Fundamentals of PID Control

An Online Continuing Education Course for Engineers

Course Number: IC-2003

Credit: 2 Hours / 2 PDH / 2 CPD

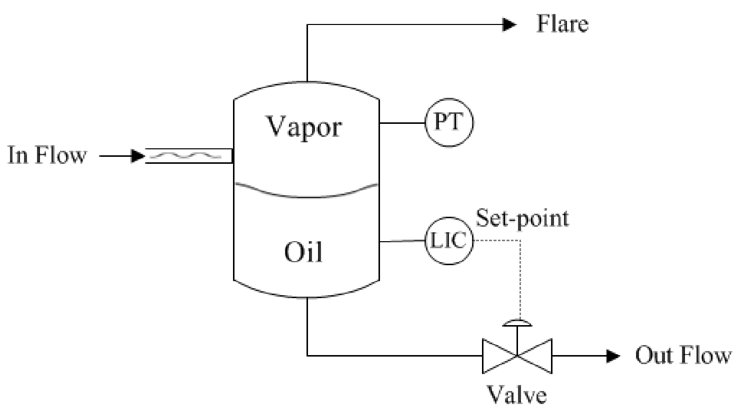
Fundamentals of PID Control

Anthony K. Ho, P.E.

1. History and Background of PID Control

The original technology for industrial PID (proportional, integral, and derivative) control was pneumatic, hydraulic, or mechanical and the controller usually had a simple interface for manual tuning. The first theoretical analysis of PID controller can be dated back to 1922 when Russian American engineer Nicolas Minorsky developed an automatic ship steering system for the US Navy, which was based on observation of the steersman steering the ship using current error, past error, and rate of change. Later, controllers with electrical systems were developed after World War II.

PID control is used to control and maintain processes. It can be used to control physical variables such as temperature, pressure, flow rate, and tank level. The technique is widely used in today's process industry to achieve accurate control under different process conditions. PID is simply an equation that the controller uses to evaluate the controlled variables. A Process Variable (PV) temperature, for example, is measured and feedback to the controller. The controller then compares the feedback to the set-point (SP) and generates an error value. The value is examined with one or more of the three proportional, integral, and derivative methodology. As a result, the controller issues the necessary commands or alters Control Variable (CV) to correct the error (E). These procedures form an iterative process. Below is a common control loop application.



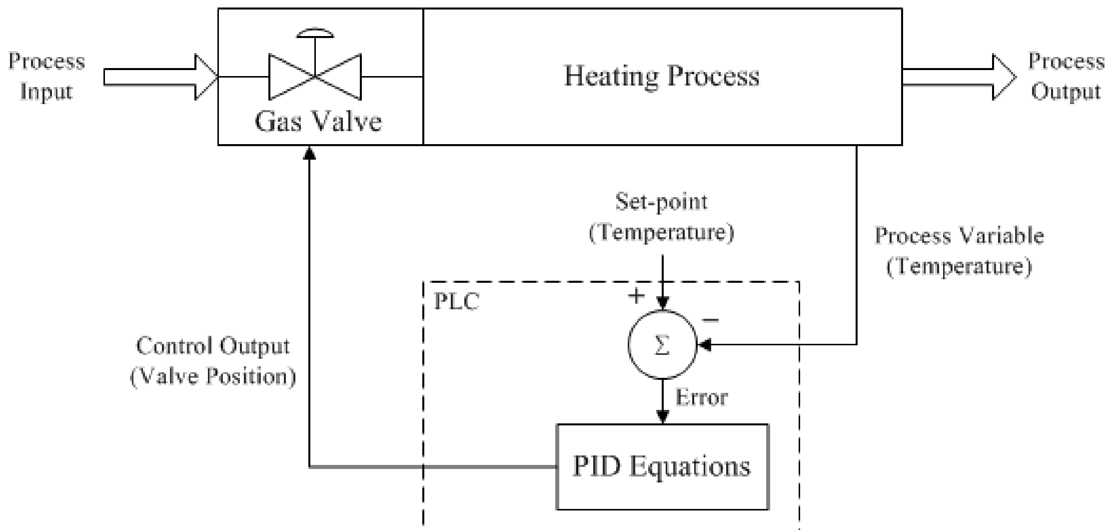
PT	Pressure Transmitter
LIC	Level Indicator Controller

In this example, oil is flowing into the tank in a non-constant rate. The oil level in the tank is the process variable which is measured and feedback to the controller. An operator entered a set-point for a desired level. The controller compares the current level with the set-point and generates a value which is examined with a PID method. The controller then adjusts the valve position to correct the error.

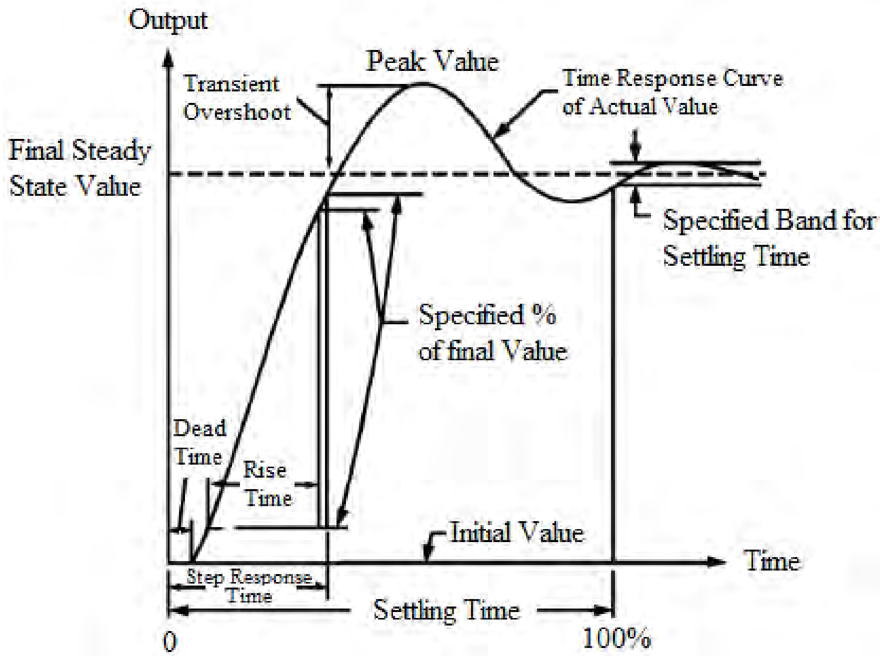
2. Theory of PID Control

PID controllers typically use control loop feedback in industrial and control systems applications. The controller first computes a value of error as the difference between a measured process variable and preferred set-point. It then tries to minimize the error by increasing or decreasing the control inputs or outputs in the process, so that process variable moves closer to the set-point. This method is most useful when a mathematical model of the process or control is too complicated or unknown. To increase performance, for example to increase the responsiveness of the system, PID parameters must be adjusted according to the specific application.

The block diagram below shows an example of a heating process that is controlled by a PID controller (PLC in this case). The temperature of the furnace is controlled by adjusting the gas valve. The operator set the desired temperature as the set-point. Temperature of the furnace is measured and feedback to the controller. The feedback is compared to the set-point and an error value is calculated. The PID equations then determine the suitable valve position to correct the error. In fact, this is an example of a PID feedback control loop.



The following figure is a typical step response curve after a controller responded to a set-point change. The curve rises from 10% to 90% of final steady state value within a period known as the rise time. The curve rises from 0% to 63.2% of peak value within a period known as the step response time. The rise time is equal to step response time minus the dead time.



One of the advantages of PID is that for many processes there are straightforward correlations between the process responses and the use and tuning of the three terms (*P*, *I*, and *D*) in the controller. There are two steps in designing a PID system. First, engineer must choose the structure of the PID controller, for example P, PI, or PID. Second, numerical values for the PID parameters must be chosen in order to tune the controller.

These three parameters for the PID algorithm are the proportional, integral, and derivative constants. The proportional constant decides the reaction based on the current error, the integral constant determines the reaction according to the total of recent errors, and the derivative constant determines the reaction using the rate at which the errors have been changing. These three actions are then used to adjust the process through control element such as the position of a valve. In simple terms, *P* depends on the current error, *I* depends on the sum of past errors, and *D* predicts future errors based on current rate of change of errors.

2.1. Proportional Control

The proportional part of PID examines the magnitude of the error and it reacts proportionally. A large error receives a large response. For example, if there is a large temperature error, the fuel valve would be opened a lot. On the other hand, a small error receives a small response. In mathematical term, the proportional term (P_{out}) is expressed as:

$P_{out} = K_p\, e$

Where:

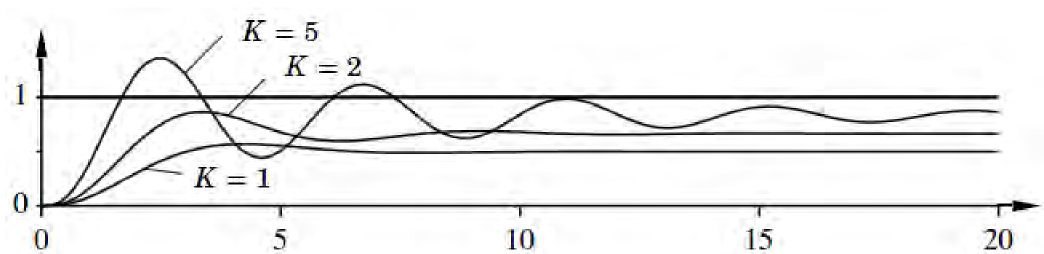
P_{out}: Proportional portion of controller output

K_p: Proportional gain

e: Error signal, e = Set-point – Process Variable

e = SP – PV here represents a reverse acting loop. When e = PV – SP, it refers to a direct acting loop. In a direct acting loop, process variable is greater than the set-point; therefore, the appropriate controller action is to increase the output. A typical example of a direct-acting system is control of temperature using cooling water. On the other hand, in a reverse acting loop, process variable is less than the set-point; therefore, the controller action is to decrease the output. An example would be control of temperature using steam.

The following figure illustrates a proportional control and shows that there is always a steady state error in proportional control. The error will decrease with increasing gain, but the tendency towards oscillation will also increase.



Let's look at a furnace as an example. Assume that between 1500 °F and 2000 °F, the system response will be proportional such that it opens the fuel valve in proportion to the error. When temperature is below 1500 °F, the valve is opened 100 %. When it is above 2000 °F, the valve is closed or 0 % opened. The proportional band (also known as the reciprocal of gain) in this case is assumed to be 500 °F. However, proportional band is usually expressed in percentage. The percentage is calculated by dividing the proportional band in engineering unit (°F) by the full controller range and multiplying by 100. The full controller range in this case is 2000 – 70 °F = 1930 °F (assume 70 °F is the room temperature.) Therefore, the proportional band in percentage is equal to (500 °F / 1930 °F) × 100 or 26 %. The engineer can adjust the proportional band and make the system more or less responsive to an error. The rule of thumb is the smaller the proportional band (or large gain) results in a large output or faster response to a given input error, and larger band (or small gain) yields to a less responsive controller. However, be cautious that an excessively large proportional gain will lead to process instability and oscillation.

You may see that there are issues with proportional control only. One of them is that proportional control cannot compensate very small errors (these errors are also known as offset.) Another issue is that it cannot adjust its output based on the rate of change in the measured variable. For example, if you are driving on the highway at 55 miles per hour and the car in front of you slows down, you may respond by applying 50 % brake pressure. This is a proportional response because the car in front is slowing down so you respond with 50 % brake. However, if the car slows down even more, you would apply much more brake pressure since the rate of stopping of the car in front is greater. In other words, we apply more brake pressure as we see the distance between our cars keeps getting smaller since we naturally respond to rate of change of error. However, proportional control systems do not do this, they only respond to the magnitude of the error, not to its rate of change.

2.2. Integral Control

To address the first issue with the proportional control, integral control attempts to correct small error (offset). Integral examines the error over time and increases the importance of even a small error over time. Integral is equal to error multiplied by the time the error has persisted. A small error at time zero has zero importance. A small error at time 10 has an importance of 10 times error. In this manner, integral increases the response of the system to a given error over time until it is corrected. Integral can also be adjusted and the adjustment is called the reset rate. Reset rate is a time factor. The shorter the reset rate the quicker the correction of an error. However, too short a reset rate can cause erratic performance. In hardware-based systems, the adjustment can be done by a potentiometer changing the time constant of a RC circuit. Most of today's applications use software based control such as PLC module where the engineer changes the parameter of reset rate. The mathematical expression of an integral-only controller (I_{out}) is:

$$I_{out} = \frac{1}{T_i} \int e\,dt = K_i \int e\,dt$$

Where:

I_{out}: Integral portion of controller output

T_i: Integral time, or reset time

K_i: Integral gain

e: Error signal, e = Set-point – Process Variable

2.3. Derivative Control

The derivative part of the control output attempts to look at the rate of change in the error signal. Derivative will cause a greater system response to a rapid rate of change than to a small rate of change. In other words, if a system's error continues to rise, the controller must not be responding with sufficient correction. Derivative senses this rate of change in the error and provides a greater response. Derivative is adjusted as a time factor and therefore is also called rate time. It is essential that too much derivative should not be applied or it can cause overshoot or erratic control. In mathematical term, the derivative term (D_{out}) is expressed as:

$$D_{out} = T_d \frac{d}{dt} e = K_d \frac{d}{dt} e$$

Where:

D_{out}: Derivative portion of controller output

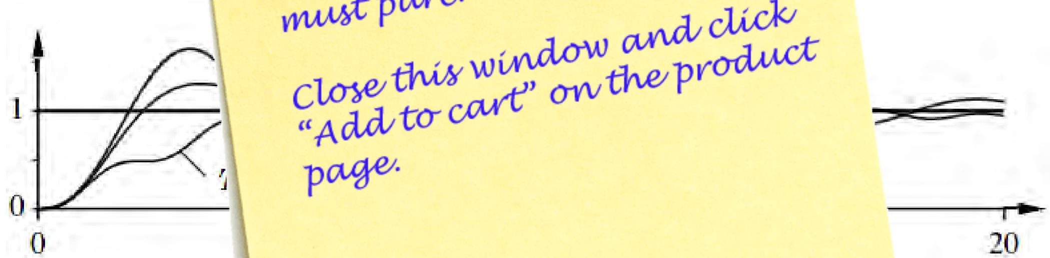
T_d: Derivative time

K_d: Derivative gain

e: Error signal, e = Set-point – Process Variable

To summarize all three controls, proportional control causes an input signal to change as a direct ratio of the error signal variation. It responds immediately to the current tracking error but it cannot achieve the desired set-point accuracy without an unacceptably large gain. Thus, proportional term usually needs the other terms. Integral control causes an output signal to change as a function of the integral of the error signal over time duration. Integral term yields zero steady-state error in tracking a constant set-point. It al ... onstant disturbances. Derivative action reduces transient err ... change as a function of the rate of change of ... ms will yield the control output, or the control

The following f ... ional, integral, and derivative control. (Cl



In practice, most P ... utomatic. In manual mode, the controll ... operator, typically by pushing buttons that increase or de ... output. A controller may also operate in combination with other controllers, such as in a cascade or ratio connection, or with nonlinear