Controller Tuning for a Slow Nonlinear Process

Louis H. Dreinhoefer

ABSTRACT: Determining the tuning constants of a classical PID (proportional-integral-derivative) controller for a slow nonlinear process, such as the temperature loops on a heat-treat or annealing furnace, is a laborious task. This paper describes the results obtained by using the Ziegler-Nichols open-loop method of determining the tuning constants for a classical PID temperature controller for a heat-treat furnace. The process being controlled is described, and open-loop test data are presented.

Introduction

This paper describes the application of the Ziegler-Nichols open-loop tuning method [1] for a PID (proportional-integral-derivative) temperature controller on a heat-treat furnace. Basically, the process is a first-order system with time delay, with the ratio of the process time constant to time delay being greater than 10. The author considered several tuning methods while looking for a method that could be applied easily in the field with limited resources. The Ziegler-Nichols method proved to be easy and effective to use, and, in a few hours, the controllers were tuned. Although this tuning method dates back to 1942 for analog controllers, it is still appropriate for modern digital control systems. The body of the paper will provide a detailed description of the tuning method, the PID controller, the process, the tuning test data, performance data, and control theory support of the results. Also, some of the nonlinearities of the process and tuning compromises will be described. The purpose of this paper is to make it easier for others to use this method by providing a simple example.

Review of Tuning Method

The Ziegler-Nichols tuning method [1] consists of stabilizing the process at a desired operating level and then plotting the open-loop process response for a small step change in the output control variable. It may be necessary to repeat this test for several operating levels and load conditions to determine the extent of the nonlinearities. The sampling frequency and resolution of the process signal measurement must be chosen to show accurate results.

The procedure consists of the following steps:

1. Stabilize the process to a normal operating level.
2. Start the data collection system.
3. Collect data until the process has reached a steady rate of change and then has a rate decrease.
4. Plot the data and graphically determine the reaction lag time (Lr) and reaction rate (K), Refer to Fig. 1 for an example.
5. Calculate the controller gains as shown in Table 1 [2]. The controller output step size equals m (percent).
6. When the error is less than the proportional band, and reaction rate is as follows, where ut(t) is the output control action or manipulated variable (scaled in percent), e(t) the error or difference between the set point and the feedback, Kp the proportional band measured in controller input units (degrees Fahrenheit), L the integral rate (repeats per minute), and Td the derivative time (minutes).

\[ u(t) = \left(\frac{100}{P}\right)e(t) + \frac{100}{K} \int e(t) dt + \frac{T_d}{P} \frac{de(t)}{dt} \] (1)

The controller is called “noninteracting” because the integral rate and the derivative time can be set independently in the time domain [3]. However, just the opposite occurs in the frequency domain, as will be shown subsequently. Other features implemented with this controller are integral windup prevention and derivative smoothing. Windup is prevented by not performing integral or derivative calculations when the error is larger than the proportional band. When the error is less than the proportional band, the proportional and derivative elements are calculated first, and then the integral portion is calculated. If the integral calculation causes an output exceeding the controller output limits, the stored integral is adjusted to make the total PID output equal to the limit. Derivative smoothing consists of using the difference between the current error and the error of a previous update rather than the most recent error.

The Laplace transformation of Eq. (1) is
shown below. It is required to analyze the controller in the frequency domain, where \( s \) is the Laplace operator \((j2\pi f)\), \( f \) the frequency (hertz).

\[
u(s) = \frac{(100/P_0)}{[1 + K/s + sT_d]} \quad (2)
\]

Equation (2) may be factored into the following form to show one pole at the origin and two zeros.

\[
u(s) = \frac{[100ahT_d(1 + s/a)(1 + s/b)]}{(P_s)s} \quad (3)
\]

where

\[
a = 1 + (1 - 4K/T_d)^{1/2}j(2T_d) \quad (4)
\]

\[
b = 1 - (1 - 4K/T_d)^{1/2}j(2T_d) \quad (5)
\]

The zeros of the controller are determined by both the values of \( K \) and \( T_d \). Therefore, the frequencies of the zeros change if either of these tuning constants are changed. Other forms of PID control allow independent adjustment of the controller zeros but are more difficult to use with the Ziegler–Nichols tuning method [3].

For the case of Ziegler–Nichols tuning with a quarter wave decay ratio, the equation becomes simpler with a pair of duplicate zeros. A quarter wave decay ratio means that the amplitude of the second overshoot is one-fourth the amplitude of the first overshoot. The following equation is obtained from Table 1 for a PID-type controller:

\[
K_T T_d = 0.25 \quad (6)
\]

By substituting Eq. (6) into Eqs. (3)–(5), the following results are obtained:

\[
a = b = 0.5/T_d \quad (7)
\]

\[
u(s) = \frac{[25(1 + 2T_d s)(1 + 2T_d s)]}{(T_d P_s s)} \quad (8)
\]

The Process

The process being controlled is a heat-treat furnace. Each zone has four gas-fired, radiant tube burners. The four burners in a zone are controlled by a common combustion air control damper. A mechanically cross-connected gas regulator keeps the burner gas on ratio with the air. A recirculation fan draws air over the radiant tubes and delivers the heated air down onto the load. A thermocouple located between the fan and the load is used as feedback for the controller. Additional thermocouples are placed in the load and underneath the load for data monitoring and practice sequencing. The control hardware consists of a small computer, a programmable logic controller (PLC), and a thermocouple input system.

The PID control resides in the computer and the burner is modulated and ignited by the PLC. The PID controller is cycled twice per minute.

The dominant characteristics of the process are in the radiant tube burners, the heat-transfer mechanics, and the furnace structure. This can be described as a first-order process with a time delay in Laplace form as shown in Eq. (9) and the block diagram of Fig. 3. An additional pole is shown at a higher frequency to approximate the response of the temperature measurement system. It is significant only to provide a more accurate frequency analysis and to serve as a reminder for other possible higher frequency poles. The transfer function of the furnace process \( c(s) \) has \( K \) steady-state gain, \( T_1 \) delay time, \( T_2 \) fundamental time constant, and \( T_3 \) time constant of the temperature measurement system.

\[
c(s) = \frac{K \exp (-sT_1)}{(1 + sT_2)(1 + sT_3)} \quad (9)
\]

The process is nonlinear as indicated by the following. The mass consists of the furnace structure plus the load. The losses, or heat transfer to these two items, change with time and temperature. The losses \( F(T) \) to the furnace structure are a function of the temperature in the furnace. There are also losses that are not a function of temperature. The heat transfer to the load is a function of the furnace temperature, load temperature, and furnace air velocity. The burner system has a time delay, primarily caused by the radiant tube burner geometry and mass.

Calculation of Tuning Constants

Tuning constants were calculated for the PID controller from the data in Table 2 using the equations of Table 1. The time base of the test data is in seconds, while the controller time base is in minutes. The conversion from seconds to minutes was part of this calculation. The results are shown in Table 3.
The step responses shown in Fig. 4 represent typical performance achieved. The extremes of an empty and a fully loaded furnace are shown. The empty furnace exhibited a quarter wave delay ratio. The gains of the controller were each varied independently ± 10 percent without noticeable change in response. The performance was also checked for steps to 950 and 1050°F, and similar responses were observed. The calculated gains from Table 3 for a loaded furnace (response not shown) did not provide an acceptable response for an empty furnace. The main performance requirements of the process are minimum overshoot and minimum steady-state control bandwidth. The response for a decrease in temperature was not checked because it is not required in this application. In actual practice, \( P_k = 8 \) seemed to be a compromise value for both the no-load and loaded furnace conditions.

**Control Results**

The Ziegler–Nichols tuning method was developed based on empirical tests. Two different controller analysis tools—Bode plot analysis and root locus—will be used to provide predictions of performance, stability, and gain sensitivity [4]. A computer program was used to make the Bode and root locus plots. To use these tools, the process and the controller equations must be completed with actual values. The process equation can be estimated only from the open-loop data of Fig. 1. Referring to Eq. (9) and Fig. 3, everything is known except for the \( F(T) \) losses. The \( F(T) \) losses are estimated to be less than 0.5 (°F/F) from Fig. 1 data. The open-loop response shows a constant rate of rise after 15°F of change and would probably not level out to a steady value before a 40°F change. Actual measurement of the \( F(T) \) losses would require collecting data for an hour or more for each process condition. Refer to the Appendix for derivations of the following results.

For an empty furnace:

\[
    c(s) = \frac{0.078 \exp(-1.9s)}{s(1 + 0.1s)} 
\]

(10)

For a loaded furnace:

\[
    c(s) = \frac{0.057 \exp(-1.3s)}{s(1 + 0.1s)} 
\]

(11)

The effects of using the preceding estimated process equations versus including the \( F(T) \) losses are minimal on the closed-loop controller analysis. Including the \( F(T) \) losses only increases the phase margin slightly for the controlled system at a lower range of process gains. Therefore, the \( F(T) \) losses are not used in the Bode and root locus analyses to reduce errors by incorrect estimates of the losses.

The controller function can be calculated accurately from Eqs. (3)–(5) using the actual tuning constants of Table 3.

\[
    u(s) = \frac{3.22(1 + 1.9s)(1 + 1.9s)}{s} 
\]

(12)

The Bode plots were obtained by plotting the product \( c(s)u(s) \), as shown in Fig. 5. Multiple curves were plotted for different values of \( P_k \) from this plot, it can be seen that the \( P_k \) can range from 42 to 12 on an empty furnace and from 28 to 8 on a loaded furnace and provide a stable response with a phase margin of at least 25 deg. Phase margin is defined as the sum of 180 plus the actual phase angle at the crossover frequency. Using common controller constants for both furnace conditions, stable operation is projected to be possible for values of \( P_k \) between 28 and 12. The minimum overshoot response should occur when the crossover frequency is at the peak on the phase-angle curve. For the empty furnace, this occurs at \( P_k = 16 \). The approximate controller closed-loop response can be estimated from the inverse of the crossover frequency \( \omega_c \). For phase margins of 25–35 deg., the overshoot is predicted to be about 30 percent of \( P_k \) or 4.8°F. Typical second-order systems with phase margins of greater than 70 deg. usually have a no-overshoot response [5]. The effect of the time delay appears in the phase curves only. It becomes a significant factor at frequencies greater than 1/T, \( \omega_c \). The loaded furnace has a greater phase margin because of a shorter time delay. This provides more damping and, thus, less overshoot. The controller integral and derivative constants provide a positive phase-margin effect at the same frequency as the inverse of the time delay.

The root locus plots were generated by plotting \( c(s)u(s) = 1 \), as shown in Fig. 6. Only the upper half of the plot is shown, since the lower half is a mirror image. Refer to Table 4 for the \( P_k \) values that correspond with the case number. The root locus plot clearly shows the effects of controller gain and time delays. For a system to be stable, the closed-loop poles must be in the left half of the \( s \) plane (\( s \) axis negative). The damping ratio can be calculated from \( \cos \theta \), where \( \theta \) is the angle between the \( s \) axis and a line.
from the origin through the dominant closed-loop pole. A dominant pole is defined as the pole farthest to the right. For the no-load furnace, the dominant pole switches from the inner circle to the outer time-delay curve at an approximate $P_d$ of 14. The dominant pole switch on the loaded furnace occurs at an approximate $P_d$ of 10. The gain at which the dominant pole switches is the gain where peak damping occurs. This is also the minimum overshoot gain. As seen in the Bode plot, the root locus plot shows greater damping for the loaded furnace. The natural frequency response, $\omega_n$, of the closed-loop system, can be graphically calculated from the length of the line from the origin to the inner-circle closed-loop pole. For a loaded furnace, the items of Table 5 were calculated from the root locus plot.

The results of both the Bode and root locus methods support the Ziegler-Nichols tuning constants. There is some difference in the projected optimum gain for minimum overshoot. Both methods indicate a wide band of gain for acceptable performance. The difference in the actual and theoretical $P_d$ indicates that a higher gain can be used in actual practice. This difference may be caused by the system mode not fully representing the actual process.

**Conclusion**

The Ziegler-Nichols open-loop tuning method is an appropriate tuning method for a furnace with time-delay and first-order characteristics. It is also simple enough, such that maintenance people can be trained to use it. Other methods suggest that better performance may be obtained; but considerable skill, knowledge, and computing tools are required to use them [6]. This method is a good starting point to get a process under control. The performance of the controller at the extreme required limits of temperature and load conditions indicates the robustness of controllers by this tuning method.

There are two primary limitations to the Ziegler-Nichols tuning method. First, the process must be controllable by manual control. Second, the process must have dominant characteristics of first order, plus time delay. The Ziegler-Nichols closed-loop tuning method could not be used because a linear mode of oscillation could not be achieved due to limit cycling. Both control theory and actual application support the validity of the open-loop tuning method. Future work could be done using the open-loop tuning method as a tool in process function identification. Then the resulting process model could be used to design other types of controllers and to select appropriate tuning constants.

**Appendix: Derivation of Furnace Model**

Refer to the furnace model portion of Fig. 3 for the following discussion. From the open-loop test data of Table 2, the reaction rates must be converted to units of minutes rather than seconds. Also, the step size of 20 percent must be used to normalize the result. The result is that the burners produce a temperature-rise rate in the furnace mass equal to 0.078 and 0.057 (°F/min %) for the no-load and loaded furnaces, respectively. Then let the $F(T)$ losses equal the feedback for the furnace model. Assuming a 40°F temperature rise for a 20 percent step, $F(T)$ equals 0.5 (%/°F).

The general closed-loop gain equation for feedback systems can be expressed as follows [5].

$$C_f = G/\left(1 + GH\right) \quad \text{(A1)}$$
The known values for the no-load furnace can be substituted into Eq. (A1).

\[ C_f = \frac{0.078/s}{1 + (0.078/s)} \]

For \( H = 0.5 \),

\[ C_f = \frac{2}{1 + s/0.039} \]

The closed-loop crossover frequency is about 10 times greater than the furnace break frequency (0.039 rad/min). Therefore, \( C_f \) can be represented by \( G \), as shown in Eq. (A2).

\[ C_f = \frac{0.078}{s} \quad \text{(A2)} \]

Then by adding the other components of the process—time delay and thermocouple conversion—the complete process model is shown in Eq. (10). Likewise, a process model can be calculated for the loaded furnace.

**Acknowledgment**

The author thanks D. W. Clark of Oxford University, who supplied a computer program for making the Bode and root locus plots, and D. J. Knapp of the Aluminum Company of America for his assistance in producing and analyzing the Bode and root locus plots.

**References**


---

**Out of Control**

"...and this is the office of our resident expert on zeros."