Pergamon

Computers chem. Engng Vol. 20, No. 4, pp. 373–386, 1996

0098-1354(95)00000-Q
0098-1354/96 $15.00 + 0.00

A PROCEDURE FOR SISO CONTROLLABILITY ANALYSIS—WITH APPLICATION TO DESIGN OF pH NEUTRALIZATION PROCESSES

S. Skogestad

Chemical Engineering Department, University of Trondheim—NTH, N-7034 Trondheim, Norway

Abstract—A procedure for analyzing the input-output controllability of single-input single-output (SISO) systems is presented. This procedure is applied to a pH neutralization process which must be redesigned to get acceptable controllability. It is found that more or less heuristic design rules given in the literature follow directly. The key steps in the analysis are to consider disturbances and to scale the variables properly. It is suggested that most of the material presented in this paper is suitable for an undergraduate control course.

1. INTRODUCTION

In process control courses the issues of controller design and stability analysis are often emphasized. However, in practice the following three issues are usually more important:

I. How well can the plant be controlled? Before attempting to start any controller design one should have some idea of how easy the plant actually is to control. Is it a difficult control problem? Indeed, does there even exist a controller which meets the required performance objectives?

II. What control strategy should be used? Another important question is to decide on the control strategy: What to measure, what to manipulate, how to pair? In textbooks one finds qualitative rules for this. For example in Seborg *et al.* (1989) one finds in a chapter called "The art of process control" the rules:

1. Control outputs that are not self-regulating.
2. Control outputs that have favorable dynamic and static characteristics, i.e. there should exist an input with a significant, direct and rapid effect.
3. Select inputs that have large effects on the outputs.
4. Select inputs that rapidly effect the controlled variables.

These rules are reasonable, but what is "self-regulating", "large", "rapid" and "direct"? One objective of this paper is to quantify these terms.

III. How should the process be changed to improve control? For example, one may want to find the required size of a buffer tank for damping a disturbance, or one may want to know how fast a measurement should be to get acceptable control.

Controllability analysis. All the above three questions are related to the inherent control characteristics of the process itself, that is, to what is denoted the *input–output controllability* of the process. We shall use the following definition:

> (Input–output) Controllability is the ability to achieve acceptable control performance, that is, to keep the outputs (y) within specified bounds or displacements from their setpoints (r), in spite of unknown variations such as disturbances (d) and plant changes, using available inputs (u) and available measurements (e.g. y_m or d_m).

In summary, a plant is controllable if there *exists* a controller (connecting measurements and inputs) that yields acceptable performance for all expected plant variations. Thus, controllability is independent of the controller, and is a property of the plant (process) only. It can only be affected by changing the plant itself, that is, by design modifications. These may include:

1. Change the apparatus itself, e.g. type, size, etc.
2. Relocate sensors and actuators.
3. Add new equipment to dampen disturbances, e.g. buffer tanks.
4. Add extra sensors for measurement (to be used in feedforward and cascade control).
5. Add extra actuators (to be used for parallel control).
6. Change the control objectives.
7. Change the structure of the lower levels of control already in place.

(It may be argued whether it is appropriate to label the last two items as design modifications, but at least they address issues which come before the actual controller design.)

Surprisingly, in spite of the fact that mathematical methods are used extensively for control system design, the methods available when it comes to controllability analysis are largely qualitative. In most cases the "simulation approach" is used. However, this requires a specific controller design and specific values of disturbances and setpoint changes. In the end one never really knows if a result is a fundamental property of the plant or if it depends on these specific choices. The objective of the paper is to present a procedure for controllability analysis for scalar systems and to apply this procedure to a few examples. Earlier work on input–output controllability analysis includes that of Ziegler and Nichols (1943), Rosenbrock (1970) and Morari (1983) who made use of the concept of "perfect control".

One shortcoming with the controllability analysis presented in this paper is that all the measures are linear. This may seem very restrictive, but usually it is not. In fact, one of the most important nonlinearities, namely that of input constraints, can be handled quite well with a linear anlysis. To deal with slowly varying changes one may perform a controllability analysis at several selected operating points. As a last step one may perform some nonlinear simulations to confirm the linear controllability analysis. Experience from a large number of case studies confirms that the agreement is generally very good.

Remarks on the definition of controllability. The above definition is in tune with most engineers' intuitive feeling about the term, and was also how the term was used historically in the control literature. For example, Ziegler and Nichols (1943) define controllability as "*the ability of the process to achieve and maintain the desired equilibrium value*". Unfortunately, in the '60s the term "controllability" became synonymous with the rather narrow concept of "state controllability" introduced by Kalman, and the term is still used in this restrictive manner in the system theory community. "State controllability" is the ability to bring a system from a given initial state to any final state (but with no regard to the quality of the response between these two states). This concept is of interest for realizations and numerical calculations, but as long as we know that all the unstable modes are both controllable and observable, it has little practical significance. For example, Rosenbrock (1970, p. 177) notes that "most industrial plants are controlled quite satisfactorily though they are not [state] controllable". To avoid confusion with Kalman's state controllability, Morari (1983) introduced the term "dynamic resilience". However, this term does not capture the fact that it is related to control, and instead it is proposed to use the term "input–output controllability" if one explicitly wants to make the distinction with "state controllability".

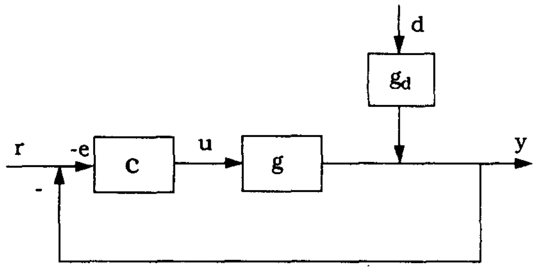

Fig. 1. Block diagram of feedback control system.

2. CONTROLLABILITY ANALYSIS

Controllability may be analyzed by formulating mathematically the control objectives, and then synthezing a controller to see whether the objectives can be met. However, in practice such an approach is difficult and time consuming. The objective of this section is to present simple rules which do not require that a detailed controller design is performed. Consider a linear process model in terms of deviation variables:

$$y = g(s)u + g_d(s)d. \tag{1}$$

Here y denotes the output, u the manipulated input and d a disturbance (may include a disturbance entering at the input which are frequently referred to as a "load change"). $g(s)$ and $g_d(s)$ are transfer function models which describe the effect on the output of the input and disturbance, and all controllability results in this paper are based on this information. The Laplace variable s is often omitted to simplify notation. The control error e is defined as:

$$e = y - r, \tag{2}$$

where r denotes the reference value (setpoint) for the output. In this paper we mostly consider feedback control as illustrated in Fig. 1 where:

$$u = c(s)(r - y) \tag{3}$$

and $c(s)$ is the controller. Eliminating u from equations (1) and (3) yields the closed-loop response:

$$y = Tr + Sg_d d; \quad e = -Sr + Sg_d d. \tag{4}$$

Here the sensitivity is $S = (1+gc)^{-1}$ and the complementary sensitivity is $T = gc(1+gc)^{-1} = 1 - S$. The transfer function around the feedback loop is denoted L. In this case $L = gc$.

In this paper *bandwidth is defined as the frequency* ω_B *where the loop gain is one in magnitude*, i.e.

$|L(j\omega_B)|=1$ (or more precisely where the low-frequency asymptote of $|L|$ first crosses 1 from above). This frequency is also called the "gain crossover frequency". Other definitions of bandwidth are used, but the difference is small. At frequencies lower than the bandwidth ($\omega<\omega_B$) feedback is effective and will affect the frequency response. However, for sinusoidal input signals (for example, a disturbance) with frequencies higher than ω_B the response will not be affected much by the feedback.

The simplest interpretation of the frequency domain is that it represents the steady-state sinusoidal response. For example, if we send an input $u(t)=u_0\sin(\omega t)$ through a stable system with transfer function $g(s)$, then the output as $t\to\infty$ is $y(t)=y_0(\sin\omega t+\phi)$ where $y_0=|g(j\omega)|u_0$ and $\phi=\angle g(j\omega)$. Here $g(j\omega)$ represents at each frequency ω a complex number obtained from $g(s)$ by setting $s=j\omega$. A common shorthand notation used in this paper to express the sinusoidal response is (phasor notation):

$$y(\omega)=g(j\omega)u(\omega), \tag{5}$$

where $y(\omega)$ and $u(\omega)$ are complex numbers (vectors) representing at each frequency the size and phase of a sinusoidal signal. For example, $u(\omega)=5$ means that $u(t)=5\sin(\omega t)$. Thus $u(\omega)$ is *not* equal to $u(s)$ evaluated at $s=\omega$, nor is it equal to $u(t)$ evaluated at $t=\omega$.

2.1. Scaling. The interpretation of most measures presented in this paper assumes that the transfer functions g and g_d are in terms of scaled variables. The first step in a controllability analysis is therefore to scale (normalize) all variables (input, disturbance, output) to be less than 1 in magnitude (i.e. within the interval -1 to 1) by normalizing each variable by its maximum value, for example, $u=u'/u'_{max}$ where u' denotes the unscaled and u the scaled variable, and u'_{max} is the largest allowed input change (in unscaled variables). For the other variables we have $d=d'/d'_{max}$, $e=e'/e'_{max}$, $y=y'/e'_{max}$ and $r=r'/e'_{max}$, where d'_{max} is the largest expected disturbance and e'_{max} the largest allowed control error. In most cases the maximum values (u'_{max}, e'_{max}, d'_{max}) are assumed independent of frequency.

Thus, in the following *we assume that the signals are persistent sinusoids, and that g and g_d have been scaled, such that at each frequency the allowed input $|u(\omega)|<1$, the expected disturbance $|d(\omega)|<1$, the expected reference signal $|r(\omega)|<r_{max}(\omega)$*. The performance requirement is that the control error $|e(\omega)|<1$, Note that e and r are measured in the same units so $r_{max}=r'_{max}/e'_{max}$ is the magnitude of the largest expected setpoint change relative to the allowed control error. We will assume that $r_{max}(\omega)$ is frequency dependent such that $|r_{max}(\omega)|=R_{max}$ up to the frequency ω_r and is 0 above this frequency. In other words, for a setpoint change $r(t)=R_{max}\sin(\omega t)$, the tracking error $e(t)=y(t)-r(t)$ should be less than one in magnitude up to the frequency ω_r, and above this frequency there are no specifications on tracking. Throughout the paper we assume $R_{max}>1$.

Remark 1. It could be argued that the magnitude of the sinusoidal disturbances should approach zero at high frequency, that is d'_{max} should be made frequency dependent. While this may be true, we really only care about frequencies up to the bandwidth ω_B, and in most cases it is reasonable to assume that we do indeed have sinusoidal disturbances of about the same magnitude up to this frequency.

Remark 2. It could also be argued that e'_{max} should be frequency dependent. For example, we may require no steady-state offset, i.e. e'_{max} should approach zero at low frequencies. Again, including frequency variations is not recommended when doing a preliminary controllability analysis (however, one may take such considerations into account when interpreting the results of the controllability analysis). Of course, if we were using e'_{max} to derive weighting functions to use for controller synthesis, then it should be made frequency dependent.

Remark 3. If more detailed information is given about the desired setpoint changes one may want to use another form for the frequency dependency of r_{max}.

2.2. Summary of controllability rules for feedback control

Scale the variables d, u, y and r as outlined above to obtain the scaled transfer functions $g(s)$ and $g_d(s)$. Let $g_m(s)$ denote the measurement transfer function and assume $g_m(0)=1$ (perfect steady-state measurement). Let ω_B denote the bandwith of the system, defined as the highest frequency where $|L(j\omega_B)|=1$ (see above). Let ω_d denote the frequency at which $|g_d(j\omega_d)|$ first crosses 1 from above. The following rules apply:

Rule 1. Speed of response to reject disturbances. Must at least require $\omega_B>\omega_d$. More specifically, we must with feedback control require $|L|=|gc(j\omega)|>|g_d(j\omega)|$ at frequencies where $|g_d(j\omega)|>1$.

Justification: Without control $y=g_d d$. Scaling has been applied such that the largest disturbance at a given frequency is $d(t)=1\cdot\sin(\omega t)$ [i.e. $|d(\omega)|=1$]. Thus, at frequencies $\omega<\omega_d$ the output y will be unacceptable ($|y|>1$) for a disturbance $|d|=1$, so control is needed at these frequencies, and we must require $\omega_B\geq\omega_d$.

With feedback control $y = Sg_d d$ where $S = (1/(1+L)) \approx (1/L)$ at frequencies where $|L| > 1$. Thus to have $|y| < 1$ for $|d| = 1$ we must require $|L| > |g_d|$ at these frequencies.

Rule 2. Speed of response to follow setpoints. Must at least require $\omega_B > \omega_r$ where ω_r is the frequency up to which tracking is required. More specifically, we must require $|L(j\omega)| > R_{max}$ up to frequency ω_r.

Unless R_{max} is close to 1, the requirement $\omega_B > \omega_r$ is not tight, and a higher bandwidth is required in practice. The exact value depends on how sharply $|L(j\omega)|$ drops off in the frequency range from ω_r (where $|L| > R_{max}$) to ω_B (where $|L| = 1$). For example, with $L(s) = \omega_B/s$ (first-order response) the required bandwidth is $\omega_B > \omega_r R_{max}$, while for $L(s) = \omega_B^2/s^2$ (not considering stability) the required bandwidth is $\omega_B > \omega_r \sqrt{R_{max}}$.

Justification: With feedback control $e = -Sr$ where $S \approx 1/L$ at frequencies where $|L| > 1$. Thus to have $|e| < 1$ for $|r| = |R_{max}|$ (up to frequency ω_r) we must require $|L| > |R_{max}|$.

Rule 3. Input constraints for disturbances. Must require $|g(j\omega)| > |g_d(j\omega)|$ at frequencies where $|g_d(j\omega)| > 1$.

Justification: This is needed to avoid input constraints when perfectly rejecting a disturbance $d(t) = 1 \cdot \sin(\omega t)$ [*i.e.* $d(\omega) = 1$]: From $y = gu + g_d d = 0$ we get $u = -(g_d/g)d$ and with $d = 1$ we need $|u| = |g_d|/|g| < 1$ to avoid input constraints.

Strictly speaking, perfect control is not required, and the minimum input needed for "acceptable" control (namely $|y| < 1$) is $|u| = (|g_d| - 1)/|g|$. (Consider $y = gu + g_d d$ with $d = 1$, then the smallest required input to get $|y| = 1$ is found when u is such that the complex vectors gu and g_d are in opposite directions, i.e. $|y| = 1 = |g_d| - |gu|$.) The difference is clearly small at low frequencies where $|g_d|$ is larger than 1. (However, for multivariable systems the difference may be large for ill-conditioned plants even at low frequencies.)

Rule 4. Input constraints for setpoints. Must require $|g(j\omega)| > R_{max}$ up to frequency ω_r where tracking is required.

Justification: This is needed to avoid input constraints for perfect tracking of $|r(\omega)| = R_{max}$: from $y = gu$ and $y = r$ (perfect control) we get $u = r/g$, and with $r = R_{max}$ (up to frequency ω_r) we need $|u| = R_{max}/|g| < 1$ to avoid input constraints.

Rule 5. Time delay θ in $g(s)g_m(s)$. Must require $\omega_B < 1/\theta$ to have acceptable control performance.

Justification: It is impossible to remove the effect of the delay and $L(s)$ must contain a term $e^{-\theta s}$. For example, the ideal controller which minimizes $J = \int_0^\infty |e(t) - r(t)|^2 \, dt$ when $r(t)$ is a step and there is no penalty on the inputs has complementary sensitivity $T = e^{-\theta s}$. The corresponding loop gain $L = T/(1-T)$ crosses 1 in magnitude at about the frequency $1/\theta$. In practice, the ideal controller cannot be realized so this value provides an upper bound on the bandwidth.

Rule 6. Real RHP-zero z in $g(s)g_m(s)$. Must require $\omega_B < z/2$ to have acceptable control performance at low frequencies.

Justification: Again, it is impossible to remove the effect of a RHP-zero. The ideal controller which minimizes $J = \int_0^\infty |e(t) - r(t)|^2 \, dt$ when $r(t)$ is a step and there is no penalty on the inputs has complementary sensitivity $T = (-s+z)/(s+z)$. The corresponding loop gain $L = T/(1-T)$ crosses 1 in magnitude at about the frequency $z/2$. In practice, the ideal controller cannot be realized so this value provides an upper bound on the bandwidth.

Remark: Strictly speaking, a RHP-zero only makes it impossible to have tight control in the frequency range close to the location of RHP-zero. If we do not need tight control at low frequencies, then we may reverse the sign of the controller gain, and instead achieve tight control at frequencies higher than z. One special example is for plants with a zero at the origin [$g(s)$ contains an isolated term s in the numerator] where one can achieve good transient control, but control has no effect at steady-state.

Rule 7. Phase lag constraint. Must require in most practical cases: $\omega_B < \omega_u$. Here the "ultimate" frequency, ω_u is where the phase of $g(j\omega)g_m(j\omega)$ is $-180°$.

This rule is given by Balchen and Mumme (1988), but without any theoretical justification. In fact, the condition is *not* a fundamental limitation, since for minimum phase plants (no delays or RHP-zeros), any phase lag may in theory be counteracted (disregarding input constraints) by placing zeros in the controller (use of "derivative action"). However, in practice this is not possible, because the controller structure may be limited and because of model uncertainty.

Justification for PID-controller: With a PID-controller the maximum phase lead is 54.9° for a controller with derivative action over one decade (the maximum phase lead for the term $(\tau_d s + 1)/(0.1\tau_d s +)$ is 54.9° at frequency $\sqrt{10}/\tau_d$). Thus, if we require a phase margin larger than 54.9° we must require $|L| \leq 1$ at frequency ω_u and the rule follows.

Rule 8. Real open-loop unstable pole in $g(s)$ at $s = p$. Need high feedback gains to stabilize the system and must require for acceptable performance $\omega_B > 2p$.

Justification: For example, to stabilize a plant $g(s) = 1/(s-p)$ with a constant gain controller $c(s) = K_c$ we need $K_c > p$, and we find that the asymptote of $|L|$ crosses 1 at frequency K_c, so we have $\omega_B > K_C = p$. This is a minimum requirement for stability. For performance the gain must be larger and the value $K_c = 2p$ places the closed-loop pole at the mirror image of $s = p$ and yields the minimum value of the input [in tems of $\int_0^\infty |u(t)|^2\,dt$] required for stabilization (Kwakernaak and Sivan, 1972, p. 289).

Another justification follows from the fact that a strictly proper plant with a RHP-zero and a single unstable real pole, e.g. $g(s) = (s-z)/((s-p)(\varepsilon s+1))$, can be stabilized by a stable controller if and only if $p < z$ (Youla *et al.*, 1974). (Combining Rules 6 and 8 yields $p < 0.25z$ because the "gap" must be larger to achieve reasonable performance.)

In addition, for unstable plants we need $|g| > |g_d|$ up to the frequency p (which may be larger than ω_d). Otherwise, the input may saturate when there are disturbances, and the plant cannot be stabilized.

Most of the rules are summarized graphically in Fig. 2. The above rules are necessary conditions ("minimum requirements") in order to achieve acceptable control performance. One reason they are not sufficient is that they are based on considering only "one effect at a time".

The rules quantify the qualitative rules from Seborg *et al.* (1989) given in the introduction. For example, the rule "Control outputs that are not self-regulating" may be quantified as: "Control outputs y for which $|g_d(j\omega)| > 1$ at some frequency" (Rule 1). The rule "Select inputs that have a large effect on the outputs" may be quantified as: "In terms of scaled variables we must have $|g| > |g_d|$ at frequencies where $|g_d| > 1$ (Rule 3), and we must have $|g| > R_{max}$ at frequencies where setpoint tracking is desired (Rule 4)".

Another important insight from the above rules is that a larger disturbance or a smaller allowed control error requires faster response (higher bandwidth).

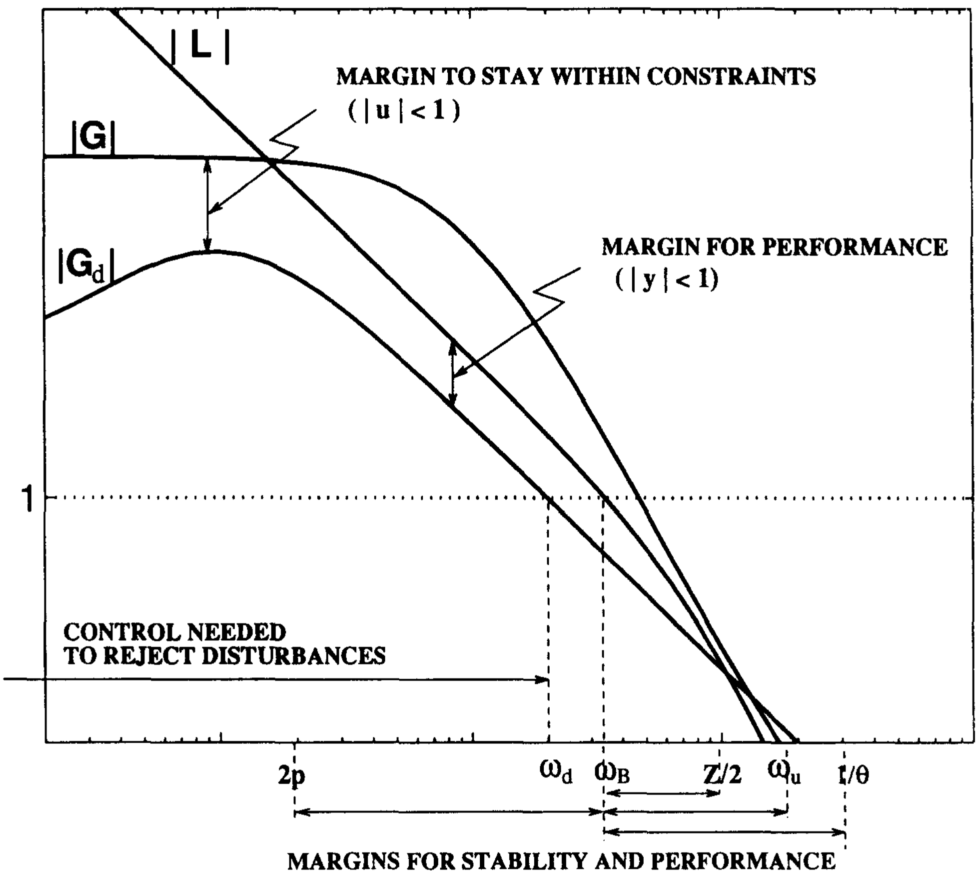

Fig. 2. Summary of controllability requirements.

In summary Rules 1, 2 and 8 tell us that we need high feedback gain ("fast control") in order to reject disturbances, to track setpoints and to stabilize the plant. On the other hand, Rules 5–7 tell us that we must use low feedback gains in the fequency range where there are RHP-zeros or delays or where the plant has a lot of phase lag. We have formulated these requiremnts for high and low gain as bandwidth requirements. If they somehow are in conflict then the plant is not controllable and the only remedy is to introduce design modifications. Often the problem is that the distrubances are too large such that we hit input constraints, or such that the required bandwidth is not achievable. To avoid the latter problem, we must at least require that the effect of the disturbance is less than 1 (in terms of scaled variables) at frequencies beyond the bandwidth, that is:

$$|g_d(j\omega)| < 1; \quad \forall \omega \geq \omega_B, \tag{6}$$

where as found above we must require (approximately) $\omega_B < 1/\theta$ and $\omega_B < z/2$ and $\omega_B < \omega_u$. Condition (6) may be used, as in the pH neutralization example below, to determine the size of buffer tanks.

2.3. *Feedforward control*

Consider a feedforward controller $u = c_f(s)d_m$ where $d_m = g_{md}(s)d$ is the measured disturbance. The disturbance response becomes:

$$y = gu + g_d d = \underbrace{(gc_f g_{md} + g_d)}_{\hat{g}_d} d, \tag{7}$$

where $\hat{g}_d(s)$ denotes the effect of the disturbance with the feedforward controller in place. We want to consider controllability (achievable performance) with feedforward control.

Rules 3 and 4 on input constraints apply directly to feedforward control, while Rule 8 does not apply since unstable plants can only be stabilized by feedback control. The remaining rules make use of the term "bandwidth" which we above defined as the frequency up to which the feedback loop gain $|L|$ is larger than one. However, if the term "bandwidth" (ω_B) is interpreted as "the frequency up to which control is effective" then the rules partly apply also to feedforward control. Rules 5 and 6 on time delay and RHB-zero must be modified by replacing gg_m by $g_d^{-1}gg_{md}$. This follows by considering the ideal feedforward controller which yields $\hat{g}_d = 0$ in (7). We get:

$$c_f^{\text{ideal}} = -g_d g^{-1} g_{md}^{-1}, \tag{8}$$

which should be stable and causal (contain no prediction) to be realizable. Note that a delay in $g_d(s)$ is an advantage for feedforward control ("it gives the feedforward controller more time to make the right action").

Model uncertainty is a more serious problem for feedforward than for feedback control because there is no correction from the output measurement. Let the actual plant models be denoted as g'', g_d'' and g_{md}''. Then the actual disturbance response with the ideal feedforward controller in (8) is (assuming that this controller is realizable(:

$$y = g''u + g_d''d = \underbrace{g_d'\left(1 - \frac{g_d}{g_d''}\frac{g''}{g}\frac{g_{md}''}{g_{md}}\right)}_{\hat{g}_d''} d. \tag{9}$$

Here g_d'' is the actual disturbance response without feedback control and $\hat{g}_d''$ with feedforward control. The effectiveness of feedforward control is determined by the ratio $|\hat{g}_d''|/|g_d''|$, (which takes the place of the sensitivity function for feedback control). Ideally it is zero, but this requires accurate models of g and g_d as well as of the measurement g_{md}. For example, a 10% error in each of these three may yield $|\hat{g}_d''|/|g_d''| = |1 - 1.1 \cdot 1.1 \cdot 1.1| = 0.33$, that is, because of uncertainty even the ideal feedforward controller removes only 67% of the disturbance effect. If the ratio is larger than 1 at some frequency (which may easily happen) then feedforward control makes control worse.

Because of the sensitivity to model uncertainty and because of the presence of unmeasured disturbances, feedforward control is usually combined with feedback control. Assume that the feedforward controller has already been designed. Then the controllability of the remaining feedback problem can be analyzed using the above rules if $g_d(s)$ is replaced by $\hat{g}_d(s)$.

3. SIMPLE EXAMPLES

3.1. *First-order process with delay*

Consider disturbance rejection for the following process:

$$g(s) = k\frac{e^{-\theta s}}{1 + \tau s}; \qquad g_d(s) = k_d\frac{e^{-\theta_d s}}{1 + \tau_d s}. \tag{10}$$

In addition there is a measurement delay θ_m for the output and θ_{md} for the disturbance. All parameters have been appropriately scaled such that at each frequency $|u| < 1$, $|d| < 1$ and we want $|y| < 1$. One interesting question is: for each of the eight parameters k, τ, θ, k_d, τ_d, θ_d, θ_m and θ_{md}, what value is preferred to for good controllability?

Qualitative results are given in Table 1. Essentially, the effect of the manipulated input should be as large and quick as possible, whereas the

Table 1. Desired value of parameters to have good controllability

	Feedback control	Feedforward control
k	Large	Large
τ	Small	Small
θ	Small	Small
k_d	Small	Small
τ_d	Large	Large
θ_d	No effect	Large
θ_m	Small	No effect
θ_{md}	No effect	Small

opposite is true for the disturbance. The main difference between feedback and feedforward control is that a delay in the disturbance has no effect for feedback control, while it is an advantage for feedforward control as it leaves more time to take the appropriate control action.

We now want to quantify the statements in Table 1. Assume $k_d > 1$ such that control is needed. From Rule 1 we need for acceptable performance ($|y| < 1$) with disturbances:

$$\omega_d \approx k_d \tau_d < \omega_B. \tag{11}$$

On the other hand from Rule 5 we must for stability (and performance) require:

$$\omega_B < 1/\theta_{tot}, \tag{12}$$

where θ_{tot} is the total delay around the loop. Combining (11) and (12) yields $\omega_d < 1/\theta_{tot}$ or:

$$\theta_{tot} = \theta + \theta_m < \tau_d / k_d. \tag{13}$$

For feedforward control any delay for the disturbance itself yields a smaller "net delay", and to have $|y| < 1$ we require:

$$\theta + \theta_{md} < \tau_d / k_d + \theta_d. \tag{14}$$

To stay within the constraints ($|u| < 1$) we must from Rule 4 require $|g(j\omega)| > |g_d(j\omega)|$ for frequencies $\omega < \omega_d$. Specifically, for both feedback and feedforward control:

$$k > k_d; \qquad k/\tau > k_d/\tau_d. \tag{15}$$

3.2. *Step response controllability analysis*

The controllability analysis presented in this paper is based in the frequency domain. However, many engineers feel more comfortable with the time domain and step responses. Consider a unit step disturbance, $d = 1$, to the first-order with delay plant in equation (10). Without control the output response for $t > \theta_d$ is:

$$y(t) = k_d(1 - e^{-(t-\theta_d)/\tau_d}). \tag{16}$$

The response is shown graphically in Fig. 3. Since $k_d > 1$ the output $y(t)$ will exceed 1 after some time. Disregarding for a moment the delay, the time where $y(t) = 1$ is at $t = -\tau_d \ln(1 - (1/k_d)) \approx \tau_d / k_d$ (the approximation holds for $k_d \gg 1$ and corresponds to the point where the initial tangent of the time response crosses 1, see Fig. 3). Assuming that we measure the disturbance (feedforward control), the "minimum reaction time" to achieve $|y| < 1$ is then (see Fig. 3) $(\tau_d / k_d) + \theta_d$.

This is then an upper bound on the allowed delay in the process (Walsh, 1993). This is the same value as was obtained in equation (14) using the frequency domain in the case of feedforward control.

From this example we see that a step response controllability analysis yields results similar to the frequency domain, at least for a first-order process and feedforward control. For feedback control a step response controllability analysis is generally less suitable. For example, one cannot simply measure the time it takes from when the disturbance enters to when the output exceeds its maximum value (which is 1 in terms of the scaled variables used in this paper) and use this as the minimum response time for disturbance rejection. As shown by Fig. 3 this time depends on the delay in the disturbance model, whereas we know that θ_d should not matter for rejecting distrubances with feedback control. In conclusion, the frequency domain should generally be used for controllability analysis, and the purpose of this example was *not* to suggest using step responses, but to provide another justification for the usefulness of the frequency domain.

4. NEUTRALIZATION PROCESS

4.1. *One tank*

Consider the process in Fig. 4 where a strong acid (pH = − 1) is neutralized by a strong base (pH = 15) in one mixing tank with volume $V = 10\ m^3$ to produce $q = 0.01\ m^3/s = 10$ l/s of "salt water". The problem is to use feedback control to keep the pH in the product stream in the range 7 ± 1 ("salt water") by

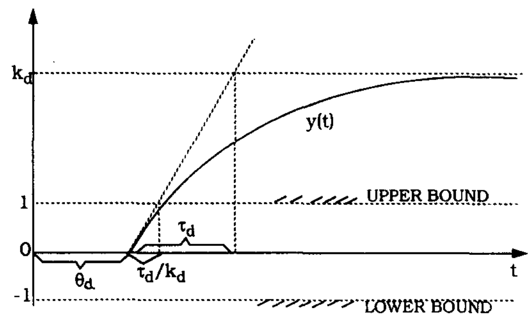

Fig. 3. Response for step distrubance, $g_d = \frac{k_g e^{-\theta_d s}}{(1 - \tau_d s)}$.

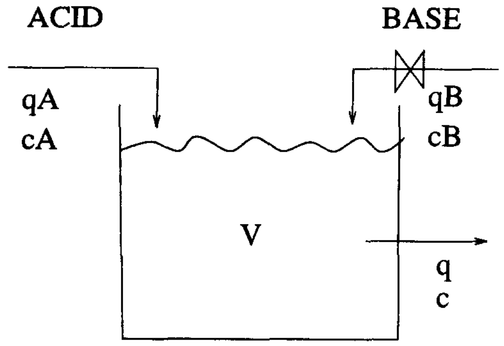

Fig. 4. Neutralization process with one mixing tank.

manipulating the amount of base q_B. The delay for the measurement of pH is $\theta = 10$ s.

To achieve the desired product with pH = 7 one must exactly balance the inflow of acid (the disturbance) by addition of base (the manipulated input). Intuitively, one might expect that the main control problem is to adjust the base accurately, and therefore that a very accurate valve is needed. However, as we shall see this "feedforward" way of thinking is misleading, and the main hurdle to good control is that very fast response times are needed.

A dynamic model is given in the Appendix. For the controlled output we introduce the excess of acid c (mol/l) defined as:

$$c = c_H - c_{OH}. \tag{17}$$

Somewhat surprisingly, we find that in terms of c the dynamic model, which is usually believed to be strongly nonlinear, is given by that of a simple mixing process:

$$\frac{d}{dt}(Vc) = q_A c_A + q_B c_B - qc. \tag{18}$$

Introduce the following scaled variables

$$y = \frac{c}{10^{-6}}; \qquad u = \frac{q_B}{q_B^*}; \qquad d = \frac{q_A}{0.5 q_A^*}, \tag{19}$$

where superscript * denotes the steady-state value. The appropriately scaled linear model with one tank then becomes (see Appendix):

$$g_d(s) = \frac{k_d}{1 + \tau s}; \qquad g(s) = -2g_d(s); \qquad k_d = 2.5 \times 10^6, \tag{20}$$

where $\tau = V/q = 1000$ s. Note that the steady-state gain in terms of scaled variables is more than a million so that the output is extremely sensitive to both u and d.

We now proceed with the controllability analysis. The frequency responses of $g_d(s)$ and $g(s)$ are shown grahically in Fig. 5. From Rule 2 input constraints do not pose a problem since $|g| = 2|g_d|$ at all frequencies. The main control problem is the high disturbance sensitivity, and from (11) (Rule 1) we find the frequency up to which feedback is needed:

$$\omega_d \approx k_d/\tau = 2500 \text{ rad/s}. \tag{21}$$

This requires a response time of 1/2500 = 0.004 s which is clearly impossible.

The small value of the response time may be explained from a step response analysis as follows: the pH in the tank should remain within 7 ± 1. At pH = 7 the concentration of H^+-ions is $c_H = 10^{-7}$ mol/l. Since the tank volume is $V = 10^4$ l, the amount of H^+-ions in the tank is $c_H V = 10^{-3}$ mol. Similarly, at the lower bound at pH = 6 the amount of H^+-ions in the tank is $10^{-6} \times 10^4 = 10^{-2}$ mol. Thus, adding about 10^{-2} mol of H^+-ions to the tank changes the pH from 7 to 6. Now, the concentration of H^+-ions in the acid stream is 10 mol/l (corresponding to pH = −1), so we only need to add 10^{-3} l of acid (about 20 droplets) to change pH from 7 to 6. The largest expected disturbance ($d = 1$) corresponds to an increase in the acid inflow from 5 to 7.5 l/s. Thus, with a step increase in acid inflow it will only take 10^{-3} l/(2.5 l/s) = 0.004 s to change the pH from 7 to 6, which agrees with the result from the controllability analysis.

4.2. *Design change: multiple tanks*

The only way to improve controllability is by design changes. The most useful change in this case is to do the neutralization in several steps. With n equal mixing tanks in series the transfer function for the effect of the disturbance becomes:

$$g_d(s) = k_d h_n(s); \qquad h_n(s) = \frac{1}{\left(\frac{\tau_h}{n}s + 1\right)^n}, \tag{22}$$

where $k_d = 2.5 \times 10^6$ is the gain for the mixing process, $h_n(s)$ is the transfer function of the mixing tanks and τ_h is the total residence time V_{tot}/q. The magnitude of $h_n(s)$ as a function of frequency is shown in Fig. 6 for one to four equal tanks in series.

From controllability Rule 5 we get that the best achievable closed-loop bandwidth ω_B is about $\omega_\theta \stackrel{\text{def}}{=} 1/\theta$. To be able to reject disturbances, we must then require from equation (6) that:

$$|g_d(j\omega_\theta)| \leq 1. \tag{23}$$

Thus, the purpose of the mixing tanks is to reduce the effect of the disturbance by a factor $k_d = 2.5 \times 10^6$ at the frequency $\omega_\theta = 0.1$ (rad/s). Combining (22) and (23) yields the following minimum value for the total residence time for n equal tanks in series:

$$\tau_h = \theta n \sqrt{(k_d)^{2/n} - 1}. \tag{24}$$

The corresponding total volume is $V_{tot} = q\tau_h$ where $q = 0.01\ m^3/s$. With $\theta = 10\ s$ we then find that the following designs have the same controllability with respect to disturbance rejection:

No. of tanks n	Total volume V_{tot} (m^3)	Volume each tank (m^3)
1	250,000	250,000
2	316	158
3	40.7	13.6
4	15.9	3.98
5	9.51	1.90
6	6.96	1.16
7	5.70	0.81

For example, with one tank we get

$$V_{tot} = q\theta\sqrt{(k_d)^2 - 1} \approx q\theta k_d$$
$$= 0.01\ m^3/s \times 10\ s \times 2.5 \times 10^6 = 2.5 \times 10^5\ m^3.$$

That is, we need a volume corresponding to that of the world's largest ship to get acceptable controllability. The minimum total volume is obtained with 18 tanks of about 203 l each—giving a total volume of $3.662\ m^3$. However, taking into the account the additional cost for extra equipment such as piping, mixing and level control we would probably select a design with 3 or 4 neutralization tanks for this example.

4.3. Control system design

This condition $|g_d(j\omega_\theta)| < 1$ in equation (23), which formed the basis for our controllability analysis, may be optimistic because it does not take into account that we must also reject the disturbance at lower frequencies. The problem is that the disturbance transfer function $g_d(s) = k_d h(s)$ is of high order when n is large. More specifically, $|g_d(j\omega)|$ has a roll-off slope of $-n$ (on a log–log Bode plot) at frequencies higher than the inverse of the residence time. It is then difficult to achieve sufficiently high roll-off in the loop transfer function $L(s)$ to get $|L(j\omega)| > |g_d(j\omega)|$ at frequencies lower than the bandwidth (as required from Rule 1), although we are able to achieve $|L(j\omega_B)| = 1 \geq |g_d(j\omega_B)|$ at the bandwidth, $\omega_B = \omega_\theta$. The reason is that a high roll-off in $L(s)$ yields a large phase lag, and we get stability problems (the Bode stability criterion requires the loop gain to be below 1 at the frequency

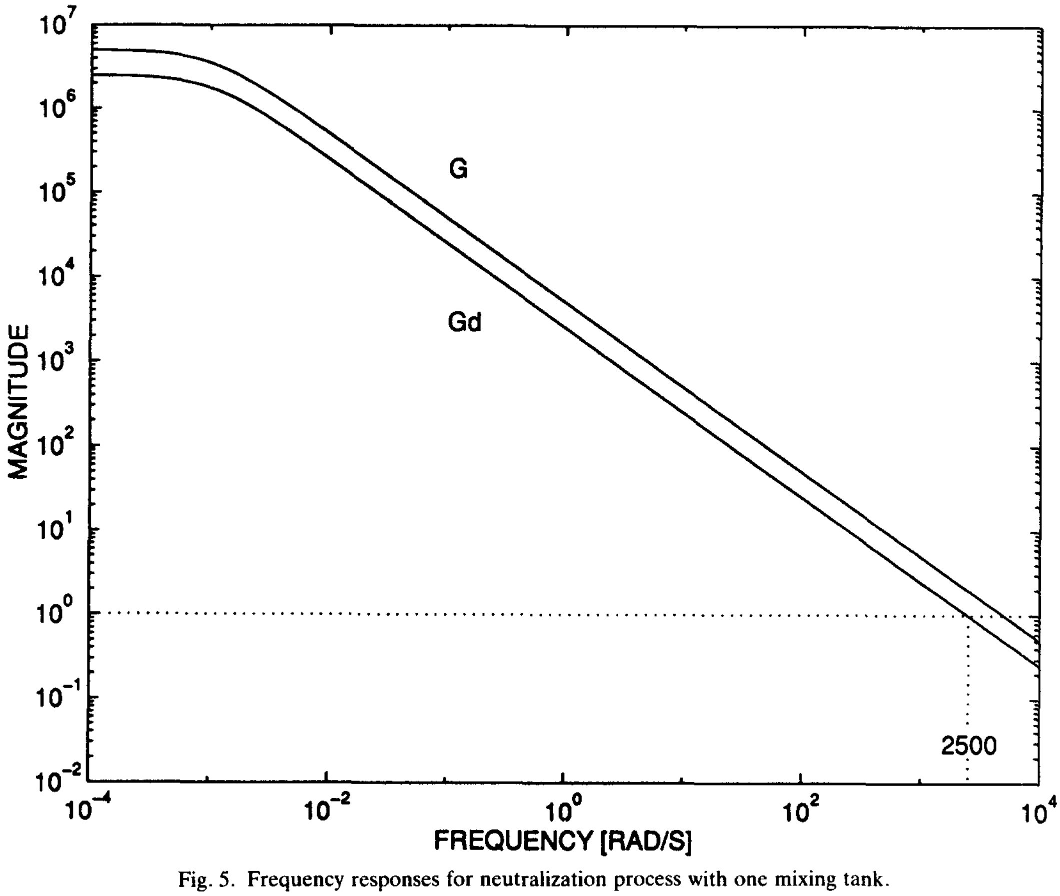

Fig. 5. Frequency responses for neutralization process with one mixing tank.

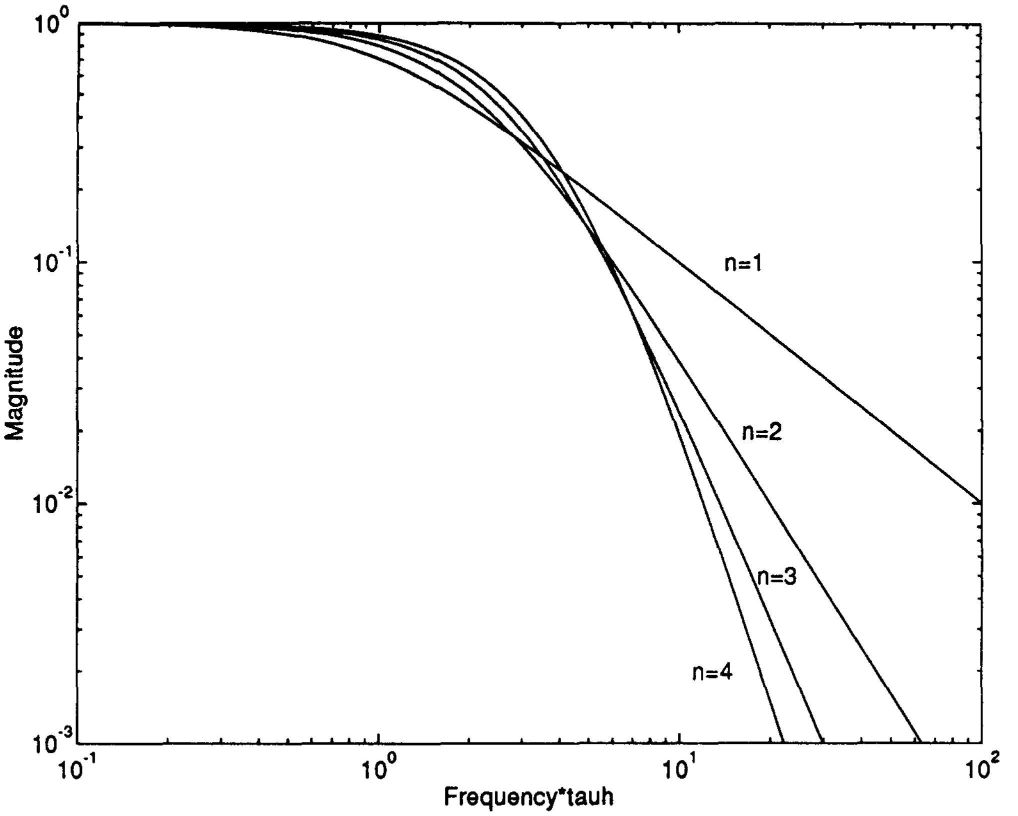

Fig. 6. Frequency responses for n tanks in series with same total residence time τ_h; $h_n(s) = 1/((\tau_h/n)s + 1)^n$, $n = 1, 2, 3, 4$.

where the phase is $-180°$). For example, a roll-off of -2 will even in the best case (minimum phase system) yield a phase lag of $-180°$. Around the bandwidth the delay also contributes to the phase, so in practice the roll-off of $|L|$ at the bandwidth cannot exceed about -1.

In conclusion, if we have two buffer tanks or more and use a control system with a single controller as shown in Fig. 7, then the above controllability analysis is optimistic. Of course, the roll-off for L may be steeper than -1 at lower frequencies, so we may achieve some benefit of using additional tanks. For example, consider the case with two buffer tanks and assume $|L| \approx |g_d|$ at low frequencies, i.e. $|L|$ has a roll-off of -2 at low frequencies. For stability, we need $|L|$ to have a roll-off of -1 around the bandwidth. To this effect, we assume that $|L|$ has a break frequency (a zero) at about $\omega_B/5$ (the phase contribution from the -2 slope will then be about $-90° + \arctan 5 = -11°$) at ω_B. Then $|L|$ will cross 1 at about frequency $\sqrt{5}\,\omega_d$ (using asymptotic values) where ω_d is the frequency where $|g_d|$ crosses 1. In other words $\omega_d \approx \omega_B/2.2$, so we must require $|g_d(\omega_0/2.2)| < 1$, and it follows that with a single controller as in Fig. 7 we must increase the total volume of the two tanks by a factor 2.2 compared with the value found previously. With three or more tanks, the required increase in volume to maintain controllability is even larger. (In this analysis we have assumed that the only limitation on the bandwidth is given by the pH measurement delay θ, and the situation would become even worse if we were to take into account the high order of $g(s)$ and the additional delay caused by incomplete mixing in the tanks.)

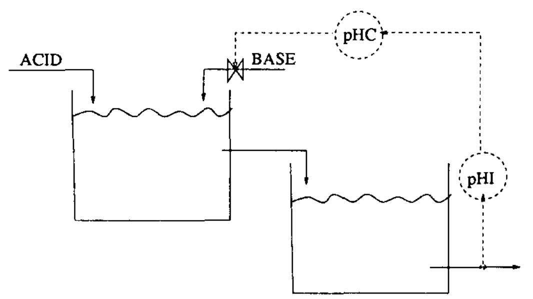

Fig. 7. Neutralization process with two tanks and one controller.

The solution is to install a pH control system on each tank and add base gradually. Consider a case with n tanks in series. The overall closed-loop response from a disturbance into the first tank to the pH in the last tank then becomes:

$$y = g_d \prod_{i=1}^{n} \left(\frac{1}{1+L_i} \right) d \approx \frac{g_d}{L} d,$$

where

$$L = \prod_{i=1}^{n} L_i$$

and L_i denotes the loop transfer function for tank i. In this case we can design each loop $L_i(s)$ with a slope of -1 and bandwidth $\omega_B = \omega_\theta$, such that the overall loop transfer functin L has slope $-n$ and achieves $|L| > g_d|$ at all frequencies. Thus, our analysis confirms the usual recommendaton of adding base gradually and having one pH-controller for each tank (McMillan, 1984, p. 208). It does not seem like any other control strategy can achieve a sufficiently high roll-off for $|L|$.

4.4. Remarks

1. Walsh (1993, p. 31) uses the following data for the capital cost of large mixing tanks (in 1000£s Sterling): $c(\text{kGBP}) = 20 + 2V^{0.7}$ where V is the tank volume in m^3. With these data 3 tanks are best for the above example (capital cost is 97 vs 101 kGBP for 4 tanks). We have not taken into account the cost of pH control systems which according to Walsh each cost about 40 kGBP in capital and 40 kGBP/yr in maintenance. This would further favor the use of three tanks.

2. The results given above compare well with those of other authors. A simple shortcut method given by McMillan (1984, p. 204) is to use one mixing tank for each 2 units change in pH. For example, with a pH change of 8, as in our example (from pH 15 to 7), four tanks are recommended.

3. McMillan (1984, p. 205) also give a more rigorous method based on estimating the peak error when using a PID controller tuned using the Ziegler–Nichols rules. This peak error is compared with the allowed error and the number of tanks is increased until acceptable control is possible. A closer inspection of this method reveals that it yields the same results as obtained with our frequency domain analysis, and is in fact identical to the contollability condition (23).

4. Traditionally, a "feedforward" approach has been taken when considering controllability of such processes, and one key argument has been that control is difficult because one needs to adjust the amount of base extremely accurately to counteract the disturbance in the acid. This is a valid argument for feedforward control, but *not* for feedback control as the feedback control action will be able to adjust the input accurately. As demonstrated above the key problem for feedback control is that the output is extremely sensitive to disturbances (k_d and ω_d are large), which demands an extremely high bandwidth.

5. Of course, feedforward control based on measuring q_A and c_A can be used in addition to feedback to improve performance. According to McMillan (1984, p. 204) one can typically save one mixing tank using a well designed feedforward controller. Actually, we can use a controllability analysis to estimate more accurately the effectiveness of applying feedforward control to the first tank. For example, consider the case with three tanks where the total required volume with feedback alone was found to be 40.7 m^3. Assume that the feedforward controller is able to remove 80% of the disturbance effect, that is, assume $\hat{k}_d = 0.2k_d = k_d/5$. This is rather optimistic [recall equation (9)] and requires accurate measurements as well as being able to add the base at the right time. From equation (24) we then find that by adding feedforward control the required volume for the 3 buffer tanks may be reduced by about a factor $5^{1/3} = 1.71$ to 23.8 m^3.

6. In terms of minimizing the total volume it is almost always optimal to have mixing tanks of equal size. (The only possible exception is for disturbances at frequencies lower than about n/τ_h, see Fig. 6, where it is slightly better to use *fewer* tanks, but in this frequency range the tanks may not have much effect.) Still, there are some suggestions in the literature regarding using tanks of different sizes. One argument is that with different sizes and with independent control of each tank it is less likely that the resonance peaks of the individual tanks are at the same frequency (McMillan, 1984, p. 208). This may have some merit, although one would expect that retuning the controllers would be simpler. There are also recommendations about having the small tank towards the end (McMillan, 1984, p. 208), but at least from a linear point of view the order makes no difference.

7. This example was motivated by the thesis of Walsh (1993), who analyzed controllability of waste water systems using an open-loop step response analysis. Walsh compared numrically (p. 150) estimates of the achievable control using an *open-loop* step response analysis, with the actual closed-loop step responses using PI control on each tank. The discrepancy was quite large, especially for large n. However, it turns out that the results compare very well with the values obtained from our frequency-domain analysis.

8. It is instructive to study in more detail the difference between a step response and frequency domain controllability analysis for the case with n tanks in series. Let us follow Walsh (1993) and use the "disturbance attenuation" as a basis of comparison. Let d denote the concentration disturbance entering the first tank (after mixing the two feed streams), and let y denote the concentration in the last tank. Then disturbance attenuation is defined as:

$$\delta_a = |y(t)|_{max}/|d| \tag{25}$$

where $|d|$ is the magnitude of the concentration disturbance and $|y(t)|_{max}$ is the largest effect this disturbance has on the product concentration.

Let us first consider a *frequency domain* analysis where we assume $d(t) = \sin \omega t$. The disturbance attenuation depends on the frequency ω, and we want to find the "worst" disturbance attenuation. For n tanks with feedback control the attenuation is given by $S(s)h_n(s)$ where $S(s)$ is the sensitivity function and $h_n(s)$ is given in (22). It is possible to make the sensitivity function small at low frequencies and thus achieve good disturbance attenuation here. However, with a delay θ in the feedback loop we will have $|S(j\omega_\theta)| \approx 1$. Thus, the disturbance attenuation at frequency $\omega_\theta = 1/\theta$ is approximately $|h(j\omega_\theta)|$, and taking this as the worst value we get:

$$\delta_a = |h_n(j\omega_\theta)| \approx \left(\frac{\theta}{\tau_h/n}\right)^n, \tag{26}$$

where the approximation applies for $\tau_h \gg \theta$, that is, for δ_a small. This value compares very well with numerical results from closed-loop step responses with PI-control on each tank given by Walsh (1993, p. 150).

Let us now consider an *open-loop step response* analysis. For n identical tanks in series the time response to a step disturbance $d(t) = 1$ is given by [e.g. Walsh (1993) p. 94]:

$$y(t) = 1 - e^{-t/\tau_h} \sum_{i=0}^{n-1} \left(\frac{t}{\tau_h/n}\right)^i i! \tag{27}$$

In the ideal case with a perfect (and unrealizable) controller which immediately detects the disturbance and takes the proper action, $y(t)$ will reach its maximum value at time $t = \theta$ (at the delay), and we have (Walsh, 1993, p. 94):

$$\delta_a = y(\theta) \approx \frac{1}{n!}\left(\frac{\theta}{\tau_h/n}\right)^n. \tag{28}$$

Recall that the expression for δ_a in (26) compares very well with the numerical closed-loop step responses using PI control (Walsh, 1993). Thus, by comparing δ_a in (26) and (28) we see that the open-loop step response analysis is optimistic by a factor $n!$. For $n = 1$ the results of the open-loop step response analysis and the frequency domain analysis are the same, but the step response analysis is optimistic for higher order systems. The main reason for the discrepancy is that the ideal controller needed to perfectly reject the step disturbance cannot be realized by a PI controller (in fact, it is not realizable with any real controller). On the other hand, a real feedback control system will have a resonance frequncy around ω_θ, and a frequency analysis based on considering the behavior at this frequency will yield good predictions of the closed-loop step response. In conclusion, our frequency-domain controllability analysis compares favorably with the results of a closed-loop step response and yields results superior to that of an open-loop step response cotrollability analysis.

5. DISCUSSION

The controllability analysis in this paper is based on a frequency-domain definition of performance, and one may question how applicable it is. Although we have already in the example made comparisons between the frequency domain and step responses, a discussion on the usefulness of the frequency domain seems in order.

First, it should be noted that the interpretation of the frequency domain analysis employed in this paper is in fact in the time domain: At each frequency $g(j\omega)$ yields information about the response to a sinusoidal input. One may also argue that sinusoidal disturbances are quite common in practice, for example, there may be slow sinusoids with period 24 h due to outdoor temperature changes, or there may be fast sinusoids caused by poorly tuned controllers at other places in the system.

Another justification for the frequency domain is that any piecewise smooth periodic function can be written as a sum of sinusoids in a Fourier series. Also non-periodic functions can be included by use of the Fourier integral transformation. This illustrates that sinusoidal signals form a broad class of signals, but one should probably be careful about mixing this "frequency-contents" interpretation with the "frequency-by-frequency sinusoidal response" interpretation used in this paper.

Another justification for using the frequency domain for performance follows from the fact that the H_∞-norm in the frequency domain is equal to the induced 2-norm in the time domain. To be more specific consider the closed-loop disturbance response $y = Sg_d d$, and assume that the d and y have been scaled such that for a sinusoidal disturbance, $d = \sin \omega t$, the performance requirement is that resulting sinusoidal output should be less than one in magnitude, $|y(t)| < 1$. This is equivalent to requiring:

$$|Sg_d(j\omega)| < 1, \forall \omega. \quad (29)$$

Since the H_∞-norm is defined as the peak value as a function of frequency, condition (29) is equvalent to $\|Sg_d\|_\infty < 1$. As stated above, it is a fact from functional analysis that the H_∞-norm is also equal to the induced 2-norm in the time domain. Thus, if (29) is satisfied then it follows for any disturbance $d(t)$ with finite energy that:

$$\frac{\|y(t)\|_2}{\|d(t)\|_2} < 1, \quad (30)$$

where the 2-norm of a signal $u(t)$ is defined as:

$$\|u(t)\|_2 = \sqrt{\int_{-\infty}^{\infty} |u(t)|^2 \, dt}. \quad (31)$$

That is, the "energy" of the output signal is always less than that of the disturbance.

In summary, our frequency domain performance requirement has strong implications for the time domain, both in terms of the magnitude of sinusoids as well as for the energy (2-norm) of arbitrary time signals.

6. CONCLUSION

The paper has presented a controllability analysis for scalar systems using the frequency domain applicable to both feedback and feedforward control. The analysis may be used to answer whether or not a given plant is controllable, and thus extends beyond the traditional use of "controllability indicators". The method has been applied to a pH neutralization process, and it is found that more or less heuristic design rules given in the literature follow directly. The key steps in the analysis are to consider disturbances and to scale the variables properly.

The tools presente in this paper may also be used to study the effectiveness of adding extra manipulated inputs or extra measurements (cascade control). It may also be generalized to multivariable plants where directionality becomes a further crucial consideration. Some results are given in Wolff *et al.* (1992) and Skogestad and Wolff (1992). A direct generalization to decentralized control of multivariable plants is given by Hovd and Skogestad (1992).

Acknowledgement—Manfred Morari (1983) was the first to consider a rigorous approach to controllability analysis by making use of the concept of "perfect control". He also directed me to the paper of Ziegler and Nichols (1943) who first introduced the term controllability in the control literature.

REFERENCES

Balchen J. G. and K. Mumme, *Process Control, Structures and Applications.* Van Nostrand Reinhold, New York (1988).

Hovd M. and S. Skogestad, Simple frequency domain tools for control structure analysis, structure selection and design. *Automatica* **28**,989–996 (1992).

Kwakernaak H. and R. Sivan, *Linear Optimal Control Systems.* Wiley, New York (1972).

McMillan G. K., *pH Control.* Instrument Society of America (1984).

Morari M., Design of resilient processing plants III. A general framework for the assessment of dynamic resilience. *Chem. Engng Sci.* **38**, 1881–1891 (1983).

Rosenbrock H. H., *State-Space and Multivariable Theory.* Nelson, London (1970).

Seborg D. E., T. F. Edgar and D. A. Mellichamp, *Process Dynamics and Control.* Wiley, New York (1989).

Skogestad S. and E. A. Wolff, Controllability measures for disturbance rejection. Preprints *IFAC Workshop on Interactions Between Process Design and Process Control,* London (Edited by J. D. Perkins), pp. 23–29. Pergamon Press (1992).

Walsh S., Integrated design of chemical waste water treatment systems. Ph.D. Thesis, Imperial College, U.K. (1993).

Wolff E. A. S. Skogestad, M. Hovd and K. W. Mathisen, A procedure for controllability analysis. Preprints *IFAC Workshop on Interactions Between Process Design and Process Control,* London (Edited by J. D. Perkins), pp. 127–132. Pergamon Press (1992).

Youla D. C., J. J. Bongiorno and C. N. Lu, Single-loop feedback stabilization of linear multivariable dynamical plants. *Automatica* **10**, 159–173 (1974).

Ziegler J. G. and N. B. Nichols, Process lags in automatic control circuits. *Trans. ASME* **65**, 433–444 (1943).

APPENDIX

Neutralization Model

Derivation of model: consider Fig. 4. Let c_H (mol/l) and c_{OH} (mol/l) denote the concentration of H^+ and OH^--ions, respectively. Material balances for these two species yield:

$$\frac{d}{dt}(Vc_H) = q_A c_{H,A} + q_B c_{H,B} - qc_H + rV,$$

$$\frac{d}{dt}(Vc_{OH}) = q_A c_{OH,A} + q_B c_{H,B} - qc_{OH} + rV,$$

where r(mol/s, m^3) is the rate for the reaction $H_2O = H^+ + OH^-$ which for completely dissociated ("strong") acids and bases is the only reaction in which H^+ and OH^- participate. We may eliminate r from the equations by taking the difference to get a differential equation in terms of the excess of acid, $c = c_H - c_{OH}$:

$$\frac{d}{dt}(Vc) = q_A c_A + q_B c_B - qc.$$

This is the material balance for mixing tank without reaction. The reason is that the quantity $c = c_H - c_{OH}$ is not affected (invariant) by the reaction. Note that c will take on negative values when pH is above 7.

We are not interested in variations in the feed concentrations, c_A and c_B, so they are assumed constant. Linearization and Laplace transformation yields:

$$c(s) = \frac{1}{1+\tau s}\left[\frac{c_A^* - c^*}{q^*} q_A(s) + \frac{c_B^* - c^*}{q^*} q_B(s)\right],$$

where $\tau = V/q^*$ is the residence time and * is used to denote steady-state values.To derive this we have made use of the total material balance $dV/dt = q_A + q_B - q$ (alternatively one may assume V is constant but this is not strictly necessary) and the corresponding steady-state balance $c_A^* + c_B^* = q^*$. We now introduce the following scaled variables:

$$y(s) = \frac{c(s)}{c_{max}}; \quad d(s) = \frac{q_A(s)}{q_{Amax}}; \quad u(s) = \frac{q_B(s)}{q_{Bmax}}$$

and get

$$y(s) = \frac{1}{\tau s + 1}\left(\underbrace{\frac{c_A^* - c^*}{c_{max}} \cdot \frac{\mathrm{Amax}}{q^*}}_{k_d} d(s) + \underbrace{\frac{c_B^* - c^*}{c_{max}} \cdot \frac{q_{Bmax}}{q^*}}_{k} u(s)\right).$$

We use the following numbers: $V = 10\,m^3$, $q_A^* = q_B^* = 0.005\,m^3/s$, $q^* = 0.01\,m^3/s$, $c_{H,A}^* = 10$ mol/l (corresponding to pH $= -1$ and $c_A = 10 - 10^{-15} \approx 10$ mol/l, $c_{OH,B}^* = 10$ mol/l (corresponding to pH $= 15$ and $c_B^* = 10^{-15} - 10 \approx -10$ mol/l), $c^* = 0$ mol/l (corresponding to pH $= 7$), $c_{max} = 10^{-6} - 10^{-8} \approx 10^{-6}$ mol/l (i.e. pH $= 7 \pm 1$), and $q_{Amax} = q_A^*/2 = 0.0025\,m^3/s$, $q_{Bmax} = q_B^* = 0.005\,m^3/s$. Note from the latter that the largest disturbance is $\pm 50\%$ of q_A^*, while the largest input is $\pm 100\%$ of q_B^*. With these values we get $\tau = 1000$ s, $k_d = 2.5 \times 10^6$ and $k = -5 \times 10^6$.