Towards Intelligent PID Control

Hang, Chang C.; Åström, Karl Johan; Persson, Per; Ho, Weng Kuen

Published in:
Automatica

1992

Automatica, Vol. 28, No. 1, pp. 1–9, 1992
Printed in Great Britain.

0005–1098/92 $5.00 + 0.00
Pergamon Press plc

Towards Intelligent PID Control*

K. J. ÅSTRÖM,†§, C. C. HANG,‡ P. PERSSON† and W. K. HO‡

Dimensionless numbers characterizing open-loop process dynamics and closed-loop PID behaviour are introduced and used in assessing performance and tuning, relations between these numbers are derived using analytical method and simulations

Key Words—Automatic tuning; conventional control; expert systems; industrial control; knowledge engineering; PID control; process control; threeterm control.

Abstract—Autotuners for PID controllers have been commercially available since 1981. These controllers automate some tasks normally performed by an instrument engineer. The autotuners include methods for extracting process dynamics from experiments and control design methods. They may be able to decide when to use PI or PID control. To make systems with a higher degree of automation it is desirable to also automate tasks normally performed by process engineers. To do so, it is necessary to provide the controllers with reasoning capabilities. This seems technically feasible with the increased computing power that is now available in single-loop controllers. This paper describes some features that may be included in the next generation of PID controllers.

1. INTRODUCTION

IN THE DESIGN of a knowledge-based feedback controller (Åström *et al.*, 1986; Årzén, 1989), it is desirable to incorporate the expert knowledge of design engineers so that the controller can make decisions on the choice of control algorithm and provide diagnostics on the effectiveness of the control system. For real-time implementation, it is desirable to have as much deep knowledge as possible. It is also desirable that the controller to a limited degree could explain its own reasoning, e.g. why derivative action was used. It should also be able to tell if PID control is appropriate and if not, suggest alternative control schemes.

This paper attempts to develop tools to assess what can be achieved by PID control of two broad classes of systems with a Ziegler–Nichols (1942) like tuning formula. Based on empirical studies and approximate analysis, we introduce two parameters, namely, the normalized deadtime and the normalized process gain to characterize the open-loop process dynamics. Two more parameters, the peak load error and the normalized rise time are also introduced to characterize the closed-loop response. Simple methods of measuring these parameters are proposed.

It is shown that the normalized deadtime and the normalized process gain can be used to predict the achievable performance of PID controllers tuned by the Ziegler–Nichols formula. Using these relations, the controller can interact with the operator and advise him on the choice of control algorithms. If desired, it can also make the choice automatically and explain its reasoning.

Useful relations, which can be used to assess whether the PID controller is properly tuned are also established. The relations give significant insight into the properties of PID control. The simplicity of the relations allows the development of a first generation of intelligent controllers using current technology.

The paper is organized as follows. Two classes of processes are introduced in Section 2. Some useful dimensionless numbers are introduced in Section 3. In Sections 4 and 5, some relations between the features are derived based on approximate analysis and simulations. The results are used in Section 6 to predict the performance that may be achieved with PID control based on Ziegler–Nichols tuning. Conclusions are presented in Section 7.

2. PROCESS CHARACTERISTICS

It is assumed that process dynamics can be described with sufficient accuracy by linear

* Received 1 September 1990; revised 28 April 1991; received in final form 19 June 1991. The original version of this paper was presented at the IFAC Workshop on AI in Real Time Control which was held in Shenjang, People's Republic of China during September 1989. The published proceedings of this IFAC Meeting may be ordered from: Pergamon Press plc, Headington Hill Hall, Oxford OX3 0BW, U.K. This paper was recommended for publication in revised form by Editor A. P. Sage.

† Department of Automatic Control, Lund Institute of Technology, Box 118, S-221 00 Lund, Sweden.
‡ Department of Electrical Engineering, National University of Singapore, 10 Kent Ridge Crescent, Singapore 0511.
§ Author to whom all correspondence should be addressed.

models. Additional constraints on system dynamics both in the time and the frequency domains are also introduced. This will lead to two classes of systems that are common in process control.

Time domain characterization

It will be assumed that the step response is monotone or essentially monotone, i.e. monotone except for a small initial part. Such processes can be divided into two broad classes. The first class corresponds to stable processes. Their dynamics can be characterized by the parameters K_p, L and T, where K_p is the static process gain, L is the apparent deadtime and T is the apparent time constant. These parameters can be obtained from a step response experiment (Fig. 1). The transfer function

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

is a crude analytical approximation of stable processes.

The second class corresponds to processes with integral action. The transfer function

$$G(s)=\frac{K_v}{s(1+sT_v)}\mathrm{e}^{-sL} \tag{2}$$

is a crude analytical approximation of such a process. An even simpler model is

$$G(s)=\frac{K_v}{s}\mathrm{e}^{-sL}. \tag{3}$$

Notice that the transfer function (3) may be regarded as the limit of (1) when K_p and T go to infinity in such a way that $K_p/T=\text{constant}=K_v$.

The systems considered were used in the classical works on Ziegler–Nichols tuning. In traditional process control literature the classes of systems are called processes with self-regulation (stable process) and processes without self-regulation (processes with integration). Notice that systems with resonant poles are not considered, such processes do not have essentially monotone step response.

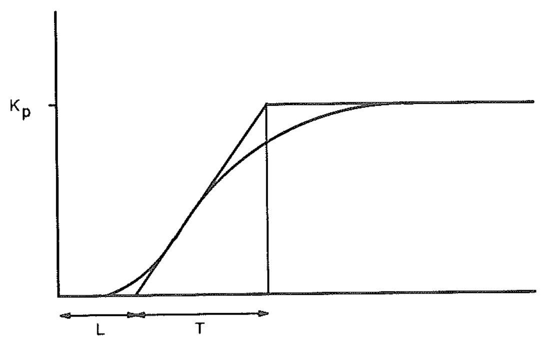

FIG. 1. Graphical determination of mathematical model for a stable process.

Frequency domain characterization

A frequency domain characterization of process dynamics will also be introduced. It is assumed that the Nyquist curve is monotone or essentially monotone, i.e. both the phase and the amplitude are monotone functions of frequency. This condition also excludes processes with resonances. The main difference between the two classes of processes is that at zero frequency, stable processes have finite gain whereas processes with integral action have infinite gain. Both systems can be characterized by the first intersection of the Nyquist curve with the negative real axis. This defines the ultimate gain, K_u, i.e. the controller gain that makes the process unstable under proportional feedback, and the ultimate period T_u. Lack of monotonicity can be accepted at frequencies where the phase lag is larger than 180°.

3. FEATURES

Dimension-free parameters, like Reynold's numbers, have found much use in many branches of engineering. They have, however, not been much used in automatic control. This section attempts to introduce some dimension-free numbers that are useful in assessing control system performance.

Normalized deadtime

A normalized deadtime can be defined for stable processes as the ratio of the apparent deadtime to the apparent time constant

$$\theta_1=\frac{L}{T}. \tag{4}$$

This number can be estimated from a record of the step response. It has been known from practical experience that the normalized deadtime can be used as a measure of the difficulty in controlling a process. Processes with a small θ_1 are easy to control and processes with a large θ_1 are difficult to control. Fertik (1975) introduced the name process controllability for the related quantity $\theta_1/(1+\theta_1)$. Parameter θ_1 was actually called the controllability ratio by Deshpande and Ash (1981). To avoid possible confusion with the standard terminology of modern control theory, the words normalized deadtime will be used.

A ratio analogous to equation (4) can be introduced for processes with integration. Let $G(s)$ be the transfer function of such a process. The transfer function $sG(s)$ belongs to the class of stable processes. Its behaviour can then be characterized by an apparent deadtime, L and an apparent time constant, T_v. The normalized deadtime for processes with integration

$$\theta_2 = \frac{L}{T_v} \tag{5}$$

can then be introduced.

Since the transfer function (3) may be regarded as the limit of the transfer function (1), processes with integration may be considered as a special case of stable processes with very small values of normalized deadtime.

Normalized process gain

Parameter K_p is the process gain for stable processes. It is not necessarily dimension-free. It can however be made dimension-free by multiplying with a factor. The ultimate gain, K_u, is a suitable normalization factor. The normalized process gain, κ_1, is defined as

$$\kappa_1 = \frac{G(0)}{|G(i\omega_u)|} = K_p K_u \tag{6}$$

where $G(s)$ is the plant transfer function and ω_u is the smallest frequency such that

$$\arg G(i\omega_u) = -\pi.$$

This number κ_1 is easily obtained from the Nyquist curve. It also has a physical interpretation as the largest process loop gain that can be achieved under proportional control. This number is useful for assessing control performance. Roughly speaking, a large value indicates that the process is easy to control while a small value indicates that the process is difficult to control.

Stable processes have a steady-state error under proportional feedback. The error obtained for a step command of size s_0 is

$$e_s = \frac{1}{1 + K_p K_c} s_0 > \frac{1}{1 + \kappa_1} s_0 \tag{7}$$

where K_c is the proportional gain used. The inequality follows because $K_p K_c < \kappa_1$. The number κ_1 can thus be used to estimate the minimum steady-state error under proportional control and also to determine if integral action is required in order to satisfy specifications on static error.

For processes with integration, the product $K_u K_v$ has dimension frequency. The dimension-free process gain therefore has to be defined differently. The normalized process gain κ_2 for processes with integration is defined as

$$\kappa_2 = \frac{\lim_{s\to 0} sG(s)}{\omega_u\,|G(i\omega_u)|} = \frac{K_v K_u}{\omega_u} = \frac{K_v K_u T_u}{2\pi} \tag{8}$$

where $G(s)$ is the plant transfer function and ω_u is the smallest frequency such that

$$\arg G(i\omega_u) = -\pi.$$

With constant set-point and no load disturbance, processes with integration will not have a steady-state error under proportional feedback. With a ramp set-point of velocity v_0, the steady-state error is

$$e_v = \frac{1}{K_v K_c} v_0 > \frac{1}{\kappa_2 \omega_u} v_0.$$

Load disturbance error

The response to step load disturbance is an important factor when evaluating control systems. The effect of a load disturbance depends on where the disturbance acts on the system. It will be assumed that the disturbance acts on the process input. With proportional control a step load disturbance of magnitude l_0 gives the static error

$$l_1 = \frac{K_p l_0}{1 + K_p K_c} > \frac{K_p l_0}{1 + \kappa_1} \tag{9}$$

for stable processes and

$$l_2 = \frac{l_0}{K_c} = \frac{K_v \omega_u l_0}{K_v \omega_u K_c} > \frac{K_v l_0}{\omega_u \kappa_2} = \frac{K_v T_u l_0}{2\pi\kappa_2} \tag{10}$$

for processes with integration. The quantities $l_1/(K_p l_0)$ and $l_2\omega_u/(K_v l_0)$ are therefore dimension-free.

When a controller with integral action is used the static error due to a step load disturbance is zero. A meaningful measure is then the maximum error. To obtain a dimension-free quantity for stable processes, the maximum error is divided by K_p. Thus the peak load error for stable processes, λ_1, can be defined as

$$\lambda_1 = \frac{1}{K_p l_0} l_{\max} \tag{11}$$

where l_0 is the amplitude of the step load disturbance and $l_{\max}$ the maximum error due to the step load disturbance. For processes with integration the corresponding quantity is

$$\lambda_2 = \frac{\omega_u}{K_v l_0} l_{\max} = \frac{2\pi}{K_v T_u l_0} l_{\max}. \tag{12}$$

Normalized closed-loop rise time

The closed-loop rise time, t_r, is a measure of the response speed of the closed-loop system. To obtain a dimension-free parameter it will be normalized by the apparent deadtime, L, of the open-loop system. Thus the normalized rise time, τ, for both classes of processes are defined as

$$\tau = \frac{t_r}{L}. \tag{13}$$

4. EMPIRICS

The Ziegler–Nichols closed-loop tuning procedure has been applied to a large number of different processes and it has been attempted to correlate the observed properties of the open-loop and closed-loop systems to the features introduced in Section 3. Results for processes with the transfer functions

$$G_1(s)=\frac{e^{-sL}}{(1+s)^2}\quad 0.1\leq L\leq 3 \tag{14}$$

$$G_2(s)=\frac{1}{(1+s)^n}\quad 3\leq n\leq 20 \tag{15}$$

$$G_3(s)=\frac{1-\alpha s}{(1+s)^3}\quad 0\leq \alpha\leq 2 \tag{16}$$

will be used to illustrate the features of stable processes and

$$G_4(s)=\frac{e^{-sL}}{s}\quad 0<L \tag{17}$$

$$G_5(s)=\frac{e^{-sL}}{s(s+1)}\quad 0.05\leq L\leq 0.8 \tag{18}$$

$$G_6(s)=\frac{1-\alpha s}{s(s+1)}\quad 0.05\leq \alpha\leq 0.75 \tag{19}$$

for unstable processes. These models cover a wide range of dynamic characteristics found in typical process control applications.

The PID controller used had the form

$$u_c=K_c\left(\beta y_r-y+\frac{1}{T_i}\int e\,dt-T_d\frac{dy_f}{dt}\right) \tag{20}$$

$$e=y_r-y \tag{21}$$

$$Y_f(s)=\frac{1}{1+sT_d/N}Y(s) \tag{22}$$

where u_c, y and y_r are the controller output, process output and set-point respectively. The noise filtering constant, N, is usually in the range of 3 to 20. The value $N=10$ was used in the simulations. In conventional PID controllers parameter $\beta=1$. It is often advantageous to have a smaller value of β to reduce the overshoot to set point changes. This is called set point weighting (Åström and Hägglund, 1988). The value $\beta=1$ was used in all simulations. Parameters of the PID controller were determined by straightforward Ziegler–Nichols closed-loop tuning.

The results obtained are summarized in Tables 1–3 for stable processes and Tables 4–6 for processes with integration. The overshoot "os" and undershoot "us", of the closed-loop step response are also given.

5. RELATIONS

Normalized deadtimes, θ_1 and θ_2, and the normalized process gains, κ_1 and κ_2, have been introduced to characterize the open-loop dynamics. Peak load errors, λ_1 and λ_2, and normalized closed-loop rise time, τ, have been introduced to characterize the closed-loop response. Relations between these numbers will now be established. In doing so an intuitive interpretation of the numbers will also be developed.

Normalized deadtime and normalized process gain

From Tables 1–3, there appears to be a relation between the normalized deadtime, θ_1, and the normalized process gain, κ_1. For specific processes it is possible to find the relation exactly (see Åström *et al.*, 1988). For example, for first order processes with deadtime we have

$$\theta_1=\frac{L}{T}=\frac{\pi-\operatorname{atan}\sqrt{\kappa_1^2-1}}{\sqrt{\kappa_1^2-1}}. \tag{23}$$

Similar expressions can also be derived for the processes given by equations (14), (15) and (16). The relations are shown in Fig. 2. It can be approximated by the expression:

$$\kappa_1=2\left(\frac{11\theta_1+13}{37\theta_1-4}\right). \tag{24}$$

(See Hang *et al.*, 1991.) This formula is important because it means that the normalized deadtime, θ_1, and the normalized process gain, κ_1, can be used interchangeably to assess process dynamics.

A relation analogous to equation (23) can be derived for processes with integration. Let $G(s)$

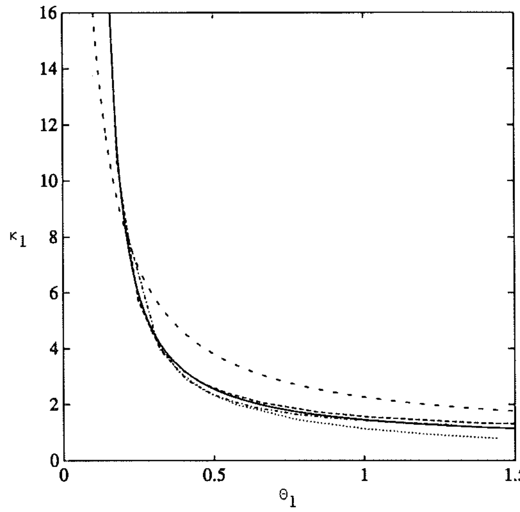

FIG. 2. Relation between θ_1 and κ_1. ——— First order process with deadtime, - - - process of (14), —•— process of (15), • • • process of (16), —— approximation.

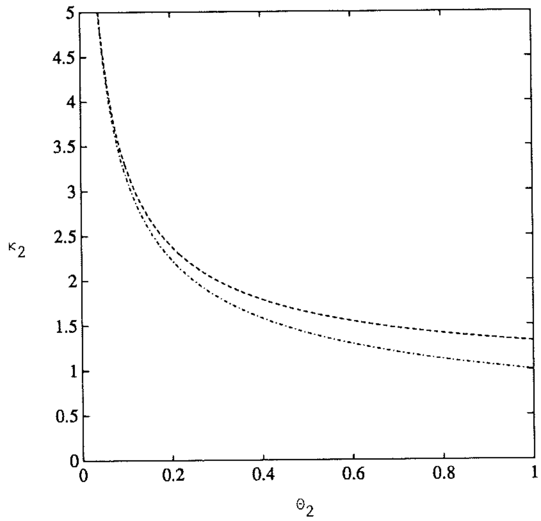

FIG. 3. Relation between θ_2 and κ_2. --- process of (18), —·— process of (19).

be the transfer function of a process with integration. The transfer function $sG(s)$ is of the class of stable processes and can be characterized by an apparent deadtime, L and an apparent time constant, T_v. For a transfer function given by (2), it can be shown that

$$\theta_2 = \frac{L}{T_v} = \frac{\pi/2 - \operatorname{atan}\sqrt{\kappa_2^2 - 1}}{\sqrt{\kappa_2^2 - 1}}. \tag{25}$$

The calculation is the same as in Åström *et al.* (1988). This means that κ_2 can be used as a measure of the normalized deadtime, θ_2 (Fig. 3).

Normalized peak load error

Consider the closed-loop system obtained with the controller $G_c(s)$ and the process $G_p(s)$. Assume that the disturbance enters at the plant input. The transfer function from the load disturbance to the output is

$$G_d(s) = \frac{1}{G_c(s)} \frac{G_p(s)G_c(s)}{1 + G_p(s)G_c(s)}. \tag{26}$$

A PID controller with Ziegler–Nichols tuning has the transfer function

$$G_c(s) = \frac{K_c(s + \alpha)^2}{2\alpha s}$$

and

$$\alpha = \frac{1}{2T_d} = \frac{4}{T_u}.$$

This choice gives good rejection of load disturbances as discussed in Hang (1989). With Ziegler–Nichols tuning the closed-loop system has a bandwidth $\omega \approx 7.4/T_u$ (Åström and Hägglund, 1988). For frequencies less than ω, the transfer function (26) can be approximated by

$$G_d(s) \approx \frac{1}{G_c(s)} = \frac{2\alpha s}{K_c(s + \alpha)^2}.$$

The response to a step of size l_0 is

$$l(t) = \frac{2\alpha t l_0}{K_c} e^{-\alpha t}.$$

It has a maximum

$$l_{\max} = \frac{2l_0}{eK_c} = \frac{0.74 l_0}{K_c}$$

at

$$t = 2T_d.$$

The normalized peak load for stable processes is thus

$$\lambda_1 = \frac{1}{l_0 K_p} l_{\max} = \frac{0.74}{K_c K_p} = \frac{1.23}{\kappa_1}. \tag{27}$$

The product $\kappa_1\lambda_1$ can thus be expected to be constant. An analogous calculation for processes with integral action shows that $\kappa_2\lambda_2$ can also be expected to be constant. This is supported by the experimental results given in Tables 1–6. Consider the range where Ziegler–Nichols tuning is applicable, i.e. $0.15 < \theta_1 < 0.6$ or $2.25 < \kappa_1 < 15$ for stable processes (Hang *et al.*, 1991), and $\theta_2 > 0.3$ or $\kappa_2 < 2$ for processes with integration. Tables 1–3 show that $\kappa_1\lambda_1$ ranges from 1.2 to 1.6 and Tables 4–6 show that $\kappa_2\lambda_2$ is close to 1.8. The following empirical relations thus hold approximately

$$\kappa_1\lambda_1 = 1.4 \tag{28}$$

$$\kappa_2\lambda_2 = 1.8. \tag{29}$$

It can also be shown that

$$\kappa_1\lambda_1 = \kappa_2\lambda_2 = \frac{K_u}{l_0} l_{\max}. \tag{30}$$

Closed-loop rise time

The experimental results given in Tables 1–6 show that the normalized rise time for Ziegler–Nichols tuning is approximately constant. In the range where Ziegler–Nichols is applicable, parameter τ ranges from 0.8 to 1.2 for stable processes and 0.4 to 0.7 for processes with integration. The following empirical relations thus hold

$$\tau \approx 1 \text{ for stable processes} \tag{31}$$

$$\tau \approx 0.5 \text{ for processes with integration.} \tag{32}$$

This means that the actual rise time with Ziegler–Nichols tuning is approximately equal to the apparent deadtime for stable processes and

TABLE 1. EXPERIMENTAL RESULTS FOR A SYSTEM WITH $G(s) = e^{-sL}/(s+1)^2$

L	T_u	os	us	θ_1	κ_1	τ	λ_1	$\kappa_1\lambda_1$
0.1	1.4	68	26	0.14	20.5	0.81	0.08	1.6
0.2	2.0	61	16	0.18	10.5	1.04	0.15	1.6
0.4	2.9	49	6	0.25	5.7	0.95	0.27	1.5
0.6	3.6	41	2	0.32	4.0	0.93	0.36	1.4
1.0	4.8	30	5	0.47	2.7	0.88	0.51	1.4
1.5	6.1	22	12	0.66	2.1	0.82	0.64	1.3
2.0	7.3	16	18	0.84	1.74	0.78	0.73	1.3
2.5	8.5	13	23	1.02	1.55	0.75	0.80	1.2
3.0	9.5	12	26	1.21	1.44	0.71	0.86	1.2

TABLE 2. EXPERIMENTAL RESULTS FOR A SYSTEM WITH $G(s) = 1/(s+1)^n$

n	T_u	os	us	θ_1	κ_1	τ	λ_1	$\kappa_1\lambda_1$
3	3.6	52	14	0.22	8.0	1.16	0.19	1.5
4	6.3	38	7	0.32	4.0	1.22	0.35	1.4
6	10.9	23	10	0.49	2.4	1.11	0.54	1.3
8	15.2	17	16	0.64	1.88	1.01	0.65	1.2
10	19.3	13	20	0.77	1.65	0.93	0.73	1.2
15	29.5	9	27	1.05	1.39	0.82	0.85	1.2
20	39.7	6	31	1.29	1.28	0.76	0.91	1.2

TABLE 3. EXPERIMENTAL RESULTS FOR A SYSTEM WITH $G(s) = (1-\alpha s)/(s+1)^3$

α	T_u	os	us	θ_1	κ_1	τ	λ_1	$\kappa_1\lambda_1$
0	3.6	52	15	0.22	8.0	1.16	0.19	1.5
0.1	4.1	48	10	0.25	6.2	1.09	0.24	1.5
0.25	4.6	42	6	0.29	4.6	1.09	0.31	1.4
0.5	5.3	33	3	0.38	3.2	1.16	0.44	1.4
1.0	6.3	18	4	0.57	2.0	0.98	0.67	1.3
1.5	6.9	4	12	0.78	1.45	0.89	0.90	1.3
2.0	7.4	−7	22	1.0	1.14	0.84	1.08	1.2

TABLE 4. EXPERIMENTAL RESULTS FOR A SYSTEM WITH $G(s) = e^{-sL}/s$

T_u	os	us	θ_2	κ_2	τ	λ_2	$\kappa_2\lambda_2$
4L	71	12	∞	1	0.69	1.82	1.8

Notice that changes in L only change the time scale. Hence, except for T_u, the rest of the values in the table do not change with L.

half the apparent deadtime for processes with integration. For stable processes, Ziegler–Nichols tuning gives a closed-loop response with the rise time equals to the deadtime. This explains why the tuning rule does not work well for processes with normalized deadtimes larger than one.

6. ZIEGLER–NICHOLS TUNING

The results obtained will now be combined with previous experience to evaluate Ziegler–Nichols tuning. First notice that Ziegler–Nichols tuning is very simple. It is based on a simple characterization of the process dynamics from the step response or from the critical point on the Nyquist curve. The Ziegler–Nichols tuning rules or modifications of them are also commonly used in the industry.

When can Ziegler–Nichols tuning be used?

The results obtained show that Ziegler–Nichols tuning gives good results under certain conditions which can be characterized by θ_1 or κ_1 for stable processes and, θ_2 or κ_2 for

TABLE 5. EXPERIMENTAL RESULTS FOR A SYSTEM WITH $G(s) = e^{-sL}/s(s+1)$

L	T_u	os	us	$\theta_2 = L$	κ_2	τ	λ_2	$\kappa_2\lambda_2$
0.05	1.4	79	44	0.05	4.6	0.31	0.39	1.8
0.1	2.0	78	37	0.1	3.3	0.39	0.54	1.8
0.2	2.9	76	30	0.2	2.4	0.42	0.74	1.8
0.4	4.2	74	21	0.4	1.80	0.63	0.98	1.8
0.6	5.3	73	16	0.6	1.54	0.68	1.14	1.8
0.8	6.4	72	13	0.8	1.41	0.69	1.25	1.8

TABLE 6. EXPERIMENTAL RESULTS FOR A SYSTEM WITH $G(s) = (1 - \alpha s)/s(1 + s)$

α	$T_u = 2\pi\sqrt{\alpha}$	os	us	$\theta_2 = \alpha$	$\kappa_2 = \dfrac{1}{\sqrt{\alpha}}$	τ	λ_2	$\kappa_2\lambda_2$
0.05	1.4	80	44	0.05	4.5	0.26	0.40	1.8
0.1	2.0	79	38	0.1	3.2	0.34	0.57	1.8
0.25	3.1	79	29	0.25	2.0	0.45	0.90	1.8
0.5	4.4	79	23	0.50	1.41	0.50	1.28	1.8
0.75	5.4	79	21	0.75	1.15	0.51	1.56	1.8

TABLE 7. CHOICE OF CONTROLLER

		Tight control is required		
	Tight control is not required	High measurement noise	Low saturation limit	Low measurement noise and high saturation limit
1. ($\theta_1 > 1$; $\kappa_1 < 1.5$)	I	I + B + C	PI + B + C	PI + B + D
2. ($0.6 < \theta_1 < 1$; $1.5 < \kappa_1 < 2.25$)	I or PI	I + A	PI + A	PI + A + C or PID + A + C
3. ($0.15 < \theta_1 < 0.6$; $2.25 < \kappa_1 < 15$)	PI	PI	PI or PID	PID
4. ($\theta_1 < 0.15$; $\kappa_1 > 15$ $\theta_2 > 0.3$; $\kappa_2 < 2$)	P or PI	PI	PI or PID	PI or PID
5. ($\theta_2 < 0.3$; $\kappa_2 > 2$)	PD + E	F	PD + E	PD + E

A = Feedforward compensation recommended, B = feedforward compensation essential, C = deadtime compensation recommended, D = deadtime compensation essential, E = setpoint weighting necessary, F = pole placement tuning.

processes with integration. The conditions are summarized in Table 7. Five cases are introduced in the table. Cases 1–4 are applicable to stable processes while cases 4–5 are applicable to processes with integration. They are classified as follows.

Case 1 (Stable processes with $\theta_1 > 1$, $\kappa_1 < 1.5$): PID control based on Ziegler–Nichols tuning is not recommended when normalized deadtime θ_1 is larger than 1. This is partly due to inherent limitations of PID controllers and partly due to the Ziegler–Nichols tuning procedure. Modified tuning rules to deal with this case were proposed by Cohen and Coon (1953), Hang *et al.* (1991) and Hägglund (1991). By choosing other tuning methods, it is possible to tune PID controllers to work satisfactorily even for $\theta_1 = 10$; see Åström (1991). Also notice that feedforward control can be very beneficial since tight feedback cannot be obtained for processes with large deadtimes.

Case 2 (Stable processes with $0.6 < \theta_1 < 1$, $1.5 < \kappa_1 < 2.25$). Although the normalized deadtime is smaller than in case 1, Ziegler–Nichols tuning still gives poor results. This is easy to understand if we recall that the tuning procedure tries to make closed-loop rise time equals to the apparent deadtime. Other tuning methods and other controller structures like Smith predictors, and pole-placement could be considered (Åström and Hägglund, 1988). Feedforward can also be beneficial.

Case 3 Stable processes with $0.15 < \theta_1 < 0.6$, $2.25 < \kappa_1 < 15$). This is the prime application area for PID controllers with Ziegler–Nichols tuning. Derivative action often improves performance significantly.

Case 4 (Stable processes with $\theta_1 < 0.15$, $\kappa_1 > 15$; processes with integration having $\theta_2 > 0.3$, $\kappa_2 < 2$). Processes with integration may be viewed as the limit of stable processes when the normalized deadtime, θ_1, goes to zero. A process with integration may be regarded as an approximation of a stable process with small apparent deadtime but a large apparent time constant. The gross behaviour of processes with integration can be characterized by θ_2, the normalized deadtime obtained when the integrator or the large time constant is removed. The behaviour of such processes can also be characterized by the normalized process gain κ_2.

Tables 4, 5 and 6 indicate that Ziegler–Nichols tuning gives systems with good damping provided that θ_2 is large or equivalently κ_2 small i.e. the dynamics with the integrator removed is deadtime dominant. This is not surprising since the Ziegler–Nichols rules were derived for processes of this type. Notice, however, that the overshoot is in general too large. Set point weighting (Åström and Hägglund, 1988; Hang *et* *al.*, 1991) is thus essential for processes of this category.

Controllers with high gains can be used for this case. For systems with moderate requirements, PI or even P control may be sufficient. With high gain, measurement noise becomes an important issue since it can result in saturation of the control signal. In some cases performance can be increased significantly by using derivative action or even more complicated control laws. Temperature control where the dynamics is dominated by one large time constant is a typical case.

Case 5 (Processes with integration having $\theta_2 < 0.3$, $\kappa_2 > 2$). This case corresponds to an integrator with an additional dynamics that is lag dominated. PD control can be used for this type of processes. PID control with Ziegler–Nichols tuning does not give good results; both damping and overshoot are unsatisfactory. Design based on direct pole placement is recommended (Åström and Hägglund, 1988). Derivative action is often essential for good performance.

Implications for intelligent controllers

There are several simple autotuners that are based on the Ziegler–Nichols tuning procedure. A drawback with them is that they are unable to reason about the achievable performance. The result of this paper indicates that performance can be predicted from knowledge of parameters θ_1 or κ_1 for stable processes and θ_2 or κ_2 for processes with integration. Furthermore, it is easy to select the controller form: P, PI, PD or PID based on Table 7 which also indicates if a more sophisticated control law would be useful.

An autotuner based on the transient response method can give θ_1 from its measurements of T and L. For the correlation-based autotuner (Hang and Sin, 1988), θ_1 and κ_1 or θ_2 and κ_2 are readily computed from the impulse response generated by the correlator. For the relay-based autotuner (Åström and Hägglund, 1984) one additional measurement has to be made to determine κ_1 or κ_2, the static gain for stable processes or the integrator gain for processes with integration. These gains can be determined in closed-loop by introducing a small setpoint change. For stable processes, the static gain and the sum of deadtime and time constants can be computed using the method of moments (Åström and Hägglund, 1988). The normalized deadtime, θ_1, can then be computed. A check on the computed values of θ_1 and κ_1 can be made using equation (24).

Hang *et al.* (1991) have used θ_1 and κ_1 to improve on the Ziegler–Nichols tuning formula for stable processes. The following modifications, expressed as a simple function of θ_1 and κ_1, have been recommended to eliminate manual fine tuning. For PID control, when $\theta_1 < 0.6$ or $\kappa_1 > 2.25$, the main drawback of the Ziegler–Nichols formula is excessive overshoot and hence setpoint weighting is recommended. When $\theta_1 > 0.6$ or $\kappa_1 < 2.25$ the integral time computed by the Ziegler–Nichols formula is too large and should be reduced. For PI control, both the controller gain and the integral time have to be modified. The refined tuning formula is given in Appendix I. These modifications are essential for obtaining high quality PID or PI control without manual fine tuning. Tuning formulas for processes with integration are given in Ho (1990).

On-line assessment of control performance

The results of this paper can also be used to evaluate performance of feedback loops under closed-loop operation. Consider the relations (31) and (32) for the normalized rise time. The rise time can be measured when the setpoint is changed. If the controller is properly tuned then the closed-loop rise time should be equal to the apparent deadtime for stable processes and half of the apparent deadtime for processes with integration. If the actual rise time is significantly different, for instance 50% larger, it indicates that the loop is poorly tuned. This is particularly useful for assessing whether the control is too sluggish. Similarly the relations (28) and (29) can be used by introducing a perturbation at the controller output. A maximum error that is significantly larger than that predicted by equation (28) and equation (29) indicates that the loop is poorly tuned.

7. CONCLUSIONS

In this paper it has been attempted to analyze simple feedback loops with PID controllers that are tuned using the Ziegler–Nichols closed-loop method. It has been shown that for processes with essentially monotone step responses there exist quantities that are useful for assessing the achievable performance and selecting suitable controllers. These quantities are the normalized deadtimes, θ_1 and θ_2, the normalized process gains, κ_1 and κ_2, the peak load errors, λ_1 and λ_2, and the normalized closed-loop rise time, τ. Simple methods determine these parameters have also been suggested.

It has been shown that θ_1 is related to κ_1 and θ_2 to κ_2 and that κ_i and θ_i can be used interchangeably to assess the control problem. There are also relations like $\tau \approx 1$ and $\kappa_1\lambda_1 \approx 1.4$ for stable processes, and $\tau \approx 0.5$ and $\kappa_2\lambda_2 \approx 1.8$

for processes with integration. These relations may be used to assess the closed-loop response time and the load rejection properties. The results indicate that it would be useful to determine at least θ_1 or κ_1 for stable processes and θ_2 or κ_2 for processes with integration when tuning controllers because these parameters can be used to predict the achievable performance.

Acknowledgement—This work has been supported by the Swedish Board for Technical Development (STU) under contract DUP 85-3084P.

REFERENCES

Årzén, K. E. (1989). An architecture for expert system based feedback control. *Automatica,* **25,** 813–827.
Åström, K. J. (1991). Assessment of achievable performance of simple feedback loops. *Int. J. Adaptive Control Signal Process,* **5,** 3–19.
Åström, K. J., J. J. Anton and K. E. Årzén (1986). Expert control. *Automatica,* **22,** 227–286.
Åström, K. J. and T. Hägglund (1984). Automatic tuning of simple regulators with specifications on phase and amplitude margins. *Automatica,* **20,** 645–651.
Åström, K. J. and T. Hägglund (1988). Automatic tuning of PID controllers. ISA, Research Triangle Park, NC, USA.
Åström, K. J., C. C. Hang and P. Persson (1988). Heuristics for assessment of PID control with Ziegler–Nichols tuning. Report CODEN: LUTFD2/TFRT-7404, Department of Automatic Control, Lund Institute of Technology, Lund, Sweden.
Åström, K. J., C. C. Hang and P. Persson (1989). Towards intelligent PID control. In preprints *IFAC Workshop on AI in Real Time Control,* Shenyang, P. R. China, pp. 38–43, Pergamon Press, Oxford.
Bristol, E. H. (1977). Pattern recognition: an alternative to parameter identification in adaptive control. *Automatica,* **13,** 197–202.
Cohen, G. H. and G. A. Coon (1953). Theoretical consideration of retarded control. *Trans. ASME,* **75,** 827–834.
Deshpande, P. B. and R. H. Ash (1981). *Computer Process Control,* ISA, Research Triangle Park, U.S.A.
Fertik, H. A. (1975). Tuning controllers for noisy processes. *ISA Trans.,* **14,** 4.
Goff, K. W. (1966). Dynamics in direct digital control, parts I and II. *ISA J.,* **13,** 44–54.
Hägglund, T. (1991). A deadtime compensating three-term controller. 9th IFAC/IFORS Symposium on Identification and System Parameter Estimation, Budapest.
Hang, C. C. (1989). "Controller Zeros", *IEEE Control Systems Magazine,* **9,** 72–75.
Hang, C. C., K. J. Åström and W. K. Ho (1991). Refinements of the Ziegler–Nichols Tuning Formula. *Proc. IEE, Pt. D,* **138,** no 2, 111–118.
Hang, C. C. and K. K. Sin (1988). On-line auto-tuning of PID controllers based on cross correlation. *Proc. Int. Conf. on Industrial Electronics,* Singapore, 441–446.
Ho, W. K. (1990). Tuning of PI controllers for processes with integration based on gain and phase margin specifications. Report CODEN: LUTFD2/TFRT-7472. Department of Automatic Control, Lund Institute of Technology, Lund, Sweden.
Kraus, T. W. and T. J. Myron (1984). Self-tuning PID controller uses pattern recognition approach. *Control Engng,* June, 108–111.
Ziegler, J. G. and N. B. Nichols (1942). Optimum settings for automatic controllers. *Trans. ASME,* **64,** 759–768.

APPENDIX: REFINED ZIEGLER–NICHOLS TUNING FORMULA

For a PID controller of the form

$$u_c = K_p\left(\beta y_r - y + \frac{1}{T_i}\int e\,dt - \frac{dy_f}{dt}\right) \tag{A1}$$

$$e = y_r - y \tag{A2}$$

$$Y_f(s) = \frac{1}{1 + sT_d/10}Y(s) \tag{A3}$$

where u_c, y, y_r are the controller output, process output and set-point respectively, the refined Ziegler–Nichols tuning formula for stable processes is as follows.
PID: *Case* 1 $(2.25 < \kappa_1 < 15; 0.16 < \theta_1 < 0.57)$

$$\beta = \frac{15 - \kappa_1}{15 + \kappa_1} \quad \text{for 10\% set-point response overshoot} \tag{A4}$$

$$\beta = \frac{36}{27 + 5\kappa_1} \quad \text{for 20\% set-point response overshoot.} \tag{A5}$$

Case 2 $(1.5 < \kappa_1 < 2.25;\ 0.57 < \theta_1 < 0.96)$. 20% set-point response overshoot

$$T_i = 0.5\mu T_u \tag{A6}$$

$$\mu = \tfrac{4}{9}\kappa_1 \tag{A7}$$

$$\beta = \tfrac{8}{17}(\tfrac{4}{9}\kappa_1 + 1) \tag{A8}$$

where μ is defined as the ratio of the modified integral time to the Ziegler–Nichols integral time.
PI: $(1.2 < \kappa_1 < 15; 0.16 < \theta_1 < 1.4)$

$$\frac{K_c}{K_u} = \frac{5}{6}\left(\frac{12 + \kappa_1}{15 + 14\kappa_1}\right) \tag{A9}$$

$$\frac{T_i}{T_u} = \tfrac{1}{5}(\tfrac{4}{15}\kappa_i + 1). \tag{A10}$$