Brief Paper

A Method for Closed Loop Automatic Tuning of PID Controllers*†

TOR STEINAR SCHEI‡

Abstract—A simple method for the automatic tuning of PID controllers in closed loop is proposed. A limit cycle is generated through a nonlinear feedback path from the process output to the controller reference signal. The frequency of this oscillation is above the crossover frequency and below the critical frequency of the loop transfer function. The amplitude and frequency of the oscillation are estimated and the control parameters are adjusted iteratively such that the closed loop transfer function from the controller reference to the process output attains a specified amplitude at the oscillation frequency.

1. Introduction

IN RECENT YEARS several "simple" methods for the automatic tuning of PID controllers have been proposed. A specific input to the process is usually generated automatically and the PID-parameters are then determined directly from the response. Many of these methods originate from the well-known work of Ziegler and Nichols (1942, 1943). In their step response method the controller design is based on a registration of the open-loop step response of the process, which is characterized by two parameters. An extension to the Ziegler–Nichols method is presented by Nishikawa *et al.* (1984). In this method the parameters of simple process models are determined directly from the step or pulse response. The model parameters are computed from certain areas related to the response and the PID parameters are determined from these model parameters. This transient response method is less sensitive to noise than the original step response method proposed by Ziegler and Nichols.

The Ziegler–Nichols frequency response method for tuning of PID controllers is based on the determination of a single point on the Nyquist curve for the process transfer function. This point is the intersection of the Nyquist curve with the negative real axis, which is traditionally called the critical point. The control parameters are then determined from the knowledge of the critical gain and the critical period. In the original Ziegler–Nichols method the critical point is determined by increasing the gain of a proportional controller until the control loop reaches the edge of stability. The critical point is then defined by the period and the amplitude of the oscillation generated in the control loop. This method is difficult to automate. Åström and Hägglund (1984a,b) approximately determine the critical point by connecting a relay in the feedback loop from the process output to the process input. Most processes will then oscillate

in a limit cycle. The critical gain and the critical period are determined from the amplitude and the frequency of this oscillation.

These and similar methods have been used by the process control industry for the development of a wide range of "autotuning" instruments that recently have appeared in the market (Morris, 1987). References to other approaches are given by Gawthrop and Nomikos (1990).

The present method may be regarded as an extension of the Åström and Hägglund method to closed loop systems. A limit cycle is generated such that the control system oscillates at a frequency that is crucial for determination of control system stability and performance. The control parameters are adjusted iteratively such that the closed loop transfer function from the controller reference to the process output attains a specified amplitude at the oscillation frequency. It is assumed that a stable (and conservative) controller is established prior to the tuning, and the purpose of the proposed method is to improve the performance. For processes that are open-loop stable, conservative controllers might often be established by simple transient response experiments.

The paper is organized as follows. In Section 2 the excitation method is described under the assumptions that the PID parameters are constant and that a limit cycle is actually generated in the control system. Section 3 describes an iterative method for determination of the PID parameters. Simulation experiments are presented in Section 4. The advantages and disadvantages of the proposed method compared to the Åström and Hägglund open-loop autotuner are discussed in Section 5.

2. The excitation method

The system is excited by connecting a relay and a linear dynamic element in a feedback path from the process output to the reference signal for the controller as shown in Fig. 1. $M_r(s)$ is the closed loop transfer function from the controller reference, y_r, to the process output, y. For most systems a limit cycle will be generated due to the nonlinear characteristic of the relay. The linear block $d(s)$ is used to influence the frequency of this limit cycle oscillation. The reference signal will vary in steps between $y_0 - \Delta y$ and $y_0 + \Delta y$, where Δy is the amplitude of the relay function. The system will oscillate at approximately the frequency ω_{lc}, where

$$\angle d(j\omega_{lc})M_r(j\omega_{lc}) = -180°. \qquad (1)$$

This follows from a describing function approximation of the relay characteristic.

In a PID controller the derivative action is normally applied to the measurement signal only, not to the reference signal. The controller transfer function from reference to controller output is then

$$C_r(s) = K_p\left(1 + \frac{1}{T_i s}\right), \qquad (2)$$

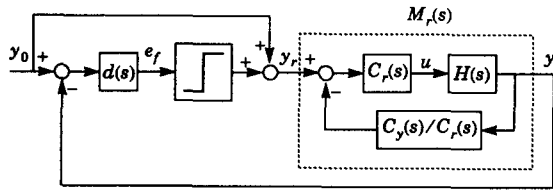and the transfer function from measurement signal to

FIG. 1. Block diagram showing the relay feedback path from the process measurement, y, to the reference signal for the controller, y_r.

controller output is

$$-C_y(s) = -K_p\left(1 + \frac{1}{T_i s} + \frac{T_d s}{1 + (T_d/N_f)s}\right),\qquad (3)$$

where K_p is the proportional gain, T_i is the integral time constant, and T_d is the derivative time constant. It is assumed that the derivative part is filtered by a first order lag with time constant T_d/N_f. The total controller output is

$$u(s) = C_r(s)y_r(s) - C_y(s)y(s).\qquad (4)$$

The _complementary sensitivity_ for the control loop is defined as

$$M(s) = \frac{G(s)}{1 + G(s)},\qquad G(s) = H(s)C_y(s),\qquad (5)$$

where $G(s)$ is the loop transfer function and $H(s)$ is the transfer function for the process. The closed loop transfer function from controller reference to process output is

$$M_r(s) = \frac{y(s)}{y_r(s)} = \frac{H(s)C_r(s)}{1 + H(s)C_y(s)} = \frac{C_r(s)}{C_y(s)} M(s).\qquad (6)$$

The filter $d(s)$ is chosen to be

$$d(s) = \frac{C_y(s)}{sC_r(s)}.\qquad (7)$$

Equations (6) and (7) are inserted in (1):

$$\angle d(j\omega_{lc})M_r(j\omega_{lc}) = \angle \frac{M(j\omega_{lc})}{j\omega_{lc}} = -180°.\qquad (8)$$

Hence, the frequency of the limit cycle, ω_{lc}, is determined by

$$\angle M(j\omega_{lc}) = -90°.\qquad (9)$$

The frequency of this oscillation is above the crossover frequency (gain equals one) and below the critical frequency (phase equals $-180°$) for the loop transfer function. This can be seen from the Nichols chart in Fig. 2. The oscillation frequency is in the frequency range that is most crucial for the determination of stability and transient properties for the control system. Some theoretical questions regarding existence and uniqueness of relay oscillations and accuracy of the describing function approximation are analyzed by Åström and Hägglund (1984a) and Tsypkin (1984).

3. Determination of PID parameters

In the previous section the excitation method is described under the assumption that the control parameters are constant and that a stable limit cycle is generated in the control system. In the proposed tuning method the amplitude and frequency of this oscillation are determined by simple peak and zero-crossing detection of the signal $e_f(t)$ in Fig. 1.

The PID control parameters are updated iteratively from the following simple scheme. The amplitude of $M(j\omega_{lc})$ is specified to be m_s. The corresponding amplitude, g_s, of the loop transfer function, $G(j\omega_{lc})$, is determined from equations (5) and (9). At time t_i the oscillation frequency is estimated to be $\hat{\omega}_{i}$ and the amplitude of $G(j\hat{\omega}_i)$ is estimated to be $\hat{g}_i = \hat{h}_i |C_{y,i}(j\hat{\omega}_i)|$, where $\hat{h}_i$ is the estimated process gain at the oscillation frequency. $C_{y,i}(s)$ is the transfer function for the controller from time t_{i-1} to t_i. Now, the integral time constant, T_i, and the derivative time constant, T_d, are set proportional to $\hat{\omega}_i^{-1}$, and the proportional gain, K_p, is determined such that $\hat{h}_i |C_{y,i+1}(j\hat{\omega}_i)| = g_s$. $C_{y,i+1}(s)$ is the transfer function for the controller with the new parameter set. The tuning is terminated when the relative change in control parameters are below certain values.

The number of oscillations used for estimation of ω_{lc} and $|G(j\omega_{lc})|$ is a compromise between stability of the autotuner and the duration of the tuning experiment. Simulations indicate that the control parameters might be updated once for each oscillation period. t_{i-1} and t_i then correspond to two zero-crossings of $e_f(t)$, with one oscillation period in
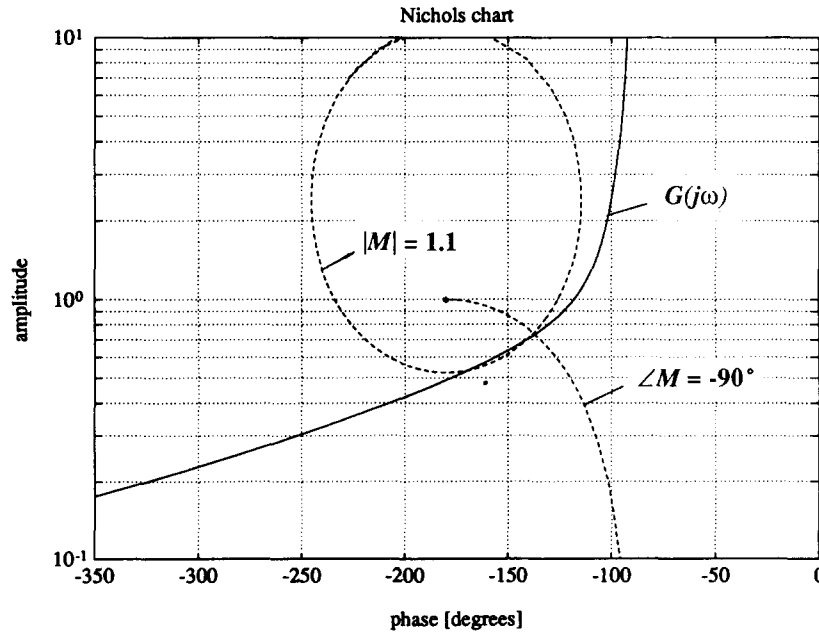

FIG. 2. Nichols chart for determination of the complementary sensitivity function. Only the curves for $|M| = 1.1$ and $\angle M = -90°$ are shown on the chart. The frequency and the amplitude of the limit cycle oscillation are determined by the intersection of the loop transfer function $G(j\omega)$ and the curve for $\angle M = -90°$.
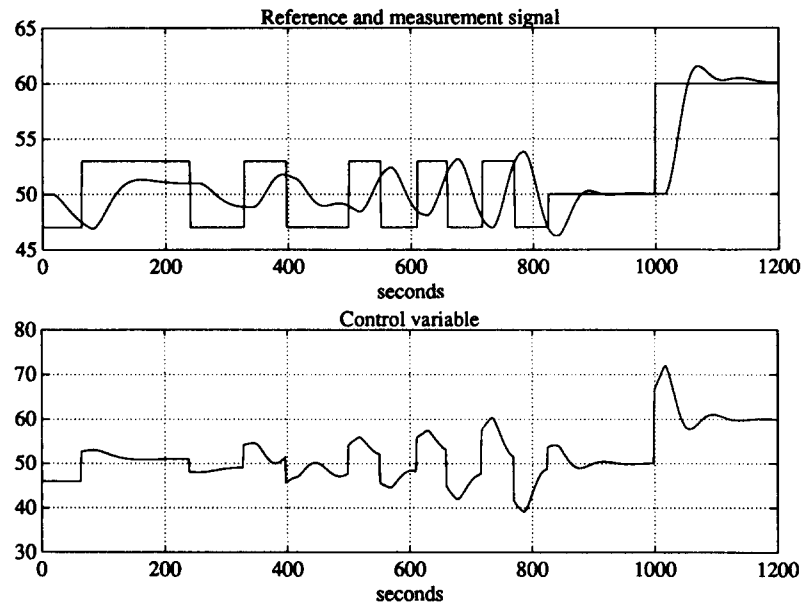
FIG. 3. Autotuning of PID controller. The tuning is terminated after about 850 sec. After 1000 sec there is a step in the reference signal for the controller.

between. The tuning algorithm might be made more robust against spurious disturbances by limiting the relative change in control parameters at each updating time.

In order to obtain the best possible tuning for a wide range of process characteristics, the autotuner has four modes. The 'tuning rules' below are determined through extensive simulation experiments with a wide range of process models.

$$
\text{Mode 1:} \quad T_i = \frac{3.0}{\omega_{lc}}, \quad T_d = \frac{0.75}{\omega_{lc}}, \quad m_s = 1.1,
$$

$$
\text{Mode 2:} \quad T_i = \frac{3.0}{\omega_{lc}}, \quad T_d = 0, \quad m_s = 1.1,
$$

$$
\text{Mode 3:} \quad T_i = \frac{1.0}{\omega_{lc}}, \quad T_d = 0, \quad m_s = 0.9, \tag{10}
$$

$$
\text{Mode 4:} \quad T_i = \frac{6.0}{\omega_{lc}}, \quad T_d = 0, \quad m_s = 1.3.
$$

Modes 1 and 2 are for PID and PI control with a medium value of the integral time constant. These modes work well for most processes. The third mode is for processes with very dominant non-minimum phase behavior. The integral time constant should be chosen relatively short for these processes. The last mode is for control of integrating processes (e.g. level control). These processes should be controlled with a long integral time constant in order to obtain sufficient phase margin for the control loop.

4. Simulation examples

A PID controller is first tuned for the process model

$$
H(s) = \frac{1}{(1 + 30s)(1 + 15s)} e^{-15s}. \tag{11}
$$

Figures 3 and 4 show the tuning experiment for a situation with no disturbances or measurement noise. The initial part
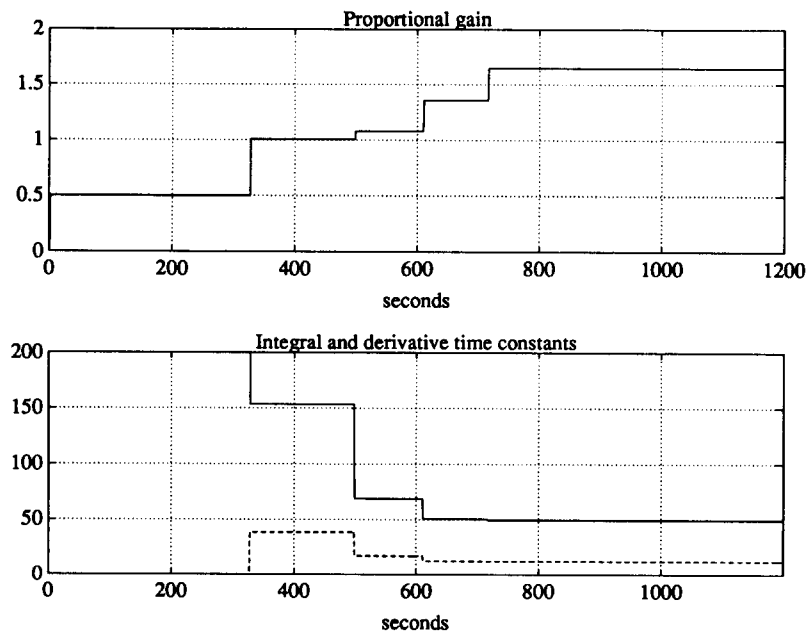

FIG. 4. Proportional gain, integral time constant and derivative time constant (dashed line) during tuning of PID controller in Fig. 3. Initially the controller has only proportional action ($T_i = \infty$ and $T_d = 0$).
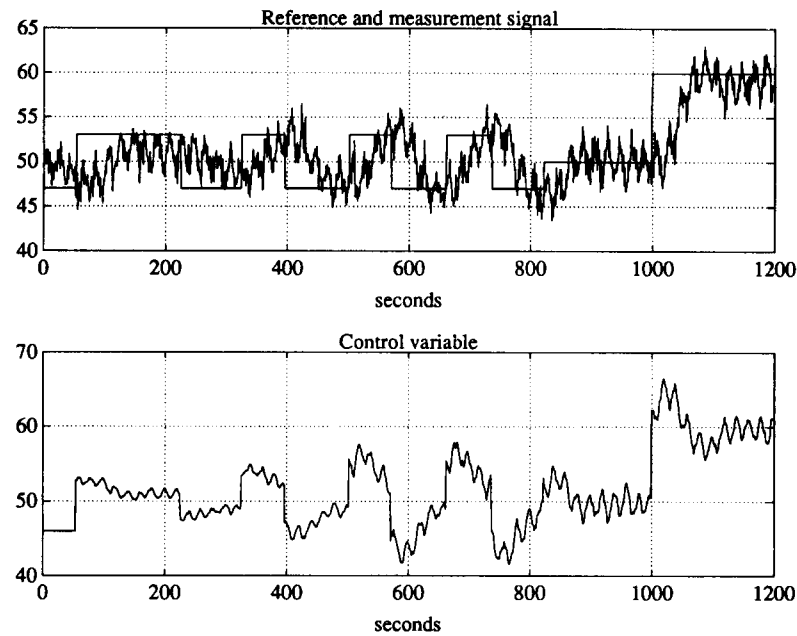
FIG. 5. Autotuning of PI controller for process with measurement noise and periodic disturbance.

of the tuning takes relatively long time due to the very conservative initial control parameters. The open-loop transfer function $G(j\omega)$ after the tuning is shown in the Nichols diagram in Fig. 2. Now, the same model is simulated with white measurement noise and a periodic disturbance added to the process output. The standard deviation for the measurement noise is 1% and the disturbance is a sine wave with a period of 20 sec and amplitude 1.5%. A PI controller is tuned as shown in Figs 5-6.

Other simulation experiments showing the performance of this method for various linear and nonlinear process models and for various disturbances are presented in Schei (1991).

5. Conclusions

A simple method for the automatic tuning of PID controllers has been proposed. The method is mainly investigated through extensive simulations with process models that are typical for the processes encountered in the process industry, and with different kinds of disturbances and measurement noise. The conclusion from these simulations is that the method performs very well if the tuning modes are reasonably chosen. Even for very difficult processes, where a PID controller is inferior to more advanced control methods, the controller tuning is often close to the "best" that can be achieved within the limitations of a PID controller.

Compared to the Åström and Hägglund open-loop autotuner the proposed method is supposed to have the following properties. The oscillation frequency is reasonably chosen with respect to obtaining sufficient gain and phase margins for the control loop. With the Åström–Hägglund autotuner the process is excited at a frequency where the phase of the process transfer function is approximately −180°. Often in process control it is not desirable to use derivative action in the controller. With a PI controller the phase lag of the loop transfer function then has to be more
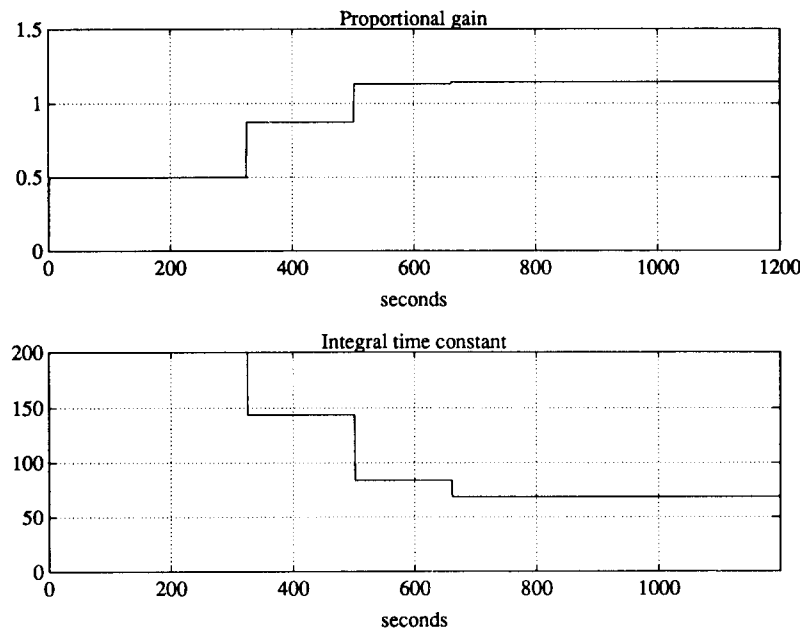

FIG. 6. Proportional gain and integral time constant during tuning of PI controller in Fig. 5.

negative than −180°. Hence, the oscillation frequency is above the critical frequency for the total control loop. In this situation it is assumed that the Åström–Hägglund autotuner has to be |more conservative than the present method in order to ensure sufficient phase margin for the control loop.

| The proposed method is also assumed to be less sensitive both to disturbances well above and well below the oscillation frequency compared to the Åström–Hägglund method. The former is due to the low pass character of the filter $d(s)$. The latter is due to the closed loop approach. In the Åström–Hägglund autotuner the process input vary in steps between two constant values. The step size then has to be chosen large enough to account for drift during the tuning.

The main disadvantage of the proposed method is assumed to be the duration of the tuning experiment. The proposed method is assumed to take longer than the Åström–Hägglund method for two reasons. First, the excitation frequency is lower, especially if the initial controller tuning is very conservative. Second, the number of oscillation periods are probably higher due to the iterative updating of the control parameters.

References

Åström, K. J. and T. Hägglund (1984a). Automatic tuning of simple regulators. *Proc. IFAC 9th World Congress*, Budapest.

Åström, K. J. and T. Hägglund (1984b). Automatic tuning of simple regulators with specifications on phase and amplitude margins. *Automatica*, **20**, 645–651.

Gawthrop, P. J. and P. E. Nomikos (1990). Automatic tuning of commercial PID controllers for single-loop and multiloop applications. *IEEE Contr. Syst. Mag.*, **10**, 1, 34–42.

Morris, H. M. (1987). How adaptive are adaptive process controllers? *Contr. Eng.*, March, 96–100.

Nishikawa, Y., N. Sannomiya, T. Ohta and H. Tanaka (1984). A method for auto-tuning of PID control parameters. *Automatica*, **20**, 321–332.

Schei, T. S. (1991). A new method for automatic tuning of PID control parameters. *Proc. Europ. Contr. Conf.*, Grenoble.

Tsypkin, Y. Z. (1984). *Relay Control Systems*. Cambridge, U.K.

Ziegler, J. G. and N. B. Nichols (1942). Optimum settings for automatic controllers. *Trans. ASME*, **64**, 759–768.

Ziegler, J. G. and N. B. Nichols (1943). Process lags in automatic control circuits. *Trans. ASME*, **65**, 433–444.