidado sobre la forma en que se pide la información de los parámetros*: Algunos controladores usan la ganancia integral K_i y otros el tiempo integral T_i.
c. *Prestar atención a las unidades de los parámetros usados por el controlador*: normalmente son ciclos/minuto para K_i y minutos para K_d. Esto significa que la escala de tiempo que debemos usar para estimar K_e / T_g debe estar en minutos.
d. *Método empírico para ajustar una sintonización*:
   La conclusión a la que se arribó en 5.2 sobre el movimiento del filtro rechazabanda hacia arriba o hacia abajo en las frecuencias sugiere una metodología para corregir sobre la marcha una sintonización que fue realizada por el método de Ziegler-Nichols y que no presenta las características que se esperaban.

   Sea por ejemplo un sistema que fue sintonizado siguiendo el procedimiento de Ziegler-Nichols para la respuesta al escalón y presenta una respuesta de lazo cerrado que es demasiado oscilatoria (más inestable de lo esperado). Es fácil concluir que se ha subestimado el valor del retardo puro, que es el que provoca la inestabilidad. Los pasos a seguir son los siguientes:

   1. Dado que la ganancia integral K_i tiene un comportamiento inverso con el retardo T_u, **disminuir un poco K_i hasta un valor K_i' .**

2. Nótese que en la sintonización de Ziegler-Nichols hay una relación fija entre K_i y K_d, la cual debe mantenerse para conservar el ancho de banda del filtro rechazabanda que constituye el controlador PID. Recuérdese de la Tabla 3 que

$$K_i = \frac{1}{2\ T_u}$$
$$K_d = 0.5\ T_u$$

Poniendo K_d en función de K_i se concluye que **debe fijarse la ganancia derivativa en**

$$K_d^{'} = \frac{0.25}{K_i^{'}}$$

3. Probar nuevamente el comportamiento del sistema. Si es necesario, disminuir un poco la ganancia proporcional K_c.

Obviamente, deberá procederse a la inversa si el comportamiento del lazo cerrado es demasiado amortiguado.

6. Variaciones del cálculo de Ziegler-Nichols

Hay una tendencia en la actualidad a no utilizar los parámetros estrictos de Ziegler-Nichols sino versiones un tanto más conservadoras en el sentido de una menor ganancia K_c y un ancho de banda un poco más amplio del filtro rechazabanda representado por el PID. Un cálculo modificado muy habitual en la práctica es, para el método de la ganancia límite por oscilación

$K_c = 0.5\ K_{cu}$ (en lugar de $0.6\ K_{cu}$)

$K_i = 1/T_i = 1/P_n$ (en lugar de $2/\ P_n$)

$K_d = Td = P_n/8$ (igual que en el cálculo estricto) (25)

De igual manera, para el cálculo por el ensayo al escalón:

$K_c = \dfrac{1}{(\frac{K_e}{T_g})T_u}$ (en lugar de $\dfrac{1.2}{(\frac{K_e}{T_g})T_u}$)

$K_i = \dfrac{1}{4\ T_u}$ (en lugar de $\dfrac{1}{2\ T_u}$) (26)

$K_d = 0.5\ T_u$ (igual que en el cálculo estricto)

El lector puede verificar que la relación entre uno y otro cálculo mantiene la dependencia ya conocida $P_n = 4T_u$. Debe prestarse atención al hacer el ajuste empírico descripto en la sección 5.2.1, una vez que se ajustó manualmente el valor K_i hasta un valor K_i', de usar

$$K_d' = \frac{0.125}{K_i'}$$

La Figura 32 muestra en azul a trazos el diagrama de Bode de un controlador PID calculado con Ziegler-Nichols estricto para un proceso caracterizado por K_e = 3; T_g = 1.5 seg ; T_u = 0.15 seg. En trazos rojos se muestra lo propio para un PID con el cálculo modificado de la expresión (26), percibiéndose una menor ganancia y un ancho de banda mayor. En líneas llenas, los diagramas del proceso controlado.

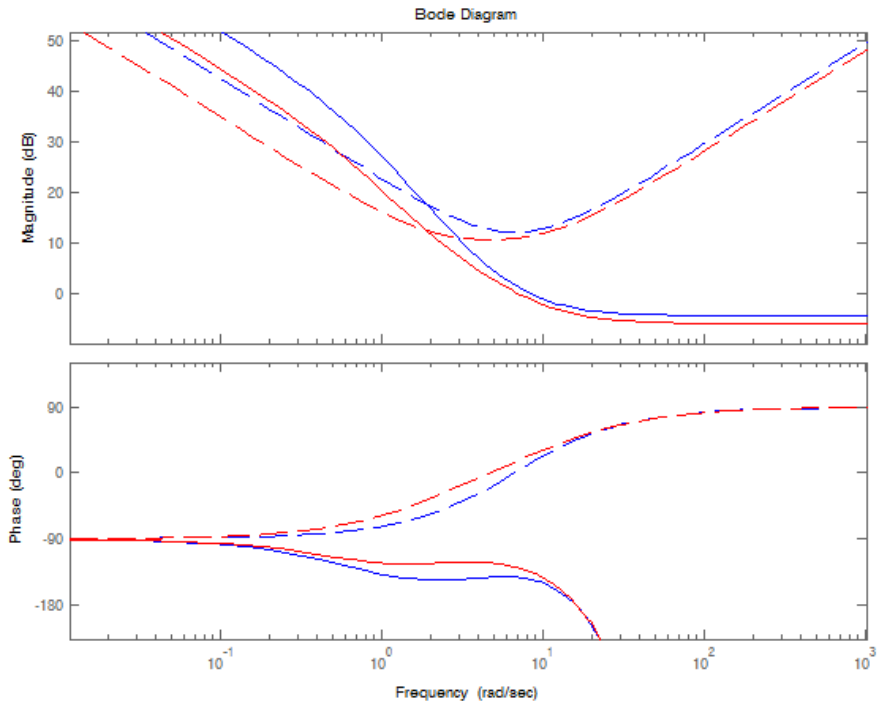


Figura 32: Diagramas de Bode de un PID sintonizado con Ziegler-Nichols estricto y modificado

7. Modificaciones del controlador PID

7.1 Filtro en la acción derivativa

El hecho de que el controlador PID (15)(16) tenga una ganancia creciente en altas frecuencias presenta el problema de la amplificación del ruido de medición, que es amplificado por el término derivativo. Para atenuar este efecto suele introducirse un filtro de primer orden en la acción derivativa:

$$\frac{U(s)}{E(s)} = K_c \left(1 + \frac{K_i}{s} + \frac{K_d\, s}{1 + \alpha K_d\, s}\right) \tag{27}$$

Siendo α < 1 el factor del filtro derivativo, que toma valores típicos entre 0.05 y 0.1. Se ve claramente en esta expresión que la ganancia del controlador en altas frecuencias será K_c / α. La Figura 33 muestra la comparación de un PID teórico (azul a trazos) con un PID filtrado según la ec. (27) (rojo a trazos). En líneas llenas, el mismo sistema con ambos controladores.

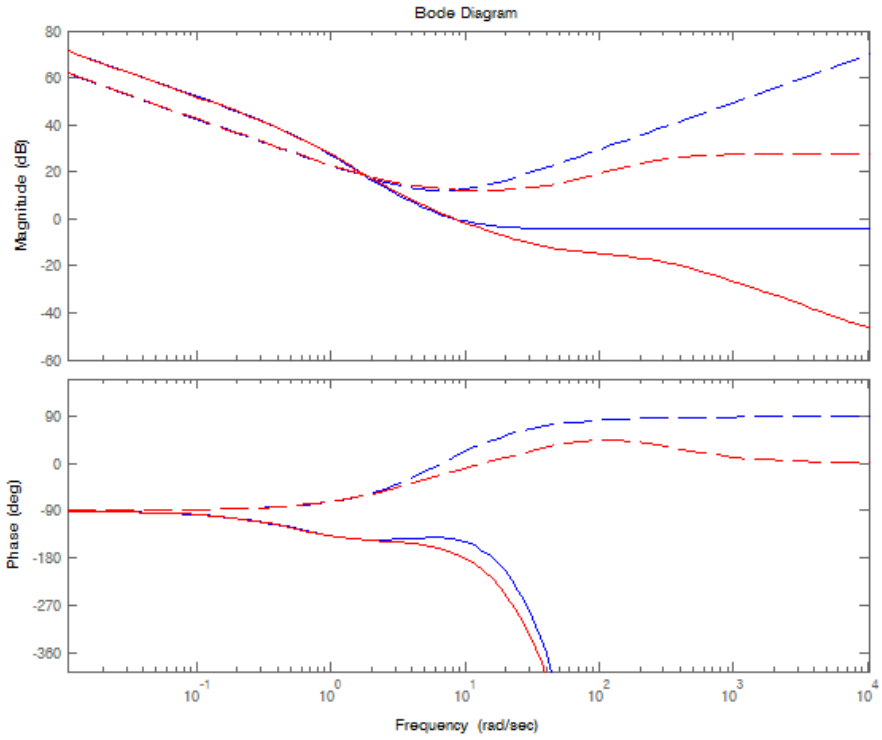


Figura 33: Comparación de un PID teórico con uno con filtrado derivativo.

Algunos fabricantes inclusive sustituyen el bloque derivativo por una red de adelanto de fase:

$$\frac{U(s)}{E(s)} = K_c \left(1 + \frac{K_i}{s} + \frac{1 + K_d\, s}{1 + \alpha K_d\, s}\right) \tag{28}$$

7.2 Antiwindup

Cuando se exigen al controlador grandes variaciones en la variables controlada, por ejemplo cuando se solicita un cambio de referencia brusco, ocurre que rápidamente el actuador llega a su nivel de saturación, pero el término integral del controlador sigue integrando, de forma tal que cuando se alcanza el Set Point, el controlador no logra reaccionar inmediatamente debido a dicho valor anteriormente integrado. El resultado es una sobreelongación mayor de la esperada, aun cuando el PID esté bien sintonizado. Este fenómeno, denominado *windup*, está graficado en la Fig. 34, en la cual a una variable medida *h2* se le solicita un cambio de referencia que provoca la saturación del actuador y la acumulación excesiva del término integral.

El problema ha recibido diversos tratamientos, todos trabajando sobre el término integral:

- Limitación del término integral
- Integración condicional
- Seguimiento integral: la mejor de todas: leyendo la posición del actuador o teniendo un modelo del mismo, se detecta cuando el mismo se ha saturado y se suspende la integración.

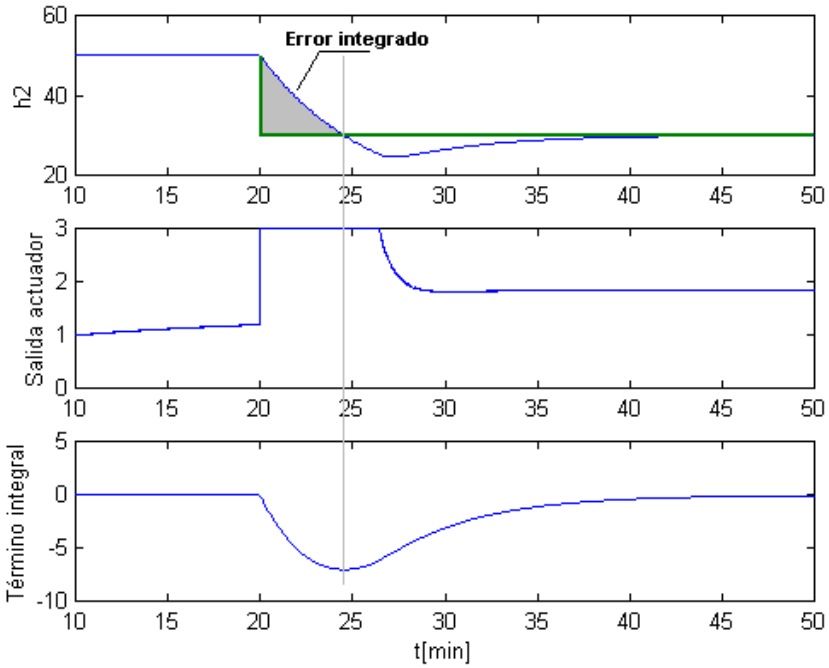


Figura 34: Windup por saturación del actuador.

7.3 Transferencia "*bumpless*"

Los controladores poseen una capacidad de paso de modo manual a automático y viceversa. En estas ocasiones ocurre algo semejante al problema del windup: durante el tiempo en que el operador maneja manualmente el sistema, el término integral sigue creciendo, y cuando se vuelve a modo automático se produce una muy brusca variación sobre el actuador y en consecuencia sobre la salida. Esto es solucionado por la capacidad de transferencia "bumpless", que produce un cambio gradual entre la salida en modo manual y el valor que es calculado en modo automático.

7.4 Autosintonía

Los controladores PID digitales más avanzados – y más caros, por cierto- poseen la capacidad de autosintonizarse, es decir de encontrar automáticamente los parámetros apropiados para el proceso en el cual está conectado. Para realizar esto cada fábrica de controladores tiene su propio sistema, que guarda celosamente, pero todos se reducen a la idea de encontrar un modelo del proceso, que normalmente se considera de primer orden con retardo puro o, en los más elaborados, de segundo orden.

Para obtener un modelo del sistema, el controlador debe excitar de alguna manera al proceso para ver su comportamiento. Esto en esencia es lo mismo que hace quien aplica un escalón para encontrar los parámetros por Ziegler-Nichols, o quien lo hace oscilar para encontrar la ganancia límite: se están encontrando parámetros que identifican al sistema. Para llevar a cabo esta identificación, algunos controladores, una vez estabilizado el proceso, lo colocan en lazo abierto y aplican un pequeño escalón en la acción de control, analizando a continuación la salida y corrigiendo los parámetros. Uno de los mejores procedimientos conocidos hasta el momento, el denominado “método del relé de Aström”, aplica una señal cuadrada en lugar de un escalón.

Referencias bibliográficas

Cohen, G. H. y Coon, G. A., 1953. Theoretical consideration of retarded control. Trans. Amer. Soc. Mech. Eng.; Vol. 76, 8p.

Ziegler, XX y Nichols, XX., 1942. Optimun Settings for Automatic Controllers, Transactions of ASME.

Astrom, K.J. y Hagglund, T.H., 1995. New tuning methods for PID controllers. Proc. 3rd European Control Conference, pp. 2456-2462.

Dormido Bencomo, S., Morilla García, F., 1995. Autosintonía y métodos de Antiwindup en los reguladores PID. Asociación Española de Informática y Automática.

Apéndice 1

Demostración de la dependencia de la frecuencia natural de oscilación ω_n del retardo puro T_u para los procesos modelados con un sistema de primer orden con retardo puro.

Los mismos responden a la conocida forma

$$G(s) = \frac{e^{-T_u s} \; K_e}{1 + T_g \, s} \tag{1}$$

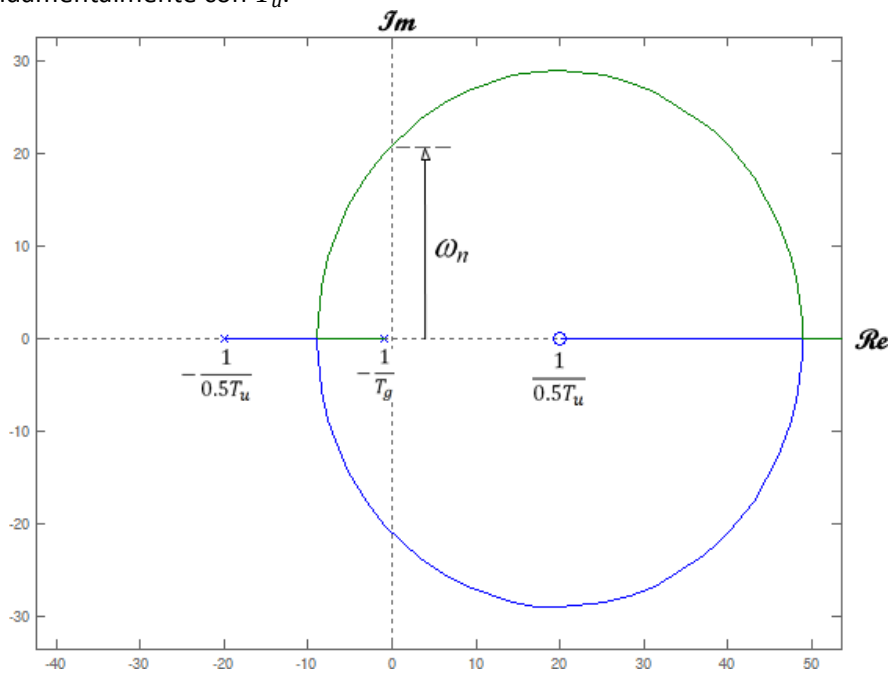
Aunque se obtiene un resultado mucho más preciso usando una aproximación de Padé de segundo orden para representar al retardo puro, por simplicidad en los cálculos se utilizará la aproximación de primer orden:

$$e^{-T_u s} \cong \frac{1 - \frac{T_u}{2} s}{1 + \frac{T_u}{2} s} \tag{2}$$

Reemplazando se aproxima $G(s)$ como

$$G(s) \cong \frac{(1 - \frac{T_u}{2} s)}{(1 + \frac{T_u}{2} s)} \frac{K_e}{(1 + T_g \, s)} \tag{3}$$

Antes de seguir adelante, es bueno tener en mente que este sistema tiene un lugar de las raíces como el de la siguiente figura, y que nuestro objetivo es encontrar el punto en el cual, aumentando la ganancia K_e, las raíces tocan el eje imaginario y el sistema se vuelve oscilatorio con una frecuencia natural ω_n, y la relación de este valor fundamentalmente con T_u.



De (3) se puede encontrar la ecuación característica del sistema:

$$2T_g T_u \; s^2 + (4T_g + 2T_u - 2\,T_u K_e)\, s + (4\,K_e + 4) = 0 \tag{4}$$

Las raíces de la misma son

$$r_{1,2} = \frac{-\left(2T_g + T_u \pm \sqrt{K_e^{\,2} T_u^{\,2} - 12 K_e T_g T_u - 2 K_e T_u^{\,2} + 4 T_g^{\,2} - 4 T_g T_u + T_u^{\,2}} - K_e T_u\right)}{2 T_g T_u} \tag{5}$$

Para que el sistema sea estable, todos los coeficientes de (4) deben ser positivos. Asumiendo que la ganancia del sistema K_e es positiva, el término independiente es positivo. El término lineal en s igualado a cero nos permite encontrar la ganancia K_{eLim} con la cual el sistema se hace oscilatorio:

$$4T_g + 2T_u - 2\,T_u K_e = 0 \tag{6}$$

Lo cual da

$$K_{eLim} = \frac{2\,T_g + T_u}{T_u} \tag{7}$$

Reemplazando el valor de K_{eLim} dado por (7) en (5) se obtendrán las raíces del sistema justo cuando las mismas se ubican sobre el eje imaginario; es decir que no tienen parte real y deben dar un par imaginario conjugado. Efectivamente:

$$r_{1,2\,lim} = \pm\, j\, \frac{2}{T_u} \sqrt{\frac{T_g + T_u}{T_g}} \tag{8}$$

Por supuesto, el valor absoluto nos da la frecuencia natural de oscilación

$$\omega_n = \frac{2}{T_u} \sqrt{\frac{T_g + T_u}{T_g}} \tag{9}$$

Si, como se ha supuesto, $T_u < T_g$, el radical de esta ecuación será un número levemente superior a la unidad y queda demostrado que ω_n depende fundamentalmente de T_u.

En esta demostración, en la que se ha recurrido por simplicidad a una aproximación de Padé de primer orden del retardo puro, el radical de (9) es aproximadamente π/3. Usando una aproximación de segundo orden, este valor es aproximadamente π/4, con lo que se demuestra lo que se había postulado empíricamente más arriba:

$$P_n \cong 4\,T_u \tag{10}$$