ISSN 0280–5316
ISRN LUTFD2/TFRT--5583--SE

Automatic Tuning Based on Transfer Function Estimation

Pablo Tapia Lopez

Department of Automatic Control
Lund Institute of Technology
August 1997

Department of Automatic Control
Lund Institute of Technology
P.O. Box 118
S-221 00 Lund Sweden

Document name
MASTER THESIS

Date of issue
August 1997

Document Number
ISRN LUTFD2/TFRT--5583--SE

Author(s)
Pablo Tapia Lopez

Supervisors
Karl Johan Åström

Sponsoring organisation

Title and subtitle
Automatic Tuning Based on Transfer Function Estimation. (Automatisk inställning baserad på skattning av överföringsfunktionen).

Abstract

This report describes a method for automatic tuning of PID controllers in closed loop. A limit cycle is generated through nonlinear feedback from the process output to the controller reference signal. The frequency of this oscillation is near the critical frequency of the loop transfer function. The amplitude and frequency of the oscillation are estimated as well as process gain in steady state. PID parameters are calculated adaptively considering factors such as load disturbances, measurement noise and process dead times. The approach is compared to methods proposed by Åström and Hägglund, Schei, and Woodyatt and Middleton.

Key words

Classification system and/or index terms (if any)

Supplementary bibliographical information

ISSN and key title
0280–5316

ISBN

Language
English

Number of pages
41

Security classification

Recipient's notes

Contents

1. Introduction

There has been a substantial increase of model based controllers in the past 10 years, see e.g. and XXX. The reason for this is that it is possible to obtain controllers that give superior performance over wide operating ranges.
When the requeriments of control system need a behavior very precise, in a wide range of performability conditions and, in certain degrees, it must be adaptive.
Tuning of parameter setting of a controller or optimum control for a long time is an objective in control engineering. The purpose is to set the controller so that optimum performance of the control system is achieved. [Ziegler and Nichols, 1942] introduced a method of tuning analogue, or continuous time, P-, PI- and PID controllers. The method was later extended by [Takahashi et al., 1971] for discrete time controller.
Many systems have properties which vary depending on some other quantity. This makes the control of such systems much more complicated since an applied controller will operate the system nicely in some point of operation but not in other. Adaptive control has been the solution in many of these problems. Adaptive control means that the controller in some way, automatically, adapts to the operating conditions. Often these control algorithms are based on models of the system and utilizes estimations of the properties of the system. [1]
By automatic tuning (or auto-tuning) we mean a method where a controller is tuned automatically en demand from a user [1]. An automatic tuning procedure consists of three steps:

1. Generation of a process disturbance.
2. Evaluation of the disturbance response.
3. Calculation of controller parameters.

One ask myself, why to generate a process disturbance? It is easy, when the process is disturbed, is possible to determine the process dynamic. This can be done in many ways, by adding steps, pulses, or sinusoids to the process input. On the other hand, the automatic tuning techniques can be divided in two kinds: model-based approaches and rule-based approaches. In this paper, it will use the first kind in which a model of process is obtained explicitly, and the tuning is based on this model. But is needed to understand always is not possible to know the model of process due to e.g. long dead times, long time constants, nonlinearities, and inverse responses. Then sensor and actuators may be poorly placed and they may carry bad dynamics.
Autotuning Model-Based methods: The models can be obtained in many ways, these are transient responses, frequency responses and parameter estimation. In this report it will apply frequency response methods, since from output and input signals knowledges, a process model is given in terms of points on the Nyquist curve can be identified on line.
Autotuning in the industry:There are several controllers based on microcomputers, the most important are the Exact of Foxboro(adjusts the parameters of PID controller through transient response analysis of closed-loop system, when it have appreciable changes on the reference, this device has certain similar with Schei's method), or the autotuner of swedish Sattcontrol company(which adjust automatically PID through an appropriate test signal that it generate). On the other hand, the Novatune Asea controller and the Electromax have

the posibility to estimate the parameters of a process model in indefinite way. The two first devices are, perhaps, which perform better philosophy of this report.

2. For Relay Tuning

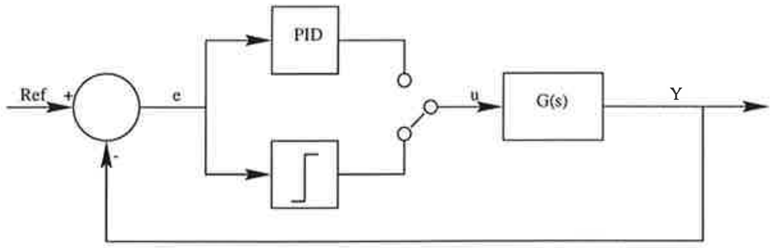

Figure 1 Control scheme corresponding to Åström and Hägglund method

The original Åström and Hägglund method relies on the idea that most system oscillate under relay control, see [2].The tuning procedure starts when the process has been driven to a suitable steady state; then control is switched to the relay, which may or may not have hysteresis, waiting until a permanent oscillation of output process. By measurasing the features of this oscillation, its amplitude and frequency, the describing function approximation lets one to identify a point on transfer functionon Nyquist diagram. As well, PID parameters are computed depending on with the control performance specifications, e.g. the desired phase margin for the closed-loop system. Since the condition on ϕ_m means that the Nyquist diagram of $G_c(j\omega)G_p(j\omega)$ must cross the unit circle in a point with the specified phase margin, a traditional tuning technique is to pick out one point of $G_p(j\omega_{lc})$,where ω_{lc} is the oscillation frequency or also ultimate frequency, and said frequecy ω_{lc}; then, the required phase margin can be obtained by solving the complex equation

$$G_p(j\omega_{lc})G_c(j\omega_{lc}) = e^{j(\phi_m - \pi)} \tag{1}$$

The phase margin is given by

$$\phi_m = \pi + arg G_p(j\omega_g)G_c(j\omega_g) \tag{2}$$

where $\mid G_p(j\omega_g)G_c(j\omega_g) \mid = 1$ .The gain margin is given by

$$A_m = \frac{1}{\mid G_p(j\omega_{lc})G_c(j\omega_{lc}) \mid} \tag{3}$$

therefore is necessary to choose the PID parameters so that $G_p(j\omega)$ is moved to the desired point in fig. 2. If the ultimate frequency (K_u) and ultimate period (T_u) are known, the following simple design formula [3] may be used to achieve a pre-specified phase margin:

$$K = K_u \cos\phi_m \tag{4}$$

$$T_i = \frac{T_u}{4\pi}\left(\tan\phi_m + \sqrt{1 + \tan^2\phi_m}\right) \tag{5}$$

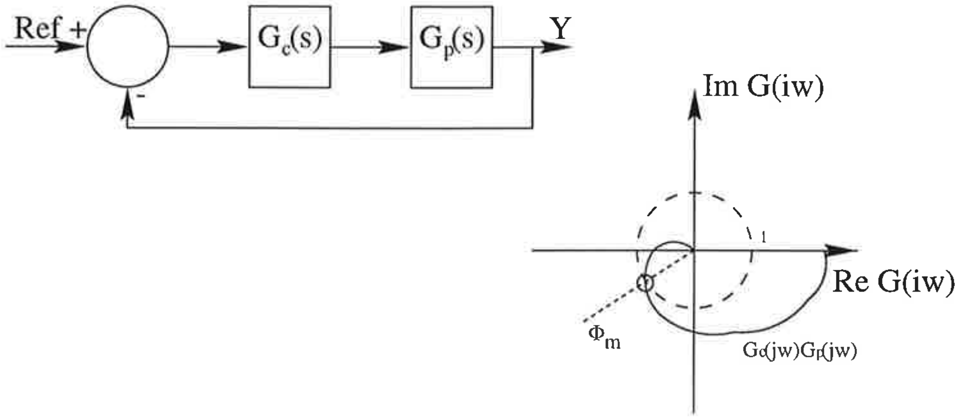

Figure 2 Control scheme in which $G_p(s)$ is process transfer function,$G_c(s)$ is PID Controller and ϕ_m is phase margin required for closed loop system

$$T_d = \frac{T_i}{4} \tag{6}$$

On the other hand, not all the process can be forced to oscillate by a relay, since its describing function lies on the real negative axis, or parallel to it the third quadrant depending on the presence or absence of hysteresis; thus, a relay without hysteresis can be used only if $G_p(j\omega)$ intercepts the real negative axis, while a relay with hysteresis is suitable if it crosses the imaginary negative axis or there are a significant mesurement noise; this is because a "litle" noise can do that the relay "switch", while that a relay with hysteresis solve this behavior with suitable choise of hysteresis width.
Finally, no relay could be used if the phase of $G_p(j\omega)$ is greater than $-\pi/2$ for every ω.

3. Schei's method for Automatic tuning

3.1 Commetary about "A Method for Closed loop Automatic Tuning of PID Controllers"

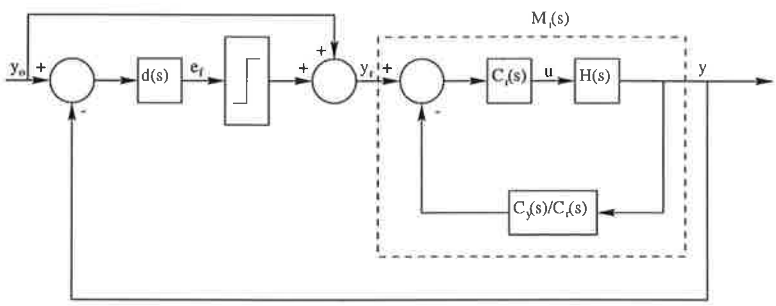

Figure 3 Control scheme corresponding to Schei method

In this study, a method for automatic tuning of PID controller based on the knowledge of the oscillation frequency and its amplitude that it give us a nonlinear-relay is suggested. Until here the base of method is the same that on Astrom and Hagglund autotuner.

Schei defines his system of the follow form:

$$U(s) = C_r(s)Y_r(s) - C_y(s)Y(s) \tag{7}$$

where $U(s)$ is the control signal, $C_r(s)$ is the controller transfer function from reference to controller output, that is

$$C_r(s) = K_p(1 + \frac{1}{T_i s}) \tag{8}$$

and the transfer function from measurement signal to controller output is

$$-C_y(s) = -Kp(1 + \frac{1}{T_i s} + \frac{T_d s}{1 + (T_d/N)s}) \tag{9}$$

On the other hand, Schei works with the complementary sensitivity defined as

$$M(s) = \frac{G(s)}{1 + G(s)}, \qquad G(s) = H(s)C_y(s) \tag{10}$$

where H(s) is transfer function, which server him in order to determinate the frecuency of the limit cycle, ω_{lc}, by phase of M(jω_{lc})=-90, instead of phase of d(jω_{lc})M_r(jω_{lc})=-180, where d(s) is the filter chosen as

$$d(s) = \frac{C_y(s)}{sC_r(s)} \qquad and \qquad M_r(s) = \frac{Y(s)}{Y_r(s)}. \tag{11}$$

This permit, according to Schei, that in process where there is not derivative action in the contoller, like a PI controller, the phase lag of loop transfer function is more negative than the -180 degrees of autotuner and the oscilation frecuency is above the critical frecuency for the control system; which give few phase margin, while his method provide more phase margin.
Contrary to, this method needs more time than the autotuner in order to calculate the oscilation frecuency, ωlc. Since, the more time of experiment, more exact is ωlc; and the less time of experiment, less unstable is the system and less exact is the ω_{lc}.

3.2 Conclusions

The Schei method base the design of PID Controller in knowledge of the period and amplitude of the oscillation generated in the control loop. Therefore, this method is similar to Ziegler-Nichols original method, where the controller parameters are determined of knowledge of G gain critical and critical period. In order to achieve it, Schei put in the control system a low-pass filter in order to work in low frequency. After Schei generate a "suitable pre-tuning" signal, in other words this signal will give, in less possible time, about 850 sec. for the simulated process, the critical frequency.

Few degrees of freedom Though Schei describe four modes of performance in order to work for a wide range of process characteristics, which is good; but the true is that tuning rules are linear (parameters Ti and Td), like Ziegler-Nichols method,

Normal mode:

$T_i = \frac{3.0}{\omega_{lc}}, \quad T_d = \frac{0.75}{\omega_{lc}}, \quad m_s = 1.1,$

$T_i = \frac{3.0}{\omega_{lc}}, \quad T_d = 0, \quad m_s = 1.1,$

Non-minimun phase processes mode:

$T_i = \frac{1.0}{\omega_{lc}}, \quad T_d = 0, \quad m_s = 0.9,$

Integrating processes Mode:

$T_i = \frac{6.0}{\omega_{lc}}, \quad T_d = 0, \quad m_s = 1.3,$

where the parameter K_p has the follow structure in all the modes:

$$K_p = \frac{g_s}{\hat{h}\frac{\sqrt{1+\omega_{lc}^2 \hat{T}_i^2}}{\omega_{lc}\hat{T}_i}}$$

and the convergent is slow, while that with Kappa-Tau method, that is non-linear, the convergent of parameters is more fast. Therefore, this is a poor result, since in Kappa-Tau method is using three parameters to characterize process dynamics instead of two parameters used by Schei and Ziegler-Nichols.

4. Woodyatt-Middleton's method for Automatic tuning

4.1 Commentary about "Auto-Tuning PID Controller Design Using Frequency Domain Approximation"

Introduction This method for auto-tuning PID controllers, as Woodyatt and Middleton say, it based on limited frequency response information, using a weighted least squares approximation to a desired open loop transfer function for a PID controller. The frequency response is calculated by injection of sine waves into the open loop system, following the next steps: "Fourier fitering is used to obtain estimates of the frequency response at a given frequency. Seven frequencies are usually used for identification of the plant.These frequency response experiments can be conducted simultaneously, thus reducing the time that is required for identification. The advantage is that estimates of variance of the noise can also be calculated".

PID Parameter Calculation The method attempts to match the actual open loop Nyquist plot to the desired open loop Nyquist plot by a suitable choice of controller parameters. This is accomplished by means of a least squares minimization. The PID cotroller is given as follow

$$G_{PID} = \frac{K_d s^2 + K_p s + K_i}{s(sT+1)} \tag{12}$$

Notice $G_{ol}(j\omega_i)$ as the desired open loop frequency response at frequency ω_i , and notice $\theta = [K_d \quad K_p \quad K_i]^T$, with K_d, K_p and K_i defined in before equation. Denote $G_{PID}(j\omega_i) = \frac{1}{j\omega_i T+1}\begin{bmatrix} j\omega_i & 1 & \frac{-j}{\omega_i}\end{bmatrix}\theta$, therefore define a vector $\phi(j\omega_i)$ such that the actual closed loop frequency response $G_{cl}(j\omega_i)$ $\left(G_{ol} = \frac{G_{cl}}{1-G_{cl}}\right)$ is given by $\phi(j\omega_i)^T\phi$. Then the objective is to choose the parameters controller such that the loss function J is minimized

$$J = \sum_{i=1}^{n} \mid G_{ol}(j\omega_i) - \phi(j\omega_i)\theta \mid^2 \tag{13}$$

The parameter vector θ which minimizes J can be found by applying the standard least-squares result

$$\hat{\theta} = (\sum_i \phi(j\omega_i)^*\phi(j\omega_i))^{-1}(\sum_i \phi(j\omega_i)^* G_{ol}(j\omega_i)) \tag{14}$$

where B^* denotes the complex conjugate transpose of B. If they assume that $G_{ol}(j\omega)$ and $G(j\omega)$ are symetric, then the solution of $\hat{\theta}$ simplifies to

$$\hat{\theta} = [\sum_i [Re(\phi(j\omega_i))^T * Re(\phi(j\omega_i))$$

$$+Im(\phi(j\omega_i))^T * Im(\phi(j\omega_i))]]^{-1}$$

$$\times \sum_i [Re(\phi(j\omega_i))^T * Re(G_{ol}(j\omega_i))$$

$$+Im(\phi(j\omega_i))^T * Im(G_{ol}(j\omega_i))] \tag{15}$$

Now, the implicit weighting function of the least squares method in [11] emphasizes the frequency regions, and in particular favours low frequencies. According to [10], the identification process gives a uniform absolute error in G(s), proposing a diagonal weighting matrix V:

$$V = diag\{\frac{\sigma^2 \mid G(j\omega_1) \mid^2}{\mid S^*(j\omega_1)T^*(j\omega_1) \mid^2}, \ldots, \frac{\sigma^2 \mid G(j\omega_n) \mid^2}{\mid S^*(j\omega_n)T^*(j\omega_n) \mid^2}\} \tag{16}$$

where σ^2 is the estimate of the noise variance, that is calculated as follow

$$\sigma^2 = \frac{2C}{u_0^2 T} \tag{17}$$

where the auto-correlation function of the noise is $\Psi_{nn}(\tau) = C\delta(\tau)$, T is the time of identification experiment, u_0 is the input sine wave, and the terms $S^*(s)$ and $T^*(s)$ are, respectively, the desired sensitivity and complementary sensitivity functions for the closed loop system. As in [10] it explains, "it is important to note that this matrix provides weightings that are close to the crossover frequency in the closed loop system. This is because $S(j\omega)$ is high pass, with a cut-off near crossover, and $T(j\omega)$ is low pass, with a cut-off near crossover". Therefore the values of the ϕ and G_{ol} matrices used in the calculation of $\hat{\theta}$ are:

$$\phi = \sqrt{V}^{-1} \varphi \tag{18}$$

$$G_{ol} = \sqrt{V}^{-1} [L^*(j\omega_1), ..., L^*(j\omega_n)]^T \tag{19}$$

where $L^*(s)$ is the desired open loop transfer function.

$$\varphi = \begin{bmatrix} \frac{j\omega_1 G(j\omega_1)}{1+j\omega_1 T} & \frac{G(j\omega_1)}{1+j\omega_1 T} & -j\frac{G(j\omega_1)}{\omega_1(1+j\omega_1 T)} \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \frac{j\omega_n G(j\omega_n)}{1+j\omega_n T} & \frac{G(j\omega_n)}{1+j\omega_n T} & -j\frac{G(j\omega_n)}{\omega_n(1+j\omega_n T)} \end{bmatrix} \tag{20}$$

where, usually, this matrix will be composed by seven rows corresponding to seven frequencies used for identification of the plant, which are selected by user.

5. Description of New Tuning Methods for PID Controllers

All the information about this subject it possible to find it in [1]. Is suitable to say that there is not one only design method with which it may carry out the choice of parameters controller. This happen because is easy to find differents types of situations like measurement noise, model requirements, load disturbances response, and model uncertainty. For this reason there is a need a for variety of tuning methods; something techniques require more knowledge about the process, while that other techniques only need little process knowledge.
In new tuning methods it tries to obtain the controller parameters across features to characterize the process dynamics.
These methods can be viewed like an extension of the Ziegler-Nichols methods. The main difference is that in this methods it use three parameters to characterize the process dynamics instead of two parameters that use with Ziegler-Nichols method.
In particular, the methods are based on the step response of process and frequency-response. This chapter will base on the second. Frequency-domain are based on to pick out the parameters ultimate frequency Ku, ultimate period Tu and gain ratio k. These parameters can be obtained from Ziegler-Nichols experiment or applying the original idea of relay-feedback together with the estimation of the static gain of the process.
The process chosen to verify the design method are the corresponding at the Test Batch [1],chapter 5, which are representative for the dynamics of typical industrial processes, for example:

$G_1(s) = \frac{e^{-s}}{(1+sT)^2}$ $\quad\quad T = 0.1, ..., 10$

$G_2(2) = \frac{1}{(1+s)^n}$ $\qquad n = 3, 4, 8$

$G_3(s) = \frac{1}{(1+s)(1+\alpha s)(1+\alpha^2 s)(1+\alpha^3 s)}$ $\qquad \alpha = 0.2, 0.5, 0.7$

$G_4(2) = \frac{1-\alpha s}{(1+s)^3}$ $\qquad \alpha = 0.1, 0.2, 0.5, 1, 2$

If we wish processes with integrator, only is necessary to add it to systems listed above. The starting point to obtain tuning rules is the Ziegler-Nichols method; which has the follow drawback:

a) The responses are too oscillatory.

b) Different tuning rules are required for setpoint response and for load disturbance response; which can be dealt using setpoint weighting to obtain the desired setpoint response.

c) The rules give poor results for systems with long normalized dead time.

d) There is no tuning parameter.

Tuning Rules Controllers for these systems are designed using the dominant pole desig method with maximun sensitivity M_s as the design parameter.Why to use M_s as design parameter? There are two reasons for to use this parameter:

a) Sensitivity to Measurement Noise,[1]in chapter 4, this is very important since in lot of process there are measurement noise of high frequency;two expresions show us this idea. The first describ the transmission from measurement noise to control actions by

$$G_{nu} = \frac{G_c}{1 + G_l} \tag{21}$$

The second describ the transmission from measurement noise to process output by

$$G_{ny} = \frac{1}{1 + G_l} = S \tag{22}$$

where G_p is the process transfer function,G_c is the controller transfer function, $G_l = G_p G_c$ is the loop transfer function, and S is called the sensitivity function. The high frequency gain of PID controller is given by

$$K_{hf} = K(1 + N) \tag{23}$$

Notice that if it do $N = 0$ it has PI control.

b) Sensitivity to Process Characteristics, [1] in chapter 4; since the process may change the controller parameters must be chosen in such way that the closed-loop system is not sensitive to variations in process dynamics.

This parameter is defined as

$$M_s = \max_{\omega > 0} \mid \frac{1}{1 + G_c(j\omega)G_p(j\omega)} \mid = \max_{\omega > 0} \mid S(j\omega) \mid \tag{24}$$

The quantity sensitivity M_s is the inverse of the shortest distance from the Nyquist curve to the critical point -1. M_s has reasonable values from 1.3 to 2.See [4], which physical interpretations is the following. Supposing that in the process there is a sinusoidal disturbance with frequency ω, it happen that the amplitude of open-loop system be a_0, and the amplitude of the controlled system, with sensitivity function S, is $a_0 \mid S(j\omega) \mid$. when the system is feedback it obtein:

- If $\mid S(j\omega) \mid < 1$ reduces the effect of the disturbances.
- If $\mid S(j\omega) \mid > 1$ amplifies the effect of the disturbances.

Therefore, it is very important that when a controller is designed to take into account the frequecies where control signal may be amplified. Others

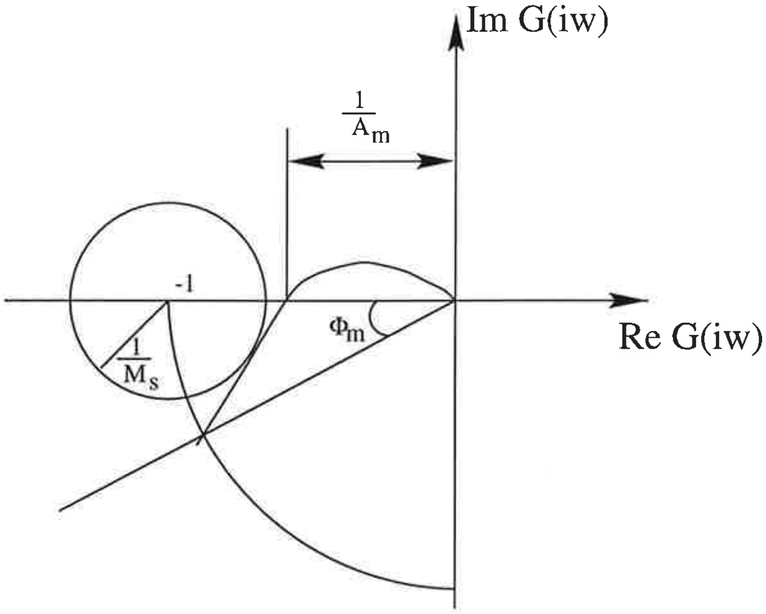

Figure 4 Nyquist intepretation of sensitivity M_s, amplitude A_m, and phase margin ϕ_m

sensitivity mesures are Amplitude margin A_m (2) and phase margin ϕ_m (3). This sensitivity mesures can be relationed related as follow

$$A_m > \frac{M_s}{M_s - 1} \tag{25}$$

$$\phi_m > 2arcsin\frac{1}{2M_s} \tag{26}$$

Typical values of A_m range from 2 to 5, while ϕ_m vary from 30° to 60°. The system will be stable if happen:

a) The gain is greater than the factor $M_s/(M_s - 1)$
b) The gain is lower than the factor $M_s/(M_s + 1)$

Even is possible to limit the stability if there is a nonlinearity characterized as follow

$$xM_s/(M_s + 1) < f(x) < xM_s/(M_s - 1) \tag{27}$$

Dominant Pole Design "What is a dominant pole? The poles closest to the origin."
The key with Dominant Pole Design (DPD) is to find a controller such that the dominant poles of the closed loop system are in specified locations, see [4]. In general it is possible to place as many poles as there are free parameters in the controller.Then the problems may arise when there are more poles in closed loop than free parameters in the controller.
PI control:
In case of a PI or PD controller there are two free parameters, thus two poles must be placed. In particular, two complex conjugated closed loops are specified. It parametrizes the PI controller as

$$G_c(s) = k + \frac{k_i}{s} = k(1 + \frac{1}{sT_i}) \tag{28}$$

The characteristic equation of a system controlled by PI controller is

$$1 + (k + \frac{k_i}{s})G_p(s) = 0 \tag{29}$$

Require that one pole of the closed system is

$$p_{1,2} = \omega_0(-\zeta_0 \pm \sqrt{1-\zeta_0^2}) = \omega_0 e^{j(\pi \mp \gamma)} = \omega_0(-\cos\gamma \pm j\sin\gamma) \tag{30}$$

where $\gamma = \arccos\zeta_0$. Define $a(\omega_o)$ and $\phi(\omega_0)$ by $G_p(\omega e^{j(\pi-\gamma)}) = a(\omega_0)e^{j\phi(\omega_0)}$, thus it introduce this notation, see [4] and [1]

$$A = a(\omega_0)\cos\phi(\omega_0) \tag{31}$$

$$B = a(\omega_0)\sin\phi(\omega_0) \tag{32}$$

set real and imaginary part of (30) to zero, solve for k and k_i as follow, depending on [4]

$$k = -\frac{\sqrt{1-\zeta_0^2}A + \zeta_0 B}{\sqrt{1-\zeta_0^2}(A^2+B^2)} \tag{33}$$

$$k_i = -\frac{\omega_0 B}{\sqrt{1-\zeta_0^2}(A^2+B^2)} \tag{34}$$

with the before definition it get

$$k = -\frac{\sin(\phi(\omega_0)+\gamma)}{a(\omega_0)\sin\gamma} \tag{35}$$

$$k_i = -\frac{\omega_0\sin\phi(\omega_0)}{a(\omega_0)\sin\gamma} \tag{36}$$

and

$$T_i = \frac{k}{k_i} = \frac{\sin(\phi(\omega_0) + \gamma)}{\omega_0 \sin\phi(\omega_0)} \tag{37}$$

Is necessary that k and k_i are reasonable limits, because when ω_0 increases both k and k_i will increase initially, and for values more biggest both parameters will decrease, so that ω_0 must be limited

$$\gamma < -\phi(\omega_0) < \pi \tag{38}$$

PID Control:
PID controllers with derivative filter, where it use the standard form for a PID controller with a derivative filter

$$G_{PID}(s) = Kp(1 + \frac{1}{T_i s} + \frac{T_d s}{1 + (Td/N)s}) \tag{39}$$

the derivative filter is used in order to limit the derivative gain, since the derivative action may give difficults results if there is measurement noise in high frequency.In particular,the amplitude of the control signal can grow arbitrarily.This can solve implementing the derivative term as follow, see [1]

$$D = -\frac{T_d}{N}\frac{dD}{dt} - kT_d\frac{dy}{dt} \tag{40}$$

this differential equation can be represented as follows:

$$D = -\frac{skT_d}{1+sT_d/N}y$$

Typical values of N are 8 to 20. Now the parameters are k, k_i, and $k_d = kT_d$ being the calculates more complicate than PI Control. The parameters ω_0 and ζ_o can be chosen in the same way as for the PI controller.
As our case is PID controller there are three free parameters, thus three poles must be placed.
Suppose It has the before PID controler and the process transfer function $G_p(s)$. Therefore it wants to place poles in

$$p_{1,2} = \omega_0(-\zeta_0 \pm \sqrt{1 - \zeta_0^2})$$

$$p_3 = -\alpha_0\omega_0$$

The real pole at $s = -p_0$ where the constant time $T_0 = 1/p_0$, with p_0 approximately equal to $1/T_i$. This value can be appoximately by the follow inequlity, see [1] chapter 5

$$T_0 > 2\left(L + \sum_{k=1}^{n} T_k\right) = 2(T_{ar} - T_0) \tag{41}$$

where T_0 is the time constant associated with the slowest pole (p_0), T_k are the time constants associated with the remaining poles, and T_{ar} is the average residence time. The inequality is obteined from analysis the root locus of the system.In the other hand, with setpoint weighting, the closed-loop system has a zero at

$$s = -z_0 = -\frac{1}{bT_i}$$

But if it choose b so that $z_0 = p_0$, it achieve that setpoint does not excite to the pole in $-p_0$ on the real axis.This works well when the dominant poles are well damped ($\zeta_0 > 0.7$); in otherwise the choise $z_0 = 2p_0$ give a system with less overshoot. Given α_0, ω_0, ζ_0, and N. Solving

$$1 + G_p(p_i)G_{PID}(p_i) = 0 \tag{42}$$

for $i = 1, 2, 3$ gives an equation of the ninth degree in k if it wants to take N into account.The problem there is not its resolution otherwise many possible solutions, but choosing the filter constant N is possible to solve the problem numerically. Now the parameters are k, k_i, and $k_d = kT_d$ being the calculates more complicate than PI Control. The parameters ω_0 and ζ_o can be chosen in the same way as for the PI controller, see [4]. Then the solution is as follow, remaining previously these functions

$$G_p(\omega e^{j(\pi-\gamma)}) = a(\omega_0)e^{j\phi(\omega_0)}$$
$$G_p(-\alpha\omega_0) = -b(\omega_0)$$

$$k = -\frac{\alpha_0^2 b(\omega_0)\sin(\gamma+\phi) + b(\omega_0)\sin(\gamma-\phi) + \alpha_0 a(\omega_0)\sin 2\gamma}{a(\omega_0)b(\omega_0)(\alpha_0^2 - 2\alpha_0\cos\alpha + 1)\sin\gamma} \tag{43}$$

$$k_i = -\alpha_0\omega_0\frac{a(\omega_0)\sin\alpha + b(\omega_0)(\sin(\alpha-\phi) + \alpha_0\sin\phi)}{a(\omega_0)b(\omega_0)(\alpha_0^2 - 2\alpha_0\cos\alpha + 1)\sin\gamma} \tag{44}$$

$$k_d = -\frac{\alpha_0 a(\omega_0)\sin\alpha + b(omega_0)(\alpha_0\sin(\alpha+\phi) - \sin\phi)}{a(\omega_0)b(\omega_0)(\alpha_0^2 - 2\alpha_0\cos\alpha + 1)\sin\gamma} \tag{45}$$

A Design Procedure:
The procedure to design a PI or PID controller will go several directions: Search ω_0 that maximizes k_i for a given ζ_0. As well as, this ζ_0 must be the smallest that satisfies $| S | \leq 2$ on the sensitivity function. In particular, two values were developed: $M_s = 2$ and $M_s = 1.4$. The inequality $K_{hf} < K_{max}$ determines if it is possible to use a PID controller or if the noise only permits use of a PI controller. The value of N is approximately 10 but is possible to reduce N if K_{hf} is too large. The parameter b is chosen with values between 1 and 0 depending on large overshoot, where the results obtained were the follows, depending on [1]

$$b = \begin{cases} \frac{0.5}{p_0 T_i} & \text{if } \zeta_0 < 0.5 \\ \frac{0.5+2.5(\zeta_0-0.5)}{p_0 T_i} & \text{if } 0.5 \leq \zeta_0 \leq 0.7 \\ \frac{1}{p_0 T_i} & \text{if } \zeta_0 > 0.7 \end{cases}$$

6. Application of new tuning methods to automatic tuning

6.1 Introduction

The new tuning methods developed by [1] go in two ways: Step-response methods and Frequency-response methods. From these two methods, only frequency-response is suitable in order to apply to Automatic Tuning; this is thus for several reasons:

- Step-response is difficult to measure on-line; as well as the information that it obtain when there is not a steady-state on the system response is incorrect in order to tuning.
- Use of the relay method is only applied with frequency-response method, which allows to determine process dynamics.

Therefore frequency response analysis can be used for on-line tuning of PID controllers. How does it achieve? First, it introduce one bandpass filters filters, in the same Schei's way, then is investigated the signal content at different frequencies. After , from this knowledge, a process in terms of points on the Nyquist curve can be identified on-line. pIn particular, the process is characterized by the ultimate gain K_u and the ultimate period T_u.Ultimate gain is defined as the gain that brings the system to the stability limit.It is given by

$$K_u = -\frac{1}{G_p(j\omega_u)}$$

where ω_u is the lowest frequency where the process phase lag is $-180°$. The ultimate period is $T_u = 2\pi/\omega_u$. As a third parameter in this method, it use the gain ratio κ. This parameter is an indicador of how difficult it is to control the process. Processes with a small κ. are easy to control. The difficulty increases with increasing κ, which is given by

$$\kappa = \left| \frac{G_p(j\omega_u)}{G_p(0)} \right| = \frac{1}{K_p K_u}$$

The controller parameters are normalized as $K/K_u, T_i/T_u$ and T_d/T_u, and computed for the differents process in the test batch, using dominant pole design with two values of the design parameter, $M_s = 2$ and $M_s = 1.4$. Then that it does is to look for relations between the normalized controller parameters and the normalized process parameters. This is achieved by plotting the normalized controller parameters as a function of κ. For example:

$$K/K_u = f_1(\kappa)$$

$$T_i/T_u = f_2(\kappa)$$

$$T_d/T_u = f_3(\kappa)$$

$$b = f_4(\kappa)$$

where the functions, $f(\kappa)$, can be approximated by this type of exponential

$$f(\kappa) = a_0 e^{a_1\kappa + a_2\kappa^2} \tag{46}$$

PID Control In [1],chapter 5 it shows the normalized parameters of a PID controller as a function of κ. In the case, $M_s = 2$, the gain of ratio varies from 0.45 to 0.9, the normalized integral time from 0.2 to 0.4. For $M_s = 1.4$ it happens something similar, with the exception that the setpoint weighting varies over a large range in this case. Therefore, it is easy to find, in this case, tuning rules that only depend on two parameters. Always depending on [1], "with tuning rules based on three parameters it is, however, possible to find tuning rules that gives an accuracy of 25%, at least for the test batch".

			$M_s = 1.4$	$M_s = 2$		
	a_0	a_1	a_2	a_0	a_1	a_2
K/K_u	0.33	-0.31	-1.0	0.72	-1.6	1.2
T_i/T_u	0.76	-1.6	-0.36	0.59	-1.3	0.38
T_d/T_u	0.17	-0.46	-2.1	0.15	-1.4	0.56
b	0.58	-1.3	3.5	0.25	0.56	-0.12

Table 1 Tuning formula for PID control obtained by the frequency-response method.The table gives parameters of functions of the form $f(\kappa) = a_0 exp(a_1\kappa + a_2\kappa^2)$ for the normalized controller parameters.

Conclusions This method is done taking into account specifications of load disturbances and mesurement noise, sensitivity to modeling errors, and setpoint response.The load disturbances is the first design criterion that is optimized by the minimizing integrated error. Modeling errors are dealed requiring that the maximum sensitivity be less than a specified value M_s. The user will choose the suitable value, that it will be between 2 and 1.4. The standard value is $M_s = 2$, but smaller values can be chosen if responses without overshoot are desired. Therefore, characterizing the process with three parameters, reasonable tuning rules can be obtained.

6.2 Automatic Tuning

Now it deal with tuning rules obtained,it must mean a method where a controller is tuned automatically.Then well, the method proposed may be regarded as an extension of the [1] method; where they approximately determine the critical point by connecting a relay in the feedback loop from the process output to the process input.

The excitation method. The system is excited by connecting a relay and a linear filter dynamic in a feedback way from the process output to the reference signal as shown in Fig. 5. $M_{sp}(s)$ is the closed loop transfer function from the process output, y, to the controller reference, y_{sp}. A limit cycle will be generated through the nonlinear characteristic of the relay. The filter $d(s)$

is used in the same way that [5],in other words as low-pass filter; from this form is possible to work with the suitable frequencies, even with measurement white noise. The reference signal will vary depending on switch the relay, in steps between $y_o - \Delta y$ and $y_o + \Delta y$, where Δy is the amplitude of relay. Generally, the relay will switch between the values 1 and -1; well, in order to achieve that the relay signal oscillates to suitable frequency, is necessary to select an appropriate test signal y_o. The phase at the oscillation frequency ω_{lc} is as follow

$$\angle d(j\omega_{lc})M_{sp}(j\omega_{lc}) = -180°. \qquad (47)$$

The PID controller is given by the expresion following

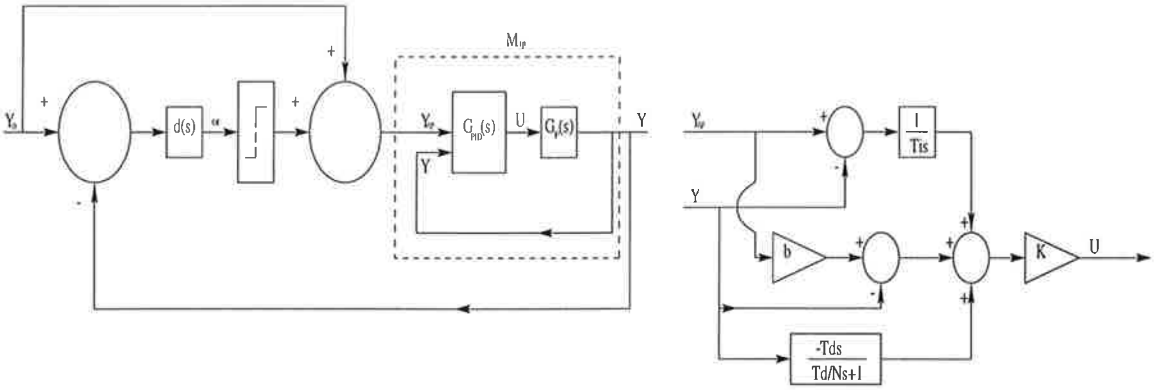

Figure 5 Control scheme corresponding to Kappa method application. The right figure show a detail from PID controller

$$U = K\left(bY_{sp} - Y + \frac{1}{sT_i}(Y_{sp} - Y) - \frac{sT_d}{1 + sT_d/N}Y\right) \qquad (48)$$

where u is the control signal, y is the process output, y_{sp} is the setpoint, K is controller gain, T_i is integral time, T_d is derivative time, and the setpoint b. The high frequency gain N is not determined, since in generally, is given a constant value with a range from 7 to 15 approximately.
Two values of ***sensitivity*** it will have, $M_s = 2$ and $M_s = 1.4$; that it will choose one or other value depending on the design specifications. For instance, the first value it will apply to system where the robust conditions were few demanding. The second value it will use to where the robust conditions are necessaries, like to load disturbances or to mesure noise.
The filter $d(s)$ is chosen to be

$$d(s) = \frac{d_1(s)}{sd_2(s)} \qquad (49)$$

where $d_1(s)$ y $d_2(s)$ are the transfer functions follow

$$d_1(s) = Kp(1 + \frac{1}{T_is} + \frac{T_ds}{1 + (Td/N)s}) \qquad (50)$$

$$d_2(s) = Kp(1 + \frac{1}{T_is}) \qquad (51)$$

Filter parameters are the same that the controller, so that this parameters are updated in same time that control parameters.

PID Parameter Calculation. Control parameters are adjusted iteratively which will be improved during the tuning process. It is assumed that a conservative and stable controller is prior to the tuning. In this case, the main tuning relations in eqns. 52 give four equations with three parameters.

The frequency response is method applied, and following the same way that Ziegler and Nichols, the process is characterized by the ultimate gain K_u, the ultimate period T_u, and the gain ratio

$$kappa = \mid \frac{G_p(i\omega_u)}{G_p(0)} \mid = \frac{1}{K_p K_u}$$

where it assumes that K_p always is in steady state, and thus is a constant value. In the same way that [1, 5], is possible to calculate the amplitude and frequency oscillation from the system by peak and zero-crossing detection of the signal $o_f(t)$ in Fig 5. where the oscillation frequency is ω_{lc} and its amplitude is $\hat{h}_i$;then it will approach parameters by

$$T_u \approx \frac{2\pi}{\omega_{lc}}$$

$$K_u \approx \frac{1}{\hat{h}_i}$$

$$\omega_u \approx \omega_{lc}$$

The number of oscillations used for estimation of ω_{lc} is much smaller than the method [5], which implies that the duration of the tuning experiment is shorter.

With all this information, the PID parameters calculation is depending on tuning formulas for PID or PI control obtained by the frequency-response method

$$K = \hat{K}_{u,i} a_0 e^{a_1 \hat{\kappa}_i + a_2 \hat{\kappa}_i^2}$$

$$T_i = \hat{T}_{u,i} a_0 e^{a_1 \hat{\kappa}_i + a_2 \hat{\kappa}_i^2}$$

$$T_d = \hat{T}_{u,i} a_0 e^{a_1 \hat{\kappa}_i + a_2 \hat{\kappa}_i^2}$$

$$b = a_0 e^{a_1 \hat{\kappa}_i + a_2 \hat{\kappa}_i^2} \tag{52}$$

The tuning is finised when is appreciable converge PID parameters to ones certain values.

6.3 Spectral Density of Control Signal

In this section it deals to verify that control signal finds in ones determinates frequencies.In order to which is necessary to make a spectral density analysis.The spectral density of energy is defined as

$$E_{xx}(j\omega) = X(j\omega)X^*(j\omega) \tag{53}$$

and the cross spectral density between two signals x and y, also with Fourier transform is defined as

$$E_{xy}(j\omega) = X(j\omega)Y^*(j\omega) \tag{54}$$

According to the Parseval relations, as it shows

$$\int_{-\infty}^{\infty} x(t)y^*(t)dt = \frac{1}{2\pi}\int_{-\infty}^{\infty} X(j\omega)Y^*(j\omega)d\omega \tag{55}$$

such that it verifies: the signal energy is independent of the choise of representation in time or frecuency, respectively. According to the Plancherel theorem the product of two Fourier transform equal the Fourier transform of the convolution of the two time domain signals.Then, the energy cross spectrum $E_{xy}(j\omega)$ is

$$E_{xy}(j\omega) = X(j\omega)Y^*(j\omega) = F\{x\} \cdot F\{y^*\} = F\{\int_{-\infty}^{\infty} x(t)y^*(t-\tau)dt\} = F\{e_{xy}(\tau)\} \tag{56}$$

See [6]. In this case, it calculates spectral density of control signal, depending

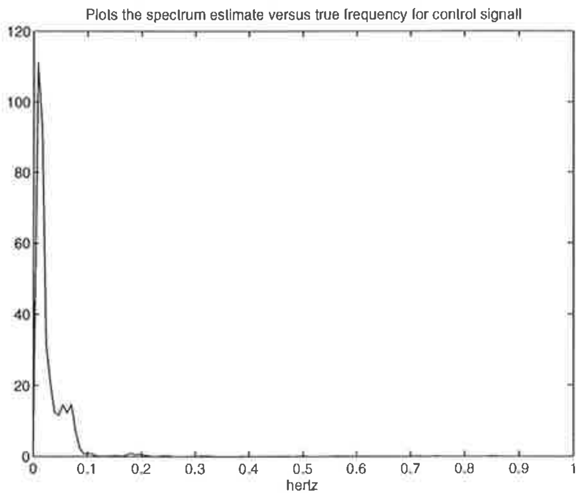

Figure 6 Plots the spectrum density of control signal

on the following experiment:

- With the autotuner presented in this report, it has simulated the tuning of process shown in [5] and that correspond exactly to the process model:

$$H(s) = \frac{1}{(1+30s)(1+15s)}e^{-15s} \tag{57}$$

Then, a PI controller is tuned as shown in Fig. 7 From this experiment,in Fig. 6,it plots this spectral density of control signal, which shows that the biggest quantity signal is to low frequency; so that some appropriates frequencies of the control signal are generated automatically, and the result is satifactory.

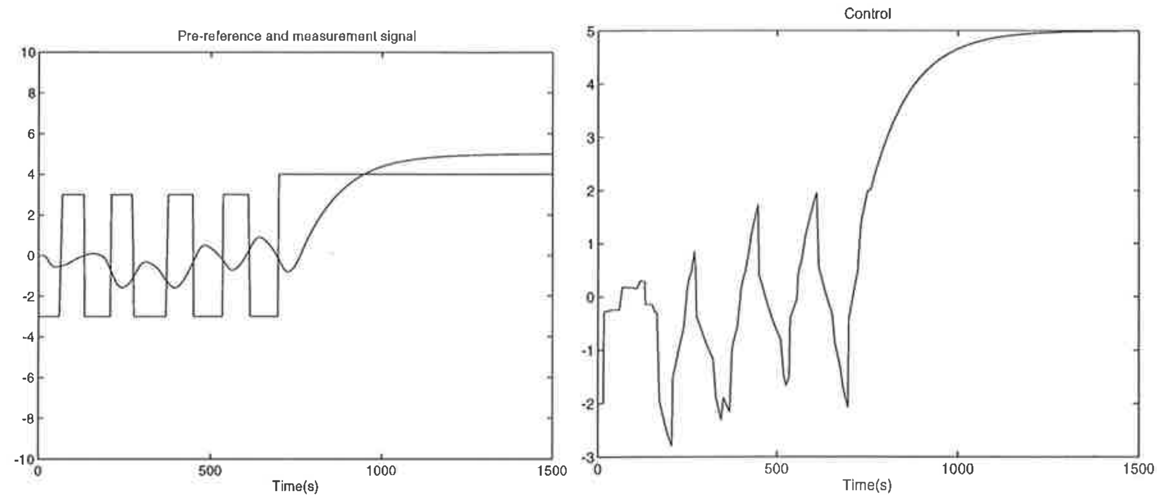

Figure 7 Autotuning of PI controller.

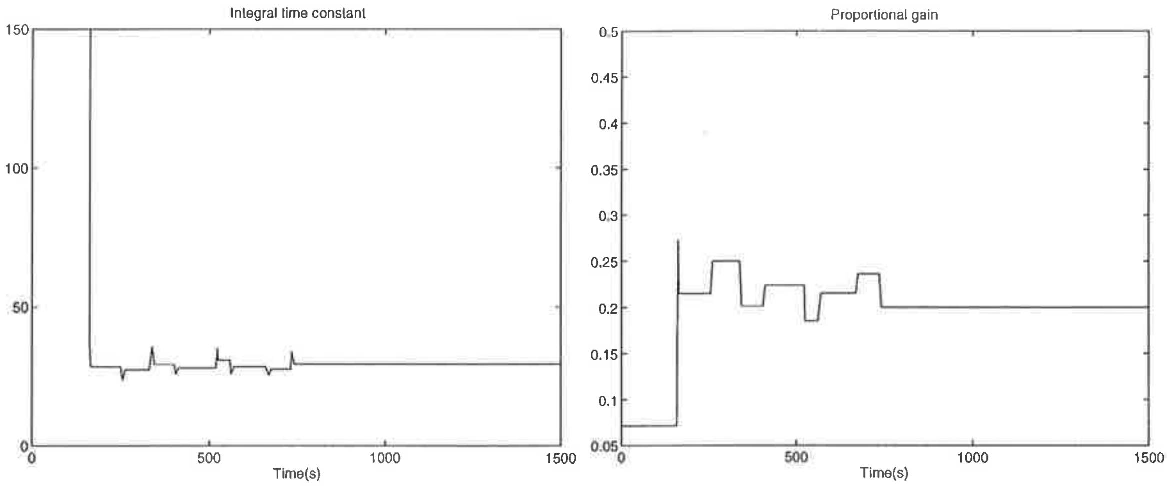

Figure 8 Integral time and proportional gain during for the experiment in Fig. 7

On the other hand, in the Fig. 8 it notices the values of PI parameters, integral time and proporcional gain. Is necessary to notice that it produces some desviations (or small peaks) in the convergy of the sames. In order to explain these desviations, is mainly to analyse the filter signal, $d(s)$, (in the Fig. 5) and consequent effect that it produces over relay. Then is possible to analyse these small peaks. Knowing that for the tuning, it mesures the oscillation frequency of Of signal, in Fig. 5, and as well as, for each zero-crossing of signal it updates the value of frequency, remaining obviously, the two zero-crossing befores. Therefore, there is two important questions:

1. It produces the peak as transient situation, forced for the new updating of frequency.
2. In the experiment, in the first mesure of frequency there is not desviations, after are appreciable some bigs peaks and others smaller.

Why does it happen?
First, in the first frequency there is not desviation because is one frequency nearby at ultimate frequency of system, and it has one period about 300

seconds. After, are appreciables some bigger peaks than the rest, this is because the frequecies at that it happens are smaller than ω_u (about 340 seconds of period); whereas the smallest peaks produce at frequencies nearby a ω_u. These explications are shown ploting in Fig. 9, where it shows the corresponding on the filter signal and integral time, each zero-crossing, since for each zero-crossing there is one desviation. Therefore, is not necessary with this method to generate this large test signal, since with the first mesure of frequency it obtains the true parameters.

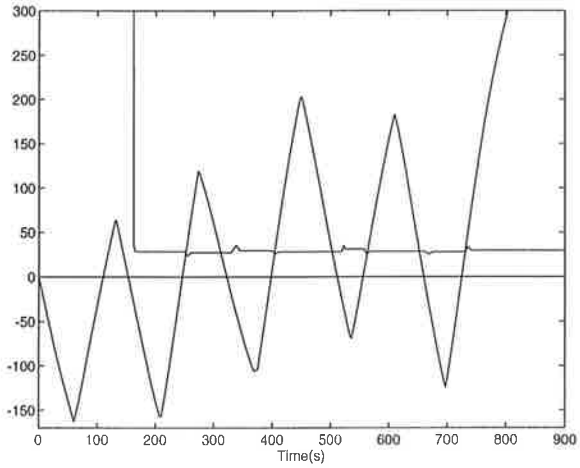

Figure 9 Relation between Integral time calculation and the filter signal

7. Comparison with existing techniques

The technique proposed in this paper will compare with the method described in [10].
It has chosen to compare with this technique because is last work done about Automatic tuning; which at the same time does a comparison with Kappa-Tau method described in [1]. Then, following the same steps than in [10], it simulats the control of the process $G(s) = \frac{1}{(s+1)^3}$. The method [10] develops the calculation of PID parameters based on frequency domain approximation using least squares methods, in particular use a weighted least squares approximation to a desired open loop transfer function. In the paper [10] it compared these technique with Kappa-Tau method choosing as sensitivity $M_s = 2.0$.In the simulation of this process have been taken into account two considerations.The first is the load disturbances and the second is noise mesurement in high frequency. Obviously the result of these simulation was not the suitable, since when the sensitivity is equal to 2.0 it amplifies a disturbance and is less robust design.Therefore is more practice to use this value when there are not disturbance. But in [1] it describes the calculate of PID parameters for $M_s = 2.0$ and $M_s = 1.4$; being suitable the last value in the presence of noise and disturbance.Then, the controller developed in method [10] was the follow

$$U(s) = 1.11\left(R(s) - Y(s) + \frac{(R(s)-Y(s))}{1.95s} + \frac{0.61s}{0.31s+1}(R(s) - Y(s))\right)$$

whereas the controller that is proposed depending on [1] is as follow

$$U(s) = 2.5\left(0.52R(s) - Y(s) + \frac{(R(s)-Y(s))}{2.2s} - \frac{0.56s}{0.28s+1}Y(s)\right)$$

where $M_s = 1.4$ and $N = 2$. In this simulation is given the noise power to 0.003 with a frequency of 628 radianes/s, and the disturbances is given with a amplitude of 0.01 and period of 1 s. The plot of the step responses of the systems controlled using the [10] technique, and the technique proposed in [1] are shown in Fig. 10 In both these simulations there is one changes in the reference signal at $t = 8s$. In an analysis of sensitivity, the proposed technique works with a $M_s = 1.4$ whereas techniques described in [10] has one of 1.35. This means that two controllers have a similar robustness against plant variations and againist noise. The controller designed with [1] method has almost the same rise time and settling time than the other method, depending on it notices in Fig. 10. But it is truly important to notice, that when closed loop system has an Overshot, depending on it notices in Fig. 10,the system developed with [1] method has more better behavior than the other method; since the other system oscillates more than proposed system.

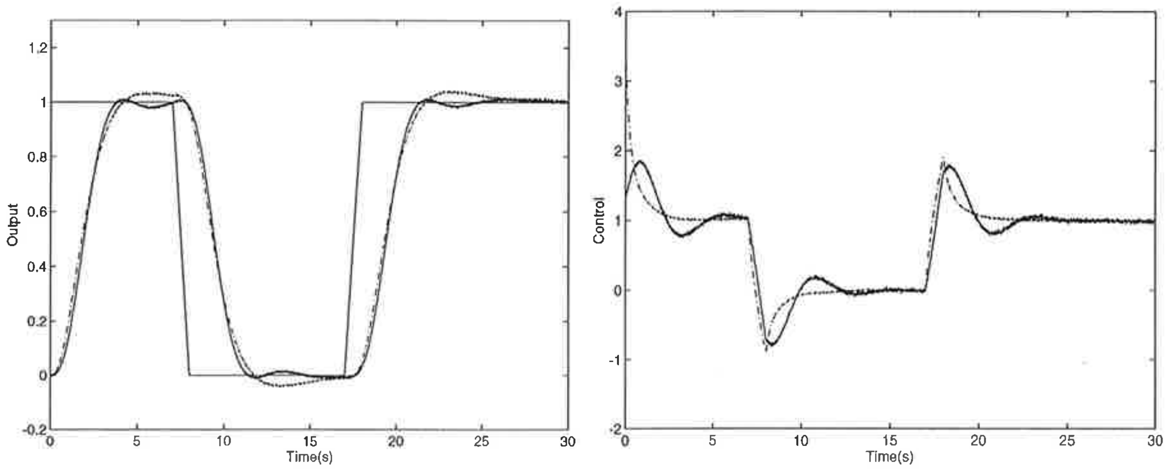

Figure 10 The output signal and the control signal using the alternative technique, marked with dashdot, and the technique proposed, marked with full line.

In the other hand, from point of view of Automatic Tuning the alternative method does not show information, for example the convergence time of parameters; but in this paper in the section before, it applies [1] method to automatic tuning.

8. On-line measurement of process transfer function

8.1 Introduction

There are several conditions in order to obtain good experiments results on the transfer function identification. The choice of input is obviously of great importance in order to carry out the identification. Another important aspect is the presence of controllers in closed-loop operation and the complexity of

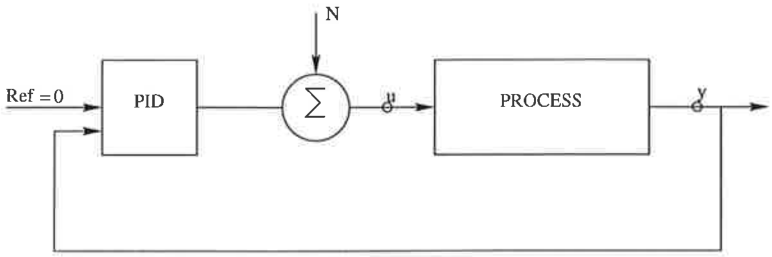

Figure 11 Identification scheme of process transfer function

such controllers, see [6].
The identification, on the before figure, of the transfer function between the control input u and the observation y of a system in closed-loop operation is called direct identification. Then the following step is to introduce pertubation signals at A, measure u and y and determine process transfer function. For instance, a suitable input spectrum to be used in system identification for the purpose of control system design is such that there is much input energy, i.e., $| U_k |^2$ is large, for frequencies ω_k around the bandwidth frequency of the investigated system. Then the form of the input is sometimes determined by the method, e.g., with sinusoids as used in frequency response analysis, step functions and impulse in transient analysis, or PRBS in correlation analysis, see [6]. Amplitude of the input is chosen as a trade-off signal-to-noise ratio and nonlinearities. A large amplitude should be avoided when fitting a linear model so that the test signal does not enter a nonlinear region of behavior. However, the amplitude should be large enough to

ensure a good signal-to-noise ratio. A statistical motivation for a large amplitude is that the attainable parameter accuracy is

$$\sigma^2 \quad \alpha \quad 1/input \quad power \tag{58}$$

A good rule of thumb for a minimum amplitude to choose is that the effect of the input on the output in a diagram need at least be perceptible to the eye. Frequency domain characteristics of test signals is that the input should be sufficiently exciting with an autospectrum $S_u(i\omega)$ that is not "too small". In addition, the estimated input spectrum $\hat{G}_u(i\omega)$ must not be "too small",see [6].

$$\hat{G}_u(i\omega) = \frac{\hat{S}_{yu}(i\omega)}{\hat{S}_{uu}(i\omega)} \tag{59}$$

8.2 Identification of interesting frequencies

In same way that obtained by relay tuning, now is necessary to choose this interesting frequencies, depending on Fig. 12

Why are interesting these frequecies?

There are several reasons that explains it, but before to get started is necessary to make an analysis about scheme of Fig. 11, where the system it can be given on terms of transfer function as follows

$$Y(s) = \frac{G(s)}{1 + G(s)G_{PI}(s)} N(s) = \frac{1}{G_{PI}(s)} \frac{G(s)G_{PI}(s)}{1 + G(s)G_{PI}(s)} N(s) \tag{60}$$

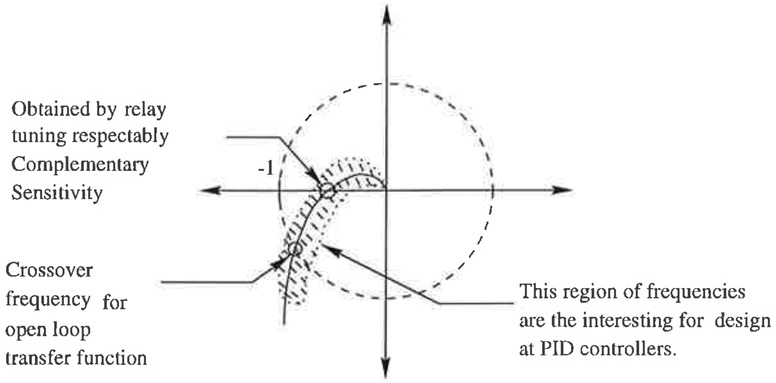

Figure 12 Nyquist diagram at $G(s)G_{PI}(s)$

where the complementary sensitivity was as follows

$$T(s) = \frac{G(s)G_{PI}(s)}{1 + G(s)G_{PI}(s)} \tag{61}$$

Then for instance, it obtains with PI controller the following Bode relations, where the transfer function of process and PI controller are the following

$$G(s) = \frac{1}{(1+s)^3}, \qquad G_{PI}(s) = \frac{1.09(1 + s1.6)}{s1.6} \tag{62}$$

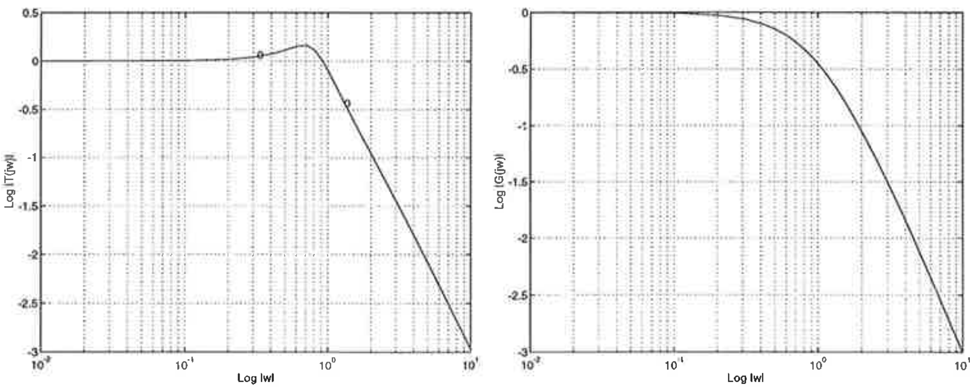

Figure 13 Modul frequency response of Complementary sensitivity and Process

In the Fig. 13 are appreciables certain important frequencies. On the process response it shows the cutoff frequency, which corresponds to dominant pole, and therefore the smaller pole, in this particular case this frequency is 0.51 rad/sec. On the Complementary sensitivity response it shows the bandwidth of system that is on 1 rad/sec., and a resonance peak that is on 0.6 rad/sec. On the other hand, in Fig. 12 is necessary to notice two points, the first it obtains when the system is tuned by relay tuning depending on Fig. 5, and according to the phase at the frequency oscillation ω_{lc} that was as follows

$$\angle d(j\omega_{lc})M_{sp}(j\omega_{lc}) = -180^\circ.$$

the second point correspond to crossover frequency of open loop transfer function.

Therefore, it deals to find the relation between this frequencies and to give answer to the initial question. The interesting frquencies have a range from dominant pole of process until the bandwidth of Complementary sensitivity. This region ,see Fig. 12, is determined by a short width.
Then, with this dates it can specify somethings:

1. Near of this frequencies is system bandwidth, which determines mainly three things:
   a) For a high bandwidth implies that rise time decreasing.
   b) For a high bandwidth implies that the system will not filter in high frequency noise.
   c) It is necessary to take one agreement between one high value of bandwidth and other lowest.
2. Near of this frequencies is resonance frequency, which is very important because if poles of T(s) can aproximate by two dominant complex poles, like the expresion following

$$T(s) = \frac{\omega_0^2}{s^2 + 2\xi\omega_0 s + \omega_0^2}$$

   where ω_0 is natural frequency and resonance frequency is as follows

$$\omega_r = \omega_0\sqrt{1 - 2\xi^2}$$

3. The frequency response of open loop transfer function determines:
   a) The gain margin and phase margin in order to decide relative stability.
   b) The crossover frequency which together with modul frequency response can be used in order to obtain the plotting slope.

   where the slope is defined as follows

$$Slope = \frac{dlog|L(i\omega)|}{dlog\omega}$$

   Then the slope shows the capacity of system in order distinguist between signal and noise. Therefore, a high speed response has a great bandwidth. It notices In the Fig. 14 that the crossover frequency is in the region of interesting frequencies.

How to identify this frequencies?
There are two possibilities:

1. Relay tuning.
2. Frequency estimation of transfer function.

The first was analyzed in before chapters. Respectably the second, it will tackle more forward, but the fundamental idea is to identify the dominant pole frequency of transfer function; since this frequency is into interesting frequencies region.
How to tackle the PID controller design?
There are a great many methods, though mainly there are two design lines:

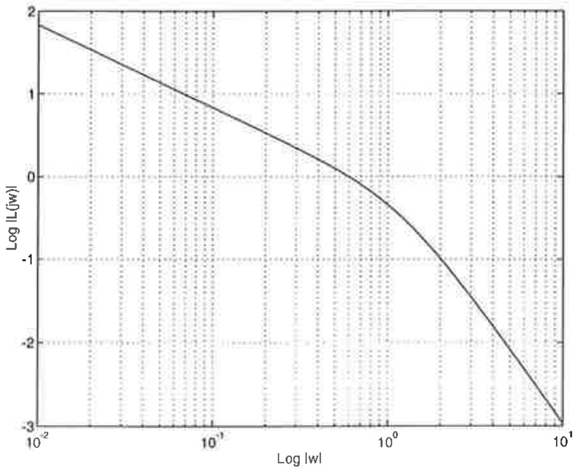

Figure 14 Frequency response of open loop transfer function

Slope	$\mid L(i\omega) \mid$	ω
-2	2	$\omega_{gc}/sqrt2$
-2	0.5	$sqrt2\omega_{gc}$
-1	2	$\omega_{gc}/2$
-1	0.5	$2\omega_{gc}$
-0.5	2	$\omega_{gc}/4$
-0.5	0.5	$4\omega_{gc}$

Table 2 Table approximate that shows the relation between the Slope and the bandwidth

1. PID Controllers: Theory, Design, and Tuning. See [1].
2. "A frequency dominant design method for PID controllers". See [11]. This method is based on frequecy response identification; in particular is based on a least squares fit of the actual to the desired open-loop, using weighting function, where the effect of this weighting function is that it will tend to emphasize the frequency regions where $\mid W(j\omega_i) \mid^2$ is large and de-emphasize the frequency regions where $\mid W(j\omega_i) \mid^2$ is small, being $W(j\omega_i) = \frac{G(j\omega_i)}{j\omega_i}$; which implies that this author wants to approximate at this frequencies.

8.3 Identification of process transfer function

The correlation method. The correlation method is a method for frequency response analysis, which can improve significantly the determination of amplitude ratios and phase shifts in the presence of disturbances. A scheme is shown on the follows figure The identification principle is based on the following idea: Assuming that the identification process is a linear time invariant dynamic system it can be described by some weighting function g(t). The Laplace Transform $L\{\cdot\}$ of the weighting g(t) provides a transfer function $G(s) = L\{g(t)\}$ that relates the Laplace-transformed input $U(s) = L\{u(t)\}$

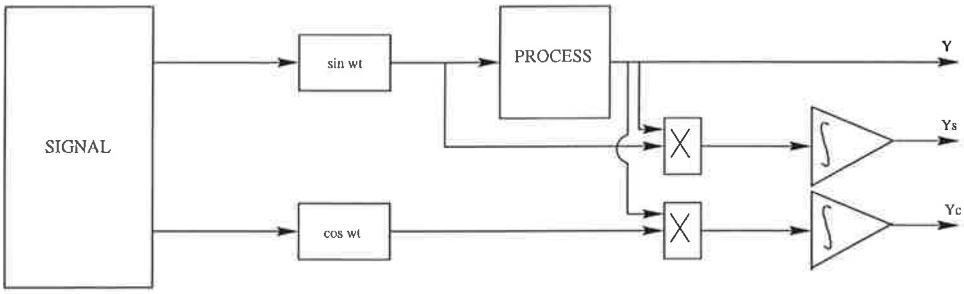

Figure 15 Frequency response analysis

with the output $Y(s) = L\{y(t)\}$,See [6].

$$y(t) = \int_0^\infty g(\tau)u(t-\tau)d\tau \tag{63}$$

$$Y(s) = G(s)U(s)$$

Then, the process is excited by a sine wave from the signal, depending on Fig. 15, which also has a synchronized cosine output. The process output is multiplied by the sine and cosine signals and the products are integrated. The signal channels Y_s and Y_c are called as the sine channel and the cosine channel, and Y(t) denote the output of process.
Assume that it can not take into account the transient from initial conditions and only consider the input-output response, as well as this method should not applied when there are larges time constants, and then it must wait to steady state in order to do the experiment;therefore if the process is stable and if the disturbances are ignored the steady state output is given by

$$y(t) = \mid G(j\omega) \mid u_1 sin(\omega t + \phi(\omega)) \qquad \phi(\omega) = argG(j\omega) \tag{64}$$

being the input signal to the process the following

$$y(t) = u_1 sin(\omega t) \tag{65}$$

and G is the process transfer function. The outputs Y_s and Y_c of sine and cosine channels are given by, see [7]

$$Y_s(t) = \int_0^T y(t)sin(\omega t)dt = \int_0^T y_0 sin(\omega t)sin(\omega t + \phi)dt =$$

$$= \frac{Ty_0}{2}cos\phi - \frac{y_0}{2}\int_0^T cos(2\omega t + \phi)dt \tag{66}$$

$$Y_c(t) = \int_0^T y(t)cos(\omega t)dt = \int_0^T y_0 = cos(\omega t)sin(\omega t + \phi)dt =$$

$$= \frac{Ty_0}{2}sin\phi + \frac{y_0}{2}\int_0^T sin(2\omega t + \phi)dt \tag{67}$$

Therefore, if the integration time is selected in such a way that ωT is a multiple of π it get

$$Y_s(T) = \frac{1}{2} y_0 T cos\phi = \frac{T}{2} u_o Re\{G(j\omega)\}, \qquad \omega T = \pi, 2\pi, ... \tag{68}$$

$$Y_c(T) = \frac{1}{2} y_0 T sin\phi = \frac{T}{2} u_0 Im\{G(j\omega)\}, \qquad \omega T = \pi, 2\pi, ... \tag{69}$$

If the integration time T is a multiple of π/ω the outputs of the integrators will be proportional to the real and imaginary parts of transfer function. From these quantites the magnitude and the phase shift of $G(j\omega)$ is computed as follows

$$\mid G(j\omega) \mid = \frac{2}{T u_0} \sqrt{Y_s^2(T) + Y_c^2(T)}$$

$$arg G(j\omega) = arctan \frac{Y_c(T)}{Y_s(T)} \tag{70}$$

The correlation method can be viewed as filtering y(t) through a band-pass filter with center frequency ω and bandwidth proportional to 1/T. The experiment is repeated for a number of frequencies ω so that a Bode-diagram is can be drawn, see [7] and [6]

Sensitivity to disturbances There are several kinds of disturbances that can be observed

- White noise acting on the output
- Sampling interference
- Unmodelled nomlinearities
- The presence of higher order harmonics

Then, with this factors it assume that the output y is changed by a disturbance v so that

$$Y(t) = \mid G(j\omega) \mid u_0 sin(\omega t + \phi(\omega)) + v(t), \qquad \phi(\omega) = arg G(j\omega) \tag{71}$$

The sensitivity to this type of disturbance give a errors that can be computed as follows

$$\Delta Y_s = \int_0^T sin(\omega t) v(t) dt$$

$$\Delta Y_c = \int_0^T cos(\omega t) v(t) dt \tag{72}$$

Therefore, the error due to a constant disturbance is zero. Another interesting case of low disturbance sensitivity is white noise and bandwidth limited noise, which is shown on the follows expresion where the relative error of the transfer function is estimated

$$\frac{\mid \Delta G(j\omega) \mid}{\mid G(j\omega) \mid} = \sigma_v \sqrt{\frac{\omega}{2k\omega_c}} \tag{73}$$

where σ_v^2 is the variance of the output disturbance v, ω_c is the bandwidth of the disturbance, and k is the number of full periods of the sinusoid completed during the measurement time, see [6].

Simulations The process transfer function purpose for this simulation is as follows

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

Experimental results with estimates of $\mid G(j\omega) \mid$, $\phi(\omega)$ over some frequency range can be presented graphically in the form of Bode, Nichols and Niquist diagrams. The form chosen is Bode diagram since this contains the gain and phase response.

Then the input is chosen as a sinusoid with unit amplitude and a range of frequenies between 0.1 rad/sec. and 10 rad/sec., where each frequency is used for each experimentused for each experiment. The estimates of the obtained

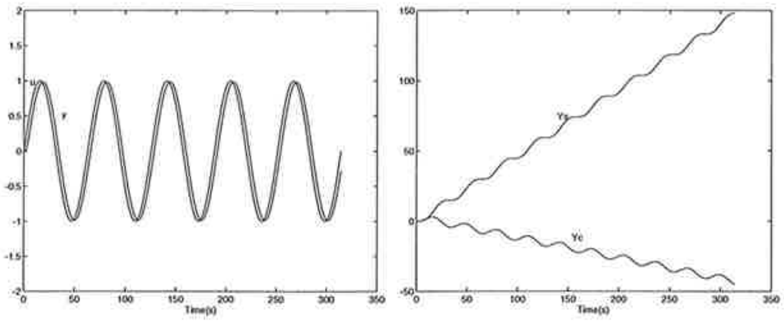

Figure 16 Results of simulation of simulink based on the correlation method.On first graphs it shows the process input u, and proces output y. On the second graphs it shows the outputs of the correlation channels, for $\omega = 0.1$, and period T equal to 5.

are give in Table 3

Experimetal results with these estimate of $\mid G(j\omega) \mid$, $\phi(\omega)$ allows to plot the following Bode diagram

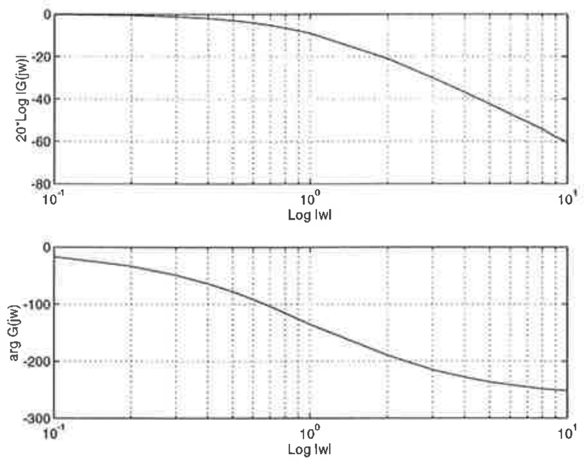

Figure 17 The graph shows the estimated magnitude and phase angle of the process transfer function.

On the Table 3, there is two types of results: The estimated on process transfer function identification, and the exact values. Knowing that the chosen process process transfer function is

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

ω	Estimated values Y_c	Y_s	Argument	Modul	Exact values Argument	Modul
0.1	-45.0099	148.0725	-16.9078	0.9852	-17.1318	0.9852
0.2	-40.5403	61.9724	-33.1914	0.9429	-33.9298	0.9429
0.3	-34.5890	30.3360	-48.7479	0.8787	-50.0977	0.8787
0.4	-28.1581	13.9672	-63.6173	0.8004	-65.4042	0.8004
0.5	-21.9708	4.7518	-77.7962	0.7155	-79.6952	0.7155
0.6	-16.5038	-0.3533	-91.2265	0.6305	-92.8913	0.6305
0.7	-11.9823	-2.9379	-103.77	0.5498	-104.9761	0.5498
0.8	-8.4486	-4.0003	-115.3369	0.4761	-115.9794	0.4761
0.9	-5.8099	-4.1987	-125.8553	0.4107	-125.9616	0.4107
1.0	-3.9072	-3.9482	-135.2942	0.3536	-135.0000	0.3536
2.0	-0.5602	-0.4238	-189.1240	0.0894	-190.3048	0.0894
3.0	0.01809	-0.0261	-214.7213	0.0317	-214.6952	0.0316
4.0	0.01060	-0.0096	-227.8445	0.0143	-227.8913	0.0143
5.0	0.0063	-0.0042	-236.3137	0.0075	-236.0702	0.0075
6.0	0.0039	-0.0021	-241.7214	0.0044	-241.6130	0.0044
7.0	0.0026	-0.0012	-245.1809	0.0028	-245.6097	0.0028
8.0	0.0017	-0.0007	-248.4921	0.0019	-248.6250	0.0019
9.0	0.0012	-0.0004	-250.1023	0.0012	-250.9794	0.0013
10.0	0.0009	-0.0003	-251.5634	0.0009	-252.8682	0.0010

Table 3 Estimates of the process transfer function obtained from the experiments

then is possible to evaluate its frequency response, which corresponding to last two columns of Table 3, in other words the exact values.

Why does it obtain diferents results?

Depending on the before theorical analysis, can be computed differents types of errors, on application this method, on presence of disturbances. But in this simulation there is not disturbances. The reason is easy to explain, since thought the process is continuos, the results of this experiment has been developed in the matrix language Matlab and Simulink, which implies that the calculations are numericals and discrete, and then it is not exacts. For instance, simulation of Simulink models generally involves the integration of sets of ordinary differential equations. Performance of the simulation in terms of speed and accuracy varies for different models and conditions; in particular the choice of minimum step size will generate a great range of values respectably the desired response.

Therefore, the process transfer function is obtained during the experiment without too much calculations. It is very interesting because measurement can be reapeted for each frequency and the experiment can be done adaptively checking previous results. There is one case when this method can not be used that is when analysing systems with very long time constants. In this case it takes a long time to wait for steady state.

8.4 Multifrequency Signals

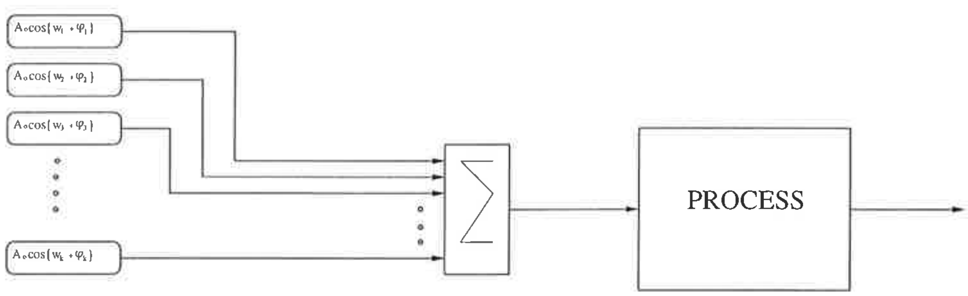

Figure 18 Scheme of Multifrequency identification.

Introduction One drawback of the correlation method is that the experiments may take considerable time since the measurements have to be reapeated for each frequency. In this section, it deals to dertermine the transfer function at several frequencies at same time by having parallel channals. The basic idea is simple. The time is reduced by introducing several sinusoids simultaneously.

Choice of sinusoidal input signals *why to choose sinusoidal signals?* The reason is simple, this signals contain a high percetage of the total signal power, but binary multifrequency signals contain only a percentage of the total signal power. On the other hand, is very important to take into account the following property: *Low Peak Factor Multifrequency Signals.*See [9]. Then, the peak factor of any signal x(t) is defined by:

$$Peak \quad Factor = \frac{x_{max} - x_{min}}{2\sqrt{2}x_{rms}} \tag{74}$$

where x_{max}, x_{min} and x_{rms} are, respectively, the maximum, minimum and rms values of x(t). The normalising factor of $2\sqrt{2}$ in the denominator means that the peak factor of a single sine wave is 1, while that of a binary signal is $1/\sqrt{2}$. Therefore, is required:

1. Multiharmonic or multisine signal.
2. Low peak factor.

This obviously has all its power in the specified harmonics, but the main design requeriment for such signals is to phase the harmonics in such way as to avoid one or more large peaks occurring at some time during the signal period, with only a small signal amplitude between the peaks.
A multiharmonic signal, with its total power normalised to 1 and with its components at specified harmonically related frequencies $f = k/T$ Hz, where each k is a (specified) integer, is given by

$$x(t) = \sum_k \quad (2P_k)^{\frac{1}{2}} \left[\frac{2\pi kt}{T} + \varphi_k\right] \tag{75}$$

where T is the period of the fundamental, and P_k is the relative power of harmonic k, i.e. $\sum_k P_k = 1$.

To ilustrate the design problem, suppose that is required, for instance, design a signal with 20 consecutive harmonics from 1 to 20, with equal power in each harmonic. Then $P_k = 0.05$ *and* $(2P_k)^{\frac{1}{2}} = 0.316$, so that if the totol signal power is 1. Therefore, the input signal is as following

$$x(t) = \sum_{k=1}^{20} 0.316 \cos\left[\frac{2\pi kt}{T} + \varphi_k\right]$$

Most of signals are not very attractive for use in system identification. The larges peaks means in each waveform that any system nonlinearities (in particular, saturation nonlinearities) could easily be reached, while the small amplitude of both signals throughout most of the rest of their time periods could well lead to inaccuracies due to quantisation if the signal is digitised.
This problem is possible to resolve it by applying a simple formula for the phases, proposed by Schroeder, with which a considerable reduction in peak factor can be achieved in most cases. This formula is as following

$$\varphi_k \quad = \quad 2\pi \sum_{i=1}^{k} i.P_i \tag{76}$$

since Schroeder noted that frequency modulated signals have a low peak factor and that Woodward's theorem states that the spectrum of a high index frequency modulated waveform is approximately the probability distribution of its instantaneous frequency. Working this through, Schroeder found that, for harmonic the phase should be depending on his formula.

Simulations The advantages of having a low peak factor in a multifrequency signal are shown in this identification. Then in this example, the use of 16 consecutive harmonics, with equal power in each harmonic, so that $P_k = 1/16$, and $(2P_k)^{\frac{1}{2}} = 0.3536$, $k = 1, 2, ..., 16$. The system in this example is as follows

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

which is a low pass system with a bandwidth frequency of 0.5098 rad/sec. Then for this example with 20 equally spaced, equal amplitude harmonics, the input signal is given as follows

$$x(t) = \sum_{k=1}^{16} 0.3536 \cos\left[\frac{2\pi kt}{T} + \varphi_k\right]$$

and by using Schroeder's formula it obtains the following φ_k

$$\varphi_k = \frac{2\pi}{16} \sum_{i=1}^{k} i, \qquad \textit{so that}$$

$$\varphi_1 = \frac{2\pi}{16}, \quad \varphi_2 = \frac{6\pi}{16}, \quad \varphi_3 = \frac{12\pi}{16}, \quad \ldots, \quad \varphi_{16} = \frac{272\pi}{16}$$

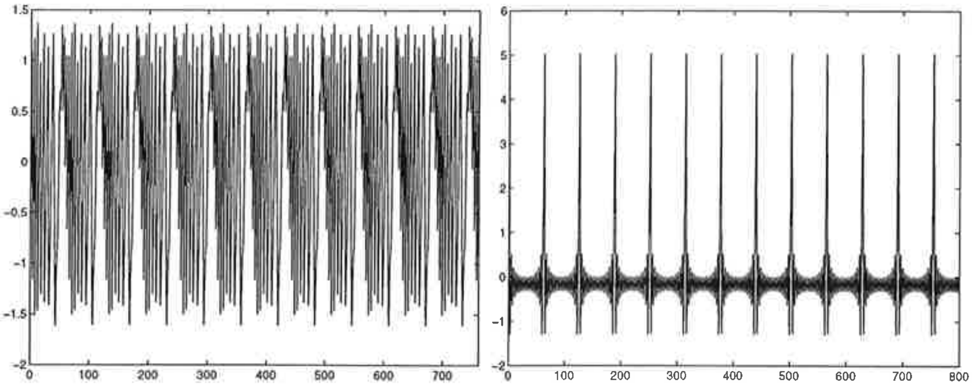

Figure 19 Input signal by using Schroeder's formula is shown on the left figure. The right hand figure is input signal has all phases set to $-90°$

The resulting signal are shown in Fig 19; by applying Schroeder's formula it obtains $x_{max} = 1.56, x_{min} = -1.85$, and the peak factor of 1.2 is much lower than for either the unphased cosine signal, where $x_{max} = 5.05, x_{min} = -1.29$, and the peak factor is 2.24.
The frequency response $G(j\omega)$ of the system is measured, using Welch's averaged periodogram method, where the signal is divided into overlapping sections, each of which is detrended and windowed. The Bode plot of the gain $\mid G(j\omega) \mid$ and the phase $ArgG(j\omega)$ are shown in Fig 20 for Schroeder signal and normal clipped signal. Schroeder's formula results in low peak factors in

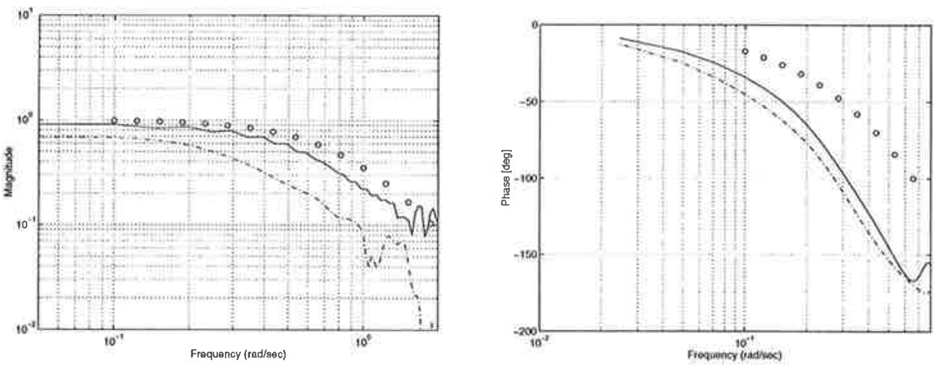

Figure 20 Bode Frequency response of process transfer function, where there is three signals for each plotting, marked with o is real response, the dashed lines correspond to unphased clipped signal, and the straight line correspond to Schroeder's signal

most cases. In this case does not work perfectly since the signal is with only a small number of harmonics; in [9] is possible to see one exact identification for each frequency, because they use 400 harmonics. On the other hand, the input signal clipped unphased offers a result worse since it has not low peak factor; in other words, its peak factor is quite low, but not as low as could have been achieved though the application of Schroeder's formula.
On the other hand, the corresponding coherence functions are shown in Fig 21. This function is defined by

$$\gamma^2(\omega) = \frac{\mid S_{xy}(j\omega) \mid^2}{S_{xx}(j\omega) \quad S_{yy}(j\omega)} \tag{77}$$

where the quadratic coherence always takes on a value in the interval

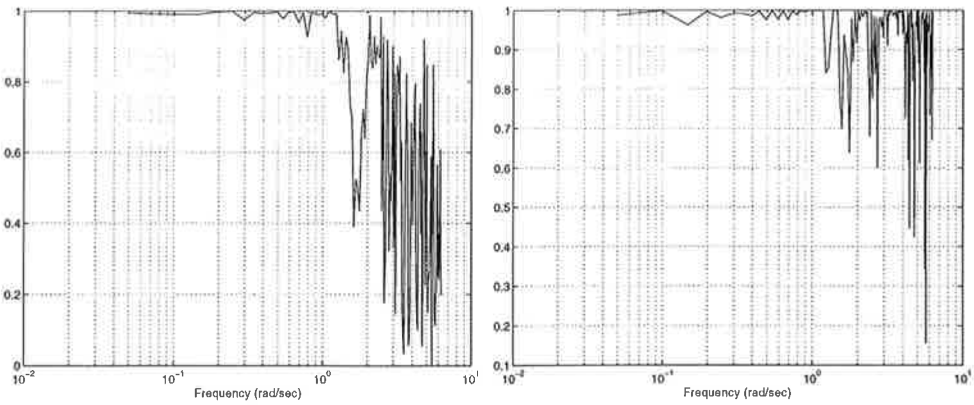

Figure 21 The quadratic coherence spectrum between the two signals x and y by using Schroeder's formula is shown on the left figure. The right hand is also the coherence function for unphased input signal

$0 \leq \gamma^2(\omega) \leq 1$ with a value close to one if the noise level is low, or also expresses the degree of linear correlation in the frequency domain between the input x and the output y. Therefore, in Fig 21 it shows while that the input signal has quite power (low frequency) the coherence function is more proximity to 1. On the other hand, Schroeder's function offers, in this frequencies, better behavior than the normal function. See [6].

8.5 Parametric function estimation with time delay

Frequency response fitting based on least squares weighting identification in complex domain is good idea. Then, in a rational approximant of process transfer function

$$\hat{G}(s) = \frac{\hat{B}(s)}{\hat{A}(s)} = \frac{\hat{b}_1 s^{n-1} + \hat{b}_2 s^{n-2} + \cdots + \hat{b}_n}{s^n + \hat{a}_1 s^{n-1} + \hat{a}_2 s^{n-2} + \cdots + \hat{a}_n} e^{-\hat{\tau} s} \tag{78}$$

To notice that with this method, a transfer function is fitted to the data $G(j\omega_k)$ obtained from experimental data and known at the frequency points $\omega_k \quad k = 1, 2, \cdots, N$. A first approximation problem is shown as follows

$$\min_{\Theta} J_1(\Theta) = \min \Theta \sum_{k=1}^{N} \mid G(j\omega_k) - \hat{G}(j\omega_k) \mid^2 = \sum_{k=1}^{N} \mid G(j\omega_k) - \frac{\hat{B}(j\omega_k)}{\hat{A}(j\omega_k)} \mid^2 \tag{79}$$

but this formulation gives a non-linearity in the error. By using the equation $\hat{B} - G\hat{A}$, it gets the loss function [Levy, 1959]:

$$J(\Theta) = \sum_{k=1}^{N} W_k^2 \mid \hat{B}(j\omega_k) - \hat{A}(j\omega_k) G(j\omega_k) \mid^2 \tag{80}$$

instead. This formulation is better than before because leads that error is linear in Θ. The weightings W_k^2 are used to de-emphasize high frequencies and to emphasize certain frequencies, see [8]. As far the choice of weightings W_k^2 there is two options, to be constant and therefore should decrease with the frequency

in order to avoid large weighting of high frequencies; Another possibility is to use an iterative scheme where the weight function W is updated after each iteration, like [Sanathan and Konerner, 1963]: $W^{k+1}(s) = 1/\hat{A}^k(s)$.
Following with the analysis method, it obtains

$$\phi = \begin{bmatrix} (j\omega_1)^{n-1} & (j\omega_1)^{n-2} & \cdots & 1 \\ \vdots & \vdots & & \vdots \\ (j\omega_N)^{n-1} & (j\omega_N)^{n-2} & \cdots & 1 \end{bmatrix}$$

and

$$\psi = \Gamma \begin{bmatrix} (j\omega_1)^n \\ \vdots \\ (j\omega_N)^n \end{bmatrix} \qquad \Phi(\tau) = \begin{bmatrix} -\Gamma\phi & D(\tau)\phi \end{bmatrix}$$

the $\phi-matrix$ will be dependent of the time delay $\hat{\tau}$. Where $\Gamma = diag(\{G(j\omega_k)\}_{k=1}^N)$, and $D(\tau) = diag(\{exp(-j\omega_k\tau)\}_{k=1}^N)$
and the corresponding loss function with unity weighting to simplify notation.

$$J(\Theta, \tau) = \sum_{k=1}^{N} W(j\omega_k) \mid^2 \mid \hat{B}(j\omega_k) - \hat{A}(j\omega_k)G(j\omega_k) \mid^2 = \mid \Phi(\tau)\theta - \psi \mid^2$$

This gives a non-quadratic minimization problem in 2n+1 variables θ and τ. Then, the problem can be separated into a quadratic minimization in θ and another non-quadratic minimization in the time delay τ. The procedure in order calculate θ was shown previously and the escalar minimization in τ is carried out as follows, see [8].
Theorem 1
Let $\Phi^*\Phi$ be non-singular and let P denote the matrix valed function

$$P(\tau) = I - \Phi(\tau)(\Phi(\tau)^*\Phi(\tau))^{-1}\Phi(\tau)^*$$

Every local minimun of J with respect to τ and θ is given by

$$\hat{\theta} = \Theta(\hat{\tau})$$

where $\hat{\tau}$ is a local minimum of the function f defined by

$$f(\tau) := J(\Theta(\tau), \tau) = \psi^* P(\tau)\psi$$

where

$$\Theta(\tau) := \min_{\theta} J(\theta, \tau) = (\Phi^*(\tau)\Phi(\tau))^{-1}\Phi(\tau)^*\psi$$

Proof: A necessary condition for J to have a local minimun at $\tau = \hat{\tau}, \theta = \hat{\theta}$ is

$$0 = \begin{bmatrix} \frac{\partial J}{\partial \theta} & \frac{\partial J}{\partial \tau} \end{bmatrix}_{\tau=\hat{\tau}, \theta=\hat{\theta}}$$

For each fixed τ, there exists a unique solution to the quadratic minimization problem in θ, given by $\Theta(\tau)$. Computing the first of f, with respect to τ gives

$$\frac{df}{d\tau} = \left.\frac{\partial J}{\partial \tau}\right|_{\theta=\Theta(\tau)} + \left.\frac{\partial J}{\partial \theta}\right|_{\theta=\Theta(\tau)} \frac{d\Theta}{d\tau}$$

$$= \left.\frac{\partial J}{\partial \tau}\right|_{\theta=\Theta(\tau)}$$

then, it show that the last step the fact that $\frac{\partial J}{\partial \theta} \equiv 0$ on the curve $(\theta, \tau) = (\Theta(\tau), \tau)$. This shows that any stationary point of J, with respect to $\theta and \tau$, corresponds to a stationary point of f with respect to τ. If a stationary point $(\hat{\theta}, \hat{\tau}) to J(\theta, \tau)$ in fact is a local minimum, then J has local minimum n every curve $(\theta(v), \tau(v))$, where u is a scalar parameter, passing through $(\hat{\theta}, \hat{\tau})$. One such is $(\theta, \tau) = (\Theta(v), v)$ which implies that f has a local minimum at $(\hat{\theta}, \hat{\tau})$, see [8].
The minimization of f with respect to τ can be done for example by the following modified Newton-Raphson algorithm:
Algorithm 1
The updating of time delay τ is according to

$$\tau_{n+1} = \tau_n - \frac{\frac{df}{d\tau}(\tau_n)}{\alpha \left| \frac{d^2 f}{d\tau^2}(\tau_n) \right| + (1-\alpha) \frac{d^2 f}{d\tau^2}(\tau_n)}$$

where $0.5 < \alpha < 1$. The modification is done in order to make local maxima repulsive and local minima attractive. Values of α less than one are chosen to get a stronger "repulsion" from a maximum. However, if the initial value of τ is chosen exactly at a maximum. Since ψ is independent of τ, the first and second derivatives of f are given by

$$\frac{df}{d\tau} = \psi^* \frac{dP}{d\tau} \psi$$

$$\frac{d^2 f}{d\tau^2} = \psi^* \frac{d^2 P}{d\tau^2} \psi$$

Taking into a account the previous relations of P,Φ and D, the first derivative of P is computed according to

$$\frac{dD}{d\tau} = diag(\{-j\omega_k exp(-j\omega_k \tau)\}_{k=1}^{N})$$

$$\frac{d\Phi}{d\tau} = \begin{bmatrix} 0 & \frac{dD}{d\tau}\phi \end{bmatrix}$$

$$\frac{dP}{d\tau} = Q + Q^* \qquad Q := -P\frac{d\Phi}{d\tau}(\Phi^*\Phi)^{-1}\Phi^*$$

and the second derivative of P is given by

$$\frac{d^2 D}{d\tau^2} = diag(\{-j\omega_k^2 exp(-j\omega_k \tau)\}_{k=1}^{N})$$

$$\frac{d^2\Phi}{d\tau^2} = \begin{bmatrix} 0 & \frac{d^2 D}{d\tau^2}\phi \end{bmatrix}$$

$$P_1 := \frac{d\Phi}{d\tau}(\Phi^*\Phi)^{-1}\Phi^*$$

$$P_2 := \frac{d^2\Phi}{d\tau^2}(\Phi^*\Phi)^{-1}\Phi^*$$

$$\frac{d^2P}{d\tau^2} = \frac{dQ}{d\tau} + \frac{dQ^*}{d\tau} \qquad \frac{dQ}{d\tau} = P(2P_1^2 - P_2) + Q^*Q + Q^*$$

More information about this method is possible to find in [8].

Simulations In many process, is very important to consider a higher order system [8] described by

$$G(s) = \frac{1}{(s+1)^8}$$

where hence, system identification can be seen as a special case of model reduction. Note that model reduction is nothing but a way to compress a description of a system. Since high order models will also imply high order controllers and high order predictors filters, which may be expensive and dificult to implement. For this reason, it will choose the desired transfer function, respectably the another process, to be

$$\hat{G}(s) = \frac{e^{-\tau s}}{s^2 + 2\zeta\omega_0 s + \omega_0^2}$$

It is important to specify a time delay in the desired transfer function performance for higher order systems in order to compensate for the phase shift of $G(j\omega)$. In this particular case, the time-delay $\tau = 2$ will select to be equal to the "apparent" delay in the system and $\zeta = 0.7$ will choose as the desirable damping factor.
The process transfer function proposed by simulations is one with the follows form

$$\hat{G}(s) = \frac{\hat{B}(s)}{\hat{A}(s)} = \frac{\hat{b}_0 s + \hat{b}_2}{s^3 + \hat{a}_1 s^2 + \hat{a}_2 s + \hat{a}_3} e^{-\hat{\tau} s}$$

In this simulation, it will analyze the follows cases

1. Delay equal to zero
2. Small delay
3. Big delay
4. Weighting effect

Delay equal to zero. In this case, when the delay is equal to zero the algorithm applied is the same than the basic least squares identification in the complex frequency domain, since the matrix $D(\tau)$ is unitary and then the regressor matrix will be $\Phi = [-\Gamma\phi \quad \phi]$. The real and estimated parameters are shown on the Table 4, in which it notices precision almost exact. The reason of this perfect estimation is the choice of identification frequencies, this are

$\omega = 0.01, 0.1, 0.25$; if the choice had been other, for instance in high frequency is possible to obtain unstable models.

Parameters	Delay equal to 0	Small delay	Big delay
a_1	1.5	1.5	1.5
a_2	0.75	0.75	0.75
Exact a_3	0.125	0.125	0.125
b_0	1	1	1
b_1	0.6	0.6	0.6
τ	0	0.01	11
$\hat{a}_1$	1.5000016	1.49985	3.7401
$\hat{a}_2$	0.7500015	0.74986	2.8085
Estimated $\hat{a}_3$	0.1250003	0.12496	0.6030
$\hat{b}_0$	0.9999996	1	-0.750
$\hat{b}_1$	0.6000017	0.59985	2.8943
$\hat{\tau}$	0	0.01001	10.4165

Table 4 Comparision between real parameters and estimated parameters by Least squares fitting of a rational function with differents delays

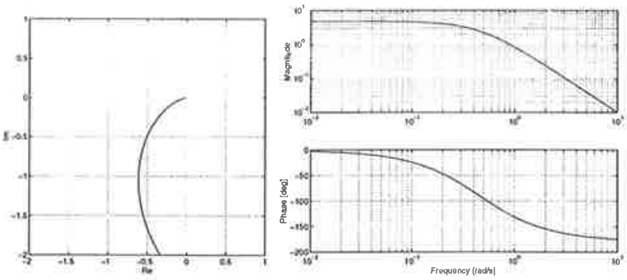

Figure 22 Nyquist and Bode plot of real and estimated frequency response for Delay equal to zero

Small delay. Why to distinguish between small and big delay? The reason is very clearly since the effect on pole placement is more important as fast as delay is bigger. For instance, for $\tau = 0.01$ and identical parameters of process transfer function, it notices, on Table 4, a good parameters estimation. This delay as is small, hardly has weight on lag phase; since the delay is a behavior of Non-minimum Phase and has a delay of excessive phase without reduction in high frequencies.

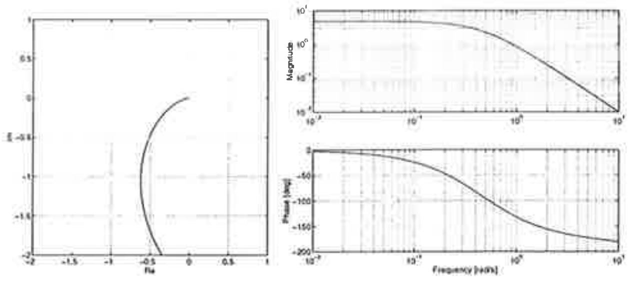

Figure 23 Nyquist and Bode plot of real and estimated frequency response for Delay equal to 0.01

Big delay. In this case, the chosen delay is $\tau = 11$, and though the process transfer function has its poles and zeros in left-half plane, the big delay produces a important lag phase and a behavior of Non-minimum Phase; even giving a gain and phase margins negatives, as it shows in Fig 24. Therefore, the algorithm used on identification can not reproduce the original parameters of process, since the delay changes the initial parameters. On the other hand, thought the identified delay is 10.4165 and the real is 11, it is not enough in order to obtein a approximate identification. A analysis of obtained parameters shows that Poles are on left-half plane but Zero is on right-half plane, see Table 4. Although there is a discrepancy between real parameters and estimated parameters, the difference of frequency response between real system and estimated system is minimum. This is because on identification it takes the frequency response on the whole, and therefore it does not take the contribution of process transfer function on the one hand and the delay on the other hand.

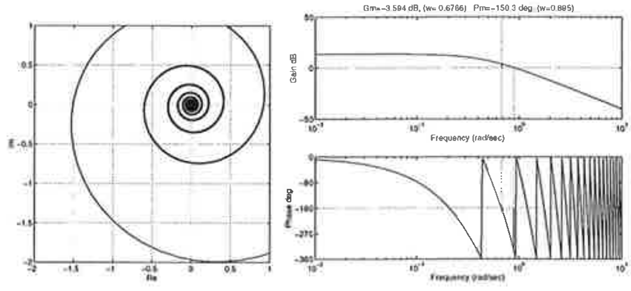

Figure 24 Nyquist and Bode plot of real and estimated frequency response for Delay equal to 11

Weighting effect. In the loss function J in equation 80, now is considered its weighting function $W(j\omega_k)$. This function provides weightings that increase the importance of frequencies that close to the crossover frequency. This is because, if it chooses W in this way $W^{k+1}(s) = 1/\hat{A}^k(s)$, it has that W is low pass with a cut-off near crossover. The shape of $\mid W(j\omega_k) \mid^2$ naturally determines the frequency range over which the minimization of J takes place. For instance, if it takes W in the following form

$$W(j\omega) = \frac{1}{(j\omega + 0.5)^3}$$

and it chooses the frequencies $\omega = 0.01, 0.1, 0.6$, the weighting effect is minimum, since only $\omega = 0.6$ escapes lightly of cut-off frequency, therefore the effect is minimun as it follows in Table 5.

Parameters	Estimated values with weighting function	Estimated values without weighting function
a_1	1.4996	1.4995
a_2	0.7496	0.7495
a_3	0.1249	0.1249
b_0	1	1
b_1	0.5996	0.5995
τ	0.01	0.01

Table 5 Comparision between estimated parameters with weighting function and with weighting function equal to one, by Least squares fitting of a rational function with a delay

9. Conclusion

A new techinique has been presented for the design of PID auto-tuners. The method is mainly investigated through extensive simulations with process transfer functions that are typical for the processes encountered in the process industry. The method can be applied when process dynamics is given only in terms of three parameters obtained from a frequency response. This method compared to [2] has the following properties: the oscillation frequency is chosen with respect to obtaining sufficient gain and phase margins for the control loop, as well as, in process without action derivative, as a PI controller the phase lag of the loop transfer function then has to be more negative than-180o respectably [2], while this method, in this way, has more phase margin. Compared this method with Schei, it improves the duration tuning and accuracy PID parameters. Compared with Woodyatt-Middleton works at same level, but need prior information, so that is not exactly a Auto-tuner.
On the other hand, other points of view in order to identify points of frequency response have been presented, but in same way that relay tuning as for the identification it refers.x

10. References

[1] K. J.ÅSTRÖM and T.HÄGGLUND. *PID Controllers: Theory, Design and Tuning.*. Instrument Society of America, Research Triangle Park, NC, second edition, 1995.

[2] K. J.ÅSTRÖM and T.HÄGGLUND. *"Automatic Tuning of Simple Regulators with Specifications on Phase and Amplitude Margins"*, Automatica **20**, pp.645-651,1984.

[3] K. J.ÅSTRÖM and T.HÄGGLUND. *"Automatic tuning of PID controllers."*ISA, Research Triangle Park, NC, USA, 1988.

[4] P. PERSSON. *Towards Autonomous PID Control.* , PhD thesis ISRN LUTFD2/TFRT–1037–SE, Department of Automatic Control, Lund Institute of Technology, Lund, Sweden, April 1992.

[5] T. S. SCHEI. *"A Method for Closed Loop Automatic Tuning of PID Controllers"* , Automatica, **Vol 28, No 3**, pp.587-591, 1992.

[6] R. JOHANSSON. *System Modeling & Identification* , Prentice Hall, 1993.

[7] K. J.ÅSTRÖM . *"Lectures on System Identification."* Chapter 3, Frequency Response Analysis.

[8] M. LILJA *Controller Design by Frequency Domain Approximation.* , PhD thesis ISRN LUTFD2/TFRT–1031–SE, Department of Automatic Control, Lund Institute of Technology, Lund, Sweden, August 1989.

[9] K. R. GODFREY. *"Introduction to non-binary signals used in system identification"* IEE International Conference on "Control 91",Heriot-Watt University, Edinburgh, UK.

[10] A. WOODYATT and R. MIDDLETON. *"Auto-Tuning PID Controller Design Using Frequency Domain Approximation"* ,

[11] T. J. D. BARNES, L. WANG and W. R. CLUETT. *"A Frequency Domain Design Method For PID Controllers"* , Proceedings of the American Control Conference, 890-894, 1993.

[12] J. G. ZIEGLER and N. B. NICHOLS. *"Optimum settings for automatic controllers"*. Trans. ASME, **64**, pp.759-768, 1942.