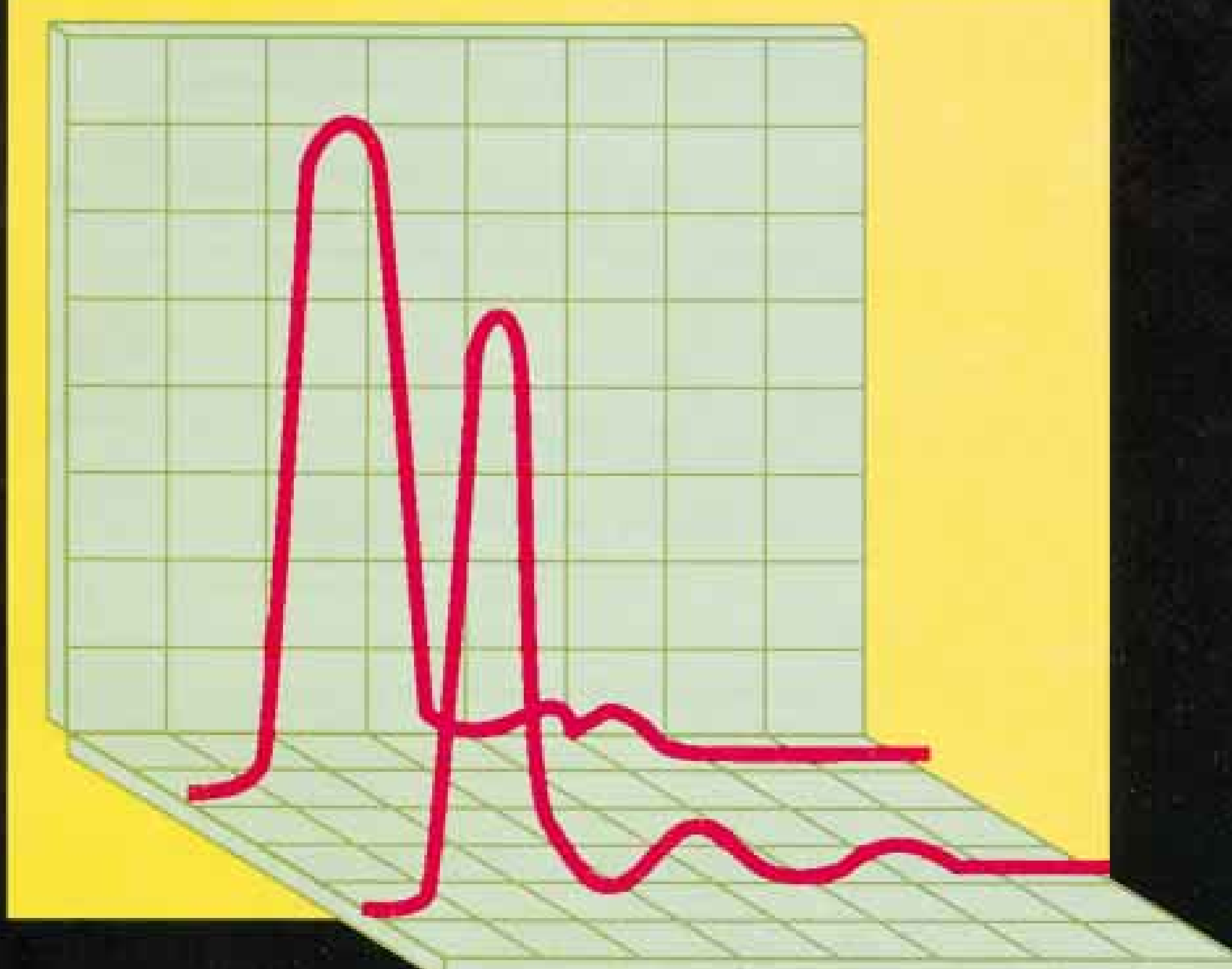
CONTROL ENGINEERING

A CAHNERS PUBLICATION

for designers and users of control and instrumentation equipment and systems world wide

Reference Guide to PID Tuning

Part 3



A collection of reprinted articles of PID tuning techniques

Tuning of PID Controls of Different Structures

A. KAYA, The University of Akron, Akron, O. and T. J. SCHEIB, Bailey Controls Co., Wickliffe, O.

Although readily available as products, PID controllers require tuning to suit a specific process. This article presents new data for optimized tuning of three PID variations.

Proportional-integral-derivative (PID) control is the most used controller type in industry and is available as a stock item. However, its use is so diversified that the control engineer must tune the PID values according to specific needs.

Studies have guided industry by providing quantitative data for tuning PID controllers for the given process, operational conditions, and performance criteria. Unfortunately, these studies considered only one kind of functional PID control structure—the most popular one. Control hardware suppliers, on the other hand, have marketed other PID variations to better meet the requirements of a wide variety of control loops.

Recently, it has become easier to build PID variations due to low-cost microprocessor chip technology. But, the user does not always have sufficient guidance on how to tune these new PID controls. Some are probably using the tuning guides published for one form of PID, while some may not even know the PID form they have purchased. The objective of this article is to provide quantitative data as a tuning guide for three PID variations and to provide the user with an awareness of related industry practices.

Tuning factors

There are three major factors in tuning a PID controller: the process, the controller's form, and the performance criterion (or response) of the feedback control loop. The process and PID controller are shown in Figure 1, while some popular performance criteria are given in equations (5)-(7).

In the published literature, the process, which is usually of high order, is approximated as a first order process with a time delay, having a transfer function as in equation (1):

$$G(s) = \frac{Ke^{-\theta s}}{\tau s + 1}$$

where, K is the process gain; τ is the process time constant, sec; θ is the process dead time (or delay), sec; and s is the Laplace domain variable.

Such approximation is made from a typical process reaction curve, as presented in Figure 2 (Reference 1). It is clear that the maximum slope of

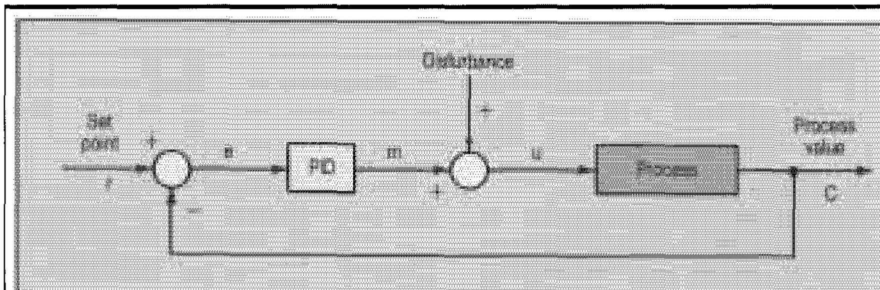


FIG. 1: Structure of the feedback control loop includes the process and PID elements.

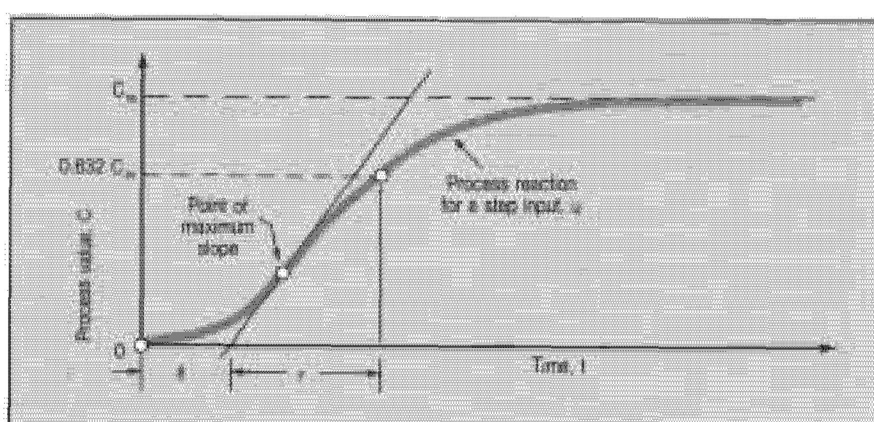


FIG. 2: The process reaction curve, a response to change of controller output, m, is often approximated as a first-order process. Delay time, θ, is found from the curve's maximum slope.

the curve determines the time delay magnitude, θ. Assuming first-order, the process reaction is (equation 2):

$$C(t) = C_{ss}(1 - e^{-t/\tau})$$

Here, for $t = \tau$, $C(\tau) = 0.632C_{ss}$, which defines the τ value from Figure 2. Note that C_{ss} is the steady-state value of the process. Furthermore, the K value is found as $K = (C_{ss}/u)$.

The controller considered in the published literature is an ideal PID control device that has an input/output relation and a transfer function relation as given by the following equations (3 and 4):

$$m = K_c\left(e + \frac{1}{T_i}\int e\,dt + T_d\frac{de}{dt}\right)$$

$$m = K_c\left(1 + \frac{1}{T_i s} + T_d s\right)e$$

where, m is the controller output; K_c is proportional gain; T_i is integral time, sec; and T_d is derivative time, sec.

Popular performance criteria

Next, three widely used controller performance criteria are introduced. Each is based on a different form of minimizing the error integral; namely: the minimum integral of square error (ISE), the minimum integral of absolute error (IAE), and the minimum integral of absolute error multiplied by time (ITAE)—as stated respectively in equations (5), (6), and (7).

$$\text{ISE} = \int_0^\infty e^2\,dt$$

$$\text{IAE} = \int_0^\infty |e|\,dt$$

$$\text{ITAE} = \int_0^\infty t\,|e|\,dt$$

In the above equations, the error is $e = r - c$ (refer to Figure 1). A graphical description of error, e, with respect to time, in Figure 3, helps to more clearly explain the error criteria.

There are other performance criteria used in the process industry. Ziegler and Nichols aimed at an overall desirable response of ¼ decay ratio, or $(b/a) = 0.25$ in Figure 3. Others used a specified closed loop response in terms of response time.

The error, e, (see Figure 1) occurs from the set point changes or the disturbances (load changes) within a process, as shown schematically in Figure 3. Thus, the PID control should be tuned according to which criterion is considered important.

Disturbance or set point tuning

Usually the process is subject to disturbances. However, in many cases a secondary (or cascade) control loop is present within the primary control loop. As the primary loop controller responds to disturbances, it adjusts the set point of the cascade control loop. This supports the idea that controller tuning for either the set point change or for the disturbance change can be important.

The popular tuning criteria, as widely used in industry, have been the ISE, IAE, and ITAE. For an ideal PID controller given by equation (3), PID tuning values of these criteria have been studied and reported about twenty years ago, in References 1 to 3.

In these publications, the user was provided the tuning constants a through f and equations (8) and (9), with which to find the PID control values. Note, that for given process parameter values of K, τ, and θ, the PID control values of K_c, T_i, and T_d are found for either the set point or disturbance type tuning. But, the validity of the published tuning constants is restricted to small and moderate time delays, or $0 < (\theta/\tau) \le 1$.

So, for disturbance tuning (eqs. 8):

$$KK_c = a(\theta/\tau)^b$$

$$\tau/T_i = c(\theta/\tau)^d$$

$$T_d/\tau = e(\theta/\tau)^f$$

and for set point tuning (eqs. 9):

$$KK_c = a(\theta/\tau)^b$$

$$\tau/T_i = c + d(\theta/\tau)$$

$$T_d/\tau = e(\theta/\tau)^f$$

Now, with the development of microprocessors, different PID control functions have come into existence besides the ideal PID control. Among these are control types named: classical, noninteracting, and industrial—as described by the next three equations (10 to 12), respectively. Each control type has transfer function relations similar to equation (4).

Classical:

$$m = K_c\left(1 + \frac{1}{T_i s}\right)\left(\frac{1 + T_d s}{1 + T_a s}\right)e$$

Noninteracting:

$$m = \left(K_c + \frac{1}{T_i s}\right)e - \left(\frac{T_d s}{T_a s + 1}\right)c$$

Industrial:

$$m = K_c\left(1 + \frac{1}{T_i s}\right)\left[r - \left(\frac{T_d s + 1}{T_a s + 1}\right)c\right]$$

In these equations, T_a is a filter time constant, usually defined as,

$$T_a = 0.1\,T_d$$

See Figure 1 for the definition of other significant terms.

There are other PID structures in use in industry whose characteristics are discussed in References 4 and 5. However, the user needs to know the

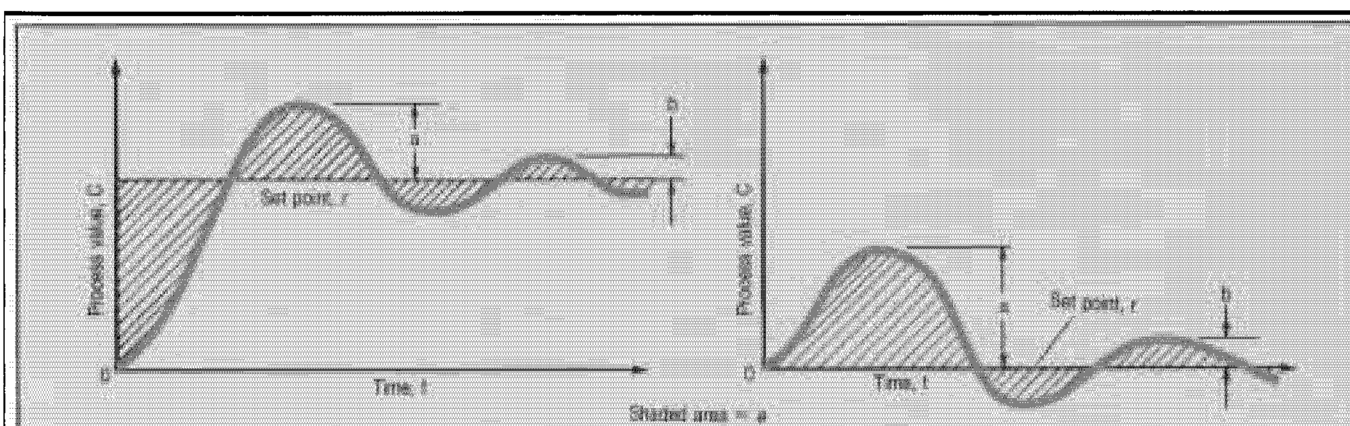


FIG. 3: Two diagrams show typical variations of process value versus time. Response to step change of the set point is illustrated at the left; while response to a step change of disturbance is given at the right. The shaded areas represent a graphical description of error.

OPERATING CONDITION

Const.	Disturbance Tuning			Set Point Tuning		
	Criteria			Criteria		
	ISE	IAE	ITAE	ISE	IAE	ITAE
a	1.11907	0.98089	0.77902	0.71959	0.65	1.12762
b	−0.89711	−0.76167	−1.06401	−1.03092	−1.04432	−0.80368
c	0.7987	0.91032	1.14311	1.12666	0.9895	0.99783
d	−0.9548	−1.05211	−0.70949	−0.18145	0.09539	0.02860
e	0.54766	0.59974	0.57137	0.54568	0.50814	0.42844
f	0.87798	0.89819	1.03826	0.86411	1.08433	1.0081

Table 1: Tuning constants for Classical controller given in equation (10) for $0<(\theta/\tau)\leq 1$.

Const.	Disturbance Tuning			Set Point Tuning		
	Criteria			Criteria		
	ISE	IAE	ITAE	ISE	IAE	ITAE
a	1.3466	1.31509	1.3176	1.26239	1.13031	0.98384
b	−0.9308	−0.8826	−0.7937	−0.8388	−0.81314	−0.49851
c	1.6585	1.2587	1.12499	6.0356	5.7527	2.71348
d	−1.25738	−1.3756	−1.42603	−6.0191	−5.7241	−2.29778
e	0.79715	0.5655	0.49547	0.47617	0.32175	0.21443
f	0.41941	0.4576	0.41932	0.24572	0.17707	0.16768

Table 2: Tuning constants for Noninteracting controller given in equation (11) for $0<(\theta/\tau)\leq 1$.

Const.	Disturbance Tuning			Set Point Tuning		
	Criteria			Criteria		
	ISE	IAE	ITAE	ISE	IAE	ITAE
a	1.1147	0.91	0.7058	1.1427	0.81699	0.8326
b	−0.8992	−0.7938	−0.8872	−0.9365	−1.004	−0.7607
c	0.9324	1.01495	1.03326	0.99223	1.09112	1.00268
d	−0.8753	−1.00403	−0.99138	−0.35269	−0.22387	0.00854
e	0.56508	0.5414	0.60006	0.35308	0.44278	0.44243
f	0.91107	0.7848	0.971	0.78088	0.97186	1.11499

Table 3: Tuning constants for Industrial controller given in equation (12) for $0<(\theta/\tau)\leq 1$.

quantitative values of K_c, T_i, and T_d of PID controls for each case (*i.e.*, the process, the PID function, and the performance criterion) to obtain an optimum operation.

The purpose of the work reported in this article was to analyze and find those optimum tuning values for the new PID control variations presented in equations (10) through (12), and to compare their performances.

Optimization procedure

The processes on which the optimum tuning constants were developed had specification parameters of $\tau = 30$ and $K = 1$. Delay time was varied from $\theta = 3$ to 30 sec. Such processes and controllers were simulated by ACSL (Advanced Continuous System Language).

The optimization procedure was conducted for $(\theta/\tau) = 0.1$, 0.5, and 1.0. A search procedure in the maximum gradient direction was chosen as the controller values were changed to find the minimum value of the Performance Index.

Control values for each (θ/τ) ratio were fitted into the curves of equations (8) and (9) to find the tuning constants *a* through *f*. In doing this, the minimum least squares calculation method was utilized.

To validate the optimum procedure, the ideal PID control was first used to generate the tuning values and compare them with those published 20 years ago. Due to potential variation in steps, up to a ten percent difference between the values of the Performance Index was considered small and such results were considered compatible. The procedure of the study was validated in this manner.

Next, the same procedure was used for the various new PID controls and their corresponding tuning constant values, *a* through *f*, were found. Then, the values of error criteria were compared, when the different PID controls were tuned by the published tuning values of the ideal controller and by those found in this study.

New, optimum tuning values

The tuning constants, *a-f*, for the new PID controls are significantly different from those published for the ideal controllers. The tuning constants developed in this work are given for each controller type, in Tables 1 to 3.

In order to substantiate the superiority of the newly-found optimum tuning constants over those previously published for the ideal controller, tests were conducted to evaluate the values of the error criteria. The tests were conducted by using the new controllers on a process with $(\theta/\tau) = 0.5$. The new PID controllers were tuned using published data—for the ideal PID controller—and by using the data found in this work; then, the corresponding values of error criteria were compared.

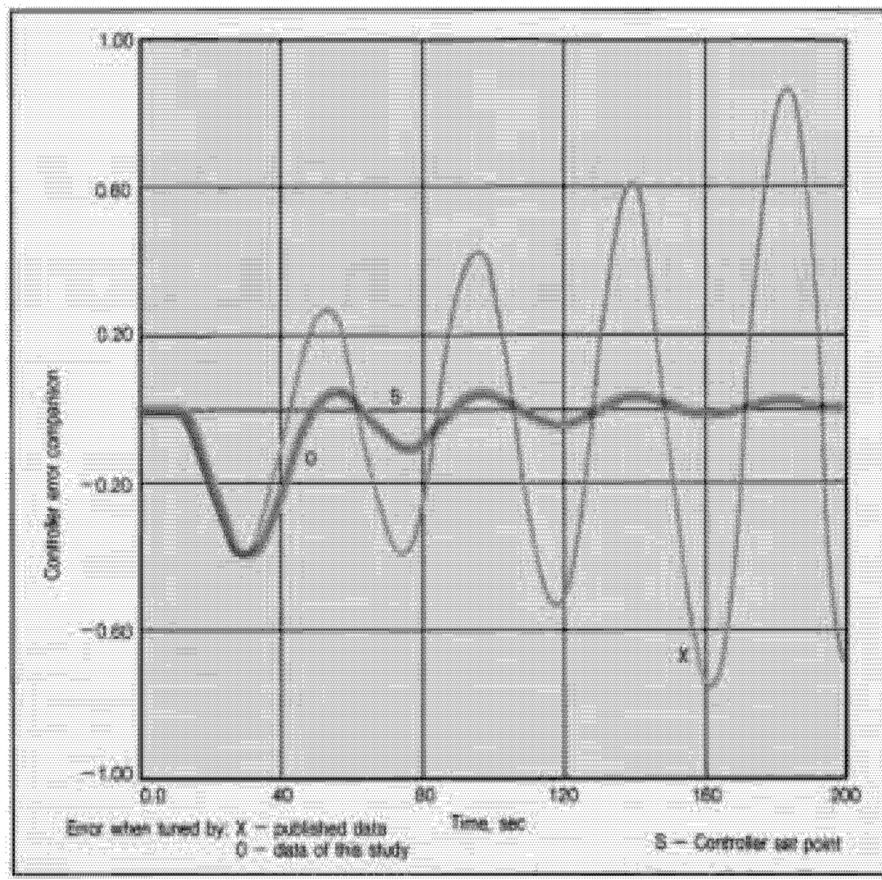


FIG. 4: The above graph is an error comparison example of the Classical controller tuned for: ISE error criterion, a step change of disturbance input, and a process with $(\theta/\tau) = 0.5$.

The comparison indicates that in most cases, the values of error criteria are significantly lower when tuned by the optimum values found in this study, as opposed to those published for ideal controllers. As an example, one graphical comparison is shown in Figure 4. Here, the classical controller was tuned for the ISE criterion and for an input that corresponds to a step change of disturbance (see Figure 3, right). Note that the process response is unstable when tuned by the published data.

Conclusions

Some suppliers of commercial controllers disclose the control functions of PID controls (References 4 and 5). Although the functions are described by different formats, they can be related to the forms of equations (4) and (10) to (12), as given earlier.

It is recommended that the new tuning constants presented here be adopted for the PID controllers with new functional relations. Furthermore, the analysis should be extended to the case of longer time delays, where $(\theta/\tau) > 1$. Since the published literature has dealt only with the case of $0 < (\theta/\tau) \leq 1$, the comparative analyses of this work are also restricted to the same interval.

It is recommended that the case of longer process time delays be studied further. As a corollary, additional guidance should be given to the users as to what PID control structures should be employed with processes having such long delay times. □

REFERENCES

1. Lopez, A. M., et al., "Tuning Controllers with Error-Integral Criteria," *Instrumentation Technology*, November 1967, pp. 57-62.
2. Miller, J. A., et al., "A Comparison of Controller Tuning Techniques," CONTROL ENGINEERING, December 1967, pp. 72-75.
3. Rovira, A. A., et al., "Tuning Controllers for Set Point Changes," *Instruments and Control Systems*, December 1969, pp. 67-69.
4. Gerry, J. P., "A Comparison of PID Control Algorithms," CONTROL ENGINEERING, March 1987, pp. 102-105.
5. "Function Code Application Manual," *IE93-900-20*, Bailey Controls Co., Wickliffe, O., 1986.

Pneumatic Instruments Gave Birth to Automatic Control

MICHAEL BABB, CONTROL ENGINEERING

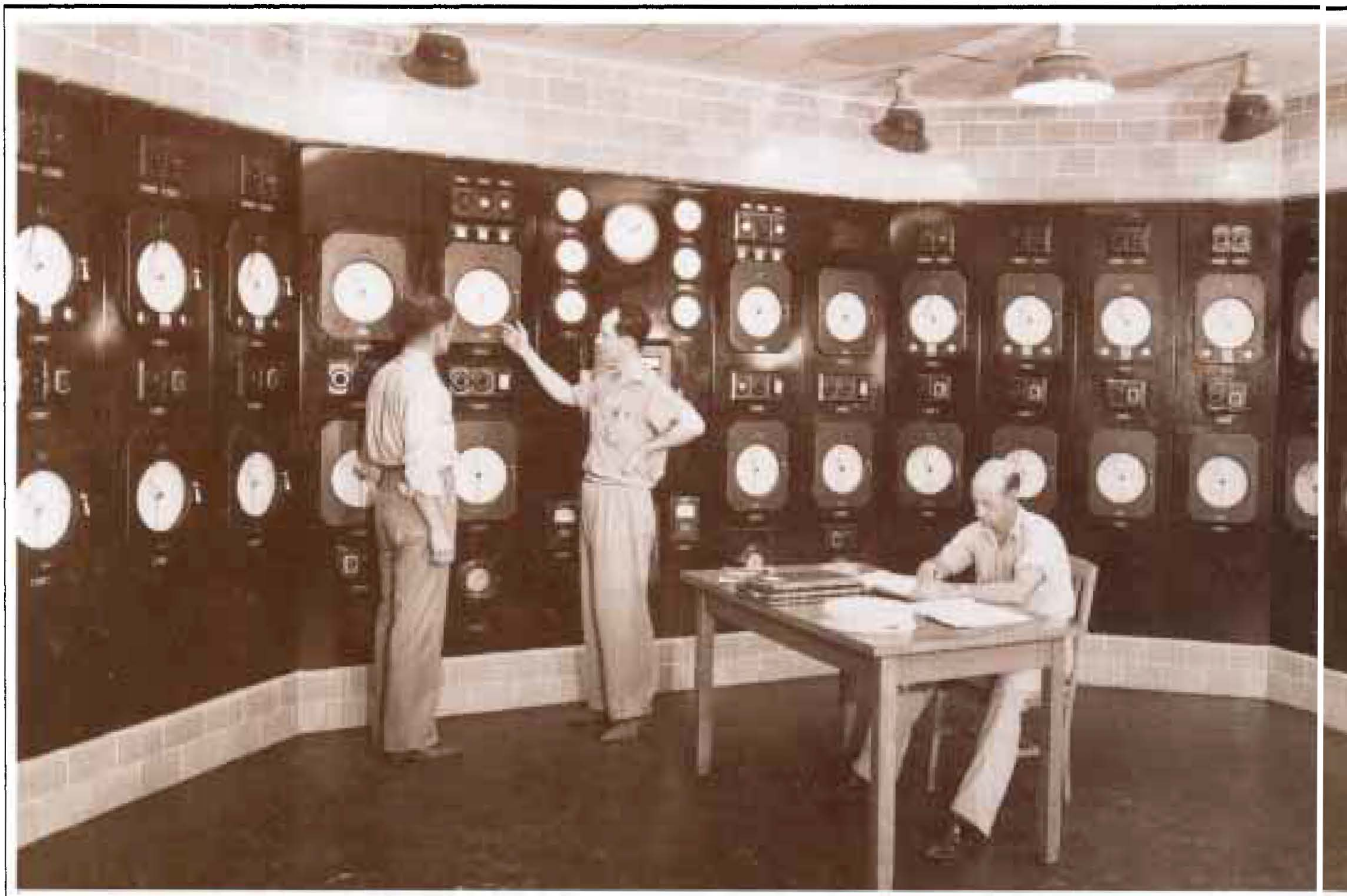
When engineers discovered that big valves could be delicately positioned with little puffs of air, they were on their way to inventing automatic control.

For most Americans, companies such as Boeing Aircraft, Union Carbide, and Dow Chemical are household words. During the heady days of WW II manufacturing, they rallied their work forces to new pinnacles of productivity, and made a contribution to the war effort that has received high recognition ever since.

However, it is unlikely that many Americans have ever heard of companies like Brown, Taylor, Foxboro, or Leeds & Northrup. Fewer still would recognize the contribution these behind-the-scenes instrument makers made, not just to WW II production, but to the very way we carry on manufacturing in the 1990s.

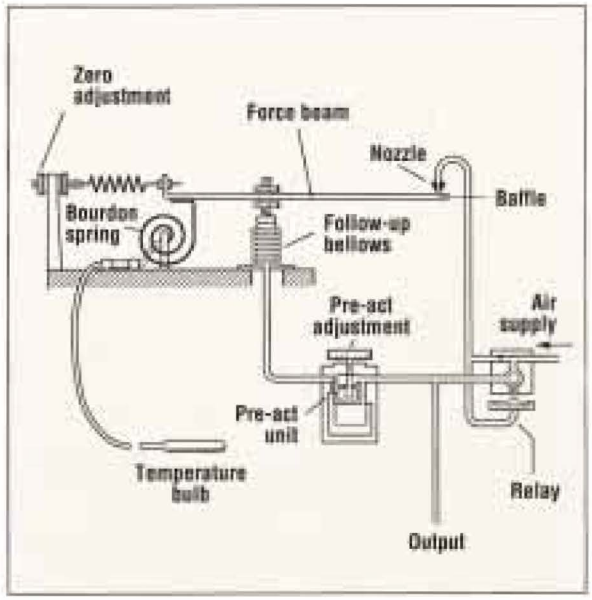
What exactly did engineers from these companies do? About 50 years ago, a handful of them, probably less than a few dozen, all told, invented what we now call automatic control.

The instruments they built—complex air-powered mechanical gadgets that mystified the layman and baffled even the operators who used them—paved the way for a new and revolutionary method for producing the basic commodities of our industrial age.



Their new method readily adapted itself to liquids and gases, or anything, for that matter, that flows through a pipe. It was called *continuous-flow processing*.

The fruits of their labors made it possible for the allied forces in WW II to obtain tremendously large quantities of synthetic rubber and high-octane aviation fuel. And tens of thousands of their mysterious instrument controllers, working in concert at a single super-secret processing site, managed to squeeze out a few pounds of U-235 from gaseous uranium hexafluoride. Just enough, in fact, to begin the atomic age.



The heart, and lungs, of early proportional control is this force-balance mechanism. A temperature increase uncoils the Bourdon spring, moving the beam down to increase the nozzle baffle gap. The movement—a few thousandths of an inch at the most—decreases the nozzle back pressure, which through the relay causes an proportional increase in output pressure. At the same time, the increase in output pressure is fed back to the follow-up bellows, which restores the force beam to its original position.

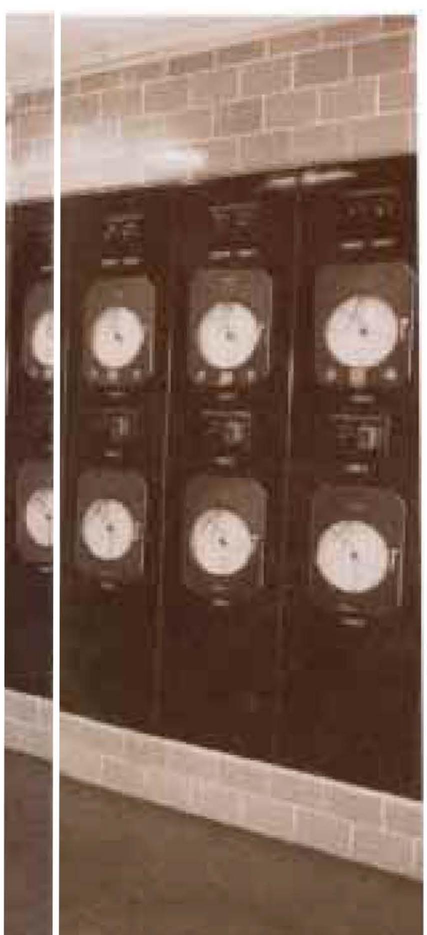
The "Pre-Act" unit was Taylor's innovation. Its effect is to delay the feedback to the follow-up bellows so that the force beam does not come immediately back into equilibrium. This gives the controller output an extra "boost" whenever a change is made, a separate control action that became known as "derivative" control.

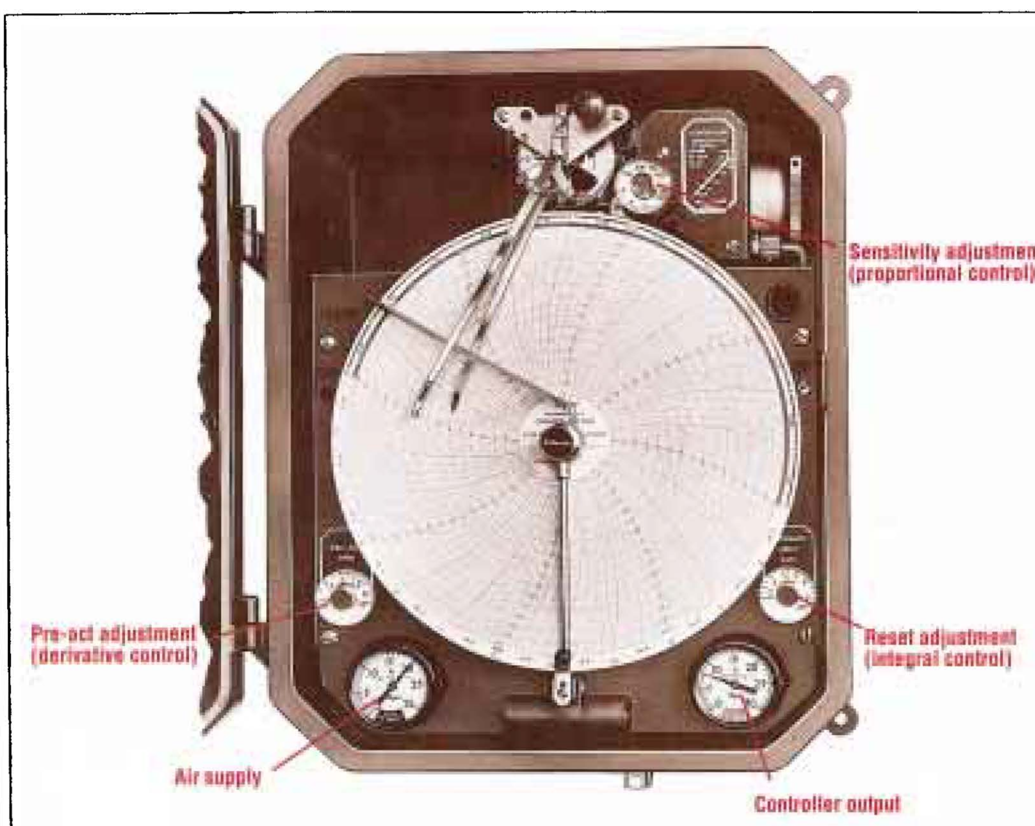
It all started with milk

Pasteurization is not considered a great engineering feat. The temperature of a batch of milk is raised to 143 deg F and held there for 30 minutes. But pasteurization was, in fact, one of the first applications for process control. It was also one of the first processes that came under government control.

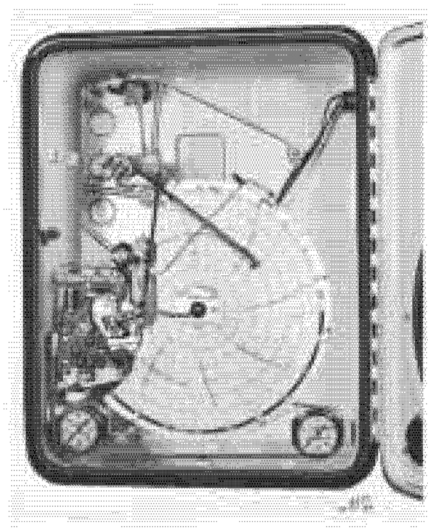
Which probably explains why, about 1912, the dairy industry figured out that it had better start keeping records of the pasteurization of batches of milk. At that time, practically no one recorded temperatures, and so thermometer manufacturer Taylor Instruments, based in Rochester, N.Y., didn't sell many of the mercury-recording thermometers they made.

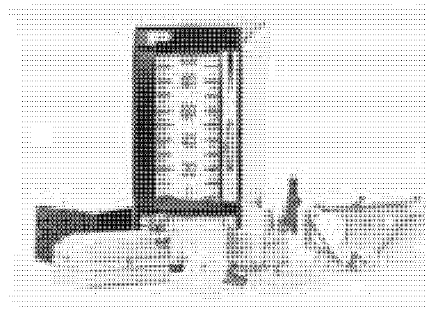
The market changed in a hurry when Iowa-based dairy equipment manufacturer Cherry-Burrel put in an order with Taylor for 50 temperature





(Above) Taylor Instruments' technology breakthrough, the Fulscope controller, was the first PID controller when it was introduced in 1940. At left, a battery of Fulscopes controls the sensitive temperature process of synthetic rubber making in Copolymer Corp.'s Baton Rouge, La., facility. A supervisor looks for perfect circles on the trend charts, an indication of good control.





(Above) Fischer & Porter did not introduce pneumatic instrumentation until the mid-1940s; shown here is their "Concept 45" line (ca. 1960). (Left) The Bailey Meter Co. combined "Pilotrol" (proportional) and "Standatrol" (integral) units into one controller. Bailey was an early controls industry leader for combustion and boiler control applications.

recorders. A windfall was in the making. Taylor engineers and salesmen recognized the vertical opportunity and moved quickly to capitalize on it.

The original order in 1912 was the beginning of decades of domination of the dairy industry by Taylor. From their factory in Rochester came all sorts of customized temperature recorders, housed in ruggedized packages to withstand the daily cleaning-in-place. Taylor's temperature recorders became the *de facto* standard of the dairy industry. Competitors Foxboro and Brown were virtually locked out. Even as late as 1940, dairy sales of $625,000 accounted for 17% of Taylor's industrial business, according to *Fortune* magazine.

With its success in the dairy business, Taylor learned an important lesson in controls marketing. Once an end-user company buys into a specific instrument line, and the operators become familiar with them, it is unlikely that they will ever switch to another brand. It is, in fact, virtually impossible to uproot an established instrument vendor; better to move in fast and lock out the competition early in the game.

But Taylor didn't rest on its laurels. Pasteurization technology changed in the mid-1930s. A new, nearly-continuous "flash" process heated the milk to 160 deg F for 15 seconds. To accommodate the speeded-up requirements placed on the operator, Taylor invented a quick-acting diverting valve to flush the milk back into the heating tank if the temperature were not precisely maintained. The instrument was, in fact, directly and automatically controlling a part of the process.

Continuous-flow processing

The dairy process had been greatly accelerated, but it still was a batch operation, as were all other chemical and refining processes in the early part of this century.

Probably the first real move towards continuous-flow processing was taken around 1925 by Carbide & Carbon Chemicals. The company was experimenting with fractionation of natural gas in their West Virginia plant, and found they could produce new synthetic chemicals, such as ethyl alhohol and ethylene glycol. However, storing natural gas is a more cumbersome affair than storing, say, crude oil, and so Carbide was on the lookout for ways to speed the natural gas through their crackers more expeditiously. Maybe they wouldn't have to store gas at all.

Taylor was hired to supply the recorders and a few controllers. But the technology team didn't work out. For one thing, Carbide was too secretive about its processes, and Taylor engineers didn't get a good grasp on the situation, as they had done in the dairy industry. Also, there was a major technical barrier: the gas pipes required bigger and more tightly packed valves. It was hard to open and shut them, and Taylor's pneumatic actuator didn't do a good job. By 1933, Taylor came up with a suitable valve positioner. But Carbide operators had become skilled in manual continuous-flow operation, and didn't want to hear about it.

A Brown Instruments advertisememt in 1944.

Meanwhile, deep in the heart of Texas, things were booming. The petroleum refiners were buying up all the crude the oil fields could produce. By 1929, they began to look at continuous-flow as a means of increasing their competitive advantage. They built continuous-cracking furnaces to handle the job, but due to the increased speed of the operation, refiners had to lean on instrument control more heavily than ever before. They also had a great need for more accurate flow measurement.

Enter Foxboro

For years, Foxboro had been making instrumentation for the oil fields. It was practically the only company with enough flowmeter technology to get the job done.

But now Foxboro had a good opportunity to expand its operation into the refineries. It didn't take long for entire complexes to become completely outfitted with Foxboro instrument rigs. Taylor didn't even make a flowmeter, so they had no opportunity to compete. Meanwhile, Foxboro's instruments and Stabilog controllers, housed in circular cases, became the *de facto* standard of the oil, gas, and refining industries.

The business was so good that by 1933, Brown introduced new instruments and had a go at it. While Foxboro dominated, there was enough business opportunity for Brown to also have a significant share.

While Brown and Foxboro went after the booming refining business, Taylor opted to go down a different path. It chose instead to concentrate on developing and improving the Fulscope controller.

In the mid-1930s, Taylor designed the "double response" unit which stabilized control action by providing for automatic valve reset. It was an independent unit attached directly to the valve stem, but when Taylor redesigned the Fulscope, it was modified and integrated into the controller housing. A later generation of control engineers would refer to the reset as "intergal mode" control.

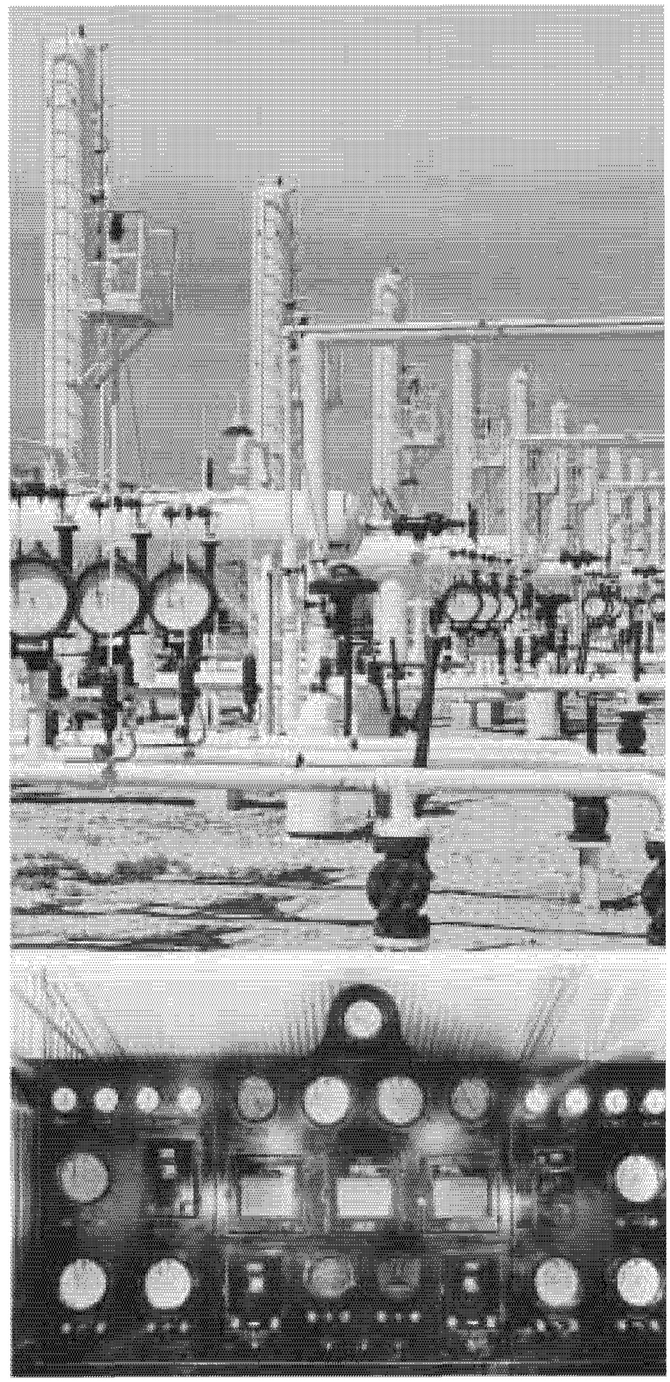
It was in 1938 that Taylor engineers were monkeying around with their controller and came across something new. John Ziegler describes the process of discovery:

"Taylor around this time was working the viscose rayon industry, trying to control the rayon shredder which was one of the god-awfulest pieces of chemical engineering equipment ever devised. The proportional Fulscope with the double response unit would not work. Someone in the research department was tinkering with Fulscopes and somehow had got a restriction in the feedback line to the capsule that made the follow-up in the bellows. He noted that this gave a strange 'kicking' action to the output. They tried this on the rayon shredders and it gave perfect control on the temperature. The action was dubbed 'Pre-Act' and was found to help the control in other difficult applications, like refinery tube stills.

"The Pre-Act was the first derivative control and was incorporated into the Model 56R. It worked great on juice units in the sugar industry, but not so great in other applications."

Taylor then set out to design the Model 100 Fulscope. It was to incorporate the automatic reset action provided by the double response unit, plus the new Pre-Act unit.

When the redesigned Fulscope went into service in 1940, it was a technology breakthrough, the first three-mode controller on the market.



Foxboro's circular instruments and Stabilog controllers were commonplace in the Texas oil and gas fields during the 1920s, and migrated into refinery control rooms and 1930s (photo: Premier Oil Refining Co.).

There was a great deal of clever mechanical engineering that went into the Fulscope. Special needle valves were designed for setting the reset and Pre-Act rates. A distinctive parallelelogram linkage was configured for continuous adjustment of sensitivity, or "gain" as it was later known. Although not advertised as such, Taylor made sure their instrument was operator-friendly. Simple knob controls were all that was needed to adjust the three modes, another industry first. By contrast, controllers from Brown and Foxboro were relatively difficult to tune and repair.

With its new controller, Taylor broke into Foxboro's Texas turf, notably at Texaco and Humble Oil refineries.

The Fulscope had one major drawback, however. Setting the tuning parameters for a two-mode controller had been difficult, but with the third mode, tuning became a major obstacle.

Taylor engineers John Ziegler and Nathaniel Nichols went to work on the problem. By 1941 they had a relatively straightforward method for tuning the three-mode, or PID, controller. Basically, the method involves increasing the sensitivity (proportional response) until a sustained oscillation is obtained. Ziegler and Nichols called this the "ultimate sensitivity." Then, with the proportional adjustment set to one-half the value that caused the ultimate sensitivity, set the reset (integral) rate equal to the ultimate frequency and set the Pre-Act (derivative) to 1/8 of this frequency. This basic tuning formula, with variations, became known as the "Ziegler-Nichols" method. It is a classic procedure that appears in every textbook on control engineering and continues to be used for PID controllers in our day.

Milk to atom bombs

WW II came and the instrument panic was on. *Fortune* magazine reported, "Synthetic rubber struggled through 1941 on a 10,000-ton-a-year budget, then doubled and redoubled until drafting boards sank below the waves of blueprints for fifty-odd units to cost $750 million and produce over a million tons of synthetic rubber a year."

Carbide took on the unlikely butadiene-from-alcohol assignment, the key ingredient of synthetic rubber. It boils only 1.9 deg C from its nearest chemical neighbor, butene-1. This required very tight temperature control to draw them apart. "Companies that had never let a Taylor or Foxboro or Brown instrument through the gates," reported *Fortune*, "now scrambled their neat panel boards with a hodgepodge of instruments. The complexity and velocity of the new processes had outrun the ability of human operators to control them."

Similarly, to provide fuel for Roosevelt's 100,000 aircraft per year program, the government poured $1 billion in petroleum refineries. *Fortune* reported that production increased from 30,000 barrels a day in 1940 to 580,000 a day by 1945. "Instruments for the high-octane and synthetic rubber programs combined ran to a total of perhaps $40 million," reported *Fortune*. "Foxboro got the biggest share of this business, probably as much as both Brown and Taylor combined, with Taylor on the short end at about 15%, or $6 million."

But the Fulscope's great day was yet to come. In May, 1943, key Taylor executives met with representatives from the Manhattan Project in the Woolworth Building in New York City, who explained what they needed: a method for controlling the flow of gaseous uranium hexafluoride as it passed through 4,000 barriers, with pumps and instrument control at every stage. Some of the instruments would have to measure the "violently corrosive" gas in ranges of 0 to 0.125 psi; the best at that time measured 0 to 5 psi. It was the "trickiest flow job in history." And Foxboro, the acknowledged leader in flow control, had been passed over, probably because it was too busy producing instruments for other industries.

The job was to be done in Oak Ridge, Tenn. Kellex was contracted to built an 11-mile long gaseous diffusion plant along the Clinch River. The plant, named K-25, would require a staggering total of 200,000 instruments, of which 40,000 were actual controllers. The numbers amounted to about 25% of all the instruments made in the U.S. during WW II.

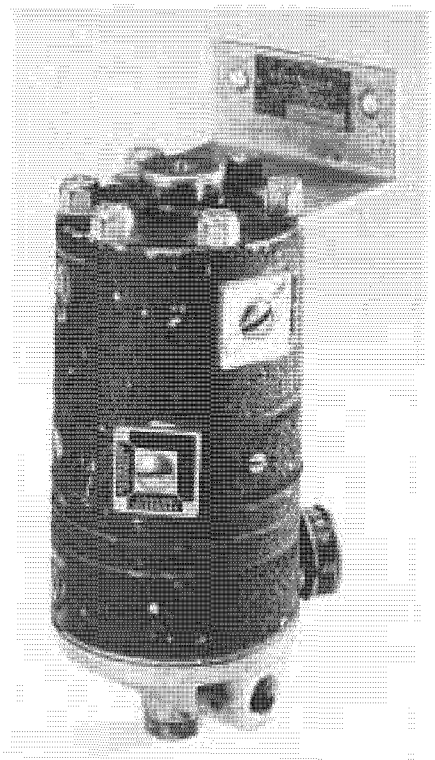
Taylor's response was to invent a new basic instrument, the pressure transmitter, and to produce them in prodigious quantities. Kellex ordered 30,000 of them, in 27 ranges. Taylor developed another twenty specials, and Kellex bought 70,000 of them.

When finished, K-25 had no less than 11 miles of instrument panels. So automated was the operation that Carbide, who ran the plant for Oak Ridge, bragged "only twenty operators per mile are needed . . . a figure that may get whittled down to ten." The end result was a few pounds of U-235, enough to build one bomb.

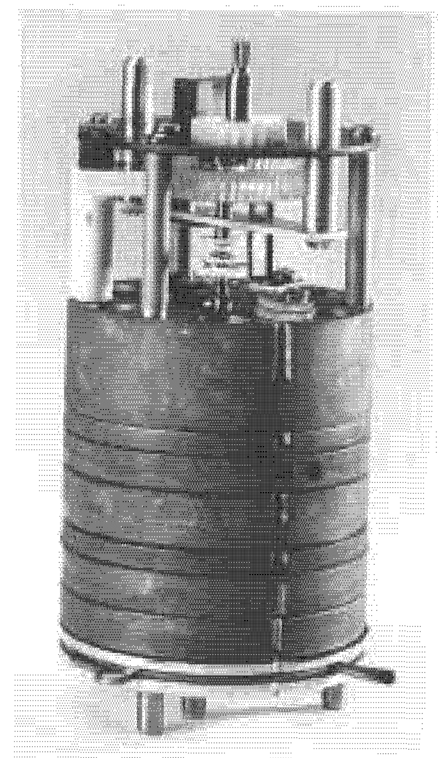
The postwar era

Taylor's monumental efforts to supply instrumentation for the Manhattan Project doubled its industrial sales volume to $8 million in the first peacetime year, and for the first time topped Brown and Foxboro. But there were new worlds to conquer.

Leeds & Northrup had been the leader in measuring temperature electrically. In 1941, Brown (purchased by Minneapolis-Honeywell in 1936) introduced its line of thermocouple-based "ElectroniK" temperature recorders and controllers. It caught Foxboro flatfooted and overturned the controls business in the refineries. At war's end, Foxboro rushed to develop its own line of electronic controllers, and the battle was on.



Moore Products created a stir in 1946 with this "miniature" 5-in. Nullmatic controller.



One of the last great pneumatic control innovations was in 1965: Moore's Syncro Station.

In the meantime, one of Taylor's wartime subcontractors, Moore Products Co., moved into the pneumatic foreground. Coleman B. Moore had left Brown Instrument in 1939 to form his own company, and by 1946, produced his first control product of fame: the Nullmatic "stack" controller. Measuring only 5 inches on a side, the Nullmatic eschewed the familiar circular pen-and-chart recorder; control engineers, C. B. Moore figured, had learned to trust their instruments, and didn't require all the record keeping. The Nullmatic allowed construction of dense control panels.

C. B. Moore burned dollar bills to prove electronic controllers were unsafe.

C. B. Moore went on to become a leading exponent of pneumatic controls. Electronic controllers started gaining ground in the 1950s, and their safety had him worried. To demonstrate the inherent danger of electronics, and to amuse ISA audiences, he would dip his hand or a dollar bill into alcohol and set it afire with an electric current.

The one final significant achievement in pneumatic controls was the introduction, by Moore Products in 1965, of the Syncro Station. It was—and still is, for many of them are still in use—a self-synchronizing controller which allowed simple, bumpless transfer from automatic to manual control. The Syncro has a "pneumatic memory" which is essentially a mechanical memory built with a fluidic device with an air turbine driven lead screw.

Today, the great era of pneumatic controls has gone by. The electronic controllers that displaced them have themselves been largely replaced by digital electronics.

But, even though research and development had dwindled to nothing, pneumatic devices are still with us, and in much larger quantities than most would suspect. Their inherent safety is still unquestioned, and their durability virtually unchallenged. Many devices operating today are literally decades old, and going strong. □

PID Controller Tuning Using Standard Form Optimization

MICHAEL J. GILBERT POLONYI, Stridiron Services, Maspeth, N.Y.

A set of algebraic equations has been developed for dynamically exact controller settings based on loop information and the choice of a preferred transient response.

Tuning a process controller can be a frustrating experience because of not knowing if and when the setting was adequate—and just turning away, hoping that the complaints would turn away also. This article proposes some guidelines that may help in the task.

PID-controller tuning has remained an obscure art rather than an exact science. The reasons for this are simple: nobody really knows what the settings should be since all criteria are qualitative in nature, and, due to the large self-regulatory capacity of most process systems, the margin of tolerance is high and accurate settings are not necessary.

A fast response is always desirable. In practice, however, you may not get good results due to the unknown nature of many parameters—the system may become unstable, or stop controlling altogether. A slow response may then be necessary. (Another good reason to choose a slow response is that it is more efficient in energy terms.)

A variety of transient responses to step changes are possible, and a number of the responses that are considered standard are described in the reference. Choosing a response from these standard forms for a specific behavior amount to optimization and is known as Standard Form Optimization, or SFO. By using SFO, a set of algebraic equations can be derived for each PID-controller. A few simple examples of settings obtained by this method are shown in the table on pg. 106. These equations can also be used for the design of a self-tuning controller by incorporating a time constant identification algorithm into the hardware of the controller.

System time constants

The SFO method recommended here requires at least approximate knowledge of the system's three largest time constants. This is not a big problem, once the basic understanding of determining a time constant is mastered. Time constant determination has remained a rather esoteric concept in process control, and its usefulness is not always duly recognized. There are three basic time constants: first-order, integration, and dead-time.

Advantages of SFO

- A broad choice of responses are available such as fast, slow, or energy efficient. These can be as described as the standard responses shown in the reference such as the integral of absolute error multiplied by time (ITAE), Butterworth, binomial, etc.
- The elimination of the uncertainty which comes with the use of empirical methods, effective only for simple control loops.
- The elimination of assumptions that do not necessarily occur, such as the number of time constants or the relations between the spectral and transient responses.
- Settings that are mathematically exact which are derived from a set of algebraic equations.
- An easy-to-method which eliminates the need for computer simulation or analysis.
- Setting for either two-term (PI) or three term (PID) controllers.
- Readily available transient response curves.

First-order and integration time constants—A first-order time constant T is 0.632 of the final value when a first-order system lag responds to a step input. A first order system is the same as an integrator on a feedback loop. Therefore, the integration time constant and a first-order time constant have the same value. Actually, an integrator is an idealization of a first-order system, because real integrators can exist only within a certain range, or else the output would grow unbound.

Even if it were physically possible for the output to grow unbound without saturating, the system would eventually force the integrator to wind down into a first-order type of response. Take for example the case of the integration time constant of a level control application—if the level were to increase unbound, eventually it would stop because the system pump pressure would equal the head of the liquid column. In other words, at some point the system will start experiencing some feedback of its own.

Dead-time time constant—Dead time is usually introduced into the control loop by the inability to measure the development of a process variable as it occurs. For example, to measure the effect on the temperature of a fluid coming from a heat exchanger, the temperature may have to be measured at a point downstream of the process to which the fluid is going, and it may take the fluid considerable time to reach that point. Another case is where chlorine is mixed in an open channel to kill bacteria. The residual chlorine sensor can be installed in the channel only at a point where the process has already been completed, and this unavoidably creates a considerable dead-time lag. Obviously, dead-time will also change if the velocity of the fluid changes, *i.e.*, at different loads. This introduces an additional complication in the model, and the controller will have to be adaptive if this effect must be accounted for.

Dead-time is fairly simple to measure in a control loop, because even when it can happen in any part of the loop, it will simply add up, and exactly where it originates is of no concern. In other words, there is only one dead-time constant per loop.

If dead-time is significant, *i.e.*, one of the three largest time constants, it must be accounted for by using appropriate equations for tuning controllers. It is not that dead time is hard to handle (especially with present controller technology) but it certainly affects adversely the speed of the response and, therefore, the quality of the control loop.

Dead time can be observed and measured by manually introducing a change on an otherwise stable operation, and by measuring how long it takes to observe a response.

Recognizing a time constant

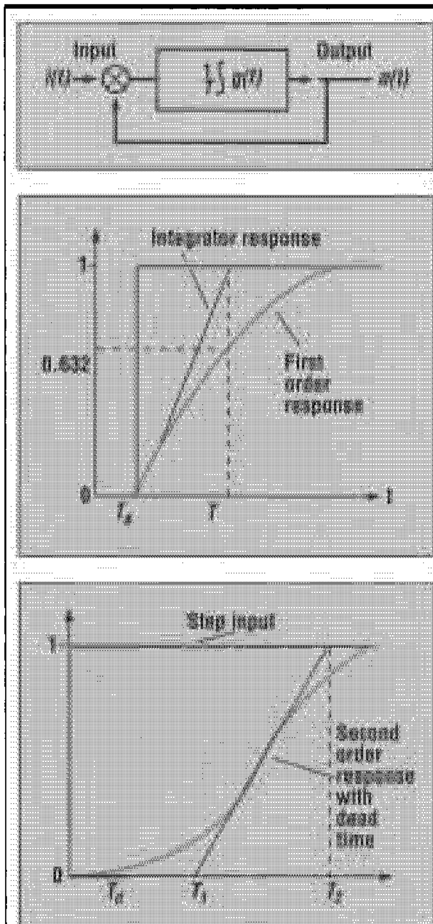
Think of it this way—anywhere there is a time delay, there is a time constant. For example, heat exchangers, boiler furnaces, and nuclear reactors typically have first-order time constants between 10 to 20 sec.

Level control applications also have typical 10- to 20-sec time constants, but they can be easily changed since they depend on a very arbitrary choice—al-lowed level change.

Pressure time constants are usually much larger. For steam boilers, evaporator main time constants range from 50 sec for once-through (supercritical), 200 to 400 sec for water-tube models, and 30 to 45 min for fire-tube models.

Moments of inertia (relative to their maximum load for large rotating machinery such as turbogenerators, pumps, and electric motors) are also integration time constants, and can range typically between 5 and 20 sec. The heavier the shaft respective to the maximum load, the higher the relative moment of inertia.

With the exception of the moment of inertia, electric motors and generators have several electric time constants that can be ignored when dealing with electromechanical variations. They usually become significant when trying to control voltage and reactive power flow.



A first-order system (block diagram at top) is an integrator with a unit feedback loop. The first-order time constant T is 63.2% of the final value, and varies from the second-order.

When controller settings and response of the system are of major concern, main time constants can be determined through some relatively simple tests if they are large enough, *i.e.*, more than 5 sec. For smaller values, electronic measurements and instruments are necessary.

Process equipment designers should be the first source to inquire when trying to determine time constant values, but usually they are themselves ignorant of this information. In any case, dynamic models and time constant information must be analyzed, described how it was obtained, and under what operating circumstances that it applies, before approving the equipment. Otherwise, the whole concept may be flawed.

Any real system (mechanical, electrical or chemical) has many time constants. Fortunately, we are interested only in the largest ones. The higher the time constants, the slower the system will respond, which is good for the control engineer because there is plenty of time to take control action. From a process point of view, this may be a limiting factor since fast load changes will not be responded to; *i.e.*, the faster the natural response of the system the more accurate the controller action, and the settings, must be.

The natural response of a boiler, reactor, and so forth, depends on its physical design characteristics. For example, a water-tube boiler will have a smaller main time constant than a fire-tube boiler, because the water content of the first is smaller than the second, for the same steam production, *i.e.*, load.

Primary design rule

A good first rule is to design with high time constants for equipment in processes that require slow changing loads. This is especially true if there are sudden bursts (of load) that average out after a while. Also, design with low time constants if the process needs to respond primarily to load ramps.

For example, to control an exothermic reaction in a batch process, temperature increases should be as slow as necessary in order to allow a stable heat dissipation. This will depend on the chemical reaction itself and the volume of the reactor. A large volume may cause hot spots and product burning. Therefore, the reaction itself can be slowed down by adding the reagents slowly enough in the large-volume reactor. The exothermic reactor time constants, in this case, will be:

- The time it takes for the temperature to reach its lower limit when addition of reactant has stopped and cooling is fully open;
- A rising temperature time constant when cooling is off and reactant valve is fully open. It must be larger (*i.e.*, slower) than above in order to have a safe control setup.

Note that a smaller time constant means here to have a larger cooling than heating capacity—absolutely necessary to avoid a runaway temperature condition. Only the smaller cooling time constant counts for the adjustment of the controller settings. The ratio between the time constants above will determine the temperature swings during the reaction process. If temperature must be kept within small boundaries, a much larger cooling capacity will be required.

Transmitter time constants

Measurement is a crucial link in control. If data coming from a transmitter are not reliable, they are useless. Data distortion can occur due to various reasons as follows:

- Poor mechanical installation;

- Environmental specifications are exceeded;
- Poor choice of instrument;
- Poor calibration stability (drift);
- Too much relative instrument error;
- Poor physical location.

Temperature sensors: Besides the measurment dead time due to installation limitations as described previously, temperature sensing elements can have significant first-order time lags (in the order of 50 sec) especially if they are installed in thermowells which introduce additional time lags. A bare thermocouple will give the fastest—practically instantaneous—response, while a filled-system bulb in a well will give probably the slowest.

Flow sensors: Although flow measurement is always technically problematic, the time lags (those associated with the instruments and installation) are small, especially where flow of liquids is involved.

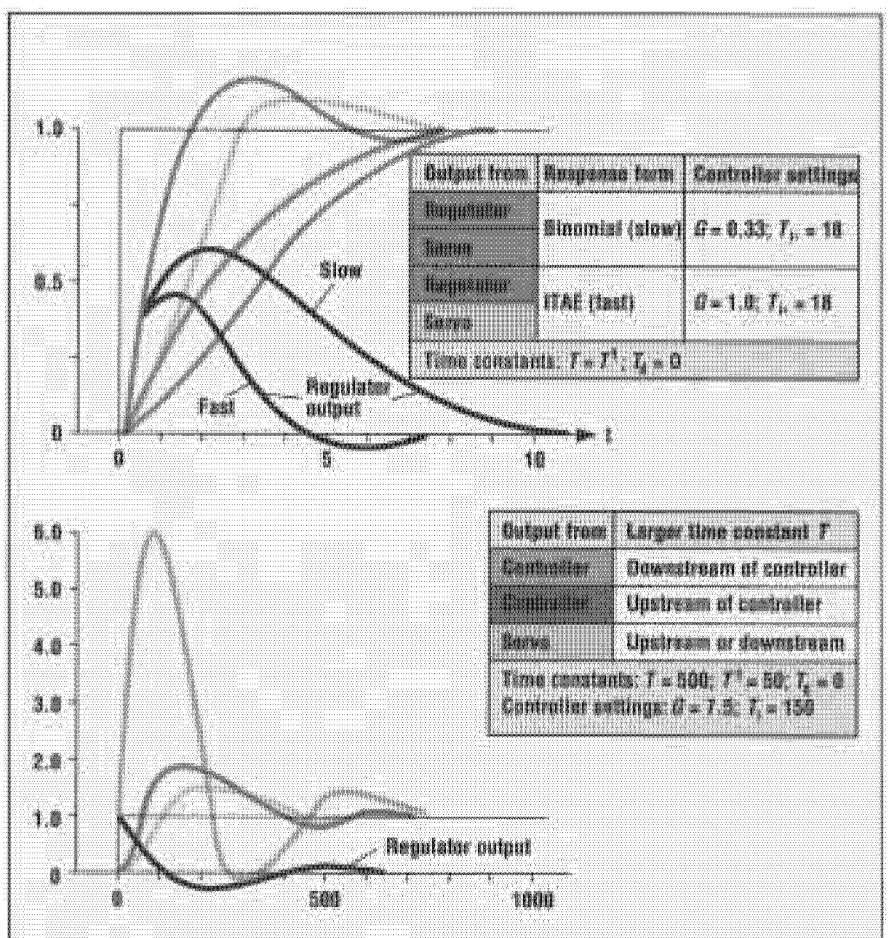
Pressure sensors: Although pressure measurement is basically simple—and pressure transmitters are rugged, accurate, fast—some of the largest system first-order time constants are due to pressure variations as in high pressure steam boiler load control.

Level sensors: Level sensors can be finicky—especially if they are of the differential pressure type which need to be compensated for product density. Their installtion requirements and specification are very important. It is too common to specify the wrong sensor for the application. This will result in erroneous readings which will translate to the rest of the loop.

Analyzer-type sensors: Analyzers (pH, O_2, redox, dissolved oxygen, and photometric types) have, in general, become much more rugged. Analyzing transmitters are routinely being included in control loops, but they still require substantial maintenance, and redundant systems are necessary if they are to be heavily relied upon. Installation requirements will unavoidably introduce some dead-time lag.

Electromechanical sensors: Sensors for rpm, displacement, position, and so forth are, in general, very reliable and do not need calibration—and this is a major advantage. Accuracy and speed of response are major issues for all mechanical measurements.

Electric power system transducers: These sensors convert ac variables such as voltage, current, active and reactive power, frequency, and power factor, into linearized dc mA signals, or electric impulses. Most types need periodic test bench calibration if accuracy and stability are of major concern which is often the case in frequency and active power measurements.



The top illustration shows the standard form of responses of Binomial (slow) and ITAE (fast) to a step input. The settings shown in the tables are those needed to obtain the standardized responses as shown by the curves. In the lower illustration, ITAE (fast) form of responses are shown. When the larger time constant is downstream of the controller, the servo and regulator responses are the same, but a large overshoot of controller output occurs. A filtering and smoothing effect takes place when the largest time constant is located upstream of the controller, beneficially limiting controller output and the associated control effort of such a control loop.

Instrument calibration

Transmitters not used for indication, but for process control purposes only, need to be sensitive rather than accurate. This can make calibration a convenient secondary priority if there are no indicating instruments connected to that same signal. Calibration-free instruments should be given a preference above all other types when available. This usually means a pulse or digital output. The resolution will be one pulse, and accuracy will be the inverse of maximum pulse output.

In most process control applications, a gain factor other than one is generated by the span of the transmitter and, to a lesser extent, by the sizing of the control valve. Rarely is a gain factor generated by the process itself. The controller then introduces only an adjustment factor to reach the dynamically optimum loop gain.

Controller settings and gains

The following relationships can be applied to all settings:

$$\text{Proportional band (\%)} = \frac{100}{\text{Gain}}$$

$$\text{Integral time (sec)} = \frac{60}{\text{Reset (repeats/min)}}$$

$$\text{Derivative action (sec)} = \frac{60}{\text{Rate (repeats/min)}}$$

The physical data of the working conditions are required for most processes, *e.g.*, pressures, maximum flow, level transmitter range, in order to be able to determine the gain of the transmitters that are a part of the control loop. The transmitter data have to belong to the particular instrument that is connected to the control loop. This is very important if the proper controller settings are to be obtained.

Since the gain of the controller will be affected by the gain of the transmitter and valve, it is necessary to determine these before attempting to adjust the controller gain. For example, in the case of a pressure transmitter, if it is calibrated for 1,500 psi to be the full output, and for the output to be zero at 600 psi, the transmitter span will be: 900 psi, (*i.e.,* 1,500 psi − 600 psi). If the system pressure setpoint is 1,000 psi, the transmitter gain will be as follows:

$$\text{Transmitter gain} = \frac{\text{Setpoint value}}{\text{Transmitter range}}$$

$$= \frac{1{,}000}{900} = 1.1$$

Transmitter gain is a major controller adjusting factor, and can easily range between 0.1 and 10. Control valve gain should not affect loop gain so dramatically. This must be verified and accounted for if other than one. Controller range, transmitter signal range, and valve signal range must all be the same; *e.g.,* 4-20 mA, 3-15 psi. Some controllers must have their output range limited as part of the adjustments; for example, three-element drum level control.

Some valve actuators may have smaller ranges. This is the case for control valve split-range designs. But what counts is the combined gain of both valves working together as one from the controller.

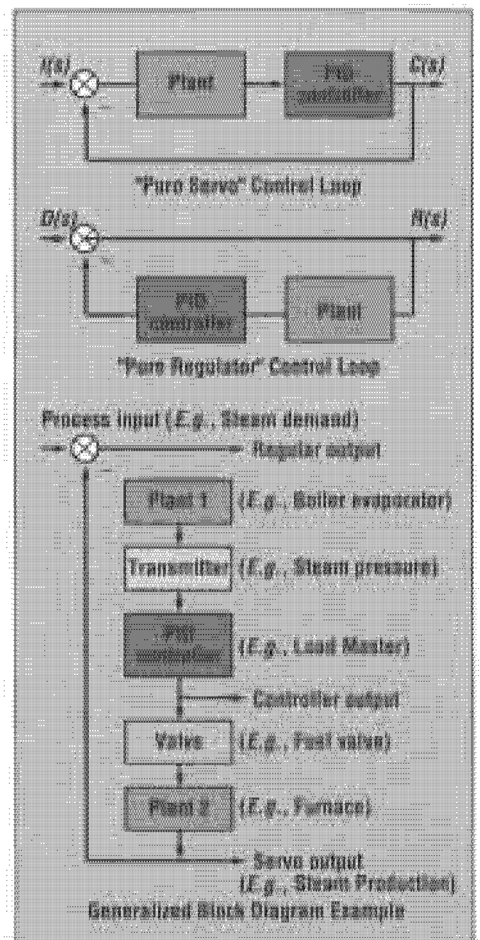
If a change in signal occurs within a loop as, for example, with the use of an electropneumatic transducer, the ranges must all correspond on a 0 to 100% scale basis. In other words, full scale and zero scale correspondence must be maintained; otherwise a new gain coefficient is introduced. The equations presented here do not make any provisions for changes in signal range.

With the exception of dead time, linearity is assumed overall. That is, transmitters and control valves have inherently linear responses. PID controllers are linear by definition.

Dynamic models

Characterization is not accounted for. All transmitters are assumed to be linearized for any flow, pressure, or temperature condition, as are valves.

Servo and Regulator Loops



In any control loop, two outputs can be defined for each input, and they are called a Servo output and a Regulator output. They are futher defined as follows:

$$\text{Servo:} \quad \lim_{t\to\infty} \frac{c(t)}{i(t)} = 1$$

$$\text{Regulator:} \quad \lim_{t\to\infty} \frac{r(t)}{d(t)} = 0$$

with i(t) = d(t), since the input d(t) to a Regulator is also called a"disturbance."

The Laplacian forms for the servo and regulator in the block diagrams are:

$$\text{Pure Servo:} \quad \frac{C(s)}{I(s)} = \frac{\text{(Plant)(PID)}}{1 + \text{(Plant)(PID)}}$$

$$\text{Pure Regulator} \quad \frac{R(s)}{I(s)} = \frac{1}{1 + \text{(Plant)(PID)}}$$

The controller settings derived and listed in the table on the next page are based on third-order dynamic models for PI controllers, and on fourth-order dynamic models for PID controllers. In this way, exact settings for gain, reset, and rate will result from the closed-loop model by equating the characteristic equation to a fast response (ITAE) or slow response (binomial).

Determining a time constant requires more ingenuity than work. quires more ingenuity than work. The following items are some general guidelines that may prove to be useful.

Empirical method—Measure an input step response as follows:

- Stabilize the process by either transferring the load fluctuations to another unit, or by arranging a period or interval at constant load with the production people;
- Set the loop to be measured in manual mode;
- Introduce a step change in the system. This step must be either a load change (*e.g.,* steam flow) or a control variable step (*e.g.,* fuel);
- Measure the response of the affected variable. Instruments for this purpose must be accurate and reliable. If the response is slow enough, take readings at fixed time intervals and plot them. If the response is too fast, a plotter must be used;
- The time constant is the time it takes for the initial slope to intersect the final value (see response curves, opposite). All readings must be converted to a percentage of the physical variable, by defining a "base" value;
- This method requires some trial and error until a satisfactory set of readings has been obtained, but has the advantages of: not disrupting the operation too much; use of minimum amount of instruments; and usually are of short duration (approximately ten minutes for each test, with one or two minutes of actual taking of readings);
- Determine the largest time constants of the system, and list them beginning with the largest one and down. Stop when the next is the controller, or if there are three time constants.

Analytical method—This method is recommended when all relevant parameters are known, which in real life is rarely the case.

There is one situation in which this is especially useful—level control. A typical question is: What is the main time constant of a tank or drum? The answer is, the maximum allowed level fluctuation times the tank surface, divided by the maximum flow.

This is a case similar to a chemical reactor temperature time constant, since two time constants can be determined: one for rising level and one for dropping level. But only one is in the control loop and that is the one that counts for the controller settings.

The maximum allowed level fluctuation is not necessarily the transmitter range, but usually a much smaller value since normal operating conditions are being considered. The level that the controller maintains before the level alarm goes off must also be considered. Therefore, a level transmitter usually introduces a gain factor much smaller than one. This transmitter gain must be compensated for by the controller by adjusting its own gain.

The largest time constants do not always come from the process or the plant. They also come from the control equipment itself; the control valve, for instance, or the pneumatic signal tubing, or the transmitter. This is very well the case when controlling flow with a flowmeter and control valve on the same pipeline. These cases require that the time lag of the controller be much lower than the time lag of the transmitter, and the error of the transmitter much lower than the error of the valve. Otherwise, the purpose of the control loop is self-defeating. Although widely used, this kind of "self-control" loop should be avoided because it slows down the response of the control system. Note that "much lower" or "much higher" means, in practical process control terms, a factor of ten or more.

On-line system identification

Virtually all process control loops can be separated into two parts: the plant and the controller. Since it is known what the controller can do, the major task is to model and identify the plant with the transmitters and control valves. With PID controllers, only three plant time constants can be handled. Therefore, a model is already specified as follows:

Plant = Gain and the three largest time constants.

This makes the identification process substantially easier. By measuring at regular intervals the signals of the controller to the plant, and reading back the plant's response, the actual values can be determined. Of course, the plant's response will be contaminated by noise coming from the process disturbance input, but this noise can be eliminated by a filtering algorithm which can easily be implemented in a digital controller.

Although some initial gain estimation by the old-fashioned paper and pencil method, and obtaining dead time by stop-watch are easy, these also can be determined by completely automatic means. A primary rule to remember is that process plant gain originates in the transmitters and the control valves. □

Controller setting examples using SFO algorithm

(These equations only for processes with a very large time constant upstream of the controller. A more general set is available from the author.)

PI controller with no dead time $T > 30T'$	$G = \frac{bT}{a^2 T'}$ $T_i = abT'$ $\omega = \frac{1}{aT'}$
PID controller with no dead time $T > 10(T' + T'')$	$G = \frac{cT(T' + T'')^3}{a^3(T'T'')^2}$ $T_v = \frac{[b(T' + T'')^2 - T'T''a^2]\,aT'T''}{c(T' + T'')^3}$ $T_i = \frac{acT'T''}{T' + T''}$ $\omega = \frac{T' + T''}{aT'T''}$
PID controller with dead time (T_d) $[T > 10(T' + T_d)]$	$G = \left[1 - a\left(\frac{T'}{bT_d}\right)^{\frac{1}{2}} + \frac{T'}{T''}\right]\frac{T}{T_d}$ $T_i = c(bT_dT')^{\frac{1}{2}} + T_d$ $\omega = \frac{1}{(bT_dT')^{\frac{1}{2}}}$ $T_v = T_d$
For all controllers	$K_c = \frac{G}{K_v K_t}$ $PID = K_c\left(1 + \frac{1}{sT_i} + sT_v\right)$

Symbols:
s = Laplace operator
G = Loop gain
T_i = Controller integral action time
T_d = Loop dead time
T_v = Controller derivative action time
ω = Loop natural frequency
T = Process main time constant
T' = Process second time constant
T'' = Process third time constant
K_t = Transmitter gain
K_v = Control valve gain
K_c = Controller gain

Coefficients for standard forms of transient responses

Binomial or slow PI (3rd order): $a = 3$; $b = 3$
Binomial PID (4th order): $a = 4$; $b = 6$; $c = 4$
ITAE or fast response PI (3rd order): $a = 1.75$; $b = 2.15$
ITAE PID (4th order): $a = 2.1$; $b = 3.4$; $c = 2.7$

REFERENCES

1. D. Graham and R. C. Lathrop, *The Synthesis of "Optimum" Transient Response: Criteria and Standard Forms,* Transactions of AIEE, Vol. 72, Part II, November 1953, Pg 281.
2. M. Polonyi, "Boiler Dynamic Load Control and Optimization," CONTROL ENGINEERING, Sept. 1985