Available online at www.sciencedirect.com
SCIENCE DIRECT
Control Engineering Practice 14 (2006) 149–163
CONTROL ENGINEERING PRACTICE
www.elsevier.com/locate/conengprac

On the automatic tuning and adaptation of PID controllers

I.J. Gyöngy, D.W. Clarke*

Department of Engineering Science, Oxford University, Parks Road, Oxford OX1 3PJ, UK

Received 16 January 2004; accepted 20 January 2005
Available online 5 March 2005

Abstract

A simple approach to the automatic tuning of PID process controllers is proposed. Like the relay-based autotuner, its objective is to attain a design-point on the Nyquist diagram. By injecting sinewaves and employing a phase/frequency estimator, closed-loop adaptive tuning is possible and there is exact convergence to the design-point without the approximations of describing-function theory. The variant discussed here achieves a required phase margin and imposes a carefully chosen constraint on the controller parameters, leading to consistent behaviour for a wide variety of generic test-cases. A real-life demonstration on a non-linear flow rig is provided.

Keywords: PID tuning; Autotuning; Adaptive control; Process control; Phase/frequency estimator

1. Introduction

Automatic tuning has been extensively studied in literature and is a well-established feature in industrial PI(D) controllers (Hang, Lee, & Ho, 1995; Åström, Hägglund, Hang, & Ho, 1993). Nevertheless, the desire to improve the applicability and robustness of existing autotuners still fuels much research (e.g. Ho, Hong, Hansson, Hjalmarsson, & Deng, 2003; Tan, Ferdous, & Huang, 2002; Yu, 1999).

Since the exact dynamics of the plant is generally unknown, the basic feature of autotuners is some experimental procedure by which plant information is obtained in order to compute the controller parameters. Autotuning techniques can therefore be classified according to this experimental procedure. Three main categories that can be established are step response, periodic excitation and relay-based schemes. The first of these approaches relies on simple open-loop step testing to characterise the process dynamics. The drawback of such methods is their sensitivity to disturbances (due to the experiment being performed open-loop) and, in many cases, the use of a parameterised process model (e.g. first-order/dead-time), which can result in a lack of generality. Alternatively, the open-loop plant may be excited using more complex (binary or multi-level) pseudo-random or multisine signals, with the frequency spectrum tailored to the particular process in question (Braun, Ortiz-Mojica, & Rivera, 2001; Barker & Godfrey, 1999). These type of signals, combined with appropriate signal processing tools, can yield accurate process models or frequency response estimates in which the effects of noise and process non-linearities are suppressed (the signals may also be used to highlight non-linearities).

By contrast, relay feedback tuning schemes use a closed-loop (CL) experiment. The test involves the replacement of the PI(D) controller by a relay with hysteresis, which for a wide range of processes induces a limit cycle in the loop. Describing-function analysis allows the frequency response corresponding to a particular phase shift (set by the size of the hysteresis) to be estimated from the frequency and amplitude of

*Corresponding author. Tel.: +44 1865 273021;
fax: +44 1865 273906.
E-mail addresses: istvan.gyongy@eng.ox.ac.uk (I.J. Gyöngy), david.clarke@eng.ox.ac.uk (D.W. Clarke).

0967-0661/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.conengprac.2005.01.007

this oscillation. The method is simple (requiring no prior information) and fast but, especially in its simplest form, not without limitations. Firstly, the approximations inherent in describing-function analysis can, in the cases of some industrially significant types of processes, lead to errors in the estimate of the frequency point; this in turn can result in less effective PI(D) parameters. Load disturbances also influence the estimation accuracy.

A further disadvantage of the relay autotuner (as well as the aforementioned open-loop schemes) is that it is only suitable for off-line tuning (i.e. no control is provided during tuning). Although modifications have been proposed in which the experiment is performed on-line (e.g. Tan, Lee, & Jiang, 2001), the upset caused by the relay is rather large. The latest developments in relay autotuning also include the simultaneous estimation (with no describing-function approximations) of multiple frequency points, at the cost of increased computation complexity (Hang, Åström, & Wang, 2002; Sung & Lee, 2000).

Here we propose a different autotuning method that retains the PID controller in the loop and uses information obtained by injecting a variable-frequency sinewave into the loop (normally undesirable, but in this case the amplitude can be reduced to match that of noise, so output quality is maintained). The concept of using sinusoidal perturbation for CL controller tuning/adaptation is by no means new. One of the early explorations of the technique was carried out by Smyth and Nahi (1963).[1] In their system, the phase and amplitude response information used to adapt the parameters of a feedback compensation network is acquired via a fixed frequency test-signal. The adaptive PID controller of Glattfelder (1969), on the other hand, applies three (fixed) frequencies and seeks to attain a 'well-shaped' CL phase versus frequency curve. More recent work includes the PI tuner of Crowe and Johnson (2002), which uses two sequentially injected, variable-frequency test signals to achieve phase margin and gain margin specifications. However, despite the long history of perturbation methods, autotuners (or indeed adaptive controllers) based on such techniques have so far failed to gain wide-scale adoption in the process control world. The aim here is therefore to have a tuner comparable to relay methods in terms of its ease of use and speed, yet, thanks to its greater flexibility, giving superior PID tuning. Robustness to noise, set-point changes and disturbances is also an important factor.

In addition to providing 'tuning on demand', the method is also capable of keeping the loop continuously in tune, at the expense of a low-amplitude perturbing sinewave signal being added to the loop's set-point. Although effective adaptive algorithms exist (e.g. Huang, Roan, & Jeng, 2002) which use just the normal operating signals in the loop to monitor the process (obtaining process models, frequency points or characteristics such as damping), these require supervision (Hägglund & Åström, 2000).

1.1. Basics

A schematic diagram of the system is shown in Fig. 1. The objective is to adapt the controller so as to achieve a carefully chosen design-point on the Nyquist diagram.

The key components are phase/frequency and plant gain estimators (PFE, GE), described in detail in (Clarke & Park, 2003; Clarke, 2003). In essence a PFE injects a test sinewave into a system and continuously adapts its frequency ω_1 until its phase shift attains a desired value θ_d (in this case that of the design-point). When applied to the CL system, we can describe this process using the equation:

$$\frac{d\omega_1(t)}{dt} = K_a(\mathrm{Arg}\{G_c(j\omega_1)\} - \theta_d), \tag{1}$$

where K_a is the adaptive gain of the PFE and $G_c(s) = CG(s)/(1 + CG(s))$ is the CL transfer function. In the above equation a 'frozen parameter' assumption has been made, in that the frequency ω_1 (and $C(s)$) are taken to vary slowly enough to allow transients to be neglected (so that the response of $G_c(s)$ to the test signal is a quasi-steady-state sinusoid).

In a concurrent operation, the instantaneous values of the test frequency and the corresponding plant gain $|G(j\omega_1)|$ (as provided by the GE) are used to adjust the controller parameters so that convergence is attained.

Also forming important parts of the tuner, but not shown in Fig. 1, are variable band-pass filters (VBPF) at the inputs of the PFE and GE. These are second-order filters centred on the current value of the test frequency.

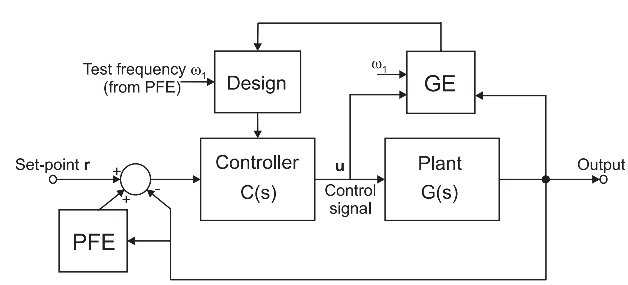

Fig. 1. Schematic diagram of system.

[1] The jet engine fuel control system of Vasu (1957) also deserves a mention. The system tries to maintain an optimal (average) fuel flow (giving maximum thrust, or equivalently, maximum engine pressure ratio) by introducing a slow sinusoidal variation (frequency f_o) in the flow. Coherent detection is used to detect the component at f_o of the resulting fluctuation in the engine pressure ratio. The detected amplitude, which should be zero at optimum fuel flow, is passed through a compensation network to generate the drive signal for the fuel servo.