Non-parametric tuning of PID Controllers: Evolution of ideas

Igor Boiko
Khalifa University of Science and Technology, 127788, Abu Dhabi, UAE
E-mail: i.boiko@ieee.org

Abstract. Non-parametric tuning was introduced by J. Ziegler and N. Nichols with the step test and continuous cycling tests and the respective tuning rules. Whereas the step test based tuning had some feature of parametric tuning, the continuous cycling test based tuning was purely non-parametric. The latter test involves bringing the system under proportional control into a sustained oscillation and measuring the frequency of these oscillations and fixing the gain value, that are called the ultimate frequency and ultimate gain. The test was intended for manual tuning of PID controllers. It is relatively hard to automate, and it is not used in auto-tuners. Another test, called relay feedback test (RFT), was proposed by K. Astrom and T. Hagglund to excite self-oscillations through the inclusion of a nonlinear element (relay) in the control loop. This test is easier to automate, and it found applications in industrial auto-tuners. In both these tests ultimate frequency coincides with the phase cross-over frequency of the process/plant. Yet, both the original continuous cycling test and the RFT suffer from such a drawback as the change of the phase cross-over frequency after PID controller introduction. A new test, the modified relay feedback test (MRFT), which can ensure that the test oscillations are excited at the frequency that becomes the phase cross-over frequency after introduction of a controller in the loop, was recently introduced. This property is attained through coordinated test and tuning, which is considered in the book chapter. Moreover, the tuning rules are optimized for a selected class of processes, which offers the paradigm of optimal non-parametric tuning.

1. Introduction
Proportional-integral-derivative (PID) control is the main type of industrial control. The control design involves the selection of the values of the three parameters as per the following equation:

\[ W_c(s) = K_c \left(1 + \frac{1}{T_i s} + T_d s \right), \]

where \( K_c \) is proportional gain, \( T_i \) integral time constant, and \( T_d \) derivative time constant. The task of tuning is for a given process to determine the values of \( K_c \), \( T_i \), and \( T_d \), which provide optimal in a certain sense or acceptable performance of the control system (loop). The quality of tuning or loop performance can be determined based on the reactions of the closed loop to external set points or disturbances. There are a number of tuning methods. Many non-parametric methods use the so-called continuous cycling. The term continuous cycling is used to define the self-excited non-vanishing oscillations generated in the loop, aimed at measuring the parameters of these oscillations and producing the controller tuning parameters on the basis of the measurements obtained.
The first continuous cycling test was proposed by J. Ziegler and N. Nichols [1]. The loop is first brought to a steady state. At the second step, the integral and derivative components of the PID controller are disabled. After that the proportional gain is incremented by steps until a self-sustained oscillation occurs. The tuning coefficients are computed from the measured values of the so-called ultimate gain and ultimate frequency (period) that are the value of the proportional gain, at which the sustained oscillation is produced, and the frequency (period) of this oscillation, respectively. The tuning rules of [1] have the following format:

\[
K_c = c_1 K_u, \quad T_i = c_2 \frac{2\pi}{\Omega_u}, \quad T_d = c_3 \frac{2\pi}{\Omega_u},
\]

where \(c_1\), \(c_2\), and \(c_3\) are constant parameters that define the tuning rule, with \(c_1 = 0\), \(c_2 = 0.5\), and \(c_3 = 1.25\) for the full PID control and other values for PI or PD controllers, \(K_u\) and \(\Omega_u\) are ultimate gain and ultimate frequency, respectively.

In 1984 K. Astrom and T. Hagglund proposed a replacement of the variable proportional gain of the Ziegler–Nichils’s test with a relay function [2] (see figure 1). The idea of the test is based on the observation that the relay feedback system (figure 1) generates oscillations of the same frequency \(\Omega_u\). The equality of the ultimate frequencies can be easily explained through the describing function (DF) method [3]. For the ideal relay nonlinearity, the describing function \(N(a)\) is given by the expression \(N(a) = 4h/(\pi a)\), where \(a\) is the amplitude of the first harmonic of the oscillations of the error signal, \(h\) is the amplitude of the relay. Once the relay is replaced with the DF, the frequency of the self-excited oscillations can be found from the harmonic balance equation:

\[
N(a_0) W_p(j\Omega_0) = -1
\]

that can be interpreted as the marginal stability of the system having the open-loop transfer function \(N(a) W_p(j\omega)\). The relay feedback system excites non-vanishing oscillations of frequency \(\Omega_0\) and amplitude \(a_0\) which are measured in the test. Frequency \(\Omega_0\) is the ultimate frequency, and the ultimate gain is computed as the value of the DF: \(K_u = N(a_0) = 4h/(\pi a_0)\). The tuning rules can be the same as in the Ziegler–Nichols method.

![Figure 1. Astrom–Hagglund relay feedback test.](image)

Until recently a popular notion had been that the most important point on system’s frequency response is where the phase characteristic of the process equals to \(-180^\circ\) (frequency \(\omega_p\)), which will be referred to as the phase cross-over frequency. However, introduction of a PID controller results in a shift of the phase cross-over frequency, which leads to the necessity of producing test oscillations in the third quadrant of the process Nyquist plot (if the designed controller is integral-dominated), so that the frequency of the test oscillations should become the phase cross-over frequency in the open-loop system with the PID controller (see [4] and [5]). This idea is implemented in the coordinated test and tuning [4, 5], which is also referred to as the holistic test and tuning. It is considered in the following section.

2. Modified relay feedback test (MRFT) by I. Boiko

Both RFT and Ziegler–Nichols’s tests are intended for producing oscillations of the frequency corresponding to the process phase lag of \(\pi\). If the test oscillations were generated in the third

---

1 This equality is only approximate, which follows from the relay systems theory.
quadrant of the complex plane, in such a way so that the subsequent inclusion of the PI controller resulted in the placement of the point corresponding to the ultimate frequency on the real axis, one would be able to design tuning rules which would provide exact value of the desired gain margin. Therefore, having a test that would allow one to generate test oscillations in the third quadrant at an arbitrary phase lag of the plant, is a necessary component of the holistic tuning mentioned above. The second element is suitable tuning rules.

There are a few tests that can be used for producing test oscillations in the third quadrant. They are a test that uses an additional time delay [6], test with an additional relay and integrator [7], test using a phase-lock loop [8]. Another continuous cycling test, named the modified relay feedback test (MRFT) proposed in [5], is an adaptation of the second-order sliding mode sub-optimal control algorithm [11], to generating non-vanishing oscillations:

\[
\begin{align*}
    u(t) &= \begin{cases} 
    h & \text{if } e(t) \geq b_1 \text{ or } (e(t) > -b_2 \text{ and } u(t-\epsilon) = h), \\
    -h & \text{if } e(t) \leq b_2 \text{ or } (e(t) < -b_1 \text{ and } u(t-\epsilon) = -h),
    \end{cases}
\end{align*}
\tag{3}
\]

where \(b_1 = \beta e_{\min}\), \(b_2 = -\beta e_{\max}\), \(e_{\max} > 0\), \(e_{\min} > 0\) are the last “singular” points of the error signal (figure 2) corresponding to the last maximum and minimum values of \(e(t)\) after crossing zero; \(u(t-\epsilon) = \lim_{\epsilon \to 0, \epsilon > 0} u(t - \epsilon)\) the control value at the time immediately preceding current time \(t\); \(h\) the amplitude of the relay; \(-1 < \beta < 1\) is a constant, \(\beta\) is positive for generating oscillations in the 3-rd quadrant and negative to generate oscillations in the 2-nd quadrant. Initial values of \(e_{\max}\) or \(e_{\min}\) can be assigned as \(e(0)\) (the choice of either \(e_{\max}\) or \(e_{\min}\) depends on the sign of \(e(0)\)). Formula (3) provides a discontinuous control algorithm (the switching is defined by the values of \(b_1\) and \(b_2\), which in turn depend on the singular points \(e_{\min}\) and \(e_{\max}\)).

Let us show that the motions in the system, where the control is given by (3) are periodic. Apply the describing function (DF) method [3] to the analysis of motions in the system with MRFT. If the motions in the system are periodic then \(e_{\max}\) and \(e_{\min}\) represent the amplitude of the oscillations: \(a_0 = e_{\max} = -e_{\min}\) and the equivalent hysteresis value of the relay is \(b = b_1 = b_2 = \beta e_{\max} = -\beta e_{\min}\). The DF of the hysteretic relay is given as the following function of the amplitude \(a\) [3]:

\[
N(a) = \frac{4h}{\pi a} \sqrt{1 - \left(\frac{b}{a}\right)^2} - j \frac{4hb}{\pi a^2}, \quad a > b.
\tag{4}
\]
However, the system under the MRFT control (3) is not a conventional relay system. The relay has the hysteresis value that is unknown a priori and depends on the amplitude value: $b = \beta a$. Therefore, (4) can be rewritten as follows:

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

(5)

The MRFT will generate oscillations in the system under control (3). Parameters of the oscillations can be found from the harmonic balance equation:

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

(6)

where $a_0$ and $\Omega_0$ are the amplitude and the frequency of the periodic motions. The negative reciprocal of the DF is given as follows:

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

(7)

Figure 3. Finding periodic solution.

Figure 3 gives a graphic interpretation for finding a periodic solution in the system with control (3). The figure shows the Nyquist plot of the process and the negative reciprocal of the describing function. The DF is the straight line (ray) in the third quadrant, making an angle of $\psi = \arcsin \beta$ with the real axis. The periodic solution corresponds to the point of intersection of the two plots. The solution exists only if the Nyquist plot of the process lies in the third quadrant of the complex plane. In the problem of analysis, the frequency $\Omega_0$ and the amplitude $a_0$ are unknown and found from the complex equation (6). In the problems of identification and tuning, $\Omega_0$ and $a_0$ are measured from the MRFT, and the controller tuning parameters are calculated directly from $\Omega_0$ and $a_0$. It should be noted that the above-given formulas are based on the approximate DF model of the oscillations in the system under the MRFT. More accurate treatment can be obtained through exact models of oscillations in relay feedback systems, such as [5, 9, 10].

3. Coordinated (holistic) test and tuning and tuning rules

As shown above, test oscillations can be generated at a point in the third or second quadrant of the complex plane using the MRFT. However, for using this advantage to ensure the desired gain or phase margin, the tuning rules must be coordinated with the test as described below.

Let the PID controller transfer function be given by formula (1) and the tuning rules by (2), where $c_1$, $c_2$ and $c_3$ are constant parameters that define the tuning rule, to which we shall refer as to the homogeneous tuning rules, because the tuning parameters are homogeneous functions of critical gain (given by $4h/(\pi a_0)$) and critical period (inverse critical frequency). The idea behind the tuning rules (2) is the scalability the tuning parameters for processes having different
gains and time constants. If the tuning rules are given by (2) then the closed-loop system characteristics are invariant to the time constants of the process, so that if all time constants of the process were increased by the factor of $\rho$ then the critical frequency would decrease by the same factor $\rho$, and the product of every time constant by the critical frequency would remain unchanged. The use of the *homogeneous tuning rules* along with the MRFT allows one to fully utilise the features of non-parametric tuning methods. If homogeneous tuning rules (2) are used then the frequency response of the PID controller at frequency $\Omega$ becomes

$$W_c(j\Omega_0) = c_1 \frac{4h}{\pi a_0} \left(1 - j \frac{1}{2\pi c_2} + j2\pi c_3\right).$$

(8)

Therefore, if the tuning rules are established through the choice of parameters $c_1$, $c_2$ and $c_3$, and the test provides self-excited oscillations of frequency $\Omega_0$, which will be equal to the phase cross-over frequency $\omega_\pi$ of the open-loop system (including the controller), then the controller phase lag at the frequency $\omega_\pi = \Omega_0$ will depend only on the values of $c_2$ and $c_3$:

$$\varphi_c(\omega_\pi) = \arctan\left(2\pi c_3 - \frac{1}{2\pi c_2}\right),$$

(9)

which directly follows from formula (8) if $\omega_\pi = \Omega_0$.

### 3.1. Tuning for specification on gain margin

Let us derive the formulas that would allow us to tune PID controllers with specification on gain margin for the open-loop system. Let the specified gain margin be $\gamma_m > 1$. Then, considering the absolute values of both sides of (8) and accounting for (2), we obtain the following equation:

$$\gamma_m c_1 \sqrt{1 + \left(2\pi c_3 - \frac{1}{2\pi c_2}\right)^2} = 1,$$

(10)

which is a constraint that can be considered together with the tuning rules (2). To provide the specified gain margin, the MRFT must be carried out with parameter

$$\beta = -\sin \varphi_c(\Omega_0) = -\sin \arctan\left(2\pi c_3 - \frac{1}{2\pi c_2}\right) = -\frac{2\pi c_3 - \frac{1}{2\pi c_2}}{\sqrt{1 + \left(2\pi c_3 - \frac{1}{2\pi c_2}\right)^2}}.$$

(11)

If, for example, we use the parameter $c_2$ value the same as in [1] ($c_2 = 0.8$), then to obtain gain margin of $\gamma_m = 2$, the tuning parameter $c_1$ for the MRFT should be selected as $c_1 = 0.49$, and parameter $\beta$ for the test should be selected in accordance with (11) as $\beta = 0.195$. For any arbitrary process, the system will have the gain margin of $\gamma_m = 2$ (6 dB) exactly (subject to the filtering hypothesis of the DF method). Therefore, the MRFT with parameter $\beta$ calculated as (11) and tuning rules (2) satisfying constraint (10) can ensure the desired gain margin. However, (10) is an equation containing three unknown variables, which gives one the freedom to vary parameters $c_1$, $c_2$ and $c_3$ within the given constraint. Optimal selection of these parameters is considered in the next section. However, simple tuning rules satisfying the gain margin specification can be easily obtained with the use of some characteristics of the tuning rules [1]. In particular, we assume that the controller should provide the same phase response on the frequency of oscillation of the MRFT $\varphi_c(\Omega_0)$ as the phase response of the corresponding controller at the critical frequency of conventional RFT tuned as per [1]. Therefore, we can use
the values of parameters $c_2$, $c_3$ equal to the corresponding values of Ziegler–Nichols tuning rules. Coefficients of the tuning rules for gain margin $\gamma_m = 2$ in the format of values of parameters $c_1$, $c_2$, $c_3$ along with the values of $\varphi_c(\Omega_0)$ and parameter $\beta$ for the test are given in table 1.

It is worth noting the difference between the values of the ultimate (critical) frequency of the conventional RFT and the frequency of oscillations in the MRFT (except for the proportional controller): even if the coefficients $c_2$, $c_3$ of table 1 have the same values as corresponding coefficients of [1], they will actually produce different values of controller parameters $T_i$ and $T_d$. In fact, due to the negative value of $\varphi_c(\Omega_0)$ for the PI controller (and consequently, lower frequency of oscillations of the MRFT), one would get a higher value of $T_i$ computed through the MRFT and data of table 1. And vice versa, due to the positive value of $\varphi_c(\Omega_0)$ for the PID controller, one would get lower values of $T_i$ and $T_d$ computed through the MRFT and data of table 1.

3.2. Tuning for specification on phase margin

Let us derive the formulas for tuning PID controllers with specification on phase margin for the open-loop system. Using the same format of the homogeneous tuning rules (2), and considering that if the parameter $\beta$ of the MRFT is calculated from the sum of $\varphi_c(\Omega_0)$ and the phase margin $\phi_m$ as:

$$\beta = \sin (\phi_m - \varphi_c(\Omega_0)) = \sin \left( \phi_m + \arctan \left( \frac{1}{2\pi c_2} - 2\pi c_3 \right) \right),$$

we can formulate the constraint for the tuning rules ensuring $\phi_m$ as follows:

$$c_1 \sqrt{1 + \left( \frac{2\pi c_3 - \frac{1}{2\pi c_2}}{2\pi c_2} \right)^2} = 1.$$

The graphical interpretation of the MRFT and tuning with specification on phase margin are presented in figure 4. Indeed, if tuning rules (2) are subject to the constraint (13) then at the frequency $\Omega_0$ of the modified RFT: (i) the absolute value of the open-loop frequency response, in accordance with (2) and (8), is

$$|W_{ol}(j\Omega_0)| = |W_c(j\Omega_0)||W_p(j\Omega_0)| = \frac{\pi a}{4h} \frac{4h}{\pi a} \sqrt{1 + \left( \frac{2\pi c_3 - \frac{1}{2\pi c_2}}{2\pi c_2} \right)^2} = c_1 \sqrt{1 + \left( \frac{2\pi c_3 - \frac{1}{2\pi c_2}}{2\pi c_2} \right)^2} = 1,$$

![Figure 4. Modified RFT and tuning with specification on phase margin.](image)
Table 1. Coefficients defining sample (non-optimal) rules for gain margin tuning with $\gamma_m = 2$.

| Control | $c_1$ | $c_2$ | $c_3$ | $\varphi_c(\Omega_0)$ | $\beta$ |
|---------|-------|-------|-------|------------------------|--------|
| P       | 0.50  | 0     | 0     | 0                      | 0      |
| PI      | 0.490 | 0.80  | 0     | $-11.2^\circ$          | 0.195  |
| PID     | 0.453 | 0.50  | 0.125 | $-11.2^\circ$          | $-0.423$ |

Table 2. Coefficients defining generic (non-optimal) rules for phase margin tuning with $\phi_m = 45^\circ$.

| Control | $c_1$ | $c_2$ | $c_3$ | $\varphi_c(\Omega_0)$ | $\beta$ |
|---------|-------|-------|-------|------------------------|--------|
| P       | 1.0   | 0     | 0     | $0^\circ$              | 0.707  |
| PI      | 0.981 | 0.80  | 0     | $-11.2^\circ$          | 0.831  |
| PID     | 0.906 | 0.50  | 0.125 | $23.5^\circ$           | 0.341  |

which constitutes the magnitude cross-over frequency, and (ii) the phase of the open-loop frequency response is

$$\arg W_{ol}(j\Omega_0) = \arg W_c(j\Omega_0) + \arg W_p(j\Omega_0) = -180^\circ + (\phi_m - \varphi_c(\Omega_0)) + \varphi_c(\Omega_0) = -180^\circ + \phi_m,$$

which proves that the specification on the phase margin is satisfied. Given that the controller at the MRFT frequency $\Omega_0$ should provide the same phase response as that occurs at the critical frequency of RFT (which, by the way, does not give optimal tuning rules, and the tuning rules are provided only for illustration), one can obtain the values of the parameters $c_1$, $c_2$ and $c_3$ given in table 2 for the phase margin specification of $\phi_m = 45^\circ$. As in tuning with specification on gain margin, one should note the difference between ultimate frequency of the conventional RFT and the frequency of oscillations in the MRFT, which provides different values of the controller parameters. Tables 1 and 2 give examples of sample tuning rules that were generated on the basis of [1] by keeping the values of $c_2$ and $c_3$ and finding $c_1$ that would satisfy (10) or (13) and, therefore, providing the required gain or phase margins.

4. Process-specific optimal tuning rules and tuning quadrotor attitude controllers

The derived relationships (10), (13) among the coefficients $c_1$, $c_2$, $c_3$ are not the best possible tuning rules yet. The problem of development of optimal tuning rules satisfying constraints (10) or (13) may be solved for a particular class of plant/process dynamics via solving the problem of parametric optimisation for $c_1$, $c_2$, $c_3$ with a certain criterion (for example, an integral performance criterion or other) and constrains (10) and (13).

The problem of designing an optimal tuning rule for a class of processes defined by a model structure with parameter variation ranges was first formulated and solved in [5] and [12]. The problem is formulated as optimization under uncertainty for the coefficients $c_1$, $c_2$, $c_3$ with constraints given by the gain margin or phase margin specifications, using the integral performance criterion. Because the process model is not fixed but a set of models is considered, a special formulation of optimization on domain of situational parameters is used [12]. Situational parameters are considered as representing an uncertainty, which must be compensated for in the computation of the cost function, so that various tuning rules can be weighed against each other. This approach is implemented through the use of a two-stage optimization process, with different cost functions at each stage. Initially the space of situational parameters $\alpha_i$ is discretized, resulting in $N$ combinations. At the first stage, for each combination of the
situational parameters, a conventional optimization problem for \(c_1, c_2, c_3\) is solved using an integral performance criterion (IAE, ISE, ITAE or ITSE)

\[
f^\star(\alpha_i) = \min\{f(\alpha_i, x)\}, \quad i = [1, 2, \ldots, N],
\]

where \(x^T = \{c_1, c_2, c_3, \beta\}\) and \(x^\star := \{x \mid f(\alpha_i, x) = f^\star(\alpha_i)\}\), \(f\) is the integral performance.

At the second stage, a matrix \(N \times N\) of the values \(g_{ij} = f(\alpha_i, x_j)/f^\star(\alpha_i)\) is produced \((i = 1, 2, \ldots, N, j = 1, 2, \ldots, N)\). The optimal set of tuning rules is found as

\[
x_{\text{opt}} := \left\{x \mid \min_{j=1,\ldots,N} \left\{ \max_{i=1,\ldots,N} g_{ij} \right\} \right\}.
\]

This approach to finding optimal tuning rules was applied to the flow process first [12]. Recently optimal tuning rules were produced for quadrotor attitude controllers [13]. The optimal tuning rules for the PD controller are found as: \(c_1 = 0.406, c_3 = 0.359, \beta = -0.720\). Testing of these tuning rules on a number of different quadcopters showed excellent performance. A video showing an example of performance verification is provided in [14].

5. Conclusions
The use of the MRFT for generating continuous oscillations for PID controller tuning is presented. It is demonstrated that this test can produce oscillations of the frequency corresponding to the desired phase lag of a given process (plant), without any knowledge of the process dynamics. It is also shown that, when coordinated with specially designed tuning rules, the MRFT can produce the tuning rules guaranteeing specified gain margin or phase margin. All this is without any knowledge of process dynamics! Design of optimal process specific-tuning rules is considered too. An example of optimal tuning rules for the class of quadrotor attitude controllers is presented.

References
[1] Ziegler J G and Nichols N B 1942 Optimum settings for automatic controllers Trans. Amer. Soc. Mech. Eng. 64 759–768
[2] Astrom K J and Hagglund T 1984 Automatic Tuning of Simple Regulators with Specifications on Phase and Amplitude Margins Automatica 20 645–651
[3] Atherton D P 1975 Nonlinear Control Engineering—Describing Function Analysis and Design (Workingham, Berks, UK: Van Nostrand Company Limited)
[4] Boiko I 2012 Loop tuning with specification on gain and phase margins via modified second-order sliding mode control algorithm Int. J. System Science 43(1) 97–104
[5] Boiko I 2013 Non-parametric Tuning of PID Controllers: A Modified Relay Feedback Test Approach (London: Springer)
[6] Tan K K, Lee T H and Wang Q-G 1996 Enhanced automatic tuning procedure for process control of PI/PID controllers AIChE Journal 42 2555–62
[7] Friman M and Waller K V 1997 A two-channel relay for autotuning Industrial and Engineering Chemistry Research 36(7) 2662–71
[8] Crowe J and Johnson M 1998 A phase-lock loop identified module and its application IChemE Conf. “Advances in Process Control” (Wales, U.K.)
[9] Tsypkin Ya Z 1984 Relay Control Systems (Cambridge: Cambridge University Press)
[10] Falldin N V 2004 Relay automatic control systems Mathematical Models, Dynamic Characteristics and Analysis of Automatic Control Systems ed K A Pupkov and N D Eguopov (Moscow: MGTU) pp 573–636
[11] Bartolini G, Ferrara A and Usai E 1997 Output tracking control of uncertain nonlinear second-order systems Automatica 33(12) 2203–22
[12] Boiko I 2014 Design of non-parametric process-specific optimal tuning rules for PID control of flow loops J. of the Franklin Institute 351(2) 964–85
[13] Chehadeh M and Boiko I 2019 Design of rules for in-flight non-parametric tuning of PID controllers for unmanned aerial vehicles J. of the Franklin Institute 356(1) 474–91
[14] Chehadeh M and Boiko I 2017 Quadrotor disturbance rejection using optimal tuning rule URL https://www.youtube.com/watch?v=jWw9Ur9NiRo