Autotune identification for systems with right-half-plane poles and zeros

Shih-Haur Shen[†], Hong-Den Yu[‡], Cheng-Ching Yu*

Department of Chemical Engineering, National Taiwan University of Science and Technology, Taipei 106, Taiwan

Abstract

Autotuning based on relay feedback tests have received a great deal of attention recently. The relay feedback tests fail to generate sustained oscillation in some cases. For systems with RHP poles and/or zeros, incorrect autotune identification procedure may give erroneous information and unstable limit cycle. Frequency domain analyses provide a criteria for prediction of the existence of stable limit cycle. Furthermore, a procedure is given to ensure the success of a relay feedback test. The analyses and procedure are illustrated through open loop stable and unstable systems. Results show that, except for some rare occasion, successful relay feedback experiments can be obtained following the proposed procedure. © 1999 Elsevier Science Ltd. All rights reserved.

Keywords: Autotuning; Limit cycle; Relay feedback; RHP pole and zero

Nomenclature

a	amplitude of continuous cycling
a	intersection of $-1/N$ and $G(j\omega)$ loci
a_-	region right before the point 'a'
a_+	region right after the point 'a'
BR	boilup ratio
D	distillate flow rate
$G(s)$	transfer function
h	height of relay
Im(·)	imaginary part of (·)
K_c	controller gain
K_u	ultimate gain
n	number of clockwise encirclements
N	describing function
p	number of open-loop unstable poles
P_u	ultimate period
R	reflux flow rate
Re(·)	real part of (·)
RHP	right half plane
u	system input
V	vapor boilup
x_B	bottoms composition
x_D	distillate composition
y	system output
y^{set}	set point of y

z	number of unstable poles for closed-loop system
ZN	Ziegler–Nichols tuning

Greek symbols

ω	frequency
ω_u	*ultimate frequency*
τ_I	*integral time for PI controller (min)*

1. Introduction

Autotuning of PID controllers based on relay feedback tests receives a great deal of attention recently [1–3]. It identifies the important dynamic information, ultimate gain and ultimate frequency, in a straightforward manner. The success of this type of autotuners lies on the fact that it is simple and reliable. The appealing feature of the relay feedback autotuning has lead to a number of commercial autotuners [2] and industrial applications [4]. Modifications [5–7] and extensions [8–10] of the Åström–Hägglund autotuner have been made.

Despite the success of relay feedback tests in most applications, a few cases of unsuccessful applications are also reported. Papastathopoulou and Luyben [11] show that a relay feedback test cannot generate sustained oscillation on a D–BR (distillate–boilup ratio) controlled C_3 splitter. Another example is that the relay feedback test fails to generate a stable limit cycle for the incorrectly paired WB [12] column [10]. Obviously, the relay feedback test cannot generate sustained oscillation in *all* systems. The failure can come from lack of process

* Corresponding author. Fax: +886-2-2737-6644; e-mail: ccyu@ch.ntust.edu.tw

† Now with CMP division, Applied Materials Taiwan Ltd.

‡ Now with CVD division, Applied Materials Taiwan Ltd.

understanding, incorrect knowledge about process system, system inherent property or, even worse, ignoring known process information.

The objective of this work is to investigate under what conditions the relay feedback experiment may fail to generate sustained oscillation. In particular, systems with right-half-plane (RHP) poles and zeros are studied. This paper is organized as follows. Section 2 describes the problems and a number of systems are analyzed according to the pole and zero location of the process transfer functions. Frequency domain analyses for the prediction of a stable limit cycle are given in Section 3. From the analyses, a workable procedure for relay feedback identification is given in Section 4 followed by the conclusion.

2. Problem description

The basic idea of Åström–Hägglund autotuner is that an ideal (on–off) relay is placed in the feedback loop (Fig. 1A) to generate a limit cycle. A typical relay feedback experiment is described next. Considering a system with a positive steady-state gain. Initially, a positive change in the manipulated input $u(u = h)$ is made (Fig. 1B) and as soon as the output y becomes higher than its set point ($e < 0$), a negative change in $u(u = -h)$ is followed. Generally, a limit cycle is generated by repeating this procedure and the ultimate gain ($K_u = 4h/\pi a$) and ultimate period (P_u) can be read off directly
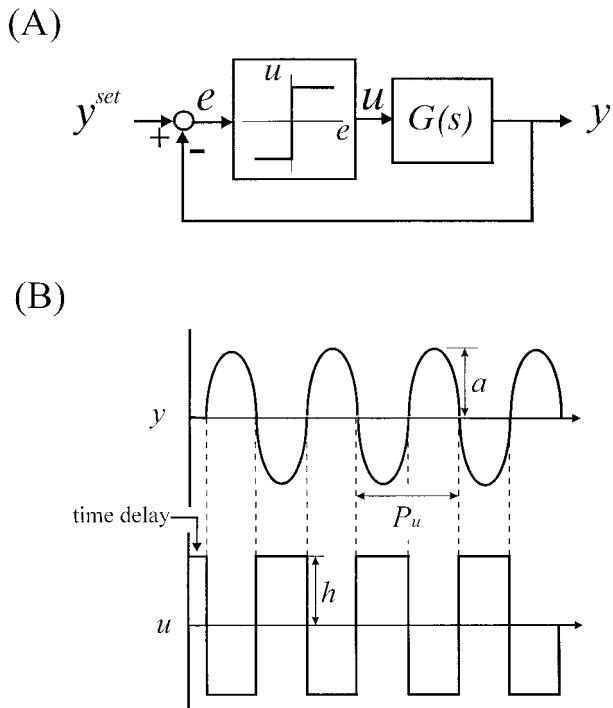
from the input–output responses (Fig. 1). Obviously, the success of a relay feedback test depends on the existence of a limit cycle. In most literature examples, sustained oscillation is observed using a relay feedback test. However, two literature examples indicate apparent failure of the relay feedback tests.

2.1. Failed examples

For a D–BR controlled C_3 splitter, Papastathopoulou and Luyben [11] reveal that the relay feedback test cannot generate sustained oscillation for the x_D–BR loop (a relay placed between x_D and BR). Consider a 190-tray C_3 splitter [11]. Table 1 gives the steady-state operating condition. Initially, a 5% increase in BR (from 26.7 to 28.1) is made. Once a decrease in x_D is observed (Fig. 2), the relay is switched to a negative position ($BR = 25.3$) until x_D crosses the set point. Fig. 2 shows clearly that the relay feedback test fails to generate a limit cycle.

A second example is the incorrectly paired WB column [12] as shown by Shen and Yu [10]. This pairing gives a negative relative gain which implies the sign of the closed-loop gain is opposite to that of the open-loop one.

$$\begin{pmatrix} x_B \\ x_D \end{pmatrix} = \begin{pmatrix} \frac{6.6e^{-7s}}{10.9s+1} & \frac{-19.4e^{-3s}}{21s+1} \\ \frac{12.8e^{-s}}{16.7s+1} & \frac{-18.9e^{-3s}}{21s+1} \end{pmatrix} \begin{pmatrix} R \\ V \end{pmatrix} \qquad (1)$$

A sequential tuning procedure is carried out form the x_B–R loop to the x_D–V loop (Fig. 3). Again, the relay feedback test fails to generate a limit cycle on the x_D–V loop as shown in Fig. 3 ($t \geq 200$ min).

These two seemingly different examples show a common characteristic: the loop to be tested shows the inverse step response characteristic. Therefore, a systematic analysis is carried out for systems with different characteristics and process understandings.

(A)

(B)

Fig. 1. (A) Block diagram for relay feedback system and (B) relay feedback experiment for a system with positive steady-state gain.

Table 1
Steady-state operating condition for C_3 splitter

Parameter	Values
Number of trays	190
Feed tray	61
Relative volatility	1.12–1.24[a]
Column pressure (atm)	7.5
Feed flow rate (kg mol/min)	16.67
Distillation flow rate (kg mol/min)	11.71
Bottoms flow rate (kg mol/min)	4.96
Reflux ratio	10.3
Boilup ratio	26.743
Feed composition (mole fraction)	0.7
Distillation composition (mole fraction)	0.9966
Bottoms composition (mole fraction)	0.0002

[a] $\alpha = 1.23644 - 0.076695x - 0.036147x^2$, where $x =$ mole fraction of light component on each tray.

2.2. System analyses

The systems studied are classified according to the pole and zero locations of the transfer functions. The level of process understanding depends on the knowledge (correct, incorrect or unknown) about the sign of the steady-state gain of the system.

2.2.1. Open-loop stable systems

2.2.1.1. Systems without right-half-plane (RHP) zeros. Obviously, this is the typical transfer function for chemical process. Again, consider the transfer function between x_D and R for the WB column.

Example 1. First order plus dead time system [12].

$$G(s) = \frac{x_D}{R} = \frac{12.8e^{-s}}{16.8s + 1}$$

Assume that we know the sign of $G(s)$, $G(0) = 12.8$ (> 0). It may sound strange that the 'sign' of the process is unknown after the instrumentation has been set up. A more likely situation is that process operators simply ignore this information when performing a relay feedback test.) One is expecting an increase in x_D when a positive change in R is made as shown is Fig. 4A. A sustained oscillation is observed when we continue the
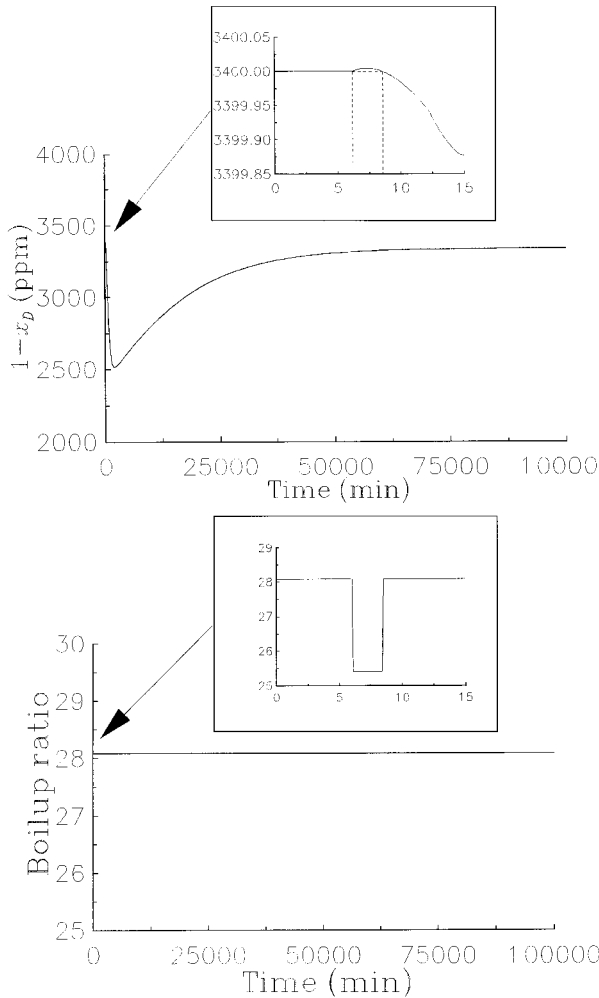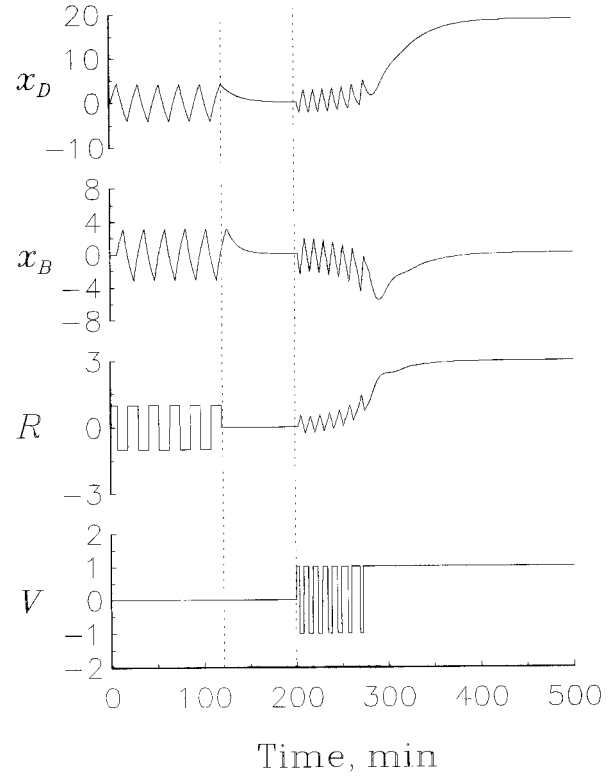

Fig. 3. Multivariable autotuning procedure for WB column with incorrect pairing (starting from x_B–R loop ($t = 0$–120 min)) followed by x_D–V loop ($t > 200$ min)).
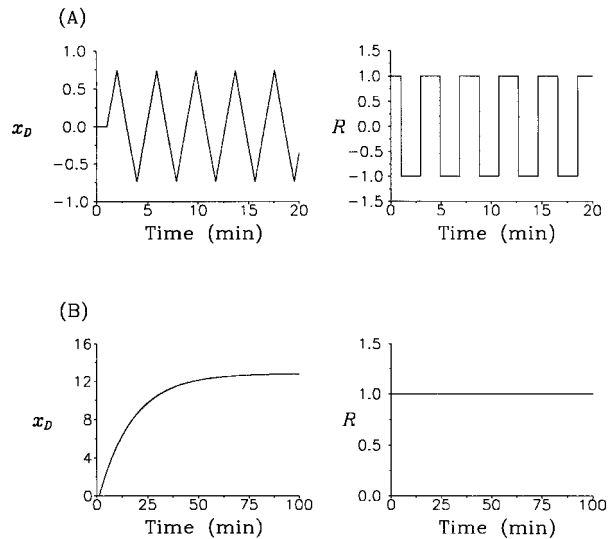

Fig. 4. Relay-feedback test for the first order plus dead time example (Example 1) with: (A) correct prior knowledge (positive sign) and (B) incorrect prior knowledge (negative sign).

Fig. 2. Failed relay feedback test for C_3 splitter.

relay feedback test. However, if we have a misconception about the sign of the process gain and the sign is thought to be 'negative'. Under this circumstance, if an initial 'positive' change in R is made, one is expecting x_D becoming negative. Clearly, this will not happen since the sign of the process is 'positive' (Fig. 4B). Therefore, the relay never switches and the experiment fails. This simple example indicates that the relay feedback test may fail if one has *incorrect* knowledge about the system.

One way to overcome this possible failure is: let the relay switches when the output starts moving (ignoring the knowledge about of sign of the process). This approach seems to work is this example as shown in Fig. 4A.   ☐

2.2.1.2. Systems with RHP zeros. Consider the following system with *one* RHP zero.

    Example 2. System with one RHP zero [13].

$$G(s) = \frac{y}{u} = \frac{(-3s + 1)e^{-0.6s}}{(5s + 1)(s + 1)}$$

This system shows an inverse response for a step change in u. Following the simple minded procedure in Section 2.2.1.1 (ignoring the knowledge about the 'sign'), the relay switches as soon as the output starts to move and the responses is shown in Fig. 5B. The results show that the relay feedback test does not generate a stable limit cycle (Fig. 5B). More importantly, this experiment looks successful for the first 20 min (the operator or the computer is very likely to stop the experiment at $t = 20$ min). One may go on to design the controller by ignoring any information about the sign, then one has (taking the averaging height between the 2nd and 4th cycling): $K_u = 4h/\pi a = 3.90$ and $P_u = 2.44$ This gives the Ziegler–Nichols (ZN) tuning of $K_c = 1.77$ and $\tau_I = 2.0$ for a PI controller. Unfortunately, this set of tuning constants will produce an unstable closed loop response, since the true ultimate gain and period are $K_u = 1.40$ and $P_u = 13.5$.

On the other hand, if we utilize the knowledge about the sign of the system, a positive sign, one is expecting an increase in y for a positive change in u and the relay switches accordingly (Fig. 5A). This results in an successful relay feedback test as shown in Fig. 5A. Again, the ZN setting becomes: $K_c = 0.636$ and $\tau_I = 11.25$. This gives a stable set point responses.   ☐

Example 2 explains the reason why the relay feedback tests fail in *D–BR* controlled C_3 splitter and the incorrectly paired WB column examples. From the system point of view, the RHP zero misleads the movement of the relay. Obviously, this can be overcome by utilizing the always available process information: the 'sign' of the process gain.

This principle can also be applied to systems with multiple RHP zeros.

    Example 3. Systems with two RHP zeros.

$$G(s) = \frac{(-s + 1)^2 e^{-0.1s}}{(0.8s + 1)^3}$$

The unit step response of this system is given in Fig. 6. It shows that we have three possible locations (a, b and c) for the relay to switch. If one recognizes that the sign of the system is 'positive', then the points a and c in Fig. 6 are two possible switching locations. Fig. 7A and
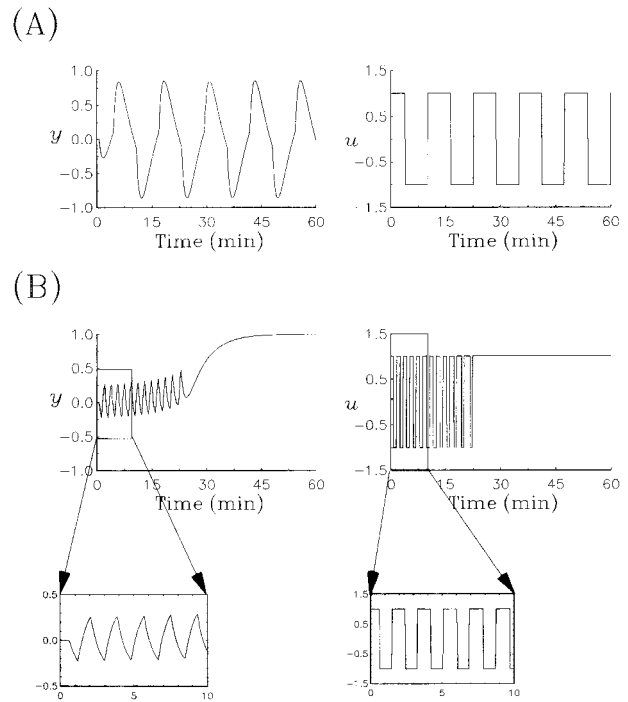

Fig. 5. Relay feedback tests for system with RHP zero (Example 2) with: (A) correct prior knowledge (positive sign) and (B) no prior knowledge (sign unknown).
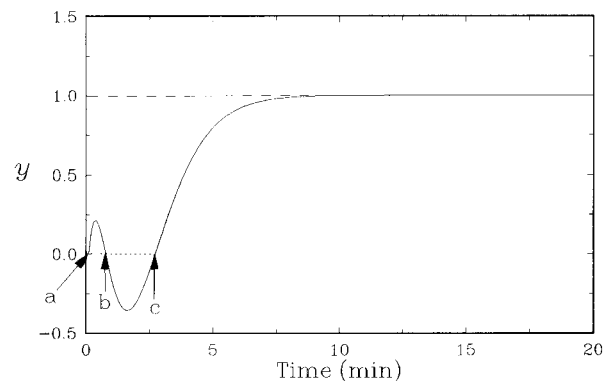

Fig. 6. Unit step set point change for system with two RHP zeros (Example 3).
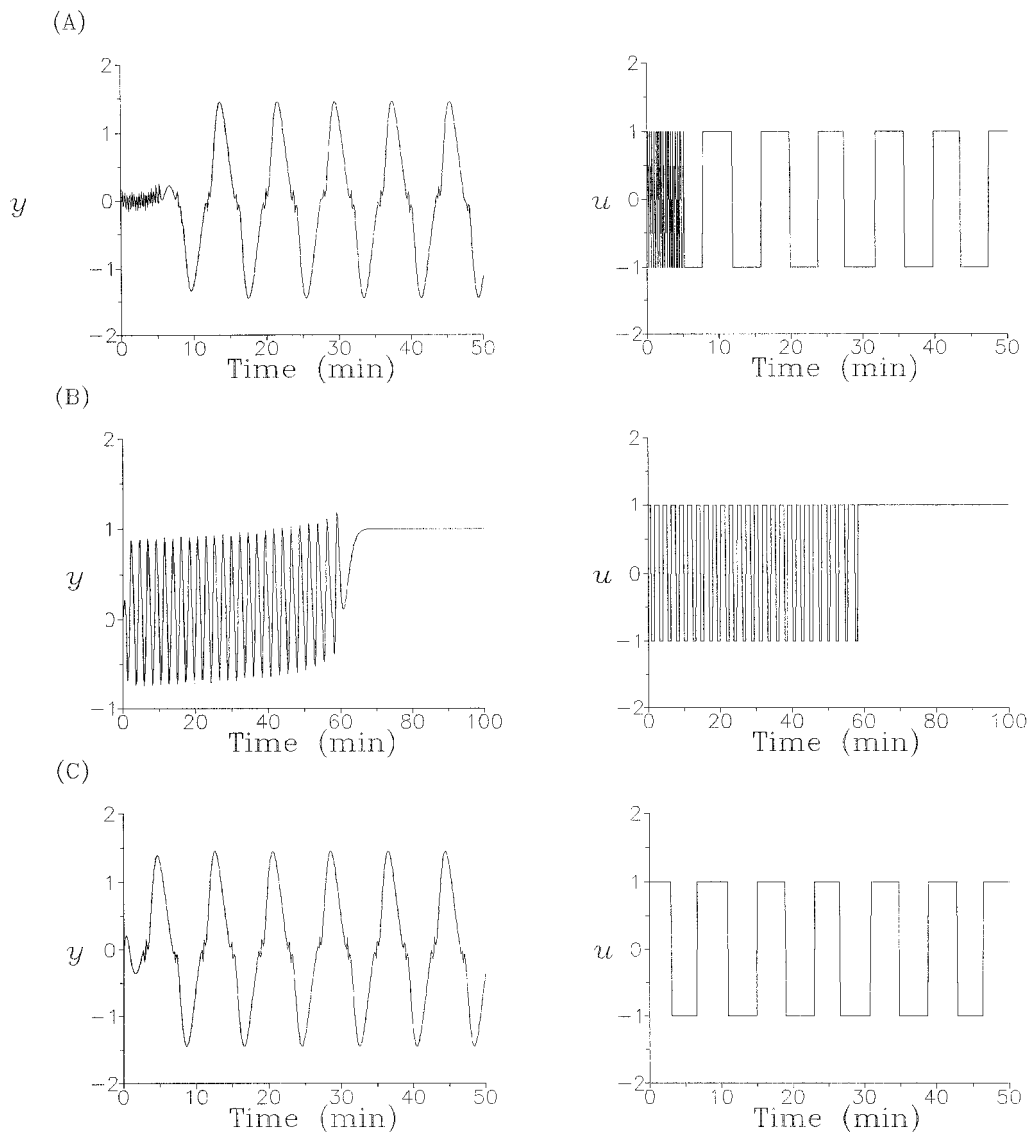
Fig. 7. Relay feedback tests for Example 3 when the relay switched at: (A) point a, (B) point b and (C) point c in Fig. 1.

C shows that either switching locations give successful relay feedback experiments and the switching point b results in a failed experiment. Example 3, again, indicates that successful relay feedback tests can be achieved by utilizing the knowledge about the sign of the system. System with more then two RHP zeros are discussed in Yu [14]. □

2.2.2. Open-loop unstable systems

Open-loop unstable systems are often encountered in exothermic chemical reactors [15]. Little is said about the applicability of the relay feedback to open-loop unstable systems. Systems with one RHP poles are classified according to the shape of the Nyquist plot. The first one is the case where the Nyquist plot starts from the negative real axis and moves toward the third quadrant. That is:

$$\lim_{\omega \to 0} \frac{d\text{Im}(G(j\omega))}{d\omega} < 0$$

where Im(·) stands for the imaginary part of a complex number. This is the typical characteristic for open-loop unstable chemical reactor systems.

Example 4. Open-loop unstable system with $\lim_{\omega \to 0}(d\text{Im}(G)/d\omega) < 0$.

$$G(s) = \frac{(3s+1)e^{-s}}{(3s-1)(4s+1)}$$

The solid line in Fig. 8 shows the Nyquist contour of Example 4 as ω goes from zero to infinity. Fig. 9 shows a relay feedback test gives sustained oscillation regardless

of the knowledge about the process gain (positive or negative sign). However, this is not always the case for open-loop unstable systems.   □

Example 5. Open-loop unstable system with $\lim_{\omega \to 0}(d\text{Im}(G)/d\omega) > 0$
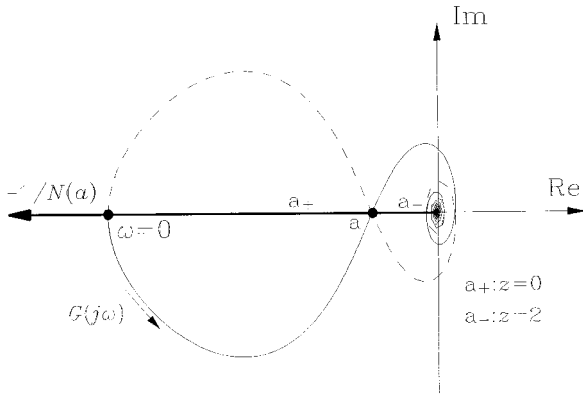
$$G(s) = \frac{(s+1)e^{-s}}{(2s-1)(10s+1)}$$

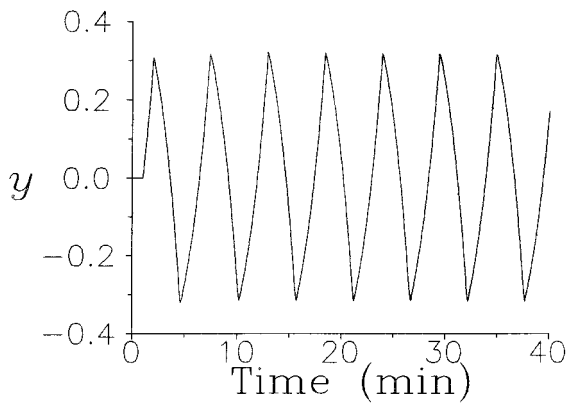Fig. 8. Nyquist plot and $-1/N$ loci for open-loop unstable system of Example 4.

Fig. 9. Relay feedback test for open-loop unstable system in Example 4.

The Nyquist plot of Example 5 shows a different characteristic where the $G(j\omega)$ moves toward the second quadrant initially and further crosses the imaginary axis into the first quadrant (solid line in Fig. 10). Furthermore, it is not possible to generate a stable limit cycle regardless of how the relay switches (Fig. 11). Example 5 clearly indicates the relay feedback test may fail to generate continuous cycling for open-loop unstable systems and the criterion, $\lim_{\omega \to 0}(d\text{Im}(G)/d\omega) > 0$, provides the necessary condition for this failure. However, this is not a sufficient condition. Let us consider the following example.   □

Example 6. Open-loop unstable system with $\lim_{\omega \to 0}(d\text{Im}(G)/d\omega) > 0$.

$$G(s) = \frac{(3s+1)e^{-s}}{(3s-1)(10s+1)}$$

Fig. 12 shows the Nyquist plot of Example 6. Despite the fact that it (solid line in Fig. 12) meets the criterion for the failure, a relay feedback test can generate a stable limit cycle.   □

From the on-going analyses, heuristics can be found for the success/failure of a relay feedback provided with different levels of process understanding. For open-loop stable systems, the relay feedback can generate a stable limit cycle if the knowledge of the 'sign' of the process gain is utilized in the experiment (Table 2). However,
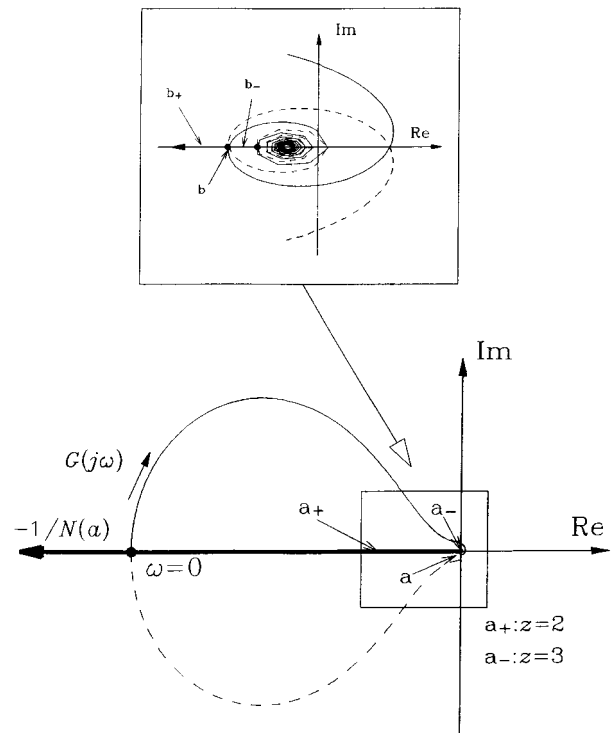

Fig. 10. Nyquist plot and $-1/N$ loci for open-loop unstable system with $\lim_{\omega \to 0}(d\text{Im}(G(j\omega))/d\omega) > 0$ (Example 5).
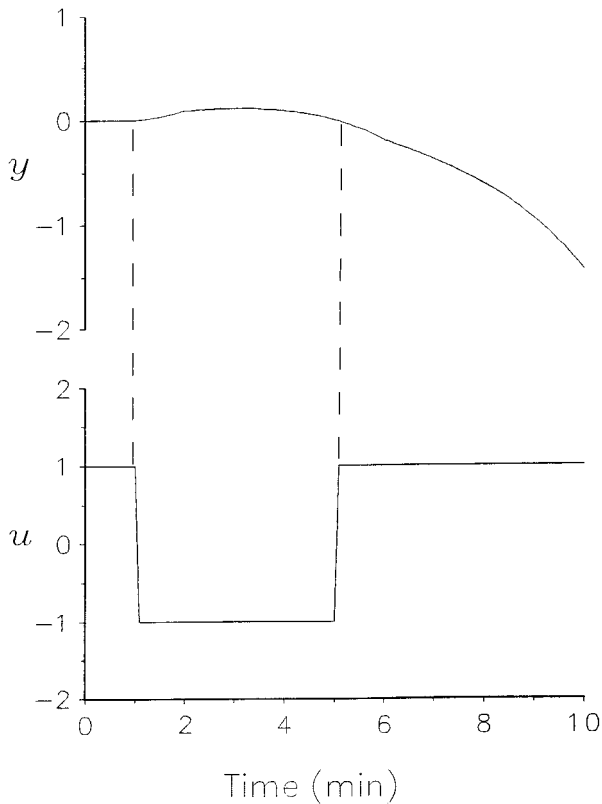
Fig. 11. Relay feedback test for open-loop unstable system in Example 5.

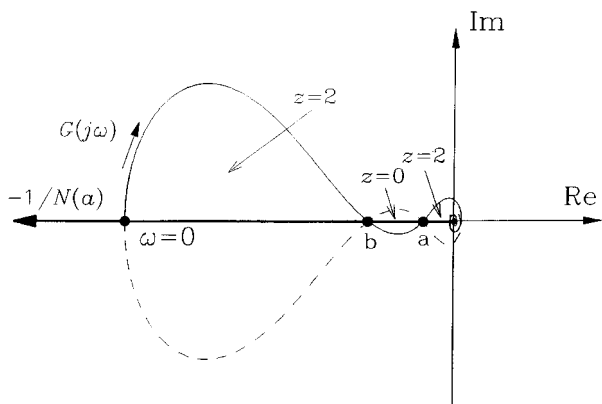Fig. 12. Nyquist plot and $-1/N$ loci for open-loop unstable system with $\lim_{\omega \to 0}(\mathrm{dIm}(G(j\omega))/\mathrm{d}\omega) > 0$ (Example 6).

for open-loop unstable system, the results are not quite as clear. For some systems, it is simply not possible to generate a stable limit cycle (e.g. Example 5) for an ideal relay. Table 2 summarizes these heuristics. A simple rule of thumb immediately follows: utilize the knowledge of process sign throughout the relay feedback experiment.

3. Frequency domain analyses

In addition to these heuristics, the condition for limit cycling can be analyzed using the describing function $N(a)$ (of an ideal relay). If the system exhibits a limit cycle with an amplitude $a(y(t) = a\sin \omega t)$, then we have [1]

$$1 + NG(j\omega) = 0$$

or

$$G(j\omega) = -\frac{1}{N}$$

Therefore, a simple way to investigate the existence of a limit cycle is to plot the loci of $G(j\omega)$ and $-1/N$ and to look for intersections. For example, the point 'a' is the intersection of the $-1/N$ and $G(j\omega)$ contours in Fig. 8. However, the existence of an intersection of $G(j\omega)$ and $-1/N$ is only the *necessary* condition for a stable limit cycle [1,16–18]. In other words, the intersection only indicates that a stable limit cycle *may* exist.

Fendrich [18], based on the Nyquist stability criterion, proposes a simple way to check the existence of a stable limit cycle. In terms of stability analysis, a limit cycle indicates we have a pair (an even number to be exact) of poles on the imaginary axis. Obviously, this corresponds to the limit of stability. Therefore, the loci of $-1/N(a)$ right before (e.g. region of a_- in Fig. 8) and after (e.g. region of a_+ in Fig. 8) the intersection (e.g. a) should indicate a change in the number of unstable poles. If the system has a pair of unstable poles in the a_- region and no unstable pole in the a_+ region, then a stable limit cycle exists (e.g. Fig. 8). The means the

Table 2
Summary of success/failure for relay feedback test with different degree of prior knowledge

Sign of steady-state gain	Open-loop stable system			Open-loop unstable system	
	No RHP zero	RHP zeros		$\lim_{\omega\to0}\frac{\mathrm{dIm}(G(j\omega))}{\mathrm{d}\omega} < 0$	$\lim_{\omega\to0}\frac{\mathrm{dIm}(G(j\omega))}{\mathrm{d}\omega} > 0$
		Odd	Even		
Known	✔	✔	✔	✔	?
Unknown	✔	×	✔	✔	?

×, Failed relay feedback test; ✔, successful relay feedback test; ?, unknown.

number of unstable poles goes from zero to two as we crosses the limit of stability. Notice that the arrow on the $-1/N(a)$ loci represents an increase in the amplitude a (Fig. 8).

This simple method utilizes the well-known Nyquist stability criterion to check for the existence of a stable limit cycle. Let us re-state the criterion. If n denotes the number of clockwise encirclements made by $G(j\omega)$ for the point of interest and p represents the number of unstable poles of $G(s)$, then the number of unstable poles (z) is:

$$z = n + p$$

This criterion is employed to find the number of unstable poles for points in the regions of a_- and a_+. Then, the rule of Fenrich can be applied directly.

Let us take the open-loop unstable system (Example 4) to illustrate this simple method. First, the system has one unstable pole: $p = 1$. The number of clockwise encirclements made by $G(j\omega)$ is 1 in a_- and -1 in a_+, i.e. $n = 1$ and $n = -1$. Therefore, we have $z = 2$ in a_- is 2 ($n = 2$) and $z = 0$ for a_+. This is exactly what shown in Fig. 8. According to Fenrich rule, we have a stable limit cycle and the relay feedback test in Fig. 9 confirms this. For the unstable system of Example 5, Fig. 11 shows that the number of unstable poles reduces from three to two as the locus of $-1/N(a)$ crosses the intersection 'a'. Since the change in z is an odd number, a stable limit cycle, or a successful relay feedback test, is not expected (Fig. 11). Fig. 12 also indicates that a stable limit cycle exists which corresponds to point the 'a' in Fig. 12.

The simple method presented here is very effective to identify cases where the relay feedback tests fail. Yu [14] gives a detailed discussion on other systems.

4. Discussion

Despite the fact that the frequency domain analysis can be used to verify the existence of a stable limit cycle, we generally do not have knowledge about the structure (e.g. pole or zero location) of the process system at this phase. However, we do have the information about the sign of process gain (it is hidden in the action of the controller and direction of the valve movement) once the instrumentation is set up. This type of process knowledge need to be utilized during the relay feedback experiment. Therefore, the following procedure is recommended in the relay feedback identification:

1. obtain the knowledge about the sign of the process gain (positive or negative);

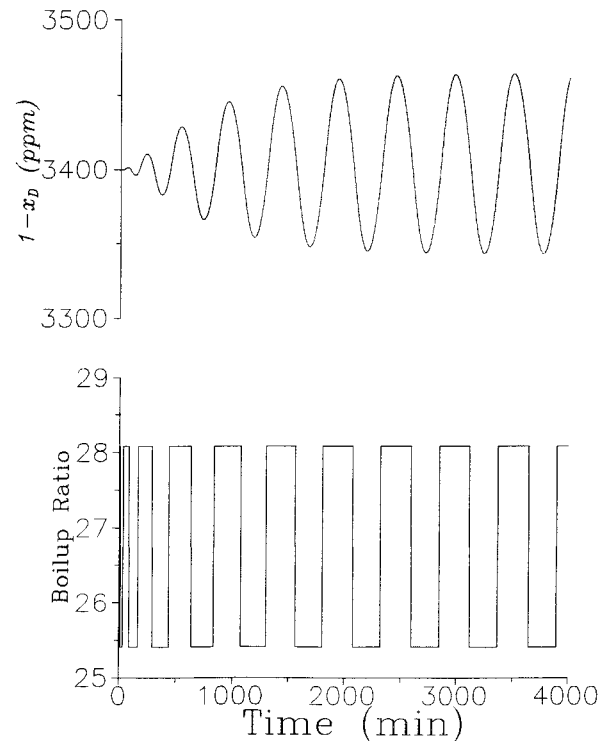2. switch the relay according the sign of the process system.

Fig. 13. Successful relay feedback test for C_3 splitter by employing the knowledge about the sign of process gain.

Using the procedure, we can generate sustained oscillation for most systems except some unusual process systems (e.g. Example 5). Let us take the C_3 splitter example (Table 1) to illustrate this. First, physically we know an increase in boilup ratio (BR) while holding D constant will result in an increase in light component concentration (x_D). That is the steady-state gain between x_D and BR is positive. An increase in BR is made initially and then x_D decreases for a while (as the result of inverse response) followed by an increase and the relay switches accordingly. Repeating this procedure, a successful relay feedback test is obtained as shown in Fig. 13. The results show that K_u and ω_u for the C_3 splitter are 91 and 0.0216, respectively.

5. Conclusions

Despite successful applications of relay feedback in system identification, it may fail to produce sustained oscillation in several occasions. In this work, the reasons for unsuccessful autotune identification are pointed out. These include: misconception about the 'sign' of the steady-state gain, ignoring the information about the 'sign' in the experiment and, for some open-loop unstable systems, the system inherent property. Furthermore, for systems with RHP zero(s), incorrect autotune identification procedure may provide misleading information about K_u and ω_u and, subsequently,

gives erroneous tuning constants and unstable responses. Frequency domain analyses are given for the prediction of the success/failure of a relay feedback test and systems with RHP poles and zeros are studied to illustrate this criterion. Finally, guidelines are given for relay feedback tests.

Acknowledgements

References

[1] K.J. Åström, T. Hägglaund, Automatic tuning of simple regulators with specifications on phase and amplitude margins, Automatica 20 (1984) 645.

[2] K.J. Åström, T. Hägglaund, Industrial adaptive controllers based on frequency response techniques, Automatica 27 (1991) 599.

[3] C.C. Yu, Autotuning of PID controllers: relay feedback approach, London: Springer-Verlag, 1998.

[4] H.S. Papastathopoulou, W.L. Luyben, Tuning controllers on distillation columns with distillate-bottoms structure, Ind Eng Chem Res 29 (1990) 1859.

[5] W. Li, E. Eskinat, W.L. Luyben, An improved autotune identification method, Ind Eng Chem Res 30 (1991) 1530.

[6] R.C. Chiang, S.H. Shen, C.C. Yu, Derivation of transfer function from relay feedback systems, Ind Eng Chem Res 31 (1992) 885.

[7] A. Leva, PID autotuning algorithm based on relay feedback, IEEE Proc D 140 (1993) 328.

[8] J.Y. Lin, C.C. Yu, Automatic tuning and gain scheduling for PH control, Chem Eng Sci 48 (1993) 3150.

[9] W.L. Luyben, E. Eskinat, Non-linear auto-tune identification, Int J Control 59 (1994) 595.

[10] S.H. Shen, C.C. Yu, Use of relay feedback test for automatic tuning of multivariable control systems, AIChE J 40 (1994) 627.

[11] H.S. Papastathopoulou, W.L. Luyben, Potential pitfalls in ratio control schemes, Ind Eng Chem Res 29 (1990) 2044.

[12] R.K. Wood, M.W. Berry, Terminal composition control of binary distillation column, Chem Eng Sci 28 (1973) 1707.

[13] Luyben W.L. Process modeling, simulation and control for chemical engineers. 2nd ed. New York: McGraw-Hill, 1990.

[14] Yu H.D. Potential problems in relay feedback systems. MS thesis, National Taiwan Institute of Technology, Taipei, 1993.

[15] N. Devia, W.L. Luyben, Reactor: size versus stability, Hydrocarbon Proc 6 (1978) 119.

[16] Ogata K. Modern control engineering. 2nd ed. Englewood Cliffs, NJ: Prentice Hall, 1990.

[17] Slotine J.J.E., Li W. Applied nonlinear control. Englewood Cliffs, NJ: Prentice Hall, 1991.

[18] O.R. Fendrich, Describing function and limit cycles, IEEE Trans Autom Contr 37 (1992) 486.