PID Control, Design, Extension, Application

Panagopoulos, Helene

2000

Document Version:
Publisher's PDF, also known as Version of record

Link to publication

Citation for published version (APA):
Panagopoulos, H. (2000). *PID Control, Design, Extension, Application*. [Doctoral Thesis (compilation), Department of Automatic Control]. Department of Automatic Control, Lund Institute of Technology (LTH).

Total number of authors:
1

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

PID Control

Design, Extension, Application

Hélène Panagopoulos

Automatic Control

PID Control

Design, Extension, Application

PID Control
Design, Extension, Application

Hélène Panagopoulos

Lund 2000

To my family

Department of Automatic Control
Lund Institute of Technology
Box 118
SE-221 00 LUND
Sweden

ISSN 0280–5316
ISRN LUTFD2/TFRT--1059--SE

Printed in Sweden by Wallin & Dalholm Boktryckeri AB
Lund 2000

Contents

Acknowledgment

I want to express my sincere gratitude

– to, above all, my supervisors Professor Karl Johan Åström and Professor Tore Hägglund, for their valuable ideas, suggestions, and examples throughout my research in design of PID controllers. It has been a great pleasure and a privilege to work with them and I am grateful for their outstanding support and encouragement during the work of this thesis.

– to Anders Wallén for fruitful discussions, cooperations and help with MATLAB 5, to Jonas Eborn as the co-author of one of the papers in this thesis, and to Professor Bo Bernhardsson, Dr. Lennart Andersson and Oskar Nordin for their help in proof reading the thesis and suggesting valuable changes.

– to Börje Eriksson and Oskar Nordin at Modo Paper in Husum, Sweden, for their assistance and patience to realize the practical evaluations of my work at Modo Paper, Husum, Sweden.

– to Eva Dagnegård, Britt-Marie Mårtensson and Eva Schildt, for all help and encouragement during my years at the department. I would also like to thank Leif Andersson for his LaTex and Unix support.

– to the Swedish foundation for strategic research (SSF) within the project CPDC.

– finally, I would like to thank all colleagues during my years at the department, family, and friends for all help and encouragement they have given me during this period of research.

Introduction

In these competitive days there are substantial benefits for process industry in improving control as it will, from several points of view, yield better economy: lowered cost of raw material, energy, equipment and workforce, and better fulfillment of the environmental conditions. As a consequence it is possible to increase the level of competition. The required investments, to achieve initial improvements on existing operations, are generally very modest compared to the obtained benefits. The only real costs are the working-hours of the workforce and, if needed, the replacement of some control hardware. According to Bialkowski (1997), the cost/benefit ratio of these investments is at least 1:10 and often, much higher.

In this thesis the subject of "How to improve control?" is treated by presenting: new design methods for proportional integral derivative (PID) controllers, the extensions of these controllers to improve their performance, and the applications of these design methods to industrial processes. The thesis consists of the following papers, and supplements:

I. ÅSTRÖM, K. J., H. PANAGOPOULOS, and T. HÄGGLUND (1998): "Design of PI Controllers Based on Non-Convex Optimization." *Automatica,* **34:5**.

II. PANAGOPOULOS, H., K. J. ÅSTRÖM, and T. HÄGGLUND (1999): "Design of PID Controllers Based on Constrained Optimization." In *1999 American Control Conference.* San Diego, California.

IIs. PANAGOPOULOS, H., K. J. ÅSTRÖM, and T. HÄGGLUND (2000): "Supplement and Errata to "Design of PID Controllers Based on Constrained Optimization"." Department of Automatic Control, Lund Institute of Technology, Lund, Sweden.

III. PANAGOPOULOS, H., and K. J. ÅSTRÖM (2000): "PID Control Design and $\mathcal{H}_\infty$ Loop Shaping." Accepted to *Robust and Nonlinear Control.*

IV. EBORN, J., H. PANAGOPOULOS, and K. J. ÅSTRÖM (1999): "Robust PID Control of Steam Generator Water Level." In *IFAC'99 14th World Congress of IFAC.* Beijing, P. R. China.

V. PANAGOPOULOS, H., A. WALLEN , O. NORDIN, and B. ERIKSSON (2000): "A New Tuning Method with Industrial Evaluations." *Submitted to Control Systems 2000.* Canada.

VI. PANAGOPOULOS, H., and T. HÄGGLUND (2000): "A New Modular Approach to Active Control of Undamped Modes." Department of Automatic Control, Lund Institute of Technology, Lund, Sweden.

Today's PID controllers

In today's process industry it is still PID controllers which are the most frequently used controllers in the distributed control systems. Estimates indicate that more than 90% of all controllers used are of the PID type. There are a number of reasons for its popularity.

– To begin with the PID controller has a long history of proven operation, and if it is properly designed and well-tuned it gives satisfactory control in many applications.

– Secondly, PID controllers are well understood in operational, technical and maintenance occupations. To implement it in present control systems is not difficult, and it requires no advanced equipments or demands of the instrument engineer. As a PID controller has fixed complexity, that is only three controller parameters has to be set, it will not require an exact process model, and the controller parameters can be easily adjusted after the installation. When developing a new design method these functions are important to have in mind.

– Thirdly, with today's cheap microprocessors complex control algorithms can easily be implemented, which are superior to the traditional PID controllers. The rapid progress of control techniques has also resulted into the packaging of control systems. Nevertheless, the PID controller maintains its popularity in the process industry. One reason is the instrument engineers habit to have considerable authority over the control of the plant. Packaged control systems such as the model predictive control in Maciejowski (1999) will then be used to support the instrument engineers, rather than replacing him/her. For example, the model predictive control is implemented on top of the conventional PID controllers to provide set points to them. Another reason for the popularity of PID controllers, are the difficulties for an engineer to understand the principles and mechanisms of complex control systems. The intuitive comprehensibility, the continuity between control methods, and the conventional experience are what engineers truly desire on automating the control systems in process industries.

– Finally, there are many extensions which make an industrial PID controller useful for controlling a process. For example, the possibility to switch between manual and automatic mode, set point tracking and anti-windup. Furthermore, all new PID controllers are based on microprocessors. This has given opportunities to provide additional features like automatic tuning, gain scheduling, and continuous adaption. Auto tuning is useful for instrument engineers to keep the control loops well tuned. The benefits are even larger for more complex loops. Tuning facilities are also starting to appear in the distributed control systems. In this case it is possible to have a very powerful interaction of the user because of the graphics and the computational capabilities available in a system.

A Perspective on Control Design

For a long time the development of design methods has been the goal of the control community. In the 1930's and 1940's the classical frequency domain methods were developed. It began with Black's breakthrough on feedback amplifiers, followed by Nyquist and Bode's work in stability theory. In the classical approach the main concern was to design feedback compensators in order to achieve a certain stability margin, see for example Truxal (1955). The emphasis was on model uncertainties and the feedback was used to decrease sensitivity to disturbances and model errors. The compensatory design was done mainly by graphical methods evolving from the Nyquist's stability criterion.

In the 1950's analytical design methods were developed, see for example Newton *et al.* (1957), where specifications on the transient performance were given. With the appearance of analog computers it was easy to check time response criteria, and performance criteria quantitatively expressed. At the same time less attention was given to robustness and sensitivity issues.

During the mid 1950's there was a renewed interest to consider ordinary differential equations as a model for control systems which typically called for the extensive use of digital computers. Much of this work was stimulated by a new field, the control of artificial earth satellites. At about the same time the study of optimal control was extended to find optimal trajectories for nonlinear dynamical systems, particularly for robots, aircrafts and spacecrafts.

In the 1960's the development of control design methods based on optimization techniques had the advantage to capture many different aspects of the design problem. During this time efficient computer methods were developed to solve these problems. A general discussion of the use of optimization for control design is found in Boyd and Barratt (1991) and, Mayne and Polak (1993).

In the 1970's, it became apparent that more attention should be given to robustness issues. This was particularly true for systems with lightly damped oscillating modes such as flexible space vehicles. During the last two decades new controller design methods, such as the $\mathcal{H}_\infty$-design, have been developed in which robustness issues have regained its importance.

Why a New PID Design?

Why do we need a new design method for PID controllers? What is wrong with the existing ones? There are a number of reasons.

– Over the past 40 years an enumerable number of books and technical articles have presented methods for the design of PID controllers. "The fact that none of these seems to have gained widespread acceptance indicates that, as in the case of razor blades, the best solution has not yet been found.", see Smith *et al.* (1975). On the other hand, there is a need for a variety of techniques to tune PID controllers, which ranges from simple tuning procedures to more elaborate ones based on process modeling. It is also necessary to be aware of the fact that there are many different types of control problems and consequently many different design methods. Only to use one method is as dangerous as only to believe in empirical tuning rules.

– Many design methods are based on simple parameter models, obtained by measuring the natural period and/or the dead time, and time, for example, Ziegler and Nichols (1942), and Cohen and Coon (1953). Even if these design methods are easy to use it is important to be aware that they are not, always, a substitute for the insight and understanding of design methods based on more process knowledge. Today it is possible to obtain, in a simple way, enough process knowledge needed for a good design, see Wallén (1999), and Van Overschee and Moor (1999).

– A lot of design methods consider only one aspect of the control problem. For example, in the classical design rule of Ziegler and Nichols (1942) the objective was to reject load disturbances, no aspect was taken to changes in set point. It would be desirable to have a design method which capture engineering criteria such as load disturbance response, robustness to model uncertainties, rejection of measurement noise, and set point tracking. Many traditional tuning rules in PID control do not consider all of these aspects in a balanced way which is a considerable drawback.

– The increased popularity of automatic tuning and of field-busses in the process industry will drastically simplify the use of controllers. Single loop controllers and distributed control systems are important application areas of the process industry where most of controllers are of the PID type. As this is a vast application area, there are millions of controllers of this type in use, new tuning rules would be highly beneficial.

These findings gives an indication of the need for modern methods which take into consideration the process dynamics, and, at the same time, several aspects of the design specifications. Concerning the particular method to use, it is too early to draw definite conclusions. There are many different ways to determine process characteristics, many methods for design of PID controllers, and many ways of combining such techniques to create auto tuners, see Van Overschee and Moor (1999). But, a structured approach to control loop design and tuning is still needed.

Important Properties of a New PID Design

The discussion in the previous section will serve as a guideline to determine an efficient design method for PID controllers.

A controller design which is based on simple model will give a straightforward strategy. On the other hand, the flexibility in the specifications, and the alternatives when an unsatisfactory design is obtained are limited. A more desirable approach would be a model based method which recognizes the importance of dynamics from higher order models. It is then possible to obtain better control of important loops, for example slow ones such as temperature control loops, or hard ones where the usual thumb rules are not enough.

When solving a control problem it is necessary to understand the primary goal of control. Two common control objectives would be: good set point tracking and fast rejection of load disturbances. It is also important to have an assessment of the major limitations. Typical specifications are then: rejection of load disturbances, set point tracking, robustness to model uncertainties, and rejection of measurement noise.

Consequently, the controller design should be applicable to a wide range of systems. It should have the possibility to introduce specifications which capture the essence of real control problems. The method should also give all parameters of the PID controller including the set point weight, the set point filter, and the measurement noise filter, if desired. Furthermore, the method should be robust in the sense that it provides controller parameters, if they exist, or, if the specifications can not be met an appropriate diagnosis should be presented. Most of these requirements are satisfied by the design methods for PI and PID controllers presented in the thesis. A short resumée of the background of the proposed design methods for the PI and PID controllers follows below.

Background of the New PI and PID Controller Designs

Commercial PID controllers with automatic tuning facilities have only been available since the beginning of the 1980's. There are several reasons for this: First, the recent development of microelectronics has made it possible to incorporate the additional program code needed for the automatic tuning at a reasonable cost. Second, projects related to automatic tuning at the universities are quite new. Most research efforts have been devoted to the related, but more difficult, problem of adaptive control.

Projects related to automatic tuning started off in the beginning of the eighties at the Department of Automatic Control, Lund Institute of Technology, Sweden by the advent of the relay method in Åström and Hägglund (1984) for the identification procedure of the process model. It is still used in industrial products, single station controllers as well as distributed control systems, produced by ABB Automation Products.

The original auto tuner used a crude and simple design method, a modification of the Ziegler-Nichols method, to determine the controller parameters, where the process information was based on one point of the Nyquist curve. An improved method, the dominant pole design method, was presented in Hägglund and Åström (1985) where the process information was based on two points of the process Nyquist curve. Thus, the two most dominant closed loop poles were positioned. In Persson (1992) this method was improved and further developed. It assumed that the process transfer function was known. The design method positioned the dominant poles of the closed loop system such that a good rejection of load disturbance would be obtained. This is equivalent to minimize the integrated error

$$IE = \int e(t)dt,$$

with a robustness constraint. Consequently, the dominant pole design leads to an optimization problem, where robustness is incorporated by searching for the relative damping of the dominant poles which obtain a given value of sensitivity. One of the main drawbacks of the dominant pole design is its difficulty to obtain good starting values and reliable numerical methods. The present work can be viewed as a continuation that is based on a direct solution of an optimization problem to minimize the IE. This is a difficult problem as the complexity of the controller is restricted compared to, for example, the $\mathcal{H}_\infty$-design. Similar work have been done by, for example, Schei (1994), Van Overschee and Moor (1999), Langer and Landau (1999), and Kristiansson and Lennartsson (1999).

Extensions

The popularity of the PID controller in today's process industry is due to a number of reasons:

- Gives satisfactory control for most industrial processes.
- Long history of proven operation.
- Easy to understand.
- Standard form, thus no requirement of exact process model.
- Easy adjustment of the controller parameters after installation.

It is surprising that such a simple controller as the PID works so well. But, the simplicity implies limitations. There are applications where significant better performance is obtained with more sophisticated controllers:

- Processes with long dead times.
- Processes with nonlinearities.
- Processes with large interactions.
- Processes with poorly damped oscillatory modes.

For example in the first case, the performance of the system is enhanced by dead time compensation. In the second case, amelioration of the controller performance is made by gain scheduling. In the third case, improvements of the systems performance are made by using cascade control, feed forward control, or ratio control. How to improve the performance in the last case is treated in this thesis. A new way to damp the oscillatory modes is proposed in Panagopoulos and Hägglund (2000).

Furthermore, with today's development of more advanced, and complex design techniques, it would be desirable for control engineers to tune the three parameters of the PID controller based on these techniques. In this thesis it is proposed how to choose the parameters of the PID controller based on the $\mathcal{H}_\infty$-design in Panagopoulos and Åström (2000).

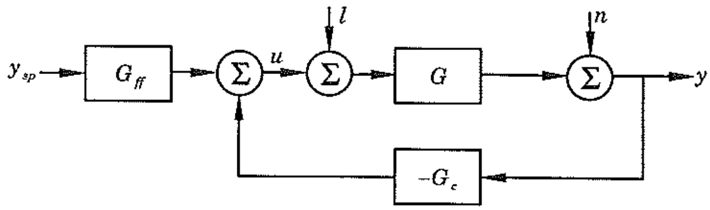


Figure 1. Block diagram describing the problem formulation of the proposed design methods for PI and PID controllers.

Design: A New PID Controller Design

In this thesis two design methods have been developed: design of PI and PID controllers respectively. The formulation of the design problem for the two methods coincide, they are therefore presented together.

Formulation of the Design Problem

Consider the design problem illustrated in Figure 1.. A process with transfer function $G(s)$ is controlled with a PID controller with two degrees of freedom, see Horowitz (1963). Transfer function $G_c(s)$ describes the feedback from process output y to control signal u, and $G_{ff}(s)$ describes the feed forward from set point y_{sp} to u. Three external signals act on the control loop, namely set point y_{sp}, load disturbance l and measurement noise n.

The design objective is to determine the controller parameters in $G_c(s)$ and $G_{ff}(s)$ so that the system behaves well with respect to changes in the three signals y_{sp}, l and n as well as in the process model $G(s)$. Therefore, the specification will express requirements on:

- Load disturbance response
- Robustness with respect to model uncertainties
- Measurement noise response
- Set point response

The formulation of the design problem can loosely be divided into two categories: specifications on performance and robustness. The first specification takes care of the rejection of load disturbances, measurement noises, at the same time as good set point following is obtained. Accordingly, the presented controller structure allows for independent tuning of the feedback term G_c and feed forward term G_{ff}. The second specification takes care of the sensitivity to model uncertainties. One of the major drawbacks of the classical Ziegler Nichols method was the absence of this specification.

Process and controller structures

The design problem is formulated to apply to a wide variety of systems. Consequently, the process is assumed to be linear, time invariant, and specified by a transfer function $G(s)$, which is analytic with finite poles and, possibly, an essential singularity at infinity. The description covers finite dimensional systems with time delays and infinite dimensional systems described by linear partial differential equations.

The limitation of the design to linear time invariant systems is a minor restriction. Many nonlinear process control plants are well modeled locally by linear time invariant systems. The operating range of the controllers can then be extended using gain scheduling or adaptation.

The controller is described by

$$u(t) = k\big(by_{sp}(t) - y(t)\big) + k_i \int_0^t \big(y_{sp}(\tau) - y(\tau)\big)d\tau + k_d\Big(-\frac{dy(t)}{dt}\Big), \quad (0.1)$$

where $k,\ k_i,\ k_d$ and b are controller parameters. (In the well known notation $k = K,\ k_i = K/T_i,\ k_d = K \cdot T_d$). If it is needed the signals y and y_{sp} can be replaced by their filtered values y^f and y^f_{sp}. The filtered signals are generated by

$$Y^f(s) = F_y(s)Y(s),$$
$$Y^f_{sp}(s) = F_{sp}(s)Y_{sp}(s),$$

where the filters are low pass filters of first or second order. The controller can thus be characterized by either three or four parameters $k,\ k_i,\ (k_d),\ b$, and two filters $F_y,\ F_{sp}$.

Load Disturbance Attenuation

The primary goal of the proposed design problem is to achieve good rejection of load disturbances, which in mathematical terms corresponds to minimize the integrated error, IE, where,

$$IE = \int (y_{sp} - y)dt.$$

The idea is illustrated in Figure 2., where the output y of the closed loop system is plotted for a unit step disturbance l applied at the process input.

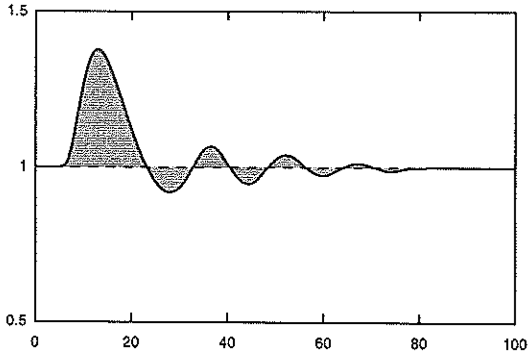


Figure 2. The output y of the closed loop system for a unit step disturbance l applied at the process input.

To minimize the IE corresponds to the minimization of the shaded area in the figure, where the controller parameters are related to the IE as

$$IE = 1/k_i.$$

Thus, the performance criterion of minimizing the integrated error, IE, is equivalent to maximize the integral gain, k_i. The problem formulation is, then, well suited to be solved with parameter optimization. Unfortunately, the criterion, alone, will not be sufficient to guarantee good performance for all systems, as undamped or even unstable responses can be obtained. Consequently, an additional constraint is needed to guarantee the stability of the closed loop system, which is discussed later on.

There are several reasons to optimize the process response to changes in load disturbance instead of set point. To begin with, load disturbances are more likely to change during operation compared to set points, which are usually kept fixed. Secondly, the rejection of a step upset is the most demanding of the load disturbances. A good set point tracking can be achieved by the use of a feed forward term. Finally, the performance criteria is economically related, the IE is proportional to operational cost, see Shinskey (1990). Still a lot of papers focus primarily on good set point tracking, even if practitioners such as Shinskey (1990) and Lopez *et al.* (1967) have emphasized the importance of optimizing the process response to changes in load disturbance.

Measurement Noise

In today's tuning methods the reduction of measurement noise is seldom incorporated in the calculations. In those cases a filter is designed, it is often only applied on the derivative part of the PID controller.

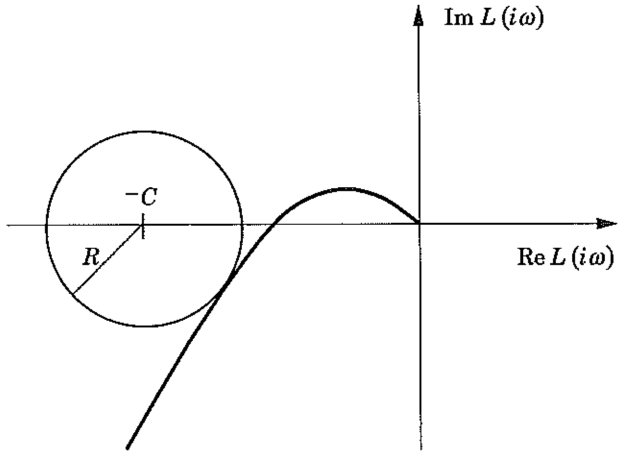


Figure 3. The robustness constraint of the proposed design methods for PI and PID controllers.

The advantages of the proposed design method are: First, the systematic approach to determine measurement noise filters , such that the noise level is reduced. Second, the possibility to filter all terms of the controller, which will reduce the high frequency gain from measurement noise to controller output. A more detailed discussion is given in Panagopoulos *et al.* (1999) and Panagopoulos *et al.* (2000), and an application is given in Eborn *et al.* (1999).

Set Point Response

The necessity of taking into consideration the set point response have been observed previously, see, for example, Shigemasa *et al.* (1987). A common way, in industry, to obtain a smooth set point response is to ramp it by using either a rate limiter, or a jump and rate limiter, see for more details in Åström and Hägglund (1995).

In the proposed design method the controller structure has two degrees of freedom, that is, a feedback term G_c and a feed forward term G_{ff}, see Figure 1.. This allows for independent tuning of the feedback and feed forward terms since both load rejection and set point tracking are important objectives. The parameters to be determined in the feed forward term are the set point weight b and, if necessary, the set point filter F_{sp}. A more detailed discussion of how to determine the feed forward term is given in Åström *et al.* (1998) and Panagopoulos *et al.* (2000).

Robustness

As was mentioned earlier the primary goal of the proposed design method was to achieve good rejection of load disturbances, that is, to minimize the IE or maximize k_i. It was realized that the criterion, by itself, would not be adequate for controller performance as it would not penalize a loop with an undamped response. The problem has been known for a while and to overcome it other tuning criteria have been developed such as the integrated absolute error (IAE), the integrated absolute time weighted error ($IATE$), the integrated square error (ISE), the integrated square time weighted error ($ISTE$), etc., see Shinskey (1990), Zhuang and Atherton (1991), Kaya and Scheib (1988) and references there in. There are drawbacks to use these criteria: some of them have not an obvious relationship to operational cost as the IE. There are criteria which uses a bad weight on the controller error either from an aspect of time or gain. For some criteria the mathematics needed to solve them may become too intrinsic. Still, it was tried in Andreas and Åström (1997) to replace the criteria of maximizing IE with the one of maximizing IAE for the proposed design method for PI controllers. The major drawback was that the optimal solution resulted into an undamped closed loop system, for more details see Andreas and Åström (1997).

In the proposed design method the drawback of not being able to guarantee a stable closed loop system have been overcome by the addition of a constraint. Consequently, the criterion of minimizing the IE is retained with the addition of a robustness constraint which guarantees the stability of the closed loop system. The robustness constraint is given in explicit terms: the Nyquist curve of the loop transfer function, $L(i\omega)$, must lie outside a specified circle with center $-C$ and radius R, that is,

$$\|C + L(i\omega)\| \geq R,$$

see Figure 3.. This establishes a nice connection between traditional design of PID controllers and the $\mathcal{H}_\infty$ control, see Panagopoulos and Åström (2000) Simple measures of the robustness are the maximum of the sensitivity function, the maximum of the complementary sensitivity function, and the weighted $\mathcal{H}_\infty$ norm, see Glover and McFarlane (1989), of the multi variable transfer function

$$\frac{1}{1+GG_c}\begin{bmatrix} G & 1 \\ -GG_c & -G_c \end{bmatrix}.$$

Thus, the robustness measure provides a transparent design variable. In the case of measuring the robustness in terms of the maximum of the sensitivity function, the Nyquist curve of the loop transfer function should

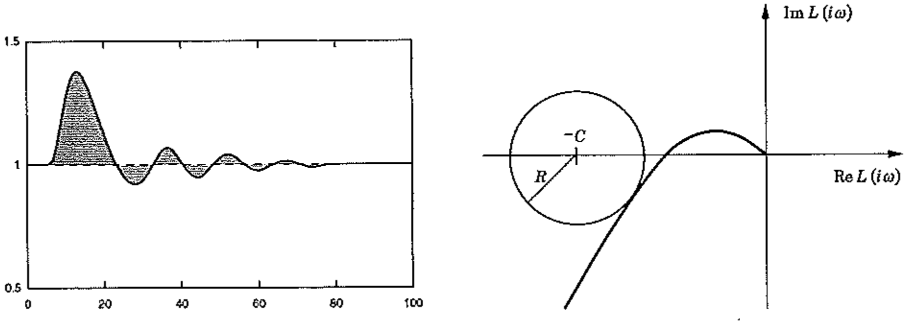


Figure 4. The problem formulation of the proposed design method.

avoid a circle with center at $C = -1$ and radius $R = 1/M_s$, where,

$$M_s = \|(1 + GG_c)^{-1}\|_\infty.$$

With a constraint on M_s it follows from the circle criterion that the closed loop system will also remain stable with a static nonlinearity in the loop, provided that the nonlinearity is bounded by $M_s/(M_s + 1)$ and $M_s/(M_s - 1)$, see Åström *et al.* (1998) and references there in.

Tuning Parameter

In almost all applications a tuning parameter is useful to adjust the trade-off between performance versus robustness, since loop tuning depends on the users knowledge and confidence of the plants dynamics, linearities and changing conditions. Consequently, the effects of the tuning parameter should be transparent to the user.

The tuning parameter of the proposed design method is directly related to the robustness for model uncertainties, that is, the radius R of the circle in Figure 3., see Åström *et al.* (1998) and Panagopoulos *et al.* (1999). The advantages of the chosen design variable are its relationship to both the robustness for model uncertainties and the performance of the closed loop system. Furthermore, the tuning parameter is dimension-free, which makes it suitable for automatic tuning. This is to be compared to the Internal Model Control, see Rivera *et al.* (1986), whose tuning parameter has dimension time. In the Internal Model Control framework, the tuning parameter is the closed loop time constant which can be used to trade performance, in terms of the time response of the process output, versus robustness. The trade-off is not optimal in the sense that some performance objective is minimized subject to a robustness constraint, which is done in the proposed design.

Parameter Optimization

The proposed design method can be formulated as a parameter optimization problem:

$$\begin{aligned} \max \quad & k_i \\ \text{subject to} \quad & \|C + L(i\omega)\| \geq R, \end{aligned} \tag{0.2}$$

which is illustrated in Figure 4.. Consequently, we have a nonlinear optimization problem to solve, which is non-convex and requires a numerical search algorithm and a method for generating good initial values.

The are several advantages when using optimization for the design of controllers. Parameter optimization is, especially, a powerful tool for the design of PID controllers where the controller structure and the parameters to optimize are given. It allows for a much greater range of performance and robustness criteria compared to the classical Ziegler-Nichols methods. Optimization gives the ability to apply the design method to a wide variety of processes, which is to be compared to the restriction of simple parameter models.

There are, also, several pitfalls when using optimization. Care must be exercised when formulating criteria and constraints; otherwise, a criterion will indeed be optimal, but the controller may still be unsuitable due to a neglected constraint. The required computations may become excessive. Numerical problems may arise if the optimization problem is non-convex. In those cases, it is important to have a systematic approach for finding "reasonable" initial guesses of the parameter values.

The Proposed Design Method for PI Controllers

The proposed design method for PI controllers is given in Part I, which contains the paper,

I. ÅSTRÖM, K. J., H. PANAGOPOULOS, and T. HÄGGLUND (1998): "Design of PI Controllers based on Non-Convex Optimization." *Automatica,* **34:5**.

In paper I. an efficient numerical method for designing PI controllers is presented. The specifications capture demands on load disturbance rejection, set point response, measurement noise, and robustness to model uncertainties, which have been explained previously in this section. The paper shows how the optimization problem in (0.2), can be reduced to the solution of nonlinear algebraic equations by exploiting its structure. It is shown how to find efficient iterative methods to solve the problem, and good initial parameter values. For special classes of systems, for example those with a monotone transfer function, it is possible to provide good initial conditions in a systematic way.

The design procedure has been applied to a variety of systems; stable and integrating, with short and long dead times, with real and complex poles, and with positive and negative zeros. It is shown that unique solutions exists for special classes of systems, but very complicated situations may occur for complicated systems. The method will give a solution if one exists, and it will indicate when there is no PI controller satisfying the specifications.

The new contributions are: the analysis of the nature of the robustness constraint, the efficient numerical procedures, the systematic approach to find good initial parameter values, and the method to determine good set point tracking.

A natural step would be to extend the work on designing PI controllers presented in paper **I.** to the design of PD and PID controllers. For the former case, the solution to the design of PI controllers can be immediately generalized to the design of PD controllers by replacing all k_i by k_d in paper **I.**. On the other hand an immediate generalization of the solution to the design of PI controllers to the design of PID controllers is not possible.

The Proposed Design Method for PID Controllers

The proposed design method for PID controllers is given in Part II, which contains the papers,

II. PANAGOPOULOS, H., K. J. ÅSTRÖM, and T. HÄGGLUND (1999): "Design of PID Controllers Based on Constrained Optimization." *1999 American Control Conference, San Diego, California.*

IIs. ÅSTRÖM, K. J., H. PANAGOPOULOS, and T. HÄGGLUND (2000): "Supplement and Errata to "Design of PID Controllers Based on Constrained Optimization"." Department of Automatic Control, Lund Institute of Technology, Lund, Sweden.

In paper **II.**, the design of PID controllers is presented. The specifications capture demands on load disturbance rejection, set point response, measurement noise, and robustness to model uncertainties, which have been explained previously in this section.

In paper **IIs.** it is explained why a direct generalization of the optimization problem in (0.2) for the design of PI controllers, does not work well for PID controllers. It is shown how the robustness constraint in (0.2) leads to an optimization problem which, in most cases, have ridges which result in poor robustness and control. With these insights it is explained what additional constraints is needed to get good robustness and control. Thus the optimization problem for the design of ID controllers becomes a constrained optimization problem. It is shown how the problem should be solved iteratively and, how to find good initial parameter values as the problem is non-convex. The design procedure has been applied to a variety of systems: stable and integrating, with long dead times and with right half plane zeros.

The new contributions are: the analysis of the nature of the derivative part of the PID controller, the additional constraints needed to solve the design problem, the methods to determine filters for measurement noise and set point weighting.

Extension: $\mathcal{H}_\infty$ Loop Shaping

Robust $\mathcal{H}_\infty$ control is quite useful for systems under parameter perturbations, and uncertain disturbances. However, the conventional output feedback designs of the robust $\mathcal{H}_\infty$ control are very complicated. In general, the order of the controller would not be lower than that of the plant. Consequently, practical control engineers may lack incentive to employ them for industrial applications. On the other hand, PID controllers have found extensive industrial applications for several decades. It would be a desirable option if control engineers could tune the three PID controller parameters such that the corresponding controller parameters obtained by the robust $\mathcal{H}_\infty$ control would be obtained. As many different methods have been proposed for the design of PID controllers which compromise between robustness and performance, see Åström and Hägglund (1995), it would be of great interest to relate these traditional methods to the one of $\mathcal{H}_\infty$ loop shaping methods.

A proposed relation which can hopefully bridge the gap between a theoretical $\mathcal{H}_\infty$ control and classical PID controls is given in Part III, which contains the paper:

III. PANAGOPOULOS, H., and K. J. ÅSTRÖM (1999): "PID Control Design and $\mathcal{H}_\infty$ Loop Shaping." Accepted to *Robust and Nonlinear Control.*

It is shown how to interpret the mathematical $\mathcal{H}_\infty$ criteria in engineering terms. In Åström *et al.* (1998) and Panagopoulos *et al.* (1999) all the robustness constraints are encapsulated in the constraint that the loop transfer function avoids a circle in the Nyquist diagram. This is a nice feature as the Nyquist plot allows to assess the influence of the modeling errors, at the same time as appropriate specifications for the controller design are derived. Consequently, it is possible to establish a nice connection between traditional design of PID controllers and the $\mathcal{H}_\infty$.

The new contributions are: how the robustness constraint for the PI/PID design in Åström *et al.* (1998) and Panagopoulos *et al.* (1999) should be

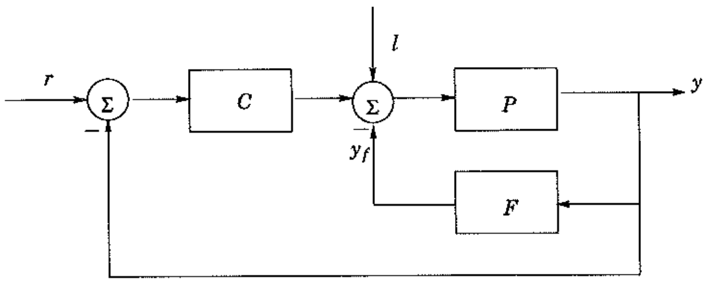


Figure 5. Control with active damping. With the inner feedback loop the active control system F rejects undamped modes of the process P. In the outer control loop the controller C gives good disturbance rejection, robustness to model uncertainties, and set point following.

chosen to guarantee that the weighted $\mathcal{H}_\infty$ norm, see Glover and McFarlane (1989), of the transfer function from load and measurement disturbances to process inputs and outputs is less than a specified value γ. Furthermore, a new way to determine for what class of systems a PID controller will be stabilizing is presented.

Application: PID Control

An efficient design method should be applicable to a wide range of systems. The proposed design method for PID controllers have been tested through extensive simulations to a variety of systems. But these results are not enough to judge the performance of a design method. Consequently, the tuning method for PID controllers have been evaluated in a benchmark for control of steam generator water level in a power plant, and at the pulp and paper company of Modo Paper, in Husum, Sweden.

The results from the evaluations are found in Part IV and V:

IV. EBORN, J., H. PANAGOPOULOS, and K. J. ÅSTRÖM (1999): "Robust PID Control of Steam Generator Water Level." In *IFAC'99 14th World Congress of IFAC.* Beijing, P. R. China.

V. PANAGOPOULOS, H., A. WALLEN , O. NORDIN, and B. ERIKSSON (2000): "A New Tuning Method with Industrial Evaluations." *Submitted to Control Systems 2000.* Canada.

Extension: Control of Undamped Process Modes

The standard PID controller have many advantages, but there are applications where significant better performance is obtained with more sophisticated controllers. For example, when the process contains long time delays, processes with undamped modes, and nonlinearities. For the case of processes with long time delays, a PID can be used for control, but slow responses will be obtained.

There are two ways to improve the performance of the closed loop system when the standard PID controller is not sufficient, either by using a modular/synthetic approach or a unimodular/analytic approach. For example, in the case of processes with long time delays a modular approach is taken when a Smith predictor is inserted into the loop to improve the performance of the closed loop system. Another example when a modular approach is employed are for processes with nonlinearities. If the characteristic of the nonlinearity is known, it is compensated by feeding signals through a function module which forms the inverse of the nonlinearity. In this thesis a modular approach is used to show how the performance of the closed loop system can be improved for processes with poorly damped modes.

In Part VI a new idea is presented for the rejection of disturbances whose frequencies lies in the same range as the undamped process modes,

VI. PANAGOPOULOS, H., and T. HÄGGLUND (2000): "A New Modular Approach to Active Control of Undamped Modes." Department of Automatic Control, Lund Institute of Technology, Lund, Sweden.

The problem formulation is illustrated in Figure 5.. It is desired to reject the disturbance l whose frequencies lies in the same range as the undamped process modes of P. The solution to the problem is to cancel out the disturbance l at the source by the insertion of the active controller F in Figure 5.. Consequently, with the inner feedback loop the active control system F rejects undamped modes of the process P, and with the outer feedback loop the controller C gives good disturbance rejection, robustness to model uncertainties, and set point following of the closed loop system. As an modular approach is taken it will not complicate the original tuning work of the outer feedback loop as the active controller F is independent of the controller structure and design of C.

The active controller is the product of a bandpass and an allpass filter. To determine the parameters of these two filters the only information needed is a few characteristics of the frequency response of the process P, see Figure 6.. The idea has been applied to simulation examples, and been compared with other methods.

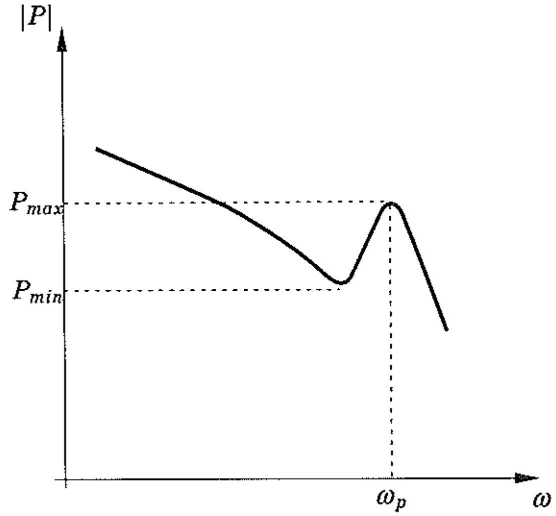


Figure 6. The information of the frequency response of process P which is needed to determine the active controller F.

Conclusions

This thesis deals with various aspects of PID control. It has considered the design of PID controllers, the extensions of these controllers to improve their performance, and the applications of these design methods to industrial processes.

New tuning methods for PI and PID controllers have been presented. These methods uses a model of the process to be controlled, given as a transfer function. The design methods captures essential requirements of a control system, such as:

- Load disturbance response
- Robustness with respect to model uncertainties
- Measurement noise response
- Set point response

The formulation of the design problems can loosely be divided into two categories: specifications on performance and robustness. The primary design goal of the proposed design methods is to achieve good rejection of a load disturbance with the constraint on robustness to guarantee the stability of the closed loop system. The presented controller tuning methods are suitable for supervisory system, since they can handle the different kinds of process models which may come up. This will improve the development of automatic tuners.

To extend the use of PID control a relation which bridge the gap between it and the theoretical $\mathcal{H}_\infty$ control have been proposed. It is shown how the robustness constraint for the PI/PID design in Åström *et al.* (1998) and Panagopoulos *et al.* (1999) should be chosen to guarantee that the weighted $\mathcal{H}_\infty$ norm of the transfer function from load and measurement disturbances to process inputs and outputs is less than a specified value γ. A new way to determine for which class of systems a PID control will be stabilizing is also presented.

Furthermore, the use of PID control have been extended to handle processes with undamped modes. A modular approach has been taken, where an active control system has been designed, which consists of an allpass filter and a bandpass filter. To determine the parameters of these two filters the only information needed is a few characteristics of the process frequency response.

The proposed design methods for PI, and PID controllers have been tested through extensive simulations to a variety of systems which is not enough to give a true judgment of their performance. Consequently, the tuning methods have been evaluated in a benchmark for control of steam generator water level in a power plant, and at the pulp and paper company Modo Paper, in Husum, Sweden.

Future Work

The design of PID controllers are at the present time going through a very interesting phase as the tuning problem of them have been recognized, at least in the pulp and paper industry, during the 1990's. According to Bialkowski (1996) three key needs have been identified: improved tuning techniques, adequate training, and suitable tools. For example, in the the pulp and paper industry the Lambda tuning technique, see Rivera *et al.* (1986), have been adopted and chosen as a standard in the control education of the Swedish pulp and paper organization "Skogsindustriernas Teknik AB". In this thesis, ideas have been presented on how to improve control, but there is, still, more to be done. A couple of ideas for future work of the presented thesis will follow below.

The two presented controller tuning methods assumed the transfer function of the process to be known. Would these design methods, still, work when less process information is available? For example, when the information of the process model is based on the results of the relay method in Åström and Hägglund (1984). Furthermore, the optimization problem of the design of the PID controller was based on three constraints.

Are there other constraints which would give a better solution? Moreover, the proposed design methods was based on single input single output systems and on a continuous time approach. Will it be possible to extend these design methods to multi input multi output systems? What are the limitations to take a discrete approach of the presented controller design methods compared to the continuous one? Finally, it would be interesting to have a systematic way to judge when a PI or a PID controller should be used. For example, for systems with large time delays PID control will not give enhanced performance compared to the PI.

The aim of this thesis was: "How to improve control?". The idea was to expand the usefulness of PI and PID controllers viewed as building blocks. It would be of interest to refine this view for the following examples: How should nonlinearities be treated? How to expand the use of the proposed active controller for undamped process modes? How to extend the relation between the presented controller designs to the $\mathcal{H}_\infty$ design for multi input multi output systems. Finally, it would be interesting to apply all of these ideas to real systems.

References

Andreas, H. and K. J. Åström (1997): "Design of PI controller by minimization of IAE." Report ISRN LUTFD2/TFRT--7565--SE. Department of Automatic Control, Lund Institute of Technology, Lund, Sweden.

Åström, K. J. and T. Hägglund (1984): "Automatic tuning of simple regulators with specifications on phase and amplitude margins." *Automatica*, **20**, pp. 645–651.

Åström, K. J. and T. Hägglund (1995): *PID Controllers: Theory, Design, and Tuning*. Instrument Society of America, Research Triangle Park, North Carolina.

Åström, K. J., H. Panagopoulos, and T. Hägglund (1998): "Design of PI Controllers based on Non-Convex Optimization." *Automatica*, **34:5**, pp. 585–601.

Bialkowski, B. (1996): "As We See It." The Entech Report. Volume 8.

Bialkowski, B. (1997): "World-Class Manufacturing: If you are not improving, odds are you are losing ground." The Entech Report. Volume 9.

Boyd, S. P. and C. H. Barratt (1991): *Linear Controller Design – Limits of Performance*. Prentice Hall Inc., Englewood Cliffs, New Jersey.

Cohen, G. H. and G. A. Coon (1953): "Theoretical consideration of retarded control." *Trans. ASME*, **75**, pp. 827–834.

Eborn, J., H. Panagopoulos, and K. J. Åström (1999): "Robust PID Control of Steam Generator Water Level." In *IFAC'99, 14th World Congress of IFAC*, vol. G, pp. 461–464. Beijing, P.R. China.

Glover, K. and D. McFarlane (1989): "Robust stabilization of normalized coprime factor plant descriptions with H_∞-bounded uncertainty." *IEEE Transactions on Automatic Control*, **34:8**, pp. 821–830.

Hägglund, T. and K. J. Åström (1985): "Automatic tuning of PID controllers based on dominant pole design." In *Proceedings of the IFAC Conference on Adaptive Control of Chemical Processes*, pp. 212–217. Frankfurt, Germany.

Horowitz, I. M. (1963): *Synthesis of Feedback Systems.* Academic Press, New York.

Kaya, A. and T. J. Scheib (1988): "Tuning of PID controls of different structures." *Control Engineering*, July, pp. 62–65.

Kristiansson, B. and B. Lennartsson (1999): "Optimal PID Controllers Including Roll Off and Schmidt Predictor Structure." In *IFAC'99, 14th World Congress of IFAC*, vol. F, pp. 297–302. Beijing, P.R. China.

Langer, J. and I. D. Landau (1999): "Combined pole placement/sensitivity function shaping method using convex optimization criteria." *Automatica*, **35:6**, pp. 1111 – 1120.

Lopez, A. M., J. A. Miller, C. L. Smith, and P. W. Murrill (1967): "Tuning controllers with error-integral criteria." *Instrumentation Technology*, November, pp. 57–62.

Maciejowski, J. M. (1999): *Predictive Control with Constraints.* To be published.

Mayne, D. Q. and E. Polak (1993): "Optimization based design and control." In *Preprints IFAC 12th World Congress*, vol. 3, pp. 129–138. Sydney, Australia.

Newton, Jr, G. C., L. A. Gould, and J. F. Kaiser (1957): *Analytical Design of Linear Feedback Controls.* John Wiley & Sons.

Panagopoulos, H. and K. J. Åström (2000): "PID Control Design and H_∞ Loop Shaping." *Robust and Nonlinear Control.* Accepted for publication.

Panagopoulos, H., K. J. Åström, and T. Hägglund (1999): "Design of PID Controllers based on Constrained Optimization." In *1999 American Control Conference*, pp. 3858–3862. San Diego, California.

Panagopoulos, H., K. J. Åström, and T. Hägglund (2000): "Supplement and Errata to Design of PID Controllers based on Constrained Optimization."

Panagopoulos, H. and T. Hägglund (2000): "A New Modular Approach to Active Control of Undamped Modes." Report ISRN LUTFD2/TFRT--7589--SE. Department of Automatic Control, Lund Institute of Technology, Lund, Sweden.

Persson, P. (1992): *Towards Autonomous PID Control.* PhD thesis ISRN LUTFD2/TFRT--1037--SE, Department of Automatic Control, Lund Institute of Technology, Lund, Sweden.

Rivera, D. E., M. Morari, and S. Skogestad (1986): "Internal model control—4. PID controller design." *Ind. Eng. Chem. Proc. Des. Dev.*, **25**, pp. 252–265.

Schei, T. S. (1994): "Automatic tuning of PID controllers based on transfer function estimation." *Automatica*, **30:12**, pp. 1983–1989.

Shigemasa, T., Y. Iino, and M. Kanda (1987): "Two degrees of freedom PID auto-tuning controller." In *Proceedings of ISA Annual Conference*, pp. 703–711.

Shinskey, F. G. (1990): "How good are our controllers in absolute performance and robustness?" *Measurement and Control*, **23**, May, pp. 114–121.

Smith, C. L., A. B. Corripio, and J. J. Martin (1975): "Controller tuning from simple process models." *Instrumentation Technology*, December, pp. 39–44.

Truxal, J. (1955): *Automatic Feedback Control System Synthesis.* McGraw-Hill, New York.

Van Overschee, P. and B. D. Moor (1999): "Optimal PID Control of Chemical Batch Reactor." In *Proceedings of the 1999 European Control Conference.* Karlsruhe, Germany.

Wallén, A. (1999): "A Tool for Rapid System Identification." In *1999 Conference on Control Applications*, pp. 1555–1560. Kohala Coast Island of Hawaii, Hawaii.

Zhuang, M. and D. Atherton (1991): "Tuning PID controllers with integral performance criteria." In *Control '91.* Heriot-Watt University, Edinburgh, UK.

Ziegler, J. G. and N. B. Nichols (1942): "Optimum settings for automatic controllers." *Trans. ASME*, **64**, pp. 759–768.

I

Pergamon

PII: S0005-1098(98)00011-9

Automatica, Vol. 34, No. 5, pp. 585–601, 1998

Printed in Great Britain
0005-1098/98 $19.00 + 0.00

Design of PI Controllers based on Non-Convex Optimization*

K. J. ÅSTRÖM†, H. PANAGOPOULOS† and T. HÄGGLUND†

Key Words—PI control; design; optimization; specifications; load disturbance rejection; set point response; robustness; sensitivity.

Abstract—This paper presents an efficient numerical method for designing PI controllers. The design is based on optimization of load disturbance rejection with constraints on sensitivity and weighting of set point response. Thus, the formulation of the design problem captures three essential aspects of industrial control problems, leading to a non-convex optimization problem. Efficient ways to solve the problem are presented. © 1998 Elsevier Science Ltd. All rights reserved.

1. INTRODUCTION

The PI controller is unquestionably the most commonly used control algorithm; see Yamamoto and Hashimoto (1991). In spite of its wide spread use there exists no generally accepted design method for the controller. PI controllers have traditionally been tuned empirically, e.g. by the method described in Ziegler and Nichols (1942). This method has the great advantage of requiring very little information about the process. There is, however, a significant disadvantage because the method inherently gives very poor damping, typically $\zeta \approx 0.2$, see Åström and Hägglund (1995b).

There are several reasons to look for better methods to design PI controllers. One reason is the significant impact it may give because of the widespread use of the controllers. Another reason is that emerging auto-tuners and tuning devices can benefit significantly from improved design methods.

There are several requirements on an efficient design method. It should be applicable to a wide range of systems and it should have the possibility to introduce specifications that capture the essence of real control problems. Furthermore, the method should be robust in the sense that it provides controller parameters if they exist, or if the specifications cannot be met an appropriate diagnosis should be presented. We believe these requirements are satisfied by the method presented in this paper. The approach gives a simple way to solve simple control problems and more difficult ones where more efforts are needed. The method will give a PI controller which satisfies the specifications, provided that such a controller exists. If not, the reasons for failure will be indicated. It can also be used to develop simpler methods for restricted classes of systems as was done in Åström and Hägglund (1995a).

The method proposed in this paper formulates the design problem as an optimization problem: optimize the load disturbance rejection with a constraint on the maximum sensitivity. By exploiting the structure of the optimization problem it is reduced to the solution of algebraic equations. Efficient iterative methods are given together with good methods for finding starting values.

The method presented assumes a linear process whose dynamics is characterized in terms of a transfer function, which does not have to be rational. Thus, it can be applied to systems described by partial differential equations. If the transfer function is not known it can be obtained by system identification.

Parts of the proposed method builds on previous works. The use of optimization was discussed in Hazebroek and van der Waerden (1950). In this and other early works the emphasis was on criteria that admitted analytical solutions. The idea of optimizing load disturbance rejection with sensitivity constraints was suggested by Shinskey (1990). He used a constraint in terms of a rectangle around the critical point but the idea to use a constraint on the maximum sensitivity, M_s, was proposed by Persson (1992) and Persson and Åström (1992). The use of both M_s and the maximum complementary sensitivity, M_p, as design parameters was suggested by Schei (1994). The new contributions in this paper are: the analysis of the nature of the sensitivity constraints, the efficient numerical procedures, and the method for determining set point weighting.

2. FORMULATION OF THE DESIGN PROBLEM

The formulation of a design problem includes a characterization of the process and its

*Received 24 March 1997; received in final form 10 December 1997. This paper was recommended for publication by Associate Editor C. P. Jobling under the direction of Editor Sigurd Skogestad. *Corresponding author* Professor K. J. Åström. Tel. +46-46 2228781; Fax +46-46 138118; E-mail kja@control.lth.se.

† Department of Automatic Control, Lund Institute of Technology, Box 118, S-221 00 Lund, Sweden.

environment, the controller structure, and specifications on the performance of the closed-loop system.

Requirements: Before going into details we will first discuss the requirements. The method should be applicable to a wide range of systems and it should be based on specifications that reflect the essence of real control problems. Furthermore, the method should give a simple solution to simple problems. With more efforts and more skills from the designer it should also be possible to sharpen the specifications.

The process: It is assumed to be linear, and specified by a transfer function $G(s)$ which is analytical with finite poles and possibly an essential singularity at infinity. This description covers finite dimensional systems with time delays and infinite dimensional systems described by linear partial differential equations.

The controller: It is a PI controller described by

$$u(t) = k(by_{sp}(t) - y(t)) + k_I \int_0^t (y_{sp}(\tau) - y(\tau))\,d\tau, \quad (1)$$

where $u(t)$ is the control signal, $y(t)$ is the process output, $y_{sp}(t)$ is the set point, and k, k_I, and b are controller parameters.

When there are substantial measurement noise, it is customary to filter the measurement signal with a filter, typically of the form

$$G_f(s) = \frac{1}{1 + sT_f}, \quad (2)$$

where T_f is the filter time constant.

Thus, the controller has three or possibly four parameters, k, k_i, b, and T_f. It is industry practice to use integration time, defined as $T_i = k/k_i$, instead of parameter k_i. However, for the computations it is more convenient to use k_i. Industrial controllers typically use either $b = 0$ or $b = 1$, but lately it has been recognized that it is advantageous to use full range of b-values, that is $0 \leq b \leq 1$. The controller given by equation (1) is said to have two degrees of freedom when $b \neq 1$. The advantage of such structures has been pointed out by Horowitz (1963) and their use in PID controllers is discussed in Shigemasa *et al.* (1987) and Åström and Hägglund (1995b).

Specifications: They express requirements on

- load disturbance response,
- set point response,
- robustness with respect to model uncertainties.

In process control applications efficient rejection of load disturbances is of primary concern, whereas set point responses are typical of secondary importance. However, set point response may be of primary importance, for example, in motion control systems. Although it has been frequently pointed out by engineers that load disturbances is of primary concern, it is interesting to note that papers on PI control traditionally focus on set point response, see e.g. Shinskey (1990).

The sensitivity to model uncertainty is of primary significance, and observe that the poor sensitivity is one of the major drawbacks of the classical Ziegler–Nichols method.

In virtually all applications it is useful to have a tuning parameter that permits adjustment of the trade off of aggressiveness versus robustness. The effects of the tuning parameter should be transparent to the user. In the method presented, we have a tuning parameter to specify the sensitivity to model uncertainties.

2.1. *A formal description*

In order to use a formal design method it is necessary to capture specifications in a suitable mathematical form. This is extensively discussed in Åström and Hägglund (1995b).

2.1.1. *Load disturbance rejection.* It can be conveniently expressed in terms of the integrated absolute error due to a load disturbance in the form of a unit step at the process input, i.e.

$$IAE = \int_0^\infty |e(t)|\,dt. \quad (3)$$

This criterion is difficult to deal with analytically because the evaluation requires computation of time functions. The integrated error defined by

$$IE = \int_0^\infty e(t)dt, \quad (4)$$

is much more convenient. In Åström and Hägglund (1995b) it is shown that $IE = 1/k_i$. Thus, the criterion IE is directly given by the integrating gain of the controller. Remember that, $IE = IAE$ if the error is positive. Furthermore, if the system is well damped the criteria will be close which, in our case, will be ensured by the sensitivity constraints.

2.1.2. *Sensitivity to modeling errors.* They can be expressed in terms of the largest value of the sensitivity function. Let the loop transfer function be $L(s) = G(s)G_c(s)$ where G_c is the controller transfer function, and let the sensitivity function be $S(s) = 1/(1 + L(s))$. The maximum sensitivity is then given by $M_s = \max|S(i\omega)|$. Keep in mind that the quantity M_s is the inverse of the shortest distance from the Nyquist curve of the loop transfer function to the critical point -1. Typical values of M_s are in the range of 1.2–2.0.

With a constraint on M_s it follows from the circle criterion that the closed loop system will also remain stable with a static nonlinearity in the loop, provided that the nonlinearity is bounded by two straight lines with slopes $M_s/(M_s+1)$ and $M_s/(M_s-1)$, see Khalil (1992).

Let $T(s) = 1 - S(s) = L(s)/(1 + L(s))$ be the complementary sensitivity function. The sensitivity can also be expressed by the largest value of the complementary sensitivity function, i.e. $M_p = \max|T(i\omega)|$. The value M_p is the size of the resonance peak of the closed loop system obtained with $b = 1$, see equation (1). Typical values of M_p are in the range of 1.0–1.5.

2.1.3. *Set point response.* The design has so far focused on the response to load disturbances, which is of primary concern. However, it is also important to have a good response to set point changes. The transfer function from set point to process output is given by

$$\frac{Y(s)}{Y_{sp}(s)} = \frac{k_i + bks}{k_i + ks}\frac{L(s)}{1 + L(s)} = G_{sp}(s). \tag{5}$$

One way to give specifications on the set point response is to specify the resonance peak of the transfer function $G_{sp}(s)$, i.e.

$$M_{sp} = \max|G_{sp}(i\omega)|. \tag{6}$$

Consequently, the b-value is determined such as it fulfills equation (6). Notice that $M_{sp} < M_p$ if $b < 1$.

2.1.4. *Measurement noise filtering.* The PI controller has a high-frequency gain k. The effects of measurement noise can be reduced, substantially, by filtering the signal with the filter given by equation (2). The specifications can be expressed in terms of the magnitude of the high-frequency gain of the controller. This specification is optional and only used in exceptional cases.

2.2. *Design parameters*

The tradeoff between performance and robustness varies between different control problems. Therefore, it is desirable to have a design parameter to change the properties of the closed-loop system. Ideally, the parameter should be directly related to the performance of the system, it should not be process oriented. There should be good default values so a user is not forced to select some value. This is of special importance when the design procedure is used for automatic tuning. The design parameter should also have a good physical interpretation and natural limits to simplify its adjustment.

The variables M_s and M_p are both possible candidates for design variables. The influence of these design parameters is illustrated in Fig. 1. The curves to the left show that M_s is a suitable design parameter. Decreasing values of M_s results in time responses that are slower but less oscillatory. The curves to the right show that M_p is not suitable. Although M_p is varied between 1.2 and 2.0, no significant variation in the time responses is noticed. The reason is, that the M_p circles are too close to each other in the frequency region around $-180°$. Therefore, we will choose M_s as a design variable.

On the other hand, it is important that the resulting M_p value is not too large. We will therefore also calculate M_p when the design is completed. If M_p is too large there are several possibilities. One is to repeat the design with a smaller M_s value. Another is to use constraints on both M_s and M_p. This will give rise to difficulties in the optimization because the set enclosed by M_s and M_p circles is not convex. This difficulty can be avoided by constructing a circle that has the M_s and M_p circles in its interiors. A straightforward calculation shows that this is a circle with a center C and a radius R where

$$C = \frac{M_s - M_sM_p - 2M_sM_p^2 + M_p^2 - 1}{2M_s(M_p^2 - 1)},$$
$$R = \frac{M_s + M_p - 1}{2M_s(M_p - 1)}. \tag{7}$$

An advantage with this is that the same optimization procedure can be used because the constraint set is a circle.

3. THE OPTIMIZATION PROBLEM

The design problem discussed in the previous section can be formulated as an optimization problem: Find controller parameters that maximize k_i subject to the constraints that the closed-loop system is stable and that the Nyquist curve of the loop transfer function satisfies the encirclement condition and that it is outside a circle with center at $s = -C$ and radius R.

Introduce $L(s) = (k + k_i/s)G(s)$ and the function

$$f(k, k_i, \omega) = |C + \left(k - i\frac{k_i}{\omega}\right)G(i\omega)|^2. \tag{8}$$

The sensitivity constraint can then be expressed as

$$f(k, k_i, \omega) \geq R^2 \tag{9}$$

and the optimization problem is to maximize k_i subject to the sensitivity constraint (9).

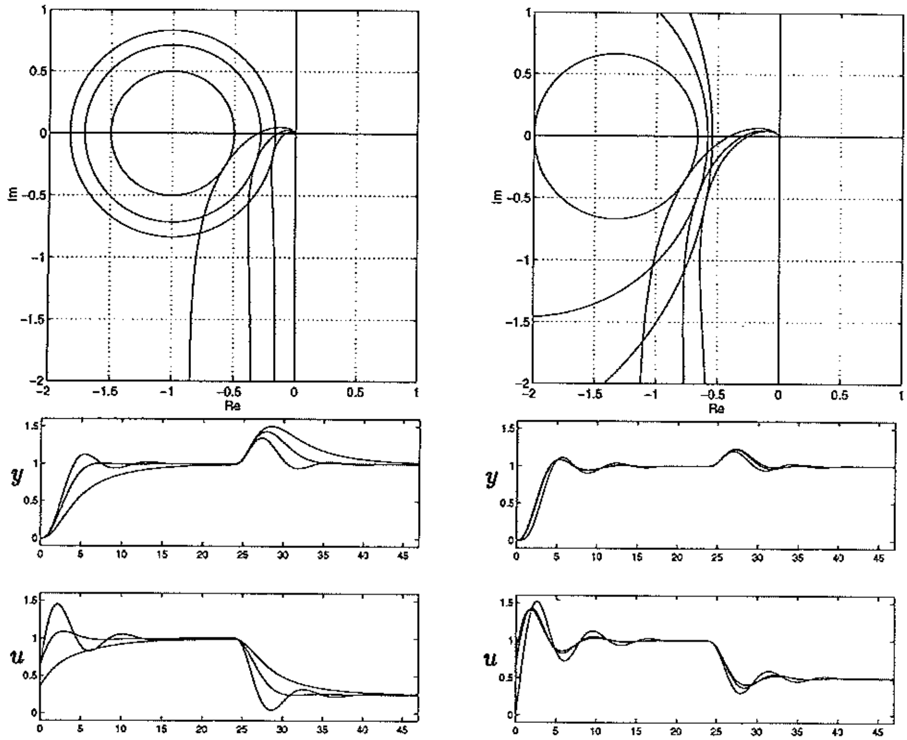


Fig. 1. Illustrates the effects of using M_s (left) and M_p (right) as a design parameter. The upper curves show the Nyquist curves of the loop transfer functions together with the M_s and M_p circles for $M_s = 1.2, 1.4, 2.0$ and $M_p = 1.2, 1.4, 2.0$. The lower curves show process outputs and control signals for the different design parameters. They give responses to a set point change followed by a load disturbance.

Let $\alpha(\omega)$ and $\beta(\omega)$ be the real and imaginary parts of the process transfer function. Hence

$$G(\mathrm{i}\omega) = \alpha(\omega) + \mathrm{i}\beta(\omega) = r(\omega)\mathrm{e}^{\mathrm{i}\varphi(\omega)},$$

where

$$\alpha(\omega) = r(\omega)\cos\varphi(\omega),$$

$$\beta(\omega) = r(\omega)\sin\varphi(\omega).$$

The function f can then be written as

$$f(k, k_i, \omega) = C^2 + 2C\alpha(\omega)k + 2C\frac{\beta(\omega)}{\omega}k_i + r^2(\omega)k^2 + \frac{r^2(\omega)}{\omega^2}k_i^2. \quad (10)$$

In the following, we will occasionally drop the argument ω in α, β, r, and φ in order to simplify the writing.

The optimization problem is nontrivial because the constraint, which is infinite dimensional, defines a set in parameter space which is not convex. There are also other subtleties which may cause problems. For a specific problem it is not difficult to solve the problem numerically with standard optimization routines because the search range can often be limited and if the optimization fails it is possible to interfere manually. Since PI controllers are very common it is, however, worthwhile to make special algorithms which are tailored for the problem. Such procedures are also required for automatic tuning where manual interaction is very inconvenient.

Before discussing the solution to the optimization problem we will investigate the sensitivity constraint which is a key difficulty of the problem.

3.1. *The sensitivity constraint*

The sensitivity constraint given by equation (9) has a nice geometric interpretation. For fixed ω, equation (9) represents the exterior of an ellipse in the k–k_i plane. The ellipse has its axes parallel to the coordinate axes. For $0 \leq \omega < \infty$ the ellipses generate envelopes that define the boundaries of the sets of parameters which satisfy the sensitivity constraints. It can be assumed that the process transfer function is such that a stable closed system is obtained with positive k_i. It is thus sufficient to consider the upper-half of the k–k_i plane. The center of

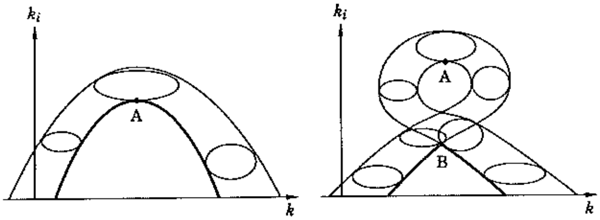


Fig. 2. Geometrical illustration of the sensitivity constraint (9) and the envelope generated by it. The envelope to the left has a continuous derivative, but the one to the right has a corner.

the ellipses generate the boundary of the stability regions. The envelopes for different systems may have different characters as shown in Fig. 2. Notice in particular that the envelope may have a corner as is illustrated in Fig. 2b.

3.1.1. *Stability.* The stability region can be expressed in terms of a condition on k and k_i which is obtained by setting $C = 1$ and $R = 0$ in equation (9), i.e.

$$k\alpha(\omega) + k_i \frac{\beta(\omega)}{\omega} + i\left(k\beta(\omega) - k_i \frac{\alpha(\omega)}{\omega}\right) = -1. \tag{10}$$

Hence

$$k\alpha(\omega) + k_i \frac{\beta(\omega)}{\omega} = -1,$$

$$k\beta(\omega) - k_i \frac{\alpha(\omega)}{\omega} = 0.$$

Solving these equations for k and k_i gives the following parametric description of the boundary of the stability region:

$$k = -\frac{\alpha(\omega)}{r^2(\omega)},$$
$$k_i = -\omega \frac{\beta(\omega)}{r^2(\omega)}, \tag{11}$$

where $r^2 = \alpha^2(\omega) + \beta^2(\omega)$, and the parameter ω ranges from zero to infinity. Notice that the stability region may consist of disjoint sets.

3.2. *Optimization*

Having understood the nature of the constraints it is now conceptually clear how to solve the optimization problem. It is simply a matter of finding the largest value of k_i on the envelope. The difficulties that may occur are due to the fact that there may be several local maxima and that the maximum may occur at a corner.

Since it is quite time consuming to generate the envelope it is desirable trying to find algorithms that can give a more effective solution. It is also of interest to characterize the situations when there is only one local minimum.

The envelope is given by

$$f(k, k_i, \omega) = R^2,$$
$$\frac{\partial f}{\partial \omega}(k, k_i, \omega) = 0. \tag{12}$$

Since the function f is quadratic in k_i the envelope has two branches. Only one branch corresponds to stable closed-loop systems. Compare with Fig. 2.

Instead of generating the envelopes and searching for the largest value of k_i on the envelope, we will characterize the points where k_i has its largest value. From the discussion of the constraint it is clear that there are two cases. The simplest case is when the largest value of k_i occurs at a point where the envelope has a continuous derivative. The other case is when the envelope has a corner.

The envelope given by equation (12) defines implicitly k_i as a function of k. To find the maximum of this function we observe that

$$\mathrm{d}f = \frac{\partial f}{\partial k}\mathrm{d}k + \frac{\partial f}{\partial k_i}\mathrm{d}k_i + \frac{\partial f}{\partial \omega}\mathrm{d}\omega = 0. \tag{13}$$

It follows from equation (12) that $\partial f/\partial\omega = 0$ on the envelope. At a local extremum we have $\mathrm{d}k_i = 0$. For arbitrary variations of $\mathrm{d}k$ we must thus require that $\partial f/\partial k = 0$. Combining this with the envelope conditions (12) we get

$$\frac{\partial f}{\partial k}(k, k_i, \omega) = 0,$$
$$\frac{\partial f}{\partial \omega}(k, k_i, \omega) = 0, \tag{14}$$
$$f(k, k_i, \omega) = R^2,$$

which corresponds to the situation when the maximum occurs at a point on the envelope where it has continuous derivatives, see Fig. 2a. In the Nyquist diagram this agrees with the case when the loop transfer function is tangent to the circle at one point.

We will now consider the case when the largest value of k_i occurs at a point where the envelope has a corner. This occurs at the intersection of ellipses corresponding to two different frequencies, ω_1 and ω_2, see Fig. 2b. The envelope condition (12) is then satisfied for both frequencies. This gives the condition

$$\begin{aligned} f(k, k_i, \omega_1) &= R^2, \\ \frac{\partial f}{\partial \omega}(k, k_i, \omega_1) &= 0, \\ f(k, k_i, \omega_2) &= R^2, \\ \frac{\partial f}{\partial \omega}(k, k_i, \omega_2) &= 0. \end{aligned} \tag{15}$$

In the Nyquist diagram this corresponds with the case when the loop transfer function is tangent to the circle at two points.

It is thus possible to characterize the point where k_i has its largest value by algebraic equations. This means that the design problem is reduced to solving algebraic equations (14) or (15), and that elaborate search procedures are avoided. Both equations can be solved by using the Newton–Raphson method. Equation (14) which represents the most common situation can, however, be simplified substantially.

3.3. *A simplification*

Equation (14) is a nonlinear equation in three variables, k, k_{ii} and ω. It is possible to solve this equation directly, but a much more efficient algorithm will be obtained by eliminating some variables.

Inserting expression (10) into equation (14), gives

$$\begin{aligned} \frac{\partial f}{\partial k} &= 2C\alpha + 2r^2 k = 0, \\ \frac{\partial f}{\partial \omega} &= 2C\left(\frac{\beta}{\omega}\right)' k_i + 2C\alpha' k \\ &\quad + \left(\frac{r^2}{\omega^2}\right)' k_i^2 + 2rr'k^2 = 0, \\ f &= R^2, \end{aligned} \tag{16}$$

where prime means differentiation with respect to ω. Solving k and k_i from the first and last equations gives

$$\begin{aligned} k &= -C\frac{\alpha}{r^2} = -C\frac{1}{r}\cos\varphi, \\ k_i &= -\frac{\omega\beta C}{r^2} - \frac{\omega R}{r} = -\frac{\omega}{r}(C\sin\varphi + R), \end{aligned} \tag{17}$$

where the positive sign is chosen to satisfy the encirclement criterion. The following condition is obtained by inserting the expressions of k and k_i in the second equation of equation (16),

$$\begin{aligned} h(\omega) &= 2R\left(\left(C\frac{\beta}{r} + R\right)\left(\frac{r'}{r} - \frac{1}{\omega}\right) - C\left(\frac{\beta}{r}\right)'\right) \\ &= 2R\left((R + C\sin\varphi)\left(\frac{r'}{r} - \frac{1}{\omega}\right) - C\varphi'\cos\varphi\right) \\ &= 0. \end{aligned} \tag{18}$$

Thus, the solution to equation (16) is reduced to a single algebraic equation (18) in ω. Solving it gives the frequency ω_0 for which we can compute the controller gains k and k_i given by equation (17).

The condition (16) does not tell if the extremum is a minimum, a maximum or saddle point, but constraint (9) implies that the function should be a minimum with respect to ω. This gives the following local condition:

$$\frac{d^2 f}{d\omega^2}(\omega_0) > 0. \tag{19}$$

Equation (18) can be solved iteratively with the Newton–Raphson method which converges very fast if suitable initial conditions are given. Notice, however, that in general there may be several solutions which can be found by starting the iteration from different initial conditions.

For special classes of systems, for example systems with monotonic transfer function, it is possible to provide good initial conditions. The following result is useful.

Theorem 1. Let ω_φ denote the frequency where the process has a phase of φ. Assume that the transfer function $G(s)$ has positive low-frequency gain and that

$$\begin{aligned} \frac{d \arg G(i\omega)}{d\omega} &< 0, \\ \frac{d \log_{10}|G(i\omega)|}{d \log_{10}\omega} &< 1. \end{aligned} \tag{20}$$

Then there exists a solution to equation (18) in the interval

$$\omega_{90} < \omega < \omega_\dagger = \omega_{180 - \arcsin R/C}. \tag{21}$$

Proof. It is assumed that the low-frequency gain of the process is positive. Then the integral gain k_i must be positive if the closed-loop system should be stable. For this reason equation (17) implies that $R + C\sin\varphi < 0$. It follows from the assumptions (20) that $\varphi' < 0$ and $r'/r - 1/\omega < 0$. As a result equation (18) and the assumptions imply that $h(\omega_{90}) > 0$ and $h(\omega_\dagger) < 0$. But h is a continuous function, therefore equation (18) must have a solution in the interval (21). □

Remark 1. Notice that since a PI controller has negative phase for all frequencies, the solution must be found in the interval $\omega_{90} < \omega < \omega_{\dagger}$. Thus, the monotonicity condition has to be valid in that interval.

Remark 2. It follows from equation (17) that the condition (21) implies that both k and k_i are positive. Systems with monotone transfer functions satisfy the condition (20). For such systems it is thus straightforward to find a good range of frequencies to solve equation (18). Also notice that as long as the phase decreases monotonically it is possible to have an increase in the amplitude curve, provided that the slope is never larger than one. Most systems encountered in process control satisfy these conditions.

A necessary condition for stability of the closed-loop system is that the parameter k_i is positive which implies that $\omega_0 > \omega_{90}$.

3.4. *Initial conditions*

Good initial conditions is a crucial factor for the computational efficiency of the Newton–Raphson iterations when solving the optimization problem. The two cases when the envelope has a continuous derivative or a corner has to be treated separately.

First we consider the case when the envelope has no corner. The solution to the optimization problem is obtained by solving the algebraic equation (18) with Newton–Raphson for a suitable search range. According to Theorem 1 the search range is chosen as $[\omega_{90}, \omega_{\dagger}]$. To obtain better initial conditions we narrow it by applying interval bisection in equation (18).

Finally, the case when the envelope has a corner is considered. This problem is more difficult to solve, because it is necessary to solve a system of equations (15), which in this case requires four initial conditions of ω_1, ω_2, k and k_i. They can be obtained with the following procedure.

For fixed values of ω the sensitivity constraint, equation (12), represents ellipses in gain space which generates envelopes. It is quite complicated to compute the envelopes, however, they can be approximated by the loci of the vertices of the ellipses. The horizontal vertices are given by

$$k = -\frac{\alpha C}{r^2} \pm \frac{R}{r}, \qquad k_i = -\frac{\omega\beta C}{r^2}, \tag{22}$$

where the left vertex corresponds to a minus sign and the right vertex to a plus sign. The loci of the vertices define curves in the gain space that enclose the envelope. Good initial values for the Newton–Raphson iteration can be obtained by finding the point on the loci where k_i has its largest value. When there is a corner this value is obtained as the intersection of the vertices given by equation (22), see Fig. 2b, giving ω_1 and ω_2. Thus, k can be computed from equation (22) for either ω_1 or ω_2. Because of the construction this value overestimates the integral gain.

Note that the same technique can be used as an alternative way of finding initial conditions in the case of no corners. In this case the envelope can be approximated by the loci of the lowest vertex of the ellipse, that is

$$k = -\frac{\alpha C}{r^2}, \qquad k_i = -\frac{\omega\beta C}{r^2} - \frac{\omega R}{r}. \tag{23}$$

Notice that this equation is identical to equation (17). The upper vertex is not of any interest because it gives an unstable closed-loop system.

4. THE DESIGN PROCEDURE

We have thus found efficient procedures to determine feedback gains k and k_i, by optimizing load disturbance rejection subject to constraints on sensitivity to model uncertainties. To complete the design procedure, it remains to determine the set point weighting, i.e. parameter b in equation (1).

4.1. *Set point weighting*

The set point response is governed by the transfer function G_{sp} given by equation (5). In order to have a small overshoot in set point response, set point weighting b will be determined so that $M_{sp} = \max|G_{sp}(i\omega)|$ is close to one. It follows from equations (5) that $M_{sp} \le M_p$ when $0 \le b \le 1$. A bound of M_{sp} is thus given indirectly through M_p.

We will make the approximation that maximum of $|G_{sp}(i\omega)|$ occurs for ω_{mp}, where $\omega = \omega_{mp}$ is the frequency where the maximum of $|L(i\omega)/(1 + L(i\omega))|$ occurs. Parameter b will be determined so that

$$|G_{sp}(i\omega_{mp})| = 1 \tag{24}$$

with the constraint $0 \le b \le 1$. Using equations (5) and (24) this implies that

$$b = \begin{cases} \dfrac{\sqrt{k^2\omega_{mp}^2 - k_i^2(M_p^2 - 1)}}{k\omega_{mp}M_p} & \text{if } (\omega_{mp}k/k_i)^2 \ge M_p^2 - 1, \\ 0 & \text{if } (\omega_{mp}k/k_i)^2 < M_p^2 - 1. \end{cases} \tag{25}$$

If $b = 0$, it is not sure that the design objective (24) will be obtained. If the set point response is important and the M_p value is large, the design can then be repeated with a lower value of M_s or using the constraint given by the circle (7).

4.2. *Measurement noise filtering*

Having performed a design of the controller we obtain the controller gain k, which is also the high-frequency gain of the controller. Combined with the specifications on roll-off we can then determine the order and the bandwidth of a suitable noise filter.

Here we will only consider the first-order filter given by equation (2). The choice of filter-time constant T_f in equation (2) is a tradeoff between filtering capacity and loss of performance. A large value of T_f provides an effective noise filtering, but it will also change the control performance. A small value of T_f means that the control performance is retained, but the noise filtering is less effective.

A nice feature of the new design procedure is that it provides a systematic way to determine T_f. The choice

$$T_f = \frac{1}{m\omega_0}, \tag{26}$$

makes it possible to determine the effects of the filter at the frequency ω_0. We get

$$|G_f(i\omega_0)| = \frac{1}{\sqrt{1 + 1/m^2}},$$

$$\arg G_f(i\omega_0) = -\arctan(1/m).$$

Reasonable values of m are in the interval 5–10. For $m = 5$ we get $|G_f(i\omega_0)| = 0.981$ and $\arg G_f(i\omega_0) = -11°$. For $m = 10$ we get $|G_f(i\omega_0)| = 0.995$ and $\arg G_f(i\omega_0) = -5.7°$. These modifications of the loop transfer functions give normally only minor changes in control loop performance. If needed, it is also possible to recalculate the parameters by determining the filter as described above, and then perform a regular design of the PI controller with the transfer function $G_f(s)G(s)$.

4.3. *The procedure*

To sum up we find that the design problem can be solved by the following procedure:

(1) Choose the design parameters M_s and/or M_p and compute C and R.

(2) Determine the search range ($\omega_l < \omega < \omega_h$). We have $\omega_l = \omega_{90}$ and $\omega_h = \omega_{180}$. For systems that satisfy the monotonicity condition (20) we have $\omega_h = \omega_{180\text{-}\arcsin R/C}$. Notice that there may be several search intervals. Narrow the search range by applying interval bisection to equation (18).

(3) *Normal case*: Use the initial values from Step 2 and solve equation (18) by Newton–Raphson. Evaluate the condition (19) and compute M_s and M_p. If both are satisfactory go to Step 5 otherwise compute new C and R and go to Step 2.

(4) *Corner case*: Compute the approximate envelope for values in the range $\omega_l < \omega < \omega_h$. Determine largest value of k_i, the frequencies of tangency ω_1 and ω_2 and compute k. These are used as initial values too solve equation (15) with Newton–Raphson. Compute M_s and M_p. If both are satisfactory go to Step 5 otherwise compute new C and R and go to Step 2.

(5) Determine the parameter b from equation (25).

(6) Evaluate the design including noise sensitivity. Modify the design parameters if required.

For simple systems it is sufficient to choose only the design parameter M_s. For special classes of systems all choices can be made automatically.

5. EXAMPLES

The design method has been tested on a large number of examples. In this section we will give a number of examples illustrating its properties.

5.1. *Typical process control problems*

To start with we will consider some representative systems which are normally encountered in process control. They have the following transfer functions,

$$G_1(s) = \frac{1}{(s+1)^3},$$

$$G_2(s) = \frac{1}{(s+1)(1+0.2s)(1+0.04s)(1+0.008s)},$$

$$G_3(s) = \frac{e^{-15s}}{(s+1)^3}, \qquad G_4(s) = \frac{1}{s(s+1)^2},$$

$$G_5(s) = \frac{1-2s}{(s+1)^3}, \qquad G_6(s) = \frac{9}{(s+1)(s^2+2s+9)}.$$

Systems G_1 and G_2 represent processes that are relatively easy to control. System G_3 has a long dead time, and G_4 models an integrating process. System G_5 has a zero in the right half plane, and system G_6 has complex poles with relative damping 0.33. Systems of type G_5 and G_6 are not common in process control, but they have been included to demonstrate the wide applicability of the design procedure. For all systems, except G_6, we can, by inspection, verify that they satisfy the monotonicity assumption. The system $G_6(s)$ will be discussed further on in Example 5 where it is shown that it also satisfies this assumption. We thus know that there exists a solution where the envelope has a continuous derivative.

Figure 3 shows the Nyquist curves of the loop transfer functions obtained for two values of the

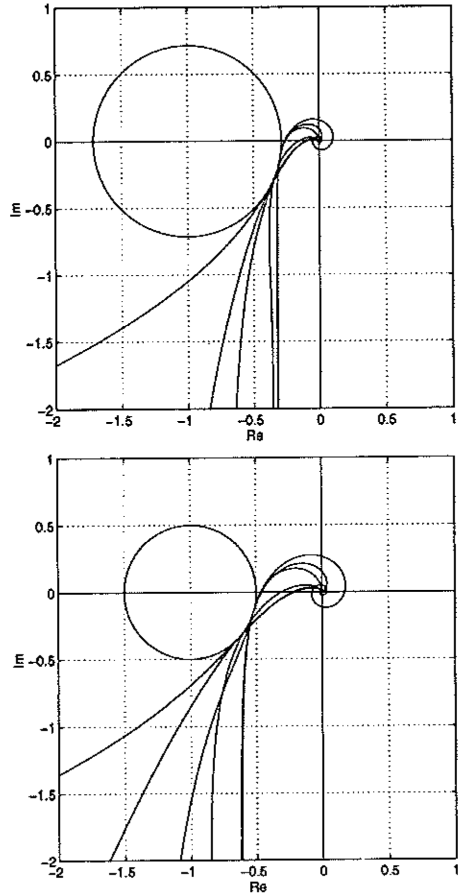


Fig. 3. The Nyquist plots of the open-loop frequency response for the systems $G_j(s)$, $j = 1, \ldots, 6$ for $M_s = 1.4$ in the upper figure and $M_s = 2.0$ in the lower one.

design parameter M_s. The responses to changes in set point and load are shown in Fig. 4, and the details of the design calculations and simulations are summarized in Table 1.

Even though the systems G_1–G_6 represent processes with large variations in process dynamics, Figure 4 shows that the resulting closed-loop responses become similar for each value of M_s. This is important because it means that the proposed design procedure gives closed-loop systems with desired and predictable properties. The fact that even integrating processes can be treated in the same way as stable processes is interesting. In many other design approaches, stable and integrating processes have to be treated separately, see Åström and Hägglund (1995b).

There is also a large similarity between the responses obtained with the different values of the tuning parameter M_s. This shows that the M_s-value is a suitable tuning parameter. Responses obtained with $M_s = 1.4$ show little or no overshoot. This is normally desirable in process control. Responses obtained with $M_s = 2.0$ give faster responses. The settling time at load disturbances, t_s, is significantly shorter with this larger value of M_s. On the other hand, these responses are oscillatory with a larger overshoot. This can be seen from the comparison between IE and the integrated absolute error IAE in Table 1.

The controller gain k varies significantly with the design parameter M_s. However, integral time T_i is fairly constant for the stable processes, i.e. all processes except G_4. This means that, for PI control, the different design specifications are mainly obtained by adjusting only the gain. This observation is made earlier, see Åström and Hägglund (1995b).

For the systems in Table 1 the values of M_p are smaller than the M_s values except for the system G_4 with $M_s = 1.4$ where M_s and M_p are equal. The constraint on the sensitivity function is thus the critical constraint for these systems.

Except for the integrating process G_4, the M_p values obtained for $M_s = 1.4$ are all close to one. Consequently, parameter b is also close to one. For $M_s = 2.0$, the M_p values are, however, larger. This means that the overshoots would have been significant if the set point weighting were chosen to $b = 1$. However, acceptable set point responses are obtained by using small values of b. In some cases, $b = 0$. It means that the procedure has failed to obtain $M_{sp} = 1$. If set point responses are important and if the overshoots are unacceptable, a redesign may be done using smaller values of M_s or optimization with constraint on M_p, see Section 4.

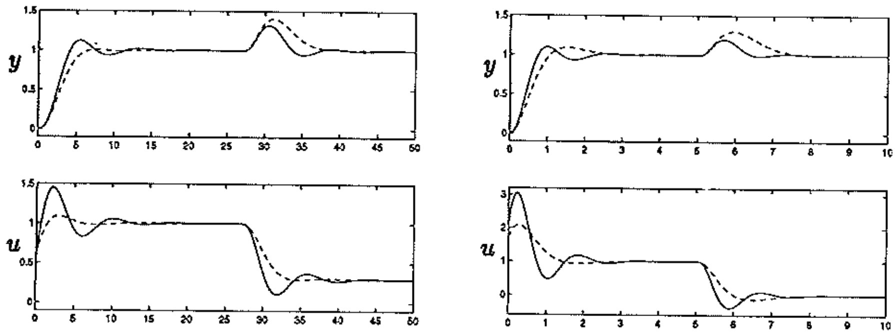
5.2. *More complex systems*

We will now discuss several examples with more complex dynamics.

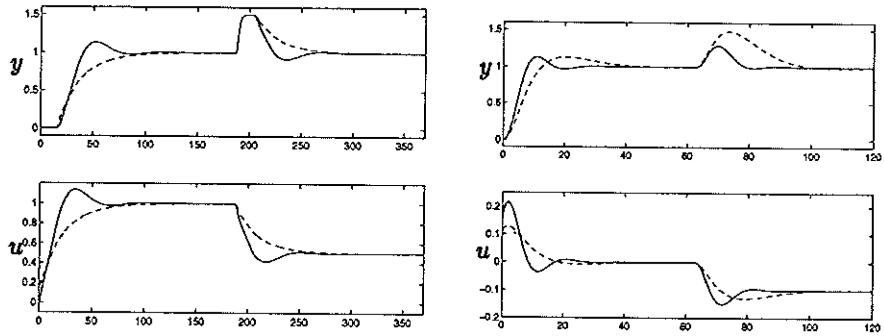
Example 1. *Pure time delay* (see Example 2). Many design methods for PI controllers perform poorly for systems with long relative time delays. The system G_3 is of this type because the ratio between the dead time and the dominating time constant is 15/3. A more extreme case is

$$G_7(s) = e^{-s}.$$

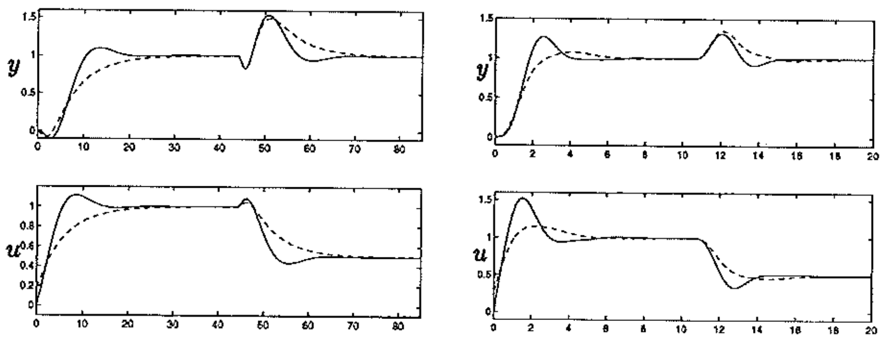
This system has $\omega_{90} = \pi/2$ and $\omega_{180} = \pi$. It satisfies the monotonicity condition (20). Making the design for $M_s = 1.4$ and 2.0, the design procedure gives the controller parameters which can be found in Table 2. The Nyquist diagrams and the set point and load disturbance responses are shown in Fig. 5. The figure shows that the design procedure manages to obtain controller parameters even for systems with extreme dead times.



Systems G_1 (left) and G_2 (right)



Systems G_3 (left) and G_4 (right)



Systems G_5 (left) and G_6 (right)

Fig. 4. Comparison between the PI controllers for $M_s = 1.4$ and 2.0. The graphs shows a step response followed by a load disturbance of the closed-loop system when designing for $M_s = 1.4$ (dashed line) and 2.0 (full line).

Example 2 (*Pure integrator with time delay*). A pure integrator with a time delay is another common model. The transfer function for such a system is

$$G_8(s) = \frac{e^{-s}}{s}.$$

This system has $\omega_{90} = 0$ and $\omega_{180} = \pi/2$. It fulfills the monotonicity condition (20). Making the design for $M_s = 1.4$ and 2.0, the design procedure gives the controller parameters which can be found in Table 2. The Nyquist diagrams and the set point and load disturbance responses are shown in Fig. 5, which demonstrates that the design procedure produces suitable controller parameters in this example too.

Table 1. Properties of controllers obtained for system G_1–G_7 for different values of the design parameter M_s

Process	M_s	k	T_i	b	IE	IE/IAE	ω_0	t_s	M_p
$G_1(s)$	1.4	0.633	1.95	1.00	3.07	1.00	0.74	10.3	1.00
	1.6	0.862	1.87	0.93	2.17	0.95	0.79	7.87	1.05
	1.8	1.06	1.82	0.70	1.72	0.86	0.82	6.77	1.24
	2.0	1.22	1.78	0.50	1.45	0.77	0.85	6.27	1.45
$G_2(s)$	1.4	1.93	0.745	0.89	0.387	1.00	3.33	2.25	1.10
	1.6	2.74	0.672	0.75	0.245	1.00	3.83	1.68	1.27
	1.8	3.47	0.625	0.63	0.180	0.96	4.25	1.39	1.46
	2.0	4.13	0.591	0.52	0.143	0.91	4.40	1.21	1.66
$G_3(s)$	1.4	0.164	6.16	1.00	37.5	1.00	0.096	91.3	1.00
	1.6	0.208	5.87	1.00	28.2	1.00	0.099	51.2	1.00
	1.8	0.241	5.66	0.92	23.5	0.88	0.101	39.8	1.02
	2.0	0.266	5.51	0.00	20.8	0.75	0.102	35.9	1.17
$G_4(s)$	1.4	0.167	14.0	0.70	84.0	0.97	0.29	35.3	1.40
	1.6	0.231	10.7	0.64	46.2	0.99	0.34	24.1	1.49
	1.8	0.286	9.00	0.57	31.5	0.99	0.38	18.5	1.62
	2.0	0.333	8.00	0.50	24.0	0.96	0.41	15.7	1.77
$G_5(s)$	1.4	0.179	1.78	1.00	9.90	0.90	0.38	28.6	1.00
	1.6	0.228	1.69	1.00	7.43	0.87	0.40	18.6	1.00
	1.8	0.265	1.64	0.87	6.18	0.80	0.41	14.7	1.04
	2.0	0.294	1.60	0.00	5.42	0.70	0.41	13.5	1.20
$G_6(s)$	1.4	0.313	0.373	0.88	1.19	0.87	1.98	4.13	1.04
	1.6	0.387	0.344	0.51	0.891	0.79	2.05	2.94	1.15
	1.8	0.441	0.325	0.00	0.739	0.70	2.05	2.69	1.26
	2.0	0.482	0.313	0.00	0.648	0.64	2.12	2.56	1.37

Table 2. Details of the design calculations of the systems G_7–G_{11} for different values of the design parameter M_s

Process	M_s	k	k_i	b	ω_0	M_p
$G_7(s)$	1.4	0.158	0.472	1.00	1.73	0.99
	2.0	0.255	0.854	0.00	1.83	1.17
$G_8(s)$	1.4	0.282	0.0418	0.66	0.54	1.45
	2.0	0.488	0.131	0.46	0.73	1.82
$G_9(s)$	1.4	2.94	11.5	0.81	7.89	1.17
	2.0	5.31	27.0	0.48	9.68	1.59
$G_{10}(s)$	1.4	1.25	1.62	0.78	3.49	1.23
	2.0	2.48	4.43	0.51	4.59	1.68
$G_{11}(s)$	1.4	1.30	2.03	0.86	3.75	1.13
	2.0	2.59	5.24	0.51	4.82	1.64

Example 3 (*A distributed parameter system*). The method also applies directly to systems described by partial differential equations. To illustrate this we consider a system described by a linear heat equation. Such a system has the transfer function

$$G_9(s) = e^{-\sqrt{s}}.$$

Hence

$$G_9(i\omega) = e^{-\sqrt{\omega/2}}e^{-i\sqrt{\omega/2}}.$$

The system satisfies the monotonicity condition (20) and we have $\omega_{90} = 4.93$, $\omega_{150} = 13.7$, and $\omega_{180} = 19.7$. Making the design for $M_s = 1.4$ and 2.0, the design procedure gives the controller parameters which can be found in Table 2. Notice that in this case there is a significant increase in performance when changing from $M_s = 1.4$ to $M_s = 2.0$.

Example 4 (*Fast and slow mode*). Consider a system with the transfer function

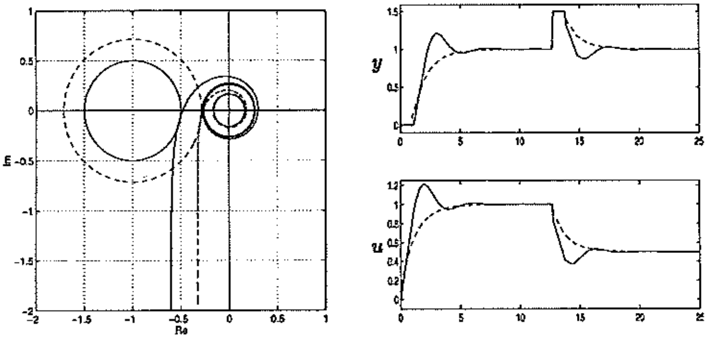
$$G_{10}(s) = \frac{100}{(s+10)^2}\left(\frac{1}{s+1} + \frac{0.5}{s+0.05}\right). \quad (27)$$

This system has two fast modes with time constants 0.1 s, one mode with a time constant of 1 s, and a slow mode with time constant 20 s. The static behavior is dominated by the slow mode which has a low-frequency gain of 10. The step response is dominated by the slow time constant, but it is the faster modes that are critical for the closed-loop system. The properties that are important for control are thus hidden in the step response. This means that most attempts to tune the system based on step response data will give poor results. Making the design for $M_s = 1.4$ and 2.0, the design procedure gives the controller parameters which can be found in Table 2. The properties of the closed loop systems obtained are illustrated in Fig. 5.

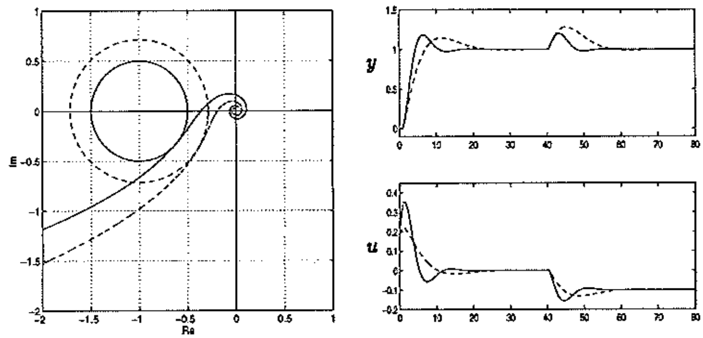
It is of interest to compare with the controller parameters obtained for the system

$$G_{11}(s) = \frac{150}{(s+10)^2(s+1)}, \quad (28)$$

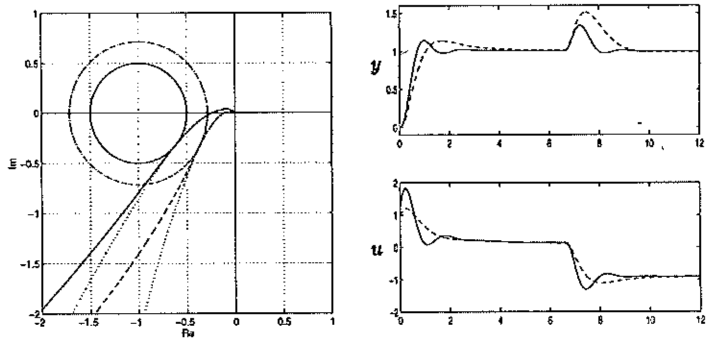
which is obtained by removing the slowest mode from the system (27). In this case when making the



System $G_7(s)$.



System $G_8(s)$.



Systems G_{10} and G_{11}.

Fig. 5. Nyquist curves of the loop transfer function and time responses of the closed-loop system when designing for $M_s = 1.4$ (dashed line) and 2.0 (full line) for systems $G_7(s)$, $G_8(s)$, $G_{10}(s)$ and $G_{11}(s)$.

design for $M_s = 1.4$ and 2.0, the design procedure gives the controller parameters which can be found in Table 2. These parameters are close to those obtained for the system (27), which shows that the design procedure manages to disregard the slow mode. The Nyquist curves for the two systems are compared in Fig. 5. Here it can be noticed that the two curves coincide for frequencies above ω_0.

Example 5 (*An oscillatory system*). Consider the system with the transfer function

$$G_{12}(s) = \frac{9}{(s+1)(s^2+as+9)}.$$

This is an interesting system from two points of view. First, the system has two oscillatory poles with relative damping $\zeta = a/6$. When parameter a is decreased it becomes more and more difficult to control. Second, depending on the value of parameter a the envelope may have a continuous derivative, $a \geq 1.0653$, or a corner, $a < 1.0653$.

For the case when the envelope has a continuous derivative a controller has already been designed for $a = 2$, see process G_6 at the beginning of this section. According to Fig. 4 the design gives a controller with good performance.

For the case when the envelope has a corner a controller was designed for $M_s = 2.0$. In Fig. 6 the Nyquist curves and the time responses are shown for the cases $a = 0.2, 0.5, 1.0$. The controller behaves reasonably well in spite of the poorly damped poles.

In Table 3 the controller parameters and the frequencies at which the loop transfer function is tangent to the M_s-circle are shown. Noticê in Table 3 how the proportional gain is negative for small values of a. This is the only way to increase the damping of the oscillatory poles with a PI controller.

Finally, we illustrate how our design method will provide a reasonable PI controller for the extreme case $a = 0$. With the design parameter $M_s = 1.4$ we obtain the following controller parameters: $k = -0.183, k_i = 0.251$ and $b = 0$. The time responses are shown in Fig. 7. We observe that the set point response is quite reasonable, even if there is a trace of poorly damped modes. The load disturbance will, however excite the oscillatory modes. The fact that the PI controller is unable to provide damping of these modes is clearly noticeable in the figure.

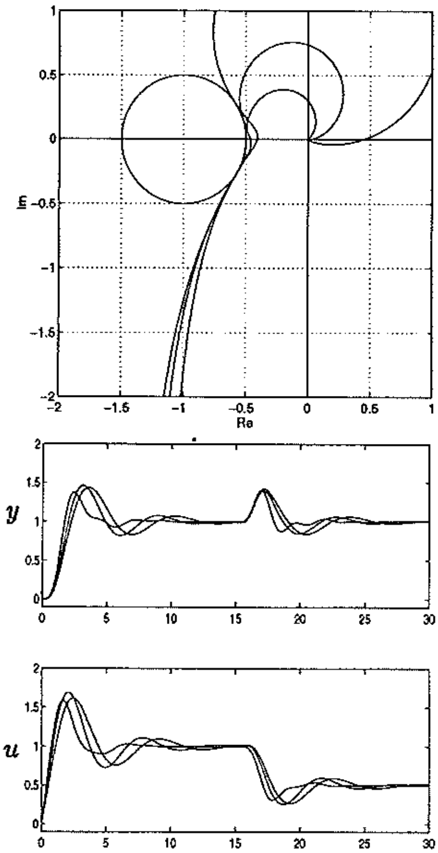


Fig. 6. Nyquist curves of the loop transfer function and time responses for Example 5 with $a = 0.2, 0.5, 1.0$, when designing for $M_s = 2.0$.

Example 6 (*A conditionally stable system*). Consider the system with the transfer function

$$G_{13}(s) = \frac{(s+6)^2}{s(s+1)^2(s+36)}.$$

This system does not satisfy the monotonicity assumption because the phase lag is not monotonic. The system is conditionally stable, since the Nyquist curve crosses the negative real axis at points $s = -0.0191$ and $s = -0.1656$. We have $\omega_{90} = 0$, and $\omega_{180} = 1.69$ and 4.17. With proportional feedback the system is stable for $k < 6.04$ or $k > 52.26$.

There are two solutions to the optimization problem for $M_s = 2.0$: $k = 0.47$, $k_i = 0.067$, $b = 0.52$, and $k = 921, k_i = 1098, b = 0.50$. The first solution gives $\omega_0 = 0.5196$ rad/s and the second gives $\omega_0 = 25.93$ rad/s. Nyquist curves and time responses to set point changes and load disturbances are shown in Fig. 8. Notice the similarities of the Nyquist curves and the differences in response speed for the solutions. The frequency scalings of the Nyquist curves are different.

Only one solution $k = 0.214$, $k_i = 0.0178$ and $b = 0.710$ is obtained for $M_s = 1.4$ with $w_0 = 0.3531$ rad/s. This illustrates that the method indicates that there is no controller that satisfies the specifications.

Table 3. Interesting parameters when designing a controller for $M_s = 2.0$ and different values of a in Example 5

a	k	k_i	ω_1	ω_2
0.0	− 0.29	0.68	0.97	2.75
0.1	− 0.25	0.82	1.08	2.71
0.2	− 0.20	0.93	1.16	2.67
0.5	− 0.09	1.17	1.37	2.55
1.0	0.09	1.38	1.65	2.30
2.0	0.48	1.54	2.79	2.79

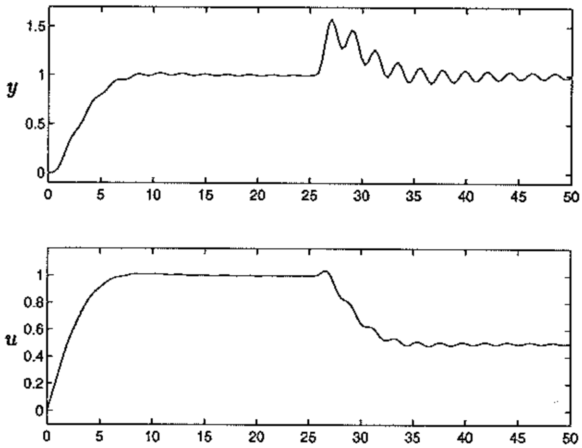


Fig. 7. Time response of the closed-loop system of Example 5 obtained for $a = 0$, when designing the PI controller for $M_s = 1.4$.

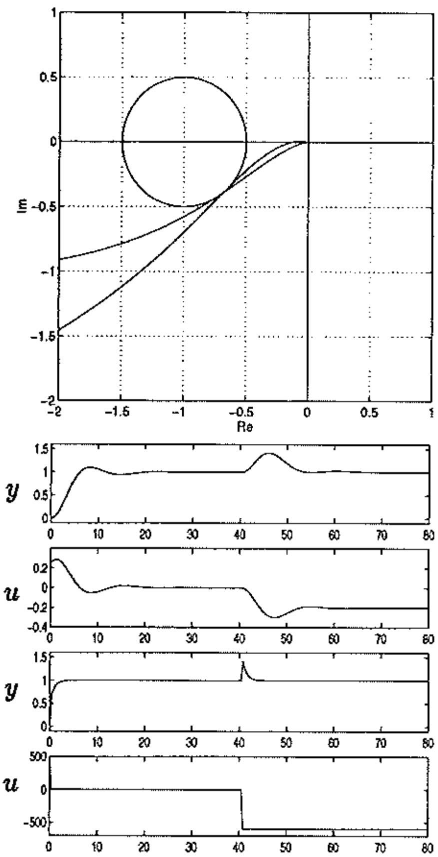


Fig. 8. Nyquist curves of the loop transfer function and time responses for the closed-loop system of Example 6 when designing for $M_s = 2.0$.

Example 7 (An unstable system). To show that PI control can also be used for unstable systems we consider a system with the transfer function

$$G_{14}(s) = \frac{a}{(s+a)(s-1)}.$$

This system does not satisfy the monotonicity assumption (20) because its phase does not decrease monotonically. The phase lag is π both at very low and at very high frequencies. The smallest phase lag is

$$\pi\text{-arctan}\,\frac{(a-1)\sqrt{a}}{2a}$$

which occurs for $\omega = \sqrt{a}$. To have a loop transfer function that lies outside a circle with center C and radius R it must be required that

$$\frac{(a-1)\sqrt{a}}{2a} \geq \frac{R}{\sqrt{C^2 - R^2}},$$

which implies that

$$a \geq \frac{(C+R)^2}{C^2 - R^2}.$$

For $C = 1$ and $R = 0.5$ we get $a \geq 3$. To solve the design problem we will therefore start with an interval of frequencies around $\sqrt{a}$. Figure 9 shows the Nyquist curves and the set point and load disturbance responses obtained for $a = 4$ and 8, respectively. The design is made for $M_s = 2.0$. For $a = 4$, the design procedure gives the controller parameters $k = 3.31$, $k_i = 0.82$, $b = 0.50$, and the frequency $w_0 = 3.04$ rad/s. For $a = 8$, we get $k = 8.70$, $k_i = 10.4$, $b = 0.5$, and the frequency $w_0 = 7.85$ rad/s. The M_p-values becomes quite large,

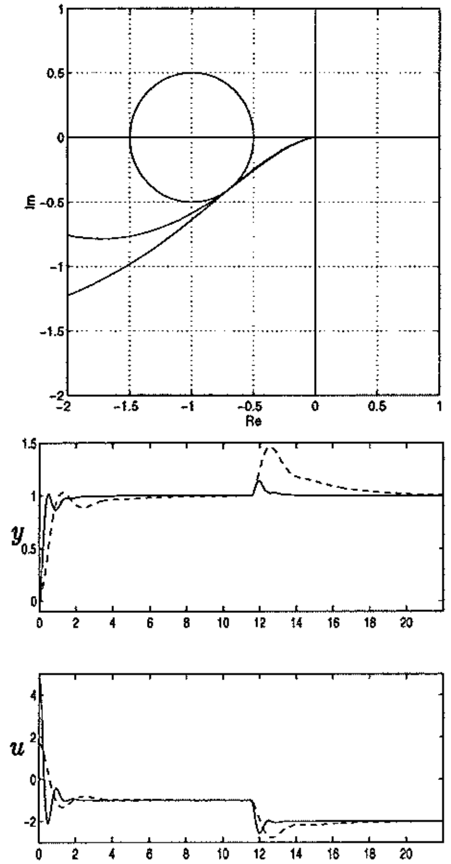


Fig. 9. Nyquist curves of the loop transfer function and time response of the closed-loop system of Example 7 with $a = 4$ and 8. The design is done for $M_s = 2.0$ in both cases.

$M_p = 1.98$ for $a = 4$ and $M_p = 1.87$ for $a = 8$. Still, we obtain set point responses with acceptable overshoots because of the set point weighting b.

Example 8 (*Filtering of measurement noise*). In most cases in PI control it is not a primary design consideration that measurement noise is fed into the control system through the controller. With reasonable sensors it is often sufficient to check the fluctuations of the control signal generated by the noise and introduce a filter on the measured signal if necessary. The design of such a filter was discussed in Section 2 and is illustrated here with the system in Example 1. Measurement noise is particularly difficult in this case because the process has constant gain for high frequencies.

We consider the case of a first-order noise filter with the transfer function

$$G_f(s) = \frac{1}{1 + sT_f}.$$

A PI controller is first designed without considering the noise filter. The design gives the frequency ω_0 from which the parameter T_f is determined as $T_f = 1/(m\omega_0)$. The design is then repeated for a system with the transfer function G_pG_f.

To evaluate the effectiveness of the noise filter we observe that the transfer function from measurement noise to control signal is given by

$$U(s) = \frac{G_cG_f}{1 + G_pG_c} N(s) \approx M_s G_c G_f N(s),$$

where N is the measurement noise. The approximation is obtained by replacing the sensitivity function with its maximum M_s. This will overestimate the control signal by at most a factor M_s. Furthermore, assume that the measurement noise is obtained by filtering white noise by a system with the transfer function

$$G_n(s) = \frac{s}{s + a}.$$

The control signal is then white noise filtered by

$$G_u(s) \approx M_s G_c G_f G_n = M_s \frac{ks + k_i}{(s + a)(1 + sT_f)}.$$

Assuming that the spectral density of the white noise is ϕ we find that the variance of the control signal is

$$\sigma_u^2 = M_s^2 \phi \frac{ak^2 + k_i^2 T_f}{2aT_f(1 + aT_f)}.$$

Table 4 shows the results of the PI controllers designed by the method mentioned above for different values of m. It has been computed for $M_s = 2.0$, $\phi = 1$, $a = 2$ and $\zeta = 0.707$. In the table we have also shown the ratio between integration time T_i and the filter time constant T_f. The tradeoffs are clear from the table. Decreasing the value of m makes the system less sensitivity to measurement noise but more sensitive to load disturbances. It is possible to introduce loss functions which capture this compromise but it is rare that the data required is available with sufficient precision to justify the analysis.

Table 4. Parameters of PI controllers with filtering of the process output discussed in Example 8

m	k	k_i	ω_0	IE	σ	T_i/T_f
2	0.31	0.73	1.48	1.36	0.89	1.25
5	0.27	0.78	1.66	1.26	1.25	2.82
10	0.26	0.81	1.74	1.21	1.67	5.51
20	0.26	0.83	1.78	1.17	2.28	11.0
∞	0.26	0.85	1.83	1.12	∞	—

6. COMPARISONS WITH OTHER METHODS

Many different methods have been proposed for tuning PI controllers. A comprehensive presentation which includes many comparisons are given in Åström and Hägglund (1995b). In this section, the new design procedure is compared with two other common methods, namely the Ziegler–Nichols frequency response method, see Ziegler and Nichols (1942), and the Lambda tuning procedure, see Rivera *et al.* (1986).

All methods differ with respect to the process knowledge required and the design specifications. In the Ziegler–Nichols method, the process model is specified in terms of the ultimate gain and the ultimate frequency. No design specification can be made. In Lambda tuning, the process model is specified in terms of the static gain, time constant, and time delay. The design variable is the desired closed-loop time constant. A common rule-of-thumb that is used in this section is that the desired closed-loop time constant should be three times the open-loop time constant, see EnTech (1993). The new design method requires that the transfer function of the process is specified, and the design is given in terms of the M_s-value. The value $M_s = 1.6$ is used in this comparison.

The design procedures have been applied to the processes G_1, G_2, and G_3 given in Section 5. The controller parameters and some performance measures are presented in Table 5.

For process G_1, the Ziegler–Nichols method gives a very oscillatory response with an overshoot of 36% and a high M_s value. This is due to the high controller gain. The Lambda tuning method gives a well damped but sluggish response. The settling time and the IAE values are three times larger than those obtained in the new design. This is due to the low controller gain.

For process G_2, the Ziegler–Nichols method gives a control loop that is close to the stability boundary. The overshoot is 47% and the M_s-value is $M_s = 11$. Again this is mainly due to the high controller gain. The Lambda tuning method gives a very sluggish response, were the IAE value and the settling time are almost a magnitude larger than in the new design method.

Process G_3 has a long dead time. For this process, the Ziegler–Nichols design results in a controller with too high gain and too long integral time. This gives a control loop with a large M_s value, and a very long settling time. Lambda tuning gives a low controller gain and an integral time that is too short. This results in oscillatory behavior with an overshoot of 21%. The rule of thumb for choosing the desired closed-loop time constant is not suitable for processes with a long dead time. Therefore, it is sometimes suggested to relate the desired closed-loop time constant to the dead time L instead of the open-loop time constant T when the dead time is long.

The Ziegler–Nichols method gives systems which are inherently poorly damped. The examples demonstrate that it is not trivial to choose the desired closed-loop time constant in Lambda tuning. The new design procedure manages to obtain consistent behavior for a wide range of processes. On the other hand, the procedure assumes that the complete transfer function of the process is known, whereas the other two methods only uses simpler models.

Many traditional methods fail to give acceptable control for several of the more difficult control problems discussed in the paper. For example, both the Ziegler–Nichols method and the Lambda tuning method gives controller gain $k = 0$ for the pure delay process $G(s) = e^{-s}$.

7. CONCLUSIONS

This paper describes a design method for PI controllers. The method assumes that the transfer function of the process is given. The specifications capture demands on load disturbance rejection, set point response, measurement noise and model uncertainty. The primary design goal is to obtain good load disturbance responses. This is done by

Table 5. Comparison between the three design procedures. The table shows controller parameters k, T_i, and k_i, and performance measures M_s, M_p, IAE, overshoot o, and settling time t_s

Process	Design	k	T_i	k_i	M_s	IAE	o %	t_s	M_p
	Ziegler–Nichols	3.60	3.02	1.19	4.93	1.40	35.9	18.0	4.39
G_1	Lambda tuning	0.278	1.92	0.145	1.17	6.90	0	25.8	1.00
	New design	0.862	1.87	0.461	1.60	2.28	8.80	8.12	1.05
	Ziegler–Nichols	13.6	0.468	29.1	11.4	0.098	47.1	2.28	11.1
G_2	Lambda tuning	0.312	1.05	0.299	1.06	3.35	0.00	14.0	1.00
	New design	2.74	0.672	4.08	1.60	0.246	9.95	1.83	1.27
	Ziegler–Nichols	0.471	30.0	0.0157	1.86	63.2	0.00	237	1.00
G_3	Lambda tuning	0.081	1.73	0.0465	2.15	37.3	21.5	118	1.39
	New design	0.208	5.87	0.0355	1.60	28.2	0.00	51.2	1.00

minimizing the integrated control error *IE*. Robustness is guaranteed by requiring that the maximum sensitivity is less than a specified value M_s. Good set point response is obtained by using a structure with two degrees of freedom. This requires an extra parameter, the set point weighting *b*, in the algorithm. The primary design parameter is the maximum sensitivity, M_s, but auxiliary design parameters such as the maximum of the complementary sensitivity, M_p, can be added.

The design problem can be formulated as a constrained optimization problem. It is shown that this problem can be reduced to a solution of nonlinear algebraic equations. Efficient ways of solving these equations are presented. There are unique solutions for special classes of systems but very complicated situations may occur for complicated systems. The method will give a solution if one exist and it will indicate when there is no PI controller that satisfies the specifications.

The design procedure has been applied to a variety of systems; stable and integrating, with short and long dead times, with real and complex poles, and with positive and negative zeros.

REFERENCES

Åström, K. J. and Hägglund (1995a). New tuning methods for PID controllers. *European Control Conference*, Rome, Italy, 2456–2462.

Åström, K. J. and T. Hägglund (1995b). *PID Controllers: Theory, Design, and Tuning*, 2nd edn. Instrument Society of America, 1995, Research Triangle Park, NC.

EnTech (1993). Automatic Controller Dynamic Specification. Number Version 1.0, 11/93. EnTech Control Engineering Inc.

Hazebroek, P. and B. L. van der Waerden (1950). Theoretical considerations on the optimum adjustment of regulators. *Trans. ASME*, **72**, 309–322.

Horowitz, I. M. (1963). *Synthesis of Feedback Systems*. Academic Press, New York.

Khalil, H. K. (1992). *Nonlinear Systems*. MacMillan, New York.

Persson, P. (1992). Towards autonomous PID control. PhD Thesis, ISRN LUTFD2/TFRT-1037-SE, Department of Automatic Control, Lund Institute of Technology, Lund, Sweden.

Persson, P. and K. J. Åström (1992). Dominant pole design—a unified view of PID controller tuning. Preprints 4th IFAC Symposium on Adaptive Systems in Control and Signal Processing, Grenoble, France, pp. 127–132.

Rivera, D. E., M. Morari and S. Skogestad (1986). Internal model control—4. PID controller design. *Ind. Engng. Chem. Process. Des. Dev.*, **25**, 252–265.

Schei, T. S. (1994). Automatic tuning of PID controllers based on transfer function estimation. *Automatica*, **30**(12), 1983–1989.

Shigemasa, T., Y. Iino and M. Kanda (1987). Two degrees of freedom PID auto-tuning controller. *Proc. ISA Annual Conf.*, 703–711.

Shinskey, F. G. (1990). How good are our controllers in absolute performance and robustness? *Measurement and Control*, **23**, 114–121.

Yamamoto, S. and I. Hashimoto (1991). Present status and future needs: the view from Japanese industry. In Arkun and Ray, Eds., *Chemical Process Control—CPCIV. Proc. 4th Inter. Conf. on Chemical Process Control*, TX.

Ziegler, J. G. and N. B. Nichols (1942). Optimum settings for automatic controllers. *Trans. ASME*, **64**, 759–768.

Errata for "Design of PI Controllers based on Non-Convex Optimization"

p. 587 Equation (7): The expressions C and R should be,

$$C = \frac{M_s(2M_p - 1) - (M_p - 1)}{2M_s(M_p - 1)},$$
$$R = \frac{M_s + M_p - 1}{2M_s(M_p - 1)}.$$

p. 589 Equation (10): Equation (10) should be labelled out.

p. 597 Figure 6: The time response in Figure 6 is wrong. The correct one is as follows,

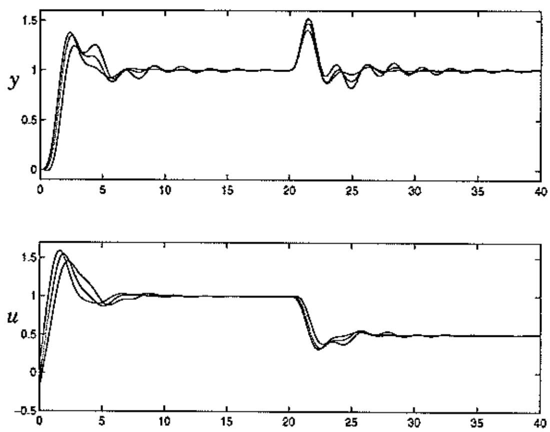


Figure 1 The time responses for Example 5 with $a = 0.2,\ 0.5,\ 1.0$, when designing for $M_s = 2.0$.

II

Proceedings of the American Control Conference
San Diego, California • June 1999

DESIGN OF PID CONTROLLERS BASED ON CONSTRAINED OPTIMIZATION

H. Panagopoulos, K. J. Åström, and T. Hägglund

Department of Automatic Control
Lund Institute of Technology
Box 118
S-221 00 Lund, Sweden
Phone: +46-46 2228789, Fax: +46-46 138118
Email: hp@control.lth.se

Abstract This paper presents a new design method for PID controllers based on optimization of load disturbance rejection with constraints on robustness to model uncertainties. The design also delivers parameters to deal with measurement noise and set point response. Thus, the formulation of the design problem captures four essential aspects of industrial control problems, leading to a constrained optimization problem which can be solved iteratively.

Keywords PID control. Design. Optimization. Specifications. Load disturbance rejection. Measurement noise filtering. Set point response. Robustness. Sensitivity.

1. Introduction

The PID controller is todays most commonly used control algorithm, see Yamamoto and Hashimoto (1991). At the moment, there exits many different methods to find suitable controller parameters. The methods differ in complexity, flexibility, and in the amount of process knowledge used. Depending on the application, there is a need to have several types of tuning methods. There are simple, easy to use methods which require little information, e.g. the method described in Ziegler and Nichols (1942), and on the other hand more sophisticated methods which requires more information and more computations. But with todays computational capacity this is no disadvantage.

There are several reasons to look for better methods to design PID controllers. One is the significant impact it may give because of the widespread use of the controllers. Another is the benefit emerging autotuners and tuning devices can derive from improved design methods.

This paper describes a new design method for PID controllers based on the assumption that the process transfer function is known. The primary design goal is to obtain good load disturbance responses. This is done by minimizing the integrated control error IE. Robustness is guaranteed by requiring that the maximum sensitivity be less than a specified value M_s. Measurement noise is dealt with by filtering. Good set point response is obtained by using a structure with two degrees of freedom.

The specifications are expressed in terms of a number of parameters for which good default values can be found. In the simplest case good default values can be given to all parameters. The user simply supplies the process transfer function and the design parameter, which is the maximum sensitivity, M_s. Consequently, the method provides all the parameters of the PID controller: controller gain k, integral time T_i, derivative time T_d and set point weight b. In addition the filters of the measured signal and the set point are delivered.

2. Problem Formulation

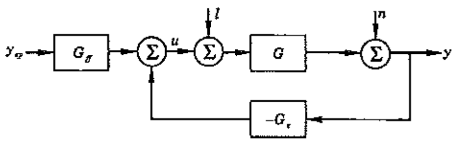


Figure 1 Block diagram describing the design problem.

The design problem is illustrated in Figure 1. A process with transfer function $G(s)$ is controlled with a PID controller with two degrees of freedom. Transfer function $G_c(s)$ describes the feedback from

0-7803-4990-6/99 $10.00 © 1999 AACC

process output y to control signal u, and $G_{ff}(s)$ describes the feed forward from set point y_{sp} to u. Three external signals act on the control loop, namely set point y_{sp}, load disturbance l and measurement noise n.

The design objective is to determine the controller parameters in $G_c(s)$ and $G_{ff}(s)$ so that the system behaves well with respect to changes in the three signals y_{sp}, l and n as well as in the process model $G(s)$. Hence, the specification will express requirements on

- Load disturbance response
- Measurement noise response
- Set point response
- Robustness with respect to model uncertainties

Process and controller structures

The design problem was formulated so as to apply to a wide variety of systems. Consequently, the process is assumed to be linear, time invariant, and specified by a transfer function $G(s)$, which is analytic with finite poles and, possibly, an essential singularity at infinity. The description covers finite dimensional systems with time delays and infinite dimensional systems described by linear partial differential equations.

Initially the controller is described by

$$u(t) = k\big(by_{sp}(t) - y(t)\big) + k_i \int_0^t (y_{sp}(\tau) - y(\tau))d\tau + k_d\left(-\frac{dy(t)}{dt}\right),$$

where k, k_i, k_d and b are controller parameters.

It proved beneficial to replace the signals y and y_{sp} with their filtered values y^f and y^f_{sp}. The filtered signals are generated by

$$Y^f(s) = F_y(s)Y(s),$$
$$Y^f_{sp}(s) = F_{sp}(s)Y_{sp}(s),$$

where the filters are low pass filters of first or second order. The controller can thus be characterized by either four parameters k, k_i, k_d, b, and two filters F_y, F_{sp}.

Load Disturbance Attenuation

The primary design goal is to achieve good rejection of load disturbances where no detailed assumptions are made about the load disturbances except that they are low frequent. By maximizing the integral gain k_i the effect of the load disturbance l on the output y is minimized. According to Åström *et al.* (1998) it is shown that this is equivalent to minimizing the integrated error (IE) for a step change in the load disturbance.

Measurement Noise

In traditional PID controllers the filters are often only applied on the derivative term. A common choice is to give the derivative term

$$D(s) = -\frac{kT_d s}{1 + s\frac{T_d}{N}}Y(s) = -\frac{k_d s}{1 + s\frac{k_d}{kN}}Y(s),$$

where N is a number in the range 2-10. This will reduce the high frequency gain to $k(1+N)$. Since the filter is only applied on the derivative term and not on the proportional term, the high frequency gain can never be made smaller than k.

In this study all terms in the controller were filtered. For first order filter with filter time constant T_f, the high frequency gain becomes k_d/T_f. For second order filters the high frequency gain goes to zero as the frequency goes to infinity. A nice feature of the design is to provide a systematic way to determine T_f.

Set Point Response

The transfer function relating set point to process output is given by

$$G_{sp}(s) = \frac{GG_{ff}}{1 + GG_c} = \frac{k_i + bks}{k_i + ks + k_d s^2} \frac{GG_c}{1 + GG_c} \frac{F_{sp}}{F_y}.$$

When the controller has been designed to give good attenuation of disturbances, the parameter b and the filter F_{sp} can be chosen to give an appropriate set point response. The maximum of G_{sp} is chosen

$$M_{sp} = \max_\omega |G_{sp}(i\omega)|.$$

Robustness

Sensitivity to modeling errors can be expressed as the largest value of the sensitivity function, i.e.

$$M_s = \max_\omega \left| \frac{1}{1 + GG_c(i\omega)} \right|.$$

The quantity M_s is simply the inverse of the shortest distance from the Nyquist curve of the loop transfer function to the critical point -1. Typical values of M_s are in the range 1 to 2.

The sensitivity can also be expressed as the largest value of the complementary sensitivity function, i.e.

$$M_p = \max_{\omega} \left| \frac{GG_c(i\omega)}{1 + GG_c(i\omega)} \right|. \tag{1}$$

Typical values of M_p are in the range 1.0 to 1.5.

Another possibility is to use the H_∞-norm

$$\gamma = \max_{\omega} \left| \frac{1 + |GG_c(i\omega)|}{1 + GG_c(i\omega)} \right|, \tag{2}$$

which is discussed in detail in Panagopoulos and Åström (1998).

Tuning Parameter

The tradeoff between performance and robustness varies among different control problems. Therefore, it is desirable to have a design parameter to change the properties of the closed loop system.

For the proposed design method the robustness constraint is a good measure of the performance of the system. It has been shown in Åström *et al.* (1998) that the variable M_s fulfills all the requirements of a good design parameter.

3. PID Design

The method for designing PID controllers will now be presented. It is given by,

$$\begin{array}{ll} \max & k_i \\ \text{such that} \quad f \geq R^2, & \kappa < 0, \quad \delta < 0, \end{array} \tag{3}$$

where κ is the curvature of the loop transfer frequency function, $L(i\omega)$, and δ is the difference in phase change of $L(i\omega)$ at two consecutive frequency points. Consequently, the first constraint in Equation (3) express the sensitivity condition, the second constraint specifies that a negative curvature of $L(i\omega)$ should be obtained and the third constraint prevents $L(i\omega)$ to have undesirable phase leads.

For systems with integral action and close to integral action, i.e. a pole relatively close to zero, the second constraint in Equation (3) is, however, too severe. Omitting it will in these cases be appropriate. Consequently, the design of PID controllers is separated into two cases depending on the considered system. This separation between integrating and non integrating processes is done in several previous design methods, see Åström and Hägglund (1995).

Measurement noise filtering

When there are substantial measurement noise, it is customary to filter the measurement signal. A nice feature of the new design procedure is to provide a systematic way to determine T_f. The choice

$$T_f = \begin{cases} \dfrac{1}{N\omega_0} & \text{for first order filter,} \\ \dfrac{1}{2N\omega_0} & \text{for second order filter,} \end{cases}$$

makes it possible to determine the effects of the filter at the frequency ω_0, where ω_0 is the frequency at which the sensitivity function is maximal. Reasonable values of N are in the interval 2–10. Because of the special choice of T_f the first and second order filter has the same amount of modification on the loop transfer function at ω_0.

Inserting a filter modifies the loop transfer function which gives minor changes in control loop performance. Consequently, adjusted controller parameters are obtained simply by repeating the design with the process G replaced by the transfer function $F_y G$.

Set point weight

In order to have a small overshoot in set point response, set point weight b and filter F_{sp} will be determined such that the resonance peak of the transfer function $G_{sp}(s)$, i.e.

$$M_{sp} = \max \left| G_{sp}(i\omega) \right|, \tag{4}$$

is close to one.

First the set point weight, b, is determined without filtering, i.e. $F_{sp} = 1$, according to Panagopoulos *et al.* (1998). Only positive values of b are allowed, since negative values of b may result in inverse step responses in the control signal. If $b = 0$, it is not sure that the design objective of Equation (4) will be obtained. Then the filter F_{sp} is determined according to Panagopoulos *et al.* (1998).

4. Examples

The design method has been tested on a number of examples which illustrate its properties. The follow-

Table 1 Properties of controllers obtained for system G_1–G_8 for different values of the design parameter M_s.

Process	M_s	k	T_i	T_d	b	T_{sp}	IE	$\frac{IE}{IAE}$	w_0	t_s	M_p
$G_1(s)$	1.4	0.291	8.03	3.15	0.00	3.68	27.6	0.71	0.90	23.5	1.37
	2.0	0.645	5.20	2.66	0.72	0.00	8.05	0.72	0.99	14.0	1.40
$G_2(s)$	1.4	0.180	2.33	2.04	0.94	0.00	12.9	0.92	0.18	27.2	1.00
	2.0	0.618	3.39	1.95	0.00	3.98	5.47	0.62	0.35	5.47	1.29
$G_3(s)$	1.4	15.92	0.322	0.15	0.39	0.00	0.020	0.81	19.4	0.72	1.26
	2.0	42.81	0.254	0.13	0.71	0.00	0.0059	0.86	26.1	-	1.64
$G_4(s)$	1.4	0.758	2.07	0.82	0.00	0.86	2.74	0.83	0.55	10.8	1.05
	2.0	1.74	1.69	0.96	0.00	1.91	0.97	0.48	0.73	15.4	1.75
$G_5(s)$	1.4	0.870	2.41	1.50	0.00	2.27	2.77	0.73	0.66	12.1	1.12
	2.0	1.57	2.13	1.42	0.00	2.85	1.37	0.60	0.68	9.05	1.63
$G_6(s)$	1.4	0.703	2.93	1.77	0.00	2.39	4.17	0.79	0.55	16.1	1.06
	2.0	1.23	2.72	1.62	0.00	2.78	2.20	0.65	0.53	11.8	1.49
$G_7(s)$	1.4	0.553	3.90	1.50	0.62	0.00	7.05	0.97	0.37	22.9	1.00
	2.0	1.067	3.26	1.85	0.00	2.85	3.06	0.60	0.45	14.1	1.40
$G_8(s)$	1.4	0.330	2.24	0.90	0.64	0.00	6.78	0.85	0.89	19.3	1.00
	2.0	0.569	2.08	0.86	0.00	2.01	3.66	0.61	0.64	17.3	1.33

ing transfer functions have been considered,

$$G_1(s) = \frac{1}{s(s+1)^3}, \qquad G_2(s) = \frac{e^{-5s}}{(s+1)^3},$$
$$G_3(s) = \frac{1}{(s+1)(1+0.2s)(1+0.04s)(1+0.008s)},$$
$$G_n(s) = \frac{1}{(s+1)^n},\ n = 4 \text{ to } 7, \qquad G_8(s) = \frac{1-2s}{(s+1)^3}.$$

The first seven models capture typical dynamics encountered in the process industry. Model G_1 is an integrating process and G_2 models a process with long dead time. Model G_8 has a zero in the right half plane, which is uncommon in process control, but it has been included to demonstrate the wide applicability of the design procedure.

Figures 2 and 3 show the responses to changes in set point and load. The details of the design calculations and simulations are summarized in Table 1. Note that the PID controller obtained is compared to the corresponding PI controller to show the amelioration of the PID design.

Although models $G_1 - G_8$ represent processes with large variations in process dynamics, Figure 2 and 3 show that the resulting closed loop responses for a load disturbance become similar for each value of M_s. This is important because it means that the proposed design procedure gives closed loop systems with desired and predictable properties.

There is also a clear similarity between the responses obtained with the different values of the tuning parameter M_s, thus indicating the suitability of the M_s-value as a tuning parameter. Responses obtained with M_s=1.4 show little or no overshoot, as is normally desirable in process control. Faster responses are obtained with M_s=2.0. The settling time at load disturbances, t_s, is significantly shorter with the larger value of M_s. On the other hand, these responses are oscillatory with larger overshoots. This can be seen from the comparison between IE and the integrated absolute error IAE in Table 1. Notice the agreement with the conclusions made for design of PI controllers in Åström *et al.* (1998).

The controller gain k varies significantly with the design parameter M_s: it is larger for designs when $M_s = 2.0$ than for those when $M_s = 1.4$. However, integral time T_i is fairly constant for the stable processes, i.e., all processes except G_1. The derivative time T_d is usually larger for designs with $M_s = 1.4$ than for those with $M_s = 2.0$. In all cases the PID design generates a controller with complex zeros for $M_s = 2.0$. Thus, the controller will not be realizable in serial form.

For M_s=2.0, the M_p-values are large. Consequently, the overshoots would be significant if the set point weight is $b = 1$. However, acceptable set point responses are obtained by suitable choices of either the set point weight b or the filter F_{sp}. According to Table 1, it is not always enough to set $b = 0$ to obtain a small overshoot filtering may also be needed.

5. Conclusions

PID controllers were designed to capture demands on load disturbance rejection, set point response, measurement noise and model uncertainty. Good load disturbance responses were obtained minimizing the integrated control error IE. Robustness is guaranteed by requiring a maximum sensitivity of less than a specified value M_s. Measurement noise is dealt with by filtering. Good set point response is obtained by using a structure with two degrees of freedom.

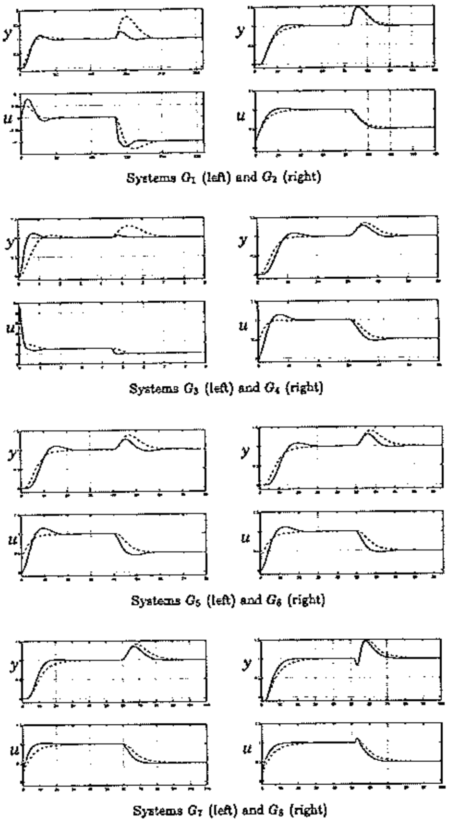


Figure 2 Comparison between the PID (full line) and the PI controller (dashed line) for $M_s = 1.4$. The graphs shows a step response followed by a load disturbance of the closed loop system.

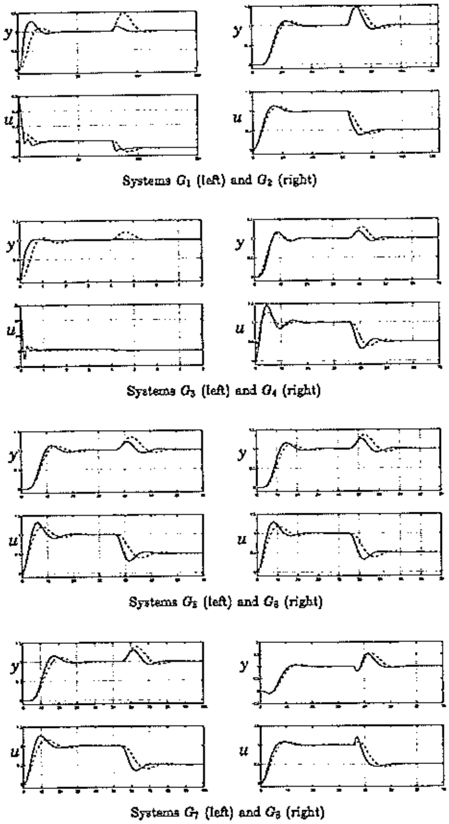


Figure 3 Comparison between the PID (full line) and the PI controller (dashed line) for $M_s = 2.0$. The graphs shows a step response followed by a load disturbance of the closed loop system.

The design procedure has been applied to a variety of systems: stable and integrating, with long dead times and with right half plane zeros.

6. References

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

ÅSTRÖM, K. J., H. PANAGOPOULOS, and T. HÄGGLUND (1998): "Design of PI controllers based on non-convex optimization." *Automatica*, 35:5.

PANAGOPOULOS, H. and K. J. ÅSTRÖM (1998): "PID control design and H_∞ loop shaping." Department of Automatic Control, Lund Institute of Technology, Box 118, S-221 00 Lund, Sweden.

PANAGOPOULOS, H., K. J. ÅSTRÖM, and T. HÄGGLUND (1998): "Design of PID controllers based on constrained optimization." Department of Automatic Control, Lund Institute of Technology, Box 118, S-221 00 Lund, Sweden.

YAMAMOTO, S. and I. HASHIMOTO (1991): "Present status and future needs: The view from Japanese industry." In ARKUN AND RAY, Eds., *Chemical Process Control—CPCIV*. Proceedings of the Fourth Internation Conference on Chemical Process Control, Texas.

ZIEGLER, J. G. and N. B. NICHOLS (1942): "Optimum settings for automatic controllers." *Trans. ASME*, 64, pp. 759–768.

IIs

Supplement and Errata to "Design of PID Controllers Based on Constrained Optimization"

The conference paper was brief due to page restrictions. In this supplement we provide additional material to the paper.

The first section of the supplement shows that a direct generalization of the problem formulation used to design PI controllers in Åström *et al.* (1998) does not work well for PID controllers. These insights motivate a reformulation of the problem for PID controller design. Section 2 and 3 show how filters can be incorporated to deal with measurement noise and set points. Finally, Section 4 presents a numerical solver, which is used to solve the parameter optimization problem. Some implementation aspects are also discussed.

1. Difficulties of the PID Design

The solution of the PID design can be formulated as the parameter optimization problem in Åström *et al.* (1998). Find controller parameters maximizing k_i subject to the constraints: 1) the closed loop system is stable, 2) the Nyquist curve of the loop transfer function is outside a circle with center at $s = -C$ and radius R, i.e.,

$$\begin{array}{ll} \max & k_i \\ \text{such that} & f(k, k_i, k_d, \omega) \geq R^2 \quad \forall\ \omega > 0, \end{array} \tag{1}$$

where f is the function

$$\left| C + \left(k - \frac{i}{\omega}(k_i - \omega^2 k_d)\right) G(i\omega) \right|^2,$$

where the constraint in (1) will be denoted the sensitivity constraint. It was shown in Åström *et al.* (1998) that the problem formulation in (1) worked well for the design of PI controllers. Unfortunately, it is not suitable for the design of PID controllers which will be explained below.

The constraint in (1) conceptually defines k_i as a function of k and k_d. Thus, the design problem is to maximize the function $k_i(k, k_d)$ which is illustrated in Figure 1. The upper and lower surface in the figure corresponds to the maximum and the minimum of the function $k_i(k, k_d)$ which fulfills the sensitivity constraint. Consider for example the contours of constant k_d. Small values of k_d give contours with a smooth optima of the upper surface, whereas the values of k_i are 0 for the lower surface. Larger values of k_d give contours where the optima appears at an edge, or in other words, where the upper and lower surface will coincide. The latter case makes the optimization difficult. Even more seriously is the sensitivity of the optimal solution to changes in the controller parameters, that is, small changes in k and k_d may give large changes in k_i. It is highly desirable that small changes in the controller parameters should not affect the performance too much. Next, these difficulties will be treated more closely. These insights will then be used to reformulate the optimization problem in (1).

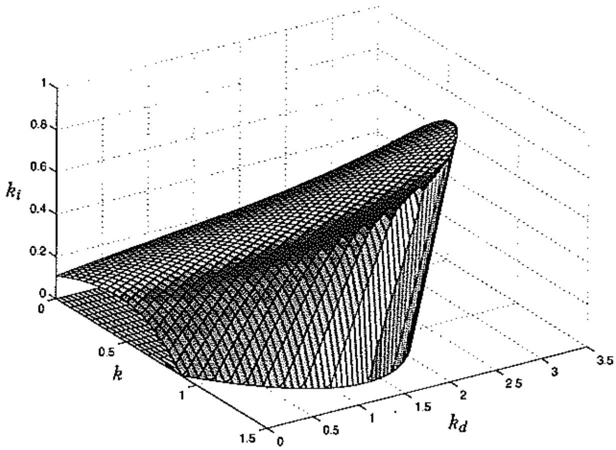


Figure 1 A geometric illustration of the sensitivity constraint in (1).

Geometric Interpretation of the Sensitivity Constraint

The sensitivity constraint in (1) has a nice geometric interpretation. This is exploited to gain insight how to reformulate the optimization problem in Equation (1). Introduce

$$G(i\omega) = r(\omega)e^{i\varphi(\omega)} = \alpha(\omega) + i\beta(\omega),$$

then, the sensitivity constraint in (1) can be written as

$$C^2 + 2C\alpha(\omega)k + 2C\frac{\beta(\omega)}{\omega}(k_i - \omega^2 k_d) + r^2(\omega)k^2 + \frac{r^2(\omega)}{\omega^2}(k_i - \omega^2 k_d)^2 \geq R^2. \tag{2}$$

In the following example, the argument ω in α, β, r, and φ will be omitted in order to simplify the writing. Rewriting Equation (2) as

$$\frac{r^2}{R^2}\Big(k + \frac{\alpha C}{r^2}\Big)^2 + \frac{r^2}{\omega^2 R^2}\Big(k_i - \omega^2 k_d + \frac{\omega\beta C}{r^2}\Big)^2 \geq 1, \tag{3}$$

we find that it represents the exterior of a cylinder in R^3 for fixed ω. In the following example the geometric characteristics of the sensitivity constraint are demonstrated. First it is shown how the cylinder's intersection with the k-k_i-plane is an ellipse, whose axes is parallel to the ones of the kk-k_i-plane. Secondly, it is demonstrated how to generate the boundary of parameter sets, k and k_i, satisfying the sensitivity constraint. Finally, it is shown how the boundary of these sets may have non smooth derivatives, even for a simple process. Recall from Åström *et al.* (1998) that similar situations could occur for the design of PI controllers. However, these cases were not very common.

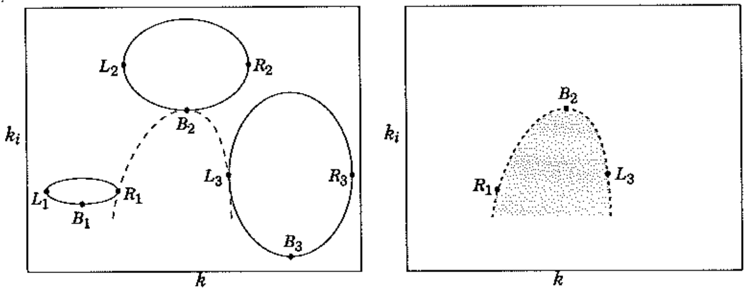


Figure 2 How to generate the set of (k, k_i)-parameters which satisfy the sensitivity constraint in (3). In the left figure the loci of the ellipses vertices (L_j, R_j, B_j) have been plotted for different (ω_j, k_d°) which forms an envelope (dashed line). In the right figure the interior of the envelope (shaded area) gives the set of (k, k_i)-parameters which satisfy the sensitivity constraint (3).

An Example: The Sensitivity Constraint

The following example shows how the boundaries of parameter sets which satisfy the sensitivity constraint in 3 have non smooth gradients, even in the simple case of a process with transfer function

$$G(s) = \frac{1}{(s+1)^4}. \tag{4}$$

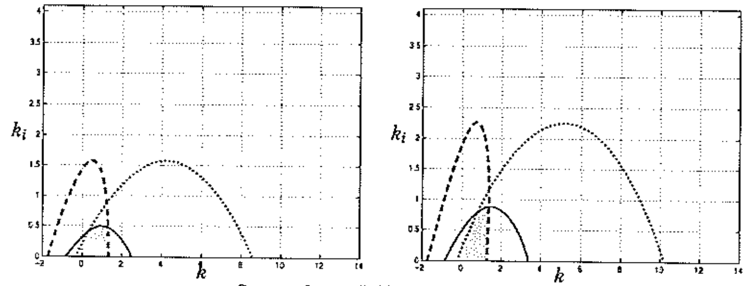
First, let ω and k_d be constant, then the sensitivity constraint (3) represents an ellipse in the k-k_i plane. As ω changes, the ellipses forms an envelope. Since it is complicated to compute the envelope of these ellipses, it is approximated by the envelope of the ellipses vertices which is explained in Figure 2, and below. In the following the left, right and bottom vertex are denoted in the following with L_j, R_j, and B_j, see Figure 2. From (3) the horizontal vertices are given by,

$$\begin{aligned} k &= -C\frac{\alpha}{r^2} \pm \frac{R}{r}, \\ k_i &= -C\frac{\omega\beta}{r^2} + \omega^2 k_d, \end{aligned}$$

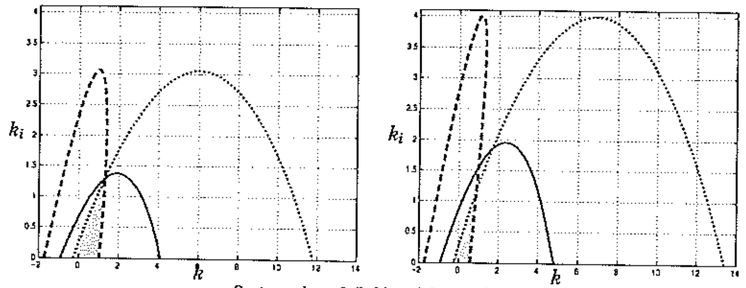
where the left vertex L_j corresponds to a minus sign and the right vertex R_j to a plus sign. The vertical vertices are similarly given by

$$\begin{aligned} k &= -C\frac{\alpha}{r^2}, \\ k_i &= -C\frac{\omega\beta}{r^2} + \omega^2 k_d \pm \frac{\omega R}{r}. \end{aligned}$$

where the bottom vertex B_j corresponds to a minus sign. In the left figure of Figure 2, the loci of the set of vertices (L_i, R_i, B_i) are plotted respectively, for (ω_i, k_d°). A new set of vertices $(L_{i+1}, R_{i+1}, B_{i+1})$ are plotted for $(\omega_{i+1}, k_d^\circ)$. When enough sets of (L_j, R_j, B_j) are generated, they will form an envelope in the k-k_i-plane, that is the dashed line in Figure 2. The interior of the envelope gives the set of (k, k_i)-parameters which satisfy the sensitivity



Systems $k_d = 1$ (left) and $k_d = 2$ (right)



Systems $k_d = 3$ (left) and $k_d = 4$ (right)

Figure 3 The envelope of the ellipses vertices for $k_d = 1,\ 2,\ 3,\ 4$ and $G(s) = 1/(s+1)^4$ with $C = 1$ and $R = 1/1.4$. The shaded areas correspond to the set of (k, k_i)-parameters which satisfy the sensitivity constraint (3). Note how small values of k_d generates a boundary whose maxima has a smooth derivative compared to large values of k_d whose maxima occurs at an edge.

constraint (3). This set corresponds to the shaded area in the right figure of Figure 2.

As an illustration, the envelope of the ellipses vertices have been plotted for $k_d = 1,\ 2,\ 3\ 4$, and the process (4) with $C = 1$ and $R = 1/1.4$ in Figure 3. The shaded areas correspond to the sets of (k, k_i)-parameters satisfying the sensitivity constraint (3). The following observations are made from Figure 3: Small values of k_d generate a boundary whose maxima has a smooth derivative. For large values of k_d the maxima occur at an edge. From a control point of view the case of a boundary whose maxima has a smooth derivative is preferable. It will provide an optimal solution of the PID design, which is not too sensitive for changes in the controller parameters.

An Example: More Insights of the Sensitivity Constraint

In the following, the effects of the characteristics of the sensitivity constraint are examined for the process (4) with respect to time responses, Nyquist diagrams, and Bode diagrams. The controller parameters are given by the upper and lower left figures in Figure 3. The choices $(k, k_i, k_d) = (1.00, 0.50, 1.00)$, and $(1.00, 0.98, 3.00)$, respectively will display the effects of a continuous and a discontinuous derivative of the sensitivity constraint.

To begin with, consider the load responses of the closed loop system in

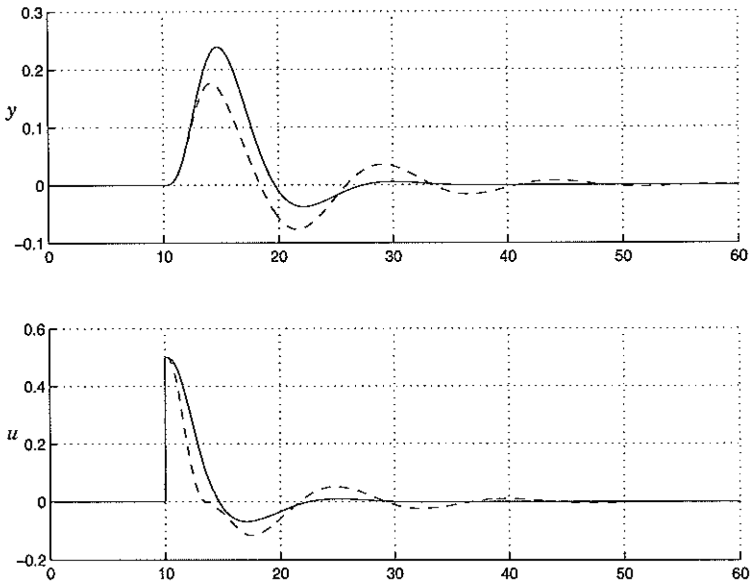


Figure 4 Comparing load responses for the process $G(s) = 1/(s+1)^4$, with the controller parameters $(k, k_i, k_d) = (1.00, 0.50, 1.00)$ (full line), and $(k, k_i, k_d) = (1.00, 0.98, 3.00)$ (dashed line).

Figure 4. To measure the controller performance the integrated absolute error, IAE, defined in Åström and Hägglund (1995), is used. The case $k_d = 1.00$, gives a well damped response, with $IAE = 2.71$. On the other hand, the case $k_d = 3.00$, gives a reduced peak, at the expense of a more undamped response, and a degradation of IAE to 2.86. The closed loop system with $k_d = 1.00$, has one real pole -2.07, and two pairs of complex poles with relative damping ζ, and frequency ω of $(\zeta, \omega) = (0.69, 1.04)$, and $(0.52, 0.47)$, respectively. For $k_d = 3.00$, the closed loop system has one real pole in -2.64, and two pairs of complex poles with $(\zeta, \omega) = (0.41, 1.41)$, and $(0.23, 0.43)$, respectively. Consequently, the second case gives poles which are in general a bit faster, but less damped compared to the first case.

Secondly, consider the Bode diagrams of the loop transfer functions L in Figure 5. The case, $k_d = 1.00$, gives a gain and phase curve which decreases for increasing values of ω. On the other hand, for $k_d = 3.00$ the gain curve is greater or equal to the prior case. The phase curve will have a phase lag and lead because of the undamped modes of L which explains the observed undamped behavior in Figure 4.

Finally, consider the Nyquist curves of L in Figure 6. The case, $k_d = 1.00$, gives a Nyquist curve corresponding to the decreasing gain, and phase curves in Figure 5. From a robustness point of view, this is a desirable feature. On the other hand, for the case $k_d = 3.00$, the phase lag, and lead in Figure 5 gives a Nyquist curve with a cusp, which is undesirable from a robustness point of view. Recall that this case corresponds to the one in Figure 1 where the upper and lower surface will coincide.

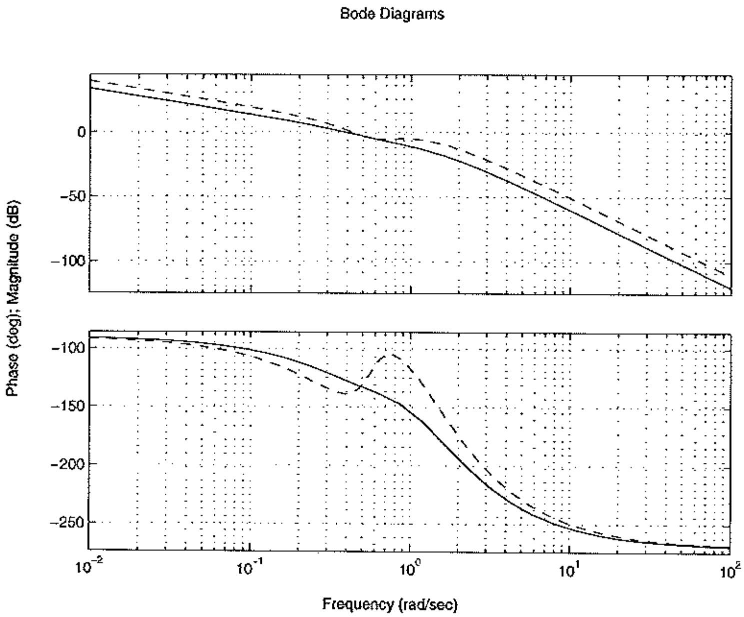


Figure 5 The Bode diagrams of the loop transfer functions for the process $G(s) = 1/(s+1)^4$, with the controller parameters $(k, k_i, k_d) = (1.00, 0.50, 1.00)$ (full line), and $(k, k_i, k_d) = (1.00, 0.98, 3.00)$ (dashed line).

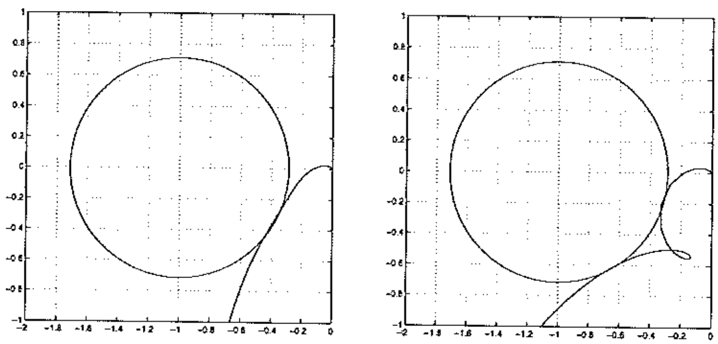


Figure 6 The Nyquist curves of loop transfer functions for the process $G(s) = 1/(s+1)^4$, with the controller parameters $(k, k_i, k_d) = (1.00, 0.50, 1.00)$ (left), and $(k, k_i, k_d) = (1.00, 0.98, 3.00)$ (right).

To sum up these observations: From a control point of view the case $k_d = 1.00$ should be chosen. It gives, a controller with good overall performance and robustness.

PID Design

The previous section explained why the direct generalization of the design method for PI controllers in Åström *et al.* (1998) is not suitable for PID controllers. Consequently, the optimization problem for the design of PID controllers in (1) has to be reformulated by inserting more constraints in addition to the sensitivity constraint such that well damped responses are obtained. This will be fulfilled with two additional constraints which prevents the open loop frequency response from having the phase lags and leads which occurred in the right figure of Figure 6. Thus the optimization problem for the design of PID controllers becomes,

$$\begin{aligned} \max \quad & k_i \\ \text{subject to} \quad & f(k, k_i, k_d, \omega) \geq R^2 \quad \forall\, \omega, \\ & \frac{\dot{\upsilon}\ddot{w} - \ddot{\upsilon}\dot{w}}{(\dot{\upsilon}^2 + \dot{w}^2)^{3/2}} < 0, \\ & \frac{\partial \arg L(i\omega)}{\partial \omega} < 0, \end{aligned} \tag{5}$$

where $L(i\omega)$ is the loop transfer frequency function, and $L(i\omega) = \upsilon(\omega) + iw(\omega)$. Note that in the second constraint $\dot{}$ and $\ddot{}$ corresponds to first and second derivative with respect to ω. The first constraint in (5) express the sensitivity condition, the second constraint specifies that a negative curvature of $L(i\omega)$ should be obtained, and the third constraint prevents $L(i\omega)$ from undesirable phase leads and lags.

For systems with integral action or close to integral action, that is a pole relatively close to zero, the second constraint in Equation (5) becomes, however, too severe. In these cases, it will be appropriate to omit it. Consequently, the proposed design method for PID controllers is separated into two cases depending on the considered system. This separation between integrating and non-integrating processes have been done in several previous design methods, see Åström and Hägglund (1995).

2. Measurement Noise Filtering

When there are substantial measurement noise, it is useful to filter the measurement signal with a filter given by,

$$F_y(s) = \frac{1}{(1 + sT_f)^n}, \tag{6}$$

with $n = 1$, or $n = 2$, see Panagopoulos *et al.* (1999). An application can be found in Eborn *et al.* (1999). The choice of filter time constant T_f in Equation (6) is a trade-off between filtering capacity and loss of performance. A large value of T_f provides an effective noise filtering, but deteriorates the control performance. On the contrary, a small value of T_f keeps the control performance, but with less efficient noise filtering. When comparing different kinds of filters the effectiveness is measured by

$$M_n = \max_{\omega} \left| \frac{G_c(i\omega)}{1 + GG_c(i\omega)} \right|, \tag{7}$$

Table 1 Properties of the PID controllers obtained for system F_yG with $G(s) = 1/(s+1)^4$ and $M_s = 1.4$ for different filter of order $n = 1,\ 2$.

n	1	1	1	2	2	2	-
N	2	5	10	2	5	10	∞
$\arg F_y(i\omega_0)$	-50°	-12.57°	-6.35°	-59°	12.69°	-6.36°	0°
K	0.7895	0.9152	1.0009	0.6496	0.8161	0.9671	1.140
T_i	2.8334	2.3322	2.2867	2.8285	2.4033	2.2963	2.227
T_d	1.2979	1.0406	1.0130	1.3083	1.0716	1.0180	0.999
ω_0	0.5091	0.6695	0.7227	0.4703	0.6226	0.7011	0.790
IAE	4.2425	3.0711	2.7821	4.9524	3.4774	2.8825	2.4209
M_n	1.1392	3.7584	7.9840	1.0726	2.0550	6.1308	∞

defined in Panagopoulos *et al.* (1999), and the control performance by the integrated absolute error (IAE), defined in Åström and Hägglund (1995).

A nice feature of the new design procedure is that it gives a systematic way to determine T_f. The choices

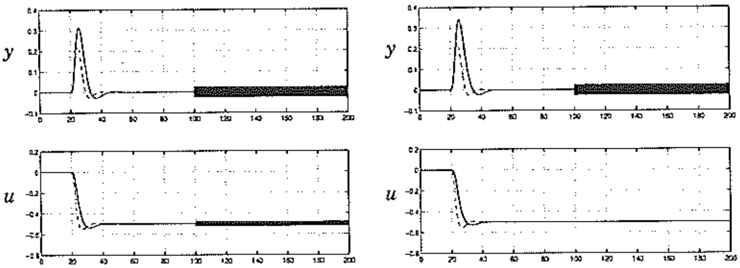
$$T_f = \begin{cases} \dfrac{1}{N\omega_0} & \text{for first order filter,} \\ \dfrac{1}{2N\omega_0} & \text{for second order filter,} \end{cases} \tag{8}$$

makes it possible to determine the effects of the filter at the frequency ω_0. Note, that the maximum of the sensitivity function occurs at the frequency ω_0. Reasonable values of N are in the interval 2–10. In Table 1 the influence of the filter on the loop transfer function at the critical frequency ω_0 is shown by calculating the $\arg F_y(i\omega_0)$. Note, how the special choices of T_f in (8) for the first, and second order filter will give the same change of the loop transfer function at ω_0. Because, the insertion of a filter will modify the loop transfer function which gives minor changes in control loop performance. Adjusted controller parameters are obtained by repeating the design with process G replaced by F_yG, .

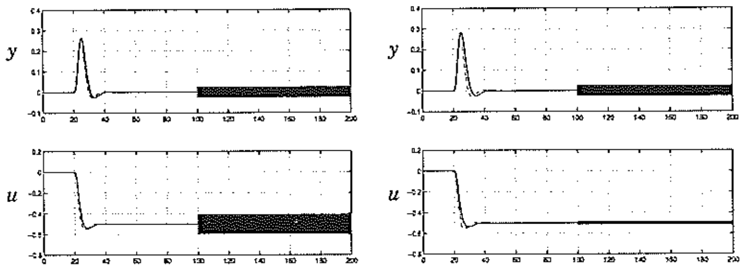
The trade-off between filtering capacity, and loss of control performance is demonstrated in Figure 7. The process G in (4) has been used. Note that the PID controller designed for F_yG is compared to the one designed for G to show the loss of control performance. However, to visualize the effects of measurement noise n on control performance, only the responses of the measurement signal y, and the controller output u for F_yG are shown. The actual noise signal used in the simulation, is a sinusoidal with frequency 40 rad/s. The details of the design calculations and simulations are summarized in Table 1. Furthermore, in Figure 8 the Bode diagrams of the transfer function from measurement noise n to control signal u, that is,

$$u = -\frac{G_c}{(1 + GG_c)}n,$$

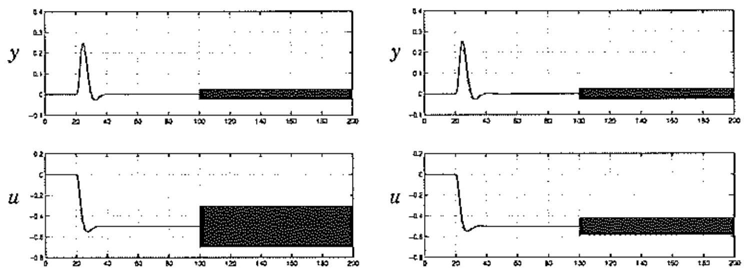
are presented.



Filter constant $N = 2$.



Filter constant $N = 5$.



Filter constant $N = 10$.

Figure 7 The response to a load disturbance followed by measurement noise using first (left) and second (right) order filters with different values of N. The graphs show responses from $F_y G$ (full line) and load disturbance responses from G (dashed line).

Figure 7 shows that a filter gives a deterioration in control performance, compared to the case without filter. Note how the deterioration increases for decreasing values of N. Compare, also, with the IAE values in Table 1. Furthermore, in Figure 7 the efficiency of filtering measurement noise with second order filters compared to first order ones is obvious. Note how the filtering efficiency decreases for increasing values of N. Compare, also, the amplitude curves of Figure 8, and the M_n values in Table 1. Finally, it is noted from Table 1 that the difference in control performance and in filtering capacity between the first and second order filter will be almost

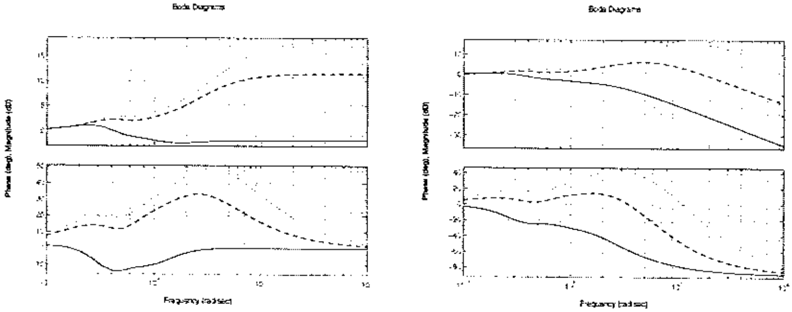


Figure 8 The amplitude curve of the transfer function from measurement noise n to control signal u for the process $F_y(s)G(s)$ with $G(s) = 1/(s+1)^4$ and $M_s = 1.4$. First (left) and second (right) order filter for filter constant $N = 2$ (full line), 5 (dashed line) and 10 (dotted line).

the same due to the choice of T_f in Equation (8).

The advantage of the proposed design method is its systematic way to determine measurement noise filters, such that the noise level is reduced.

3. Set Point Response

The design has so far focused on the response to load disturbances, which is of primary concern. However, it may also be important to have a good response to set point changes. One way to give specifications on the set point response is to consider the transfer function from set point to process output given by

$$G_{sp}(s) = \frac{G(s)G_{ff}(s)}{1+G(s)G_c(s)} = \frac{k_i + bks}{k_i + ks + k_d s^2} \frac{G(s)G_c(s)}{1+G(s)G_c(s)} \frac{F_{sp}(s)}{F_y(s)}, \quad (9)$$

which is defined in Panagopoulos *et al.* (1999),where

$$F_{sp}(s) = 1/(1+sT_{sp}). \quad (10)$$

In order to have a small overshoot, set point weight b and filter F_{sp} will be determined such that the resonance peak of the transfer function $G_{sp}(s)$,

$$M_{sp} = \max \left|G_{sp}(i\omega)\right|, \quad (11)$$

is close to one.

As Equation (11) is a function of the set point weight b and filter F_{sp}, the following approximation is made: the maximum of $|G_{sp}(i\omega)|$ occurs for $\omega = \omega_{mp}$. Note that the maximum of $|L(i\omega)/(1+L(i\omega))|$ occurs at the frequency ω_{mp}. First, the set point weight b, is determined without filtering, that is, $F_{sp}(s) = 1$. By choosing parameter b such that

$$|G_{sp}(i\omega_{mp})| = |\frac{k_i + ibk\omega_{mp}}{k_i + ik\omega_{mp} - k_d\omega_{mp}^2}|M_p = 1, \quad (12)$$

gives

$$b = \begin{cases} \sqrt{A - k_i^2 M_p^2}/k\omega_{mp}M_p & \text{if } A \geq k_i^2 M_p^2, \\ 0 & \text{if } A \leq k_i^2 M_p^2, \end{cases} \quad (13)$$

where $A = k^2\omega_{mp}^2 + (k_i - \omega_{mp}^2 k_d)^2$. Only positive values of b are allowed as negative ones may result in inverse step responses in the control signal, which is an undesirable way to reduce overshoots.

In those cases when $b > 0$, Equation (12) holds, but it may happen that $|G_{sp}|$ becomes large for values of $\omega \neq \omega_{mp}$. This is avoided if b is restricted to values such that

$$|\frac{k_i + ibk\omega}{k_i + ik\omega - k_d\omega^2}| \leq 1, \quad (14)$$

for all frequencies ω. This, gives the following additional constraints on b,

$$b \leq \begin{cases} \sqrt{T_d/T_i} & T_i \leq 4T_d, \\ \dfrac{2T_d/T_i}{1 - \sqrt{1 - 4T_d/T_i}} & T_i \geq 4T_d. \end{cases} \quad (15)$$

Notice, that for set point weight of PI controllers, ($k_d = 0$), it follows from Equation (14) that the corresponding upper limit of b is $b \leq 1$.

If $b = 0$, it is not sure that the design objective (12) will be obtained. If the set point response is important and the M_p value is large, the output of $G_{sp}(s)$ may be filtered by the first order filter in Equation (10). The filter time constant T_{sp} is chosen as

$$T_{sp} = \sqrt{k_i^2 M_p^2/A - 1)/\omega_{mp}^2}, \quad (16)$$

such that $|F_{sp}(i\omega_{mp})G_{sp}(i\omega_{mp})|^2 = 1$.

Improving the Set Point Weighting from a Practical Aspect

Background: Experimental tests of the proposed PID design have been made at the pulp and paper plant of Modo Paper in Husum, Sweden. To set the set point weight to values different from 1 was not possible which resulted in unacceptable overshoots. The problem was overcome by using $b = 1$ and the set point filter in (10). The solution was proposed by Nordin (1999)

The filter time constant T_{sp} is chosen such that the design objective (12) is obtained. In the PI case, the following time constant was obtained,

$$T_{sp} = \begin{cases} \sqrt{M_p^2 - 1}/\omega_{mp} & \text{if } M_p^2 \geq 1, \\ 0 & \text{if } M_p^2 \leq 1, \end{cases}$$

and for the PID case,

$$T_{sp} = \begin{cases} \sqrt{(k_i^2 - k^2\omega_{mp}^2)M_p^2 - A}/\omega_{mp}A^2 & \text{if } A \leq (k_i^2 - k^2\omega_{mp}^2)M_p^2, \\ 0 & \text{if } A \geq (k_i^2 - k^2\omega_{mp}^2)M_p^2. \end{cases}$$

The insertion of the filter results in slower rise time compared to the case with no filtering. But the settling time, defined in Åström and Hägglund (1995), is the same. Consequently, there is no loss in performance for a set point change. The results when trying it on a real application can be seen in Panagopoulos *et al.* (2000).

4. Numerical Solution of the Design Problem

Finally, a brief discussion of the numerical solution of the optimization problem for the design of PID controllers in Equation (5) is given. The optimization problem is non-convex and the solution is obtained numerically with the Optimization Toolbox in MATLAB 5, which requires substantial calculations. But with today's personal computers, their are no major limitations.

As for most numerical optimization routines it is important to have good initial conditions, and a suitable search interval. A natural choice of initial conditions will be the controller parameters k, and k_i from the PI design in Åström *et al.* (1998), that is,

$$[k^0\ k_i^0\ k_d^0] = [\bar{k}\ \bar{k}_i\ 0],$$

and a suitable search interval is given by,

$$\begin{aligned}\omega_{start} &= \bar{\omega}_0/2,\\ \omega_{stop} &= (\bar{\omega}_{180} + \bar{\omega}_{270})/2.\end{aligned}$$

$\bar{\omega}_0$ is the frequency at which the sensitivity function from the PI design has its maximum. $\bar{\omega}_{180}$ and $\bar{\omega}_{270}$ are the frequencies where the argument of the loop transfer function from the PI design is $-180°$ and $-270°$ respectively.

The following procedure is used to solve the design problem:

1. Give the transfer function of the process. Choose the design parameter expressed by the maximum of the sensitivity function M_s, the maximum of the complementary sensitivity function M_p, or some other norm.

2. Determine the number of constraints, as it differs depending on the considered system.

3. Make a PI design to obtain the initial values $[\bar{k}\ \bar{k}_i\ 0]$, and the frequencies $[\bar{\omega}_0\ \bar{\omega}_{180}\ \bar{\omega}_{270}]$.

4. Solve the design problem with the Optimization Toolbox in Matlab 5.

5. Verify that the resulting controller parameters fulfills the constraints. If not, adjust the initial values or settings in the accuracy of the numerical routine.

Errata to Design of PID Controllers Based on Constrained Optimization

Table 1 should be replaced by

Table 2 The properties of the obtained controllers for system G_1-G_8 with different values of the design parameter M_s.

Process	M_s	k	T_i	T_d	b	T_{sp}	IE	$\frac{IE}{IAE}$	w_0	t_s	M_p
$G_1(s)$	1.4	0.324	6.59	2.35	0.00	0.17	20.4	0.67	0.77	15.3	1.53
	2.0	0.680	4.50	2.27	0.00	0.08	6.61	0.69	0.91	37.8	1.50
$G_2(s)$	1.4	0.325	3.55	1.69	0.69	0.00	10.9	0.99	0.26	48.2	1.00
	2.0	0.555	3.21	1.74	0.00	1.61	5.80	0.64	0.29	39.2	1.32
$G_3(s)$	1.4	15.96	0.212	0.15	0.00	0.26	0.013	0.57	19.1	2.99	1.52
	2.0	43.13	0.189	0.13	0.81	0.00	0.0044	0.75	25.6	3.35	1.65
$G_4(s)$	1.4	1.14	2.23	1.00	0.00	0.27	1.95	0.81	0.79	17.4	1.09
	2.0	2.27	1.91	0.98	0.00	0.53	0.84	0.61	0.95	15.3	1.62
$G_5(s)$	1.4	0.784	2.68	1.24	0.00	0.51	3.41	0.84	0.54	23.9	1.04
	2.0	1.47	2.33	1.25	0.00	0.72	1.59	0.56	0.63	21.0	1.55
$G_6(s)$	1.4	0.635	3.12	1.47	0.00	0.44	4.92	0.88	0.42	27.5	1.01
	2.0	1.15	2.74	1.49	0.00	0.97	2.38	0.56	0.48	23.3	1.50
$G_7(s)$	1.4	0.552	3.57	1.69	0.00	0.25	6.44	0.90	0.34	32.7	1.00
	2.0	0.982	3.14	1.73	0.00	1.24	3.20	0.57	0.39	28.2	1.47
$G_8(s)$	1.4	0.312	2.25	0.80	0.60	0.00	7.22	0.86	0.51	30.6	1.00
	2.0	0.542	2.07	0.79	0.00	1.03	3.82	0.62	0.58	22.4	1.31

5. References

Åström, K. J. and T. Hägglund (1995): *PID Controllers: Theory, Design, and Tuning.* Instrument Society of America, Research Triangle Park, North Carolina.

Åström, K. J., H. Panagopoulos, and T. Hägglund (1998): "Design of PI Controllers based on Non-Convex Optimization." *Automatica*, **34:5**, pp. 585–601.

Eborn, J., H. Panagopoulos, and K. J. Åström (1999): "Robust PID Control of Steam Generator Water Level." In *IFAC'99, 14th World Congress of IFAC*, vol. G, pp. 461–464. Beijing, P.R. China.

Nordin, O. (1999): Private communication. Modo Paper, Husum, SE-890 35 Husum, Sweden.

Panagopoulos, H., K. J. Åström, and T. Hägglund (1999): "Design of PID Controllers based on Constrained Optimization." In *1999 American Control Conference*, pp. 3858–3862. San Diego, California.

Panagopoulos, H., A. Wallén, O. Nordin, and B. Eriksson (2000): "A New Tuning Method with Industrial Evaluations." In *Preprints Control Systems 2000.* Pulp and Paper Technical Association of Canada, Victoria, British Colombia, Canada. Accepted for publication.

III

PID Control Design and $\mathcal{H}_\infty$ Loop Shaping

Hélène Panagopoulos and Karl J Åström

Department of Automatic Control
Lund Institute of Technology
Box 118
S-221 00 Lund, Sweden
Phone: +46 46 2228789, Fax: +46 46 138118
Email: hp@control.lth.se

Abstract This paper shows that traditional methods for design of PID controllers can be related to robust $\mathcal{H}_\infty$ control. In particular it shows how the specifications in terms of maximum sensitivity and maximum complementary sensitivity are related to the weighted $\mathcal{H}_\infty$ norm introduced by Glover and McFarlane (1989). The paper also shows how to use the Vinnicombe metric to classify those classes of systems which can be stabilized by the presented design methods in Åström *et al.* (1998) and Panagopoulos *et al.* (1999).

Keywords $\mathcal{H}_\infty$ controller design. PID controller design. Specifications. Robustness.

1. Introduction

In PID control it is attempted to control a complex process by using a controller with a simple, fixed structure. This differs from main stream control theory where no constraints are imposed on the controllers complexity. Surprisingly the problem of finding a controller of restricted complexity is often more complicated. For this reason, a number of specialized methods have been developed for PID control, see Åström and Hägglund (1995). In this paper it is attempted to show the similarities between some of the methods developed for PID control and main stream control theory which has also been done previously in Grimble (1990). At the general level the problems are very similar because compromises between robustness and performance have to be dealt with.

The first aim of the paper is to show how the design methods for PI and PID controllers presented in Åström *et al.* (1998) and Panagopoulos *et al.* (1999), are related to the $\mathcal{H}_\infty$ loop shaping method developed in Glover and McFarlane (1989), and Vinnicombe (1998).

Traditional design methods for PID controllers make a compromise between robustness and performance. For example, in Shinskey (1990) the constraints on robustness and performance are expressed as, the Nyquist curve of the loop transfer function must lie outside a rectangle which encloses the critical point, and maximize the integral gain, respectively. The design methods for PI and PID controllers in Åström *et al.* (1998) and Panagopoulos *et al.* (1999) are based on this idea, but the robustness constraint is expressed as circles of constant sensitivity, constant complementary sensitivity or a combination of these two. Consequently, the robustness constraint of the proposed design methods can be viewed as a mixed sensitivity problem. In Kwakernaak (1985) various aspects of the performance of linear single input single output feedback systems were translated into required bounds on the sensitivity function and its complement.

The idea of the $\mathcal{H}_\infty$ loop shaping method is to design a controller which provides robustness to process uncertainties, and minimizes the signal transmissions from load disturbance and measurement noise to process input and output. This can be expressed quantitatively by requiring the $\mathcal{H}_\infty$ norm, γ, of a two-by-two transfer function matrix to be small. The value of γ is then viewed as a design variable which determines the performance and robustness of the closed loop system.

The paper shows how the condition that γ is small can be expressed in terms of requirements on the Nyquist curve of the loop transfer function. In particular the curve should be outside a contour which encloses the critical point. An explicit formula is given for the contour of the region which can be bounded internally and externally by circles which are related to the maximum of the sensitivity function and the complementary sensitivity function. This establishes the relation between classical design conditions as in Åström *et al.* (1998) and Panagopoulos *et al.* (1999) to the one of $\mathcal{H}_\infty$ robust control. In Bányász and Keviczky (1999) a frequency interpretation of the $\mathcal{H}_\infty$ criteria is given where the sensitivity and robustness shaping techniques are formulated on Nyquist and Bode plots.

The second aim of the paper is to show how the previous relation between classical design conditions and the $\mathcal{H}_\infty$ robust control given in Vinnicombe (1998) can be used to classify the class of systems which a PID controller will stabilize. In particular it is shown how the specifications for the PID design in Åström *et al.* (1998)and Panagopoulos *et al.* (1999) should be chosen to guarantee that the weighted $\mathcal{H}_\infty$ norm of the transfer function from load disturbance and measurement noise to process input and output is less than a specified value γ. This problem has also been treated in Bombois *et al.* (1999) where the results in Vinnicombe (1998) have been used and the controller design is based on LQG.

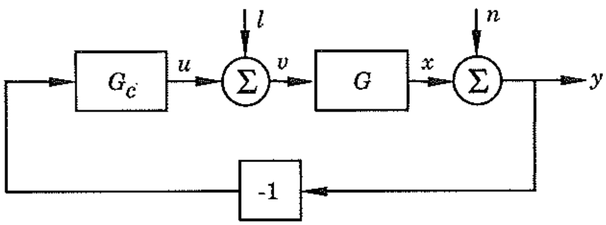


Figure 1 Block diagram describing the design problem of $\mathcal{H}_\infty$ control.

2. $\mathcal{H}_\infty$ Control

The design objectives of a typical process control problem are expressed as requirements on:

- Load disturbance response
- Measurement noise response
- Robustness with respect to model uncertainties
- Set point response

The first three design objectives are well captured by $\mathcal{H}_\infty$ loop shaping which makes it suitable for the design of process controllers, since they mostly operate as regulators. Set point weighting and filtering is then used to obtain a good set point response using the two degrees of freedom structure proposed in Horowitz (1963).

The $\mathcal{H}_\infty$ theory was originally developed by Zames (1981) to emphasize plant uncertainties, but it is also well suited to treat the issues of load disturbance and measurement noise. Consider the block diagram in Figure 1 where the inputs are the load disturbance l and the measurement noise n. The outputs of interest are the process output x, and the signal v which represents the combined action of the load disturbance l and the control signal u. It is assumed that the controller G_c gives a stable closed loop systems. The signals are related through

$$\begin{bmatrix} x \\ v \end{bmatrix} = H \begin{bmatrix} n \\ l \end{bmatrix},$$

where

$$H = \begin{bmatrix} G \\ I \end{bmatrix} (I + GG_c)^{-1} \begin{bmatrix} -G_c & I \end{bmatrix}. \tag{1}$$

The block diagonal elements in Equation (1) are the complementary sensitivity function

$$T = -G(I + GG_c)^{-1} G_c,$$

and the sensitivity function

$$S = (I + GG_c)^{-1}.$$

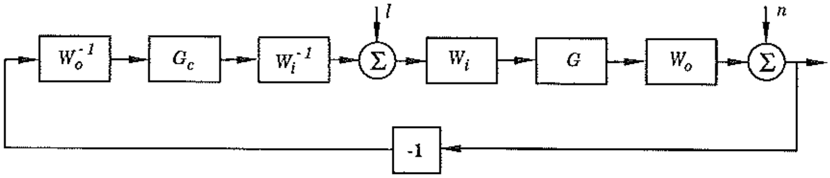


Figure 2 Block diagram describing the design problem of $\mathcal{H}_\infty$ control with frequency weighting.

The off-diagonal elements are the transfer functions

$$G(I + GG_c)^{-1},$$
$$-(I + GG_c)^{-1}G_c.$$

A good closed loop performance requires both x and v to be small with respect to the disturbances l and n which implies making the transfer function H small in the $\mathcal{H}_\infty$-norm, that is,

$$\gamma = \|H\|_\infty, \tag{2}$$

should be small. The fact that γ is small will also guarantee the closed loop system to be robust with respect to model uncertainties. Consequently, the $\mathcal{H}_\infty$ loop shaping is understood as minimizing the complementary sensitivity function T and the sensitivity function S indirectly rather than directly which is to be compared to the design methods for PI and PID controllers in Section 3.

The use of the parameter γ as a criterion for loop shaping was suggested by Glover and McFarlane (1989) and McFarlane and Glover (1992). They recommended values of γ in the range $[2, 10]$. Furthermore, they showed that their design method had nice properties, such as:

- It gives a good controller if one exists.
- The obtained closed loop system is stable against coprime factor uncertainties, Vidyasagar (1985).
- The parameter γ is a good design variable.

Frequency Weighting

Frequency weighting may be introduced in $\mathcal{H}_\infty$ loop shaping to emphasize the response at certain frequencies, see Figure 2. In this case the design of the $\mathcal{H}_\infty$-controller is solved for the transformed system $\bar{G}$ and the transformed controller $\bar{G}_c$, that is,

$$\bar{G} = W_o G W_i,$$
$$\bar{G}_c = W_i^{-1} G_c W_o^{-1},$$

where W_i is the input weight, and W_o is the output weight. Consequently, the transformed system $\bar{H}$ becomes

$$\bar{H} = \begin{bmatrix} W_o G W_i \\ I \end{bmatrix} W_o (I + GG_c)^{-1} W_o^{-1} \begin{bmatrix} -W_i^{-1} G_c W_o^{-1} & I \end{bmatrix}. \tag{3}$$

Note, that the transformation is equivalent to design for the disturbances

$$\bar{l} = W_i l,$$
$$\bar{n} = W_o^{-1} n.$$

Single Input Single Output Systems

In the case of single input single output systems the transfer function H in (1) becomes

$$H = \frac{1}{1 + GG_c} \begin{bmatrix} -GG_c & G \\ -G_c & 1 \end{bmatrix}. \tag{4}$$

The matrix H is of rank 1, and its largest singular value is given by

$$\sigma^2(H) = \frac{(1 + G_c G_c^*)(1 + GG^*)}{(1 + GG_c)(1 + G^* G_c^*)}.$$

It follows from Equation (2) that

$$\gamma^2 = \sup_{\omega} \frac{(1 + G_c G_c^*)(1 + GG^*)}{(1 + GG_c)(1 + G^* G_c^*)}. \tag{5}$$

In the case of single input single output systems frequency weighting will also be simplified, since it is sufficient to use only one weight, that is, $W_i = W$ and $W_o = 1$. The transformed system matrix $\bar{H}$ in (3) becomes

$$\bar{H} = \frac{1}{1 + GG_c} \begin{bmatrix} -GG_c & GW \\ -G_c W^{-1} & 1 \end{bmatrix} \tag{6}$$

3. PID Control

Design methods for PI and PID controllers based on constraints of the maximum sensitivity and complementary sensitivity are described in Åström *et al.* (1998) and Panagopoulos *et al.* (1999). A short review of the controller design is given in this section. The formulation of the design problem for the two methods coincide, they are therefore presented together.

The PID controller is described by the transfer function

$$G_c(s) = k + \frac{k_i}{s} + k_d s,$$

where k is the controller gain, k_i is the integral gain, and k_d is the derivative gain. In reality its structure is more complicated, because of the set point weight and filter, and the measurement noise filter, see Panagopoulos *et al.* (1999). The problem formulation of the design methods are: Maximize the integral gain k_i subject to the robustness constraint that the Nyquist curve of the loop transfer function should lie outside a specified circle. This idea, which goes back to Shinskey (1990) is discussed in detail in Åström *et al.* (1998) and Panagopoulos *et al.* (1999), where the robustness is measured in the classical terms of the maximum of the sensitivity function, M_s, and the maximum of the complementary sensitivity function,

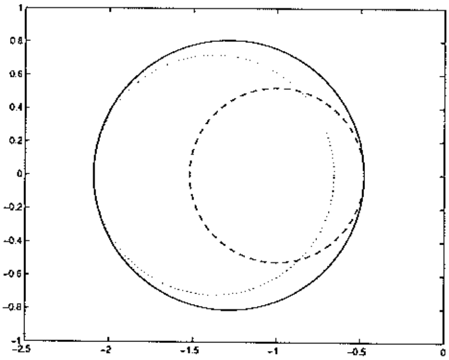


Figure 3 The M_s-circle (dashed line), the M_p-circle (dotted line) and the combined M_sM_p-circle (full line) for $M_s = M_p = 2.0$.

M_p. Thus, the robustness measure provides a transparent design variable which is expressed as

$$M_s = \|(1 + GG_c)^{-1}\|_\infty,$$

that is, the Nyquist curve of the loop transfer function avoids the circle with center at $C = -1$, and radius $R = 1/M_s$, compare with Figure 3, or

$$M_p = \|GG_c(1 + GG_c)^{-1}\|_\infty,$$

that is, the Nyquist curve of the loop transfer function avoids the circle with center at $C = -M_p^2/(M_p^2 - 1)$, and radius $R = |M_p/(M_p^2 - 1)|$, compare with Figure 3. It is possible to replace the constraints on M_s and M_p with a combined one which reduces the computational effort substantially, see Åström *et al.* (1998). For the combined constraint the Nyquist curve of the loop transfer function avoids the circle with center at

$$C = -\frac{2M_sM_p - M_s - M_p + 1}{2M_s(M_p - 1)},$$

and radius at

$$R = \frac{M_s + M_p - 1}{2M_s(M_p - 1)},$$

see Figure 3. The circle will have its diameter on the interval $[-M_p/(M_p - 1),\ -(M_s - 1)/M_s]$. For practical purposes the constraint is not much more stringent than combining the two of M_s and M_p respectively.

The combined constraint is simplified if both the sensitivity and the complementary sensitivity function are less than or equal to M, that is, the Nyquist curve of the loop transfer function avoids the circle with center

$$C = -\frac{2M^2 - 2M + 1}{2M(M - 1)},$$

and radius

$$R = \frac{2M - 1}{2M(M - 1)}.$$

The circle will have its diameter on the interval $[-M/(M - 1), -(M - 1)/M]$.

4. Comparisons

In $\mathcal{H}_\infty$ loop shaping the frequency weights are first specified to obtain desired properties, and then a controller is determined which minimizes γ. For example, for low frequencies the input weight is increased to obtain a high controller gain, and for high frequencies the output weight is increased to achieve high frequency roll-off in the controller. No restrictions are imposed on the controller complexity.

For the design methods in Åström *et al.* (1998) and Panagopoulos *et al.* (1999) for PI and PID controllers the complexity of the controller is fixed, and the low frequency gain of the controller is optimized subject to a robustness constraint expressed in terms of either the maximum of the sensitivity function M_s, or the maximum of the complementary sensitivity function, M_p or a combination of these two.

Below, it is shown how the specifications of the $\mathcal{H}_\infty$ loop shaping and the design methods for PI and PID controllers are related. In particular, it is shown how to choose the specifications of the design methods for PI and PID controllers to guarantee the $\mathcal{H}_\infty$-norm of the frequency weighted transfer function $\tilde{H}$ in (6) to be less than a specified value γ and the determination of its weights W.

The minimization of the robustness measure γ in the $\mathcal{H}_\infty$ design gives a controller that compromises between attenuation of the disturbances l and n. By introducing frequency weights it is possible to emphasize the weighting of the two disturbances by a proper choice of the weight function W, which will serve as a design variable. For reasonable choices of the weight W the largest value of γ will occur in the neighborhood of the crossover frequency. Note that the rejection of low frequency disturbances can be influenced by the weight function but it is not particularly critical.

Consequently, W is regarded as a design variable. What will be the best choice of it? For the frequency weighted transfer function $\tilde{H}$ in (6), the robustness measure γ is given by

$$\gamma^2 = \sup_\omega \frac{(1 + G_c G_c^* W^{-1} W^{-1*})(1 + GG^* WW^*)}{(1 + GG_c)(1 + G^* G_c^*)}. \tag{7}$$

The most favorable frequency weighting is the one that minimizes the numerator of Equation (7). Let $X = WW^*$, then the numerator of Equation (7) becomes

$$1 + GG^* X + G_c G_c^* X^{-1} + GG^* G_c G_c^*,$$

which has its minimum for $X = \sqrt{G_c G_c^* / GG^*}$. The weight factor then becomes,

$$W = \sqrt[4]{G_c G_c^* / GG^*}. \tag{8}$$

Consequently, the weight will typically enhance low and high frequencies. Recall that the low frequency gain of the PID controller is proportional to $1/\omega$.

EXAMPLE 1—DETERMINATION OF THE WEIGHT W
To illustrate the above result the following process is considered,

$$G(s) = \frac{1}{(s+1)^4}. \tag{9}$$

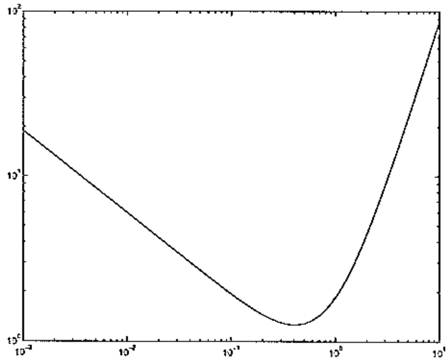


Figure 4 The amplitude function of the computed weight W for the process $G(s) = 1/(s+1)^4$ with PI control ($k = 0.76$ and $k_i = 0.36$).

A PI controller is designed using the design method in Åström *et al.* (1998) with the combined M-circle $M = 2.00$. The controller parameters are given by $k = 0.76$ and $k_i = 0.36$, for which the input weight W given in Equation (8) can be calculated. In Figure 4 the amplitude function of $W(i\omega)$ is shown, which looks quite reasonable as it emphasizes the high and low frequencies. □

The determination of the weight W in (8) makes it now possible to show how the specifications of the design methods for PI and PID controllers should be chosen to guarantee the $\mathcal{H}_\infty$-norm of the frequency weighted transfer function $\bar{H}$ in (6) to be less than a specified value γ. By introducing the weight W in (8) into (7) we find that γ is given by

$$\gamma^2 = \sup_\omega \frac{(1+|GG_c|)^2}{|1+GG_c|^2}. \tag{10}$$

This implies that

$$\gamma = \max_\omega(|S(i\omega)| + |T(i\omega)|). \tag{11}$$

Consequently, Equation (11) shows that the robustness measure γ is related to the values M_s and M_p which are also used as robustness constraints for the design methods presented in Section 3. Notice that even if the weight W in Equation (8) depends on the transfer functions G_c and G, the quantity γ depends only on the loop transfer function GG_c.

Although the robustness measure of the $\mathcal{H}_\infty$ design, and the robustness constraint of the design methods for PI and PID controllers are related, there are some fundamental differences between them. For example, in the $\mathcal{H}_\infty$ design compared to the design methods for PI and PID controllers, the tuning parameter γ is insensitive to k_i for low frequencies, and the requirements on the transfer functions $G_c/(1+GG_c)$ and $G/(1+GG_c)$ appear explicitly in Equation (4). On the other hand, the design methods in Section 3 attempt to maximize k_i explicitly.

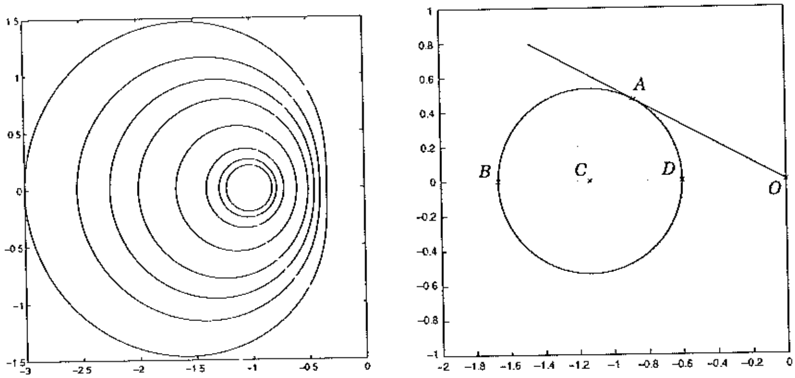


Figure 5 In the left figure the loci of $(1+|L|)/(|1+L|) = \gamma$ for $\gamma = 2$ (outer curve), 2.3, 2.6, 3, 4, 6, 8, 10 (inner curve) are shown. In the right figure points of interest are marked on the γ contour.

The Gamma Contour

Below, a graphical interpretation of the robustness measure γ in Equation (10) will be given. Let $L = GG_c$ be the loop transfer function, then it follows from Equation (10), that

$$\gamma = \sup_{\omega} \frac{1+|L|}{|1+L|}. \tag{12}$$

The condition $\gamma \leq \gamma_0$ says that the Nyquist curve of the loop transfer function should lie outside the contour $(1+|L|)/(|1+L|)$. Therefore, it is suitable to continue the analysis in the complex plane. If $L = re^{i\varphi}$ is introduced straight forward calculations of Equation (12) give the expression of the γ-contour, that is,

$$r(\varphi) = -\frac{\gamma^2\cos\varphi - 1}{\gamma^2 - 1} \pm \sqrt{\left(\frac{\gamma^2\cos\varphi - 1}{\gamma^2 - 1}\right)^2 - 1}.$$

Typical contours for different values of γ are given in the left figure of Figure 5. Straight forward calculations shows that the following relations hold for the distances between the points in the right figure of Figure 5, that is,

$$OA = 1, \qquad OB = -(\gamma+1)/(\gamma-1),$$
$$OC = \gamma^2/(2-\gamma^2), \qquad OD = -(\gamma-1)/(\gamma+1),$$
$$AC = 2\sqrt{(\gamma^2-1)/(\gamma^2-2)^2}, \qquad OB \cdot OD = 1.$$

Relations Between M and γ

Several relations exist between M and γ. If the relation $S + T = 1$ and the triangular inequality are used on Equation (11), the following inequality is obtained,

$$2M - 1 \leq \gamma \leq 2M.$$

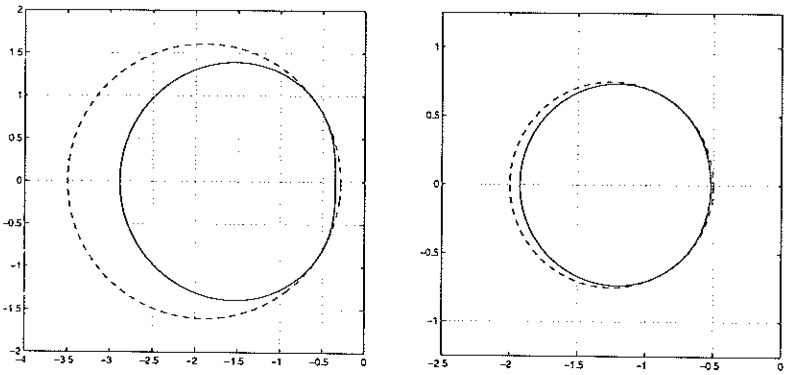


Figure 6 γ-curve (full line) for $\gamma = 2.06$ (left), 3.16 (right) enclosed by the combined M-circle (dashed line) for $M = 1.40$ (left), 2.00 (right).

Table 1 Numerical values of corresponding M- and γ-values.

M	1.37	1.40	1.50	1.60	1.65	1.70	1.80	1.90	1.91	2.00
γ	2.00	2.06	2.24	2.42	2.50	2.60	2.79	2.97	3.00	3.16

This gives an indication of what values of M are suitable to guarantee a certain γ. Sharper results can be obtained: Given a value of γ, the required value of M such that the combined M-circle encloses the γ-curve, should be chosen as,

$$M(\gamma) = \frac{\gamma^2 + 2\sqrt{\gamma^2 - 1}}{2 + 2\sqrt{\gamma^2 - 1}}. \tag{13}$$

The inverse relation is given by

$$\gamma(M) = \sqrt{4M^2 - 4M + 2}. \tag{14}$$

In Figure 6 the contours for constant γ, and the combined M-circles which encloses them are shown. Note how close the contours are for $\gamma = 3.16$ and $M = 2.00$. The figure indicates that the designs based on combined M-curves are not much more conservative than the one based on γ, particularly for $M = 2.00$. However, the calculations for constraint on γ are much more complicated. Consequently, a controller designed for the combined constraint $M_s = M_p = M$ will guarantee a value of γ smaller than the one given by Equation (14). In Table 1 the corresponding numerical values of M and γ are given. The choice $M = 2.00$ guarantees a $\gamma < \sqrt{10}$ which gives good robustness according to Vinnicombe (1998). Lower values of M give even better robustness.

Classification of Stabilizable Systems

In Vinnicombe (1998) much insight was given into the $\mathcal{H}_\infty$ loop shaping procedure. He introduced the generalized stability margin b_{G,G_c} as a performance measure of the feedback system comprising the process G and the controller G_c defined as

$$b_{G,G_c} := \begin{cases} \dfrac{1}{\gamma} & \text{if } [G, G_c] \text{ is stable,} \\ 0 & \text{otherwise,} \end{cases}$$

where b_{G,G_c} lies in the range $[0, 1]$. Recall that γ was the $\mathcal{H}_\infty$ norm of the transfer function (1) that describes how the disturbances l and n propagate through the system. Consequently, a value of b_{G,G_c} which is close to 1 implies a controller with good disturbance attenuation. On the other hand, a value of b_{G,G_c} which is close to 0 implies a controller with poor disturbance attenuation. The solutions of many design examples, see Vinnicombe (1998), and McFarlane and Glover (1992), indicate that a value of $b_{G,G_c} > 1/\sqrt{10}$ or $\gamma < \sqrt{10}$ gives a reasonable robustness and performance, compare with the values in Table 1.

In Vinnicombe (1998) the ν-gap metric was introduced to measures the distance between two systems in terms of their frequency responses $G_1(i\omega)$ and $G_2(i\omega)$. For the scalar case it is defined as,

$$\delta_\nu(G_1, G_2) := \sup_\omega \frac{|G_1(i\omega) - G_2(i\omega)|}{\sqrt{1 + |G_1(i\omega)|^2}\sqrt{(1 + |G_2(i\omega)|^2}},$$

subject to a winding number condition (see Vinnicombe (1998), page 111 and 114). Introduce the two scalar systems $G_1(s) = B_1(s)/A_1(s)$ and $G_2(s) = B_2(s)/A_2(s)$, which are not necessarily stable and where $A_j(s)$ and $B_j(s)$ are polynomials. The winding number condition will then be satisfied if $B_1(s)B_2(-s)+A_1(s)A_2(-s)$ have the same number of roots as $\deg A_2(-s)$ in the right half plane. If the winding condition is not satisfied then $\delta_\nu(G_1, G_2)$ is defined to be 1. Thus, the ν-gap metric, δ_ν, take values in the range $[0, 1]$. It is an important measure, since it means that the distance between two linear plants can be estimated directly from their frequency responses.

Vinnicombe has derived very interesting results relating the generalized stability margin to robustness and model uncertainty. He showed that if the closed loop system (G, G_c) consisting of the controller G_c and the process G is stable with a generalized stability margin $b_{G,G_c} > \beta$, and $\bar{G}$ is a process which is "close" to G in the sense that $\delta_\nu(G, \bar{G}) \leq \beta$, then G_c is also guaranteed to stabilize $\bar{G}$.

A PID controller which is designed to satisfy the constraints $M_p < M_0$ and $M_s < M_0$ thus guarantees that $\gamma \leq \gamma(M_0)$. The results of Vinnicombe then gives a complete characterization of the class of systems which are stabilized by the controller.

EXAMPLE 2—CLASSIFICATION OF STABILIZABLE SYSTEMS

To illustrate the results above the process and controller in Example 1 are considered. The generalized stability margin for this design is $b_{G,G_c} = 0.32$. It is now possible to investigate the effects of model uncertainties. Assume for example that the true process is instead governed by the following model,

$$\bar{G}(s) = \frac{1}{(2s + 1)^2}.$$

To verify if the controller designed for the process G in Example 1, also works for $\bar{G}$, the ν-gap metric is calculated between G and $\bar{G}$, and compared to the generalized stability margin. The winding number condition of the ν-gap metric is fulfilled. Straight forward calculations gives that the largest value of $\delta_\nu(G, \bar{G})$ (≈ 0.25) is less than the smallest value of the generalized stability margin (≈ 0.32). Consequently, the controller designed for G will stabilize the system $\bar{G}$. □

5. Conclusions

Traditional methods for designing PID controllers were related to robust $\mathcal{H}_\infty$ control, in particular the specifications of design methods for PI and PID controllers in Åström *et al.* (1998) and Panagopoulos *et al.* (1999) to the one of the weighted $\mathcal{H}_\infty$ norm in Glover and McFarlane (1989), denoted γ. The requirement of a sufficiently small γ was expressed as, the Nyquist curve of the loop transfer function should lie outside a certain region. The region showed to be bounded internally and externally by circles closely related to the ones of constant sensitivity and constant complementary sensitivity, which is of use for efficient computations, see Åström *et al.* (1998). Furthermore, it is shown how to use the Vinnicombe metric in Vinnicombe (1998) to classify those classes of systems which can be stabilized by the presented design methods for PI and PID controllers in Åström *et al.* (1998) and Panagopoulos *et al.* (1999).

6. References

Åström, K. J. and T. Hägglund (1995): *PID Controllers: Theory, Design, and Tuning.* Instrument Society of America, Research Triangle Park, North Carolina.

Åström, K. J., H. Panagopoulos, and T. Hägglund (1998): "Design of PI Controllers based on Non-Convex Optimization." *Automatica*, **34:5**, pp. 585–601.

Bányász, C. and L. Keviczky (1999): "Geometric tools to design robust control systems." In *IFAC'99, 14th World Congress of IFAC*, vol. G, pp. 277–282. Beijing, P.R. China.

Bombois, X., M. Gevers, and G. Scorletti (1999): "Controller validation for a validated model set." In *European Control Conference'99.* Karlsruhe, Germany.

Glover, K. and D. McFarlane (1989): "Robust stabilization of normalized coprime factor plant descriptions with H_∞-bounded uncertainty." *IEEE Transactions on Automatic Control*, **34:8**, pp. 821–830.

Grimble, M. J. (1990): "$\mathcal{H}_\infty$ Controllers With a PID Structure." *Transactions of the ASME Journal of Dynamic Systems, Measurement, and Control*, **112**, pp. 325–336.

Horowitz, I. M. (1963): *Synthesis of Feedback Systems.* Academic Press, New York.

Kwakernaak, H. (1985): "Minimax Frequency Domain Performance and robustness optimization of linear feedback systems." *IEEE Transactions on Automatic Control*, **10:30**, pp. 994–1004.

McFarlane, D. and K. Glover (1992): "A loop shaping design procedure using H_∞ synthesis." *IEEE Transactions on Automatic Control*, **37:6**, pp. 759–769.

Panagopoulos, H., K. J. Åström, and T. Hägglund (1999): "Design of PID Controllers based on Constrained Optimization." In *1999 American Control Conference*, pp. 3858–3862. San Diego, California.

Shinskey, F. G. (1990): "How good are our controllers in absolute performance and robustness?" *Measurement and Control*, **23**, May, pp. 114–121.

Vidyasagar, M. (1985): *Control System Synthesis: A Factorization Approach*. MIT Press, Cambridge, Massachusetts.

Vinnicombe, G. (1998): *Uncertainty and Feedback. H_∞ Loop-Shaping and the ν-Gap Metric*. To be published.

Zames, G. (1981): "Feedback and optimal sensitivity: Model reference transformations, multiplicative seminorms and approximate inverses." *IEEE Transactions on Automatic Control*, **AC-26**, pp. 301–320.

IV

G-2e-18-2
14th Triennial World Congress, Beijing, P.R. China

ROBUST PID CONTROL OF STEAM GENERATOR WATER LEVEL

Jonas Eborn, Hélène Panagopoulos and Karl Johan Åström

Department of Automatic Control
Lund Institute of Technology, Lund, Sweden
jonas@control.lth.se

Abstract: This paper presents a new approach to a classical solution of the level control problem. It is a bottom up procedure based on PID control and gain scheduling. The controllers are tuned with a novel optimization approach. The performance of the designed controller is very good considering the low complexity. All of the specifications are met within the range of what is possible. The effort required to design the controller using this approach is also discussed.

Keywords: PID control, Model reference, Windup, Gain scheduling, Level control, Steam generators.

1. INTRODUCTION

Control of the water level in a steam generator is one of the most important control loops in power plants. The difficult shrink & swell phenomenon, see Bell and Åström (1996), gives rise to an inverse response which limits control performance. Failure to keep a tight level control is a major cause of plant shutdowns in nuclear and conventional power plants, see Ambos et al. (1996).

Very often, PID is used for water level control in both conventional and nuclear power plants. However, tuning of PID controllers such that satisfactory performance is maintained over a wide operating range, while maintaining robustness against disturbances and plant uncertainty is a subject that has not been thoroughly treated in the literature. Since the 60's and 70's the interest in the research community has been directed towards optimal and adaptive control, and later on control design based on robustness optimization.

In this paper it is shown that the simple PID control structure can be used with good results on an industrial plant. With a design method based on constrained optimization both performance and robustness is obtained.

2. CONTROL STRUCTURE

The chosen control structure is a PID controller with feedforward from the disturbance, Q_v, and a reference model. It is similar to the so called *three-point controller* in Åström and

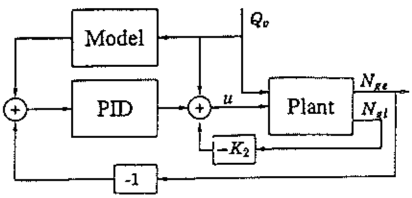


Fig. 1. Structure of the PID controller with proportional feedback, feedforward and reference model.

[1] This work has been supported by the Swedish Research Council for Engineering Science, contract 95-759.

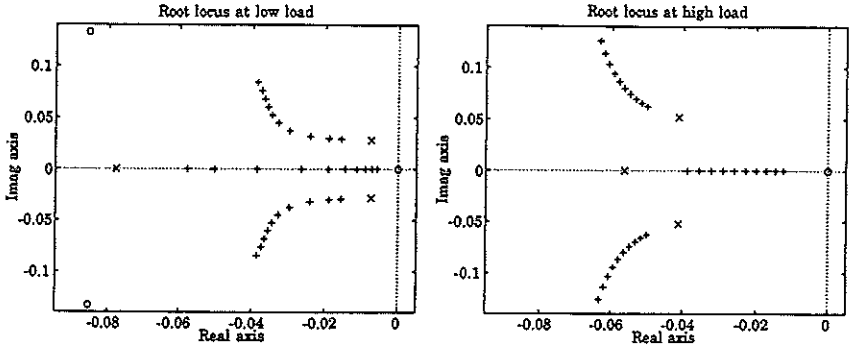


Fig. 3. Root loci with respect to $K_2 \in [5, 50]$ at two different loads. There are more roots farther away from the origin.

Hägglund (1995), which is commonly used for drum boiler level control.

To exploit the second available measurement, N_{gl}, simple proportional feedback was used. This is useful at low load where it makes the PID controller considerably faster, T_i can be lowered from 80 seconds to 25. However, at high load there are no benefits with feedback from N_{gl} and so $K_2 = 0$.

The reference model is used to "hide" the shrink & swell effect from the controller and thus avoid excessive undershoot. The model is a simple second order transfer function that gives the same transient response as the plant to a disturbance but settles to zero. The model used is

$$G_{model}(s) = \frac{2\tau_e s}{(1 + 2\tau_e s)(1 + 2\tau_g s)}$$

where τ_e and τ_g are time constants of the nominal plant model at low load. The process

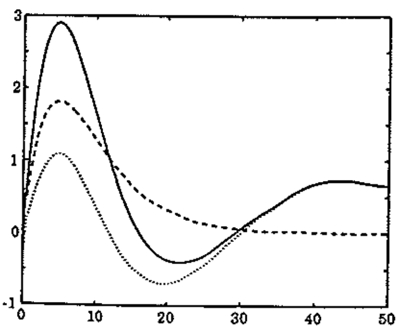


Fig. 2. Water level(—) compared to reference model output(– –) and the difference(···) for a step disturbance in Q_s at low load.

output compared to the reference model and the difference seen by the PID controller is shown in Figure 2.

The implemented PID algorithm is discrete with sampling period 0.5 second, derivative filtering and anti-windup. The reference model is sampled with the same period as the PID.

3. CONTROL DESIGN

The design of the PID controller and extra proportional feedback was made in two steps; first K_2 was chosen from a root locus plot, secondly the PID parameters were calculated using the nominal transfer functions with the proportional feedback. The root loci at different loads is shown in Figure 3. At low load K_2 was chosen as 20, which is just above the "knee" on the root locus. At high load $K_2 = 0$ since it did not give any additional damping nor increased the bandwidth.

The tuning of the controllers is based on a new numerical method for design of PI controllers, see Åström et al. (1998); Panagopoulos (1998), that results in a constrained optimization problem. The primary design goal is to obtain good load disturbance responses. For non-oscillatory systems this is the same as minimizing the integrated control error, IE. It can be shown that this is equivalent to maximizing integral

Table 1. PID parameters used at different loads. Anti windup tracking time is $T_r = T_i/2$. The N_{gl} fault only affects integral time at low load.

Case	High load	Low load	N_{gl} fault
K	0.8	0.35	-
T_i	45	25	[45, 80]
T_d	10	5	-
K_2	0	20	-

Table 2. Achieved results for the different specifications. Results with sensor fault and delay margins were evaluated on the simplified model, the other results are from the EdF evaluation on the detailed model. The recommended delay margins are shown in the table, recommended settling time is 100 s.

Case	High load		Low load		Saturation		Fault handling	
	Step	Ramp	Step	Ramp	Step up	Step down	Step	Ramp
Time $\lvert N_{gt}\rvert > 5$	15 s	-	-	197 s	-	11 s	-	210 s
Settling time	198 s	194 s	11 s	386 s	155 s	138 s	50 s	490 s
Delay margin	21 s (10 s)		23 s (25 s)		-		33 s (10 s)	

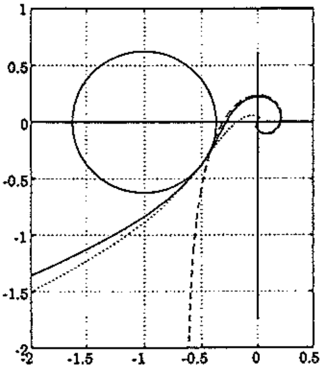


Fig. 4. Nyquist plots of the open loop frequency response designed for $M_s = 1.6$ in three cases; low(- - -) and high load(···) and low load with sensor fault(—).

gain, $k_i = K/T_i$. Robustness is generated by requiring that the maximum of the sensitivity is less than a specified value M_s. The new numerical algorithm thus maximizes k_i with the constraint on K that the sensitivity peak should be less than M_s.

The design was made for nominal models on the three cases; 80% load, 10% load and 10% load with sensor fault ($K_2 = 0$). The tuning parameter was chosen as $M_s = 1.6$. The resulting Nyquist curves of the loop transfer functions are shown in Figure 4.

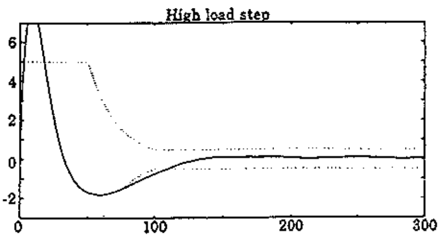


Fig. 5. Water level for step at high load. The dotted lines are the *preferred* magnitude limits on the error. The *strict* limit is ±20.

The numeric method used for the tuning gives PI parameters, but in the implemented controller also derivative action was used by setting $T_d \approx T_i/4$, see Table 1. After EdF completed the evaluation also a full algorithm for optimizing K, T_i, T_d simultaneously has been developed by Panagopoulos (1998). Using this algorithm on the benchmark problem gives PID parameters similar to the ones used in this article.

3.1 *Gain scheduling*

The advantages of using a classical control structure like the PID is the immediate intuition of the control parameters and the controller's continuous behaviour with smoothly varying parameters. Consequently, it is easy to use gain scheduling to smoothly change controller parameters between the two operating points at low and high load. For N_{gl} sensor faults, T_i is changed immediately when a sensor fault is indicated. Naturally, a bumpless algorithm is used and, as can be seen in Figure 6, there is no visible difference between the step responses until 20 seconds after the fault occurs ($t_{fault} = 20$s).

4. CONTROLLER EVALUATION

The benchmark specifications give eight normal "cases" that should be handled by the controller. These include steps and ramps at different loads and additional steps to check saturation

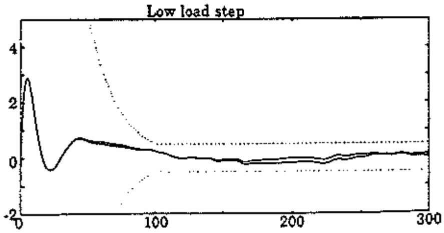


Fig. 6. Water level for step at low load. Also the case with fault on the N_{gl} measurement is shown, but there is no significant difference.

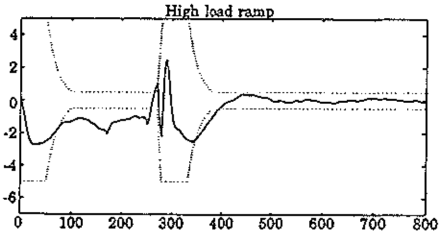


Fig. 7. Water level for ramp at high load.

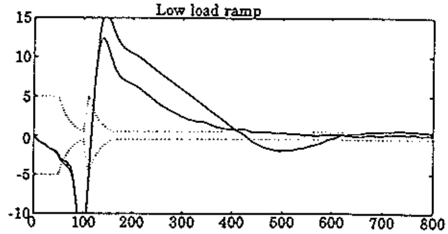


Fig. 9. Water level for ramps at low load. Also the case with fault on the N_{gl} measurement (top curve) is shown. The controller still stabilizes the system, although there is less damping.

and sensor fault handling capabilities. The results from the EdF evaluation of our controller are summarized in Table 2.

Most of the specifications are fulfilled by the proposed controller. The settling times at high load are a little long, but the output only goes a little outside the ±0.5 lines. The low load ramp case is very difficult since saturation of the feedwater flow makes it impossible to remove the level peak caused by excessive shrink & swell at only 5% load. Recommended delay margins are shown in parenthesis in the table.

Plots of all eight cases are given in Figures 5-10. Responses to step disturbances at high and low load are shown in Figures 5-6, ramp disturbances are shown in Figures 7-9 and plots of the saturation test cases at only 5% load are shown in Figures 8-10. All the plots were made in Matlab on the simplified S-function model supplied with the benchmark problem. The strange-looking response to the high load ramp at $t = 280$s in Figure 7 is due to bumpy parameter switching in the simplified model of the plant. Results from the evaluation on the detailed model are more regular.

5. CONCLUSIONS

A solution to the water level control benchmark using a PID structure with feedforward has been presented. The proposed controller designed with constrained optimization meet most of the specifications well. Cases with actuator saturation and sensor fault are handled with almost no degradation of performance.

REFERENCES

AMBOS, P., G. DUC, and C.-M. FALINOWER (1996): "Loop shaping H_∞ design applied to the steam generator level control in EdF nuclear power plants." In *Proceedings of the 5th IEEE Conference on Control Applications*, pp. 1299–1304. Dearborn, Michigan.

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

ÅSTRÖM, K. J., H. PANAGOPOULOS, and T. HÄGGLUND (1998): "Design of PI controllers based on non-convex optimization." *Automatica*, 35:5.

BELL, R. D. and K. J. ÅSTRÖM (1996): "A fourth order non-linear model for drum-boiler dynamics." In *IFAC'96, Preprints 13th World Congress of IFAC*, vol. O, pp. 31–36. San Francisco, California.

PANAGOPOULOS, H. (1998): *PID Controller Design*. Lic Tech thesis ISRN LUTFD2/TFRT--3222--SE, Department of Automatic Control, Lund Institute of Technology, Lund, Sweden.

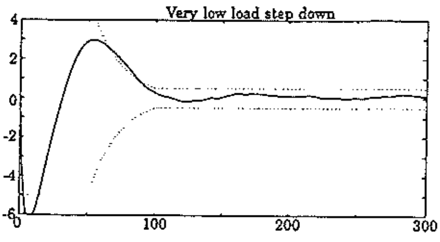


Fig. 8. Water level for step down at very low load, from 10% to 5% load.

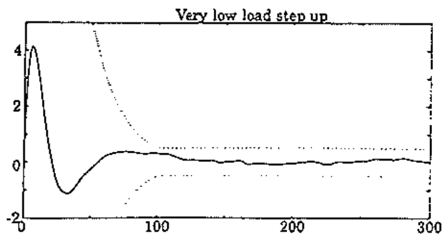


Fig. 10. Water level for step up at very low load, from 5% to 10% load.

V

A New Tuning Method with Industrial Evaluations

H. Panagopoulos* **A. Wallén***
O. Nordin† **B. Eriksson†**

Abstract The paper presents applications of a new interactive tool for system identification and controller design. The tool has been applied to a starch boiler and to a steam pressure control in a drying section of a paper machine at the pulp and paper company Modo Paper in Husum, Sweden. The tool has a graphical user interface for rapid system identification and PI/PID controller design.

Keywords Application. Interactive tool. System Identification. PI/PID controller design. Pulp and paper. Starch boiler. Temperature control. Drying section of a paper machine. Steam pressure control.

1. Introduction

In this paper the applications of a new interactive tool for system identification and controller design are presented. The tool has been applied to a starch boiler and to a steam pressure control in a drying section of a paper machine at the pulp and paper company Modo Paper in Husum, Sweden. It has a graphical user interface for rapid system identification, see Wallén (1999) and controller design, see Åström *et al.* (1998) and Panagopoulos *et al.* (1999). The tool is implemented using the graphics and numerics in MATLAB.

The identification procedure of the tool is based on open loop step response analysis. Through the graphical user interface the user can manipulate the step response for a given model structure directly and then perform a least squares fit to data.

The controller design of the tool determines a PI or a PID controller upon the user's request, where a tuning parameter has to be set which gives a trade-off between robustness and performance. The controller is determined such that it captures the engineering criteria: load disturbance re-

* Department of Automatic Control, Lund Institute of Technology, Box 118, S-221 00 Lund, Sweden, Email: hp@control.lth.se, andersw@control.lth.se
† Modo Paper Husum, SE-890 35 Husum, Sweden, Email: Oskar.Nordin@Modopaper.com, Borje.Eriksson@Modopaper.com

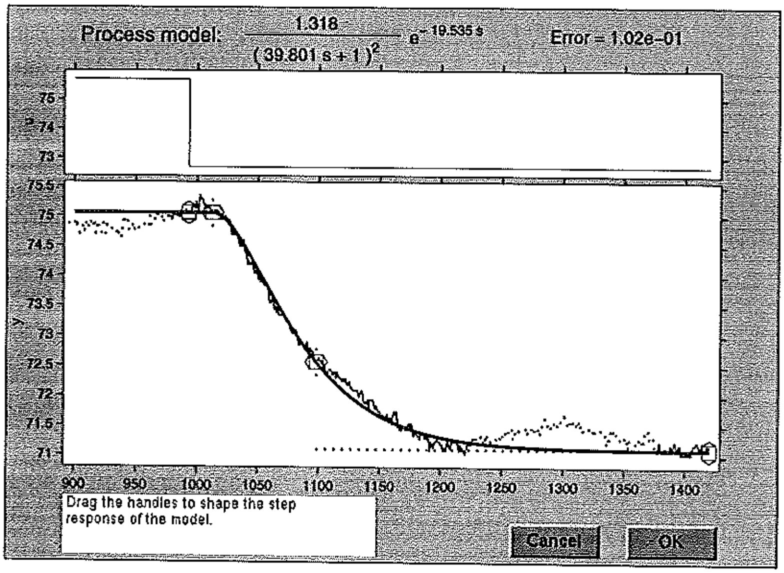


Figure 1 The graphical user interface for editing the open loop step response for a given model structure.

jection, robustness to model uncertainties, rejection of measurement noise and set point tracking. These engineering specifications are taken into account through the solution of an optimization problem.

In the first part of Section 2 a review of the graphical user interface for rapid system identification in Wallén (1999) is given. The second part of the section gives a review of the design of PI controllers in Åström *et al.* (1998) and of PID controllers in Panagopoulos *et al.* (1999). The application of the new tool is presented in Section 3 for the temperature control of a starch boiler and in Section 4 to the steam pressure control of the drying section in a paper machine.

2. The Tool

In this section a brief description of the theory behind the interactive tool is explained. As with almost all design methodologies, it consists of a characterization of the process dynamics followed by an optimal control design step. A short review of the process modeling is given in Subsection 2.1 and of the controller design in Subsection 2.2.

2.1 Process Modeling

The process modeling is an important part of the design methodology. For PID design methods, these models typically consist of two or three parameters, for instance, the dead time, the time constant, and the static gain. When using simple design methods as Ziegler and Nichols (1942), these simplified models will probably suffice. However, when the design of the

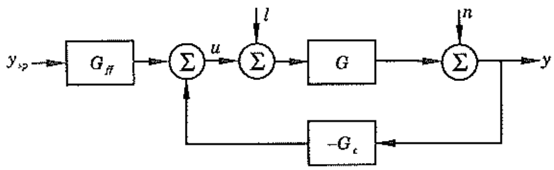


Figure 2 Block diagram describing the problem formulation of the controller design in Subsection 2.2.

controllers includes more sophisticated objectives as in Subsection 2.2, a more detailed model is needed.

This paper uses the system identification tool presented in Wallén (1999). The tool is based on analysis of open loop step responses, or sequences of step responses. The main reason for this is that step responses are easy to generate and that they are intuitive for the user. The tool provides a number of pre-defined model structures formulated as continuous-time rational transfer functions with time delays. When the user has selected a model structure, he may edit the model parameters interactively by dragging handles in the graphical user interface, shown in Figure 1. In this case there are handles for the time delay, the time constant, and the initial and final levels of the step response. Both the process model and the plotted model response change continuously when a handle is moved. The tool can also find the process model which minimizes the mean square error between the model output and the experimental data. This is an efficient way to obtain a process model based on step response data.

The tool also includes other features, for example a possibility to remove data points such as outliers and other disturbances interactively. In Figure 1 this is indicated by the dotted parts of the experimental data. Further, the tool is able to calculate PI and PID controller parameters according to Subsection 2.2 and to simulate the resulting closed loop system.

2.2 Control Design

An efficient controller design should reflect the essence of a real control problem. It should capture engineering criteria such as load disturbance rejection, robustness to model uncertainties, rejection of measurement noise, and set point tracking. The problem formulation for the design of PI controllers in Åström *et al.* (1998) and of PID controllers in Panagopoulos *et al.* (1999) captures these specifications which is illustrated in Figure 2. A process with transfer function G is controlled with a PID controller with two degrees of freedom. The design objective is to determine the controller parameters in G_c and G_{ff} so that the system output y behaves well with respect to changes in the set point y_{sp}, load disturbance l, and measurement noise n. The controller can thus be characterized by two transfer functions: the feedback G_c and the feed forward G_{ff} transfer functions,

$$G_c(s) = K(1 + \frac{1}{sT_i} + T_d s)F_y(s),$$
$$G_{ff}(s) = K(b + \frac{1}{sT_i})F_{sp}(s),$$

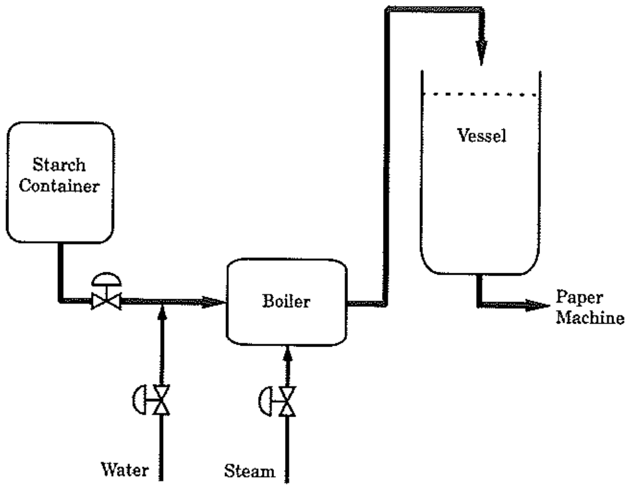


Figure 3 An illustration of the control problem of the starch boiler.

where K, T_i, T_d and b are the controller parameters and F_y, F_{sp} are two low pass filters of first or second order.

The primary design goal is to achieve good rejection of load disturbances, with a constraint on the robustness with respect to model uncertainties, to guarantee a stable closed loop system. A constrained optimization problem has to be solved, where the transfer function of the process G is needed. The solution results in a PI/PID controller, where measurement noise n is treated by filtering the output y with F_y, and good set point following is obtained by using the set point weight b or if needed the set point filter F_{sp}, for more details see Panagopoulos *et al.* (1999). The controller design has one tuning parameter, the maximum of the sensitivity function M_s, to be set by the user which gives a trade-off between robustness and performance, see Panagopoulos *et al.* (1999).

3. Application to a Starch Boiler

The interactive tool was applied to a starch boiler at the pulp and paper company Modo Paper in Husum, Sweden, with the goal of improving the existing control. Below, the control problem is described and in Figure 3 it is illustrated.

At start up the boiler is first heated with water which is passed to the vessel. This will lower the concentration of the starch in the vessel which is undesirable but unavoidable. When the correct temperature is reached, the flow into the boiler is switched from water to starch. The starch entering the vessel must be kept at about 140°C. The medium is heated by injecting steam into the boiler. To influence the starch concentration in the vessel as little as possible it is desirable to heat the medium as fast as possible.

The control problem is to regulate around the operating point 140°C of the boiler which corresponds to 69% of the signal range. The main dis-

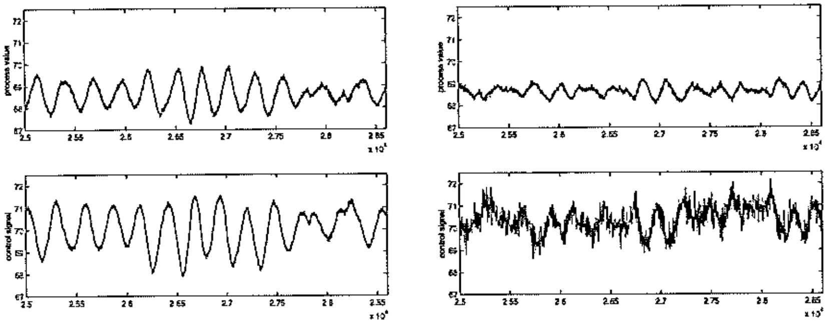


Figure 4 Validation of the PI controller (left) compared to the PID controller (right) for the temperature control of a starch boiler.

turbances are varying steam pressure and change in starch flow due to changes in viscosity. Since we have a temperature control loop it is suspected that a PID controller may give better performance compared to the previously used PI controller.

Modeling

In order to design the PID controller with the method in Section 2, a model of the starch boiler is required. With an open loop step response, the tool described in Section 2 gives the transfer function

$$G(s) = \frac{1.34e^{-12.9s}}{(43.3s+1)^2}.$$

Figure 1 shows the tool after the least squares fit has been performed. The data points that have been deselected in the figure correspond to load disturbances during the experiment.

Controller Design

A PID controller was designed for the tuning parameter $M_s = 1.6$, which provides a good compromise between performance and robustness, with the following controller parameters: $K = 1.35$, $T_i = 40$ and $T_d = 19$. The existing controller had a measurement noise filter on the derivative term, whose filter constant was chosen to $N = 8$. Furthermore, the control system did not allow for values of the set point weight parameter b other than 1. The controller parameters of the previous PI controller were: $K = 0.80$, $T_i = 60$. They were obtained by either trial and error, or from the given settings at start up of the plant.

Validation

The PID controller was compared to the original PI for the temperature control of the starch boiler. A performance validation was made for set point following and rejection of load disturbances. The validation of the PI and PID controller is shown in the left figure respectively in the right figure of Figure 4. The variance of the controller error, $e = y_{sp} - y$, was

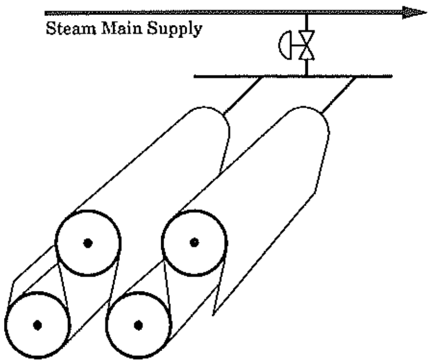


Figure 5 An illustration of the control problem of the steam pressure control in drying cylinders of the drying section.

used as a performance measure. It is defined as,

$$ISE = \frac{\sum_{i=1}^{M} e(k)^2}{M},$$

where M is the number of measured data points. The ISE of the PI and the PID controller became 0.308 and 0.0499 respectively. Consequently, the PID controller gives a better performance compared to the PI which is also seen in Figure 4. The variations around the operating point 69% are reduced in the right figure compared to the left one because of the well-tuned PID controller. Because of the valve hysteresis problems, together with periodic disturbances, for example in the steam pressure, which affects the system, the variations around the operating point will remain. The noise level in the control signal for the PID controller is increased due to the derivative action. But the use of the measurement noise filter on the derivative term reduced the noise enough for not wearing out the valve.

Note that the mean of the process value in Figure 4 is not exactly 69%. This is due to the data logger which measures a voltage/current over a resistor with a given tolerance which induces an improper measurement interval.

Conclusion

The PID design was appreciated by the operators as it gave a faster start up of the process, a shorter waiting time and lowered production costs.

4. Application to Steam Pressure Control in Drying Cylinders of a Drying Section

The design method for PID controllers in Section 2 has been applied to a pressure control in a drying section of a paper machine at Modo Paper in Husum, Sweden, with the goal of improving the existing control. Below, the control problem is described and in Figure 5 it is illustrated.

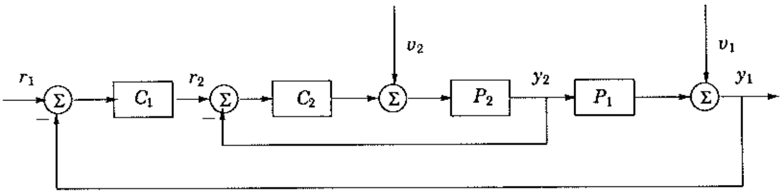


Figure 6 A block diagram description of the steam pressure control in drying cylinders of the drying section.

The drying section of a paper machine normally consists of two rows of totally 40-90 drying cylinders. In normal operation paper surrounds the drying cylinders in the drying section. The cylinders are divided into 4-5 groups, where all the cylinders in one group keeps the same steam pressure. The controllers control on the one hand the pressure difference over each group, that is, the moisture content in the paper, and on the other hand the pressure in the steam main supply to each group. The control of the latter case has been considered here.

The control problem is described in Figure 6 where the goal is to keep the moisture content in the paper y_1 constant despite the disturbances v_1 and v_2. The former one represents the moisture content of the paper along the paper web in the machine direction and the latter one the pressure in the steam main supply. The control is realized with cascade control, where the outer loop's controller C_1 controls the moisture content in the paper y_1 and the inner loop's controller C_2 controls the pressure in the group represented by $P_2(s)$. Every 30 second a new set point r_1 is generated and as a consequence a new set point of the steam pressure r_2. Because, 30 seconds is the time it takes for the scanner of the moisture sensor to sweep over the transverse paper web to measure the moisture content in the paper.

The goal is to design the inner loop controller C_2 to obtain good set point following of the steam pressure r_2 and good rejection of disturbances in the steam main supply v_2.

Design

In order to design a PID controller for the method in Section 2 a transfer function of the process P_2, in Figure 6, was needed. It was modeled by the transfer function

$$P_2(s) = K_p \frac{T_2 s + 1}{s(T_1 s + 1)},$$

where $K_p = 0.008$, $T_1 = 5.0$, and $T_2 = 30$.

A PID controller was designed for the process model $P_2(s)e^{-sL}$ with $L = 2$ seconds because of the delay caused by the sampling. The tuning parameter $M_s = 1.4$ was chosen for the controller design, which give good robustness to model uncertainties. The controller parameters were: $K = 5.21$, $T_i = 3.16$, $T_d = 1.02$. As large overshoots on the output were undesirable and the existing control system would not allow for values of the set point weight other than 1, a first order set point filter was determined according to Panagopoulos *et al.* (2000), with filter time constant

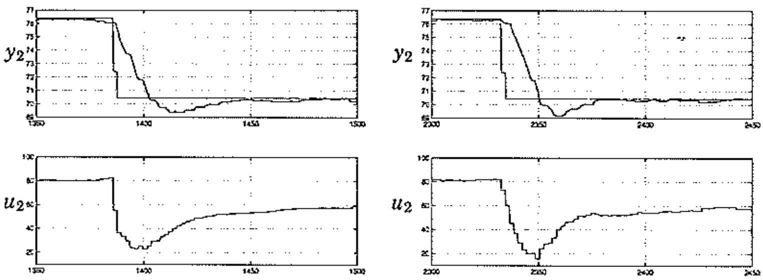


Figure 7 Validation of the PI controller (left) compared to the PID controller (right) for a set point change of the pressure control in the drying section of the paper machine.

$T_f = 5.32$. A PI controller existed already with the controller parameters $K = 5.0$, $T_i = 5.0$.

Validation

The PID controller was compared to the original PI for the pressure control of a drying section in a paper machine. Two performance validations were made: Firstly, a performance validation was made for a set point change, see Figure 7. Secondly, a performance validation was made for set point following and rejection of load disturbances see Figure 8 and 9.

In the first case the settling time t_s, defined in Åström and Hägglund (1995), is used as a performance measure for a set point change in r_2. It is the time it takes before the process output remains within $p\%$ of its steady state value, here the value $p = 2\%$ is used. The settling time t_s of the PI respectively of the PID controller becomes13.6 seconds respectively 12.8 seconds. Consequently, the output of the PID controller is faster than the one of the PI which is also seen in Figure 7. Note, how the output of the PID controller is slower in the beginning of the set point change than the one of the PI, because a set point filter is used on the former one to give acceptable overshoot/undershoot.

In the second case a performance validation was made for set point following and rejection of load disturbances. The validation of the PI controller is shown in Figure 8 and of the PID controller in Figure 9. The variance of the controller error, $e = r_2 - y_2$, was used as a performance measure. It is defined as,

$$ISE = \frac{\sum_{i=1}^{M} e(k)^2}{M},$$

where M is the number of measured data points. The ISE of the PI and the PID controller became 0.0137 and 0.0066 respectively. Consequently, the PID controller gives a better performance compared to the PI.

4.1 Conclusions

In the first case, the PID design resulted into a faster controller such that the settling time of the pressure y_2 was shorter for changes in set point r_2. The system from set point r_2 to moisture content y_1 can then be viewed as a fast process, in other words, a first order system with small time constant

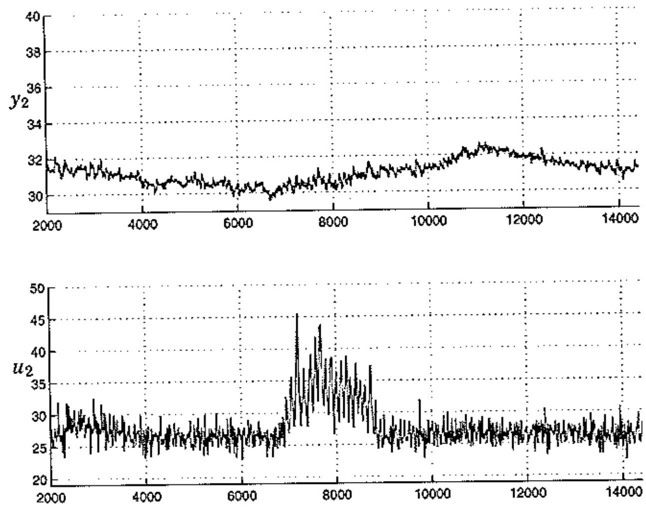


Figure 8 A performance validation for set point following and rejection of load disturbances of the PI controller for the pressure control in a drying section of a paper machine.

(about 60 s after tuning). A tighter control of the moisture content y_1 can then be used, as the rejection of the disturbance v_1 is improved.

In the second case, the PID design gave better rejection of the disturbance in steam main supply v_2, which is to be considered as the severest disturbance in trying to keep a constant moisture content y_1. If the pressure loops succeed in rejecting v_2 quickly, this implies less variations in the y_1.

5. Conclusions

In this paper a new tool is presented which gives a systematic way to optimize PID controllers. A process model formulated as a transfer function is needed in order to do a control design. In a first step, a transfer function of the process is constructed with the rapid tool for process identification in Wallén (1999) by using step response data. The simplicity of the graphical interface is of great importance.

In a second step, a PI/PID controller is designed with the methods presented in Åström *et al.* (1998), and Panagopoulos *et al.* (1999). The controller designs reflect the essence of a real control problem. They capture engineering criteria such as load disturbance rejection, robustness to model uncertainties, rejection of measurement noise, and set point tracking. The controller design methods also have one tuning parameter which gives a trade-off between robustness and performance.

Finally, the applicability of the tool has been illustrated in two industrial examples at the pulp and paper company Modo Paper, in Husum, Sweden: temperature control of a starch boiler, and steam pressure control

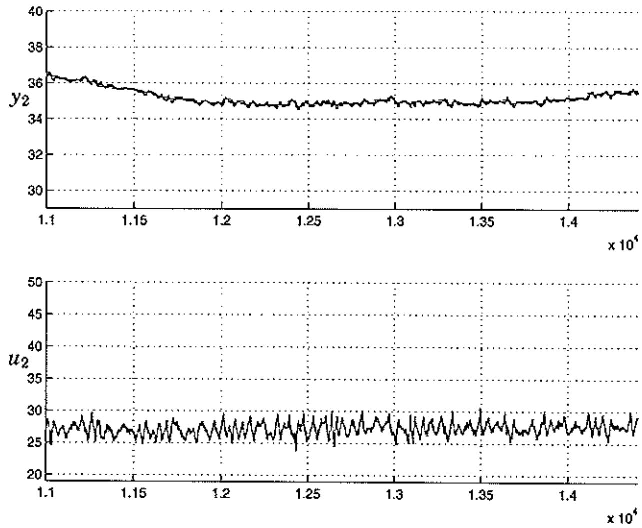


Figure 9 A performance validation for set point following and rejection of load disturbances of the PID controller for the pressure control in a drying section of a paper machine.

in a drying section of a paper machine.

References

Åström, K. J. and T. Hägglund (1995): *PID Controllers: Theory, Design, and Tuning.* Instrument Society of America, Research Triangle Park, North Carolina.

Åström, K. J., H. Panagopoulos, and T. Hägglund (1998): "Design of PI Controllers based on Non-Convex Optimization." *Automatica*, **34:5**, pp. 585–601.

Panagopoulos, H., K. J. Åström, and T. Hägglund (1999): "Design of PID Controllers based on Constrained Optimization." In *1999 American Control Conference*, pp. 3858–3862. San Diego, California.

Panagopoulos, H., K. J. Åström, and T. Hägglund (2000): "Supplement and Errata to Design of PID Controllers based on Constrained Optimization."

Wallén, A. (1999): "A Tool for Rapid System Identification." In *1999 Conference on Control Applications*, pp. 1555–1560. Kohala Coast Island of Hawaii, Hawaii.

Ziegler, J. G. and N. B. Nichols (1942): "Optimum settings for automatic controllers." *Trans. ASME*, **64**, pp. 759–768.

VI

A New Modular Approach to Active Control of Undamped Modes

H. Panagopoulos, and T. Hägglund

Department of Automatic Control
Lund Institute of Technology
Box 118
S-221 00 Lund, Sweden
Phone: +46-46 2228789, Fax: +46-46 138118
Email: hp@control.lth.se

Abstract In this paper, an active control system is designed to handle processes with undamped modes. A modular approach has been taken, where an active control system has been designed, which consists of an allpass filter and a bandpass filter. To determine the parameters of these two filters the only information needed is a few characteristics of the process frequency response.

Keywords Undamped modes. Load disturbance. Oscillations. Vibrations. Allpass filter. Bandpass filter. Design.

Introduction

Undamped systems occur in many areas, for example in mechatronics, in industrial robots, in drive systems as steel mills, in HVDC, in combustion control, in cranes, in automotive systems, and in ship movements. They are classified into two cases: those who are intentionally made to contain undamped modes and those who are not. For example, cranes are intentionally made to have undamped modes on the contrary to motors which are affected by vibrations. Furthermore, todays' trends show that future machines will be faster, lighter, have softer materials and be more flexible. This implies that future machines will be unintentionally made to contain undamped modes. Examples are industrial robots and vehicles.

The standard PID controller have many advantages, but there are cases when it does not suffice to meet the required control performance. For example, when the process contains long time delays, undamped modes, and nonlinearities. There are two ways to improve the performance of the closed loop system, by either using a modular/synthetic approach or a unimodular/analytic one. For example, in the case of processes with long time delays a modular approach is taken when a Smith predictor is inserted into the loop to improve the performance of the closed loop system. In this paper the former approach is used to improve the performance of the closed loop system for processes with undamped modes.

The control structure proposed in this paper has two degrees of freedom. A controller $C(s)$ stabilizes the plant, and an active control system F rejects harmonic disturbances. In the conventional method the controller $C(s)$ must simultaneously play both roles, leading to performance limitations.

Background

Due to physical, economical and safety constraints the oscillation and vibration control have become a prime consideration in the general manufacturing industry. For example, a ships roll motion due to the waves is a highly resonant system. It is then, necessary for ships carrying passengers or weapon platforms to control this undesirable motion. Another example is the use of cranes to move containers. In this case it is required that the payload does not undergo excessive swing, if not damage to personnel and cargo may occur. Moreover, in the automotive industry recent trends are towards smaller and lighter vehicles for better fuel economy and market expansion, see Karkosch and Svaricek (1999). At the same time, there has been a major increase in consumer awareness of long term health impacts of exposure to high noise and vibration levels.

Due to these physical, economical and safety constraints, conventional techniques such as passive sound absorptive materials, mass tuned dampers, modification of the system design, etc., by themselves would not satisfy the requirements of producing a minimum weight system with an optimal fuel economy and in those cases needed a high comfort level. Consider for example, an engine mount; in order to limit engine movement, it is desired to have a very stiff mount. However, to minimize transmission of engine vibrations into the passenger compartment, a very soft engine mount is required. These contradictory requirements have made engine mount manufactures to design passive vibration absorbers that provide an optimal compromise. But, passive noise and vibration treatments at low frequencies are physically not realizable. Consequently, active control technologies have been investigated for low frequency noise and vibration control. Consequently, the complete solution to the noise and vibration problems can be obtained by integration of passive and active systems.

The problem of active control of noise and vibrations has been a subject of much research in recent years, see Karkosch and Svaricek (1999) and references there in. The main part of the published literature makes use of adaptive signal processing and filtering techniques. According to Ohmori *et al.* (1999), the most popular techniques for harmonic disturbance rejection are (i) the feedback controller design method based on internal model principle; and (ii) the feed forward controller design method based on external model principle.

In the feedback control structure of (i) approach the requirement is

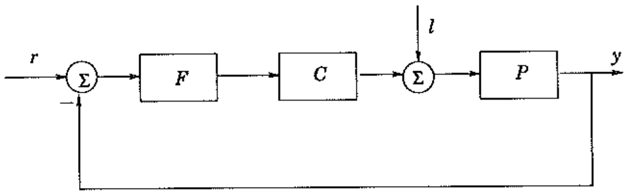


Figure 1 The use of a notch filter F to handle undesired excitation of undamped modes of the process P.

an asymptotic disturbance rejection. This can be realized by the insertion of notch filters. The advantages of this type of algorithm are that it is linear making analysis easier, and that convergence is very rapid. The disadvantages are that there are some performance limitations because the algorithm must satisfy both closed loop stability and disturbance rejection.

On the other hand, in external model controllers of (ii) approach, the disturbance model is placed outside the basic feedback loop. The disturbance model is adjusted adaptively to match the actual disturbance. The advantage of this approach is that compensation is like feed forward, then the effect on the nominal open loop gain can be made small. The disadvantages are that the analysis and implementation are somewhat more complex than for the internal model based algorithms.

Notch Filtering Limitations

A classical method to avoid unnecessary excitation of undamped modes is illustrated in Figure 1. The process P which consists of undamped modes is controlled by a controller C. To avoid unnecessary excitation of the undamped modes in P a notch filter F is introduced. Unfortunately, the notch filter F will not provide any additional damping of these modes, when the disturbance l in Figure 1 excites them.The paper will present a new method to overcome this problem.

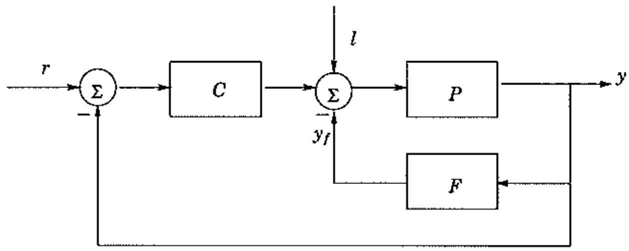


Figure 2 Control with active damping. In the inner feedback loop the active control system F rejects undamped modes of the process P. In the outer control loop the controller C gives good disturbance rejection, robustness to model uncertainties, and set point following.

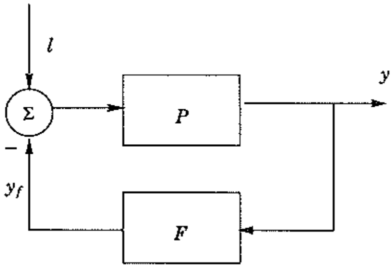


Figure 3 The reduced block diagram, describing the relation from load disturbance l to measurement output y.

The New Approach

Consider the two degree of freedom controller in Figure 2. The controller consists of an inner feedback loop where the purpose of the active control system F is to increase the damping of the plant. In the outer control loop the purpose of controller C is to achieve good disturbance rejection, robustness to model uncertainties and set point following. There are no requirements on the controller structure and design. The objective of the active control system F is to damp the oscillations in the output y, when the disturbance l excites the undamped modes of the process P. Thus, the design problem of F given in Figure 2 is reduced to the block diagram in Figure 3.

Assume that the disturbance l is a sine wave whose frequency corresponds to the one of the undamped modes of the process P. The purpose of F is then to synthesize a waveform y_f which is identical in magnitude, phase and frequency to the original signal l. This is realized if the filter F is the product of an allpass filter F_a with transfer function,

$$F_a(s) = \frac{s^2 - 2\zeta_a\omega_a s + \omega_a^2}{s^2 + 2\zeta_a\omega_a s + \omega_a^2}, \tag{1}$$

and a bandpass filter F_b with the transfer function,

$$F_b(s) = K_b \frac{s}{s^2 + 2\zeta_b\omega_b s + \omega_b^2}. \tag{2}$$

The allpass filter F_a makes it possible to obtain the right amount of phase lag of the output y at a specific frequency without affecting the gain, because, F_a gives only a phase and no gain contribution for all frequencies, compare with the left figure of Figure 4.

The bandpass filter F_b makes it possible to obtain the right amplitude of the output y at a specific frequency without affecting the phase, because, F_b gives only a gain but no phase contribution at a specific frequency, compare with the right figure of Figure 4. A good overview of different filters is found in Wie and Byun (1989).

In the next section it is shown how to determine the parameters of the allpass filter F_a: ζ_a, ω_a, and of the bandpass filter F_b: K_b, ζ_b, ω_b based on a few characteristics of the process frequency response.

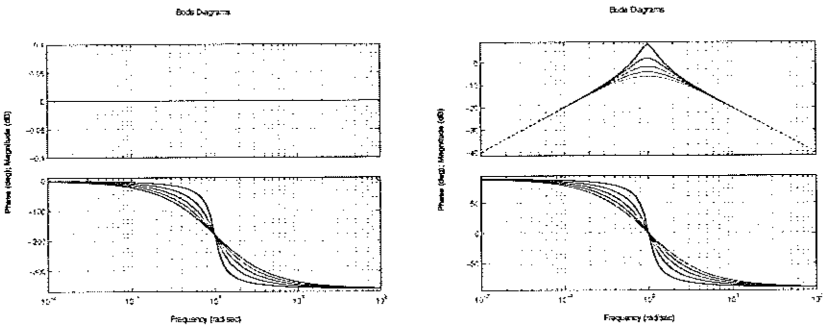


Figure 4 The Bode diagram of the allpass filter F_a (left) with $\omega_a = 1$, $\zeta_a = 0.2, 0.4, 0.6, 0.8$, and the bandpass filter F_b (right) with $\omega_b = 1$, $K_b = 1$, $\zeta_b = 0.2, 0.4, 0.6, 0.8$.

Design of the Active Control System F

For the design of the active control system F it is assumed that the undamped process P has the same structure as the flexible shaft system. This is no severe restriction, since a lot of undamped systems occurring in different areas may be modeled as a flexible shaft system. The system consists of two flywheels which are coupled with a spring, where the spring coefficient is assumed to be constant. It is driven by a torque input on the driving fly wheel, and the speed of the load flywheel is measured. The transfer function of the system is given by,

$$P(s) = \frac{\omega_p^2}{(s+\omega_f)(s^2+2\zeta_p\omega_p s+\omega_p^2)}, \tag{3}$$

where ζ_p and ω_p are the damping and the resonance frequency of the undamped modes of P.

The active control system F is the product of the allpass filter F_a in (1), and the bandpass filter F_b in (2). To determine the parameters of the two filters the following characteristics of the process frequency response $P(i\omega)$ is needed: the resonance frequency ω_p, and the amplitudes P_{max} and P_{min} defined in Figure 5.

The Allpass Filter Design: The purpose of the allpass filter F_a is to shift the phase of the output y at the resonance frequency ω_p such that,

$$\arg\{P(i\omega_p)\} + \arg\{F_a(i\omega_p)\} = -360°,$$

which is equal to

$$\arg\left\{\frac{s^2-2\zeta_a\omega_a s+\omega_a^2}{s^2+2\zeta_a\omega_a s+\omega_a^2}\right\}\Bigg|_{s=i\omega_p} = -\arg\{P(i\omega_p)\} - 360°, \tag{4}$$

where the damping coefficient ζ_a determines the rate of decrease of $\arg F_a(i\omega)$, compare with the left figure of Figure 4. It was noted from simulations that the

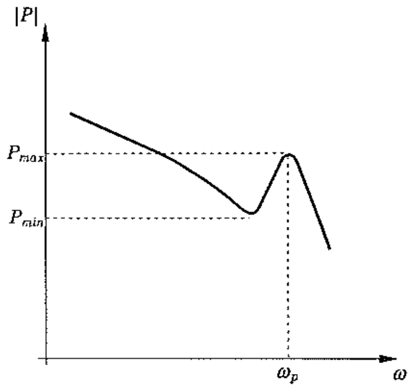


Figure 5 Characteristics of the process frequency response $P(i\omega)$ needed to design the active control system F.

choice of ζ_a is quite insensitive to performance. The choice of $\zeta_a = 0.5$ was made based on Wie and Byun (1989).

If Equation (4) is solved for ω_a, with $\zeta_a = 0.5$ the filter F_a will give the desired phase lag at ω_p.

The Bandpass Filter Design: The purpose of the bandpass filter is to change the amplitude of the output y at the frequency ω_p, such that $|P(i\omega_p)| \cdot |F_b(i\omega_p)| = 1$ at $\omega_b = \omega_p$, that is,

$$|P(i\omega_p)| \cdot \left| K_b \frac{s}{s^2 + 2\zeta_b \omega_p s + \omega_p^2} \right|_{s=i\omega_p} = 1, \tag{5}$$

where Equation (2) have been used. If Equation (5) is solved for K_b, then

$$K_b = \frac{2\zeta_b \omega_p}{|P_{max}|}, \tag{6}$$

gives the right amount of gain reduction at $\omega = \omega_p$.

In the right figure of Figure 4 it is illustrated how the damping coefficient ζ_b affects the amplitude and phase curve of $F_b(i\omega)$. In this case, the effect of ζ_b on the amplitude curve is the interesting one. A small value will give a narrow width of the notch gain at ω_p compared to larger values of ζ_b. It was noted from the simulations that best performance was obtained by using relative large values of ζ_b. The choice $\zeta_b = 0.9$ gave a desirable performance.

The Characteristics of the Resulting Design To analyze the effects of the determined active control system F, consider the transfer function

$$\frac{Y(s)}{L(s)} = \frac{P(s)}{1 + P(s)F(s)},$$

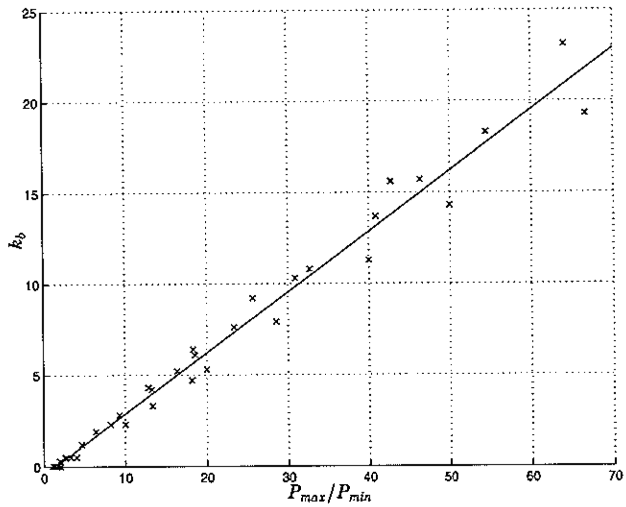


Figure 6 The gain k_b of the bandpass filter F_b plotted as a function of the characteristics P_{max}/P_{min} of the process frequency function $P(i\omega)$.

given by Figure 3. A performance measure of the design of the active filter F is its damping at the resonance frequency ω_p, that is,

$$\left|\frac{P(i\omega_p)}{1+P(i\omega_p)F(i\omega_p)}\right|. \tag{7}$$

To estimate the magnitude of Equation (7), the following results are needed: $|P(i\omega_p)| = P_{max}$ and $|F_a(i\omega_p)| = 1$. Furthermore, from Equation (1) and (6) it follows that $|F_b(i\omega_p)| = K_b/2\zeta_b = 1/P_{max}$. Then, Equation (7) gives,

$$\left|\frac{P(i\omega_p)}{1+P(i\omega_p)F(i\omega_p)}\right| \geq \frac{|P(i\omega_p)|}{1+|P(i\omega_p)|\,|F(i\omega_p)|} = \frac{P_{max}}{2}, \tag{8}$$

where the triangle inequality has been used. Consequently, the filter F will reduce the gain at the most with half the magnitude of the process P at the resonance frequency ω_p.

The performance of the designed active filter F was evaluated on a test batch for the process (3) with $\omega_f = 1$, and different values of ω_p and ζ_p. As a performance measure the integrated absolute error IAE, defined in Åström and Hägglund (1995), was computed for $e = r - y$ with $r = 0$. The calculated IAE for the different processes in the test batch showed that in some cases smaller values could be obtained. In other words, the use of K_b in Equation (6) will not give the optimal gain reduction at ω_p for all processes in the test batch. For some processes K_b should be larger for other ones it should be smaller. Consequently, the performance of the active filter F would improve if its gain K_b depends on the process characteristic

$$\frac{P_{max}}{P_{min}},$$

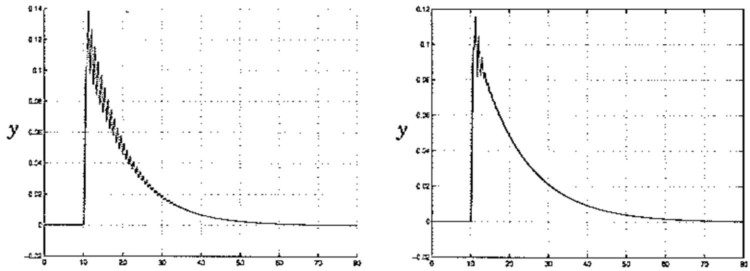


Figure 7 The time response for a short pulse on the input of the flexible shaft system system for the cases with no active control (left) and with active control (right).

given in Figure 5. Equation (6) should then be refined to,

$$K_b = k_b \frac{2\zeta_b w_b}{|P(i\omega_b)|}, \tag{9}$$

where k_b is a function of P_{max}/P_{min}. The function $k_b(P_{max}/P_{min})$ is obtained in the following way: for each process in the test batch calculate the IAE values of the closed loop system in Figure 3 for a set of k_b and choose the k_b which gives the minimum IAE. These values of k_b have been marked with a cross (x) in Figure 6 together with corresponding values of P_{max}/P_{min} . A rough approximation of $k_b(P_{max}/P_{min})$ is given by,

$$k_b = -0.46 + 0.33\frac{P_{max}}{P_{min}}, \tag{10}$$

which corresponds to the full line drawn in Figure 6.

Concluding Remarks: In this section an active control system F, has been designed. It rejects the excitation of undamped process modes in the output y, which are excited by the disturbance l in Figure 3. The active filter F is the product of the allpass filter F_a in (1), and the bandpass filter F_a in (2). The filter parameters of F_a, and F_b are determined in the following way:

ω_a: solve Equation (4).

ζ_a: ζ_a=0.5

ω_b: $\omega_b = \omega_p$

ζ_b: ζ_b=0.9

k_b: $k_b = -0.46 + 0.33P_{max}/P_{min}$

Example

The proposed active control system has been applied to an example found in the literature, the flexible shaft system in Farrugio (1998).

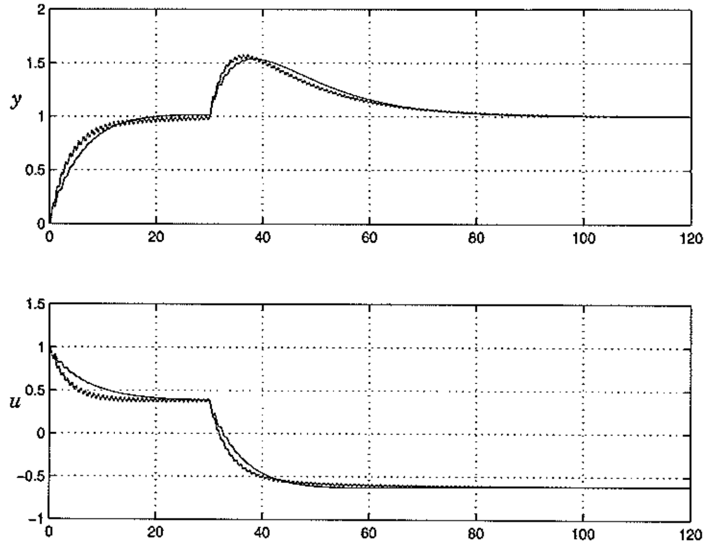


Figure 8 The time responses of the closed loop system when controlling the flexible shaft system system with a PI controller for the cases with no active control (undamped response) and with active control (well damped response).

The Flexible Shaft System

The flexible shaft system from the paper Farrugio (1998) consists of two flywheels which are coupled by a spring and can be approximated by the transfer function,

$$P(s) = \frac{\kappa}{(s+\omega_f)(s^2+2\zeta_p\omega_p s+\omega_p^2)}.$$

The proposed active control system has been tested on the nominal model in Farrugio (1998), with $\kappa = 15.21$, $\zeta_p = 0.02$, $\omega_p = 7.66$, $\omega_f = 0.10$. To determine the active filter F the parameters needed are: $\omega_b = 7.66, P_{max} = 0.85$, $P_{min} = 0.088$, $\omega_a = 7.64$, and $k_b = 2.75$. In Figure 7 the time response for a short pulse on the input of the flexible shaft system system is shown for the cases with no active control in the left figure and with active control in the right figure. Consequently, the active control system gives an improvement in performance.

Furthermore, the active control system has been tested in closed loop where a PI controller with controller parameters $K = 1.0$ and $T_i = 12.5$ is used. In Figure 8 the time response of the closed loop system is shown for the cases with no active control and with active control. The well damped response is obtained with the active controller which gives an improved performance of the closed loop system.

Conclusions

In this paper, an active control system is designed to handle processes with undamped modes. The objective of the active controller is to damp the oscillations in the process output, when a disturbance excites the undamped modes of the process. A modular approach has been taken, where the active filter consists of an allpass filter and a bandpass filter. The parameters of the allpass and bandpass filter are determined in a systematic way, where the only information needed is a few characteristics of the process frequency response. There are no requirements on the controller structure and design. The active filter has been applied to a variety of typical undamped systems, and it works well.

References

Åström, K. J. and T. Hägglund (1995): *PID Controllers: Theory, Design, and Tuning.* Instrument Society of America, Research Triangle Park, North Carolina.

Farrugio, D. (1998): "Control of a Flexible Shaft using H_∞." In *Theory and Practice of Control Systems*, 6. IEEE, IEEE, Mediterranean Conference, Alghero, Italy.

Karkosch, H.-J. and F. Svaricek (1999): "Automotive Applications of Active Vibration Control." In *Proceedings of the 1999 European Control Conference.* Karlsruhe, Germany.

Ohmori, H., H. Miyamoto, and A. Sano (1999): "Sinusoidal Disturbance Rejection by Plug in Adaptive Controller." In *IFAC'99, 14th World Congress of IFAC*, vol. I, pp. 289–294. Beijing, P.R. China.

Wie, B. and K.-W. Byun (1989): "New generalized structural filtering concept for active vibration control synthesis." *J. Guidance,* **12:2,** pp. 147–154.

LUND INSTITUTE OF TECHNOLOGY

Lund University

Department of Automatic Control

ISSN 0280–5316
ISRN LUTFD2/TFRT--1059--SE