PID CONTROL

Araki M.
Kyoto University, Japan

Keywords: feedback control, proportional, integral, derivative, reaction curve, process with self-regulation, integrating process, process model, steady-state error, overshoot, decay ratio, rise time, settling time, gain margin, phase margin, action mode, ultimate sensitivity test, step response test, Ziegler–Nichols' tuning methods, Chien–Hrones–Reswick's tuning method, modulus-optimum tuning method, symmetrical-optimum tuning method, anti-windup, two-degrees-of-freedom controller

Contents

Summary

The PID controller, which consists of proportional, integral and derivative elements, is widely used in feedback control of industrial processes. In applying PID controllers, engineers must design the control system: that is, they must first decide which action mode to choose and then adjust the parameters of the controller so that their control problems are solved appropriately.

To that end, they need to know the characteristics of the process. As the basis for the design procedure, they must have certain criteria to evaluate the performance of the control system. The basic knowledge about those topics is summarized in this article.

1. Introduction

"PID" is an acronym for "proportional, integral, and derivative." A PID controller is a controller that includes elements with those three functions. In the literature on PID controllers, acronyms are also used at the element level: the proportional element is referred to as the "P element," the integral element as the "I element," and the derivative element as the "D element." The PID controller was first placed on the market in 1939 and has remained the most widely used controller in process control until today. An investigation performed in 1989 in Japan indicated that more than 90% of the controllers used in process industries are PID controllers and advanced versions of the PID controller.

"PID control" is the method of feedback control that uses the PID controller as the main tool. The basic structure of conventional feedback control systems is shown in Figure 1, using a block diagram representation. In this figure, the process is the object to be controlled. The purpose of control is to make the process variable y follow the set-point value r. To achieve this purpose, the manipulated variable u is changed at the command of the controller.

As an example of processes, consider a heating tank in which some liquid is heated to a desired temperature by burning fuel gas. The process variable y is the temperature of the liquid, and the manipulated variable u is the flow of the fuel gas. The "disturbance" is any factor, other than the manipulated variable, that influences the process variable. Figure 1 assumes that only one disturbance is added to the manipulated variable. In some applications, however, a major disturbance enters the process in a different way, or plural disturbances need to be considered.

The error e is defined by $e = r - y$. The compensator $C(s)$ is the computational rule that determines the manipulated variable u based on its input data, which is the error e in the case of Figure 1. The last thing to notice about Figure 1 is that the process variable y is assumed to be measured by the detector, which is not shown explicitly here, with sufficient accuracy instantaneously that the input to the controller can be regarded as being exactly equal to y.

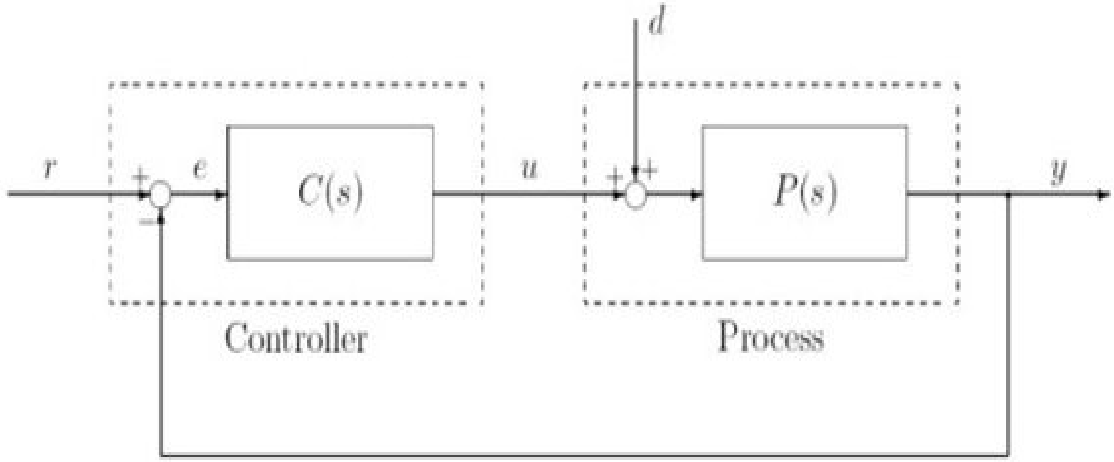

Figure 1. Conventional feedback control system

Early PID control systems had exactly the structure of Figure 1, where the PID controller is used as the compensator $C(s)$. When used in this way, the three elements of the PID controller produce outputs with the following nature:

- P element: proportional to the error at the instant *t*, which is the "present" error.
- I element: proportional to the integral of the error up to the instant *t*, which can be interpreted as the accumulation of the "past" error.
- D element: proportional to the derivative of the error at the instant *t*, which can be interpreted as the prediction of the "future" error.

Thus, the PID controller can be understood as a controller that takes the present, the past, and the future of the error into consideration. After digital implementation was introduced, a certain change of the structure of the control system was proposed and has been adopted in many applications. But that change does not influence the essential part of the analysis and design of PID controllers. So we will proceed based on the structure of Figure 1 up to Section 6, where the new structure is introduced.
The transfer function *C(s)* of the PID controller is

$$C(s) = K_P \left\{ 1 + \frac{1}{T_I s} + T_D\ D(s) \right\} \tag{1}$$

provided that all the three elements are kept in action. Here, K_P, T_I and T_D are positive parameters, which are respectively referred to as "proportional gain," "integral time," and "derivative time," and as a whole, as "PID parameters." *D(s)* is the transfer function given by

$$D(s) = \frac{s}{1 + (T_D/\gamma\ )\, s} \tag{2}$$

and is called the "approximate derivative." The approximate derivative *D(s)* is used in place of the pure derivative *s*, because the latter is impossible to realize physically. In (2), γ is a positive parameter, which is referred to as "derivative gain." The response of the approximate derivative approaches that of the pure derivative as γ increases. It must be noted, however, that the detection noise, which has strong components in the high frequency region in general, is superposed to the detected signal in most cases, and that choosing a large value of γ increases the amplification of the detection noise, and consequently causes malfunction of the controller.

This means that the pure derivative is not the ideal element to use in a practical situation. It is usual practice to use a fixed value of γ, which is typically chosen as 10 for most applications. However, it is possible to use γ as a design parameter for the purpose of, for instance, compensating for a "zero" of the transfer function of the process.

In applying PID controllers, engineers must "design" the control system. In other words, they must first decide which element(s) to keep in action and then adjust the parameters so that their control problems are solved appropriately. To that end, they need to know the characteristics of the process. As the basis for this design procedure, they must have certain criteria to evaluate the performance of the control system. Those topics will be treated in the following four sections. (See *Elements of Control Systems*.)

2. Process Models

Define the unit step function $f_{step}(t)$ by

$$f_{step}(t) = \begin{cases} 0 & t < 0 \\ 1 & t \geq 0 \end{cases} \tag{3}$$

The response $y_{u,\ step}(t)$ of the process variable to the unit-step manipulated variable $u(t) = f_{step}(t)$ directly added to the process at rest is called the "step response" or "indicial response" of the process. The term "reaction curve" is also used, essentially with the same meaning but focusing on the graphical representation. If the step response converges to a finite value K when $t \to \infty$, as exemplified in Figure 2, the process is said to be "with self-regulation" and K is called "stationary gain."

If the step response diverges when $t \to \infty$, the process is said to be "without self-regulation." If a process is without self-regulation and its step response approaches a straight line with the slope R, as exemplified in Figure 3, it is said to be "with a single integrator" or simply "integrating."

It has been observed that step responses of many processes to which PID controllers are applied have monotonically increasing characteristics as shown in Figures 2 and 3, so most traditional design methods for PID controllers have been developed implicitly assuming this property. However, there exist some processes that exhibit oscillatory responses to step inputs. This topic will be treated later (see Section 6.3).

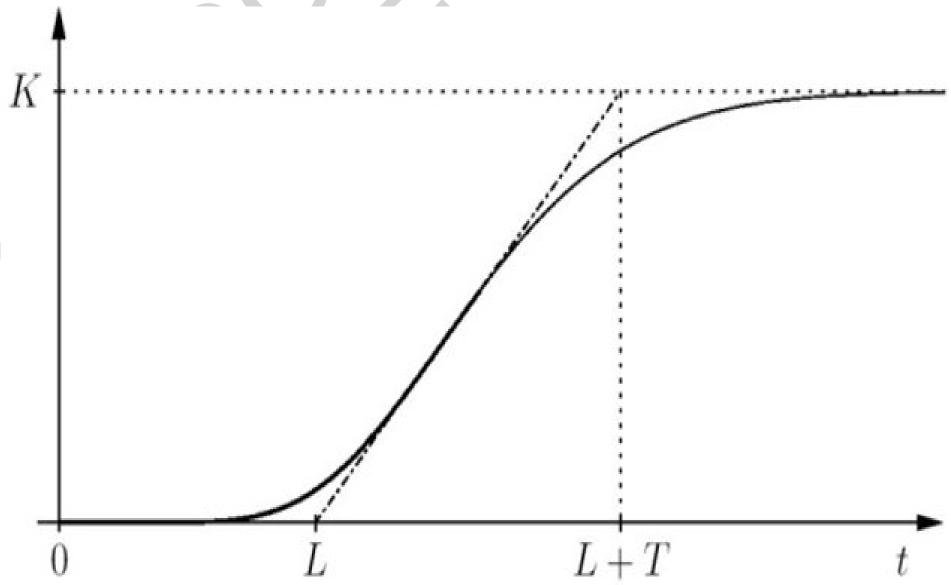

Figure 2. Step response of a process with self-regulation

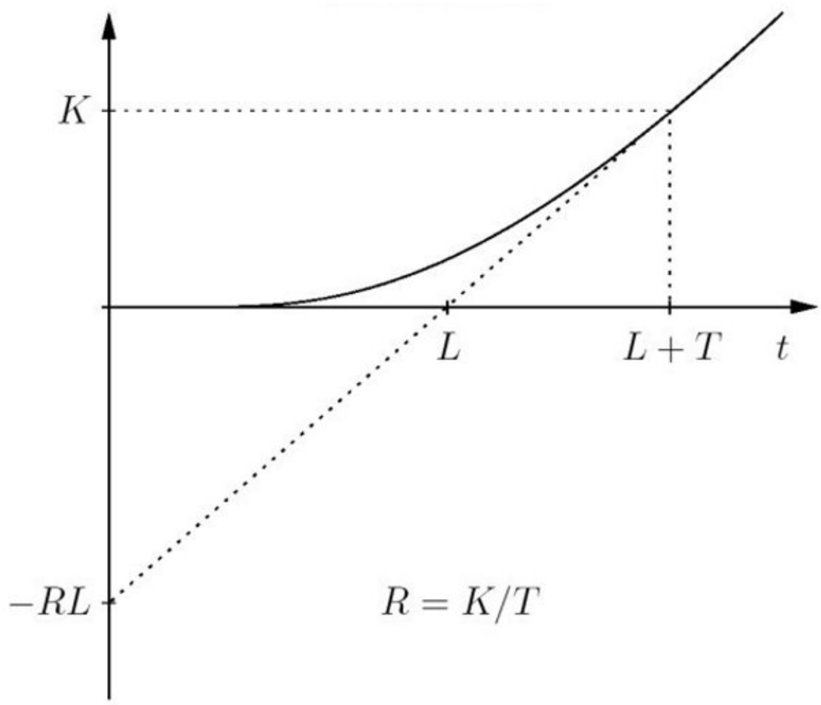

Figure 3. Step response of an integrating process

A more basic assumption employed in the design methods explained in the following is "linearity." "Linearity" means that, if the responses of the process variable to inputs $u_1(t)$ and $u_2(t)$ are, respectively, $y_1(t)$ and $y_2(t)$, then its response to the summed-up input $u_1(t)+u_2(t)$ becomes $y_1(t)+y_2(t)$, all under the condition that the process is at rest at the initial instant.

In systems theory, it is generally expected that linearity approximately holds true in a small range of variables, while the approximation error increases as the range increases. This expectation is met in some processes but upset in others. There are processes, for instance, such that the response to the negative step is largely different from the inverse of the response to the positive step.

In spite of such reality, the linearity assumption has been employed widely, first because it is difficult to establish a practically tractable general method without this assumption, second because experience shows that the designed results work approximately well for many processes, and third because the results obtained from desk work are in any case insufficient, so that trial-and-error adjustment at actual processes is always needed, and the nonlinear property can be considered in that procedure.

Under the above assumptions, the following transfer functions can be used to model the process. For a process with self-regulation,

$$P(s)=\frac{K}{1+Ts}e^{-Ls} \tag{4}$$

is the simplest model. This model is referred to as the “first-order-lag + pure-delay” model, because $K/(1+Ts)$ is the transfer function of the first-order-lag element whose stationary gain is K and time constant is T, and e^{-Ls} is that of the pure delay whose delay time is L. The simplest model for an integrating process is

$$P(s)=\frac{R}{s}e^{-Ls}=\frac{K}{Ts}e^{-Ls} \qquad R=\frac{K}{T} \tag{5}$$

This model is referred to as the “integrator + pure-delay” model. The parameters K and T of the second expression are redundant by one and so there is no way, mathematically speaking, to determine them uniquely.

However, this expression is sometimes used, with understanding that the parameter T is the time constant of the process, first in order to make the denominator Ts dimensionless so that the time scale of the reaction curve is standardized, and second in order to make the equation giving the steepest slope of the reaction curve the same as that for the “first-order-lag + pure-delay” model. The latter makes the turning formulae of PID parameters applicable without confusion (see Section 5.3).

The above two models have long been used as the basis of design methods for PID control systems, because their parameters can easily be determined from simple tests (see Section 5.2), and the designed results are very often sufficient as the initial values to start the trial-and-error adjustment procedure.

But recently there has been a move to make full use of the capability of modern computers and sensing systems for adjusting the controller as exactly as possible, based on the initial test or on-line data. For that purpose, the above models are too simple, so more sophisticated models are considered (see Section 6.3).

-
-
-

Bibliography

Åström K.J. and Hägglund T. (1995). *PID Controllers: Theory, Design and Tuning*, 343pp. Instrument Society of America [This book provides detailed explanation on the most topics treated in the text except the classical tuning method of the two-degree-of-freedom PID controller.]

Bennett S. (1993), *A History of Control Engineering 1930–1955*, pp.28–69. Peter Peregrinus. [This book contains a short history of early PID controllers.]

Taguchi H. and Araki M (2000). Two-degree-of-freedom PID controllers. *Proceedings of the IFAC Workshop on Digital Control: Past, Present and Future of PID Control*, pp. 91–96. Elsevier. [This paper

provides a brief survey of the properties of the two-degree-of-freedom PID controller as well as its tuning method.]

Biographical Sketch

Mituhiko Araki received his B.E., M.E., and Ph.D. degrees, all in electronic engineering, from Kyoto University, Japan, in 1966, 1968, and 1971. Since 1971 he has been with the Department of Electrical Engineering, Kyoto University, where he is currently a Professor. He visited Imperial College, London from 1973 to 1975, Santa Clara University, California in 1979, and Waterloo University, Ontario in 1982. His research interests are in systems and control theory and its application to industrial and medical problems. As theoretical topics, he has been engaged in research on stability of composite systems, M-matrices, the Nyquist array method for design of multivariable controllers, state predictive controllers for plants with pure delays, multirate digital control systems, two-degree-of-freedom optimal controllers, and frequency responses of sampled-date systems. As industrial applications, he proposed two-degree-of-freedom PID controllers, studied their optimal tuning, and applied them to temperature controllers and servo-pulser controllers. In addition to the above, he has been engaged in researches on control of synchronous generators, control and scheduling of steel producing systems, and scheduling of elevators.

Since 1991 he has been engaged in co-operative topics with medical doctors such as the control of blood pressure of patients under surgery, the control of blood sugar after surgery, the control of intraocular pressure during eye surgery, and clinical stage classification. Currently he is placing more time and efforts on these medical applications, while he is still much interested in industrial applications of control theory and scheduling techniques.

He is a member of the IEEE and of many Japanese academic societies. He was a council member of IFAC, an associate editor of the *IEEE Transactions on Automatic Control*, the editor-in-chief of *Systems, Control and Information* (a Japanese journal), and the editor-in-chief of the *Transactions of the Society of Instrument and Control Engineers in Japan*. He is currently the editor for control system applications of *Automatica*.