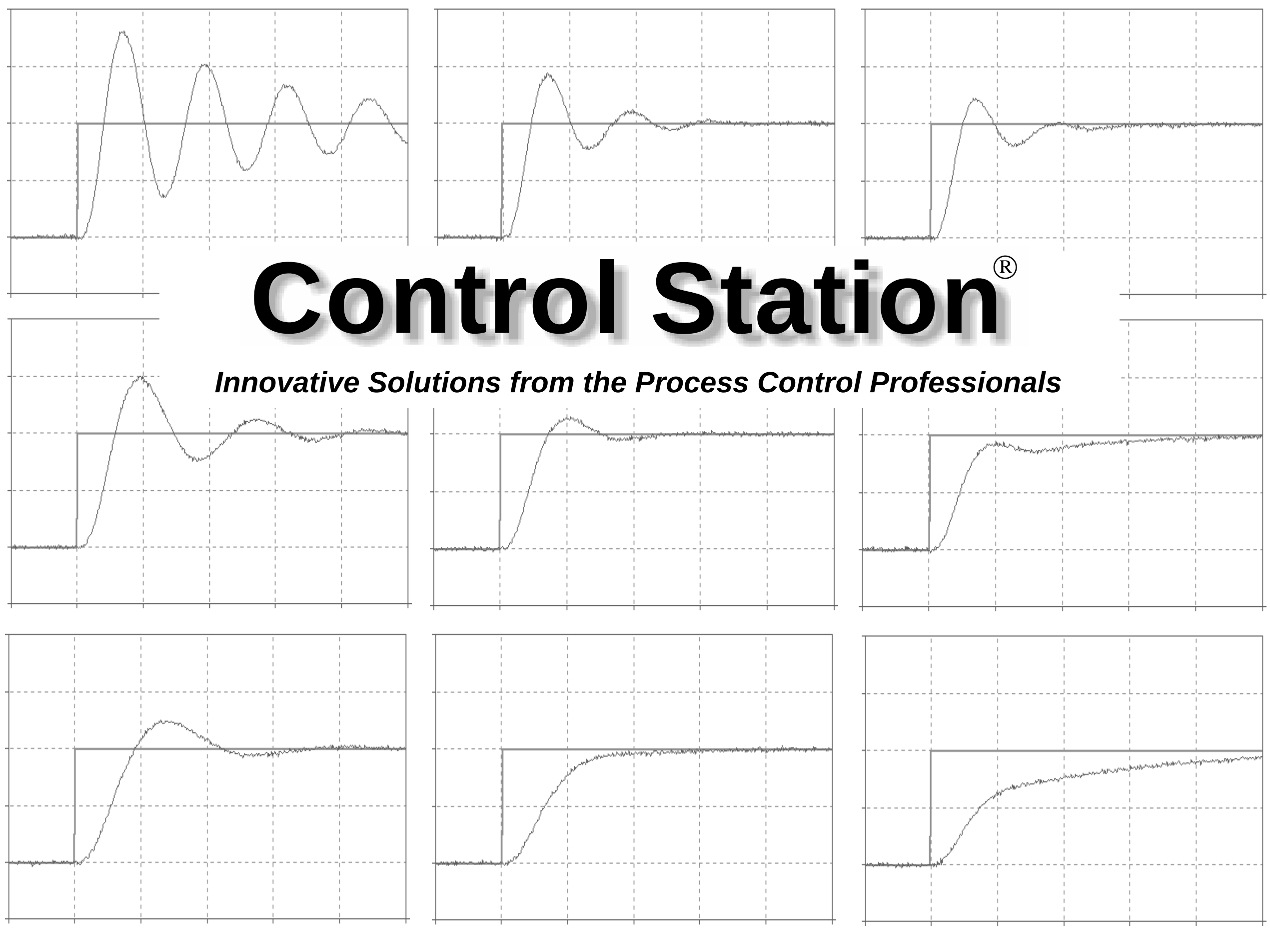


Control Station®

Innovative Solutions from the Process Control Professionals

Practical Process Control©

"Fundamentals of Instrumentation and Process Control"

Practical Process Control

"Fundamentals of Instrumentation and Process Control"

Table of Contents

1. Introduction to Process Control

Objectives:

In this chapter you will learn:

- Why Do We Need Process Control?
- What is a Process?
- What is Process Control?
- What is Open Loop Control?
- What is Closed Loop Control?
- What are the Modes of Control?
- What are the Basic Elements of Process Control?

Introduction

Why do we need Process Control?

Effective process control is required to maintain safe operations, quality products, and business viability.

Safety

The primary purpose of a Process Control system is safety: personnel safety, environmental safety and equipment safety. The safety of plant personnel and the community is the highest priority in any operation. An example of safety in a common heat exchanger process is the installation of a pressure relief valve in the steam supply. Other examples of safety incorporated into process control systems are rupture disks and blow out panels, a pressure switch that does not allow a pump to over pressurize a pipe or a temperature switch that does not allow the fluid flowing through a heat exchanger to overheat.

Quality

In addition to safety, process control systems are central to maintaining product quality. In blending and batching operations, control systems maintain the proper ratio of ingredients to deliver a consistent product. They tightly regulate temperatures to deliver consistent solids in cooking systems. Without this type of control, products would vary and undermine quality.

Profit

When safety and quality concerns are met, process control objectives can be focused on profit. All processes experience variations and product quality demands that we operate within constraints. A batch system may require +- 0.5% tolerance on each ingredient addition to maintain quality. A cook system may require +- 0.5 degrees on the exit temperature to maintain quality. Profits will be maximized the closer the process is operated to these constraints. The real challenge in process control is to do so safely without compromising product quality.

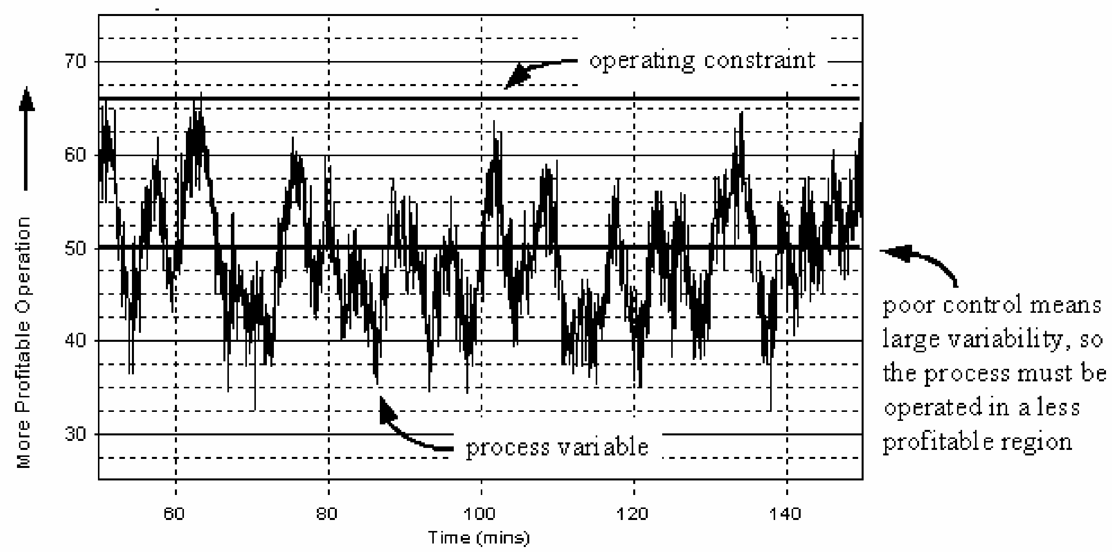


Figure 1-1 Copyright Control Station

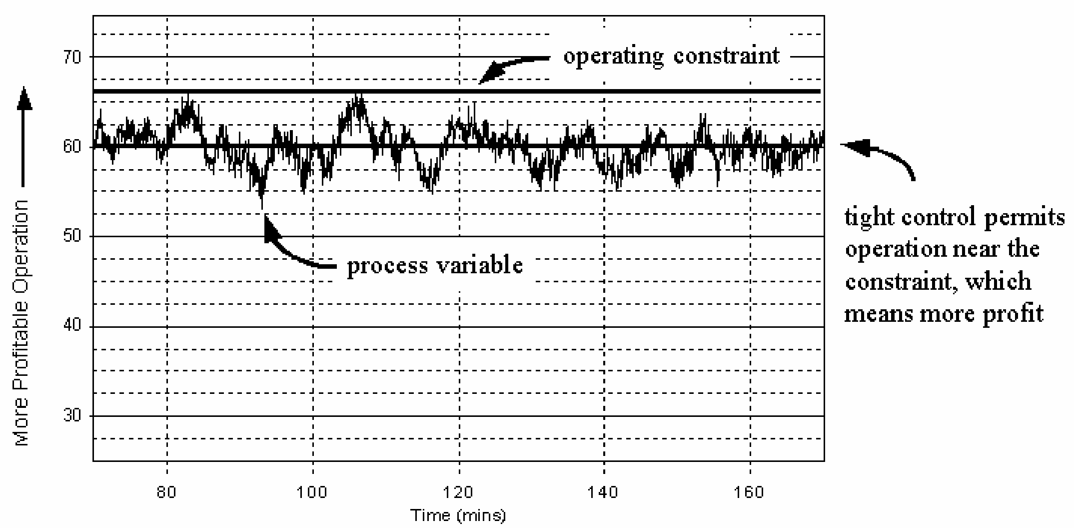


Figure 1-2 Copyright Control Station

What is a Process?

☞ A ***process*** is broadly defined as an operation that uses resources to transform inputs into outputs. It is the resource that provides the energy into the process for the transformation to occur.

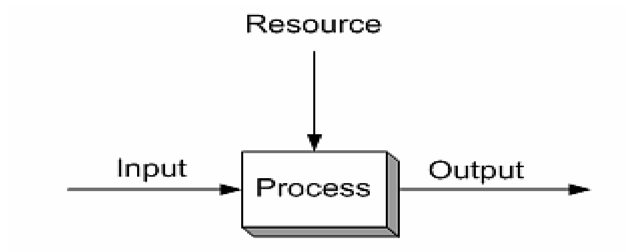


Figure 1-3

Figure 1-4 shows a hot water generation process commonly found in plants. The input to this process is cold water and the output of the process is hot water. Steam is the resource that provides energy for the transformation to occur within the heat exchanger plates.

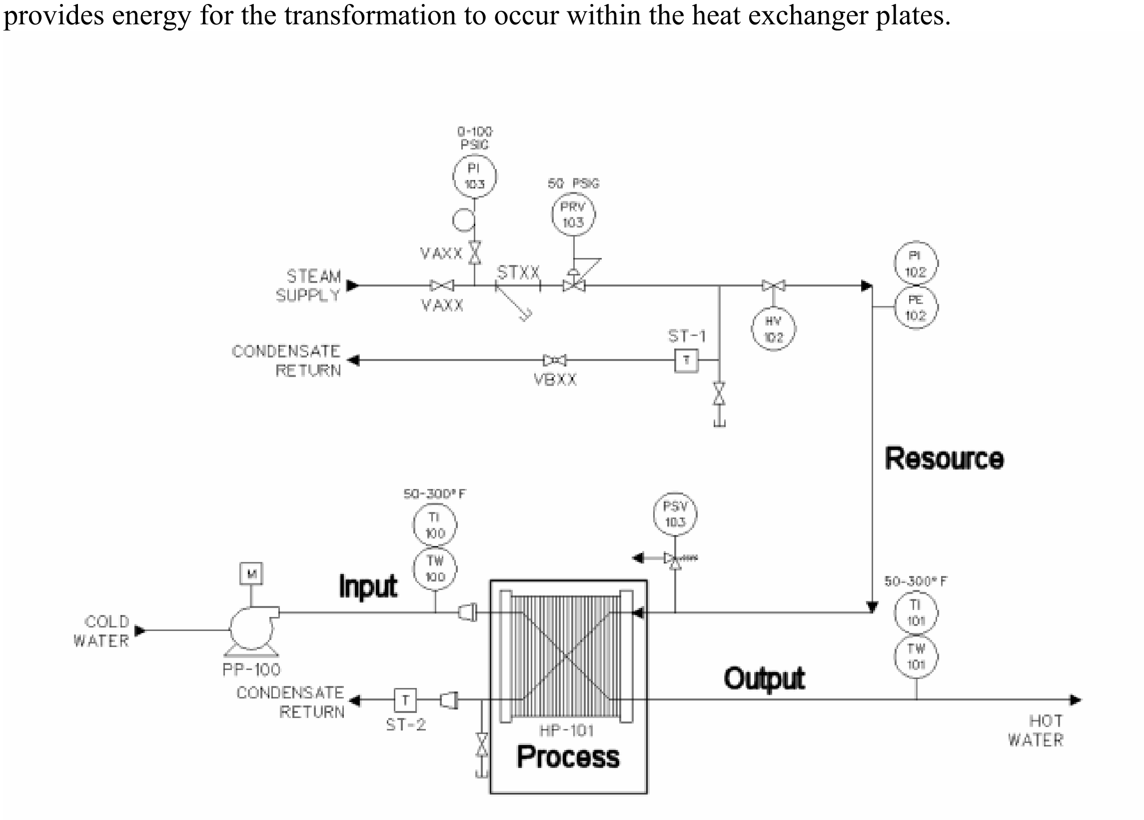


Figure 1-4

Most plants operate multiple types of processes, including separation, blending, heating, and cooling to name a few. Each process exhibits a particular dynamic (time varying) behavior that governs the transformation, that is, how do changes in the resource or inputs over time affect the transformation. This dynamic behavior is determined by the physical properties of the inputs, the resource and the process itself. A typical heat exchanger process contains a plate and frame heat exchanger to transfer the heat from the steam to the incoming water. The properties of the incoming water (temperature), the steam (pressure) and properties of the specific heat exchanger used (surface area, efficiency of heat transfer) will determine the dynamic behavior, that is; how will the output be affected by changes in water temperature or steam pressure (flow)?

What is Process Control?

☞ ***Process control*** is the act of controlling a final control element to change the manipulated variable to maintain the process variable at a desired Set Point.

A corollary to the definition of process control is a controllable process must behave in a predictable manner. For a given change in the manipulated variable the process variable must respond in a predictable and consistent manner. Following are definitions of some terms we will be using in out discussion of process control:

☞ The ***manipulated variable (MV)*** is a measure of resource being fed into the process, for instance how much thermal energy.

☞ A ***final control element (FCE)*** is the device that changes the value of the manipulated variable.

☞ The ***controller output (CO)*** is the signal from the controller to the final control element.

☞ The ***process variable (PV)*** is a measure of the process output that changes in response to changes in the manipulated variable.

☞ The ***Set Point*** (***SP)*** is the value at which we whish to maintain the process variable at.

Figure 1-5 shows a block diagram of a process with a final control element and sensors to measure the manipulated variable and process variable. In single loop control systems the actual value of the manipulated variable is often not measured, the value of the process variable is the only concern.

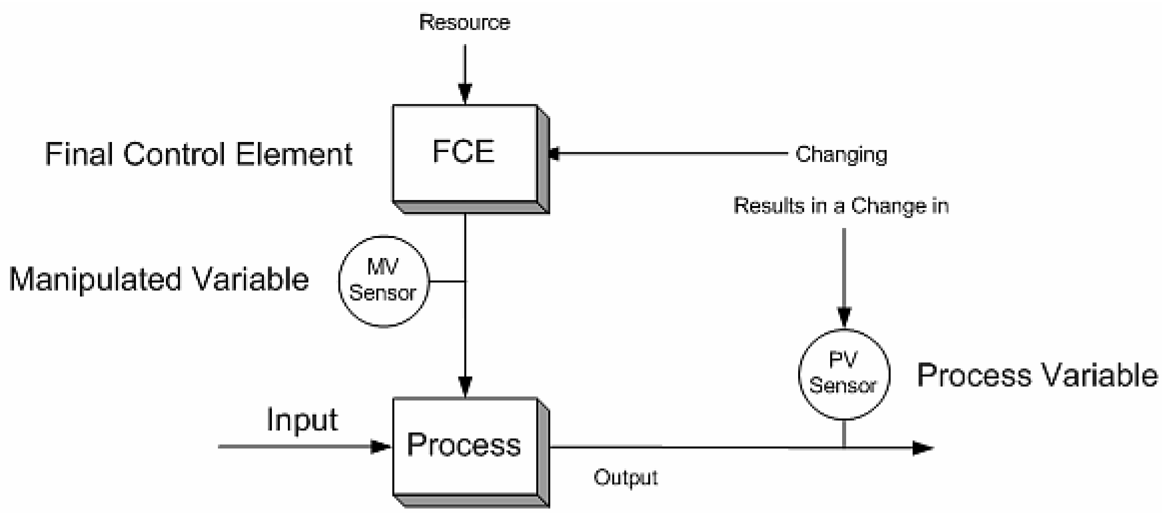


Figure 1-5

Figure 1-6 shows a heat exchanger. We see that the manipulated variable (MV) is steam pressure. The final control element is the valve, by changing the valve opening we are changing the flow of steam which we can measure by its pressure. The process variable (PV) is the temperature of the water exiting the heat exchanger; this is the measure of the process output that responds to changes in the flow of steam.

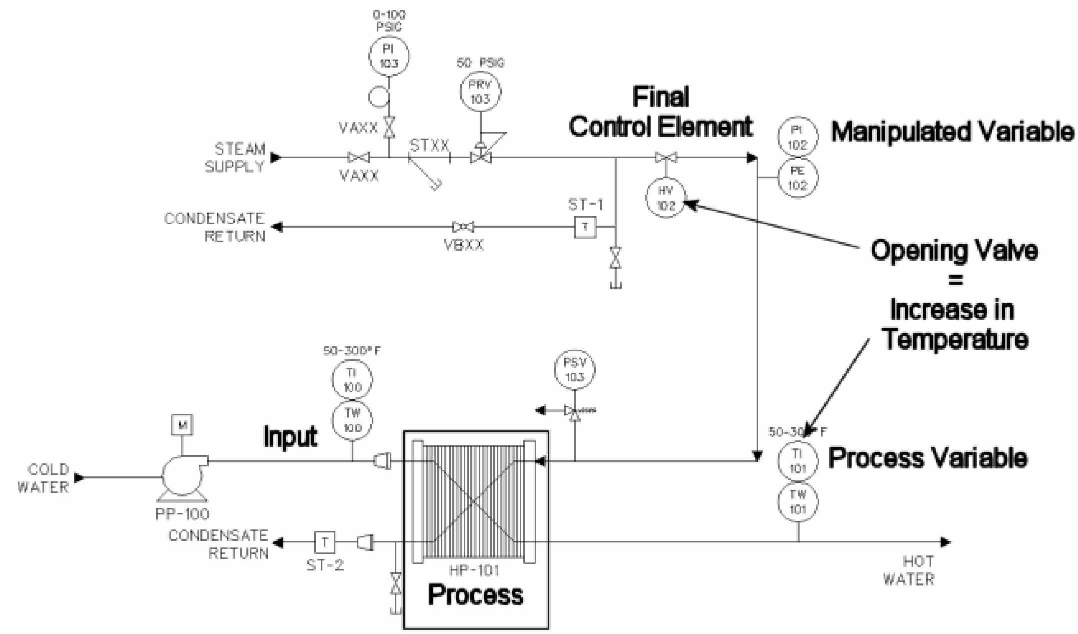


Figure 1-6

This is a controllable process because opening the valve will always lead to an increase in temperature, conversely closing the valve will always lead to a decrease in temperature. If this were not true, if sometimes on closing the valve we had an increase in temperature, the process would not be controllable.

Basics of Process Control

What is Open Loop Control?

☞ In ***open loop control*** the controller output is not a function of the process variable.

In open loop control we are not concerned that a particular Set Point be maintained, the controller output is fixed at a value until it is changed by an operator. Many processes are stable in an open loop control mode and will maintain the process variable at a value in the absence of a disturbance.

☞ ***Disturbances*** are uncontrolled changes in the process inputs or resources.

However, all processes experience disturbances and with open loop control this will always result in deviations in the process variable; and there are certain processes that are only stable at a given set of conditions and disturbances will cause these processes to become unstable. But for some processes open loop control is sufficient. Cooking on a stove top is an obvious example. The cooking element is fixed at high, medium or low without regard to the actual temperature of what we are cooking. In these processes, an example of open loop control would be the slide gate position on the discharge of a continuous mixer or ingredient bin.

Figure 1-8 depicts the now familiar heat exchanger. This is a stable process, and given no disturbances we would find that the process variable would stabilize at a value for a given valve position, say 110°F when the valve was 50% open. Furthermore, the temperature would remain at 110°F as long as there were no disturbances to the process.

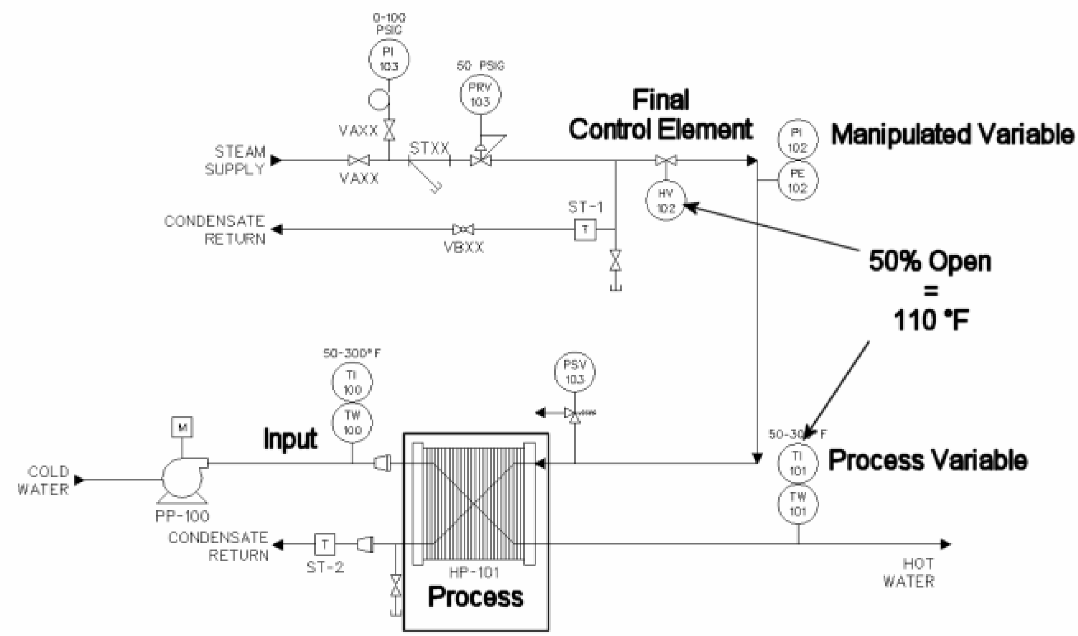


Figure 1-7

However, if we had a fluctuation in steam supply pressure, or if the temperature of the water entering the heat exchanger were to change (this would be especially true for recirculation systems with a sudden change in demand) we would find that the process would move to a new point of stability with a new exit temperature.

What is Closed Loop Control?

☞ In ***closed loop control*** the controller output is determined by difference between the process variable and the Set Point. Closed loop control is also called feedback or regulatory control

The output of a closed loop controller is a function of the error.

☞ ***Error*** is the deviation of the process variable from the Set Point and is defined as E = SP - PV.

A block diagram of a process under closed loop control is shown in figure 1-9.

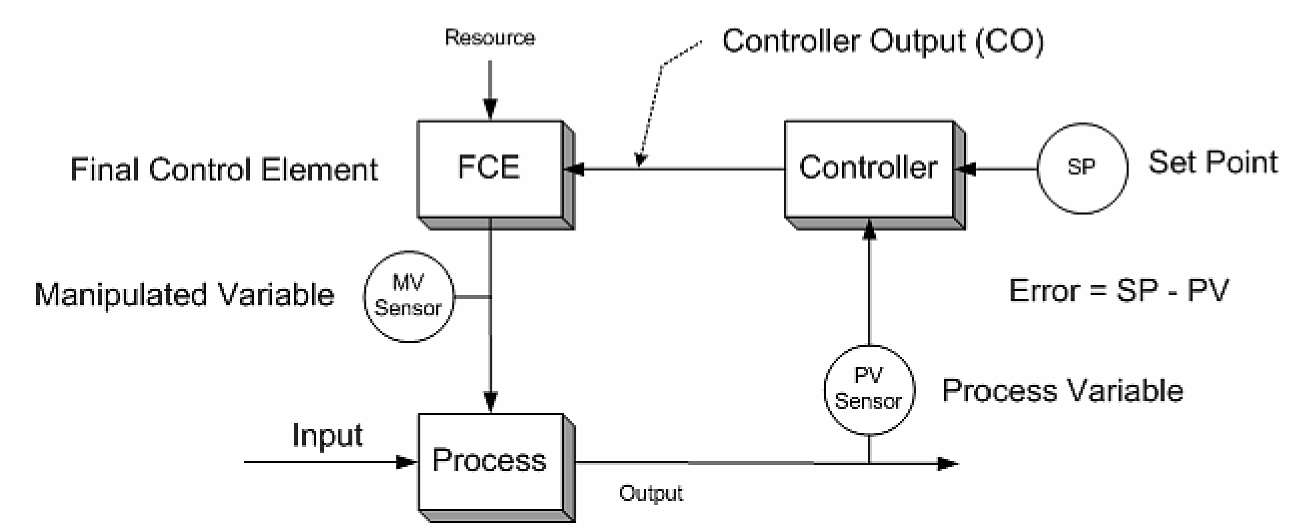


Figure 1-8

Figure 1-10 depicts the heat exchanger under closed loop control.

☞ An important point of this illustration is that the process, from the controller's perspective, is larger than just the transformation from cold to hot water within the heat exchanger. From the controllers perspective the process encompasses the RTD, the steam control valve and signal processing of the PV and CO values.

How the valve responds to the controller output and its corresponding effect on the manipulated variable (steam pressure) will determine the final effect on the process variable (temperature). The quality and responsiveness of the temperature measurement directly effects how the controller sees its effect on the process. Any filtering to diminish the effects of noise will paint a different picture of the process that the controller sees.

The dynamic behaviors of all of the elements in a control loop superimpose to form a single image of the process that is presented to the controller. To control the process requires some understanding of each of these elements.

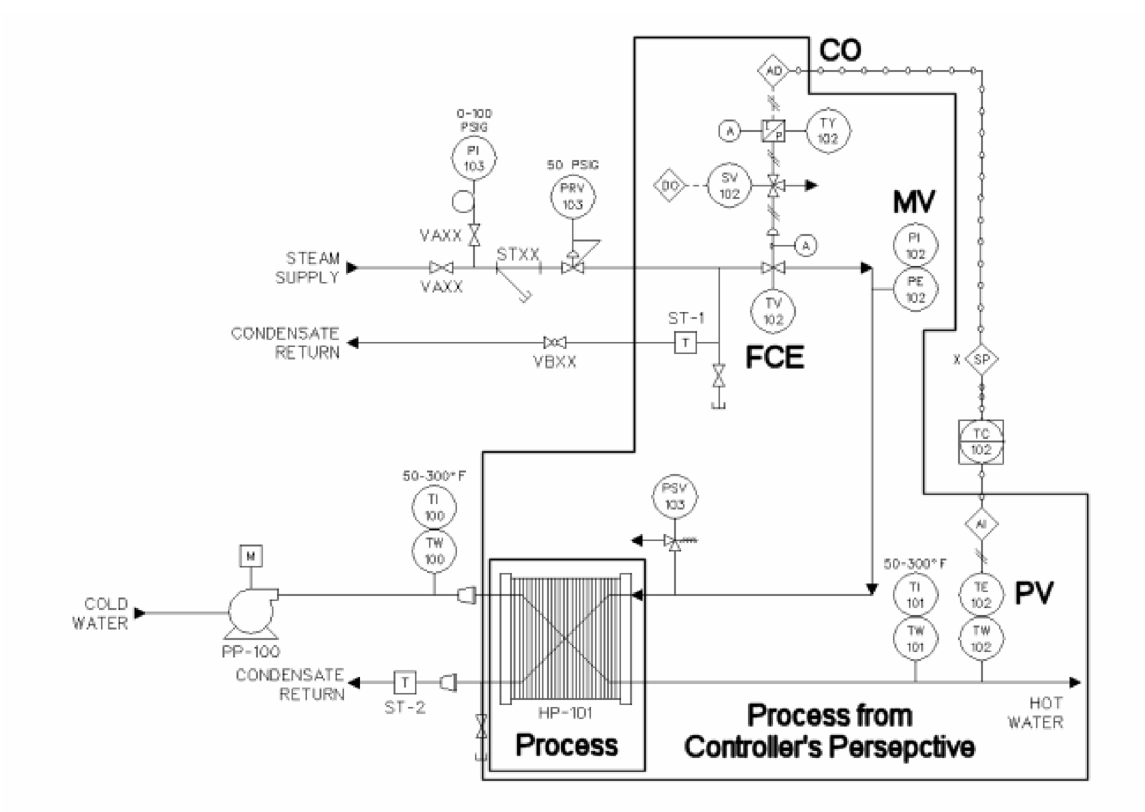


Figure 1-9

What are the Modes of Closed Loop Control?

Closed loop control can be Manual, On-Off, PID, Advanced PID (ratio, cascade, feed-forward) or Model Based depending on the algorithm that determines the controller output based on the error.

Manual Control

> ☞ In ***manual control*** an operator directly manipulates the controller output to the final control element to maintain a Set Point.

In Figure 1-11 we have placed an operator at the steam valve of the heat exchanger. Their only duty is to look at the temperature of the water exiting the heat exchanger and adjust the steam valve accordingly; we have a manual control system.

While such a system would work, it is costly (we're employing someone to just turn a valve), the effectiveness depends on the experience of the operator, and as soon as the operator walks away we are in open loop.

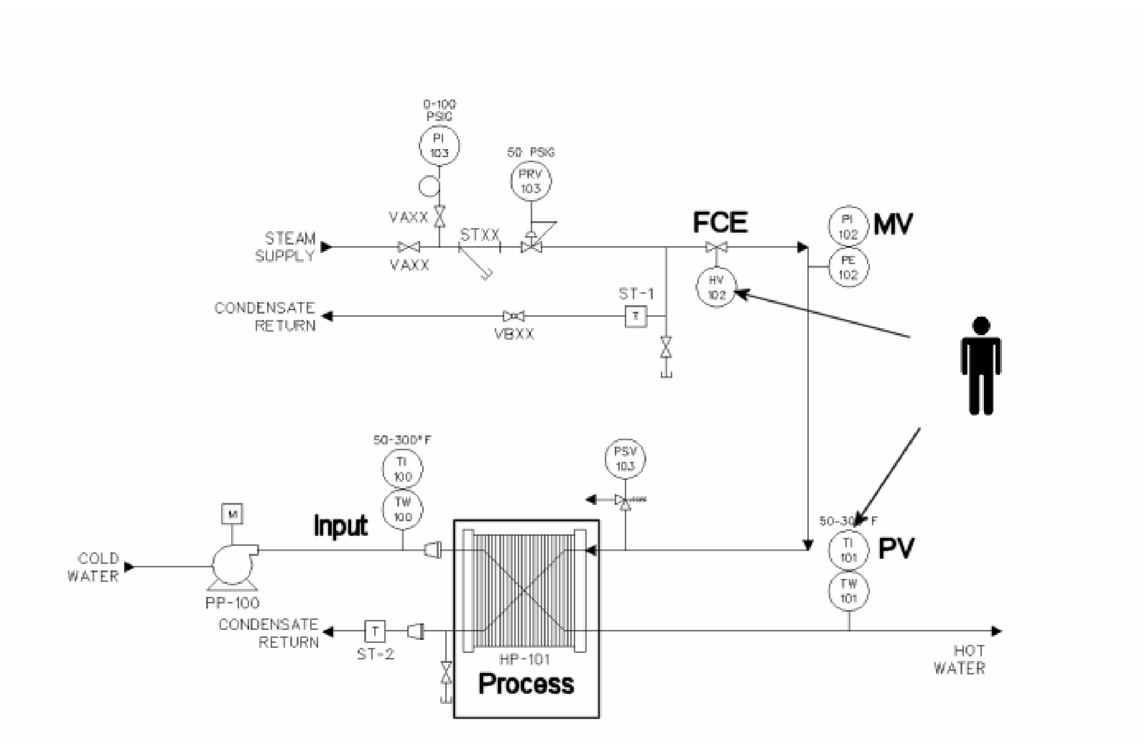


Figure 1-10

On-Off Control

> ☞ ***On-Off control*** provides a controller output of either on or off in response to error.

As an on-off controller only proves a controller output hat is either on or off, on-off control requires final control elements that have two command positions: on-off, open-closed.

In Figure 1-12 we have replaced the operator with a thermostat and installed an open-close actuator on the steam valve, we have implemented on-off control.

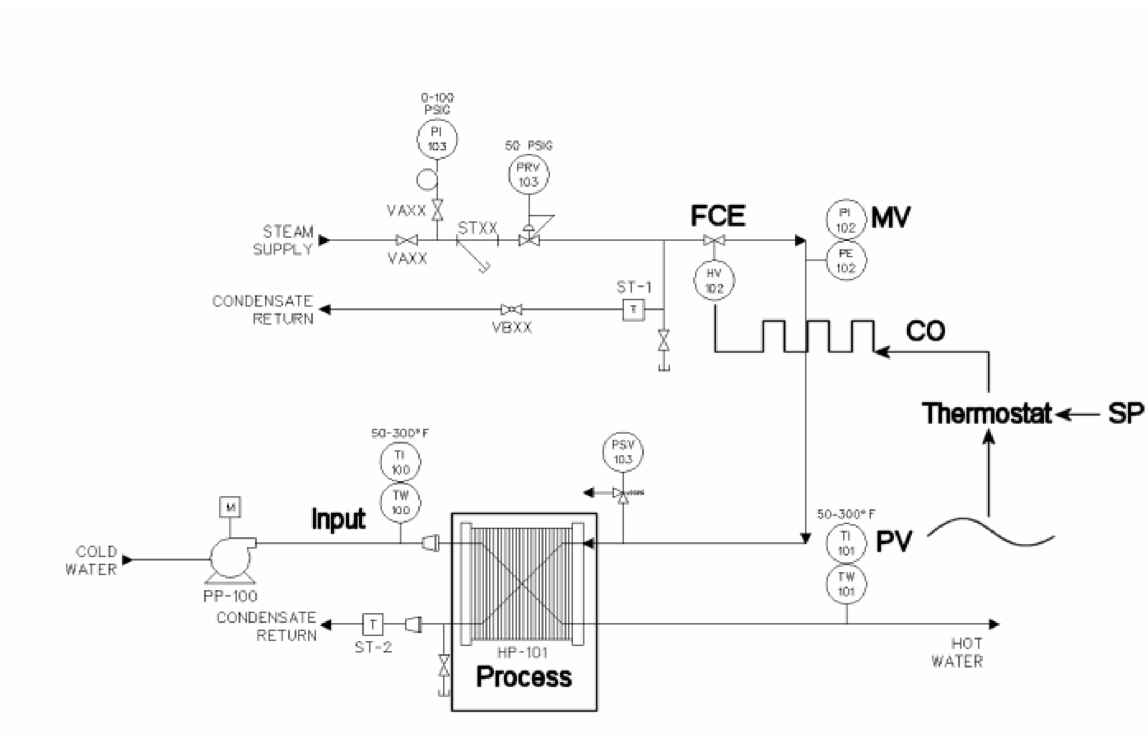


Figure 1-11

As the controller output can only be either on or off, the steam control valve will be either open or closed depending on the thermostat's control algorithm. For this example we know the thermostat's controller output must be on when the process variable is below the Set Point; and we know the thermostat's controller output must be off when the process variable is above the Set Point.

But what about when the process variable is equal to the Set Point? The controller output cannot be both on and off.

On-off controllers separate the point at which the controller changes its output by a value called the deadband (see Figure 1-13).

> ☞ Upon changing the direction of the controller output, ***deadband*** is the value that must be traversed before the controller output will change its direction again.

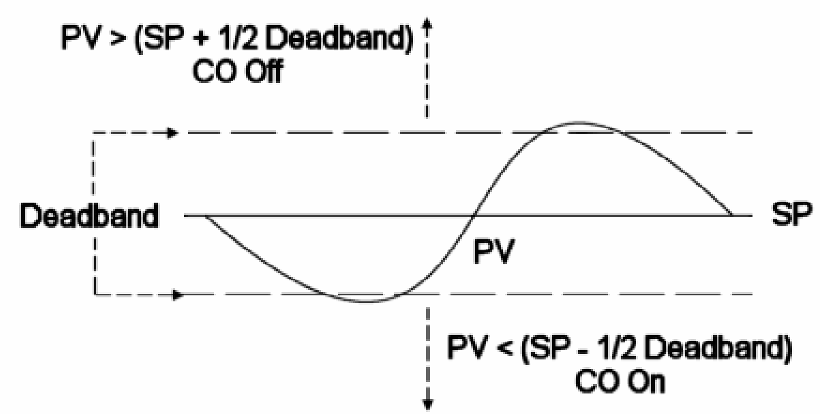


Figure 1-12

On the heat exchanger, if the thermostat is configured with a 110°F Set Point and a 20°F deadband, the steam valve will open at 100°F and close at 120°F. If such a large fluctuation from the Set Point is acceptable, then the process is under control.

If this fluctuation is not acceptable we can decrease the deadband, but in doing so the steam valve will cycle more rapidly, increasing the wear and tear on the valve, and we will never eliminate the error (remember, the thermostat cannot be both on and off at 110F).

PID Control

> ☞ ***PID control*** provides a controller output that modulates from 0 to 100% in response to error.

As an on-off controller only proves a controller output that is either on or off, on-off control requires devices that have two command positions: on-off, open-closed.

As a PID controller provides a modulating controller output, PID control requires final control elements that have can accept a range of command values, such as valve position or pump speed.

> ☞ To ***modulate*** is to vary the amplitude of a signal or a position between two fixed points.

The advantage of PID control over on-off Control is the ability to operate the process with smaller error (no deadband) with less wear and tear on the final control elements.

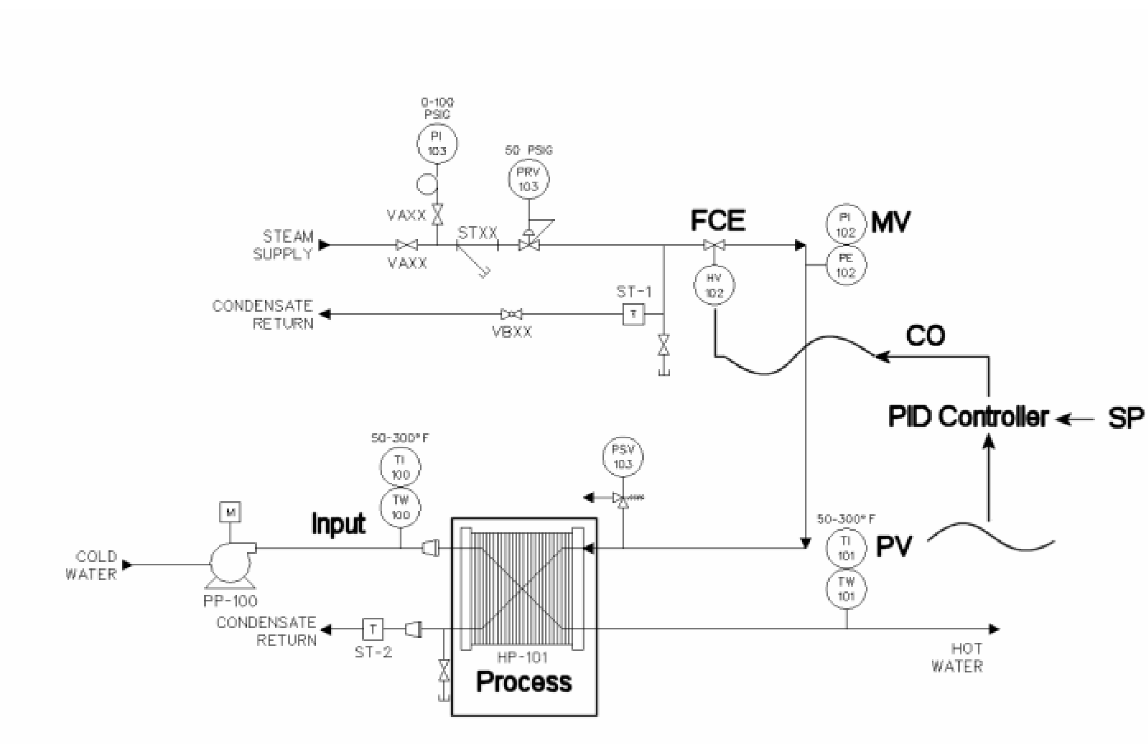


Figure 1-13

Time Proportion Control

☞ ***Time proportion control*** is a variant of PID control that modulates the on-off time of a final control element that only has two command positions.

To achieve the effect of PID control the switching frequency of the device is modulated in response to error. This is achieved by introducing the concept of cycle time.

Cycle Time is the time base of the signal the final control element will receive from the controller. The PID controller determines the final signal to the controller by multiplying the cycle time by the output of the PID algorithm.

In Figure 1-15 we have a time proportion controller with a cycle time of 10 seconds. When the PID algorithm has an output of 100% the signal to the final control element will be on for 10 seconds and then repeat. If the PID algorithm computes a 70% output the signal to the final control element will be on for 7 seconds and off for 3 and then repeat.

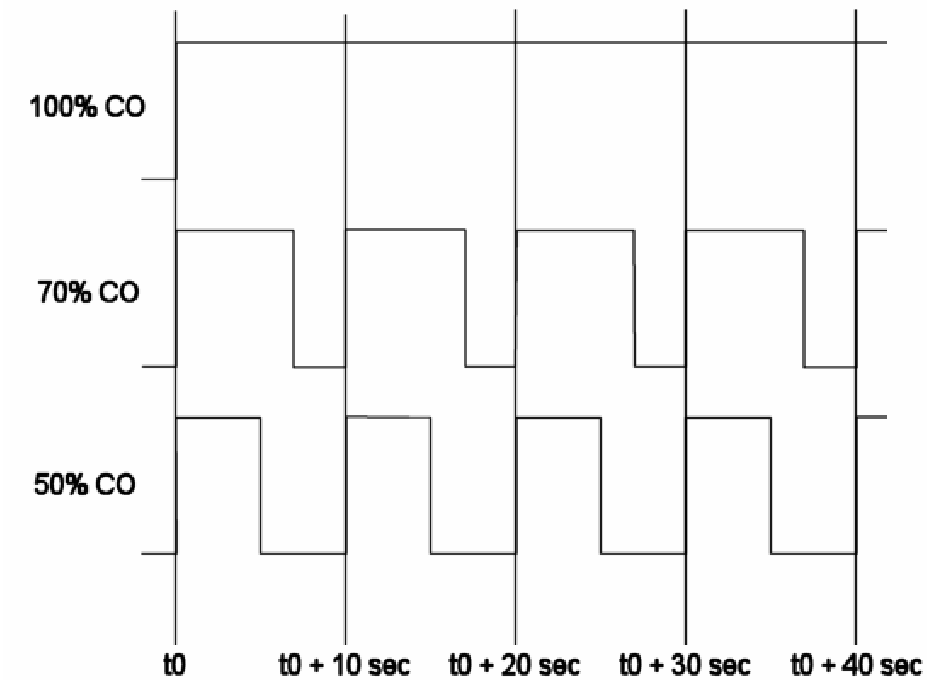


Figure 1-14

While time proportion control can give you the benefits of PID control with less expensive final control elements it does so at the expense of wear and tear on those final control elements. Where used, output limiting should be configured on the controller to inhibit high frequency switching of the final control element at low controller outputs.

What are the Basic Elements of Process Control?

☞ Controlling a process requires knowledge of four basic elements, the ***process*** itself, the ***sensor*** that measures the process value, the ***final control element*** that changes the manipulated variable, and the ***controller***.

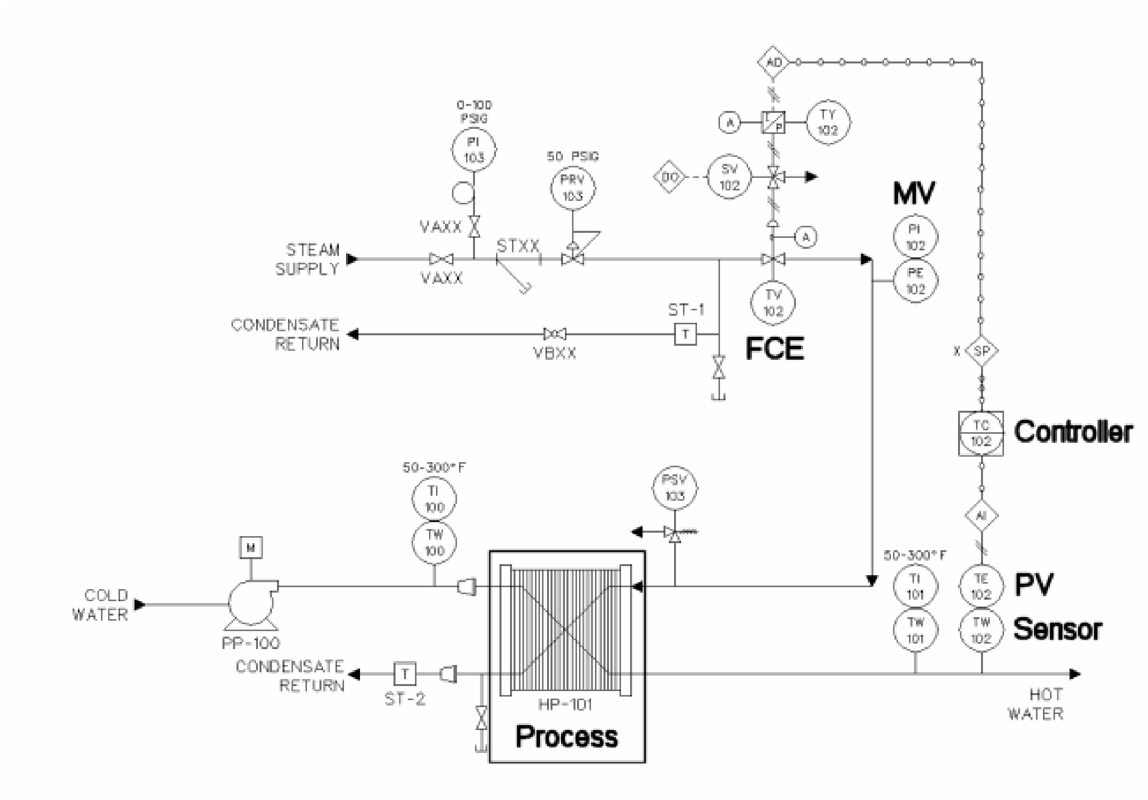


Figure 1-15

The Process

We have learned that processes have a dynamic behavior that is determined by physical properties; as such they cannot be altered without making a physical change to the process. We will be learning more about process dynamics in Chapter 2.

Sensors

Sensors measure the value of the process output that we wish to effect. This measurement is called the Process Variable or PV. Typical Process Variables that we measure are temperature, pressure, mass, flow and level. The Sensors we use to measure these values are RTDs, pressure gauges and transducers, load cells, flow meters and level probes. We will be learning more about sensors in Chapter 3.

Final Control Elements

A Final Control Element is the physical device that receives commands from the controller to manipulate the resource. Typical Final Control Elements used in these processes are valves and pumps. We will be learning more about final control elements in Chapter 4.

The Controller

A Controller provides the signal to the final element. A controller can be a person, a switch, a single loop controller, or DCS / PLC system. We will be learning more about PID controllers in Chapter 5. We will be learning about tuning PID controllers in Chapter 6.

Process Characteristics

Objectives:

In this chapter you will learn:

- What is a First Order Process?
- What is Process Dead Time?
- What is the Process Time Constant?
- What is Process Gain?
- What is Process Action?
- What are Higher Order Processes?
- What is a Linear Process?
- What is a Nonlinear Process?
- What are Self-Regulating Processes?
- What are Integrating Processes?

Introduction: Process Order

Process control theory is based on the insight gained through studying mathematical models of processes. A branch of mathematics called differential equations is used to build these models. Differential equations are equations that contain derivatives of variables. The order of a differential equation is the highest number of derivatives of a variable that is contained within the equation. The order of a process is the order of the differential equation that is required to model it.

Process order is an important concept because it is a description of how a process will respond to controller action. Fortunately we do not need to delve into the world of mathematics to gain a practical knowledge of process order; we simply perform a step test on the process and let the process reaction curve tell us.

> ☞ A ***reaction curve*** is a graph of the controller output and process variable with respect to time.

Reaction curves are obtained after the process variable has stabilized by making a step change in the controller output. The properties of the reaction curve will tell us all we need to know about controlling a particular process.

First Order Processes

Lesson 1. What is a First Order Process?

☞ A ***first order process*** has an exponential response to a process step change and can be completely characterized by three parameters: dead time, time constant and gain.

Figure 2-1 is a reaction curve of a first order process. This reaction curve shows the PV response to a 5 percent change in the controller output.

Understanding the FOPDT (First Order Plus Dead Time) process model is the foundation for understanding PID control. Becoming knowledgeable of dead time, time constant and gain will greatly aid your tuning efforts of PID controllers.

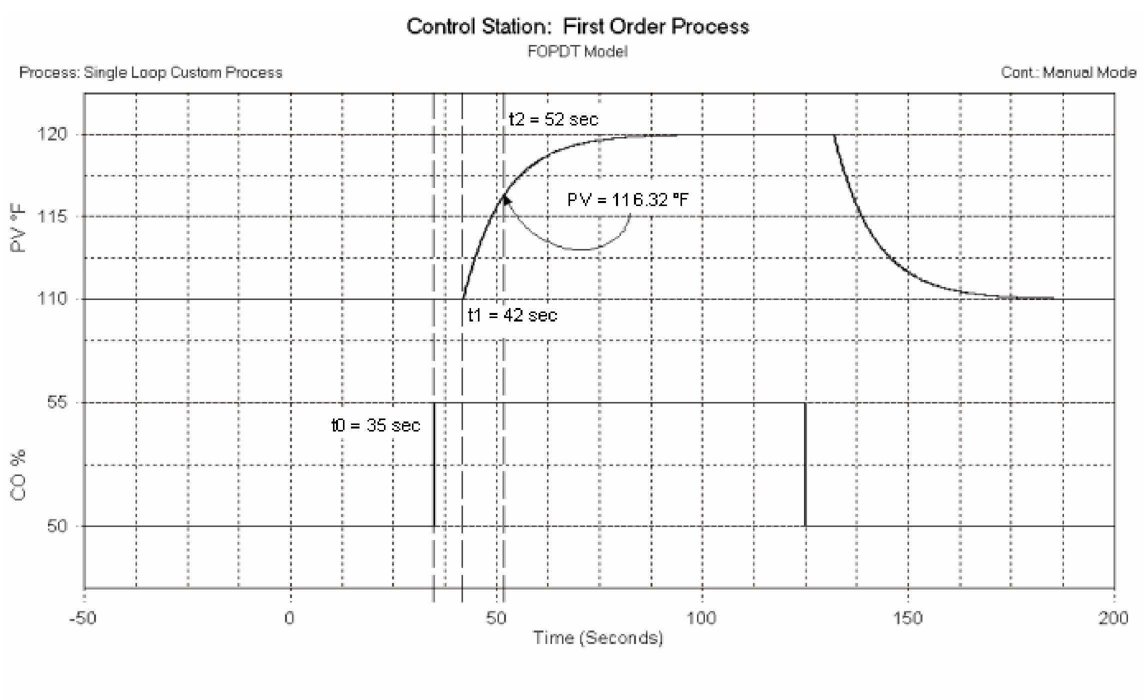


Figure 2-1

What is Process Dead Time?

☞ ***Process dead time*** is the period of time that passes between a change in the controller output and a change in the process variable being measured.

Dead time is often the result of transportation delays (material on a belt, compressible material in a pipe) although sensors and final control elements may add to process dead time. Dead time is the enemy of loop tuning, the amount of dead time in a process will determine how "tightly" the process can be tuned and remain stable.

Measuring Dead time

In Figure 2-2 we see that the controller output was changed at t_0 = 35 seconds. It was not until t_1 = 42 seconds that the process variable started to change. The dead time in this example is

$$\text{Deadtime} = t_1 - t_0 = 42\,\text{seconds - }35\,\text{seconds} = 7\text{seconds}$$

☞ Process dead time as seen by a controller is a function of the dead times of the sensor, the final control element and the process itself.

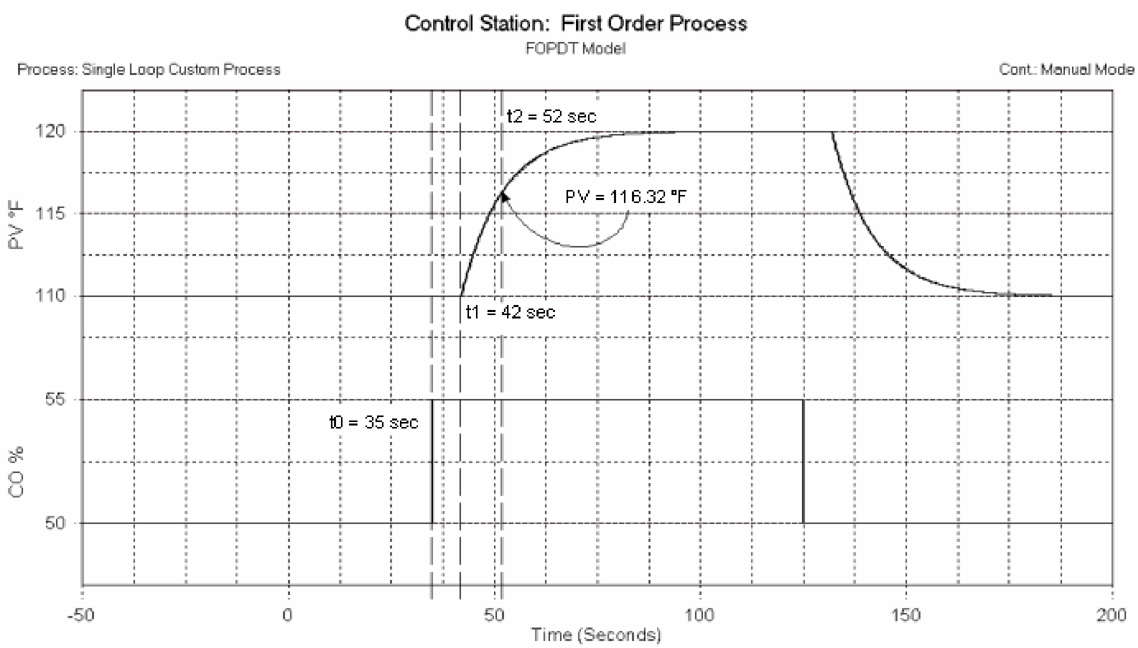


Figure 2-2

What is the Process Time Constant?

☞ A ***process time constant*** is the amount of time for the process variable to reach 63.2 percent of its final value in response to a step change in a first order process.

For those familiar with RC circuits in electronics will recall that voltage in a capacitor is an exponential function of time, and the time constant RC is the time required for a capacitor to reach 63.2 percent of the applied voltage. In fact, mathematically a first order process and an RC circuit are identical in behavior; a first order process has an ability to store energy just as a capacitor has an ability to store charge.

We will find in Chapter 6 that the process time constant will determine the amount of integral action that should be configured in a PID controller.

Measuring the Time Constant

To find the time constant of this example FOPDT process, we must find the process value that represents a 63.2 percent change in response to the step change, and from the trace determine the time that this value of the PV occurred.

From Figure 2-1 we see that the process value was stable at 110°F prior to the step change in the controller output. After the step change the process variable stabilized at 120°F. The change in the process variable is

$$PV_{Final} - PV_{Initial} = 120\,°F - 110\,°F = 10\,°F$$

$$63.2\%\; x\, 10\,°F = 0.632\; x\, 10\,°F = 6.32\,°F$$

The PV value after one time constant will be

$$110\,°F + 6.32\,°F = 116.32\,°F$$

From the reaction curve in figure 2-2 we see this value for the PV occurred at t_2 = 52 seconds. The time constant will be this time minus the time at which the PV started to change.

$$TC = t_2 - t_1 = 52\,\text{seconds} - 42\,\text{seconds} = 10\,\text{seconds}$$

The process time constant is often referred to as a lag, and sometimes the process order is included in the lag. When a first order lag is mentioned it is referring to the time constant of a first order process, a second order lag would refer to the time constant of a second order process.

☞ The Process time constant as seen by a controller is a function of the time constants of the sensor, the final control element and the process itself.

Controllability of a Process

☞ The relationship between dead time and process lag, in general, determine the controllability of the process. Processes where the dead time is less than the time constant (dead time ÷ time constant < 1) are considered easier to control. Processes where the dead time is greater than the lag (dead time ÷ time constant > 1) are more difficult to control as the controller must be detuned to maintain stability.

In the example:

$$\frac{\text{deadtime}}{\text{time constant}} = \frac{7\ \text{seconds}}{10\ \text{seconds}} = 0.7 < 1$$

Therefore the process should be relatively easy to control.

What is Process Gain?

☞ ***Process gain*** is the response of the process variable to a change in the controller output, or the change in the process variable divided by the change in the controller output.

Measuring Process Gain

From Figure 2-1 the change in the process variable is

$$\Delta PV = PV_{Final} - PV_{Initial} = 120\,°F - 110\,°F = 10\,°F$$

☞ We will be using the delta symbol (Δ) to represent change, so ΔPV is the change in PV.

The change in the controller output is

$$\Delta CO = CO_{Final} - CO_{Initial} = 55\% - 50\% = 5\%$$

The gain for the process is

$$\frac{\Delta PV}{\Delta CO} = \frac{10°F}{5\%} = 2\,\frac{°F}{\%}$$

☞ Process gain as seen by a controller is the product of the gains of the sensor, the final control element and the process itself.

$$\text{Process Gain} = \text{Gain}_{\text{Process}}\ x\ \text{Gain}_{\text{Sensor}}\ x\ \text{Gain}_{\text{Final Control Element}}$$

☞ The gain of a controller will be inversely proportional to the process gain that it sees

$$\text{Controller Gain} \propto \frac{1}{\text{Process Gain}}$$

Making Gains Unitless

There is one important caveat in this process; the gain we have calculated has units of °F/%. Real world controllers, unlike most software simulations, have unitless gain values.

When calculating the gain for a real controller the change in PV needs to be expressed in percent of span of the PV as this is how the controller calculates error.

In this example, the particular simulation that generated this reaction curve had in input range of 0 to 500°F, giving us an input span of 500°F. The change in PV as a percent of span would then be

$$\frac{\Delta PV}{PV_{Span}} \, x \, 100\% = \frac{PV_{Final} - PV_{Initial}}{PV_{Span}} \, x \, 100\%$$

$$= \frac{120\,°F - 110\,°F}{500\,°F} \, x \, 100\% = \frac{10\,°F}{500\,°F} \, x \, 100\% = 0.02 \, x \, 100\% = 2\%$$

and

$$\frac{\Delta CO}{CO_{Span}} \, x \, 100\% = \frac{CO_{Final} - CO_{Initial}}{CO_{Span}} \, x \, 100\%$$

$$= \frac{55\% - 50\%}{100\%} \, x \, 100\% = \frac{5\%}{100\%} \, x \, 100\% = 0.05 \, x \, 100\% = 5\%$$

Our process gain would now be

$$\frac{2\%}{5\%} = 0.4 \frac{\%}{\%}$$

☞ A shortcut method to convert the gain into the proper units is to use the following calculation:

$$\text{Gain} = \frac{\Delta PV}{\Delta CO} \, x \, \frac{CO_{Span}}{PV_{Span}} = \frac{10°F}{5\%} \, x \, \frac{100\%}{500°F} = 0.4\,\% \,/\, \%$$

Values for Process Gain

As we have seen the process gain that the controller sees is influenced by two factors other than the process itself, the size of the final control element and the span of the sensor.

In the ideal world you would use the full span of both final control element and the sensor which would give a process gain of 1.0.

☞ As a rule of thumb, scaled process gains that are greater than 1 are a result of oversized final control elements. Process gains less than 1 are a result of sensor spans that are too wide. For the heat exchanger to achieve a process gain of 1, we would need a sensor with a span of 200°F, say a range of 30 to 230°F.

The result of a final control element being too large (high gain) is:

1. The controller gain will have to be made correspondingly smaller, smaller than the controller may accept.
2. High gains in the final control element amplify imperfections (deadband, stiction), control errors become proportionately larger.

If a sensor has too wide of a span:

1. You may experience problems with the quality of the measurement.
2. The controller gain will have to be made correspondingly larger making the controller more jumpy and amplifying signal noise.
3. An over spanned sensor can hide an oversized final element.

The general rule of thumb is the process gain for a self regulating process should be between 0.5 and 2.0.

What is Process Action?

> ☞ ***Process action*** is how the process variable changes with respect to a change in the controller output. Process action is either direct acting or reverse acting.

The action of a process is defined by the sign of the process gain. A process with a positive gain is said to be direct acting. A process with a negative gain is said to be reverse acting.

On the hot water system, if we open the control valve 10% more from its current position and the temperature increases by 20°F the process gain is 20°F/10% or 2°F/%.

If this were a cooling application we could expect the temperature to change by -20F by opening a glycol valve by 10% more. The process gain in this case would be -2°F/%.

Another way to think of process action is with a direct acting process the process variable will increase with an increase in the controller output. In a reverse acting process the process variable will decrease with an increase in the controller output. This is illustrated in Figure 2-3.

Figure 2-3

Process Action and Controller Action

> ☞ The action of a process is important because it will determine the action of the controller. A direct acting process requires a reverse acting controller; conversely a reverse acting process requires a direct acting controller.

The controller must be a mirror of the process; if you put a direct acting controller on a direct acting process you will have a runaway condition on your PV.

Process Orders

Higher Order Processes

What are Higher Order Processes?

☞ ***Higher order processes***, unlike first order processes, can exhibit an oscillatory response to a step change. The oscillatory behavior a process exhibits on its own places it into one of three process types: Over-damped, Under-damped or Critically Damped.

Reaction curves for the three order types are shown in Figure 2-4.

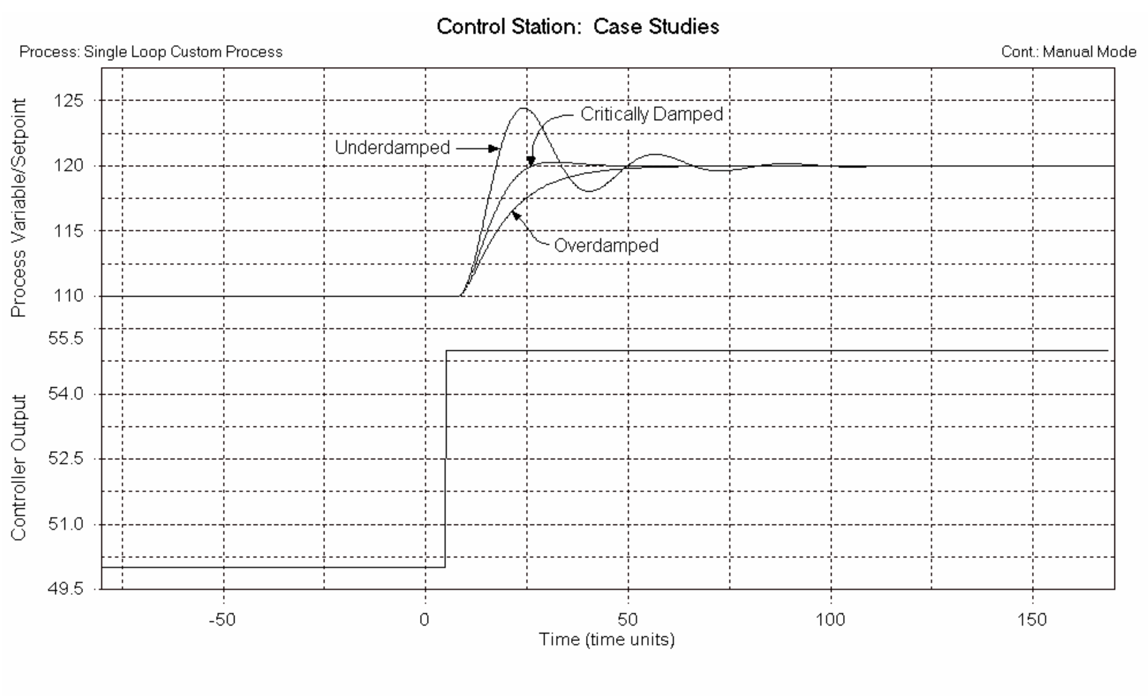


Figure 2-4

Over-damped Response

Higher order processes that are over-damped look very much like FOPDT processes. The difference between first order and higher order over-damped processes is the initial response to a step change. A first order process has a crisper response to a controller step change after the dead time has passed compared to higher order processes.

In general, the higher the process order the more "S" shaped the reaction curve will be and the initial response to a step change will be more sluggish.

Figure 2-5 compares over-damped second and third order processes to a first order process.

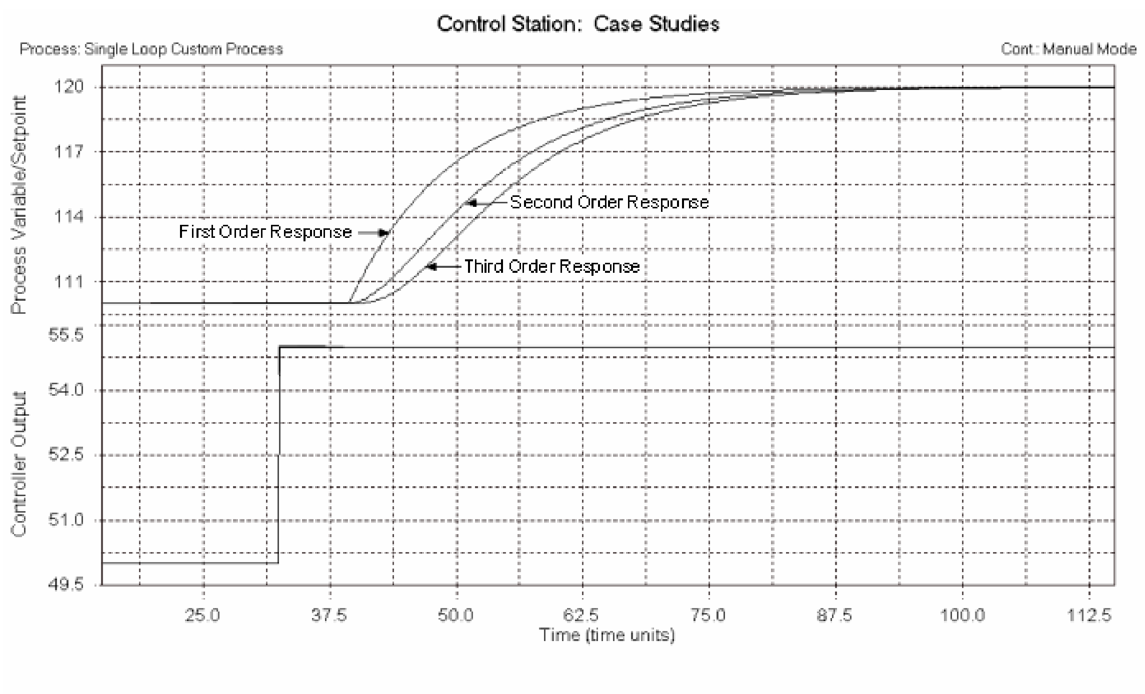


Figure 2-5

First Order Fit of Higher Order Over-Damped Processes

For tuning purposes over-damped higher order process are treated like FOPDT processes. The sluggish response to a controller step change is treated as additional dead time. Figure 2-6 shows a reaction curve of a Second Order Plus Dead Time process. Overlaid with the trace is a FOPDT model fit of the process data.

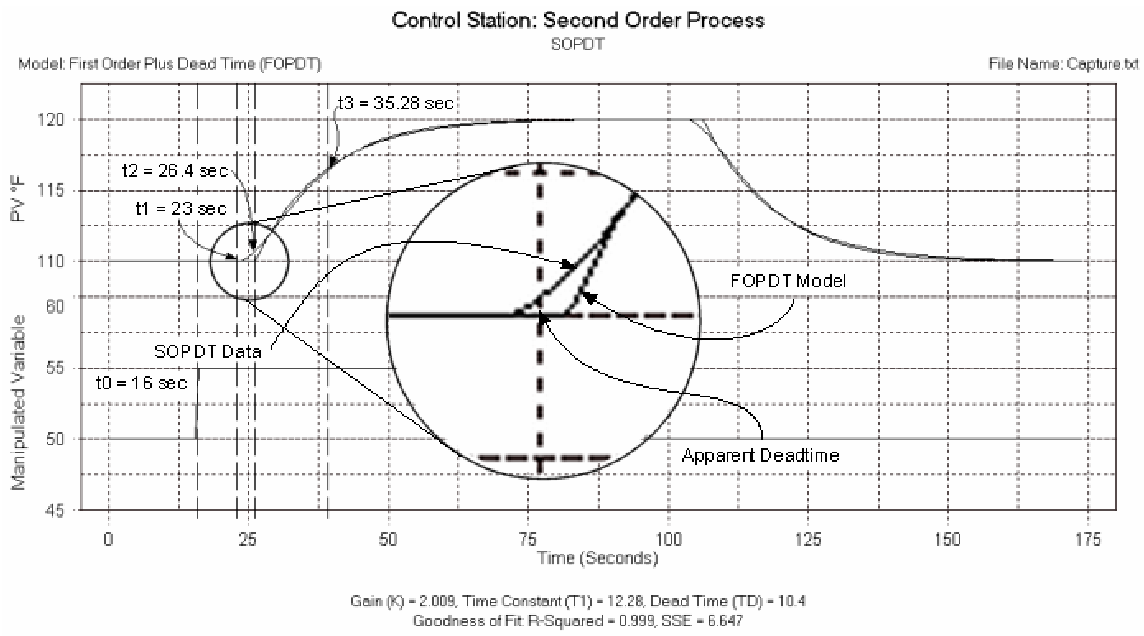


Figure 2-6

This SOPDT process in Figure 2-6 has the same 7 second dead time as the FOPDT in Figure 2-2. The sluggish response to the controller step change adds 3.4 seconds of apparent dead time to the process. Whereas the (dead time) ÷ (time constant) was 0.7 in the FOPDT process, the SOPDT has a (dead time) ÷ (time constant) value of 0.85 and will not be able to tuned as aggressively.

First Order Fit of Higher Order Under-Damped Response

Higher order processes that are under-damped will oscillate on their own in response to a controller step change. Figure 2-7 is a reaction curve of an under-damped process and the FOPDT fit to obtain the process dead time, time constant and gain.

We see that the FOPDT model is not a very good fit for an under-damped response and therefore would expect rules based tuning parameters may require a good deal of "tweaking" to bring this process under control. Also, you will find that there are no tuning values that will remove the oscillations from an under-damped process.

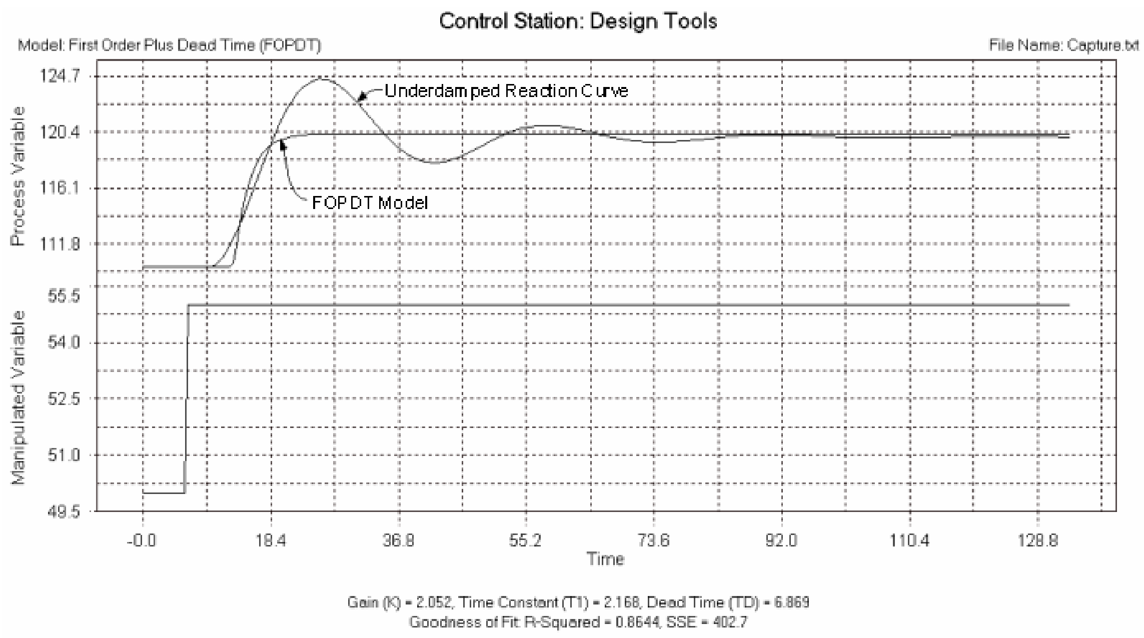


Figure 2-7

Critically Damped Response

Higher order processes that are critically damped will overshoot and then settle in to their final value but they will not oscillate. Figure 2-8 is a reaction curve of a critically damped process and the FOPDT fit to obtain the process dead time, time constant and gain. Like the over-damped process, the FOPDT model is a good fit for a critically damped process.

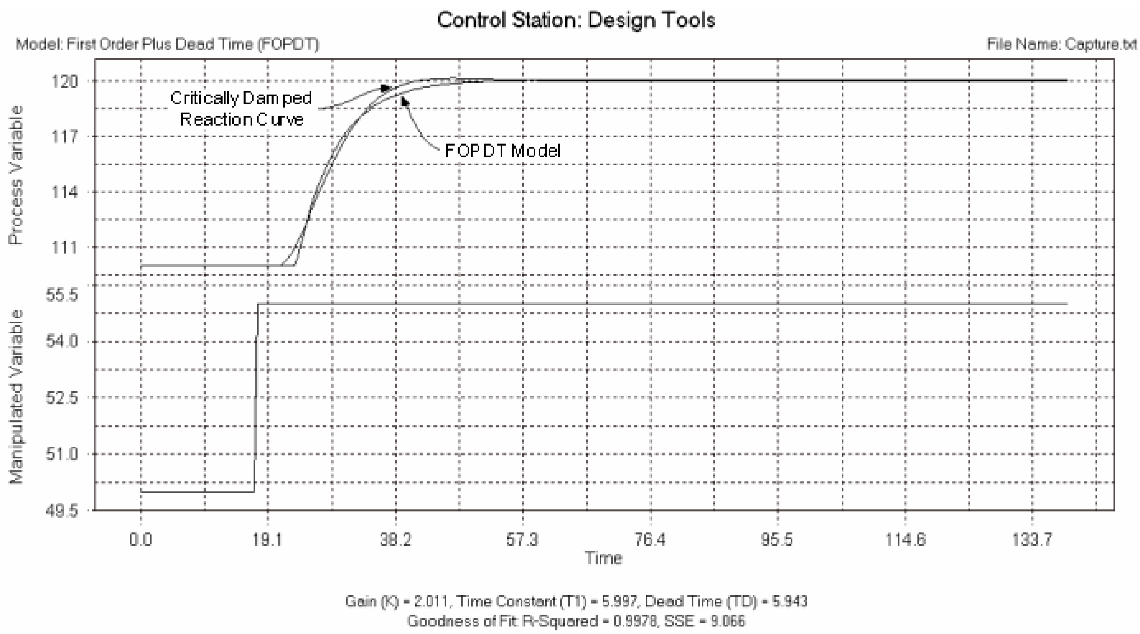


Figure 2-8

What is a Linear Process?

☞ Linear means in a line, non-varying. A ***Linear process*** is one that has non-varying process characteristics over the range of the process variable.

No matter what the current value of the process variable, a step change to the process will produce an identical reaction curve. Figure 2-9 illustrates a linear process. As the controller output is stepped in equal increments from 0 to 100% the process reacts identically to each step.

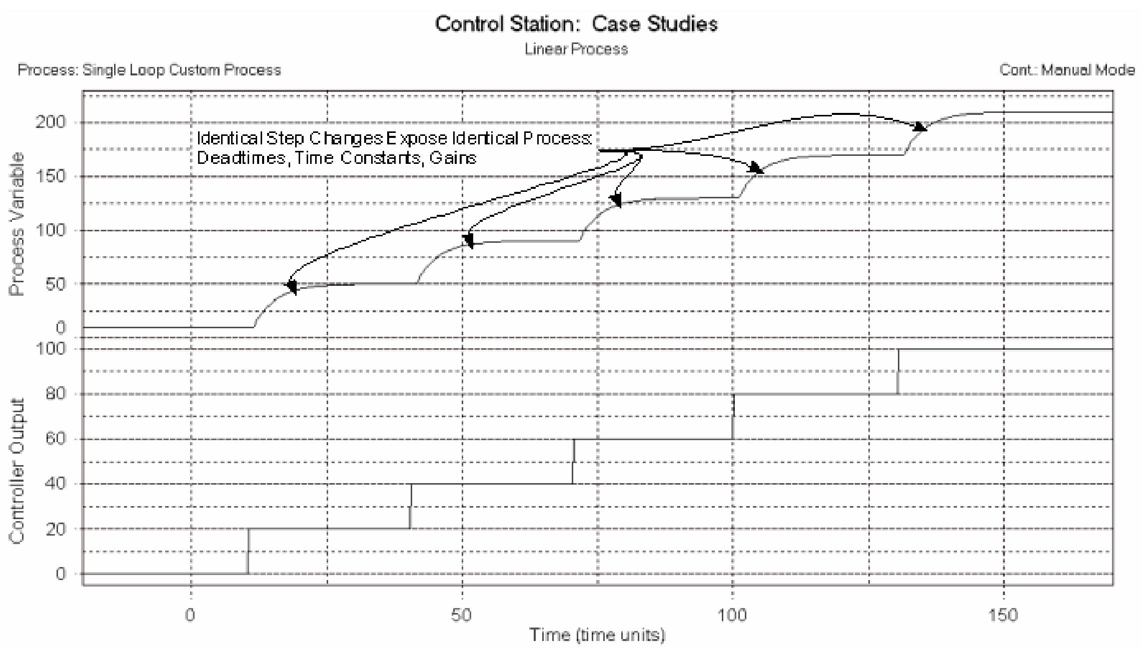


Figure 2-9

Linear processes are the goal of process design for they are the easiest to control and tune. A properly tuned linear process will handle process disturbances and Set Point changes equally well.

We know that from the controller's perspective a process is comprised of the sensor, final control element and the process itself. To achieve a linear process all of these elements must be linear over the range of their operation.

What is a Nonlinear Process?

☞ A ***nonlinear process*** is one that has varying process characteristics over the range of the process variable.

The reaction curve produced by a process step change will depend on the current value of the process variable. . Figure 2-10 illustrates a nonlinear process. Unlike the linear process, as the controller output is stepped in equal increments from 0 to 100% the process reacts differently to each step.

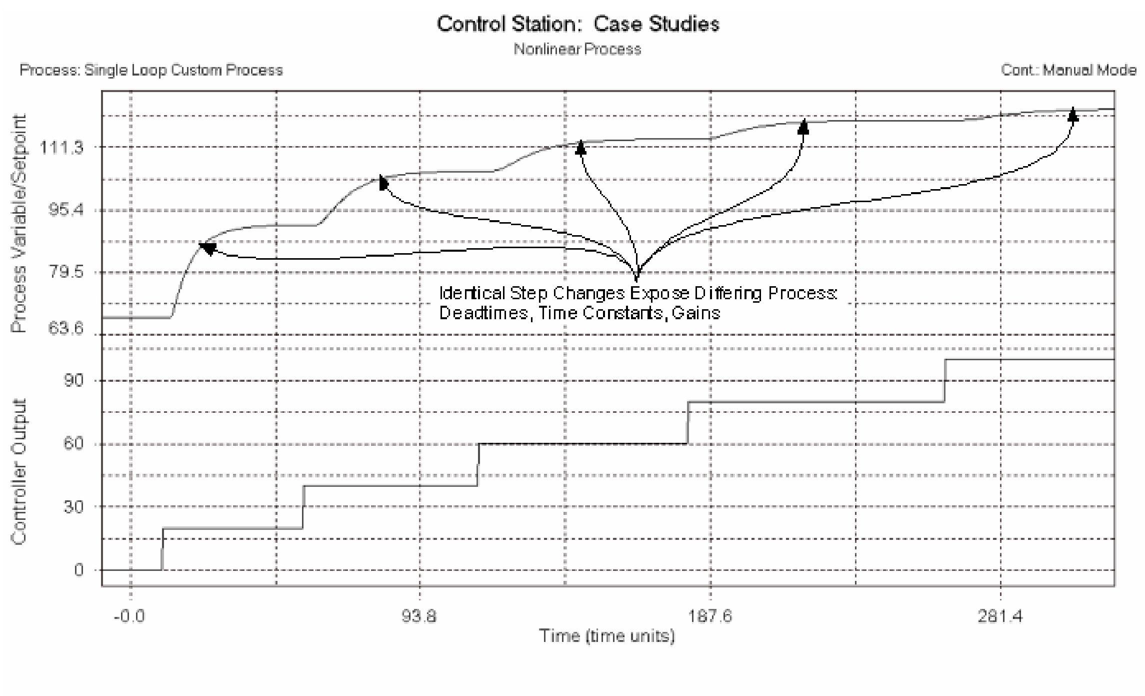


Figure 2-10

While linear processes may be the design goal of most process engineers, the reality is that most processes are nonlinear in nature due to nonlinearity in the final control element or the process itself.

For example, a heating process is nonlinear because the rate at which heat is transferred between two objects depends on the difference in temperature between the objects, or a valve that is linear in the middle of its operating range may become very nonlinear towards its limits.

Dealing with Nonlinearity

The impact of controlling a nonlinear process is that you must choose you operating range for which your tuning parameters are to be valid. Many times this means choosing between tuning for Set Point response and disturbance rejection.

Disturbance Rejection

Figure 2-11 is of a nonlinear process. If we were to control a process that had a fixed 110°F Set Point and only expected disturbances of +- 5°F in the input stream we would tune the controller for optimal disturbance in this region, realizing is the Set Point is changed we may have an unstable system.

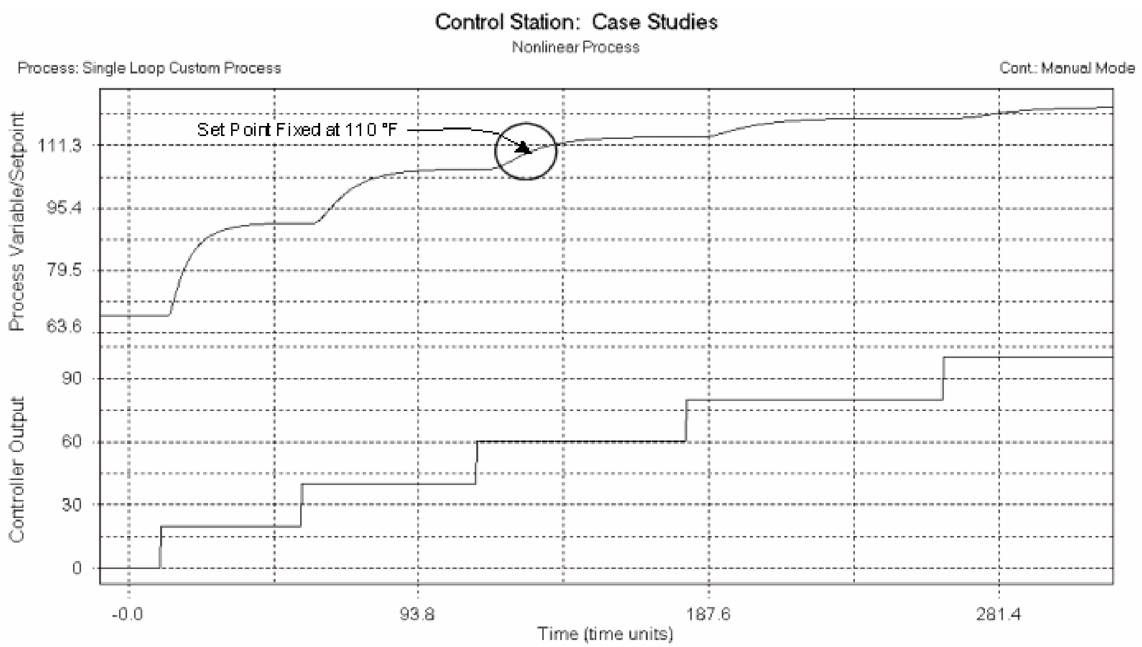


Figure 2-11

Set Point Response

Figure 2-12 is of the same process, only now we wish to tune to handle a range of Set Points form 90°F to 110°F. In this case we must tune the controller differently. To have a stable process we would have to tune the controller to operate where the process has its highest gain and accept sluggish response where the process has lower gains.

In this example we would tune the controller at 90°F as we have the highest process gains in this region and accept sluggish response at higher Set Points. To improve controller response across wide operating ranges in a nonlinear process requires advanced techniques such as gain scheduling or a controller with adaptive tuning.

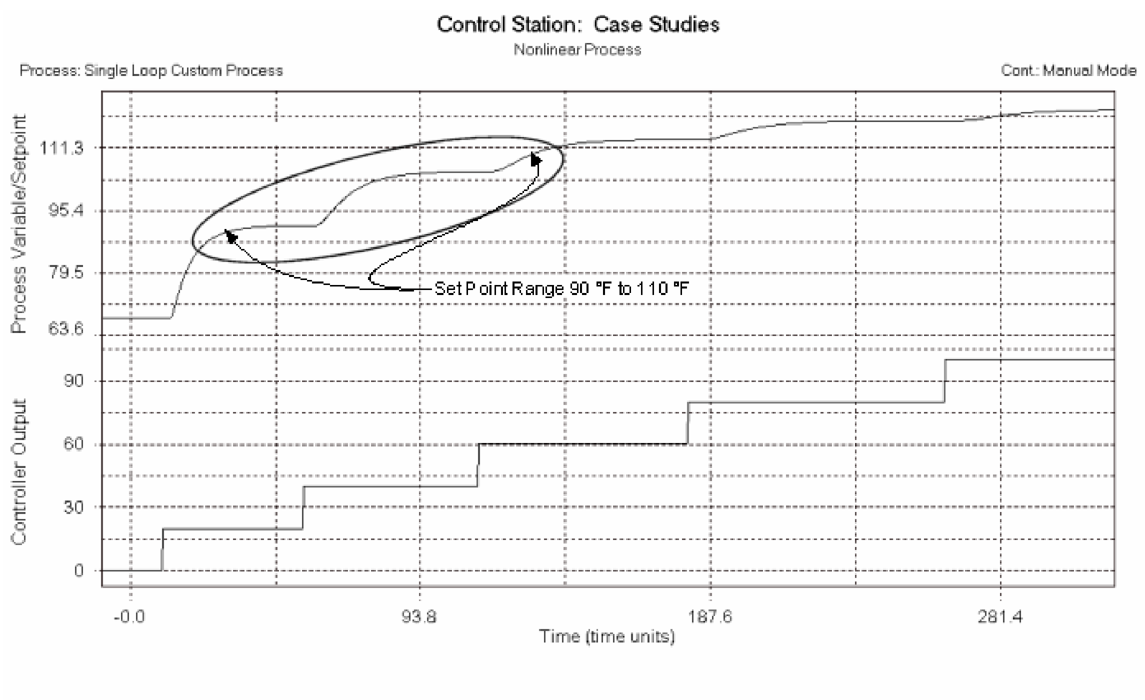


Figure 2-12

> ☞ The robustness of a controller is a measure the range of process values over which the controller provides stable operation. The more nonlinear a process is, the less aggressive you must be in you tuning approach to maintain robustness.

> ☞ Normally you would not want to disturb a process by stepping a controller through such a wide range as was done in Exercise 2. The safer approach would be to assume the process is nonlinear and determine the range at which the process must operate. Do a step test at the extremes of your range and tune where the process gain is the highest. This will give you your lowest controller gain and most robust response.

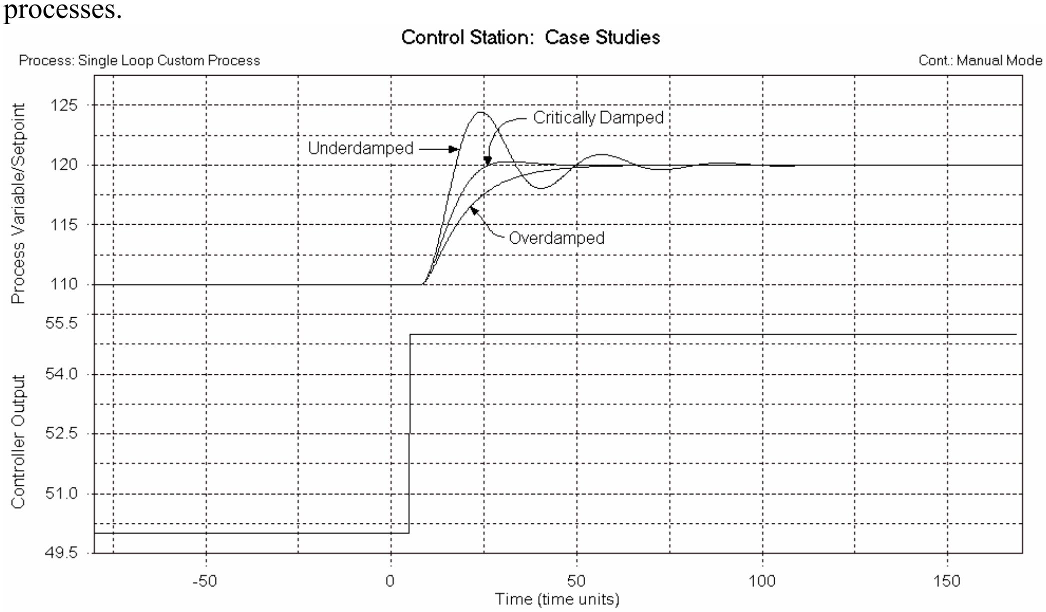
Process Type

Processes fall into one of two types, Self-Regulating and Integrating processes. Your approach to tuning will be dependent on the type of process.

What are Self-Regulating Processes?

> ☞ In a ***self-regulating process*** the process variable is stable in an open loop configuration.

A self-regulating process that is stable at a particular process variable will stabilize at a new process variable in response to a step change in the controller output. As illustrated in Figure 2-13, all the processes we have discussed so far are self-regulating. We will find later on that most tuning rules are only valid for self-regulating processes.



What are Integrating Processes?

☞ In an ***integrating process*** the process variable is stable in an open loop configuration only at its balance point.

An integrating process relies on balancing the process inputs and outputs to remain stable. Making a step change in the controller output of a stable integrating process will cause the process variable to move increasingly in one direction.

Figure 2-14 shows the reaction curve of an integrating process. The process was stable at a controller output of 50% with a process variable of 90. The controller output was changed to 51% but a new stabilization point was never attained, the process variable eventually went to 0.

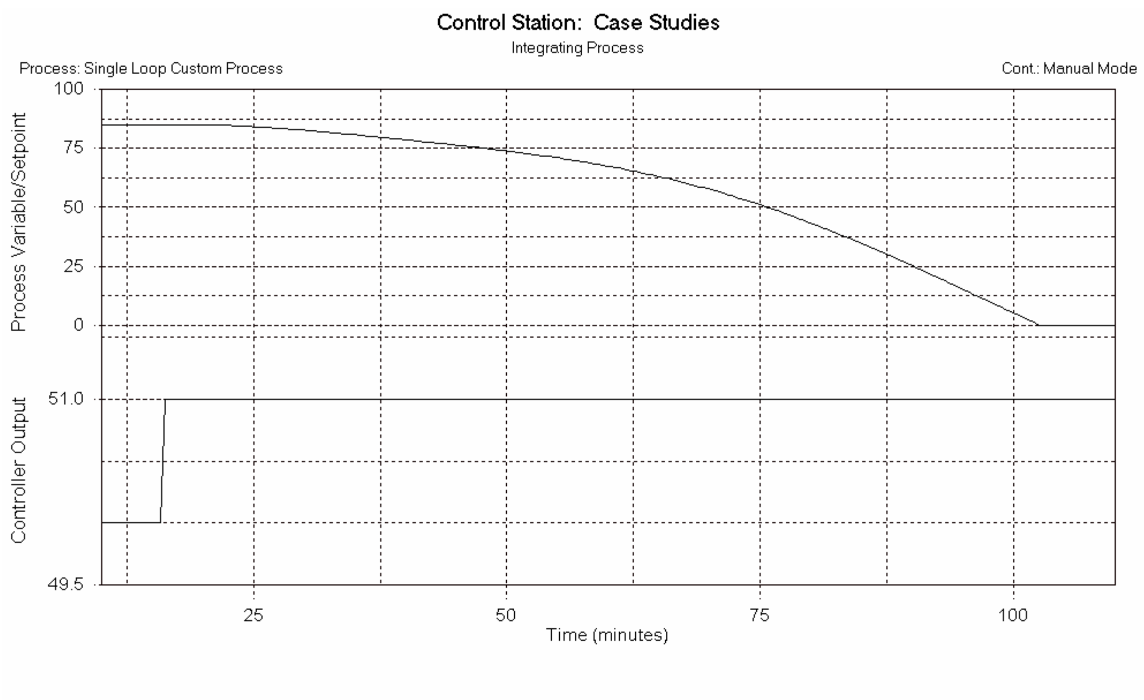


Figure 2-14

Level control is a typical integrating process. When the amount of material flowing into a tank matches the flow going out of the tank the process inputs and outputs are in balance, the tank level will remain constant. If we were to make a step change increasing the flow of material into the tank the process would be out of balance, the tank level would rise until the tank overflowed.

An integrating process, if left in manual, will tend to run away at some point.

Dead time, Time Constants and Gain in an Integrating Process

Since the reaction curve of an integrating process is different from a self-regulating process, is there still a process dead time, time constant and gain? Yes, integrating processes have these characteristics as well.

Dead Time in an Integrating Process

Dead time is still measured in the same way, measure the time lag between the change in controller output and the change in the process variable.

Time Constants in an Integrating Process

Although integrating processes have time constants, they cannot be easily determined from a reaction curve. But as it turns out, time constants are not that important in the tuning of an integrating process. We will find in Chapter Six that the time constant of a self- regulating process will give you a good starting value for the integral term, and integrating processes require no or very little integral effect, it is already in the process.

Gain in an Integrating Process

The definition of gain for an integrating process is different from a self-regulating process as the reaction curve is different. The process gain for an integrating system is the slope of the process variable.

$$\text{Gain} = \frac{\left(\Delta PV / \Delta t\right)}{\Delta CO} = \frac{\left(PV_{Final} - PV_{Initial}\right) / \left(time_{Final} - time_{Initial}\right)}{CO_{Final} - CO_{Initial}}$$

While process gain was unitless for a self-regulating process, for an integrating process the gain has units of 1/time.

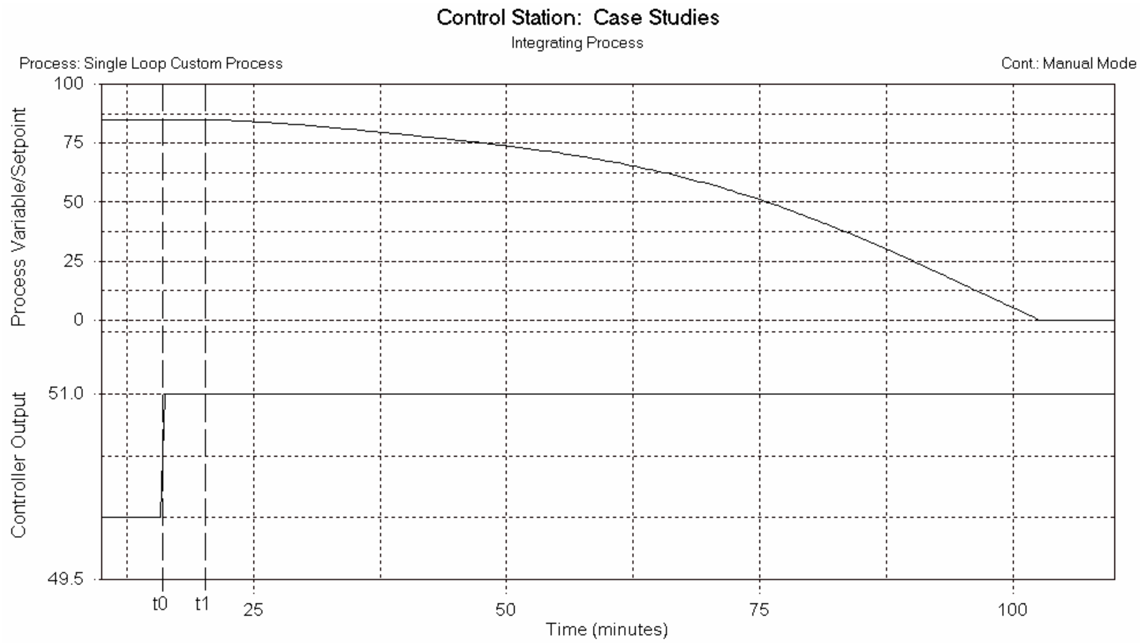


Figure 2-15

Figure 2-15 is the reaction curve of an integrating process in response to a 1% step in the controller output. The points t_0 and t_1 have been marked on the curve. The controller output was changed at t_0 = 16 minutes. The PV did not start to change until t_1 = 20 minutes. The dead time for this process is:

$$\theta = t_1 - t_2 = 20 - 16 = 4 \text{ minutes}$$

From figure 2-15 we see the reaction curve is nonlinear and therefore the gain will have different values depending on where we measure it. We will examine three portions of out reaction curve, from time = 25 to 37.5 minutes, time = 50 to 62.5 minutes and time = 75 to 87.5 minutes.

For t = 25 to 37.5 minutes

$$\Delta t = \left(37.7 - 25\right) \text{minutes } x \frac{60 \text{ seconds}}{\text{minute}} = 750 \text{ seconds}$$

$$\Delta PV = \left(80 - 84\right)\% = -4\%$$

$$\Delta CO = \left(51 - 50\right)\% = 1\%$$

$$Gain = \frac{-4\% \big/ 750 \text{ seconds}}{1\%} = -0.00533 \big/ \text{second}$$

For t = 50 to 62.5 minutes

$$\Delta t = (62.5 - 50)\,\text{minutes } x \frac{60\,\text{seconds}}{\text{minute}} = 750\,\text{seconds}$$

$$\Delta PV = (66 - 74)\% = -8\%$$

$$\Delta CO = (51 - 50)\% = 1\%$$

$$Gain = \frac{-8\% \big/ 750\,\text{seconds}}{1\%} = -0.0101 \big/ \text{second}$$

For t = 75 to 87.5 minutes

$$\Delta t = (87.7 - 75)\,\text{minutes } x \frac{60\,\text{seconds}}{\text{minute}} = 750\,\text{seconds}$$

$$\Delta PV = (30 - 51)\% = -21\%$$

$$\Delta CO = (51 - 50)\% = 1\%$$

$$Gain = \frac{-4\% \big/ 750\,\text{seconds}}{1\%} = -0.0289 \big/ \text{second}$$

Introduction to Instrumentation

Objectives:

In this chapter you will learn:

- What are Sensors and Transducers?
- What are the Standard Instrumentation Signals?
- What are smart Transmitters?
- What is a Low-Pass Filter?
- What Instrument Properties Affect a Process?
- What is Input Aliasing?
- What is Instrument Noise?
- How Do We Measure Temperature?
- How Do We Measure Level?
- How Do We Measure Pressure?
- How Do We Measure Flow?

Introduction

The intent of this chapter is not to teach how to select a particular instrument nor to familiarize the student with all of the available types of instruments. The intent of this chapter is to provide an introduction to some of the more commonly measured process variables in these processes: basic terminology and characteristics relevant to their role in a control loop. Detailed information and assistance on device selection is readily available from most instrumentation suppliers.

Instrumentation Basics

You cannot control what you cannot measure. Sensors are the foundation of Feedback Control. Sensors measure the process variable and transmit a signal that represents the measurement to the Controller. The quality of performance of the system is directly related to the performance of the sensor.

The process variables we measure most often are temperature, pressure, level, flow and mass. No sensor measures a process variable directly. Each sensor measures the effect of the process variable by physical position, force, voltage or some other more easily measured property.

In this chapter we will be examining some common sensors used in Process Control and the effect they have on the performance of the system.

What are Sensors and Transducers?

Sensors

☞ A ***sensor*** is a device that has a characteristic that changes in a predictable way when exposed to the stimulus it was designed to detect.

When making process measurements we are not really measuring the value of the process variable. We are inferring the value of the process variable by measuring the response of a sensor to the process. Sensors have a physical property that changes in a measurable and consistent fashion.

For example, when measuring temperature we are not directly measuring the temperature of an object, we are measuring a sensors change in resistance or the amount of voltage it produces from being exposed to a temperature.

Transducers

☞ A ***transducer*** is a device that converts one form of energy into another.

Transducers are used to convert the output of a sensor into a signal that a controller can use. The output of a sensor may be a mechanical movement, or a change in size or position, or a nonstandard electrical signal. The output of a sensor may even be nonlinear. A transducer will convert the output of the sensor into a standard signal that a controller can use.

A sensor and transducer may be packaged together as shown in Figure 3-1.

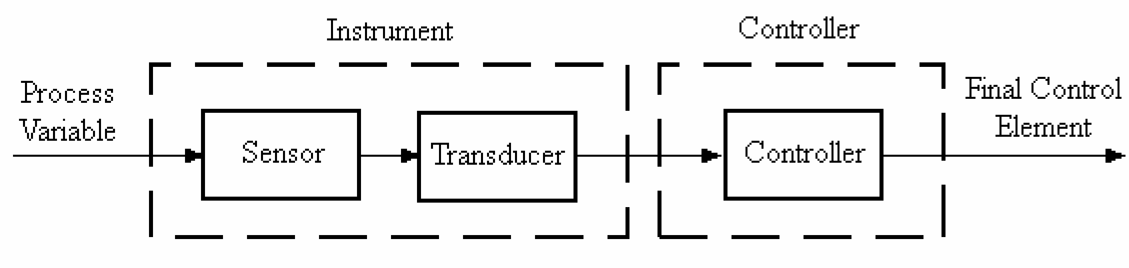


Figure 3-1

Or the transducer may be part of the controller as shown in Figure 3-2.

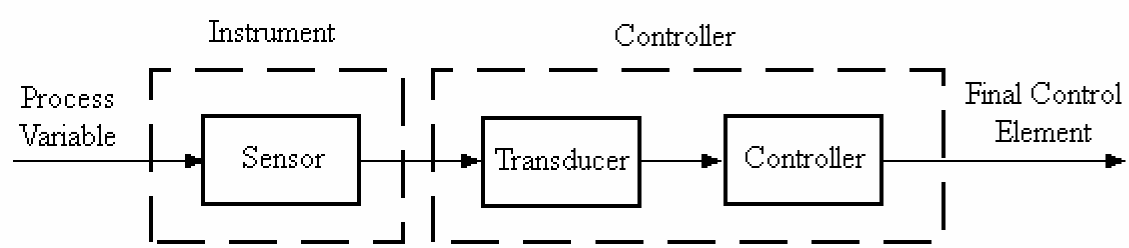


Figure 3-2

☞ For the sake of convenience we will refer to any device that measures a process variable as an instrument, understanding that some signal processing will take place.

What are the Standard Instrumentation Signals

☞ **Standard instrument signals** for controllers to accept as inputs from instrumentation and outputs to final control elements are ***pneumatic***, ***current loop*** and ***0 to 10 volt***.

Pneumatic

Before 1960 pneumatic signals were used almost exclusively to transmit measurement and control information. Today we still commonly find 3 to 15 psig used as the final signal to a modulating valve.

Where the final control element requires a pneumatic signal, in most cases the controller outputs a standard electrical signal and a transducer between the controller and the final control element converts the signal to 3 to 15 psig.

Most often an I/P (I to P) transducer is used. This converts a 4-20 mA signal (I) into a pressure signal (P).

This conversion process is normally linear where the pneumatic signal is given by

$$\text{Signal psig} = \left(\%\,\text{Controller Output } x\, 12 \text{ psig}\right) + 3\text{psig} .$$

If the controller output is 40%, then the pneumatic signal from the I/P transducer is

$$\text{Signal psig} = \left(40\%\; x\, 12 \text{ psig}\right) + 3\text{psig} \;=\; 4.8\text{psig} + 3\text{psig} \;=\; 7.8\text{psig} .$$

Current Loop

4-20 milliamp current loops are the signal workhorses in many processes. A DC milliamp current is transmitted through a pair of wires from a sensor to a controller or from a controller to its final control element. Current loops are used because of their immunity to noise and the distances that the signal can be transmitted. Since the signal being transmitted is current the voltage drop that occurs across conductors does not affect the signal, it just limits the length of the signal cable; and induced voltages do not affect the signal.

Loop Scaling

The level of the current in the loop is related to the value of the process variable or the controller output. How the current level and process value are related is the loop scaling.

Output Scaling

Scale outputs for a one to one correspondence. That is the controller output is configured for 0% to correspond to a 4mA signal and 100% to correspond to a 20mA signal. The final control element is calibrated so that 4mA corresponds to its 0% position or speed and 20mA corresponds to its 100% position or speed.

Input Scaling

Scale inputs for a one to one correspondence as well. If we were using a pressure transducer with a required operating range of 0 psig to 100 psig we would calibrate the instrument such that 0 psig would correspond to 4mA output and 100 psig would correspond to a 20mA output. At the controller we would configure the input such that 4mA would correspond to an internal value of 0 psig and 10mA would correspond to an internal value of 100 psig.

0 - 10 V

0 to 10 volt is not commonly used in many control systems because this signal is susceptible to induced noise and the distance of the instrument or final control element is limited due to voltage drop.

You may find 0-10 volt signals used in control systems providing the speed reference to variable speed drives. For this application the cabling from the controller to the drive is typically confined to a control panel, meaning the cabling distances are short and electrical noise is more easily controlled.

What are Smart Transmitters?

Pneumatic, current loop and 0-10 volt signals are all analog signals. They are capable of transmitting a signal that is continuous over its range, but the signal that is transmitted can only represent a single value and the communication is only one way.

> ☞ A smart transmitter is a digital device that converts the analog information from a sensor into digital information, which allows the device to simultaneously send and receive information and transmit more than a single value.

Smart transmitters, in general, have the following common features:

- Digital Communications
- Configuration
- Re-Ranging
- Signal Conditioning
- Self-Diagnosis

Digital Communications

Smart transmitters are capable of digital communications with both its configuration device and a process controller. Digital communications have the advantage of being free of bit errors, the ability to monitor multiple process values and diagnostic information and the ability to receive commands. Some smart transmitters use a shared channel for analog and digital data (Hart, Honeywell or Modbus over 4-20mA), others use a dedicated communication bus (Profibus, Foundation Fieldbus, DeviceNet, Ethernet).

> ☞ Most smart instruments wired to multi-channel input cards require isolated inputs for the digital communications to work.

Configuration

Smart transmitters can be configured with a handheld terminal and store the configuration settings in non-volatile memory.

Signal Conditioning

Smart transmitters can perform noise filtering and can provide different signal characterizations.

Self-Diagnosis

Smart transmitters also have self-diagnostic capability and can report malfunctions that may indicate erroneous process values.

What Instrument Properties Affect a Process?

Some instrument properties that can affect the performance of your control system include:

- The instrument's range and span.
- The resolution of the measurement.
- The instrument's accuracy and precision.
- The instrument's dynamics.

Range and Span

☞ The ***range*** of a sensor is the lowest and highest values it can measure within its specification.

An RTD may have a specified range of -200°C to +560°C.

A temperature transducer with an RTD sensor may have a specified range of -10°C to +65°C.

☞ The ***span*** of a sensor is the high end of the Range minus the low end of the Range.

The RTD with a range of -200°C to +560°C would have a span of 760°C.

The RTD transducer with a range of -10°C to +65°C would have a span of 75°C.

Match Range to Expected Conditions

Instruments should be selected with a range that includes all values a process will normally encounter, including expected disturbances and possible failures.

In Chapter Two we learned that an instrument with too wide of an operating span will lower the process gain, possibly hiding an oversized final control element and requiring higher controller gains. Also, for most instruments operating ranges larger than required reduce the accuracy of the measurement.

In the heat exchanger example we could reasonably expect that the water temperature would always be between 0°C and 100°C (frozen water will not flow and the system is designed so that we cannot superheat the water). Matching a temperature transmitter closely to this range will give us the best measurement performance.

Measurement Resolution

> ☞ ***Resolution*** is the smallest amount of input signal change that the instrument can detect reliably.

Resolution is really a function of the instrument span and the controller's input capability. Most controllers convert their analog input signals into a digital equivalent. The resolution of the measurement is determined by the span of the measurement and the number of bits in the digital conversion process.

Most current controllers resolve their analog inputs into 16 bits of digital information. 16 bits of information allows for 65,535 values with which to represent the input signal.

The resolution of a 16 bit conversion is $\frac{\text{Input Span}}{65{,}535}$.

If we have a 4-20 mA signal that represents a 0 to 100 psig process value, the input resolution is $\frac{100\text{ psig}}{65{,}535} = 0.0015\text{ psig}$.

Some older control systems may have input processing that only resolves their inputs into 12 bits of digital information. 12 bits will only allow for 4095 values with which to represent the process signal which is 16 times less resolution than a 16 it conversion.

The resolution of a 12 bit conversion is $\frac{\text{Input Span}}{4{,}095}$.

The same 0 to 100 psig process with a 12 bit conversion has an input resolution of $\frac{100\text{ psig}}{4{,}095} = 0.0244\text{ psig}$.

No matter how many bits are used for the conversion process, the larger your instrument span, the lower the resolution of your measurement.

Accuracy and Precision

☞ ***Accuracy*** of a measurement describes how close the measurement approaches the true value of the process variable.

Accuracy is often expressed as a % error over a range or an absolute error over a range.

% Error Over a Range

Accuracy may be specified as ± a percentage over a range, span or full scale of an instrument. The uncertainty in your measurement will be the percentage times the range specified by the manufacturer.

For example, Manufacturer A specifies that their pressure instrument has an accuracy of ±0.4% of full scale. The full scale of their instrument is 500 psig. We can expect the measurement signal from this instrument to be accurate to 2 psig for all pressures in the instruments range.

$$0.4\% \; x \; 500 \text{ psig} = 0.004 \; x \; 500 \text{ psig} = 2 \text{ psig}$$

Even if we span the instrument for the application to 0 to 100 psig the instrument is still only accurate to 2 psig.

Absolute Over a Range

Accuracy may also be specified as ± an absolute value over a range.

Manufacturer B specifies that their pressure instrument has an accuracy of ±1 psig over the full operating range. The full scale of their instrument is also 500 psig. We can expect the measurement signal from this instrument to be accurate to 1 psig for all pressures in the instruments range.

In any case, the sensors we select must be more accurate then the degree to which we want to control the process. There is no amount of control loop tuning you can do on a process to maintain 30 psig ± 1 psig pressure instrument is only accurate to ± 2 psig.

> ☞ ***Precision*** is the reproducibility with which repeated measurements can be made under identical conditions.

Precision may also be referred to as stability or drift. Precision is always required for good control, even when accuracy is not required.

Accuracy vs. Precision

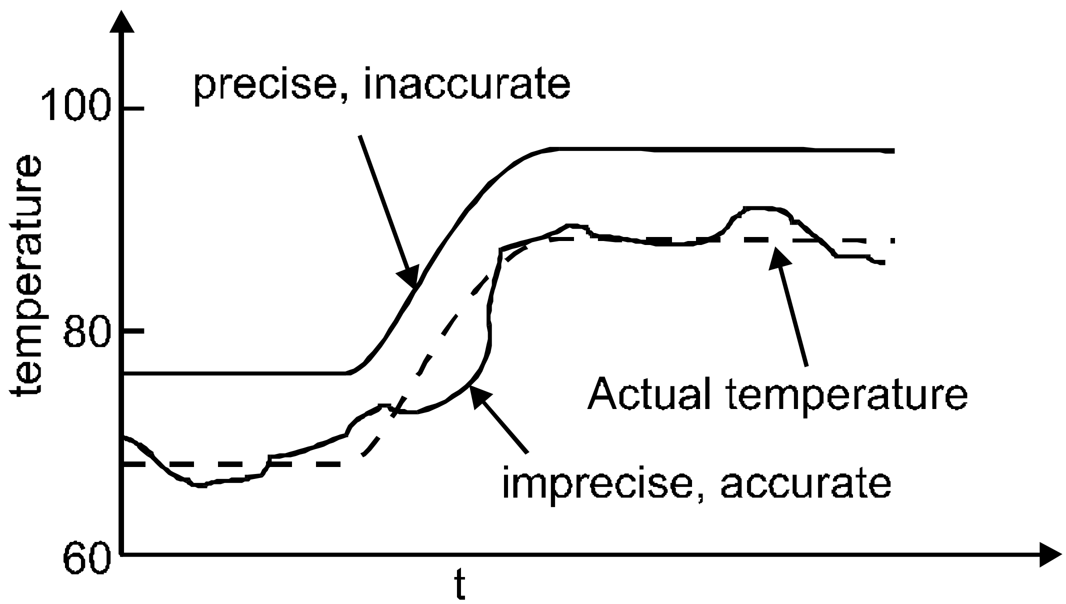


Figure 3-3

The distinction between accuracy and precision is illustrated in Figure 3-3. The dashed line represents the actual temperature being measured. The upper line represents a precise but inaccurate value from an instrument; the lower line represents an accurate but imprecise measurement from an instrument.

Precision is the more important characteristic of an instrument.

Instrumentation Dynamics

Instruments have dynamic properties just as process do. The dynamic properties instruments posses are identical to process dynamics: gain, time constants and dead time. All of these instrument dynamics contribute to the process dynamics that the controller sees.

Instrument Gain

☞ The gain of an instrument is often call ***sensitivity***. The sensitivity of a sensor is the ratio of the output signal to the change in process variable.

For a thermocouple the typical sensitivity is 5 mV per °C. This means for every °C change in the process variable the thermocouple will change its output by 5 mV.

Instrument Time Constants

As for processes, one time constant for an instrument is the time it takes to provide a signal that represents 63.2% of the value of variable it is measuring after a step change in the variable. Instrument manufacturers may sometimes specify the rise time instead of the time constant.

☞ Rise time is the time it takes for an instrument to provide a signal that represents 100% of the value of the variable it is measuring after a step change in the variable. The rise time of an instrument is equal to 5 time constants.

Instrument Dead Time

☞ The ***dead time*** of an instrument is the time it takes for an instrument to start reacting to process change.

What is Input Aliasing?

Input aliasing is a phenomenon that occurs from digital processing of a signal. When a signal is processed digitally it is sampled at discrete intervals of time. If the frequency at which a signal is sampled is not fast enough the digital representation of that signal will not be correct. Input aliasing is an important consideration in digital process control. Processor inputs that have configurable sample rates and PID loop update times must be set correctly.

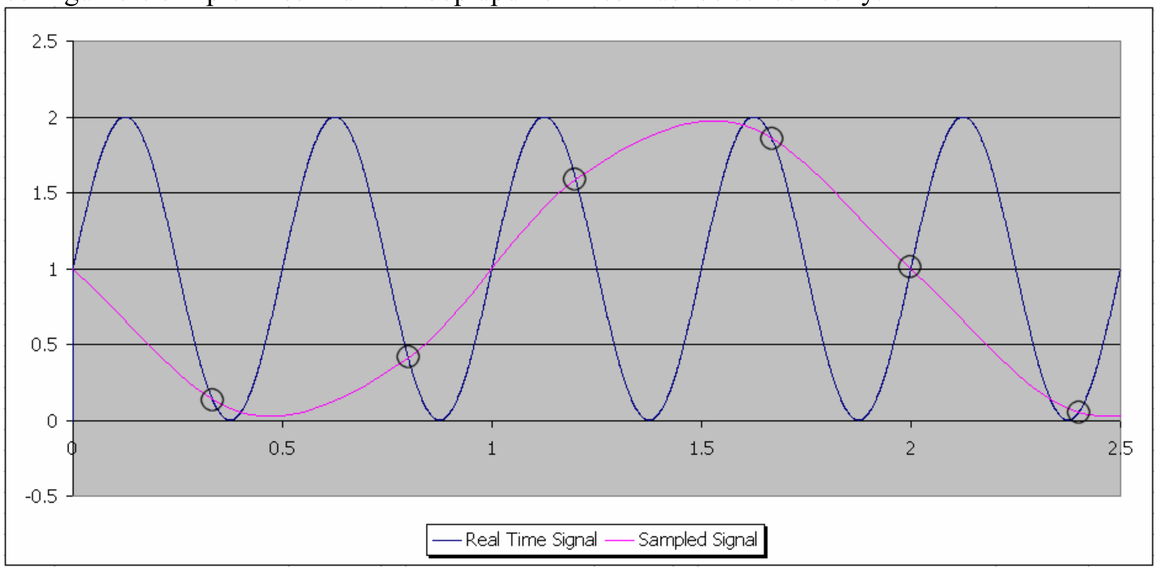


Figure 3-4

Figure 3-4 shows a low frequency 2 Hz signal that was sampled every 0.4 seconds. The resulting curve from the sampled points looks nothing like the original. If this was the signal for the process variable then we would be unable to achieve control.

Correct Sampling Frequency

Fortunately establishing the correct sampling rate for a signal is not a trial and error procedure. There is a well established theorem (Nyquist Frequency Theorem) that tells us to correctly sample a waveform it is necessary to sample at least twice as fast has the highest frequency in the waveform. In the digital world this means sampling at 1/20th of the period of the waveform. The 2 Hz signal in Figure 3-4 has a period of 1/2Hz = 0.5 seconds per cycle. To sample this in the digital world the sampling interval would be 0.5/20 seconds = 0.025 seconds. Figure 3-5 shows the same waveform sampled at this interval. Notice the match between the real and sampled signal.

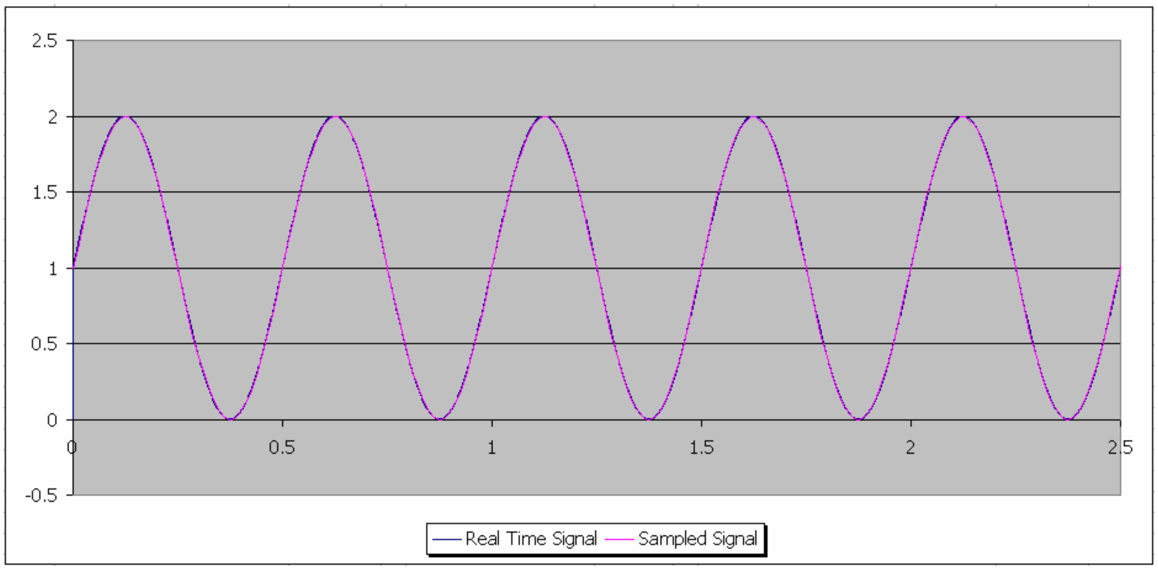


Figure 3-5

Determining the Correct Sampling Interval

While it's nice to know there is guidance on how to set the sample interval for a waveform based on its frequency, how does one know what the frequency of a process variable is?

When it comes to instrumentation, it's not the frequency that's important, it's the time constant. Figure 3-6 is a graph of the response of an instrument with a 5 second time constant (25 second rise time). The signal from the instrument was sampled at 1 second intervals.

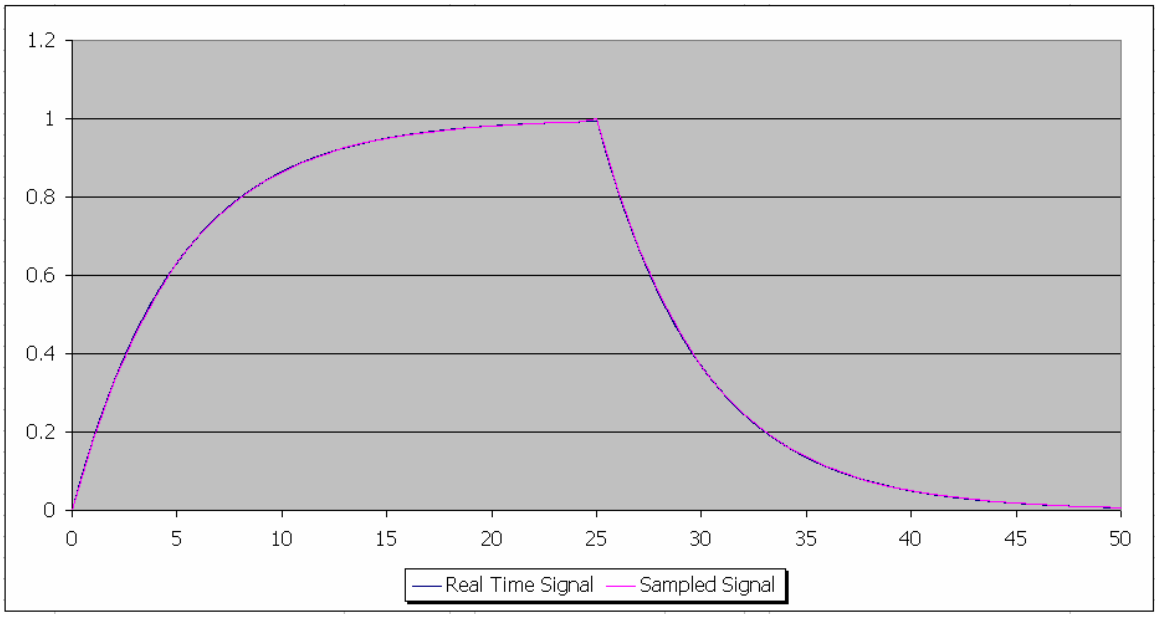


Figure 3-6

☞ One rule of thumb would be to set the sample interval for an instrument at 1/10th to 1/20th of the rise time (1/2 to 1/4th of the time constant).

Another rule of thumb would be to set the sample interval to 1/10th to 1/20th of the process time constant.

Temperature instrumentation (RTDs and thermocouples in thermowells) typically have time constants of several seconds or more. For these processes sampling intervals of 1 second are usually sufficient.
Pressure and flow instrumentation typically have time constants of ½ to 1 second. For these processes sampling intervals of 0.1 second are usually sufficient.

What is Instrument Noise?

> ☞ ***Noise*** is a variation in a measurement of a process variable that does not reflect real changes in the process variable.

A signal from a sensor can have many components. This signal will always have as one of its components the process value that we are measuring, but it may also contain noise. Noise is generally a result of the technology used to sense the process variable. Electrical signals used to transmit instrument measurements are susceptible to having noise induced form other electrical devices. Noise can also be caused by wear and tear on mechanical elements of a sensor.

Noise may also be uncontrolled random variations in the process itself. Whatever the source, noise distorts the measurement signal.

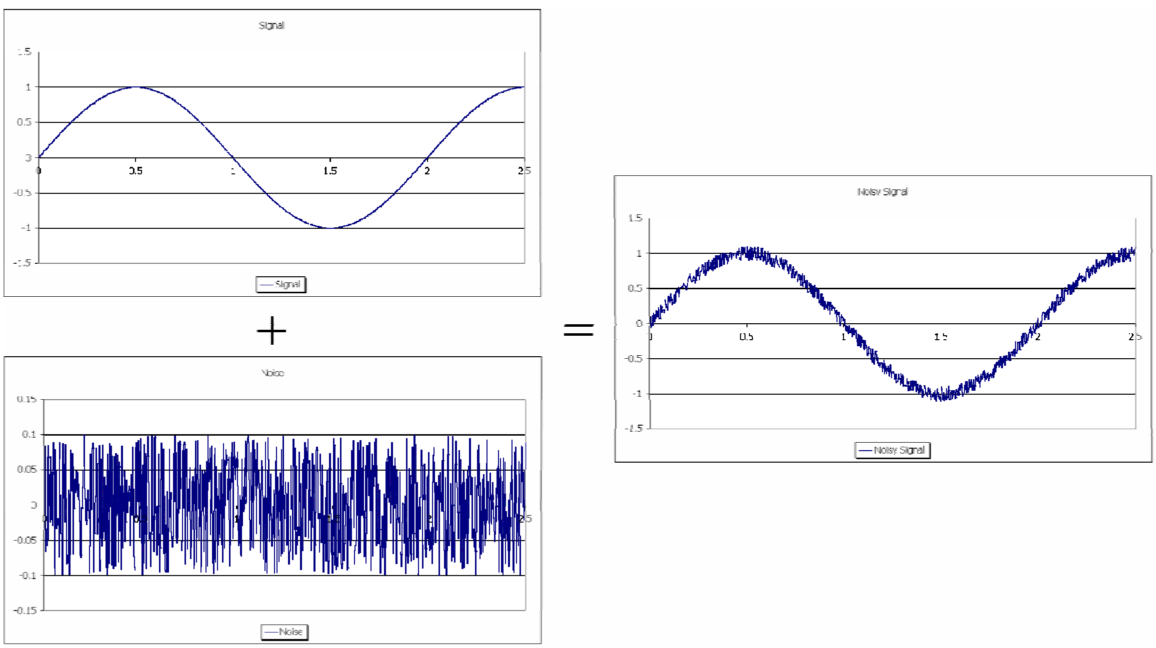


Figure 3-7

Effects of Noise

Noise reduces the accuracy and precision of process measurements. Somewhere in the noise is the true measurement, but where? Noise introduces more uncertainty into the measurement.

Noise also introduces errors in control systems. To a controller fluctuation in the process variable from noise are indistinguishable from fluctuations caused by real disturbances. Noise in a process variable will be reflected in the output of the controller.

Eliminating Noise

The most effective means of eliminating noise is to remove the source. Reduce electrically induced noise by following proper grounding techniques; using shielded cabling and physical separation of signal cabling form other electrical wiring. If worn mechanical elements in the sensor are causing noise repair or replace the sensor.

When these steps have been taken and excessive noise is still a problem in the process variable a low pass filter may be used.

Low Pass Filters

Smart instruments and most controllers have noise dampening features built in. Most of these noise dampeners are actually low pass filters.

> ☞ A ***low-pass filter*** allows the low frequency components of a signal to pass while attenuating the higher frequency components.

Fortunately for us, noise tends to fall into the higher end of the frequency spectrum while the underlying process value tends to lie in the lower end.

Selecting a Filter by Cut-off Frequency

Attenuation of a signal is a reduction in its strength, or amplitude. Attenuation is measured in decibels (dB).

$$\text{dB of attenuation} = 20\log_{10}\left(\frac{\text{Amplitude Out}}{\text{Amplitude In}}\right)$$

For example: let's say we have an amplitude ratio of 0.95 (the value of the signal out is 95% of value of the signal in), the dB of attenuation would be:

$$20\log_{10}(.95) = -0.45$$

An attenuation of 0 dB would mean the signal would pass with no reduction in amplitude while a large negative dB would indicate a very small amplitude ratio (at -10 dB of attenuation we would have an amplitude ratio of 0.32).

The ideal filter would be designed to pass all signals with 0 dB of attenuation below a cut-off frequency and completely attenuate all frequency components above the cut-off frequency. This ideal filter does not exist in the real world. -3 dB of signal attenuation has been established as the cut-off frequency in filter selection. Figure 3-8 illustrates the effect of a filter with a 3 Hz cut-off frequency on a noisy 1.2 Hz signal. Where a filter is selected by choosing cut-off frequency, select that is above the frequency of your process value.

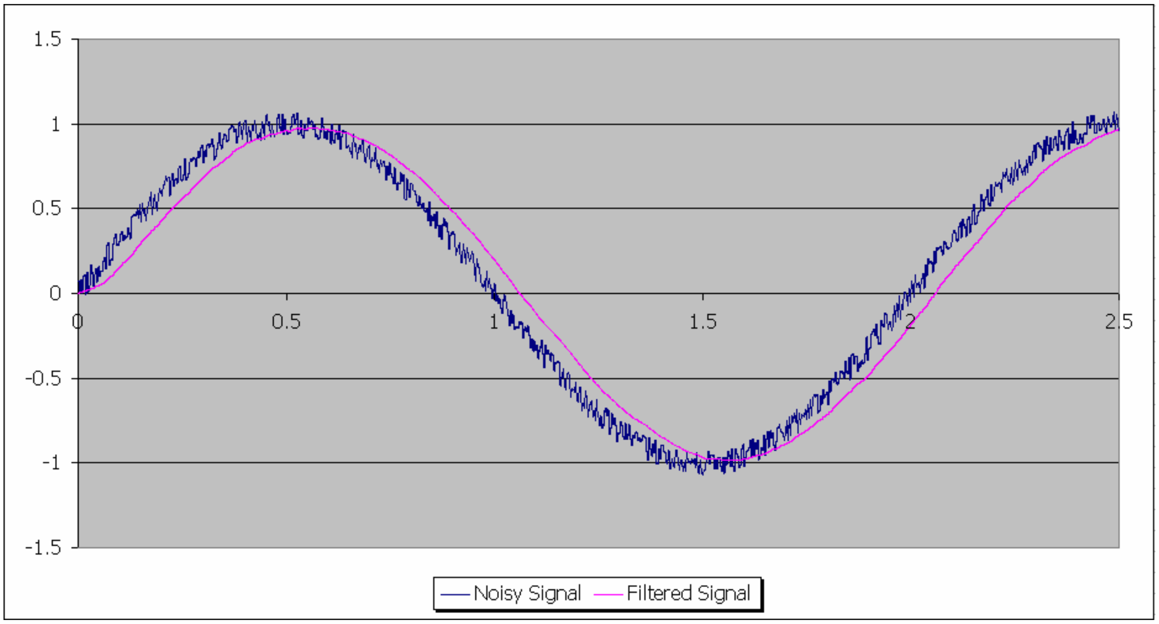


Figure 3-8

Selecting a Filter by Time Constant

The effect of a low pass filter is to introduce a first order lag in the process variable response.

> ☞ Low pass filters may sometimes be referred to as first order lag filters.

Some filters are configured by selecting a time constant for the lag response of the filter. The relationship between the cut-off frequency and the time constant of a low pass is approximately given by:

$$\text{Cut - Off Frequency} \approx \frac{1}{5\ \text{Time Constants}}.$$

For example, to configure a filter for a cut-off frequency of 60 Hz specify a filter time constant of

$$\text{Time Constant} \approx \frac{1}{5\ x\ \text{Cut - Off Frequency}} = \frac{1}{5\ x\ 60} = 0.0033\ \text{seconds}.$$

Figure 3-9 shows the step response of the 3 Hz filter illustrating the filter time constant of 0.068 seconds.

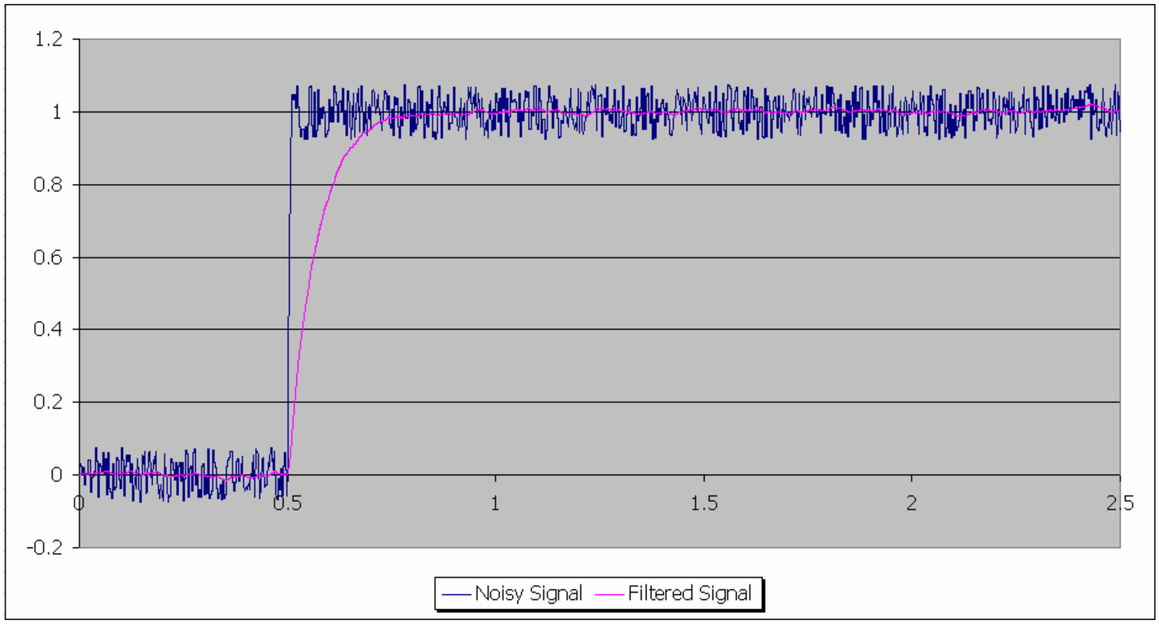


Figure 3-9

Selecting a Filter by Alpha Value

Some filters are selected by a value called alpha (α), notably the derivative filter in a PID controller. α is an averaging weighting term used in controller calculations to impart a first order lag on the measured variable. It is just another way of specifying a low pass filter.
α generally has a value between 0 and 1. A filter with a α = 0 would pass the signal through unfiltered. A filter with a α = 1 would filter everything; nothing would pass through the filter.

☞ When a filter is specified by cut-off frequency, the lower the frequency the greater the filtering effect.

☞ When a filter is specified by time constant, the greater the time constant the greater the filtering effect.

☞ When a filter is specified by α, when α=0 no filtering is done, when α = 1 no signal passes through the filter.

What is Temperature?

☞ Temperature is a measure of degree of hotness or coldness of an object.

The concept of temperature may be more easily expressed by heat will not flow between two objects at the same temperature in contact with one another. Heat energy will always flow from a higher temperature object to a lower temperature object. So temperature is a relative term.

Units of Temperature

There is several temperature scales in use today. Temperature scales are defined by assigning temperatures to two points of reproducible physical phenomena. The two most common temperature scales are Fahrenheit (°F) and Celsius (°C).

The reference points for Fahrenheit are the freezing point and boiling point of water. Water freezes at 32°F and boils at 212°F. The Celsius scale uses the same reference points only it defines the freezing point of water as 100°C and the boiling point as 100°C.

We see from these definitions that a Fahrenheit degree is 9/5 of a Celsius degree and is offset by 32°. To convert between these scales we use the following formulas:

$$°F = \left(\frac{9}{5} x \, °C \right) + 32$$

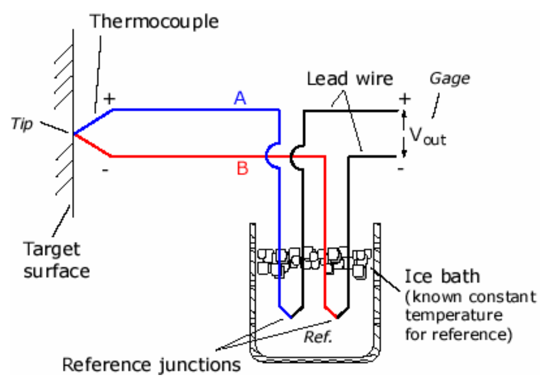
$$°C = \frac{5}{9} x \left(°F - 32 \right)$$

What Temperature Instruments Do We Use?

☞ Temperature is most commonly measured by Resistance Temperature Devices (RTD) and Thermocouples, and to a lesser degree Infrared (IR).

Temperature is a very common process variable and temperature control is critical to maintaining product quality and ensuring safe and reliable operation of many processes.

Thermocouples



Thermocouples are fabricated from two electrical conductors made of two different metal alloys. At one end of the cable the two conductors are electrically shorted together by crimping, welding, etc. This end of the thermocouple, the hot or sensing junction, is thermally attached to the object to be measured. The other end, the cold or reference junction is connected to a measurement system. Thermocouples generate an open-circuit voltage, called the Seebeck voltage that is proportional to the temperature difference between the sensing (hot) and reference (cold) junctions:

$$V_S = V\left(T_{Hot} - T_{Ref}\right)$$

$$\text{Where : } V_S = \text{Seebeck Voltage}$$

$$V = \text{Proportionality Constant}$$

$$T_{Hot} = \text{Temperature of the Sensing Junction}$$

$$T_{Ref} = \text{Temperature of the Reference Junction}$$

The relationship between the temperature difference and the voltage produced has been documented for several commonly used combinations of metals.

Junctions

Since thermocouple voltage is a function of the temperature difference between junctions, it is necessary to know both voltage and reference junction temperature in order to determine the temperature at the hot junction. Consequently, a thermocouple measurement system must either measure the reference junction temperature or control it to maintain it at a fixed, known temperature. Most controllers today measure the temperature of the reference junction.

Junction Misconceptions

There is a misconception of how thermocouples operate. The misconception is that the hot junction is the source of the output voltage. This is wrong. The voltage is generated across the length of the wire; if the entire wire length is at the same temperature no voltage would be generated. If this were not true we could connect a resistive load to a uniformly heated thermocouple inside an oven and use additional heat from the resistor to make a perpetual motion machine.

Another misconception is that junction voltages are generated at the cold end between the special thermocouple wire and the copper circuit; hence, a cold junction temperature measurement is required. This concept is wrong. The cold end temperature is the reference point for measuring the temperature difference across the length of the thermocouple circuit.

Lead Wires

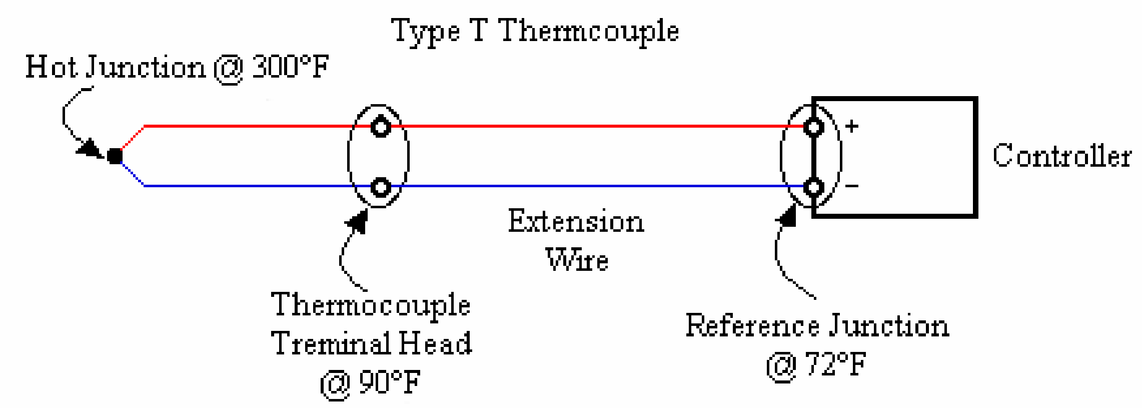


Figure 3-10

Figure 3-10 illustrates a typical thermocouple installation. In this example we have a hot junction at 300°F, a terminal head at 90°F and a reference junction at the controller at 72°F. The thermocouple will generate a voltage proportional to the temperature difference between the hot junction and the terminal head, the voltage will be proportional to 300 – 90°F, or 210°F.

The extension wire will generate a voltage proportional to the difference between its junctions, a voltage that is proportional to 90 – 72°F or 18°F. The generated by the temperature differential in the extension wires will be added to the voltage from the thermocouple so that the voltage at the reference junction is proportional to a temperature differential if 210 + 18 °F, or 228°F. The controller will process this voltage correctly interpreting a 228°F temperature differential, add in its reference junction temperature and display a process value of 300°F.

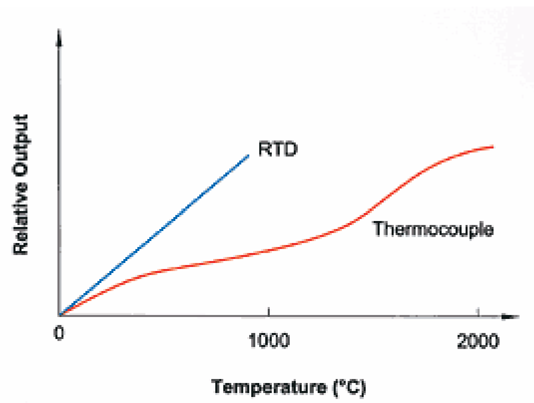
If the extension wire is replaced with copper wire a voltage that is proportional to the temperature differential between the terminal head and the controller will not be added to the thermocouple signal. The controller will display a process value of 282°F.

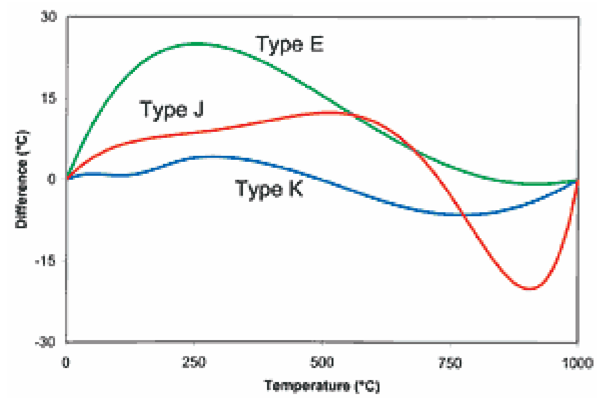
Use of the correct extension wire is critical in thermocouple applications; otherwise the temperature differential across the extension leads will be introduced as measurement error.

Does this mean you cannot use copper terminal blocks for thermocouples? No, as it is highly unlikely there will be a temperature differential across the terminal block it will introduce no errors.

Linearization

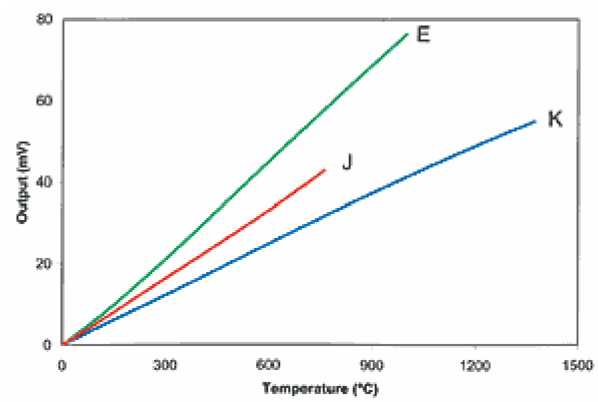
Within the useable temperature range of any thermocouple, there is a proportional relationship between thermocouple voltage and temperature. However, this relationship is extremely nonlinear over the operating range of a thermocouple.





Nor is the degree on nonlinearity the same for all thermocouple types.

Fortunately, the electrical properties of thermocouples are well defined. The non-linear signals generated by thermocouples are converted into process values by software lookup tables or polynomial conversion within the controller or a transmitter.



Gain

Not all thermocouples have the same gain (sensitivity) either. Higher gain is better for noise immunity. Of types J, K and E, type K has the lowest gain but is still most popular because of its linearity.

Thermocouple Types

About 13 "standard" thermocouple types are commonly used. Eight have been given an internationally recognized letter type designator. The letter type refers to defined voltage characteristics, not the composition of the metals, so any thermocouple that matches the defined voltage characteristics may receive the letter type designator.

The four most common types are J, K, T and E. There are high temperature thermocouples R and S. Each type has a different sensitivity (gain), linearity, temperature range and environment, although the maximum temperature varies with the diameter of the wire used in the thermocouple.

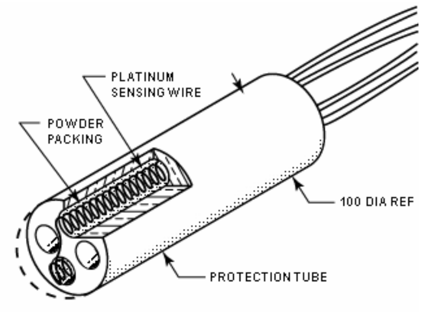
While the thermocouple type dictates the temperature range, the maximum range is also limited by the diameter of the thermocouple wire. A very thin thermocouple may not reach the full temperature range. Six of the most common types are listed in Table 3-1.

T/C Type	*Metals*	*Color Code*	*Application Limits*	*Accuracy*
J	Iron (+) Constantan (-)	White Red	32 to 1382°F TC Grade 32 to 392°F Ex Grade	Greater of: 3.962°F or 0.75%
K	Chromel (+) Nickel (-)	Yellow Red	-328 to 2282°F TC Grade 32 to 392°F Ex Grade	Greater of: 3.962°F or 0.75%
T	Copper (+) Constantan (-)	Blue Red	-328 to 662°F TC Grade -76 to 212°F Ex Grade	Greater of: 1.8°F or 0.75%
E	Chromel (+) Constantan (-)	Purple Red	-328 to 1652°F TC Grade 32 to 392°F Ex Grade	Greater of: 3.06°F or 0.5%
R	Pt -13% RH (+) Platinum (-)	None Established	32 to 2642°F TC Grade 32 to 300°F Ex Grade	Greater of: 2.7°F or 0.25%
S	Pt – 10% Rh (+) Platinum (-)	None Established	0 to 2642°F TC Grade 32 to 300°F Ex Grade	Greater of: 2.7°F or 0.25%

Table 3-1

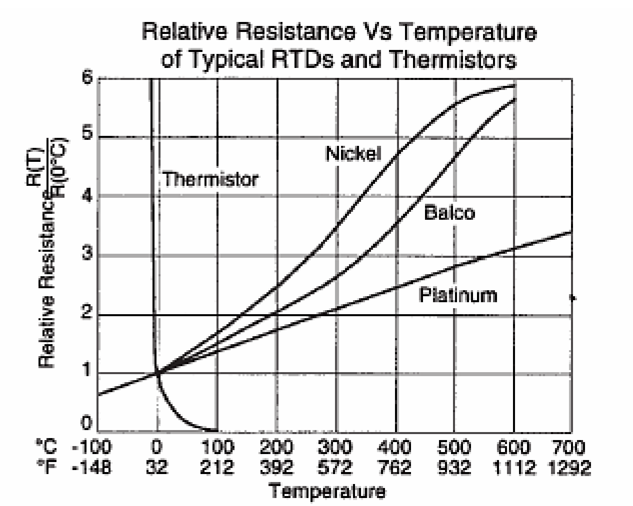
Resistive Temperature Devices

A resistance-temperature detector (RTD) is a temperature sensing device whose resistance increases with temperature. An RTD consists of a wire coil or deposited film of pure metal whose resistance at various temperatures has been documented.



RTDs are used when applications require accuracy, long-term stability, linearity and repeatability. RTDs can work in a wide temperature range; some platinum sensors handle temperatures from 165°C to 650°C.

Common materials for RTDs include platinum, nickel, copper, balco and tungsten.



Platinum is the most popular because it has a very stable and nearly linear resistance vs. temperature function, especially when compared to other metals and temperature sensors.

A 100Ω platinum RTD has a nominal resistance of 100 ohms at 0°C, a 1000Ω platinum RTD has a nominal resistance of 1000 ohms at 0°C.

RTDs are known for their excellent accuracy over a wide temperature range. Some RTDs have accuracies as high as 0.01 ohm (0.026°C) at 0°C. RTDs are also extremely stable devices. Common industrial RTDs drift less than 0.1° C/year, and some models are stable to within 0.0025° C/year.

The Importance of the Temperature Coefficient alpha

Platinum is the most popular RTD, in part, because of its nearly linear change in resistance with respect to temperature. However, it is not exactly linear and the change in resistance is more accurately modeled by polynomial called the Callendar-Van Dusen Equation.

$$R_t = R_0\left[1 + At + Bt^2 - 100Ct^3 + Ct^4\right].$$

While accurate, the Callendar-Van Dusen equation is neither convenient nor meaningful to the average person. After some manipulation (and introduction of some new terms) the temperature coefficient alpha (α can be derived from the Callendar-Van Dusen equation. Although various manufacturers may specify alpha differently, alpha is most commonly defined as the resistance at 100°C minus the resistance at 0°C; all divided by 100 times the resistance at 0°C.

$$\alpha = \frac{R_{100} - R_0}{100\ x\ R_0}$$

For example, a 100Ω platinum RTD with α=0.003911 will have a resistance of 139.11 ohms at 100°C. A 100Ω platinum RTD with α=0.003850 will have a resistance of 138.50 ohms at 100°C.

Common values for the temperature coefficient α are:

- 0.003850, DIN 43760 Standard (also referred to as European curve)
- 0.003911, American Standard
- 0.003926, ITS-90 Standard

While RTDs are nearly linear, they are not exactly linear. Most transmitters and RTD input cards have internal algorithms or lookup tables to convert the measured resistance into a process value.

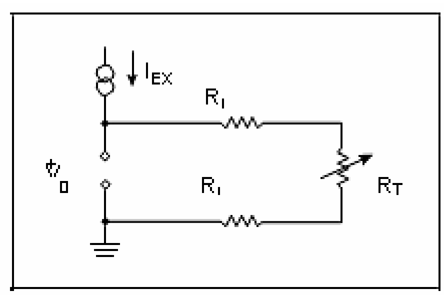
The accuracy of the conversion is dependent upon matching the Temperature Coefficient of the transmitter/input device to the connected RTD. Connecting a 0.003911 RTD to a 0.003850 input can introduce errors of 0.5°C.

Lead Wire Resistance

Because the RTD is a resistive device, you must drive a current through the device and monitor the resulting voltage. However, any resistance in the lead wires that connect your measurement system to the RTD will add error to your readings.

2 Wire RTDs

In a two wire RTD, the measurement system supplies a constant current source, I_{EX}, to drive the RTD. The voltage drops across the lead wires is indistinguishable from the voltage drop across the sensor.

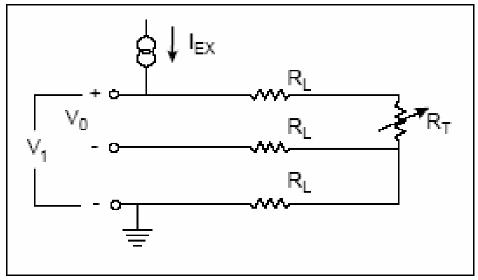


If we had a lead resistance of 0.3 Ω in each wire, R_l, a 0.6 Ω error would be added to the resistance measurement (this is approximately 40 feet of 18 gauge 2 conductor cable).

For a platinum RTD with α= 0.00385, the resistance equals 0.6 Ω/(0.385 Ω/°C) = 1.6°C of error.

3 Wire RTDs

Three wire RTDs are commonly used to remove the lead wire resistance error.
With three leads two voltage measurements can be made. When the voltages are subtracted the result is the voltage drop that would have occurred through R_T alone.



Self Heating

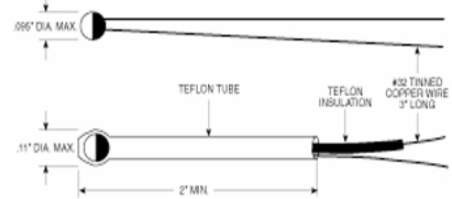
Because an RTD is a passive resistive device, you must pass a current through the device to produce a measurable voltage. This current causes the RTD to internally heat, which appears as an error. Self heating is typically specified as the amount of power that will raise the RTD temperature by 1° C, or 1 mW/ °C.

Self heating can be minimized by using the smallest possible excitation current, but this occurs at the expense of lowering the measurable voltages and making the signal more susceptible to noise form induced voltages.

The amount of self heating also depends heavily on the medium in which the RTD is immersed. An RTD can self heat up to 100 times higher in still air than in moving water.

The self heating error can also be calculated and corrected in the measuring algorithm.

Thermistors



A thermistor is similar to an RTD in that it is a passive resistance device; however thermistors are generally made of semiconductor materials giving them much different characteristics.

Thermistors usually have negative temperature coefficient, their resistance decreases as the temperature increases. Thermistors also exhibit a large change in resistance proportional to a small change in temperature (large gain or high sensitivity) that can be as high as several percent per degree Celsius.

The increased sensitivity of a thermistor comes at the expense of linearity. Thermistors are extremely non-linear instruments, and the degree on non-linearity is highly dependent on the manufacturing process. Owing to this thermistors do not have standardized electrical properties like thermocouples and RTDs.

Thermistors are accurate temperature sensing devices capable of measuring to ± 0.1°C. However the useful sensing range of a thermistor is fairly limited compared to RTDs and thermocouples. The typical sensing range is - 80°C to 250°C.

Infrared

Objects radiate electromagnetic energy. The higher the temperature of an object the more electromagnetic radiation it emits. This radiation occurs within the infrared portion of the electromagnetic spectrum (see Figure 3-11).

Infrared temperature sensors work by measuring the amount of infrared energy emitted by an object, allowing us to measure the temperature of an object without being in contact.

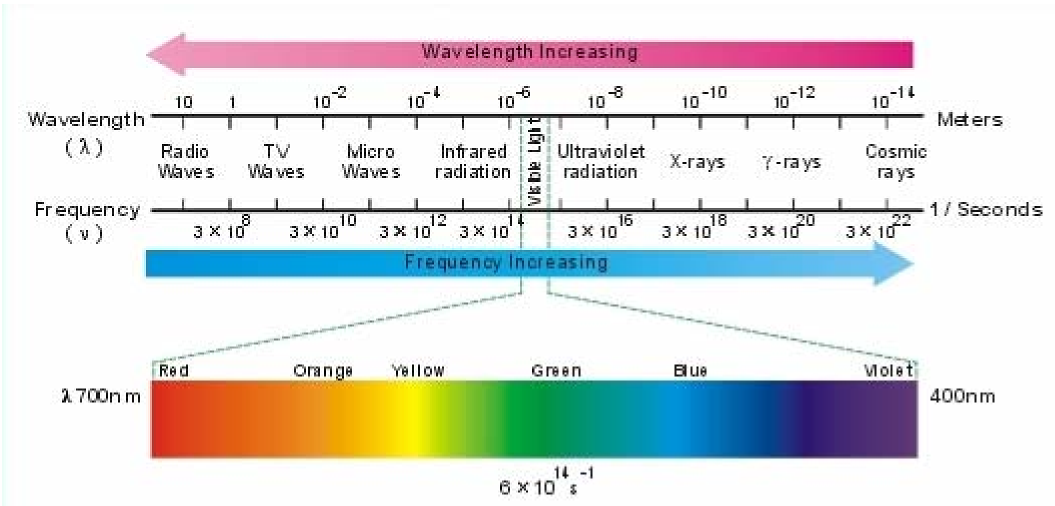


Figure 3-11

Emittance

The principles of infrared temperature measurement are grounded in quantum mechanics, which gives rise to the concept of emittance. Within quantum mechanics we have the concept of a black body. A black body is a surface that neither reflects nor transmits, but absorbs all incident radiation independent of wavelength and direction. In addition to being a perfect absorber of all radiation, a black body is a perfect radiator as well. This ideal black body is the basis of infrared temperature measurement and gives rise to the emittance setting (although different in meaning, emissivity is often used interchangeably in discussing infrared temperature measurement).

The emittance of a real surface is the ratio of the thermal radiation emitted by the surface at a given temperature to that of the ideal black body at the same temperature. By definition a black body has an emittance of 1. Another way to think of emissivity is the efficiency at which an object radiates thermal energy. Emittance is a decimal number that ranges between 0 and 1 or a percentage between 0% and 100%.

To function properly an infrared temperature instrument must take into account the emittance of the surface being measured. Emittance values can often be found in reference tables, but such table will not take into account local conditions of the surface.

A more practical way to set the emittance is to measure the temperature with an RTD or thermocouple and set the instrument emissivity control so that both readings are the same.

As knowing emittance of an object is fundamental to obtaining accurate temperature readings, instrument manufacturers have attempted to design sensors that would provide accurate measurements independently of this value. The best know and most commonly used design is the Two-Color-Ratio Thermometer.

Polished surfaces can also pose a problem. Emittance is often an inverse indicator of the reflective properties of an object. Objects with high emittance values are typically not good reflectors; they will not reflect that much incident energy. Conversely, objects with low emittance values are typically good reflectors. For example, machined stainless steel has an emittance value of 0.14, temperature readings from an infrared instrument could be influenced from reflected radiation from nearby hotter object reaching the sensor. Objects with emittance values below 0.2 can be difficult to apply infrared temperature sensing technology.

Field of View

Infrared temperature instruments are like an optical system in that they have a field of view. The field of view basically defines the target size at a given distance. Field of view may be specified as

- Angle and focal range (2.3° from 8" to 14')
- Distance to spot size ratio and focal range (25:1 from 8" to 14')
- Spot size at a distance (0.32" diameter spot at 8")

All of these specifications are equivalent, at 8 inches the field of view will be a 0.32" diameter spot (8"/25), at 14 feet the field of view will be 6.72" (14' / 25).

An infrared temperature sensor measures the average temperature of everything in its field of view. If the surface whose temperature we are measuring does not completely fill the field of view we will get inaccurate results.

Spectral Response

The infrared portion of the electromagnetic spectrum covers a range of frequencies. The range of frequencies that an infrared sensor can measure is called its spectral response. Infrared temperature sensors do not measure the entire infrared region; different frequencies (wavelengths) are selected for detection to provide certain advantages. If you wanted to use an infrared temperature sensor to measure the temperature of an object behind glass you would need to select an instrument with a spectral response in the infrared region in which glass is "transparent", that is the infrared radiation emitted by the glass is outside the instruments spectral response. Such an instrument would be limited to high temperature use though because of the selected spectral response.

What is Pressure?

Pressure is found anywhere you have contained liquids and gasses. Pressure is the ratio between the force acting on a surface and the area of that surface. Pressure is measured in units of force divided by area. Pressure is controlled to provide safe operations and pressure often influences key process operations such as heat transfer, fluid flow and vapor-liquid equilibrium.

Units of Pressure

As pressure is force acting over an area, its units are of force per area. Some of the common units for pressure that you will find in process systems are:

- psi (pounds per square inch)
- bar, 1 bar = 14.5 PSI, common for pump ratings
- inH20 (inches of water), 27.680 inH20 = 1 psi, common for vacuum systems and tank levels
- mmHg (millimeters of mercury), 760 mmHg = 14.7 psi, common for vacuum systems

Absolute, Gauge and Differential Pressure

Depending on the reference pressure used, a pressure instrument can measure absolute, gauge or differential pressure.

> ☞ Gauge pressure is defined relative to atmospheric conditions, that is, it does not measure the current atmospheric pressure. The units of gauge pressure are psig; however gauge pressure is often denoted by psi as well.

Our atmosphere has a weight that constantly pushes down on us, about 14.7 psi; this is the barometric pressure we hear about from the weatherman. When we measure the air in automobile tires we are actually measuring gauge pressure. The air pressure inside the tire must also overcome the current barometric pressure, and yet when the barometric pressure changes the gauge pressure that we measure does not because it is relative to the barometric pressure. When an instrument is measuring gauge pressure the units are referred to as psig (pounds per square inch gauge). Most pressure measurements are made with gauge pressure instruments.

> ☞ Absolute pressure is defined as the pressure relative to an absolute vacuum. The units of absolute pressure are psia.

If we were to measure the air pressure in an automobile tire with an absolute pressure gauge we would get a reading about 14.7 psi higher than we would expect. Absolute pressure gauges are commonly used in vacuum systems as we are trying to control the atmospheric conditions that the process is exposed to negate the effects of daily barometric pressure changes.

☞ Differential pressure uses a reference point other than full vacuum or atmospheric pressure.

A differential pressure gauge will have two sensing elements and the process variable will be the difference between the two measurements.

Differential pressure gauges are commonly used in tanks or vessels in which the pressure in the head space varies. A differential pressure gauge will subtract the head pressure from the total pressure giving an accurate measurement of the liquid pressure.

As shown in Figure 3-13, the type of pressure sensor is defined by its reference point.

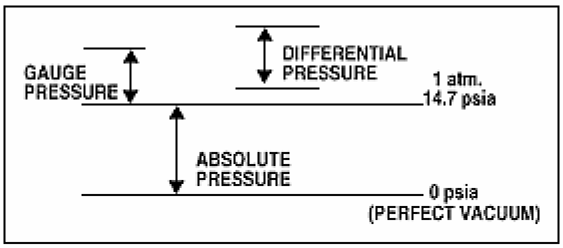


Figure 3-13

What is Level?

Material level in a tank or vessel is a common process variable in plants. We measure level for filling control, for inventory control, and for providing safe operations to personnel and equipment. Level is often measured in percent of span but may be displayed in engineering units of feet or inches, or even pounds if the tank geometry and material density is taken into account. Level measurement technologies can be placed into one of two groups, contact and non-contact level measurement; level sensors in each of these groups can be of the continuous or point variety.

Point and Continuous Level

Point sensing level probes only sense tank level at a discrete level. Point sensing probes are typically used for high-high or low-low level sensing to prevent plant personnel and/or process equipment from being exposed to harmful conditions. Point level probes are also used in pairs in processes in which we do not particularly care what the exact level in a tank is, only that it is between two points.

Continuous level probes sense the tank level as a percent of span of the probes capabilities. Continuous level probes are typically used where we need some type of inventory control, where we need to know with some degree of confidence what the particular level in a tank is.

Common Level Sensing Technologies

Non-Contact Level Measurement

Non-contact level measurement, as the name implies, requires that the sensing element not be in contact with material being measured. The use of non-contact level sensors eliminates agitator interference. Non- contact level sensing is primarily either ultrasonic or radar/microwave.

Ultrasonic Measurement

Ultrasonic makes use of sound waves in the 20 - 200 kHz range (above the range for human hearing). A transducer mounted in the top of a tank transmits sound waves in bursts onto the surface of the material to be measured. Echoes are reflected back from the surface of the material to the transducer and the distance to the surface is calculated from the burst-echo timing.

The key points in applying an ultrasonic transducer are:

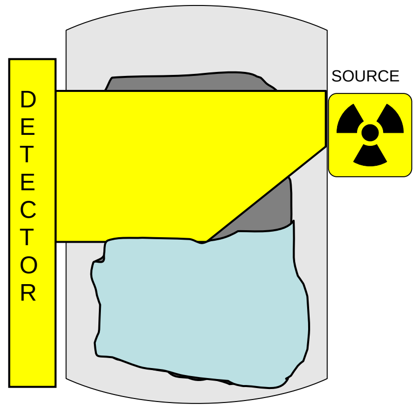
- The speed of sound varies with temperature. If the transducer does not use temperature compensation and the temperature of the air space in the vessel varies your level readings will not be correct.
- Heavy foam on the surface of the material can absorb the sound wave bursts resulting in no echo or an echo that is too weak to process.
- An irregular material surface can cause false echoes resulting in irregular readings.
- Heavy vapor in the air space can distort the sound waves resulting in false reading.

Radar / Microwave

Radar, or microwave level measurement operates on similar principles to ultrasonic level probes, but instead of sound waves electromagnetic waves in the 10GHz range are used. When properly selected, radar can overcome many of the limitations of ultrasonic level probes.
Radar can:

- Be unaffected by temperature changes in the tank air space.
- See through heavy foam to detect the true material level.
- See through heavy vapor in the tank air space to detect true material level.

Nuclear Level Sensor



Continuous nuclear level detection is typically used where most other technologies are unsuccessful. Different radioactive isotopes are used, based on the penetrating power needed to pass through the tank. Radiation from the source is detected on the other side of the tank. Its strength indicates the level of the fluid. Point, continuous, and interface measurements can be made.

As no penetration of the vessel is needed there are a number of situations that cause nucleonic transmitters to be considered over other technologies. These applications generally involve high temperatures / pressures or where toxic or corrosive materials are within the vessel. Placing the source and / or detector in wells within the vessel can reduce source sizes. An extension of this is to use a moving source within the vessel; this facilitates the unique ability to combine density profiling with accurate tracking of a moving interface.

Nuclear level detection has some drawbacks. One is high cost, up to four times that of other technologies. Others are the probable requirement for licenses, approvals, and periodic inspections; and the difficulty and expense of disposing of spent radiation materials. Another factor to consider is that the radiation symbol found on these devices can cause concern to plant personnel. From a psychological standpoint, the radiation symbol found on these controls is frequently the cause of unfounded concern with uninitiated plant personnel. Plant Management is usually required to ensure that appropriate education is given to any staff likely to be involved with this measurement technology.

Contact Level Measurement

Contact level measurement, as the name implies, requires that the sensing element be in contact with material being measured. Continuous level measurements can either be direct or indirect with contact level sensors. With direct level measurement the sensor is in contact with the material over the entire span that we wish to measure. Direct contact continuous level measurement is commonly used in bulk powder storage silos and un-agitated tanks. RF capacitance/resistance level sensors are the typical direct contact level sensors for both the point and continuous level variety.

With indirect level measurement the sensor is in contact with the material at a single point and the tank level is inferred from this point measurement. Pressure transmitters measuring hydrostatic head are typically used for indirect contact continuous level measurement. For agitated tanks indirect contact continuous level measurement is advantageous.

Pressure Measurement

Measurement of level by pressure relies on hydrostatic principles. Pressure is a unit force over a unit area. A cubic foot (12"L x 12"W x 12"H) of water weighs 62.4796 pounds. The area that a cubic foot of water occupies is 144 square inches (12"L x 12"H); therefore a cubic foot of water exerts a force of 62.4796 pounds over 144 square inches, or 0.4339 psig for a 12" water column.

It would not matter how many cubic feet of water were placed side by side, the pressure would still be 0.4399 psig. Hydrostatic pressure is only dependent on the height of the fluid, not the area that it covers.

Going back to the cubic foot of water, every inch of water we add or subtract to the level will change the pressure measurement by 0.03616 psig. The relationship between psig and level is linear so pressure is easily converted to a level process variable.

☞ 1 inch H20 = 0.0316 psig

While this is all well and good for water, how would we handle measuring other fluids using pressure - compare the densities.

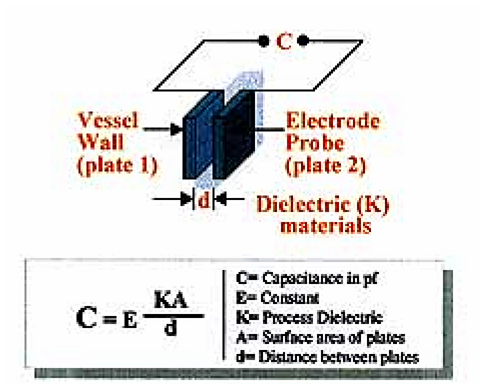
For instance, chocolate weighs somewhere around 80 pounds per cubic foot, so 80 divided 62.5 times 0.0316 = 0.0404 psig. For a change in level of one inch in a tank filled with liquid chocolate the pressure measurement will change by 0.0404 psig.

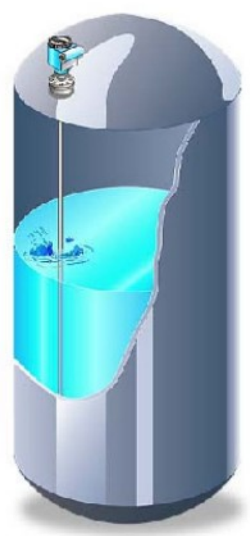
This also points to one of the constraints of level measurement by pressure; the material requires a constant density for accurate measurements.

Another consideration for using pressure to infer level is the pressure found in the head space of the tank. As we are using gauge pressure our reference is atmospheric pressure, we have an open tank! If this is not the case, if the head pressure in the tank can be other than atmospheric, we must use a differential pressure sensor to measure level.

RF Capacitance / Resistance

RF (radio frequency) Capacitance level sensors make use of electrical characteristics of a capacitor to infer the level in a vessel. A capacitor is essentially two metal plates separated by a dielectric (insulator) and acts as a storage vessel for electrons. The size of the metal plates, the distance between the plates and dielectric filling the space all determine the capacitance value, or how many electrons can be stored. In an RF capacitance probe one of the metal plates is the tank wall, the other metal plate is the sensor probe and the dielectric is either air (where the tank is empty) or the material in the tank.





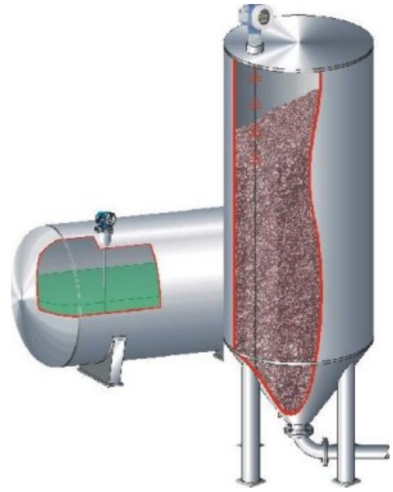
As the material rises in the vessel, the capacitance changes. The level transducer measures this change, linearizes it and transmits the signal to the process control system. A point level probe will look for a specific change in capacitance to determine whether it is on or off.

A variation on the RF capacitance probe is the RF admittance probe. The RF admittance probe measures total impedance rather than just capacitance. RF admittance probes can be adapted to measure material buildup on the probe and provide a signal that is only proportional to the material level. RF admittance probes may also be referred to as RF impedance probes.

As we can see, RF Capacitance probes are dependent on the electrical properties of the material being measured and having a sufficient amount of probe surface area immersed in the material. Proper selection requires informing the probe vendor of the material to be measure, especially in applications where you are measuring conductive materials or have a nonmetallic tank. Point probe sensitivity can be increased by welding a plate on the sensor tip to increase the capacitance, and therefore the sensitivity (gain).

Guided Wave Radar

Guided wave radar is similar to the non-contact radar probes, only a rod or cable is immersed into the material like an RF capacitance probe. The rod or cable is used to guide the microwave along its length, where the rod or cable meets the material to be measured a wave reflection is generated. The transit time of the wave is used to calculate level very precisely. Unlike an RF capacitance probe, a guided wave radar probe can measure extremely low dielectric material.



What is Flow?

☞ Flow is the motion characteristics of constrained fluids (liquids or gases).

The motion characteristics we are most interested in measuring are velocity and mass. As with most other process measurements, these characteristics are most often not measured directly. We measure process variables that are more readily accessible and use established relationships between the velocity or mass of the flow and the variable that we are measuring.

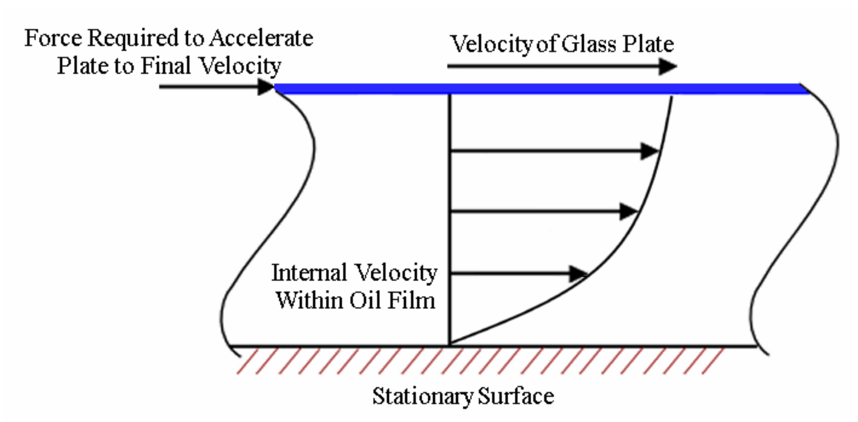
Factors Affecting Flow Measurement

Since in most cases we are not measuring flow directly, it is important to understand the factors of your particular application that affect the selection and installation of flow instrumentation. The critical factors in flow instrumentation selection and application are viscosity, fluid type, Reynolds number and flow irregularities.

Viscosity

☞ Dynamic or absolute viscosity (η) is measure of the resistance to a fluid to deformation under shear stress, or an internal property of a fluid that offers resistance to flow.

Viscosity is commonly perceived as "thickness" or resistance to pouring. Water is very "thin" having a relatively low viscosity, while molasses is very thick having a relatively high viscosity. The concept of viscosity may also be visualized by picturing a flat sheet of glass on a film of oil on top of a flat surface. If we apply a parallel force to the sheet of glass it will accelerate to velocity, the final velocity being reached dependent only on the amount of force applied.



The velocity of the oil will not be uniform through the film layer. The oil that is next to the sheet of glass will have a velocity close to that of the glass; while the oil that is next to the stationary surface will have a velocity near zero. This internal distribution of velocities is due to the internal resistance of the fluid to shear stress forces, its viscosity.

The viscosity of a fluid may then be thought of as a ratio between the per unit force to accelerate the plate and the distribution of the velocities within the film, or

$$\text{Dynamic Viscosity} = \frac{\left(\frac{F}{A}\right)}{\left(\frac{v}{x}\right)} = \frac{Fx}{Av}$$

Where : F = Force required to accelerate the plate

A = Area of the plate

v = Final velocity of the plate

x = Thickness of the film

Temperature and Pressure Effects on Viscosity

The dynamic viscosity of a fluid varies with its temperature. In general, the viscosity of a liquid will decrease with increasing temperature while the viscosity of a gas will increase with increasing temperature. Viscosity measurements are therefore associated with a particular temperature.

For both liquids and gases, pressure has very little effect on viscosity between 0 and 1500 psi. Extremely high pressures tend to increase viscosity.

Units of Viscosity

Viscosity is measured in units of poise or Pascal·seconds or stokes. When viscosity is expressed in poise or Pascal·seconds it is meant that we are measuring dynamic viscosity. When viscosity is expressed in units of stokes it is meant that we are measuring kinematic viscosity.

☞ Kinematic viscosity is the dynamic viscosity of a fluid divided by the density of the fluid.

The poise (P) has units of $\frac{\text{kg}}{\text{m} \cdot \text{s}}$ and is commonly expressed as centipoise (cP);
1 centipoise = 0.01 poise.

The Pascal-second has units of $\frac{\text{N} \cdot \text{s}}{\text{m}^2}$ and is commonly expressed as milliPascal·seconds (mPa·s);
1000 milliPascal·seconds = 1 Pascal·second.

The stoke (S) has units of $\frac{\text{m}^2}{\text{s}}$ and is commonly expressed as centistokes (cS);
1 centistoke = 0.01 stokes.

Viscosities and Densities of Common Household Fluids

Fluid	Temperature (°C)	Viscosity (cP)	Density (lb/ft^3)
Corn Syrup	25	2,000 – 3,000	88
Corn Syrup	50	160	88
Chocolate Syrup	85	54	82
Milk	25	3	64
Molasses	20	5,000	87
Peanut Butter	20	150,000 – 250,000	68
Water	20	1	62.4

☞ The fluid streams in processes can include both high and low viscous fluids. The viscosity of the fluid under process conditions must be taken into account when selecting a flow instrument for optimum performance.

Conversion Tables

VISCOSITY

poise P	centipoise cP	Pascal·seconds Pa·s	milliPascal·sec mPa·s	stokes S	centistokes cS
1	100	0.1	100	x 1/density	x 100/density
0.01	1	.001	1	x 0.01/density	x 1/density
10	1000	1	1000	x 10/density	x 1000/density
0.01	1	0.001	1	x 0.01/density	x 1/density
x density	x 100 density	x 0.1 density	x 100 density	1	100
x 0.01 density	x density	x 0.001 density	x density	0.01	1

VOLUMETRIC FLOW RATE

ft^3/sec	ft^3/min	liters/min	m^3/min	m^3/hr	gal/min
1	60	1699	1.699	101.95	448.83
0.01667	1	28.32	0.02832	1.699	7.481
5.855×10^{-4}	0.03531	1	0.001	0.06	0.2642
0.5885	35.31	1000	1	60	264.2
0.00981	0.5885	16.667	0.01667	1	4.403
0.002228	0.1337	3.786	0.003785	0.2271	1

GRAVIMETRIC FLOW RATE

lb/sec	lb/min	lb/hr	gm/sec	gm/min	kg/hr
1	60	3600	435.6	27220	1633
0.01667	1	60	7.560	453.6	27.22
00.0002778	0.01667	1	0.1260	7.560	0.4536
0.002205	0.1323	7.938	1	60	3.600
3.675×10^{-5}	0.002205	0.1323	0.01667	1	0.600
6.125×10^{-4}	0.03675	2.205	0.2778	16.67	1

VOLUME

gallons (US)	ft^3	in^3	barrels (oil)	cm^3	liters
1	0.1337	231	0.02381	3785	3.785
7.481	1	1728	0.1781	28320	28.32
0.004329	0.0005787	1	1.031×10^{-4}	16.39	0.01639
42	5.615	9702	1	158940	15.894
0.000264	3.53×10^{-5}	0.06102	6.29×10^{6}	1	0.001
0.264	0.0353	61.03	0.0629	1000	1

VELOCITY

ft/sec	ft/min	ft/hr	meter/sec	meter/min	meter/hr
1	60	3600	0.3048	18.29	1097.3
0.01667	1	60	0.005080	0.3048	18.29
0.0002778	0.01667	1	8.4667×10^{-5}	0.005080	0.3048
3.281	196.85	11811	1	60	3600
0.0547	3.281	196.85	0.01667	1	60
0.0009117	0.0547	3.281	0.0002774	0.01667	1

Fluid Type

Fluids may generally be divided into two types: Newtonian and Non-Newtonian fluids.

Newtonian Fluids

☞ When held at a constant temperature, the viscosity of a Newtonian fluid will not change regardless of the size of the shear force.

Water, glycerin, corn syrup and liquid sugar are examples of Newtonian fluids.

Non-Newtonian Fluids

☞ When held at a constant temperature, the viscosity of a Non-Newtonian fluid will change with relation to the size of the shear force, or will change over time under a constant shear force.

Non-Newtonian fluids typically have viscosities above 500 centipoise and into one of two classes.

Class I: Time-independent non-Newtonian fluids. These fluids have a viscosity that does not vary with time at a given shear stress. However, their viscosity does vary with increases or decreases in the shear stress. Class I non-Newtonian fluids include pseudo-plastic, dilitant and Bingham fluids.

- Pseudo-plastic, or shear thinning fluids undergo a decrease in viscosity at constant temperature and increasing shear forces.
  Example: chocolate paste
- Dilitant, or shear thickening fluids undergo an increase in viscosity at constant temperature and increasing shear forces.
  Example: starch/water solutions, peanut butter
- Bingham, or plastic, fluids require a minimum shear force to be applied before flow will occur and then behave in a Newtonian manner.
  Example: cocoa butter

Class II: Time-dependent non-Newtonian fluids. These fluids have a viscosity that does vary with time at a given shear stress. Class II non-Newtonian fluids include thixotropic and rheopectic fluids.

- Thixotropic, or time thinning, fluids undergo a decrease in viscosity at constant temperature and constant shear forces.
  Example: lecithin
- Rheopectic, or time thickening, fluids undergo an increase in viscosity at constant temperature and constant shear forces.
  Example: milk, molasses

> ☞ Many fluid streams exhibit non-Newtonian behavior. That is, the viscosity of the fluid will change with time or the amount of shear stress (agitation or pumping action). Flow instruments that are not sensitive to viscosity should be selected for measuring non-Newtonian flows.

Reynolds Number

> ☞ The Reynolds number is the ratio of inertial forces to viscous forces of fluid flow within a pipe and is used to determine whether a flow will be laminar or turbulent.

The Reynolds number is given by:

$$Re = 124\frac{Dv\rho}{\eta}$$

Where : Re = Reynolds number

D = Pipe diameter in inches

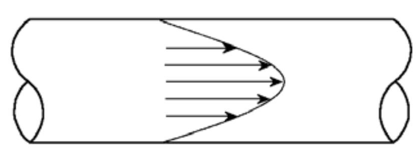
v = Fluid velocity in ft/sec

ρ = Fluid density in lb/ft^3

η = Fluid viscosity in cP

Laminar Flow

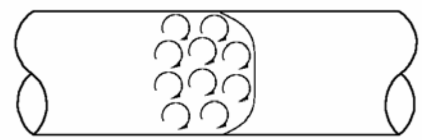
Laminar flow occurs at low Reynolds numbers, typically $Re < 2000$, where viscous forces are dominant. Laminar flow is characterized by layers of flow traveling at different speeds with virtually no mixing between layers. The velocity of the flow is highest in the center of the pipe and lowest at the walls of the pipe.



Laminar Flow

Turbulent Flow

Turbulent flow occurs at high Reynolds numbers, typically *Re* > 4000, where inertial forces are dominant. Turbulent flow is characterized by irregular movement of the fluid in the pipe. There are no definite layers and the velocity of the fluid is nearly uniform through the cross-section of the pipe.
The flow is turbulent.



Turbulent Flow

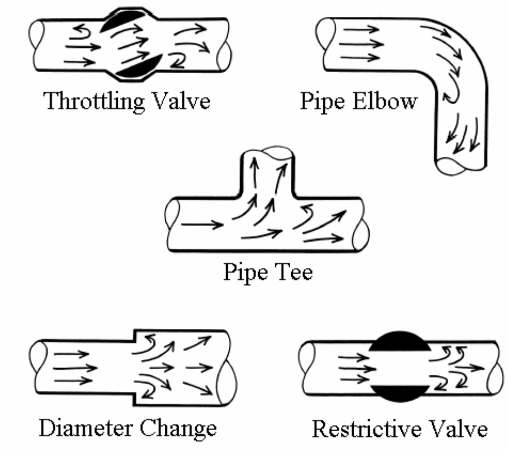
Transitional Flow

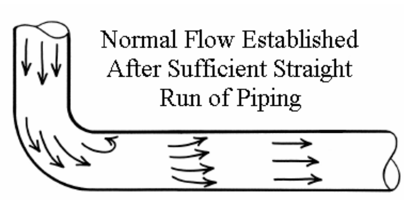
Transitional flow typically occurs at Reynolds numbers between 2000 and 4000. Flow in this region may be laminar, it may be turbulent or it may exhibit characteristics of both.

> ☞ Many fluid streams are laminar. That is, they have a low Reynolds number. Flow instruments must be selected to accurately measure laminar flow.

Flow Irregularities

Flow irregularities are changes in the velocity profile of the fluid flow caused by installation of the flow instrument. Flow irregularities are typically caused by throttling valves, restrictive valves, pipe elbows, pipe tees and changes in pipe size.





Normal velocity profiles are established by installing a sufficient run of straight piping after valves or fittings that will cause a flow irregularity.

> ☞ Some flow instrumentation is sensitive to flow irregularities and require minimum lengths of straight piping upstream and downstream of the instrument. These minimum distances are often expressed in terms of pipe diameters.

Common Flow Instruments

Flow meters come in many different technologies but can be categorized by the variable that they measure: the velocity of the fluid flow or the mass of the fluid flow.

We have learned that process streams typically involve fluids coving a wide range of viscosities, involving many non-Newtonian fluids and laminar flows. The ideal flow meters to use for process measurements would then be able to handle a wide range of viscosities, be unaffected by changes in fluid viscosity and be suitable for flows with low Reynolds numbers.

From the chart on the following page we see that two fit these criteria very well: magnetic flow meters for volumetric measurements and coriolis flow meters for mass measurements. For those occasions where fluids with extremely high viscosities must be measured positive displacement meters are used.

FLOWMETER SELECTION GUIDE

Technology	Typical *Re* or viscosity	Expected Error from Viscosity	US/DS Pipe Diameters	Viscous Fluids	Clean Fluids	Dirty Fluids	Slurry	Steam	Pressure Drop	Sizes Available (inches)	Flow Rangeability	Typical Accuracy R=Rate FS=Full Scale	Relative Cost
VOLUMETRIC FLOWMETERS													
Positive Displacement	*Re* No Limits *<= 8,000 cS*	Low	None	√	√				High	1-16	15:1	0.25 – 0.5% R	Medium
Ultrasonic Doppler	*Re > 4,000*	None	20 / 5	?		√	?		None	≥ ½	10:1	3% R	High
Ultrasonic Transit Time	*Re* > 10,000	None	20 / 5	?	√	?			None	≥ ½	20:1	1% R	High
Magnetic	*Re* > 4,500	None	5 / 3	√	√	√			None	1/10 - 104	30:1	0.25 – 1%R	High
Orifice Plate	*Re* > 10,000	High	20 / 5	?	√	?		√	Medium to High	All	4:1	2-4% FS	Low
Vortex Shedding	*Re* > 10,000 < 30 cP	Medium	20 / 5	?	√	?		√	Medium	½ - 12	40:1	1%R	High
Venturi	*Re* > 75,000	High	15 / 5	?	√	?		√	Low	3 - 72	4:1	2-4% FS	Medium
Flow Nozzle	*Re* > 50,000	High	20 / 5	?	√	?		√	Medium	3 - 48	4:1	2 – 4% FS	Medium
Pitot	*Re* > 100,000	Low	30 / 5		√	?		√	Low	All	3:1	3 – 5% FS	Low
MASS FLOWMETERS													
Coriolis	*Re* No Limits	None	None	√	√	?	?		Medium	1/10 - 8	20:1	< 0.5% R	High
Thermal Mass	*Re* No Limits	None	20 / 5		√	?		√	Low	All	20:1	2% FS	High

Volumetric Flow

Because the cross-sectional area of a pipe is a known quantity, the velocity of the fluid moving through the pipe cross-section is often measured and the volumetric flow is inferred. The basic relationship for determining volumetric flow is:

$$Q = V \times A$$

Where

Q = volumetric flow through a pipe

V = average velocity of the fluid flow

A = cross sectional area of the pipe where the velocity is measured.

Units of Volumetric Flow

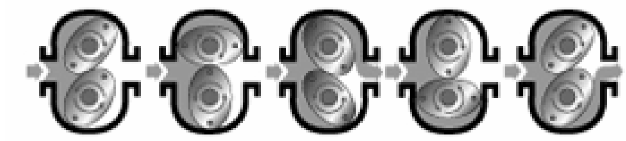
Volumetric flow meters provide a signal that is in units of volume per unit time.

Typical units of volumetric flow are:

- gpm (gallons per minute)
- cfm (cubic feet per minute)
- l/hr (liters per hour)

Positive Displacement Flow Meters

Positive displacement flow meters operate by capturing the fluid to be measured in rotating cavities of a known volume. The volume of the cavity and the rate of rotation will give the volumetric flow value.

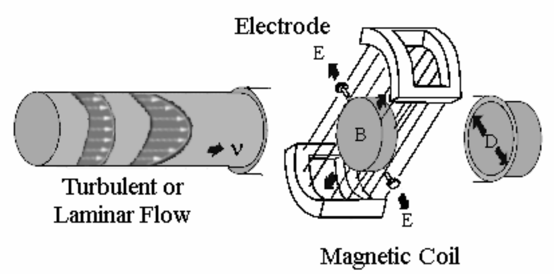


Positive displacement meters are unaffected by Reynolds number and work with laminar, turbulent and transitional flow. However, low viscosity fluids will slip past the gears decreasing the accuracy of the flow meter.

Magnetic Flow Meters

Magnetic flow meters infer the velocity of the moving fluid by measuring a generated voltage.

Magnetic flow meters are based on Faraday's law of electromagnetic induction- a wire moving though a magnetic field will generate a voltage.



$E = kBDv$

k = Constant

B = Magnetic Flux Density

D = Pipe Diameter

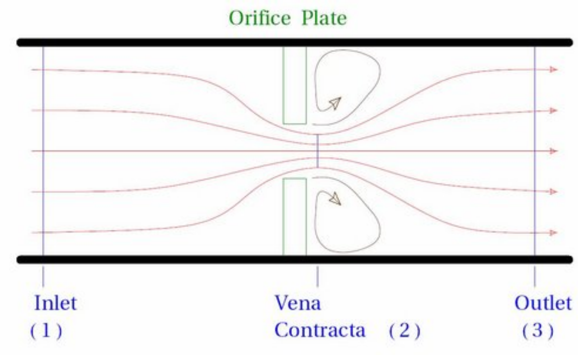
v = Fluid Velocity

In a magnetic flow meter the fluid acts as the moving wire and a magnetic field is place around the flow path. The voltage generated by the moving fluid is proportional to its velocity. Magnetic flow meters require a conductive liquid to operate.

$$Flow = Pipe\ Area \cdot velocity$$
$$= \pi r^2 v = \frac{\pi D^2}{4} v$$
$$= \frac{\pi D^2}{4} \frac{E}{kBD} = \frac{\pi DE}{4kB}$$

Orifice Plate**

An orifice plate is a restriction with an opening smaller than the pipe diameter which is inserted in the pipe; the typical orifice plate has a concentric, sharp edged opening. As fluid flows through the pipe it has a certain velocity (which we want to measure) and a certain pressure (which is quite easily measured). When the fluid reaches the orifice plate, with the hole in the middle, the fluid is forced to converge to go through the small hole; the point of maximum convergence is actually just after the physical orifice, at the so-called "vena contracta" point (see diagram). As it does so, the velocity and the pressure changes. By measuring the difference in fluid pressure between the normal pipe section and at the vena contracta, we can find the velocity of the fluid flow by applying Bernoulli's equation.



With negligible frictional losses or change of elevation, Bernoulli's equation reduces to an equation relating the conservation of energy at two points in the fluid flow:

$$P_1 + \frac{1}{2} * \rho * V_1^2 = P_2 + \frac{1}{2} * \rho * V_2^2$$

or

$$P_1 - P_2 = \frac{1}{2} * \rho * V_2^2 - \frac{1}{2} * \rho * V_1^2$$

with

$Q = A * V$ or $V = Q / A$

and $Q_1 = Q_2$

$$P_1 - P_2 = \frac{1}{2} * \rho * (\frac{Q1}{A2})^2 - \frac{1}{2} * \rho * (\frac{Q1}{A1})^2$$

Solving for Q_1:

$$Q_1 = A_2 * \left(\frac{2 * (P_1 - P_2)}{\rho * (1 - A_2^2 / A_1^2)} \right)^{0.5}$$

And finally introducing terms to account for fluid compressibility and orifice geometry:

$$Q_1 = A_1 * V_1 = C_{meter} * Y * A_2 * \left(\frac{2 * (P_1 - P_2)}{\rho * (1 - A_2^2 / A_1^2)} \right)^{0.5}$$

Q_1 is the volumetric flow (e.g. gallons/hour)
V_1 is the fluid velocity
A_1 is the pipe area
A_2 is the orifice area
A_2 / A_1 is known as the β ratio.
P_1 is the inlet pressure
P_2 is the outlet pressure
ρ is the density of the fluid

The meter coefficient, C_{meter}, accounts for all non-idealities, including friction losses, and depends on the type of meter, the ratio of cross sectional areas and the Reynolds number. Equations for this exist, but for rough calculations C_{meter} can be approximated to 0.62.

The compressibility factor, Y, accounts for the expansion of compressible gases; it is 1.0 for incompressible fluids. These two factors can be estimated from correlations.
This type of meter only works well when supplied with a fully developed flow profile. This is acheived by a long upstream length (20 to 40 diameters, depending on Reynolds number) or the use of a flow conditioner.

Orifice plates are small and cheap to install, but impose a significant energy loss on the fluid due to friction

** This entry is from Wikipedia, the leading user-contributed encyclopedia.

Mass Flow
Mass flow meters measure mass flow directly by measuring the inertial effects of the fluids moving mass.

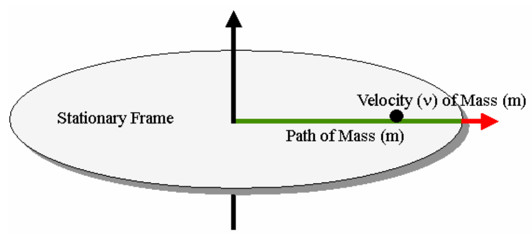
Units of Mass Flow

Mass flow meters provide a signal that is in units of mass per unit time.
Typical units of mass flow are:

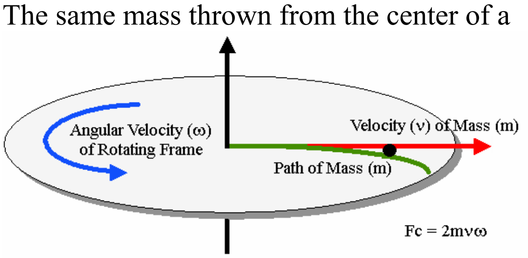
- lb/hr (pounds per hour)
- kg/s (kilograms per second)

Coriolis Flow Meters

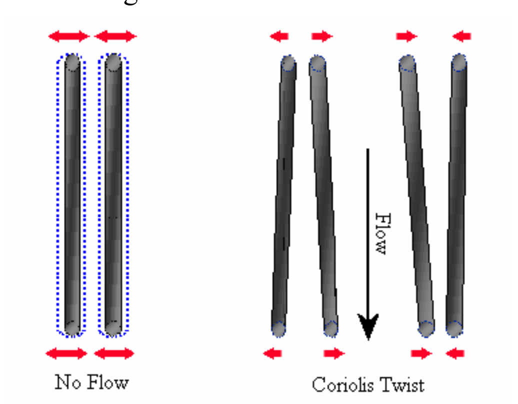
Coriolis flow meters are based on the effects of the Coriolis force. Simply, when observing motion from a rotating frame the trajectory of motion will appear to be altered by a force arising form the rotation.



A mass thrown from the center of a non-rotating disk will appear to travel in a straight line in the direction of its velocity when viewed both from the disk and outside of the disk's reference.
The same mass thrown from the center of a rotating disk will appear to curve its trajectory counter to the rotation of the disk when viewed from the disk's frame of reference. This apparent change in motion is accounted for by a term called the Coriolis force which is proportional to the mass of the moving object.



Coriolis mass flow meters work by applying a vibrating force to a pair of curved tubes through which fluid passes, in effect creating a rotating frame of reference. The Coriolis Effect creates a force on the tubes perpendicular to both the direction of vibration and the direction of flow, causing a twist in the tubes.



Under no-flow conditions the tubes will vibrate parallel to each other. As the tubes vibrate away from each other under flow conditions the Coriolis force will cause the tubes to deflect towards each other at the inlet. As the tubes vibrate towards each other under flow conditions the Coriolis force will cause the tubes to deflect away each other at the inlet. The amount of twist generated is proportional to the mass flow though the tubes.

Turndown

A specification commonly found with flow sensors is turndown or rangeability. Turndown is the ratio of maximum to minimum flow that can be measured with the specified accuracy.

For example, a flow meter with a full span of 0 to 100 gallons per minute with a 10:1 turndown could measure a flow as small as 10 gallons per minute. Trying to measure a smaller flow would result in values outside of the manufacturer's stated accuracy claims.

Installation and Calibration

As in any instrument installation, optimal flow meter performance and accuracy can only be achieved through proper installation and calibration. Requirements will vary between type and manufacturer by some general considerations are:

- The length of straight piping required upstream and downstream of the instrument. This is usually specified in pipe diameters.
- The properties of the fluid/gas to be measured.
- Mounting position of the flow meter. Flow meters must be kept full during operation. Also many flow meters must only be "zoered" when full.

☞ A final control element is the last system element that responds to a control signal and performs the actual control action.

Every process control loop contains a final control element, a device that enables the process variable to be manipulated. There are many types of final control elements, valves, feeders and variable speed devices that are commonly found in production processes.

The function of a final control element is to follow the command of the controller output to eliminate the error in the process variable. In this chapter we will be learning a little about two of the most common final control elements: valves and pumps.

The intent of this chapter is not to teach how to size a valve or select a pump nor to familiarize the student with all of the available types of valves and pumps. The intent of this chapter is to provide an introduction to some of the most commonly used pumps and valves: basic terminology and characteristics relevant to their role in a control loop. Detailed information and assistance on device selection is readily available from most instrumentation suppliers.

Valves

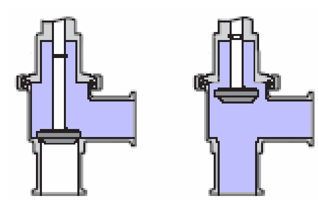
What is a Control Valve?

The most common final control element in process industries is the control valve. Control valves manipulate the flow of utilities (such as steam, air, water or glycol), ingredients or in process streams, compensating for load or Set Point changes to keep the regulated process variable as close as possible to the desired Set Point.

> ☞ A control valve is an inline device in a flow stream that receives commands from a controller and manipulates the flow of a gas or fluid in one of three ways:
> 1) Interrupt flow (shut-off service)
> 2) Divert flow to another path in the system (divert service)
> 3) Regulate the rate of flow (throttling service)

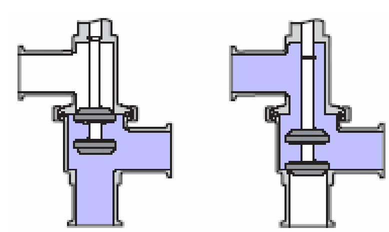
Shut-Off Service

Control valves for flow shut-off service have two positions. In the open position flow is allowed to exit the valve. In the closed position flow is blocked from exiting the valve.



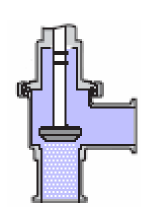
Divert Service

Control valves for divert service also have two positions, however flow is never blocked. In the one position flow is allowed from the common port to port A. In the other position flow is allowed form the common port to port B.



Throttling Service

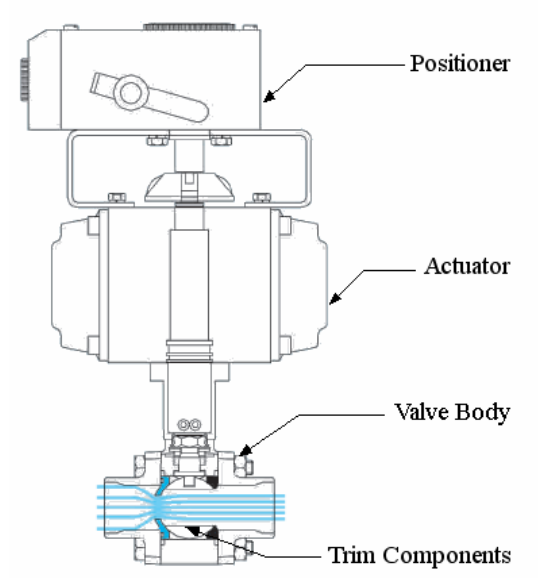
Control valves for throttling service have many positions. The position of the valve determines the rate of flow allowed through the valve.



Parts of a Control Valve

When we talk about control valves we are actually referring to assemblies. A control valve consists of an actuator, the internal trim parts and the valve body itself. Control valve can also have additional accessories such as supply pressure regulators, positioners, I/P transducers, manual operators and limit switches.

It is the particular combination of actuator, internal trim parts and valve body that will determine the type of service the control valve will be suited for.

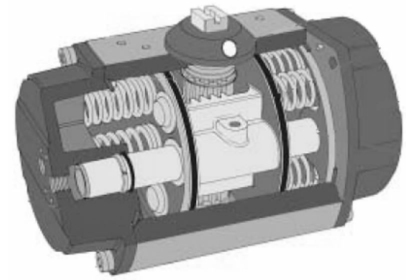


☞ Trim parts: the internal parts of a valve that are in contact with the manipulated fluid.

What is an Actuator?

☞ Actuators are pneumatic, electrical or hydraulic devices that provide the force and motion to open and close a valve.

Many common valve actuators found in industry are on/off pneumatically operated at 60 to 80 psi by solenoid control. An on/off pneumatic actuator can be air driven in both directions or can be pneumatic in one direction and spring return in the other.



The configuration of the actuator determines the fail position of the valve. An actuator that is pneumatically driven in both directions will fail in its last position. An actuator with a spring return will fail to the pneumatically de-energized position. If this is the closed position on the valve then we have a fail closed configuration. If it is the open position then we have a fail open configuration.

The interface to a control system for an on/off actuator is typically through a solenoid valve.

Some actuators can place a valves at any position between the on and off points. These actuators typically accept a 3-15 psi signal to move a diaphragm, which in turn moves a connected valve stem. A pneumatic positional actuator will fail in to the pneumatically de-energized position.

The interface to a control system for a positional actuator is typically through an I/P transducer. An I/P transducer converts a 4-20 mA signal form a controller to a pneumatic signal the positioner can accept.

Just as processes have a time constant and dead time, valves have a time constant and dead time which is largely determined by its actuator. Valve manufacturers measure their valve response by a parameter called T_{63}, which is the time it takes for a valve to reach 63% of its final position in response to a command change after the dead time has passed.

No matter what type of actuator is used it is important that it is sized correctly. Not only to minimize T_{63} and dead time, but too small of an actuator will not have sufficient force to reliably position the valve at its command position.

What is a Positioner?

☞ A valve positioner is an accessory to a positional actuator that provides closed loop control of the valve's position.

A positioner is mechanically linked to the valve stem and compares the command signal to the valve with the actual stem position and corrects for error. If a positioner receives a command signal of 50 percent it will maintain the valve at its 50 percent position.

Since a positioner is a feedback controller it too has tuning parameters. While a position can improve valve performance by reducing deadband, a poorly tuned positioner can make your control loop unstable and hard to tune.

In effect, a control loop with a positioner is a cascade control loop. As in cascade loops start by tuning the inner loop, in this case the positioner. Proper tuning for a positioner is where the position response is crisp without overshoot. Once your positioner is tuned you may proceed with tuning of the loop.

What is C_v?

☞ C_v is the symbol for valve coefficient, a measure of the flow capacity of a valve at a set of standard conditions.

☞ The flow capacity of a valve is the amount of fluid it will pass per unit of time. Flow capacity is usually expressed in gallons per minute (GPM).

The valve coefficient is defined as flow of water that will pass through the valve when fully open with a pressure drop of 1 psi.

$$C_V = Q\sqrt{\frac{sg}{\Delta P}}$$

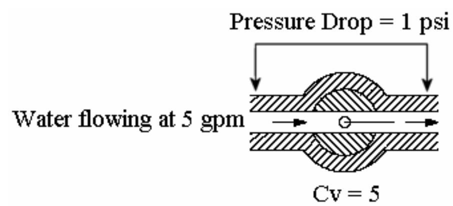
$$\text{Where}: Q = \text{Fluid Flow in GPM}$$

$$sg = \text{Specific Gravity of the Fluid}$$

$$\Delta P = \text{Pressure Drop in PSI}$$

☞ The specific gravity of a fluid is the weight of a fluid divided the weight of an equal volume of water.

A valve must be of sufficient size to pass the flow required to satisfy the process under all possible production scenarios at an acceptable pressure drop.



What are Valve Characteristics?

Valves are characterized by how their C_V, or flow varies with respect to the position of its closure member. This characteristic is classified as either inherent or installed.

☞ A closure member is the internal part of a valve that manipulates the fluid flow.

Inherent Characteristics

☞ The inherent characteristic of a valve is the relationship between the flow rate through the valve and the travel of the closure member as the closure member is moved from the closed position to its rated travel with a constant pressure drop across the valve.

Inherent valve characteristics are measured by the valve manufacturer in a test stand under a specified set of process conditions, particularly a constant differential pressure across the valve.

The three most common valve characterizations are equal percentage, linear and quick opening (Figure 4-2).

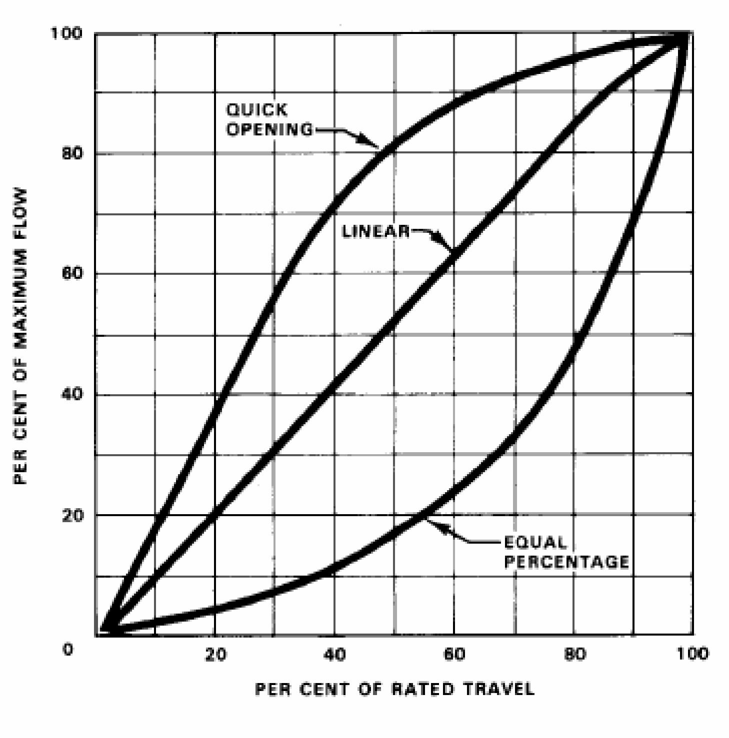


Figure 4-2

Rangeability

Associated with inherent valve characteristics is the inherent rangeability of a valve.

> ☞ Inherent rangeability is the ratio of the largest flow coefficient (Cv) to the smallest flow coefficient (Cv) of a valve as its closure member travels through its range without deviating beyond specified limits for its characteristic.

A valve with a C_V of 100 and a rangeability of 50 will perform within its characteristic from a C_V of 2 to a C_V of 100.

Gain

Also associated with the characteristic of a valve is the gain of a valve. As gain is the $\frac{\%\text{ change in output}}{\%\text{ change in input}}$, the gain of a valve is slope of its characteristic curve $\frac{\%\text{ Max Flow (Cv)}}{\%\text{ Rated Travel}}$, which is proportional to the C_V of the valve.

☞ If you double the CV of a valve, it will have twice the gain.

While a valve must be sufficiently sized for flow, an oversized valve will lead to large process gains (the gain of a process as seen by a controller is the product of the gains of the sensor, final control element and the process itself). Processes with a large gain amplify valve problems (deadband and stiction) and can be oscillatory or even unstable.

Valves that are oversized for a process will also operate close to their seat under normal process conditions. The problem with this is most valves become very nonlinear in this operating region, or the process may saturate at a low controller output leading to windup.

Equal Percentage Valves

☞ The inherent flow characteristic of equal percentage valves is for equal increments of rated travel the flow characteristic (C_v) will change by equal percentages.

Equal percentage valves are the most commonly used control valves. The gain of an equal percentage valve is nonlinear. The gain of an equal percentage valve increases from its smallest value near closure to its largest value at full open.

Table 4-1 shows how the C_V and gain will vary for an equal percentage valve in which the flow changes by 100% for each equal valve increment.

Valve Opening	% of Maximum Flow (C_V)	Gain (% Flow / % Open)
0	0	
20	6.25	0.3125
40	12.5	0.3125
60	25	0.417
80	50	0.625
100	100	1.0

Table 4-1

Linear Valves

☞ The inherent flow characteristic of linear valves is for equal increments of rated travel the flow characteristic (C_v) will change by equal increments.

The gain of a linear valve is linear. The gain of a linear valve remains constant through its full operating range.

Table 4-2 charts how the C_V and gain of a linear valve may vary over its stroke.

Valve Opening	% of Maximum Flow (C_V)	Gain (% Flow / % Open)
0	0	
20	20	1.0
40	40	1.0
60	60	1.0
80	80	1.0
100	100	1.0

Table 4-2

Quick Opening Valves

☞ The inherent flow characteristic of quick opening valves is for maximum flow to be achieved with minimum travel.

The gain of a quick opening valve is nonlinear. The gain of an opening valve decreases from its largest value near closure to its smallest value at full open.

Table 4-3 charts how the C_V and gain of a quick opening valve may vary over its stroke.

Valve Opening	% of Maximum Flow (C_V)	Gain (% Flow / % Open)
0	0	
20	40	2
40	70	1.75
60	90	1.5
80	95	1.1875
100	100	1

Table 4-3

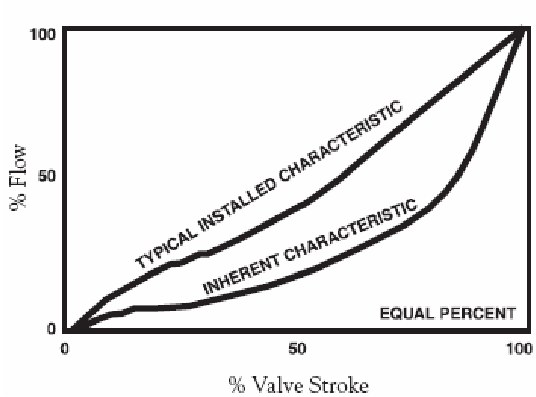
Installed

If the gain of a valve is related to its characteristic, and the gain of a process (as seen by a controller) is a product of all of the gains, and we wish to have a process with a linear gain, why don't we just use valves with inherent linear characteristics?

The answer is that manufacturers test their valves and publish their characteristics under a specified set of conditions, namely a fixed pressure drop across the valve. In real processes conditions are often much different, the pressure drop across the valve is typically influenced by process conditions.

☞ The installed characteristic of a valve is the relationship between the flow rate through the valve and the travel of the closure member as the closure member is moved from the closed position to its rated travel under actual process conditions.

In many process applications the pressure drop across a valve varies with the flow. In these instances an equal percentage valve will act to linearize the process, thus an equal percentage valve would have a linear installed characteristic for these processes. Figure 4-3 shows a linear and equal percentage valve with their installed characteristics in a typical pump installation.



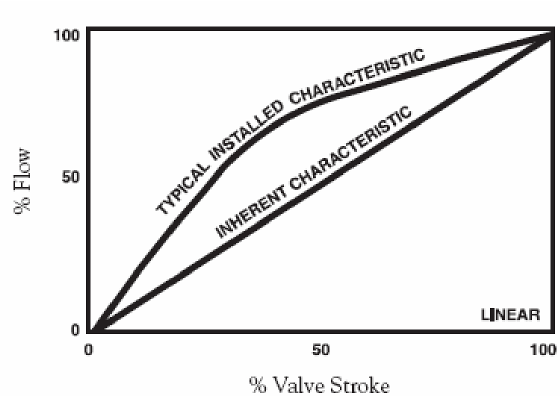


Figure 4-3

How do you know what inherent valve characteristic to choose to get a linear installed characteristic? Most times this selection is through experience, guesswork or the valve manufacturer's recommendation.

The correct selection of valve characteristic to linearize the process gain will ease the tuning process and make for a robust system.

What is Valve Deadband

> ☞ Valve Deadband is the range through which the controller output signal can be varied, upon reversal of direction, without effecting a change in the valve's stem movement.

Valve deadband is a major contributor to process variance. Anytime the controller output reverses direction the valve must travel through its deadband before any corrective action is seen by the controller. Deadband is usually a result of excessive movement in the mechanical linkages of an actuator.

For a valve with a positioner 1% or less is a generally accepted value for allowable deadband, without a positioner this increases to 3%. In practice it may be difficult to measure actual stem movement to 3% of travel (on a valve with 1.5 inches of stem travel this equates to 0.045 inches of offset!), the effects of excessive valve deadband will likely be more apparent in the process.

Figure 4-4 is a reaction curve testing the deadband of a valve. From this curve we can determine that the valve has a 2% deadband, it is not until controller steps of 5% are executed that the valve faithfully follows the direction of the controller output.

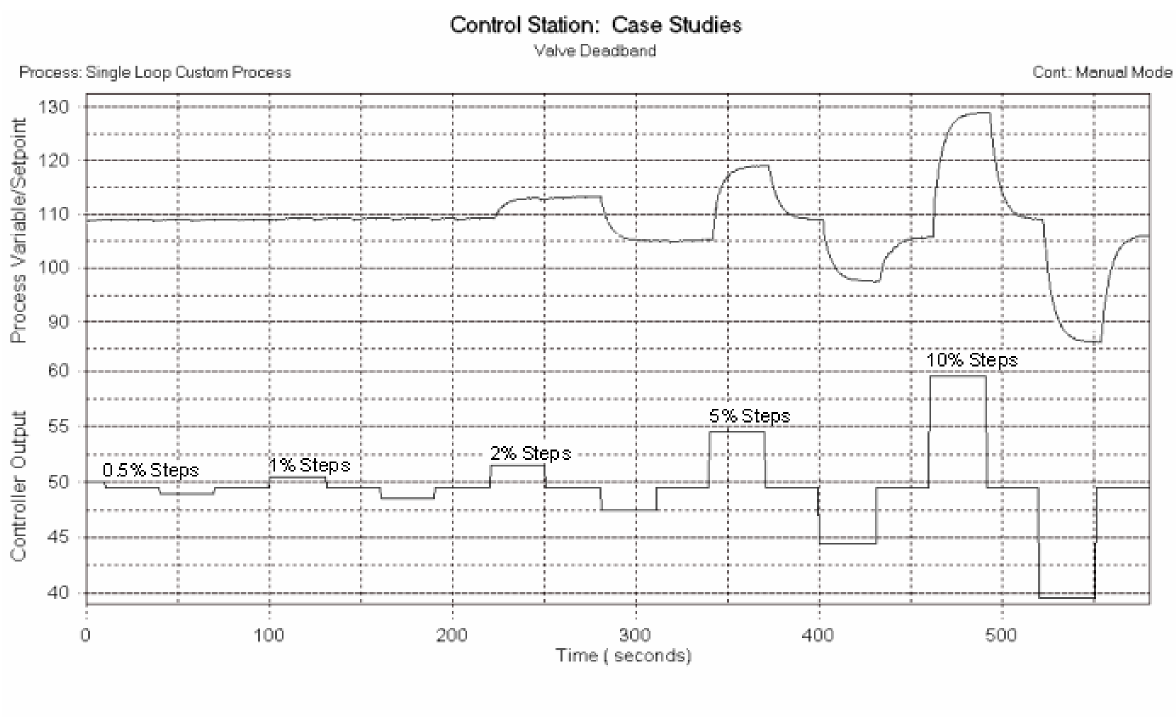


Figure 4-4

Testing for Deadband

The amount of deadband in a valve is tested by placing the control loop in manual and making a series of steps in the controller output, waiting a proscribed amount of time between each step. How much time should you wait between steps? I would recommend 5 time constants plus a dead time. There are two methods of stepping the controller output.

Method A

Method A is illustrated in Figure 4-3 is as follows:

When the process has stabilized execute the following series of steps, waiting 5 time constants plus a dead time between each step.

1. Step the controller output by -0.5%, wait, step the controller output by -0.5% and wait.
2. Step the controller output by +0.5%, wait. Step the controller output by +1.0%, wait.
3. Step the controller output by -1.0%, wait. Step the controller output by -1.0%, wait.
4. Step the controller output by +1.0%, wait. Step the controller output by +2.0%, wait.
5. Step the controller output by -2.0%, wait. Step the controller output by -2.0%, wait.
6. Step the controller output by +2.0%, wait. Step the controller output by +5.0%, wait.
7. Step the controller output by -5.0%, wait. Step the controller output by -5.0%, wait.
8. Step the controller output by +5.0%, wait. Step the controller output by +10.0%, wait.
9. Step the controller output by -10.0%, wait. Step the controller output by -10.0%, wait.
10. Step the controller output by +10.0%, wait.

The amount of deadband is the step size of the step preceding the one where the controller action is followed consistently.

Method B

Method B may be used when you do not want to step the controller through such a wide range of values.

1. Step the controller in one direction such that you know all deadband will be removed and wait.
2. Step the controller in the opposite direction by a small amount (0.5%) and wait.
3. Step the controller again in the same direction by the same amount (0.5%) and wait.
4. Repeat step 3 until the process variable shows a response to the controller step.

The amount of deadband is the change in controller output from step 1 to step 4.

Effects of Deadband

In a self-regulating process excessive valve deadband leads to increased settling time (the deadband must be integrated out), but in an integrating process with integral control excessive valve deadband leads to cycling.

What is Stiction?

☞ Stiction is the combination of the words stick and friction, it occurs when the force required to start valve movement is much greater than the force required to keep the valve moving.

When stiction is present it will keep a valve from moving for small changes in its position command, and then when enough force is applied the actuator overcomes the initial resistance and the valve jumps to a new position. Stiction is often the result of an actuator that is undersized or excessive friction in the valve packing.

Figure 4-5 is a reaction curve testing the stiction of a valve. From this curve we can determine that the valve has 1% of stiction, it is not until the fifth controller step of 0.2% is executed that the valve moves to a new position.

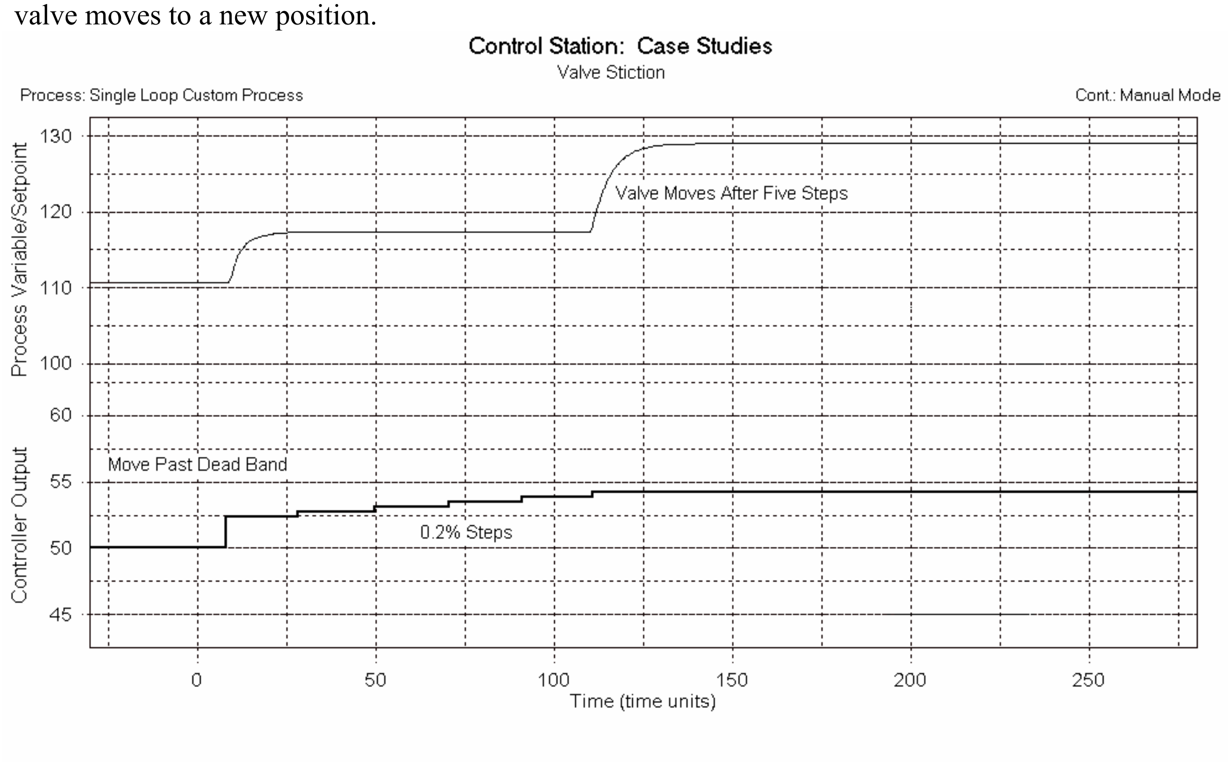


Figure 4-5

Testing for Stiction

The amount of stiction in a valve is tested by placing the control loop in manual and making a series of steps in the controller output, waiting a proscribed amount of time between each step. How much time should you wait between steps? As in testing for deadband I would recommend 5 time constants plus a dead time.

Method

1. Step the controller in one direction such that you know all deadband will be removed and wait for the process to stabilize.
2. Step the controller in the same direction by a small amount (0.2%) and wait.
3. Step the controller again in the same direction by the same amount (0.2%) and wait.
4. Repeat step 3 until the process variable shows a response to the controller step.

The amount of stiction is the change in controller output from step 1 to step 4.

Effects of Stiction

Stiction in a valve leads to increased variability and cycling of the process variability. Stiction is more of a problem than deadband because deadband effects only occur on valve reversal and can be "integrated out". Stiction occurs on all valve movements and cannot be integrated out.

What are the Types of Valves?

We have learned that valves are used to either shut-off flow, change the flow path or throttle flow, and that valves are characterized by how their flow coefficient varies with relation to closure member position, but valves are typed according to the motion of their closure member. The two types of valves are linear motion and rotary motion.

Linear Motion

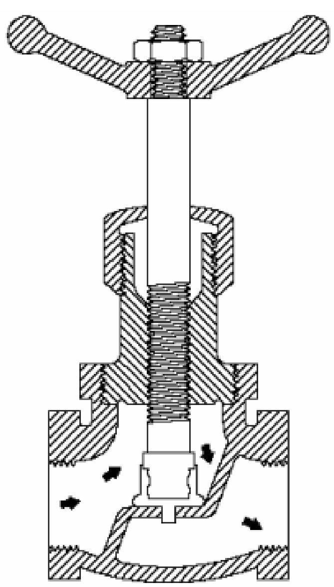
Linear motion valves have a closure member that moves with a linear motion to modify the rate of flow through the valve. Linear motion valves are generally named for the shape of their closure member. Common linear motion valves include globe, gate, diaphragm and pinch valves.

Globe Valve

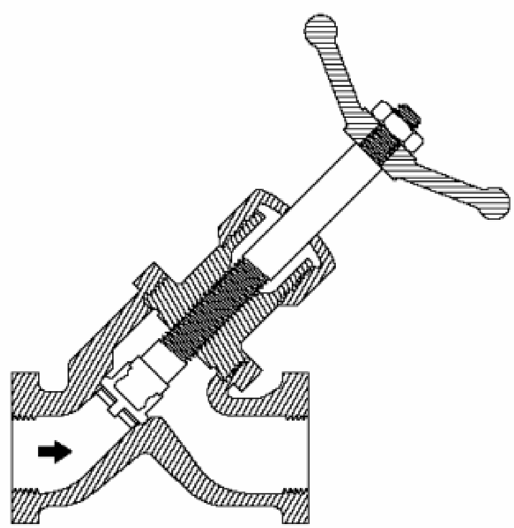
Globe valves are so named for their globular shaped cavity around the valve seat area. The closure member of a globe valve is a plug with a flat or convex bottom that is lowered onto a matching horizontal seat located in the center of the valve. Raising the plug opens the valve, allowing fluid flow.

Globe valves have good throttling characteristics but because the flow path is not linear they have a relatively high pressure drop across the valve. Globe valves are used in throttling and shut-off applications where this pressure drop is acceptable.

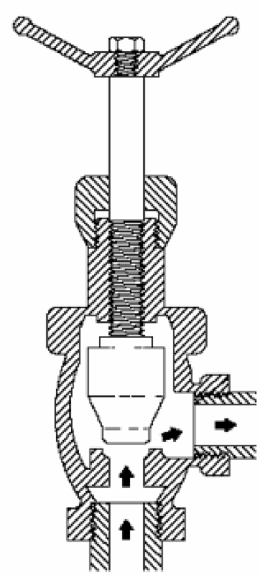
The three primary body designs for globe valves are Z body, Y body and angle.



Z Body Globe Valve

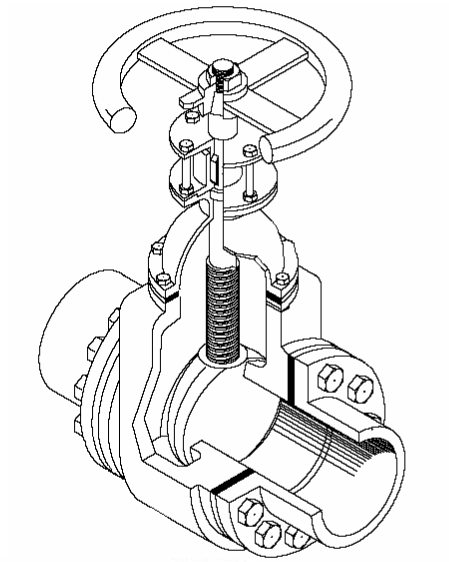


Y Body Globe Valve



Right Angle globe Valve

Gate Valve

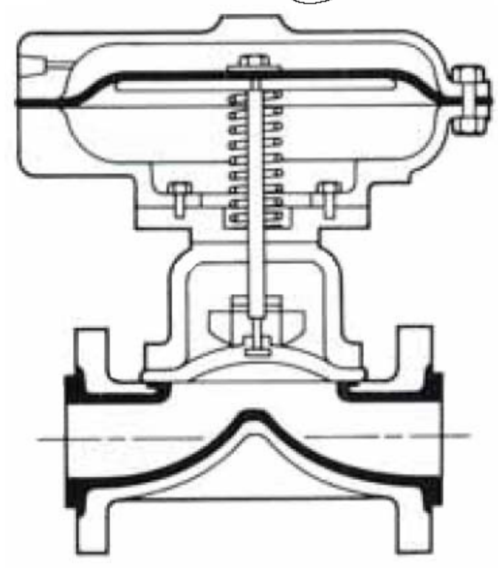


The closure member of a gate valve is a flat face, vertical disc, or gate that slides down through the valve to block the flow.

Gate valves are designed to operate in their fully open or fully closed position and therefore are found only in flow shut-off applications. When fully open the disc is removed completely from the flow stream. This offers virtually no resistance to flow when the valve is fully open, therefore gate valves operate with little pressure drop across the valve.

Gate valves have very poor flow throttling characteristics and are not used for throttling purposes.

Diaphragm Valve



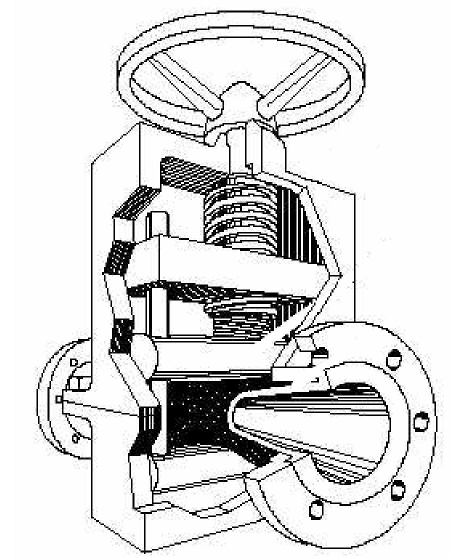
The closure member of a diaphragm valve is a flexible surface (the diaphragm) that is deformed.

The main advantage of a diaphragm valve is that the stem seal is eliminated. Diaphragm valves are used mostly for shut-off service of slurries, corrosive or viscous fluids but may also be used in flow throttling applications as well.

Diaphragm valves may be used in pumping applications with a set constant pressure on the diaphragm. This allows flow to be stopped in the absence of a motive force (pump), but when a sufficient pressure is generated in the pipe to overcome the force on the diaphragm flow is allowed.

Pinch Valve

A pinch valve is similar to a diaphragm valve, however in a pinch valve the entire valve body is flexible and the closure member pinches the valve shut closing off flow. As a pinch valve has no internal obstructions it has a very low pressure drop and is well suited for applications of slurries or liquids with large amounts of suspended solids.



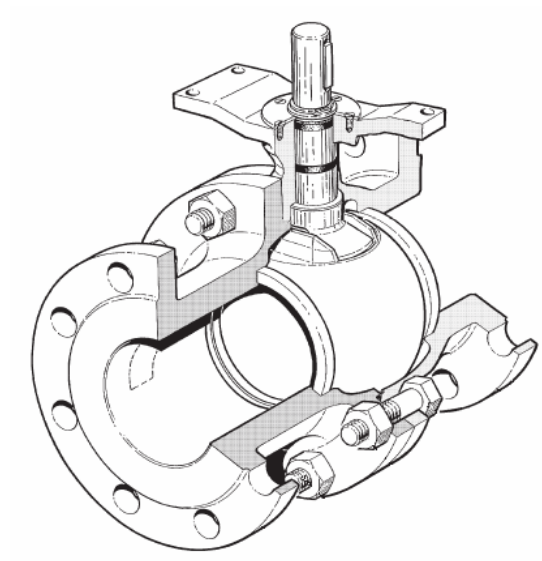
Rotary Motion

Rotary valves have a closure member that moves with a rotary motion to modify the rate of flow through the valve. Like linear motion valves, rotary motion valves are generally named for the shape of their closure member. Common rotary motion valves include ball, butterfly and plug valves.

Ball Valve

The closure member of a ball valve is shaped like a ball with a port for fluid flow.

A ball valve allows straight-through flow in the open position and shuts off flow when the ball is rotated 90 degrees. Because of their quarter turn actuation and low pressure drop ball valves are commonly found in flow shut-off applications. Depending on the particular flow port configuration of the ball they may be used in flow throttling applications as well.

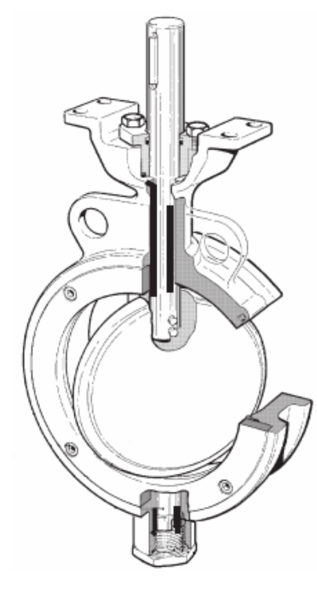


Butterfly Valve

The closure member of a butterfly valve is a circular disc or vane with its pivot axis at right angles to the direction of flow in the pipe.

Like ball valves, a butterfly valve allows straight-through flow in the open position and shuts off flow when the ball is rotated 90 degrees. Because of their quarter turn actuation and low pressure drop butterfly valves are commonly found in flow shut-off applications. Unlike ball valves, butterfly valves are generally not used for flow throttling applications.

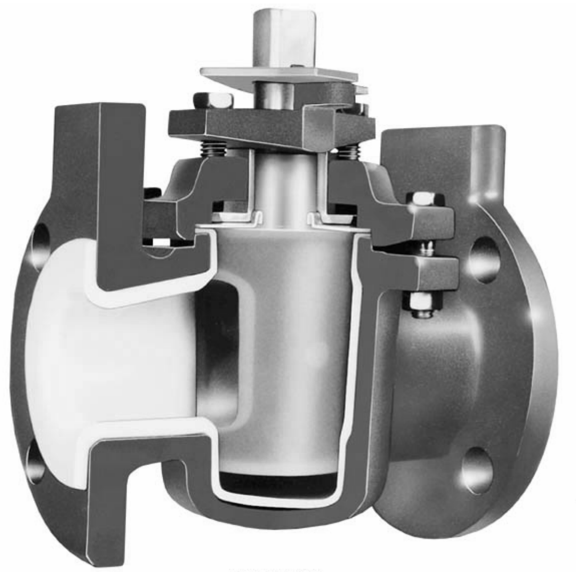
The advantage of a butterfly valve over a ball valve is its relative compactness.



Plug Valve

The closure member of a plug valve is a cylindrical or tapered cylindrical shaped plug with a flow port.

Like a ball valve, a plug valve allows straight-through flow in the open position and shuts off flow when the ball is rotated 90 degrees. Like ball valves, plug valves are found mostly in flow shut-off applications. However plug valves are available in much larger sizes that ball valves.

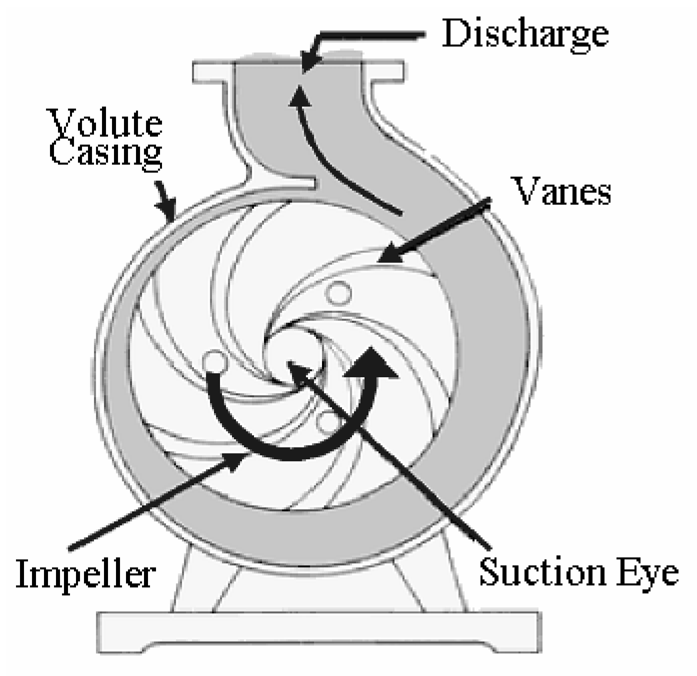


Pumps

Most pumps are used to transport fluid materials from one point to another, and generally they fall into one of two types: Centrifugal and Positive Displacement Pumps.

What is a Centrifugal Pump?

A centrifugal pump is so named because it relies on centrifugal forces to create the kinetic energy in the fluid to impart flow. A centrifugal pump has two main parts, the impeller and the volute casing. Process fluids enter the pump through the suction nozzle and then into the suction eye of the impeller. The fluid is captured between the vanes and spun outward, as liquid leaves the vanes a low pressure area is created in the suction eye causing more fluid enter.



It is the spinning impeller which provides the centrifugal acceleration. The volute casing converts that kinetic energy into pressure energy.

The kinetic energy of a liquid exiting a centrifugal pump is proportional to the velocity of the impeller at its edge. The faster the impeller rotates or the bigger the impeller is in diameter the higher the velocity of the liquid at the impeller edge, and the higher the kinetic energy of the liquid.

As the liquid leaves the impeller the volute casing provide a resistance decelerating the fluid. The fluid further decelerates in the discharge nozzle and the fluid velocity is converted to pressure.

What is Pump Head?

> ☞ Head is a term used to measure the kinetic energy created by a centrifugal pump.

Head is usually expressed in feet and is a measure of how high a column of fluid would go upon exiting a pump with no resistance to flow. Head is approximately related to fluid velocity at the impeller tip by the formula:

$$\text{Pump Head} \approx \frac{v^2}{2g}$$

$$\text{Where}: v = \text{Velocity of the Impeller Tip}$$

$$g = \text{Gravitational Constant } 32\ \frac{\text{ft}}{\text{sec}^2}$$

The velocity of an impeller tip is given by:

$$v = \frac{\text{RPM}\ x\ \text{D}^2}{229}$$

$$\text{Where D} = \text{Impeller Diameter in inches}$$

$$\text{RPM} = \text{Impeller Speed in Revolutions per Minute}$$

Thus

$$\text{Pump Head} \approx \frac{\text{RPM}^2\ x\ \text{Diameter}^2}{229^2\ x\ g}$$

For example, a centrifugal pump with an impeller diameter of 7 inches and an impeller speed of 1800 RPM is capable of generating approximately 94 feet of head.

> ☞ The important thing to remember is the static head a centrifugal pump is capable of generating is a function of its RPM and impeller diameter.

Why Do We Use Head and Not PSI?

If the fluid velocity in a centrifugal pump is converted to pressure at the pump discharge, why don't we use units of pressure, like psi, to measure pump performance?

Pressure is measured in units of force, pounds per square inch. A 32 foot column of water will exert a pressure of 13.85 psi at its base while a 32 foot column of glycerin would exert a pressure of 17.44 psi at its base

And yet a pump that is rated at 32 feet of head will pump both of these fluids 32 feet into the air even though it requires different pressures to do so.

> ☞ A pump with a given impeller size and speed will raise a liquid to a specific height (the head distance) regardless of the weight of the liquid.

Since a centrifugal pump is capable of pumping many different fluids with different specific gravities it is more convenient to discuss the pump's head (and easier for manufacturers to publish performance data).

To convert head to psi the following formula is used:

$$\text{psi} = \frac{\text{head}\; x\; sg}{2.31}$$

$$\text{Where } sg = \text{specific gravity of the fluid}$$

If the specific gravity does not affect the pump head, what does it affect? Even though a pump that generates 32 feet of head will pump both water and glycerin 32 feet in the air, it will require more horsepower to pump the glycerin as it is heavier than water.

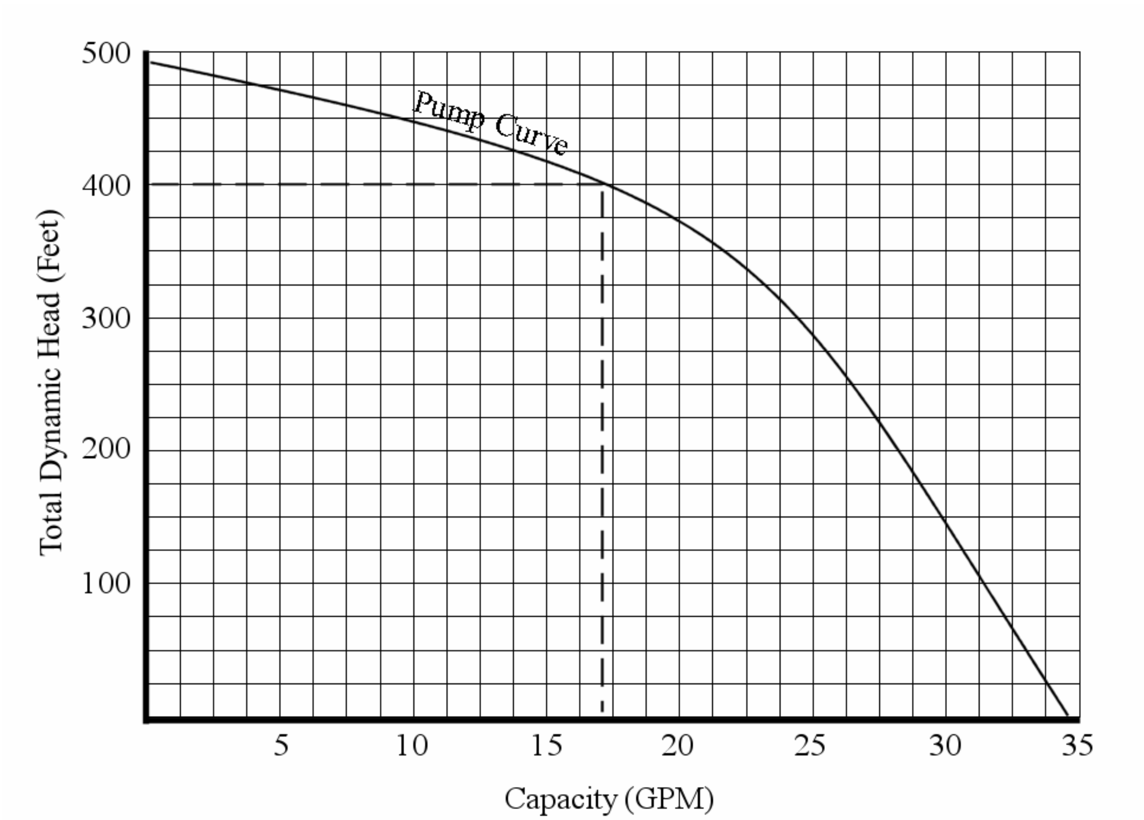
> ☞ Fluids with a specific gravity greater than 1 will require more horsepower to deliver its rated head.

What is a Pump Curve?

☞ A pump curve is a graph that represents a pump's water flow capacity at any given pressure resistance

☞ Capacity is the rate at which the pumped fluid can be delivered to the desired point in the process. Capacity is commonly measured in gallons per minute (gpm) or cubic meters per hour (m^3/hr).

Figure 4-6 depicts a pump curve for a particular impeller size and motor horsepower. The Y axis represents total head in the system and the X axis represents pump capacity. This particular pump configuration will deliver 17 GPM against 400 feet of head, and will always run somewhere along its curve.

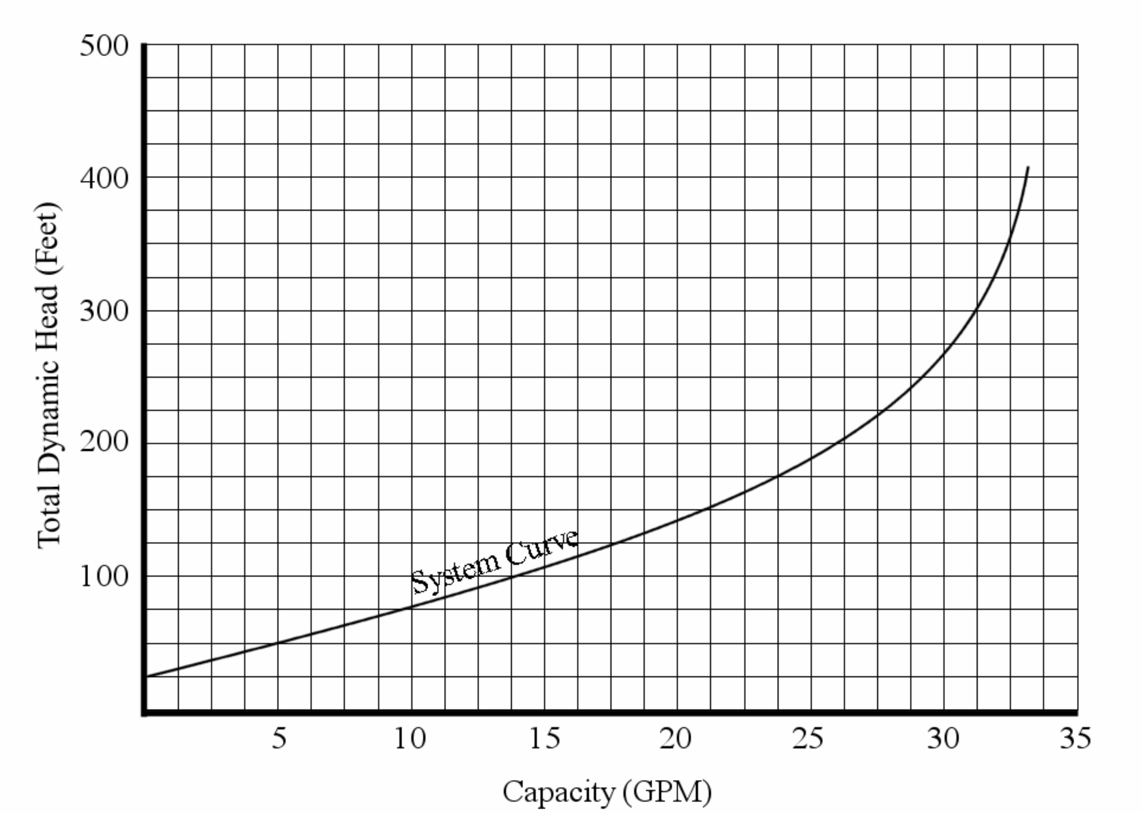


Other data pertaining to pump performance may be shown on a pump curve as well, such as efficiency, braking horsepower and NPSH.

What is a System Curve?

> ☞ A system curve is a graph that represents the pressure resistance of a process system at any given flow.

Figure 4-7 depicts a system curve for a particular process. The X axis represents the flow through the process system and the Y axis represents the resistance the process will generate (expressed in units of head).

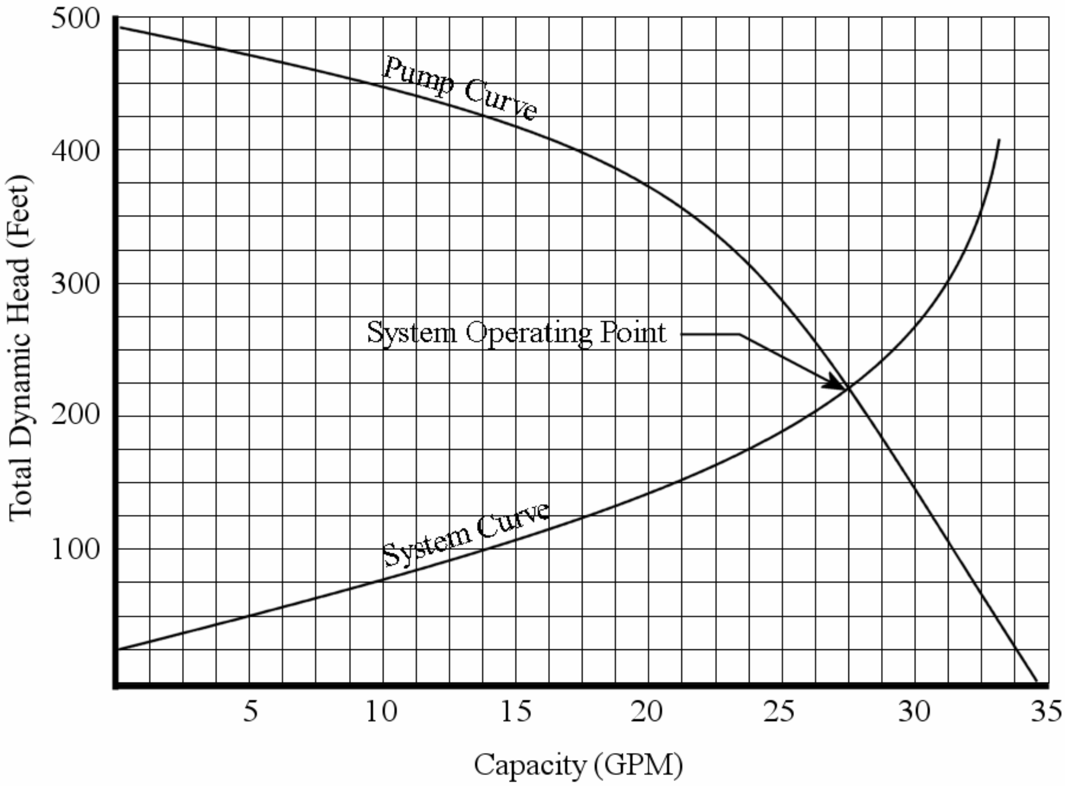


A system curve is influenced by many factors, the amount of suction lift or suction head, the amount of static head, the amount of frictional head, and the pressure differential between the suction side tank and the discharge.

What is the System Operating Point?

☞ Where a centrifugal pump operates on its curve is at the intersection of the system curve and the pump curve, this is the system operating point.

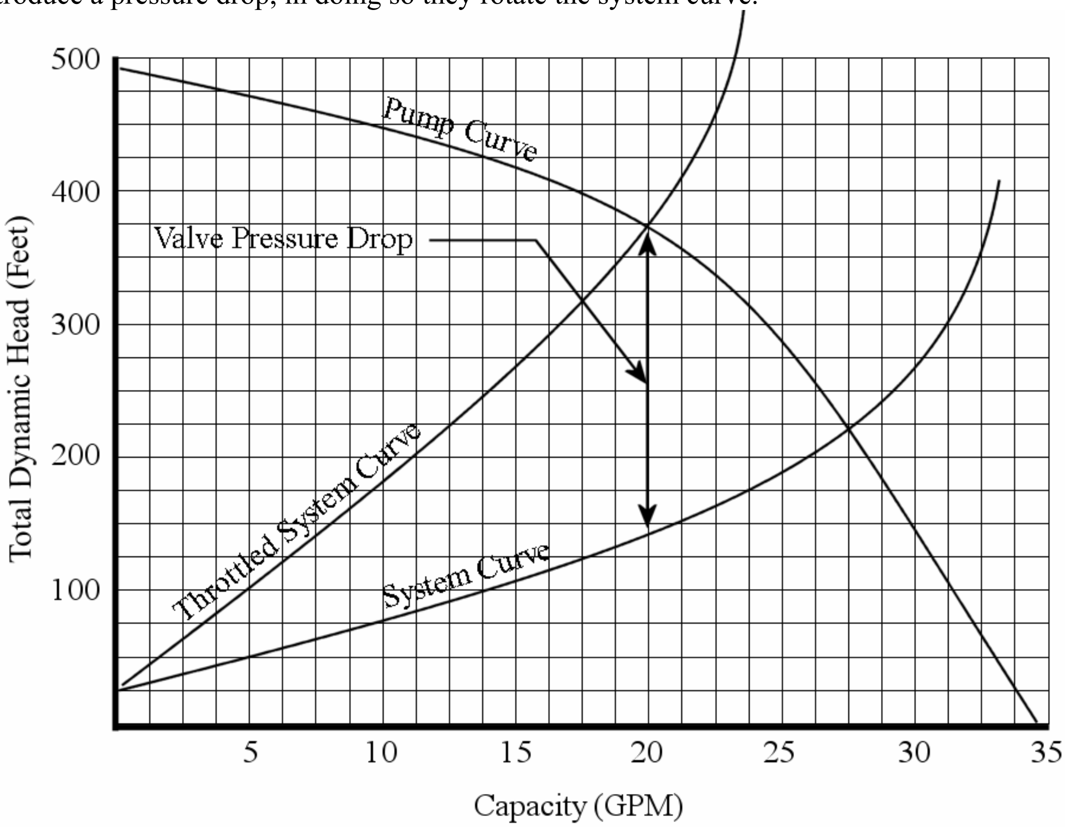
Figure 4-8 shows the pump curve and system curve overlaid on the same graph. The point at which this system will operate is 27.5 GPM with a head of about 225 ft, or 97.4 psi if we were pumping water.



If we actually wanted process to operate at 27.5 GPM all would be well, but if we don't then somehow the system curve or the pump curve must be changed to change the system operating point.

Throttling Valves

The system curve can be changed by using a throttling valve. Throttling valves are designed to introduce a pressure drop, in doing so they rotate the system curve.



In Figure 4-9 we can move the operating point to 20 GPM by introducing an additional 235 feet of head through a throttling valve, or an additional pressure drop of 102 psi (for water).

Variable Frequency Drives

We remember that the head developed by a centrifugal pump is determined by the impeller size and the impeller rotational speed. Changing either one of these will shift the pump curve. Therefore using a variable frequency drive (VFD) is an effective way of controlling the system operating point. There are three properties of the pump curve that are affected by speed: capacity, head and horsepower.

Speed - Capacity Relationship

The capacity of a pump is proportional to speed.

$$Q_2 = Q_1 \, x \left(\frac{N_2}{N_1} \right)$$

$$\text{Where } Q_1 = \text{Initial Capacity (GPM)}, \; Q_2 = \text{Final Capacity (GPM)}$$
$$N_1 = \text{Initial Speed (RPM)}, \quad N_2 = \text{Final Speed (RPM)}$$

For example, if we have a pump with a system operating point of 30 GPM at 1800 RPM, what would the system operating point be at 1170 RPM?

$$Q_2 = 30 \, x \left(\frac{1170}{1800} \right) = 30 \, x \, 0.65 = 19.5 \, \text{GPM}$$

Speed - Head Relationship

The head of a pump is proportional to speed^2.

$$H_2 = H_1 \, x \left(\frac{N_2}{N_1} \right)^2$$

$$\text{Where } H_1 = \text{Initial Head (Feet)}, \quad H_2 = \text{Final Head (Feet)}$$
$$N_1 = \text{Initial Speed (RPM)}, \; N_2 = \text{Final Speed (RPM)}$$

For example, if we have a pump operating at 150 feet of head at 1800 RPM, what would the pump head be at 1170 RPM?

$$H_2 = 150 \, x \left(\frac{1170}{1800} \right)^2 = 150 \, x \, 0.42 = 63 \, \text{feet of head}$$

Speed - Horsepower Relationship

The horsepower required is proportional to speed[3].

$$P_2 = P_1\, x \left(\frac{N_2}{N_1}\right)^3$$

$$\text{Where } P_1 = \text{Initial Horsepower}, \quad H_2 = \text{Final Horsepower}$$

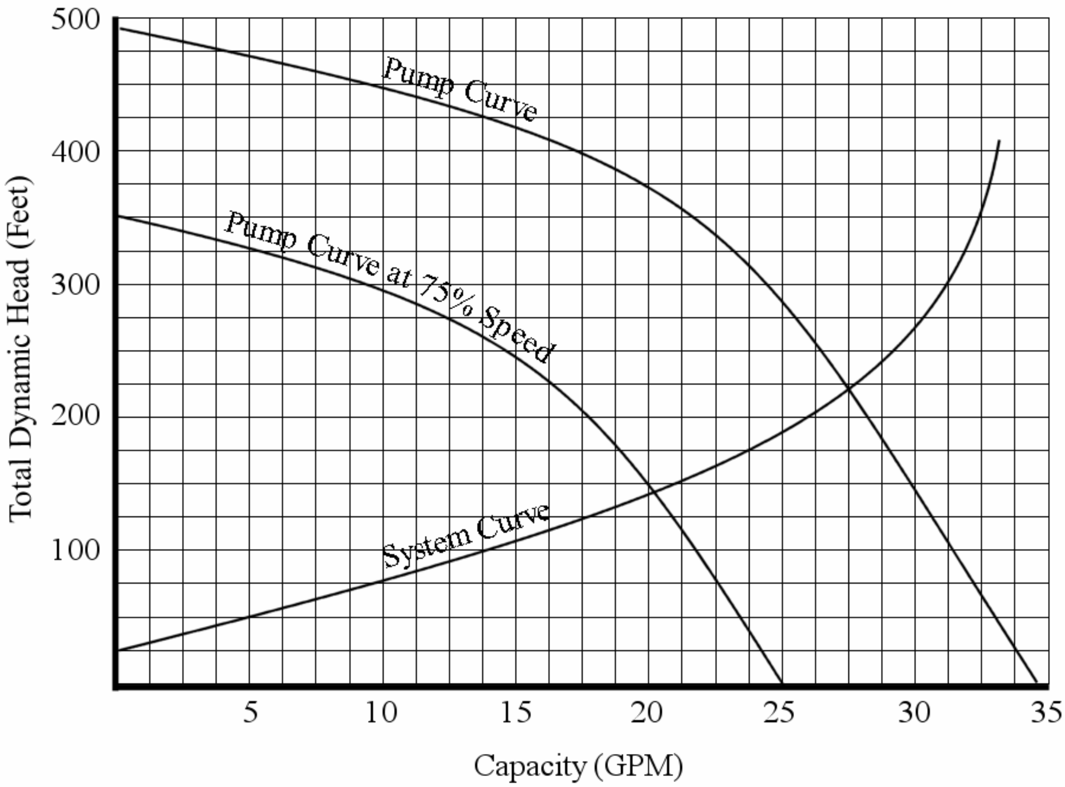
$$N_1 = \text{Initial Speed (RPM)},\ N_2 = \text{Final Speed (RPM)}$$

For example, if we have a pump operating at 2 HP at 1800 RPM, what would the required HP be at 1170 RPM?

$$P_2 = 2\, x \left(\frac{1170}{1800}\right)^3 = 2\, x\, 0.27 = .55\,\text{HP}$$

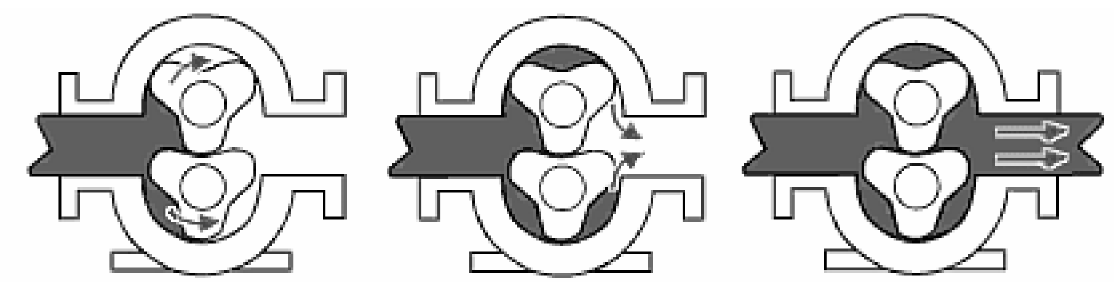
In many instances, process staff are asked to "over speed" pumps in order to extend the operating range of a process. Bear in mind that speeding a motor up from a base frequency of 60 Hz to 90 Hz will require 3.4 times the horsepower the pump was using at 60 Hz!

Figure 4-10 shows the effect of speed control on a centrifugal pump, shifting the pump curve so that a new system operating point is defined.



What is a Positive Displacement Pump?

Positive displacement pumps operate by forcing a fixed volume of fluid form the inlet section of the pump into the discharge section of the pump. Positive displacement pumps come in many forms but all have a common feature, they will produce the same flow at a given RPM no matter what the system head is. A typical pump curve for a positive displacement pump may look like the following diagram, but because of the simplicity of the curve they are usually not given.



How Does a PD Pump Differ From a Centrifugal Pump?

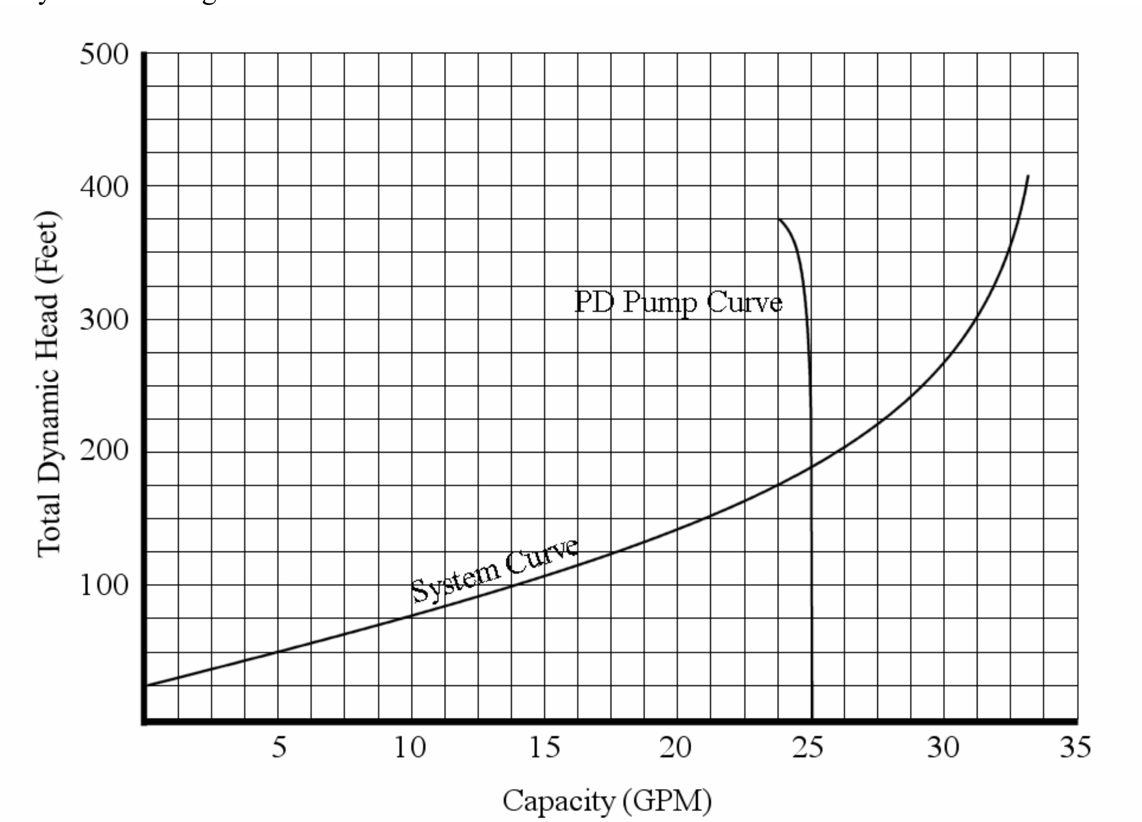
Pump Head

Unlike centrifugal pumps, positive displacements pumps are rated in psi, not head. In a centrifugal pump the maximum amount of head a pump could deliver was a function of its impeller diameter and speed. A positive displacement pump does not have a maximum head. The pump will continue to build pressure until either the motor overloads or a component is damaged. Thus positive displacement pumps are rated in terms of pressure, psi or bar, the manufacturer's rating for maximum safe operating pressure.

Also, we learned that the head of a centrifugal pump was proportional to its impeller speed^2. Changing the speed of a positive displacement pump has little or no effect on the discharge pressure. Discharge pressure is determined by the resistance in the piping.

Pump Curve

Positive displacement pumps do not have published pump curves. If they did they would look very much like figure 4-11.



A positive displacement pump would have a near vertical pump curve. The fall-off in GPM at higher pressures is a result of slip in the lobes. The implication of this is that you cannot change the system operating point by throttling the outlet flow. Doing so will increase the pressure at the pump without affecting flow to any great degree.

Changing the System Operating Point

To change the system operating point of a positive displacement pump requires the use of a variable frequency drive or a recycle loop.

Variable Frequency Drives

As in centrifugal pumps, there are relationships between pump speed and capacity, and pump speed and horsepower. We have already learned there is no relationship between pump speed and head (or discharge pressure) in a positive displacement pump.

Speed - Capacity Relationship

The capacity of a positive displacement pump is proportional to speed. This is identical to the capacity relationship for a centrifugal pump.

$$Q_2 = Q_1 \, x \left(\frac{N_2}{N_1} \right)$$

$$\text{Where } Q_1 = \text{Initial Capacity (GPM)},\ Q_2 = \text{Final Capacity (GPM)}$$

$$N_1 = \text{Initial Speed (RPM)}, \quad N_2 = \text{Final Speed (RPM)}$$

For example, if we have a pump with a system operating point of 30 GPM at 1800 RPM, what would the system operating point be at 1170 RPM?

$$Q_2 = 30 \, x \left(\frac{1170}{1800} \right) = 30 \, x \, 0.65 = 19.5 \, \text{GPM}$$

Speed - Horsepower Relationship

The horsepower required is proportional to speed. This is different form centrifugal pumps where the horsepower was proportional to the speed^3.

$$P_2 = P_1 \, x \left(\frac{N_2}{N_1} \right)$$

$$\text{Where } P_1 = \text{Initial Horsepower}, \quad H_2 = \text{Final Horsepower}$$

$$N_1 = \text{Initial Speed (RPM)},\ N_2 = \text{Final Speed (RPM)}$$

For example, if we have a pump operating at 2 HP at 1800 RPM, what would the required HP be at 1170 RPM?

$$P_2 = 2 \, x \left(\frac{1170}{1800} \right) = 2 \, x \, 0.65 = 1.3 \, \text{HP}$$

The PID Controller

Objectives:

In this chapter you will learn:

- What are the PID Equations?
- What are the Modes of Operation?
- What is Proportional Control?
- What is Integral Control?
- What is Derivative?
- What is Loop Update Time?
- What Combinations of Control Action Can I Use?

The Many Faces of PID

The output of all PID controllers is determined by an equation that has three terms that respond to error, a proportional term, an integral term and a derivative term. Hence the acronym PID controller: Proportional, Integral and Derivative control. Owing to the fact that three terms work on the error to calculate a new controller output, PID control is often referred to as three term control.

The PID controller is the most popular and widely used feedback controller in the process industries. PID has been successfully used for over 50 years providing robust control for many processes. And yet for a controller having just three terms its inner workings appear to be a mystery to many.

Part of the confusion in PID control is than there are no standard definitions for the PID control algorithms embedded in the various controllers, or for the terminology manufactures use. Manufacturers may even use different algorithms for different controllers within a product family. Given this environment, we will review the most generally accepted and commonly found PID algorithms and terminology.

What are the PID Equations?

At the heart any controller is its algorithm, the set of instructions and procedures that determine the controller output for any given input. The algorithm in a PID controller is an equation that calculates the output of the controller based on the error and the values of the three tuning terms. While there is no industry standard for the PID equation or for what they are called, most fit within one of three generally accepted types: Series, Dependent and Independent.

Series

$$CO = K_c\left[(E) + \frac{1}{T_i}\int_0^t (E)dt\right] x \left[1 + T_d\,\frac{d(E)}{dt}\right] + Bias$$

$$\begin{aligned} \text{Where}: CO &= \text{Controller Output} \\ E &= \text{Error} \\ K_c &= \text{Controller Gain} \\ T_i &= \text{Integral Gain} \\ T_d &= \text{Derivative Gain} \end{aligned}$$

This form of the PID equation is most commonly called the series because the derivative term is in series with both the proportional term and the integral term. The original pneumatic and electronic controllers used this form of the PID equation because they were easier to build this way. You may also see this form of the PID equation referred to as the Classical, Real or Interacting form.

Most controller tuning rules are based on this form, however it is not prevalent today with digital control systems (this equation is not an option in Allen Bradley PLC equipment).

Dependent

$$CO = K_c \left[(E) + \frac{1}{T_i} \int_0^t (E) dt + T_d \frac{d(E)}{dt} \right] + Bias$$

$$\begin{aligned} Where: CO &= \text{Controller Output} \\ E &= \text{Error} \\ K_c &= \text{Controller Gain} \\ T_i &= \text{Integral Gain} \\ T_d &= \text{Derivative Gain} \end{aligned}$$

This form of the PID equation is most commonly called the dependent form because the integral and derivative terms are dependent on the controller gain value. If no derivative is used ($T_d = 0$) the dependent and series forms of the PID equation become identical; the same is true if no integral is used ($1/T_i \approx 0$). Therefore the tuning rules developed with the series form are valid for PD and PI controllers using the dependent form of the equation. You may also see this form of the PID equation referred to as the ISA, Ideal or Non-Interacting form.

In the dependent form the integral period is the amount of time in which the controller will output will change by an amount equal to the proportional action.

> ☞ Because the dependent form may be tuned using rules based formulas and the intuitive nature of the integral action the dependent form of the equation is the preferred selection when you have a choice of equations.

Independent

$$CO = K_p(E) + K_i \int_0^t (E)dt + K_d \frac{d(E)}{dt} + Bias$$

$$\begin{aligned} Where: CO &= \text{Controller Output} \\ E &= \text{Error} \\ K_p &= \text{Proportional Gain} \\ K_i &= \text{Integral Gain} \\ K_d &= \text{Derivative Gain} \end{aligned}$$

This form of the PID equation has been called the independent gains form because the integral and derivative terms are independent of the proportional gain. There is no overall controller gain.

> ☞ Allen Bradley offers this form of the equation in their PLCs. The results of tuning rules must be adjusted for use in controllers using the independent form of the equation.

What are the Modes of Operation?

PID controllers have two modes of operation: automatic mode and manual mode. In the automatic mode the controller output is determined by the PID parameters and the type of controller (P, PI or PID).

In manual mode the operator directly manipulates the controller output.

When going from automatic mode to manual mode the transfer should always be bumpless, that is the controller output does not change at the mode transition.

When going from manual to auto the transfer may or may not be bumpless. In a non-bumpless transfer occurs the controller output will start changing to return to the prior Set Point.

In a bumpless transfer the Set Point is set equal to the current process value when the mode transition occurs. This is often a selectable action called PV tracking. When PV tracking is on the transition from manual to automatic will be bumpless, when off the output will immediately start changing to return to the prior Set Point.

☞ PV tracking is an option in the Honeywell UDC configuration and Allen Bradley SLC 5, PLC 5 and ControlLogix processors.

What is Proportional Control?

When the integral and derivative terms are set to zero in the ISA and series equations we are left with:

$$CO = K_c(E) + Bias$$

and in the Dependent form:

$$CO = K_p(E) + Bias$$

In both PID forms the controller's proportional action is identical, $K_c = K_p$.

In both controllers the output is proportional to the error. If the bias value is zero then the controller output is the product of the gain and the error as a percent of the process variable span.

Controller Output = Gain x Error (as a percent of PV span)

If we have zero error, the proportional term will contribute zero to the controller output and the controller output will be zero. Without error a proportional controller does nothing! Proportional action acts in the now, what is the error now?

For this reason a proportional controller will not maintain a Set Point. As the proportional gain in a controller is increased the steady state error will be reduced but the amount of cycling will be increased.

Bias

The bias value is used to remove the steady state error at a particular Set Point. Since a zero error means zero output the bias value is the controller output required to maintain a particular Set Point, proportional action is only used to remove error. Proportional + bias control provides a very robust easy to tune control loop, however, a steady state error will develop when the Set Point is changed or a sustained disturbance is encountered.

Proportional only control is used for fast responding processes that do not require offset free operation (some level, pressure).

Controller Gain, Proportional Gain or Proportion Band

All three of these terms are used to refer to the tuning parameter that adjusts a controller's proportional response to error.

When the term controller gain is used it is generally recognized as referring to K_c, implying that we are working with the series or dependent gain form of the PID equation.

When the term proportional gain is used it is generally recognized as referring to K_p, implying that we are working with the independent form of the PID equation.

Some controllers have a proportion band setting instead of gain while some controllers give you a choice. Proportion band is just another way of expressing the proportional response of a controller to error. Proportion band is commonly defined as the percentage of the range of the measured variable for which a proportional controller will produce a 100% change in its output.

Let's take the heat exchanger with its 110°F Set Point with a proportional controller. In a proportional controller with no bias added to the output, if the PV is 110°F then the controller output is 0%. Lets say we would want the proportional controller to have an output of 100% when the PV is 90°F; let's also say that the temperature input to the controller has a configured range of 0°F to 250°F giving us a span of 250°F. The proportion band for this controller would be (110°F - 90°F) /250°F x 100% = 8%.

In this example we placed the Set Point at one end of the proportion band, often the proportion band is illustrated with the Set Point in the middle giving us a 50% controller output with zero error. The proportion band is shifted relative to the Set Point by adjusting the bias term. Regardless of where the Set Point is placed with respect to the proportion band, the proportion band is error as a percentage of the process variable span over which the controller will vary its output linearly from 0 to 100%.

Back to the heat exchanger, the process variable input span is 250°F and the proportion band is 8%, 0.08 x 250°F = 20°F, a proportional controller will be effective only over a 20°F range, outside of this range the controller output will no longer change.

The relationship between gain and proportion band is given by:

Proportion Band [%] = 100 / K(gain) or K(gain) = 100 / Proportion Band [%].

Our proportion band of 8% is equal to a gain of 12.5 (100/8).

When the error is 8 percent of the process variable input span the controller output will be 100% (12.5 x 8% = 100%).

As we can see, the proportion band and gain values relate to the span of process variable which is a controller configuration. If we were to change the temperature range from 0-250°F to 0-500°F we would need to change the proportion band to 4% to maintain the same control effect. What would be the new gain value? 25!

Changing the process variable configuration of a controller affects the proportional gain setting. If we double the span we have to halve the proportion band, or double the gain.

Controller Action

The gain of a controller can be either positive or negative, just as the gain of a process can be positive or negative. When the gain of a controller is positive ($K_c > 0$) the controller output increases as process variable decreases, thus the controller is reverse acting. When the gain of a controller is negative ($K_c < 0$) the controller is said to be direct acting because the controller output will increase as the process variable increases.

$$CO = K_c(E) + Bias$$

or

$$CO = K_c\left(SP - PV\right) + Bias$$

Some controllers will only accept positive numbers for gain values. In these cases the gain of the controller is selected by choosing the equation to use for error. For these controllers selecting E = SP - PV selects a reverse acting controller, selecting E = PV - SP selects a direct acting controller.

Getting the controller action right is essential. A direct acting controller must be selected for a reverse acting process; a reverse acting controller must be selected for a direct acting process.

Process Nonlinearity

The proportional term of a tuned controller is determined by the process gain. If the process under control is nonlinear then the proportional term would only be good at the particular Set Point that we established it at. If the Set Point were changed to operate where the process had a higher gain the tuning could cause cyclic or even unstable control, and if the Set Point were changed to operate where the process had a smaller gain the tuning could cause a sluggish response. Tuning is often a tradeoff, trading fast response at a particular Set Point for robustness of tuning over a range of Set Points by tuning the process at the operating point with the largest process gain.

What is Integral Control?

When the derivative term is set to zero in the dependent and series equations we are left with the following:

$$CO = K_c(E) + \frac{K_c}{T_i}\int_0^t (E)dt + Bias$$

and in the independent form:

$$CO = K_p(E) + K_i\int_0^t (E)dt + Bias$$

We recognize the first term as the proportional control term and the last term as the bias value. The middle term in both forms is the integral term.

The integral operator is mathematical shorthand for area under a curve. Figure 5-1 demonstrates the principal of integral.

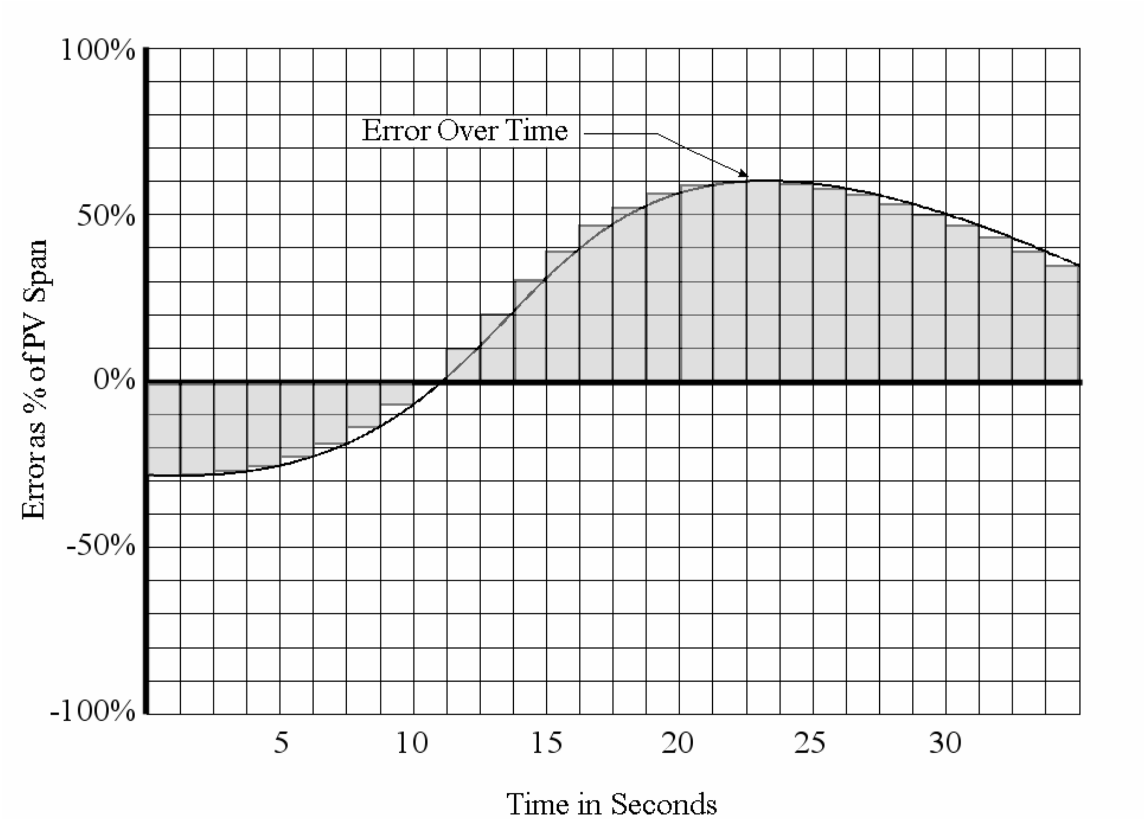


Figure 5-1

In figure 5-1 we have a curve which represents the value of the error (SP – PV) over time. In a digital controller the area under the curve is approximated by a series of rectangles whose height is the value of the error and whose width is the controller's loop update time. The area of the rectangle is the error times the loop update time and the integral is the sum of the areas of the individual rectangles.

Whereas proportional action works in the present (what is the error now), integral action works in the past for integral control action sums the error over all past time and multiplies this summed error value times the tuning parameter.

With integral action, when the error is zero the integral term stops accumulating but still has a value, contrasting to the proportional term which has no value when the error is zero. Integral action eliminates the proportional offset and does not require a bias. However, too much integral action will lead to loop instability and oscillation.

While integral can be used alone for process control, it is normally used with proportional control. PI control used for fast responding processes that require offset free operation.

It is in the integral term that we first see the difference between the dependent and independent forms of the PID equation. The dependent tuning parameter is $\frac{K_c}{T_i}$ while the independent tuning parameter is K_i. The implications of this difference in terms are:

1. To turn off integral action in a dependent form controller we must set T_i to a large value, while to turn off integral action in a dependent controller we set K_i to 0.
2. Tuning rules will help you find a value for T_i. To use the tuning rules for an independent form controller set $K_i = \frac{K_p}{T_i}$.

> ☞ When converting, you must be careful of units. Allen Bradley PLCs express K_i in inverse seconds and T_i in minutes. The actual conversion for Allen Bradley is
>
> $$K_i = \frac{K_p}{60T_i}$$

3. When tuning an independent form controller, if you make an adjustment to K_p you must also make a proportional adjustment to K_i for the controller to behave according to the tuning rules. If you increase K_p by 10% you must increase K_i by 10% as well.
4. Integral only control can only be accomplished with dependent form controller.

Repeats, Integral or Reset?

All three of these terms are used to refer to the tuning parameter that adjusts a controller's time response to error. When the term repeats is used it is generally recognized as referring to $\frac{1}{T_i}$, implying that we are working with the dependent gain form of the PID equation with units of repeats per minute or repeats per second.

When the term integral gain is used it is generally recognized as referring to K_i, implying that we are working with the independent form of the PID equation with units of $\frac{1}{\text{seconds}}$ or $\frac{1}{\text{minutes}}$.

The term reset is often used to refer to the integral tuning parameter without reference to a particular PID equation type or is simply referred to as gain.

Integral Windup

Early PID controllers exhibited a phenomenon known as integral windup. This would occur when the controller output would reach 0% or 100% but a deviation from the Set Point still existed. Under these conditions the integral term will still accumulate even though the controller output was clamped.

If and when the error would change sign the controller would be slow to respond, waiting for the integral term to accumulate back to 0, or unwind. This process became known as integral windup and has been largely eliminated in modern PID controllers by stopping accumulation of the integral term when the controller output becomes saturated (0 or 100%).

However, you may still run across this phenomenon in some Allen Bradley SLC and Honeywell 620 PID implementations, or, to a lesser extent, if your final control element does not affect the process value over the entire range of the controller output (a valve that is oversized). In these cases restrict the controller output to the range over which the final control element is effective or repair/replace the final control element.

What is Derivative?

The derivative term in the dependent form of the PID equation is $K_c\, T_d\, \frac{d(E)}{dt}$

The derivative term in the dependent form of the PID equation is $K_d\, \frac{d(E)}{dt}$

The derivative operator $\frac{d(\ )}{dt}$ is mathematical shorthand for slope of a curve, or the rate at which the points on a curve are changing at a particular point. Figure 5-2 demonstrates the principal of derivative.

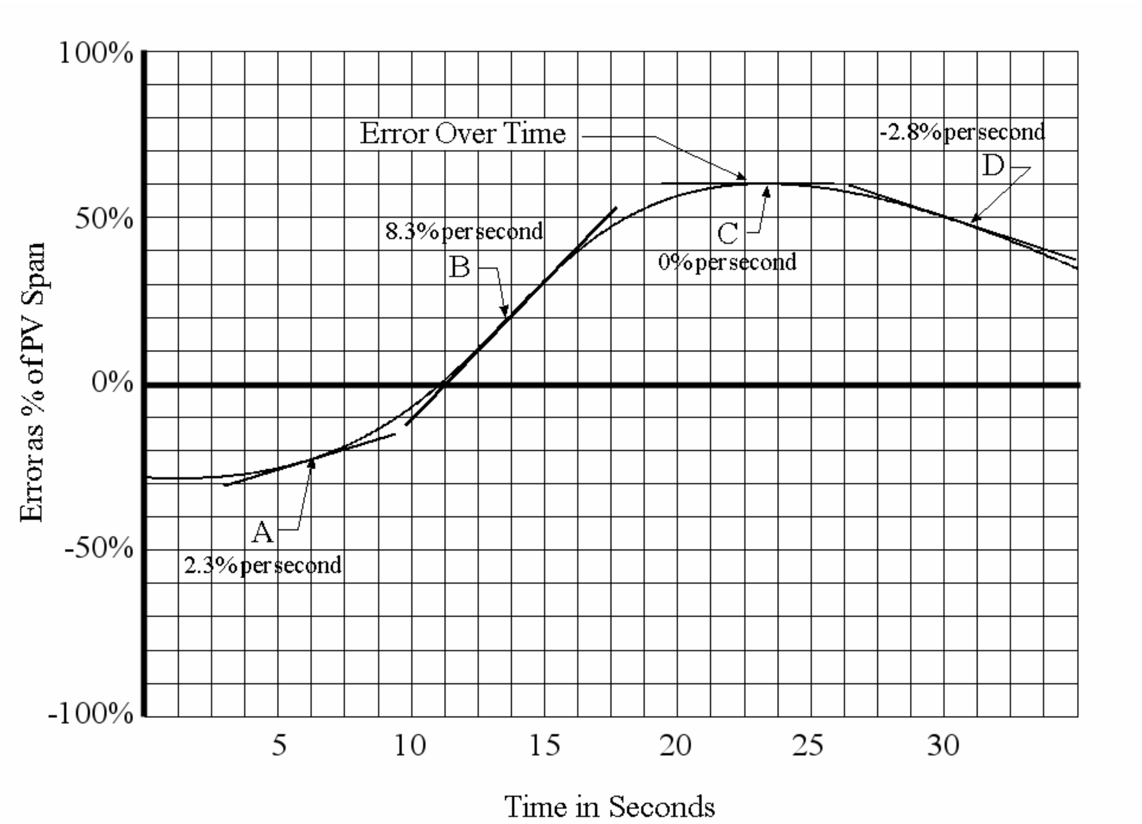


Figure 5-2

In figure 5-2 we have a curve which represents the value of the error (SP – PV) over time. The slope of the curve has been drawn at four points, A, B, C and D. The value of the slope at these points is derivative at these points.

We see at point A the error is changing at a rate of 2.3% per second, at point B it is changing at a rate of 8.3% per second, at point C the error is not changing and at point D the error has changed direction and is changing at -2.8% per second.
In a digital controller the slope of the error curve is approximated by subtracting the previous error value from the current error value and dividing that result by the time between the error points: controller's loop update time.

$$\text{Digital Derivative} = \frac{\text{Error}_{\text{Time } t} - \text{Error}_{\text{Time } t-1}}{\text{Loop Update Time}}$$

We found that proportional action works in the present (what is the error now) and integral action works in the past (what has the error been), derivative action works in the future (where is the error going to be). Derivative works by trying to anticipate the future behavior of the error by looking at its rate of change, the effect is to speed up control response in slow processes and slow down control in fast processes.

Derivative action is never used by itself, it is always used with proportional or proportional plus integral. Derivative action will allow the use of higher controller gains or can smooth integral instability.

As is in the integral term, we see a difference between the dependent and independent forms of the PID equation with respect to derivative action. The dependent tuning parameter is $K_c T_d$ while the independent tuning parameter is K_d. The implications of this difference in terms are:

1. Tuning rules will help you find a value for T_d. To use the tuning rules for an independent form controller set $K_d = K_p T_d$.

> ☞ When converting, you must be careful of units. Allen Bradley PLCs express K_d in seconds and T_d in minutes. The actual conversion for Allen Bradley is $K_d = 60 K_p T_d$.

2. When tuning an independent form controller, if you make an adjustment to K_p you must also make a proportional adjustment to K_d for the controller to behave according to the tuning rules. If you increase K_p by 10% you must increase K_d by 10% as well.

Derivative Filter

When derivative is applied to loops with a noisy process variable signal excessive controller movement often results. This is because derivative is slope and a vertical line has infinite slope, a near vertical line has a very large slope.

Noise in a process signal is often seen as very abrupt changes in the signal of very short duration, that is, a near vertical line. This abrupt change in the process variable due to signal noise is mirrored in the error. The derivative calculation block sees a near vertical change in the error, the slope at this point is very large resulting in a very large derivative contribution to the controller output.

Manufacturers of digital controllers have lessened the PV noise effect by converting the step response seen by the derivative calculation block into a lag response by using a low pass derivative filter characterized by its alpha (α) value.

As in the rest of process control, there are no standards for the implementation of the derivative filter α but in general:

1. If $\alpha = 0$ the derivative term would act as if no filtering were present.
2. If $\alpha = 1$ the controller would function as if there were no derivative tuning.
3. Some controllers may let you set α to a value greater than 1. At values greater than 1 the derivative effects are amplified making the controller response sluggish.
4. For derivative filtering α should have a value between 0.125 and 0.250.

☞ In the Allen Bradley PLC product family the derivative filter is enabled by selecting derivative smoothing on the PID configuration tab but you do not get to select the value for alpha. Allen Bradley determines the value of α by

$$\alpha = \frac{1}{16\left(\frac{dt}{T_d}\right)+1}$$

$$\text{Where } Td = \text{Derivative Gain}$$
$$dt = \text{Loop Update Time}$$

The effect is the larger you make T_d the more aggressive the filter becomes. The ideal values for alpha result when T_d is between two to five times the loop update time.

As you go beyond five loop update times you increasingly filter out the derivative action you are trying to achieve

When T_d is less than two times the loop update times you are increasingly diminishing the filter effect you are trying to achieve.

Derivative Kick

A problem with PID controllers that use the rate of change of the error to determine derivative action is that a Set Point change will be seen as a sudden increase in the error (E = PV - SP). This sudden increase in the error will cause the derivative term to momentarily become very large, adding to the controller output and causing a derivative kick to the final element. This derivative kick is avoided in modern controllers by looking at the rate of change of the process variable, not the error.

The disadvantage to derivative working on the PV is that you cannot tune this controller. Tuning a controller requires introducing a disturbance. A Set Point change is a disturbance you can control, it is much harder (and often not desirable) to introduce a process disturbance during the tuning process.

> ☞ Where you have a choice in the controller configuration it is better to select derivative action on error to tune your controller, then select derivative action on PV when you are done. This will eliminate the derivative kick and yet allow you to test your tuning parameters with a Set Point change.

What is Loop Update Time?

☞ In a digital controller the loop update time is the rate at which the controller solves its PID equation, it is the time that passes between calculations of the PID block.

Establishing the correct loop update time is to a digital controller's performance. In the chapter on instrumentation we learned about the input aliasing phenomena. A digital loop controller will experience the same phenomena if the loop update time is too slow, the snapshot of the error may not

When it comes to setting the loop update time faster is always better. The only downside to a loop update time that is faster than it needs to be is that it consumes processor resources.

☞ The rule of thumb is to set the loop update time 8 to 20 times faster than the loop time constant.

Temperature control is generally a slow process. RTDs in a thermowell have typical time constants of 20 or more seconds. A loop update time of 1 second would be sufficient in this application.

Flow and pressure control are generally fast processes. Instrument time constants for these processes range from 0.5 to 1 seconds. $1/10^{th}$ of a second would be a minimum loop update time for these processes.

☞ Most analog I/O in Allen Bradley control systems communicate with the processor via block transfers. The block transfer trigger time of the analog I/O the provides information for a PID control block must be less than or equal to the PID block trigger time.

What Combinations of Control Action Can I Use?

Proportional Only

Proportional only is inherently stable and simplest to tune. Stable and dynamic response is achieved with minimal effort. Its use is recommended when an offset with a sustained disturbance or Set Point change is allowed, or when used with integrating processes.

Proportional + Derivative

The addition of the derivative term allows for higher proportional gain, giving less Offset than proportional alone. Like proportional only control, use when an offset with a sustained disturbance or Set Point change is allowed, or with integrating processes..

Integral Only

Although rare in practice, an I only controller will operate without offset but the response will be sluggish. Small values of T_i will cause oscillations.

Proportional + Integral

PI control is the most commonly used control. PI control will eliminate offset but introduces instability. PI control will work for most processes.

Full PID

Full PID is the most complicated to tune but can give better performance than a PI controller. The derivative term will allow the use of higher gains without sacrificing stability. Watch for noise in the PV as it will be reflected in the controller output causing excessive movement of the final control element.

Feature	Controller Term	AB SLC 500	AB PLC 5	AB ControlLogix	Honeywell UDC 3000
Proportional	*Controller Gain*	[1]K_C = 0.1 to 25.5	[2]K_c = 0 to 32,767	K_c = Floating Point	K_c = 0.1 to 9999
	Proportional Gain		[2]K_p = 0 to 32,767	K_p = Floating Point	
	Proportion Band				PB = 0.1% to 999.9%
Integral	*Repeats*	1/T_i, [1]T_i = 0.1 to 25.5 minutes / repeat	1/T_i [2]T_i = 0 to 32,767 minutes / repeat x 100		RSET MIN = 0.00 to 50.00 minutes / repeat RSET RPM = 0.00 to 50.00 repeats / minute
	Integral Gain		[3]K_i = 0 to 32,767 1/seconds x 1000		0.06 – 60 0.001 – 1
Derivative	*Derivative Time*	[1]T_d = 0.01 to 2.55 minutes	[2]T_d = 0 to 32767	T_d	RATE MIN = 0.08 to 10.0 minutes
	Derivative Rate		[2]K_d = 0 to 32767		
	Derivative Filter		Yes/No		
Bias		Yes	Yes	Yes	MAN RSET = -100 to +100%
Controller Action	*Reverse Acting*	E = SP - PV	E = SP - PV	E = SP - PV	
	Direct Acting	E = PV - SP	E = PV - SP	E = PV - SP	
Update Time		0.01 to 10.23 seconds	[2]1 to 32,767 seconds x 100		

Notes: [1] 0 to 3276.7 in SLC 5/03 and higher processors when the RG bit is set.
[2] Divided by 100 for calculations
[3] Divided by 1000 for calculations

Fundamentals of Loop Tuning

Objectives:

In this chapter you will learn:

- How Do You Tune a Dead Time Dominant Process?
- How Do You Tune a Cascade Loop?
- How Do You Tune by Trial and Error?

Introduction

What is the Goal of Tuning?

Although this question may seem trivial, how you answer it may well determine your approach to tuning. The obvious answer to the question is to maintain a Set Point, however there are many other things to consider.
Ideally, a properly tuned control loop will:

- Operate within safe constraints of the process.
- Maximize operating profit.
- Eliminate offset from Set Point.
- Be stable over the normal operating range.
- Avoid excessive control action (not overstress the final control element)

How you address each of these items will, in some part, determine your tuning approach and the desired outcome.

Operate Within Safe Constraints of the Process

For any control loop, part of the tuning process is establishing safe operating parameters and ingraining them into the loop control.
One item to address is limits on the controller output. If you are controlling the speed of a pump that is only rated for a 5:1 speed turndown, place a lower output limit of 20% on the controller to avoid overheating the pump motor. If you are controlling temperature use high temperature alarms to place your final control element in a safe position on process over-temperature.

Maximize Operating Profit

What this really means is spend your time tuning those loops that really matter. Temperature control on a tempering machine is critical to maintaining product quality, temperature control on potable hot water for cleaning is not.
We also have many processes that undergo start-up sequences where the Set Point for a process variable may change during the sequence. Controller response at the intermediate steps should provide stability, but spend your time optimizing performance under production conditions.

Eliminate offset from Set Point.

This goes along with maximize operating profit. In general, process temperatures are more critical then levels. Spend your time tuning process temperature loops to maintain Set Point. Most level loops will operate satisfactorily with proportional control alone and a Set Point offset.

Be stable over the normal operating range.

Control loops provide error correction due to process disturbances and Set Point changes. Most processes we encounter are nonlinear. The process variable range over which you choose to tune will determine how aggressive your tuning may be.
A process that requires small disturbance rejection can be tuned more tightly than a process that requires operating over a range of Set Point. For stability (robustness) you must tune your process at the point in the operating range that it has its highest gain and accept sluggish response where the gain is lower.

Tuning is in many cases a tradeoff between performance and reliability.

Avoid excessive control action (not overstress the final control element)

Some tuning methods are designed to give a quarter wave decay response, meaning the final control element will reverse directions several times before Set Point is achieved. Each reversal of a final control element can cause stress, wear and tear, wearing out your final control element prematurely. And in the case of valve deadband, can actually increase the time it takes to achieve Set Point. Sometimes the best tuning may be achieved with a damped response with no overshoot.
Also beware of derivative action as it will reflect noise in the process variable signal to the controller output, resulting in excessive controller movement.

The Approach

Once you have decided what your tuning goals are, what your destination should look like, you must choose your approach. There are basically three paths to your destination: rules based, trial and error and software assisted. Each path will get you there, the difference may be in the time required to achieve acceptable tuning parameters or your path may be chosen by the type of process data you have available.
But remember, there is no methodology that will give you the precise tuning parameters for your process. Also your process has inherent dynamic characteristics, you can not make a slow process respond faster or eliminate overshoot while maintaining rapid response by changing tuning parameters.

How Do You Tune by Trial and Error?

Eventually you will be in the position of having to tune a loop without the assistance of a reaction curve or a software package, and the question will be: What do I do now? Hopefully these cases are limited to non-critical loops in your facility, there is no trial and error method that will give you fast optimized results.
When in these situations there are two methods you can use, start with proportional or start with integral.

Trial and Error, Proportional First

By now you know that all controllable processes can be controlled by proportional action alone. Proportional control is stable and robust, the downside is a Set Point offset.
You also know by now that a process variable will behave in a predictable way to increasing proportional action. A self-regulating that is over-damped will progress through critically damped to under-damped with increasing proportional gain.
What this means is that you can dial in the shape of your final response by starting with proportional action. When your proportional gain is high enough to affect an under-damped response in the process variable, there is no amount of integral gain that can be added to remove the oscillation. The general method to tune in this manner is:

1. Place the controller in manual and allow the process to stabilize. Turn off integral and derivative action.
2. Set the proportional gain to a small value and place the controller in automatic.
3. Change the Set Point and watch the response of your process variable. Is it over-damped or under-damped? If over-damped increase your gain, if under-damped decrease your gain.
4. Repeat step 3 until you have a crisp response in your process variable with no or little overshoot.
5. Check your controller output. If there is excessive movement in the output reduce your gain until it is eliminated or use PV filtering.
6. Turn on integral action with a relative large value for T_i (or small value for K_i).
7. Change the Set Point and watch the response of your process variable. If you have oscillations increase T_i (decrease K_i). If no oscillation is present decrease T_i (increase K_i).
8. Repeat step 7 until you have the desired response.
9. If you must use derivative start at $T_d = T_i \div 8$. If PV noise is reflected in the output reduce T_i or enable derivative filtering.

Trial and Error, Integral First

According to the IMC tuning rules, the reset time is set to the process time constant. The question here is, is there an easy way to get to the time constant without a graph?

We remember that the time constant is a characterization of a first order process. After one time constant a first order process will reach 63.2 percent of its final value. Another property of the time constant is that after four time constants a process will reach 98 percent of its final value. An approximation of the time constant can be obtained by doing a step test and diving the time it takes for the process to reach its final value by 4.

For those that want to go the extra step, the process gain could be calculated as well from the step and a modification of IMC tuning model equations applied (omit dead time if it is insignificant or measure it as well).

The general method to tune in this manner is:

1. Place the controller in manual and allow the process to stabilize. Record the value of the process variable at this point.
2. Change the output by a small amount and let the process stabilize. Record the value of the process variable at this point, this is how we will know the step test is complete.
3. Bring the process variable back to the value it had in step one. If this is not close to the original controller output value then your final control element has an excessive deadband.
4. Change the controller output by the same amount as in step two. Using a stopwatch record the time it takes for the process variable to reach the value it had in step two.
5. Divide this time by 4, this will be your time constant. Enter this value in your controller for T_i.

> ☞ Be careful of units! If Ti is in minutes make sure you record the time from the stopwatch in minutes.

6. Set the proportional gain to a small value and place the controller in automatic.
7. Change the Set Point and watch the response of your process variable. Is it over-damped or under-damped? If over-damped increase your gain, if under-damped decrease your gain.
8. Repeat step 7 until you have a crisp response in your process variable with no or little overshoot.
9. If you must use derivative start at $T_d = T_i \div 8$. If PV noise is reflected in the output reduce T_i or enable derivative filtering.

Rules of Thumb

1. The controller action must be opposite the process action.
2. If the oscillation goes away when the controller is placed in manual, the loop is the cause of the oscillation.
3. If you halve the span of a process variable or double the gain of your final control element you must halve the controller gain to get the same controller action.
    a. A corollary to this rule concerns the independent gain PID equation. If you halve the proportional gain of the controller you must halve the integral and derivative gain as well.
4. Tune for a damped response when to minimize valve wear.
5. When using derivative, use a derivative filter to minimize the noise that is reflected in the output. Try α values between 0.1 and 0.25.
6. If your controller has high gains use a filter on the process variable to minimize noise being reflected in the controller output.
7. For temperature loops set the PID block update time to 1.0 seconds or less. For pressure and flow loops set the PID block update time to 0.1 seconds or less.
8. For analog I/O that requires block transfers set the transfer rate equal to the PID block update time.

Good Practice and Troubleshooting

Common Control Loops

Flow Control

Because the flow sensor and the process are typically fast, the dynamics of the control loop is governed by the dynamics of the control valve. Almost always use a PI controller.

Level Control

In a level loop the dynamics of the sensor and actuator are fast compared to the process. Use P only control unless controlling to a Set Point is desired, then use PI with a small amount of integral.

Pressure Control

In a pressure loop the dynamics of the sensor and process are fast compared to the actuator. Use P only control unless controlling to a Set Point is desired, then use PI.

Temperature Control

In a temperature loop the dynamic of the actuator is usually faster than the dynamics of the process and sensor. Use a PI controller, if the dead time is sufficiently large use PID.

Troubleshooting

Check each subsystem separately

- Final Control Elements
- Sensors
- The Controller
- The Process

Final Control Elements

Common Valve Problems

- Check the regulator settings supplying instrument air to the valve actuator. Look for pinched air lines and air leaks. If the proper air supply is not being supplied to the actuator good control will be difficult to achieve. With today's emphasis on air instrument air conservation and the lowering of plant wide air pressures some actuators may need to be upsized to provide sufficient force to operate the valve.
- Linkages are often used to connect the actuator to the valve stem. Loose linkages can result in process cycling. If a positioner is used also check the linkage used to connect the positioner feedback.
- Check that the actuator is properly calibrated. Calibration involves stroking the valve through its full travel to verify that the valve position corresponds with the controller output. To calibrate a valve:

    1. Place the loop in manual with a 0% output. Adjust the valve zero until the valve is at its full de-energized position.
    2. Set the manual output at 100%. Adjust the valve span until the valve is at its full energized position.
    3. Set the manual output at 50%. Verify that the valve is at its 50% position.

    If a loop has been tuned with an improperly calibrated valve, recalibration may change the process gain requiring retuning of the control loop.
- Check the valve deadband. Excessive deadband will cause an integrating process to oscillate and increase the stabilization time of a self-regulating process.
- Check for valve stiction. Stiction in a valve will cause oscillations in a self-regulating process.
- Check the gain of the valve. An oversized valve will magnify deadband and stiction problems.
- Check the tuning of the positioner. An aggressively tuned positioner can cause valve cycling.

Sensors

Common sensor Problems

Smart Transmitters

- Check the configuration. Make sure the variable being transmitted is the one you want.
- Check the calibration, do the units and range match the controller's units and range configuration for the process variable.
- Check the filtering. Is an excessive dampening time (time constant) specified.

Temperature Sensors

- Check the calibration, does the thermocouple type or RTD curve match the configuration of the controller.
- Check the installation, is the thermowell or sensing element sufficiently immersed in the medium. Is the thermocouple or RTD properly fitted for the thermowell. Is the temperature sensor in the right place. Is there buildup on the thermowell.

Pressure Sensors

- Check for plugged lines to the sensor.

Flow Sensors

- Verify that the mounting conforms to the sensor manufacturer's recommendations (e.g. straight runs of piping, environment, etc.)

The Controller

Common Controller Problems

- Does the controller have a high gain? Filtering on the process variable can help.
- Does the rate at which the PID block is being triggered match the loop update time.
- Has the controller output or input become saturated.

The Process

Common Process Problems

- What's changed? Ask the operator what is different today.
- Understand the Process.
- Is the process is non-linear, a change in process conditions may require different tuning.
- Set output limits to avoid integral windup or operating in low gain regions of your final control element.