CEAI, Vol. 4, No. 3, pp. 39-46, 2002 Printed in Romania

DEVELOPMENT AND EVALUATION OF A PID AUTO-TUNING ALGORITHM BASED ON RELAY FEEDBACK

Ioan NAŞCU

Technical University of Cluj Napoca, Automation Department

E-mail: Ioan.Nascu@aut.utcluj.ro

Abstract: *A method for automatic tuning of the PID controllers, usually called auto-tuning, is presented. The method is based on a frequency domain model of the process to be controlled and uses a single point on the Nyquist curve. The identification of this point is performed using a relay connected in a feedback loop with the process, forcing the process output to oscillate. The controller design is based on the dominant closed-loop poles method. No apriori information on the process parameters is required. Simulations compare the proposed method with two other tuning methods: Ziegler-Nichols method and internal model method applied to PID control. A variant using two points on the Nyquist curve is also presented. Some experimental results are shown for auto-tuning control of a thermal process.*

Keywords: *auto-tuning, PID controllers, dominant pole design, frequency domain design, relay experiment.*

1. INTRODUCTION

Despite the development of more advanced control strategies, the majority of controllers used in industrial instrumentation still are of the PID type. Their popularity is easy to understand - they have a simple structure, their principle is well understood by engineers and their control capabilities have proven to be adequate for most control loops. Moreover, due to process uncertainties, a more sophisticated control scheme is not necessarily more efficient than a well-tuned PID controller. However, it is common that PID controllers are often poorly tuned because the choice of controller parameters requires professional knowledge by the user.

For industrial process control there are now many PID controllers with features like self-tuning and auto-tuning. These features provide easy-to-use controller tuning and have proven to be well accepted among process engineers.

For the automatic tuning of the PID controllers, several different methods have been proposed. Some of these methods are based on identification of one point of the process frequency response, while the others are based on the knowledge of some characteristic parameters of the open-loop process step responses. The identification of a point of the process frequency response can be performed either using a proportional regulator, which brings the closed-loop system to the stability boundary or, by a relay forcing the process output to oscillate. Aström (1984) reports an important and interesting approach. The method is based on the Ziegler and Nichols frequency domain design formula. A relay connected in a feedback loop with the process is used in order to determine the critical point.

The performance of the controllers tuned according to ZN rules depend strongly on the value of the process normalized dead-time (the normalized dead-time is defined for stable processes as the ratio of the apparent dead-time to the apparent time constant [3]). ZN rules often give poor damping and excessive overshoot in response to setpoint change for processes with small values for normalized dead-time [1]. For this type of processes we developed a method based on dominant pole design.

Dominant pole design methods find controller parameters which place the dominant poles of the closed-loop system (the complex conjugate closed-loop poles closest to the jω axis) in specified locations.

2. PROCESS

In many practical cases the process model is a first, second or a third order with no delay (tank level control, temperature control in: stirred tank heating processes, heat exchangers, thermal treatment furnaces).

Strictly applying theory, not all of these processes can be forced to oscillate by a relay. A relay without hysteresis can be used only if process Nyquist curve crosses the negative real axis, while a relay with hysteresis is suitable if it crosses the negative imaginary axis. In fact, considering that in any digital controller implementation the sampling process itself introduces a phase lag and that in real situations the process output is filtered it can be assumed that all the processes in practical cases will oscillate when a relay controller is connected. Unfortunately, the relay experiment gives small values for the amplitude and the period of oscillation in absence of hysteresis. In these conditions the tuning obtained by the ZN methods can often be improved significantly by using other methods.

In this paper, a continuous time process described by a rational transfer function is considered, and this transfer function is assumed to have essentially only real, stable poles:

$$H_p(s) = \frac{k_p}{\prod_{j=1}^{n}(T_j s + 1)} \tag{1}$$

where $T_j > 0$ $(j=1, \ldots, n)$. Satisfactory control performances are obtained even if the transfer function has some complex, stable poles.

No a priori information about the value of any model parameter is supposed to be known.

3. AUTOMATIC TUNING METHOD

A standard PID control system with single input, single output, as shown in figure 1 is considered.

The PID controller has a non-interacting structure

$$H_c(s) = K_c\left(1 + \frac{1}{sT_i} + sT_d\right)\frac{1}{T_f s + 1} \tag{2}$$

The transfer function of the closed loop system from the reference signal to the output is given by

$$H_0(s) = \frac{H_1(s)}{1 + H_1(s)} \tag{3}$$

with:

$$H_1(s) = H_p(s) \cdot H_c(s). \tag{4}$$

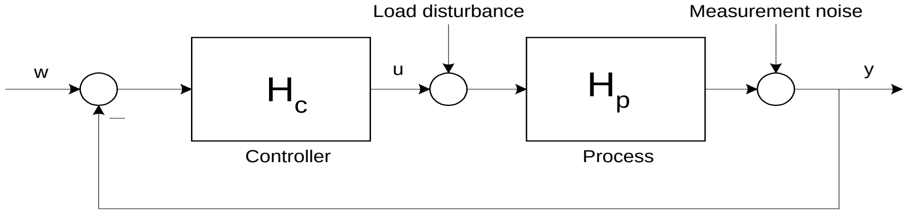


Fig. 1. Control system structure.

A simple method will be proposed for approximate determination of the controller parameters so that the transfer function of the closed-loop system has desired dominant poles, using some knowledge of the Nyquist curve of the process transfer function $H_p(s)$.

Before continuing, some preliminaries are useful to define the context in which the tuning method is derived.

The pole placement design methods are usually used to assign all closed-loop poles. One difficulty with these methods is that complex process models lead to complex controllers. Several papers on PID control are based on the idea of positioning a few closed-loop poles. The reason for this is that the dynamics of complex systems can often be characterized by a few poles. If a pair of complex poles is dominating, a second-order model can be used as a reasonable approximation of the closed-loop system.

We assume that the closed-loop has a standard second order model

$$H_0(s) = \frac{\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2} \quad . \tag{5}$$

The closed-loop pole locations are parameterized with

$$p_{1,2} = \omega_n\left(-\zeta \pm j\sqrt{1-\zeta^2}\right) = \sigma_d \pm j\omega_d \tag{6}$$

where:

$\sigma_d = -\zeta\omega_n, \quad \omega_d = \omega_n\sqrt{1-\zeta^2}, \quad 0 < \zeta < 1,$

ζ represents the damping ratio, ω_n - the undamped natural frequency, σ_d - the exponential decay and ω_d - the damped natural frequency of the step transient response term.

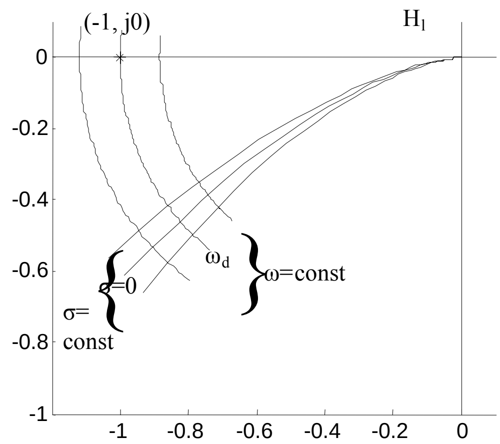


Fig. 2. H_l plane.

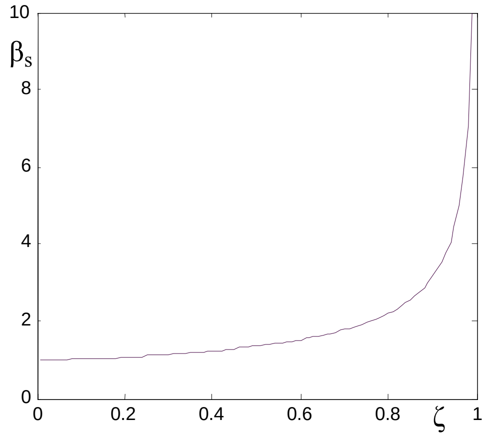


Fig. 3. The ratio $\beta_s=\omega_s/\omega_d$.

Consider the open-loop transfer function $H_l(s)$ as a mapping from the s-plane to the H_l plane. The constant-σ lines and constant-ω lines in the s plane map into curves shown in figure 2. The constant-σ lines map into curves that are similar to the Nyquist plot and are in a sense parallel to the Nyquist plot. The σ=0 line (the jω axis) in the s-plane is mapped into the Nyquist plot in H_l plane. The closed-loop poles are mapped into the (-1, j0) point in the H_l plane. Since the constant-ω line that passes trough the (-1, j0) point in the H_l plane actually corresponds to the damped natural frequency ω_d, we can approximate that the point $H_l(j\omega_d)$ is the closest point of the $H_l(j\omega)$ locus to the (-1, j0) point. Regarding the error involved in this approximation, considering ω_s as the exact frequency value at the point of nearest approach of the $H_l(j\ \omega)$ locus to the (-1, j0) point, figure 3 shows the ratio $\beta_s=\omega_s/\omega_d$ as a function of the relative damping ζ. For small ζ, ω_s does not differ significantly of ω_d so no correction is necessary.

In order to tune the PID controller it is assumed that the value of the process transfer function at frequency ω_1 is known, i.e.,

$$H_p(j\omega_1) = a_1 + jb_1 \tag{7}$$

The parameters a_1 and b_1 can be obtained using a relay connected in a feedback loop with the process, forcing the process output to oscillate [2]. The condition for oscillation is given by

$$N(a_r) \cdot H_p(j\omega) = -1 \tag{8}$$

or

$$H_p(j\omega) = -\frac{1}{N(a_r)} \tag{9}$$

where

$$-\frac{1}{N(a_r)} = -\frac{\pi}{4d_r}\sqrt{{a_r}^2 - {\varepsilon_r}^2} - j\frac{\pi\varepsilon_r}{4d_r} \tag{10}$$

is the negative inverse of relay describing function, d_r is the relay amplitude, a_r the amplitude of the oscillation in the process output and ε_r is the relay hysteresis width.

For the given values d_{r1} and ε_{r1} we will obtain oscillation with amplitude a_{r1} and period T_1.

With a pure relay (no hysteresis, $\varepsilon_{r1}=0$) it is possible to determine the point $(-\pi a_{r1}/4d_{r1}, j0)$ where the Nyquist curve of the process transfer function intersects the negative real axis at the frequency $\omega_1 = 2\pi/T_1$. Having a relay with hysteresis ($\varepsilon_{r1}\neq 0$), we can determine the point $(-\pi\sqrt{{a_{r1}}^2 - {\varepsilon_{r1}}^2}/4d_{r1},\ -j\pi\varepsilon_{r1}/4d_{r1})$ where the Nyquist curve of the process transfer function intersects a straight line parallel to the real axis at the frequency $\omega_1 = 2\pi/T_1$.

Hence

$$\omega_1 = \frac{2\pi}{T_1}, a_1 = \frac{-\pi\sqrt{{a_{r1}}^2 - {\varepsilon_{r1}}^2}}{4d_{r1}}, b_1 = -\frac{\pi\varepsilon_{r1}}{4d_{r1}} \tag{11}$$

The auto-tuning method proposed in this paper has three variants.

4.1. *Fixing one point on the Nyquist curve (NP1).*

The design problem is then to determine a controller so that the frequency response function of the open-loop system (process and controller) has a desired value at a given frequency, i.e.

$$H_1(i\omega_1) = H_p(i\omega_1)\, H_c(i\omega_1) = c_1 + id_1 \tag{12}$$

Let the transfer function of the PID controller be parametrized as

$$H_c(s) = K_c\left(1 + \frac{1}{sT_i} + s\alpha T_i\right) \tag{13}$$

To simplify the calculus, the noise filter time constant is neglected. In the simulation it will be set at $0.2T_d$. Also, it was assumed that $T_d = \alpha T_i$.

Substituting (7) and (13) into (12) we obtain:

$$\begin{aligned} c_1 &= K_c(a_1 - b_1\alpha\omega_1 T_i + b_1/\omega_1 T_i) \\ d_1 &= K_c(b_1 + a_1\alpha\omega_1 T_i - a_1/\omega_1 T_i) \end{aligned} \tag{14}$$

If α is specified a priori, then the parameters K_c and T_i of the PID controller can be determined. Using c_1/d_1 ratio, equations (14) give a second-order equation from which T_i is solved. The controller gain K_c is then obtained.

The method is thus based on the idea to determine a controller that moves the point (a_1, jb_1) to the point (c_1, jd_1). It is the same idea as for the ZN frequency response method but using different values for the point co-ordinates. The ZN method uses a pure relay ($\varepsilon_r=0$) experiment and moves the ultimate point to the point (-0.6, -0.28j). We suggest to chose the point (c_1, jd_1) the closest point of the H_l(j ω) locus to the (-1, j0) point (A point in figure 4). Moreover, the shortest distance from the Nyquist curve to the critical point (-1, j0) represent the inverse of the sensitivity.

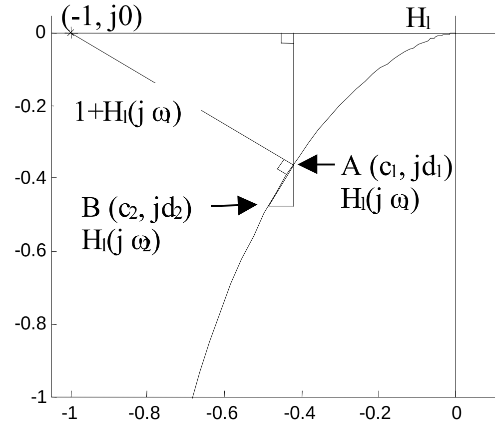


Fig. 4. Placement of (c_1, jd_1) point.

Putting the frequency ω_s at the A point is equal to ω_1 obtained using the relay feedback experiment (with or without hysteresis), and taking into consideration the β_s ratio given in figure 3, it is possible to calculate the desired frequency of the closed-loop dominant poles ω_n

$$\omega_n = \frac{\omega_s}{\beta_s\sqrt{1-\zeta^2}} = \frac{\omega_1}{\beta_s\sqrt{1-\zeta^2}} \tag{15}$$

Since the second-order system (5) has the following open-loop transfer function

$$H_1(s) = \frac{{\omega_n}^2}{s^2 + 2\zeta\omega_n s} \tag{16}$$

for $s=j\,\omega_1$, we obtain the A point co-ordinates

$$c_1 = \frac{-1}{\beta^2(1-\zeta^2) + 4\zeta^2},$$

$$d_1 = \frac{-2\zeta}{\beta\sqrt{1-\zeta^2}\left[\beta^2(1-\zeta^2) + 4\zeta^2\right]}. \tag{17}$$

The design method is based on specification of only one parameter, the relative damping ζ. For ζ=0.7 we have c_1= -0.28 and d_1=-0.31.

Let us now note that $H_p(j\,\omega_1)$, $H_c(j\,\omega_1)$ and $H_l(j\,\omega_1)$ are complex quantities and can be written as follows:

$$H_p(j\omega_1) = r_p e^{j\phi_p}, \quad H_c(j\omega_1) = r_c e^{j\phi_c}, \quad H_l(j\omega_1) = r_l e^{j\phi_l} \tag{18}$$

Thus, from (12) we obtain

$$r_l e^{j\phi_l} = r_p e^{j\phi_p}\; r_c e^{j\phi_c} \tag{19}$$

and finally we have another possibility, different from (14) to determine the controller parameters [1]:

$$Kc = \frac{r_l \cos(\phi_l - \phi_p)}{r_p}, \quad Ti = \frac{1}{2\alpha\omega_l}\left(\tan(\phi_l - \phi_p) + \sqrt{4\alpha + \tan^2(\phi_l - \phi_p)}\right) \tag{20}$$

4.2. *Fixing one point and the slope (NPS), or two points (NP2)of the Nyquist curve.*

In order to tune the PID controller it is assumed that the values of the process transfer function at two neighboring frequencies ω_1 and ω_2 are known, i.e.,

$$H_p(i\omega_1) = a_1 + ib_1, \quad H_p(i\omega_2) = a_2 + ib_2 \tag{21}$$

The frequency of the second point ω_2 and his co-ordinates (a_2, jb_2) are determined in the same manner like in the first point (a_1, jb_1) case (11), using a relay feedback experiment with $\varepsilon_{r2} > \varepsilon_{r1}$.

Fixing the value of the open-loop transfer function at one frequency the method gives only two unique parameters of the controller. This is the reason why the condition $T_d = \alpha T_i$ was introduced in (13). The transfer function of the PID controller will be now parametrized as

$$H_c(s) = K_c\left(1 + \frac{1}{sT_i} + sT_d\right) \; . \tag{22}$$

To obtain unique parameter values one possibility is to position one point and to fix the slope of the Nyquist curve at this point. As we discussed, the fixed point (c_1, jd_1) is the closest point of the $H_l(j\omega)$ locus to the (-1, j0) point (figure 4). A natural requirement is that the slope at frequency ω_1 should be orthogonal to the line $1+H_l(j\omega_1)$.

Differentiating the open-loop transfer function with respect to $j\omega$ from (4) we obtain

$$\frac{dH_l(j\omega)}{d(j\omega)} = H_p(j\omega)\frac{dH_c(j\omega)}{d(j\omega)} + \frac{dH_p(j\omega)}{d(j\omega)}H_c(j\omega) \tag{23}$$

Approximating the term $dH_d(j\omega)$ by a difference between the two closed points on the Nyquist curve the following relation is obtained

$$H'_l(j\omega_1) = H_p(j\omega_1)\frac{dH_c(j\omega)}{d(j\omega)} + \frac{H_p(j\omega_1) - H_p(j\omega_2)}{j(\omega_1 - \omega_2)}H_c(j\omega) \tag{24}$$

or

$$H'_l(j\omega_1) = (a_1 + jb_1)K_c(T_d + \frac{1}{\omega_1^2 T_i}) - \frac{j(a_1 - a_2) - (b_1 - b_2)}{(\omega_1 - \omega_2)}K_c[1 + j(\omega_1 T_d - \frac{1}{\omega_1 T_i})] \tag{25}$$

Differentiating the open-loop transfer function with respect to $j\omega$ from (16), for $\omega=\omega_1$ and ζ *=0.7* we obtain

$$H'_l(j\omega_1) = 24.097 - j \cdot 10.295, \quad \delta = \frac{imag(H'_l(j\omega_1))}{real(H'_l(j\omega_1))} = -0.427 \tag{26}$$

From (25) and (26) we obtain

$$q_1 T_d + q_2\frac{1}{T_i} = q_3 \tag{27}$$

Where

$$q_1 = b_1 + \frac{\omega_1(b_2 - b_1)}{\omega_2 - \omega_1} - \delta\frac{\omega_1(a_2 - a_1)}{\omega_2 - \omega_1} - \delta a_1,$$

$$q_2 = \frac{b_1}{\omega_1^2} - \frac{b_2 - b_1}{\omega_1(\omega_2 - \omega_1)} + \delta\frac{a_2 - a_1}{\omega_1(\omega_2 - \omega_1)} - \frac{\delta a_1}{\omega_1^2},$$

$$q_3 = \delta\frac{b_2 - b_1}{\omega_2 - \omega_1} + \frac{a_2 - a_1}{\omega_2 - \omega_1}.$$

From (14) and (27) we obtain

$$c_1 = K_c(a_1 - b_1\omega_1 T_d + b_1/\omega_1 T_i), \quad d_1 = K_c(b_1 + a_1\omega_1 T_d - a_1/\omega_1 T_i), \quad q_3 = q_1 T_d + q_2/T_i \tag{28}$$

and the parameters K_c , T_i and T_d of the PID controller can be now determined.

From the figure 4 we notice that is possible to approximate δ by the value $\delta = -d_1/(-1-c_1)=-0.43$. Notice also that is possible to obtain negative values for T_i . In this case we have to increase iteratively the value of δ and to recalculate T_i until a positive value is obtained. This will increase the computational effort considerably.

The calculus is simplified fixing two neighboring points of the Nyquist curve: (c_1, jd_1) and (c_2, jd_2). In this case the parameters K_c , T_i, T_d and T_f are determined from:

$$\begin{aligned} c_1 - d_1 T_f \omega_1 &= K_c (a_1 - b_1 \omega_1 T_d + b_1/\omega_1 T_i) \\ d_1 + c_1 T_f \omega_1 &= K_c (b_1 + a_1 \omega_1 T_d - a_1/\omega_1 T_i) \\ c_2 - d_2 T_f \omega_2 &= K_c (a_2 - b_2 \omega_2 T_d + b_2/\omega_2 T_i) \\ d_2 + c_2 T_f \omega_2 &= K_c (b_2 + a_2 \omega_2 T_d - a_2/\omega_2 T_i) \end{aligned} \tag{29}$$

4. SIMULATION

The proposed auto-tuning algorithms have been tested by simulated examples. The performances are evaluated and compared with two other tuning methods: Ziegler-Nichols method (ZN) and internal model method applied to PID control (IMC). Notice that IMC can be considered as a kind of reference but this is not an auto-tuning method (it requires the full process model). Different aspects, such as process dynamics, setpoint changes and load disturbances are analyzed.

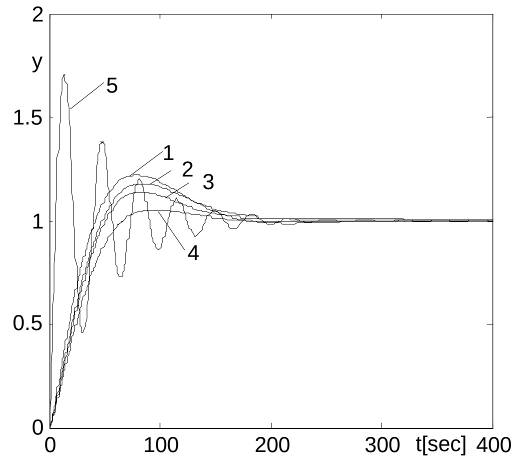
The simulations were carried out in MATLAB environment. The amplitude of the setpoint and the load disturbance is 1. For the first relay experiment the amplitude of the relay was set to $0.1u_0$ and the hysteresis to $0.01\ y_0$ (u_0, y_0 - input and output steady-state values). For the second experiment the relay hysteresis was doubled.

Example I

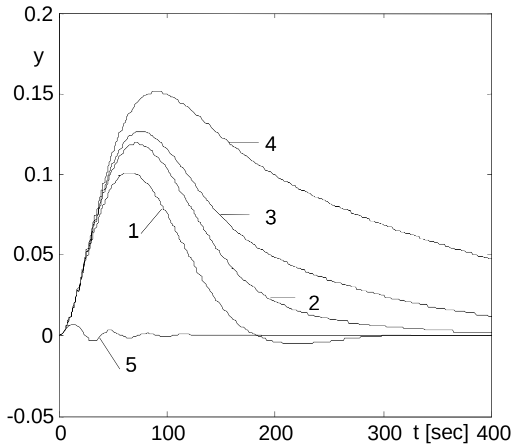
The process is described by a second-order transfer function with a large time constant specific to the process and a small time constant due to the sensor response time:

$$H_p(s) = \frac{2}{(270s+1)\ (30s+1)} \tag{30}$$

The closed loop step response using ZN method is not well damped (Figure 5). The step responses obtained using the proposed methods are comparable with IMC response. The fastest response for load disturbance is obtained with ZN method; the proposed methods lead to a slow response but not so slow like the IMC method. The load disturbance rejection, using NP1 method is quite satisfactory. The design parameter has been set at $\zeta = 0.7$.



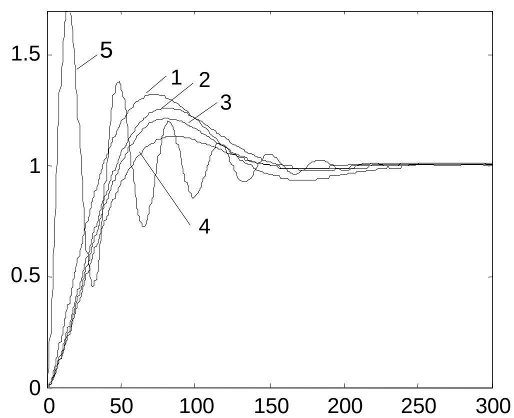
a)



b)

Fig. 5. Comparison of the NP1(1), NP2(2), NPS(3), IMC (4), and ZN(5) methods, a) - step response, b) - load disturbance response.

For $\zeta = 0.5$ faster responses for load disturbance are obtained but the overshoot of the step response increased (Figure 6).



a)

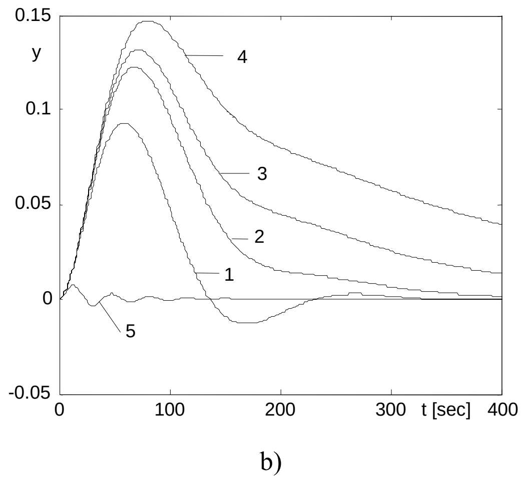


Fig. 6. Comparison of the NP1(1), NP2(2), NPS(3), IMC (4), and ZN(5) methods, a) - step response, b) - load disturbance response.

Example II

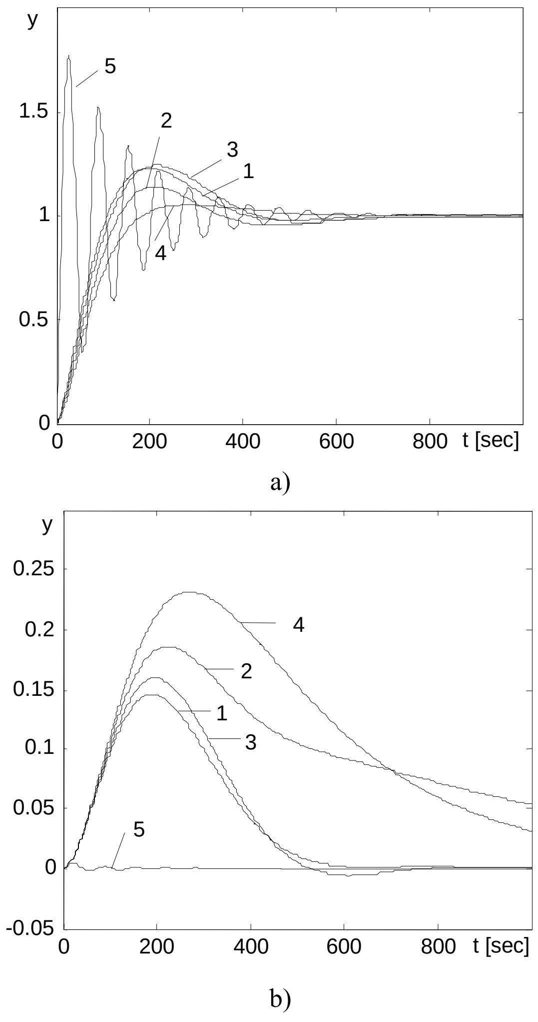


Fig. 7. Comparison of the NP1(1), NP2(2), NPS(3), IMC (4), and ZN(5) methods, a) - step response, b) - load disturbance response.

The process is described by a second order transfer function with two not too different time constants:

$$H_p(s) = \frac{2}{(270s+1)\ (200s+1)} \tag{31}$$

The step response obtained using ZN method is not well damped, as in the previous example (Figure 7). The NP methods give much better response. The load disturbance rejection, using NP1 and NPS is not so fast like ZN but is quite satisfactory (ζ =0.7).

Example III

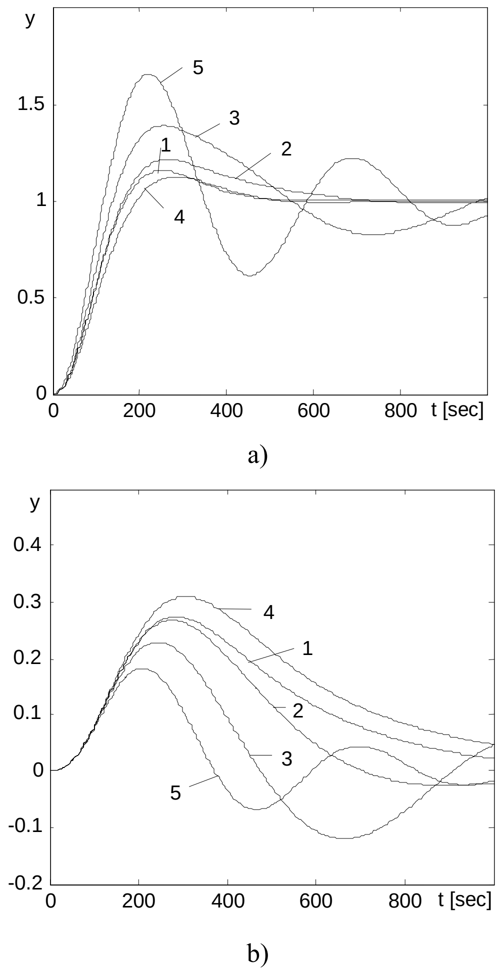


Fig. 8. Comparison of the NP1(1), NP2(2), NPS(3), IMC (4), and ZN(5) methods, a) - step response, b) - load disturbance response.

For more complex processes, third order for example:

$$H_p(s) = \frac{2}{(270s+1)(200s+1)(90s+1)} \tag{32}$$

the NP1 gives better results than NP2 and NPS because the last two design methods are too restrictive. (Figure 8; $\zeta = 0.5$)

5. EXPERIMENTAL RESULTS

For experimental results the auto-tuning algorithm was implemented on a PC computer equipped with a data acquisition board and simple laboratory processes were controlled. Figure 9 shows the results when the auto-tuner was applied to temperature control of a soldering hammer.

The soldering hammer was first brought to steady-state conditions in manual control (θ_0=200ºC, u_0=0.18). A sampling period of 1 second was used in all of the experiments. For the first experiment the amplitude of the relay was set to $0.1u_0$ (d_{r1}=0.2) and the hysteresis to 0.01 θ_0 (ε_{r1}=2ºC). After a permanent oscillation of θ(t) occured we measured the period and amplitude of this oscillation and obtained T_1=150 sec and a_{r1}=3.1ºC. For the second experiment the relay hysteresis was increased to ε_{r2}=3ºC and T_2=270 sec, a_{r2}=3.5ºC were obtained. Finally, PID parameters were computed according to the control performance specifications, e.g. the desired damping ratio ($\zeta =0.7$) using the three methods presented in this paper.

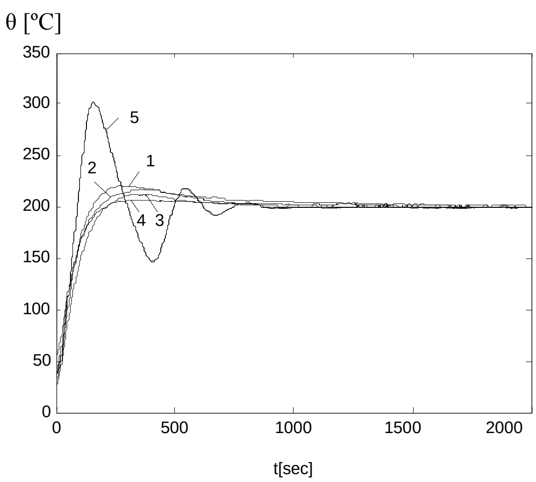


Fig. 9. Comparison of the NP1(1), NP2(2), NPS(3), IMC (4), and ZN(5) methods, temperature control of a soldering hammer, step response (200ºC).

6. CONCLUSIONS

Some relay based algorithms for auto-tuning of PID controllers have been presented, assuming a process model structure and achieving the regulator tuning by identifying one or two points of the process frequency response.

The auto-tuner is an extension of the Ziegler-Nichols relay approach. As seen in simulation and experimental results, the performances are comparable with those obtained using IMC and better than those obtained using the Ziegler-Nichols method. That means small overshoot in response to setpoint change for processes specified in section 2. If the main goal of the controller design is to obtain a faster response for load disturbances then the value of the design parameter ζ can be decreased.

This auto-tuning method yield PID parameters only for a restricted class of process models. It is not a general methodology for arbitrary process models.

Concerning the complexity of the method, the proposed methods involve calculations more complicated than ZN, but the experiments are easy to be performed.

Experiments and simulation studies have indicated that the presented self-tuners perform well and can be easily used even by people who are not specialist in automatic control.

7. REFERENCES

[1] Aström,K.J. *et.all.* – “Automatic tuning of PID controllers”, Instrument Society of America, 1988. [2] – “Automatic tuning of simple regulators with specifications on phase and amplitude margins”, *Automatica*, Vol.20, pp. **645**, 1984. [3] – “Towards intelligent PID control”, *Automatica*, Vol.28, pp. **1**, 1992.

[4] Ogata,K. – “Modern Control Engineering”, Prentice-Hall, 1990.

[5] Ho, W.K., Hang, C.C. and Cao,L.S. – “Tuning of PID controllers based on gain and phase margin specifications”, *Automatica*, Vol.31, No.3, pp. **497**, 1995.

[6] Poulin, E., Pomerleau, A., Desbiens, A. and Hodouin, D. – “Development and evaluation of an auto-tuning and adaptive PID controller”, *Automatica*, Vol.32, pp. **71**, 1996.

[7] Schei,T.S. – “Automatic tuning of PID controllers based on transfer function estimation”, *Automatica*, Vol.30, pp. **1983**, 1994.

[8] Weng, K.H. and Chang, C.H. – “Tuning of PID controllers based on gain and phase margins specifications”, *Automatica*, Vol.31, pp. **497**, 1995.