THREE TERM CONTROL

Elke Laubwald: Visiting Scientist, control systems principles.co.uk

ABSTRACT: This is one of a series of white papers on systems modelling, analysis and control, prepared by Control Systems Principles.co.uk to give insights into important principles and processes in control. In control systems there are a number of generic systems and methods which are encountered in all areas of industry and technology. These white papers aim to explain these important systems and methods in straightforward terms. The white papers describe what makes a particular type of system/method important, how it works and then demonstrates how to control it. The control demonstrations are performed using models of real systems designed by our founder - Peter Wellstead, and developed for manufacture by TQ Education and Training Ltd in their CE range of equipment. This white paper is about the most widely used type of controller – the three term or PID controller. The white paper explains the three term controller using the CE2000 Control Software tool and gives tips on the practical tricks that are used that make PID controllers work better.

1. Why is Three Term Control Popular?

Three term controller is an amazing control method. Although it was developed in the first quarter of the last century, it is still the most widely used form of control system in the world. Before electronic control systems there were mechanical, hydraulic and pneumatic versions of the three term controller. You can look at these wonderful mechanisms if you visit the *Deutsches Museum* in Munich. Of course in many other Museums there are also collections and many university departments have collections of these old controllers. It is worthwhile to inspect them and understand the beginning of practical control instruments. Even though digital computer control has replaced electronic controllers, it is the three term controller that is most widely used, even though a computer can do much more sophisticated control. Natural systems also use forms of three term control, and in medical physiology examples of integral and proportional control action can be found.

Why is an old, very simple and naturally occuring controller still used? Clearly, is it because it is simple and functions well. My boss at control systems principles always reminds me that a KISS solution to a technical solution is the best solution. KISS means 'Keep It Simple, Stupid' and the three term controller is a KISS solution.

The training and ability in factory technicians can be quite bad. Also they have little time, so that controllers may not be properly tuned. After my graduation I was employed in a milk product company. My first job was a control loop audit of all company plants. It was a dirty job in big dusty factories and I wanted very much to leave the city and return to my home village and my *Bergführer* boyfriend! It was a big test for a new lady engineer, but I stayed in the company and I checked <u>every</u> control loop in <u>every</u> factory of the company. In some control loops the sensors and actuators had problems, but it was surprising for me to find <u>all</u> controllers not working or set to the values that came from the instrument factory. The tuning of the control systems in the factories was a big job, but we did it and I learnt a lot about three term controllers. So pay special attention to the ideas in this white paper - Elke worked hard to learn these tips, and you must work to.

2. What is Three Term Control?

The basic idea of three term control is wonderfully simple and I explained it to the factory technician engineers as follows. It has to do with knowledge of the past, knowledge of the present, and predictions of the future. Please consider how you make decisions in your life and you will see that past experience and present knowledge and a 'guess' on the future are the components of information that we use in human decisions.

The three term controller is wonderful because it does exactly the same thing for automatic control in an elegant way. It gives three parts to the control signal:

- A part proportional to the current error (control action given by today's information).
- A part proportional to the integral of past error (control action given all past information).
- A part proportional to the rate of change or derivative of error (control action given the direction of future changes).

We can make a mathematical version of this description. A three term controller in basic form has the error signal $e(t)=r(t)-y(t)$ as its input and produces the control signal u(t) as its output. It is defined in the equation:

$$u(t)=k_p e(t)+k_i \int_0^t e(\tau)d\tau + k_d \frac{de(t)}{dt} \quad \textbf{(1)}$$

Where, k_p is known as the proportional gain, k_i is the integral gain and k_d is the derivative gain.

The first term in Equation (1) is the proportional block of the controller, the second term is the integral block and the third term is the derivative block. A block diagram of the basic implementation is shown in Figure 1.

Equation (1) is how electrical engineers write the controller equation. I am from the chemical direction of control engineering and we use a different version:

$$u(t)=K\left\{e(t)+\frac{1}{T_i}\int_0^t e(\tau)d\tau + T_d \frac{de(t)}{dt}\right\} \quad \textbf{(2)}$$

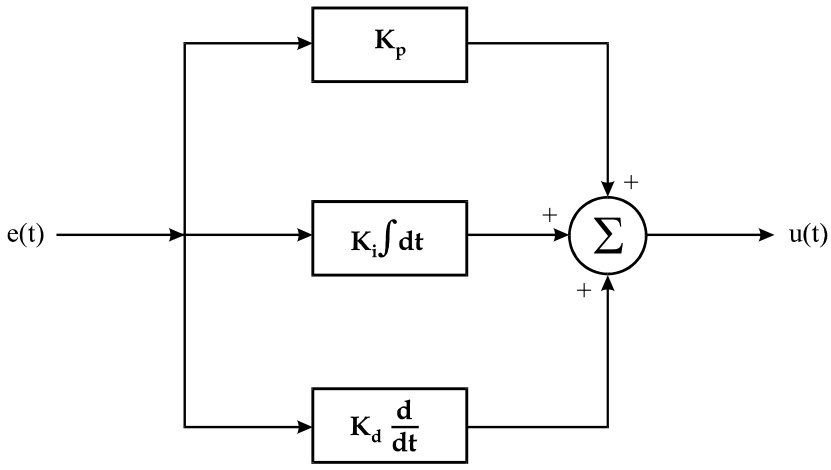

Figure 1 Block Diagram of the Three Term Controller

In this form the gain *K* operates on all terms. The parameter T_i is the **integral time** in seconds and the parameter T_d is the **derivative time** in seconds.

Sometimes the proportional gain *K* is included in the individual terms and the following definition used:

$$u(t)=Ke(t)+\frac{1}{T_i}\int_0^t e(\tau)d\tau + T_d \frac{de(t)}{dt} \quad \textbf{(3)}$$

Comparing this form with Equation (1), it is clear that the integral gain k_i is the inverse of the integral time T_i, and the derivative gain k_d is equal to the derivative time T_d. This form of the three term controller is shown in Figure 2.

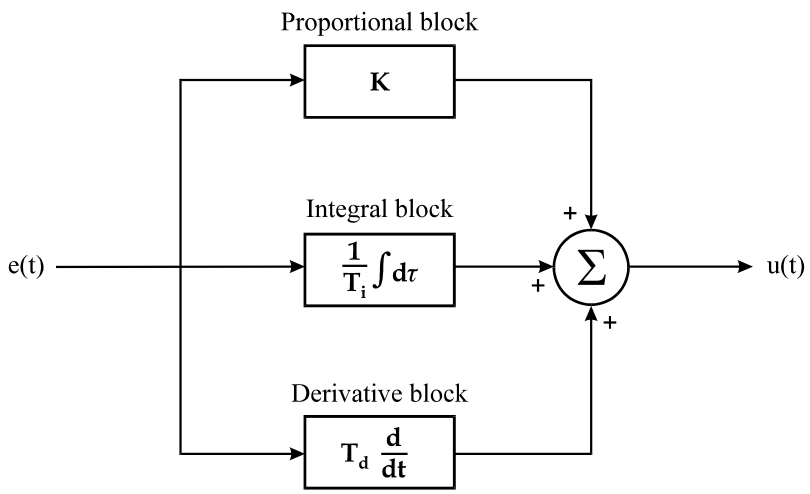

Figure 2 Block Diagram of the Three Term Controller

Some controllers companies do not use the name proportional gain, instead the name proportional band, PB is used on the front panel:

$$PB=\frac{100\%}{K_p} \qquad \textbf{(3)}$$

We will normally use the proportional gain form

You can see that there are some variants in terminology for three term control – so be careful and pick one form and ensure that everybody in your team has the same definition. I worked in a project for a control department at one time. It was chaotic - control systems were not documented and different engineers were using different forms of controller. I had to be very strict to them and insist on a standard form.

Derivative Washout

A problem with derivatives is that the output of a differentiator increases in proportion to the rate of change of the error signal. The faster the rate of change of error then the larger the output of the derivative term will be. This is actually one of the functions of the derivative term. However, it has the negative side effect of amplifying any unwanted high frequency noise contained in the error signal. This causes big difficulties because it puts rapidly changing signals on actuators and makes them vibrate and causes damage. This can be avoided if a **wash-out filter** is included in the derivative term. The output of the continuous time differentiator block, $u_d(t)$ is given by:

$$u_d(t)=T_d\frac{de(t)}{dt}$$

With the wash-out filter, the derivative term is:

$$\frac{T_d}{G}\frac{du_d(t)}{dt}+u_d(t)=T_d\frac{de(t)}{dt} \qquad \textbf{(4)}$$

The coefficient *G* is the wash-out term and is the upper limit on the gain that the differentiator output can have as the frequency of the input signal increases. For example, the basic differential block has the transfer function:

$$\frac{u_d}{e} = j\omega T_d$$

Thus as the frequency $\omega = 2\pi f$ of a sine wave input signal increases then the gain of the differential block increases in direct proportion to the frequency. This causes rapidly changing noise components of the differentiator input signal to be amplified. With the wash-out term the maximum gain that the derivative block can have is G. Specifically, if we consider the transfer function of the derivative block with the wash-out filter to be:

$$\frac{u_d}{e} = \frac{j\omega T_d}{1 + j\omega \frac{T_d}{G}}$$

we see that the wash-out term has introduced first order filtering action which limits the high frequency gain of the differential block. As the frequency $\omega = 2\pi f$ of the input becomes larger the transfer function becomes:

$$\frac{u_d}{e} = G$$

The wash-out gain *G* can be thought of as the maximum gain of the differential block for high frequency input signals. Almost all three term controller instruments have a washout term, but it is set in the instrument factory and you do not normally need to change it.

3. The Digital Form of Three Term Control

Sometimes you will want to make a digital three term controller and this is how to do it. The proportional block stays the same in digital form. The sampled error sequence *e(n)* at sampling step *n* is gives the proportional component of the actuation sequence $u_p(n)$:

$$u_p(n) = k_p e(n) \qquad \textbf{(5)}$$

The discrete integration block is derived from the continuous time version by approximating it with a discrete equivalent to the integration. The output of the continuous time integrator block, $u_i(t)$, is:

$$u_i(t) = \frac{1}{T_i}\int_0^t e(\tau)d\tau \qquad \textbf{(6)}$$

Where the parameter T_i is the integral time. Replace the integral by a summation gives the discrete equivalent with a sampling interval of *T* seconds gives:

$$u_i(n) = \frac{T}{T_i}\sum_{j=0}^{n} e(j) = u_i(n-1) + \frac{T}{T_i}e(n) \quad \textbf{(7)}$$

A constant error of 1 volt causes the control signal increase by T/T_i or K_iT volts in at every *T* seconds sample interval.

The discrete time differentiator with washout is discretised as follows:

$$u_d(n) = \frac{T_d}{T_d + GT}u_d(n-1) + \frac{T_d G}{T_d + GT}\left[e(n) - e(n-1)\right] \qquad \textbf{(8)}$$

Where : *e(n)* is the value of the error at sampling interval n; *e(n-1)* is the value of the error at sampling interval n-1; T_d is the derivative time; *T* is the sampling interval; *G* is the wash-out filter gain.

The advantage of washout is to reduce aggressive control signals. A basic derivative block turns a step change into a rapid and large impulse signal that could break a controlling actuator or damage the drive circuits. However, because of the wash-out filter term, the output signal will be a series of smaller decaying exponentials. The tail of the differential will become longer as the wash-out gain G is decreased. As the wash-out gain is increased the decaying exponentials will become more like impulses and hence more like the ideal derivative block. Look at figures 3 and 4. They show the damping down of the response to a square wave in a differentiator with different wash out gains.

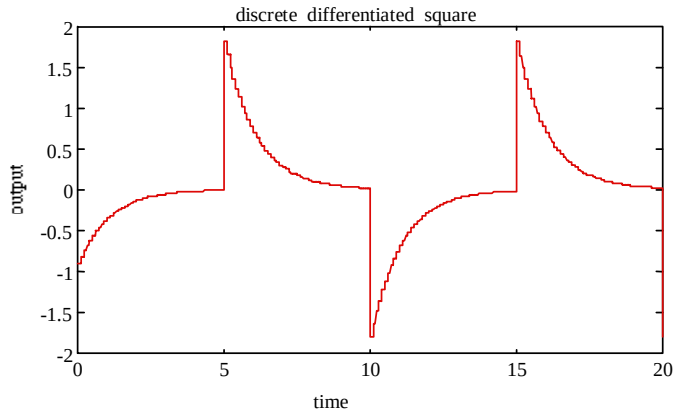

Figure 3. Differential Action on a Square Wave with Wash-out Gain G=1

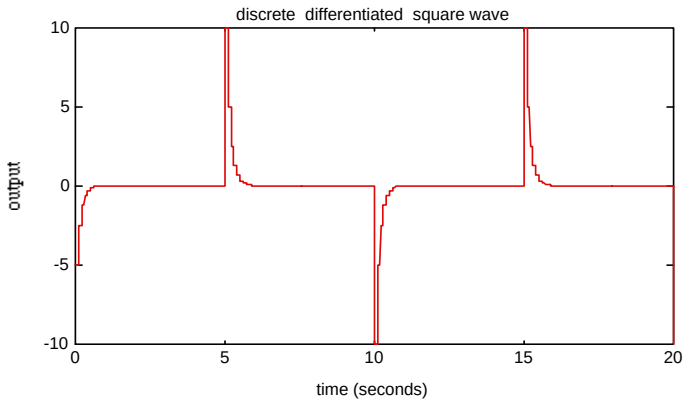

Figure 4. Differential Action on a Square Wave with Wash-out Gain G=10

4. Practical Features of Three Term Controllers

The theoretical form of the three term controller must be changed in a real application to make them work properly. I have described the 'wash-out' filter for derivative action. This is usually 'invisible' in commercial control instruments. There are these other important features also:

1. Integral controllers 'wind-up' when the system being controlled has limited actuation signals. There are ways to avoid this in a practical system.
2. Derivative controller 'kick' the actuators when fast changes happen – the wash out filter helps but a change in the controller structure can also help.
3. The response to the reference signal (also called the setpoint) can be given a special weighting to improve the output response.

Integral Windup

All systems have limits on the maximum and minimum control signals that the actuators can accept. This is normal, but most control system design methods ignore this. Happily, in three term control saturation can be compensated in a simple way. For example, when integral action is used in a system with saturation a phenomenon of 'wind-up' will happen. When the integrator output signal is greater than the saturation level, then the system output is also limited and the error signal that is driving the controller cannot reduce. The integrator 'thinks' it needs to do more and continues to increase, or wind-up, with no limit on how big the control signal can become. This can continue for a long time and cause the controller to behave badly. One way to stop this is to reset the output of the integrator to zero at periodic times. Most commercial controllers will have an 'integral reset' feature for this – the other way is to use anti-windup. Look at figure 5 to see what windup can do in a proportional plus integral controller (PI) applied to a system which has a saturating actuator. You can see the saturation block in the CE2000 simulation between the proportional block and the simulated second order system block. The saturation limits the maximum input to the system to between -10 volts and 10 volts. Between these limits the actuator supplies an output which is proportional to the input voltage. The maximum actuator output is reached for an input voltage of 10 volts and for inputs greater than 10 volts the actuator output remains at 10v. The minimum actuator output is –10v. Thus for actuator inputs less than -10 volts the actuator output remains at –10v. Limiting actuators or input saturation occurs in all control systems because there is always a maximum and minimum limit on control signals. The problem is a common problem in process systems such as fluid level and flow control where valves and flow are one-way, but also electro-mechanical Servo Systems have this problem.

What is the problem in figure 5? The limits on the output of a saturating actuator introduces a non-linearity into the control system so when the control error increases and the actuator output reaches the upper operating limit then changes in the error signal have no effect upon the control signal and the integrator output will continue to grow. Because of the saturation this growth is not fed through the system and hence the controller cannot supply a corrective action. The growth in the integral block output is called 'wind-up' and is a problem because the output of the integral block can become very large before the error sign changes. In such cases the error must stay negative for a long time to 'unwind' the integrator and the system settling time is large.

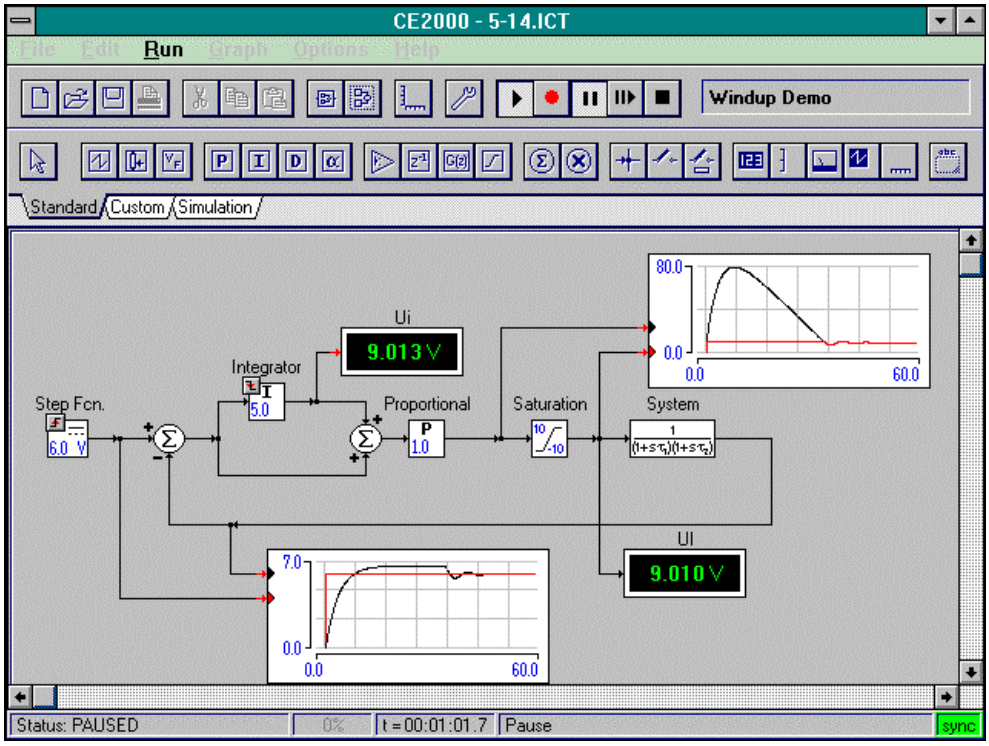

Figure 5. Wind-up in a PI control system

Integrator windup can be avoided with limits on the integrator output directly within the integrator block and some logical instructions in the integrator algorithm.

For a digital integrator this is done as follows:

$$IF \quad \{u_i(n-1)>u_{\max}\} O \{u_i(n-1)<u_{\min}\}$$
$$THEN \quad u_i(n)=u_i(n-1)$$
$$ELSE \quad u_i(n)=u_i(n-1)+\frac{T}{T_i}e(n)$$

This does not completely prevent wind-up, because it only influences the integral part of controller output. The alternative is the anti-wind-up system that I will describe next.

In the simulation block diagram (figure 6) I have added an anti-wind-up loop to a PI controller. The anti-wind-up loop uses a saturation block with the same limits as the saturation in the real system and generates a signal e_{sat}. The function of e_{sat} is to prevent the integrator increasing when the actuator is saturated. When the actuator is not saturated the signal e_{sat} is zero and does not effect the controller. When the actuator is saturated then e_{sat} is non-zero and is fed back in a way to reduce the input to the integrator until e_{sat} is zero, this reduces the controller output until it equals the saturation limit and stops wind-up. The response of the anti-wind-up loop is determined by the value of the gain in the block labelled **anti-w/u** in figure 6.

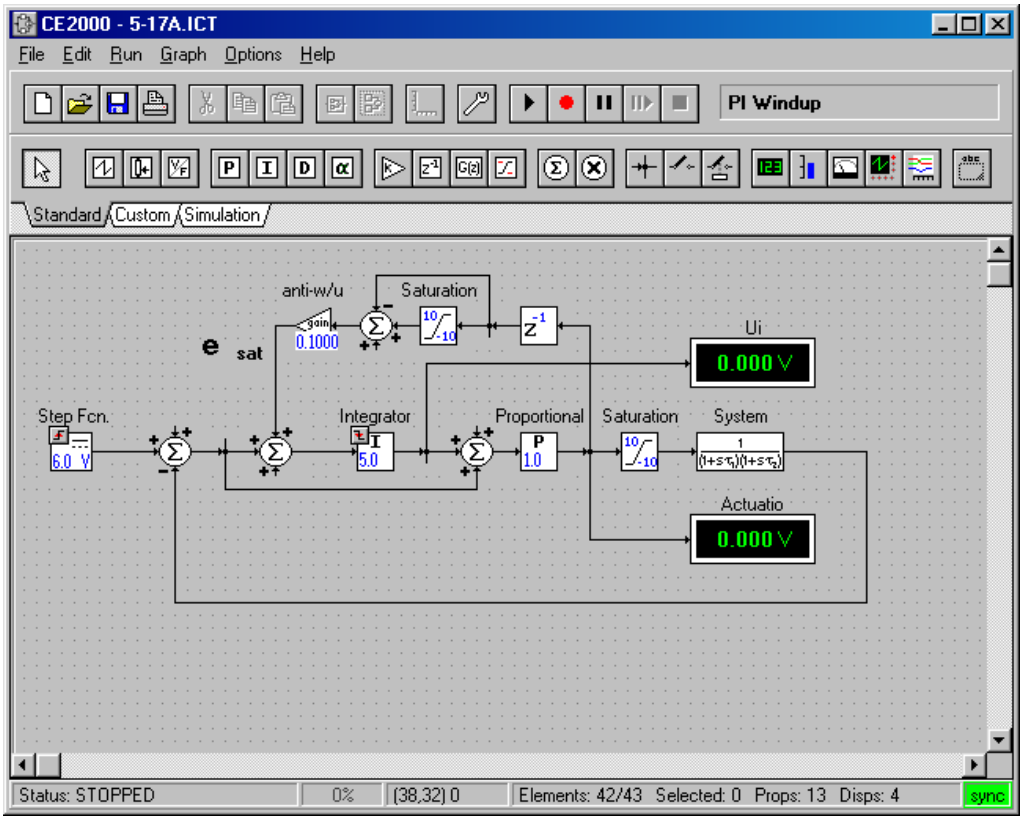

Figure 6. A PI control system with anti-wind-up loop

Figure 7 shows the response of the simulation in figure 6 with an anti-windup gain of 0.1. Without anti-wind-up the controller output 'winds up' to almost 80 (see the chart at the top right of figure 5) and the

output reaches the reference only after between 40 and 50 seconds. In figure 7 the maximum control signal is 30 and the output reaches the reference value between 10 and 15 seconds.

You can see that the problems of wind-up are reduced by this mechanism. The anti-wind-up loop improves the controller performance by reducing the amount of wind-up in the integrator. By increasing the value of the anti-wind-up gain it is possible to reduce the wind-up some more. Be careful of a big anti-wind-up gain because stability can be influenced – try to use the smallest value that gives a reasonable response.

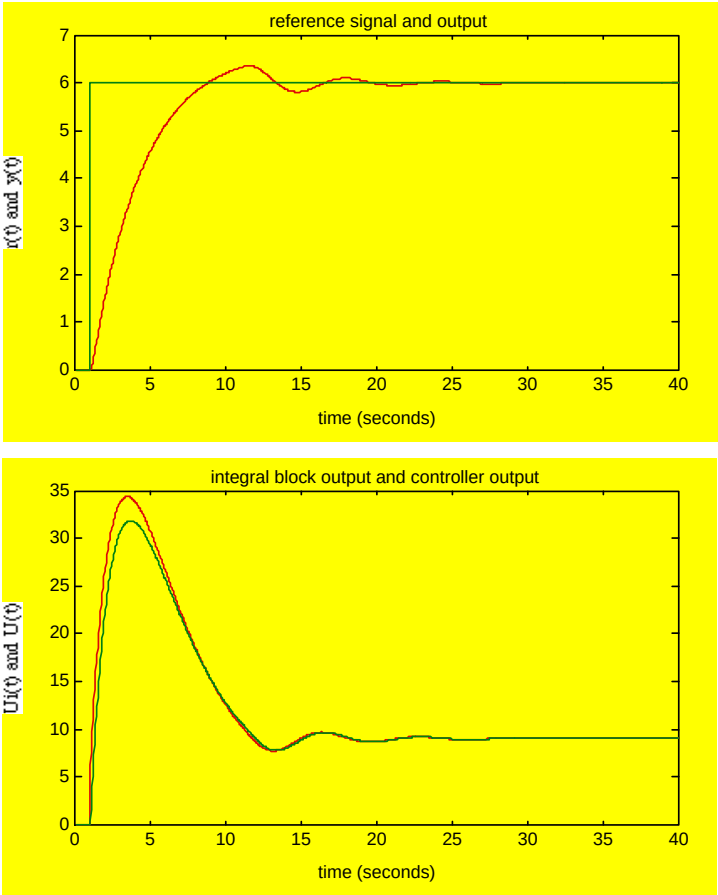

Figure 7 Anti-Wind-up for a PI Controller with anti-wind-up gain 0.1

Derivative Kick

I have said before that derivative action amplifies high frequencies. The wash out filter is a good idea to reduce this and is in most commercial control instruments. Another idea to reduce rapid control action from the reference input is to modify how the three term controller is connected up. This is done to avoid the phenomenon of derivative-kick caused by rapid changes in the reference signal $r(t)$ which are magnified by the derivative action and become a high amplitude transient component in the control signal. The simple way to stop derivative kick is to rearrange the controller so that the derivative block has the signal $y(t)$ at its input, <u>not</u> the error signal. This prevents the rapid changes in reference or setpoint signal coming to the derivative action, but does not change the closed loop stability of the control system.

The figure 8 shows a proportional plus derivative controller arranged in a normal way, where a square wave reference is applied to a PD control of a double integrator, (the famous Ball and Beam System can be modelled by a double integrator – so this is a nice example – check for a Ball and Beam download on our site). The control signal is to the bottom right of the simulation diagram and the reference is the blue

trace in the plot at the middle bottom. Notice that the controller output plot of U(t) shows big control signals when the reference square wave changes sign. These big control signals are the 'kicks' in the derivative actions. I have also heard them called derivative spikes in some control applications.

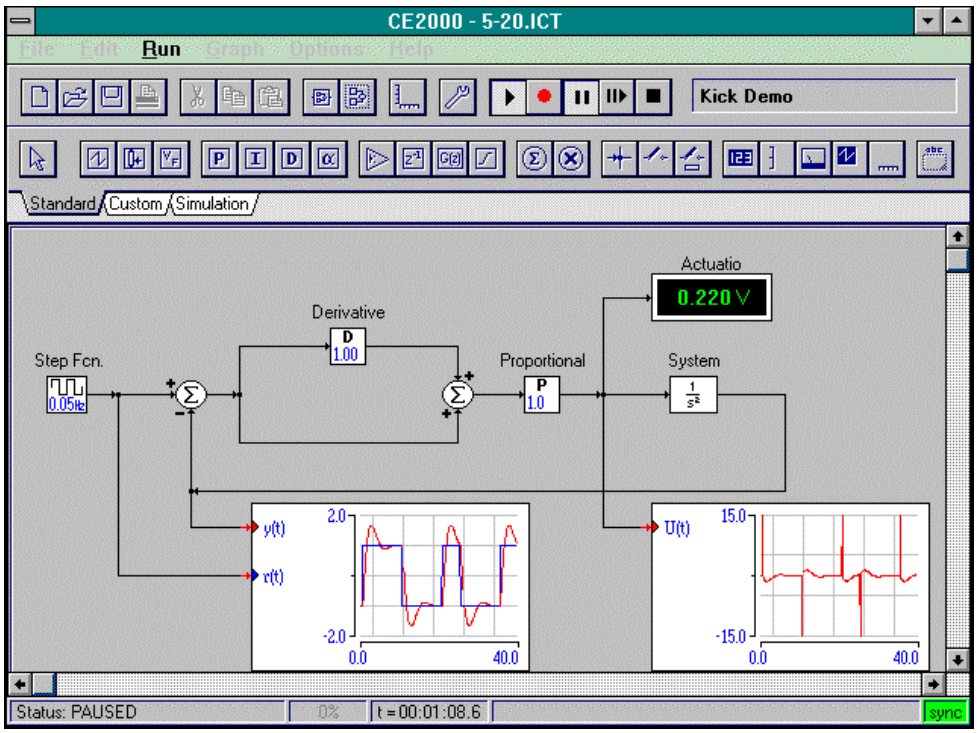

Figure 8. PD Control of a Double Integrator with Derivative Kick

In Figure 9 the PD system is altered to avoid derivative kick. In this configuration only the output signal is used in the derivative block and the rapid changes in the reference do not appear in the control signal, (see the plot of U(t) to the bottom right of the simulation plot). Altering the system in this way does not influence the system stability – to prove this it is necessary to write the closed loop transfer function and see that the closed loop characteristic equation is not altered.

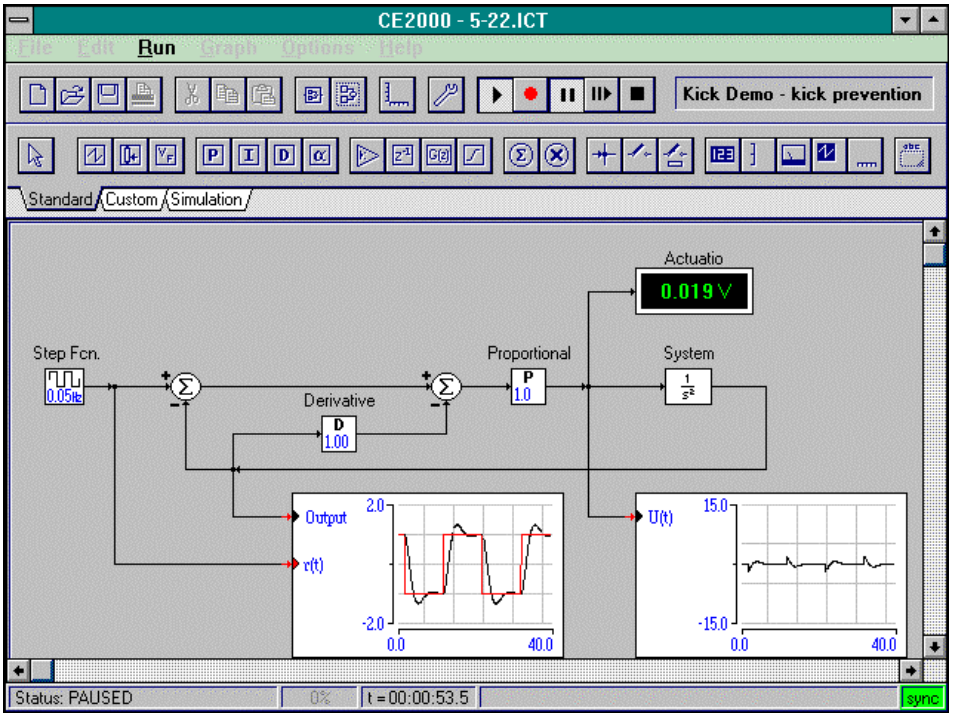

Figure 9. PD Control of a Double Integrator with Derivative Kick Prevention

All three term controller instruments that I use have derivative kick prevention built-in to the hardware, but do check the specifications. Also all process plants that I have worked on have the derivative action switched out.

So why use derivative action? Remember my first description of three term control? I said that derivative action is about putting knowledge of the future in a controller. This is called prediction and gives phase advance which can stabilise a control system that is not very stable or has time delays. Elke's practical tip on derivative action is this – only use derivative action when phase advance and predictive action is needed in the controller. This happens in systems that have oscillations in their open loop responses, or are unstable (like the Ball and Beam System) or systems with time delays.

Reference/Setpoint Weighting

A good tip for getting better reference response is to use feedforward of the reference signal. In three term control this is called reference/setpoint-weighting. In setpoint weighting the input to the proportional block is the modified error signal $e_w(t)$:

$$e_w(t)=K_r\, r(t)-y(t)$$

The gain K_r is a weighting on the setpoint which can improve the control systems response to reference signals but does not alter the stability of the system or the way in which disturbances in the feedback loop are treated. Technically this is called feedforward and it alters the transient response to the reference by changing the closed loop system zeros.

The figure 10 shows reference weighting in a system with a proportional plus integral (PI) controller.

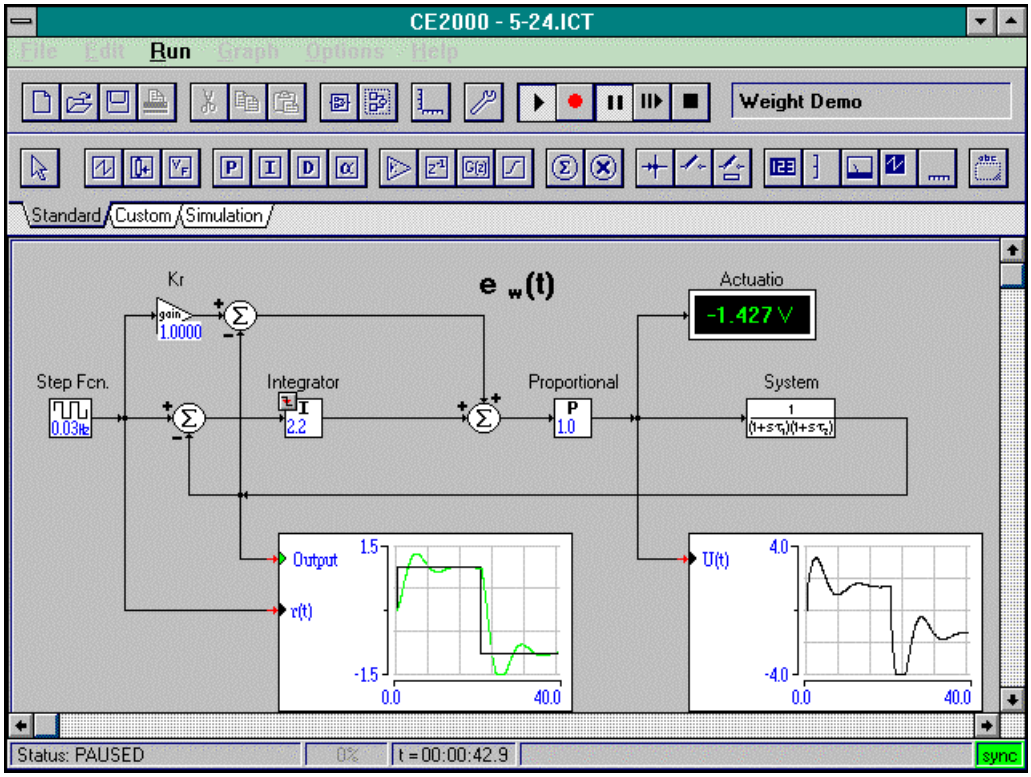

Figure 10. PI controller with reference-weighting

The behaviour of the control system shown in figure 10 is given for various setpoint-weighting gains (K_r) in Figure 11, 12. The main idea from these pictures is that by increasing K_r the output reaches the setpoint more quickly. but without changing the loop gain.

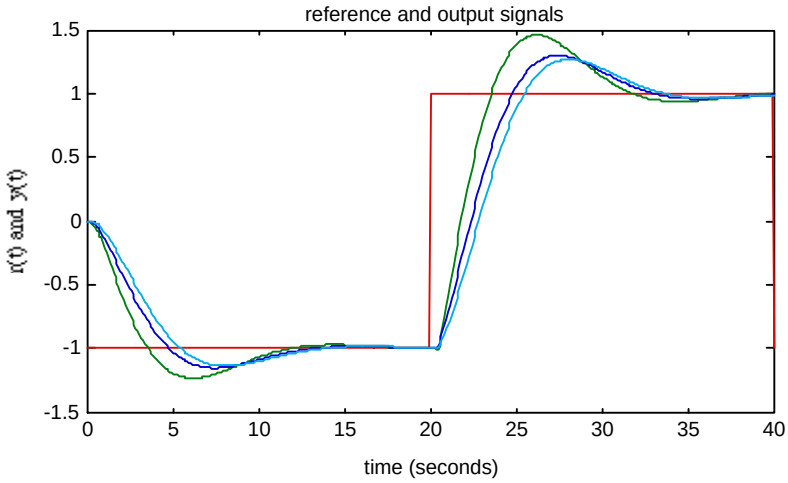

Figure 11. Reference and Output of a PI Controller with reference-weighting for values of Kr =2 (green trace), 1(dark blue trace), 0.5(light blue trace).

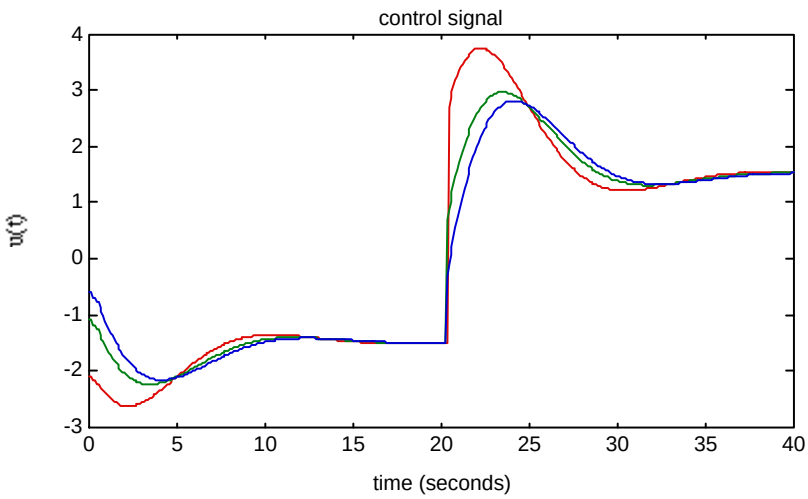

Figure 12. Control Signal of a PI Controller with reference-weighting for values of Kr =2(red trace), 1(green trace), 0.5(blue trace).

I have not seen many commercial three term controllers with reference weighting in them, and I do not use it often in control software that I build. The main idea of modern control instruments and software is to reduce the number of user controls. This is for the reason that I have said – control technician engineers are sometimes not well trained and are very busy in a big plant. The simpler the control unit the better. As my boss said – everybody wants a KISS controller!

5. Final Remarks

The three modifications to the basic three term controller can be used together in a general three term controller, but all of the terms will be necessary to solve a control problem. Many of my controllers operate with the derivative block switched out – I only use derivative action as a last chance when the

system is oscillatory or has time delays. Controllers which require derivative action to improve the system stability may have also the integral block switched out. If there are big time delays in a system then three term control may not be the best solution - some form of predictive control maybe required if the investment in time, complexity and money are acceptable.

I have not written about controller tuning because this is a *Teufelskreis* (a 'vicious circle' I think in English). Let me explain - there are many different techniques for three term controller tuning and people do not agree on the best because each application requires different tuning. I tested many tuning methods in my plant audits and found advantages and disadvantages with all tuning methods. Only if the control system is 'difficult' do I use special tuning rules. Normally, my habit is to use a laptop computer with a simulation of an approximate model of the target plant. I hand-tune a simulated plant and three term controller and check signal levels and saturation problems on the simulation. I then transfer this to the real plant for final testing. Finally, the technician engineers in the plant are trained to use the simulation model so that they can do the hand tuning also.

6. A Final Word from Elke

I hope that you have got some ideas about the importance of three term control from this white paper. I am sorry to say that it is not possible to answer general questions about the contents of our white papers, unless we have a contract with your organisation. For more information about three term control and the control simulation features of CE2000 Control Software go to the TQ Education and Training web site using the links on our web site www.control-systems-principles.co.uk or use the email info@tq.com.

I have implemented many three term controllers and I think that this white paper is a good overview over three term control. There are many detail research papers and text books on three term control – too much detail I sometimes think, but the book by Tore Hagglund and Karl Astrom is very good and not to long.

Aufwiedersehen!

Elke Laubwald