

An Estimation Approach for Process Control based on Asymmetric Oscillations

José Sánchez Moreno, María Guinaldo Losada, Sebastián Dormido, José Luis Fernández Marrón
Departamento de Informática y Automática, UNED, C/ Juan del Rosal 16, 28040 Madrid
{jsanchez, mguinaldo, sdormido, jlmarron}@dia.uned.es

Antonio Visioli
Department of Mechanical and Industrial Engineering, University of Brescia, via Branze 38, 25123
Brescia, Italy {Antonio.visioli@unibs.it}

Abstract

An estimation procedure for process control has been developed based on the information obtained from the oscillations that a non-linear element as a simple relay introduces in the feedback loop. Features of the method are: (1) the procedure does not demand a priori process information, (b) non-iterative algorithms are needed to derive the process parameters, (c) only one test is needed, and (d) it allows identifying the process at a user-specified phase lag in the third quadrant. The method is presented for estimation of most common transfer functions found in chemical and process industry: integrators, first-, second- as well as processes with non-minimum-phase dynamics.

Keywords: Describing function, relay, estimation, limit cycle, oscillations.

1 Introduction

The main result presented in this paper is an enhanced estimation procedure based on a non-linear element as a relay. The proposed approach allows the identification of model parameters from measurements derived of the asymmetric oscillations generated by a relay and a bias added to the output. These measures are the oscillation frequency, the main harmonics, and the steady gain (if the process is without integration). The identification procedure is relevant for four reasons:

- (1) Unlike previous works based on relay-induced oscillations, the model parameters do not need to be calculated using a-priori information of the process (i.e., static gain, velocity gain, or dead-time),
- (2) Iterative methods to solve non-linear equations are not necessary,
- (3) The number of tests to obtain the measures is always one regardless the transfer function structure or order, and
- (4) The procedure can be extended to estimate any type of transfer function.

The structure of the paper is as follows. First, the basis of the estimation approach and the problems that have been detected in similar methods are described in Section 2. In Section 3, the procedure is explained and solutions to the problems found are given; also, expressions to estimate the parameters of common transfer functions are derived and simulation results are presented. Section 4 discusses a way to estimate with user-defined phase margin specifications. The paper finishes with conclusions and recommendations about the use of the procedure.

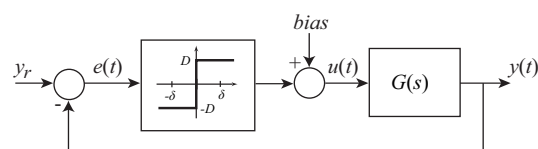


Figure 1: Basic relay feedback control loop.

2 Estimation based on relay-induced oscillations: basis and problems

The pioneering works on the use of relay feedback for identification purposes are from the 80's [1]. The basis of the method is that a linear system under an ideal relay control (Figure 1) oscillates, approximately, at its ultimate frequency, that is, $\omega_{osc} \approx \omega_u$ and the critical gain K_u is derived from the describing function (DF) of the ideal relay. That is,

$$N(A) = \frac{4D}{\pi A} = K_u \quad (1)$$

where A is the oscillation amplitude and D is the relay output. Knowing that the oscillation happens when the expression

$$\frac{-1}{N(A)} = G(j\omega_{osc}) \quad (2)$$

is satisfied, it is feasible to derive the parameters of a transfer function model. One way to do that it is to use the magnitude and argument of the critical point

$G(j\omega_{osc})$ in the Nyquist plot that corresponds to the intersection of $G(s)$ and $-1/N(A)$, separating both components to solve them. That is,

$$\left| \frac{-1}{N(A)} \right| = |G(j\omega_{osc})| \quad (3)$$

$$\arg \frac{-1}{N(A)} = \arg G(j\omega_{osc}) \quad (4)$$

To separate both parts is key to the estimation approach presented here. It is possible to apply iterative methods to solve directly (2), but when the transfer function to estimate has four or more unknowns (for example, K , T_1 , T_2 , and L), most of the time the iterative methods do not converge [7]. It is due primarily to the inaccuracy of the describing function and the extreme nonlinearity of the equations with respect to the parameters. However, by separating (2) in magnitude and argument, it is possible to solve (3) to get the zeros and poles, and (4) to obtain the time delay.

Following this approach, for estimating the parameters of a first order process with time delay $Ke^{-Ls}/(Ts+1)$, the following two equations are obtained

$$\frac{\pi A}{4D} = \frac{K}{\sqrt{1+(T\omega_{osc})^2}} \quad (5)$$

$$-\pi = -\arctan(T\omega_{osc}) - L\omega_{osc} \quad (6)$$

Once the steady gain K is known, it is possible to calculate first the lag T and next the time delay L (or observing L from the process response, T is first estimated and next K). However, the result will be an approximation as the critical gain and the phase lag of -180° are just approximations based on the principal harmonics of the relay output.

Example 1: The system

$$G(s) = \frac{e^{-s}}{s+1} \quad (7)$$

under relay feedback control with $D=1$ oscillates at $\omega_{osc} = 2.105$ with $A=0.6327$, $|G(j\omega_{osc})| = 0.4289$, and $\arg G(j\omega_{osc}) = -174.609^\circ$. However, the describing function theory says that

$$|G(j\omega_{osc})| = \frac{1}{K_u} = \frac{\pi A}{4D} = 0.4969 \quad (8)$$

and $\arg G(j\omega_{osc}) = -180^\circ$. Supposing that K is known, and using the information from the DF to estimate $G(s)$, it produces $T=0.8003$ and $L=1.9842$. Obviously, the estimation is not very accurate.

As the previous example confirms, authors report that the use of the describing function information can introduce errors in K_{-180° of 5%-20% [3]. For this reason, there has been a lot of research work to improve the information that the describing function provides and obtain additional critical points in the Nyquist plot at different phase margins to estimate transfer functions of order higher than one.

The first try in using the describing function for estimating different types of transfer function is found in [8] with the AutoTune Variation method (ATV). This method needs a priori process information obtained from inspecting its temporal response (the steady-state gain and the dead-time) and the oscillation frequency produced by an ideal relay in a test; with all these parameters, the method allows fitting five different transfer functions models. For each of these models, expressions of magnitude and argument are derived, that is, (3) and (4). As the method uses only the ultimate gain and the ultimate frequency from the test, that is, just one critical point, to get solutions for some models is difficult and, according to the author, "there is no guarantee that any of the models will fit the data".

The Luyben's method [8] was improved by the ATV2 method [7] being only necessary to know the dead-time. In the ATV2 method, analytical expressions were obtained for the steady-state gain and time constants of transfer function models from first to third order. In this work, the estimation is afforded by working with (2) as the nonlinear part is eliminated because the dead-time is a known parameter. For each model, two equations are obtained by separating the real and imaginary parts of (2) but the number of unknowns is two, three or four, depending of the model's order. To obtain an additional critical point in the Nyquist plot, a known dead-time is added during a second test. In this way, four equations can be posed for each model and solved applying a linear least-square method. The bi-ATV method is introduced in [11]; the main difference with respect to the previous ones is the use of a biased-relay to obtain the steady-state gain from the experiment; the approach to solve the equations is similar to the original ATV. Another improvement, known as ATV+, was introduced in [10]; to avoid the prior knowledge of the dead time as in previous ATV versions, the ATV+ proposes to find an estimate of the delay through the determination of minimum and maximum bounds; using dead time values located inside the bounds, calculations for different candidate

models are repeated and, by analyzing the residual errors, the best model is then chosen. Other modification of the ATV2 method is described in [13] and better results are obtained but just in limiting cases. The approaches presented in [12] and [15] provide a simple procedure to estimate a FOPTD system with a single symmetrical relay but it has not being extended to higher order processes and an iterative method has to be applied to obtain the system parameters. Another method, known as phase deviation, is proposed in [16]; it obtains the parameters with only a relay test plus experimental measurements of the first-, third- and fifth-order harmonics of the process. Ideas of these last two references have been used in this paper. Recent contributions can be found in [9] and [6]; in both works, the information provided by the describing function is used to know the critical point and the dead-time is obtained by measuring the initial difference in the response of the process output with respect to the process input. More references related to identification based on the relay's describing function are [14] and [4]; the main contributions of these works is the design of the experiments for obtaining additional critical points at different frequencies to solve the equations.

After reviewing the literature on DF-based identification, the main problems found are summarized:

- The reciprocal of the DF provides an approximation of the Nyquist point at the frequency where the process oscillates. It is fundamental to obtain an exact value of $G(j\omega_{osc})$ during a test.
- In process without integration, it is necessary to obtain $G(0)$ as a first parameter. If the transfer function template to fit is exactly equal to the real process to identify, it is not necessary, as the identification will provide exact results at any frequency. However, if the true process has higher order or different structure to the template and $G(0)$ is not known, this will produce that the result is good around the critical frequency $\omega_{\approx-180^\circ}$ but with discrepancies at $\omega = 0$ (see Example 2). It is due to the fact that the template is fitted with lesser degrees of freedom than the true process. So, the fitting will be exact around ω_{osc} but will present discrepancies in frequencies close to zero.
- If the process has an integrator, as the identification is based on the behavior in the third quadrant, the results around $\omega_{\approx-180^\circ}$ will be good even when the transfer function has different structure than the real process. However, if the structure is different, the results will differ at low frequencies, improving the results if the identification is done at frequencies around ω_{-135° .

- It is necessary to design a procedure to get in just one test as many points of the Nyquist plane as unknowns $|G(s)|$ owns to solve the system of equations. One point will always be the oscillation point situated in the third quadrant of the Nyquist map. A second point must be $G(0)$ when needed depending of the template to fit.

- To be able of identifying at a user-defined phase lag. For example, recommendations on the margin phase of the Nyquist point to use in the identification depending of the process features are given in [4]. In [2], authors recommend estimating the process at ω_{-135° or ω_{-180° depending if PI or PID control is applied, respectively.

Example 2: An example of not taking into account $G(0)$ can be found in [16]. As the identification is done using a critical point close to ω_{-180° and

$G(s) = [(-s+1)/(s+1)^4]e^{-4s}$ is estimated using a second order model, the fitting is good enough around this Nyquist point but with an error in the steady-state gain of 24.6%. The estimated transfer function is $\hat{G}(s) = [1/(2.77s^2 + 2.28s + 1.24)]e^{-6.7s}$. This process is employed in the simulations as a study case (see Table II) and the results are improved with respect to [16].

3 The basis of the approach

As said before, the reciprocal of the relay DF is just an approximation and can introduce errors in the estimation. The solution adopted to get accurately $G(j\omega_{osc})$ during a test is presented in [15]. As $y(t)$ and $u(t)$ are periodic and piecewise, using the Laplace transform of both, it can be written

$$G(j\omega_{osc}) = \frac{Y(j\omega_{osc})}{U(j\omega_{osc})} = \frac{\int_0^{2\pi} y(t)e^{-j\omega_{osc}t} dt}{\int_0^{2\pi} u(t)e^{-j\omega_{osc}t} dt} \quad (9)$$

and following (9), and as indicated in [16], it is possible to obtain the harmonics

$$G(jn\omega_{osc}) = \frac{\int_0^{2\pi} y(t)e^{-jn\omega_{osc}t} dt}{\int_0^{2\pi} u(t)e^{-jn\omega_{osc}t} dt}, n = 1, 2, 3, \dots \quad (10)$$

Expression (10) lets solving two problems: (i) to get the value of $G(j\omega_{osc})$ that represents the exact point in the Nyquist plot where the intersection with $-1/N(A)$ is happening and, (ii) to obtain the additional points $G(jn\omega_{osc})$ needed in one test to work out the equations.

The last problem to solve is the estimation of the steady gain. Expression (9) cannot be applied as the oscillations produced by a relay are symmetric and the integration of the semi periods will be zero. However, if an asymmetry is introduced in the oscillations by adding a small bias to the relay output (see Figure 1), the process static gain can be derived from (9) [11] as

$$G(0) = \frac{\int_0^{2\pi} y(t) dt}{\int_0^{2\pi} u(t) dt} \quad (11)$$

It is important to notice that without bias, the signals are symmetric and even harmonics become zero.

3.1 Models

The identification procedure can be adapted to most of the process found in process industry just obtaining the magnitude and argument expressions of the transfer function to fit. Expressions for a first order plus time delay (FOPTD), an overdamped second-order with time delay (SOPTD-1), a second-order with time delay (SOPTD-2) to identify underdamped processes, and an integrating process with inverse response and time delay (IPIRTD) are given.

Model 1: FOPTD

$$G_1(s) = \frac{Ke^{-Ls}}{Ts+1} \quad (12)$$

Model 2: SOPTD-1

$$G_2(s) = \frac{Ke^{-Ls}}{(Ts+1)^2} \quad (13)$$

Model 3: SOPTD-2

$$G_3(s) = \frac{Ke^{-Ls}}{as^2 + bs + 1} \quad (14)$$

Model 4: IPIRTD

$$G_4(s) = \frac{K(1-T_1s)e^{-Ls}}{s(T_2s+1)} = \frac{(1-T_1s)e^{-Ls}}{as^2 + bs} \quad (15)$$

where $K = 1/b$ and $T_2 = a/b$.

From each of these models, the expressions of magnitude and argument of $G(j\omega_{osc})$ are derived.

Model 1: FOPTD

$$|G_1(j\omega_{osc})| = \frac{K}{\sqrt{T^2\omega_{osc}^2 + 1}} \quad (16)$$

$$\arg G_1(j\omega_{osc}) = -\arctan(T\omega_{osc}) - \omega_{osc}L \quad (17)$$

where $\arg G_1(j\omega_{osc}) \in [-2\pi, 0]$.

As K , ω_{osc} and $G_1(j\omega_{osc})$ are obtained directly from the test using (10) and (11) by adding a bias, T can be obtained directly from (16). After that, L is derived from (17).

Model 2: SOPTD-1

$$|G_2(j\omega_{osc})| = \frac{K}{T^2\omega_{osc}^2 + 1} \quad (18)$$

$$\arg G_2(j\omega_{osc}) = -2\arctan(T\omega_{osc}) - \omega_{osc}L \quad (19)$$

The unknowns are the same that in the previous one, so the procedure to apply is the same.

Model 3: SOPTD-2

$$|G_3(j\omega_{osc})| = \frac{K}{\sqrt{a^2\omega_{osc}^4 + b^2\omega_{osc}^2 - 2a\omega_{osc}^2 + 1}} \quad (20)$$

$$\arg G_3(j\omega_{osc}) = \arctan(-b\omega_{osc}, -a\omega_{osc}^2 + 1) - \omega_{osc}L \quad (21)$$

There are three unknowns K , a , and b in (14). As K is obtained directly from the test using (11) by adding a bias, to get the other two unknowns is necessary the expression corresponding to the second harmonics, that is, $|G_3(j2\omega_{osc})|$. This expression is derived from (20) just replacing ω_{osc} by $2\omega_{osc}$ and its experimental value is obtained from the test (see (10)). Once a and b are known, the dead-time is derived from (21).

Model 4: IPIRTD

$$|G_4(j\omega_{osc})| = \sqrt{\frac{T_1^2\omega_{osc}^2 + 1}{\omega_{osc}^2(a^2\omega_{osc}^2 + b^2)}} \quad (22)$$

$$\arg G_4(j\omega_{osc}) = -\arctan\left(\frac{-(T_1a\omega_{osc}^2 - b)}{-\omega_{osc}(T_1b + a)}\right) - \omega_{osc}L \quad (23)$$

The unknowns in (22) are T_1 , a , and b . As it is a process with integration, the velocity gain can directly be obtained solving the system of equations. So, the expressions of $|G_4(j2\omega_{osc})|$ and $|G_4(j3\omega_{osc})|$ are needed and its experimental values are got from the test (see (10)). The dead-time is directly got from (23) once the other three unknowns are obtained.

The following expressions are the result of solving the previous equations for the four models. For the

sake of simplicity, C_n represents $|G(jn\omega_{osc})|$ and $\arg C_1$ corresponds to $\arg G(j\omega_{osc})$.

Model 1: FOPTD

$$T = \frac{\sqrt{K^2 - C_1^2}}{\omega_{osc} C_1} \quad (24)$$

$$L = -\frac{\arg C_1 + \arctan(\omega_{osc} T)}{\omega_{osc}} \quad (25)$$

Model 2: SOPTD-1

$$T = \frac{\sqrt{C_1(K - C_1)}}{\omega_{osc} C_1} \quad (26)$$

$$L = -\frac{2 \arctan(T\omega_{osc}) + \arg C_1}{\omega_{osc}} \quad (27)$$

Model 3: SOPTD-2

$$\begin{aligned} aux &= \sqrt{(3C_1^2 - 12C_2^2)K^2 + 9C_1^2 C_2^2} \\ a &= \frac{aux}{6\omega_{osc}^2 C_1 C_2} \end{aligned} \quad (28)$$

$$b = \frac{\sqrt{-3C_1^2 K^2 + 48C_2^2 K^2 - 45C_1^2 C_2^2 + 12C_1 C_2 aux}}{6\omega_{osc} C_1 C_2} \quad (29)$$

$$L = \frac{\arctan(-b\omega_{osc}, -a\omega_{osc}^2 + 1) - \arg C_1}{\omega_{osc}} \quad (30)$$

Model 4: IPIRTD

$$T_1 = \frac{1}{\omega_{osc}} \sqrt{\frac{C_1^2 C_2^2 - 6C_1^2 C_3^2 + 15C_2^2 C_3^2}{-9C_1^2 C_2^2 + 24C_1^2 C_3^2 - 15C_2^2 C_3^2}} \quad (31)$$

$$a = \frac{1}{\omega_{osc}^2} \sqrt{\frac{-5C_1^2 + 32C_2^2 - 27C_3^2}{-108C_1^2 C_2^2 + 288C_1^2 C_3^2 - 180C_2^2 C_3^2}} \quad (32)$$

$$b = \frac{1}{\omega_{osc}} \sqrt{\frac{5C_1^2 - 128C_2^2 + 243C_3^2}{-108C_1^2 C_2^2 + 288C_1^2 C_3^2 - 180C_2^2 C_3^2}} \quad (33)$$

$$L = -\frac{\arctan(-(T_1 a \omega_{osc}^2 - b), -\omega_{osc}(T_1 b + a)) + \arg C_1}{\omega_{osc}} \quad (34)$$

3.2. Outline of the procedure

The procedure is based on producing a limit cycle oscillating at ω_{osc} as consequence of the relay control actions generate when $G(s)$ crosses zero. As shown before, mathematically, the limit cycle is explained as the consequence of the intersection of

$G(s)$ with $-1/N(A)$ at ω_{osc} . The intersection point corresponds to the value of $G(j\omega_{osc})$.

The estimation procedure can be divided into the following steps:

1. Add a small bias, i.e, $\approx 0.1D$ to the relay output.
2. Once the oscillation is stable, annotate ω_{osc} , and the harmonics $G(jn\omega_{osc})$ needed to solve the equations. In models without integration, it is necessary K .
3. Use the expressions:
 - (24) and (25) for FOPTD model.
 - (26) and (27) for SOPTD-1 model.
 - (28), (29), and (30) for SOPTD-2 model.
 - (31), (32), (33), and (34) for IPIRTD model.

Remark 1: Model 3 lets detecting if the true process corresponds to a dynamics of first order instead of second order. In this case, the procedure will generate a value for a close to zero.

Remark 2: Model 2 and 3 can produce the same result if the true process corresponds to an overdamped second order system. In case of being underdamped (non-monotone), models 1 and 2 can produce a complex value for the time lag or an overdamped solution, and the correct results are produced only by Model 3.

3.3 Simulation examples

The simulations have been run in Matlab/Simulink with a sampling time $h=0.001$. The parameters used in the tests have been $D=1$ and $bias=0.1D$.

Table I shows the results when the structure of the true process and the model are the same. Models 2 and 3 produce similar solutions for Cases 2 and 3 as both real processes correspond to overdamped systems. However, in Case 4, model 2 produces complex values as the true process is underdamped. The identification in Case 5 provides a complex value for T_1 close to zero (0.08i) due to the numerical errors; and it is ignored and equated to zero.

Table II corresponds to the results when the structure of the true process is different to the transfer function to fit. Cases 7 and 8 correspond to the fitting of FOPTD models; cases 9, 10, and 11 to second order processes, and cases 12 and 13 to processes with integration and inverse response. In Case 9, it is not possible to present solutions for models 1 and 2 as the true process is an underdamped system (non-monotone). Case 10 corresponds to the process used in Example 2 but the steady gain is correctly

estimated (see Figure 2). Although the DF theory establishes that the oscillation produced by a simple relay happens with a phase lag of -180° , it must be noticed in the simulations how the phase margin changes for each process. The higher discrepancies happen in processes with non-minimum phase dynamics as the process output becomes less sinusoidal.

Table I: Solutions for processes with same order and structure than the transfer function fitted.

	Real process	Proposed method
1	$\frac{e^{-2s}}{10s+1}$	$\frac{1.003e^{-2.0008s}}{10.0364s+1}$ $\varphi_m \approx -0.78^\circ$
2	$\frac{e^{-2s}}{(2s+1)^2}$	$\frac{1.0003e^{-2.004s}}{4.024s^2+4.007s+1}$ $\frac{1.0003e^{-1.996s}}{(2.0043s+1)^2}$ $\varphi_m \approx -0.61^\circ$
3	$\frac{e^{-0.2s}}{0.5s^2+1.5s+1}$	$\frac{1.001e^{-0.2s}}{0.5004s^2+1.502s+1}$ $\frac{1.001e^{-0.189s}}{(0.7188s+1)^2}$ $\varphi_m \approx 2.27^\circ$
4	$\frac{e^{-s}}{2s^2+s+1}$	$\frac{1.002e^{-0.98s}}{2.013s^2+0.9946s+1}$ $\varphi_m \approx 0.6^\circ$
5	$\frac{e^{-10s}}{s(20s+1)}$	$\frac{0.999e^{-10.0034s}}{s(19.994s+1)}$ $\varphi_m \approx 2.88^\circ$
6	$\frac{(-2.4s+0.6)}{s(s+1)}$	$\frac{(-2.399s+0.6002)e^{-0.00269s}}{s(0.9995s+1)}$ $\varphi_m \approx 21.5^\circ$

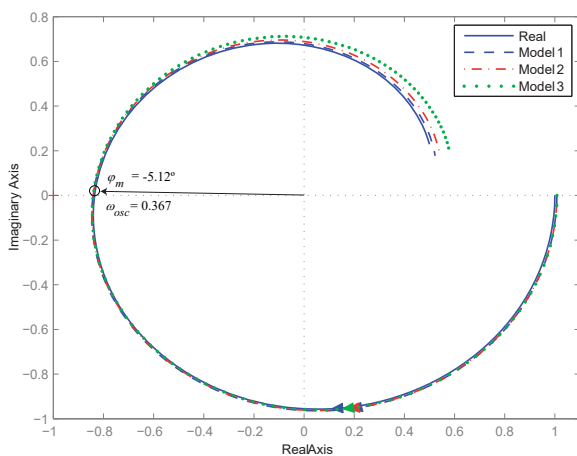


Figure 2: Nyquist curves of real process and models for Case 10.

Table II: Solutions for processes with higher order or different structure than the transfer function fitted.

	Real process	Proposed method
7	$\frac{e^{-2s}}{(10s+1)(s+1)}$	$\frac{1.002e^{-2.8632s}}{11.6998s+1}$ $\frac{1.002e^{-1.2824s}}{(4.129s+1)^2}$ $\frac{1.002e^{-2.0011s}}{10.013s^2+11.02s+1}$ $\varphi_m \approx 0.62^\circ$
8	$\frac{e^{-0.5s}}{(s+1)^5}$	$\frac{1.002e^{-3.1787s}}{3.3228s+1}$ $\frac{1.002e^{-2.254s}}{(1.8206s+1)^2}$ $\varphi_m \approx -0.01^\circ$
9	$\frac{e^{-6s}}{(s+3)(s^2+s+1)}$	$\frac{0.3336e^{-6.28s}}{0.9992s^2+1.05s+1}$ $\varphi_m \approx 3.58^\circ$
10	$\frac{(-s+1)e^{-4s}}{(s+1)^4}$	$\frac{1.008e^{-7.164s}}{1.855s+1}$ $\frac{1.0008e^{-6.305s}}{1.934s^2+2.609s+1}$ $\frac{1.0008e^{-6.4526s}}{(1.2479s+1)^2}$ $\varphi_m \approx -5.12^\circ$
11	$\frac{(-10s+1)e^{-s}}{s(s+1)^5}$	$\frac{(-10.36s+0.9824)}{s(2.646+1)} e^{-3.12}$ $\varphi_m \approx 12.5^\circ$
12	$\frac{(-5s+1)e^{-0.5s}}{s(s+1)(s^2+s+1)}$	$\frac{(-5.167+0.9768)}{s(0.5523s+1)} e^{-1.83s}$ $\varphi_m \approx 14^\circ$

4 Selecting the identification phase margin

The identification approach using a relay without hysteresis produces a model that corresponds exactly with the dynamics of the true process at the oscillation frequency, that is $G(j\omega_{osc}) = \hat{G}(j\omega_{osc})$. If the order of the process is equal to the transfer function model to fit, the identification will be exact and the behaviour of $\hat{G}(s)$ will be equal to $G(s)$ in all the frequencies range. However, if the order of process and model is different, it can be necessary to modify the oscillation frequency to find a more suitable approximation to the pursued control purposes. It is known that for PI control the identification should be done at an oscillation frequency that corresponds to a phase margin $\varphi_m = 45^\circ$ due to the phase lag that the PI control

introduces. For PID control the phase margin should be 0° as the controller provides phase lead [2].

In the previous examples, the phase margin is not 0° and changes depending of the features of the process. Theoretically, the bias added to the output cannot be used for this purpose as it does not affect to the phase margin as not provide phase shifting (see [5]),

$$\frac{-1}{N(A)+bias} = -\frac{\pi A}{\pi A \cdot bias + 4D} \quad (35)$$

To obtain a point in the Nyquist curve different to the obtained with the simple relay can be done by introducing hysteresis in the relay ([1]). The negative reciprocal of the DF of a biased relay with hysteresis is

$$\frac{-1}{N(A,\delta)} = -\frac{\pi}{4D} \sqrt{A^2 - \delta^2} - i \frac{\pi\delta}{4D} \quad (36)$$

where A is the oscillation output, D the relay amplitude, and δ the hysteresis (it must be noticed that the reciprocal is independent of the bias). As happens in the simple relay, this function is represented in the Nyquist map as a straight line parallel to the real axis. However, by increasing δ , the line can be move down along the imaginary axis of the Nyquist map. As the intersection of $G(s)$ with this line means the existence of an oscillation, the critical point where the intersection happens can be modified, increasing the phase margin to fulfill the user specifications.

From (36), the theoretically phase margin obtained by applying hysteresis to the relay is defined by

$$\varphi_m = \arctan\left(\frac{\delta}{\sqrt{A^2 - \delta^2}}\right) = \arcsin\left(\frac{\delta}{A}\right) \quad (37)$$

where $A \geq \delta$. Obviously, this expression provides just an approximation that depends on the features of the true process.

Example 3: The result of the identification of $G(s) = e^{-s}/s(s+1)$ using a relay with $D=1$, $\delta = 0$, and $bias=0.1D$ is

$$\hat{G}1(s) = \frac{(-0.1614s + 1.051)e^{-0.781s}}{s(1.139s + 1)}$$

with $\varphi_m \approx 2.66^\circ$. If the hysteresis is increased to $\delta = 2$, the result of the identification is

$$\hat{G}2(s) = \frac{(-0.08321s + 1.001)e^{-0.929s}}{s(1.005s + 1)}$$

with $\varphi_m \approx 46^\circ$. In Figure 3, it can be appreciated the differences between the two results. The differences are explained by the fact that the identification algorithm is based on the fitting of the curve that represents a transfer function template to the curve that corresponds to the true process in the Nyquist map using the critical point as linkage between both curves. If model and process are similar (same degrees of freedom), the fitting is exact as the transfer function curve can be adjusted to the true process curve. However, if model and process own different structures, then the model curve cannot be adjusted exactly in all the frequency range to the process. In the Example 3, as the true process is with integration, the identification with a phase margin around 45° or higher will produce better results as the model curve adapts better to the real process in the third quadrant.

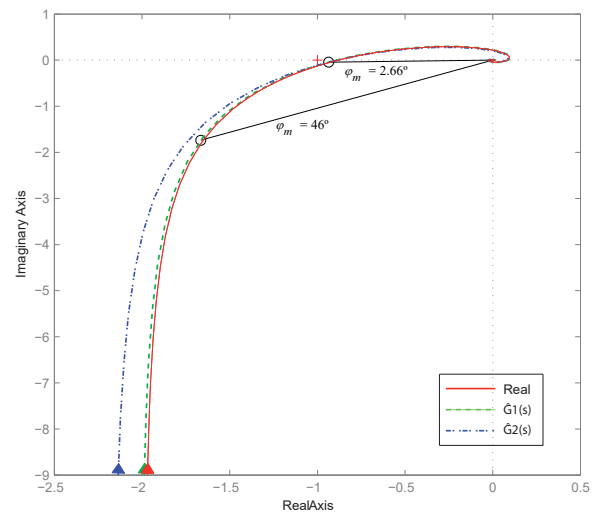


Figure 3: Example of the differences in the identification results depending of the phase margin.

4.1. Automatic generation of the hysteresis

In the previous example, the hysteresis was fixed by trial and error in successive simulations. A solution proposed in [1] for autotuning of PID controllers can be adapted to adjust iteratively the hysteresis during the tests using the information obtained from the experimental measure of $G(j\omega_{osc})$. The formula is

$$\delta_{n+1} = \delta_n + \lambda(\varphi_{ref} - \varphi_n) \frac{\delta_n - \delta_{n-1}}{\varphi_n - \varphi_{n-1}} \quad (38)$$

where φ_{ref} is the desired phase margin and $\lambda \leq 1$ is a convergence speed factor.

Example 4: Applying the iterative method to the identification of the process of Example 3 using (38), fixing $\varphi_{ref} = 45^\circ \pm 1^\circ$ and $\lambda = 0.5$, it is obtained $\delta = 1.897$ and the new identified model is

$$\hat{G}_3(s) = \frac{(-0.02286s + 1)e^{-0.979s}}{s(s+1)} \quad (39)$$

with $\varphi_m \approx 45.32^\circ$.

Conclusions

An enhanced method for the identification of stable transfer functions models has been presented and explained. The basis of the method is the oscillations that a relay generates in the feedback control loop. Instead of using the approximations of the intersection points that the describing function of the relay provides, on-line measurements of the process and control signals are used to obtain the harmonics needed to solve the linear equations and generate the estimations. In this way, the number of tests to make the estimations is always one regardless the order or structure of the model transfer function to fit. Moreover, the identification procedure is not iterative so the computational cost is very low. With the simple relay, the identification procedure is done near the ultimate frequency. By using a relay with hysteresis and modifying its value, the identification will be possible at a user-specified phase lag in the third quadrant.

Acknowledgements

This work has been funded by Spanish Ministry of Economy and Competitiveness under contracts DPI2012-31303 and DPI2014-55932-C2-2-R.

References

- [1] Åström, K.J., Hägglund T. (1984) "Automatic tuning of simple regulators with specifications on phase and amplitude margins", *Automatica* 20(5), pp. 645-651.
- [2] Åström, K.J., Hägglund, T. (2006) *Advanced PID Control*, ISA, 460 pages.
- [3] Chang, R.C., Shen, S.H., and Yu, C.C. (1992) "Derivation of transfer function from relay feedback systems", *Industrial & Engineering Chemistry Research* 31, pp. 855-860.
- [4] Friman, M., Waller, K.V. (1997) "A two-Channel relay for autotuning", *Industrial & Engineering Chemistry Research* 36(7), pp. 2662-2671.
- [5] Gelb, A., Van der Velde, W.E. (1968) *Multiple-Input Describing Functions and Nonlinear System Design*, McGraw-Hill, New York, USA.
- [6] Ghorai, P., Majhi, S., and Pandey, S. (2016) "Dynamic model identification of a real-time simple level control system", *Journal of Control and Decision* 3(4).
- [7] Li, W., Eskinat, E., and Luyben, W.L. (1991) "An improved autotune identification method", *Industrial & Engineering Chemistry Research* 30(7), pp. 1530-1541.
- [8] Luyben, W.L. (1987) "Derivation of transfer functions for highly nonlinear distillation columns", *Industrial & Engineering Chemistry Research* 26(12), pp. 2490-2495.
- [9] Pandey, S., Majhi, S. (2016) "Limit cycle based identification of second order processes with time delay", *Indian Control Conference* (2016), pp. 438-443
- [10] Scali, C., Marchetti, G., and Semino, D. (1999) "Relay with additional delay for identification and autotuning of completely unknown processes", *Industrial & Engineering Chemistry Research* 38 (5), pp. 1987-1997.
- [11] Shen, S.H., Wu, J.S., and Yu, C.C. (1996) "Use of biased-relay feedback for system identification", *AIChE Journal* 42(4), pp. 1174-1180.
- [12] Srinivasan, K., Chidambaram, M. (2003) "Modified relay feedback method for improved system identification", *Computers & Chemical Engineering* 27(5), pp. 727-732.
- [13] Srinivasan, K., Chidambaram, M. (2004) "An improved autotune identification method", *Chemical and Biochemical Engineering Quarterly* 18(3), pp. 249-256.
- [14] Tan, K.K., Lee, T.H., and Wang, Q.G. (1996) "An enhanced automatic tuning procedure for PI/PID controllers for process control". *AIChE Journal* 42(9), pp. 2555- 2562.
- [15] Vivek, S., Chidambaram, M. (2005) "Identification using single symmetrical relay feedback test". *Computers & Chemical Engineering* 29(7), pp. 1625-1630
- [16] Wang, P., Gu, D., and Zhang, W. (2007) "Modified relay feedback identification based on describing function analysis". *Industrial & Engineering Chemistry Research* 46(5), pp. 1538-1546.