Integrals of Relay Feedback Responses for Extracting Process Information

Jietae Lee and Su Whan Sung
Dept. of Chemical Engineering, Kyungpook National University, Daegu 702-701, Korea

Thomas F. Edgar
Dept. of Chemical Engineering (C0400), The University of Texas, Austin, Austin, TX 78712

DOI 10.1002/aic.11256
Published online July 27, 2007 in Wiley InterScience (www.interscience.wiley.com).

Relay feedback methods have been widely used in tuning proportional-integral-derivative controllers automatically. Previous approaches usually use specific point data such as the oscillation amplitude and period of the relay response. In this article, we propose new identification methods, which use integrals of the relay response instead of the point data. The proposed methods guarantee better accuracy and advantages in obtaining the ultimate information of process as well as parametric models compared with previous approaches, because effects of the high order harmonic terms are suppressed significantly using the integrals of the relay responses. © 2007 American Institute of Chemical Engineers *AIChE J,* 53: 2329–2338, 2007
Keywords: relay feedback, integrals, identification, ultimate information, PID, autotuning

Introduction

Since Astrom and Hagglund[1] introduced the autotuning method, which used the relay feedback test, many variations have been proposed for autotuning of PID controllers.[1–3] Several methods such as a saturation relay,[3] relay with a P control preload,[4] and a two-level relay[5] were introduced to obtain more accurate ultimate information of the process by suppressing the effects of the high order harmonic terms. To obtain a Nyquist point other than the critical point, a relay with hysteresis or a dynamic element such as time delay has been used.[6–9] Recently, a two-channel relay has been proposed to obtain a Nyquist point information corresponding to a given phase angle.[10,11] Methods to reject unknown load disturbances and restore symmetric relay oscillations have been available.[12–14] A biased relay has been used to obtain the process steady state gain as well as the ultimate information from only one relay test.[15] Huang et al.[16] used the integral of the relay transient to obtain the steady state gain of the process.

Many Nyquist points of the process dynamics can be extracted from only one relay experiment by applying the FFT (fast Fourier transformation) technique to the whole transient responses from the start to the final cyclic steady-state part of the relay responses.[17] However, the computations are somewhat complex and the complete transient responses must be stored. For the same purpose, Laplace transformation of a periodic function has been used to obtain many frequency responses from one relay test.[18]

The shape factor has been used to extract a three-parameter model from the cyclic steady state part of the relay response.[19] Several authors derived exact expressions relating the parameters of the FOPTD process to the measured data of the relay response.[20,21] They used the analytic expressions to extract parameters of the FOPTD model. However, the methods based only on the cyclic steady state data cannot provide acceptable robustness for uncertainty such as process/model mismatches and nonlinearity. They may provide poor model parameter estimates such as negative gain when the model structure is different from that of the process.[22] The second-order plus time delay (SOPTD) model can also

Correspondence concerning this article should be addressed to S. W. Sung at suwhansung@knu.ac.kr.

be extracted from the cyclic steady state part of the relay response using analytic equations. However, as in the FOPTD model case, the method is also not robust. A relay experiment with a subsequent P control experiment or another relay feedback test can be used to obtain an SOPTD model robustly.[23,24]

In this research, identification methods to extract more accurate frequency response information and parametric models from a single conventional relay feedback test are proposed. We use various integrals of the original relay feedback responses to enhance identification performances without modifying the relay feedback system. As in the step response method,[6] areas (equivalently, integrals) will have merits over point data and they are investigated here. Since the conventional relay feedback method is used, the proposed method shares its practical and theoretical merits. Integrals of the relay responses let the fundamental frequency term to be dominant compared to the high harmonic terms, resulting in better accuracy in estimating frequency response information and model parameters. Because it is not required to store the whole trajectories and computations are simple, the proposed methods can be incorporated easily in commercial PID controllers.

Conventional Relay Feedback Method

We consider a classical relay feedback system as shown in Figures 1 and 2 to derive the required equations in this research. We start the relay feedback system at a steady state condition. The relay is first kept on until the process output rises up to a given level and then is set to the normal mode of switching at the instant that the process output crosses a given set point. This relay feedback system will produce a stable oscillation as shown in Figure 2. It is notable that we should set the given level in the beginning of the relay feedback to a significantly large value if we want to extract the zero-frequency information of the process. Otherwise, we cannot guarantee acceptable robustness in extracting zero frequency information from the relay responses.

Astrom and Hagglund[1] used this oscillation to extract approximate ultimate information and tune the proportional-integral-derivative (PID) controllers automatically. Let the input and output trajectories be $u_a(t)$ and $y_a(t)$ for the conventional relay feedback system, respectively. At the time t_1, $u_a(t)$ and $y_a(t)$ are assumed to be fully developed (cyclic steady state). Then, it can be represented by the Fourier series as

$$u_a(\tilde{t}) = \frac{4h}{\pi}\left(\sin(\omega\tilde{t}) + \frac{1}{3}\sin(3\omega\tilde{t}) + \frac{1}{5}\sin(5\omega\tilde{t}) + \cdots\right) \quad (1)$$

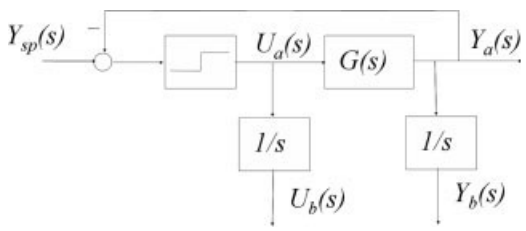

Figure 1. A conventional relay feedback system with integrals of responses.

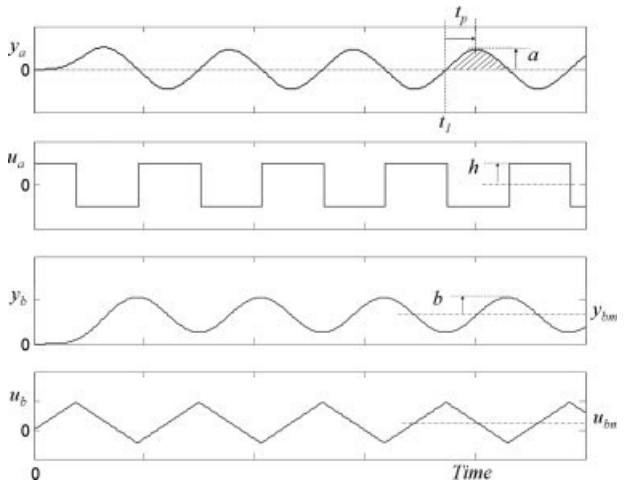

Figure 2. Typical relay feedback responses and their integrals.

where $\tilde{t} = t - t_1$ and h is the relay amplitude. Let p and $\omega = 2\pi/p$ are the period and the frequency of the relay feedback oscillation. Equation 1 is valid for

$$\tilde{t} = t - t_1 \geq 0 \quad (2)$$

The output corresponding to $u_a(\tilde{t})$ is

$$y_a(\tilde{t}) = \frac{4h}{\pi}\left(|G(j\omega)|\sin(\omega\tilde{t} + \angle G(j\omega)) + \frac{1}{3}|G(j3\omega)|\sin(3\omega\tilde{t} + \angle G(j3\omega)) + \cdots\right) \quad (3)$$

where $G(s)$ is the process transfer function. Neglecting the high harmonic terms and assuming $\angle G(j\omega) \approx -\pi$, we obtain the ultimate frequency ω_u, and $u_a(\tilde{t}) \approx 4h\sin(\omega_u\tilde{t})/\pi$, $y_a(\tilde{t}) \approx 4h|G(j\omega_u)|\sin(\omega_u\tilde{t} + \angle G(j\omega_u))/\pi \approx -4h|G(j\omega_u)|\sin(\omega_u\tilde{t})/\pi$. We obtain the following approximate ultimate period P_u and ultimate gain K_{cu} as

$$P_u = p \quad (4)$$

$$K_{cu} = \frac{4h}{\pi a} \quad (5)$$

where a is the measured amplitude of $y_a(t)$. It should be noted that the estimated ultimate information is approximate because we neglect the high harmonic terms. As a result, the ultimate period of Eq. 4 and the ultimate gain of Eq. 5 shows relative errors up to 5 and 18%, respectively, for the first-order plus time delay (FOPTD) process. In this case, the ultimate gain error may not be acceptable.

Proposed Methods for Ultimate Data Estimation

In this research, we use integrals of the process input and output instead of the point data to obtain more accurate process frequency information and parametric process models by suppressing the effects of high harmonic terms.

Proposed Eq. 1

Let $u_b(t)$ and $y_b(t)$ be integrals of the relay responses as

$$u_b(t) = \int_0^t u_a(t)dt \tag{6}$$

$$y_b(t) = \int_0^t y_a(t)dt \tag{7}$$

Then, from Eq. 1, the response $u_b(t)$ after t_1 is

$$u_b(\tilde{t}) = u_{bm} - \frac{4h}{\pi\omega}\left(\cos(\omega\tilde{t}) + \frac{1}{9}\cos(3\omega\tilde{t}) + \frac{1}{25}\cos(5\omega\tilde{t}) + \cdots\right) \tag{8}$$

where u_{bm} is the mean value of $u_b(t)$

$$u_{bm} = \frac{1}{p}\int_{t_1}^{t_1+p} u_b(t)dt \tag{9}$$

From Eq. 3, the response $y_b(t)$ after t_1 can be represented as

$$y_b(\tilde{t}) = y_{bm} - \frac{4h}{\pi\omega}\left(|G(j\omega)|\cos(\omega\tilde{t} + \angle G(j\omega)) + \frac{1}{9}|G(j3\omega)|\cos(3\omega\tilde{t} + \angle G(j3\omega)) + \cdots\right) \tag{10}$$

where y_{bm} is the mean value of $y_b(t)$.

$$y_{bm} = \frac{1}{p}\int_{t_1}^{t_1+p} y_b(t)dt. \tag{11}$$

Figure 2 shows typical plots of these responses.

By neglecting the high harmonic terms and assuming $\angle G(j\omega) = -\pi$ (equivalently, the relay period is the ultimate period), we obtain $u_b(\tilde{t}) - u_{bm} \approx -4h\cos(\omega_u\tilde{t})/\pi\omega_u$ and $y_b(\tilde{t}) - y_{bm} \approx -4h|G(j\omega_u)|\cos(\omega_u\tilde{t})/\pi\omega_u$. Then, we have the following approximate ultimate gain K_{cu}.

$$K_{cu} = \frac{2hp}{\pi^2 b} \tag{12}$$

where b is the measured amplitude of $y_b(\tilde{t}) - y_{bm}$. That is, $b = 4h|G(j\omega_u)|/\pi\omega_u = 2ph|G(j\omega_u)|/\pi^2$. Then, (12) is obtained directly from $1/|G(j\omega_u)| = 2ph/\pi^2 b$. The quantity b is physically the half of the shaded area of $y_a(\tilde{t})$ in Figure 2. The proposed method of Eq. 12 will be superior to Eq. 5 because the ratios of the high harmonic terms to the fundamental frequency term in $y_b(\tilde{t})$ are much smaller than those in $y_a(\tilde{t})$ as shown in Eqs. 1, 3, 8, and 10.

Proposed Eq. 2

From the relay feedback responses in Figure 2, we can construct the following responses as

$$u_c(\tilde{t}) = u_a(\tilde{t}) + \frac{6\pi}{p}(u_b(\tilde{t} + p/4) - u_{bm}) = \frac{16h}{\pi}\sin(\omega\tilde{t}) + \frac{52h}{25\pi}\sin(5\omega\tilde{t}) + \cdots \tag{13}$$

$$y_c(\tilde{t}) = y_a(\tilde{t}) + \frac{6\pi}{p}(y_b(\tilde{t} + p/4) - y_{bm}) \tag{14}$$

We should remark that $u_a(\tilde{t})$ is a rectangular wave, $u_b(\tilde{t})$ is a triangular wave and $u_c(\tilde{t})$ is their combination such that the third harmonic term vanishes. The forcing functions of $u_b(\tilde{t})$ and $u_c(\tilde{t})$ are closer to a sinusoidal wave than $u_a(\tilde{t})$. Hence, $y_b(\tilde{t})$ and $y_c(\tilde{t})$ are closer to the sinusoidal wave than $y_a(\tilde{t})$. Figure 3 shows typical plots of these responses.

Approximating the maximum of $y_c(\tilde{t})$ as $\max(y_a(\tilde{t})) + \frac{6\pi}{p}\max(y_b(\tilde{t}) - y_{bm}) = a + 6\pi b/p$, we have

$$K_{cu} = \frac{16h}{\pi(a + 6\pi b/p)} = \frac{1}{\frac{1}{4}\left(\frac{\pi a}{4h}\right) + \frac{3}{4}\left(\frac{\pi^2 b}{2hp}\right)} \tag{15}$$

Proposed Eq. 3

Consider the following quantity.

$$q_a \equiv \frac{2}{p}\int_{t_1}^{t_1+p} y_a(t)^2 dt = \frac{16h^2}{\pi^2}\left(|G(j\omega)|^2 + \frac{1}{9}|G(j3\omega)|^2 + \cdots\right) \tag{16}$$

The right-hand side of Eq. 16 is derived by applying orthogonality of sine and cosine functions (Appendix A) to Eq. 3. It is remarked that, to compute the above quantity, the trajectory of $y_a(t)$ does not need to be stored. By ignoring high harmonic terms in q_a, we obtain

$$K_{cu} = \frac{4h}{\pi\sqrt{q_a}}. \tag{17}$$

When $y_a(\tilde{t})$ is sinusoidal, $a = \sqrt{q_a}$ and Eqs. 5 and 17 provide the same results.

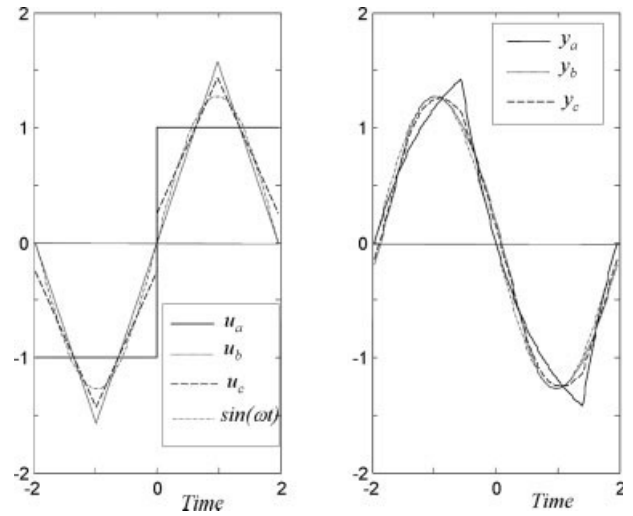

Figure 3. Normalized wave forms of fully developed relay feedback responses.

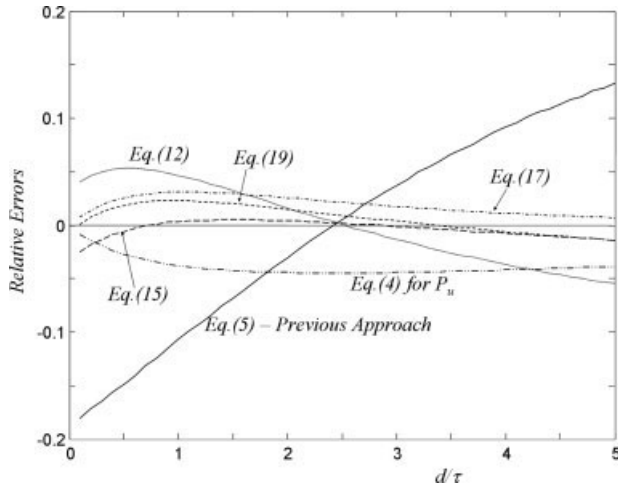

Figure 4. Relative error in estimation of the ultimate period and gain for FOPTD processes ($G(s) = \exp(-ds)/(\tau s+1)$).

Proposed Eq. 4

Consider the following quantity.

$$q_b \equiv \frac{2}{p}\int_{t_1}^{t_1+p} y_b(t)^2 dt - 2{y_{bm}}^2 = \frac{16h^2}{\pi^2\omega^2}\left(|G(j\omega)|^2 + \frac{1}{81}|G(j3\omega)|^2 + \cdots\right) \quad (18)$$

As for Eq. 16, the right-hand side of Eq. 18 is derived by applying orthogonality of sine and cosine functions to Eq. 10. By ignoring high harmonic terms in q_b, we have

$$K_{cu} = \frac{2hp}{\pi^2\sqrt{q_b}} \quad (19)$$

When $y_b(\tilde{t})$ is sinusoidal, $b = \sqrt{q_b}$ and Eqs. 12 and 19 provide the same results.

Figure 4 shows the relative errors in estimated ultimate period and ultimate gains. For FOPTD processes with ratios of time delays to time constants between 0.1 and 5, Eqs. 12, 15, 17, and 19 have relative errors below about 6%. We can see that errors in Eq. 5 for the ultimate gain can be improved considerably.

Proposed Methods for Nyquist Point Data Estimation

In the previous section, we estimate the ultimate gain on the assumption that the period of relay feedback oscillation is the ultimate period. We omit that assumption in this section.

Our goal is to find the following amplitude ratio and the phase lag of the process at the frequency of relay oscillation.

$$G(j\omega) = A_\omega \exp(j(-\pi + \phi_\omega)), \quad \omega = \frac{2\pi}{p} \quad (20)$$

where $A_\omega = |G(j\omega)|$ and $\phi_\omega = \pi + \angle G(j\omega)$. From Eqs. 3, 10, 16, and 18, we can obtain the approximate amplitude ratio at the frequency $\omega = 2\pi/p$ by neglecting high harmonic terms. Among them, we consider

$$A_\omega = \frac{\pi^2\sqrt{q_b}}{2hp}, \quad (21)$$

which is the most accurate equation for an amplitude ratio at the frequency $\omega = 2\pi/p$. When $|G(jn\omega)| < |G(j\omega)|$, $n = 2, 3, \cdots$, its approximation error is

$$|(A_\omega - |G(j\omega)|)/|G(j\omega)|| \le \sqrt{1 + 1/3^4 + 1/5^4 + \cdots} - 1 = \frac{\pi^2}{4\sqrt{6}} - 1 = 0.0073 \quad (22)$$

Now, consider the following quantity

$$q_c = \left[\frac{2}{p}\int_{t_1}^{t_1+p} u_a(t)y_b(t)dt\right] \Big/ \left[\frac{2}{p}\int_{t_1}^{t_1+p} u_b(t)y_b(t)dt - 2u_{bm}y_{bm}\right]$$
$$= \frac{\frac{8ph^2}{\pi^3}\left[\sin(\angle G(j\omega)) + \frac{|G(j3\omega)|}{27|G(j\omega)|}\sin(\angle G(j3\omega)) + \cdots\right]}{\frac{4p^2h^2}{\pi^4}\left[\cos(\angle G(j\omega)) + \frac{|G(j3\omega)|}{81|G(j\omega)|}\cos(\angle G(j3\omega)) + \cdots\right]}$$
$$= \frac{2\pi}{p}\left[\frac{\sin(\phi_\omega)}{\cos(\phi_\omega)} + \frac{|G(j3\omega)|}{|G(j\omega)|} \times \left(-\frac{\sin(\angle G(j3\omega))}{27\cos(\phi_\omega)} + \frac{\sin(\phi_\omega)\cos(\angle G(j3\omega))}{81\cos^2(\phi_\omega)}\right) + \cdots\right] \quad (23)$$

Ignoring high harmonic terms, we have

$$\phi_\omega = \arctan\left(\frac{p}{2\pi}q_c\right) \quad (24)$$

Then, the approximate phase angle of process is $\angle G(j\omega) \approx \phi_\omega - \pi$.

Figure 5 shows the relative errors of the amplitude ratio estimates of Eq. 21 and phase lag estimates of Eq. 24 for the FOPTD process. We can see that the estimates of Eqs. 21 and 24 have relative errors below 0.5%.

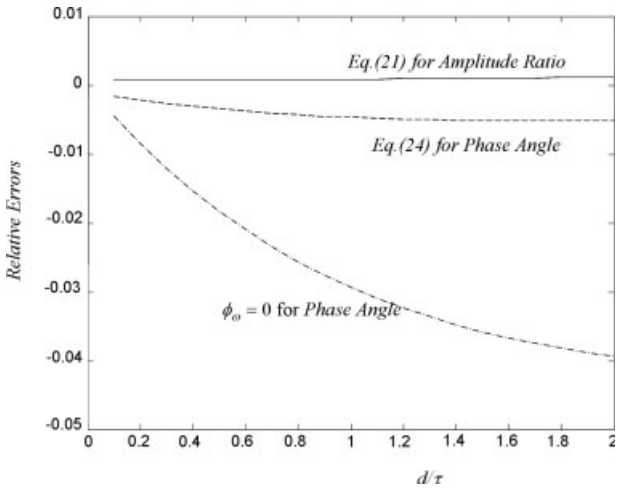

Figure 5. Relative error in estimation of the amplitude ratio and phase angle for FOPTD processes ($G(s) = \exp(-ds)/(\tau s+1)$).

Steady State Data Estimation

Process dynamic information contained in the cyclic steady state part of the relay responses corresponds to the frequency of $\omega = 2\pi/p$ and its multiples. Therefore, to estimate the process information at the frequency zero, we need the transient relay response from the start to the cyclic steady state. Here, a method to extract the process steady state information without storing the relay transient and complex computations is introduced. In Ma and Zhu,[18] it is shown that

$$\begin{aligned} U_a(s) &= \int_0^{t_1} u_a(t)\exp(-st)dt + \int_{t_1}^{\infty} u_a(t)\exp(-st)dt \\ &= \int_0^{t_1} u_a(t)\exp(-st)dt \\ &\quad + \frac{1}{1-\exp(-ps)}\int_{t_1}^{t_1+p} u_a(t)\exp(-st)dt \\ Y_a(s) &= \int_0^{t_1} y_a(t)\exp(-st)dt + \int_{t_1}^{\infty} y_a(t)\exp(-st)dt \\ &= \int_0^{t_1} y_a(t)\exp(-st)dt \\ &\quad + \frac{1}{1-\exp(-ps)}\int_{t_1}^{t_1+p} y_a(t)\exp(-st)dt \end{aligned} \tag{25}$$

Let

$$\begin{aligned} \tilde{U}_a(s) &= (1-\exp(-ps))U_a(s) \\ \tilde{Y}_a(s) &= (1-\exp(-ps))Y_a(s) \end{aligned} \tag{26}$$

Since $G(s) = \tilde{Y}_a(s)/\tilde{U}_a(s)$, we have (Appendix B)

$$G(0) = \lim_{s\to 0}\frac{\tilde{Y}_a(s)}{\tilde{U}_a(s)} = \tilde{Y}'_a(0)/\tilde{U}'_a(0) \tag{27}$$

where

$$\begin{aligned} \tilde{U}'_a(0) &= \int_{t_1}^{t_1+p} u_b(t)dt = p\times u_{bm} \\ \tilde{Y}'_a(0) &= \int_{t_1}^{t_1+p} y_b(t)dt = p\times y_{bm} \end{aligned}$$

The above Eq. 27 for the process steady state gain is equivalent to that in Huang et al.[16]

As in the moment analysis of Astrom and Hagglund,[6] we can obtain $G'(0)$ (Appendix B)

$$\begin{aligned} G'(0) &= \lim_{s\to 0}\frac{\tilde{Y}'_a(s) - G(s)\tilde{U}'_a(s)}{\tilde{U}_a(s)} \\ &= \lim_{s\to 0}\frac{\tilde{Y}''_a(s) - G'(s)\tilde{U}'_a(s) - G(s)\tilde{U}''_a(s)}{\tilde{U}'_a(s)} \\ &= -G'(0) + \frac{\tilde{Y}''_a(0) - G(0)\tilde{U}''_a(0)}{\tilde{U}'_a(0)} \\ &= \frac{\tilde{Y}''_a(0) - G(0)\tilde{U}''_a(0)}{2\tilde{U}'_a(0)} \end{aligned} \tag{28}$$

where

$$\begin{aligned} \tilde{U}''_a(0) &= 2p\int_0^{t_1} u_b(t)dt - 2\int_{t_1}^{t_1+p} tu_b(t)dt \\ \tilde{Y}''_a(0) &= 2p\int_0^{t_1} y_b(t)dt - 2\int_{t_1}^{t_1+p} ty_b(t)dt \end{aligned}$$

This process information is very useful in obtaining higher order process models.

Proposed Methods for First-Order Plus Time Delay Model Estimation

From the relay feedback responses, we can obtain the following first-order plus time delay (FOPTD) model.

$$G_m(s) = \frac{k\exp(-ds)}{\tau s + 1} \tag{29}$$

Since it has three unknowns, three experimental quantities are needed. From the steady state gain k of Eq. 27, the peak value a of $y_a(\tilde{t})$, and the relay oscillation period p, the FOPTD model can be determined analytically by the following previous approach.[20,22,25]

$$k = \frac{y_{bm}}{u_{bm}} \tag{30}$$

$$\tau = \frac{p}{2}\frac{1}{\ln\left(\frac{kh+a}{kh-a}\right)} \tag{31}$$

$$d = \tau\ln\left(\frac{\exp(p/(2\tau))+1}{2}\right) \tag{32}$$

These equations are exact for the FOPTD process.

Proposed Eq. 5

Equation 31 (consequently, Eq. 32) does not provide acceptable accuracy for the case of a large time delay with measurement errors. We overcome this problem using the integrals of the relay feedback responses. From the oscillation amplitude b of $y_b(\tilde{t})$ instead of the peak value a of $y_a(\tilde{t})$, we can obtain estimate of the time constant as

$$\begin{aligned} \tau &= \left(\frac{b}{kh} + \frac{p}{4}\right)\Big/\ln(\chi_1) \\ \chi_1^{c_1} &- 2\chi_1 + 1 = 0 \\ c_1 &= p\Big/\left(\frac{2b}{kh} + \frac{p}{2}\right) \end{aligned} \tag{33}$$

This equation can be obtained by rearranging the analytic equations for p and b in Table 1. Here, the proposed last five equations in Table 1 are derived from the first three equations.[20,22] Equation 33 is exact for the FOPTD process, but requires solving a nonlinear algebraic equation for χ_1. For a simpler application, we use the following approximate solution:

$$\tau = \frac{p}{2}\frac{1-1/c_1}{\ln\left[2 + \frac{1-6(1-c_1/2)^4}{1-2(c_1/2)^{-1/(c_1-1)}}\right]}. \tag{34}$$

Table I. Analytic Equations for Relay Feedback Responses

Process	Analytic Equations
$\frac{k\exp(-ds)}{\tau s+1}$	$y_a(\hat{t}) = kh\left(1-\frac{2}{1+\beta}\exp(-\hat{t}/\tau)\right)$,
	$\hat{t} = \tilde{t} - d,\ \beta = \exp(-p/(2\tau)),\ h =$ relay amplitude
	$p = 2\tau\ln(2\exp(d/\tau)-1)$
	$a = kh\frac{1-\beta}{1+\beta}$
	$q_a = (kh)^2\left(2-\frac{8\tau(1-\beta)}{p(1+\beta)}\right)$
	$y_b(\hat{t}) - y_{bm} = kh\left(\hat{t}-\tau-\frac{p}{4}+\frac{2\tau}{1+\beta}\exp(-\hat{t}/\tau)\right)$
	$b = kh\left(\tau\ln\left(\frac{1+\beta}{2}\right)+\frac{p}{4}\right)$
	$q_b = (kh)^2\left(\frac{p^2}{24}-2\tau^2+\frac{8\tau^3(1-\beta)}{p(1+\beta)}\right)$
	$F = \frac{q_a}{a^2} = 2\left(\frac{1+\beta}{1-\beta}\right)^2-\frac{8\tau}{p}\frac{1+\beta}{1-\beta}$

This equation can be obtained by applying perturbation analysis and numerical technique to the exact nonlinear equation.

Proposed Eq. 6

From the analytic equation for q_a in Table 1, we can obtain

$$\begin{aligned} \tau &= p/(4\chi_2) \\ \chi_2 &= \frac{1}{c_2}tanh(\chi_2) \\ c_2 &= 1-\frac{q_a}{2k^2h^2} \end{aligned} \tag{35}$$

Instead of solving the nonlinear equation for χ_2, we use an approximate solution:

$$\tau = \frac{p}{4}\frac{c_2}{tanh\left(\sqrt{15}\frac{1-c_2}{6c_2-1}\right)}. \tag{36}$$

This approximation is obtained by utilizing[26]

$$tanh(x) \approx \frac{x}{1+\frac{x^2}{3+\frac{x^2}{5}}}. \tag{37}$$

Proposed Eq. 7

From the analytic equation for q_b in Table 1, we can obtain

$$\begin{aligned} \tau &= p/(4\chi_3) \\ -c_3{\chi_3}^3+\chi_3 &= tanh(\chi_3) \\ c_3 &= \frac{1}{3}-\frac{8q_b}{p^2k^2h^2} \end{aligned} \tag{38}$$

Instead of solving the nonlinear equation for χ_3, we use an approximate solution:

$$\tau = \frac{p}{4}\frac{\sqrt{c_3}}{\sqrt{1-\sqrt{c_3}\,tanh(1/\sqrt{c_3})}}. \tag{39}$$

This approximation is obtained by applying numerical technique together with the approximation of Eq. 37.

Proposed Eq. 8

From Eq. 21 and $|G(j\omega)| = k/\sqrt{\tau^2\omega^2+1}$, we can also obtain

$$\tau = \frac{p}{2\pi}\sqrt{\frac{4p^2k^2h^2}{\pi^4q_b}-1}. \tag{40}$$

In summary, we estimate the time constant of the FOPTD model using one of the proposed methods of Eqs. 34, 36, 39, and 40. The static gain and time delay can be estimated by Eqs. 30 and 32, respectively. Figure 6 shows the relative approximation errors of equations in estimating the time constant τ. It is seen that errors are all within 1% except for approximation (40). Figure 7 shows relative error changes when the measurements of a, b, q_a, and q_b are not accurate. We can see that the estimated time constant based on a (previous approach of Eq. 31) is very sensitive to the measurement error when the time delay is large. This sensitivity is because the time constant change does not affect the amplitude of $y_a(t)$ much when the time delay is large. One should be cautious in using the Eq. 31 when the time delay is expected to be large. Figure 7 shows that the proposed methods can relieve this disadvantage.

The curvature factor by Luyben[19] can be used to check how large d/τ is. Without measuring an additional data for the Luyben's curvature factor, as an alternative, a quantity

$$F = q_a/a^2 \tag{41}$$

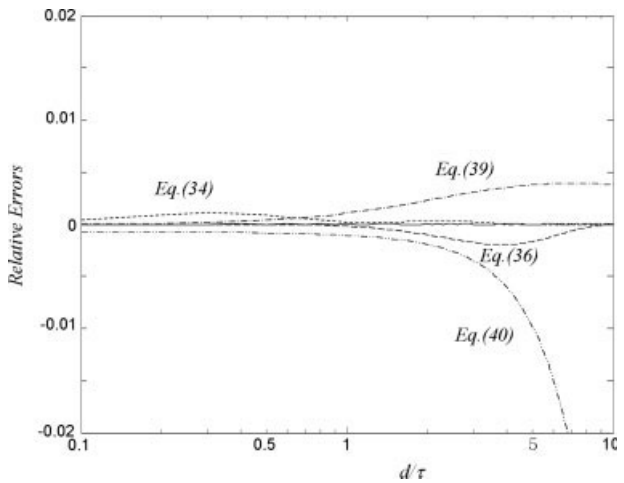

Figure 6. Relative error in estimations of the time constant τ for FOPTD processes ($G(s) = \exp(-ds)/(\tau s+1)$).

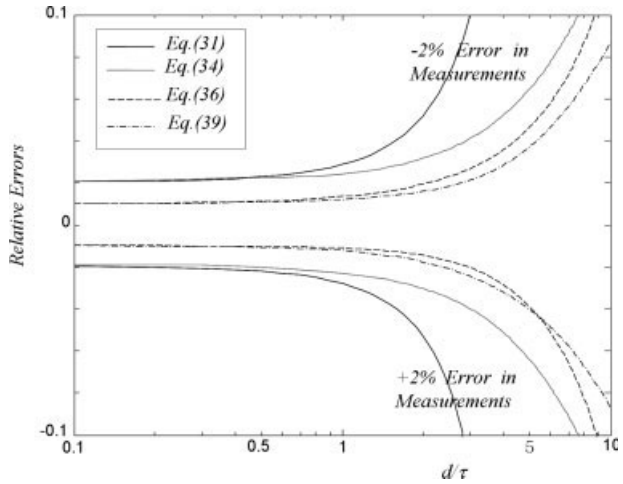

Figure 7. Sensitivities of the estimated time constant τ for FOPTD processes ($G(s) = \exp(-ds)/(\tau s+1)$).

is used here. Figure 8 shows the curvature factor by Luyben[19] and Eq. 41 for FOPTD processes. The curvature factor is useful to select an appropriate modeling method before we estimate the model parameters. For example, the method of Eq. 31 is not appropriate if the proposed curvature factor is larger than about 0.7 (which corresponds to $d/\tau = 1.0$), as shown in Figure 7.

A FOPTD model can also be obtained without process steady state gain information.[19,22] However, because information in the relay oscillation responses is concentrated around ultimate frequencies, model parameters can be very inaccurate. So, such methods are not recommended in general.

Figure 9 shows the integral of absolute error (IAE) in the frequency domain for high order processes, critically damped second-order plus time delay (SOPTD) processes, underdamped SOPTD processes, and processes with inverse responses.

$$\text{IAE} = \int_0^{2\pi/p} \frac{|G_m(j\tilde{\omega}) - G(j\tilde{\omega})|}{|G(j\tilde{\omega})|} d\tilde{\omega} \tag{42}$$

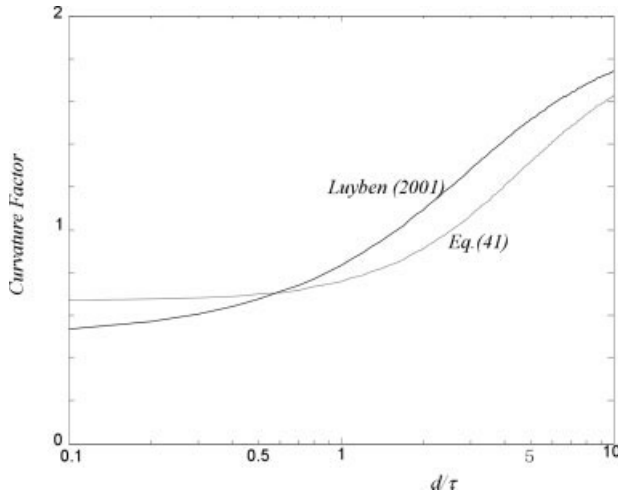

Figure 8. Curvature factor plot for FOPTD processes ($G(s) = \exp(-ds)/(\tau s+1)$).

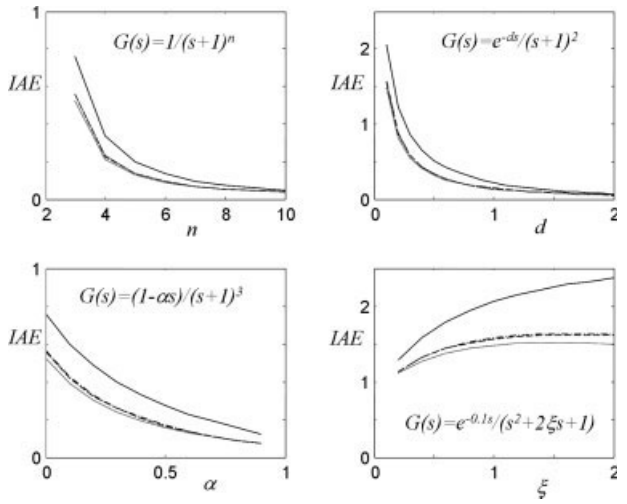

Figure 9. IAE plots of FOPTD models for several higher order processes (solid line: Eq. 31, dotted line: Eq. 34, dashed line: Eq. 36, dash-dotted line: Eq. 39).

It is seen that the proposed estimations of Eqs. 34, 36, and 39 improve the application of the previous approach of Eq. 31.

Proposed Method for Critically Damped Plus Time Delay Model Estimation

Panda and Yu have shown that a three-parameter model of

$$G_m(s) = \frac{k \exp(-ds)}{(\tau s + 1)^2} \tag{43}$$

is applicable to a wide range of processes.[22] Model parameters for this critically damped plus time delay (CDPTD) model are estimated. For the steady state gain k, Eq. 30 is used. From Eq. 21 and $A_\omega = |G_m(j\omega)| = k/((\tau\omega)^2 + 1)$, we have

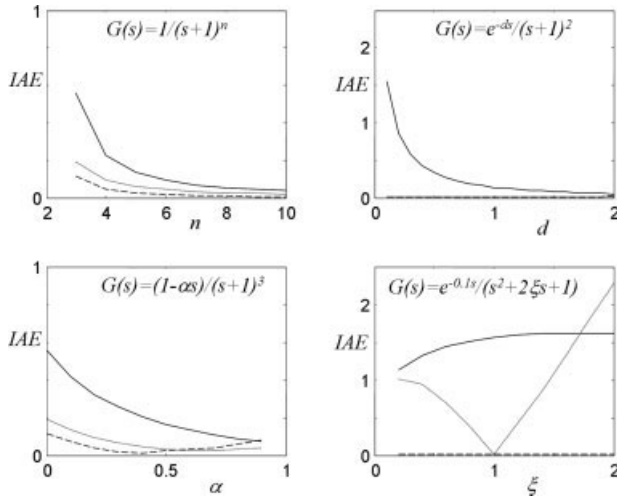

Figure 10. IAE plots for several higher order processes (solid line: FOPTD model (Eqs. 30, 32, and 39), dotted line: CDPTD model (Eqs. 43–46), dashed line: SOPTD model (Eqs. 47–51)).

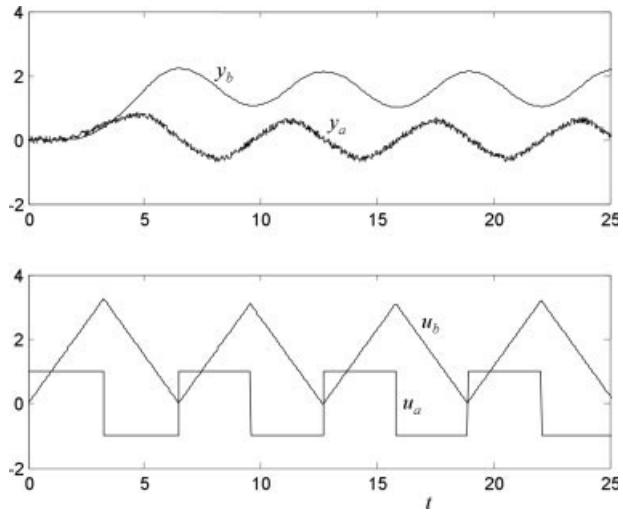

Figure 11. Relay feedback responses subject to measurement noise.

$$\tau = \frac{p}{2\pi}\sqrt{\frac{2khp}{\pi^2\sqrt{q_b}} - 1} \tag{44}$$

An analytic equation for the oscillation period[3,22] is

$$\exp\left(-\frac{d}{\tau}\right) + c_3\frac{d}{\tau} - c_4 = 0 \tag{45}$$

where

$$c_3 = \frac{2\beta}{1+\beta}$$
$$c_4 = \frac{2\left((1+p/(2\tau))\beta + \beta^2\right)}{(1+\beta)^2}$$
$$\beta = \exp(-p/(2\tau))$$

Equation 45 can be solved for d/τ by the Newton-Raphson method with an initial estimate

$$\frac{d}{\tau} = \frac{p}{2\tau}\left(1 - \frac{2}{\pi}\arctan\left(\frac{2\pi\tau}{p}\right)\right) \tag{46}$$

Proposed Method for Second-Order Plus Time Delay Model Estimation

A general second-order plus time delay (SOPTD) model

$$G_m(s) = \frac{k\exp(-ds)}{a_2s^2 + a_1s + 1} \tag{47}$$

has four unknowns; thus four experimental quantities are needed. We use the process steady state gain information of Eq. 27, its derivative of Eq. 28, q_b of Eq. 18, and the oscillation period p. We solve

$$k = G_m(0) = \frac{y_{bm}}{u_{bm}} \tag{48}$$

$$k(-d - a_1) = G_m'(0) \tag{49}$$

$$q_b = \frac{16h^2}{\pi^2\omega^2}\left(|G_m(j\omega)|^2\right) = \frac{16h^2}{\pi^2\omega^2}\frac{k^2}{(1-a_2\omega^2)^2 + (a_1\omega)^2} \tag{50}$$

$$d = \frac{p}{2}\left(1 - \arctan(q_c/\omega) + \arctan2(1 - a_1\omega^2, a_2\omega)\right) \tag{51}$$

The computational procedure is as follows:

- Step 0: From relay feedback response, obtain $G_m(0)$, $G_m'(0)$, q_b, and the oscillation period p. Let $d = p/4$.
- Step 1: Calculate a_1 from Eq. 49.
- Step 2: Calculate a_2 from Eq. 50.
- Step 3: Adjust d and repeat Steps 1 and 2 so that Eq. 51 is satisfied.

Figure 10 shows the IAE values in the frequency domain for various processes. For most processes tested, the CDPTD model by Eqs. 43–46 shows better results than the FOPTD model by Eqs. 30, 32, and 39. We can see that the SOPTD model by Eqs. 47–51 is the best. However, the SOPTD model needs more measurements and iterative computations.

Simulations

Simulations are performed to investigate the performances of methods under sampling, discretization errors, and noisy environments. The following process is considered.

$$G(s) = \frac{\exp(-1.2s)}{(s+1)^2(0.5s+1)}. \tag{52}$$

Table 2. Identification Results of the Proposed Method

Process	Run No.	Output Noise Size	$G(0)$ (1.00)*	$G'(0)$ (−3.70)*	Identified Models	IAE
$\frac{\exp(-1.2s)}{(s+1)^2(0.5s+1)}$	1	0	1.009	−3.780	$\frac{1.009\exp(-2.210s)}{1.968s+1}$	0.093
					$\frac{1.009\exp(-1.574s)}{(1.107s+1)^2}$	0.019
					$\frac{1.009\exp(-1.565s)}{1.191s^2+2.215s+1}$	0.013
	2	0.1	1.025	−4.114	$\frac{1.025\exp(-2.0961s)}{2.011s+1}$	0.062
					$\frac{1.025\exp(-1.465s)}{(1.112s+1)^2}$	0.039
					$\frac{1.025\exp(-1.657s)}{0.974s^2+2.457s+1}$	0.085
	3	0.1	1.002	−3.432	$\frac{1.002\exp(-2.076s)}{2.030s+1}$	0.076
					$\frac{1.002\exp(-1.443s)}{(1.117s+1)^2}$	0.040
					$\frac{1.002\exp(-1.317s)}{1.724s^2+2.115s+1}$	0.053

*Exact value.

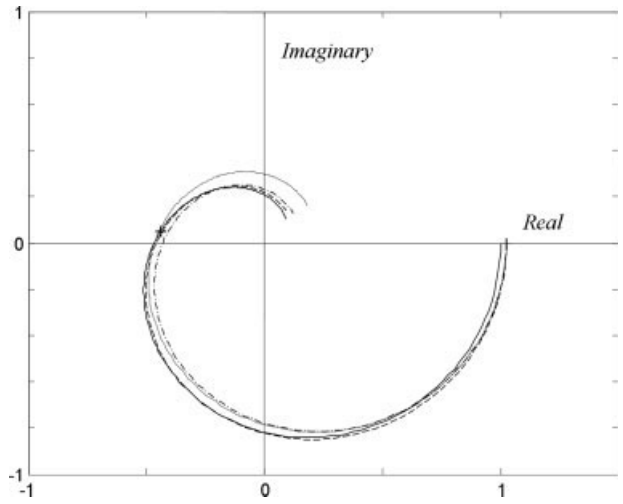

Figure 12. Nyquist plots of the process of Eq. 52 and its estimates (solid line: FOPTD model (Eqs. 30, 32, and 39), dotted line: CDPTD model (Eqs. 43–46), dashed line: SOPTD model (Eqs. 47–51)).

Uniformly distributed noise $n(t)$ with mean value and magnitude of 0, and 0.2 is introduced in the output. For noisy responses, q_a should be adjusted[6] as

$$q_a = \frac{2}{p}\int_{t_1}^{t_1+p} y_a(t)^2 dt - 2\bar{\sigma}_a{}^2,\ \bar{\sigma}_a{}^2 = \frac{1}{T}\int_0^T n(t)^2 dt \qquad (53)$$

The noise characteristics, $\bar{\sigma}_a{}^2 = \int_0^T n(t)^2 dt/T$, can be estimated before starting the relay feedback test. Similarly, q_b needs to be adjusted as in q_a when noise is involved.

Simulation results are shown in Figure 11. Models obtained are in Table 2, and their Nyquist plots are in Figure 12. For the FOPTD model, estimation Eqs. of 30, 32 with 39 are used. Because the process (52) is an overdamped process with a moderate time delay, all models estimated have excellent agreement with the true model, as shown by the Nyquist plot of Figure 12.

Conclusions

Without modifying the original relay feedback system, simple and accurate estimates of process ultimate information are easily obtained. We use various integrals of the relay responses instead of point data to reduce the effects of the high harmonic terms. For FOPTD processes, errors over 15% of the previous approaches in estimating the ultimate gain can be reduced to below 5%. We derived the equations to extract a FOPTD, CDPTD, and SOPTD models from the process information at the steady state and near the ultimate frequency. They are very simple and can be applied easily to commercial PID controllers.

Literature Cited

1. Astrom KJ, Hagglund T. Automatic tuning of simple regulators with specifications on phase and amplitude margins. *Automatica*. 1984; 20:645–651.
2. Hang CC, Astrom KJ, Wang QG. Relay feedback auto-tuning of process controllers—a tutorial review. *J Process Control*. 2002;12:143–162.
3. Yu CC. *Autotuning of PID Controllers: A Relay Feedback Approach*. London: Springer, 2006.
4. Tan KK, Lee TH, Huang S, Chua KY, Ferdous R. Improved critical point estimation using a preload relay. *J Process Control*. 2005;16:445–455.
5. Sung SW, Park JH, Lee I. Modified relay feedback method. *Ind Eng Chem Res*. 1995;34:4133–4135.
6. Astrom KJ, Hagglund T. *PID Controllers*. NC: Instrument Society of America, 1995.
7. Kim YH. PI controller tuning using modified relay feedback method. *J Chem Eng Jpn*. 1995;28:118–121.
8. Tan KK, Lee TH, Wang QG. Enhanced automatic tuning procedure for process control of PI/PID controllers. *AIChE J*. 1996;42:2555–2562.
9. Chiang R, Yu CC. Monitoring procedure for intelligent control: online identification of maximum closed-loop log modulus. *Ind Eng Chem Res*. 1993;32:90–99.
10. Friman M, Waller KV. A two channel relay for autotuning. *Ind Eng Chem Res*. 1997;36:2662–2671.
11. Sung SW, Lee J. A two-channel relay feedback method under static disturbances. *Ind Eng Chem Res*. 2006;45:4071–4074.
12. Hang CC, Astrom KJ, Ho WK. Relay auto-tuning in the presence of static load disturbance. *Automatica*. 1993;29:563–564.
13. Shen SH, Wu JS, Yu CC. Autotune identification under load disturbance. *Ind Eng Chem Res*. 1996;35:1642–1651.
14. Sung SW, Lee J. Relay feedback method under large static disturbances. *Automatica*. 2006;42:353–356.
15. Shen SH, Wu JS, Yu CC. Use of biased relay feedback method for system identification. *AIChE J*. 1996;42:1174–1181.
16. Huang HP, Jeng JC, Luo KY. Auto-tune system using single-run relay feedback test and model-based controller design. *J Process Control*. 2005;15:713–727.
17. Wang QG, Hang CC, Bi Q. Process frequency response estimation from relay feedback. *Control Eng Practice*. 1997;5:1293–1302.
18. Ma MD, Zhu XJ. A simple auto-tuner in frequency domain. *Comput Chem Eng*. 2006;30:581–586.
19. Luyben WL. Getting more information from relay feedback tests. *Ind Eng Chem Res*. 2001;40:4391–4402.
20. Kaya I, Atherton DP. Parameter estimation from relay autotuning with asymmetric limit cycle data. *J Process Control*. 2001;11:429–439.
21. Panda RC, Yu CC. Analytic expressions for relay feedback responses. *J Process Control*. 2003;13:489–501.
22. Panda RC, Yu CC. Shape factor of relay response curves and its use in autotuning. *J Process Control*. 2005;15:893–906.
23. Sung SW, O J, Lee I, Lee J, Yi SH. Automatic tuning of PID controller using second-order plus time delay model. *J Chem Eng Jpn*. 1996;29:990–999.
24. Sung SW, Lee I. Enhanced relay feedback method. *Ind Eng Chem Res*. 1997;36:5526–5530.
25. Wang QG, Hang CC, Zou B. Low-order modeling from relay feedback. *Ind Eng Chem Res*. 1997;36:375–381.
26. Abramowitz M, Stegun IA. *Handbook of Mathematical Functions*. New York: Dover, 1972.
27. Kreyszig E. *Advanced Engineering Mathematics*. New York: Wiley, 1999.

Appendix A: Orthogonality of Trigonometric Functions[27]

For a set of trigonometric functions, $\{1, \cos(\omega t), \cos(2\omega t), \ldots, \sin(\omega t), \sin(2\omega t), \ldots\}$, we have

$$\begin{aligned}
&\frac{2}{p}\int_t^{t+p} \sin(n\omega t)\sin(m\omega t)dt = \begin{cases} 0, & n \neq m \\ 1, & n = m \end{cases} \\
&\frac{2}{p}\int_t^{t+p} \sin(n\omega t)\cos(m\omega t)dt = 0 \\
&\frac{2}{p}\int_t^{t+p} \cos(n\omega t)\cos(m\omega t)dt = \begin{cases} 0, & n \neq m \\ 1, & n = m \neq 0 \\ 2, & n = m = 0 \end{cases}
\end{aligned} \qquad (A1)$$

for $p = 2\pi/\omega$. Applying these relationships, we can derive Eqs. 16 and 17. For example, for a function of $f(t) = \alpha_0 + \alpha_1 \cos(\omega t) + \beta_1 \sin(\omega t) + \alpha_2 \cos(2\omega t) + \beta_2 \sin(2\omega t) + \cdots$,

$$\frac{2}{p}\int_t^{t+p} f(t)^2 dt = 2{\alpha_0}^2 + {\alpha_1}^2 + {\beta_1}^2 + {\alpha_2}^2 + {\beta_2}^2 + \cdots \quad \text{(A2)}$$

This is known as the Parseval theorem.

Appendix B: Moment Analyses[6]

Since

$$\tilde{U}_a(s) \equiv (1 - \exp(-ps)) \int_0^{t_1} u_a(t) \exp(-st) dt + \int_{t_1}^{t_1+p} u_a(t) \exp(-st) dt \quad \text{(B1)}$$

we have

$$\begin{aligned}
\tilde{U'}_a(s) &\equiv pe^{-ps} \int_0^{t_1} u_a(t) e^{-st} dt \\
&\quad + (1 - e^{-ps}) \int_0^{t_1} (-t) u_a(t) e^{-st} dt \\
&\quad + \int_{t_1}^{t_1+p} (-t) u_a(t) e^{-st} dt \\
\tilde{U''}_a(s) &\equiv -p^2 e^{-ps} \int_0^{t_1} u_a(t) e^{-st} dt \\
&\quad - 2pe^{-ps} \int_0^{t_1} t u_a(t) e^{-st} dt \\
&\quad + (1 - e^{-ps}) \int_0^{t_1} t^2 u_a(t) e^{-st} dt \\
&\quad + \int_{t_1}^{t_1+p} t^2 u_a(t) e^{-st} dt
\end{aligned} \quad \text{(B2)}$$

Then applying integration by parts:

$$\begin{aligned}
\tilde{U'}_a(0) &= p \int_0^{t_1} u_a(t) dt - \int_{t_1}^{t_1+p} t u_a(t) dt \\
&= p u_b(t_1) - t u_b(t)|_{t_1}^{t_1+p} + \int_{t_1}^{t_1+p} u_b(t) dt \\
&= \int_{t_1}^{t_1+p} u_b(t) dt \\
\tilde{U''}_a(0) &= -p^2 \int_0^{t_1} u_a(t) dt - 2p \int_0^{t_1} t u_a(t) dt \\
&\quad + \int_{t_1}^{t_1+p} t^2 u_a(t) dt \\
&= 2p \int_0^{t_1} u_b(t) dt - 2 \int_{t_1}^{t_1+p} t u_b(t) dt
\end{aligned} \quad \text{(B3)}$$

where $u_b(t) = \int_0^{t_1} u_a(t) dt$ and $u_a(t)$ and $u_b(t)$ are periodic after $t = t_1$.

By applying the same procedure to

$$\tilde{Y}_a(s) \equiv (1 - \exp(-ps)) \int_0^{t_1} y_a(t) \exp(-st) dt + \int_{t_1}^{t_1+p} y_a(t) \exp(-st) dt \quad \text{(B4)}$$

we can obtain equations for $\tilde{Y'}_a(0)$ and $\tilde{Y''}_a(0)$:

$$\begin{aligned}
\tilde{Y}'_a(0) &= \int_{t_1}^{t_1+p} y_b(t) dt \\
\tilde{Y}''_a(0) &= 2p \int_0^{t_1} y_b(t) dt - 2 \int_{t_1}^{t_1+p} t y_b(t) dt
\end{aligned} \quad \text{(B5)}$$

Manuscript received Oct. 12, 2006, and revision received Jun. 13, 2007.