

Modern Control Started with Ziegler-Nichols Tuning

GEORGE J. BLICKLEY, CONTROL ENGINEERING

When two engineers at Taylor Instrument Co. decided to document the work they had done in finding ways to tune process controllers, they changed the whole control industry.

Up until 1940, most tuning of process controllers was an art conducted by seat-of-the-pants methods on controllers that were a hodge-podge of techniques or add-on components that defied any rigid rules that could be universally applied.

One of the engineers at Taylor was John G. Ziegler, the practical one of the pair with a lot of experience in process applications, and who performed all the simulator tests that led to the methods they were seeking. The other was Nathaniel B. Nichols who was the mathematician and who reduced all of the math to a few simple relationships that could be understood by technicians and operators.

The result was the now famous "Ziegler-Nichols" method of tuning controllers—a method that survived the slings and arrows of its early de-

tractors, withstood the test of time, and works just as well as many of the later, more sophisticated optimizing forms on a great majority of process applications. Most of the work was done in 1940, a paper entitled "Optimum Settings for Automatic Controllers" was formulated and presented in December 1941 at the annual meeting of the American Society of Mechanical Engineers.

It must be remembered that all of this was done before the theory of servomechanisms was developed, and the paper presented some new terms to the industry, as well as the first—but eminently practical—rules for tuning.

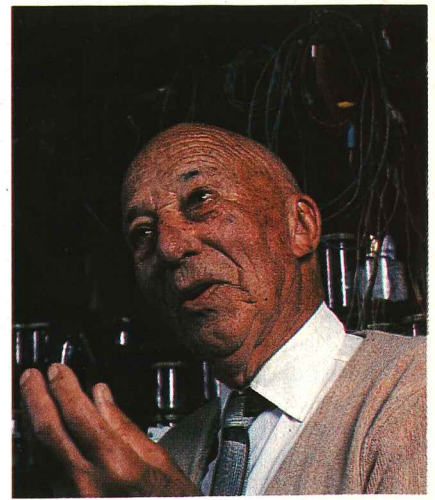
John Ziegler worked for 36 years for Taylor—now known as ABB Kent-Taylor, and still headquartered in Rochester, N.Y., and retired in 1972. We tracked him down to his little piece of

real estate on the side of a hill in Walnut Creek, Calif., with a great view of Mount Diablo. His shop and lab, as well as his considerable vegetable garden, makes one wonder what the word "retired" means. We asked Mr. Ziegler a few questions, the answers to which may give us some feel for how the roots of present-day automatic control were formed.

What were the predecessors to the integrated three-term PID controller? What were some of the principles used?

The earliest Taylor controller was the Model 10R which was a nonindicating, or blind, proportional controller. It used either a pressure capsule for pressure control or a vapor temperature bulb for temperature control. The set point was changed mechanically by a cam which operated a poppet valve air relay.

This evolved into a recorder, but some people insisted that you could not use the same thermal tube system



John G. Ziegler (above) and Nathaniel B. Nichols (not shown) were the founders of the Ziegler-Nichols method of tuning controllers.

for recording and controlling. The thinking was that, with two separate thermal systems, if the controller system broke, you could operate manually from the recorder, and vice versa. One of our competitors put both in one case, and we followed suit.

There were several developments using air valves, linkages, etc. that resulted in very high sensitivity proportional controllers that were essentially on-off controllers. The first truly ad-

"All of these things were developed in the 'golden years of control' from 1935 to 1940. It was an interesting time and I'm glad I was in on it."

justable proportional controller by Taylor was the Model 56R which was made in 1933. It was called the Fulscope and was the first proportional controller with any range to speak of. Sensitivity (or proportional band) was adjustable by a knob from 1,000 psi/in. (on-off) to about 2 psi/in.

About that time, Leeds & Northrup had the mechanically balanced potentiometer called the Micromax, and Taylor made a Fulscope unit to go inside to provide pneumatic proportional control. It was called the Fulscope Micromax. L&N subsequently came out with the Speedomax which was the first chopper-balanced potentiometer. They used a carbon microphone as the chopper. In about 1939, Brown Instrument (Honeywell) came out with the electronic chopper.

Foxboro came out with their Model 40 about 1934-1935. It was probably the first proportional plus reset recorder/controller. It was used principally in flow control in the petroleum industry. The trick here was to dampen the mercury manometer so that the sensitivity or gain could be set high. This was one of the things that had to be explained to get the Ziegler-Nichols approximation. The reset action was caused by spools of capillary which had to be changed for different reset rates. Some of the problems with this was slow response, overshoot on startup, and offset.

It was not too long after this, in 1935, that Taylor designed the "double response unit" which was a gadget that mounted on top of the diaphragm valve and provided automatic reset on top of the proportional control from the Fulscope. It had feedback from the valve stem and later became the first valve positioner in the indus-

try, in 1936. The double response unit was very complicated, had poor instructions, and was difficult to adjust causing it to be nicknamed the "dubious response unit." I have to say that I am the last living man who could walk up to a double response unit and align it perfectly the first time. When properly set, it would compensate for load changes, and it could be set to undercompensate or overcompensate. It provided stable control with no offset.

Taylor around this time was working the viscose rayon industry, trying to control the rayon shredder which was one of the god-awfullest 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 Fulscofes 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.

All of these things were developed in the "golden years of control" from 1935 to 1940. It was an interesting time and I'm glad I was in on it. The Pre-Act was the first derivative control and was incorporated into the Model 56R and worked great on juice units in the sugar industry, but not so great in other applications. We bravely set out to design the Model 100 Fulscope in 1939, in order to incorporate the automatic reset action provided by the double response unit. The Pre-Act was not too popular, but I persisted in

"I have to say that I am the last living man who could walk up to a double response unit and align it perfectly the first time."

getting a more stable version of it incorporated into the Fulscope 100.

Ken Tate was the head of the Taylor engineering department when the 100 series was being designed, and he was the genius responsible for the ingenious parallelogram linkage changing the mechanical feedback to make a continuous adjustment of proportional response sensitivity. And Bill Vogt, Tate's second in command, designed the reproducible needle valves for setting reset rate and pre-act time.

Only the
I/O PLEXER
lets you mix
both analog and
digital I/O
in one small
package.



You can achieve *any* mix of analog/digital input and output with the **I/O PLEXER**. The unit's compact case houses a complete **remote I/O controller** powerful enough to greatly reduce your computer's workload. Base your control system on one of many third party software packages, or save money by creating a simple control program in BASIC: most of the capabilities you need are built right into the **I/O PLEXER**.

Modem control signals which permit operation as an RTU are a standard feature.



Local Control Functions allow the **I/O PLEXER** to implement control decisions without host involvement.

Call toll-free for our brochure:

800/248-1632

I/O PLEXER
THE SMART SOLUTION

For information circle 8

dutec FOR INNOVATIVE CONTROL

This was then the very first proportional-plus-reset-plus derivative control integrated in one unit.

What led to the ASME paper on optimum settings, and what determined the parameters of the Ziegler-Nichols method?

We did not know how to set this new controller and I realized that we had to get some way of determining the controller settings rather than cut-and-try. I was out on a still in a chemical plant and it was almost a life's work getting the settings. I finally got it stable, but I wasn't sure I had the right settings. We had a unit in our factory demonstration room which consisted of a series of tanks and capillaries to simulate a multicapacity system for a fairly typical process to control pressure.

I was a very poor mathematician, so any sinusoidal oscillation was way beyond me mathematically. One of the few people at that time who were working on trying to understand control were Rufus Oldenberger of Woodruff Governor who used entirely different language from us, and as a result, we never got the two schemes together. The ASME had a "sewing circle," a bunch of old timers in the Instrument and Regulator Division who met twice a year to discuss automatic control. Their theory of automatic control was that it was equivalent to a sprung mass with a dashpots—that is, mass, spring, and damping—because if you

"I was a very poor mathematician, so any sinusoidal oscillation was way beyond me mathematically."

yanked the spring you would get an S-shaped curve. In control, they figured, if you made a change you would get an S-shaped curve and they likened all control to this. It was a fool's game, of course, because it is not representative of very many processes.

Rufus Oldenberger was not too far off from this because it was just what he faced in his turbine governors. But the ASME decided that this was typical of all processes. Now, such a process does not have an ultimate sensitivity. You could keep raising the gain of the controller and there is a setting that would give you a damped oscillation. If you raise the gain, it oscillates more. If you raise it still more, the period gets shorter and shorter and finally

you come to a zero period and an amplitude ratio of one. Any two capacity system will do that. That's why Nick, (Nathaniel B. Nichols, co-author of the paper) and I made such a to-do in our setting paper about using the ultimate sensitivity for determining a period of oscillation which gave you a setting of the reset rate and the pre-act time. We emphasized that you raised the proportional response until you got a sustained oscillation. That was called the ultimate sensitivity.

When we gave that paper, there was a great hue and cry. The preprints came out late, when the old timers got it and read it, they said it was heresy and we were damned to the deepest hell because we did not know what we were talking about. They almost rejected the paper and were not going

"We did not know how to set this new controller and I realized that we had to get some way of determining the controller settings rather than cut-and-try."

to let us go on the program with this sacrilege of ultimate sensitivity. They decided to let us give the paper anyhow, and Nick, being the sensitive, opted to let me deliver the paper.

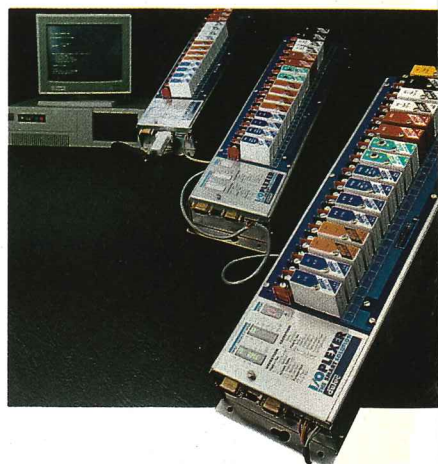
The questions at the end were pretty bitter because they could not stomach this ultimate sensitivity. The questions got worse and worse and I was answering them. Finally, a little guy in the back of the room got up. He was from Goodyear, since he was on the committee he had received an advance copy of the paper. He stuttered some, and he stammered out for all to hear, "We had one process in our plant, a very bad one, and so I tried this method and it just worked perfectly." That broke up the meeting.

Where did the concept of the reaction curve come from?

One of the old timers in the sewing circle (J. C. Peters) had written an ASME paper "Experiments in Process Control" on tests on a tank within a tank and the settings he listed came close to the best settings just by cut-and-try. He showed an uncontrolled curve which we called the reaction curve. We called the slowness at which the curve goes up the capacity, and the curve also showed what was a lag of some kind.

The Ziegler-Nichols approximation

It's easy to measure, monitor, or control almost anything with your computer, an **I/OPLEXER** and a little BASIC know-how!



Data acquisition and control can be both easy and affordable! Built-in firmware enables the **I/OPLEXER** to provide averages, out-of-range detection, maximums, minimums, thermocouple compensation/linearization, pulse counting, pulse durations, and edge detection. Analog and digital outputs include one-shot and delayed ON or OFF signals, pulse trains, and several waveforms.

I/OPLEXERS allow you to mix analog and digital I/O in one small package.



duTec products comply 100% with the industry standard serial communications protocol which is supported by over a dozen third-party software packages. Call toll-free for our free brochure:

800/248-1632

I/OPLEXER
THE SMART SOLUTION

For information circle 10

duTec FOR INNOVATIVE CONTROL

was simply to say that a process which follows this curve is very little different from one which has a dead period lag and a constant rate of change in response to an input. From these curves we used to Z-N approximation to determine settings and sent them to J. C. Peters, and they turned out to be better than the ones he had.

How did you and Mr. Nichols decide on 25 percent decay ratio?

Nick came to Taylor in the research department about the time the Model 100R was developed. I was playing on this analog simulator trying to figure what determined the sensitivity, the reset rate, and the pre-act time. Nick was put with me about the time I determined that the reset rate was dependent on the period of oscillation, as was the pre-act time. It turned out that when you set the proportional to about half of what caused the ultimate sensitivity, you would have about a 25% amplitude ratio. It's not perfect, but very close. So that's what we said—get an ultimate sensitivity and note the period. Any moron can do that. Then set the reset rate at one over the period and set the pre-act time to 1/6 to 1/8 of the period.

"When we gave that paper, there was a great hue and cry. The preprints came out late, when the old timers got it and read it, they said it was heresy and we were damned to the deepest hell because we did not know what we were talking about."

I had determined by a very crude method of what determined the sensitivity—I guess they call it gain now—of a process, and Nick leaped on this and from that we came to what determined the reset and pre-act. Nick wanted to use a 37% decay for some mathematical reason, but I insisted on the 25% because it was very easy for someone to see that the second wave is half as big as the first wave, and the third wave is half as big as the second wave. This gives about the least area under the curve.

Nick was cranking out these curves for me for a lot of different processes. We were checking areas under curves for different amplitude ratios and trying to find the best settings for differ-

ent process reaction curves. To speed it up, Nick rented out the differential analyzer at MIT and got into discussion with people at MIT on fire control servos. They were having trouble keeping the systems stable, and Nick believed that even though their math was correct, there was another little time constant they were missing in the loop somewhere. He guessed it was the compressibility of the hydraulic fluid, which they denied. He convinced them to use Taylor's pre-act, or derivative action, and when they put it in, the guns were stable.

As a result of all this, they asked him to come to the Radiation Lab at MIT to help win World War II. Taylor would not give him a leave of absence, so he left. I said that he would be back, and sure enough, after the war he returned to Taylor. He designed a nice electronic recording potentiometer which Taylor management rejected because they did not believe it was in their line. Because of this, Nick left Taylor again and went to Raytheon.

What was the market response when you introduced the Fulscope 100?

Enthusiastic as hell! We knocked our prime competitor right out of major chemical plants, such as Dow and Monsanto. They thought it was such a wonderful mechanism with responses labeled with calibrated dials.

It was a brand new instrument and was ahead of anything on the market. This is because it had the following features:

- It had the three responses of sensitivity, automatic reset, and pre-act (PID). No other controller had this;
- It had any combination of the three responses (P, PI, PID, or PD), including fixed high sensitivity for on-off control;
- All responses had adjustments with calibrated dials;
- All responses had continuous wide ranges.

Some process control manufacturers claim to have had PID controllers before the Fulscope 100. Is this true?

I think that is a crock of manure. They may have had other gadgets. Some of the instrument companies did make various new types of controllers for special applications. Some of them had PI controllers before the Fulscope 100, but it was the first three-response controller. □



If your control computer goes down, the **I/O PLEXER** knows how to think for itself.



Local Control Functions allow the **I/O PLEXER** to take over in the event of communications or host computer failure—your process continues to operate safely.

LCF also reduces the number of commands the host computer must transmit: many control tasks are handled by the **I/O PLEXER** itself.

Because analog and digital I/O are available in the same unit, **LCF** can perform closed-loop operations. Often, these functions eliminate the need for additional hardware expense.



Local Control: one of many reasons **I/O PLEXER** is the smart solution for remote I/O control.

Call toll-free for a free brochure:
800/248-1632

I/O PLEXER
THE SMART SOLUTION

For information circle 14

dutec FOR INNOVATIVE CONTROL