PID tuning rules—Appendices

By Gregory K. McMillan

Since the PID controller is a key part of almost every control loop and has significantly untapped capability, the following appendices are offered to help users get the most out of their PID controllers. The appendices also offer an insight and understanding that are useful beyond the tuning of controllers. The appendices are designed to help users see through the considerable complexity presented in the literature and get to the essence of process control opportunities feasting on the knowledge gained. Bon Appétit.

Appendix A offers a short cut tuning method enabling a reduction in open loop test time of over 80% for process with large process time constants, such as temperature loops. The resulting dynamics identified have extensive uses beyond tuning. The estimated open loop gain, time constant, and deadtime can be used to create tieback models with high fidelity dynamics for prototyping innovative control strategies, exploring PID options, trying simple solutions for optimization and coordination of loops mentioned in the article, and operator training. For slow loops (i.e., temperature), the model time constant and deadtime can be made faster by the same factor without drastically changing tuning and performance. The limit as to how fast the model dynamics can be speeded up is the module execution time. The total loop deadtime after speed up should be greater than 10 times the module execution time to keep the increase in integrated error less than 10%.

PID flexibility is enhanced by an extensive set of features to provide the ability to handle a wide spectrum of industrial applications and opportunities. However, many of the parameters for these features are not fully understood or optimally adjusted. Appendix B offers a PID checklist to provide application guidance for the setting of frequently used and often neglected parameters. The effect of each parameter is concisely summarized.

The objective of Appendix C is to sort out fact from fiction by deriving the equations that clearly show the effect of dynamics and tuning on loop performance. The greatest value of the equations is their simplicity enabling the fundamental understanding of key relationships. The equations provide a guide as to how to change plant dynamics and PID tuning to achieve specific performance objectives. The PID structure choices and the functional role of external reset feedback (dynamic reset limit) are illustrated. This appendix is an excerpt from *PID Control in the Third Millennium: Lessons Learned and New Approaches,* Editors Ramon Vilanova and Antonio Visioli, Springer 2012.

Appendix D offers an overview and perspective of the three major types of process response, including the positive feedback process not commonly discussed in the literature. The elegant conciseness of the short cut method is seen in the responses and block diagram.

Appendix E shows how major tuning methods converge to the same expression for the controller gain. The appendix shows how the short cut method, thought to be mostly applicable to near-integrating or integrating processes, provides the correct PID gain for deadtime dominant processes if a deadtime block is used to compute the ramp rate.

Appendix A – Short cut tuning method

Auto and adaptive tuning is recommended for all loops. The short cut tuning method can be used as a quick check of dynamics and tuning settings in about five dead times. The reduction in test time afforded by this method is particularly dramatic for processes with a large process time constant (continuous temperature and composition control of vessels and columns). The dynamics identified can be used to provide plant-wide experimental models or adapt parameters in first principle models per Appendix F in the ISA book *Advanced Temperature Measurement and Control*, and tune controllers via the generalized approach discussed in the *InTech* article "PID tuning rules."

The changes in controller mode and output listed in the method should be reviewed by operations and done by a control room operator. The final control element deadband noted in the method can be estimated as 0.2% for a sliding stem valve (globe) and 0.4% for a good rotary valve (ball, plug, or butterfly) with a digital positioner. For a variable frequency drive (VFD), the deadband in the setup of the drive for speed control is used. The method can be automated in a DCS module to perform the calculations whenever the change in controller output in a short time frame is large enough. The trend chart used to visually identify dynamics should have a time scale of about 20 dead times and a process variable scale of about 50 times the noise. Consider the minimum noise to be about 0.05% of full scale.

A ball park estimate of total loop dead time before any test is the sum of ½ module execution time, valve pre-stroke dead time (e.g., 0.2 sec), transmitter damping setting (e.g., 0.2 sec), and the process dead time. The process dead time for level includes the effect of sensor sensitivity limit and noise.

Process type	Process dead time	Process time constant
Pipeline liquid pressure	0.1 sec	0.1 sec
Pipeline gas pressure	0.5 sec	10 sec
Vessel pressure	1.0 sec	50 sec
Column pressure	5.0 sec	20 sec
Liquid flow	0.2 sec	0.5 sec
Vessel level	50 sec	0.0001/sec (integrating gain)
Column level	10 sec	0.001/sec (integrating gain)
Vessel temperature	2 min	100 min
Column temperature	40 min	200 min

Vessel composition	1 min	50 min
Column composition	40 min	200 min

For at line analyzers, you need to add the sample transportation delay, 1.5 times the analysis cycle time, and the multiplex time to the total loop dead time. The module execution time should be less than 2% of the total loop dead time so the increase in integrated error is less than 10%.

Short cut tuning method procedure:

1. On the trend chart, find the measurement noise amplitude by momentarily putting the PID in manual and estimating the amplitude of deviations over a period about equal to the dead time.
2. With the PID in auto, add a process variable (PV) filter to keep the fluctuations in the PID output within the final control element deadband.
3. On the trend chart, find the largest positive and negative deviation of the PID output that compensates for process disturbances. Propose to operations a test step size depending on direction that is comparable in size to the largest deviation and at least five times larger than the final control element deadband. The objective is the largest test step size that is not too disruptive to process conditions. Explain to operations that the PID will be returned to automatic after a time period that corresponds to about five dead times (state actual estimated time rather than using the term "dead time"). Choose for the first test a step in the least disruptive and safest direction.
4. Put the PID manual and make the step change in the PID output with the agreed upon size and direction. Put the PID back in last mode (e.g., Auto) after about five dead times.
5. Note the time from the change in manual output to a change in the PV output of the noise band as the dead time (θ_o).
6. Note the maximum % change in the process variable ($\Delta\%PV_{max}$) in one dead time interval. Look for the maximum change over at least four dead times. Divide this maximum % change in one dead time by the dead time to get the ramp rate and then by the % change in controller output ($\Delta\%CO$) to get the "pseudo integrator" or "near integrator" gain ($K_c = \Delta\%PV_{max}/\theta_o/\Delta\%CO$).
7. If further testing is permitted, do a step in the opposite direction with a step size that is the sum of the allowed positive and negative step sizes using the procedure listed in items 4-6. If further testing is permitted, repeat this step in the opposite direction, and then do a step to get back to the normal operating point (initial controller output before tests).
8. If the complete change in the PV occurs within the first dead time interval after the start of the response, the process is dead time dominant. For temperature loops, if the PV change in the first dead time interval is 5% to 20% of the maximum PV change in subsequent dead time intervals, a secondary time constant can be considered to exist that is approximately equal to the dead time ($\tau_2 = 0.5*\theta_o$), otherwise the secondary time constant can be considered to be negligible ($\tau_2 = 0$). Note that the primary time constant (τ_1) also known as the open loop time constant (τ_o) for a "near-integrating" or "pseudo-integrating" process can be estimated by Equation A-6.

If the first step is positive (S_1) and the second step is negative ($-S_2$), Step 1 is S_1, Step 2 is $-S_1 - S_2$, Step 3 is $S_1 + S_2$, and Step 4 is $-S_2$ (back to original controller output).

The PID controller gain (K_c), integral time (T_i), and derivative time (T_d), can be estimated as:

$$K_c = \frac{0.5}{\theta_o * K_i} \quad \text{(A-1)}$$

To avoid slow oscillations in an integrating process from too much reset action:

$$T_i = \frac{2}{K_c * K_i} \quad \text{(A-2)}$$

If we substitute Equation A-1 into A-2, we have an integral time useful for 99.99% of the loops:

$$T_i = 4 * \theta_o \quad \text{(A-3)}$$

If the process variable can develop a runaway acceleration from positive feedback, such as temperature control of highly exothermic polymerization reactions, or if the controller gain is detuned by a factor of ten for an integrating process (a common occurrence), the integral time is increased by a factor of 10:

$$T_i = 40 * \theta_o \quad \text{(A-4)}$$

If the loop is clearly dead time dominant ($\theta_o >> \tau_1$), such as sheet thickness control by manipulation of die bolt position, the integral time is reduced by a factor of 10 (rare case):

$$T_i = 0.4 * \theta_o \quad \text{(A-5)}$$

For temperature control in vessels and columns where there is a secondary lag from heat transfer surfaces or stages ($\tau_2 = 0.5*\theta_o$), the derivative time is:

$$T_d = 0.5 * \theta_o \quad \text{(A-6)}$$

For a "near-integrating" or "pseudo-integrating" process, the open loop time constant (τ_o) is:

$$\tau_o = \frac{K_o}{K_i} \quad \text{(A-7)}$$

For a "near-integrating" or "pseudo-integrating" process, the open loop gain (K_o) can be estimated as the initial percent process variable ($\%PV_i$) divided by the initial percent controller output ($\%CO_i$):

$$K_o = \frac{\%PV_i}{\%CO_i} \quad \text{(A-8)}$$

The Equation A-1 setting for gain is about half and the Equation A-3 setting for integral time is about twice what would be used for maximum disturbance rejection. The lower gain and higher integral time setting provides robustness and a smoother response. In general, there is a tradeoff between robustness and performance where higher performance comes at a price of lower robustness.

Appendix B – PID checklist

There is an incredible offering of PID features and options. To help utilize the full potential of the PID, a checklist is offered here as a guide. While the full aspects of the PID capability are book worthy, the following overview can get you started on the right path.

If you do not get the valve action and control action right, nothing else matters. The controller output will ramp off to an output limit. The valve action (inc-open and inc-close) can be set in many different places, such as the PID block, Analog Output (AO) block, Splitter block, Signal Characterizer block, Current to Pneumatic (I/P) transducer, or the Positioner. Make sure the valve signal is not reversed in more than one location for an inc-close (fail open) valve. Once the valve action is set properly, the control action is set to be the opposite of the process action. The control action is reverse and direct if an increase in the PID output causes the PID process variable (PV) to increase or decrease, respectively. Verify with process engineer the valve action, process action, and resulting control action required. Deferring considerations, here is the checklist without delay.

1. Does the measurement scale cover entire operating range including abnormal conditions?
2. Is the valve action correct (inc-open for fail close and inc-close for fail open)?
3. Is the Control action correct (direct for reverse process and reverse for direct process if the valve action is set)?
4. Is the best “Form” selected (ISA standard)?
5. Is the “obey setpoint limits in cascade and remote cascade mode” option selected?
6. Are the “back calculate” signals correctly connected between blocks for bumpless transfer?
7. Is the “PV for back calculate” selected in secondary loop PID?
8. Is the best “Structure” selected (PI action on error, D action on PV for most loops)?
9. Is the “setpoint track PV in manual” option selected to provide a faster initial setpoint response unless setpoint must be saved in PID?
10. Are setpoint limits set to match process, equipment, and valve constraints?
11. Are output limits set to match process, equipment, and valve constraints?
12. Are anti-reset windup (ARW) limits set to match output limits?
13. Is the module scan rate (PID execution time) less than 10% of minimum reset time?
14. Is the signal filter time less than 10% of minimum reset time?
15. Is the PID tuned with auto tuner or adaptive tuner?
16. Is rate time less than ½ deadtime (rate typically zero except for temperature loops)
17. Is external-reset feedback (dynamic reset limit) enabled for cascade control, analog output (AO) setpoint rate limits, and slow control valves or variable speed drives?
18. Are AO setpoint rate limits set for blending, valve position control, and surge valves?
19. Is integral deadband > limit cycle PV amplitude from deadband and resolution?
20. Can an enhanced PID be used for loops with wireless instruments or analyzers?

The setting of all options and parameters must be verified as applicable. Simulations representative of the dynamic behavior of the process and the field automation system along with the actual configuration to form a virtual plant is advisable for testing and confirmation plus training and opening the door to process control improvement (see [Exceptional Opportunities in Process Control – Virtual Plants]).

Most PID controllers use the ISA "Standard" form. Analog controllers used the "Series" form where the derivative calculation was done on the rate of change of the process variable (PV) in series with proportional and integral calculations. This form was principally the result of analog circuitry limitations. An advantage of the "Series" form is the inherent prevention of the effective rate time from exceeding the effective reset time preventing instability from excessive rate action. The inherent protection is the result of an interaction factor that reduces the effective rate time and controller gain and increases the controller reset time as shown in Equations 3-6 thru 3-9 in the chapter on Basic Control in the ISA book *Advanced Temperature Measurement and Control*. The interaction factor becomes 1.0 if the rate time is zero. These equations and interaction factors can be used for the conversion of an analog controller's "Series" PID tuning to a DCS "Standard" PID tuning. This chapter also provides figures and a discussion of the forms and structures commonly offered in a modern distributed control system (DCS).

The PID structure mode commonly used is "PI on error and D on PV." This structure provides a step change in the PID output from the proportional mode for a step change in the PID setpoint. Operator or sequence initiated changes in operating point are step changes in the setpoint. This step change in the output from proportional action helps get the PV to setpoint faster, which is important for reducing startup, transition, and batch times. However, this step in the output (particularly large for a large controller gain) is more likely to cause overshoot. The selection of "I on error and P and D on PV" can eliminate this overshoot, but the time to reach setpoint (rise time) may be painfully slow.

There is a unified approach where tuning for maximum disturbance rejection and the structure "PI on error and D on PV" can be used to minimize rise time with no overshoot. The approach is to add a lead-lag to the setpoint change. The lag time is set equal to the reset time and the lead time is set less than or equal to ¼ of the lag time. The lead time is reduced to provide a slower approach to setpoint.

A primary controller output can be prevented from going beyond the setpoint limits of a secondary loop driven by the output by obeying setpoint limits in cascade and remote cascade mode. Stopping the output at the setpoint limit allows a more immediate recovery when the output reverses direction. The proper use of the back calculate signal enables a bumpless transfer for mode changes and a responsive transition in override control and prevents bursts of oscillations for slow secondary loops and slow valves. The use of PV for the back calculate combined with the dynamic reset limit or "external reset feedback" limits the primary loop output from changing faster than the secondary loop or a slow valve can respond preventing a burst of oscillations for large disturbances or setpoint changes. A fast readback of actual valve position is needed by a dedicated signal or primary HART variable for the valve PV. A secondary HART variable for valve readback may not be fast enough for the dynamic reset limit.

If the setpoint tracks the PV in the manual mode, then the setpoint change in the auto mode will provide a step in the controller output from the structure "PI on error and D on PV." If the setpoint is left at the last operating point in auto (no setpoint change in auto), the approach to setpoint is extremely slow unless the output is prepositioned by an ROUT mode because the change in controller output to achieve the setpoint is a ramp from integral action instead of a step from the proportional mode. For temperature loops, the integral time setting is large causing a slow rise time. If the setpoint must be retained in the PID loop, setpoint tracking of PV is not used. For primary loops used in traditional basic control of continuous operations, there are few setpoint changes, and controller outputs are prepositioned for startup. When grade transitions, flexible manufacturing, model predictive control, and real-time optimization result in setpoint changes, the use of setpoint tracking PV in manual enables a smoother transition to advanced control besides cascade control.

The signal filter time and the PID module execution time should each be less than 10% of the smallest integral time to prevent the integrated error from an unmeasured disturbance increasing more than 20% per Equation (2) in the *InTech* article "PID tuning rules."

Directional velocity limits on the setpoint in an Analog Output (AO) block used in conjunction with a dynamic reset limit and the use of AO PV for the back calculate signal can provide intelligent coordination and regulation of control loop speed without the need for retuning the PID.

Directional AO setpoint velocity limits can provide a slow approach to an optimum and a fast getaway from trouble for valve position control (slow approach to optimum valve position when inside position constraint and fast recovery from valve position outside constraint).

Directional AO setpoint velocity limits offer quick recovery from surge conditions and a lower chance of re-entry in surge by a fast opening and slow closing of the surge valve. In the old days, directional slewing rate was done on the fail open surge valve by quick exhaust valves or boosters with a higher vent rate than pressurization rate. The action of these devices was disruptive and unrepeatable posing operational, maintenance, and tuning problems.

Directional AO setpoint velocity limits offer the opportunity for loops to have the same speed of response despite different dynamics. The greatest need is commonly seen in coordination of flow loop response. For blending operations, flows are set in ratio to each other. If the setpoints of the flow loops are simultaneous driven and directional velocity limits ensure the speed of the flow response is nearly identical, the blend composition will not be upset by an unbalance in flows. The same strategy is useful for minimizing the upset from load changes and analyzer corrections of ratios for reactor feeds and for using the same model for flow response in model predictive control (MPC), particularly advantageous for minimizing duplicate MPC setup and maintenance in parallel equipment trains.

The limit cycles from deadband (e.g., valve backlash with two or more integrators in loops and process), and resolution or threshold sensitivity limits (e.g., valve stiction with one or more integrators in loops or process) can be killed by setting the integral deadband equal to the PV amplitude of the limit cycle. The PV amplitude depends upon operating point and usually gets larger as a valve approaches the closed position or as the product or corrosion builds up on sealing or seating surfaces and stems. The enhanced

PID developed for wireless will inherently kill the oscillations if there is no noise to trigger an update. A filter or threshold sensitivity setting used to screen out noise can prevent unnecessary updates.

For analyzer loops where the dead time is much larger than the sum of the process time constant and dead time, the enhanced PID enables the use of a PID gain that is the inverse of the open loop gain. This PID gain provides a single correction for a setpoint change that puts the PV at the setpoint for the next update (see *InTech* article "Wireless – Overcoming challenges of PID control & analyzer applications").

Appendix C – Control loop performance

The following is an excerpt from Chapter 14 that I contributed to *PID Control in the Third Millennium: Lessons Learned and New Approaches*, Editors Ramon Vilanova and Antonio Visioli, Springer 2012.

Special algorithms can be designed to deal with measured load disturbances at the process input, setpoint changes, and disturbances at the process output (e.g., noise). Often neglected is the overriding requirement that controllers in industrial applications must be able to deal with unmeasured and unknown load disturbances at the process input. Fortunately, the PID controller excels at this load disturbance rejection. An estimate of the current and best possible load rejection as a function of the process and automation system dynamics and controller tuning provides the information on what can be done to improve plant design and tuning. A simple set of equations can be developed that estimates the integrated error and peak error for a step change in a load disturbance. The value is more in helping guide decisions on improvements rather than predicting actual errors because of the uncertainty of the size and speed of load disturbances and the nonlinear and non-stationary nature of industrial processes. The equations are simple enough to provide key insights as the relative effects of the controller gain and integral time and the first order plus dead time (FOPTD) approximation of the process and automation system dynamics. In FOPDT model, a fraction of each of the time constants smaller than the largest time constant is taken as equivalent dead time and summed with the pure dead times to become the total loop dead time (θ_o) termed a process dead time (θ_p) in the literature. The fraction of the small time constants not taken as dead time is summed with the largest time constant to become the open loop time constant (τ_o). While the equations for tuning and estimation of errors is based on the open loop time constant, we will assume the largest time constant is in the process so we have the more common term of process time constant (τ_p) seen in the literature. In reality, fast loops, such as liquid flow and pressure, have a time constant in the FOPDT model much larger than the flow response dead time due to a transmitter damping setting and signal filter time constant. Similarly, the equations seen in the literature use a process gain (K_p) rather than the open loop gain (K_o) that is the product of the final control element, process, and measurement gain. For improving dynamics, a distinction of the location of nonlinearities, dead time, and the largest time constant are important. By avoiding the categorization of dynamics as being solely in the process, a better understanding of the effect of the final control element size, installed characteristic, stick-slip, and backlash, the effect of measurement noise, lag, delay, calibration span, and the effect of PID filter and execution time is possible. The nomenclature used in the quantification of these effects is defined at the end of the chapter.

Since a controller cannot compensate for an unmeasured load disturbance before the loop dead time, the peak error (E_x) (maximum error for a disturbance) is the excursion of the first order response to the step disturbance (E_o) based on the open loop time constant for a time duration of the loop dead time (Equation C-1). The open loop error is the final error seen at the PID from an unmeasured load disturbance if the PID was in manual. The terms "open loop" and "closed loop" are used for a response without and with feedback correction, respectively.

$$E_x = [1 - e^{-\theta_o / \tau_o}] * E_o \qquad \text{(C-1)}$$

If the total loop dead time is much larger than the open loop time constant, then the peak error is basically the open loop error. If the dead time was less than the time constant, then Equation C-1 can be simplified to Equation C-2 eliminating the exponential term.

$$E_x = \frac{\theta_o}{(\theta_o + \tau_o)} * E_o \qquad \text{(C-2)}$$

The minimum integrated error (E_i) can be approximated as the area of two right triangles with the altitude equal to the peak error and the base equal to the dead time. Taking the area of each triangle as ½ the base multiplied by the altitude we obtain Equation C-3 where the integrated error is simply the peak error multiplied by the dead time and consequently proportional to the dead time squared.

$$E_i = \frac{\theta_o^2}{(\theta_o + \tau_o)} * E_o \qquad \text{(C-3)}$$

Equations C-2 and C-3 are for the minimum possible errors determined by the open loop process and system automation system dynamics. It is not possible to do better than what is permitted by the dynamics. Thus, these are the ultimate limits to loop performance for unmeasured load disturbances. What is achieved in feedback control depends upon the tuning. In practice, controllers are not tuned aggressively enough to achieve the ultimate limit because the response tends to be too oscillatory especially for large setpoint changes and the controller lacks robustness. A 25% increase in loop dead time or open loop gain or 25% decreases in the open loop time constant can result in oscillations that do not sufficiently decay. We can develop the equations that set the practical limit in terms of controller tuning settings from the equations for the ultimate limit based on open loop dynamics. We will also see that we can independently arrive at the same equation for the integrated error from the response of the PI algorithm to a step disturbance.

If we divide through by the dead time term in Equation C-2, we have Equation C-4 where the peak error depends upon the ratio of the open loop time constant to total loop dead time.

$$E_x = \frac{1}{(1+\frac{\tau_o}{\theta_o})} * E_o \qquad \text{(C-4)}$$

Most tuning methods for maximum disturbance rejection use a controller gain (K_c) that is proportional to the ratio of the open loop time constant to total loop dead time and inversely proportional to the open loop gain (Equation C-5).

$$K_c = \frac{\tau_o}{\theta_o * K_o} \qquad \text{(C-5)}$$

If we solve for the open loop time constant to total dead time ratio, we see this ratio is simply the product of the controller gain and open loop gain ($K_c * K_o$). If we substitute the product for the ratio in Equation C-3, we have Equation C-6, which is the practical limit to the peak error. Peter Harriott developed the same form of the equation but with a numerator of 1.5 for the peak error from a proportional only controller tuned for quarter amplitude decaying response.

$$E_x = \frac{1}{(1+K_c * K_o)} * E_o \qquad \text{(C-6)}$$

For time constant to dead time ratios that are much larger than one, which is the case for pressure and temperature control of vessels and columns, the product of the controller gain and open loop gain is much greater than one leading to the peak error being simply inversely proportional to the product. Since the controller gain used in practice is about half of the gain for maximum disturbance rejection, we end up with Equation C-7 for the peak error.

$$E_x = \frac{2}{K_c * K_o} * E_o \qquad \text{(C-7)}$$

Equation C-7 corresponds to a peak error reached in about two dead times. If we approximate the integrated error as the area of two right triangles each with a base equal to two dead times and consider the integral time (T_i) setting as being 4 dead times, we end up with Equation C-8 for the integrated error.

$$E_i = \frac{T_i}{K_c * K_o} * E_o \qquad \text{(C-8)}$$

We can derive Equation C-8 from the equation for a PI controller's response to an unmeasured load disturbance. The change in controller output from time *t1* to time *t2* is the sum of the contribution from the proportional mode and the integral mode (Equation C-9a). The module execution time (Δt_x) is added to the reset or integral time (T_i) to show the effect of how the integral mode is implemented in digital controllers. An integral time of zero ends up as a minimum integral time equal to the execution time so there is not a zero in the denominator for the integral mode. For analog controllers, the execution time is effectively zero.

$$CO_{t2} - CO_{t1} = K_c * (E_{t2} - E_{t1}) + \left[\frac{K_c}{T_i + \Delta t_x} \right] * \int_{t1}^{t2} E_t * \Delta t_x \qquad \text{(C-9a)}$$

The errors before the disturbance (E_{t1}) and after the controller has completely compensated for the disturbance (E_{t2}) are zero ($E_{t1} = E_{t2} = 0$). Therefore, the long-term effect of the proportional mode, which is first term in Equation C-9a, is zero. Equation C-9a reduces to Equation C-9b.

$$\Delta CO = \left[\frac{K_c}{T_i + \Delta t_x} \right] * \int_{t1}^{t2} E_t * \Delta t_x \qquad \text{(C-9b)}$$

The integrated error is the integral term in Equation C-9b giving Equation C-9c. For over-damped response the integrated error and the integrated absolute error (IAE) are identical.

$$E_i = \int_{t1}^{t2} E_t * \Delta t_x \qquad \text{(C-9c)}$$

If we substitute Equation C-9c into Equation C-9b, we have Equation C-9d.

$$\Delta CO = \left[\frac{K_c}{T_i + \Delta t_x} \right] * E_i \qquad \text{(C-9d)}$$

The change in controller (*ΔCO*) multiplied by the open loop gain (K_o) must equal the open loop error (E_o) for the effect of the disturbance to be eliminated. We can express this requirement as the change in output being equal to the open loop error divided by the open loop gain (Equation C-9e).

$$\Delta CO = \frac{E_o}{K_o} \qquad \text{(C-9e)}$$

If we substitute Equation C-9e into Equation C-9d and solve for the integrated error, we end up with Equation C-9f, which is the same as Equation C-8 except for the addition of the execution time interval for the digital implementation of the PI algorithm.

$$E_i = \left[\frac{(T_i + \Delta t_x)}{K_o * K_c} \right] * E_o \qquad \text{(C-9f)}$$

Recently, Greg Shinskey added a term to the numerator to include the effect of a signal filter time constant on the integrated error (Equation C-10). In Shinskey's presentation of the equation, the change in controller output rather than the open loop error is used, which eliminates the open loop gain in the denominator. Equation C-10 is applicable regardless of tuning settings. The additional equivalent dead time from the filter time and execution time interval may necessitate a decrease in controller gain and increase in integral time further degrading performance.

$$E_i = \left[\frac{(T_i + \Delta t_x + \tau_f)}{K_o * K_c} \right] * E_o \qquad \text{(C-10)}$$

To summarize, in the process industry, automation system and process dynamics, and in particular the loop dead time, set the ultimate limit to loop performance but controller tuning sets the practical limit for unmeasured disturbances. For example, a loop with a small dead time will perform as badly as a loop with a large dead time if the PID has sluggish tuning. On the other hand, a PID with fast tuning may have an excessive oscillatory response for increases in the loop dead time or process gain. Equation C-6 shows the practical limit to the peak error (E_x) is inversely proportional to 1 plus the product of the PID gain (K_c) and the process gain (K_p). Equation C-9f indicates the integrated error (E_i) is proportional to the ratio of the PID integral time to gain (T_i/K_c). For small filters (τ_f) and PID execution time (ΔT_x), the maximum controller gain is decreased, and the minimum integral time is increased based on the increase in loop dead time. The filter and execution time should also be added to the integral time for the integrated error to show the increase in the practical limit (Equation C-10). For a PID tuned for maximum disturbance rejection, Equation C-3 reveals the ultimate limit to the peak error depends upon the ratio of the total loop dead time (θ_o) to the open loop time constant (τ_o). Equation C-8 and C-10 indicates the integrated error depends upon the ratio of the loop dead time squared to open loop time constant. A PID controller tuned for maximum disturbance rejection has a controller gain proportional to the ratio of the largest open loop time constant to loop dead time (τ_o/θ_o), and an integral time proportional to the loop dead time. Note the controller tuning depends upon the largest open loop time constant and not the process time constant. If the largest time constant is in the measurement path, the observed peak error in the measurement predicted by Equation C-2 will be smaller than the actual peak error in the process because of the signal filtering effect of the measurement time constant.

The peak error is important for preventing: shutdowns from reaching trip settings of safety instrumentation systems (SIS), environmental emissions and process losses from reaching the relief settings of rupture discs and relief valves, off-spec paper sheet and plastic web from exceeding permissible variation in thickness and clarity, compressor shutdowns from crossing surge curve, and recordable incidents by exceeding environmental limits.

The integrated error is a good indicator of the quantity of liquid product off-spec in equipment with back mixing. In these volumes, positive and negative fluctuations in concentration are averaged out unless irreversible reactions are occurring.

An important emerging consideration is the realization that initial open loop response in the FOPDT approximation of a self-regulating process is a ramp seen in the response of an integrating process such as level and batch temperature. The ramp is more persistent in a self-regulating process with a large open loop time constant. The process is termed “near integrating” or “pseudo integrating.” An equivalent integrating process gain (K_i) can be approximated as the open time constant divided by the open loop gain (Equation C-11). For processes where the open loop time constant is more than 10 times larger than the dead time, the identification of this near integrator gain in three dead times can reduce the time required for process identification by more than 90% compared to those techniques that go to the 98% response time. The process variable (PV) is passed through a dead time block to create an old PV that is subtracted from the new PV to create a ∆PV and then an integrating gain by dividing by the dead time and the change in controller output. The maximum of a continuous train of these "near integrating" process gains updated every execution of the PID module can be used for tuning controllers on all types of processes.

$$K_i = \frac{K_o}{\tau_o} \tag{C-11}$$

If we substitute the near integrating gain for the time constant to dead time ratio in Equation C-5, we have Equation C-12. Recently, this method was found to even work on processes where the dead time was greater than the time constant. To provide a smoother response, less setpoint overshoot and more robust settings, the controller gain in both Equation C-5 and C-12 is cut in half.

$$K_c = \frac{1}{\theta_o * K_i} \tag{C-12}$$

The optimum integral time depends upon the type of process. The integral time ranges from about four times the dead time to integrating and "near integrating" processes to ½the dead time for severely dead time dominant process ($\theta_o >> \tau_o$). Equation C-13 provides a reasonable curve fit to the required relationship for self-regulating processes. For a dead time less much less than the time constant ($\theta_o < 0.1\ \tau_o$), the ultimate period is about four times the dead time, and the denominator is about one giving an integral time that is about four times the dead time. For a dead time much greater than the time constant ($\theta_o > 10\ \tau_o$), the ultimate period is about two times the dead time, and the denominator is about four, giving an integral time that is ½ of the dead time.

For self-regulating processes:

$$T_i = \frac{T_u}{Min(4, 3 * (\frac{4 * \theta_o}{T_u} - 1)^2 + 1)} \tag{C-13}$$

For a dead time dominant process, the combination of Equation C-13 for integral time and Equation C-5 for controller gain results in almost an integral-only controller. Since the controller gain is so low, this process is a candidate for setpoint feedforward to reduce the setpoint response rise time.

For an integrating process, the product of the controller gain and integral time must be greater than twice the inverse of the integrating process gain to prevent slowly decaying oscillations from the integral mode dominating the proportional mode. If the user is confident in the knowledge of the integrating process gain, this relationship can be used to find the integral time (Equation C-14a). Since the maximum controller gain allowable on many level and batch temperature loops is greater than 100 and the actual controller gain used is often less than 10, the integral time must be increased to prevent the slow rolling oscillations. Consequently, while an integral time of four dead times is possible for an integrating process, in practice an integral time of 40 dead times is more appropriate because the maximum controller gain is beyond the user's comfort level (Equation C-14b).

To prevent slowly decaying oscillations integrating processes from excessive integral action:

$$T_i \geq \frac{2}{K_c * K_i} \qquad \text{(C-14a)}$$

The positive feedback in the runaway processes necessitates an integral time 10 times larger than the integral time for a "near integrating" self-regulating process. The integral time should be 40 dead times or larger for a runaway process (Equation C-14b). Some highly exothermic polymerization batch reactors have gone to proportional plus derivative control to avoid the problem of a user setting too small of an integral time.

For integrating processes with controller gains less than 10 times, the maximum permissible controller gain and for runaways (processes with positive feedback):

$$T_i = 40 * \theta_o \qquad \text{(C-14b)}$$

Too small of a controller gain besides too large of a controller gain can cause a runaway. There is a window of allowable controller gains for positive feedback processes. Any changes in tuning settings particularly for runaway reactions must be closely monitored.

Common metrics for a setpoint response are rise time (time to reach setpoint), overshoot (maximum error after crossing setpoint), and settling time (time settle out within a specified band around the setpoint). The ultimate limit for rise time is proportional to the loop dead time. The ultimate limit for overshoot and settling time is theoretically zero. The practical limit to rise time is similar to the practical limit for peak error for fast tuning settings but degrades to the relationship for the integrated error for sluggish tuning settings. Fortunately, there are many features that can be used to readily help achieve the ultimate limit to the rise time. The practical limits for overshoot and settling time depend upon a balance between the contributions from the integral and proportional modes. In general, the controller gain for maximum disturbance rejection can be used to minimize rise time, and the integral time can be increased to minimize overshoot and settling time.

The minimum rise time (T_r) can be approximated as the change in setpoint (ΔSP) divided by the maximum rate of change of the process variable. For an integrating or "near integrating" process, the maximum PV ramp rate is the integrating process (K_i) gain multiplied by the change in controller output as detailed in the denominator of Equation C-15a. If the step change in controller output from the proportional mode for a structure of proportional action on error is less than the maximum available output change (difference between current output and output limit), Equation C-15a simplifies to Equation C-15b for feedback control. The output change must be corrected for methods used to make the setpoint response faster. For setpoint feedforward, the step change in output is a combination of the feedforward and feedback action. For smart bang-bang logic, the step output change is the maximum available output change.

$$T_r = \frac{\Delta SP}{K_i * \min(|\Delta CO_{\max}|, (K_c + K_{ff}) * \Delta SP)} + \theta_o \qquad \text{(C-15a)}$$

For a maximum available output change larger than the step from the proportional mode

($|\Delta CO_{\max}| > K_c \Delta SP$) the change in setpoint in the numerator and denominator cancel out yielding a simpler equation:

$$T_r = \frac{1}{(K_i * K_c)} + \theta_o \qquad \text{(C-15b)}$$

For the "near integrating" process response seen in vessel and column temperature loops where the process time constant is significantly larger than the total loop dead time, the integrating gain is the open loop gain (K_o) divided by the open loop time constant (τ_o) yielding Equation C-15c.

$$T_r = \frac{\tau_o}{(K_o * K_c)} + \theta_o \qquad \text{(C-15c)}$$

The practical and ultimate limit to loop performance can be reconciled by realizing there is an implied dead time (θ_i) from the tuning. Equation C-16 shows the implied dead time that can be approximated as the original dead time (θ_o) multiplied by a factor that is 0.5 plus Lambda (λ). Lambda is the closed loop time constant for a setpoint change. For a PID tuned for maximum disturbance rejection, Lambda is set equal to the original dead time. The implied dead time is then equal to the original dead time.

$$\theta_i = 0.5 * (\lambda + \theta_o) \qquad \text{(C-16)}$$

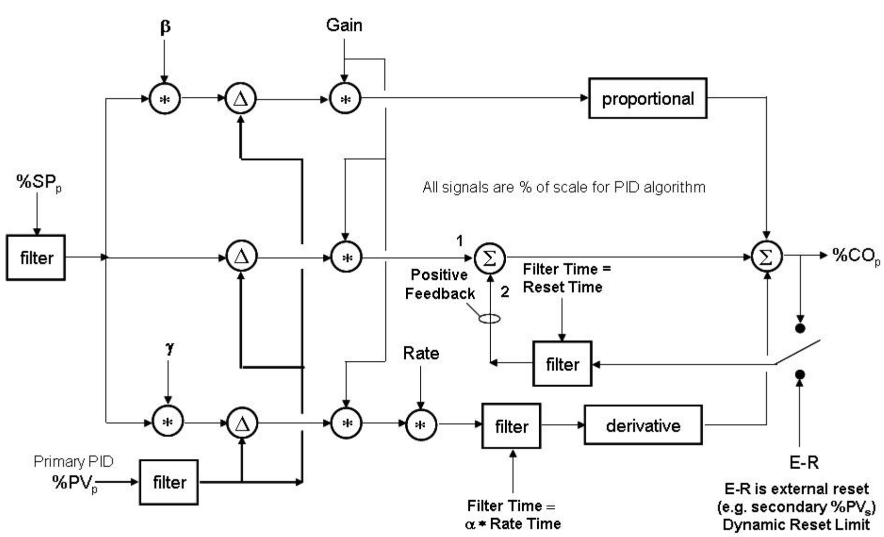
The peak and integrated errors for unmeasured step disturbances represents the worst case. Step disturbances originate from manual actions, safety, switches, and sequential operations. If discrete actions (e.g., the opening and closing of on-off valves and the starting and stopping of pumps) are replaced by control loops with modulated final control elements (throttling valves and variable speed drives) or are attenuated by intervening volumes, the step disturbances are smoothed. The attenuated load disturbance has a time constant (τ_L) that is the residence time of the volume or closed loop time constant of the upstream control loop. To include the effect of a load time constant, the process excursion in the first dead time, which is the key time for determining minimum peak error, can be computed by Equation C-17. The open loop error (E_o) in the equations for peak and integrated error can be replaced with the load disturbance (E_L) that is the open loop error multiplied by the exponential response of the disturbance in one dead time The effect is mitigated by a reset time that is slow relative to the disturbance time constant.

$$E_L = (1 - e^{-\theta_o / \tau_L}) * E_o \qquad \text{(C-17)}$$

PID controllers tuned too fast can introduce process variability from an oscillatory response, PID controllers tuned too slowly can make a loop with good dynamics perform as badly as a loop with poor dynamics. In other words, money invested to reduce process dead time or to get faster measurements and valves is wasted unless the PID controller tuning is commensurate with the speed of the process so that the practical limit approaches the ultimate limit to loop performance.

Since industrial processes have valve, process, and measurement dynamics that vary with time, operating point, and step size, it is important to have automated methods of tuning.

Instead of integrating the error, the feeding back of the controller output or external reset signal through a filter block and adding it to the contribution of the proportional and derivative modes creates an integral mode action where the filter time constant is the integral time setting. When the error is zero, the output of the filter block is simply the controller output or external reset signal and integral action stops. The positive feedback implementation illustrated in Figure C-1 enables several important PID options, such as dynamic reset limit, enhancement for wireless, and dead time compensation. Figure C-1 is for the ISA standard form for the PID controller. The eight structures commonly used in industrial processes are obtained by setting the setpoint weight factor β for the proportional and the setpoint weight factor γ for the integral mode in Figure C-1. If the factor is zero, a setpoint change does not affect the contribution to the output from respective mode (action is on PV only). If the factor is one, the full effect of a setpoint change is included (full action is on error). A factor between zero and one provides the ability to include but moderate the effect of a setpoint change (balanced action on setpoint change and PV change). In this figure the multiplication symbol "$*$" in a circle is used to denote the multiplication by the β or γ weight factor.



Positive Feedback Integral Mode Enables Key PID Features (external reset, velocity limiting of AO setpoint for optimization and coordination, wireless enhancement, and dead time compensation)

β and γ are setpoint multiplication factors for the proportional and derivative modes, respectively, to determine how much proportional and derivative action occurs on setpoint changes. These factors do not affect the ability of the PID to reject disturbances. For the fastest possible setpoint response, structures 1 and 2 are used. If preventing overshoot is more important than minimizing rise time, structure 3 is used. If the ability to customize the balance between fast rise time and minimum overshoot for a setpoint response is needed, structure 8 is used. This structure also offers the ability to achieve both good load and setpoint responses.

The eight PID structures commonly used in industrial processes are:

1. PID action on error ($\beta = 1, \gamma = 1$)
2. PI action on error, D action on PV ($\beta = 1, \gamma = 0$)
3. I action on error, PD action on PV ($\beta = 0, \gamma = 0$)
4. PD action on error ($\beta = 1, \gamma = 1$) (no I action)
5. P action on error, D action on PV ($\beta = 1, \gamma = 0$) (no I action)
6. ID action on error ($\gamma = 1$) (no P action)
7. I action on error, D action on PV ($\gamma = 0$) (no P action)
8. Two degrees of freedom controller (β and γ adjustable 0 to 1)

When an external signal is used as the input to a “Filter” block in the positive feedback implementation of the integral mode, the integral action will not drive the controller output faster than the external reset signal is changing. This capability known as “dynamic reset limit” and “external reset feedback” is particularly important for slow final control elements (large valves and variable frequency drives), cascade control, and override control.

If the external reset signal is the actual valve position or variable frequency drive (VFD) speed, the PID controller output will not ramp faster than the valve or VFD can respond. Control valves and dampers have a slewing rate that increases with actuator size and stroke length. Damper slewing rate is particularly slow due to the need to prevent positive feedback from negative torque requirement. VFDs have velocity limiting of the command signal to prevent overloading the motor. If the external reset signal is the secondary loop process variable (PV) for cascade control, the primary PID cannot ramp the setpoint of the secondary PID faster than the secondary PID PV can respond. This capability is important for inherently preventing severe oscillations from breaking out for large setpoint changes or large disturbances [24,32]. The use of the selected PID output as an external reset signal for override control also inherently prevents the unselected PID controllers from ramping off-scale. PID algorithms without the positive feedback implementation of integral action add a “Filter” block to the external reset signal with a filter time equal to the PID reset time to prevent the ramping off-scale of the unselected PID output. The dynamic reset limit is a key feature that enables the development of an enhancement of the PID for wireless measurements that also has the ability to eliminate oscillations from threshold sensitivity and resolution limits and feedforward timing errors.

The dynamic reset limit can open opportunities important for sustainable manufacturing and in particular abnormal situation management and optimization. If a setpoint velocity limit is set in the Analog Output block, the dynamic reset limit prevents the PID from going faster than the velocity limit. The PID can achieve a slow approach to an optimum and a fast recovery upon encroachment of a constraint such as encountered in the prevention of compressor surge, exothermic reactor runaway, RCRA pH violation, and Bioreactor biomass starvation. Previously, an open loop back-up (kicker) has been used for these applications because the tuning of the controller for drastically different speeds of actuation is problematic. The dynamic reset limit option eliminates the need to tune the controller based on direction and the concern about the exact value of the velocity limit. The tuning is set for the fastest recovery. The velocity limit is adjusted for the slowest approach to the optimum.

There are many more examples where an intelligent adaptation of the speed of actuation of the final control element or secondary loop could be beneficial. In general, you want to approach optimums slowly to minimize disruption, but as you operate close to the edge, you depend upon a fast recovery to prevent going over the edge. With compressor surge control, the edge is literally a cliff. While other applications might not be as dramatic, the technique opens a wide spectrum of PID techniques for sustainable manufacturing, which in its broadest definition includes efficiency, flexibility, operability, maintainability, safety, and profitability.

Wireless measurement devices have a “default update rate” (time interval for periodic reporting) and a “trigger level” (threshold sensitivity limit for exception reporting) set as large as possible to conserve battery life. The integral mode in the traditional PID will continue to ramp while the PID is waiting for an updated measurement from a wireless device. Also, when an update is received, the traditional PID considers the entire change to have occurred within the PID execution time interval (ΔT_x). If derivative mode is used, the rate of change of the measurement is the difference between the new and old measurement divided by the PID execution time interval. The result is a spike in the controller output.

The enhanced PID for wireless executes the PID algorithm as fast as wired devices. A change in setpoint, feedforward signal, and remote output translates immediately (within PID execution time interval) to a change in PID output. However, integral action does not make a change in the output until there is an update. When an update occurs, the elapsed time between the updates is used in an exponential calculation that mimics the action of the filter block in the positive feedback implementation of integral action. If derivative action is used, the elapsed time rather than the PID execution time interval, is used to calculate the rate of change of the process variable. The integral and derivative calculations are executed only once upon a change in setpoint or measurement. A threshold sensitivity setting is used to prevent an update from noise.

A traditional PID will have to be detuned to prevent instability for a large increase in the time between updates. The enhanced PID will continue to be stable for even the longest update time interval. For a measurement update time interval larger than the process response time, the enhanced PID controller gain can be set equal to the inverse of the open loop gain (product of valve, process, and measurement gain) to provide a complete correction for setpoint change or update. Subsequent sections show the enhanced PID can suppress oscillations from a wide variety of sources. This reduction in variability results from the suspension of integral action and the wait in feedback correction till there is a more complete response is beneficial. To achieve these benefits, the user simply enables the enhanced PID option in the PID block, which automatically enables the dynamic reset limit option. No retuning is necessary to achieve a smooth response, but if the update time is larger than the process response time, the enhanced PID can be tuned with a much higher gain.

Nomenclature

$\%CO_{t1}$ = controller output at time *t1* before correction for load disturbance (%)
$\%CO_{t2}$ = controller output at time *t2* after correction for load disturbance (%)
E_i = integrated error (%*sec)
E_L = open loop error corrected for load disturbance time constant (%)
E_o = open loop error for unmeasured step disturbance (%)
E_x = peak error (%)
K_c = PID gain (dimensionless)
K_i = integrating process gain (% per sec per %)
K_o = open loop gain for self-regulating processes (dimensionless)
$\Delta\%CO$ = change in controller output (%)
$\Delta\%CO_{\max}$ = maximum available change in controller output to output limit (%)
$\Delta\%PV$ = change in controller process variable (%)
$\Delta\%PV_{\max}$ = maximum change in controller process variable in one dead time time interval (%)
$\Delta\%SP$ = change in controller setpoint (%)
T_i = PID integral time (sec/repeat)
T_r = rise time of setpoint response (sec)
T_u = ultimate period (sec)
λ = closed loop time constant for setpoint change (sec)
θ_i = implied total loop dead time (sec)
θ_o = original total loop dead time (sec)
τ_f = signal filter or volume attenuating time constant (sec)
τ_L = load disturbance time constant (sec)
τ_o = open loop time constant (sec)
τ_P = process time constant (sec)
Δt_x = controller execution time (sec)

Resources

1. Vilanova R. and Visioli A., (2012), *PID Control in the third Millennium: Lessons Learned and New Approaches*, Springer
2. Bohl A. and McAvoy T., (1976) Linear feedback vs. time optimal control. II. The regulator problem, *Ind. Eng. Chem.*, Process Des. Dev., Vol. 15, No. 1
3. Harriott P., (1964) *Process Control*, McGraw Hill
4. McMillan G., (2005) *Good Tuning: A Pocket Guide*, ISA
5. McMillan G., (2011) *Advanced Temperature Measurement and Control (Appendix C – Unification of Controller Tuning Relationships)*, ISA
6. McMillan G., (Oct. 2010) How to get the most out of your PID, ISA Automation Week, Houston

7. McMillan G., (July-Aug. 2010) Wireless - Overcoming challenges of PID control & analyzer applications, *InTech*
8. Shinksey F., (June 1993) The effect of scan period on digital control loops, *InTech*
9. Shinskey F., (1994) *Feedback Control for Industrial Processes*, McGraw Hill
10. Shinskey F., (1996) *Process Control Systems*, McGraw Hill
11. Shinskey F., (May 2006) The power of external reset-feedback, *Control*
12. Shinskey F., (May 2010) Maximizing control-loop performance, *Control*

Appendix D – Dynamics

The following is an excerpt from an ISA book on reactor control available winter 2012. Key insights are in bold type and italics. While written for reactors, the dynamics are generalized to be applicable to industrial processes.

The three types of process responses encountered in industry are defined based on an open loop test where the PID is in manual or remote output so there is no response of the PID to the process (no closed loop response). A step change is made in the controller output. The process response is observed until the process can be identified. During the test, there should be no disturbances so the process response seen is entirely the result of the step change in PID output.

The three types of processes are self-regulating, integrating, and runaway. A self-regulating process will decelerate to a new steady state operating point (Figure D-1). An integrating process will continually ramp (Figure D-2). A runaway process will accelerate until hitting a relief or interlock setting (Figure D-3). For estimating loop performance and tuning settings the parameters used to identify each type of response are gain, time constant, and dead time. The definition of the parameters depends upon the type of response. The terms have alternate names in industry. For example, “lag” is used for time constant, “delay” is used for dead time, and “sensitivity” is used for gain.

The response observed in these tests includes the response of the analog output, final control element (e.g., control valve or variable frequency drive), process, sensor, transmitter, analog input, and the process variable (PV) input to the PID. The observed response includes the effect of velocity limits, dead times, time constants, and gains in the automation system. Better terminology would be “open loop response” than “process response” because the observed response includes almost everything in the loop response. Also, the source of the individual parameters that create the particular dynamic in the response should precede the term (e.g., valve dead time and measurement dead time).

All processes have a dead time that is the time interval between the step change in output and the first recognizable change in process. Noise can delay the recognition until the excursion is beyond the noise band creating a longer dead time. The observed dead time is frequently called the process dead time. The observed dead time is really a total loop dead time (θ_o) that is the sum of all the pure dead times and the equivalent dead times from all time constants smaller than the largest time constant in the loop for a first order (one time constant) plus dead time approximation. For a second order plus dead time approximation that includes a secondary time constant, all time constants smaller the largest time constants create an equivalent dead time. The secondary time constant creates the bend in the initial response right after the dead time. The equivalent dead time increases from 30% to 99% of the time constant as the ratio of the time constant to the largest time constant gets smaller. Time constants small compared to the largest time constants are summed as being essentially 100% dead time.

We will see in the section on loop performance that ultimate limit to the peak and integrated errors are proportional to the dead time and dead time square, respectively. ***If there was no dead time and no noise, perfect control would be possible.***

Not seen in these open loop responses is the additional delay experienced in closed loop operation from the time it takes for the output signal to pass through deadband (backlash), threshold sensitivity (stiction), and resolution limits of the final control element. The test uses a step change larger than these limits. In closed loop operation, a step change in output can occur for a step change in setpoint but subsequent closed loop action to recover from overshoot or disturbances involves gradual changes in the output. The additional dead time from these limits can be approximated as the limit divided by the rate of change of the signal (e.g., valve deadband divided by the rate of change of the PID output).

The open loop time constant (τ_o) is the largest time constant plus any the portion of smaller time constants not taken as effective dead time. While often called the process time constant, the largest time constant can occur anywhere in the loop. For liquid pressure and flow loops, the largest time constant is usually somewhere in the automation system (e.g., valve, sensor, transmitter, or DCS). Ideally, the largest time constant called is the primary time constant of the process downstream of where the disturbances and manipulate flow enter the process. We will see in the section on loop performance, the effect of such a primary time constant is beneficial. The ultimate limit of the peak and integrated is inversely to this primary time constant. The primary process time constant slows down an excursion from a disturbance giving time for the PID to catch up with it. If the largest time constant is in the measurement, the trend chart oscillation may look better because the amplitude is attenuated by the filtering effect of the measurement time constant. In an open loop test, you cannot discern the location of the largest time constant.

The open loop gain (K_o) is the product of the final control element, process, and measurement gain. Consider a loop with a control valve. The final control element gain (change in flow divided by the change in % PID output) is the slope of the valve's installed characteristic curve. The process gain (change in process variable divided by the change in valve flow) is the slope of a plot of the process variable versus valve flow. The measurement gain (change in % PID input divided by the change in process variable) is the 100% divided by the measurement span. ***Since the PID algorithm uses % signals, calculations of the open loop gain must involve % signals despite the fact that the PID block and graphics show the PID process variable and in some DCS the PID output in engineering units.***

Self-regulating processes are defined by a total loop dead time, an open loop time constant, and an open loop gain called a steady gain (Figure D-1). A secondary time constant may be used to describe the initial lag (bend) in the response. The open loop gain is a steady state gain that is the % change in PID input from its initial to its final value divided by the % change in PID output giving a dimensionless gain. Liquid pressure and flow loops have a self-regulating response. Continuous composition, pH, and temperature loops have a self-regulating process response, but the process time constant for large well mixed vessels is so large that in the time frame of interest for the PID (two to four dead times), the response resembles the ramp of an integrating process. ***For tuning and analysis, it is useful to treat self-regulating processes with a time constant order(s) of magnitude than the total loop dead time as "near integrating."***

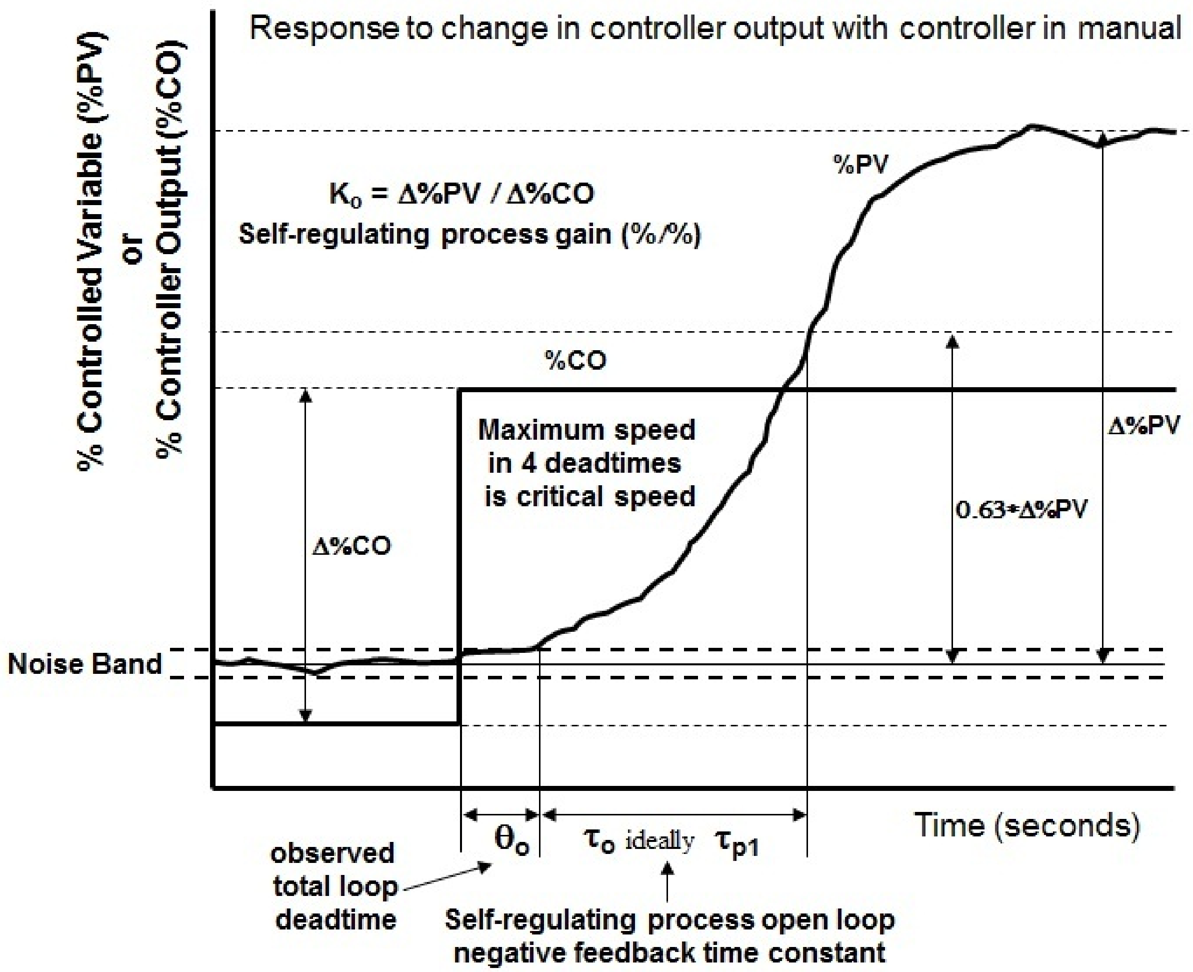


Figure D-1. Open loop self-regulating processes decelerate and line out at a new operating point (steady state) below a physical limit or Safety Instrumentation System (SIS) or relief device setting.

Integrating processes are defined by a total loop dead time and an integrating process gain (Figure D-2). A secondary time constant may be used to describe the initial lag (bend) in the response. The integrating process gain is the ramp rate in % of PID input per second divided by the % change in PID output giving a gain of reciprocal time units (1/sec). Level is the most common loop with an integrating response. Gas pressure has a near integrating or true integrating response depending on the size of the pressure drop across the manipulated valve relative to the changes in pressure during an open loop test. Batch composition, pH, and temperature loops have essentially an integrating response unless altered by a reaction. An integrating process will exhibit self-regulating closed loop response for a proportional only controller. The distance of the new from the initial operating point decreases as the PID gain increases. Integrating and "near integrating" processes require aggressive proportional action. The steady-state gain divided by the open loop time constant of a near integrating process is effectively an integrating process gain. The maximum PID gain is inversely proportional to the process time constant or integrating process gain. ***Most integrating processes are so slow (integrating gain so small) and the dead time is so relatively small that the maximum PID gain so large that the primary limit to how high you set this gain is user knowledge and noise.***

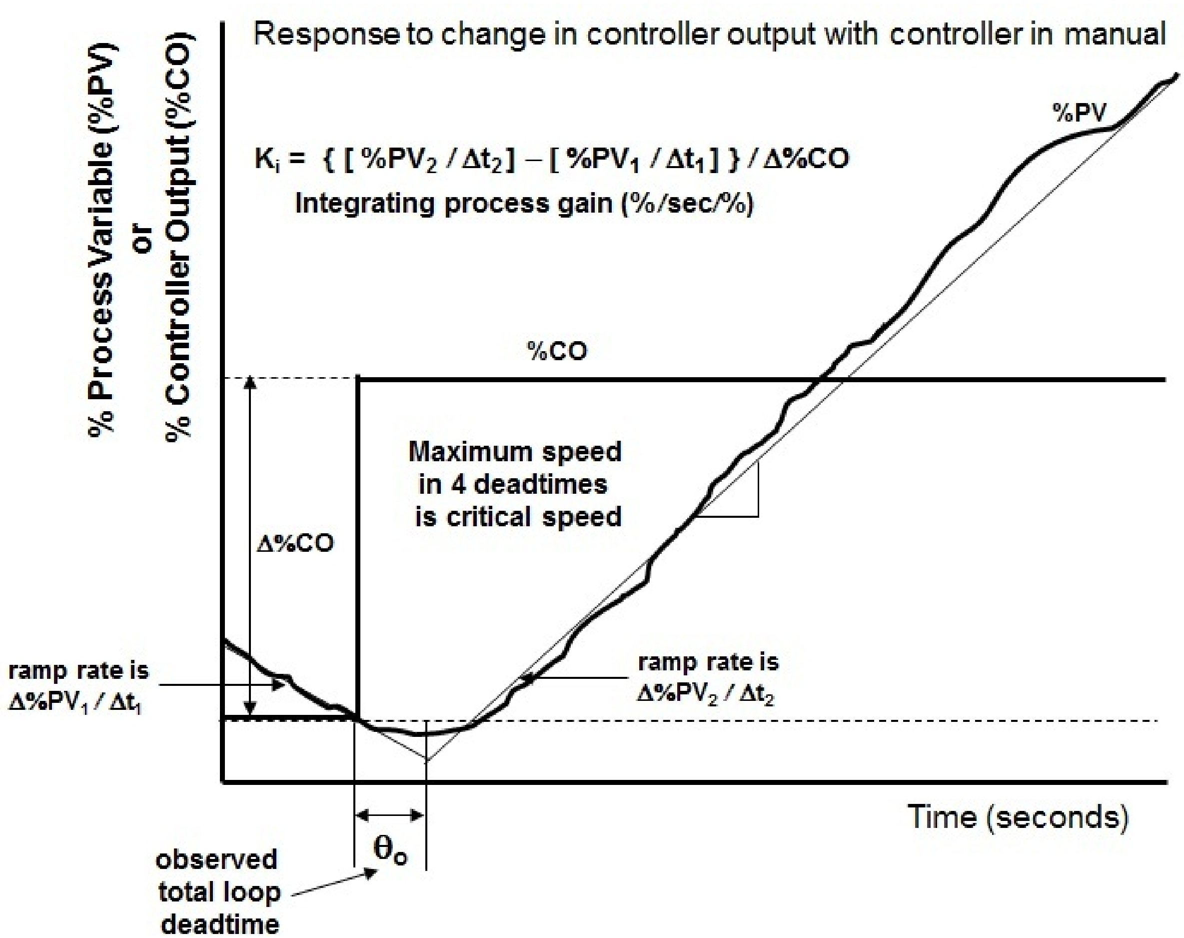


Figure D-2. Open loop integrating processes continue to ramp until hitting a physical limit or triggering the activation of a Safety Instrumentation System (SIS) or relief device

Runaway processes are defined by a total loop dead time, a <u>positive feedback</u> open loop time constant, and an open loop gain called a steady gain (Figure D-3). The positive feedback time constant causes the continual acceleration. A secondary time constant may be used to describe the initial lag (bend) in the response. The open loop gain is a steady-state gain that is the % change in PID input from its initial to its final value divided by the % change in PID output giving a dimensionless gain. Open loop tests are rarely done in runaway processes because of the safety concerns from the acceleration. Manual control of a runaway process is extremely difficult. ***Most true runaway processes are always operated in automatic or a higher mode. Tuning tests are done with the PID loops in auto.*** The large controller gains and integral times used for these loops provide a step change in controller output from the proportional mode and negligible ramping from the integral mode in four dead times. Polymerization and specialty chemical reactors with a heat release from an exothermic reaction that can exceed the cooling rate can develop a runaway response for an increase in temperature.

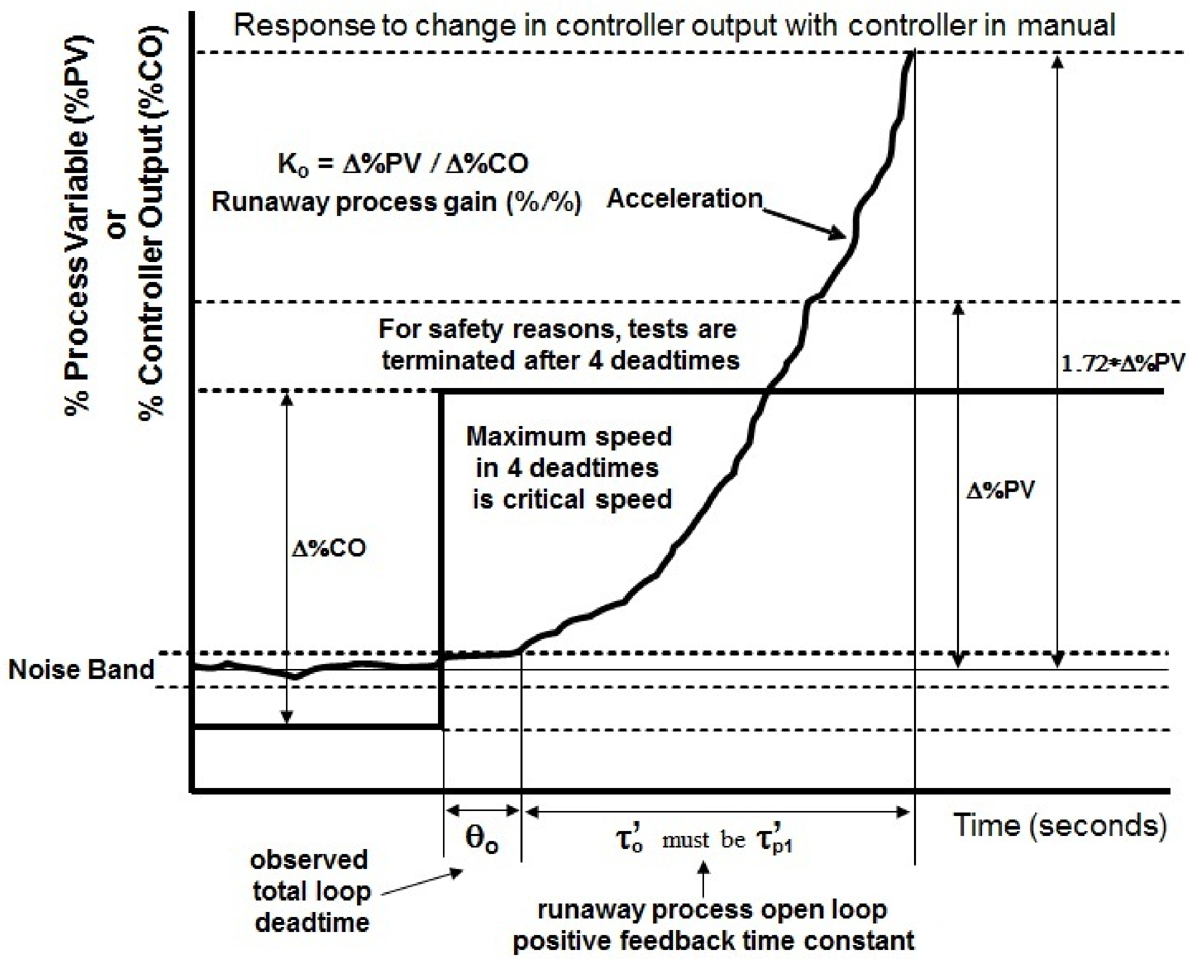


Figure D-3. Runaway processes continue to accelerate until hitting a physical limit or triggering the activation of a Safety Instrumentation System (SIS) or relief device.

A short cut method can easily and rapidly identify the dynamics for slow self-regulating, integrating, and runaway responses because these processes exhibit a similar response in the first four dead times. The method is consistent with the action of a well-tuned PID that reacts and turns around an excursion from a disturbance in 2-4 dead times. A short cut method for response analysis and tuning identifies just the dead time and the maximum ramp rate during the next two-four dead time intervals. By choosing the maximum ramp rate, the effect of a secondary time constant is ignored. The %/sec ramp rate divided by the % change in PID output is the integrating or "near integrating" gain (1/sec). The test can be done with the PID in auto if the structure has the proportional mode on error and the controller gain is high enough to provide a significant step change in PID output. The parameters can generally be identified in less than five dead time intervals, making the test much faster for self-regulating processes that require four-five time constants to recognize the steady state. The fact that the test can be done in auto is fast, and that the parameters are easy to identify, make it attractive for routinely checking dynamics and tuning whenever there step change in the controller output. A short cut tuning method for maximum disturbance rejection simply sets the maximum PID gain equal to 0.5 divided by the product of the integrating gain (1/sec) and dead time (sec). The PID reset time is four times the dead time for disturbance rejection. To minimize overshoot or protect against runaway or to deal with a PID gain set much less than the maximum due to convention or noise, the reset time is increased to 40 times the dead time. Recent evaluations of the tuning equations for dead time dominant processes reveal the short cut method provides a reasonable PID gain making the method more universal than originally expected.

The logic and computation for identifying the dynamics in the short cut method is elegantly concise. When there is a significant step change in PID output, the dead time is identified as the time till when the PV is changing faster in the direction consistent with direction of the change in PID output, valve action, and process action. The PV is sent to a dead time block to create an old PV, which is then subtracted from the new PV to get a delta PV. Every execution of the dead time block provides a new delta. The largest delta PV over four dead times is divided by the dead time and change in PID output. The result is the integrating process gain (1/sec). The PV and output must be in %. See Appendix A for short cut method.

The block diagram (Figure D-4) shows the source of the dynamics depicted in Figures D-1 through D-3.

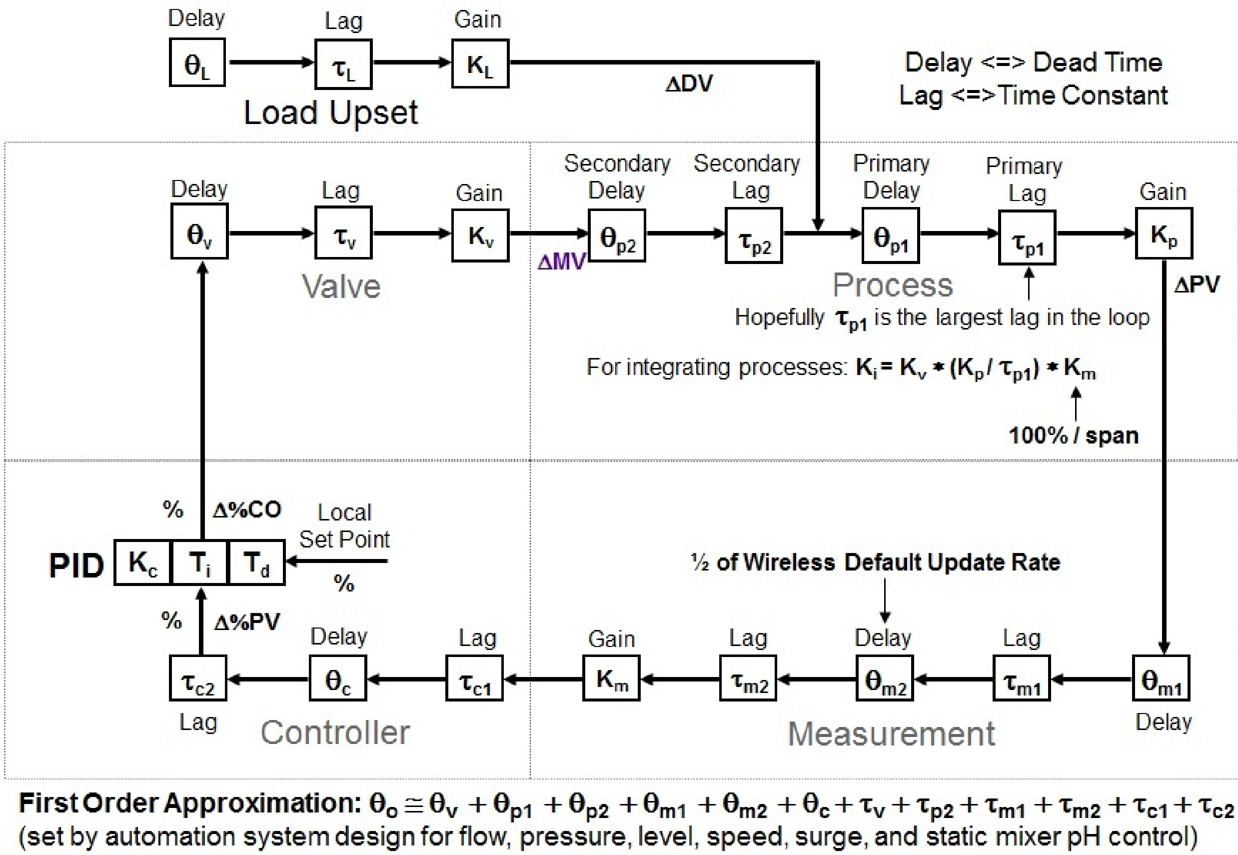


Figure D-4. The open loop gain is the product of the valve, process, and measurement gains, the loop dead time is the sum of all the delays and small lags, and the open loop time constant is the largest lag.

Appendix E – Unification of tuning methods

Appendix C in the ISA book *Advanced Temperature Measurement and Control* shows how the Lambda tuning for self-regulating and integrating processes gives the same equation for controller gain used in the short cut method if Lambda is the set equal to the open loop dead time ($\lambda = \theta_o$). The Ziegler Nichols ultimate oscillation and reaction curve methods yield a similar result except the detuning factor (K_x) used by Ziegler Nichols was 1.0 to provide maximum disturbance rejection ($K_x = 1.0$).

Here we see a further unification by the use of a dead time block delaying the process variable by one loop dead time in the identification of the integrating process gain. Even though the integrating process gain is not applicable to a self-regulating process, the use of the dead time block provides a controller gain that is good for dead time dominant processes extending the utility of the short cut method described in Appendix to essentially all types of process dynamics. For processes with inverse response, the dead time interval should include the time the process is going in the wrong direction so the maximum change in controller output is identified when the process is going in the correct direction.

Identification of integrator gain:

$$K_i = \frac{(\Delta\% PV_{\max} / \Delta t)}{\Delta\% CO_{\max}} \tag{E-1}$$

PID gain for disturbance rejection with detuning factor (K_x):

$$K_c = K_x * \frac{1}{K_i * \theta_o} \tag{E-2}$$

Equation E-1 substituted into Equation E-2:

$$K_c = K_x * \frac{\Delta\% CO_{\max}}{(\Delta\% PV_{\max} / \Delta t) * \theta_o} \tag{E-3}$$

If the time interval is equal to the observed dead time ($\Delta t = \theta_o$) and the $\Delta PV_{\max}$ and $\Delta CO_{\max}$ are created by passing the process variable and control output through a dead time block with the dead time parameter set equal to the observed dead time θ_o, then dead time is canceled out.

Equation E-3 with $\Delta t = \theta_o$:

$$K_c = K_x * \frac{\Delta CO_{\max}}{\Delta PV_{\max}} \tag{E-4}$$

The detuning factor used in the short cut method to provide a practical degree of smoothness and robustness is 0.5 ($K_x = 0.5$), which corresponds to a Lambda equal to the observed dead time ($\lambda = \theta_o$). The factor K_x can be increased to 1.0 with no sacrifice in robustness for the enhanced PID developed for wireless when the update interval is much larger than the process dead time and time constant.

Note that check for maximum change in process variable would begin one dead time after the change in controller output is above a trigger level and continue for four more dead time intervals. The process dead time is identified as the time interval from the start of a controller output change to an observed change in the process variable beyond the noise band. The noise band can be set or automatically identified. The test works for a change in the controller output in manual or for a set point change if the controller gain is high. To ensure a fast update and minimal reaction to noise, it is critical that dead time blocks are used to create a continuous train of delayed process variables and delayed controller output variables for computing the change in these variables over the dead time interval. The method provides experimental models that can be used for rapid deployment of plant-wide dynamic simulations for training and process control improvement. For self-regulating processes, Equation A-7 and A-8 are used.

The method is applicable to dead time dominant self-regulating processes ($\theta_o >> \tau_o$), such as sheet lines, where the controller gain is simply the factored inverse of the open loop process gain.

The open loop gain for a self-regulating process:

$$K_o = \frac{\Delta\%PV_{\max}}{\Delta\%CO_{\max}} \tag{E-5}$$

E-5 substituted into E-4 gives the PID gain for a dead-time dominant self-regulating process ($\theta_o >> \tau_o$):

$$K_c = K_x * \frac{1}{K_o} \tag{E-6}$$