Abstract—The relay feedback test is modified the way, so that the ultimate frequency in the test coincides with the phase cross-over frequency in the open-loop system having a PID controller. This allows for formulation simple non-parametric tuning rules for PID controllers that provide desired gain margins exactly. Tuning rules for a PI controller and simulation examples are provided.

I. INTRODUCTION

D espite the success of some types of advanced process control, the PID control still remains the main type of control used in the process industries. PID controllers are usually implemented as stand-alone controllers or configurable software modules within the distributed control systems (DCS). The DCS software is constantly evolving providing a number of new features, among which the controller autotuning functionality is one of most useful. It is found in the latest releases of such popular DCS as Honeywell Experion PKS® and Emerson DeltaV®. The practice of the use of of autotuning algorithms shows that many of them do not provide a satisfactory performance if the process is subject to noise, variable external disturbance or non-linear. On the other hand the simplest algorithms such as Ziegler-Nichols’s closed-loop tuning method [1] and Astrom-Hagglund’s relay feedback test (RFT) [2] provide a satisfactory performance in those conditions despite the inherent relatively low accuracy of those methods. Apparently, the use of an underlying model of the process in a parametric method (not fully matching the actual process dynamics) that usually has three or higher number of parameters may result in the significant deterioration of the identification-tuning accuracy if the test conditions are affected by noise, disturbances or nonlinearities. Only the most basic characteristics of the system, such as the ultimate gain and ultimate frequency [1], remain nearly unchanged in those conditions. However, the use of only ultimate gain and frequency cannot ensure sufficient accuracy of tuning. Therefore, a trade-off between the accuracy and reliability of tuning (which also translates into accuracy) is apparent. The cause of the relatively low accuracy of [1], [2] and other non-parametric methods is well known. This is the use of only two measurements of the test over the process, which would be equivalent to the use of a parametric method with a model having only two parameters, which is insufficient for obtaining an acceptable precision [3].

There is one more factor that also contributes to the issue of precision. This is a popular statement that says that the most important test point in the closed-loop test is the one in which the phase characteristic of the process is –180° (frequency ωc). However, this approach does not account for the change of frequency ωc due to the controller introduction, which is the factor that contributes to the precision deterioration (it is analyzed below).

The paper is organized as follows. At first the problem of selection of the test point on the frequency response of the process is analyzed. After that a modified RFT that provides generation of the oscillations at a given point of the phase response of the process is proposed. Finally, the tuning rules that provide a higher precision of non-parametric tuning than the rules of [1] and [2] are derived, and the performance of the test and tuning is analyzed.

II. EFFECT OF CONTROLLER INTRODUCTION ON STABILITY IN ORIGINAL PROCESS AND ITS APPROXIMATION

It has been a popular notion that the most important point on the frequency response of the system is the point where the phase characteristic of the process is equal to –180° (frequency ωc). We shall also refer to this point as the phase cross-over frequency, as the phase characteristic of the plant (or plant and controller) crosses the line –180°. However, this point remains the most important one only in the system with the proportional controller. This circumstance is often neglected, and this rule is applied to all types of PID control. Let us consider the following motivating example and analyze how the introduction of the controller may affect the results of identification and tuning.

Example 1. Let us assume that the process is given by the following transfer function (which was used in a number of works as a test process):

\[ W_p(s) = e^{-2s} \frac{1}{(2s + 1)^3}, \]

(1)

Find the first order plus dead time (FOPDT) approximating model \( \tilde{W}_p(s) \) to the process (1) based on matching the values of the transfer functions at frequency \( \omega_c \):

\[ \tilde{W}_p(s) = \frac{K_p e^{-\tau s}}{T_p s + 1}, \]

(2)

where \( K_p \) is the process static gain, \( T_p \) is the time constant, and \( \tau \) is the dead time. Let us note that both (1) and (2)
produce the same ultimate gain and ultimate frequency in the Ziegler-Nichols closed-loop test [1] or the same values of the amplitude and the ultimate frequency in the RFT [2]. (Note: strictly speaking, the values of the ultimate frequency in tests [1] and [2] are slightly different, as the frequency of the oscillations generated in the RFT does not exactly correspond to the phase characteristic of the process –180°; this fact follows from the relay systems theory [4], [5]). Obviously, this problem has infinite number of solutions, as there are three unknown parameters of (2) and only two measurements obtained from the test. Assume that the value of the process static gain is known: \( K_p = 1 \), and determine \( T_p \) and \( r \). Those parameters can be found from equation
\[
\hat{W}_p(j\omega_r) = W_p(j\omega_r),
\]
where \( \omega_r \) is the phase cross-over frequency for both transfer functions. Therefore, \( \arg W_p(j\omega_r) = -\pi \). The value of \( \omega_r \) is 0.283, which gives \( W_p(j\omega_r) = (-0.498, 0) \), and the FOPDT approximation is, therefore (found via solution of the set of two algebraic equations):
\[
\hat{W}_p(s) = \frac{e^{-7.393s}}{6.153s + 1}.
\] (3)

The Nyquist plots of the process (1) and its approximation (3) are depicted in Fig. 1. The point of intersection of the two plots (denoted as \( \Omega_0 \)) is also the point of intersection with the real axis. Also \( \Omega_0 = \omega_0 \) for both process dynamics (1) and (3), and therefore \( \hat{W}_p(j\Omega_0) = W_p(j\Omega_0) \). If the designed controller is of proportional type then the gain margins for processes (1) and (3) are the same. However, if the controller is of PI type then the stability margins for (1) and (2) are different. Let us illustrate that. Design the PI controller given by the following transfer function:
\[
W_c(s) = K_i \left(1 + \frac{1}{T_c s}\right),
\]
using the Ziegler-Nichols tuning rules [1]. This results in the following transfer function of the controller:
\[
\hat{W}_p(s) = W_p(s)\hat{W}_p(j\Omega_0),
\]
where \( \hat{W}_p(j\Omega_0) = (-0.408, 0) \). The Nyquist plots of the open-loop systems containing the controller are shown in Fig. 2. The mapping of point \( \Omega_0 \) in Fig. 1 into point \( \Omega_0 \) in Fig. 2 is done via clockwise rotation of vector \( \hat{W}_p(j\Omega_0) \) by the angle \( \psi = \arctan\left(1/(0.8 \cdot 2\pi)\right) = 11.25° \) and multiplication of its length by such value, so that its length becomes equal to 0.408. However, for the open-loop system containing the PI controller, the points of intersection of the Nyquist plots of the system and of the real axis are different for the system with process (1) and with process approximation (3). They are shown as points \( \Omega_1 \) and \( \Omega_2 \) in Fig. 2. The mapping of those points to the Nyquist plots of the process and its approximation is shown in Fig. 1. Therefore, the stability margins of the systems containing a PI controller is not the same any more. It is revealed as different points of intersection of the plots and of the real axis in Fig. 2. In fact the position of vector \( \hat{W}_p(j\Omega_0) = \hat{W}_c(j\Omega_0)\hat{W}_p(j\Omega_0) \) is fixed, but this vector does not reflect on the stability of the system. As one can see in Fig. 2, the gain margin of the system containing the FOPDT approximation of the process is higher than the one of the system with the original process.

The considered example illustrates a fundamental problem of all methods of identification-tuning based on the measurements of process response in the critical point (\( \Omega_0 \)). This problem is the shift of the critical point due to the introduction of the controller. The question that follows from the above analysis is whether the test point can be selected in a different way, so that the introduction of the controller would be accounted for in the test itself. And if this is possible then what kind of test it should be to ensure the measurements in the desired test point.

Address the first question now. Assume that we can design a certain test, so that we can assign the test point at...
the desired phase lag of the process $\arg W_p(j\Omega_0) = \varphi$, where $\varphi$ is a given quantity, and measure $W_p(j\Omega_0)$ in this point. Consider the following example.

**Example 2.** Let the plant be the same as in Example 1. Assume that the introduction of the controller will be equivalent to the mapping similar to the mapping described above – the vector of the frequency response of the open-loop system in the point $\Omega_0$ will be a result of clockwise rotation of the vector $\tilde{W}_p(j\Omega_0)$ by a known angle and multiplication by a certain known factor: $\tilde{W}_p(j\Omega_0) = \tilde{W}_c(j\Omega_0) \tilde{W}_p(j\Omega_0)$. Also assume that the controller will be the same as in Example 1 (for illustrative purpose - because the tuning rules are not formulated yet). Therefore, let us find the values of $T_p$ and $\tau$ for the transfer function (2) (we still assume $K_c=1$) that ensure that the equality $\tilde{W}_p(j\Omega_0) = W_p(j\Omega_0)$ holds, where $\arg W_p(j\Omega_0) = -180^\circ + 11.25^\circ = -168.75^\circ$ (the angle is selected considering the subsequent clockwise rotation by 11.25°). Therefore, $\Omega_0 = 0.263$, and $W_p(j\Omega_0) = (-0.532, -j0.103)$. The corresponding FOPDT approximation of the process is

$$\hat{W}_p(s) = \frac{e^{-7.293s}}{5.897s + 1}, \quad (6)$$

One can notice that both the time constant and the dead time in (6) are smaller than in (3). Application of controller (5) shifts the point $\Omega_0$ of intersection of $W_p(j\Omega_0)$ and $\tilde{W}_p(j\Omega_0)$ to the real axis. This point remains the point of intersection of the two Nyquist plots. Therefore, the gain margin of both systems: with the original process and with the approximated process are the same. Consider now the problem of the design of the test that can provide matching the points of the actual and approximating processes in the point corresponding to a specified phase lag.

### III. MODIFIED RFT

Consider the following discontinuous control:

$$u(t) = \begin{cases} 
  h & \text{if } \sigma(t) \geq \Delta_1 \text{ or } (\sigma(t) > -\Delta_2 \text{ and } u(t-) = h) \\
  -h & \text{if } \sigma(t) \leq \Delta_2 \text{ or } (\sigma(t) < \Delta_1 \text{ and } u(t-) = -h)
\end{cases} \quad (7)$$

where $\Delta_1 = \beta \sigma_{\max}$, $\Delta_2 = -\beta \sigma_{\min}$, $\sigma_{\max}$ and $\sigma_{\min}$ are last “singular” points of the error signal (Fig. 3) corresponding to last maximum and minimum values of $\sigma(t)$ after crossing the zero level, $\beta$ is a positive constant. The algorithm (7) is similar to the so-called “generalized sub-optimal” algorithm used for generating a second-order sliding mode in systems of relative degree two [7], [8]. The difference is that the generalized sub-optimal algorithm involves an advance switching of the relay but the proposed algorithm involves a lagged switching of the relay.

![Fig. 3. Relay feedback test](image)

![Fig. 4. Finding periodic solution](image)

Let the reference signal $r(t)$ be zero in Fig. 3. Let us show that in the steady mode, the motions in the system Fig. 3, where the control is given by (7) are periodic. Apply the describing function (DF) method [9] to the analysis of motions in Fig. 3. Assume that the steady mode periodic, and prove that this is a valid assumption by finding parameters of this periodic motion. If the motions in the system are periodic then $\sigma_{\max}$ and $\sigma_{\min}$ represent the amplitude of the oscillations: $a = \sigma_{\max} = -\sigma_{\min}$, and the equivalent hysteresis value of the relay is $\Delta = \Delta_1 = \Delta_2 = \beta \sigma_{\max} = -\beta \sigma_{\min}$. The DF of the hysteretic relay is given as follows

$$N(a) = \frac{4h}{\pi a} \sqrt{1 - \frac{\Delta^2}{a^2}} - j \frac{4h \Delta}{\pi a^2}, \quad a > \Delta \quad (8)$$

However, system Fig. 3 with control (7) is not a conventional relay system. This system has the hysteresis value that is unknown a-priori and depends on the amplitude value: $\Delta = \beta a$. Therefore, (8) can be rewritten as follows:

$$N(a) = \frac{4h}{\pi a} \sqrt{1 - \beta^2 - j\beta}, \quad (9)$$

The RFT will generate oscillations in the system under control (7). We shall further refer to that test as to “modified relay feedback test”. Parameters of those oscillations can be found from the harmonic balance equation:

$$W_p(j\Omega_0) = -\frac{1}{N(a_0)}, \quad (10)$$

where $a_0$ is the amplitude of the periodic motions, and the negative reciprocal of the DF is given as follows:

$$-\frac{1}{N(a)} = -\frac{\pi a}{4h} \sqrt{1 - \beta^2 + j\beta} \quad (11)$$

Finding a periodic solution in system Fig.3 with control (7) has a simple graphic interpretation (Fig. 4) as finding the point of intersection of the Nyquist plot of the process and of the negative reciprocal of the DF, which is a straight line.
that begins in the origin and makes a counterclockwise angle \( \psi = \arcsin \beta \) with the negative part of the real axis.

In the problem of analysis, frequency \( \Omega_0 \) and amplitude \( a_0 \) are unknown variables and are found from the complex equation (10). In the problems of identification and tuning, \( \Omega_0 \) and \( a_0 \) are measured from the modified RFT, and on the basis of the measurements obtained either parameters of the underlying model are calculated (for parametric tuning) or tuning parameters are calculated immediately from \( \Omega_0 \) and \( a_0 \) (for non-parametric tuning).

Reviewing again Example 2, we can note that, for example, Ziegler-Nichols tuning rules are meant to be applied, and the subsequent transformation via introduction of the PI controller involving clockwise rotation by angle \( \psi = \arctan \frac{1}{0.8 \cdot 2\pi} = 11.25^\circ \) is going to be applied, then parameter \( \beta \) of the controller for the modified RFT should be \( \beta = \sin 11.25^\circ = 0.195 \). The modified RFT also allows for the exact design of the gain margin (assuming the DF method provides exact model). Since the amplitude of the oscillations \( a_0 \) is measured from the test, the process gain at frequency \( \Omega_0 \) can be obtained as follows:

\[
\left| W_p(j\Omega_0) \right| = \frac{a_0}{4h},
\]

which after introduction of the controller will become the process gain at the critical frequency. Therefore, if the tuning rules are given in the format:

\[
K_c = c_1 \frac{4h}{a_0}, \quad T_c = c_2 \frac{2\pi}{\Omega_0},
\]

where \( c_1 \) and \( c_2 \) are parameters that define the tuning rule, then the frequency response of the controller at \( \Omega_0 \) becomes

\[
W_c(j\Omega_0) = c_1 \frac{4h}{a_0} \left(1 - j \frac{1}{2\pi} \right),
\]

and for obtaining the gain margin \( \gamma \) \((\gamma > 1)\) parameter \( c_1 \) should be selected as

\[
c_1 = \frac{1}{\gamma \sqrt{1 + 1/4\pi^2 c_2^2}}
\]

In the considered example, if we keep parameter \( c_2 \) the same as [1]: \( c_2 = 0.8 \), then to obtain, for example, gain margin \( \gamma = 2 \) tuning parameter \( c_1 \) for the modified RFT should be selected \( c_1 = 0.49 \). Any process regardless of the actual dynamics will have gain margin \( \gamma = 2 \) (6dB) exactly (within the framework of the filtering hypothesis of the DF method). Therefore, the modified RFT can ensure the stability of the closed-loop system with a controller, and also provide the desired gain margin.

IV. Precise treatment of modified RFT

The DF model provides a simple but approximate model of the periodic motions. A more accurate model would still be desirable. Yet, the relay system Fig. 3 is not a conventional relay system, as the value of the hysteresis is unknown. Therefore, methods [4], [5] and others cannot be directly applied to that system. To provide an exact treatment of the periodic problem in system Fig. 3 with controller (7), introduce the following function:

\[
\Phi(\omega) = \pm \left[ a_\varphi^2 (\omega) - y^2 \left( \frac{\pi}{\omega}, \omega \right) + jy \left( \frac{\pi}{\omega}, \omega \right) \right],
\]

where \( a_\varphi \) is the amplitude of the oscillation of \( y(t) \) of the process output when a periodic square-pulse control of frequency \( \omega \) and amplitude \( h \) is applied to the process, \( y \left( \frac{\pi}{\omega}, \omega \right) \) is the value of the process output at the time instant corresponding to the relay switch from \(-h \) to \( h \) \((\pi \omega \text{ is half the period in the periodic motion, and } t = 0 \text{ is the time of the relay switch from } h \text{ to } -h)\).

In the problem of analysis, frequency \( \omega \) occurs within the time interval \( t \in \left(0; \frac{\pi}{\omega}\right) \), and the sign "-" is applied if the maximum of \( y(t) \) occurs within the time interval \( t \in \left(\frac{\pi}{\omega}; \frac{2\pi}{\omega}\right) \). \( y(t, \omega) \) can be computed by means of its Fourier series as follows:

\[
y(t, \omega) = 4h \sum_{k=1}^{\infty} \frac{1}{\pi} \sin \left( \frac{\pi k}{2} \right) \sin[k\omega t + \varphi(k\omega)] L(k\omega)
\]

\[
= 4h \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{2k-1} \sin[(2k-1)\omega t + \varphi(2k-1)\omega)] L((2k-1)\omega)
\]

where \( \varphi(\omega) = \arg W_p(j\omega) \), \( L(\omega) = \left| W_p(j\omega) \right| \) are the phase and magnitude of \( W_p(j\omega) \) at frequency \( \omega \).

The frequency-dependent variable \( a_\varphi(\omega) \) can be computed via (17) and (18), and \( y(\pi/\omega, \omega) \) can be evaluated via the Tsypkin locus [4] or the LPRS [5]. This formula can also be obtained immediately from formula (18) if time in (18) is set to \( t = \pi/\omega \):

\[
y \left( \frac{\pi}{\omega}, \omega \right) = 4h \sum_{k=1}^{\infty} \frac{1}{2k-1} \text{Im} W_p[(2k-1)\omega]
\]

Fig. 5. Finding exact periodic solution

The use of function \( \Phi(\omega) \) for finding the parameters of
the periodic motions is not different than the use of the frequency response $W_p(j\omega)$. It is plotted in the complex plane, the straight line that begins in the origin and makes a counterclockwise angle $\psi = \arcsin \beta$ with the negative part of the real axis is drawn, and the point of intersection of the two plots is found (Fig. 5), which provides the exact periodic solution. The frequency of the oscillations $\Omega_0$ is determined by the frequency point of $\Phi(\omega)$ in the point of intersection:

$$\Phi(\Omega_0) = -\pi + \arcsin \beta,$$

and the amplitude is computed as $a_0 = |\Phi(\Omega_0)|$.

With respect to the identification and tuning problems, the use of the exact model of the oscillations under the modified RFT can be advantageous in the parametric identification.

V. RULES FOR NON-PARAMETRIC TUNING

It seems that the most appealing application of the presented test is a non-parametric tuning. However, given a large variety of possible process dynamics, it is difficult to formulate certain universal rules for tuning. In practice of process control, tuning rules that provide a less aggressive response than the one provided by IAE, ITAE criteria or Ziegler-Nichols formulas (and others) are widely used. This approach is motivated by the consideration of safety, which chosen versus to high performance. This trend is reflected in the review of the modern PID control given in [10].

Let us consider the PI controller only, and only the rules given in the format of the proportional dependence of the controller gain on the ultimate gain and of the integral time constant on the period of the oscillations in the modified RFT – as given by formula (13). Considering the fact that the frequency-domain characteristics of all loops tuned via the modified RFT are going to be very consistent (the gain margin is the same), let us analyze the time-domain characteristics of the loops with different process dynamics and generate the tuning rules that provide the best or at least good consistency of the time-domain characteristics.

Let us use the FOPDT model as the implied process dynamics for the purpose of optimal selection of the coefficients $c_1$ and $c_2$. It serves a different purpose than an underlying process model in the parametric tuning. This model provides a combination of minimum-phase and non-minimum-phase dynamics, which is typical of real processes. Analysis of the time-domain performance of FOPDT processes with different ratios between the dead time and the time constant (subject to the same value of the gain margins) would allow us to find the optimal tuning rules. Within the time domain, it would be difficult to compare such characteristics as settling time or other measures of speed of response – due to the difference in time constants of different processes. Therefore, let us use overshoot in the step response as a relatively “universal” characteristic. Let us find the overshoot values of the step responses of a series of FOPDT dynamics with dead time to time constant ratio ranging $\tau/T_p = [0.3;1.5]$, subject to equal gain margins in those loops, by varying gain margin and parameter $c_2$ values. The noted dependence is presented in Fig. 6, where gain margin $\gamma \in [2;4]$ and parameter $c_2 \in [0.3;3.3]$.

Fig. 6. Difference between maximum and minimum overshoot [%] for modified RFT

A similar plot for the case of the conventional RFT is presented in Fig. 7 for comparison. In that case the gain
margins are not equalized by respective selection of parameter $\beta$, and the difference between the maximum and the minimum overshoots is about three times of the former (note the scale difference along the vertical axis). Therefore, the modified RFT has an equalizing effect for the time-domain characteristics too.

Analysis of the data presented in Fig. 6 shows that for satisfactory consistency of the step response (difference between maximum and minimum overshoots is lower than 10%) the gain margin and the value of $c_2$ should not be smaller than certain values. In particular, for $\gamma=2$ $c_2 \geq 1.1$; $\gamma=2.5$ $c_2 \geq 0.7$; $\gamma=3$ $c_2 \geq 0.6$; $\gamma=3.5$ $c_2 \geq 0.5$, and $\gamma=4$ $c_2 \geq 0.5$.

Therefore, the recommended settings for non-aggressive tuning with expected overshoot 0-3.3% might be $\gamma=3$ $c_2=0.7$, which results in the following tuning rules:

$$K_c = 0.33 \frac{4h}{\pi a}, \quad T_c = 0.7 \frac{2\pi}{\Omega_0}. \quad (21)$$

As follows from formula (14) the PI controller would introduce the lag at the frequency $\Omega_0$ equal to

$$\arctan \frac{1}{2\pi c_2^2} \approx 12.81^\circ.$$ Therefore, the parameter $\beta$ of the modified RFT should be $\beta = \sin\left(\arctan \frac{1}{2\pi c_2^2}\right) = 0.222$.

![Fig. 8. Step response of the processes of Example 3.](image)

**VI. EXAMPLES**

*Example 3.* Consider the process transfer function that was used in Example 1. Apply the modified RFT with parameter $\beta=0.222$ and tuning rules (21), which correspond to $\gamma=3$, and $K_c = 0.49 \frac{4h}{\pi a}$ and the same rule for $T_c$, which corresponds to $\gamma=2$, to this process. The step responses of the tuned loops to the set point change (denoted as “s.p.”) and to the disturbance application (denoted as “disturbance”) are presented in Fig. 8. The graphs presented demonstrate a satisfactory loop performance in a conservative approach ($\gamma=3$) and in a more aggressive loop tuning ($\gamma=2$). The performance of the loops is in agreement with the design criteria selected: the desired gain margin. A higher gain margin ensures a smoother but slower response. The choice of the approach is motivated by the considerations of the particular application, whether this being a servo (tracking) or a regulatory control, possibility of process dynamics change, convenience of operation, etc.

**VII. CONCLUSION**

A modified RFT and a method of non-parametric tuning of a PI controller based on this test are presented in the paper. The modified RFT can easily be implemented via including some simple logic in the conventional relay test. It is proved that the proposed method provides the desired value of the gain margin exactly (subject to the assumptions of the DF method). An exact model of the modified RFT that may be useful for the parametric tuning is developed too. It is shown that the proposed approach ensures an equalizing effect with respect to overshoot in step response of a variety of possible process dynamics – due to the same gain margin. Examples that demonstrate the proposed method and illustrate the conclusions are provided. Despite the consideration of only PI controller, the proposed method can easily be extended to the case of the PID control. The proposed method was also implemented in a Honeywell DCS TPS®, tested on a number of process control loops at an oil plant, and demonstrated performance matching the above analysis.

**REFERENCES**


