RELAY with HYSTERESIS for MONITORING and CONTROLLER DESIGN

C. Scali * and **R. Costaggiu**

* *Dipartimento di Ingegneria Chimica, Università di Pisa, Via Diotisalvi 2, 56126 Pisa, ITALY.*
e-mail: scali@ing.unipi.it

Abstract. Characteristics of techniques based on relay with hysteresis are analysed in the perspective of developing a complete methodology to assist the three stages of the problem: initial design, monitoring of performance and retuning of the control system for unknown processes. The technique, when used for monitoring of performance, through the evaluation of the peak of the complementary sensitivity function, can give low accuracy even for simple classes of processes. A promising application is proposed for a fast identification of a starting point on the Nyquist curve in the third quadrant, which can be used to address the three stages of the global problem. *Copyright © 1998 IFAC*

Key Words. Relay Feedback, Hysteresis, Identification, Monitoring, PI Tuning

1. INTRODUCTION

Chemical processes are often difficult to model because they are poorly known or their parameters change when the process is operated at different conditions. Therefore the approach to the design of the control system based on a good model of the process can find severe limitations: modelling can be too time consuming and there is the need of updating the controller parameters to maintain good performance in different conditions.

In these cases a viable alternative is through the use of techniques which are simple to apply on line and allow a quick, though not optimal, determination of the controller parameters. Among these techniques, the adoption of a relay in the feedback loop to bring the system in conditions of limit cycle with controlled amplitude is one of the most used methods (Åström and Hägglund, 1984). In its simplest formulation it allows to compute the parameters k_u (ultimate gain) and ω_u (ultimate frequency) from the critical oscillation as:

$$k_u = \frac{4h}{\pi a}; \qquad \omega_u = \frac{2\pi}{p_u}. \tag{1}$$

where: h is the relay amplitude, a is the limit cycle amplitude, p_u is the limit cycle period.

From this minimum amount of information on the process, it is possible to estimate the values of the parameters of PID controllers able to give reasonable performance of the controlled system in many cases. Simplicity of application, short duration of tests, ease of repetition in the case of frequent parameters changes are the attractive features of this method. Extensions to the design of decentralized control systems for multivariable processes have been proposed by several authors; Loh *et al.* (1993), Shen and Yu (1994) use a sequential procedure, which consists in the relay identification of one loop at a time with all other loops closed; Friman and Waller (1994), and Palmor *et al.* (1995) study the application of simultaneous multirelay tests in the plant.

Obviously, due to the non linearity of the relay, there are some approximations in the determination of the parameters and the limited information about the process can be responsible for unacceptable results in some cases. For example, starting from the knowledge of k_u and ω_u, the ZN technique (Ziegler and Nichols 1942) is a forced choice for the tuning of PID controllers, but this can bring to poor performance and even instability in some cases.

Several modifications of the original technique have been proposed in order to overcome this drawback: Hang *et al.* (1991) propose a refinement of the ZN technique which allows to extend the stability region but requires to know the gain of the process; Shen and Yu (1994) propose the adoption of larger detuning factors, without requiring any other knowledge on the process, but the closed loop response can be slow in many cases. In a previous paper (Semino *et al.*, 1996) we showed that no one of the techniques can work satisfactorily for different classes of processes when the only available information is the critical point (k_u, ω_u).

Li *et al.* 1991 illustrate an improved identification of the process (ATV method) leading to a model with up to 4 parameters, which requires two additional relay tests with (known) delay. Certainly, if we can afford the larger times and complexity of a series of identification runs, it can be worth to go through a further improvement of the ATV method and to exploit this added knowledge on the process with an appropriate controller design method (Semino and Scali, 1998). On the contrary, if we want to maintain as simple as possible the identification procedure, the relay feedback tests are very attractive and it is worth to try to correct the basic defect which derives from having the position of the starting point on the real negative axis.

The use of a relay with hysteresis permits to identify a point belonging to the third quadrant, as illustrated in Figure 1.

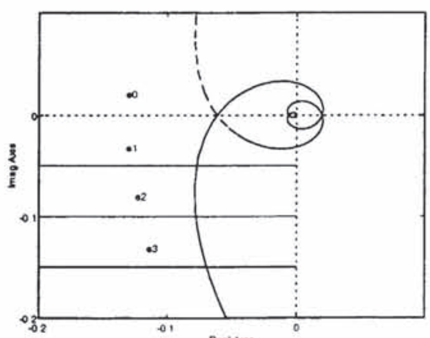

Fig. 1. Identification via relay with hysteresis

The real and imaginary part of of each point at limit cycle conditions, can be computed with some approximations, inherent in the describing function analysis, as:

$$Re = -\frac{\pi}{4h}(a^2 - \varepsilon^2)^{0.5}; Im = -\frac{\pi}{4h}\varepsilon \qquad (2)$$

By changing the amount of hysteresis it is possible to identify different points with phase lag between -180^o and -90^o. The knowledge of these points can be exploited in the stage of design of the regulator as it allows to give some margin of robustness (Åström and Hägglund, 1995) and for a monitoring of the control system performance, as illustrated by Chiang and Yu (1993).

This paper analyses the possibility of application of the relay with hysteresis for the two purposes of monitoring and of control system design. For different classes of processes the main features and relative accuracy of the method will be evaluated in the perspective of developing a complete methodology, able to assist the global process of initial design, monitoring and redesign of the control system.

2. RELAY with HYSTERESIS for MONITORING: THEORY

Chiang and Yu (1993) illustrate the application of the relay with hysteresis for monitoring the performance of the controlled system. The method can be extended to two by two and multivariable processes, as reported in Ju and Chiu (1996, 1997).

The experimental setup is the one illustrated in Figure 2: from time to time the relay is activated and the plant perturbed and brought to limit cycle conditions; the controller, already designed in a previous stage, is always active. By the analysis of the oscillating signal u, an identification of the open loop transfer function $G = PC$ can be carried out. As index of performance of the controlled system the peak of the complementary sensitivity function $\hat{\eta}$ is chosen:

$$\hat{\eta} = \max_{\omega} |\eta(\omega)| = \max_{\omega} \frac{|G|}{|1+G|} \qquad (3)$$

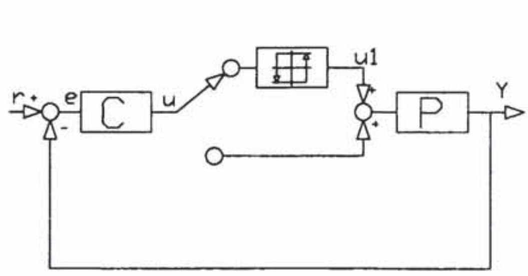

Fig. 2. Setup for monitoring via relay with hysteresis

The original contribution of Chiang and Yu (1993) is that they give a method to shift the formulation of the problem from the frequency (ω) domain to the hysteresis (ε) domain and to compute the value of $\hat{\eta}$, by means of only two experimental runs. The steps of the procedure are the following:

1. Switch from the control mode to the monitoring mode (Figure 2).

2. Perform an ideal relay experiment to identify a point on the real negative axis (amplitude of oscillation a_o).
3. Perform a second relay feedback experiment with hysteresis width ($\varepsilon \neq 0$) to identify a point in the third quadrant (amplitude of oscillation a_1).
4. Assume to approximate the relationship $a = f(\varepsilon)$ with a linear law: $a = a_o + m\varepsilon$ and compute the slope as: $m = (a_1 - a_o)/\varepsilon$.
5. Compute the maximum modulus of the complementary sensitivity function η, which now depends only on ε:

$$\hat{\eta}(\varepsilon) = \max_{\varepsilon} |[1 - 2\alpha(1-\beta^2)^{\frac{1}{2}} + \alpha^2]^{-\frac{1}{2}}| \quad (4)$$

$$\alpha = \frac{4h}{\pi(m\varepsilon + a_0)}; \; \beta = \frac{\varepsilon}{m\varepsilon + a_0} \quad (5)$$

Further details are given in the paper. The same authors point out that the method can have several limitations; among them:

- as the relay hysteresis must be a positive value, the search for $\hat{\eta}$ is confined to the third quadrant, which is not always the case,

- the linear assumption between the relay hysteresis width ε and the amplitude a of the oscillation is not always verified,

- the choice of the hysteresis width ε is not immediate, as it depends on the process noise.

As general comments at this point, we can say that the choice of $\hat{\eta}$ (also known as maximum closed loop modulus, $L_{C,max}$), as a measure of robustness of the control system, can be accepted for a quick assessment of the situation, even if there may be cases in which it is not meaningful. But we must point out that the method does not solve the global problem because it does not take into consideration the stages of controller design and retuning.

3. RELAY with HYSTERESIS for MONITORING: APPLICATIONS

Several different classes of processes were analysed (first column of Table 1): first, second, second plus lead lag, third order (all with a multiplicative time delay element), are indicated with numbers $1 \div 3$; $4 \div 6$; $7 \div 9$; $10 \div 12$, respectively; for the 3 cases of each class, parameters were changed to cover a wide range of conditions.

A first difficulty in obtaining general results stands from the fact that, according to the initial design of the controller (algorithm and tuning rule), a peak in the complementary sensitivity function may not be present; the aspect of first design of the controller has not been faced by Chiang and Yu (1993). To avoid this inconvenience in our simulation results, the choice of the controller was limited to PI algorithm and the initial design was made according to the following rule:

- choose the integral time as: $\tau_I = \sum_{i=1}^{n}(-1/p_i) + \theta/2$, where p_i, θ are the poles and delay of the process transfer function;

- find the proportional gain that corresponds to a peak value $\hat{\eta} = \hat{\eta}^o = 1.26 \cong 2dB$

The rational behind the method has been illustrated in Semino and Scali (1998) and found to give acceptable performance for PI controllers also starting from a not very precise identification, as in the case of relay tests. To be stressed that it is necessary to assume the knowledge of the process transfer function only because the controller must be tuned, in order to verify the closed loop performance test proposed by Chiang and Yu (1993).

Results for the 4 different classes of processes are synthesized in the columns two and three of Table 1, where the values of $\hat{\eta}$ computed via the hysteresis method are reported together with the percentage absolute value of the error $|\hat{e}| = |\hat{\eta} - \hat{\eta}^o|/\hat{\eta}^o$.

Table 1 Results for monitoring: accuracy in the determination of $\hat{\eta}$

N	$\hat{\eta}$	$\lvert\hat{e}\rvert\%$	$\lvert e_P\rvert\%$	$\lvert e_G\rvert\%$	$\lvert e_\eta\rvert\%$
1)	2.34	88.20	14.62	23.03	50.
2)	2.18	76.16	10.86	21.07	52.4
3)	2.2	76.34	22.41	20.65	50.5
4)	2.76	122.20	6.87	10.12	11.26
5)	1.67	34.20	5.54	6.50	10.74
6)	1.19	4.75	18.69	0.92	2.2
7)	1.68	35.10	7.60	14.84	18.99
8)	1.72	36.00	4.13	6.25	13.83
9)	1.26	1.09	55.10	1.03	2.27
10)	2.25	79.00	1.20	4.02	5.4
11)	1.69	35.10	2.06	4.23	8.86
12)	2.06	65.00	16.50	13.53	49.5

It can be seen that the error shows to be too large in most of the cases: $|\hat{e}| > 35\%$ in 10/12 cases. In the original paper (Chiang and Yu, 1993), the two cases for which results are presented, show a small error in the comparison between the values obtained through the method and values obtained via extensive simulation with different hysteresis. By comparing the correct values computed in the frequency domain: $\eta^o = 0.986, 2.17$, with the values estimated via the method : $\hat{\eta} = 1.076, 3.07$, the errors are: $\hat{e} = 9.1\%, 40\%$. By considering that a retuning of the controller should be made on the basis of the values of $\hat{\eta}$, it is evident that large errors can make the indications derived from the relay test completely useless.

An analysis of the errors in the different stages of the procedure and of its propagation has then been performed, with the aim of investigating which of different steps and approximations of the method can be considered more determinant in this respect. Therefore three other different errors in the evaluation of some related variables have been computed:

- e_P: error in the value of the ultimate gain k_u, as computed via pure relay;

- e_G: error in the value of the closed loop function $G(0)$; this is obtained experimentally at $\varepsilon = 0$ for the system with the regulator active;

- e_η: error in the value of the complementary sensitivity function at hysteresis $\varepsilon = 0$; $\eta(0) = G(0)/(1 + G(0))$.

The three errors are reported in the last three column of Table 1. The following observations can be drawn from an examination of the global table:

1. A small error on G ($|e_G| < 10\%$) is necessary to have small errors on $\hat{\eta}$ (cases 6 and 9). This condition is not sufficient because $\hat{e}$ can be large even when e_G is small (cases 4,5,8,10 and 11). To be noted that the initial determination of G, by means of a relay test without hysteresis $\varepsilon = 0$, can not have larger accuracy; in fact the error e_G shows to be of the same order of magnitude of the error e_P (only in cases 6 and 9 $|e_G| < |e_p|$). This confirms results already presented in Li *et al.* (1991), where for a large number of investigated processes, error $|e_P|$ up to 25% is reported.
2. There is an amplification of the error in passing from e_G to e_η, at least of a factor 2. This depends on the particular functional dependence $\eta(G)$ and is illustrated in Figure 3, where the error $|e_\eta|$ is reported as function of the value $G(0)$, with parameter the error $|e_G|$. The amplification of the error depends strongly on the value of $|G(0)|$: by limiting our analysis to the range $|G(0)| \leq 0.6$, to indicate regulators with a gain margin $GM \geq 1.66$, it can be seen that acceptable errors on G (e.g. $|e_G| \leq 20\%$), become unacceptable on η (e.g. $|e_\eta| \geq 70\%$) and thereafter on the performance indicator $\hat{\eta}$.
3. Only in 2 cases (10,11) starting from an error $|e_P| < 10\%$, the error $|\hat{e}|$ is much larger and this might be caused by approximations introduced in the method of Chiang and Yu (1993).

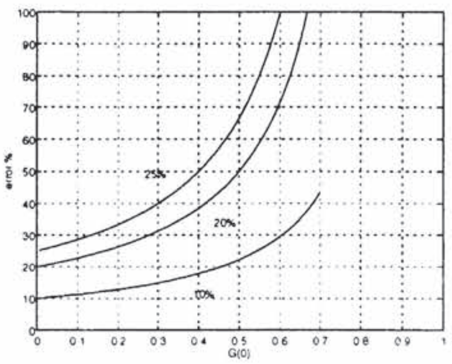

Fig. 3. Amplification of the error $|e_\eta|$ as a function of the value $G(0)$; parameter: $|e_G|$

As conclusive remarks:

- the method does not seem to be reliable to be used for monitoring of closed loop performance,

- the scarce accuracy in the determination of the index $\hat{\eta}$ does not depend on the specific method, but it derives from an amplification of errors inherent to relay techniques.

4. RELAY with HYSTERESIS for CONTROL DESIGN: THEORY

The possibility of identification of a point in the third quadrant given by the presence of hysteresis can be successfully exploited for the design of the control system. A point on the real negative axis, as identified by a pure relay, is not a very good starting point for design: for example, when a PI regulator is used, it is not possible to specify any phase margin for the controlled system which therefore may result unstable.

This fact can be better understood by referring to the interpretation of the effect of the three different components of a PID regulator, as given by Åström and Hägglund (1995) and reproduced in Figure 4. A starting point, representing the open loop process, can be moved in the three directions by changing the action of the three components P, I, D: therefore it becomes clear that a point on the negative real axis does not allow any phase margin under PI control.

Instead, if the starting point is Q, having a predetermined phase lag (identification angle ϕ_Q, in the third quadrant), it is possible to give performance specification on both the phase and the gain margin, even for a PI controller. In particular to move the point $Q(r_Q, \phi_Q)$, representative of open loop conditions, to the point $S(r_S, \phi_S)$, representative of performance specifications (Figure 4), the pa-

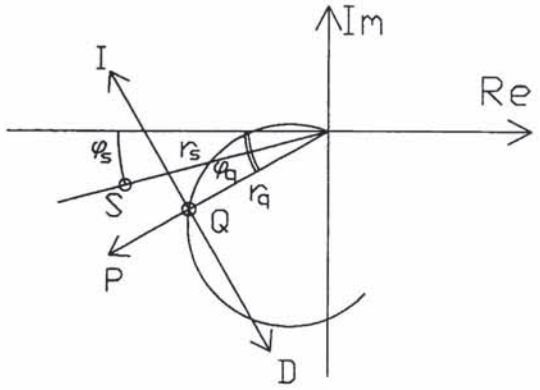

Fig. 4. The effect of the three components of a PID.

rameters of a PI controller must be set as:

$$K_c = \frac{r_S cos(\phi_S - \phi_Q)}{r_Q}; \tag{6}$$

$$\tau_I = \frac{1}{\omega_Q tan(\phi_Q - \phi_S)} \tag{7}$$

being ω_Q the frequency of oscillation of the point Q, identified from the relay test.

The difficulty in the relay with hysteresis derives from the fact that the initial phase angle cannot be specified a priori by the user. Other types of arrangement can be used for this purpose. Friman and Waller (1997) propose the adoption of a Two Channel Relay structure, in which a second relay operating on the integral of the error is added in parallel to the standard one. The technique is a bit more complicated, as the values of the two relay amplitudes h_p, h_i must be determined for a desired identification angle ϕ_Q. In their paper they give the relation $\phi_Q(h_p, h_i)$ and suggest that for a process with unknown dynamics, under PI control, a reasonable choice for the identification angle and for the specification point can be: $\phi_Q = 30^o$, and $r_S = 0.5, \phi_S = 15^o$; obviously other choices are possible and can lead to better performance. Being interested to have stability guaranteed for all cases and acceptable performance, the same specifications have been assumed here and the aspects correlated with the application of the relay with hysteresis have been explored.

5. RELAY with HYSTERESIS for CONTROL DESIGN: APPLICATION

The suggested procedure which covers the three stages of initial design, monitoring of performance and controller retuning is the following:

1. In the design stage: perturb via the relay with hysteresis the completely unknown process; in this case the setup changes with respect to the one reported in Figure 2: the relay with hysteresis acts in alternative to the controller. The scope is to obtain an identification angle $\phi_Q = 30^o$. this may require more than one run, but, as illustrated in the sequel, after the first run an estimate of the value of the hysteresis can be done to orient the second one.
2. Tune the controller according to (6,7), in order to impose the required performance specifications.
3. In the monitoring stage: from time to time exclude the controller and apply on the process the relay with the same hysteresis ($\bar{\varepsilon}$). If some changes happened in the process, the representative point moves from Q to Q', while the identification angle can be equal or different to the desired value (30^o), depending on the type of variations. In the case that $\phi_{Q'} \cong 30^o$, introduce slight changes (6,7) in the controller parameters on the basis of the new identification with $\bar{\varepsilon}$. In the case that the identification angle is far from the desired value ($\phi_Q \neq 30^o$), it is necessary to change the hysteresis to identify a new point Q'', for which the angle is close to 30^o.
4. On the basis of the new identification point $Q''(\phi_{Q''}, r_{Q''}, \omega_{Q''})$ retune the controller according to (6,7).

The critical point of the procedure is how to change the hysteresis after the first identification run, a common situation in step 1 and 3. The identification angle ϕ_Q depends on the hysteresis width (ε) and the amplitude of the oscillation (a), according to:

$$\phi_Q = atan\left[\frac{\varepsilon}{(a^2 - \varepsilon^2)^{\frac{1}{2}}}\right] \tag{8}$$

After the first run with hysteresis ε_1, which brings the system to the angle ϕ_1, a second one at hysteresis ε_2 is necessary, in general. An estimate of ε_2 can be based on the assumption of a linear dependence $\phi_Q(\varepsilon)$. Being $\phi_Q(0) = 0$, we have:

$$\varepsilon_2 = \varepsilon_1\left(\frac{30^o}{\phi_1}\right) \tag{9}$$

The second run with $\varepsilon = \varepsilon_2$ causes an oscillation with amplitude a_2 and can give $\phi_2 = 30^o$; if $\phi_{2'} \neq 30^o$, compute ε_3 as:

$$\varepsilon_3 = \varepsilon_2 + \frac{\phi_{2'} - \phi_{1'}}{\varepsilon_2 - \varepsilon_1}(30 - \phi_{2'}) \tag{10}$$

The procedure is illustrated in Figure 5. Point 1: (ϕ_1, ε_1) represents the situation after the first run, far from the desired identification angle. Point 2: $(\phi_2 = 30^o, \varepsilon_2)$ gives the value of the necessary hysteresis on the basis of the linear assumption (9).

Point 2': $(\phi_{2'} \neq 30^o, \varepsilon_2)$ gives the true value of the identification angle corresponding to ε_2. Point 3: $(\phi_3 = 30^o, \varepsilon_3)$ indicates the value of hysteresis to which a value $\phi = 30^o$ would correspond from (9). Point 3'$(\phi_{3'} \neq 30^o, \varepsilon_3)$: gives the true value of the angle.

In most cases the procedure converges in 2 steps, within an approximation of $\pm 3^o$. Simulation results show that acceptable performance, not optimal though, are obtained for the controlled system, (Costaggiu, 1998). Validation for a wider class of processes is under study, but the procedure seems to be a promising application of relay with hysteresis to address the 3 stages of the global problem stated in the introduction.

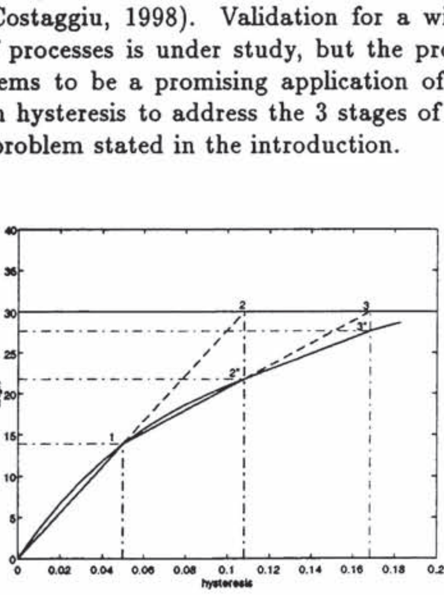

Fig. 5. Procedure to estimate the value of hysteresis ε

6. CONCLUSIONS

The analysis of characteristics of application of techniques based on the relay with hysteresis shows that it can be proposed to solve the three stages of the complete problem: initial design, monitoring and retuning of the control system.

The use of the relay with hysteresis for a monitoring of performance can give scarce accuracy, for relatively simple classes of processes. This depends on the inherent amplification of errors, common to all the relay feedback techniques, in the computation of the maximum modulus of the closed loop sensibility function, used as performance index.

A promising application of the relay with hysteresis is in the identification of a point in the third quadrant, which can be used as a starting point for the initial design, for detecting changes in the process parameters and for a retuning of the controller. The procedure for estimating the relay hysteresis reaches convergence after few runs and allows a tuning which gives acceptable performance for PI controllers for a wide range of processes.

7. REFERENCES

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 (1995). PID Controllers: Theory, Design and Tuning; 2nd Ed. *Instrument Society of America*, 139-142.

Chiang R.C., and C.C. Yu (1993). Monitoring Procedure for Intelligent Control: On-Line Identification of Maximum Closed-Loop Log Modulus. *Ind. Eng. Chem. Res.*, **32**, 90-99.

Costaggiu R. (1998). Ms. Thesis (in Italian). *Chem. Eng. Dept.*, University of Pisa, Italy.

Friman M., and K. V. Waller (1994). Autotuning of Multiloop Control Systems. *Ind. Eng. Chem. Res.*, **33**, 1707-1717.

Friman M., and K. V. Waller (1997). A Two-channel Relay for Autotuning. *Ind. Eng. Chem. Res.*, **37**, 2662-2671.

Hang C. C., K. J. Åström, and W. K. Ho (1991); Refinements of the Ziegler Nichols Tuning Formula. *IEEE Proc. Part D*, **138**, 111-118.

Ju J., and M.S. Chiu (1996). On-line Performance Monitoring of Two-inputs Two-outputs Multiloop Control Systems. *Comp. Chem. Engng.*, **20**, S829-S834.

Ju J., and M.S. Chiu (1997). Relay-based On-line Monitoring Procedure for 2x2 and 3x3 Multiloop Control Systems. *Ind. Eng. Chem. Res.*, **36**, 2225-2230.

Li W., E. Eskinat, and W. L. Luyben (1991); Improved Autotune Identification Method. *Ind. Eng. Chem. Process Des. Dev.*, **30**, 1530-1541.

Loh A. P., C. C. Hang, C. K. Quek, and V. U. Vasnani (1993). Autotuning of Multiloop Proportional Integral Controllers Using Relay Feedback. *Ind.Eng. Chem. Res.*, **32**, 1102-1107.

Palmor Z. J., Y. Halevi, and N. Krasney (1995). Automatic Tuning of Decentralized PID Controllers for TITO Processes. *Automatica*, **31** (7), 1001-1010.

Semino D., I. Mazzanti, and C. Scali (1996). Design of Decentralized Controllers by a Relay Technique: Extensions of Tuning Rules. *Proc. UKACC Int Conf. Control'96* , **427**(2), 1190-1195.

Semino D., and C. Scali (1998). Improved Identification and Autotuning of PI Controllers for MIMO Processes by Relay Techniques. *J. Proc. Contr.*, in press.

Shen S. H., and C. C. Yu (1994). Use of Relay-Feedback Test or Automatic Tuning of Multivariable Systems. *AIChE J.*, **40**, (4), 627-646.

Ziegler J. G., and N. B. Nichols (1942). Optimum Setting for Automatic Controllers. *Trans. ASME*, **65**, 433-444.