96-4009

A96-38726

AIAA-96-4009-CP

Tuning PID Controllers Using Error-Integral Criteria and Numerical Optimization

Masood Sahraian*

United Defense LP, Santa Clara, CA

Srinivas Kodiyalam**

Lockheed Martin Advanced Technology Center, Palo Alto, CA

Abstract

A method for tuning PID controllers based on an error measure and numerical optimization is presented. A state-space realization of the controller and the plant is used. A response surface approximation based optimization approach is used to minimize the number of detailed controller response analyses.

A heuristic algorithm is used to select designs that are close to D-optimal designs to construct the initial linear approximation. The tuning methodology is directly applicable to multi-input, multi-output control systems and is shown to obtain improved results compared to the traditional tuning techniques such as Ziegler-Nichols method. Realistic applications are provided to illustrate the method outlined in this paper.

1. Introduction

Proportional, integral, and derivative (PID) control modes are linear control actions which are implemented in the majority of commercial controllers because of their relative ease of implementation and adequate performance. A PID control law is also known as a three mode controller; the user can modify the dynamic properties of this controller by adjusting the three proportional, integral, and derivative parameters. A judicial selection of the controller parameters such that the control system performance requirements are met is known as PID tuning.

There are many techniques for tuning PID controllers, among them (1) open-loop techniques, and (2) closed-loop techniques [1]. In the open-loop method the process is characterized by its response to a unit step in the controller output. Normally a very simple first order system plus a time delay is used as a model of the process and the process response is fitted to this model. The controller settings for proportional, integral, and derivative gains are then calculated as a function of process characteristics. In a closed-loop method known as the Ziegler-Nichols method, an ultimate period of oscillations for the system response is determined by increasing the proportional gain while keeping integral and derivative gains at their minimum values. The controller settings for proportional, integral, and derivative gains are determined such that the process response produces a quarter decay ratio (the difference between each successive peak in the response and the steady state value is one fourth of the previous difference). The closed-loop methods like open-loop methods rely on a simple mathematical model of the process in the tuning procedure.

In another closed-loop technique known as the error-integral method a more systematic procedure is devised to find the controller settings based on minimizing the error in the response of the controller [2]. A variety of error-integral measures such as integrated square error (ISE) or integrated absolute error (IAE) are used. We refer to Figure 1 to describe this method. In this schematic a simple process (first order system plus time delay) is controlled

* Senior Analysis Engineer; Engineering Analysis, Simulation, and Test Department.

** Senior Staff Scientist, Structures Department, Senior Member AIAA.

by a PID controller. The controller computes the control signal *U(s)* based on its error input *E(s)*. In error-integral method of tuning a controller, an error criterion such as the following is defined:

$$\Phi = \int_0^{\infty} L(e(t))\, dt \qquad (1.1)$$

where L is a suitable function of the error *e(t)* such as a quadratic function (ISE measure). Note that the error *e(t)* is a function of the PID parameters, namely K_p, T_i, and T_d. These parameters are found by setting the Jacobian of Φ equal to zero. It is assumed that the function *L(e(t))* is zero only if the error is identically zero and that it possesses a minimum value.

The above description follows the method given in [2] and assumes a very simple process and relies on a procedure which is not suitable for fast, reliable, and automatic tuning of PID controllers. Note that the function Φ depends on the controller parameters and that this function is not normally known in a closed form since that would require having available an inverse Laplace transform of E(s) which is not readily available. Therefore even for the simple system of Figure 1 treated in [2], the minimum of equation (1.1) would have to be solved using a numerical procedure.

In what follows a method is described based on state-space methods and numerical optimization techniques. Numerical optimization techniques have long been used in integrated control and structural design of space systems [3]. The PID tuning method proposed in this paper does not make simplifying assumptions on the type of the process and is meant for implementation in a Computer-Aided Control Engineering (CACE) environment such as Matlab* or Matrix-X**.

* Matlab is a registered tradmark of Mathworks, Inc.
** Matrix-X is a registered trademark of Integrated Systems, Inc.

2. Formulation of the Problem

Figure 2 shows a general structure for the control system. It is assumed that the plant *P(s)* has the following state-space realization:

$$\dot{x} = F\,x + G\,u(t) \qquad (2.1)$$

$$y = Hx \qquad (2.2)$$

where x is a n x 1 state vector, u is the control vector, and y is the output vector. Without loss of generality we assume a single input, single output (SISO) system in equations (2.1) and (2.2). In that case u and y are both scalar functions of time. The treatment in this paper is, however, general and with simple modifications may be applied to a multi-input, multi-output (MIMO) system. The model representing the plant can be either a linear frequency domain model or a general nonlinear state-space model normally known as a simulation model. In both cases it is possible to transform the model of the plant into a linear, time-invariant form given by (2.1) and (2,2). If the system we start with is a general nonlinear simulation model, a linearization about a suitable operating point can be performed to transform the system to the linear, time-invariant form given above. All the transformation operations described in this paragraph can be performed using CACE tools.

Referring to Figure 2 again, the controller *D(s)* is assumed to be a PID controller given by

$$D(s) = \frac{U(s)}{E(s)} \qquad (2.3)$$

where

$$u(t) = K_p\,[e(t) + T_d\,\dot{e}(t) + \frac{1}{T_i}\int_t e(t)\,dt\,] \qquad (2.4)$$

The error input to the controller is given by

$$e(t) = y_{ref}(t) - H\,x \qquad (2.5)$$

where y_{ref} (or *r(t)*) is a reference signal that the output of the plant *y*(t) has to follow, normally a constant value.

We use the following strategy to tune the controller parameters as given in (2.4): the control law (2.4) is used in the plant system (2.1). The error output of the plant *e(t)* is then used to minimize a performance index given by (1.1). The form of the function *L(e(t))* depends on the performance specifications of the control system but it is customary to assume a quadratic function of *e(t)*. The following section describes the solution method.

3. Solution Method

In this section we first present the details of the equations arising from implementing the PID controller (2.4) in the state-space system (2.1-2). Then the optimization technique to solve for the tuning parameters is presented.

3.1 Details of the Equations

Before using (2.4) in the state-space system we need to introduce an additional state to properly represent the integral part of the controller. Let

$$\dot{x}_I = e(t) \tag{3.1.0}$$

which implies

$$x_I = \int_t e(t)\, dt \tag{3.1.1}$$

Now the augmented state vector is

$$x_a = \begin{bmatrix} x \\ x_I \end{bmatrix} \tag{3.1.2}$$

where x is a $n \times 1$ vector and x_I is a scalar since we have only a single-output system in this case. For a multi-output system x_I would be a column vector and the appropriate changes have to be made. The method described below, however , applies just the same. The augmented state-space system is then given by combining (3.1.0) and (2.1) to get

$$\dot{x}_a = F_a x_a + G_a u(t) + G_1 y_{ref}(t) \tag{3.1.3}$$

$$y_a = H_a x_a = y \tag{3.1.4}$$

where

$$F_a = \begin{bmatrix} F & 0 \\ -H & 0 \end{bmatrix} \tag{3.1.5}$$

$$G_a = \begin{bmatrix} G \\ 0 \end{bmatrix} \tag{3.1.6}$$

$$G_1 = \begin{bmatrix} 0_{n\times 1} \\ 1 \end{bmatrix} \tag{3.1.7}$$

$$H_a = \begin{bmatrix} H & 0 \end{bmatrix} \tag{3.1.8}$$

We now re-write the PID control law (2.4) by using the error equation (2.5) and its first time derivative. Note that

$$\dot{e}(t) = \dot{y}_{ref}(t) - H\dot{x} = \dot{y}_{ref}(t) - H_a\dot{x}_a \tag{3.1.9}$$

$$x_I = G_2 x_a \tag{3.1.10}$$

$$G_2 = \begin{bmatrix} 0_{1\times n} & 1 \end{bmatrix} \tag{3.1.11}$$

After substituting (3.1.9) and (2.5) into (2.4) we get

$$u(t) = K_p(y_{ref} - H_a x_a) + K_p T_d(\dot{y}_{ref} - H_a \dot{x}_a) + \frac{K_p}{T_i} G_2 x_a \tag{3.1.12}$$

The state-space equations (3.1.3) is re-written by substituting for u from (3.1.12)

$$\dot{x}_a = J(F_a - K_p G_a H_a + \frac{K_p}{T_i} G_a G_2) x_a + J(G_1 + K_p G_a) y_{ref} + K_p T_d J G_a \dot{y}_{ref} \tag{3.1.13}$$

where

$$J = (I + K_p T_d G_a H_a)^{-1} \quad (3.1.14)$$

It is assumed that the matrix inside parenthesis in (3.1.14) is nonsingular.

Note that what we have done so far is to represent the PID controller in a form similar to state feedback form in equation (3.1.12). We should note, however, that this has been an ad-hoc procedure and that: (a) the control law as presented in (3.1.12) does not necessarily have any desirable stability property, and (b) the state-space system (3.1.13) does not necessarily have any desirable controllability-observability properties

To summarize this section: the PID control law (2.4) was substituted into the modified state-space equations (3.1.3) to arrive at the closed-loop equations (3.1.13). In the following section this state-space system is used along with the optimality index (1.1) to find tuning parameters that will optimize the controller (2.4) in the sense defined by (1.1).

3.2 Tuning by Optimization

The closed-loop system (3.1.13) matrices are dependent on the three tuning parameters symbolically presented as

$$\xi = \begin{bmatrix} K_p \\ T_i \\ T_d \end{bmatrix} \quad (3.2.1)$$

The system equations (3.1.13) are re-written in the following form

$$\dot{x}_a = A(\xi)\, x_a + g(\xi, t) \quad (3.2.2)$$

$$y = H_a\, x_a \quad (3.2.3)$$

where

$$A(\xi) = J(F_a - K_p G_a H_a + \frac{K_p}{T_i} G_a G_2) \quad (3.2.4)$$

$$g(\xi, t) = J(G_1 + K_p G_a)\, y_{ref} + K_p T_d J G_a \dot{y}_{ref} \quad (3.2.5)$$

The optimization problem is formulated in the following way: Find the set of design variables, ξ, such that

$$\Phi(\xi) = \int_0^{\infty} L(e(t))\, dt \quad (3.2.6)$$

is minimized, with the side constraints,

$$\xi^l \le \xi \le \xi^u$$

where,

$$e(t) = y_{ref}(t) - H_a\, x_a \quad (3.2.7)$$

The choice of the function L depends on the performance requirements and control system. A quadratic function is normally used, i.e.,

$$\Phi(\xi) = \int_0^{\infty} e^T e\, dt \quad (3.2.8)$$

The solution ξ determines the optimal tuning parameters.

4. Optimization based on Response Surfaces

The optimization problem defined by equation (3.2.6-3.2.8) is solved using numerical optimization techniques [4] along with a response surface methodology [5] for approximating the controller responses as a function of the tuning parameters, ξ. A flowchart of the tuning optimization procedure is shown in Figure 3. The use of response surfaces for approximation of the controller responses is primarily to minimize the number of detailed solutions of the state-space equations during optimization. The implementation of response surface approach is similar to that described in Reference [6].

A combination of experiment design and regression analysis methods are used with this implementation of response surface methodology. The layout of the response surface is defined using a second order approximation of the form:

$$\phi = \beta_0 + \sum_{i=1}^{ndv} \beta_i \xi_i + \sum_{i=1}^{ndv} \beta_{ii} \xi_i^2 + \sum_{ij(i<j)} \beta_{ij} \xi_i \xi_j \quad (4.1)$$

Using a least squares procedure, the polynomial coefficients, β_{ij}, are determined. The first step in the response surface construction is to generate and analyze (ndv +1) designs for a linear approximation. Here, ndv corresponds to the number of design variables, ξ. A heuristic algorithm for generating experimental designs using Fedorov's exchange algorithm [7], is used for generating (ndv +1) designs that are close to being D-optimal designs. This algorithm attempts to maximize the determinant of $X^T X$ for the (ndv+1) selected design points, from a set of candidate design points. Matrix X is referred to as a design matrix of candidate points from which the design points for the linear approximation would be selected. For the determinant evaluation, X consists of only those rows of the original matrix of candidate points that correspond to the selected points.

If the linear model does not meet the requirements, and as each of the additional designs are analyzed, the response surface approximation is sequentially refined to a full quadratic approximation. When more number of designs than what is required for a quadratic approximation is available, the best (ndv+1 + (ndv+1)*ndv/2) are retained for constructing the response surface.

Finally, the approximate optimization problem is solved using the BFGS method programmed in ADS optimizer [8].

5. Applications

In this section two applications are presented which illustrate the method. The first application provides a proof of concept. The second application provides a slightly different formulation of the problem where an optimal plant parameter is optimized along with the PID tuning parameters.

5.1 First Application

In this section the tuning method presented above is applied to the following plant [1].

$$P(s) = \frac{1}{(s+1)^3} \tag{5.1}$$

This transfer function corresponds to the following state-space realization of equations (2.1) and (2.2).

$$F = \begin{bmatrix} -0.9999 & -394.8590 & 394.8590 \\ 0 & -1 & -3.2069e-06 \\ 0 & 3.2069e-06 & -1 \end{bmatrix}$$

$$G = \begin{bmatrix} 0 & 394.8590 & 394.8590 \end{bmatrix}^T$$

$$H = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}$$

The structure of the control system is as in Figure 2. In this figure *D(s)* is the transfer function of the PID controller as in equation (2.4).

The search for an optimal design started with a baseline PID design parameters as in Table 1. The baseline design results in a stable system. The dynamic performance of this design, however, is very poor as can be seen in Figure 4. This figure shows the closed-loop response of system (3.2.2) to a unit disturbance and a zero reference input. This design results in very large settling time and overshoot giving rise to a relative large performance measure for Φ (a value of 1.1977). The PID parameters and the value of performance measure Φ for the final optimal design is shown in Table 1. This table also shows the controller parameters as computed by Ziegler-Nichols and Shinskey methods [1].

The closed-loop response based on the second optimizer iteration is also shown in Figure 4. Despite the fact that the optimizer started with a linear approximation the solution is improved considerably only after two iterations. An optimization iteration history is shown in Figure 6. The closed-loop response for the final optimal design and for Ziegler-Nichols and Shinskey tuning methods are shown in Figure 5. The ISE measure for these three solutions along with the time to reach steady state (t_{ss}) and the maximum overshoot (y_{max}) are shown in Table 2. Finally the closed-loop poles (eigenvalues of matrix A in equation (3.2.4)) for these designs are presented in Table 3. The control system performance based on the optimal design is better than that of the Ziegler-Nichols method. The latter method has the best results of many different traditional tuning methods as discussed in [1].

Note that in the Ziegler-Nichols method an approximation of the plant transfer function (equation (5.1)) is required to arrive at a first order plus dead-time model as in Figure 1. The reason for this approximation is that the ultimate period and gain which are used to compute the PID parameters in Ziegler-Nichols method are based on this type of model [2]. In the method presented in this paper, however, no such approximations are necessary and the original model of the plant is used to compute an optimal PID design.

Although the application presented here is for a single-input, single-output (SISO) system, the method is directly applicable to a multi-input, multi-output (MIMO) system such as a cascaded control used in many processing plants. In fact we expect the method to be more powerful in such cases where the traditional methods break down.

5.2 Second Application

Our present focus is on applying the present methodology to an integrated control-process design. In this application the PID tuning parameters and a process gain are optimized simultaneously. The process gain can incorporate such effects as the actuator characteristics.

Summary

A methodology for optimal tuning of PID controllers was presented. An integrated square error (ISE) measure containing the closed-loop response error in the controlled variable was minimized by using response surface optimization. The result of this optimization is a tuned controller which results in superior dynamic performance. Other error measures such as integrated absolute error (IAE) can be used as well. The method was applied to a SISO control system. A MIMO control system, however, can be tackled directly using the method presented in this paper.

References

1. L. Martins and D.E. Carvalho, *Dynamic Systems and Automatic Control,* 1994.

2. A.M. Lopez, et. al., *Tuning Controllers with Error Integral Criteria,* Instrumentation Technology, November 1967.

3. A. Messac, et. al., *Control-Structure Integrated Design: a Computational Approach,* Proceedings, AIAA/ASME/ASCE/AHS /ASC 31st Structures, Structural Dynamics and Materials Conference, AIAA, Washington, D.C., April 1991.

4. G.N. Vanderplaats, *Numerical Optimization Techniques for Engineering Design with Applications,* McGraw-Hill, New York, USA,

5. J. Krottmaier, *Optimizing Engineering Designs,* McGraw-Hill, Berkshire, UK, 1993.

6. H. Thomas, *Optimization of Structures Designed using Nonlinear FEM Analysis,* AIAA/ASME/ASCE/AHS/ASC 37th Structures, Structural Dynamics and Materials Conference, AIAA, Washington, D.C., April 1996. AIAA CP: 96-1385.

7. N.K. Nguyen and A. J. Miller, *A Review of Some Exchange Algorithms for Constructing Discrete D-optimal Designs,* Comput. Statistics and Data Analysis, vol. 14, 1992, pp. 489-498.

8. G.N. Vanderplaats, *ADS Users Manual, Version 2.01,* VMA Engineering, Colorado Springs, Colorado, 1987.

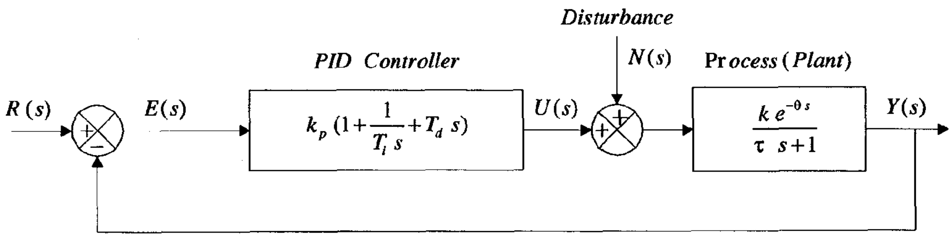

Figure 1: A typical PID Controller for Simplified Plant

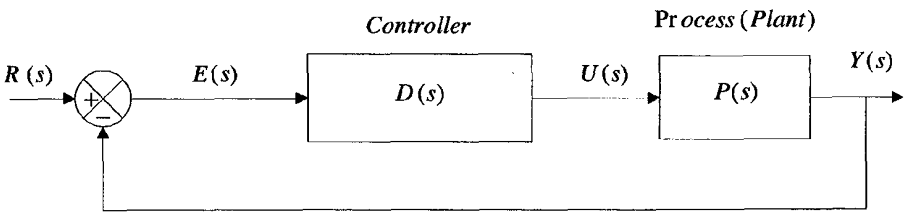

Figure 2: A Generic Structure for a Control System

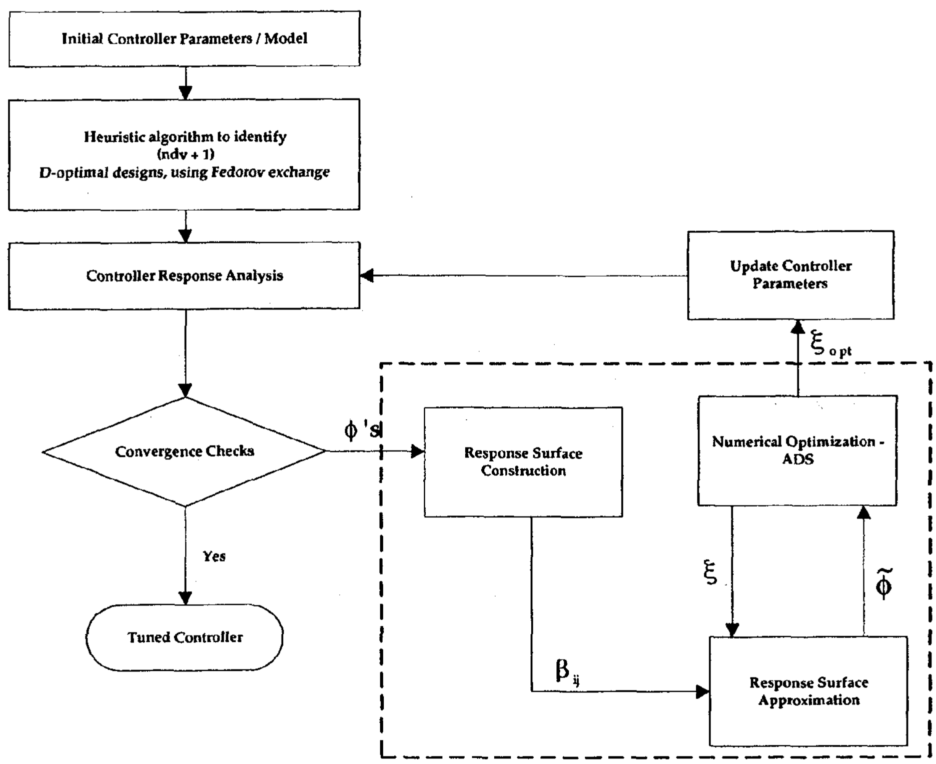

Figure 3: Flow of the Tuning Optimization Procedure

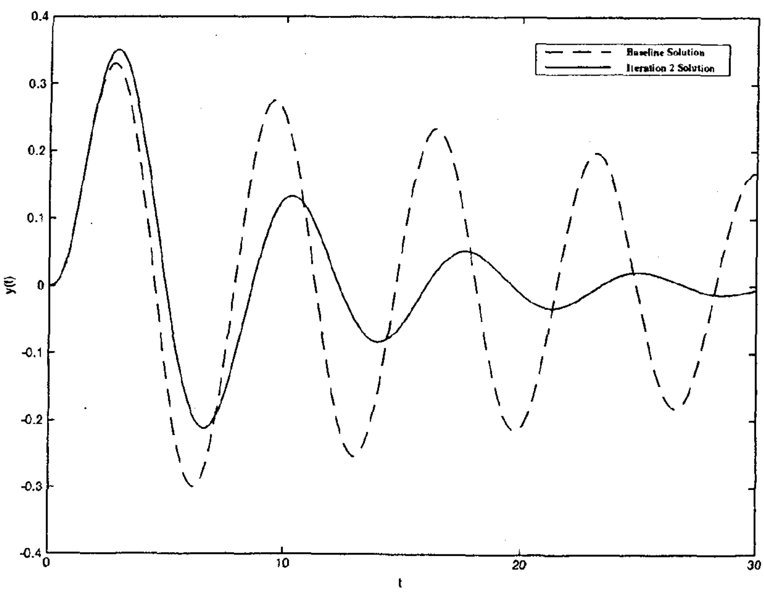

Figure 4: Comparison of Closed-loop Responses

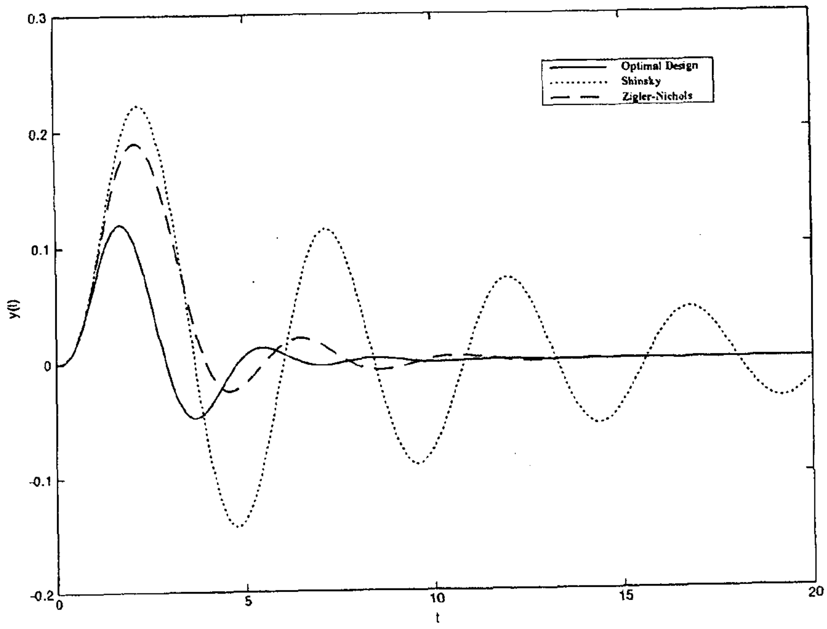

Figure 5: Comparison of Closed-loop Responses

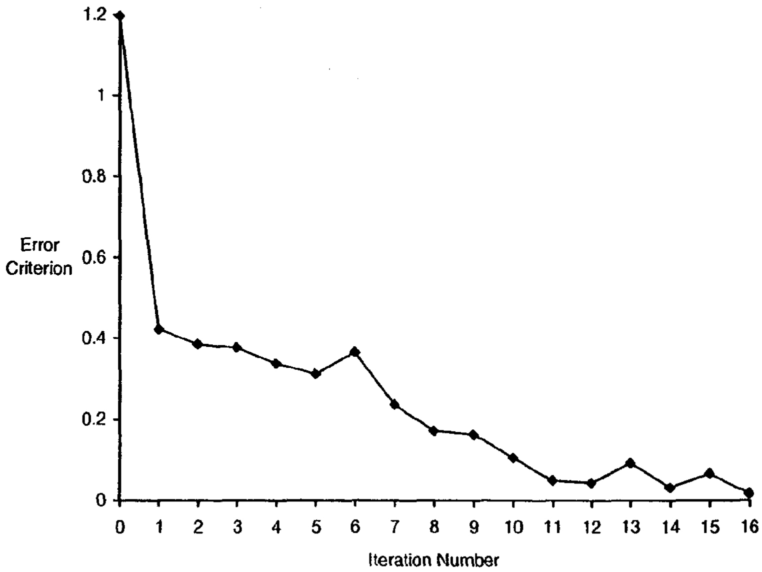

Figure 6: Optimization Iteration history for Error Criterion Φ

Design	K_p	T_i	T_d	Φ
Ziegler-Nichols	4.8000	1.8138	0.4534	0.0558
Shinskey	4.0000	1.2334	0.2902	0.1383
Baseline Design	1.600	0.7255	0.3628	1.1977
Optimal Design	7.0048	0.9178	0.6631	0.0184

Table 1: PID Parameters for Different Designs

Design	Φ	t_{ss}	y_{max}
Ziegler-Nichols	0.0558	14	0.19
Shinskey	0.1383	60	0.22
Optimal Design	0.0184	12	0.12

Table 2: Dynamic Performance Results for Different Designs

Design	First Pair	Second Pair
Ziegler-Nichols	$-1.0957 \pm 0.0205\,j$	$-0.4043 \pm 1.4283\,j$
Shinskey	$-0.0970 \pm 1.2791\,j$	$-1.4030 \pm 0.0487\,j$
Optimal Design	$-0.5278 \pm 1.4492\,j$	$-0.9721 \pm 1.5110\,j$

Table 3: Closed-loop Poles for Different Designs