Assessment of Achievable Performance of Simple Feedback Loops

Åström, Karl Johan

1989

Document Version:
Publisher's PDF, also known as Version of record

Link to publication

Citation for published version (APA):
Åström, K. J. (1989). *Assessment of Achievable Performance of Simple Feedback Loops*. (Technical Reports TFRT-7411). Department of Automatic Control, Lund Institute of Technology (LTH).

Total number of authors:
1

CODEN: LUTFD2/(TFRT-7411)/1-18/(1989)

Assessment of Achievable Performance of Simple Feedback Loops

Karl Johan Åström

Department of Automatic Control
Lund Institute of Technology
January 1989

Department of Automatic Control
Lund Institute of Technology
P.O. Box 118
S-221 00 Lund Sweden

Document name
Report

Date of issue
January 1989

Document Number
CODEN: LUTFD2/(TFRT-7411)/1-18/(1989)

Author(s)
Karl Johan Åström

Supervisor

Sponsoring organisation

Title and subtitle
Assessment of Achievable Performance of Simple Feedback Loops.

Abstract

This paper develops techniques for qualitative and quantitative assessment of the performance that can be achieved with a simple feedback loop. Methods which give order of magnitude estimates for exploring the achievable performance are presented. The results are of interest for control design, auto-tuning and expert control.

Key words
Feedback, PID Control, Control Design, Auto-tuning, Expert Control, Intelligent Control.

Classification system and/or index terms (if any)

Supplementary bibliographical information

ISSN and key title

ISBN

Language
English

Number of pages
18

Recipient's notes

Security classification

Assessment of Achievable Performance of Simple Feedback Loops

K. J. Åström

Department of Automatic Control
Lund Institute of Technology
Lund, Sweden

Abstract. This paper develops techniques for qualitative and quantitative assessment of the performance that can be achieved with a simple feedback loop. Methods which give order of magnitude estimates for exploring the achievable performance are presented. The results are of interest for control design, auto-tuning and expert control.

Keywords. Feedback, PID Control, Control Design, Auto-tuning, Expert Control, Intelligent Control.

1. Introduction

In the vigorous development of control theory that has taken place during the past decades there has been a strong emphasis on developing exact methods for well posed problems. Comparatively little work has been devoted to develop more qualitative methods that will give some insight and orders of magnitude with a small effort. Such techniques are needed to develop knowledge-based control systems. See Åström et al. (1986) and Årzén (1987). The interest in such techniques have received considerable interest lately, because the hardware needed to implement such systems is now becoming available.

In this paper we will consider one of the simpler industrial control problems namely control at the single loop level from the point of view of expert control. There are several reasons for looking at such a problem. First, there is a need, single loop control is very common and will remain so in the future. Many single loop controllers are poorly selected and poorly tuned. Second, single loop control is the backbone of most sophisticated control systems. If control at the loop level does not work well, it will be reflected directly in the performance of the higher level coordinating loops.

Lately there has been interesting development of single loop controllers. Auto-tuning and adaptation is now becoming standard features in such systems. See, e.g., Kraus and Myron (1984), Åström and Hägglund (1984, 1988). This paper indicates that there are potentials for significant improvement of auto-tuners by incorporation of knowledge-based methods. This is based on the observation that good operators and instrument engineers have an important role to play. Their possibility to interact with the systems are, however, severely hampered by the interaction facilities available in current systems. Current systems also lack capabilities of extracting very much information about the process they are controlling. Neither can they communicate the knowledge they extract to the operator and the process engineer. This paper shows that there are possibilities to obtain systems with much stronger capabilities in this respect.

A key issue is the interplay between system complexity and control performance. To understand this it is essential to have tools for assessment of the achievable performance. For example, in design of auto-tuning it is desired to develop systems that can automatically select regulator structures and tune parameters of simple regulators. To do this it is essential to assess the control performance that can be achieved. It is also important to find out if it is worthwhile to include derivative action and to assess if performance can be improved significantly by a more complex control law.

The problem of assessing the dynamics also appears when working with large systems. To control complexity it is often necessary to have submodels of different complexity. The assessment methods are useful to select the appropriate model complexity for a given task.

This paper can be viewed as an attempt to develop simple methods to assess the dynamics of a system. We will try to answer questions such as: What bandwidth can be achieved with a simple regulator? What are the gains required to do so? What are the benefits of PID control as compared to PI control? What are the factors that limit the dynamical response?

The paper is organized as follows. Some preliminary background is given in Section 2. Assessment based on static process characteristics is discussed in Section 3. Section 4 deals with techniques for performance assessment. The particular case of systems whose dynamics do not pose any performance constraints is also discussed in that section. Crude assessment methods which require little information are covered in Section 6 and Section 7 deals with more accurate techniques that require more information about the process. Some examples are given in Section 9, and the paper ends with conclusions and suggestions for further work.

2. Preliminaries

We will consider problems where the process can be characterized by dynamics that can be described by linearized models with actuators that saturates. Key factors in assessment of such a control problem are:

- Process dynamics
- Actuator saturation limits
- Disturbances
- Regulator complexity
- Specifications

There is an interplay between several of these factors. Dynamics is in principle no limitation for linear systems that are strictly positive real (SPR) or with first and second order dynamics. For such systems the speed of response is limited by measurement noise and actuator saturations. Large pole excess and nonminimum phase dynamics, like time delays and inverse response, impose severe limitations on the achievable performance. It is thus essential to find methods to determine if the performance is limited by dynamics or other factors. It is also essential to try to characterize the complexity of the dynamics, e.g., the presence of oscillatory modes, the order of the dynamics etc. For systems with difficult dynamics it can be attempted to change the system so that the dynamics becomes simpler. Time delays can be reduced by repositioning sensors and actuators. Dynamics can be improved by replacing sensors and actuators with devices having faster response. It can be attempted to use local feedback to make the dynamics simpler and more reproducible.

The disturbances include set point changes, load disturbances and measurement noise. It is essential to find the ranges and the character of these disturbances. The range of set point changes and the required precision in the controlled variable and the maximum loop gain indicate if proportional control is sufficient or if integral action is needed. The magnitude of the error due to load disturbances depend on the amplitude and the frequency characteristics of the disturbance and of the loop gain.

There are several actions that could be contemplated with respect to the disturbances. Disturbances can be reduced at the source. Feedforward control can be considered if there is a measurable signal, which is correlated with the disturbance and appropriately located. Filtering methods can also be used to reduce disturbances and possibly also to reconstruct signals that can be modeled.

Measurement noise results in variations in the control signal. Together with actuator saturation this limits the achievable regulator gain and thus also the achievable bandwidth. If an actuator saturates because of measurement noise and high gain it can be attempted to reduce the gain, to reduce the disturbance level by filtering or to replace the actuator with a more powerful device.

Model uncertainties is another limiting factor. It can to some extent be dealt with by having a high loop gain at those frequencies where the uncertainty is large. To maintain a high loop gain it is, however, necessary to know the phase reasonably well at the cross-over frequency. Uncertainties in time delay, which gives very large phase uncertainties at high frequencies, will thus give a severe limitation on the achievable bandwidth.

Several of the issues discussed above have to do with selection and positioning of sensor and actuators, particular their sizing and their resolution. An important task of an expert control system is also to help to assess if good design choices are made. Capabilities to help in auditing control systems can therefore be very valuable. Useful knowledge for this purpose can be derived by observing the operation of a control system. Investigation of static process characteristics gives important information for this. It is also useful to have diagnosis systems that indicate if some component of the control loop is degrading.

3. Static Characteristics

A theme of this paper is that much knowledge required for control system assessment can be derived by analyzing signals available in the feedback loop.

The static input-output characteristics is an important system property. It can be described simply as a function. This function will tell the ranges of the input and output signals. It will also indicate the degree of nonlinearity. By observing the inputs and the outputs of a system during a stationary condition we can also derive useful information about the system.

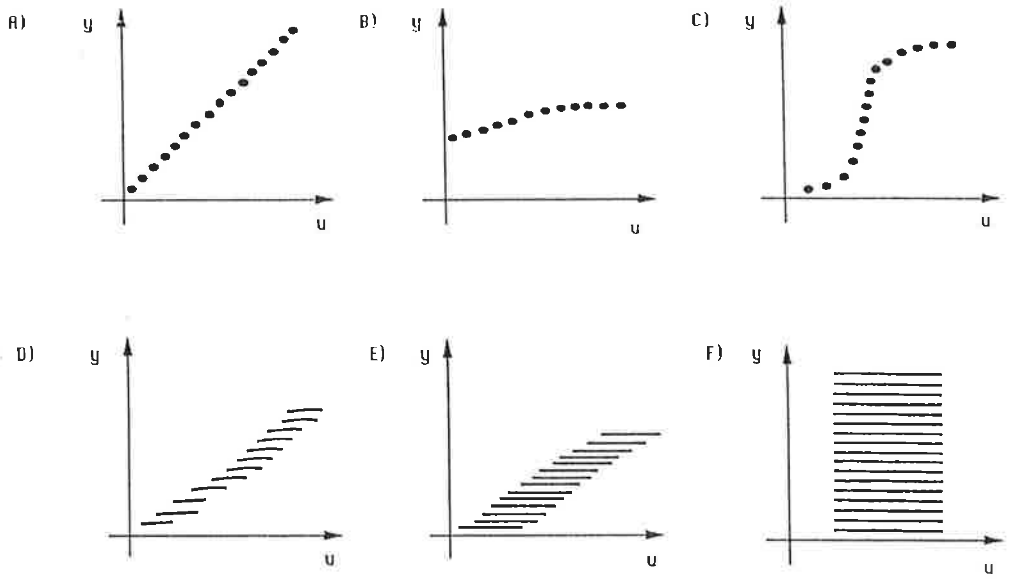

Figure 1. Examples of static input-output data logged during normal operation. The results shown in A, B, and C indicate a pure servo problem. The results in F indicate a pure regulation problem. Cases D and E are mixed cases. Case B indicate poor resolution of the sensor and case F indicates poor actuator sizing.

Preconditions

To determine stationary characteristics it is necessary to first have some criterion to decide that a system is in stationary operation. In typical process control problems this means that we would like to determine cases when there are set point changes and large process upsets. Since the set point is available, it is easy to find out when it changes. It is also useful to have information about the *time scale* of the process to know how long a set point upset lasts. Load disturbances are more difficult to determine, but criteria can be based on the magnitude and the frequency content of the signals. To obtain good data it is useful to low-pass filter the signals. To do this properly it is again necessary to know the time scales of the closed loop system.

Signal Ranges

Observation of the signal ranges and calculation of simple statistics, like mean value, variance, maximum and minimum deviations, will tell if the actuators are properly sized and if sensors and actuators have the proper resolution. If the variations are only a small part of the signal span it is an indication that a poor selection is made. It could, e.g., be indicated that a system with parallel actuators, one for large deviations and one for fine control should be used.

The Static Input-output Relation

If a detector for stationarity is available it is simple to keep a statistic for the fraction of time that the system is in stationary. A simple case is, e.g., to say that the conditions are stationary if the set point changes are sufficiently small. The static input-output relation can then be obtained simply by logging the process input and output. To obtain good data the signals should be filtered with respect to time scale of the closed loop. Curves like the ones shown in Figure 1 are then obtained. From these curves it can be determined if the major variations in the output are due to set point changes or load disturbances, i.e., if we are dealing with a servo problem or a regulation problem. We have a servo problem if the experimental data gives a well defined curve and a regulation problem if there is no definite relation between inputs and outputs. A simple statistic of the fraction of the total time when there are set-point changes or transients due to set point changes is also a useful indicator. There are of course also systems, which are mixtures of servo and regulation problems. It may be useful to let the operators participate in the assessment. For a regulation problem it may be useful to request the operator to look for candidates for feedforward signals by looking for signals that are related to the control signal.

For a servo problem the variations in the static gain of a system can also be determined. This gives a valuable indication if gain scheduling is required. The static gain curve can also be used for diagnostic purposes. Changes in the curve indicates changes in the process. By comparing the slope of the static gain curve with the incremental process gain measured during tuning or adaptation we can also get indications if there is some hysteresis in the loop.

To perform the operations it is useful to represent signals in such a way that statistical data over different time ranges are available. This can be done as follows:

Basic signal processing. Let us assume that each signal is associated with four numbers, the mean, the variance, the maximum and the minimum. These are called the signal characteristics. Each signal is also associated with a time scale T_s. This can, e.g., be the ultimate period of the control loop associated with the signal. The characteristics of each signal are first averaged over T_s. The average is then stored in a ring buffer. Each time the signal has circled the buffer the mean buffer value is transferred to another ring buffer etc. The buffers are chosen so that they correspond to intervals like minute, hour, day etc. The primary buffer can respond in the primary loop. The other may conveniently be located at higher levels in the system hierarchy.

4. Process Dynamics and Disturbances

In this section it is attempted to characterize process dynamics. We start with crude characteristics and proceed to descriptions that require more details.

Qualitative Features

The following are essential system features:

- Stable / Unstable
- Monotone / Oscillatory
- Essentially monotone, minimum phase

These features can be determined from simple experiments on the process. The assessment can be made by a properly trained operator or a neural network. Some of the features may also be known from design data. Experiments are necessary to make the assessment or to verify estimates obtained from design data. Two methods, step response and frequency response, are simple to apply and commonly used. See Åström and Hägglund (1988a).

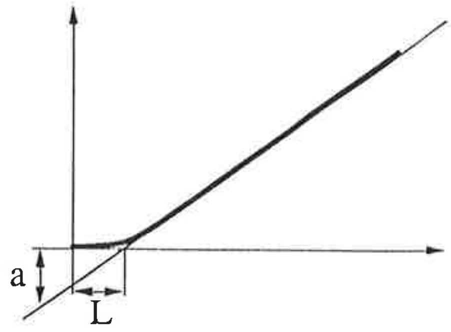

Figure 2. Determination of parameters a and L from the initial part of a unit step response.

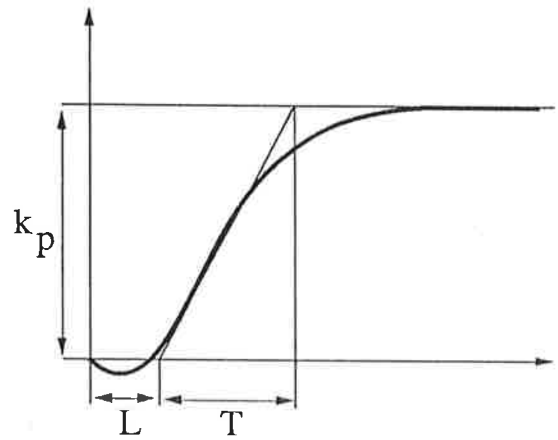

Figure 3. Determination of parameters k_p, L and T from the unit step response of a stable system.

Step Response

The step test is a simple experiment that yields useful information about a dynamical system. The test is performed by having the system in equilibrium with a constant input signal. The input signal is then suddenly changed to a new value and the response is recorded. A visual inspection of the step response gives the crude classification discussed above.

Parameters a and L can also be determined from the initial part of the step response as is shown in Figure 2. These parameters capture the high frequency behavior of the system. They can be used for simple performance assessment and simple regulator tuning as is discussed in Section 6.

For processes that are stable with monotone or essentially monotone step responses, see Figure 3, it is possible to determine three parameters: *process gain* k_p, *apparent dead-time* L and *apparent time constant* T. Knowing these parameters it is possible to asses the suitability of P, PD, PI, and PID control, and to tune the regulators. The parameters can be determined graphically by an operator. It is also possible to provide simple computational aids.

For processes with oscillatory step responses it is possible to determine period T_p and damping d of the oscillation. Use of step response data for regulator tuning is discussed in Bristol (1977) and in Kraus and Myron (1984).

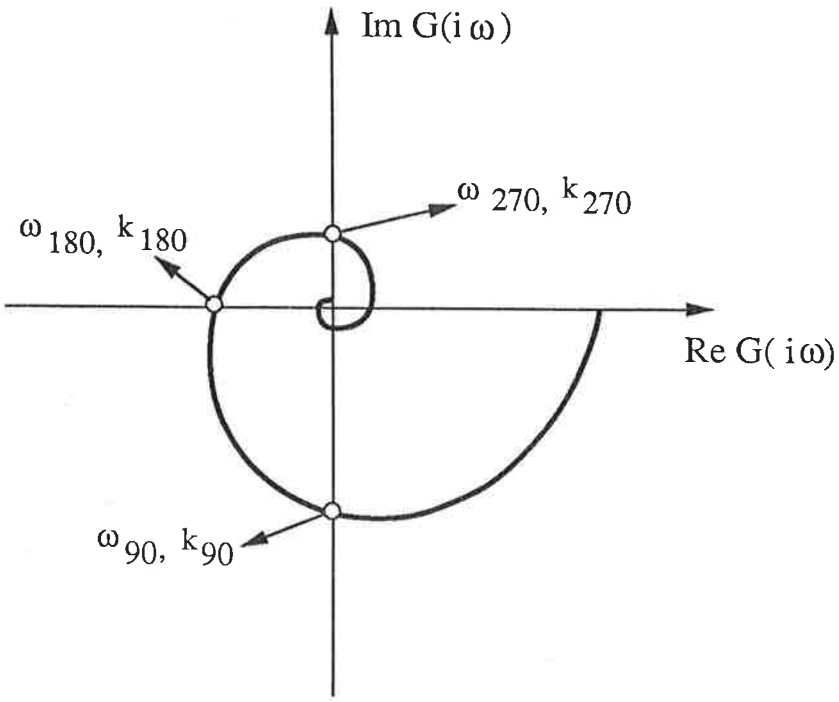

Figure 4. Determination of ω_{90}, ω_{180}, k_{90} and k_{180} from a frequency response curve.

Frequency Response

Frequency response is another simple way to characterize dynamics. A typical frequency response is shown in Figure 4. The frequency response can be determined in many different ways. It is of particular interest to note that the intersections of the Nyquist curve with the coordinate axes can be determined from simple experiments with relay feedback. See Åström and Hägglund (1984, 1988b), Hang and Åström (1988a).

With a Nyquist curve it is possible to make a crude classification of the dynamics into monotone or essentially monotone frequency responses.

Ultimate gain and ultimate period. The intersection of the frequency response with the negative real axis is of particular interest. It can be described with the parameters k_{180} and ω_{180}. The equivalent parameters $k_u = 1/k_{180}$ and $T_u = 2\pi/\omega_{180}$, called *ultimate gain* and *ultimate period* are sometimes used for historical reasons. The parameters can be determined approximately by applying relay feedback to the process. The period of the limit cycle obtained is the ultimate period (T_u) and the process gain is approximately given by

$$k_{180} = \frac{\pi a_m}{4d}$$

where a_m is the amplitude of the limit cycle and d is the relay amplitude.

Knowledge of T_u and k_{180} is sufficient for crude design of a PID regulator. If an additional parameter, e.g., k_p, is known it is also possible to improve the tuning and to assess the suitable regulator type.

The characterization of the Nyquist curve can be gradually refined by including more points like k_{90}, ω_{90}, k_{270}, and ω_{270}. The parameters k_{90} and ω_{90} can be determined by relay feedback, where the processes is cascaded with an integrator.

Mathematical Models

A complete mathematical model is a well known representation of dynamics. Simple cases that are common in process control are:

$$G(s) = k_p \frac{exp^{-sL}}{1 + sT} \tag{1}$$

and

$$G(s) = k_p \frac{exp^{-sL}}{(1 + sT_1)(1 + sT_2)} \tag{2}$$

More elaborate models are of course also possible. When specifying models it is also desirable to give a validity region. When such detailed specifications are given, there is a lot of control theory that can be used. Such models can be determined using system identification methods. Notice in particular that there are simple methods to determine the model (1) from a relay experiment, see Åström and Hägglund (1988).

Levels of Knowledge about the Process

When developing a knowledge based system it is useful to define different levels of process knowledge. The following classification is useful:

Level 0 Qualitative characterization.

Level 1 Level 0 and a and L or k_{180} and ω_{180}.

Level 2 Level 1 and k_p.

Level 3 Level 2 and more points on Nyquist curve, possibly with uncertainty regions.

Level 4 Complete mathematical model with uncertainty region.

Level 4A Process with known dynamics that is SPR or of first or second order with known model.

Disturbances

Disturbances are important aspects of a control problem. In some cases the disturbances are the key factors in control system design. There are unfortunately no simple rules similar to the Ziegler-Nichols rules to determine regulator parameters in this case.

It is important to know the origin of the disturbances, i.e., if they are due to measurement noise, load disturbances, set point changes or parameter variations.

Qualitative classification. Disturbances can be classified as transient, stationary or a combination. The transient disturbances are occasional upsets like steps, pulses, ramps and drift. The stationary disturbances can be periodic, narrow band or wide band.

Quantitative description. To describe disturbances quantitatively it is necessary to give both their amplitude and time characteristics. A simple description of the amplitude distribution can be given in terms of mean, variance, max and mean. A more elaborate description is to give the amplitude distribution.

The time variations can be described in many ways, e.g., as a spectral distribution or in terms of a filter. Crude properties of the filter, like time

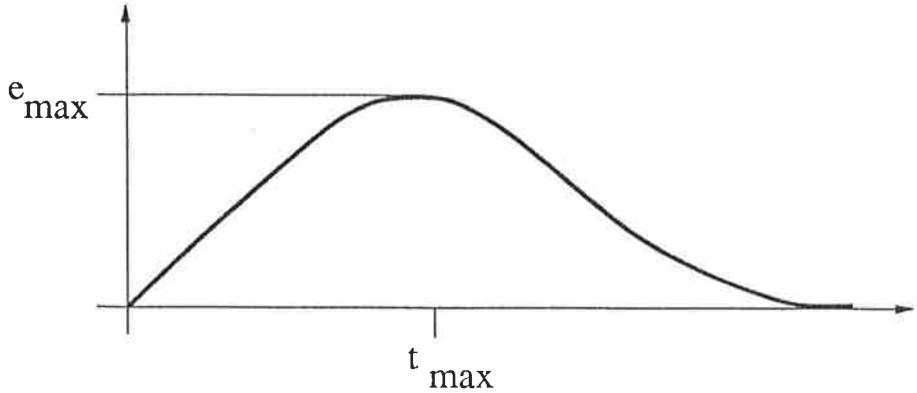

Figure 5. Determination of e_{max} and t_{max} from the response to a unit step at the process input.

constants or frequences, can also be used. To make a useful assessment it is necessary to know the disturbance levels below and above the bandwidth of the system. This means that it is necessary to know the *time scale* for a proper classification. For simplicity we label the high frequency disturbances measurement noise and the rest load disturbances.

Levels of Knowledge about Disturbances

Different levels of knowledge of disturbances can be summarized as follows:

Level 0 Qualitative knowledge.

Level 1 Level 0 and magnitudes of measurement noise and load disturbances.

Level 2 Level 1 and time constants associated with the disturbances.

Level 3 Mathematical models of disturbances.

5. Techniques for Performance Assessment

The technique for performance assessment it straightforward. A suitable class of systems is first established. In this case we will consider systems with monotone or essentially monotone step or frequency responses. A set of performance measures for the closed loop system is then established. Typical examples are

Bandwidth ω_b

Peak error e_{max}

Max time t_{max}

Integral gain k_i

The bandwidth indicates servo performance. The peak error and the max time refers to response to load disturbances as does the integral gain. See Figure 5. The integrated error due to a step load disturbance is approximately proportional to $1/k_i$. If e is the error due to a unit step disturbance at the process input we have

$$\int_0^\infty e(t)dt = \frac{1}{k_i} \tag{3}$$

The assessment is then based on semi-empirical relations between the open loop characteristics and the closed loop system properties. To establish these relations it is necessary to find suitable regulator parameters, which can be done in many ways.

To select a regulator it is essential to know if the key issue is to follow command signals or to reject load disturbances. The specifications typically include requirements on static error and some measure of dynamics. For example, it is essential to know if the fastest possible response is required, or if it is a slow loop that is only required to keep a good average. It is also essential to know if a response with no overshoot is desired.

Performance not Limited by Dynamics

We will first discuss the cases where the process dynamics does not impose any performance limitations. This is the case when the dynamics is positive real as is the case for systems with the transfer function

$$G_p(s) = \frac{b}{s+a} \tag{4}$$

Systems which are strictly positive real are easy to control, because they will be stable under feedback with infinitely large gain, even under relay feedback.

There are several ways to detect such processes. In step tests the normalized dead time $\theta = L/T$ is zero or practically much less than one. In experiments with relay feedback there will be an oscillation if the relay has hysteresis. The limit cycle obtained has maximum when the relay switches. The period of the relay oscillation will also be proportional to the hysteresis of the relay.

PI Control of a First Order System

Zero steady state and arbitrary pole positions can be obtained with PI control. Assume that the plant transfer function is given by (4).

Straightforward calculations show that a PI regulator, which gives the closed loop characteristic equation

$$s^2 + 2\zeta\omega s + \omega^2 = 0 \tag{5}$$

has the parameters

$$\begin{aligned} bk_c &= 2\zeta\omega - a \\ T_i &= \frac{2\zeta}{\omega} - \frac{a}{\omega^2} \end{aligned} \tag{6}$$

We thus get very simple tuning formula, where the regulator parameters are given in terms of ω and ζ.

The desired bandwidth ω can be selected arbitrarily high. It is thus not limited by the process dynamics. The measurement noise and the level of actuator saturation are instead the factors that limit the performace. Let e_{max} be some measure of the disturbance level, e.g., the maximum error or three times the standard deviation of the measurement noise. If u_{max} is the saturation limit we then find that the actuator saturates when

$$k_c = \frac{u_{max}}{e_{max}} \tag{7}$$

It then follows from equation (6) that the achievable bandwidth is limited by

$$\omega = \frac{1}{2\zeta}\left(b\,\frac{u_{max}}{e_{max}} + a\right) \approx b\,\frac{u_{max}}{e_{max}} \tag{8}$$

The bandwidth is thus proportional to the saturation and inversely proportional to the magnitude of the measurement noise.

Processes with Higher Order Dynamics

Processes whose transfer function is given by

$$G(s) = \frac{b}{s^2 + a_1 s + a_2} \tag{9}$$

can similarly be controlled arbitrarily well by a PID regulator. If the regulator complexity is increased even further, a process that is controllable and observable can be controlled with arbitrary dynamics using a regulator based on state feedback and an observer. Drastic increases in response speed is, however, often associated with complex regulators and accurate process knowledge. The performance achievable with a PI or PID regulator is often a good indicator of the performance achievable with a reasonable process knowledge.

To assess achievable performance it is thus essential to determine if the process dynamics is such that it does not limit performance. This means that we have to find methods of separating processes with transfer functions given by (4) or (9).

6. Crude Assessment of Control Performance

For processes, where performance is not limited by dynamics, crude controller settings can be obtained using only Level 1 information about process dynamics, i.e., the parameters a and L or ω_{180} and k_{180}. To asses controller performance and to select regulators it is, however, necessary to have at least Level 2 information. This means that the static process gain must be known in addition. These parameters can be obtained from a step test or from an experiment with relay feedback provided that the process has (essentially) monotone step responses or (essentially) monotone frequency responses.

Assessment Based on ω_{90} and ω_{180}

A simple way to assess achievable control performance is simply to determine ω_{90}, ω_{180}, i.e., the frequencies where the plant has 90 and 180 phase lag. A simple rule of thumb is that it is possible to achieve a bandwidth of ω_{90} with PI control and ω_{180} with PID control. The achievable loop gains can also be estimated from k_p, k_{90} and k_{180}. For processes, whose dynamics is a pure dead time, we have $\omega_{180} = 2\omega_{90}$. The gain in bandwidth by using derivative action is thus moderate in this case. The improvement achievable by derivative action is considerable if ω_{180} is much larger than ω_{90}.

Assessment Based on Maximum Loop Gain

The dimension-free parameters $\theta = L/T$ or $\kappa = k_p/k_{180}$ are useful for assessment of achievable performance for processes with monotone step responses or monotone frequency responses. Heuristics for regulator tuning of such processes is developed in Åström et al. (1988). In that paper it is shown that $\kappa\theta \approx 1.3$. The parameter θ can therefore be used instead of κ.

It follows from the Ziegler-Nichols closed loop tuning rules that the parameter $\kappa = k_p/k_{180}$ can be interpreted as the *maximum loop gain* with proportional control. With Ziegler-Nichols tuning the loop gain is approximately $\kappa/2$. Knowing the demands on steady state error it is thus possible to determine if proportional control is sufficient to cope with steady state errors or if integral action is necessary.

Similarly it follows from the Ziegler-Nichols open loop tuning rules that the number $k_p/a = T/L = 1/\theta$ can be used in a similar way. The loop gain under proportional control is actually k_p/a.

In Åström et al. (1988) it is shown that with Ziegler-Nichols tuning the maximum error due to a unit step load disturbance applied to the process input is approximately given by

$$e_{max} \approx \frac{0.4}{k_c}, \quad t_{max} \approx \frac{T_i}{2} \tag{10}$$

A similar analysis for PI control gives

$$e_{max} \approx \frac{0.6}{k_c}, \quad t_{max} \approx T_i \tag{11}$$

It is also shown that the closed loop rise time obtained with Ziegler-Nichols tuning is approximately equal to the apparent dead-time L.

The following heuristic assessment rules are also developed in Åström et al. (1988):

Case 1, $\kappa > 20$: A loop gain of about 10 can be used for proportional or PD control. Proportional control may be used if the steady state error is acceptable. Significant improvements may be possible with derivative action or with more complex control laws. Ziegler-Nichols tuning may not give the best results in this case.

Case 2, $2 < \kappa < 20$: This is the prime application area for PID control with Ziegler-Nichols tuning. It works well in this case. Derivative action is often very useful.

Case 3, $1.5 < \kappa < 2$: PID control is possible if the specifications are not too demanding. The Ziegler-Nichols tuning rules must be modified to get good responses. Other regulator structures like Smith predictors, pole placement, or feedforward should be considered.

Case 4, $\kappa < 1.5$: PI control can be used if the specifications are not too demanding. Derivative action is of little use. The Ziegler-Nichols tuning rules do not give good responses. Other regulator structures are recommended.

Rules for tuning the regulators are given in Ziegler and Nichols (1942), Cohen and Coon (1953), Deshpande and Ash (1981), Hang and Åström (1988a,b), and Åström et al. (1988).

7. Accurate Assessment of Control Performance

A more accurate assessment of control performance can be made if the transfer function of the process is known. A convenient way to do this is to determine appropriate P, PD, PI and PID regulators using the dominant pole design method, see Åström (1988).

Dominant Pole Design

The dominant pole design method can be described as follows. A regulator with a given structure and adjustable parameters is chosen. A number of desired closed loop poles $p_1, p_2, \ldots, p_k$ are chosen. The regulator parameters are then chosen in such a way that the closed loop system has the desired poles. The range of regulator parameters that give stable systems, such that the selected poles are dominating, are then determined. The design method will thus give not only regulator parameters but a set of regulators with associated performance ranges. The method is described in detail in Åström (1988).

To illustrate the ideas we will show how it is applied to design of PI regulators. Let the process to be controlled have the transfer function $G(s)$. The characteristic equation of the closed loop system under PI control is

$$F(s) = 1 + \left(k + \frac{k_i}{s}\right) \quad G(s) = 0 \tag{12}$$

Two dominant poles can be specified when there are two adjustable parameters. It is natural to choose these poles as

$$p_{1,2} = -\sigma \pm i\omega_1 = -\zeta\omega \pm i\omega\sqrt{1-\zeta^2} = \omega e^{i(\pi \pm \alpha)} \tag{13}$$

Requiring that the function F given by (12) has zeros at $s = p_1$ and $s = p_2$ we get

$$1 + \left(k - k_i e^{\pm i\alpha}\right) G\left(-\omega e^{\pm i\alpha}\right) = 0$$

or

$$1 + \left(k - k_i e^{\pm i\alpha}\right) r(\omega) e^{\pm i\phi(\omega)} = 0$$

where

$$\begin{aligned} r(\omega) &= \operatorname{abs}\big(G(-\omega e^{-i\alpha})\big) \\ \phi(\omega) &= -\arg\big(G(-\omega e^{-i\alpha})\big) \end{aligned} \tag{14}$$

Equation (13) is linear and has the solution

$$\begin{aligned} k(\omega) &= \frac{\sin\big(\phi(\omega) - \alpha\big)}{r(\omega)\sin\alpha} \\ k_i(\omega) &= \frac{\omega \sin\big(\phi(\omega)\big)}{r(\omega)\sin\alpha} \end{aligned} \tag{15}$$

The integration time is given by

$$T_i = \frac{k}{k_i} = \frac{\sin\big(\phi(\omega) - \alpha\big)}{\omega \sin\big(\phi(\omega)\big)} \tag{16}$$

Achievable performance. The conditions for pole domination can normally only be satisfied if the closed loop bandwidth is chosen in a certain range, otherwise the gains k and k_i will not be positive. Equation (15) always exist if $r(\omega)\sin\alpha \neq 0$. Under this condition regulator parameters can always be found, such that the closed loop characteristic equation has the zeros p_1 and p_2 given by (15). This does, however, not imply that the closed loop system is stable or that the chosen poles are dominating. These issues have to be investigated by other methods. If the process gain is positive a necessary condition for stability is that the integrator gain k_i is also positive. This gives an upper bound of the achievable bandwidth. For typical systems that appear in process control it has been found that a good estimate of the upper bound is the bandwidth ω_0, where k_i has its largest value. This bandwidth is called ω_{PI}. This choice corresponds to the integral given by Equation (4) being minimal.

The dominant pole design can also be carried out for P, PD, PI and PID regulators. In this way we will obtain ω_P, ω_{PD}, ω_{PI}, and ω_{PID}. These parameters give reliable estimates of achievable performances with the different regulator structures.

8. Examples

The techniques for performance assessment will now be illustrated with a few examples.

EXAMPLE 1
Consider a system with the transfer function

$$G(s) = \frac{1}{1+sT} e^{-sL} \tag{17}$$

Consider first the case $T = L = 1$. We get $k_p = 1$, $\omega_{90} = 0.86, k_{90} = 0.76$ and $\omega_{180} = 2.0, k_{180} = 0.44$. Since $\kappa = k_p/k_{180} = 2.3$, we see immediately that integral action is necessary to get reasonable steady state errors. A crude assessment indicates that a bandwidth of about 0.7 may be achieved with a PI regulator and that it may be doubled with PID control.

Straightforward but tedious calculations give

$$\begin{aligned}
k &= \Big((2\sigma T - 1)\cos\omega_1 L + \frac{\sigma + \omega_1^2 T - \sigma^2 T}{\omega_1}\sin\omega_1 L\Big)e^{-\sigma L} \\
k_i &= \frac{(\sigma^2+\omega_1^2)\big(\omega_1 T\cos\omega_1 L + (1-\sigma T)\sin\omega_1 L\big)e^{-\sigma L}}{\omega_1} \\
T_i &= \frac{\omega_1(2\sigma T - 1)\cos\omega_1\big(CL + (\sigma + \omega_1^2 T - \sigma^2 T)\sin\omega_1 L\big)}{(\sigma^2+\omega_1^2)\big(\omega_1 T\cos\omega_1 L + (1-\sigma T)\sin\omega_1 L\big)}
\end{aligned}$$

With a relative damping of $\zeta = 0.707$ we get $\omega_I = 0.55(k_i = 0.25)$, $\omega_{PI} = 1.1(k = 0.50, k_i = 0.51, T_i = 0.99)$, and $\omega_{PID} = 1.7(k = 0.92, k_i = 0.76, T_i = 1.2, T_d = 0.22)$. We thus find that pure integral control gives the bandwidth $\omega_I = 0.55$. PI control gives a bandwidth in the range $0.55 < \omega < 1.1$ and with PID control the range of bandwidths is $0.9 < \omega < 1.7$. The agreement with the crude assessment is reasonable. Since $\kappa = 2.3$, the empirical rules given in Section 6 indicates that PID control is a reasonable choice. Using Equation (4) we find that PI control reduces the error integral due to a load disturbance by 50% and that derivative action gives an additional reduction by 33%.

Figure 6 shows the responses to command inputs and load disturbances for some of the regulator designs. Notice that the conclusion drawn from the analysis is well supported by the time behavior.

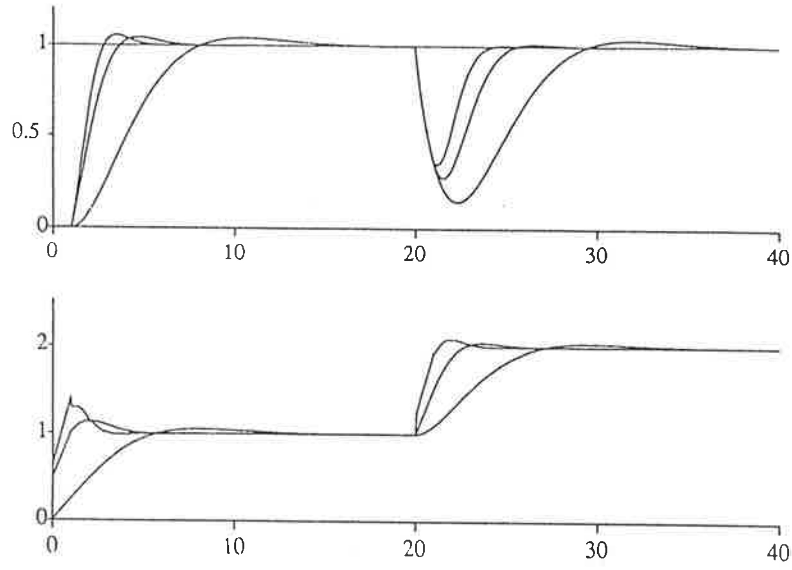

Figure 6. Responses to step changes in command inputs and load for an integrating regulator, a PI regulator with $\omega = 1.1$ and a PID regulator with $\omega = 1.7$.

EXAMPLE 2

Consider a process with the transfer function

$$G(s) = \frac{1}{(1+s)(1+0.2s)(1+0.05s)(1+0.01s)} \tag{18}$$

With pure integral control we get a bandwidth $\omega_0 = 0.67$ rad/s.

We get $k_p = 1$, $\omega_{90} = 1.9, k_{90} = 0.43$, $\omega_{180} = 10.0$, and $k_{180} = 0.04$. Since $\kappa = 25$, it follows that P or PD control can be used if a static error less than 8% is acceptable. The crude assessment indicates that a bandwidth of $\omega = 2$ may be achieved by PI control but that it may be significantly improved to $\omega = 10$ by PID control.

Applying dominant pole design with a relative damping of $\zeta = 0.707$ we get $\omega_I = 0.62(k_i = 0.40)$, $\omega_{PI} = 2.5(k_c = 1.76, k_i = 2.35, T_i = 0.75)$, $\omega_{PD} = 11(k_c = 9.72, T_d = 0.15)$, and $\omega_{PID} = 7.5(k_c = 12.2, k_i = 27, T_i = 0.45, T_d = 0.12)$. We thus find that pure integral control gives the bandwidth $\omega_I = 0.62$. With PI control bandwidths in the range $0.62 \leq \omega \leq 2.5$ can be achieved. A PID regulator gives bandwidths in the range $3.7 < \omega < 7.5$ and PD control gives $3.7 < \omega < 11$. In this case it is thus possible to obtain significant improvements in response speed by incorporating derivative action.

To consider load disturbances we will analyze the integral of the error due to a step disturbance in the load given by Equation (4). By switching from an integration regulator the error integral is reduced from 2.5 to 0.42. With PID control it is further reduced to 0.037. If load disturbances are important there are thus significant improvements by using derivative action.

Figure 7 shows responses to step changes in command and load signals for the regulators.

Notice that the crude assessment gives reasonable estimates. Also notice that this case is quite different from the previous example, where the gain by introducing derivative action was moderate.

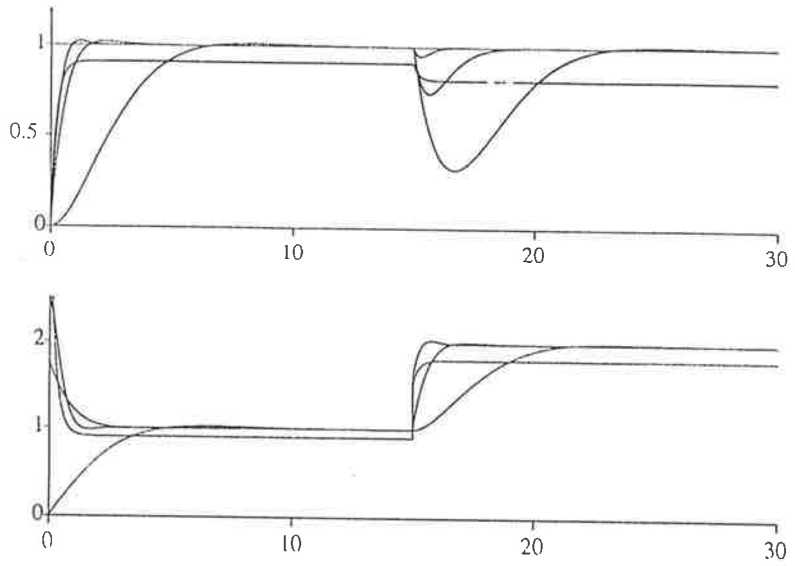

Figure 7. Responses to step changes in load and command signals for an integrating regulator with $\omega = 0.62$, a PD regulator with $\omega = 11$, a PI regulator with $\omega = 2.5$, and a PID regulator with $\omega = 7.5$.

9. Conclusions

Techniques for assessment of the performance achievable for control of simple systems have been explored. A collection of criteria and heuristics based on simple process characteristics, have been presented. The results give the performance ranges that can be achieved using regulators of the PID type.

While automatic tuning is now becoming a standard feature of simple PID regulators, these systems have limited capability of judging their own performance. Using the techniques of this paper it seems possible to introduce a higher level into the regulators. The results is thus a step towards the goal of an intelligent PID controller.

The results can also be used in systems for diagnosis, auditing and expert control.

Acknowledgement

I have benefited considerable from discussions with Tore Hägglund, Lars Bååth, Per Persson and C. C. Hang. This work has been supported by the Swedish Board of Technical Development under contract DUP 85-3084P.

10. References

Årzén, K. E. (1987): "Realization of expert system based feedback control," PhD dissertation CODEN: LUTFD2/TFRT-1029, Department of Automatic Control, Lund Institute of Technology, Lund, Sweden.

Åström, K. J. (1988): "Dominant pole placement design of PI regulators," Internal report CODEN: LUTFD2/TFRT-7381, Department of Automatic Control, Lund Institute of Technology, Lund, Sweden.

Åström, K. J., and T. Hägglund (1984): "Automatic tuning of simple regulators with specifications on phase and amplitude margins," *Automatica*, **20**, 645–651.

Åström, K. J., and T. Hägglund (1988a): *Automatic Tuning of PID Regulators*, ISA, Research Triangle Park, NC, USA.

Åström, K. J., and T. Hägglund (1988b): "A new auto-tuning design," *Preprints IFAC Int Symposium on Adaptive Control of Chemical Processes ADCHEM '88*, Copenhagen, Denmark.

Åström, K. J., J. J. Anton, and K.-E. Årzén (1986): "Expert control," *Automatica*, **22**, 277–286.

Åström, K. J., C. C. Hang, and P. Persson (1988): "Heuristics for assessement of PID ontrol with Ziegler-Nichols tuning," Internal report CODEN: LUTFD2/TFRT-7404, Department of Automatic Control, Lund Institute of Technology, Lund, Sweden.

Bristol, E. H. (1977): "Pattern recognition: An alternative to parameter identification in adaptive control," *Automatica*, **13**, 197–202.

Cohen, G. H., and G. A. Coon (1953): "Theoretical consideration of retarded control," *Transactions of the ASME*, **75**, 827–834.

Deshpande, P. B., and R. H. Ash (1981): *Computer Process Control*, ISA, Research Triangle Park, NC, USA.

Hang, C. C., and K. J. Åström (1988a): "Practical aspects of PID auto-tuners based on relay feedback," *Preprints IFAC Int Symposium on Adaptive Control of Chemical Processes ADCHEM '88*, Copenhagen, Denmark.

Hang, C. C., and K. J. Åström (1988b): "Refinements of the Ziegler-Nichols tuning formula for PID auto-tuners," *Proceedings ISA Annual Conference*, Houston, USA.

Kraus, T. W., and T. J. Myron (1984): "Self-tuning PID controller uses pattern recognition approach," *Control Engineering*, June, 106–111.

Ziegler, J. G., and N. B. Nichols (1942): "Optimum settings for automatic controllers," *Transactions of the ASME*, **64**, 759–768.