

Principles of Process Control

Third Edition

About the Author



Dipak Patranabis is presently Professor Emeritus, Department of Applied Electronics and Instrumentation Engineering, Heritage Institute of Technology, Kolkata. Prior to that, he was Professor, Head, and then Emeritus Fellow at Jadavpur University, Kolkata, in the Department of Instrumentation and Electronics Engineering. After he completed M. Sc (Tech.) from Calcutta University, Prof. Patranabis had a brief stint in teaching Physics. Then he joined Damodar Valley Corporation as an Electrical

Engineer. Subsequently he was in Guest Keen Williams Limited, Howrah, as an Instrument Engineer for over four years only to return to teaching and research at Jadavpur University, taking charge of the newly formed department of Instrumentation and Electronics Engineering. He obtained a Ph.D. from the University of Calcutta at the early period of his teaching and research career. Dr Patranabis has authored over 150 research papers and six books in Instrumentation and Electronics. He has guided many Ph.D. scholars and is still active in research and teaching. He was the President of IIST for two years, edited *Journal of IIST* for six years, and was honorary editor of the *Journal of IETE* for two years for the Circuits and Systems group. He was also the summary editor for Springer-Verlag for over 7 years and is a reviewer of research papers of many international- and national-level research journals. He has also received the Lifetime Achievement Award from the International Society for Automation (ISA), USA.

Principles of Process Control

Third Edition

D. PATRANABIS

Professor Emeritus

Department of Applied Electronics and Instrumentation Engineering

Heritage Institute of Technology

Kolkata



Tata McGraw Hill Education Private Limited

NEW DELHI

McGraw-Hill Offices

New Delhi New York St Louis San Francisco Auckland Bogotá Caracas
Kuala Lumpur Lisbon London Madrid Mexico City Milan Montreal
San Juan Santiago Singapore Sydney Tokyo Toronto



Tata McGraw-Hill

Published by Tata McGraw Hill Education Private Limited,
7 West Patel Nagar, New Delhi 110 008

Principles of Process Control, 3e

Copyright © 2012, by Tata McGraw Hill Education Private Limited

No part of this publication may be reproduced or distributed in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise or stored in a database or retrieval system without the prior written permission of the publishers. The program listings (if any) may be entered, stored and executed in a computer system, but they may not be reproduced for publication.

This edition can be exported from India only by the publishers,
Tata McGraw Hill Education Private Limited.

ISBN (13 digits): 978-0-07-133325-2

ISBN (10 digits): 0-07-133325-8

Vice President and Managing Director—MHE: Asia Pacific Region: *Ajay Shukla*
Head—Higher Education Publishing and Marketing: *Vibha Mahajan*

Publishing Manager—SEM & Tech. Ed.: *Shalini Jha*

Editorial Researcher: *Harsha Singh*

Executive—Editorial Services: *Sohini Mukherjee*

Sr Production Manager: *Satinder S Baveja*

Production Executive: *Anuj Kr Shriwastava*

Marketing Manager—Higher Education: *Vijay Sarathi*

General Manager—Production: *Rajender P Ghansela*

Production Manager: *Reji Kumar*

Information contained in this work has been obtained by Tata McGraw-Hill, from sources believed to be reliable. However, neither Tata McGraw-Hill nor its authors guarantee the accuracy or completeness of any information published herein, and neither Tata McGraw-Hill nor its authors shall be responsible for any errors, omissions, or damages arising out of use of this information. This work is published with the understanding that Tata McGraw-Hill and its authors are supplying information but are not attempting to render engineering or other professional services. If such services are required, the assistance of an appropriate professional should be sought.

Typeset at Text-o-Graphics, B-1/56, Arawali Apartment, Sector 34, Noida 201301 and printed at Pushp Print Services, B-39/12 A, Gali No. 1, Arjun Mohalla, Moujpur, Delhi - 110 053

Cover Printer Name: SDR Printers

*To
Kakaji*

Preface

The third edition of the book is being released fifteen years after the release of the second edition. Over the last decade and a half, there has been a large-scale change in automation as a discipline, of which process control is a natural part. Development of semiconductor technology leading to production of newer micro- and nano-level sensors and advancement in computer and communication hardware and software have commingled, connoting individual core subjects like Process Control, to undergo meaningful updation. However, the fundamentals have not altered and the changes proposed are primarily in the implementation logistics.

Aim

Copies of the second edition were sent out to experts and academicians of repute across the country and a thorough research in the syllabi of different institutes was carried out by the publishers. Depending on the outcome of the above, some chapters have been revised giving coverage to new topics. Suggestions from experts also include addition of mathematical preliminaries like Laplace transforms and Z-transforms. Although these are requisites for this course, the topics have been introduced to brush up the knowledge of the readers on the subject. These make the text comprehensive for students and researchers alike.

Target Readers

The updated work is expected to enjoy the support of UG and PG students of engineering and science in the disciplines of Instrumentation and Electronics, and Chemical Engineering. The coverage is good enough to meet the needs of engineering disciplines of allied fields. Practicing technocrats and teachers would also be benefitted by the book.

Roadmap for Engineering and other Disciplines

Students of allied disciplines would find the entire book useful to them when they become judicious to choose the topics of their interest. For Chemical Engineering, Metallurgical Engineering and Instrumentation and Electronics Engineering, the entire book is of interest. For Electrical and Electronics Engineering students, some parts of chapters 3, 8 and 11 may be left out, while for students of Mechanical Engineering, Chapter 8 is very much a part of study and parts of chapters 3 and 11 may be omitted. Courses in the science faculty such as Instrumentation Science, Biomedical Science, Food Processing Science and such others would find the entire book very useful.

Salient Features

The presentation format and chapter names have not been changed. It was already made in a manner that provides in-depth theoretical/analytical methods to help students build up the concepts and also practical-oriented materials for further clarification. These make the book a comprehensive text. The new salient features are the following:

- Crisp and complete coverage of Process Control.
- Addition of Laplace transforms and Z-transform preliminaries brushes up the requisites for analysis.
- Modelling of process has been given extensive coverage with the methods followed in practice and the present-day approach with computer support. Discrete modeling has also been discussed.
- Compensators have been discussed in a more rational way with examples supporting the discussion.
- Analysis of state variable approaches of controllability has been included.
- Responses of PID control action in a process have been revised and elaborated in graphs.
- Discrete control algorithms have been presented with flowcharts.
- OCS and OPC have been given coverage.
- PLC has been given wider coverage with elements of programming.
- Bus technology in process control has been covered with support of protocols to be considered.
- Nonlinear process and its control has been included as also batch process control.
- MATLAB solutions of examples have been included.
- Inclusion of newer trends in process automation (SCADA, OCS and DCS Vendors)
- Detailed coverage of Process Modeling and Nyquist Plot.

Organization

There has not been any chapterwise organizational change but some chapters have been updated in the process of revision. **Chapter 1** now includes mathematical requisites, and in **Chapter 2** process modeling has been made comprehensive adding new methods and discrete modeling methods. **Chapter 3** remains unchanged. Analysis of state variable approach in controllability and compensator techniques have been updated in **Chapter 4**. In **Chapter 5**, new practical schemes of on-off control, responses with PID actions, PLC's and programming elements have been added. **Chapter 6** has also been given an additional complex control scheme, while in **Chapter 7** some analysis and symbolic presentation of control valves have been given. **Chapter 8** remains as it is. **Chapter 9** has been given a facelift specially in the latter part where bus technology and associated material such as OCS and OPC's have been included. **Chapter 10** remains as it is, while **Chapter 11** has been expanded including batch process and nonlinear process and their control strategies. No changes have been made in the appendices.

Acknowledgements

I want to thank the reviewers, who have sent in suggestions to make the text more student friendly, and the coordinators of TMH for their cordial cooperation. I also thank Ms. Sampa Maity for preparing the additional pages with equations which were incorporated in the book. The names of the reviewers are given below.

T Panda

Indian Institute of Technology (IIT) Madras, Chennai

Tanmay Basak

Indian Institute of Technology (IIT) Madras, Chennai

Sushil Kumar

Birla Institute of Technology and Science (BITS) Pilani, Rajasthan

Gopinath Halder

National Institute of Technology (NIT) Durgapur, Paschimbanga

Y Pydi Setty

National Institute of Technology (NIT) Warangal, Andhra Pradesh

Somak Jyoti Sahu

Haldia Institute of Technology, Haldia, Paschimbanga

R P Ugwekar

*Priyadarshini Institute of Engineering and Technology,
Nagpur, Maharashtra*

Mausumi Mukhopadhyay

Sardar Vallabh Bhai National Institute of Technology (SVNIT)

Surat, Gujarat

Feedback

Since there is always scope of improvement that can be made with suggestions received from the readers, I would request them to provide feedback to my email id: dcpatranabis@yahoo.co.in.

D. PATRANABIS

Publisher's Note

Remember to write to us! Send in your comments, and suggestions at tmh.elefeedback@gmail.com.

List of Symbols

A	Actuator (usually used in diagrams)
C	Controller (usually used in diagrams)
c	Controlled output, usually function of s
c_o	Offset
D	Pipe diameter
E, e	Error function, error
F	Feedpoint
G, G_i	Transfer function of blocks, usually functions of s which is not often mentioned, suffix i stands for the different blocks; c : controller, p : process a : actuator, m : measurement; s : process (second part), etc.
H	Transfer functions of blocks usually in the feedback path
h	Head, liquid level
K, k	Gain parameter
K_c	Proportional action gain
k_c	Controller gain
I	Valve lift
M	Measured variable, Measurement system (usually used in diagrams)
m	Manipulated variable
PB	Proportional band
p	Pressure
q	Flowrate, variable
R	Ratio set, Ratio controller (usually in diagrams)
r	Reference or set point
S	Setter (usually in diagrams)
s	Laplace variable

xii *List of Symbols*

S_c	Steam consumption (also in diagrams)
$T(s)$	Transfer function
T_i	$i = 1, 2$ Temperature, etc. Time, temperature
T_r, T_R	Reset time, Integral time
T_d, T_D	Rate time, Derivative time
u	Upset, load disturbances
V, v	Volume
w	Mass rate of flow
X	Vectors, Matrices
x	Input quantity
y	Output quantity
α	Area
δ	Decay ratio, deviation
ε	Efficiency
χ	Concentration
λ	Weighting factor, Ratio
ρ	Subsidence ratio, density, weighting factor
σ	Real part of s (Laplace variable)
τ_d	Dead time
τ_i, τ_s	Process time constants
τ_m	Measurement system time constants
ω	Frequency (circular)
ω_n	Natural frequency of oscillation
ω_o, ω_{er}	Critical frequency
ζ	Damping ratio

Contents

<i>Preface</i>	<i>vii</i>
<i>List of Symbols</i>	<i>xi</i>
1. Basic Considerations	1
1.1 Introduction	1
1.2 Notes on Processes	2
1.3 Control-Loop Study	3
1.4 Sources of Disturbances	9
1.5 Control Actions	11
1.6 Z-Transforms	13
1.7 General Comments	17
Review Questions	18
2. Processes: Transfer Functions and Modelling	20
2.1 Introduction	20
2.2 Some Typical Simple Processes and their Transfer Functions from Analysis	21
2.3 Limitations on Process Equation Formulations	36
2.4 Process Modelling	37
2.5 Process Modelling Via Experimental Tests	41
2.6 Discrete Modelling	51
2.7 Scale Modelling Technique	55
2.8 Process Model from Frequency Response Studies	58
2.9 Further Comments on Parameter Evaluation Process Testing	61
2.10 Conclusion	62
Review Questions	63

3.	Block Diagrams: Transient Response and Transfer Functions	65
3.1	Block Diagram Representation	65
3.2	Step, Frequency and Impulse Response of Systems	70
3.3	Controlled Process Block Diagrams and Transfer Functions	79
3.4	System Analysis and Studies of System Response	89
3.5	Generalization with Load Changes at Arbitrary Points	96
	<i>Review Questions</i>	98
4.	Controllability and Stability	101
4.1	Introduction	101
4.2	Controllability	101
4.3	Self-Regulation	117
4.4	Stability Studies	119
4.5	Compensators	141
	<i>Review Questions</i>	144
5.	Basic Control Schemes and Controllers	149
5.1	Introduction	149
5.2	On-Off Control	149
5.3	Time Proportional Control	154
5.4	Typical PID Controller Characteristics and Related Terminology	161
5.5	Comparison of Control Actions: PID	170
5.6	Controller Tuning or Controller Parameter Adjustment	173
5.7	Pneumatic Controllers	181
5.8	Electronic Controllers	189
5.9	Hydraulic Controllers	200
5.10	Programme Controllers	201
5.11	Programmable Logic Controller	206
	<i>Review Questions</i>	223
6.	Complex Control Schemes	229
6.1	Introduction	229
6.2	Ratio Control Systems	229
6.3	Split Range Control	233
6.4	Cascade Control	233
6.5	Feedforward Control	242
6.6	Selector Control	248
6.7	Inverse Derivative Control	252
6.8	Antireset Control	253

6.9	Multivariable Control Systems	255
	<i>Review Questions</i>	263
7.	Final Control Elements	266
7.1	Introduction	266
7.2	The Pneumatic Actuator	266
7.3	Electrical Actuators	297
	<i>Review Questions</i>	303
8.	Connecting Elements and Common Control Loops	305
8.1	Introduction	305
8.2	RLC Elements	306
8.3	Flow Control	313
8.4	Pressure Control	322
8.5	Level Control	326
8.6	Temperature Control	330
	<i>Review Questions</i>	335
9.	Computer Control of Processes	336
9.1	Introduction	336
9.2	Control Computers	337
9.3	Progress in Computer Control in Process Industries	345
9.4	Control on Level Basis	348
9.5	Algorithms for Digital Control	353
9.6	Digital Control Via Z-Transform Technique	360
9.7	Distributed Control Systems	372
9.8	The Newer Trends in Process Automation	376
9.9	General Comments	391
	<i>Review Questions</i>	394
10.	Adaptive Control Systems	396
10.1	Introduction	396
10.2	Standard Approaches	397
10.3	Self-Adaptive Systems	402
10.4	Predictive Approach	407
10.5	Self-Tuning Control	409
	<i>Review Questions</i>	416
11.	Process Control Systems	417
11.1	Introduction	417
11.2	Boiler Control	418

11.3	Steel Plant Instrumentation/Control System	431
11.4	Control in Paper Industry	443
11.5	Distillation Column	453
11.6	Belt Conveyor Control	461
11.7	pH Control	464
11.8	Batch Process Control	473
	<i>Review Questions</i>	477
 <i>Appendix I</i>		 478
<i>Appendix II</i>		481
<i>Appendix III</i>		482
<i>Bibliography</i>		484
<i>Index</i>		492

Basic Considerations

1.1 INTRODUCTION

During the last few decades the science and technique of automation have evolved considerably to keep pace with rapid industrial growth. Earlier automation was applied in process industries somewhat arbitrarily. Not much effort was there to make real time analysis of the processes to ensure the requisite control. Process engineers, on the basis of their experience, evolved certain rules which guided the design of the control part of the process control systems. However, during the last fifty years, control systems in processes have been gradually evolved on an analytical footing and today the range in the control equipment for any kind of process is commendable.

A process control system basically consists of two parts: (i) the process and (ii) the control equipment. The process is 'given' to a process control engineer on the basis of which he has to design, choose, make a layout, etc., of the control equipment. The control equipment broadly consists of (i) measurement system, (ii) comparator, (iii) controller, and (iv) actuator. The actuator is driven to provide the process with the resources (or the raw materials) at a rate determined by the 'control action' set by the controller in response to a comparison function called the error (Fig. 1.1).

This function is actually a deviation of the function of the process from the desired one and is obtained by measuring certain variables and comparing these with a fixed, predetermined set. This preliminary description of the 'process control loop' is to a certain extent oversimplified. Nevertheless, this is the basic principle. The process and the control equipment are interconnected by what are known as 'lines'. These lines are of different kinds: (i) the ones through which energy relating to the process product

2 Principles of Process Control

flows are called process lines, (ii) the ones through which measurement and control signals flow are known as impulse lines, and (iii) the ones through which power to the control equipment and process gadgets flows are known as power lines.

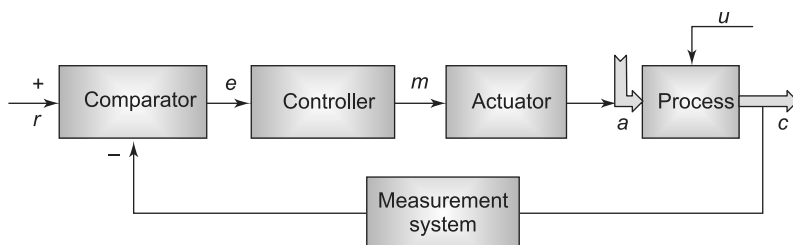


Fig. 1.1 Block diagram of the basic control loop

Control systems are basically of two different types: (i) the set-point or the reference follower also called the position control systems or servo-systems, and (ii) the regulatory systems. The generalized name given to the regulatory systems is the process control systems. These systems operate on a fixed reference (set-point), but the process conditions are such that disturbances or upsets are present which tend to deviate the operation of the process unless taken care of. In general, the processes have large capacity elements which make them “integrating types”, so that disturbances are dampened in the process itself. The remainder of the disturbances can be tackled by the control equipment.

1.2 NOTES ON PROCESSES

In general, while talking of processes, we necessarily mean processes of any kind and these are often the most complex in the system of process control. In industrial plants, where control of machinery is more important with variable references, there is a conglomeration of second-order or higher-order elements, whereas in the processes of chemical and associated plants, first-order elements with distributed parameters and interaction are dominant. Mathematical modelling of processes, which is often made for analysis purposes, will also be different in these two types. Chapter 2 has been devoted to the present trends in process modelling with a few typical examples. The other important and disturbing factor in the common processes is transportation (velocity) lag. It will subsequently be seen that this considerably affects the controllability of the processes.

When one pauses to think of a process, one really gets lost with its complexity in large time constants (both in number and value), nonlinearity,

functional delays, etc. Close control of such processes with insufficient process data (particularly dynamic) is not possible. However, reasonably good control can always be effected with approximate modelling. In fact, it often is not even advisable to go through the tedious processes of accurate modelling and consequent design of sophisticated control systems due to the prohibitive initial investment and also the further expenses involved in running the system.

Despite the best efforts of process control or instrumentation engineers, system design is to a certain extent limited owing to inadequate knowledge of the process. When process modelling is adequate, analysis for appropriate control equipment selection is also better. Control of processes is primarily required for (i) product quality, and (ii) productivity. An improvement in either one or both of these would boost up process economy. However, over-rating in quality may also not be always desirable.

1.3 CONTROL-LOOP STUDY

The process control considered here is the closed-loop control or the feedback control. For the desired product quality and productivity this loop has to be maintained and operated under suitable conditions. This dynamic process data obtained via modelling or testing should initially be checked to ensure that appropriate control action can be found for this purpose. Also, when the loop is completed with the desired controller and other control equipment, the operation should be checked with regard to transient disturbances. For control quality, that is, for successful and quick completion of the operation, the loop is tuned with reference to a fixed set-point and minor upsets are not given any consideration during the design stage. This means that the system could become unstable at some operating conditions. This is definitely not acceptable because an unstable control system is useless.

In short, the initial studies that are to be made in a process and a loop are: (i) controllability of the process, (ii) system stability, and (iii) control quality. These studies are dealt with in Chapter 4.

Study of the process behaviour with the loop closed as shown in Fig. 1.1 in the block diagram is best done by analysis of the control loop. For simplicity's sake, the control engineer often simplifies the loop structure with only one block in the forward path and one block in the feedback path. Also, the analysis is done by converting the system equation, which basically is a differential equation, into an algebraic one using the Laplace transform. This transform is universally used for linear time invariant (LTI) system to represent the system by a simple equation in terms of the Laplace operator.

1.3.1 The Laplace Transform

A transform is a mathematical tool by which a problem is shifted to a different mathematical domain for reducing the complexity of the problem. The solution is made in the new domain and by inverse transform, the solution is brought back to the domain of the original problem.

Laplace transform is an integral transform, and it is a transform of a function. For a function $f(t)$ in the time domain, its Laplace transform is denoted by $F(s)$ and is defined as

$$F(s) = L[f(t)] = \int_0^{\infty} f(t)e^{-st} dt \quad t > 0 \quad (1.1)$$

where L means ‘the Laplace transform of’. The variable s is the new variable and domain. This variable s is a complex variable having a real and an imaginary part. For the transformation to be meaningful, it is necessary that the real part of s be greater than α , if $f(t)$ is of the form $e^{\alpha t}$. However, s is not required to be evaluated for testing the convergence. Usually, a form $e^{\alpha t}$ is hardly encountered in control systems and hence the convergence condition is easily met. The systems considered are generally causal for which $t > 0$ and $f(t)|_{t \geq 0}$ are usually the functions to be considered. The conditions existing for $f(t)|_{t < 0}$ will be taken as initial conditions.

Laplace transforms are linear in characteristic implying that

$$L[k \cdot f(t)] = kL[f(t)] \quad (1.2)$$

and

$$L[f_1(t) + f_2(t)] = L[f_1(t)] + L[f_2(t)] \quad (1.3)$$

The definition made above would be used to find Laplace transforms of a few simple functions.

Case 1. Step Function $u(t)$

$$\begin{aligned} f(t) &= 0 & t < 0 \\ &= u & t > 0 \end{aligned} \quad (1.4)$$

Using Eq. (1.1),

$$\begin{aligned} F(s) &= \int_0^{\infty} e^{-st} \cdot u \cdot dt = -\frac{u}{s} e^{-st} \Big|_0^{\infty} = \frac{u}{s} \\ L[f(t)] &= L[u(t)] = \frac{u}{s} \end{aligned} \quad (1.5)$$

Case 2. Exponential Function

$$f(t) = e^{at} \quad t > 0$$

In this case a may be complex, real or imaginary.

Using Eq. (1.1),

$$\begin{aligned} F(s) = L[f(t)] &= \int_0^{\infty} e^{at} e^{-st} dt = \int_0^{\infty} e^{-(s-a)t} dt \\ &= \frac{e^{-(s-a)t}}{a-s} \bigg|_0^{\infty} = \frac{1}{s-a} \end{aligned} \quad (1.6)$$

If s and a are both real, $F(s)$ converges if $s > a$.

Case 3. Trigonometric Function; Taking a *cosine* Function

$$f(t) = \cos \alpha t \quad t > 0$$

$$\begin{aligned} F(s) &= \left[\int_0^{\infty} f(t) e^{-st} dt \right] = \int_0^{\infty} \cos \alpha t \cdot e^{-st} dt \\ &= \frac{1}{2} \int_0^{\infty} e^{-st} (e^{j\alpha t} + e^{-j\alpha t}) dt \\ &= \frac{1}{2} \left[\frac{e^{(-s+j\alpha)t}}{-s+j\alpha} + \frac{e^{(-s-j\alpha)t}}{-s-j\alpha} \right]_0^{\infty} \\ &= \frac{1}{2} \frac{s}{s^2 + \alpha^2} \end{aligned} \quad (1.7)$$

Case 4. Transform of Derivatives

Differential equations are now extensively solved using Laplace transform and hence besides transforms of functions, transforms of derivatives are also important. Thus, for first derivatives of a function $f(t)$, Laplace transform is

$$L\left[\frac{df(t)}{dt}\right] = \int_0^{\infty} \left[\frac{df(t)}{dt} e^{-st}\right] dt \quad (1.8)$$

It is assumed that $f(t)$ as well as $f'(t) = df(t)/dt$ are piecewise regular and of exponential order so that Laplace transform of the function exists. We now let $u = e^{-st}$ and $dv = f'(t)dt$ and then integrate by parts.

Hence,

$$\begin{aligned} L\left[\frac{df(t)}{dt}\right] &= e^{-st}f(t)\Big|_0^\infty + s\int_0^\infty f(t)e^{-st}dt \\ &= -f(0) + sF(s) = sF(s) - f(0) \end{aligned} \quad (1.9)$$

$f(0)$ being the value of $f(t)$ at $t = 0$

Similarly, the transformation of the second derivative is obtained.

$$\begin{aligned} L[f''(t)] &= \int_0^\infty f''(t)e^{-st}dt \\ &= e^{-st}f'(t)\Big|_0^\infty + s\int_0^\infty f'(t)e^{-st}dt \\ &= -f'(0) - sf(0) + s^2F(s) \\ &= s^2F(s) - sf(0) - f'(0) \end{aligned} \quad (1.10)$$

Continuing in this manner, the Laplace transform of higher derivatives can be obtained. For the n^{th} derivative

$$L[f^n(t)] = s^n F(s) - s^{n-1}f'(0) - \dots - f^{n-1}(0) \quad (1.11)$$

In Eqs (1.9) to (1.11), except the first terms $sF(s)$, $s^2F(s)$, $s^nF(s)$, all others are denoted as due to initial conditions.

1.3.2 Inverse Laplace Transform

The inverse transform of $F(s)$ brings back the time functions $f(t)$.

This is given as

$$L^{-1}F(s) = f(t) \quad (1.12)$$

Inverse Laplace transform of $F(s)$ is given by

$$f(t) = \frac{1}{2\pi j} \int_{c-j\infty}^{c+j\infty} F(s)e^{st}ds \quad t > 0 \quad (1.13)$$

Direct evaluation of $f(t)$ is difficult, but if $F(s)$ approaches zero as $s \rightarrow \infty$, i.e. $F(s) \rightarrow K/s^n$, $n \geq 1$, the path of integration is closed along a semicircle of large radius in the left half s -plane. Consider Fig. 1.2 (a) and (b). If radius R of the semicircle (Fig 1.2 (b)) is so large that the integrand is negligible around this contour then the two paths in Fig 1.2 (a) and (b) are equivalent. The integral (1.13) can then be evaluated using the theory

of residues which states that the integral of a function around a closed contour in the complex plane (s -plane) is $2\pi j$ times the sum of the residues in the poles within the contour. If r_v 's are the residues of $F(s)e^{st}$ of the poles in the left half plane

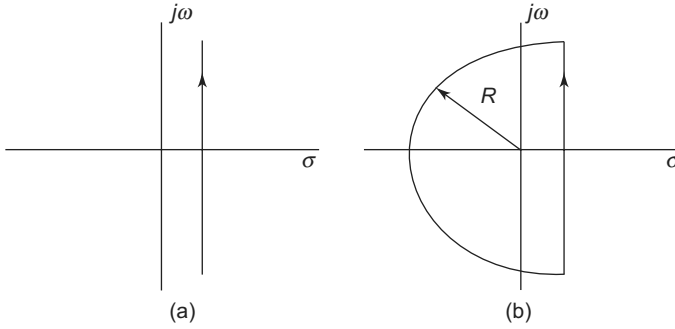


Fig. 1.2 (a) Path of integration for inverse Laplace transform (b) Equivalent path

$$f(t) = 2\pi j \sum_v r_v \quad (1.14)$$

Considering a function $F(s) = 1/[(s + 2)(s + 3)]$, the integrand of Eq. (1.13) is $F(s)e^{st}$. $F(s)$ has two simple poles at -2 and -3 and the residues are (by inspection)

$$K_{-2} = \frac{e^{-2t}}{2} \quad \text{and} \quad K_{-3} = \frac{e^{-3t}}{2}$$

Hence,

$$f(t) = \frac{1}{2}[e^{-2t} + e^{-3t}] \quad (1.15)$$

However, it is more convenient to refer to the Laplace transform table as the $F(s)$ is derived from the $f(t)$ and the inverse is then evident from the deduction. Tables are prepared assuming all initial conditions zero. It is available in Appendix 1 of the book. It would, thus, be seen from Eq. (1.6) and (1.7),

$$L^{-1}\left[\frac{1}{s-a}\right] = e^{at}$$

and
$$L^{-1}\left[\frac{s}{s^2 + \alpha^2}\right] = \cos \alpha t$$

1.3.3 Solution of Differential Equations

As in the text that follows in this book differential equations would be solved often, an illustrative example at this stage would be in order. Taking the equation

$$\frac{d^2 y}{dt^2} + Ky(t) = 0 \quad (1.16)$$

with initial conditions $y(0) = \alpha$ and $y'(0) = \beta$, with α and β constants, the Laplace transformation would be

$$s^2 y(s) - sy(0) - y'(0) + Ky(s) = 0 \quad (1.17)$$

which with given initial conditions yields

$$s^2 y(s) + Ky(s) - \alpha s - \beta = 0 \quad (1.18)$$

or,

$$y(s) = \frac{\alpha s}{s^2 + (\sqrt{K})^2} + \frac{\beta}{s^2 + (\sqrt{K})^2} \quad (1.19)$$

From the table, one easily writes $y(t)$ as

$$y(t) = \alpha \cos \sqrt{K}t + \frac{\beta}{\sqrt{K}} \sin \sqrt{K}t \quad (1.20)$$

1.3.4 Transfer Function

In this book, transfer function is frequently used which is a function of the Laplace operator s which in effect, is the transformation of d/dt with zero initial conditions. If a system is represented by a block and the output–input relation is represented by a differential equation like

$$a \frac{d^2 y(t)}{dt^2} + b \frac{dy(t)}{dt} + cy(t) = x(t) \quad (1.21)$$

where a , b , c are constants, $x(t)$ input and $y(t)$ output, then the Laplace transform of the output divided by the Laplace transform of the input, with all initial conditions zero, is the transfer function of the block. From Eq. (1.21), the transfer function $T(s)$ is

$$T(s) = \frac{y(s)}{x(s)} = \frac{1}{as^2 + bs + c} \quad (1.22)$$

Finding the roots of $as^2 + bs + c = 0$ gives

$$s = \frac{-b \pm \sqrt{b^2 - 4ac}}{2}$$

Hence,

$$\begin{aligned}
 T(s) &= \frac{1}{\left\{ s + \left(\frac{b + \sqrt{b^2 - 4ac}}{2} \right) \right\} \left\{ s + \left(\frac{b - \sqrt{b^2 - 4ac}}{2} \right) \right\}} \\
 &= \frac{1}{\sqrt{b^2 - 4ac}} \left[-\frac{1}{s + \frac{b + \sqrt{b^2 - 4ac}}{2}} + \frac{1}{s + \frac{b - \sqrt{b^2 - 4ac}}{2}} \right]
 \end{aligned}$$

from which the time function $T(t)$ can be easily obtained. However, $x(t)$ is a specific input and to find $y(t)$, one starts with

$$y(s) = \frac{x(s)}{as^2 + bs + c}$$

and for specific $x(t)$, $x(s)$ is found and used in the above equation and then solved for $y(t)$. The partial fraction method can give the right-hand side of the equation in three first-order terms and the inverse transform can be chosen for each such term easily from the table.

Example 1

Find the inverse transform of

$$F(s) = \frac{1}{s(s+2)}$$

Solution Writing right-hand side

$$\frac{A}{s} + \frac{B}{s+2} = \frac{A(s+2) + Bs}{s(s+2)}$$

But one has $As + 2A + Bs = 1$

Hence, $A = -B$ and $A = 1/2$

$$\therefore F(s) = \frac{1}{2s} - \frac{1}{2(s+2)}$$

$$\text{Hence, } f(t) = \frac{1}{2} - \frac{1}{2}e^{-2t}$$

1.4 SOURCES OF DISTURBANCES

Disturbances occupy an important place in process control studies. As these systems are regulatory in nature both load-side and supply-side changes

have to be considered. As has been pointed out, the process itself acts as a regulator. But a knowledge of the types and sources of disturbances is of prime importance for a given regulatory action to be made suitable. Sources can rarely be arbitrarily mentioned. They are process-dependent as also line-dependent. Some typical examples in relation to typical processes are mentioned here.

- (i) Disturbance due to an increase in the demand of the process product: This is a general case and may be seen in a heat exchanger in the form of a change in water rate, in a rerolling furnace in the form of a change in material flow-in rate, and so on.
- (ii) Disturbance due to a change in the fuel efficiency when a fresh stock is connected in furnace processes: For steam-heated processes, similar disturbances may occur because of a change in the steam supply pressure or dryness factor either due to a steam demand elsewhere or due to the steam line being exposed to changed atmospheric conditions.
- (iii) In heat exchangers and boilers, etc. scales build up on plant walls despite best efforts to hinder their growth. This effect actually appears as a disturbance in the heat flow and consequently in the temperature change.

It should be appreciated that one can never make a list of all the possible sources of disturbances even in a single process. It is very difficult to specify their exact locations, magnitudes and types. Fortunately, when dominant disturbances are considered and accounted for in the control system design, the control can be made reasonably adequate.

The effect of a disturbance in the process is also very important. We have mentioned above that a change in the demand side causes certain disturbances and a change in the supply side causes different disturbances when considered by the controller. Since a process is a combination of a number of units, the effects of disturbances at various points are different and should be considered separately. This aspect has been dealt with in Chapter 3.

In process control, the dominant changes that are to be considered are generally predetermined load changes which result when the through-put to the process changes. These changes which are made effective mostly on economic considerations, affect the product output in the market and supply of raw materials.

In other cases, when the output of one unit acts as the raw material to another, any disturbance to the former may affect the supply to the latter, producing load changes in it. These are often not predetermined but provision for such sudden changes is always pre-worked.

1.5 CONTROL ACTIONS

In a process control system, the choice of the control equipment is determined by the process itself. In simple processes where process time-constant is dominant over the transportation lag, the choice of control equipment and action poses no problem.

A few basic control actions that are used as such or with combination and/or modification are: (i) on-off, (ii) proportional, (iii) integral and, (iv) derivative. A proper selection of the type of action to be used can be made when at least, the process reaction curve is known. The process reaction curves are the step-response curves of the processes.

1.5.1 On-Off Action

The controller with this action is used in many common situations, such as air-coolers, water reservoirs, batch annealing furnaces, etc. Its action is very simple. When the process variable exceeds the set-point (reference) the controller gives no output signal, i.e., it is OFF, and vice versa. Physical limitations of the equipment, such as friction, inertia, etc. force the controller to be ON and OFF over a band around the set-point, as shown in Fig. 1.3. This band is the dead zone, sometimes called the differential gap, and is often chosen by the designer. The larger this gap, the less the number of times the contacts close or break and, therefore, the less is the wear and tear. But then the control limits are also more and control accuracy less.

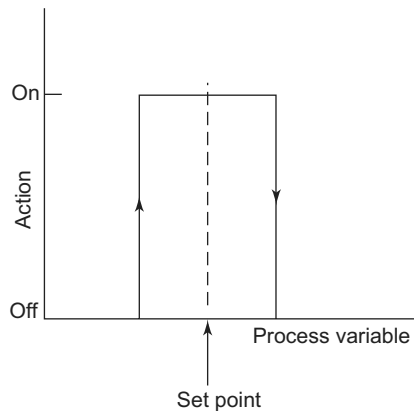


Fig. 1.3 Block diagram of the basic control loop

1.5.2 Proportional Action (P)

A controller of this type has its output, y , proportional to the error, e . Thus

$$y = K_c e + y_0 \quad (1.23)$$

where y_0 is the bias necessary for the actuator and is the value of y at $e = 0$. This is given to avoid process shut down.

Often, in practice, Eq. (1.23) is expressed as

$$y = (100/PB)e + y_0 \quad (1.24)$$

where k_c is called the proportional gain and PB is known as the proportional band expressed in percentage.

1.5.3 Integral Action (I)

In this action, the controller output is proportional to the integral of the error. Thus

$$y = (1/T_I) \int e \cdot dt + y_0 \quad (1.25)$$

Here, T_I , the reciprocal of the constant of proportionality, is known as the integral time. Integral action is often used with proportional action and this integral action time is multiplied by the proportional gain. This increased time is often referred to as the reset time and denoted by T_r .

1.5.4 Derivative Action (D)

In the derivative action, the output of the controller is proportional to the rate of change of the error. Thus

$$y = T_D de/dt + y_0 \quad (1.26)$$

However, this action is fast acting and is rarely used without proportional action, in which case, T_D is divided by the proportional gain to yield the rate time T_d . Here, T_D is known as the derivative time.

Laplace transformation equations (1.23), (1.25) and (1.26) are replaceable by transfer functions

$$\frac{y(s)}{e(s)} = K_c \quad (1.27a)$$

$$\frac{y(s)}{e(s)} = \frac{1}{sT_I} \quad (1.27b)$$

and

$$\frac{y(s)}{e(s)} = sT_D \quad (1.27c)$$

where initial conditions are ignored. Also, the combined PID controller transfer function is given by

$$\frac{y(s)}{e(s)} = K_c \left[1 + \frac{1}{sT_R} + sT_d \right] \quad (1.28)$$

Often, situations demand that all these actions, P , I and D , should be simultaneously used. In Chapter 5, the applicability of different control actions, choice of K_c , T_r and T_d and other related topics are discussed.

For very complex processes, the application of controllers on a single loop basis may prove inadequate. Depending on the type of complexity, different control strategies are considered. A challenge is often met by adopting multi-loop control systems without changing the types of control actions. A few important cases have been considered in detail in Chapter 6.

1.6 Z-TRANSFORMS

In digital control, the mathematical tool used for loop and general analysis is the Z-transform as has been demonstrated in Chapter 9. Discrete analysis starts with sampling and a continuous signal when sampled has, ideally, a magnitude at the sampling time and no area covered, as given in Fig. 1.4. The continuous signal, a time function, has now impulse approximations represented as $f^*(t)$ which with the help of the unit impulse function would be represented as

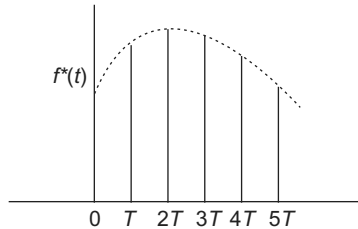


Fig. 1.4 Sampled signal

$$\begin{aligned} f^*(t) &= f^*(0)\delta t + f^*(T)\delta(t-T) + f^*(2T)\delta(t-2T) + \dots \\ &= \sum_{n=0}^{\infty} f(nT)\delta(t-nT) \end{aligned} \quad (1.29)$$

It is assumed that unit impulse at instant $t = kT$ is given by $\delta(t - kT)$.

The Laplace transform of $f^*(t)$ is given by

$$F^*(s) = L[f^*(t)] = f(0) + f(T)e^{-sT} + f(2T)e^{-2sT} + \dots \quad (1.30)$$

If in Eq. (1.30), e^{sT} is replaced by z or s is replaced by $\ln z/T$, we get the Z -transform of $f^*(t)$ given as

$$\begin{aligned} Z[f^*(t)] = F(z) &= f(0) + f(T)z^{-1} + f(2T)z^{-2} + \dots \\ &= \sum_{n=0}^{\infty} f(nT)z^{-n} \end{aligned} \quad (1.31)$$

This is the formal definition of Z -transform. However, it is subjected to constraint of convergence which implies that $|z|$ lies within certain limits.

A very common method of obtaining the Z -transform of a time function is to use the residue method. The relevant relation is

$$\begin{aligned} Z\text{-transform of } [f^*(T)] = F(z) &= \sum \text{residues of } F(s) \frac{z}{z - e^{sT}} \\ &\text{at poles of } F(s) \end{aligned} \quad (1.32)$$

If the denominator of $F(s)$ has a factor $(s - \alpha)$ so that $F(s)$ has only one pole at α , the residue is

$$R = \lim_{s \rightarrow \alpha} (s - \alpha) \left[F(s) \frac{z}{z - e^{sT}} \right] \quad (1.33)$$

If the denominator has multiplicity in poles $(s - \alpha)^r$, then

$$R = \frac{1}{(r-1)!} \lim_{s \rightarrow \alpha} \frac{d^{r-1}}{ds^{r-1}} \left[(s - \alpha)^r F(s) \frac{z}{z - e^{sT}} \right] \quad (1.34)$$

Example 2

Find the Z -transform of

$$F(s) = \frac{1}{(s + \alpha)}$$

Solution $F(s)$ has a pole at $-\alpha$

$$R = \lim_{s \rightarrow -\alpha} (s + \alpha) \left[\frac{1}{(s + \alpha)} \frac{z}{z - e^{-sT}} \right] = \frac{z}{z - e^{-\alpha T}}$$

which is also $F(z)$.

For unit step function, $F(s) = 1/s$ so that residue can be calculated at pole $s = 0$ to be $z/(z - 1)$ which is $F(z)$ of $f(t) = 1$.

Example 3Find $F(z)$ for

$$F(s) = \frac{1}{(s + \alpha)(s + \beta)}$$

Solution Two poles are at $-\alpha$ and $-\beta$. The residues are obtained as

$$\begin{aligned} R_1 + R_2 &= \lim_{s \rightarrow -\alpha} (s + \alpha) \left[\frac{1}{(s + \alpha)(s + \beta)} \frac{z}{z - e^{-sT}} \right] \\ &\quad + \lim_{s \rightarrow -\beta} (s + \beta) \left[\frac{1}{(s + \alpha)(s + \beta)} \frac{z}{z - e^{-sT}} \right] \\ &= \frac{1}{\beta - \alpha} \frac{z}{z - e^{-\alpha T}} + \frac{1}{\alpha - \beta} \frac{z}{z - e^{-\beta T}} \\ &= \frac{z}{\alpha - \beta} \frac{e^{-\beta T} - e^{-\alpha T}}{(z - e^{-\beta T})(z - e^{-\alpha T})} \end{aligned}$$

Example 4

Find the Z-transform of

$$F(s) = \frac{\omega}{s^2 + \omega^2}$$

Solution Writing $F(s) = \frac{\omega}{(s + j\omega)(s - j\omega)}$ and using the residue theorem,

$$\begin{aligned} R_1 + R_2 &= \lim_{s_1 \rightarrow j\omega} \left\{ (s - j\omega) \left[\frac{\omega}{(s - j\omega)(s + j\omega)} \frac{z}{z - e^{sT}} \right] \right\} \\ &\quad + \lim_{s_2 \rightarrow -j\omega} \left\{ (s + j\omega) \left[\frac{\omega}{(s - j\omega)(s + j\omega)} \frac{z}{z - e^{sT}} \right] \right\} \\ &= \frac{\omega}{2j\omega} \frac{z}{z - e^{j\omega T}} - \frac{\omega}{2j\omega} \frac{z}{z - e^{-j\omega T}} \\ &\quad = \frac{z(e^{j\omega T} - e^{-j\omega T})}{2j} \\ &= \frac{2j}{z^2 - 2z \left[\frac{e^{j\omega T} + e^{-j\omega T}}{2} \right] + 1} \\ &= \frac{z^{-1} \sin \omega T}{1 - 2z^{-1} \cos \omega T + z^{-2}} \end{aligned}$$

The term z^{-1} is called the backward shift operator and used conveniently in transformation of different equations. Some useful theorems have been given in Section 9.6.1 without proofs. Readers may consult appropriate texts if they are interested in the proofs.

1.6.1 Inverse Z-Transforms

For obtaining the inverse Z-transforms, again the residue method can be adopted. The alternative is the partial fraction method. In the residue method, the inverse $f(nT)$ is found as

$$f(nT) = \sum \text{residues of } F(z) \cdot z^{n-1} \text{ at poles of } F(z)z^{n-1}$$

Taking Example 3 above,

$$F(z) = \frac{z}{\alpha - \beta} \frac{e^{-\beta T} - e^{-\alpha T}}{(z - e^{-\beta T})(z - e^{-\alpha T})}$$

$$\begin{aligned} \sum \text{Residues} &= \lim_{z \rightarrow e^{-\beta T}} \left[\frac{z^n}{\alpha - \beta} \frac{e^{-\beta T} - e^{-\alpha T}}{z - e^{-\alpha T}} \right] + \lim_{z \rightarrow e^{-\alpha T}} \left[\frac{z^n}{\alpha - \beta} \frac{e^{-\beta T} - e^{-\alpha T}}{z - e^{-\beta T}} \right] \\ &= \frac{e^{-\beta n T}}{\alpha - \beta} \frac{e^{-\beta T} - e^{-\alpha T}}{e^{-\beta T} - e^{-\alpha T}} + \frac{e^{-\alpha n T}}{\alpha - \beta} \frac{e^{-\beta T} - e^{-\alpha T}}{e^{-\alpha T} - e^{-\beta T}} \\ &= \frac{e^{-\beta n T}}{\alpha - \beta} - \frac{e^{-\alpha n T}}{\alpha - \beta} \\ &= \frac{1}{\alpha - \beta} [e^{-\beta n T} - e^{-\alpha n T}] \\ &= f(nT) \end{aligned}$$

For the same example in partial-fraction mode, one gets

$$F(z) = \frac{1}{\alpha - \beta} \left[\frac{z}{z - e^{-\beta T}} - \frac{z}{z - e^{-\alpha T}} \right]$$

Referring to the standard table given in Appendix I, one gets

$$f(nT) = \frac{1}{\alpha - \beta} [e^{-\beta n T} - e^{-\alpha n T}]$$

In control, modified Z-transform is often used to account for response in between the samples. The basics of such a transform is given in Section 9.6.2.

1.7 GENERAL COMMENTS

As has already been pointed out, the major part in the process control system is the process itself. The control strategy depends almost entirely on it. In chemical processes the stress is on regulation, which means that control is to be ensured against disturbances for a fixed reference. In process control systems there are only a few tracking problems; two common examples described in Chapter 5 and 6 are the programmed type of control and cascade control.

A control loop is incomplete without an actuator or a final control element. In fact, final control element forms a major part of study in process control. Still the single most important final control element used in process industries is the pneumatic actuator cum control valve. With the advent of digital systems in process automation, electrical actuators like stepper motors are also being used increasingly. Sizing, selection and types of common pneumatic valves as also electrical ones have been discussed in Chapter 7 to a certain extent.

In many process control problems, four very important variables demand special attention. These are flow, pressure, level and temperature. Flow processes are very fast and temperature processes are very sluggish and pressure and level appear in that order in the series of variables given above. In fast processes, the response of the loop elements other than process and equipment, with regard to time/frequency is also very important. Further, in flow, pressure, level and temperature control systems, nonlinearity in modelling appears, which is to be linearized when the control strategy is to be determined in that light. Types of variations in flow, pressure, level and temperature in interconnected systems also deserve due consideration and a control system structure is to be accordingly ascertained. Chapter 8 is concerned with the development in this direction.

In the control of large and complex processes, recent trends are towards sophistication since digital computers are readily available for use in the processes. In very large plants, the use of digital computers improves the performance and simultaneously reduces the cost as well. Additionally, optimizing the control of processes and making the processes adaptive and self-adaptive in response to unpredictable disturbances have been found to be possible. Small processes and units of bigger ones have, however, been found to be more fittingly controlled in the optimized but programmed sense by adapting microprocessors. In a sense, industrially we are in real automation. Chapter 9 and 10 have been designed to cover a few important aspects of the modern trends towards this necessary sophistication. While Chapter 9 covers the general digital control aspects pointing out the advantages of using mathematical tools such as Z-transformations and including a description on present-day distributed digital control systems, Chapter 10 has been designed to discuss adaptive control systems and more recent offshoot of them—the self-tuning control.

It is not enough to discuss the principles of process control alone. The application aspect of the same is equally, if not more, important. Plants and processes have characteristics that need careful consideration for choosing a specific strategy for their control. The techniques can be demonstrated by taking examples. Chapter 11 is designed to include a few such examples as stated below.

A very common and important but difficult process is the boiler. This is a mixed process and sufficiently complex too. In this process, temperature, pressure, flow and level are all to be simultaneously controlled and the process is a glowing example of interconnected, interacting control. Non-interacting control systems for such complex processes are being developed. Taking the boiler as an example, the existing methods and the methods in the offing are briefly considered here. Also considered in Chapter 11 are some aspects of steel plant and paper-making plant controls. Control of soaking pits has been discussed in a little detail when different types of fuels are used simultaneously. In the part of paper mill control, control during stock/pulp preparation as also the drive controls in different stages of paper making have been considered. Different control strategies for a distillation column have also been included in this chapter*.

Review Questions

1. What are the studies required in a process for completing the loop? What are the studies made with the loop closed?
2. What are the different interconnecting lines in a process control system? How are they represented?
3. How do you differentiate a regulatory system from a reference follower system?
4. What are the three common different control actions used in process control systems? How do they act?
5. How do you differentiate reset and rate actions from integral and derivative actions respectively?
6. The proportional action gain of a controller is 2.25. What would be the value of the proportional band?
7. Find the time function of (a) $\frac{2s^2 + 15s + 29}{s^2 + 7s + 12}$

[Hint: Divide numerator by denominator to get

$$\begin{aligned} F(s) &= 2 + (s + 5)/(s^2 + 7s + 12) \\ &= 2 + (s + 5)/[(s + 3)(s + 4)] \end{aligned}$$

*Control of pH is highly nonlinear—inclusion of such a system in this chapter enriches its worth. Finally, some aspects of batch process control have been given coverage as complementing items.

$$= 2 + 2/(s+3) - 1/(s+4)$$

From this, the inverse operation is tried to get

$$f(t) = 2 + 2e^{-3t} - e^{-4t}]$$

8. If $L[f(t)] = F(s)$, prove that

$$L\left[\int_0^t f(t) dt\right] = \frac{1}{s} F(s)$$

[Hint:

$$L\int_0^t f(t) dt = \int_0^\infty e^{-st} \left(\int_0^t f(t) dt \right) dt$$

Integrate by parts, $u = \int_0^t f(t) dt$, $dv = e^{-st} dt$

This gives

$$L\left[\int_0^t f(t) dt\right] = -\frac{e^{-st}}{s} \int_0^t f(t) dt \Big|_0^\infty + \frac{1}{s} \int_0^\infty f(t) e^{-st} dt$$

This leads to the solution.]

9. Find the Z-transforms of

(a) $f(t) = t$, (b) $F(s) = 1/(s+1)$

[Hint: (a), $f(t) = t$, gives $F(s) = 1/s^2$ which shows multiple (a pair) poles at origin. Using Eq. (1.32)

$$F(z) = \frac{Tz}{(z-1)^2}]$$

10. Find the Z-transform of $F(s) = \frac{1}{s(s+1)(s+2)}$

[Hint: Use the residue theorem.]

11. Why is Laplace transform used in engineering science and specially in control engineering almost always? What domain does it work in?

Find the Laplace transform of (1) $F(t) = t^3$, and (2) $f(t) = te^{-4t}$

12. The input and output of a block are x and y respectively, the block characteristic is determined by a second-order linear time invariant differential equation. Write the transfer function of the block specifying the system parameters.

Processes: Transfer Functions and Modelling

2.1 INTRODUCTION

The major part in process control is the plant or the process itself. For adequate control, a knowledge of the plant characteristics, both qualitative and quantitative, is important. Analysis of the plant, both theoretical and experimental, is, therefore, necessary. Theoretical analysis can only be approximate because many of the factors contributing to the dynamics of any process are approximately known. In linear systems, the process is generally split into separate units, each of which contributes a single, a pair or a number of time constants. The transfer function of these is obtained or other mathematical modelling is easily carried out with sufficient quantitative accuracy. Obviously, the question arises whether to choose this procedure for plants of extreme complexity which may not be so easily split, and if it is adopted for these plants, approximate modelling is bound to distort the true picture severely. Such specific mathematical models cannot provide adequate quantitative information regarding the actual process. Lack of such information, uncertainty regarding the disturbances and plant complexity are the prime causes of avoiding quantitative models. Qualitative models are more popular with the system designers and also previous experience with different processes is considered useful for the design.

In the next few sections transfer functions are derived for some typical processes from theoretical concepts.

2.2 SOME TYPICAL SIMPLE PROCESSES AND THEIR TRANSFER FUNCTIONS FROM ANALYSIS

In view of the importance of the mathematical model of the process itself in process control analysis, a few simple processes are given here with the derivation of their performing equations, specifically, the transfer functions. Thus, for the behavioural study of the systems that are important from amongst a few common types of processes in chemical plants such as distillation, flow of fluid, absorption, mass transfer, extraction, mixing, evaporation, flow of heat, material handling, etc., are considered and it is shown that the equations can be derived for such cases with some approximations.

A few of the derivations are now given in brief. These linearized equations are basic to many unit operations. In the units considered dead time or multilags are assumed absent. It should be remembered that this is a theoretical approach and, therefore, approximate, in so far as derivations are concerned. Even the process complexities have been ignored and simple cases have only been considered.

2.2.1 Heat Transfer

Heat input to a tank is shown in Fig. 2.1. If heat transfer is given by H_t , overall heat transfer coefficient by H_0 , tank mass by m , specific heat of the material by C and temperature difference between the heat donor and heat receiver by T_ϕ then the heat balance equation is

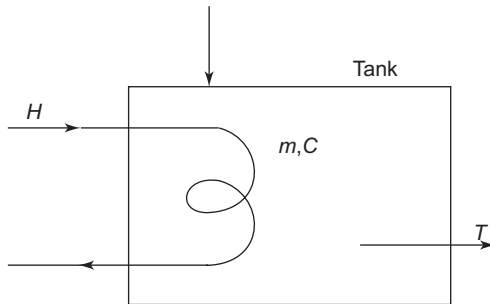


Fig. 2.1 Schematic diagram of the tank of heat transfer process

$$H_t = \alpha H_0 T_d + mC(dT_d/dt) \quad (2.1)$$

where α is the heat transfer area and

$$H_0 = 1/\left(\sum_{j=1}^n \frac{1}{h_j}\right) \quad (2.2)$$

The term h_j represents the heat transfer coefficient, for radiation, convection, conduction, etc., is temperature dependent and Eq. (2.1) is, therefore, valid only for small variations. From Eq. (2.1), the transfer function $T_H(s)$ is obtained as

$$T_H(s) = \frac{T_d(s)}{H_i(s)} = \frac{1/(\alpha H_0)}{s \left(\frac{mC}{\alpha H_0} \right) + 1} = \frac{k}{s\tau + 1} \quad (2.3)$$

A single-lag approximation with a time constant of $\tau = mC/\alpha H_0$ of a heat transfer process is thus made. The block diagrammatic representation of the process is shown in Fig. 2.2. A little more practical heat transfer process is shown in Fig. 2.3. Steam line with a fixed temperature T_s heats up the tank fluid which has an incoming fluid flow and temperature q_i and T_i respectively and an outgoing fluid flow and temperature q_0 and T_0 respectively. Usually for continuity of an incompressible fluid of density ρ , $q_i = q_0 = q$ such that the heat balance equation is

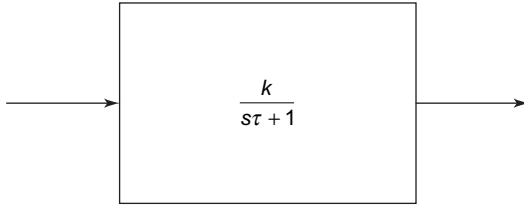


Fig. 2.2 Block diagrammatic representation of Fig. 2.1

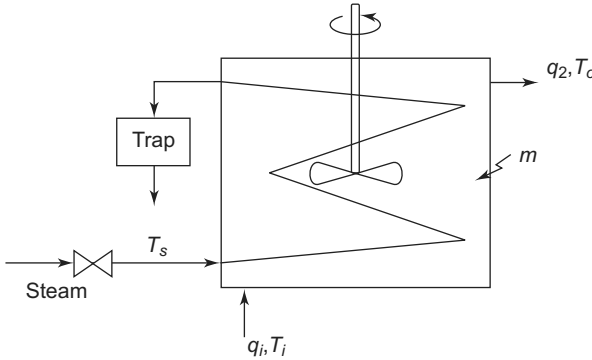


Fig. 2.3 Schematic representation of a practical heat transfer process

$$q\rho C(T_i - T_0) + \alpha H_0 (T_s - T_0) = mCdT_0/dt \quad (2.4)$$

which gives the transfer function in terms of T_i and T_s . The relevant equation is

$$\left[\left(\frac{mC}{q\rho C + \alpha H_0} \right) s + 1 \right] T_0 = \frac{q\rho C}{q\rho C + \alpha H_0} T_i + \frac{\alpha H_0}{q\rho C + \alpha H_0} T_s \quad (2.5)$$

i.e.,

$$(s\tau + 1)T_0 = k_1 T_i + k_2 T_s \quad (2.6)$$

The block schematic representation of such a system is given in Fig. 2.4.

2.2.2 Mixing Process

Mixing is a direct process for achieving either material balance or thermal balance. Depending on its application, the control system design will change. Here a material balance principle is considered first. Figure 2.5 shows the scheme of the process. The q 's are volume flow rates and χ 's are concentrations. If the m 's are masses, the material balance equation is

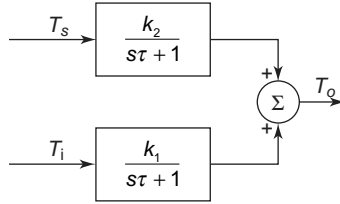


Fig. 2.4 Block diagrammatic representation of the process of Fig. 2.3

$$(dm/dt)_{\text{nett}} = (dm/dt)_{\text{influent}} - (dm/dt)_{\text{effluent}} \quad (2.7a)$$

giving

$$v(d\chi_0/dt) = q_i\chi_i - q_0\chi_0 \quad (2.7b)$$

where v is the nett volume of the tank. The above relations can be drawn when one assumes that mixing is ideal such that effluent concentration is identical with tank concentration and the influent relation is

$$q_i\chi_i = q_1\chi_1 + q_2\chi_2 + \dots + q_r\chi_r \quad (2.8a)$$

$$= q_1\chi_1 \left(1 + \sum_{k=2}^r f_k \phi_k \right) \quad (2.8b)$$

where

$$f_k = q_k/q_1, \text{ and, } \phi_k = \chi_k/\chi_1$$

In a control system χ_0 is changed by changing the q_j 's, $j = 1, 2, \dots, r-1$, i.e., by changing the f_k 's. The f_k 's can, however, be changed by changing q_1 alone, as shown by

$$sv\chi_0 + q_0\chi_0 = q_i\chi_i = q_1\chi_1 \left(1 + \sum_{k=2}^r f_k\phi_k \right) \quad (2.9a)$$

so that

$$\frac{\chi_0(s)}{q_1(s)} = \frac{\frac{x_1}{q_0} \left(1 + \sum_{k=2}^r f_k\phi_k \right)}{\frac{sV}{q_0} + 1} = \frac{k_m}{s\tau_m + 1} \quad (2.9b)$$

The block representation of this is similar to that shown in Fig. 2.5(a).

If the thermal balance is considered, fluids at different temperatures and at different weight rates of flow are allowed to mix in a tank of mass m for an effluent at a specified temperature and weight rate of flow. Thus, from Fig. 2.5(b).

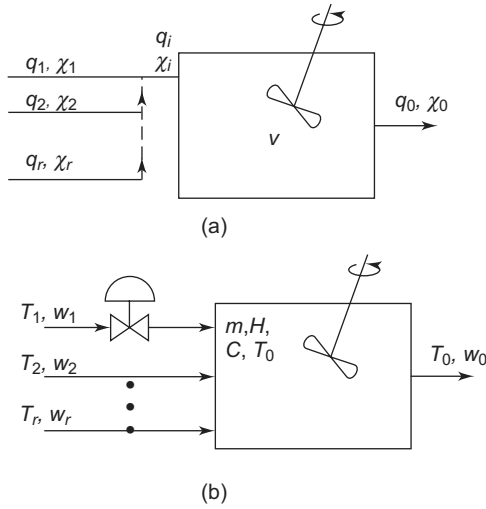


Fig. 2.5 (a) Scheme of a mixing process with material balance
(b) Scheme of a mixing process with thermal balance; m : mass, H : heat, C : specific heat, T 's: temperatures, w 's: weight rates of flow

$$H - H_R = mCd(T_0 - T_R)/dt \quad (2.10)$$

where suffix R denotes reference. The thermal balance is given as

$$(dH/dt)_{\text{nett}} = (dH/dt)_{\text{infl.}} - (dH/dt)_{\text{effl.}} \quad (2.11a)$$

Hence

$$mCdT_0/dt = w_1(T_1 - T_R)C + w_2(T_2 - T_R)C + \dots + w_r(T_r - T_R)C - w_0(T_0 - T_R)C \quad (2.11b)$$

For control of T_0 , T_1 may be controlled, so that writing Eq. (2.11b) as

$$mC dT_0/dt + w_0T_0C = w_1T_1C + \sum_{j=2}^r w_jT_jC + T_R C \left(w_0 - \sum_{i=1}^r w_i \right) \quad (2.12)$$

one gets

$$\frac{m}{w_0} \frac{dT_0}{dt} + T_0 = T_1 \left[w_1/w_0 + \frac{\sum_{j=2}^r w_jT_j + T_R \left(w_0 - \sum_{i=1}^r w_i \right)}{w_0T_1} \right] \quad (2.13a)$$

Taking transformation now

$$T_0(sm/w_0 + 1) = T_1 \left[w_1/w_0 + \frac{\sum_{j=2}^r w_jT_j}{w_0T_1} \right] \quad (2.13b)$$

which presents a little difficulty in obtaining the transfer function. However, assuming $T_j/T_1 = \psi_j$, one gets as in Eq. (2.9b)

$$T_0(s)/T_1(s) = 1/(s\tau_t + 1) \left[1/w_0 \left(w_1 + \sum_{j=2}^r w_j\psi_j \right) \right] = K_t/(s\tau_t + 1) \quad (2.13c)$$

where τ_t is the residence or hold time for the tank.

Equation (2.9b) shows a transfer function between concentration and flow-rate, whereas Eq.(2.13c) shows a transfer function between two temperatures for the same mixing-process which are arrived at by considering material balance and thermal balance respectively.

Likewise for a heat exchanger transfer functions between temperature and heat flow and between temperatures can be obtained, as already shown.

2.2.3 Stirred Tank Reactor

Continuously stirred tank is another very common chemical process where a transfer function between temperatures may be written for use in the

control loop. The equation is derived from the heat balance of the reactor, schematic diagram of the same being given in Fig. 2.6. The heat balance equation is

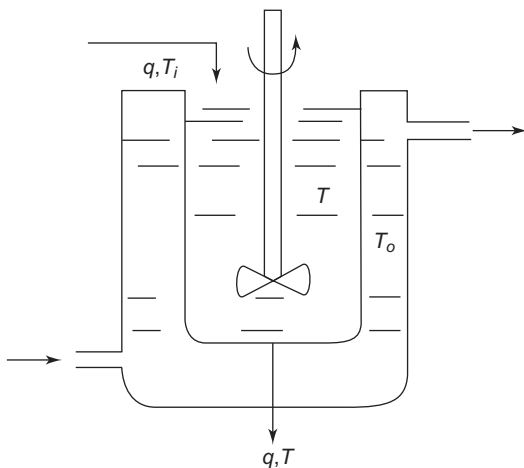


Fig. 2.6 Stirred tank reactor

$$V\rho C_p(dT/dt) = q\rho C_p(T_i - T) - h\alpha(T - T_0) + (\partial Q/\partial T)_x T \quad (2.14a)$$

where, V = volume in the reactor, ρ = density, C_p = heat capacity, q = feed rate, h = overall heat transfer coefficient, α = area of heat transfer, $(\partial Q/\partial T)_x$ = change in heat generation with temperature and T = temperature. Equation (2.14a) can be rearranged to obtain temperature functions

$$T(s) = K_1 T_i(s)/(s\tau + 1) + K_2 T_0(s)/(s\tau + 1) \quad (2.14b)$$

where time constant τ is $V\rho C_p/[q\rho C_p + h\alpha - (\partial Q/\partial T)_x]$, and constants K_1 and K_2 are given respectively by

$$K_1 = \frac{q\rho C_p}{q\rho C_p + h\alpha - \left(\frac{\partial Q}{\partial T}\right)_x}, \text{ and } K_2 = \frac{h\alpha}{q\rho C_p + h\alpha - \left(\frac{\partial Q}{\partial T}\right)_x}$$

The transfer function may now be obtained between T and T_i or T and T_0 keeping the other temperature constant.

One important aspect of such a reactor is the effect of rate constant defined as $K_r = K_0 \exp(-E/RT)$, where E is the activation energy, and the system order on it. The heat generated is calculated as

$$Q = K_r V \chi_q (1 - x)(-\Delta H) \quad (2.15a)$$

where, χ_q = feed concentration, x = fraction of conversion and can be calculated from $x/(1 - x) = K_r V/q$ and ΔH = heat of reaction.

Hence,

$$(\partial Q/\partial T)_x = (dK_r/dT)V\chi_q(1-x)(-\Delta H) \quad (2.15b)$$

Thus the time constant is given by

$$\tau = \frac{V\rho C_p}{q\rho C_p + h\alpha - (dK_r/dT)V\chi_q(1-x)(-\Delta H)} \quad (2.16)$$

For stability, $q\rho C_p + h\alpha > (dK_r/dT)V\chi_q(1-x)(-\Delta H)$. With increase in temperature the rate constant usually becomes less so that the time constant tends to be smaller provided converted factor and heat of reaction do not change. For higher order reaction, the rate constant is first evaluated and then the effect on system time constant, as has been done above.

2.2.4 Tubular Reactor

A tubular reactor is rather a difficult process involving flow kinetics. In this a dead time arises and makes the process controllability difficult (Cf. Ch. 4). If the reaction is irreversible with reactant, R and product P with the reaction rate ρ_r , then $R \xrightarrow{P_r} P$. Considering a small length Δl of the reactor over which the product concentration χ_p changes to $\chi_p - \Delta\chi_p$, and if the reactant concentration is χ_R , the material balance equation can easily be derived (Fig. 2.7) as

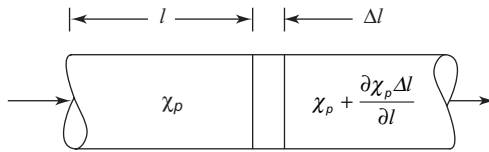


Fig. 2.7 Scheme of a tubular reactor

$$\partial\chi_p/\partial t + v\partial\chi_p/\partial l = \rho_r\chi_R \quad (2.17)$$

where v is the stream speed. At $l = 0$, product concentration is zero and the reactant and product concentration can be related by a difference equation

$$\chi_R(l, t) = \chi_{R0}\{0, (t - l/v)\} - \chi_p(l, t) \quad (2.18)$$

χ_{R0} being the initial reactant concentration. Combining Eqs (2.17) and (2.18), one gets

$$\frac{\partial\chi_p}{\partial t} + v\left[\frac{\rho_r}{v}\chi_p(l, t) - \frac{\partial\chi_p}{\partial l}\right] = \rho_r\chi_{R0}\{0, (t - l/v)\} \quad (2.19)$$

Taking Laplace transform and integrating between $l = 0$ and $l = l$, the transfer function is

$$T(s) = \chi_p(s)/\chi_{R0}(s) = \exp(-sl/v)[1 - \exp(-\rho_rl/v)] \quad (2.20)$$

Dead time or the transportation lag is thus l/v which can be seen to be obtained qualitatively. Equation (2.20) can thus be written as

$$T(s) = [1 - \exp(-\rho_r \tau_d)] \exp(-s\tau_d) \tag{2.21}$$

Figure 2.8 shows the block diagrammatic representation of this equation.

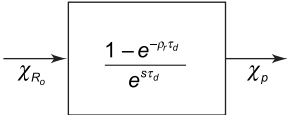


Fig. 2.8 Block diagrammatic representation of Fig. 2.7 with material balance, χ_R 's: concentrations of the reactants

2.2.5 Distillation Column

Distillation column is a very important chemical process and is an example of a mass transfer process. It has a number of stages, each stage consisting of four lags, namely: (i) the concentration lag due to the liquid volume (capacity) held by the plate—this is the largest lag, (ii) the liquid flow rate lag—this occurs due to a change in the hold-up by the plate with flow rate, (iii) the vapour flow rate lag—this occurs due to a change in the hold-up with pressure, i.e., flowrate, and (iv) the vapour concentration lag—this is the smallest in value and can often be neglected.

The schematic representation of a distillation column is shown in Fig. 2.9 with a reboiler and condenser. Top product P_t is the distillate with composition χ_t , the bottom product is P_b with composition χ_b . The parameters L_ρ and V_ρ represent the internal liquid and vapour rates respectively. The feed rate is given by ρ_f and the feed composition by χ_f . It is extremely difficult to obtain lags in n -stage columns because of interaction between the stages. The concentration lags are also dependent on flow rates in columns and feed rates besides hold-up. Another important consideration is the composition gradient, i.e., how the vapour composition and liquid composition are graded along the column over the given range of composition. The relationship between χ_t (vapour) and χ_b (liquid) is important for evaluating the transfer function between χ_t and χ_f . For a single stage column, for a hold-up of h per stage

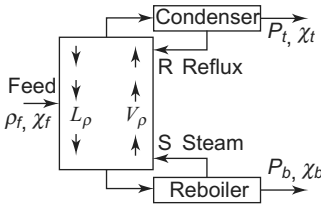


Fig. 2.9 Schematic diagram of a distillation column

$$hd\chi_b/dt = \rho_f\chi_f - P_b\chi_b - P_t\chi_t \quad (2.22)$$

If χ_t and χ_b are linearly related, i.e., assuming that the vapour composition is linearly related to liquid composition

$$\partial\chi_t/\partial\chi_b = \chi_t/\chi_b = b \quad (2.23)$$

Then Eq. (2.22) changes to

$$\frac{\chi_t(s)}{\chi_f(s)} = \frac{b\rho_f(P_b + bP_t)}{sh/(P_b + bP_t) + 1} = \frac{K}{s\tau + 1} \quad (2.24)$$

It may be noted that $P_b + P_t = \rho_f$ and both K and τ are dependent on b . If b increases τ decreases and K also decreases indicating a less steady state change of χ_t with χ_f .

In the above we have assumed zero reflux. Considering a two-stage column with reflux R , the following material balance equations are obtained:

$$h_1d\chi_1/dt = V_\rho b\chi_2 - R\chi_1 - P_t\chi_t \quad (2.25)$$

$$= V_\rho b\chi_2 - R\chi_1 - P_t b\chi_1 \quad (2.26)$$

and

$$h_2d\chi_2/dt = \rho_f\chi_f - P_b\chi_2 + R\chi_1 - V_\rho b\chi_2 \quad (2.27)$$

Equations (2.25) and (2.26) are for the top plate and Eq.(2.27) is for the bottom plate. All these three equations are obtained using Eqs (2.22) and (2.23) and in these equations suffixes t and b are for the top and bottom stages respectively.

Equations (2.25), (2.26) and (2.27) may be combined to obtain transfer function $\chi_1(s)/\chi_f(s)$ or $\chi_2(s)/\chi_f(s)$. Thus

$$\frac{\chi_1(s)}{\chi_f(s)} = \frac{bV_\rho\rho_f}{h_1h_2s^2 + \{h_1(P_b + bV_\rho) + h_2(R + bP_t)\}s + (P_bR + bP_bP_t + b^2V_\rho P_t)} \quad (2.28a)$$

$$= \frac{K_1}{\alpha s^2 + \beta s + 1} = \frac{K_1}{(s\tau_1 + 1)(s\tau_2 + 1)} \quad (2.28b)$$

and

$$\frac{\chi_2(s)}{\chi_f(s)} = \frac{K_1}{\alpha s^2 + \beta s + 1} \frac{h_1s + R + bP_t}{bV_\rho} = \frac{(R + bP_t)K_1}{bV_\rho} \left\{ \frac{h_1s}{R + bP_t} + 1 \right\} \quad (2.29a)$$

$$= \frac{K_2(s\tau_3 + 1)}{(s\tau_1 + 1)(s\tau_2 + 1)} \quad (2.29b)$$

where

$$K_1 = \frac{bV_\rho\rho_f}{P_bR + bP_bP_t + b^2V_\rho P_t}, \alpha = \frac{h_1h_2}{P_bR + bP_bP_t + b^2V_\rho P_t},$$

$$\beta = \frac{h_2(R + bP_t) + h_1(P_b + bV_\rho)}{P_bR + bP_bP_t + b^2V_\rho P_t} \quad (2.30a)$$

and

$$\tau_{1,2} = \frac{\beta \pm \sqrt{\beta^2 - 4\alpha}}{2}, \tau_3 = \frac{h_1}{R + bP_t}, K_2 = \frac{(R + bP_t)K_1}{bV_\rho} \quad (2.30b)$$

The above relations show the effect of b , h_1 , h_2 , as also feed rates V_ρ , P_b , P_t and ρ_f on K_1 , K_2 and τ_1 , τ_2 and τ_3 from which the responses can be evaluated.

The effect of the liquid flow rate can be calculated by the material balance equation

$$A d\lambda/dt = \Delta(L_\rho)_j - [d(L_\rho)_{j+1}/d\lambda]\lambda \quad (2.31)$$

where A = effective plate area, λ = clear liquid level in the plate. Equation (2.31) will now yield

$$A \left(\frac{d\lambda}{d(L_\rho)_{j+1}} \right) \frac{d\lambda}{dt} + \lambda = \left[\frac{d\lambda}{d(L_\rho)_{j+1}} \right] \Delta(L_\rho)_j$$

or

$$\frac{\Delta(L_\rho)_j(s)}{\lambda(s)} = \frac{d\lambda/d(L_\rho)_{j+1}}{As[d\lambda/d(L_\rho)_{j+1}] + 1} \quad (2.32)$$

But one easily notes that, when linearized,

$$\frac{d\lambda}{d(L_\rho)_{j+1}} = \frac{1}{A} \frac{dh}{dL_\rho}$$

such that time constant due to liquid flow rate is

$$\tau_{fl} = dh/dL_\rho \quad (2.33)$$

Similarly, for a vapour flow rate the time constant can be calculated from the material balance equation

$$h_v \left(\frac{dp_j}{dt} \right) + \eta h_L \left(\frac{dT_B}{dt} \right) \left(\frac{dp_j}{dt} \right) = \frac{p_{j+1} - p_j}{r} - \frac{p_j - p_{j-1}}{r} \quad (2.34)$$

where

h = hold-up, suffix v for vapour space and L for liquid phase

p_j = vapour pressure in the interplate space

T_B = boiling point

r = flow resistance to vapour, and

η = liquid heat capacity/heat of condensation

The terms on the left hand side of Eq. (2.34) denote total accumulation between the $(j + 1)$ th and the $(j - 1)$ th plates from the top; the first term on the right-hand side denotes vapour inflow, the second term denotes vapour outflow.

Again assuming linearization, the time constant is calculated as

$$\tau_{fv} = 2r(h_v + \eta h_L dT_B/dP) \quad (2.35)$$

2.2.6 Nuclear Reactor

This is a very complicated process and the systematic development of a set of dynamic equations for use by the control engineers is not easy. So far attempts have been made with a simple schematic diagram of the reactor. Figure 2.10 shows such a diagram with the heat exchanger. A radial design has been assumed and basic output is considered to be related to heat such that temperature is the real credential. At the central inner place is the fuel which is canned. Outside the canning are the coolant channels whose outlets are at the top and inlets are at the bottom. Then finally moderators are provided. The heat exchange scheme is also shown complete with the pump. In a simplified scheme the coolant and moderator may be considered identical.

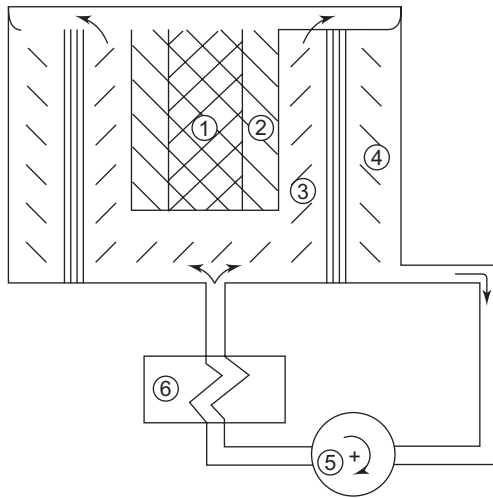


Fig. 2.10 Schematic diagram of a nuclear reactor—1: fuel, 2: canning, 3: coolant channels, 4: moderators, 5: pump, 6: heat exchanger

The inputs to the process are (i) supplied reactivity, ρ , (ii) coolant mass flow, w_c , and (iii) coolant inlet temperature, T_{ci} . The outputs are (i) temperature of fuel, cladding, coolant including moderator, (T_f , T_k and T_c , respectively), and (ii) coolant outlet temperature, T_{co} . System equations can be written following the heat balance on the assumption that heat accumulated = heat produced – drawn off heat.

Let C = specific heat, m = mass, h = heat transfer coefficient, α = heat transfer surface area, R = reactor output due to reaction, λ = decay constant, β = fraction of power given to the moderators and the suffixes are as used previously.

Also, assume that heating is mainly due to radiation such that an equation analogous to Stefan-Boltzman law may be written

$$R = h\alpha T^4 \quad (2.36)$$

Following the heat balance equation, the relevant equations are derived as

$$C_f m_f \frac{dT_f}{dt} = (\lambda - \beta)R - h_{fk}\alpha_{fk}(T_f - T_k) \quad (2.37)$$

$$C_k m_k \frac{dT_k}{dt} = h_{fk}\alpha_{fk}(T_f - T_k) - h_{kc}\alpha_{kc}(T_k - T_c) \quad (2.38)$$

Assuming an identical moderator and coolant as proposed

$$C_c m_c \frac{dT_c}{dt} = \beta R + h_{kc}\alpha_{kc}(T_k - T_c) + C_c w_c (T_{ci} - T_{co}) \quad (2.39)$$

In the formulations of the above equations activity laws have not been dealt with. The relations from Eqs (2.36) to (2.39) may now be combined to obtain the required relationship for determining the control strategy. This additionally requires certain assumptions about the process, such as the coolant rate, coolant outlet temperature to be constant, etc.

The problem has been considered in an oversimplified way but this approach coupled with further subdivisions in elements may produce better modelling.

2.2.7 Distributed Parameter Systems

It has been mentioned later in the chapter that a linear distributed system can be represented by an equation of the form

$$\frac{\partial y_i}{\partial t} + \alpha_i \frac{\partial y_i}{\partial x} + \beta_{ij}(y_i - y_j) = \lambda_i \frac{\partial^2 y_i}{\partial x^2}, i, j = 1, \dots, n \quad (2.40)$$

Each such system, however, requires to be treated along with its individual boundary conditions. Since basically the derivation of the

solution of equations of the above form is very complex, a generalized solution can hardly be attempted. Even for the cases where such complex representation is necessary as in heat exchangers, distillation apparatus, packed bed reactors, etc., equivalent lumped models are considered for convenience. For other simpler cases, an initial reduction in the above equation is possible. A typical example is that of a one-dimensional heat transfer problem through a solid material for which relation (2.40) may be reduced to

$$\frac{\partial y}{\partial t} = \lambda \frac{\partial^2 y}{\partial x^2} \quad (2.41)$$

y being the temperature at point x at time t and λ a constant representing thermal diffusivity. A solution of Eq. (2.41) may be attempted by applying the possible boundary conditions. By way of example, let us put the conditions as

$$y = y_0 \text{ at } x = 0$$

$$\partial y / \partial x = 0 \text{ at } x = L (\text{length, thickness, etc.})$$

which are quite in conformity with an idealized physical system of a thick tube. The solution of Eq. (2.41) is then obtained as

$$y = y_0 \frac{\exp(2L\sqrt{s/\lambda}) \exp(-x\sqrt{s/\lambda}) + \exp(x\sqrt{s/\lambda})}{\exp(2L\sqrt{s/\lambda}) + 1} \quad (2.42)$$

where s is the Laplace operator.

At $x = L$, Eq. (2.42) gives

$$y = 2y_0 \{ \exp(L\sqrt{s/\lambda}) + \exp(-L\sqrt{s/\lambda}) \}^{-1} \quad (2.43)$$

Hence across the length L , the function (equivalent transfer function) is obtained as

$$\frac{y_L(s)}{y_0(s)} = \frac{2}{\exp(L\sqrt{s/\lambda}) + \exp(-L\sqrt{s/\lambda})} = \frac{1}{\cosh \sqrt{\left(\frac{L^2 s}{\lambda}\right)}} \quad (2.44)$$

from which the magnitude and phase may be easily evaluated. For example, if $s = j\omega$ and a substitution is made as

$$\sqrt{\frac{\omega L^2}{2\lambda}} = \phi \quad (2.45)$$

then

$$\left| \frac{y_L}{y_0} \right| = (\cosh^2 \phi \cos^2 \phi + \sinh^2 \phi \sin^2 \phi)^{-\frac{1}{2}} \quad (2.46)$$

and

$$\angle \frac{y_L}{y_0} = -\tan^{-1}(\tan \phi \tanh \phi) \quad (2.47)$$

Not much of a simplifying assumption is made in the above deduction and it would appear that L^2/λ has the dimension of time. In fact, if the distributed system is made into an equivalent model with resistance and capacitance per unit length as shown in Fig. 2.11, one can write

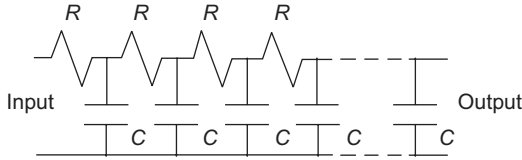


Fig. 2.11 Electrically analogous model of a distributed parameter system

$$L^2/\lambda = RC = \tau \quad (2.48)$$

Obviously then

$$\phi = \omega RC/2 = \omega\tau/2 \quad (2.49)$$

If each unit length is non-interacting with its adjacent ones, formulae (2.46) and (2.47) may be simplified to

$$A_r = |y_L/y_0| = (\cosh^2 \sqrt{\omega\tau/2} - \sin^2 \sqrt{\omega\tau/2})^{-\frac{1}{2}} \quad (2.50a)$$

and

$$y_L/y_0 = -\tan^{-1}(\tan \sqrt{\omega\tau/2} \tanh \sqrt{\omega\tau/2}) \quad (2.50b)$$

Further if $\omega \rightarrow \infty$, i.e., at high frequencies, amplitude ratio and phase are

$$A_r|_{\omega \rightarrow \infty} = \cosh \sqrt{\omega\tau/2} \quad (2.51a)$$

$$y_L/y_0|_{\omega \rightarrow \infty} = -\sqrt{\omega\tau/2} \quad (2.51b)$$

If now the frequency response is plotted it will be seen that at low frequencies the response in a distributed and lumped system with a total time constant is half the value of the unit length. At higher frequencies an exact solution following the usual procedure should be adopted.

In the above example not only the transfer function but also the solution of a typical system of distributed parameter type is given for understanding the problem in its proper perspective.

In a later chapter (Ch. 8) some more examples of process models and derivation of their transfer functions are given. These are given in connection

with typical control schemes of the common industrial variables such as flow, pressure, level and temperature and the processes are, therefore, marked as such. The readers may refer to these derivations for general understanding.

2.2.8 DC Motor

A last example of mathematical modelling and transfer function derivation via theoretical approach is from engineering industry—a dc motor. In many industries speed control of dc motors forms a unit of a plant and requires special consideration. A representative of this type of dc motor with load is shown in Fig. 2.12 in detailed form where applied field is e_f , field resistance and inductance are R_f and L_f , armature resistance and inductance are R_a and L_a with armature voltage and current being e_a and i_a ; J_m and B_m are motor inertia and coefficient of viscous friction for the motor shaft respectively, J_L and B_L are those for the load output shaft. θ_m and θ_L are motor shaft and output shaft rotation angles; n is the gear ratio. Such a motor can be controlled by a varying field with armature current constant or with a varying armature current keeping field voltage constant. In either case the air gap flux is proportional to field current, i_f and the developed torque in the motor shaft is proportional to the air gap flux and armature current. If field is controlled, armature current is constant. Thus

$$T(t) = K_1 \phi(t) i_a(t) = K_2 \phi(t) = K_3 i_f(t) \quad (2.52)$$

or

$$T(s) = K_3 I_f(s) = K_3 E_f(s) / (R_f + sL_f) \quad (2.53)$$

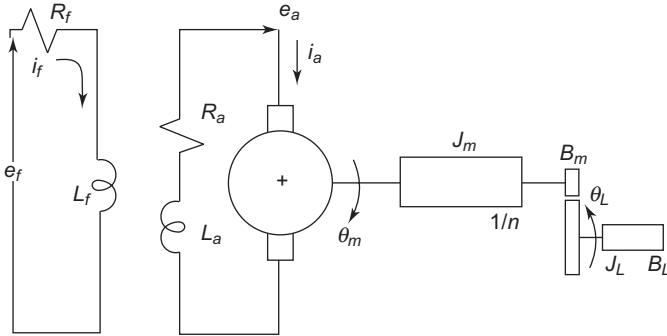


Fig. 2.12 The loaded dc motor

This torque developed must equal the torque demanded by the motor rotor and load. The motor rotor torque transform is $(J_m s^2 + B_m s) \theta_m(s)$ and the load torque on the motor shaft is $\frac{1}{n} (J_L s^2 + B_L s) \theta_L(s)$, so that, using again $\theta_m(s) = n \theta_L(s)$,

$$\frac{K_3 E_f(s)}{R_f + sL_f} = \left[(J_m s^2 + B_m s)n + \frac{1}{n}(J_L s^2 + B_L s) \right] \theta_L(s) \quad (2.54)$$

yielding the transfer function

$$\begin{aligned} \frac{\omega_L(s)}{E_f(s)} &= \frac{s\theta_L(s)}{E_f(s)} = \frac{K_3 n / \{R_f(B_L + n^2 B_m)\}}{(sL_f/R_f + 1) \left(\frac{s(J_L + n^2 J_m)}{B_L + n^2 B_m} + 1 \right)} \\ &= \frac{k_e}{(s\tau_1 + 1)(s\tau_2 + 1)} \end{aligned} \quad (2.55)$$

where $\omega_L(s)$ is the transform of the load shaft angular velocity.

2.3 LIMITATIONS ON PROCESS EQUATION FORMULATIONS

While considering the quantitative model, the general formulation is based on the nonlinear distributed system, a typical case of which is represented by

$$\frac{\partial x_i}{\partial t} + a_x \frac{\partial x_i}{\partial z} = \sum_{j=1}^n \phi_j(x_1, x_2, \dots, x_m, y) + b_x \frac{d^2 x}{dz^2}, i = 1, \dots, m \quad (2.56a)$$

$$\frac{\partial y}{\partial t} + a_y \frac{\partial y}{\partial z} = \sum_{j=1}^n c_j \phi_j(x_1, x_2, \dots, x_m, y) + b_y \frac{\partial^2 y}{dz^2} + N(y) \quad (2.56b)$$

There are, therefore, n simultaneous reactions with m varieties with the variables denoted by x_i and output y . ϕ_j is a nonlinear function and is different in different cases and $N(y)$ is another non-linear function of y . The quantities a , b and c are considered specific parameter values which should be known.

Typical linear distributed systems are often represented by the form

$$\frac{\partial \alpha_i}{\partial t} + a_i \frac{\partial \alpha_i}{\partial x} + b_{ij}(\alpha_i - \alpha_j) = c_i \frac{\partial^2 \alpha_i}{\partial x^2}, i, j = 1, \dots, n \quad (2.57)$$

Both equations (2.56) and (2.57) may be examples of packed bed reactors. It can be seen from Eqs (2.56) and (2.57) what the difference in complexity is and also how they can be adopted for system design. As already stated, the process complexity often makes it impossible to derive these equations and this is mainly due to the following reasons: (i) Enough of sufficiently high quality empirical data are not available for deriving the equations statistically; (ii) Many important uncontrollable variables are not known and cannot be properly cared for; (iii) Many controllable state and

product variables are variable over only a limited range; and (iv) There are physical limitations as well. Large lags and dead times arising due to small flow rates in large capacity vessel also limit controllability and prevent the formulation of a simple linear model of the process.

2.4 PROCESS MODELLING

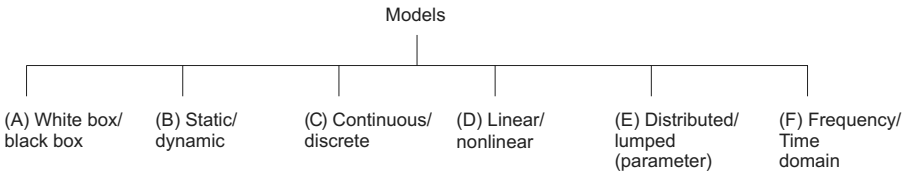
Modelling has been considered as an ‘abstraction of reality’. Modelling can help understand and explain observation made in a system. Besides it can help limit elaborate experimentation. Modelling has been defined as “a representation of the necessary and essential aspects of a system which can present knowledge of that state in a usable form.”

Usable approaches of modelling as mentioned already are (1) theoretical, and (2) experimental. Theoretical approach is by forming the mathematical equation (see Section. 2.2). For chemical processes, mass, energy and momentum balances are the basic principles. There are subsidiary and supporting relations for completing the models. In the experimental approach, only the inputs and outputs are considered which are then related by a suitable technique.

The model complexity varies depending on the application aspect of it. In general five aspects are segregated for this purpose. These are

- (1) Planning and scheduling
 - (2) Design
 - (3) Research and development
 - (4) Optimization (and operation)
 - (5) Control and prediction
1. The model may be simple, either static or dynamic, but operation time base must be large.
 2. Modelling should be able to yield the design parameters covering aspects of safety and economy which again must be compatible to each other. It is not that simple.
 3. Models developed for R and D works are initially prototype models and the data available from such models are ‘scaled up’ for developing a full-fledged system model. Initial parameters are generated by simulation or by measurement from associated types of process.
 4. This modelling process is related to design type but for optimization design, parameters are considered in optimization format often simplifying the model structure.
 5. Process control incorporating reference change or disturbance can be studied with models of individual process for which models of simple nature can be produced. Special models are developed for prediction of some variables even if indirectly.

Obviously, for the applications as listed above, types of models also vary. In fact, modelling is structurally based on simple considerations as listed in the following chart.



- (A) White-box models are based on laws and principles and are developed theoretically without any experimental data. Starting principles are laws of conservation and laws of physics for which such models are often called first principle models or mechanistic models.

In contrast, black-box models are based on input/output data of a process considered as a black box. In absence of any physical insight to develop these models the mathematical representations of these models are given in series forms. This procedure has received utmost recognition in recent times. In between the two types—that is white- and black-box types, a grey-box model is proposed where partially physical laws are known to be applicable but no clear-cut knowledge of the entire process is available. Process control (item 5) modelling is largely black-box type while *R* and *D* modelling (item 3) above is the white-box type. Biological systems or simple kinetic processes are modelled on grey-box philosophy. Series form modelling (black-box type) are many—some recent types are (a) Autoregressive moving average (ARMA) type, and (b) Autoregressive exogenous (ARX) type which will be briefly outlined later in the chapter. Fuzzy logic and artificial neuron network modelling are examples of grey-box modelling approach.

- (B) Models which are basically time independent and depend only on the recent values of the independent variables are called static models. These are useful to judge optimization and for representing continuous processes. When the independent variable are functions of time, the model is called dynamic. Most process control models are dynamic and these are useful for prediction work as well.
- (C) Models with continuous variables are continuous models and discrete models are the ones made with discrete variable—that is the variable values at given time intervals are considered. Modelling in the latter case is best done by *Z*-transform technique while for continuous type, Laplace transform technique is good enough. For obtaining the models, often the difference equations

are considered. With discrete step of time Δt , variable changes from y_k to y_{k+1} , so that one can start with the equation

$$dy/dt = (y_{k+1} - y_k)/\Delta t$$

- (D) Linear models are represented by functional equations which observe the superposition principle. As long as the operating range is limited around a given value, reasonably accurate modelling is done by linear equations. Nonlinear models are quite complicated and numerical methods are usually adopted for the description of such systems. Solution softwares are available for such systems. A very common software is MATLAB, gPROMS is also used conveniently.
- (E) When the independent variables vary in space, distributed parameter models are of consideration. In chemical engineering, tubular reactor is such a system where there is variation of the variable along the axis of the reactor. Besides, there are variations in the radial directions as well making the model very complicated. Often the process is divided into small segments/sections and it is considered that over these small sections, the variable properties remain constant and a lumped parameter modelling is made with these small segments. Depending on the process, some criteria are evolved for making this approximation.
- (F) The reference variable is generally 'time' in process modelling and then the modelling is in the time domain. However, alternatively, analysis can be done in frequency domain and modelling also can be done with frequency as the reference variable. In process control frequency domain, analysis is not common and if at all analysis is done in that domain, appropriate transform may be used to convert the result in time domain for real-time performance analysis.

2.4.1 Model Development

While developing or building the model of a system/process, consideration is given to the fact that whether this can be done either as a mechanistic one, i.e. it can be built using the first principle or not. The next consideration is if this can be considered as a lumped parameter model or distributed type, preference goes for the former one because of its simplicity, both in formulation and in analysis. Then comes the checking for the linearity, if not, move to nonlinear type mode. Other important considerations are (1) how much different would it be from the ideal one, i.e. its accuracy and thence its utility in real 'life' study, (2) the model needs to be verified and evaluated by appropriate measurements; how far is this possible to be made. Also if the model appears to be too complex, there is a possibility of dividing this into subsystems (submodels) keeping provision for convenience in analysis

and fault diagnosis. A very general procedure for building a model is given in the block diagram of Fig. 2.13.

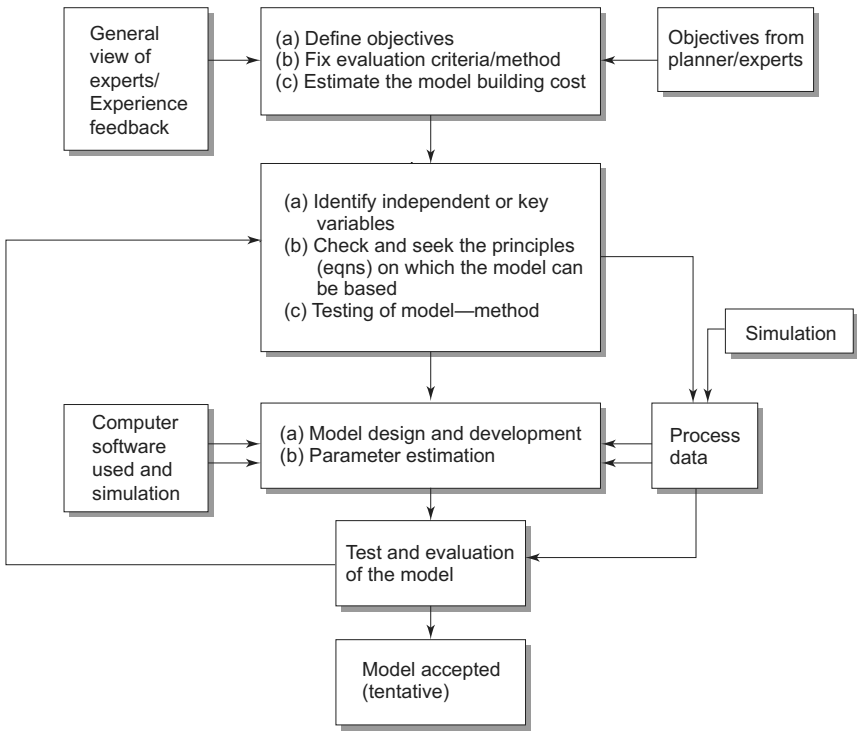


Fig. 2.13 General procedure of model building

There are three distinct steps in the modelling process. The problem statement/definition and available resources as also their identification; the design part with available process data (or simulated data) and required computer software, parameter estimation is included in this part and finally the model verification, test and evaluation as stipulated resulting in its validation.

Model is actually a representation of the system in mathematical form—like equation. In practice, algebraic equation is used to model a lumped parameter steady state model, difference equation is used for discrete system or discretized system and integral/differential equation is used for continuous process. In general, continuous process are dominantly predominant and for this modelling procedure, in so far as equational representation is concerned, divisions can be made starting with differential equations as main propositions. Figure 2.14 shows a brief chart of this division.

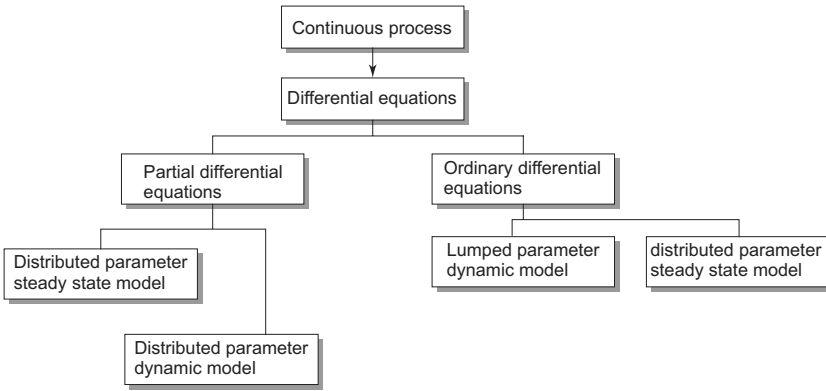


Fig. 2.14 Chart showing guiding equations of models

Some examples of modelling using test data as also identification of parameters are given in the next section.

2.5 PROCESS MODELLING VIA EXPERIMENTAL TESTS

Development of dynamic mathematical models of some specific process systems has been considered above from which the transfer functions have been derived. The approach has been through a theoretical analysis. In the following the experimental approach is presented for the development of the mathematical model and identification of the process parameters. In this approach a simple change in the input is introduced into the process and output response is recorded, from which, by data analysis the approximate process transfer function is obtained.

Specifically, the process is modelled from step-input test curve and data therefrom. In this a step disturbance is given to the process operating under steady state conditions with the controller in manual position (or the system in open loop condition) and the transient response is recorded in an appropriate recorder. During the test there should not be other disturbances like load upsets, etc. The resulting curve known as the process reaction curve, is now under consideration for modelling.

A typical process reaction curve without dead time is shown in Fig. 2.15 (for definition of dead time see Chapter 3) and is considered first. From the response, $c(t)$, we now write the per unit incomplete response $1 - c(t)/u(t) = z(t)$ as the ordinate variable. Here $u(t)$ is the step input given. For the first order systems one obtains

$$z(t) = \exp(-t/\tau) \quad (2.58)$$

Transforming, this can be written as

$$\ln z(t) = -t/\tau \quad (2.59)$$

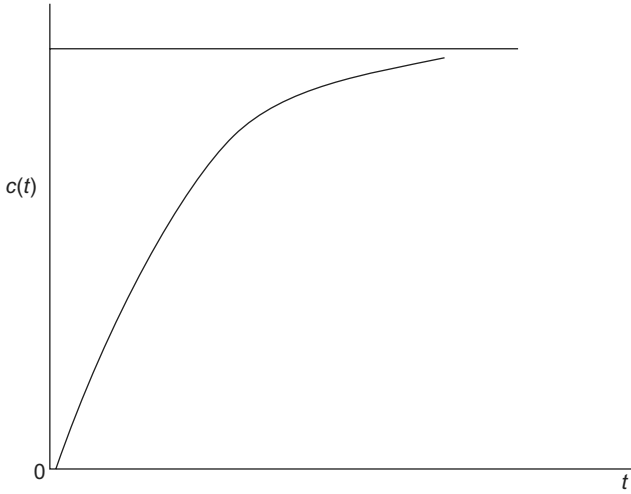


Fig. 2.15 A process reaction curve without dead time

Thus with values of $z(t)$ for different t a plot of Eq. (2.59) is obtained as shown in Fig. 2.16, the slope of the curve is the reciprocal of the time constant of the system. The transfer function of the system is then obtained as

$$T_1(s) = 1/(s\tau + 1)$$

For a second order system with time constants τ_1 and τ_2 the time response in $z(t)$ may be written as

$$z(t) = \frac{\tau_1}{\tau_1 - \tau_2} \exp(-t/\tau_1) - \frac{\tau_2}{\tau_1 - \tau_2} \exp(-t/\tau_2) \quad (2.60)$$

With $z(t)$ along the ordinate in log scale and t along abscissa in linear scale we now get a curve as shown in Fig. 2.17, curve 1. $z(t)$ is obtained from the experimental curve of Fig. 2.15 as stated earlier. If it were a first order system the curve would have been a straight line (Fig. 2.16) with, perhaps, the first term of Eq.(2.60), so that a linear curve, curve 2, asymptotically constructed would cut the ordinate at 'b' (say) which is the zero time ordinate and has a value $\log [\tau_1/(\tau_1 - \tau_2)]$. The difference of the above two curves also gives another straightline (curve 3) which has the zero time ordinate as 'c' = 'b' - 'a' and is thus equal to $\log \left(\frac{\tau_2}{\tau_1 - \tau_2} \right)$. By a simple

calculation now τ_1 and τ_2 can be evaluated, or else, at $t = \tau_1$, response is 63.2 per cent of 'b' for curve 2 while for curve 3, at $t = \tau_2$ response is 63.2 per cent of 'c'. τ_1 and τ_2 are thus evaluated and the transfer function is written as

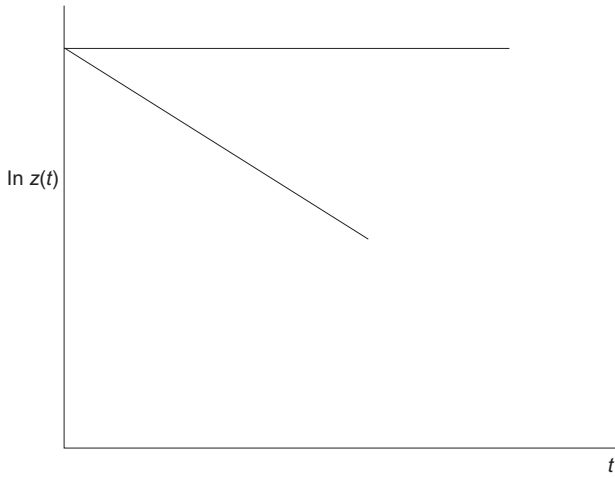


Fig. 2.16 Transformed plot of the process reaction curve of a first order process

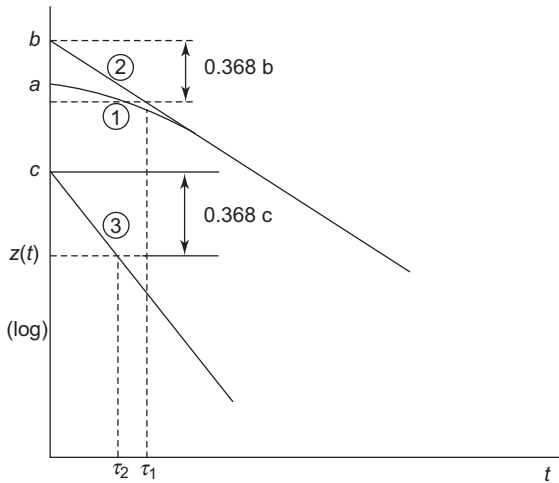


Fig. 2.17 Transformed plots of the process reaction curve of a second order process

$$T_2(s) = k / [(s\tau_1 + 1)(s\tau_2 + 1)]$$

For a second order system with monotonic response, the method of Oldenbourg and Sartorius is by identifying two times, T_α which is the projection on the asymptote of the segment of the tangent at the point of inflexion of the response curve which is included between time axis and asymptote and T_β is the projection on the asymptote of the segment between the point of inflexion and asymptote as shown in Fig. 2.18. Considering the response equation as given by Eq. 2.60, one can find T_α and T_β in terms of τ_1 and τ_2 of Eq. 2.60 as

$$T_{\alpha} = \tau_1 \left(\frac{\tau_2}{\tau_1} \right)^{\frac{\tau_2}{\tau_2 - \tau_1}} \quad (2.61)$$

and

$$T_{\beta} = \tau_1 + \tau_2 \quad (2.62)$$

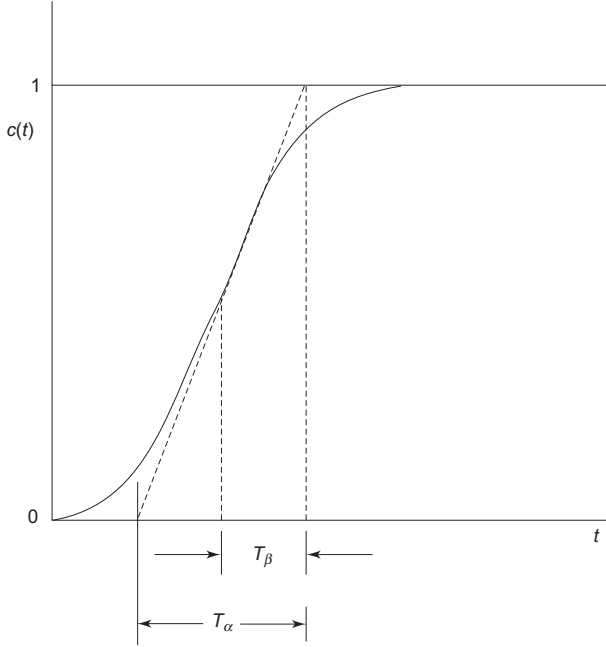


Fig. 2.18 Curves for identifying time T_{α} and T_{β} in the process reaction curve

which are normalized with T_{α} as

$$\left(\frac{\tau_1}{T_{\alpha}} \right) \left(\frac{\tau_2/T_{\alpha}}{\tau_1/T_{\alpha}} \right)^{\frac{\tau_2/T_{\alpha}}{\tau_2/T_{\alpha} - \tau_1/T_{\alpha}}} = 1 \quad (2.63)$$

and

$$\frac{T_{\beta}}{T_{\alpha}} = \frac{\tau_1}{T_{\alpha}} + \frac{\tau_2}{T_{\alpha}} \quad (2.64)$$

As T_{β}/T_{α} is known from the response record both τ_1 and τ_2 can be found from the above equations. Plots in coordinates of τ_2/T_{α} and τ_1/T_{α} of Eqs (2.63) and (2.64) are made in Fig. 2.19 such that the curved plot is the plot of Eq. (2.63) and the parallel straight lines are the plots of τ_{β}/T_{α} . The two intersection points of the two plots indicate the values of τ_1/T_{α} and τ_2/T_{α} from which τ_1 and τ_2 are found and the transfer function written.

For an oscillatory response without dead time, one has

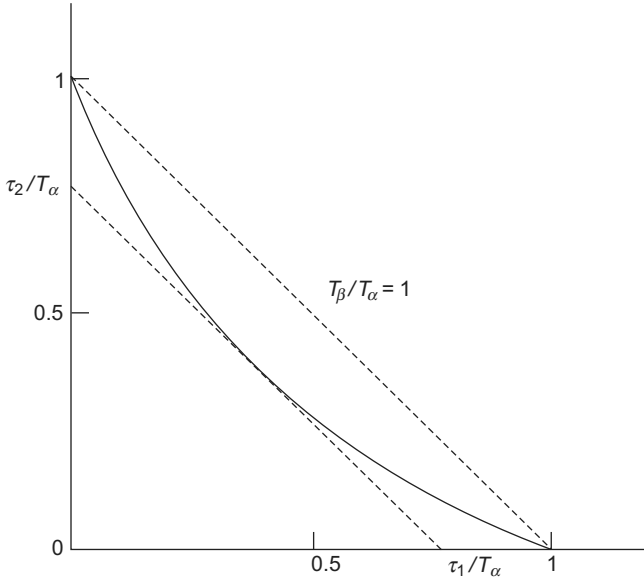


Fig. 2.19 Plots of τ_1/T_α versus τ_2/T_α for varying T_β/T_α

$$z(t) = \frac{\exp(-\zeta\omega_n t)}{\sqrt{1-\zeta^2}} \sin(\sqrt{1-\zeta^2}\omega_n t + \phi) \quad (2.65)$$

where ζ = damping constant and ω_n = natural frequency of oscillation and $\phi = \sin^{-1}\sqrt{1-\zeta^2}$. From this time t_p as recorded in the PRC, for the first peak, is given by

$$t_p = \frac{\pi}{\omega_n \sqrt{1-\zeta^2}} \quad (2.66)$$

Also the peak overshoot $\lambda = c_m(t) - u(t)$ is given as

$$\lambda/u(t) = \exp(-\pi\zeta/\sqrt{1-\zeta^2}) \quad (2.67)$$

from which

$$\zeta = \frac{1}{\sqrt{\pi / \left(\ln \frac{\lambda}{u(t)} \right)^2 + 1}} \quad (2.68)$$

Using this ζ , ω_n is evaluated as

$$\omega_n = \frac{\pi}{t_p \sqrt{1 - \zeta^2}} = \frac{2\pi}{T \sqrt{1 - \zeta^2}} \quad (2.69)$$

where T = time period of oscillation.

With dead time the PRC changes predominantly at the starting side. An extension of the linear part downward as shown in Fig. 2.20 identifies the dead time and the rest of the curve on the right-hand side is treated as above.

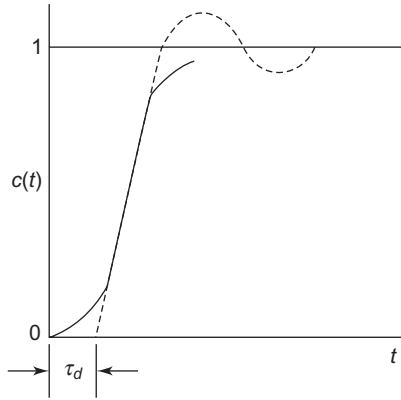


Fig. 2.20 Process reaction curve with construction for identifying process dead time

Effectively this means shifting the origin of the curve to $(\tau_d, 0)$ for the portion which does not have a dead time. Many people have tried to workout the model of a system that has a dead-time but the attempt due to Sundaresan et al, is described here as it takes care of the uncertain location of the inflexion point on the PRC for drawing the slope necessary for parameter evaluation. The model basically is considered to have two process times and a dead time, i.e., a second order model with a dead time and the process may be overdamped or underdamped so that it can be represented either by

$$G_p(s) = \frac{\exp(-s\tau_d)}{(1 + s\tau_1)(1 + s\tau_2)} \quad (2.70)$$

or by

$$G_p(s) = \frac{\exp(-s\tau_d)}{s^2/\omega_n^2 + \frac{2\zeta s}{\omega_n} + 1} \quad (2.71)$$

The parameters τ_d , τ_1 , τ_2 in the former case and τ_d , ω_n , ζ in the latter are to be determined from the process reaction curves.

Considering first the overdamped case with its PRC drawn in Fig. 2.21, the shaded area is given by the equation

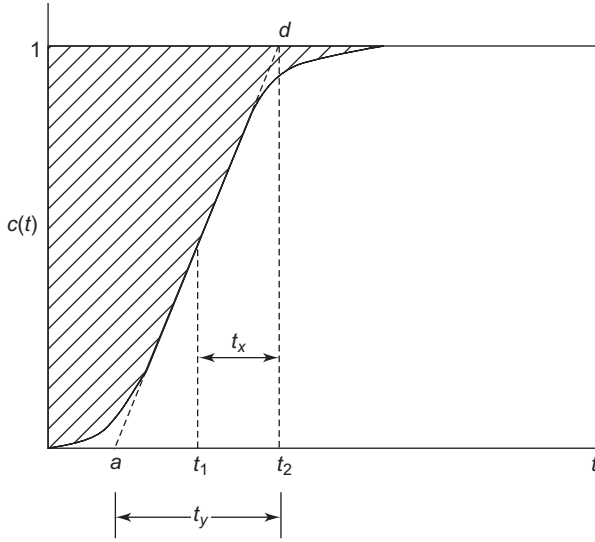


Fig. 2.21 Plot of fraction $c(t)$ versus t for an overdamped system

$$m_1 = \int_0^{\infty} (1 - c(t)) dt \quad (2.72)$$

With the response in fractional form, as shown, m_1 is also expressible as

$$m_1 = -dG_p(s)/ds|_{s=0} = \tau_d + \tau_1 + \tau_2 \quad (2.73)$$

For a step change the time response solution of Eq. (2.70) is given by

$$c(t) = \left[1 - \frac{\tau_1}{\tau_1 - \tau_2} \exp(-(t - \tau_d)/\tau_1) + \frac{\tau_2}{\tau_1 - \tau_2} \exp(-(t - \tau_d)/\tau_2) \right] u(t - \tau_d) \quad (2.74)$$

The point of inflexion is found by taking double derivative of the above equation and equating it to zero. Thus at inflexion point, time t_1 is

$$t_1 = \tau_d + \lambda \ln \beta \quad (2.75)$$

where

$$\lambda = \frac{\tau_1 \tau_2}{\tau_1 - \tau_2} \quad (2.76)$$

and

$$\beta = \tau_1/\tau_2 \quad (2.77)$$

The slope at the point of inflexion is (i.e., $dc(t)/dt|_{t=t_1}$)

$$S_i = \frac{\frac{1}{\beta^{1-\beta}}}{\lambda(\beta-1)} \quad (2.78)$$

which is the slope of the tangent at time t_1 of the $c(t)$ curve and this tangent cuts the final value at time t_2 given by (point d in the figure)

(i.e., $c(t) = 1$)

$$t_2 = \tau_d + \lambda \left[\ln \beta + \frac{\beta^2 - 1}{\beta} \right] \quad (2.79)$$

Combining Eqs (2.73), (2.78) and (2.79), one gets

$$(t_2 - m_1)S_i = \frac{\beta^{\beta/(1-\beta)}}{\beta - 1} \ln \beta \quad (2.80)$$

In Eq. (2.80), putting $\beta = \frac{1}{\tau_2/\tau_1}$, no change is obtained and hence $\tau_1/\tau_2 = \beta$ needs be considered in the range 0 to 1. Now writing

$$(t_2 - m_1)S_i = \alpha = A \exp(-A) \quad (2.81)$$

A is given as

$$A = \ln \beta / (\beta - 1) \quad (2.82)$$

This makes the maximum values of α as $\exp(-1)$ which occurs when the system is critically damped, i.e., $\beta = 1$ or $A = 1$. For overdamped system $\beta < 1$ and hence the value of A lies between 0 and $\exp(-1)$. When $\beta = 0$ and hence $\alpha = 0$, the model is that of a first order system. For a true second order system the value of α should lie between 0 and $\exp(-1)$. A plot of α vs β is shown in Fig. 2.22.

For evaluating τ_1 , τ_2 and τ_d the parameters of the model, Fig. 2.21 is constructed as shown. By numerical integration now the shaded area is obtained. Then a tangent through the linear portion, i.e., joining 'ad' is obtained and from this straight line t_2 as also S_i are obtained. From Eq. (2.81) now α is calculated and using Fig. 2.20, β is obtained. Using these values λ is obtained from Eq. (2.78). Equations (2.73), (2.76) and (2.77) can then be used to obtain τ_1 , τ_2 and τ_d as

$$\tau_1 = \frac{\beta^{\beta/(1-\beta)}}{S_i}, \tau_2 = \frac{\beta^{1/(1-\beta)}}{S_i}, \text{ and} \quad (2.83)$$

$$\tau_d = m_1 - \frac{\beta^{1/(1-\beta)}}{S_i} \left(\frac{1+\beta}{\beta} \right)$$

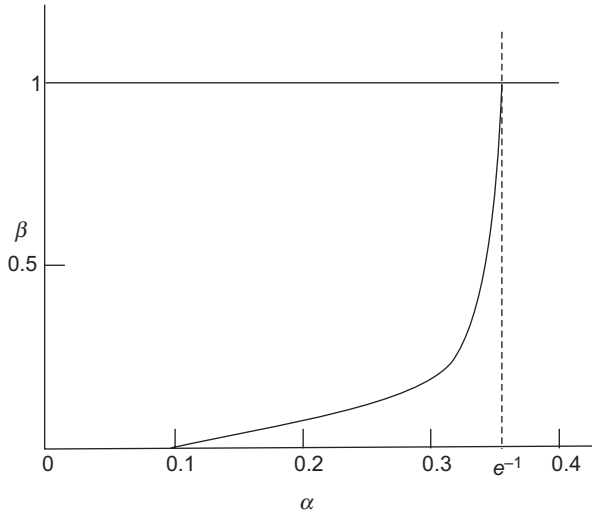


Fig. 2.22 Plot of $\alpha - \beta$ curve

For an underdamped system $\alpha > 1/e$. The time response characteristic curve of such a system is shown in Fig. 2.23. For a step input, its time domain response is given by the equation

$$c(t) = [1 - \exp(-(t - \tau_d)\omega_n\zeta)\{\frac{\zeta}{\sqrt{1-\zeta^2}}\sin(\omega_n(t - \tau_d)\sqrt{1-\zeta^2}) + \cos(\omega_n(t - \tau_d)\sqrt{1-\zeta^2})\}]u(t - \tau_d) \quad (2.84)$$

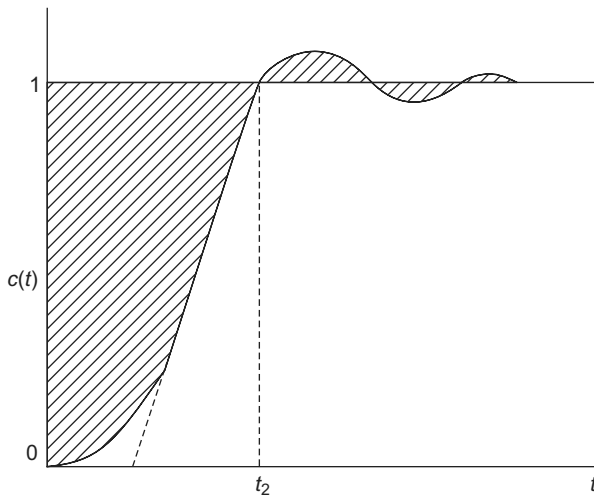


Fig. 2.23 Plot of fractional $c(t)$ versus t for an underdamped system

A procedure similar to the above is now followed with the exception that the portions hatched below $c(t) = 1$ would be positive and above would be negative. This procedure yields

$$\alpha = (t_2 - m_1)S_i = \frac{\cos^{-1} \zeta}{\sqrt{(1 - \zeta^2)}} \exp \left(- \left(\frac{\zeta}{\sqrt{1 - \zeta^2}} \cos^{-1} \zeta \right) \right) \quad (2.85)$$

$$\omega_n = \frac{\cos^{-1} \zeta}{\sqrt{(1 - \zeta^2)}} \frac{1}{t_2 - m_1} \quad (2.86)$$

and

$$\tau_d = m_1 - 2\zeta/\omega_n \quad (2.87)$$

Equation (2.85) provides infinite number of roots for ζ for a given value of α . However in the region of interest $0 \leq \zeta \leq 1$, α is a monotonic function of ζ decreasing from $\pi/2$ at $\zeta = 0$ to $1/e$ at $\zeta = 1$. The schematic plot is shown in Fig. 2.24. For the numerical integration a digital computer would be needed and this makes the evaluation easy. For critically damped case $\beta = 1$, one gets from Eq. (2.84)

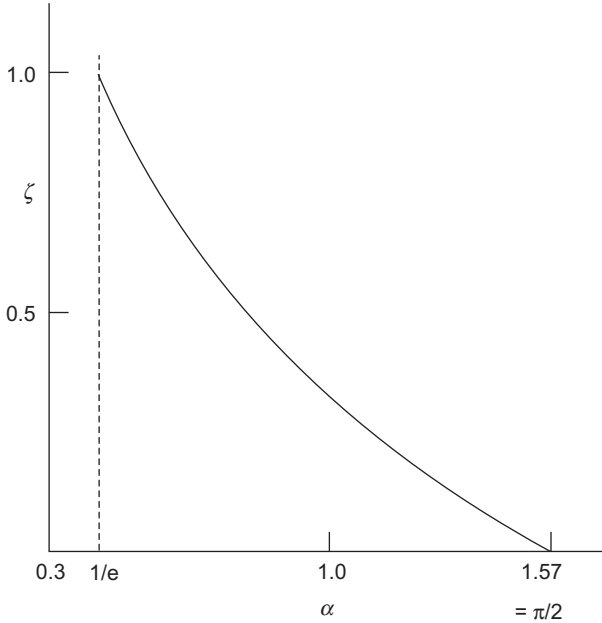


Fig. 2.24 Values of ζ for varying values of α

$$\tau_1 = \tau_2 = \tau = 1/(S_i e),$$

and

$$\tau_d = m_1 - 2/(S_i e) \quad (2.88)$$

2.6 DISCRETE MODELLING

With computers helping to control process, discretization of process equations is also standard practice now. Continuous processes (or even batch type processes) are modelled as lumped dynamic type which can be used for optimization as well. Such models have two different classes:

(1) parametric models, and (2) nonparametric models.

In parametric forms, equations can be written expressing a set of quantities which are explicit functions of a number of independent variables called parameters and therefore such models require more or less correct informations about the inner structure representable by a finite (limited) number of parameters. In other words, process knowledge is necessary for framing such a model and effectively these turn out to be white-box models.

In contrast, for nonparametric models, information about the inner structure is not known, neither is the order of the system stated *a priori*. There can be any number of parameters. However, the time span is to be known. These are effectively black-box models and are formed from experimental data.

A process with input $u(k)$, output $y(k)$ and disturbance $\delta(k)$, where k represents the discrete time value is represented as

$$y(k) = G(z^{-1})u(k - \gamma) + \delta(k) \quad (2.89)$$

where γ is the time delay in the process, z^{-1} is the backshift operator representable as

$$y(k - 1) = z^{-1}y(k)$$

In process control system, the disturbance is often describable by filtered white noise so that

$$\delta(k) = H(z^{-1})e(k)$$

where $e(k)$ is white noise with a variance, say λ , and hence Eq. (2.89) can be written as

$$y(k) = G(z^{-1})u(k - \gamma) + H(z^{-1})e(k) \quad (2.90)$$

The generalized form of Eq. (2.90) is

$$A(z^{-1})y(k) = [B(z^{-1})/D(z^{-1})]u(k - \gamma) + [C(z^{-1})/F(z^{-1})]e(k) \quad (2.91)$$

where the polynomials A, B, C, D, F are of the form given by

$$A(z^{-1}) = 1 + a_{1a}z^{-1} + a_{2a}z^{-2} + \cdots + a_{na}z^{-na}, \text{ etc.} \quad (2.92)$$

The generalized structures have different special cases and are given different names depending on their development (functional). The names

AR standing for autoregressive, MA standing for moving average, X standing for exogenous or extra input are used mostly in combination. The models encountered are ARMA, ARMAX, ARX. There is one known as Box–Jenkins structure.

The generalized difference equation models of the above types are represented as

ARMA(p, q)

$$y(k) = \sum_{i=1}^q \alpha_i y(k-i) + \sum_{j=1}^p \beta_j e(k-j) \quad (2.93)$$

ARMAX(p, q, n)

$$y(k) = \sum_{i=1}^q \alpha_i y(k-i) + \sum_{j=0}^p \beta_j e(n-j) + \sum_{m=0}^n r_m u(k-m) \quad (2.94)$$

ARX(q, n)

$$y(k) = \sum_{i=1}^q \alpha_i y(k-i) + \sum_{m=0}^n r_m u(k-m) + e(k) \quad (2.95)$$

and the Box–Jenkins model is

B J(n, p)

$$y(k) = \sum_{n=0}^n r_m x(k-m) + \sum_{j=0}^p \beta_j e(n-j) \quad (2.96)$$

In Eq. (2.93), the first summation term on the RHS along with $e[k]$ represents AR while the second term represents MA.

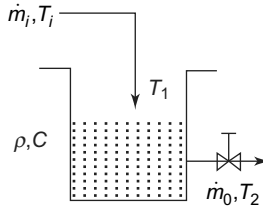
In this context, difference-equation representation of a very simple process of tank-filling with unconstrained inlet flow and constrained outlet flow as shown in Fig. 2.25 may be considered. The inlet mass flow rate is $\dot{m}_i$ and temperature T_i , tank volume is V , liquid density ρ , liquid specific heat C . Temperature T_1 at liquid level is taken different from T_i as the inlet line is considered very long and T_i is the initial temperature. Outlet temperature is T_2 and mass flow rate out is $\dot{m}_0$.

The energy balance equation is

$$C\rho \frac{d(VT_2)}{dt} = C\dot{m}_i T_1 - C\dot{m}_0 T_2 \quad (2.97)$$

The mass balance equation is

$$\rho \frac{dV}{dt} = \dot{m}_i - \dot{m}_0 \quad (2.98)$$

**Fig. 2.25** Tank as a process

If V is constant $\dot{m}_i = \dot{m}_0 = \dot{m}$

Also, from Eq. (2.98),

$$\rho V \frac{dT_2}{dt} = \dot{m}(T_1 - T_2) \quad (2.99)$$

or,

$$\tau \frac{dT_2}{dt} = (T_1 - T_2) \quad (2.100)$$

where $\tau = \frac{\rho V}{\dot{m}}$ = residence time in the tank.

Equation (2.100) can be discretized as

$$\tau \frac{T_{2,k} - T_{2,k-1}}{\Delta t} = T_{1,k-1} - T_{2,k} \quad (2.101)$$

The time index of T_1 is $k - 1$ for the system to be physically realizable. Rearranging, one gets

$$\begin{aligned} T_{2,k} &= \frac{\tau}{\tau + \Delta t} T_{2,k-1} + \frac{\Delta t}{\tau + \Delta t} T_{1,k-1} \\ &= \frac{\tau}{\tau + \Delta t} T_{2,k-1} + \left(1 - \frac{\tau}{\tau + \Delta t}\right) T_{1,k-1} \end{aligned} \quad (2.102)$$

If $\Delta t \ll \tau$, as τ is the process lag and Δt is derivative time interval, and substituting $K_1 = \frac{\tau}{\tau + \Delta t}$, a constant,

$$T_{2,k} = K_1 T_{2,k-1} + (1 - K_1) T_{1,k-1} \quad (2.103)$$

As Δt is very small, $K_1 = e^{-\Delta t/\tau}$, so that Eq. (2.103) can be written as

$$T_{2,k} = e^{-\Delta t/\tau} T_{2,k-1} + \left(1 - e^{-\Delta t/\tau}\right) T_{1,k-1} \quad (2.104)$$

Using generalized notations, output y , input u ,

$$y_k = K_1 y_{k-1} + K(1 - K_1) u_{k-1} \quad (2.105)$$

where K is a constant to take account of the modelling discrepancy. If we use elegant Z-transform, specially the backward shift operator ($z^{-1} = e^{-sT}$) that makes $y_{k-1} = z^{-1} y_k$ as expressed earlier, then Eq. (2.105) is

$$y_k = K_1 z^{-1} y_k + K(1 - K_1) z^{-1} u_k \quad (2.106)$$

From Eq. (2.91), these models can be obtained by choosing the polynomial $A, B, C, \dots$, etc., properly in the backshift operator horizon. Thus in ARX structure

$$C(z^{-1}) = D(z^{-1}) = F(z^{-1}) = 1$$

giving

$$A(z^{-1})y(k) = B(z^{-1})u(k - \gamma) + e(k) \quad (2.107)$$

ARMAX structure is

$$A(z^{-1})y(k) = B(z^{-1})u(k - \gamma) + C(z^{-1})e(k) \quad (2.108)$$

i.e.,

$$D(z^{-1}) = F(z^{-1}) = 1$$

Similarly, for B.J. structure $A(z^{-1}) = 1$ giving

$$y(k) = \frac{B(z^{-1})}{D(z^{-1})} u(k - \gamma) + \frac{C(z^{-1})}{F(z^{-1})} e(k) \quad (2.109)$$

MATLAB is a very convenient 'software' for identifying the parametric models. Identification tool box with theta format is specially suited for the purpose.

Nonparametric models do not assume any structure *a priori*. The modelling procedure considers step weights of the input-output relationship. This requires more number of parameters to be identified as the number of step weights may be many for proper modelling. With the input-output relationship available as shown in Fig. 2.26 where discrete steps of magnitudes are known with inputs, the relationship can be mathematically represented as

$$y_k = \sum_{j=1}^{\infty} a_j \Delta u_{k-j} \quad (2.110)$$

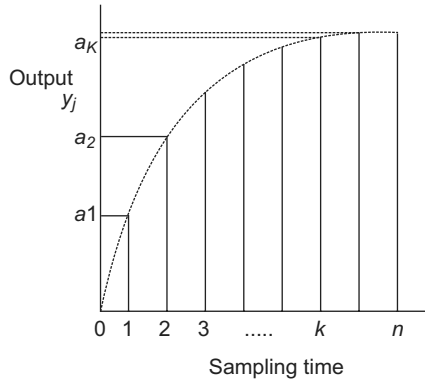


Fig. 2.26 Sampled output

The infinite series is shortened by taking difference. Thus, at the first stage, one writes

$$y_k = a_1 \Delta u_{k-1} + a_2 \Delta u_{k-2} + a_3 \Delta u_{k-3} + a_{ss} \Delta u_{k-4} + a_{ss} \Delta u_{k-5} + \text{bias} \quad (2.111)$$

where a_{ss} is the steady state coefficient and bias represent the remaining terms summed up. Again,

$$y_{k-1} = a_1 \Delta u_{k-2} + a_2 \Delta u_{k-3} + a_3 \Delta u_{k-4} + a_{ss} \Delta u_{k-5} + \text{bias} \quad (2.112)$$

Subtracting Eq. (2.112) from Eq. (2.111), one gets

$$\begin{aligned} y_k &= y_{k-1} + a_1 \Delta u_{k-1} + (a_2 - a_1) \Delta u_{k-2} + (a_3 - a_2) \Delta u_{k-3} \\ &\quad + (a_{ss} - a_3) \Delta u_{k-4} \\ &= y_{k-1} + b_1 \Delta u_{k-1} + b_2 \Delta u_{k-2} + b_3 \Delta u_{k-3} + b_4 \Delta u_{k-4} \end{aligned}$$

Step weights are now replaced by impulse weights b_j s and the number of parameters also becomes limited, making the estimation easy.

2.7 SCALE MODELLING TECHNIQUE

In spite of the fact that a very hazy picture is often obtained by modelling a process, it is popular because of the nonexistence of a suitable alternative. Modelling is known to aid process designers because it is (i) a research tool, (ii) a cheap and reliable method of studying processes for improving operation, and (iii) suitable for a large-scale design procedure. Different approaches to modelling have been tried and it has now been agreed upon that a representative model of a process can be obtained when the model theory and similarity principle are applied. In a particular process, the

phenomenon involved should be considered as the primary factor. For example, heat transfer is essentially a thermal process. While obtaining a representative model, it should be ensured that this model and the actual process (large-scale process) should perform similarly at least with respect to the essential phenomenon. The states maintained by the actual process and the scaled (down) model with respect to the particular phenomenon are known by the corresponding similarity states. A few of the important types of states are:

- (i) Shape of geometrical similarity state,
- (ii) Mechanical similarity state,
- (iii) Thermal similarity state, and
- (iv) Chemical similarity state.

Each of these states has its substates. For example, the mechanical similarity state has (a) static, (b) dynamic, and (c) kinematic states.

The above similarity principles are briefly discussed here for a basic understanding of the procedure of modelling via the model theory.

(i) Geometrical similarity

Two bodies are said to be geometrically similar if the points on one body have a one to one correspondence with points on another body. The ratio of corresponding lengths in the two bodies is known as the linear scale factor. This may vary axiswise when the area and volume scale factors are given as $n_1 \times n_2$ and $n_1 \times n_2 \times n_3$ respectively.

(ii) Mechanical similarity

- (a) Static** Two bodies are said to have static similarity if, when both of them are stressed, they maintain geometric similarity even after deformation. If the displacement ratio developed is given by, say m , then the forces acting are given, in elastic bodies, by

$$F_1/F_2 = m^2 \text{ or } 1/m^2 \quad (2.113)$$

In elastic bodies, if stresses are S_1 and S_2 and Young's moduli are Y_1 and Y_2 , then

$$S_1/S_2 = Y_1/Y_2 \quad (2.114)$$

However, for plastic bodies the force-deformation is related as

$$F_1/F_2 = (x_1/x_2)m^2 \text{ or } (x_1/x_2)/m^2 \quad (2.115)$$

the x 's being suitable deformation variables.

- (b) Dynamic** In the dynamic similarity cases, the forces inducing acceleration or retardation in moving masses are also considered.
- (c) Kinematic** In the kinematic similarity an additional dimension of time is introduced in the formalism developed for the static case.

If two systems trace geometrically similar paths in corresponding intervals of time the two systems are said to be kinematically similar.

(iii) Thermal similarity

Besides length, force and time it introduces the additional parameter, temperature. Heat flows in a process by three different ways—conduction, convection and radiation—and the rate of heat transfer is thus proportional to T or ΔT directly or functionally. In the same process other phenomena involving bulk movement of matter may be present when similarity in flow pattern—both dynamic and kinematic for example—is required. Several similarity criteria thus get interlinked and the case becomes reasonably complex.

(iv) Chemical similarity

In a continuous process if chemical composition varies from point to point maintaining a one-to-one correspondence in two systems, then these two systems are said to be chemically similar. In batch processes the composition similarity involves instant-to-instant one-to-one correspondence. Chemical concentration depends on temperature and bulk transportation implying that all chemically similar systems must be geometrically and thermally similar and in a flowing process kinematic similarity is also necessary.

From what has been said above it will be seen that scale modelling of an actual process for design purposes is quite involved and costly. In many instances, however, mathematical modelling, following the required similarity principles, can be attempted with success. It is easier to simulate the state mathematical models in an analogue computer directly. However, the points in favour of the use of a scale-model, as listed below, are worth considering.

- (i) Data obtained from the scale models (i.e., small scale plants) are more reliable and accurate than those obtained by simulation straight away.
- (ii) Widely different conditions are easily and quickly inducible in the models.
- (iii) The pilot plant practice for a new system yields convincing test results.
- (iv) Models may be interconnected when necessary and actual interdependence is reflected in the test results.

It should be remembered here that scale-modelling of processes for test results can be afforded by big companies as the cost involved is high. When proficiency is more important than the question of money, the R&D section of established houses may resort to such primary stages in plant-design as would be beneficial to everybody concerned.

2.8 PROCESS MODEL FROM FREQUENCY RESPONSE STUDIES

Frequency response method may be used for finding τ for a first order system or τ_1 and τ_2 or ζ and ω_n for a second order overdamped and underdamped systems respectively. In fact description of the dynamic behaviour in terms of frequency response can be made suitably for systems of arbitrary orders. After obtaining the amplitude ratio and phase curves for a wide range of frequency (given as sinusoidal input to the open loop system) one can obtain the sinusoidal transfer function in numerical form via curve fitting. The straightline asymptotes of the amplitude ratio plot can be quite conveniently used to obtain the corner frequencies and hence the time constants and natural frequencies, particularly when they are not very close to each other in the frequency axis. Figures 2.27(a), (b) and (c) show the techniques of evaluation of the parameters for an overdamped, overdamped/underdamped and an underdamped case respectively. While the evaluations are quite straightforward in the cases (a) and (b) of Fig. 2.27, computations of ζ and ω_n for the case of (c) are done from the ratio of peak amplitude ratio of A_p to the zero frequency amplitude ratio A_0 and from ω_p as

$$1/(2\zeta\sqrt{1-\zeta^2}) = A_p/A_0$$

or

$$\zeta = \sqrt{\left[\frac{1 \pm \sqrt{1 - (A_0/A_p)^2}}{2} \right]} \quad (2.116a)$$

and

$$\omega_n = \frac{\omega_p}{\sqrt{1 - 2\zeta^2}} \quad (2.116b)$$

However, frequency response testing by providing sinusoidal input to the process is not only very expensive and time consuming but also often improbable in many situations. For this reason short duration imperfect impulse testing or pulse testing technique is used to obtain the frequency response diagrams via the Fourier transform method of test data analysis. The method is briefly given as follows.

The process under the steady state conditions is subject to a pulse (or an impulse of very small but finite amplitude and duration) of arbitrary shape and the transient response is recorded. From the record of both the input and output time domain data, the frequency response diagrams are obtained. The output $y(t)$ in the time domain is converted to frequency domain as

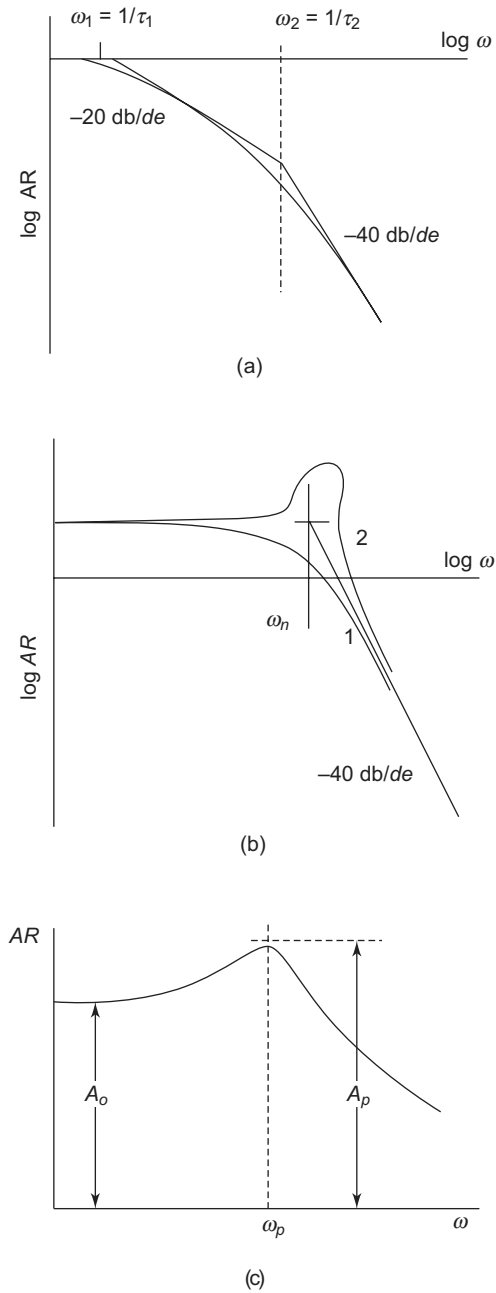


Fig. 2.27 Frequency response plots to evaluate system parameters: (a) overdamped case (b) common case (c) underdamped case

$$y(j\omega) = \int_0^{t_0} y(t) \exp(-j\omega t) dt = \int_0^{t_0} y(t) \cos(\omega t) dt - j \int_0^{t_0} y(t) \sin(\omega t) dt = a - jb \text{ (say)} \quad (2.117)$$

where t_0 = time interval over which change in output occurs because of pulse input.

Similarly input $x(j\omega)$ in the frequency domain is obtained as

$$x(j\omega) = \int_0^{t_0} x(t) \cos(\omega t) dt - j \int_0^{t_0} x(t) \sin(\omega t) dt = c - jd \quad (2.118)$$

where t_i = input pulse duration.

Frequency transformation of the process is thus

$$T(j\omega) = \frac{a - jb}{c - jd} = \frac{ac + bd}{c^2 + d^2} + j \frac{ad - bc}{c^2 + d^2} \quad (2.119a)$$

$$= \text{Re}T(j\omega) + j\text{Im}T(j\omega) \quad (2.119b)$$

The amplitude ratio $|T(j\omega)|$ and the phase angle ϕ are thus given by

$$|T(j\omega)| = [\{\text{Re}T(j\omega)\}^2 + \{\text{Im}T(j\omega)\}^2]^{1/2} \\ = \sqrt{\left(\frac{ac + bd}{c^2 + d^2}\right)^2 + \left(\frac{ad - bc}{c^2 + d^2}\right)^2} \quad (2.120a)$$

$$\phi = \tan^{-1} (ad - bc)/(ac + bd) \quad (2.120b)$$

Above equations and the input output records are used for parameter determination by drawing the $|T(j\omega)| - \omega$ and $\phi - \omega$ plots as is done in the case of Bode plots. A specific frequency ω is chosen, integrals a, b, c, d are evaluated at that frequency and with these $a, b, c, d, |T(j\omega)|$ and ϕ are calculated to obtain a point on the plot proposed. Taking numerous values of ω , the plots are completed. The integration should be performed graphically or using numerical techniques in a computer as the function $x(t)$ and $y(t)$ are given graphically and not by any analytical function.

Once these plots are ready, they are fitted to the standard file diagrams or to a specific proposed transfer function by following the procedure as indicated. If a process transfer function of the form $G_p(s) = K_p \exp(-s\tau_d)/(s\tau + 1)$ is desired, first K_p is obtained by taking the ratio of the area of the output curve to the area of the input curve (in time domain). From the amplitude ratio vs ω plot $K_p/\sqrt{2}$ point is found and the corresponding

frequency is determined which is corner frequency ω_c , from which τ is obtained as $1/\omega_c$. With this τ , a second phase vs frequency plot is superposed on the already obtained phase-frequency plot. If the two curves are identical $\tau_d = 0$. If not, the total phase lag from the pulse test data should equal that due to the process time t and dead time τ_d . Thus at corner frequency

$$\phi = \text{phase angle due to dead time} |_{\omega = \omega_c} - \tan^{-1} \omega_c t = \phi_d + \phi_\tau$$

Hence, $\phi_d = \phi_c - \phi_\tau$ and the dead time

$$\tau_d = \frac{\phi_c - \phi_\tau}{\omega_c} \times \frac{\pi}{180} \quad (2.121)$$

For a second order case, gain K_p is similarly found or from the asymptote on the amplitude ratio plot (at zero frequency). For an underdamped second order system the breakpoint frequency is the value of ω_n (see Fig. 2.27(b)). For finding the value of ζ , the resonant peak obtained is compared with the data of known second order systems (file data).

For an overdamped second order model, a first order curve (amplitude-frequency) is tried to fit with the obtained data and the corner frequency of the first order curve that fits this would give the first time constant τ_1 (as $1/\omega$). After this the first order file curve is subtracted from the obtained amplitude ratio-frequency plot which would give another curve which in turn is curve-fitted with another file first order lag curve from which the other time constant $\tau_2 (= 1/\omega_2)$ is obtained. In either case (underdamped or overdamped) the obtained curve is thus fitted to known transfer function curve whose phase angles are plotted and compared with the obtained phase angle plots. The difference, if any, is used to calculate the dead time in much the same way as in the first order case.

An important consideration is the width of the pulse which should be carefully chosen so that amplitude of Fourier integral transform remains nonzero till a sufficiently high frequency. On similar consideration the interval for numerical integration is to be chosen. Two or three trials in either case serve the purpose well.

2.9 FURTHER COMMENTS ON PARAMETER EVALUATION BY PROCESS TESTING

So long open loop test techniques for process modelling have been described. Parameter evaluation/estimation may also be done by testing on a plant consisting of two trials, one with an open loop and the other with a closed loop. The procedure is quite simple.

- (i) Set the controller in manual position. The control valve position is suddenly changed such that a step function is introduced. The time this change is made and the time the process starts responding, as

obtained in a recorder, are noted. The difference of the two is the dead time, τ_d .

- (ii) Now set the controller in automatic position with no derivative and reset action. The proportional band is gradually adjusted such that response is almost undamped oscillatory (see Chapter 4). Let the time period of oscillation be T_u and the proportional band PB (for details of PB see later (Chapter 5, for example).

From the values of τ_d and T_u it is possible to evaluate the process characteristics as regards its dynamic behaviour in terms of time constants and their number as also in terms of process controllability (see Chapter 4). The results are empirical and derived from experience.

If

- (i) $T_u/\tau_d = 2$, process has a pure dead time only;
- (ii) $T_u/\tau_d = 4$, process has a single dominant process lag (capacity).
- (iii) $T_u/\tau_d > 4$, process is a multilag one;
- (iv) $2 < T_u/\tau_d < 4$, dead time in the process is dominant.

Obviously, the proportional band of the controller, when undamped oscillation occurs in the system, is given by the relation

$$PB = G_0 \quad (1.122)$$

where G_0 is the overall gain of the other elements in the loop.

2.10 CONCLUSION

Although mathematical modelling of process helps in the analysis of process control systems and also in instrumentation system design, it is important to note that processes are never exactly represented by mathematical modelling because of large-scale approximation. Scale-modelling partially solves the situation but it does not simplify the method of control loop design; in addition to this, its economic viability has often been put to question particularly for small processes with less number of units and interconnections. However, in the absence of a better alternative, scale-modelling is sometimes resorted to. As has been shown for small processes or units, theoretical modelling is quite useful as long as physico-chemical equations can be written to represent the process operation close to the actual case. For slightly larger or bigger cases experimental test procedure to develop the mathematical models has found more favour, particularly in recent times, because of help a digital computer can extend to such cases. Such models can then be used to formulate a comprehensive design procedure of the control loop as a whole.

Review Questions

1. What are similarity states? How are they used in modelling of processes?
2. Derive the transfer function of a mixing process involving thermal balance when one of the inputs is controllable.
3. Obtain the expression for the transfer function of a tubular reactor. What are the assumptions required for deriving this transfer function?
4. How is a distillation column modelled for use in the process control analysis?
5. How is a linear distributed parameter system tackled for process control analysis especially at low fluctuation frequency of the disturbance? Obtain the temperature transfer function of a one dimensional heat transfer process which is a thick tube of length L and D is the thermal diffusivity.
6. A gasoline stabilizer consists of the following parts:
 - (1) stabilizer column,
 - (2) reboiler
 - (3) overhead condenser,
 - (4) reflux drum and reflux pump,
 - (5) feed pump,
 - (6) feed to the column (input) in the form of unstabilized gasoline.

Draw the schematic diagram of the stabilizer and obtain the transfer function between the output gasoline and the input.

7. The process transfer function of a system is given by $G_p(s) = 2/((2s + 1)(s + 1)(s + 2))$. Approximate the process by a second order model having dead time.
 (*Hint: Step change in input $x(t)$ is given and output $c(t)$ plotted as fractional response with time as in Fig. 2.21 and then the shaded area m_1 , is computed. Drawing the tangent line as in Fig. 2.21, its slope is identified as S_p , as also its intersection point with $c(t) = 1$ line wherefrom time t_2 is found, from Eq. (2.82) α is found; using Fig. 2.22, β is then computed. Finally using Eq. (2.84), τ_d , τ_1 , and τ_2 are obtained).*)
8. Approximate the third order transfer function of problem 2.7 by a second order one without dead time.
 (*Hint: As in Fig. 2.21, identify t_x with T_β and t_y with T_α of Fig. 2.18 and then use Eqs (2.62) and (2.63) and follow the subsequent procedure to obtain τ_1 and τ_2).*)

9. The frequency response plot of the transfer function of a process obtains a curve of the type shown in Fig. 2.27(c) with A_p being 30 per cent more than A_0 occurring at a frequency of 0.707 Hz. Calculate the process parameters and write the transfer function. (*Hints:* Use Eqs (2.116 a) and (2.116 b)].
10. Classify the process models structurally and indicate which type is suitable for what purpose. Describe the general procedure of building a process model in block diagrams.
11. How do you distinguish between parametric and nonparametric models? Obtain the difference equation representation of a tank which is being filled and emptied at the same rate but has different temperatures at the tank and its outflow. How do you convert this into a discrete model?
12. Describe the techniques of obtaining the discrete parametric models of processes.

Block Diagrams: Transient Response and Transfer Functions

3.1 BLOCK DIAGRAM REPRESENTATION

For control system studies the block diagram representation approach has now been universally accepted. Each equipment or component in the system is represented by a block and its transfer function is written inside the block or black box as it is sometimes known. The advantages are quite clear for such a representation; the primary one amongst those is that as transfer functions are used for interconnecting the blocks, an algebraic procedure may be adopted for the analyses of the systems and simplification of block diagrams is also possible. The transfer function is a dynamic relationship between the input x and output of the block with the initial conditions specified. The other component which is extensively used in block diagram representation is shown in Fig. 3.1 (a). This is known as a comparator and also sometimes as a differencer. It compares two signals x_1 and x_2 and the comparison is just an algebraic summation. The other representation which many authors resort to is shown in Fig. 3.1(b). The error after comparison is obtained as

$$e = x_1 - x_2 \quad (3.1)$$

Figure 3.1(c) shows the block that is used in the representation of a transfer function. In special comparators multiplication or division is assumed to be performed, as for example, in the ratio control system. These are shown by the symbols given in Fig. 3.1(d) or Fig. 3.1(e) where S stands for the multiplication $\times$ or division $\div$ sign and the operation is

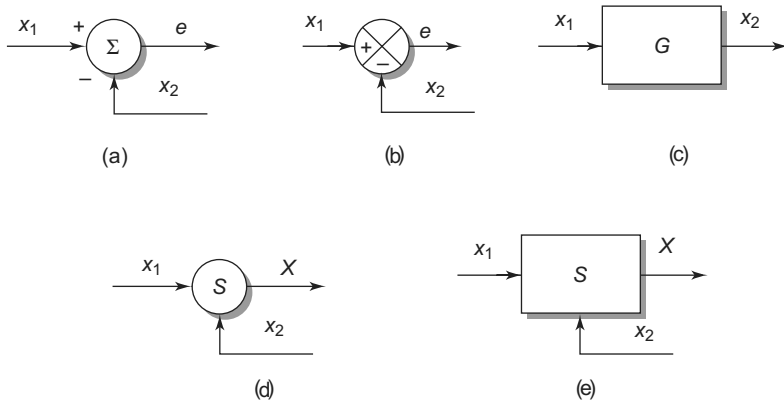


Fig. 3.1 Block diagram representation of (a) comparator, (b) alternative of Fig. 3.1(a), (c) transfer function, (d) and (e) generalized operation: operator S for multiplication or division

$$x_1 S x_2 = X \quad (3.2)$$

$$S = \times \quad \text{or} \quad \div$$

In Fig. 3.1(c) the following relationship is implied

$$x_2(s)/x_1(s) = G(s) \quad (3.3)$$

This is a purely algebraic relation and (s) represents the transformation. Often, this symbol is omitted. By way of example, the simple single feedback system is shown in Fig. 3.2. Omitting the operator notation, (s) , of equation (3.3), the algebraic equations are

$$c/e = G \quad (3.4a)$$

$$m/c = H \quad (3.4b)$$

$$r - m = e \quad (3.4c)$$

Eliminating m and e , the overall system transfer function is given by

$$c/r = G/(1 + GH) \quad (3.5)$$

The other relevant transfer functions that are often used in system analysis and design are the loop transfer function, LTF , actuating transfer function, ATF ($= e/r$), also called the error transfer function, open loop transfer function, forward transfer function, FTF . In the notations of Fig. 3.2, these are given as

$$LTF = GH \quad (3.6a)$$

$$ATF = 1/(1 + GH) \quad (3.6b)$$

$$FTF = G \quad (3.6c)$$

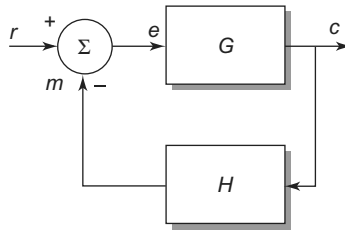


Fig. 3.2 Block diagram representation of simple feedback system

For analysis of the control system it may often be necessary to make an equivalent simplification of the block diagram obtained from the process. Some useful equivalent representations are shown in Figs 3.3(a), (b), (c) and (d).

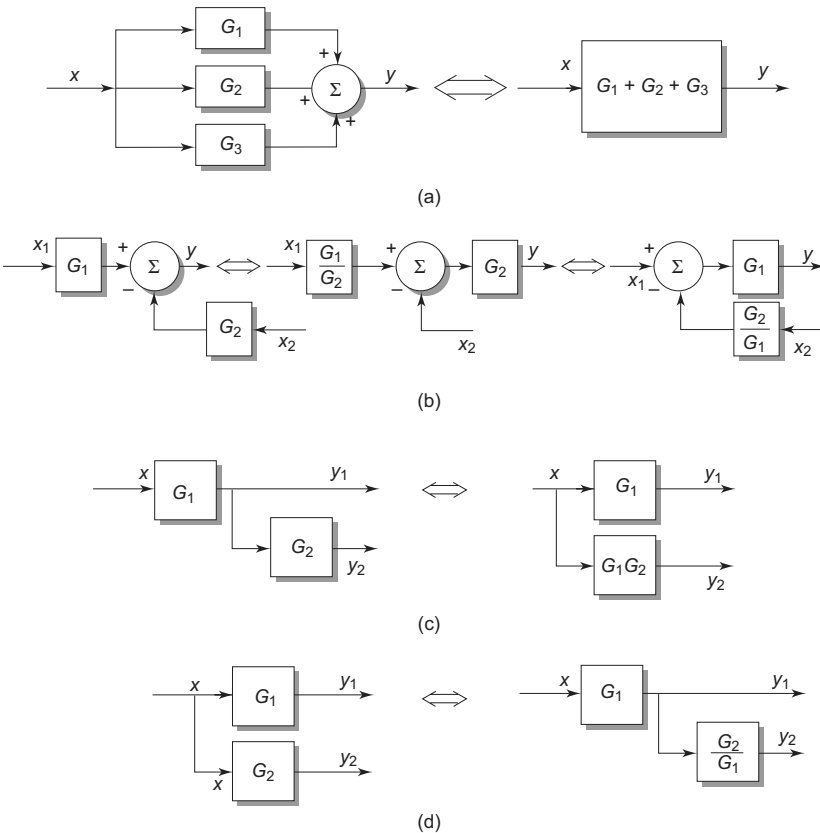


Fig. 3.3 (a), (b), (c) and (d): Schemes showing some equivalent representations

Multiloop systems are amenable to simplification by the long-hand procedure or the overall transfer function x can be deduced through simple

rules. An n -loop system shown in Fig. 3.4(a) can thus be simplified by taking the first loop first, such that

$$T_1 = c/r_1 = G_1/(1 + G_1H_1) \quad (3.7a)$$

Then the second loop analysis gives

$$T_2 = c/r_2 = T_1G_2/(1 + T_1G_2H_2) \quad (3.7b)$$

Proceeding likewise, one obtains

$$T_n = c/r_n = T_{n-1}G_n/(1 + T_{n-1}G_nH_n) \quad (3.7c)$$

T_{n-1} is then gradually replaced by the T_j 's of the preceding equations. If, however, the reciprocal of T_n is first obtained, then

$$T_n^{-1} = H_n + T_{n-1}^{-1}G_n^{-1} \quad (3.8a)$$

Replacing T_{n-1}^{-1} from the preceding stage

$$T_n^{-1} = H_n + H_{n-1}G_n^{-1} + T_{n-2}^{-1}G_n^{-1}G_{n-1}^{-1} \quad (3.8b)$$

Proceeding likewise, one gets

$$T_n^{-1} = H_n + H_{n-1}G_n^{-1} + H_{n-2}G_n^{-1}G_{n-1}^{-1} + \dots + G_n^{-1}G_{n-1}^{-1}\dots G_1^{-1} \quad (3.8c)$$

The order of appearance of the terms in Eq. (3.8c) of the system can be written by an inspection of the system block representation. The last term is the reciprocal of the *FTF*, the last but one term is the product of the reciprocal of the *FTF* and the *LTF* of the first loop, previous to the last but one term is the product of the reciprocal of the *FTF* and the *LTF* of the next higher loop, and so on. Thus

$$T_n^{-1} = (FTF)^{-1}[(LTF)_n + (LTF)_{n-1} + \dots + (LTF)_1 + 1] \quad (3.8d)$$

For a positive feedback the corresponding *LTF* would be associated with a negative sign.

Even for multiloops, with the output c not appearing in all the loops the same procedure gives the transfer function or rather its reciprocal by inspection.

By way of example we take three loop system of Fig. 3.4(b). There are three *LTF*'s given as

$$(LTF)_1 = G_3G_2G_1H_1H_2H_3$$

$$(LTF)_2 = G_3G_2J_1H_2H_3$$

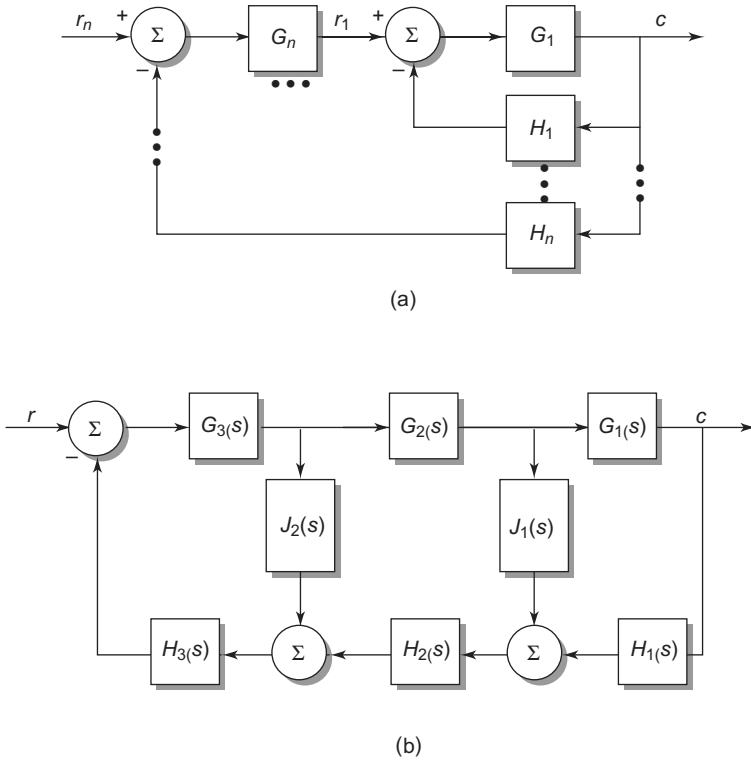


Fig. 3.4 (a) Representation of an n -loop system,
(b) representation of a 3-loop system without a common output point

$$(LTF)_3 = G_3 J_2 H_3$$

and one FTF whose inverse can be written as

$$(FTF)^{-1} = G_3^{-1} G_2^{-1} G_1^{-1}$$

The overall transfer function $c(s)/r(s)$ is obtained by the following relation

$$\begin{aligned} T(s)^{-1} &= [c(s)/r(s)]^{-1} = (FTF)^{-1} [1 + (LTF)_3 + (LTF)_2 + (LTF)_1] \\ &= G_3^{-1} G_2^{-1} G_1^{-1} + G_2^{-1} G_1^{-1} J_2 H_3 + G_1^{-1} J_1 H_2 H_3 + H_1 H_2 H_3 \end{aligned}$$

This method actually led to the derivation of the signal flowgraph technique of Masson for complex multiloop systems for transmittance realization.

3.2 STEP, FREQUENCY AND IMPULSE RESPONSE OF SYSTEMS

When the block representation of any arbitrary system has been made, its response to different types of inputs also requires to be discussed. The block basically represents a system, the plant, process, the measurement system, controller and so on. Each such block, in practice, would receive varied types of inputs and correspondingly give out outputs which would depend on the block, i.e., the actual system in the block, the type of the input and the initial condition of the system. To make a general assessment of all possible systems, inputs and initial conditions would only increase confusion. Besides, when the system order becomes high (greater than two), analytical approach fails to be incorporated easily. Order of the system is the same as the order of the differential equation by which a system is modelled mathematically. For example a linear system of non-time varying type can be modelled by an n th order equation relating the input x and output y as

$$\sum_{i=0}^n a^i \frac{d^i y}{dt^i} = bx \quad (3.9)$$

In process control systems, the individual system order is hardly more than two, although plants and processes may be of higher order but on the dominant response characteristics they may be reduced to first or second order models. Also, the system is considered to be initially at rest and finally the types of inputs which really matter are step, ramp, sinusoidal (frequency) and impulse. A combination of these four would give rise to any type of input that may be expected in a practical system. We shall make some relevant studies of response characteristics of processes up to second order types. It is available, at least a major part of it, in many standard texts, specifically in *Principles of Industrial Instrumentation* by the same author.

3.2.1 First Order System

Zero order systems are ones which are guided by Eq.(3.9) when $n = 0$, there is no dynamic error or system lag in such cases and such systems are, therefore, not considered in details here.

The types of the inputs that would be considered as already mentioned, are

$$\begin{aligned} \text{(i) Step: } x &= 0, \quad \text{when } t \leq 0^- \\ x &= x_0, \quad \text{when } t \geq 0^+ \end{aligned} \quad (3.10a)$$

$$\begin{aligned} \text{(ii) Ramp: } x &= 0, \quad \text{when } t \leq 0^- \\ x &= x_r t, \quad \text{when } t \geq 0^+ \end{aligned} \quad (3.10b)$$

$$\begin{aligned}
 \text{(iii) Sinusoidal:} \quad x &= 0, \quad \text{when } t \leq 0^- \\
 x &= x_s \sin \omega t, \quad \text{when } t \geq 0^+
 \end{aligned} \tag{3.10c}$$

and

$$\begin{aligned}
 \text{(iv) Impulse:} \quad x &= 0, \quad \text{when } 0^- > t > 0^+ \\
 x &= A \rightarrow \infty, \quad \text{when } t = 0
 \end{aligned} \tag{3.10d}$$

Impulse function is, in actuality, a derivative of the step function.

If $n = 1$ in Eq. (3.9), the equation becomes that of a first order system. Replacing the d/dt by operator s , the first order representative equation becomes

$$(s\tau + 1)y = kx \tag{3.11}$$

where $\tau = a_1/a_0$ and $k = b/a_0$, the parameter τ , having the dimension of time, is termed as the system time constant. Solving Eq. (3.11), one gets,

$$y = kx_o(1 - \exp(-t/\tau)) \tag{3.12}$$

for $y = 0$ at $t \leq 0$

This is the transient response equation and from this the instantaneous dynamic error is $kx_o \exp(-t/\tau)$ which is dependent on τ , the system time constant, which, in turn, is a function of several physical properties of the system. The steady state error, i.e., error when $t \rightarrow \infty$, is, however, zero for such a system. Figures 3.5 (a) and (b) show the input and response for two different time constants. With $\tau_1 < \tau_2$, response for a system with τ_1 is obviously better.

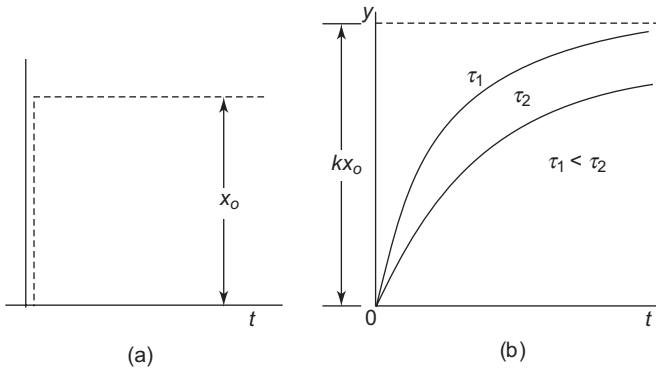


Fig. 3.5 (a) Step input represented,
(b) Response curves for a first order process
for two different time constants with step input

For a ramp input or an input which changes at a constant rate, the system equation in operator notation is

$$(s\tau + 1)y = kx_r t$$

giving a solution

$$y = C \exp(-t/\tau) + kx_r(t - \tau) \quad (3.13a)$$

which transforms into

$$y = kx_r t [1 - \tau/t(1 - \exp(-t/\tau))] \quad (3.13b)$$

for the initial condition $y = x = 0$ at $t = 0$. Figure 3.6 shows the input and output curves where the transient error is obviously (also obtained from Eq. (3.13b)) $kx_r \tau(1 - \exp(-t/\tau))$ giving a steady state value of $kx_r \tau$. The system has a time lag of τ as is seen from Eq. 3.13b and Fig. 3.6. The value is the same as the system time constant.

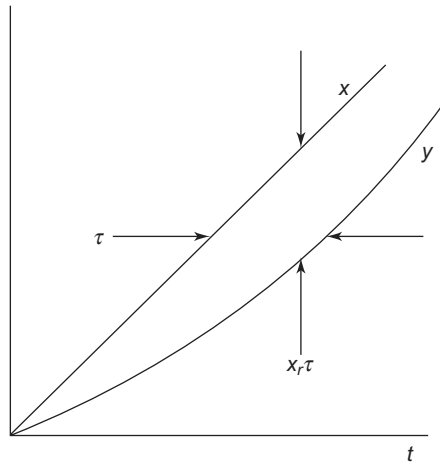


Fig. 3.6 Ramp input to and response curves of a first order process

With an input $x = x_s \sin \omega t$, the system equation is

$$(s\tau + 1)y = kx_s \sin \omega t \quad (3.14)$$

which with initial conditions gives a solution

$$y = \left[kx_s / \sqrt{1 + \omega^2 \tau^2} \right] \sin(\omega t - \phi), \phi = \tan^{-1} \omega \tau \quad (3.15a)$$

$$= y_s \sin(\omega t - \phi) \quad (3.15b)$$

The transient or dynamic error and the system time lag are given respectively by

$$E_d = kx_s (1 - 1/\sqrt{1 + \omega^2 \tau^2}), T_i = (\tan^{-1} \omega \tau) / \omega$$

which increase as ω increases but is decreased with lower τ . Figures 3.7(a) and (b) plot the normalized amplitude frequency and phase-frequency

curves for the system which are what we call the frequency response characteristics of the system. For a periodic input with a number of frequencies, errors and time lags at different frequencies are obtained and are superposed according to the superposition principles.

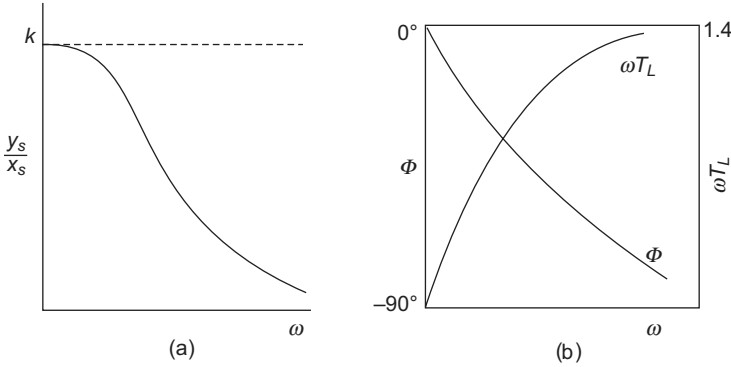


Fig. 3.7 Frequency response curves for a first order process:
(a) amplitude ratio-frequency plot and (b) phase-frequency plot

Impulse function is considered to be of infinite magnitude over zero duration time, often its magnitude time product is considered to be unity so that it can be called an unit impulse function. It is an idealized function and its response should also be idealized. An impulse function of strength A obtains the output from a first order system as

$$y = (kA/\tau)(\exp(-t/\tau)) \quad (3.16)$$

One would note that the solution is the time derivative of step input response of Eq.(3.12) with x_0 replaced by A . It has a dynamic error $kA \left(1 - \exp\left(-\frac{t}{T}\right)/\tau \right)$ from which a steady state value is kA .

3.2.2 The Second Order System

Second order systems are defined by the equation ($n = 2$ in Eq. (3.9))

$$a_2 d^2 y/dt^2 + a_1 dy/dt + a_0 y = bx \quad (3.17a)$$

A mass spring damping system is a typical second order system. Equation (3.17a) can be written in the operator form as

$$(s^2/\omega_n^2 + 2\zeta s/\omega_n + 1)y = kx \quad (3.17b)$$

where

$$\omega_n = \sqrt{a_0/a_2}, \quad \zeta = (1/2)(a_1/\sqrt{a_0 a_2}) \quad \text{and} \quad k = b/a_0 \quad (3.18)$$

The parameter ω_n is the natural frequency of oscillation of the system, ζ its damping factor and k is the sensitivity or conversion factor, all of which depend on the various physical properties of the system. With the prescribed inputs to the systems, the responses are now studied. When a step input is given, the transient part of the solution of Eq. (3.17b) may be of three different types depending on the roots of the characteristic equation $s^2/\omega_n^2 + 2\zeta s/\omega_n + 1 = 0$. The roots may be (1) real unrepeatd, $\zeta > 1$, (2) real and repeated, $\zeta = 1$, and (3) complex, $0 < \zeta < 1$. The three conditions give rise to the cases which are known as overdamped, critically damped and decaying oscillatory cases respectively. The solutions in the three cases are

$$(1) \quad y = kx_0 \left[1 \pm \frac{-\zeta \mp \sqrt{\zeta^2 - 1}}{2\sqrt{\zeta^2 - 1}} \exp(-\zeta \mp \sqrt{\zeta^2 - 1})\omega_n t \right] \quad (3.19a)$$

$$(2) \quad y = kx_0 [1 - (1 + \omega_n t) \exp(-\omega_n t)] \quad (3.19b)$$

$$(3) \quad y = kx_0 \left[1 - \frac{\exp(-\zeta \omega_n t)}{\sqrt{1 - \zeta^2}} \sin(\omega_n t \sqrt{1 - \zeta^2} - \phi) \right] \quad (3.19c)$$

$$\phi = \sin^{-1} \sqrt{1 - \zeta^2}$$

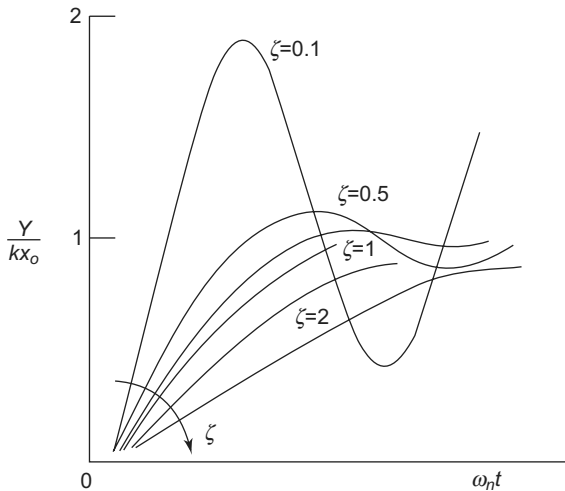


Fig. 3.8 Response curves for a second order process with step input

The plots are shown in Fig. 3.8 in normalized coordinates y/kx_0 and $\omega_n t$. With ζ very low response is oscillatory and with ζ very high response time is also high. However, if ζ is fixed, a large value of ω_n would mean a small time lag if $\omega_n t$ is constant. Transient errors are found from Eqs. (3.19) and curves of Fig. 3.8. Except when $\zeta = 0$, steady state error is zero.

For a ramp input, when initial conditions $y = dy/dt = 0$ at $t = 0$, the solutions of Eq. (3.17b) for the three different cases are

$$(1) \quad y = kx_r[t - 2\zeta/\omega_n]$$

$$\left\{1 \pm \frac{2\zeta(-\zeta \mp \sqrt{\zeta^2 - 1}) - 1}{4\zeta\sqrt{\zeta^2 - 1}} \exp((-\zeta \mp \sqrt{\zeta^2 - 1})\omega_n t)\right\} \quad (3.20a)$$

$$(2) \quad y = kx_r[t - 2/\omega_n][1 - \exp(-\omega_n t)(1 + \omega_n t/2)] \quad (3.20b)$$

and

$$(3) \quad y = kx_r \left[t - 2\zeta/\omega_n \left\{ 1 - \frac{\exp(-\zeta\omega_n t)}{2\zeta\sqrt{1-\zeta^2}} \sin(\omega_n t\sqrt{1-\zeta^2} - \phi) \right\} \right]$$

$$\phi = \tan^{-1} \frac{2\zeta\sqrt{1-\zeta^2}}{2\zeta^2 - 1}$$

The response curve in y/k and t coordinates along with the input curve are shown in Fig. 3.9. While the transient errors in the three different cases are obtained from the above equations or the plots of Fig. 3.9, there are steady state error and system time lag of values $2\zeta kx_r/\omega_n$ and $2\zeta/\omega_n$ respectively. The system is often specified by what is known as characteristic time $T_c = 1/(\zeta\omega_n)$.

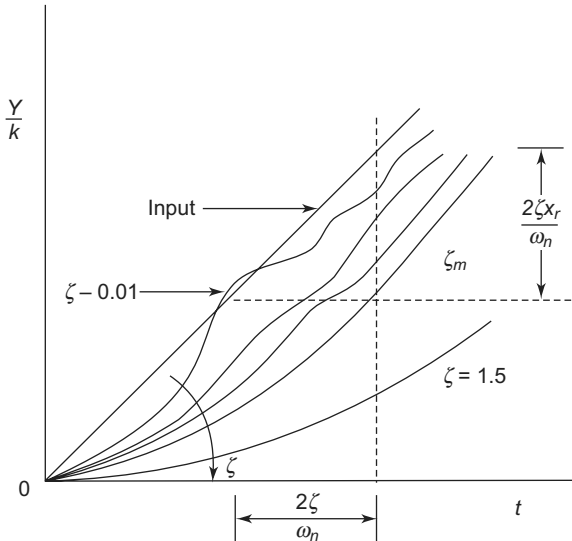


Fig. 3.9 Response curves for a second order process with ramp input

Response to the sinusoidal input is given as

$$y = kx_s / \sqrt{\left[\left\{ 1 - \left(\frac{\omega}{\omega_n} \right)^2 \right\}^2 + 4\zeta^2 \omega^2 / \omega_n^2 \right]} \sin(\omega t - \phi)$$

$$\phi = \tan^{-1} 2\zeta / \left[\frac{\omega}{\omega_n} - \frac{\omega_n}{\omega} \right] \quad (3.21)$$

The frequency response curves in $|y/kx_s| - \omega/\omega_n$ and $\phi - \omega/\omega_n$ coordinates are shown in Figs 3.10(a) and 3.10(b) while Fig. 3.10(c) shows the $\omega T_l - \omega/\omega_n$ plot, T_l being the time lag. System parameters ζ and ω_n may have values such that the system output becomes larger than the input. Specifically, at $\omega/\omega_n = 1$, and $\zeta \rightarrow 0$, output tends to be very high. Thus is the case of forced resonance. For a specific ζ , the peak output (at $\omega/\omega_n = 1$) is $kx_s/(2\zeta)$. For $\zeta = 0.707$, the response curve is maximally flat and phase maximally linear. The system time lag increases with ζ increasing, but with ω_n decreasing.

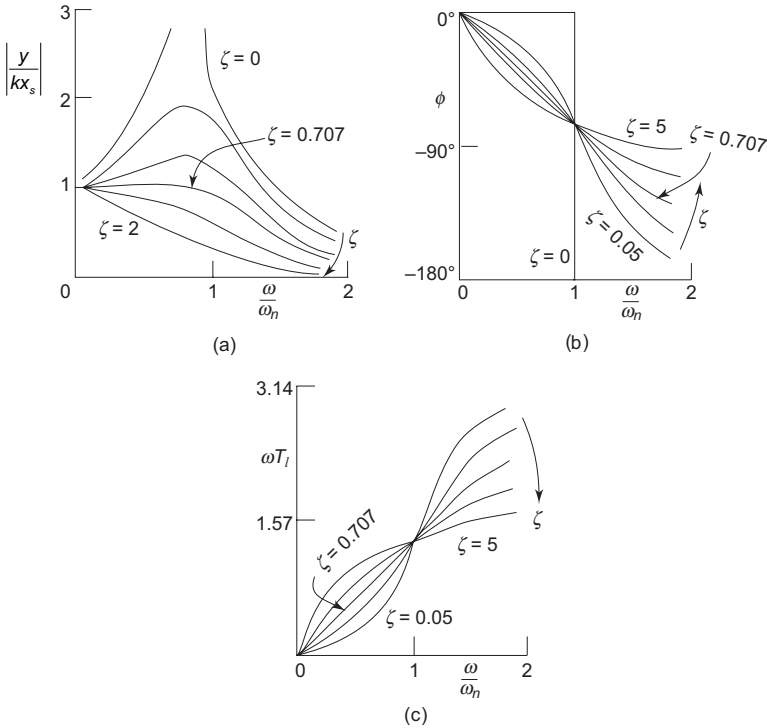


Fig. 3.10 Frequency response plots of a second order process:
 (a) amplitude ratio versus normalized frequency
 (b) phase-versus normalized frequency
 (c) normalized lag versus normalized frequency

For an impulse input of strength A , the solution of Eq. (3.17b) is given for the three cases of $\zeta > 1$, $\zeta = 0$ and $0 < \zeta < 1$ respectively as

$$y = [kA\omega_n/(2\sqrt{\zeta^2 - 1})](\exp((- \zeta + \sqrt{\zeta^2 - 1})\omega_n t) - \exp((- \zeta - \sqrt{\zeta^2 - 1})\omega_n t)) \quad (3.22a)$$

$$y = kA\omega_n^2 t \exp(-\omega_n t) \quad (3.22b)$$

and

$$y = (kA\omega_n/\sqrt{1 - \zeta^2})\exp(-\zeta\omega_n t)\sin(\omega_n t\sqrt{1 - \zeta^2}) \quad (3.22c)$$

The steady state error in all the cases is zero while the transient error depends on ζ , ω_n , A and of course on t .

Of specific importance in this category of response studies is the step input response as will be discussed at many points in the sequel of the text. Specifically when $\zeta < 0.707$, response produces some overshoots and undershoots at different instances of time. The peak overshoot and the time of peak overshoot as also the decay ratio of the oscillatory response are some important parameters often required in control system analysis. These are now evaluated.

If y_m is the peak output, then $y_m - kx_o =$ peak overshoot, λ_1 , the second overshoot is, say, λ_2 , then $\lambda_2/\lambda_1 =$ decay ratio. The peak overshoot occurs at a time t_p (say). Using Eq. (3.19c), by differentiating and using appropriate conditions one easily obtains

$$t_p = \pi/(\omega_n\sqrt{1 - \zeta^2})$$

$$\lambda_1 = \exp(-\pi\zeta/\sqrt{1 - \zeta^2})$$

$$\delta = \lambda_2/\lambda_1 = \exp(-2\pi\zeta/\sqrt{1 - \zeta^2})$$

and finally, the time period of the decaying oscillation, T ,

$$T = 2\pi/(\omega_n\sqrt{1 - \zeta^2})$$

First and second order systems with system lag, natural frequency of oscillation and damping have been considered above for specific inputs. Higher order systems have not been considered because they are not easily solved by simple analysis. Besides higher order systems are often approximated to second order ones with dominant pole or time constant consideration.

A system may also have what is known as dead time element. A system is said to have a dead time when its output is exactly of the same form as the input but occurs after a specific time τ_d known as the dead time.

Thus for an input $x(t)$ and output $y(t)$, one gets

$$y(t) = kx(t - \tau_d), \quad t \geq \tau_d \quad (3.23)$$

This type of element would change the response of the system with standard inputs but they can be obtained easily by considering the response Eq. (3.23). Figures 3.11 (a), (b) and (c) show the response for step, ramp and impulse inputs while those for sinusoidal case the frequency response plots are given in Figs 3.12 (a) and (b). For sinusoidal input, $x = x_s \sin \omega t$, $y = kx_s \sin(\omega(t - \tau_d))$, hence $y/x = k \angle \phi$; ϕ is actually $-\omega\tau_d$.

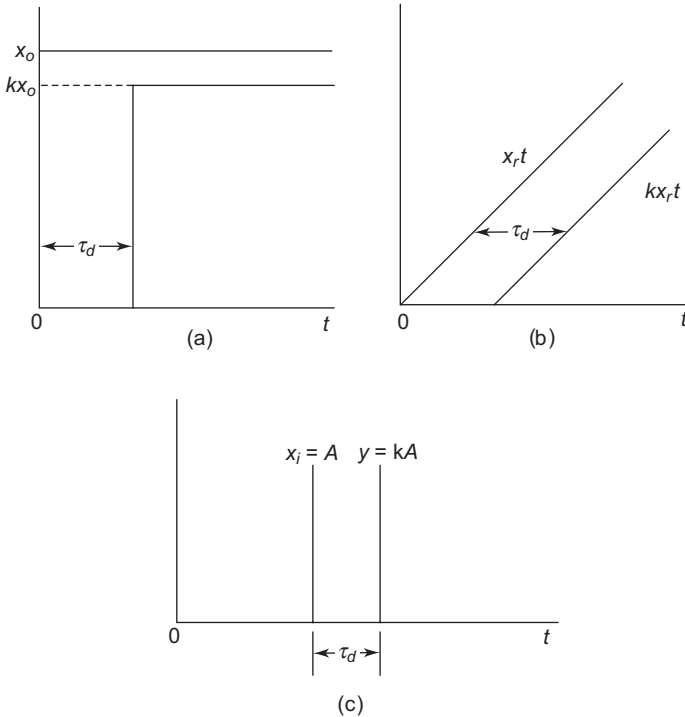


Fig. 3.11 Response of a process with dead time for (a) step input, (b) ramp input, and (c) impulse input

Dead time element is a very important factor in almost all processes. It would be pertinent to derive its system function here after a formal definition for the same is given.

Delay that occurs between two related actions is known as dead time. When an event occurs at a place which is measured at a distance downstream of the event, a delay determined by the flow rate and the distance occurs and is called the dead time or the transportation lag. Dead time occurs for some other reasons also. For example, in some chemical processes, a finite

time may elapse before a reaction starts even though the process operation has already started. In time function, the input and the output may thus be written as $f(t)$ and $f(t - \tau_d)$.

In Taylor's series one gets

$$f(t - \tau_d) = \sum_0^{\infty} (-1)^n f^n(t) \frac{\tau_d^n}{n!}$$

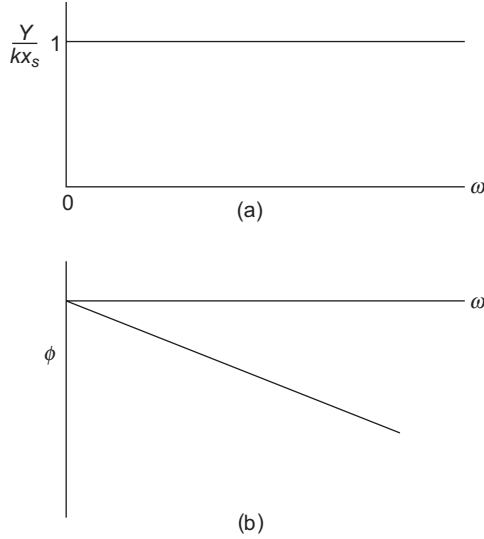


Fig. 3.12 (a) and (b): Frequency response plots for a process with dead time

where $f^n(t)$ represents the n th derivative of $f(t)$ with respect to time. In operator notation this is written as

$$\begin{aligned} f(t - \tau_d)(s) &= f(t)(s) \sum_0^{\infty} (-1)^n (s\tau_d)^n / n! \\ &= \exp(-s\tau_d) f(t)(s) \end{aligned}$$

Thus the system function for dead time is

$$G_d(s) = \exp(-s\tau_d)$$

Typical values of dead time can be quoted here for a process like pneumatic tubing. A pressure signal at one end of a 330 m tubing would travel at the speed of sound till the other end, i.e., at a speed of roughly 330 m/sec and hence the tubing has a dead time of 1 sec and its system function is $\exp(-s)$

3.3 CONTROLLED PROCESS BLOCK DIAGRAMS AND TRANSFER FUNCTIONS

A standard block diagram of a single loop linear control element consists of four blocks in general—those of the process, the actuating element and the control valve, the controller in the forward and the measurement system in the feedback path. There is only one comparator. The representation is as shown in Fig. 3.13. The variables after each block and before and after the comparator are separately marked. On the forward line these are reference, error, manipulating, actuating and controlled. On the feedback path M is the measured controlled variable. One or all of these variables are subjected to disturbances. Two major sources of disturbances are power supply fluctuation and load disturbances in the process. These will be called upsets. A schematic representation showing how the disturbances upset the system is given in Fig. 3.14. Following the symbols we have already described in an earlier section we can represent Fig. 3.14 in two different ways if we assume that the disturbances upset the block (i) completely or (ii) only partially.

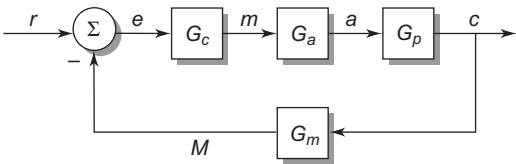


Fig. 3.13 Schematic representation of a single loop feedback system showing controller and actuator separately

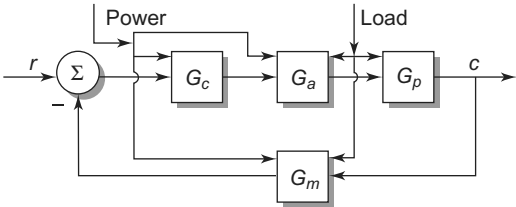


Fig. 3.14 Schematic representation of a feedback system showing the disturbances

However, since the power supply fluctuation can be easily compensated and only the load disturbances affect the process and subsequent feedback element more severely, the load disturbances are of major importance in process control systems. These disturbances affect the process operations so seriously that they affect the controllability as well. Two ways to represent these disturbances in block schematics are shown in Figs 3.15 (a) and (b). However, from the equivalences of Fig. 3.3(b) one notes that if $G_{u2} = G_{u1}/G_p$, then the two systems are identical. This also shows where

to assume the disturbances so that the process may be involved fully or partially by a proper choice of G_{u1} or G_{u2} .

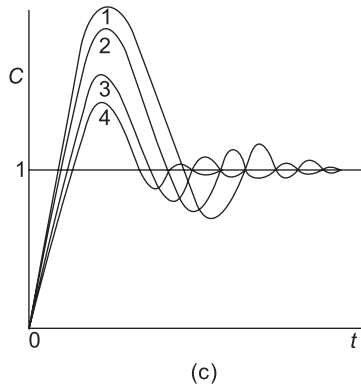
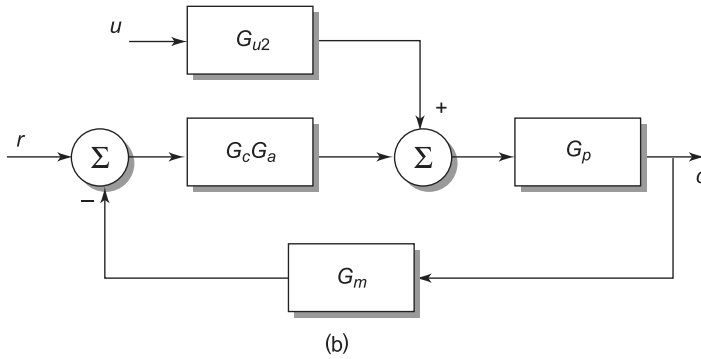
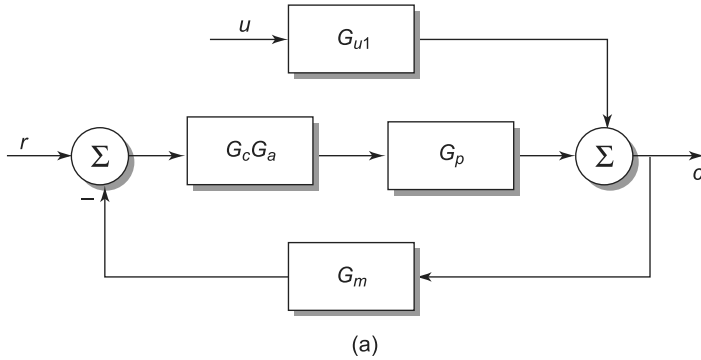


Fig. 3.15 (a) and (b): Two different ways of bringing the load disturbance in the block diagram of the system, (c) response curves for varying K_c and τ_m

curve 1:	$K_c = K_{c1} > K_{c2}$	$\tau_m = \tau_1 = \tau_2$
curve 2:	$K_c = K_{c2} > K_{c4}$	$\tau_m = \tau_2 > \tau_4$
curve 3:	$K_c = K_{c3} < K_{c1}$	$\tau_m = \tau_3 = \tau_1$
curve 4:	$K_c = K_{c4} = K_{c2}$	$\tau_m = \tau_4 < \tau_2$

In an example of deriving the block schematic of a process we include two disturbances both of which completely affect the process when there is a sudden change in either of them. The process is a heat-exchanger (Fig. 3.16a) in which $q_i = q_o = q$ is the flow rate of a fluid that is heated up from temperature T_i to T_o by a controlled flow of steam.

The block diagram is easily drawn, as shown in Fig. 3.16(b), if the reference temperature is T_r . Obviously, two comparators before T could be replaced by a single one by properly choosing G_q and G_T .

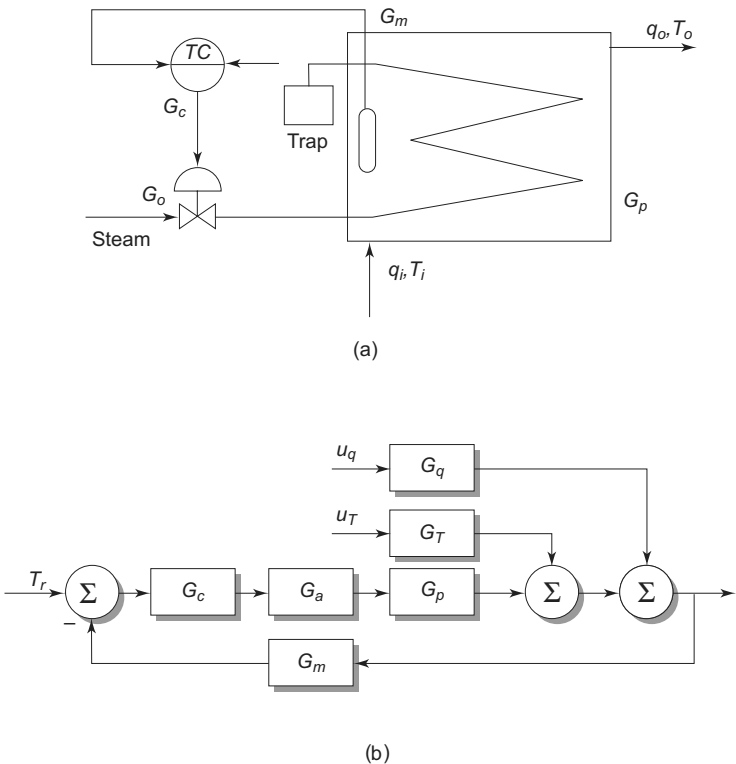


Fig. 3.16 (a) Schematic diagram of a heat exchanger
(b) equivalent block diagrammatic scheme of Fig. 3.16(a) with temperature as reference

In the case of multi-time constant processes if all the processes are not fully affected by load disturbance, G_p may be broken as $G_p = G_{p1}G_{p2}$ and disturbance may be allowed to enter between G_{p1} and G_{p2} . The analysis follows an analogous approach, as discussed in the earlier paragraphs.

Before taking up general problems of process control, simple cases are first treated to exemplify the approach of block diagram representation and show that some important results follow.

From Fig. 3.15(b), the transfer function between the output and set point is obtained as

$$T_s(s) = c(s)/r(s) = G_c G_a G_p / (1 + G_c G_a G_p G_m)$$

If proportional action is alone considered with $G_c = K_c$, the proportional gain, and further if $G_m = G_a = 1$ and $G_p = K_p/(s\tau + 1)$, where K_p is the process gain and τ is the process time constant, one easily gets

$$T_s(s) = K_c K_p / (1 + K_c K_p + s\tau) = K_e (1 + s\tau_e) \quad (3.24)$$

where

$$K_e = K_c K_p / (1 + K_c K_p), \text{ and } \tau_e = \tau / (1 + K_c K_p)$$

If this system is now met with a unit step change in the input, i.e., $r(t)$ is a unit step function, one can easily show that c approaches K_e with time $\rightarrow \infty$, i.e., $s \rightarrow 0$ and not the unit value. The discrepancy so obtained for $t \rightarrow \infty$ in the system is known as offset and is given by

$$\text{Offset} = r(t \rightarrow \infty) - c(t \rightarrow \infty) = 1 - K_e = 1/(1 + K_c K_p) \quad (3.25)$$

With $K_c K_p \rightarrow \infty$, offset approaches zero. In fact for proportional control alone a large proportional gain is necessary for this; however the other important aspects of stability and response speeds are also to be simultaneously checked.

For fixed r , but a load change (change in u_2), the transfer function for the above case with $G_{u1} = 1/(s\tau + 1)$ is given by

$$\begin{aligned} T_1(s) &= c(s)/u_2(s) = G_{u2} G_p / (1 + G_c G_p) = G_{u1} / (1 + G_c G_p) \\ &= 1 / (1 + K_c K_p + s\tau) = K_l / (1 + s\tau_e) \end{aligned}$$

where $K_l = 1/(1 + K_c K_p)$

Now if there is a load change by unit step, output changes but there is no change in the set point so that the offset now will be given by

$$\text{Offset}|_{\text{load}} = 0 - 1/(1 + K_c K_p) = -1/(1 + K_c K_p) \quad (3.26)$$

This offset is generally very important in industrial processes because in such process load disturbances are quite frequent and the means of keeping this load offset low should be given due consideration. Its reduction, however, follows a similar argument as was proposed for set point offset.

3.3.1 Static Error, Rate Static Error and Load Static Error

As mentioned in process control system the two modes of operations, one due to set point change and the other due to load disturbance are prevalent. Often in industrial processes both set point and load change. A typical example of this is a reheating furnace in a cogging mill. The temperature of the soaking zone may have to be varied depending on the type of material

being soaked, whereas load variation occurs depending on the rolling-rate variation. Considering the typical block diagram of a generalized process control system, as shown in Fig. 3.17, one obtains the relation between c , r and u as

$$c = G_f r / (1 + G_f G_m) + G_L u / (1 + G_f G_m) \quad (3.27)$$

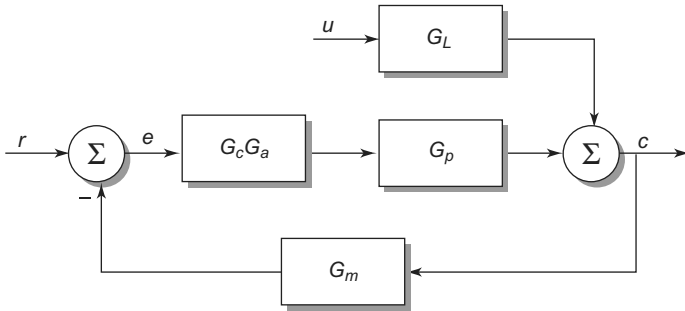


Fig. 3.17 Typical block schematic diagram of a generalized system, G_L : disturbance block transfer function

where $G_f = G_c G_a G_p$ is the forward-path transfer function between r and c , i.e., between e and c . Also, one should remember that all G 's are functions of s .

Two things may now be considered:

- (i) A change in the set point occurs whatever the condition of the upset or the load disturbance is. The change in the set point may be a function of s . But in a majority of cases, the change occurs from a given value to another given value such that a step change follows. Also, in such a case it is assumed that variation in u is zero. If r changes from r_1 to r_2 (say) such that $r_1 - r_2 = \Delta r$, one obtains the deviation (e) in a steady state condition as

$$e_s = \lim_{s \rightarrow 0} \frac{\Delta r [1 - G_f (1 - G_m)]}{1 + G_f G_m} \quad (3.28)$$

This quantity is also known as static error and is determined by magnitude of the set point changes, also the steady state values of system functions.

Another variation in the mode of the set point change occurs where the set point increases linearly with time. For a properly adjusted control system the control variable follows the set point but even in such a case a steady deviation might result. Assuming once again that $\Delta u = 0$, for such a case, from Eq. (3.27), the steady deviation is

$$e_R = \lim_{s \rightarrow 0} R [1 - G_f (1 - G_m)] / s (1 + G_f G_m) \quad (3.29)$$

where R is the rate of the linear increase of set point as mentioned. This is also known as velocity error, rate static error or simply the rate error. (ii) The second possibility is the set point is kept unaltered but the load changes and this change may be considered to be a function of s . However, in this case also, step disturbance is considered as is general and usual, and from Eq. (3.27), the steady state deviation in such a case, with $r = 0$, is given by

$$e_u = \lim_{s \rightarrow 0} -\Delta u G_L / (1 + G_f G_m) \quad (3.30)$$

where Δu is the amount of change in u . The parameter e_u is also known as offset or the load static error as has already been discussed above.

3.3.2 Elimination of Offset

It has been shown in Sec. 3.3 that with proportional action offset is not likely to be eliminated completely however large its value may be. The introduction of an integral action, however, improves the picture considerably. Now assume

$$G_c = K_c(1 + 1/sT_R)$$

so that the transfer function for the system considered in Sec. 3.3 for set point change changes to

$$\begin{aligned} T_{Rs}(s) &= \frac{K_c K_p (1 + sT_R)}{sT_R(1 + s\tau) + (1 + sT_R)K_c K_p} \\ &= \frac{sT_R + 1}{s^2/\omega_{ns}^2 + 2\zeta s/\omega_{ns} + 1} \end{aligned} \quad (3.31)$$

where $\omega_{ns} = [K_c T_p / \tau T_R]^{1/2}$ and $\zeta = [(1 + K_c K_p)/2] [T_R / \tau K_c K_p]^{1/2}$

Response to unit step change is now obtained in a routine manner indicating that $c_{Rs}(t \rightarrow \infty)$ for such a case is also unity. In fact, for unit step set point change

$$\begin{aligned} c_{Rs}(t) &= 1 - \frac{\exp(-\omega_{ns}\zeta t)}{\sqrt{1-\zeta^2}} \sin \left(\omega_{ns} t \sqrt{1-\zeta^2} + \tan^{-1} \frac{\sqrt{1-\zeta^2}}{\zeta} \right) \\ &\quad + \frac{\omega_{ns} T_R \exp(-\omega_{ns}\zeta t)}{\sqrt{1-\zeta^2}} \sin(\omega_{ns} t \sqrt{1-\zeta^2}) \end{aligned}$$

This shows that the offset in such a case is zero. Similarly, for the load change, the transfer function is given by

$$\begin{aligned} T_{Rl}(s) &= sT_R K_p / [s^2 \tau T_R + sT_R(1 + K_c K_p) + K_c K_p] \\ &= K_R s / (s/\omega_{nl}^2 + 2\zeta s/\omega_{nl} + 1) \end{aligned}$$

where $K_R = T_R/K_c$, $\omega_{nl} = [K_c K_p / (\tau T_R)]^{1/2}$ and ζ is as given above. A unit step change in the load will thus give the response as

$$c_{RI}(t) = \frac{\omega_{nl} K_R \exp(-\omega_{nl} \zeta t)}{\sqrt{1 - \zeta^2}} \sin(\omega_{nl} t \sqrt{1 - \zeta^2}) \quad (3.32)$$

which becomes zero as $t \rightarrow \infty$. Since now the set point change is also zero, offset is zero.

Thus it is clear that $P + I$ action eliminates offset. This is a very important result in the process control system.

Another instructive result is obtained as a consequence of a change in the measurement lag. If in the adopted system, $G_m = 1/(s\tau_m + 1)$ and proportional gain alone is considered with $G_a = 1$, the transfer function with a set point change is easily obtained as

$$T_{ms}(s) = K_c(s\tau_m + 1)/(s^2/\omega_m^2 + 2\zeta s/\omega_m + 1)$$

where $\omega_m = [(1 + K_c K_p)/(\tau\tau_m)]^{1/2}$, and $\zeta_m = (\tau + \tau_m)/[2\sqrt{\tau\tau_m(1 + K_c K_p)}]$

The time response of this type of equation is already given for a unit step change. To study the effect of varying τ_m and K_c , curves may be plotted. However, one can easily show that by increasing τ_m , while keeping K_c fixed, or vice versa, the transient peak increases, thereby deteriorating the system performance. It is thus necessary to choose an optimum K_c and τ_m should be as small as possible. Illustrative curves are given in Fig. 3.15c.

3.3.3 Transfer Functions of Control Equipment

The block diagram approach is quite standard a practice for analysing the overall system. However, for a generalized analytical approach, additionally, the transfer functions of the equipment are equally necessary. In the previous chapter transfer function calculation or evaluation from practical tests has been presented. Here, examples of transfer function calculation for the control equipment are given as a supplementary measure.

Take, for example, the 3-term parallel pneumatic controller (a series type is discussed in a later chapter) shown in Fig. 3.18. The process variable, via a link, moves the flapper of the flapper-nozzle assembly of the controller at point a , making gap d between the flapper and nozzles small or large and creating a large or small output pressure, p . Thus the flapper-nozzle gain can be stated in the simple form as $k_n = -p/d$, where the negative sign indicates that the closer the flapper is to the nozzle, the greater is the pressure. And the smaller the value of d is, the larger is the output pressure p . Usually the value of k_n is very large and d is very small. The order of the value of d is about a few thousandth of an inch. In such a case, a linear relationship between d and back pressure p may be assumed. The pressure enters the integral action bellows through the needle valve R

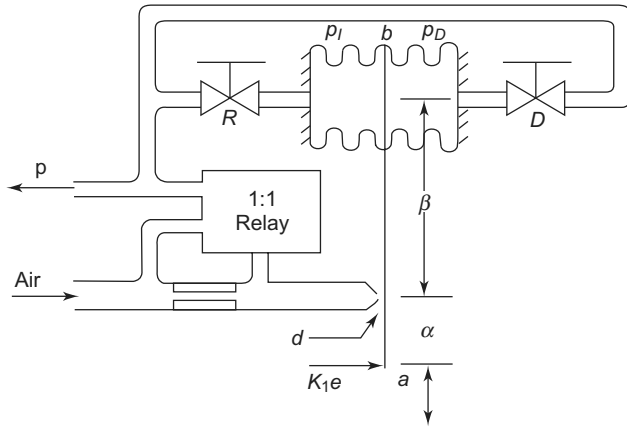


Fig. 3.18 Schematic diagram of a three term parallel type pneumatic controller

to create the pressure p_I . This pressure tries to move the other end of the differentially arranged flapper to the right. A sudden change in p , due to a process variable change, changes p_I exponentially with a time constant τ_R , where τ_R is the product of the resistance of R and capacitance of the integration bellows element. The transfer function between p_I and p is thus obtained as

$$p_I = p / (s\tau_R + 1) \quad (3.33)$$

In a similar manner, p changes p_D , the pressure in the derivative action bellows element and the transfer function

$$p_D = p / (s\tau_D + 1) \quad (3.34)$$

is easily obtained, where τ_D is the derivative time constant, being the product of resistance of valve D and capacitance of the derivative bellows element. Point b moves because of the difference in p_D and p_I . When a deviation of process variable e occurs, moving point a of the flapper by $k_1 e$, disturbance in d is given by $k_1 e \beta / (a + \beta)$ and the pressure p is changed to

$$p = k_1 k_n e \beta / (a + \beta) \quad (3.35)$$

Here, k_1 is just a conversion factor. The joint action of p_I and p_D disturbs point b which similarly affects pressure and its corresponding value is

$$p = k_3 k_n \alpha (p_I - p_D) / (\alpha + \beta) \quad (3.36)$$

where k_3 is a constant of proportionality indicating the amount of movement of point b due to the differential pressure. The equivalent displacement d of the flapper at the nozzle head is thus

$$d = k_1 e \beta / (\alpha + \beta) + k_3 (p_I - p_D) \alpha / (\alpha + \beta) \quad (3.37)$$

which when multiplied by k_n would give the pressure p .

Thus

$$\frac{k_1 k_n e \beta}{\alpha + \beta} + \frac{k_3 k_n \alpha p}{\alpha + \beta} \left(\frac{1}{s \tau_R + 1} - \frac{1}{s \tau_D + 1} \right) = -p \quad (3.38)$$

or

$$\frac{k_1 k_n e \beta}{\alpha + \beta} = -\frac{k_3 k_n}{(\alpha + \beta)} \frac{\alpha p}{(s \tau_R + 1)} \frac{s(\tau_D - \tau_R)}{(s \tau_D + 1)} - p \quad (3.39)$$

Now since k_n is very large, $\tau_R \neq \tau_D$ and the frequency is not on extreme sides, for process control applications Eq. (3.39) changes to

$$\frac{k_1 k_n e \beta}{\alpha + \beta} = -p \frac{k_3 k_n \alpha s(\tau_D - \tau_R)}{(\alpha + \beta)(s \tau_R + 1)(s \tau_D + 1)} \quad (3.40)$$

which on simplification yields the transfer function

$$\frac{p(s)}{e(s)} = \frac{k_1 \beta}{k_3 \alpha} \frac{1 + \tau_D / \tau_R}{1 - \tau_D / \tau_R} \left(1 + \frac{1}{s \tau_R \left(1 + \frac{\tau_D}{\tau_R} \right)} + \frac{s \tau_D}{\left(1 + \frac{\tau_D}{\tau_R} \right)} \right) \quad (3.41)$$

Putting $k_1 \beta / k_3 \alpha = k_c$, $1 + \tau_D / \tau_R = \phi$ and as $\tau_D \ll \tau_R$, one gets

$$\frac{p(s)}{e(s)} = k_c \phi \left(1 + \frac{1}{s \tau_R \phi} + \frac{s \tau_D}{\phi} \right) \quad (3.42)$$

The quantity ϕ is a function of τ_D and τ_R and its presence in both the second and third terms within parentheses makes the controller an interacting one, ϕ is usually called the interaction factor and k_c the gain or the proportional action gain as in the absence of integral and derivative control actions provided by τ_R and τ_D adjustments, the flapper-nozzle system only provides a proportional action. However, the stabilization of the proportional action can be effected by appropriate feedback. The details of such a system is considered in a later chapter.

The transfer function of a measurement system has been derived elsewhere (Patranabis, D., *Principles of Industrial Instrumentation*, TMH, 1976) that of a spring type actuator is derived here. The schematic diagram shown in Fig. 3.19 also shows the basic parameters involved in the derivation. Input is m , output x is the stem movement, effective area is A , effective mass of the moving system M , spring constant K and overall damping constant B . When the stem moves by an amount x due to change in input pressure m , the mass spring damping system would obviously be governed by the equation

$$M\ddot{x} + B\dot{x} + Kx = mA \quad (3.43)$$

from which the transfer function is

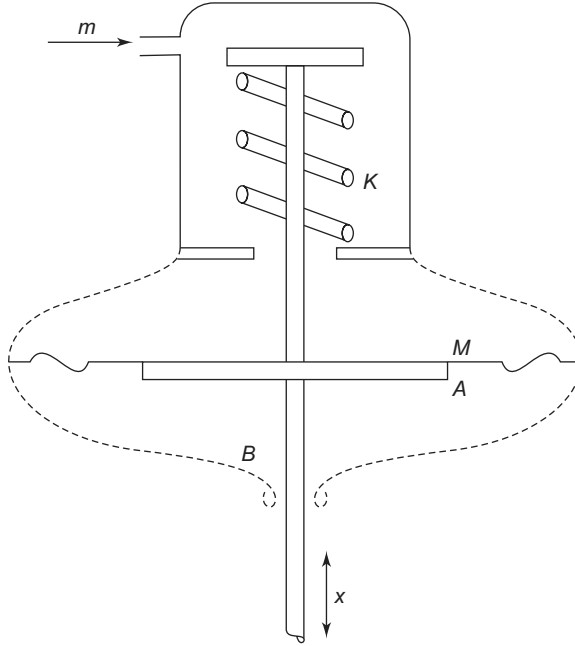


Fig. 3.19 The schematic diagram of a part of a spring type actuator

$$x(s)/m(s) = (A/K)/(s^2/\omega_n^2 + 2\zeta s/\omega_n + 1)$$

where $\omega_n = \sqrt{K/M}$, and $\zeta = (1/2)B/\sqrt{MK}$

Thus the basic transfer function is a second order one. However, in process control applications, the actuator is designed such that its natural frequency of oscillation is very large compared to the system fluctuation rate so that $\omega^2/\omega_n^2 \ll 1$, and often an approximation of the transfer function of the type is considered.

$$x(s)/m(s) = (A/K)/(2\zeta s/(\omega_n + 1)) = K_a/(1 + s\tau_a)$$

3.4 SYSTEM ANALYSIS AND STUDIES OF SYSTEM RESPONSE

It will be of interest to discuss some special control situations from the generalized approach. Consider the block diagram of Fig. 3.20 which easily simplifies to that of Fig. 3.21. From Fig. 3.21

$$(r - cG_m)G_cG_aG_p + uG_L = c \quad (3.44)$$

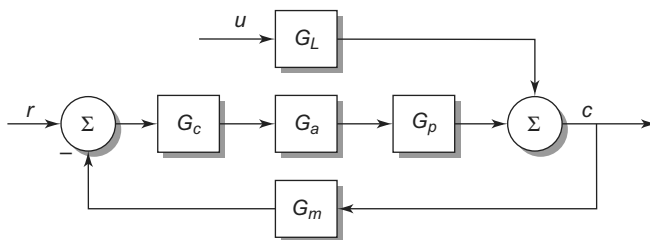


Fig. 3.20 Block diagram of a single loop system showing the different blocks separately

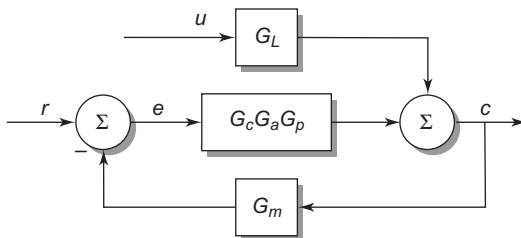


Fig. 3.21 Simplified representation of Fig. 3.20

so that

$$c = r G_c G_a G_p / (1 + G_c G_a G_p G_m) + u G_L / (1 + G_c G_a G_p G_m) \quad (3.45)$$

Usually all the G 's are functions of s . In the most general but simple case, we assume the transfer functions of the measurement system block, actuator block and controller block, respectively, as

$$G_m = k_m / (s\tau_m + 1), \quad G_a = k_a / (s\tau_a + 1), \quad \text{and,}$$

$$G_c = k_c [(s\tau_R + 1) / (s\tau_R)] (1 + s\tau_D)$$

where the controller transfer function is a somewhat simplified form of the expression Eq. (3.42).

For a multicapacity or a multitime constant process, the analytic approach to the control problem becomes quite difficult. However, for low frequency operation, as is usually the case in regulatory problems (see below), the measurement lag may be quite small, $\omega\tau_m \ll 1$. Also, if $\tau_d = 0$, and $\omega\tau_R \ll 1$, the controller transfer function becomes $G_c = \frac{k_c}{s\tau_R}$. Control

with such a controller is called integral control. With integral control there is no offset in the process. The offset or droop is the amount of difference between the output signal and the control point in the steady state. Offset occurs when load disturbance occurs but the control point is not altered to suit this.

From Eq. (3.45) it is easily seen that

$$c/u = (r/u)[G_c G_a G_p]/[1 + G_c G_a G_p G_m] + G_L/(1 + G_c G_a G_p G_m) \quad (3.46)$$

At steady state, c/u is seen to be dependent on the ratio of r/u so that other things remaining the same initially when $u = 0$, $c/r = \frac{G_c G_a G_p}{1 + G_c G_a G_p G_m}$. With

u changing to bring the same ratio of c/r , the ratio u/r should be altered in relation to G_L and the type of u . However, it is possible to show that for a step disturbance, an integral control action can reduce the offset to zero. Assume for brevity $G_m = 1$, and $G_a = 1$, and then a third order process given by $G_p = k_p / \left(\prod_{i=1}^3 (s\tau_i + 1) \right)$ is considered. When r does not change, and

$k_L / \left(\prod_{i=1}^3 (s\tau_i + 1) \right)$ is written for G_L , we have

$$\frac{c(s)}{u(s)} = \frac{K_L / \left(\prod_{i=1}^3 (s\tau_i + 1) \right)}{1 + k_c k_p / \left(s\tau_R \prod_{i=1}^3 (s\tau_i + 1) \right)} = \frac{k_L s\tau_R}{s\tau_R \prod_{i=1}^3 (s\tau_i + 1) + k_c k_p} \quad (3.47)$$

such that when time approaches infinity or $s \rightarrow 0$, $c/u \rightarrow 0$, indicating no offset. Even if proportional and integral actions are considered, a similar result is obtained.

Figures 3.20 and 3.21 also help to evaluate the response of a system to set point change or load disturbance. In a control system, therefore, there are two types of operation, namely the servo operation which occurs when there is no load change but the control point (set point) changes, and the regulator operation which occurs when the set point is fixed but there is load disturbance in the process. The response in either operation may be of two different types, namely, the transient response and the steady state response. Analytical study of these has been made for systems represented by a single block up to an order of two in Sec. 3.2. Study of the same with a general configuration of the process and other control equipment is quite difficult particularly as the system order becomes large. But as it is only necessary to see how the system is affected by controller configuration, reset and rate time variations and proportional band variation since mainly the regular operation is important in process control systems, a brief discussion to that effect may be appended here.

In the block diagram of Fig. 3.21, assume $G_a = G_m = 1$. Load disturbance is considered to affect the entire process but it provides a different steady state gain. Also for a disturbance, the general form of step disturbance is

considered with unit or other magnitude. The regulator operation transfer function (TF) is then

$$c/u = G_L/(1 + G_c G_p) \quad (3.48a)$$

and the servo operation TF

$$c/r = G_c G_p/(1 + G_c G_p) \quad (3.48b)$$

As assumed, $G_L = k' G_p$, so that

$$c/u = k' G_p/(1 + G_c G_p) \quad (3.48c)$$

It has been discussed earlier how the transient response can be obtained for systems upto an order of two. The main approach and interest here is however, to control the system behaviour by τ_R , τ_D and k_c so that prescribed conditions of performance are obtained. For a higher order system the dominant parameter-based approximation is made use of to obtain second order TF in terms of ζ and ω_n as functions of τ_R , τ_D and k_c and then the study of the transient behaviour is made. It should, however, be mentioned that from frequency response the transient behaviour is also predictable.

By way of example, we first take a two-time constant process with proportional action control only. The disturbance is assumed at the intermediate stage of the process as shown in Fig. 3.22. The regulatory TF is easily obtained as

$$\begin{aligned} \frac{c(s)}{u(s)} &= \frac{k_L/(s\tau_2 + 1)}{1 + \frac{k_c k_1 k_2}{(s\tau_1 + 1)(s\tau_2 + 1)}} \\ &= \frac{k_L(s\tau_1 + 1)}{s^2 \tau_1 \tau_2 + s(\tau_1 + \tau_2) + (1 + k_c k_p)} \end{aligned} \quad (3.49a)$$

$$= \frac{k_L}{1 + k_c k_p} \frac{s\tau_1 + 1}{\frac{s^2 \tau_1 \tau_2}{1 + k_c k_p} + s \frac{\tau_1 + \tau_2}{1 + k_c k_p} + 1} \quad (3.49b)$$

Expressing $\omega_n = \sqrt{((1 + k_c k_p)/(\tau_1 \tau_2))}$ and $\zeta = \frac{\tau_1 + \tau_2}{2\sqrt{\tau_1 \tau_2 (1 + k_c k_p)}}$

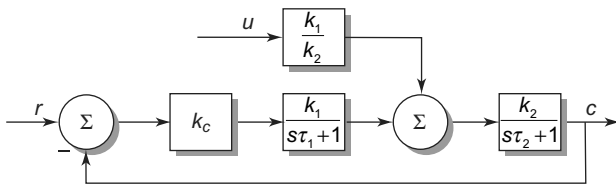


Fig. 3.22 Block diagram of a process control system showing the load disturbance at an intermediate stage of the process

$$\frac{c(s)}{u(s)} = \frac{k_L}{I + k_c k_p} \left[\frac{(s\tau_1 + 1)}{s^2/\omega_n^2 + 2\zeta s/\omega_n + 1} \right] \quad (3.49c)$$

Obviously the proportional control cannot eliminate the offset $k_L/(1 + k_c k_p)$ and its value is dependent on k_c , k_p and k_L .

For a step input, $u(s) = k_s/s$; putting this in Eq. (3.49c) and obtaining the inverse transform for $\zeta < 1$

$$c_1(t) = \frac{k_L k_s}{I + k_c k_p} \left[1 + \sqrt{\frac{1 - 2\zeta\omega_n\tau_1 + \omega_n^2\tau_1^2}{1 - \zeta^2}} \exp(-\zeta\omega_n t \sqrt{1 - \zeta^2} + \phi) \right] \quad (3.50)$$

where $\phi = \tan^{-1} \frac{\omega_n\tau_1\sqrt{1 - \zeta^2}}{1 - \zeta\omega_n\tau_1} + \tan^{-1} \frac{\sqrt{1 - \zeta^2}}{\zeta}$

and for $\zeta = 1$

$$c_2(t) = (k_L k_s / (1 + k_c k_p)) [1 + \{(\omega_n\tau_1 - 1)\omega_n t - 1\} \exp(-\omega_n t)] \quad (3.51)$$

and for $\zeta > 1$

$$c_3(t) = \frac{k_L k_s}{1 + k_c k_p} \left[1 + \frac{T_1 - \tau_1}{T_2 - T_1} \exp(-t/T_1) - \frac{T_2 - \tau_1}{T_2 - T_1} \exp(-t/T_2) \right] \quad (3.52)$$

where $T_{1,2} = \frac{1}{\omega_n} (-\zeta \pm \sqrt{\zeta^2 - 1})^{-1}$

Knowing the process time constants τ_1 , τ_2 the static process gain and k_L , the transient plots may be made. However, k_c may be allowed to vary so that both ω_n and ζ change. As k_c increases, ω_n increases and ζ decreases and oscillatory tendency may be observed in the response. Relative responses are shown in Fig. 3.23 for $\zeta < 1$, cases, when for ζ increasing, offsets are main considerations.

As ζ increases and k_c decreases, offset is seen to increase. Following the normal procedure, the maximum peak can be obtained as

$$c_{1\max}(t) = \frac{k_L k_s}{1 + k_c k_p} \left[1 + \sqrt{\frac{1 - 2\zeta\omega_n\tau_1 + \omega_n^2\tau_1^2}{1 - \zeta^2}} \exp(-\zeta\omega_n t_{\max}) \right] \quad (3.53)$$

where $t_{\max} = \left(\tan^{-1} \frac{\sqrt{1 - \zeta^2}}{\zeta} - \phi \right) / (\omega_n \sqrt{1 - \zeta^2})$

and the decay ratio is

$$\delta = \exp(-2\pi\zeta/\sqrt{1 - \zeta^2}) \quad (3.54)$$

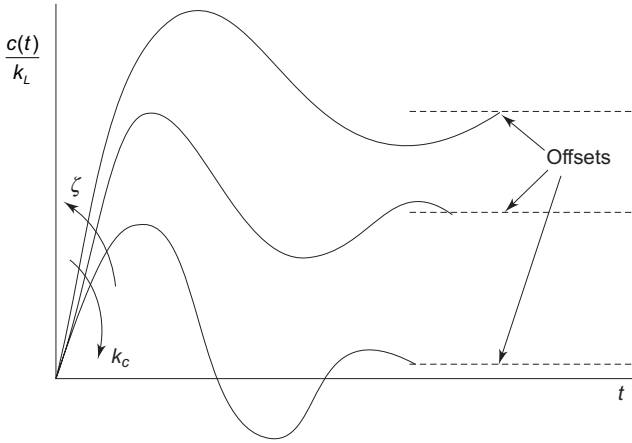


Fig. 3.23 Relative response curves for Fig. 3.22

Exercise 1

If $\tau_2 = \tau_1/4$ and $k_L = k_s = k_p = 1$, obtain how τ_1 and k_c are related for 20 per cent maximum overshoot (show graphically, if necessary).

Replacing the proportional controller with a proportional plus integral controller with $G_c = k_c(s\tau_R + 1)/(s\tau_R)$, the load-output TF is changed as

$$\frac{c(s)}{u(s)} = \frac{k_L[s\tau_R(s\tau_1 + 1)]}{k_c k_p (s\tau_R + 1) + s\tau_R(s\tau_1 + 1)(s\tau_2 + 1)} \quad (3.55)$$

which again shows that the system has no offset. This is a third order system and it is possible to arrange Eq. (3.55) with $u(s) = k_s/s$ as

$$c(s) = \frac{k_L k'_s}{k_c k_p} \frac{sT_2 + 1}{(sT_1 + 1) \left(\frac{s^2}{\omega_n^2} + \frac{2\zeta s}{\omega_n} + 1 \right)} \quad (3.56)$$

where $k'_s = k_s \tau_R$ and $T_2 = \tau_1$,

For $\zeta < 1$, from Eq. (3.56), one obtains

$$\begin{aligned} c_1(t) = & \frac{k_L k'_s}{k_c k_p} \left[\frac{1 - 2T_2 \zeta \omega_n + T_2^2 \omega_n^2}{1 - 2T_1 \zeta \omega_n + T_1^2 \omega_n^2} \right]^{1/2} \\ & \frac{\omega_n T_2}{\sqrt{1 - \zeta^2}} \exp(-\zeta \omega_n t) \sin(\omega_n t \sqrt{1 - \zeta^2} + \phi) \\ & + \frac{(T_1 - T_2)^2 \omega_n^2 \exp(-t/T_1)}{1 - 2T_1 \zeta \omega_n + T_1^2 \omega_n^2} \end{aligned} \quad (3.57)$$

where
$$\phi = \tan^{-1} \left(\frac{\omega_n T_2 \sqrt{1 - \zeta^2}}{1 - \zeta \omega_n T_2} \right) - \tan^{-1} \left(\frac{\omega_n T_1 \sqrt{1 - \zeta^2}}{1 - \zeta \omega_n T_1} \right)$$

In this case, if $\zeta \rightarrow 0$, the system tends to be unstable. For specific ζ , which again is system dependent, the maximum overshoot and decaying ratio can be evaluated following the usual procedure, and by controlling k_c and τ_R , these can be controlled to the desired values. It is possible to make the gain very large in some cases where τ_R provides the necessary control and τ_R has a limiting value on the basis of system stability.

Exercise 2

In the earlier exercise, if $\tau_1 = 1$ min, obtain the relationship between k_c and τ_R for 20 per cent overshoot.

By adding a derivative action such that $G_c = k_c(1 + 1/s\tau_R + s\tau_D)$, the system order does not alter. But since derivative action increases controllability, it is customary in multiple lag processes to use derivative action with $\tau_D < \tau_j$ (smallest) where τ_j denotes process time constant. The peak error (without integral action) and oscillation period are both reduced as k_c can be increased when derivative action is incorporated in a controller. The system under discussion can be shown to have the *TF*

$$\frac{c}{u} = \frac{k_L}{k_c k_p} \frac{s\tau_R(s\tau_1 + 1)}{\frac{\tau_1 \tau_2 \tau_R s^3}{k_c k_p} + \frac{\tau_R(\tau_1 + \tau_2 + k_c k_p \tau_D)s^2}{k_c k_p} + \frac{k_c k_p + 1}{k_c k_p} \tau_R s + 1} \quad (3.58)$$

so that a transient solution of the form of Eq. (3.57) is obtained and the appropriate choice of k_c , τ_R and τ_D can be made for the desired transient response of the process. If $\tau_R \rightarrow \infty$.

$$\begin{aligned} \frac{c}{u} &= \frac{k_L}{k_c k_p} \frac{s\tau_1 + 1}{\frac{\tau_1 \tau_2}{k_c k_p} s^2 + \left(\frac{\tau_1 + \tau_2 + k_c k_p \tau_D}{k_c k_p} \right) s + \frac{k_c k_p + 1}{k_c k_p}} \quad (3.59) \\ &= \frac{k_L}{1 + k_c k_p} \frac{s\tau_1 + 1}{\{\tau_1 \tau_2 / (1 + k_c k_p)\} s^2 + (\tau_1 + \tau_2 + \tau_D k_c k_p) s / (1 + k_c k_p) + 1} \quad (3.60) \end{aligned}$$

The similarity between Eqs (3.59) and (3.49b) is obvious, except that Eq. (3.59) has a larger ζ because of the extra term $k_c k_p \tau_D$ in the numerator of the second term of the denominator. From Eqs (3.53) and (3.54) it will be seen that this increase in ζ will reduce the peak and increase the decay

ratio, resulting in an improvement in controllability, as has already been mentioned.

3.5 GENERALIZATION WITH LOAD CHANGES AT ARBITRARY POINTS

Large processes with a greater number of time constants will be similarly treated but then direct analytic approach fails there and normally a numerical analysis is applied. However, often a dominant second order equivalence is considered and analysis follows as such. If the response to load disturbances at every 'capacitive' block is to be evaluated, the Nichols' diagram (more commonly called the Black Nichols chart in process control) is made use of. This is a chart in coordinates of $|G|$ and $\angle G$ of different $|G/(1+G)|$ and $\angle G/(1+G)$. For use of this chart the transfer function has to be a little modified to include $G/(1+G)$. Referring to Fig. 3.24, one easily obtains

$$c/u_q = G_{uq}/(1+G) = (G_{uq}/G)(G/(1+G)) = (G_{uq}/G) (c/r)$$

where

$$G_{uq} = \frac{\prod_{j=q}^n k_j}{\prod_{j=q}^n (s\tau_j + 1)} \quad \text{and} \quad G = \frac{k_c \prod_{i=1}^n k_i}{\prod_{i=1}^n (s\tau_i + 1)}, \quad \text{such that}$$

$$\frac{c}{u_q} = \frac{\prod_{j=1}^{q-1} (s\tau_j + 1)}{k_c \prod_{j=1}^{q-1} (s\tau_j + 1)} \frac{G}{1+G}$$

Thus using the chart, $G/(1+G)$ is known both in magnitude and phase (however, the magnitude and phase of G are obtained from Bode plot) and then for magnitude of c/u_q , $|G/(1+G)|$ is divided by $k_c \prod_{j=1}^{q-1} k_j$ and multiplied by $\prod_{j=1}^{q-1} (s\tau_j + 1)$. The phase of c/u_q is not of much consequence. One can now see the differences in the magnitudes of c/u_q and c/u_{q-1} or c/u_{q+1} , i.e., generally in magnitudes of c/u_q , $q = 1, 2, \dots, n$. Considering two consecutive points of load change, one gets

$$m_c = |c/u_{q+1}|/|c/u_q| = (s\tau_q + 1)/k_q$$

and two arbitrary points of load change, the ratio of magnitudes is

$$m_a = |c/u_{q+1}|/|c/u_{q-w}| = \frac{\prod_{q-w}^q (s\tau_j + 1)}{\prod_{q-w}^q k_j}$$

A discussion pertaining to the magnitudes of c/u_q and m_c , m_a is now in order. At low frequencies, i.e., when $\omega\tau_j \ll 1$

$$|c/u_q| \approx \frac{\prod_{j=q}^n k_j}{\left(1 + k_c \prod_{j=1}^n k_j\right)}$$

which may have values larger or smaller than unity. For increasingly higher frequencies if the response is initially steady and then monotonically decreases, k_c has been adjusted adequately or more than adequately. Alternatively, there may be a resonant condition at a particular frequency and in such a case the τ_j 's and k_j 's are important for evaluating the most sensitive point of disturbance, m_a and m_c help in this evaluation.

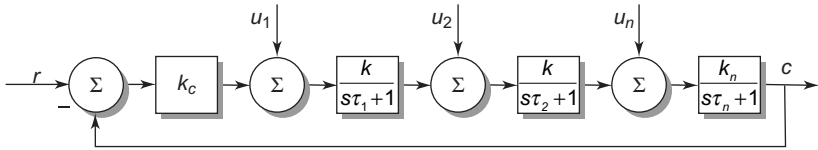


Fig. 3.24 Representation of a feedback system with upsets at n different stages

Example 1

Refer to Fig. 3.24. There are three points of upsets such that $n = 3$, and $\tau_1 = 1$ sec, $\tau_2 = 2$ sec, $\tau_3 = 3$ sec, $k_1 = 2$, $k_2 = 1$, $k_3 = 2$ and $k_c = 3$. Obtain the zero frequency closed loop gains for the load changes and closed loop gain for changes in u_2 , evaluating its maximum gain and the corresponding frequency.

Solution For zero frequency cases:

$$|c/u_q| = \frac{\prod_{j=q}^3 k_j}{\left(1 + k_c \prod_{j=1}^3 k_j\right)}$$

$$\text{For } q = 1, |c/u_1| = (2 \cdot 1 \cdot 2)/(1 + 3 \cdot 2 \cdot 1 \cdot 2) = 4/13$$

$$\text{For } q = 2, |c/u_2| = (1 \cdot 2)/13 = 2/13$$

$$\text{For } q = 3, |c/u_3| = 2/13$$

for gain in general cases with disturbance at point 2

$$\begin{aligned} c/u_2 &= \frac{s+1}{3.2} \cdot \frac{4.3}{(s+1)(2s+1)(3s+1)} \left/ \left[1 + \frac{4.3}{(s+1)(2s+1)(3s+1)} \right] \right. \\ &= 2(s+1)/(6s^3 + 11s^2 + 6s + 1) \end{aligned}$$

Therefore, putting $s = j\omega$, $G/(1+G) = 12/[(13 - 11\omega^2)^2 + 36\omega^2(1 - \omega^2)]^{1/2}$ and $(s+1)/6 = (1 + \omega^2)^{1/2}/6$.

Now for different ω the magnitudes can be calculated and the results plotted. Magnitude would be maximum either at $\omega = (13/11)^{1/2} = 1.087$ or at $\omega = 1$, when gain is 2.51 or 1.41 respectively. Gain at $\omega = 0$ rad/sec is 0.153; hence maximum gain is 2.51 at a frequency 1.087 rad/sec.

Review Questions

- 1. A system with direct negative feedback and acceleration positive feedback is used for controlling a certain parameter. Obtain its overall transfer function and draw the block schematic representation of the system.
If the process transfer function is $k/(s(sT + 1))$, obtain the gain constant of the system.
- 2. Obtain the transfer function of the control system shown in Fig. Q-3.1.

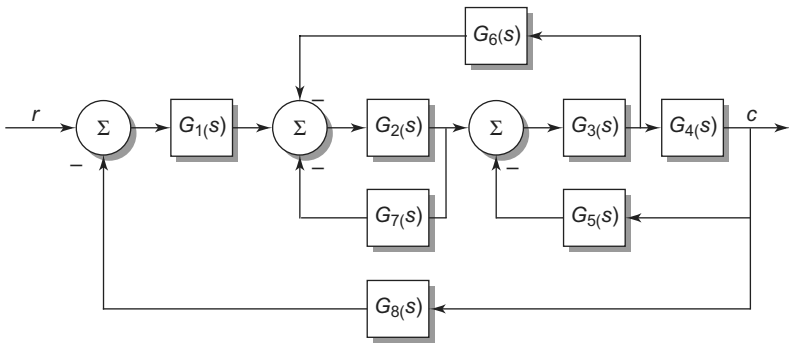


Fig. Q-3.1

- 3. The block diagram of a process control system is shown in Fig. Q-3.2. Obtain its transfer function. If the load disturbance occurs at points x

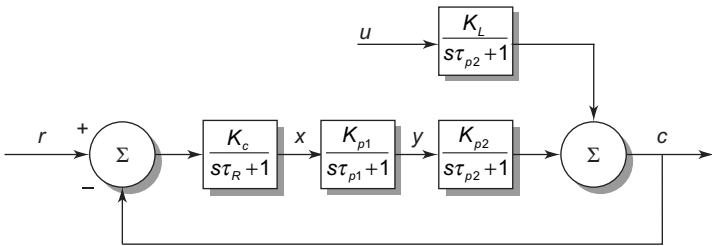


Fig. Q-3.2

or y , obtain the corresponding block diagrams with the individual block transfer functions. If $\tau_R = 0$, obtain the response characteristics for $\tau_{p1} = 4\tau_{R2}$. Calculate the offsets in the individual cases.

4. For the system of Fig. Q-3.3, if $K_{p1} = K_{p2} = 1$, $K_L = 0$ and $\tau_R = 0$, obtain the relation between K_c and $\rho = \tau_{p1}/\tau_{p2}$ for critical damping.
(Ans: $K_c = (1 - \rho)^2/4\rho$)

5. In a first order temperature process, process time constant is 5 min and actuator time constant is 5 sec. Calculate the proportional band for proportional action only to produce critical damping.

(Hint: $\tau_{p1} = 5$ sec, $\tau_{p2} = 5$ min = 300 sec, $K_c = (1 - 1/60)^2/(4/60) \cong 15$, hence $PB = 100/15\% = 6.67$ per cent)

6. A unity feedback closed-loop control of a process $G_p = 1/(s\tau + 1)$ is performed by a controller $G_c = K_c(1 + 1/sT_R + sT_D)$ with $G_a = 1$. Calculate the overall system damping for $K_c = 0.2$ and 2, offset in the two conditions for unit step change in the load acting before the process. Assume $\tau = 2$ min and $T_R = T_D = 1$ min.

7. A unity feedback system has $G_c = K_c$, $G_{p1} = 1/(2s + 1)$, $G_{p2} = 3/(2s + 1)$ and $G_a = 1$. With unit step change in the load, obtain the offset, considering that load change occurs before (i) G_{p1} and (ii) G_{p2} .

8. What is the interaction factor in a three-term controller? If derivative time increases by 10 per cent in its original value of derivative time to integral time ratio of 1/5, by how much does the reset action change?

(Ans: 1.66% dec.)

9. In a single loop control, the process is represented by a first-order model with a lag of 1 min and static process gain of 0.4. For an offset to be limited to 1% of a set point change of 1 unit step, what should be the controller gain when only P-action is given?

[Hint: Refer to Eq. (3.25): Offset = $0.01 = 1/(1 + K_c \times 0.4)$, $K_c = 247.6$]

If the set point is kept fixed, what would be the offset for the above controller setting of a unit step disturbance?

[Hint: Refer to Eq. (3.26): Offset = $-1/(1 + 247.5 \times 0.4) = -0.04$]

10. In a 3-term interactive type PID controller, the controller parameters are adjusted as $K_c = 13$, $\tau_R = 10s$, $\tau_D = 0.2s$, what is the interaction factor?

[Hint: Refer to Eq. (3.42): $\phi = 1 + \tau_D/\tau_R = 1 + 0.25/10 = 1.025$]

If the integral action time is increased by 22%, what would be the changes in the proportional gain and derivative time?

$[\tau_R$ changes to

(i) $10 (1 + 0.22) = 12.2s$, changing ϕ to $1 + 0.25/12.2 = 1.0205$

(ii) $10 (1 - 0.22) = 7.8$, changing ϕ to $1 - 0.25/12.2 = 1.032$

Hence, $K_c \phi$ changes from 13×1.025 , i.e., 13.325 to

(i) $13 \times 1.0205 = 13.26$, i.e. -0.487%

(ii) $13 \times 1.032 = 13.416$ i.e. 0.682%

τ_D changes to

(i) $\tau_D/\phi = 0.25/1.0205 = 0.2449$, i.e. 0.410% , the original τ_D/ϕ being 0.2439%

(ii) $\tau_D/\phi = 0.25/1.032 = 0.2422$, i.e. 0.677%

Controllability and Stability

4.1 INTRODUCTION

In any control system, problem of controllability is of prime importance. Only when the process is controllable, can an adequate control scheme be prepared for it. Stability of the process follows, to a large extent, from considerations of controllability. However, given a process control loop, the stability of the loop should be considered using an appropriate prescribed method. Two important methods are the root-locus technique and Bode-plot technique. Another one often used for ready reference in some typical cases is the Routh criterion. Root-locus technique is suggestive in nature while the Bode-plot technique is very appropriate for process control systems from which the choice of the controller and its parameters are logically obtained.

4.2 CONTROLLABILITY

When a particular control scheme is adopted, it is necessary that by adjusting the controller parameters the plant is made controllable. Here “controllable” means that the desired results from the plant/process are obtained within reasonable tolerable limits in a specified time. Whenever a disturbance or load variation occurs, the process should return to the set conditions within the prescribed time and with prescribed peak error.

The usefulness of a control system is marked by two quantities, (i) the largest deviation of the controlled variable from its set point, and, (ii) duration of the deviated response. The most favourable situation would

mean minimum of both these quantities. Combining the two, one derives that the area enclosed between the response to a step disturbance and the straight line through its expected rebalanced state should be minimum. For normalized response-time plot, this area is of the dimension of time. Thus, if the response follows essentially an aperiodic path and if the area as stated above is minimum, then such a control is optimum. Controllability is not exactly what is called optimal controlled condition or control effectiveness but is basically interpreted by the offset condition as mentioned below. Prime factor is the deviation or the peak deviation, reduction of this also reduces the duration of deviation and hence deviation reduction is the main objective.

If we consider the system to be split up into two parts, the process and the control equipment, and make the loop open at a convenient point to give a sinusoidal input to the open loop system (Fig. 4.1), then the system open loop gain is

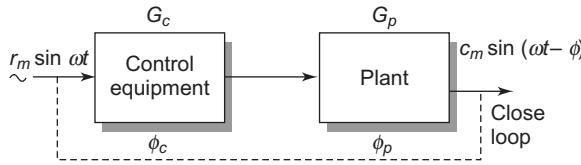


Fig. 4.1 Representation of an open loop system to feed sinusoidal input

$$G = G_c G_p \quad (4.1)$$

and the system phase lag is

$$\phi = \phi_c + \phi_p \quad (4.2)$$

One can also express G as

$$G = G_c G_p = G_c K_p / A \quad (4.3)$$

where K_p is the static gain of the process/plant and A is a dimensionless attenuation factor obtainable from the frequency response curves of the open loop system. It must be remembered that all the while the controller has been assumed to be a combined proportional, integral and derivative action type; ϕ_c in such a case is the phase lag introduced by the other parts like the actuator, measurement system, etc. along with the integral action and partly a phase lead due to derivative action. The control equipment gain can be split up as

$$G_c = K_c G_q \quad (4.4)$$

where K_c is the contribution due to proportional action only and G_q due to the rest of the control equipment. Thus from Eqs (4.1) to (4.4)

$$G = K_c K_p G_p / A = K_f G_q / A \quad (4.5)$$

where K_f is known as the 'proportional control factor' and is defined as the idealized system gain when other control equipment have a gain of unity.

Consider that the system loop is closed and a step disturbance occurs, generally, in the plant affecting the loop; the controlled condition should be such that a damped sinusoidal is obtained at the output and no offset results. The recovery time is determined by the number of oscillations before the normal control point is regained, as also its period. This also means that offset will be zero and deviation negligible. If such a controlled condition is obtained, then the plant is controllable.

If the disturbance of magnitude, $u(\text{step})$, occurs at the starting of the process, the idealised output is $c_p = K_p u$ but because of the filtering effect, the time constants of the system and the control applied, the maximum value of c is only $c_m < c_p$, as shown in Fig. 4.2(a). From this figure a few more important terms are defined.

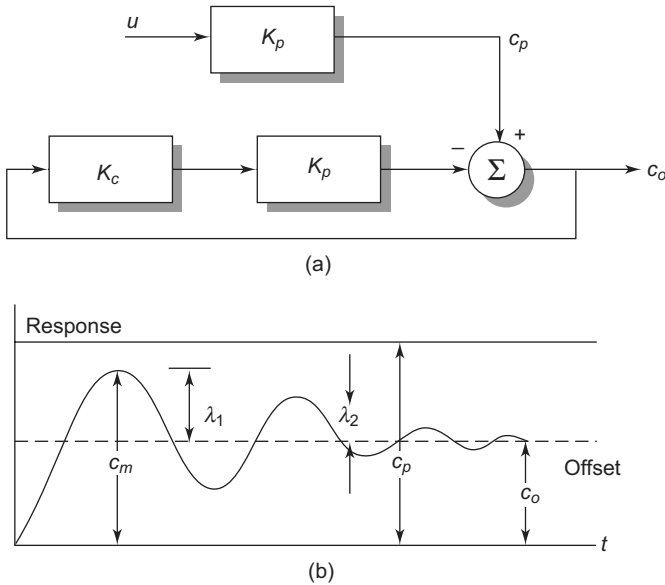


Fig. 4.2 (a) Block schematic of a system showing offset (output) with load change, (b) Time response curve of a process with load change

- (i) The deviation reduction factor (DRF), δ_f is the ratio of the deviation occurring in the plant output in the absence and presence of control action when a step disturbance occurs in the plant; obviously, this is

$$\delta_f = c_p / c_m \quad (4.6)$$

For an offset c_o , this is given by

$$\delta_f = (c_p - c_o) / (c_m - c_o) \quad (4.7)$$

(ii) The subsidence ratio (SR) is defined as

$$\rho = \lambda_1/\lambda_2 = \lambda_n/\lambda_{n+1} \quad (4.8)$$

As seen from the above Fig. 4.2(a), c_p is the deviation to be reduced if c_o is the offset (because of proportional control only); for the present actually only proportional control action is considered and the filtering effects of the loop are disregarded. This offset c_o has been obtained by control action with a gain K_c and from Fig. 4.2(b), one can write

$$K_c K_p c_o = c_p - c_o \quad (4.9)$$

giving offset as

$$c_o = c_p / (1 + K_c K_p) = c_p / (1 + K_f) \quad (4.10)$$

Also, with an equivalent second order form of response, the time response is given as (see Chapter 3),

$$c = c_o [1 - \{\exp(-\zeta \omega_n t) / \sqrt{1 - \zeta^2}\} \sin(\omega_n t \sqrt{1 - \zeta^2} + \phi)] \quad (4.11)$$

from which c_m is derived as

$$c_m = c_o [1 + \exp(-\zeta \omega_n t_m)] \quad (4.12)$$

and ζ is obtained from

$$1/\rho = \exp(-2\pi\zeta/\sqrt{1 - \zeta^2}) \quad (4.13)$$

The natural frequency of oscillation ω_n of the system is obtained from

$$\omega_n \sqrt{1 - \zeta^2} / (2\pi) = 1/T \quad (4.14a)$$

or

$$\omega_n \sqrt{1 - \zeta^2} / \pi = 1/t_m \quad (4.14b)$$

where t_m is the time at which the peak deviation c_m occurs.

Hence, from Eqs (4.7), (4.10) and (4.12),

$$\begin{aligned} \delta_f &= c_p K_f / [(1 + K_f) c_o \exp(-\zeta \omega_n t_m)] \\ &= \frac{K_f / (1 + K_f)}{\exp(-\zeta \omega_n t_m) / (1 + K_f)} \\ &= K_f \exp(\zeta \omega_n t_m) \end{aligned} \quad (4.15)$$

Also, from Eqs (4.6), (4.10) and (4.12)

$$\delta_f = c_o (1 + K_f) / [c_o (1 + \exp(-\zeta \omega_n t_m))] \quad (4.16)$$

Using Eqs (4.13) and (4.14), Eqs (4.15) and (4.16) can be recast as

$$\delta_i = K_f \exp(-\pi\zeta/\sqrt{1-\zeta^2}) = K_f \sqrt{\rho} \quad (4.17)$$

and

$$\begin{aligned} \delta_f &= (1 + K_f) / [1 + \exp(-\pi\zeta/\sqrt{1-\zeta^2})] \\ &= (1 + K_f) / (1 + \sqrt{1/\rho}) \end{aligned} \quad (4.18)$$

It is thus clear that the deviation-reduction factor, δ_f , of the control system will be large and correspondingly the system is more controllable if K_f and the subsidence ratio, ρ , are large.

In analysing the controllability of a process, a step type disturbance or signal variation is considered as this is the most difficult type because of sudden changes in the value of the process condition. For example, if one considers frequency, this would mean that an infinite band of frequencies is contained in this type of disturbance. Once controllability is ensured for this type of disturbance—because of the infinite band of frequency content, the resonant condition should also be taken care of—other types of disturbances can be automatically controlled. As controllability is primarily determined by c_p/c_m , it will be of interest to see how c_p/c_m is actually related to the process reaction curve (PRC) which is defined as the response curve of the process in the open loop condition with a step disturbance at the start of the process. If a step disturbance occurs at the demand or the load side, the response curve is termed as the load reaction curve (LRC).

When we consider the controllability of the process without referring to the control equipment, the c_m value has to be redefined with reference to the process reaction curve. For obvious reasons the load reaction curve is rarely considered in such an analysis. The response curve of the plant/process because of step disturbance $u = c_p/K_p$ is shown in Fig. 4.3. Till the point (c_o, τ_d) is reached in this curve, the value c_o being quite small, the measuring unit will not respond to send signals to the controller-actuator unit. Even when the actuator unit provides corrective action after this point is reached and the measuring block gives out the signal, the action-signal has to pass through the process following a similar curve. The actual controlled operation starts after a time $2\tau_d$ and only at that instant, or afterwards, can a variation in controlled variable be observed. Time τ_d here is referred to as the dead time, transportation lag or distance/velocity lag. The value of c after a time $2\tau_d$ i.e., c_m is the minimum deviation of the controlled condition before the control action actually starts. Thus the largest deviation reduction factor for this process is given by c_p/c_m . This can be improved by improving the sensitivity and resolution of the measuring unit. After the control action starts, recovery also begins. A part of the recovery curve is also shown in Fig. 4.3 in which the peak deviation

is denoted as c'_m . The usual deviation reduction factor is, therefore c_p/c'_m which is less than c_p/c_m . This analysis shows that as the transportation lag increases, DRF decreases and the controllability correspondingly decreases. This is determined, in actuality, by the slope of the curve as well as the initial response. The uncertainties that are apparent in the system are the values of c_o , and, therefore, those of c_m . However, approximation is often possible by drawing the mean slope of the curve shown by ab in which case oa is the transportation lag.

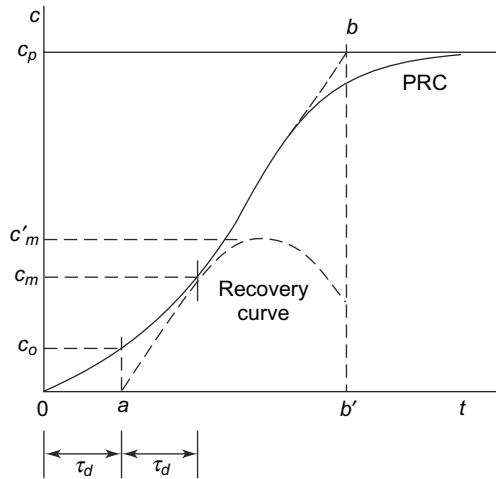


Fig. 4.3 Process reaction curve (PRC)

If disturbance occurs at an intermediate stage of the process, then the response to the disturbance will obviously be steeper while the response to actuation will be flatter. This has the effect of decreasing the controllability as is evident from Fig. 4.4. If, however, the disturbance occurs at the end of the last stage, that is, before the measuring unit the controllability is $c_p/c_m = 1$. This is easy to see from the foregoing consideration. In such cases, c_p becomes identical with c_m . An intermediate stage disturbance means a demand side disturbance and this obviously deteriorates the controllability. This should provide a clue to good design for better controllability to the process designer.

The above qualitative study of controllability has been extended by many to check how dead time affects control quality even when the loops are properly tuned (see Chapter 5 next for tuning). For the system with a dead time a controllability ratio

$$r = \tau_d/\tau_p \quad (4.19)$$

has been defined, where τ_p is the process transfer lag of a first order model or summation of process times of higher order model of the process. As

discussed in the next chapter there are some empirical tuning schemes of the controller. Of these prominent are those by Ziegler and Nichols and Cohen and Coon. It has been demonstrated how the proportional control factor, peak deviation and oscillation period vary with r under these tuning conditions. In fact, K_f is shown to decrease steadily as r increases for a large value of which K_f assumes a value of 0.45 for Ziegler Nichol's method and 0.08 for Cohen-Coon's tuning of the loop. For both these cases, the peak overshoot approaches 1 as $r \gg 1$ amounting to saying that it has virtually become open loop where as the period of oscillation approaches nearly three times the dead time as r becomes larger and larger. If it is assumed that the controlled variable returns to the set value in three periods of oscillations, for larger r , the settling time becomes 9 times the dead time. The overall picture thus emerges as the dead time increases, control becomes very poor. This, however, follows straightaway from the process reaction curve as shown earlier.

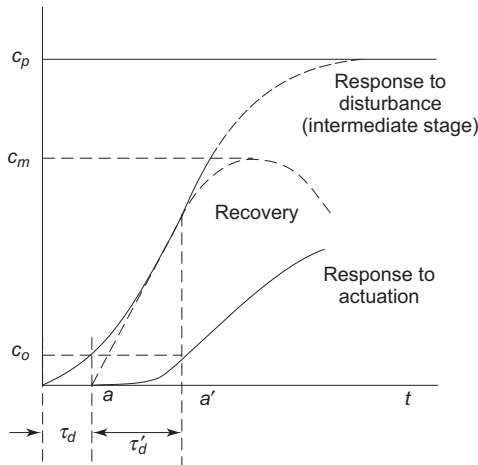


Fig. 4.4 PRC illustrating the controllability with disturbances at intermediate stage

4.2.1 Controllability from Gain-Bandwidth Product

Controllability is checked through a performance criterion of a closed loop system in the time domain which, in turn, is obtained from the error integral

$$E = \int_0^{\infty} |e(t)| \cdot dt \quad (4.20)$$

where $e = r - c$. E is often referred to as integral of absolute error (IAE) and this should be minimum. The deviation is considered for a step change in this specification, time t being measured from the time this change occurs.

Another criterion, often favoured but difficult to evaluate, is the integral time absolute error (ITAE) which weights deviation more and more as time increases and consequently provides rapid 'line-out', i.e., response falls in line with the expected value.

$$E_t = \int_0^{\infty} |e(t)| \cdot t \cdot dt$$

Another criterion is the integral squared error (ISE) criterion which also has its use in performance evaluation in control systems. This is given as

$$E_s = \int_0^{\infty} e^2(t) dt$$

Usually ISE criterion is considered for elimination of large error, as square of large error becomes even more large and the integral value is also large.

If the error persists over a long time, ITAE is a good choice as this would give a large value of the integral while for smaller error values, the IAE criterion is selected which gives adequate weightage to the error, positive or negative.

The frequency domain specification is analogously stated as, the gain bandwidth product should be a maximum. But the gain and bandwidth are a little different in this case from the conventional definition. If the loop is opened and a sinusoidal input is given whose frequency is varied, then the frequency, ω_n , at which the loop phase changes by -180° is called the critical frequency and the band of frequencies from 0 to this value is known as the bandwidth. At this frequency the loop will oscillate when closed if the controller gain is gradually increased from zero to unity. The gain at which the oscillation just starts, is known as the ultimate gain (that can be given for stable operation), K_m . The gain bandwidth product is then given as $K_m \omega_n$. The product $K_m \omega_n$ is, in fact a measure of controllability. Consider the block representation of the system shown in Fig. 4.5 where the steady state gains are marked inside the blocks. For a step input u , the steady state response may be given as

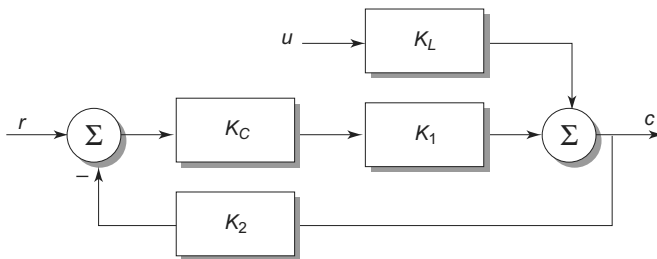


Fig. 4.5 Block diagram of a simple feedback system

$$c = K_L u / (1 + K) \quad (4.21)$$

where $K = K_1 K_2 K_c$. The transient response, assuming a second order characteristic of the loop function, is given as

$$c = (1 - (\exp(-\zeta \omega_n t) / \sqrt{1 - \zeta^2}) \sin(\omega_n t \sqrt{1 - \zeta^2} + \phi)) u$$

$$\phi = \tan^{-1}(\sqrt{1 - \zeta^2} / \zeta) \quad (4.22)$$

From the 25 per cent decay ratio which is the reciprocal of the subsidence ratio, ζ can be evaluated as follows.

Decay ratio, DR , is $\exp(-2\pi\zeta/\sqrt{1 - \zeta^2})$. With $DR = 1/4$, ζ becomes 0.22. Also, one has

$$\frac{\text{Maximum overshoot}}{\text{Final value}} = \exp(-\pi\zeta/\sqrt{1 - \zeta^2}) = (DR)^{1/2} = 1/2 = 0.5$$

so that the peak error is 1.5 times the steady state error.

For a load change if the steady state error is $K_L/(1 + K)$, then the peak error is $1.5 K_L/(1 + K)$ (Cf. Eqs 3.26 and 4.12). Integral action eliminates the offset but not the peak error, since it occurs before the integral action starts.

Now, the optimum controller setting would be done to have damped oscillation with the occurrence of a limited number of peaks. The frequency of this oscillations is $\omega_n \sqrt{1 - \zeta^2}$. Depending on ζ its value is 10 to 30 per cent less than ω_n which is the critical (resonance) frequency. It should be remembered that this ω_n is dependent on the system and is found from a 'dominant' second order equivalence. Also, the value of DR is assumed to be 1/4, the error integral $E = \int_0^\infty |e(t)| \cdot dt$ is roughly proportional to the

area under the first 1/2-wave which may be computed with reference to the response curve shown in Fig. 4.6 as follows;

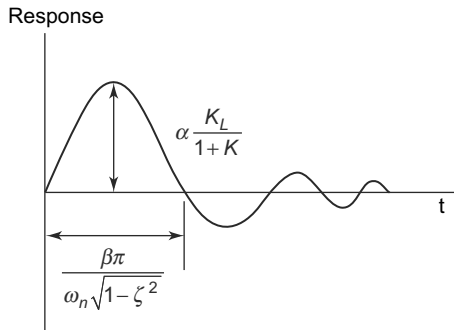


Fig. 4.6 Response to a steep disturbance to the system

The first half-wave is considered to be a triangle so that area under it would be $1/2$ (base) (altitude). As obtained already the altitude is $\alpha K_L/(1 + K)$, where α is a constant dependent on ζ and the base $\beta\pi/(\omega_n\sqrt{1-\zeta^2})$, where β is a constant whose value is easily determined so that the areas is $\alpha\beta\pi K_L/(2(1 + K)(\omega_n\sqrt{1-\zeta^2}))$ approximately. If, however, all the possible parts of the wave are considered and each being taken as a triangle, the consecutive peak values are in the ratio of $1 : 1/2 : 1/4 : 1/8 \dots$ and so on whereas there is substantially no change in the base length. This gives the total area as

$$A = \frac{\alpha\beta\pi K_L}{2(1 + K)\omega_n\sqrt{1-\zeta^2}} (1 + 1/2 + 1/4 + 1/8 + \dots) \quad (4.23)$$

$$= \alpha\beta K_L \pi / (1 + K)(\omega_n\sqrt{1-\zeta^2})$$

In a given system K_L is fixed and by proper choice and design of the system both K and ω_n are maximised. If now ζ is also fixed and $K \gg 1$, the performance can be judged from a comparison of two cases marked with suffixes 1 and 2. Thus one obtains

$$\frac{\left[\int_0^\infty |e(t)| dt \right]_1}{\left[\int_0^\infty |e(t)| dt \right]_2} = \frac{(K\omega_n)_2}{(K\omega_n)_1} \quad (4.24)$$

or, in other words, maximising both K and ω_n the error integral can be minimised. It is evident from Eq. (4.24) that $E \propto 1/(K \cdot \omega_n)$ and minimization of E would actually connote maximization of $K\omega_n$.

4.2.2 Generalized Definition of Controllability

In the above, controllability has been studied in terms of process reaction curve and in terms of gain-bandwidth product, for single input and single output systems. For many of the industrial processes, controllability is not that easily ascertained. More recently, concept of controllability (and observability) has been introduced from a little different angle and for this some mathematical tests have been prescribed. These tests involve what are known as system states and state variables. Before going into these tests of controllability we introduce state variables very briefly. It is, however, assumed that the reader does have the requisite background. Whatever is mentioned here, serves to freshen up the memory.

So long we have been dealing with two types of variables, input and output; for modelling or representation of the dynamics of a system, a third type is introduced called the state variables. When a body of mass m is acted on by a force f to move on a frictionless track, the system dynamics is represented by two basic equations:

- (i) mass $\times$ acceleration = force, and,
- (ii) rate of change of displacement = velocity, i.e.,

$$\frac{d(v(t))}{dt} = f(t)/m$$

or,

$$v(t) = (1/m) \int f(t)dt + v(t_0) \quad (4.25)$$

and $dx(t)/dt = v(t)$
or

$$x(t) = \int v(t)dt + x(t_0) = \int ((1/m) \int f(t)dt)dt + v(t_0) \int dt + x(t_0) \quad (4.26)$$

Obviously, displacement can be calculated at any time $t \geq t_0$ if applied force $f(t)$ is known after $t = t_0$ when initial velocity $v(t_0)$ and displacement $x(t_0)$ are given. These $v(t_0)$ and $x(t_0)$ are termed as the states of the system at $t = t_0$ and for this system $x(t)$ and $v(t)$ would be called the state variables with $f(t)$ as the input variable. The variable $x(t)$ is also the output of the system here.

A minimal set of variables called state variables can express the state of a dynamical system. These variables at $t = t_0$ along with the inputs at $t \geq t_0$ can completely determine the behavior of the system for $t > t_0$. In control system the usual notation for state variables is $x(t)$, for input or control variable it is $u(t)$ and output variable is $y(t)$.

As seen in Eqs (4.25) and (4.26), one state variable $x(t)$ is representable by another state variable $v(t)$ and the input variable $f(t)$; the case may be extended for the general case of n state variables and m inputs as

$$dx_j/dt = \dot{x}_j = \phi_j(x_1, x_2, x_3, \dots, x_j, \dots, x_n; u_1, u_2, \dots, u_m) \quad (4.27a)$$

or,

$$x_j(t) = x_j(t_0) + \int_{t_0}^t \phi_j(x_1, x_2, \dots, x_j, \dots, x_n; u_1, u_2, \dots, u_m)dt \quad (4.27b)$$

$$j = 1, 2, \dots, n.$$

n such equations are possible so that a vector form representation is

$$\dot{\mathbf{x}}(t) = \phi(\mathbf{x}(t), \mathbf{u}(t)) \quad (4.28)$$

where

$$\mathbf{x}(t) = [x_1(t), x_2(t), \dots, x_n(t)]^T$$

and

$$\mathbf{u}(t) = [u_1(t)u_2(t),..., u_m(t)]^T \quad (4.29)$$

As is evident, output vector is also a function of the state vectors and the input and control vectors, so that

$$y(t) = \psi[\mathbf{x}(t), \mathbf{u}(t)] \quad (4.30)$$

where $y(t) = [y_1(t), y_2(t), ..., y_p(t)]^T$ is a $p \times 1$ vector.

For linear time invariant systems, one can express the variable $\dot{\mathbf{x}}$ as a linear combination of system state variables and input variables, i.e.,

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) \quad (4.31)$$

where $\mathbf{A}$ is $n \times n$ system matrix

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

and $\mathbf{B}$ is $n \times m$ input matrix

$$\begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1m} \\ b_{21} & b_{22} & \cdots & b_{2m} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ b_{n1} & b_{n2} & \cdots & b_{nm} \end{bmatrix}$$

Figure 4.7 shows the block representation of Eq. (4.31). From the same scheme, the output can be written as

$$\mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t) \quad (4.32)$$

where $\mathbf{C}$ is a $p \times n$ output matrix and $\mathbf{D}$ is $p \times m$ transmission matrix. If $\mathbf{A}$ is a diagonal matrix, the system representation is called the canonical variable form or normal form. If the system transfer function is, with $a_o = 1$,

$$y(s)/u(s) = \sum_{j=0}^n b_j s^{n-j} \bigg/ \left(\sum_{i=0}^n a_i s^{n-i} \right) \quad (4.33a)$$

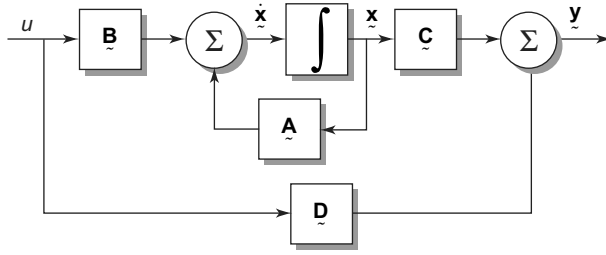


Fig. 4.7 Representation of a multivariable process control system with state space modelling

and the characteristic polynomial is factorable with poles at λ_i 's, then Eq. (4.33a) is written as

$$y(s)/u(s) = b_o + \sum_{i=1}^n c_i / (s - \lambda_i) \quad (4.33b)$$

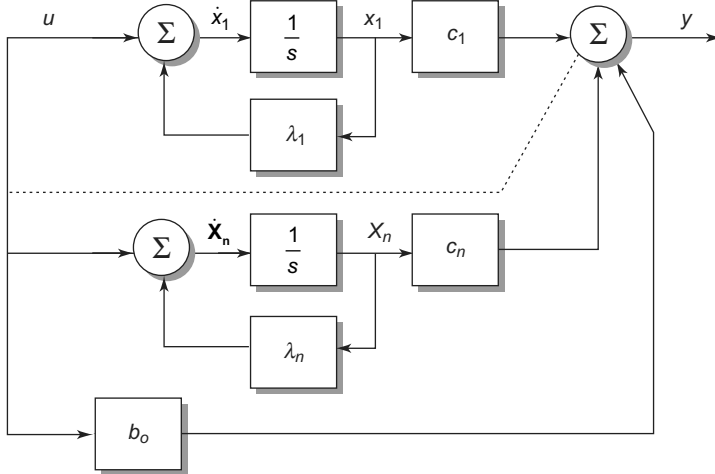


Fig. 4.8 Schematic representation of the transfer function of Eq. (4.33b)

with the block diagram representation shown in Fig. 4.8. The state equations are obviously written as

$$\dot{x}_j = \lambda_j x_j + u \quad j = 1, 2, \dots, n \quad (4.33c)$$

In the vector matrix form the state and the output equations are

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \vdots \\ \dot{x}_n \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} u$$

and

$$y = [c_1 \quad c_2 \quad c_3 \quad \cdots \quad c_n] \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ x_n \end{bmatrix} + b_o u$$

Controllability and an associated property called observability are considered in terms of state variables or states as has already been mentioned. In fact, such controllability is often called the state controllability and a system is said to be state controllable if it is possible to transfer the state of the system $x(t)$ from any initial value $x(t_o)$ to a desired value $x(t_d)$ in a specified finite time by the application of control (input) vector $u(t)$. Proceeding similarly, a system would be called completely observable if every state of the system is completely identified from the measurement of the output $y(t)$ over a finite interval of time.

A time-invariant system represented by (see Eq. (4.31))

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} \quad (4.34)$$

with $\mathbf{x}$ being an n -dimensional state vector and $\mathbf{u}$, the control input, $\mathbf{A}$ and $\mathbf{B}$ are $n \times n$ and $n \times 1$ matrices respectively with A having distinct eigenvalues. The system described by Eq. (4.34) would be controllable if it is possible to obtain control input which in the interval of time $0 < t \leq t_d$ would transfer the system from the initial state $\mathbf{x}(0)$ to the desired state $x(t_d)$.

Defining a new state vector $\mathbf{z} = \mathbf{Q}^{-1}\mathbf{x}$, where $\mathbf{Q}$ is a nonsingular constant matrix, this transformation would modify the original model of Eq. (4.34) into

$$\dot{\mathbf{z}} = \mathbf{Q}^{-1}\mathbf{A}\mathbf{Q}\mathbf{z} + \mathbf{Q}^{-1}\mathbf{B}\mathbf{u} \quad (4.35)$$

If $\mathbf{Q}$ is so selected that $\mathbf{Q}^{-1}\mathbf{A}\mathbf{Q} = \mathbf{P}$ is a diagonal matrix, then the model represented by Eq. (4.35) is a canonical state model and $\mathbf{P}$ is thus a diagonal matrix. Thus, for $\mathbf{Q}^{-1}\mathbf{B} = \mathbf{W}$, Eq. (4.35) writes as

$$\dot{\mathbf{z}} = \mathbf{P}\mathbf{z} + \mathbf{W}\mathbf{u} \quad (4.36)$$

which in the expanded form is written as

$$\begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \\ \cdot \\ \dot{z}_n \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 & \cdot & \cdot & 0 \\ 0 & \lambda_2 & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \cdot & \lambda_n \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ \cdot \\ z_n \end{bmatrix} + \begin{bmatrix} w_1 \\ w_2 \\ \cdot \\ w_n \end{bmatrix} u \quad (4.37)$$

giving the component form representation as

$$\dot{z}_j = \lambda_j z_j + w_j u, \quad j = 1, 2, \dots, n$$

This equation is solved as follows:

Multiply both sides by $e^{-\lambda_j t}$ to get

$$e^{-\lambda_j t} \dot{z}_j - e^{-\lambda_j t} \lambda_j \dot{z}_j = e^{-\lambda_j t} w_j u \quad (4.37a)$$

but

$$\frac{d}{dt} [e^{-\lambda_j t} z_j] = e^{-\lambda_j t} \dot{z}_j - \lambda_j e^{-\lambda_j t} z_j$$

Hence Eq. (4.37a) becomes

$$\frac{d}{dt} [e^{-\lambda_j t} z_j] = e^{-\lambda_j t} w_j u$$

Integrating, Eq. (4.37b) given below is obtained.

$$z_j(t) = \exp(\lambda_j t) z_j(0) + \exp(\lambda_j t) \int_0^t \exp(-\lambda_j \tau) w_j u(\tau) d\tau \quad (4.37b)$$

Thus, the system is controllable if a control $u(t)$ is found out which satisfies the relation

$$\frac{z_j(t_d) - \exp(\lambda_j t_d) z_j(0)}{\exp(\lambda_j t_d)} = \int_0^t \exp(-\lambda_j \tau) w_j u(\tau) d\tau \quad (4.38)$$

As long as $w_j \neq 0$, there are numerous values of $u(\tau)$ that satisfies Eq. (4.38). Thus the condition comes to be stated that the vector $\mathbf{W}$ should not have any zero element. If $\mathbf{u}$ is an m -dimensional vector, $\mathbf{W}$ is an $n \times m$ matrix, the necessary and sufficient condition for controllability is that the matrix $\mathbf{W}$ should not have any row with all zero elements as this would mean that in such a situation it would not be possible to influence the corresponding state variable by the control input and hence, the particular state variable is uncontrollable. This, however, may mean that other states are controllable. As a system is defined to be completely state controllable if every state is controllable, the above system may be defined as partially state controllable when the system may still be output controllable. This is obviously in contrast to a system where the condition of total controllability is satisfied with regard to the nonzero entries of $\mathbf{W}$, as already stated.

If $\mathbf{A}$ does not possess distinct eigenvalues, diagonalization is no longer possible and $\mathbf{A}$ is then transformed into Jordan canonical form. For example, let the system eigenvalues are $\lambda_1, \lambda_1, \lambda_3, \lambda_3, \lambda_3, \lambda_6, \dots, \lambda_n$, the Jordan matrix would be

$$J = \begin{bmatrix} \boxed{\lambda_1} & \boxed{1} & 0 & 0 & 0 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & \lambda_1 & 0 & 0 & 0 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & 0 & \boxed{\lambda_3} & \boxed{1} & 0 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & 0 & 0 & \lambda_3 & \boxed{1} & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & 0 & 0 & 0 & \lambda_3 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & 0 & 0 & 0 & 0 & \boxed{\lambda_6} & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdot & \cdot & \cdot & \boxed{\lambda_n} \end{bmatrix} \quad (4.39)$$

in

$$\dot{\mathbf{z}} = \mathbf{J}\mathbf{z} + \mathbf{W}\mathbf{u} \quad (4.40a)$$

where the dashed rectangles in the matrix represent that are known as Jordan blocks, and the controllability condition is that the elements of any row of $\mathbf{W}$ that correspond to the last row of each Jordan block are not all zero.

A criterion of controllability in terms of the matrices $\mathbf{A}$ and $\mathbf{B}$ suggested by Kalman is given here. It states that the system is completely controllable if the rank of the composite matrix

$$\mathbf{M} = [\mathbf{B}, \mathbf{A}\mathbf{B}, \dots, \mathbf{A}^{n-1}\mathbf{B}]$$

is n , where $\mathbf{A}$ is an $n \times n$ matrix and $\mathbf{B}$ is an $n \times m$ matrix.

Its proof is given starting from the general equation, Eq. (4.34).

Multiplying both sides by e^{-At} , one gets

$$e^{-At}[\dot{x} - Ax] = e^{-At}Bu$$

or,

$$\frac{d}{dt}[e^{-At}x] = e^{-At}Bu$$

or,

$$e^{-At}x(t) = x(0) + \int_0^t e^{-A\tau}Bu(\tau)d\tau$$

or,

$$x(t) = e^{At}x(0) + \int_0^t e^{A(t-\tau)}Bu(\tau)d\tau$$

Assuming now that the initial state (disturbed zero state) returns to origin (zero state) with control action so that $x(t) = 0$.

Hence,

$$x(0) = \int_0^t e^{-A\tau} Bu(\tau) d\tau$$

Putting the expansion of $e^{-A\tau} = -\sum_{k=0}^{n-1} \alpha_k(\tau) A^k$ in the above equation

$$\begin{aligned} x(0) &= \int_0^t -\sum_{k=0}^{n-1} \alpha_k(\tau) A^k Bu(\tau) d\tau \\ &= -\sum_{k=0}^{n-1} A^k B \int_0^t \sum_{k=0}^{n-1} \alpha_k(\tau) u(\tau) d\tau \\ &= -[B : AB : A^2B : \dots : A^{n-1}B] \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_{n-1} \end{bmatrix} \end{aligned} \quad (4.40b)$$

where

$$\int_0^t \sum_{k=0}^{n-1} \alpha_k(\tau) u(\tau) d\tau = \beta_k, \quad k = 0, 1, \dots, n-1$$

If Eq. (4.40b) is to be satisfied $[B : AB : A^2B : \dots : A^{n-1}B]$, called the composite matrix, must have a rank of the system matrix otherwise $u(t)$ will not couple as discussed in the case given by Eq. (4.38).

In the practical design of a control system, the output control may appear to be more important than the state control. Complete state controllability is, therefore, may seem to be neither necessary nor sufficient for controlling the output. Output controllability is then separately defined as follows. The system defined by Eqs (4.31) and (4.32) is completely output controllable if it is possible to construct a control vector $\mathbf{u}(\mathbf{t})$ that would transfer any initial output $y(t_o)$ to a final output $y(t_f)$ in a finite interval of time $t_o \leq t \leq t_f$. Kalman's condition of complete output controllability is that the matrix

$$\mathbf{N} = [\mathbf{CB}, \mathbf{CAB}, \mathbf{CA}^2\mathbf{B}, \dots, \mathbf{CA}^{n-1}\mathbf{B}, \mathbf{D}]$$

is of rank p .

4.3 SELF-REGULATION

When the controllability of a plant/process is examined, it is tentatively assumed that the process has a sort of inherent regulation called self-regulation, although this is not essential at least theoretically. Self-regulation

is a characteristic of the process which helps in limiting the deviation of the controlled variable. This characteristic is obtainable directly from the process reaction curves. The process time constants are made up of process capacities and process resistances. While process capacities may not be allowed to decrease, proper design may be initiated such that the process resistances fall, thereby decreasing the time constants and improving regulation. As the multi-time constant processes have interactions in capacities and resistances, one cannot very simply state that self-regulation improves when resistances are less. In addition, it is necessary that the ratio of the supply side capacity to the demand side capacity should be as small as possible. It would mean that the disturbance in the demand side would not affect the process as it should in comparison with the supply side disturbance. This ratio determines the process reaction rate as also its ability to withstand disturbances. From the resistance-capacitance viewpoint, the situation is explained with reference to Fig. 4.9. Here, one notes that the system of Fig. 4.9(b) has better self-regulation than the system of Fig. 4.9(a) because as h increases in Fig. 4.9(b), inflow q_i will tend to lessen and outflow will tend to increase—this will bring equilibrium sooner. Figure 4.9(c) shows a two capacity process and for better self-regulation in this C_1/C_2 should be small.

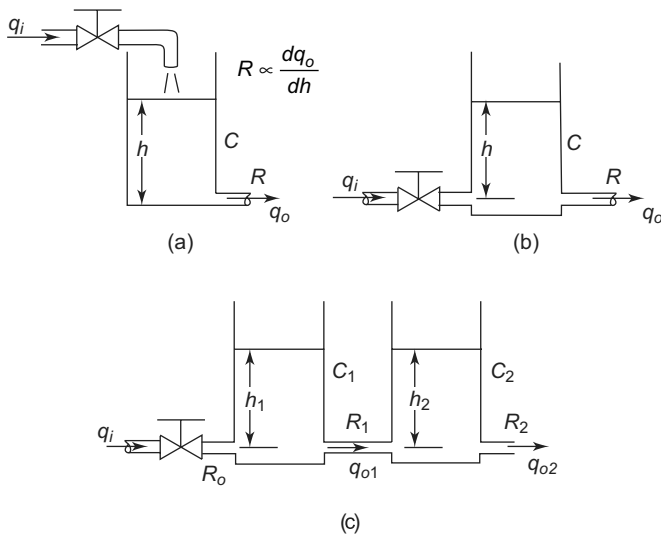


Fig. 4.9 (a) and (b) Two schemes of head regulation in a tank (c) cascaded tanks for regulation of tank level—an example of self-regulation

A process is said to be non-self-regulating if it does not have a steady state. This implies that for a second order system, its damping factor ζ must be zero or negative. Therefore, a system, that has $\zeta \geq 0$, has a self-regulation which improves with increasing ζ . Self-regulation, whatever

be its extent in a process, is a desirable feature as the automatic control is made easier for it. In that sense, perhaps $\zeta = 1$ would be a good choice and the corresponding self-regulation may be considered optimal.

For the process of Fig. 4.9(c) an electrical analogy can be drawn and a corresponding equivalent circuit from which one can obtain

$$h_2(s)/x(s) = (R_2/R_0)[s^2 C_1 C_2 R_1 R_2 + s(C_1 R_1 + C_2 R_2 + C_1 R_2 + C_2 R_1 R_2/R_0) + (R_0 + R_1 + R_2)/R_0] \quad (4.41)$$

For optimal self-regulation of this process

$$\zeta = 1 = (1/2)[C_1(R_1 + R_2) + C_2(R_0 R_2 + R_1 R_2)/R_0]/[C_1 C_2 R_1 R_2 (R_0 + R_1 + R_2)/R_0]^{1/2} \quad (4.42)$$

If Eq. (4.42) has to be satisfied, C_1/C_2 becomes imaginary, solving for non-imaginary C_1/C_2 with ζ closest to 1 gives a value of ζ as $1 + R_0 R_2/(R_1(R_0 + R_1 + R_2))$ using which the value of C_1/C_2 is obtained as

$$C_1/C_2 = 1 - R_1(R_0 - R_2)/[R_0(R_1 + R_2)]$$

Thus, for $C_1/C_2 < 1$, one must have $R_0 > R_2$.

The characteristics of non-self-regulating and self-regulating processes are shown graphically in Fig. 4.10 for a step disturbance. Temperature processes are self-regulating because an increase in heat input in it would eventually produce a new steady state temperature. Exothermal chemical reactors can, in contrast, have negative self-regulation as increase in temperature here would further increase the rate of heat evolution. For this reason their control part should be cautiously designed with the reactor temperature controller cascaded to the coolant temperature controller.

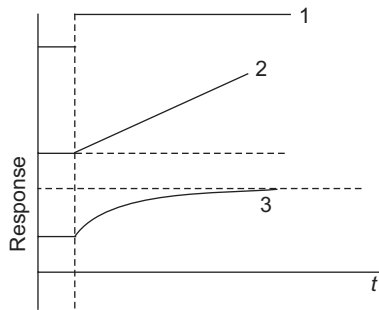


Fig. 4.10 Response characteristics of non-self regulating (2) and self-regulating (3) processes

4.4 STABILITY STUDIES

A study of the stability of a system is as, if not more, important as the study of controllability. If a system is not stable, it naturally is of no use.

A clear cut and obvious definition of stability of a system is: a system is stable if for any bounded input to it, the output from it is also bounded. This definition forms the basic concept and for evaluation of stability or its criteria different techniques are known. The common methods for linear systems are

- (i) Nyquist criterion,
- (ii) Routh-Hurwitz criterion,
- (iii) Root-locus technique, and
- (iv) Bode-plot technique.

The Nyquist criterion, as a stability criterion, is obtained through a graphical plot of transfer function for all frequencies ($0 < \omega < \infty$). If the plot encloses the $-1 + j0$ point, the system is unstable; otherwise it is stable. In some cases the plotted curve turns out to be so complicated and roundabout that it becomes difficult to ascertain whether it has enclosed the $-1 + j0$ point or not. In such a case, a certain procedure may be followed after the curve is drawn. From the point $-1 + j0$ a vector to the curve is allowed to make an excursion over the entire frequency $0 < \omega < \infty$. If the nett angle travelled by the vector is zero, then the system is stable.

Nyquist plot is a polar plot, a plot of the loop function $G(j\omega) H(j\omega)$ (see Fig. 4.11(b)) shown in Fig. 4.11(a) for a higher order system where a unit circle has also been drawn. The plot does not enclose the

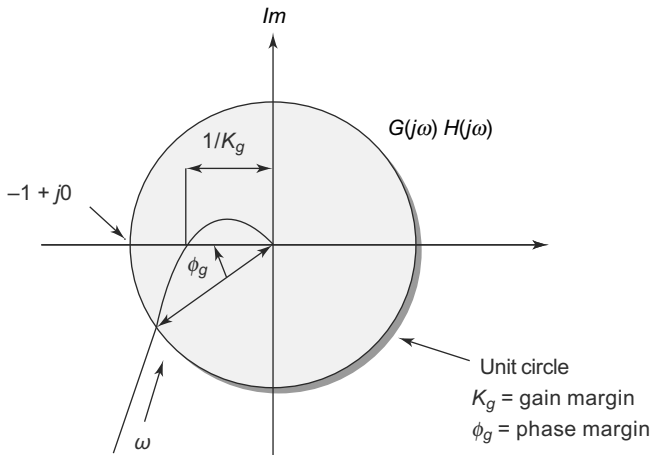


Fig. 4.11 (a) The Nyquist plot

point $(-1 + j0)$ and hence is stable and from this, the gain and phase margins are evaluated as shown. Gain and phase margins are detailed out later in Bode plot technique (Sec. 4.4.4).

However, the Nyquist criterion is not commonly applied to process control systems. In comparison to this, the Routh-Hurwitz criterion is

easier to apply. A straightforward outcome of the definition of stability is that the roots of the characteristic equation of the system must all lie on the left half of the s -plane implying that the roots must have negative real parts. It is rather difficult to evaluate the higher order roots from polynomial form. Stability can still be studied without actually evaluating the roots. A criterion to this effect was established and is known as the Routh-Hurwitz criterion. For a characteristic polynomial

$$P(s) = \sum_{j=0}^n b_j s^{n-j} \quad (4.43)$$

the Hurwitz determinants are

$$D_j = \begin{vmatrix} b_1 & b_3 & b_5 & \cdots & b_{2j-1} & \cdots & 0 \\ b_0 & b_2 & b_4 & \cdots & b_{2j-2} & \cdots & 0 \\ 0 & b_1 & b_3 & \cdots & b_{2j-3} & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & b_j & \cdots & b_n \end{vmatrix} \quad j = 1, 2, \dots, n \quad (4.44)$$

The criterion now states that for the roots of polynomial $P(s)$ to lie on the left half of s -plane, $b_0 > 0$ and $D_j > 0$.

There is a simplified way to apply this criterion through what is known as Routh's algorithm. Both Routh-Hurwitz and Routh criteria are indicative of the sufficient conditions of stability.

Routh's algorithm is based on ordering the coefficients of the characteristic equation into an array, the Routh array, and the criterion follows from this array. Considering the n th degree polynomial of Eq. (4.43), the array is formed with its coefficients b_j as

$$\begin{array}{l|llllll} s^n & b_0 & b_2 & b_4 & b_6 & b_8 & \cdots \\ s^{n-1} & b_1 & b_3 & b_5 & b_7 & b_9 & \cdots \\ s^{n-2} & c_1 & c_2 & c_3 & c_4 & c_5 & \cdots \\ s^{n-3} & d_1 & d_2 & d_3 & d_4 & \cdots & \\ & \cdots & \cdots & \cdots & \cdots & & \\ s^1 & m_1 & m_2 & \cdots & & & \\ s^0 & n_1 & & & & & \end{array}$$

where

$$c_1 = (b_1 b_2 - b_0 b_3)/b_1, \quad c_2 = (b_1 b_4 - b_0 b_5)/b_1, \quad \text{and so on}$$

$$d_1 = (c_1 b_3 - c_2 b_1)/c_1, \quad d_2 = (c_1 b_5 - c_3 b_1)/c_1, \quad \text{and so on}$$

The array will evidently have $n + 1$ rows, the first row consisting of $(n + a)/2$ elements, where $a = 1$ for n odd and $a = 2$ for n even. The $(n + 1)$ th row will contain only one element. The degree of the polynomial formed with the coefficients in the j th row is given by $(n - j + 1)$ and each subsequent term in the polynomial has two degrees less. With the array thus completed the stability test criterion via the algorithm is as follows.

- (i) If the array can be completed as such and if each of the elements in the first column is finite non-zero, then for stability, these elements should be of the same sign.
As a corollary, one derives that the system is unstable for a sign difference in all finite non-zero elements in the first column; the number in changes in sign indicates the number of roots with positive real parts.
- (ii) If any row has its first element vanishing but not all others, the array is completed by taking a small quantity ϵ in its place. This ϵ may be positive or negative, such a zero indicates that the system is not stable.
- (iii) If any one or more of the rows above the last one were to vanish, the array could still be completed. For the i th row vanishing, a subsidiary polynomial with the coefficients of the $(i - 1)$ th row is formed and differentiated. The coefficients of the derived polynomial will now replace the zeros in the i th row and the process is continued for the completion of the array. This discontinuity occurs because of the existence of 'equal and opposite roots'—the roots of the same magnitude but with a 180° phase shift between them. A pair of imaginary roots can produce these results and the system is conditionally stable. If more rows have all their elements equal to zero, it indicates that such roots are in multiplicity and the system is unstable. The order of multiplicity is given by the number of rows obtained with all the elements zero. These roots may be evaluated from a polynomial formed by the coefficients of the previous row. An example will illustrate the application of this criterion.

Let the characteristic equation be

$$s^5 + s^4 + 3s^3 + 3s^2 + 2s + 2 = 0$$

The Routh array is then

$$\begin{array}{c|ccc} s^5 & 1 & 3 & 2 \\ s^4 & 1 & 3 & 2 \\ s^3 & 0 & 0 & \end{array}$$

Since the entries corresponding to s^3 row vanishes, the subsidiary equation with the entries of the row corresponding to s^4 is formed as

$$s^4 + 3s^2 + 2 = 0$$

Differentiating, we get

$$4s^3 + 6s = 0, \text{ i.e., } 2s^3 + 3s = 0$$

The array is now completed as

$$\begin{array}{c|ccc} s^5 & 1 & 3 & 2 \\ s^4 & 1 & 3 & 2 \\ s^3 & 2 & 3 & \\ s^2 & 3/2 & 2 & \\ s^1 & 1/3 & & \\ s^0 & 2 & & \end{array}$$

From the subsidiary equation, the roots are also evaluated. There are two pairs of imaginary axis roots,

$$S = \pm j, \text{ and, } s = \pm j\sqrt{2}$$

Since there is no change of sign in the first entries of the array after completion, the system is conditionally stable.

There are specific relations between the elements of the first column of the Routh array and the Hurwitz determinants. In fact, a close look would give the following correlations

$$b_0 = b_0$$

$$b_1 = D_1$$

$$c_1 = D_2/D_1$$

$$d_1 = D_3/D_2, \text{ and so on.}$$

Thus the two criteria, the Routh-Hurwitz determinant criterion and the Routh array criterion are identical. Taking a third order example, now, with the characteristic equation

$$3s^3 + 3s^2 + 2s + 2 = 0$$

the determinants are

$$D_1 = 3$$

$$D_2 = \begin{vmatrix} 3 & 2 \\ 3 & 2 \end{vmatrix} = 0$$

$$D_3 = \begin{vmatrix} 3 & 2 & 0 \\ 3 & 2 & 0 \\ 0 & 3 & 2 \end{vmatrix} = 3(4) - 4 \times 3 = 12 - 12 = 0$$

Here $b_0 = D_1 = 3$, and $D_2 = D_3 = 0$ which may be considered positive zero. The system is stable but only conditionally as established by Routh's criterion for the fifth order case. Relative stability, which is essential in process control system studies, is not indicated by the above procedure.

The root-locus technique is a graphical method to obtain the excursion of the roots when an appropriate adjustable parameter in the system is varied. For some values of this parameter roots would lie on the left half of the s -plane and the system would then be stable.

Bode-plot technique is again a graphical approach to determine stability and is a simpler and modified form of the Nyquist plot but is used for some special types only. Fortunately, these types are more common in process control systems and hence the Bode-plot technique is very convenient for the stability studies of process control systems.

Root-locus and Bode-plot techniques are more popular for the stability studies of the process control systems, particularly when the degree of stability is also to be adjudged. In the following these methods are considered in a slightly more detail.

4.4.1 Root-locus Technique

Root-locus technique can readily furnish information pertaining to (i) relative stability, (ii) transient response, and (iii) frequency domain characteristics for simple processes having closed loop control configurations, if the transfer function is known. In majority of process control problems where this technique is applied, it is primarily used for controller parameter settings and the system design. It has effectively been used in automated system design.

A plot of the loci of the poles of the closed loop transfer function, i.e., roots of the characteristic equation, when a parameter of the open loop transfer function is varied (from 0 to ∞), is known as the root-locus diagram. If the parameter varies from $-\infty$ to 0, the plot is termed as the inverse root-locus diagram and if two or more parameters are varied, the corresponding plot is called a root-contour.

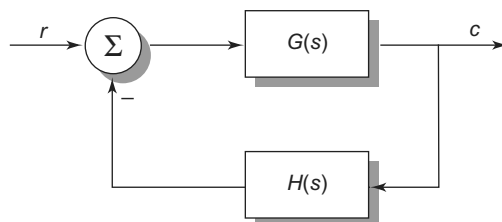


Fig. 4.11 (b) Simple closed loop system

The closed loop transfer function for the system of Fig. 4.11 is given as

$$c(s)/r(s) = G(s)/(1 + G(s)H(s)) \quad (4.45a)$$

giving the characteristic equation

$$1 + G(s)H(s) = 0 \quad (4.45b)$$

The roots of the characteristic equation are obtained when

$$|G(s)H(s)| = 1 \quad (4.46a)$$

and

$$\angle G(s)H(s) = (2m + 1)\pi \quad (4.46b)$$

where, $m = 0, \pm k, k = 1, 2, \dots$ integers.

If $G(s)H(s)$ is obtained in a factorized form

$$G(s)H(s) = \frac{K(s + \alpha_1)(s + \alpha_2)(s + \alpha_3) \dots (s + \alpha_n)}{(s + \beta_1)(s + \beta_2)(s + \beta_3) \dots (s + \beta_r)}, r > n \quad (4.47)$$

then Eq. (4.46) can be written as

$$|G(s)H(s)| = \frac{|K| \prod_{i=1}^n |s + \alpha_i|}{\prod_{j=1}^r |s + \beta_j|} = 1 \quad -\infty < K < \infty \quad (4.48a)$$

and

$$\begin{aligned} \angle G(s)H(s) &= \angle K + \sum_{i=1}^n \angle(s + \alpha_i) - \sum_{j=1}^r \angle(s + \beta_j) \\ &= (2m + 1)\pi \end{aligned} \quad (4.48b)$$

Hence one can state that any point s_1 in the s -plane that satisfies Eq. (4.48) for any value of K ($0 < K < \infty$) should be a point on the root-locus. It is, thus, a technique of evaluating closed loop roots with a knowledge of the open-loop roots.

Finding out the s_1 's for all the K 's in the s -plane that satisfy Eq. (4.48) seems to be uphill task, particularly if the order of the characteristic equation is quite large. The use of computers has greatly aided this technique for process control system studies. This requires designing of algorithms that will enable efficient use of the digital computer for this purpose. However, for the construction of the root loci, a few rules are known which will now be stated, without proof. The construction follows certain geometric procedures and proofs for them follow almost from an inspection of the construction and Eq. (4.48). It should be noted that s_1 is actually a complex

quantity as this occurs in $(\sigma + j\omega)$ -plane, and s_1 is a root of $G(s)H(s)$ only if $G(s)H(s) = -1 + j0$. The rules stated below are only aids to the construction of the root loci and the exact plots are not obtained this way. The rules are readily obtained from a consideration of the roots of the characteristic equation (Eq. (4.45b)) and the poles and zeros of $G(s)H(s)$.

(i) Starting

When $K = 0$, the points on the root loci are the poles of $G(s)H(s)$. This is easily obtained from Eq. (4.48a).

(ii) Ending

When $K = \pm\infty$, the points on the root loci are zeros of $G(s)H(s)$. This also readily follows from Eq. (4.48a).

(iii) Separate parts

If m_z denotes the number of zeros and n_p , number of poles of $G(s)H(s)$, then the number of separate parts of the loci are obtained as

$$\begin{aligned} N &= m_z \quad \text{if } m_z > n_p \\ &= n_p \quad \text{if } m_z < n_p \end{aligned} \quad (4.49)$$

This follows from the fact that each pole-zero pair represents one segment of the root loci.

When $n_p > m_z$, the order of the characteristic equation gives N .

(iv) Symmetry

Root loci and inverse root loci are symmetrical to each other with respect to the real axis of the s -plane.

(v) Asymptotes

The root loci are asymptotes with angles

$$\phi_m = \frac{(2m+1)\pi}{n_p - m_z}, \quad m = 0, 1, 2, \dots, (n_p - m_z - 1) \quad (4.50a)$$

for s being very large. Similarly, for inverse root loci

$$\phi_m = \frac{2m\pi}{n_p - m_z} \quad (4.50b)$$

There should, therefore, be $n_p - m_z$ asymptotes.

(vi) Centroid (Intersection of asymptotes)

The intersection of $n_p - m_z (= M)$ asymptotes lies on the real axis and its coordinate σ_c is given by

$$\sigma_c = -\frac{\Sigma \text{ poles of } G(s)H(s) - \Sigma \text{ zeros of } G(s)H(s)}{n_p - m_z}$$

$$\left(= -\frac{b_1 - a_1}{M} \right) \quad (4.51)$$

A very simple proof exists for items (v) and (vi) which is given here briefly.

From Eq. (4.48a), if $r = m + n$ and assuming the numerator and denominator in polynomial forms, one can write

$$G(s)H(s) = \frac{KP(s)}{Q(s)} = \frac{K(s^m + a_1s^{m-1} + \dots + a_m)}{s^{m+n} + b_1s^{m+n-1} + \dots + b_m} = -1 \quad (4.52)$$

where

$$\sum_{i=1}^m \alpha_i = a_1, \quad \sum_{j=1}^{m+n} \beta_j = b_1, \quad \prod_{i=1}^m \alpha_i = a_m \quad \text{and} \quad \prod_{j=1}^{m+n} \beta_j = b_m$$

From Eq. (4.52) one easily derives

$$-1 = \frac{K}{Q(s)/P(s)} = \frac{K}{s^n + (b_1 - a_1)s^{n-1} + \dots + R(s)/P(s)} \quad (4.53)$$

where $R(s)$ is the remainder. For the asymptotes $s \rightarrow \infty$ such that $s^n + (b_1 - a_1)s^{n-1} \gg$ rest of the terms, from which

$$s^n + (b_1 - a_1)s^{n-1} = -K \quad (4.54)$$

which yields after a slight manipulation

$$s \left(1 + \frac{b_1 - a_1}{ns} \right) = (-K)^{1/n} \quad (4.55)$$

Expanding and accepting higher order terms only

$$s + (b_1 - a_1)/n = |(K)^{1/n}| \{ \cos[(2k+1)\pi/n] + j \sin[(2k+1)\pi/n] \} \quad (4.56a)$$

$$0 < k < \infty$$

$$= |(K)^{1/n}| [\cos 2k\pi/n + j \sin 2k\pi/n] \quad (4.56b)$$

$$0 > k > -\infty$$

with $k = 0, \pm 1, \pm 2, \dots$

Writing $s = \sigma + j\omega$ and equating real and imaginary parts

$$\sigma + \frac{b_1 - a_1}{n} = (K)^{1/n} \cos \left[\frac{(2k+1)\pi}{n} \right] \quad (4.57a)$$

and

$$\omega = (K)^{1/n} \sin \left[\frac{(2k+1)\pi}{n} \right] \quad (4.57b)$$

from which

$$\omega = \left\{ \tan \left[\frac{(2k+1)\pi}{n} \right] \right\} [\sigma + (b_1 - a_1)/n] \quad (4.58a)$$

$$= m_i(\sigma - \sigma_i) \quad (4.58b)$$

Figure 4.12 shows the representation of these properties.

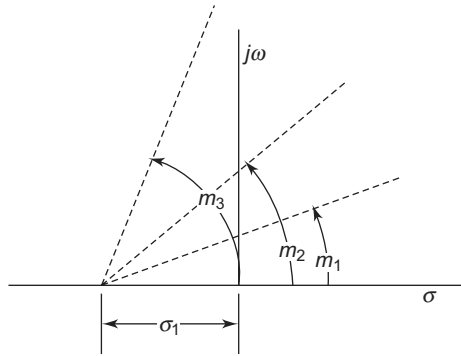


Fig. 4.12 The intersection points of root loci

(vii) Real axis root loci

Root loci on a section of the real axis are obtained only when the number of poles and zeros to the right of this section is odd. For the inverse case this should be even.

(viii) Angle of departure and arrival

The angle of departure of the root locus from a pole or the angle of arrival at a zero of $G(s)H(s)$ can be determined by fixing a point, s_k , close to the singularity (pole or zero) and on the locus associated with this singularity and then by applying Eq. (4.48b).

(ix) Intersection of the locus with the $j\omega$ -axis

For the intersection, values of ω and K are obtained by using the Routh criterion. Bode-plots can be used for complex cases.

(x) Breakaway points

The breakaway point corresponds to root multiplicity and for root locus or inverse root locus, it is determined by finding the roots of $dK/ds = 0$

or $d(G(s)H(s))/ds = 0$. Also, an algorithm formulation is possible using the coefficients of the characteristic equation of the closed-loop systems $F(s)$ and $F'(s)$. The last method involving algorithm formulation is due to Remec and is very useful for higher order systems.

(xi) Value of K on any point of the root locus

The value of K at any point s_k on the root locus or the inverse root locus (or loci) is obtained, following Eq. (4.48a), as

$$|K| = 1/|G(s)H(s)| = \frac{\text{vector lengths from the poles to } s_k}{\text{vector lengths from the zeros to } s_k} \quad (4.59)$$

An example illustrating the use of the rules will now be appropriate. Consider the process control loop whose block diagram is shown in Fig. 4.13. The loop-transfer function is given by

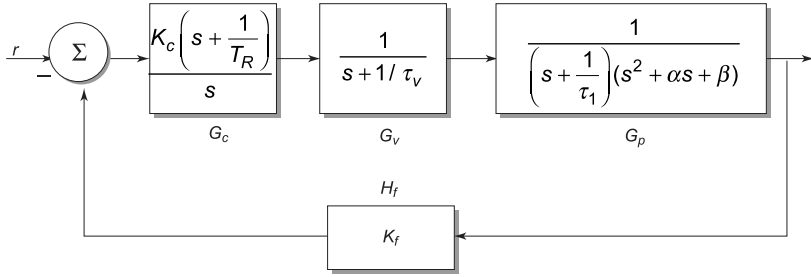


Fig. 4.13 Block diagram of a process-control loop

$$G(s)H(s) = \frac{K_f K_c (s + 1/T_R)}{s(s + 1/\tau_v) \left(s + \frac{1}{\tau_1} \right) (s^2 + \alpha s + \beta)} \quad (4.60)$$

Considering that the proportional gain is the variable parameter, if other parameters are chosen as $K_f = 1$, $T_R = 1/3$, $\tau_v = 1/5$, $\tau_1 = 1/6$, $\alpha = \beta = 2$ and $K_c = K$, the loop transfer function then becomes

$$G(s)H(s) = \frac{K(s + 3)}{s(s + 5)(s + 6)(s^2 + 2s + 2)} \quad (4.61)$$

Then from the rules laid out, the poles of $G(s)H(s)$ are at 0, -5 , -6 and $-1 \pm j1$ which are the starting points of the root loci. One finite zero is at -3 and the others (4 zeros) are at infinity and these are the end-points of the root loci. Now, $n_p = 5$ and $m_z = 1$; therefore, there are 5 segments of the root loci. Asymptote angles are

$$\phi_m = \frac{(2m + 1)\pi}{n_p - m_z}$$

and

$$\phi_m = \frac{2m\pi}{n_p - m_z}$$

Since $n_p - m_z = 4$, and with $m = 0, 1, 2, 3$

$$\phi_0 = \pi/4, \phi_1 = 3\pi/4, \phi_2 = 5\pi/4, \phi_3 = 7\pi/4 \text{ for root locus with } s \rightarrow \infty.$$

For inverse root locus with $s \rightarrow -\infty$

$$\phi_0 = 0, \phi_1 = \pi/2, \phi_2 = \pi \text{ and } \phi_3 = 3\pi/2$$

$$\sigma_1 = \frac{0 - 5 - 6 - 1 + j - 1 - j - (-3)}{4} = -2.5$$

Figure 4.14(a) shows asymptote diagrams for the above case. There are root loci on the real axis on $-3 \leq s \leq 0$, $-6 \leq s \leq -5$. Figure 4.14(b) shows the segments of the root loci on the real axis.

Angles of departures at $s = -1 + j$ are obtained easily as $-403^\circ.8$ and $-43^\circ.8$, respectively. From the characteristic equation Routh algorithm can

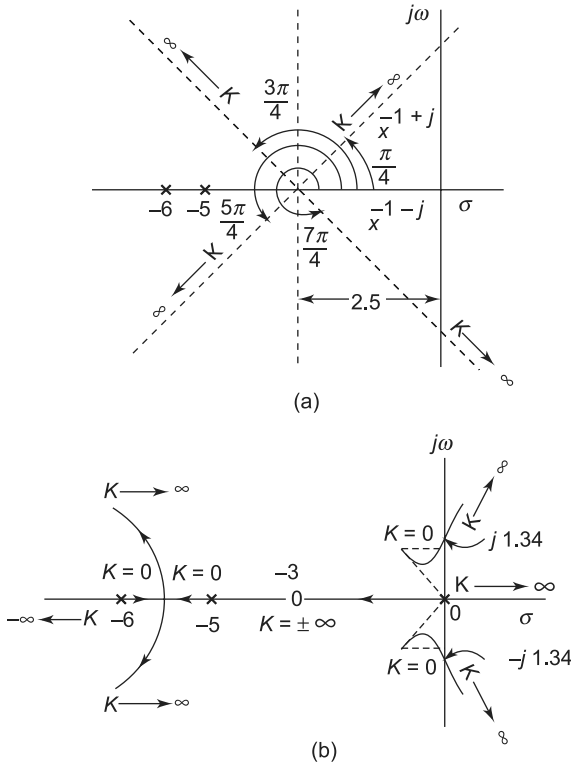


Fig. 4.14

(a) Asymptotes for the example of Fig. 4.13. (b) Sketch of the root loci in parts and indication of the number of breakaway points

be formed, from which the gain at the cross-over points are easily obtained as $K = 35$. Using this value of K the cross-over points are obtained as $s_c = \pm j1.34$ (see Fig. 4.14).

As the characteristic equation is of the 5th order, Remec's method may be applied to determine the breakaway points. There is, however, only one breakaway point as is easily understood from Fig. 4.14(b). Remec's method is briefly explained in Appendix III. The value of K at any point can be obtained by the simple procedure already specified.

4.4.2 Root-locus Properties for Use in Process Control Systems

- (i) The addition of poles to $G(s)H(s)$ reduces the relative stability of the closed-loop system. Figure 4.15 illustrates the situation clearly for 2-pole and 3-pole systems.
- (ii) The addition of zeros to $G(s)H(s)$ tends to give a more stabilizing effect to the closed-loop system. Fig. 4.16 shows this effect clearly for 2-pole-0-zero, 2-pole-1-zero and 2-pole-2-zero cases.

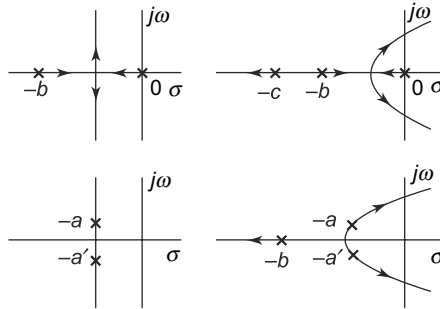


Fig. 4.15 Sketches of root loci to show decrease in stability with increase in pole

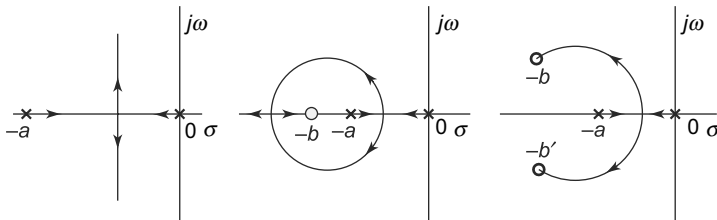


Fig. 4.16 Sketches of root loci to show increase in stability with increase in zeros

- (iii) Another important point is the effect of variation in the location of zero(s) [equivalent to change in the reset time of an integral action controller, see Fig. 4.13 and Eq. (4.60)] or pole(s) of $G(s)H(s)$, on the closed-loop system. For a 3-pole system this is explained in Fig. 4.17 with a single zero whose position is altered and the permissible

variation in K is also shown here. The loop-transfer function of such a system is

$$G(s)H(s) = \frac{K(s + z_1)}{s(s^2 + 2s + 2)}$$

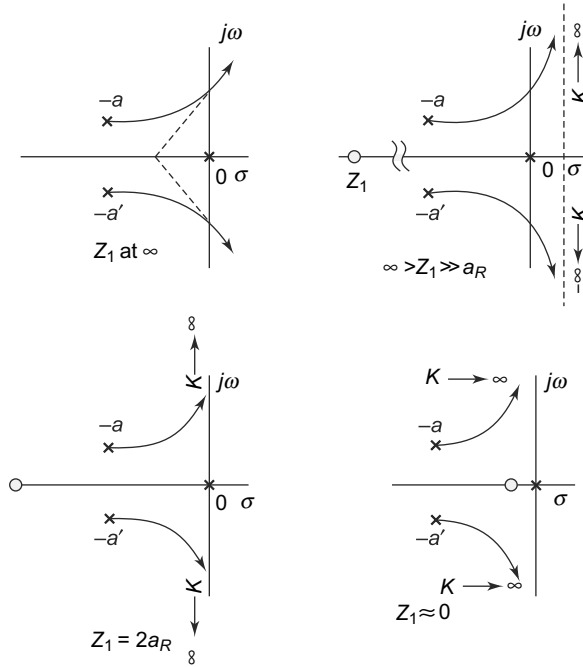


Fig. 4.17 Sketches for the relative stability studies for a 3-pole system

4.4.3 Root Contours and their Applications

Any variable parameter other than K can also be accommodated by taking K as fixed and a set of loci is, therefore, obtained for different K 's which are root contours. The transformation of the equation is also simple.

Let $G(s)H(s) = KP(s)/Q(s)$, then for $1 + G(s)H(s) = 0$, one has $KP(s) + Q(s) = 0$

If now T_p is a parameter in $P(s)$ which varies from $-\infty$ to $+\infty$, one keeps K constant and writes

$$T_p(s)P_p(s) + Q(s) = 0, \quad \text{i.e.} \quad G_p(s)H_p(s) = T_pP_p(s)/Q(s)$$

Similarly, for $Q(s) = T_qQ_q(s)$, one obtains

$$G_q(s)H_q(s) = (1/T_q)P_k(s)/Q_q(s)$$

where $P_k(s)$ is $KP(s)$.

T_p or $1/T_q$ can be varied similar to K holding K constant at a desired value. In a proportional integral derivative (PID) controller setting, the proportional band, the reset time and the rate time ($1/K_c$, T_r and T_d respectively) are the parameters that need be varied but they are not independent. The root contours can be obtained in such a situation for the optimum choice of these parameters. For the effective implementation of this approach a digital computer is of help.

4.4.4 Bode-plot Technique

In process control systems the degree of stability required is to be known for the system *a priori*. The general requirement, from the frequency characteristics, are

- (i) Gain margin (GM) of 2, i.e., 6 dB.
- (ii) Phase margin (PM) of 30° .

For servo systems these values are 4 and 45° respectively.

If the gain is 1.0, i.e., 0 dB at 150° phase lag, the phase margin is 30° and the corresponding frequency known as, sometimes, the gain cross-over frequency is very nearly the frequency of the damped oscillation $\omega_n \sqrt{1 - \zeta^2}$. Critical frequency is the frequency where the phase is -180° .

Peaking in the response curve is also considered for stability. The peak magnitude of the frequency response when the loop is redrawn as a unity feedback system, is given as

$$M_p = \max \left| \frac{c(j\omega)}{r(j\omega)} \right|_{\omega} \quad (4.62)$$

For process control system $3 > M_p > 2$, whereas for servo-systems $1.6 > M_p > 1.2$.

However, large overshooting is not always permissible even in process control system, e.g., (i) in annealing, overheating will adversely affect grain growth and (ii) in some chemical reactors, overshooting may start the reaction prematurely. Critical damping adjustment is then preferred but the 'line-out' time may be too long.

In the following the discussion pertaining to Bode-plot technique of stability analysis is presented in a little detail.

Also known as corner plots, Bode-plots are the logarithmic plots of the loop-transfer function. For a closed loop the transfer function is given as

$$T(s) = \frac{G(s)}{1 + G(s)H(s)} \quad (4.63)$$

From the characteristic equation [Eq. (4.45b)] again the stability can be adjudged as follows:

$$G(s)H(s) < -1 \quad (4.64)$$

An equality sign in Eq. (4.64) would mean that both gain and phase margins are critical.

Gain margin is defined as the ratio of the gain at which the system becomes unstable to the actual system gain assuming no phase change from $\phi = -180^\circ$.

Phase margin is the amount of negative phase shift which must be added to make the system unstable assuming no gain change from $|GH| = 1$.

Bode plot is usually made from the loop-transfer function and it consists of the gain plot in dB and the phase plot. As the loop-transfer function, in general, is available in factored form, the method seems to be quite useful because the product factors in $F(s) = G(s)H(s)$, ($s = j\omega$), become additive terms and by drawing straight line asymptotes, the approximate function is easily drawn.

Expressing

$$F(s) = \frac{K_1 \prod_{j=1}^m (s + \alpha_j) \prod_{k=m+1}^{\mu} \left(\frac{s^2}{\omega_{ck}^2} + \frac{2\zeta_k}{\omega_{ck}} s + 1 \right)}{\prod_{i=1}^n (s + \beta_i) \prod_{p=n+1}^{\eta} \left(\frac{s^2}{\omega_{cp}^2} + \frac{2\zeta_p}{\omega_{cp}} s + 1 \right)} \quad (4.65)$$

The magnitude in dB of $F(j\omega)$ is obtained as

$$\begin{aligned} 20 \log_{10} |F(j\omega)| &= 20 \log_{10} |K| \\ &+ 20 \sum_{j=1}^m \log_{10} |1 + j\omega\tau_j| + 20 \sum_{k=m+1}^{\mu} \log_{10} \left| 1 + j \frac{2\zeta_k \omega}{\omega_{ck}} - \frac{\omega^2}{\omega_{ck}^2} \right| \\ &- 20 \sum_{i=1}^n \log_{10} |1 + j\omega\tau_i| - 20 \sum_{p=n+1}^{\eta} \log_{10} \left| 1 + \frac{j2\zeta_p \omega}{\omega_{cp}} - \frac{\omega^2}{\omega_{cp}^2} \right| \end{aligned} \quad (4.66)$$

where

$$K = K_1 \frac{\prod_{j=1}^m \alpha_j}{\prod_{i=1}^n \beta_i}, \quad \tau_j = \frac{1}{\alpha_j}, \quad \tau_i = \frac{1}{\beta_i} \quad (4.67)$$

The phase of $F(j\omega)$ is obtained as

$$\begin{aligned} \angle F(j\omega) = \angle K + \sum_{j=1}^m \angle(1 + j\omega\tau_j) + \sum_{k=m+1}^{\mu} \angle\left(1 + j\frac{2\zeta_k\omega}{\omega_{ck}} - \frac{\omega^2}{\omega_{ck}^2}\right) \\ - \sum_{i=1}^n \angle(1 + j\omega\tau_i) - \sum_{p=n+1}^{\eta} \angle\left(1 + j2\zeta_p\frac{\omega}{\omega_{cp}} - \frac{\omega^2}{\omega_{cp}^2}\right) \end{aligned} \quad (4.68)$$

$\angle K$ is zero for K positive and is π for K negative; in both the cases, however, this is independent of frequency. All other arguments are dependent on frequency. Another type of term that may appear, particularly in servomechanism, is $(j\omega)^{\pm r}$ $r = 1, 2, \dots$. Thus in general four different types of terms appear both in $|F(s)|$ and $\angle F(s)$. The techniques of drawing the Bode plots for these terms are now considered briefly.

- (i) The simplest is the constant term K . It has amplitude and phase characteristics, as shown in Fig. 4.18. The values are obtained as

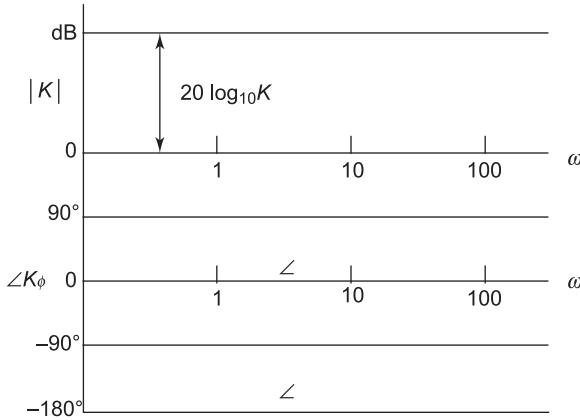


Fig. 4.18 Gain and phase plots for K

$$K_{dB} = 20 \log_{10} |K| = \text{constant}$$

$$\angle K_\phi = 0^\circ \text{ or } -180^\circ \text{ (constant)} \quad (4.69)$$

- (ii) Poles or zeros at the origin: $s^{\pm m}$, for which the dB magnitude (M_{dB}) and phase magnitude (M_ϕ) are given by

$$\begin{aligned} M_{dB} &= 20 \log_{10} |(j\omega)|^{\pm m} \\ &= \pm 20m \log_{10} \omega \end{aligned} \quad (4.70)$$

$$\angle M_{o\phi} = \pm m\pi/2$$

where suffix o refers to the origin.

These give a magnitude curve which will rise or fall (+ or – sign) at the rate of $20 m$ dB per decade or $20 m \times 0.301 = 6 m$ dB per octave.

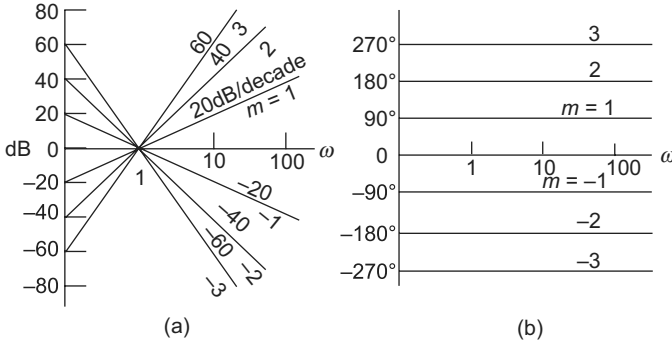


Fig. 4.19 Gain (a), phase (b) plots for the terms $(s)^{\pm m}$

(iii) Simple pole or zero: $(1 + s\tau)^{\pm 1}$

The magnitude in dB at a point s is given by

$$M_{s\text{dB}} = \pm 20 \log_{10} |1 + j\omega\tau| = \pm 20 \log_{10} \sqrt{1 + \omega^2 \tau^2} \quad (4.71a)$$

and the phase by

$$\angle M_{s\phi} = \pm \tan^{-1} \omega\tau \quad (4.71b)$$

This type occurs most frequently. The magnitude curve is easily drawn by linear asymptote approximation. When $\omega\tau \ll 1$, $M_{s\text{dB}} \approx 0\text{dB}$ and when $\omega\tau \gg 1$, $M_{s\text{dB}} = \pm(20 \log \omega + 20 \log \tau)$. The curves are similarly obtained as in the case above. As ω increases, $\omega\tau$ also increases and this aspect similarly gives a curve of 20 dB per decade rise or fall depending on the + or – sign. The asymptotes for $\omega\tau \ll 1$ and $\omega\tau \gg 1$ intersect at a frequency given by

$$20 \log_{10} \omega_c \tau = 0$$

$$\text{i.e.,} \quad \omega_c = 1/\tau \quad (4.72)$$

on the 0 dB line. The quantity ω_c , is known as the corner frequency. This approximation yielding straight-line asymptotes differs from the actual results slightly. Standard tables may be made for correcting these errors and in the plots discussed the error is symmetrical with respect to ω_c . At this frequency it is 3 dB, at $\omega_c \pm 1$ octave, 1 dB and, at $\omega_c \pm 1$ decade 0.3 dB. The phase can also be approximated by straight lines by choosing ± 1 decade about ω_c and drawing a straight line from 0° to $\pm 90^\circ$ between them. The maximum error obtained in doing so is $\pm 6^\circ$. Figure 4.20 shows the plots of magnitude and phase in this case.

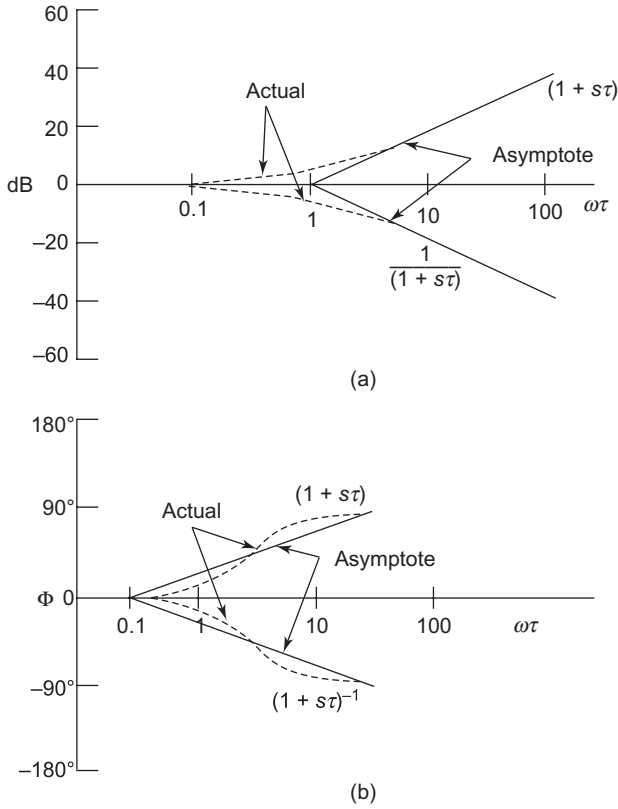


Fig. 4.20 Gain (a), phase (b) plot for the terms $(1 + s\tau)^{\pm 1}$

(iv) Quadratic poles and zeros:

$$\left(1 + 2\frac{\zeta s}{\omega_n} + \frac{s^2}{\omega_n^2}\right)^{\pm 1}$$

For the poles

$$F_q(j\omega) = \frac{1}{1 - (\omega/\omega_n)^2 + 2j\frac{\omega\zeta}{\omega_n}} \quad (4.73)$$

with magnitude in dB and phase given, respectively, by

$$M_{q\text{dB}} = -20 \log_{10} \left[\left\{ 1 - \left(\frac{\omega}{\omega_n} \right)^2 \right\}^2 + 4\zeta^2 \omega^2 / \omega_n^2 \right]^{1/2} \quad (4.74a)$$

and

$$\angle M_{q\phi} = -\tan^{-1} \frac{2\zeta}{\omega_n/\omega - \omega/\omega_n} \quad (4.74b)$$

When

$$\omega/\omega_n \ll 1, M_{q\text{dB}l} \cong 0 \quad (4.75a)$$

and when

$$\frac{\omega}{\omega_n} \gg 1$$

$$\begin{aligned} M_{q\text{dB}h} &\cong -20 \log_{10} \{ [1 - (\omega/\omega_n)^2]^2 + (2\zeta\omega/\omega_n)^2 \}^{1/2} \\ &\cong -20 \log_{10} \sqrt{(\omega/\omega_n)^4} = -40 \log_{10} \omega/\omega_n \end{aligned} \quad (4.75b)$$

i.e., there is again a fall that asymptotically approaches a straight line with a slope of -40 dB per decade and the intersection point is again obtained from $-40 \log_{10} \omega/\omega_n = 0$, giving the corner frequency as

$$\omega_c = \omega_n \quad (4.76)$$

In Eqs (4.75a) and (4.75b) suffixes l and h stand for low and high respectively.

Around the corner frequency, however, the magnitude plots will be substantially different because of the presence of the term ζ , the damping ratio. For $\zeta = 0.707$ the approximation is more valid. The dB magnitude and phase plots with ω/ω_n as the x -axis in log-scale are shown in Fig. 4.21(a) and (b).

For a pair of quadratic zeros these curves would be reversed. It is imperative that for use of these curves the ζ and ω_n values be known a priori. The quadratic factors are not very common in process control systems.

Example 1

Consider now the closed-loop system of Fig. 4.22 such that one derives

$$\begin{aligned} F(s) = G(s)H(s) &= \frac{K}{(1+s)(1+0.5s)(1+0.2s)(1+0.1s)} \\ &= \frac{K}{Q(s)} \end{aligned} \quad (4.77)$$

Solution The gain and phase plots are superposed in Fig. 4.23. For $K = 1$, the gain margin is shown to be 24 dB and the phase margin 90° . The parameter K can now be increased for faster response (for proportional

control as shown in Fig. 4.22; this means a decrease in the proportional bandwidth) such that GM is 6 dB and PM is 30° . The first requirement means that the amplitude plot has to be raised by 18 dB such that

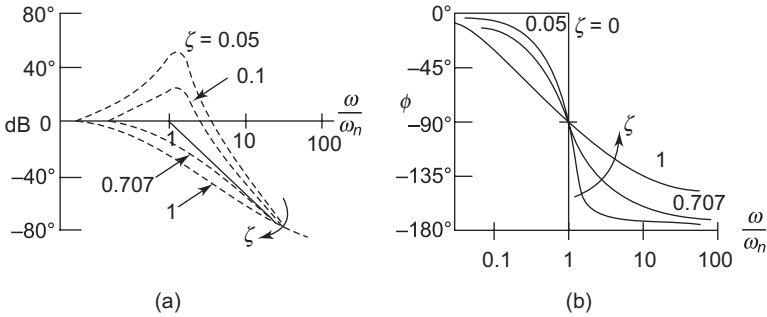


Fig. 4.21 Gain (a), phase (b) plots for quadratic terms

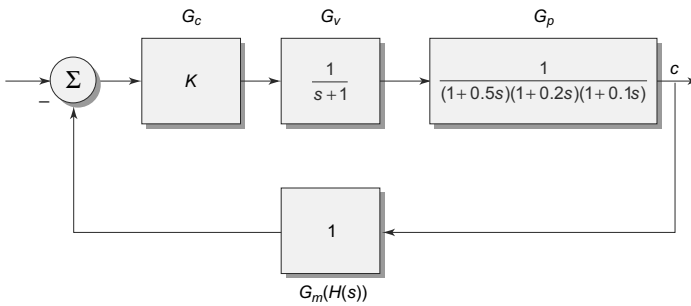


Fig. 4.22 Block representation of Example 4.1

$$20 \log_{10} K = 18, \text{ giving } K = 7.95$$

This gives a phase margin of nearly 60° and is well within the limits of specification. Since provision of separate control of phase margin is kept with the integral and derivative time constants of the controller, no further elaboration is required for this case.

When $K = 1$, the gain margin is called gain limit. In that case

$$\text{Gain margin} = \text{Gain limit} - \text{Gain } K \quad (4.78)$$

all expressed in decibels. In absolute values

$$\text{Gain margin} = \text{Gain limit}/K \quad (4.79)$$

The frequencies at phase and gain cross overs are easily obtained from the plot of Fig. 4.23.

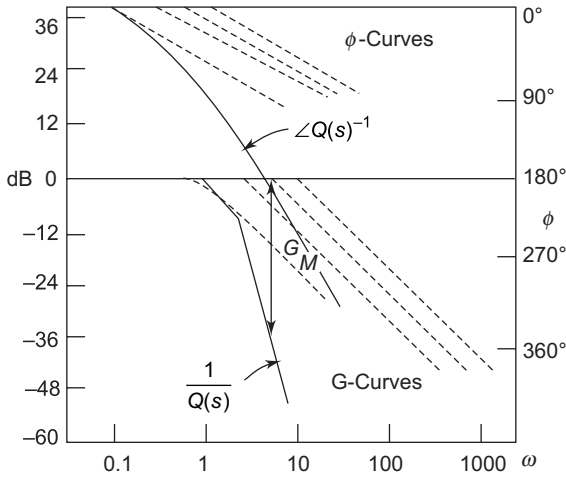


Fig. 4.23 Gain and phase plots for Example 1

Example 2

In a closed loop system, load block has a transfer function $K_L(s) = 2/(2s + 1)$ (see Fig. 4.5), $K_I(s) = 2/(2s + 1)$, $K_2 = 1/2$ and $K_c(s) = 10/(1 + sT_R)$. Calculate the damping ratio and natural frequency of oscillation for $T_R = 4$ sec. Obtain the error integral for this condition. If $T_R = 0.2$ sec. by what ratio does it change?

Solution The transfer function is

$$\frac{c(s)}{u(s)} = \frac{2(2s + 1)}{1 + \frac{2}{2s + 1} \cdot \frac{1}{2} \cdot \frac{10}{1 + sT_R}} = \frac{2(1 + sT_R)}{2s^2T_R + s(2 + T_R) + 11}$$

Hence,

$$\omega_n = \sqrt{11/(2T_R)}, \quad \zeta = (1/2)(2 + T_R)/\sqrt{(2T_R \cdot 11)}$$

For $T_R = 4$ sec., $\omega_n = \sqrt{11/8} = 1.17$, and $\zeta \approx 0.33$

Therefore, $\int_0^\infty |e_1| dt = 0.5(2/(1 + 11))(3.14/1.17\sqrt{1 - 0.1}) \approx 0.23$

For $T_R = 0.2$ sec., $\omega_n = \sqrt{11/0.4} \approx 5.25$ and $\zeta = (1/2)2.2/\sqrt{4.4} \approx 0.51$

Therefore, $\int_0^\infty |e_2| dt = 0.5(2/(1 + 11))(3.14/5.25\sqrt{1 - 0.26}) \approx 0.059$

The change is roughly in the ratio of 4:1.

4.5 COMPENSATORS

After necessary studies on transient response, controllability and stability have been made, a brief note on a special aspect of system design is appended here which when incorporated in the system enhances the overall system performance. This is regarding compensation. Often a system is seen to have poor relative stability and sluggish transient response. In fact, from the viewpoint of pole-zero locations, it is now well known that introduction of a zero at appropriate location in the complex plane can improve the transient response and stability; also, poles are to be properly located for good steady state behaviour. Locating the poles and zeros in appropriate places in the complex plane is done by what is known as compensation technique. A zero too close to the $j\omega$ -axis gives rise to high peaking in transient response and therefore it is avoided. Also a single zero cannot be brought in the transfer function because of constraint on physical realizability. Hence a pole-zero pair of the type

$$G_k(s) = (s + z_k)/(s + p_k) = (s + 1/\tau)/(s + \alpha/\tau),$$

$$\alpha = p_k/z_k > 1, \tau > 0 \quad (4.80)$$

is introduced ensuring that the pole is to the left of the zero in the left half plane to have its effect to be as little as possible. A compensation with transfer function of the form of Eq. (4.80) is known as a lead compensator. In addition to increasing speed of response and relative system stability, it helps to increase the system error constant to a certain extent. For a sinusoidal input given to this network shows that its output leads the input under steady state condition and hence this name.

If the steady state error is large, i.e., offset is large, inspite of a satisfactory transient response; it must be reduced, ideally, by adding a pole at the origin as has been discussed in the previous chapter. But addition of a pole at the origin, however, degrades the transient response. This, in turn, is remedied by adding a compensating zero very close to the pole at the origin in the left half plane. The changed situation can be well visualized in a root-locus plot. This situation for specified transient response, i.e., with given damping constant ζ and gain K identified, would actually show that a closed loop pole appears on the real axis very close to the compensating zero but on the left of it. This of course, has to be ensured with proper design of the compensator. A cascaded compensator with a transfer function

$$G_k(s) = (s + z_k)/s \quad (4.81)$$

would do the job with proper choice of z_k . But the physical realizability of the compensator requires that there is no pole at the origin and the pole is, consequently, shifted on the real axis slightly towards the left of the origin with the transfer function changing to

$$G_k(s) = (s + z_k)/(s + p_k) = (s + 1/\tau)/(s + 1/\beta\tau),$$

$$z_k/p_k = \beta > 1, \tau > 0 \quad (4.82)$$

This is called a lag compensator which improves the steady state performance without affecting the transient response characteristics. For a sinusoidal input, the output of such a network lags the input in phase in the steady state condition and hence the name.

Often a combination of lead-lag compensator is used. This is when improvement both in transient and steady state responses is required. They are connected in series. The lead, lag or lead-lag compensator is very easily designed as electrical circuit using resistances and capacitances, and, if required, active blocks like operational amplifiers. Typical two such passive R - C lead and lag networks are shown in Figs. 4.24 (a) and (b) respectively. The transfer function in the two cases are as given by Eqs (4.80) and (4.82) respectively, where $\tau = RC$, $\alpha = 1 + R/R'$ and $\beta = 1 + R'/R$.

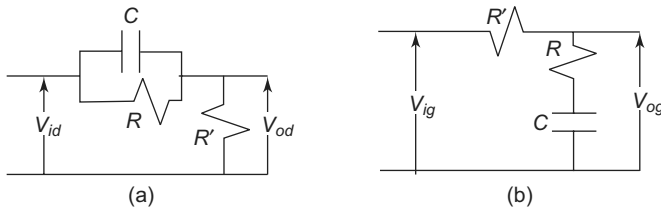


Fig. 4.24 (a) Lead circuit (b) Lag circuit

Earlier, compensators were designed mostly on trial and error basis. Specific design techniques are now available by which such trials can be avoided for given ζ , i.e., dominant root location specifications. One such method is the Warren-Ross method for lead compensator design. It must be emphasised that compensation design can be carried out both in the time as well as in frequency domains equally conveniently. Standard texts on control system engineering covers such designs fairly elaborately.

As told already, compensation is provided to enhance system performance. This connotes consideration where to locate the compensator. Different views are there and by permutation it can be located in series of the process called cascade compensator, in the feedback path called the parallel compensator, in the input side or the output side. It may be of interest to note that the controller itself is a cascade compensator. It would be seen in the above that compensator provides additional poles and/or zeros to decrease peak deviation, settling time, steady state error and also peak-occurring time. In frequency domain it is used to decrease peak frequency response and increase resonant frequency, bandwidth, phase and gain margins. Root locus technique, Nyquist criteria, Bode plot technique may be considered for appropriate design of the compensators.

In servomechanism or reference follower systems, compensating devices were used for compensation extensively at some points of time. In process control the controller is adequately designed for the purpose. In some cases additional controllers are used as in case of feed-forward control (2 loop control) a 'load-compensator' is used to compensate disturbance produced in the process (see Chapter 6, Section 6.5), which basically is a lead compensator/derivative control function generator.

Introducing poles and zeros, existing zero or pole can be cancelled to provide new zero and/or pole for improved performance. Since the objective is to have better controllability without hampering stability, gain and phase margin conditions of the existing system are considered and the compensator is designed accordingly for its gain and phase around the 'ultimate' (see Chapter 5) frequency.

Considering a cascade compensator of the transfer function

$$G_c(s) = \frac{K_c(s + \alpha_1)}{s + \beta_1} \quad (4.83a)$$

Putting $s = j\omega$

$$G_c(j\omega) = \frac{K_c \alpha_1 \left(1 + \frac{j\omega}{\alpha_1}\right)}{\beta_1 \left(1 + \frac{j\omega}{\beta_1}\right)} \quad (4.83b)$$

For unity feedback one has the characteristic equation for process $G_p(s)$,

$$1 + G_c(s)G_p(s) = 0 \quad (4.83c)$$

so that

$$G_c(j\omega) = -\frac{1}{G_p(j\omega)} \quad (4.83d)$$

From Nyquist plot (see earlier in the chapter) one can see that the compensator phase angle $\angle \theta = \angle G_c(j\omega)$ may be written in terms of the phase margin ϕ_m at the frequency ω_c where the unit circle crosses the plot as

$$\theta_c = -180^\circ + \phi_m - \angle G_c(j\omega_c) \quad (4.84a)$$

so that at $\omega = \omega_c$

$$G_c(j\omega_c) = r \angle \theta_c \quad (4.84b)$$

where

$$r = \left| \frac{1}{G_c(j\omega_c)} \right| \quad (4.84c)$$

Combining Eqs (4.83b) and (4.84b)

$$\frac{\angle \theta_c}{|G_c(j\omega_c)|} = \frac{K_c \alpha_1 \left(1 + j\omega_c / \alpha_1\right)}{\beta_1 \left(1 + j\omega_c / \beta_1\right)} \quad (4.85)$$

or,

$$\frac{\cos \theta_c + j \sin \theta_c}{|G_c(j\omega_c)|} = \frac{K_c \alpha_1 \left(1 + j\omega_c / \alpha_1\right)}{\beta_1 \left(1 + j\omega_c / \beta_1\right)} \quad (4.86)$$

From which the design parameters are obtained as

$$\alpha_1 = \omega_c \frac{K_c |G_c(j\omega_c)| - 1}{K_c \sin \theta_c |G_c(j\omega_c)|} \quad (4.87a)$$

and

$$\beta_1 = \omega_c \frac{K_c |G_c(j\omega_c)| - \cos \theta_c}{\sin \theta_c} \quad (4.87b)$$

The equations can be used for obtaining the parameters for both lead and lag compensators.

Review Questions

1. How can the controllability of a process assessed from the process reaction curves? What is the effect of disturbance on plant controllability when it occurs at different intermediate points?
2. Explain the terms (a) deviation reduction factor, (b) subsidence ratio, (c) proportional control factor. How is proportional control factor related to plant controllability?

Two systems are to have same controllability having same subsidence ratio of 4. One has an offset, the other does not. How are their proportional control factors related?

(Ans: $K_{p0} = 4(1 + K_p)/3$)

3. In a process control system, $G_p = K_p/(s + 1/\tau_p)$, $G_v = K_v$, $G_m = 1$ and $G_c = K_c(1 + 1/sT_R)$, obtain the roots for $T_R \rightarrow \infty$ first and then show how the system performance changes with T_R decreasing?
4. What is the measure of controllability in the frequency domain for a process control system? How is it evaluated?

Responses to a step disturbance to two systems show peak values of 20 and 30 per cent of the step respectively and oscillation

periods of 0.21 and 0.17 secs respectively. Estimate their relative controllability.

5. The characteristic equation of a closed loop system is given by

$$s^6 + 5s^5 + 4s^4 + 3s^3 + 2s^2 + s + 1 = 0$$

Examine the system stability in details using Routh criteria and Routh-Hurwitz criteria.

6. Use root-locus technique to study the stability problem when (a) poles are added and (b) zeros are added to the loop transfer function. The loop transfer function of a system is given by $K(s+1)/(s(s+2)(s^2+2s+1))$. Obtain (i) the breakaway points, (ii) the intersection points of root loci and (iii) the gain for limiting stability.

7. Show how the Bode-diagram technique can be used in the design of a process control system. What is the significance of gain and phase margins? Relate these to the gain-bandwidth product of the system.

8. A process of transfer function $\exp(-0.5s)/(2s+1)$ is controlled with a measurement feedback of $3/(1.5s+1)$. For an upset at the start of the process, find the controller settings using root-locus technique. (*Hint: Expand $\exp(-0.5s)$ and choose only upto the first order terms and then proceed.*)

9. What do you mean by self-regulation of a process?

In a two-tank process having three identical valves, the input valve is half-open, the output valve is full-open and the interconnection valve is half-closed, show how the regulation changes from that, when all the valves are equally open.

10. Make the Bode-plots of a practical PID controller and an ideal PID controller and show the difference in their performances. Assume equal values of K_c , T_D and T_R in the two cases.

(*Hint: The transfer functions in the two cases are*

$$G_c(s)_p = K_c(1 + sT_R)(sT_D + 1)/(sT_R(sT_D/10 + 1)) \text{ and}$$

$$G_c(s)_i = K_c(1 + \frac{1}{sT_R} + sT_D)$$

11. Obtain the block diagram and the transfer function between the tank level and valve lift in a level process. Sketch the schematic of the process and the corresponding block diagram.

(*Hint: The diagrams are shown in Figs Q-411(a) and (b). Direct material balance principle may be applied to obtain*

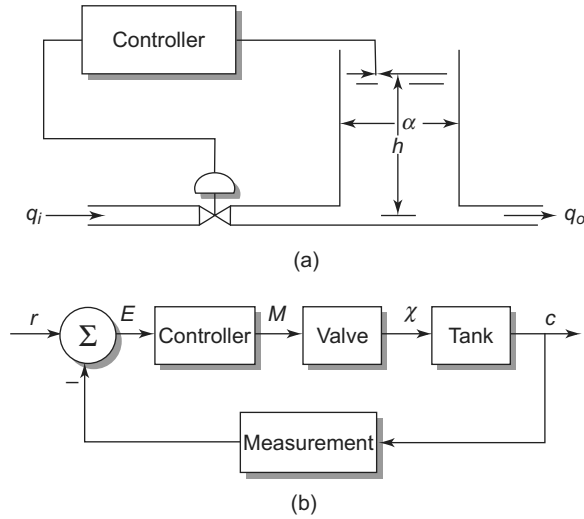


Fig. Q.4.11 (a) Schematic diagram of a level control system
(b) Block diagram of Fig. Q.4.11 (a)

$$\begin{aligned}\alpha(dh/dt) &= q_i - q_o \text{ giving } \alpha(d^2h/dt^2) = dq_i/dt - dq_o/dt \\ &= (\partial q_i / \partial h)_x dh/dt \\ &+ (\partial q_i / \partial x)_h dx/dt - (dq_o/dh)(dh/dt), \text{ or, } \alpha s^2 h = (\partial q_i / \partial h)_x sh \\ &+ (\partial q_i / \partial x)_h sx - (dq_o/dh)sh,\end{aligned}$$

$$\begin{aligned}\text{or, } h(s)/x(s) &= \frac{(\partial q_i / \partial x)_h}{\alpha s + \frac{dq_o}{dh} - \left(\frac{\partial q_i}{\partial h}\right)_x} \\ &= \left[\frac{(\partial q_i / \partial x)_h}{dq_o/dh - \left(\frac{\partial q_i}{\partial h}\right)_x} \right] \bigg/ \left[\frac{\alpha s}{\frac{dq_o}{dh} - \left(\frac{\partial q_i}{\partial h}\right)_x} + 1 \right] = \frac{K}{s\tau + 1}\end{aligned}$$

where, $\left(\frac{\partial q_i}{\partial x}\right)_h \bigg/ \left[\frac{dq_o}{dh} - \left(\frac{\partial q_i}{\partial h}\right)_x \right]$ and $\tau = \alpha \bigg/ \left[\frac{dq_o}{dh} - \left(\frac{\partial q_i}{\partial h}\right)_x \right]$

12. A system is represented by the state equation

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -4 & -5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u$$

Show that it is controllable.

(Hint: The characteristic equation is $|A - \lambda I| = \begin{vmatrix} -\lambda & 1 \\ -4 & -5-\lambda \end{vmatrix} = 0$

giving $\lambda_{1,2} = 4, 1$ so that according Eq. (4.35), $\begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \end{bmatrix}$

$$= \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} + Q^{-1}Bu$$

Now $Q = \begin{bmatrix} 1 & 1 \\ \lambda_1 & \lambda_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 4 & 1 \end{bmatrix}$, so that $Q^{-1} = \begin{bmatrix} -1/3 & 4/3 \\ 1/3 & -1/3 \end{bmatrix}$ and

$$Q^{-1}B = \begin{bmatrix} -1/3 & 1/3 \\ 4/3 & -1/3 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 4/3 \\ -1/3 \end{bmatrix}.$$

Since Q^{-1} has all non zero elements, the system is completely controllable.)

13. Why do we use IAE criterion for controllability study of process control systems? Discuss by explaining the nature of the process and their controlled conditions.
14. Prove that for state controllability, the composite matrix $[B : AB : A^2B : \dots : A^{n-1}B]$ must have a rank of the system matrix A .
15. How do you justify that a controller is just a compensator? How do you design this compensator using Nyquist criterion? Take a sample example and show the steps of design.
16. Obtain the DRF in a system if its proportional control factor is 22 and subsidence ratio 4 for the cases (1) the system does not have an offset, and (2) it has an offset. Which is better controllable, by what %?

[Hint. Refer to Eq.(4.18) for case (1) and from Eq. (4.17), if it has an offset, Thus, $K_f = 22$, $\rho = 4$ give

$$\delta_{f(1)} = 23/(3/2) = 15.33$$

$$\delta_{f(2)} = 22 \times 2 = 44$$

Case (2) is better controllable by $\{(44 - 15.33)/15.33\} \times 100 = 187\%$

17. Two systems have their natural frequencies of oscillation as 4 r/s and 3.1 r/s and the damping factor 0.22 and 0.31 respectively. If the peak deviations in the two cases for a unit step disturbance are 31% and 22% respectively, what are their relative performance indices?

[Data are $\omega_{n1} = 4 \text{ r/s}$, $\omega_{n2} = 3.1 \text{ r/s}$

$$\zeta_1 = 0.22, \zeta_2 = 0.31$$

$$e_{p1} = 0.31, e_{p2} = 0.22$$

Referring to Eq. (4.23), $I_1 = 0.31/2 \cdot \pi / \{4\sqrt{1 - (0.22)^2}\} = 0.1243$

and
$$I_2 = 0.22/2 \cdot \pi / \{3.1\sqrt{1 - (0.31)^2}\} = 0.1166$$

Hence, the system 2 is better controllable by a percent of $|(0.1166 - 0.1243)/0.1166| = 6.6\%$

Basic Control Schemes and Controllers

5.1 INTRODUCTION

In a closed loop process control system the components are classified as (i) Process equipment and (ii) Control equipment. Control equipment consists of the controller, the measurement unit, the comparator and the actuator with the power unit. For efficient control, each of these units should be appropriately designed or chosen. In this chapter controllers are discussed at certain length along with the basic control schemes that are commercially important. Such control schemes comprising the more simple and direct ones are introduced here. These are (1) On-off control, (2) Duration adjusting control or time-proportional control, (3) Proportional control, (4) Integral control or proportional speed floating control or reset action control, (5) Derivative control or rate control, and (6) Programme control.

These are obtained directly following the controller forms and hence while discussing the control schemes controllers of the specific variety are also discussed with their practical adaptability as far as possible.

5.2 ON-OFF CONTROL

When the control valve or the final control element has only two positions, fully closed or fully open, as actuated by the controller, the control scheme is known as on-off control.

When the process reaction rate (PRR) is low but the process has high demand side capacity (DSC), on-off control is recommended. PRR is usually understood as the change in the process function per unit time. Demand side capacity is defined as the ability of the process to withstand

demand changes. Low DSC would mean a large (and often sudden) change in the process function when demand from and by the process increases and for high DSC, change in process function is small.

In on-off control the controller output shuts off the control valve when the controlled process variable just exceeds the set point. This closure means that the control variable now decreases, reaches the set point and goes down below it. When the variable crosses the set point mark in this downward journey, the control valve opens fully again. This opening allows the variable to increase gradually, and in this way, the valve opening and closing, and the process variable go on cycling continuously, the speed of which depends primarily on the process lag. This cycling of the control valve, is, assuming no lag in the control equipment, as shown in Fig. 5.1 along with the process variable. Because of the time delay and what is known as the differential gap in the control equipment elements, the actual

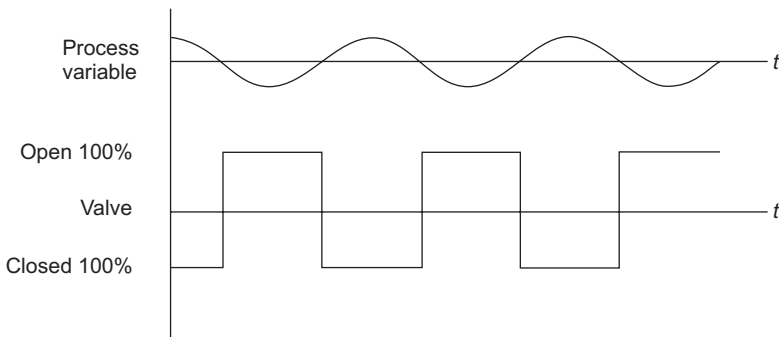


Fig. 5.1 Sketch showing the cycling in the control valve with change in the process variable

response curve changes. This time delay in this system is known as the dead time. The differential gap is analogous to the minimum input hysteresis in the overall transducer-controller-actuator system in the sense that this is defined as the smallest change in the process variable that would change the state of the control valve. Differential gap allows the controlled process variable to be oscillatory. The process variable grows or decays from the set point in different ways depending on the type of process. Assuming equal hysteresis on both sides of the set point, one can show that the diagram of response would be as shown in Fig. 5.2. The exponential nature comes because of the process. The effect of the unequal differential gap on the two sides is to make the closing and opening times of the valve unequal. Besides this, if the measurement lag is such that the exponential nature is virtually straightened out, one can easily show that by doubling the differential gap the frequency of opening and closing of the valve may be reduced to one half its value. The effect of the dead time is to increase

the differential gap. In Fig. 5.2, the dotted line response curve is drawn with a dead time t_d superposed over the differential gap δ_1 . The effective differential gap is then δ_2 . The period of oscillation changes from T_1 to T_2

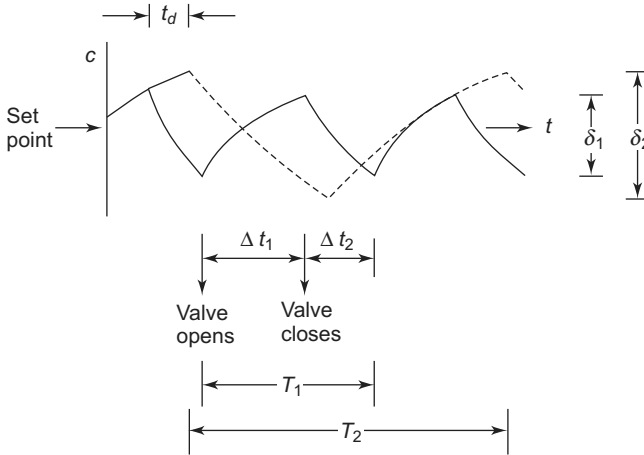


Fig. 5.2 Sketches showing the effect of the dead time on the differential gap

which is in the same ratio as δ_1 to δ_2 as long as the process lag is large. Quantities Δt_1 and Δt_2 are known as the rising and falling times of the cycle. For getting a linear approximation of the exponential nature (shown in Fig. 5.2) of the response curves, one can calculate by how much the controlled process variable increases or decreases as a result of dead time. It must be remembered that this rise or fall may not be the same if the rise and fall rates of the process variable are different as dictated by the process. The temperature of a furnace or the level of a liquid in a tank with different inflow and outflow pipe diameters will show different rise and fall rates. If the process variable is v , the process capacity is C and the input and output (energy or equivalent) to the system that change the process variable are E_i and E_o , then the approximately linear response equation for rise or fall can be written as

$$\pm(E_i - E_o) = Cdv/dt \quad (5.1)$$

This relation can be used to calculate any of the quantities marked in Fig. 5.3(b) when the controller characteristics are known. One can also obtain the change in the time period of oscillation due to the dead time. To calculate the parameters marked such as oscillation amplitude and time period, for simplicity, the process is considered ideally integrating in nature so that the response characteristics will be linear. From Fig. 5.3(a), the system equation is

$$(u - m)k/s\tau = c \quad (5.2a)$$

i.e.,

$$u - m = (\tau/k)(dc/dt) \quad (5.2b)$$

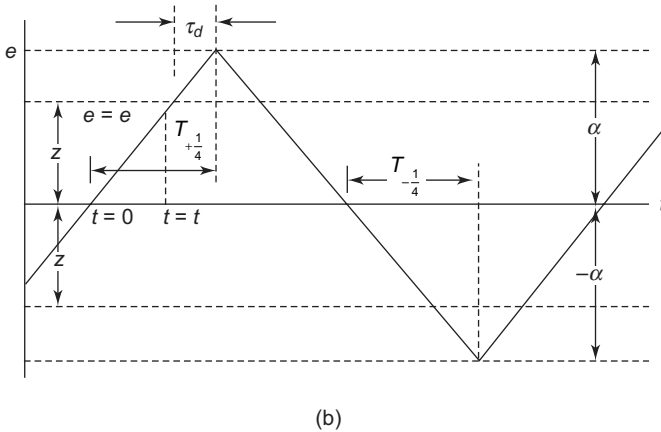
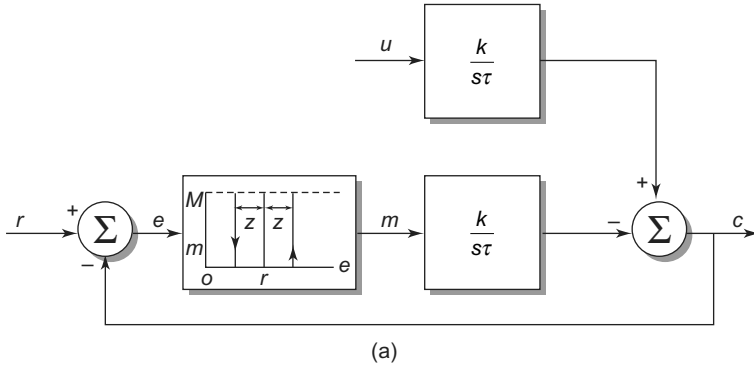


Fig. 5.3 (a) On-off control system for an ideal integrating process
(b) The control curve

For a controller with a dead zone $2z$, during the phase $c < r + z$, $dc/dt > 0$; output is M , and during the phase $c > r - z$, $dc/dt < 0$; output is zero (see Fig. 5.3b). This is because of hysteresis z on both sides of the controller actuating point (set point), actuator does not get the output M from the controller when c crosses r downwards, neither output is zero when c crosses r upwards. A dead time further deteriorates the situation. τ_d may be considered such a time. From Eq. (5.2a) with the above condition fulfilled

$$u - M = (\tau/k)(dc/dt), \quad \text{for } c < r + z, \quad dc/dt > 0, \text{ and}$$

$$u = (\tau/k)(dc/dt), \quad \text{for } c > r - z, \quad dc/dt < 0 \quad (5.3)$$

Now, since $e = r - c$, $de/dt = -dc/dt$, so that Eq. (5.3) changes to

$$u - M = (\tau/k)(-de/dt), \quad \text{and,} \quad u = (\tau/k)(-de/dt)$$

which on integration yield

$$e = t(M - u)/(\tau/k) + k', \quad \text{and} \quad e = -tu/(\tau/k) + k'' \quad (5.4)$$

Applying boundary conditions, for a causal system, $k' = k'' = 0$. The response is represented in the curve of Fig. 5.3(b), this is the error curve. For specific instantaneous time t , error is e which becomes $\alpha - z$ for a time τ_d . τ_d here, is the dead time.

Hence, from the curve

$$e/t = (\alpha - z)/\tau_d \quad (5.5)$$

From Eqs (5.4) and (5.5),

$$(M - u)/(\tau/k) = (\alpha - z)/\tau_d \quad (5.6a)$$

for the rising curve from the set point, and,

$$-u/(\tau/k) = -(\alpha - z)\tau_d \quad (5.6b)$$

for the falling part from the set-point line.

Subtracting Eq. (5.6b) from Eq. (5.6a), one gets

$$(k/\tau)(M) = 2(\alpha - z)/\tau_d \quad (5.7)$$

giving

$$\tau_d = 2\tau(\alpha - z)/(kM) \quad (5.8)$$

and

$$\alpha = kM(\tau_d + 2\tau z/(Mk))/(2\tau) = \tau + \tau_d kM/(2\tau) \quad (5.9)$$

showing the dead-time-amplitude relationship in terms of known parameters. The oscillation period may be known similarly. Thus, from the curve, rising and falling parts from the set point line being considered separately, with time till maximum marked as $T_{+\frac{1}{4}}$ and minimum $T_{-\frac{1}{4}}$,

$$e/t = \pm \alpha/T_{\pm\frac{1}{4}} \quad \text{so that}$$

$$T_{+\frac{1}{4}} = \alpha/(e/t) = \alpha/[M - u]/(\tau/k),$$

$$\text{and} \quad T_{-\frac{1}{4}} = -\alpha/(e/t) = -\alpha/[(-u)/(\tau/k)] = \alpha\tau/(uk)$$

Total time period is

$$T_p = 2 \left(T_{+\frac{1}{4}} + T_{-\frac{1}{4}} \right) = 2\alpha\tau(1/M - u) + 1/u)/k \\ = 2\alpha\tau M/(ku(M - u)) \quad (5.10)$$

Using Eq. (5.9), it comes out to be

$$T_p = (\tau_d + 2\tau_z/(Mk))M^2/(M - u)u \quad (5.11)$$

If however, the process is represented by $k/(s\tau + 1)$, the nature of the rising and falling curves would be as shown in Fig. 5.2 which are exponential in nature and the analysis to be done as such.

It should be stressed here that even if the differential gap or the dead time are fixed, the demand side capacity change would change the process reaction rate (Cf. Eq. 5.1) as also the period and amplitude of oscillation. A higher dv/dt would then mean a low time period and a large amplitude and vice versa.

A very common practical example of an on-off control is the microsen system which is shown schematically in Fig. 5.4. The coil pair are made in the form of discs or vanes; a disc or a vane on the pointer balance arm moves in them when desired temperature is reached so the oscillator starts oscillating and emitter gets a current because the transistor becomes active. This current makes the potential across relay R_1 nearly zero so that the contact in the control circuit R_1 opens. In consequence relay R_2 is deenergized and in turn contact R_2 opens putting off the supply to the heater coils/elements. As the temperature falls below set point, the balance arm moves out of the disc-coil-pair and oscillations stop resulting in a potential drop between points B and A , A being effectively at ground point whereas potential of B is $V_{cc} R_4/(R_3 + R_4)$. This allows relay R_1 to energize closing contact R_1 and in turn energizing relay R_2 thereby closing contact R_2 putting the furnace supply on.

5.3 TIME PROPORTIONAL CONTROL

In the time proportional control action the final control element takes either an on or off position but the ratio of the on time to off time is proportional to the value of the controlled variable, the on plus off time remaining constant. Depending on the deviation of the controlled variable from the set-point this ratio changes. The control scheme is also known as duration adjustable control. Electrical controllers provide this facility easily. In such situations the final control elements connected directly to the controller are either contractors, solenoid valves or high speed two position motors.

A simple scheme of a duration adjustable controller is shown in Fig. 5.5(a). V_i corresponds to the value of the controlled variable with a maximum of

V_p the reference, with which it is compared, the resultant being compared again with a recycling time base of duration T and peak value V_p so that

$$V_0 = (V_p - V_i) - V_p t/T \quad (5.12)$$

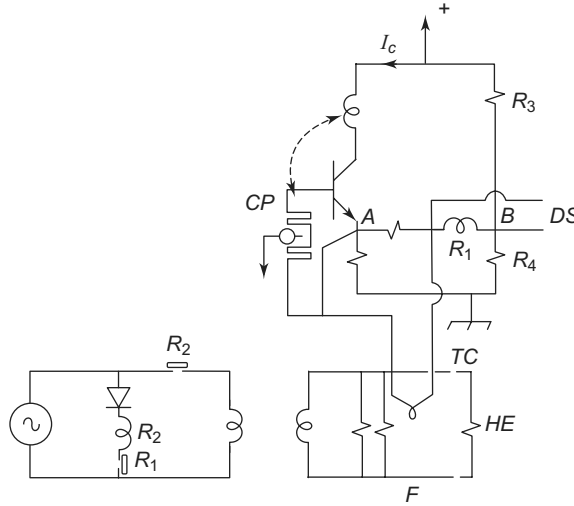


Fig. 5.4 Practical scheme of an on-off controller; TC: thermo couple, HE: heating element, F: furnace, DS: detector signal, CP: coil pair

As long as $V_p - V_i > V_p t/T$, the emitter follower gives an output which is amplified and fed to the contactor coil for putting on the process. Over the period when $V_p t/T > V_p - V_i$, the emitter follower does not give any output. With a small V_i , $V_p - V_i$ is large and thus the on time becomes larger than the off time. The converse is also true. As the correction of error is taking place during the operation, there is further change in the durations of the on and off conditions. The total on-off period however, remains same which is the duration of recycling time base T . The moderate period recycling time base, which could be obtained by a motorized system, is presently obtainable through a microprocessor-based design.

The design flexibility of the circuit of Fig. 5.4, can be increased by connecting a capacitor across the coil-pair which can be made adjustable. This capacitor made of sections of concentric cylinders may be driven by a synchronous motor for adjustment. The T_{on} and T_{off} may now be adjusted by adjusting the value of this capacitor made of the coil-pair. Just before the beginning of the off period during which the relay R_1 is to be deenergized (Fig. 5.4), this capacitance adds to the tuning of the oscillator circuit which, in turn, controls the relay. The vane position (on the pointer arm) relative to the oscillator coil pair and a certain position of the cylindrical capacitor plates (driven by the motor) must coincide to produce the off period.

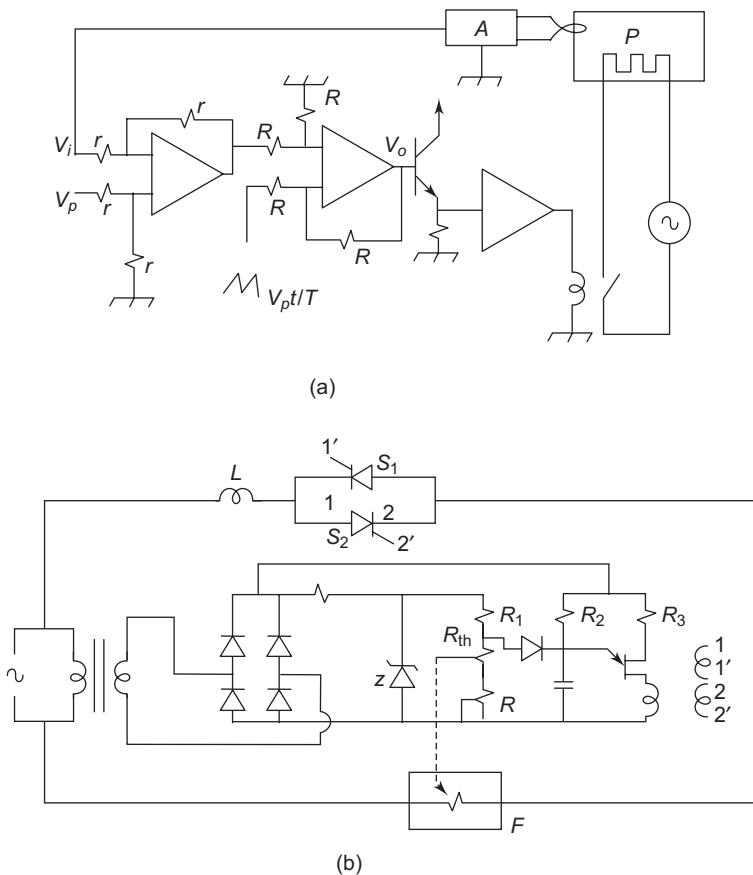


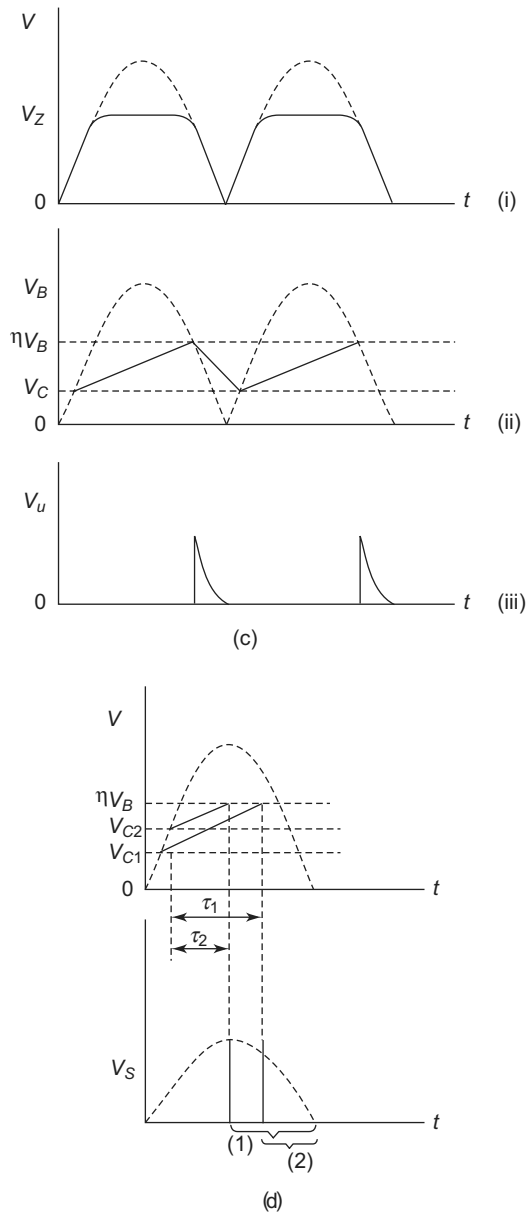
Fig. 5.5 (a) Scheme of a duration adjustable controller, P: process, A: amplifier, (b) A temperature control scheme, R_{th} thermistor, F: furnace

As the controlled variable approaches the set point, the vane covers an increasingly larger area between the coil pair resulting in an increase of the off periods.

Another temperature controller of the time proportional type depending on the furnace condition is shown in Fig. 5.5(b). Figure 5.5(c) shows in steps the generation of *UJT* firing pulses whose time intervals are dependent on the voltage V_c across the capacitor and is given by

$$V_c = V_z(R + R_{Th})/(R + R_{Th} + R_1) \quad (5.13)$$

This is the pedestal height and as the sensor is a thermistor of resistance R_{Th} , V_c falls with increase of temperature. The capacitor starts charging through the resistor R_2 from the initial voltage V_c and as soon as the voltage rises to ηV_B , where η is the intrinsic stand off ratio of the *UJT*, the *UJT* fires


Fig. 5.5

(c) The UJT firing curves (i) zoner output, (ii) UJT firing, (iii) UJT output, (d) The SCR firing curves, (i) SCR on with pedestal V_{C2} (cold condition), (2) SCR on the pedestal V_{C1} (hot condition)

and a spike voltage is induced in the secondary of the *UJT* load (a pulse transformer). Coils 1-1' and 2-2' are connected with proper polarity to the *SCR*'s such that one of these (S_1 or S_2) fires on each half cycle. The control action is executed but over the control of the pedestal height V_c through thermistor. Higher the V_c , lower is τ , the firing interval, and vice versa. For lower temperature, higher is V_c , capacitor charges upto ηV_B quicker and *UJT* fires early. Figure 5.5(d) shows the firing schedule for two V_c 's.

Every *SCR* has its own di/dt value. It may so happen that the off-*SCR* is turned on at the peak value of voltage resulting in a large di/dt . To prevent this an inductance, L , is connected in series with the *SCR*. The fired *UJT* would send two spikes to start the two *SCR*'s but the one whose anode is at the positive potential at that cycle would only turn on. Triggering voltage usually lies between 1 and 2 volts and during conduction drop across the *SCR* is about 2 volts. Diode D is used so that the capacitor may discharge only through the *UJT*. Two *SCR*'s are used as high power triacs are still not easily available. The circuit of Fig. 5.5(b) also shows the bridge rectifier scheme and the zener stabilizer for obtaining the pedestal voltage, as also the furnace heater schematically. Although this is theoretically a time proportional control, in practice it becomes a continuous control with variable heat supply depending on temperature of the furnace because of 50 hertz operation. If the operation cycle can be drastically reduced to, say 10 cycle/hour, it works effectively as a duration adjustable controller.

A practical scheme of on-off control where dead zone can be controlled by adjusting a resistor is given in Fig. 5.5(e).

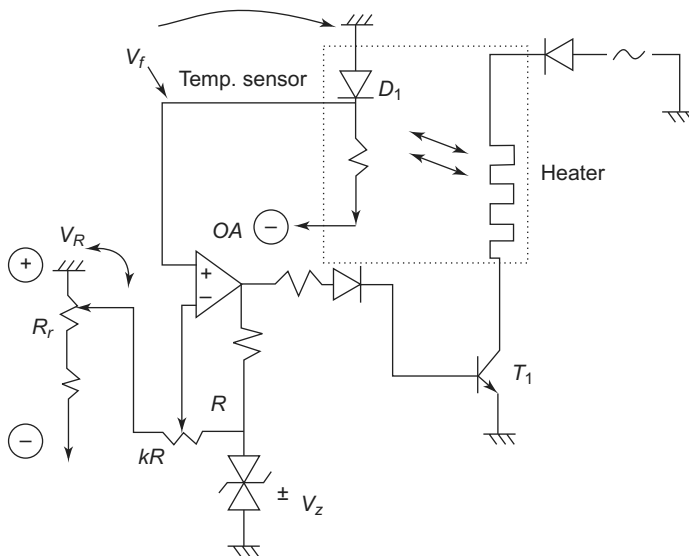


Fig 5.5 (e) On-off control scheme with adjustment dead zone

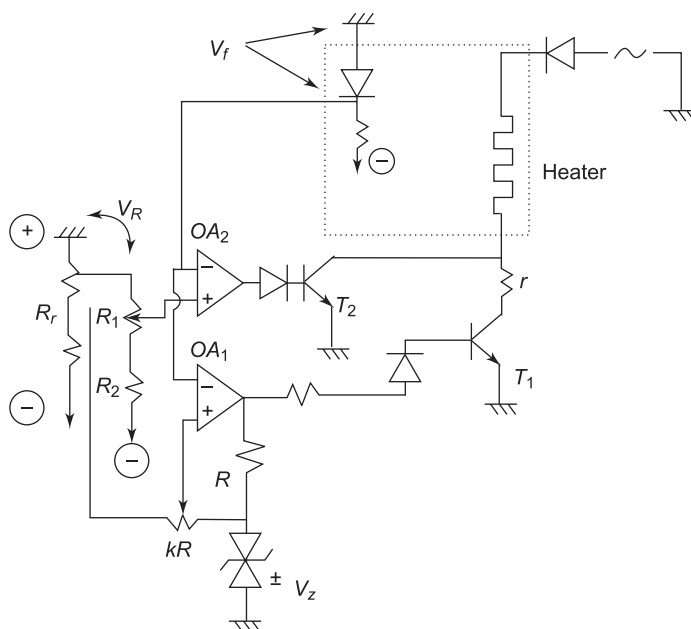


Fig. 5.5 (f) On-off control scheme with adjustable dead zone and reducible overshoot

The set point comparison and hysteresis are provided by the *OA* where a heater is controlled by switching transistor T_1 on or off. Diode D_1 senses the temperature ($T < 150^\circ\text{C}$) and with forward bias its linearity is good. Sensitivity is $-2 \text{ mV}/^\circ\text{C}$. Temperature changes output V_f of the diode which is compared against the set point voltage V_R by the comparator designed by the *OA*. The potentiometer R can be used in positive feedback mode here to produce desired hysteresis with V_f reaching the reference plus the hysteresis voltage. The comparator switches to put the transistor off when temperature T is given by

$$T = T_R + \frac{V_R - V_f(T_R) \pm kV_z}{-2 \text{ mV}/^\circ\text{C}} \quad \text{for } V_R < V_f(T_R) \text{ and } R_r \ll R$$

There occurs, however, an overshoot, making the differential gap large because of system delay. To make this small, a multilevel on-off controller shown in Fig. 5.5(f) is used. It is a two-level controller. Both the comparators drive the heater, supply transistor switches T_1 and T_2 but with the additional resistance r in series with the heater switching produces different heat conditions. Comparator OA_2 is given the bias chain R_1 - R_2 and with this added bias OA_2 allows the higher level switch T_2 to turn off before the actual set point (V_R) is reached. The resistance r limits the

heater current in the condition so that there is low level heating after T_2 is off and the overshoot is less.

A sustaining (proportional) heater control formed with two OA 's is shown in Fig. 5.5(g). It is basically a PWM circuit with feedback from heater-sensing diodes as shown. The changes in diode voltages due to heating are integrated by OA_1 and the resulting voltage change in R_2 alters the charging current in C_2 .

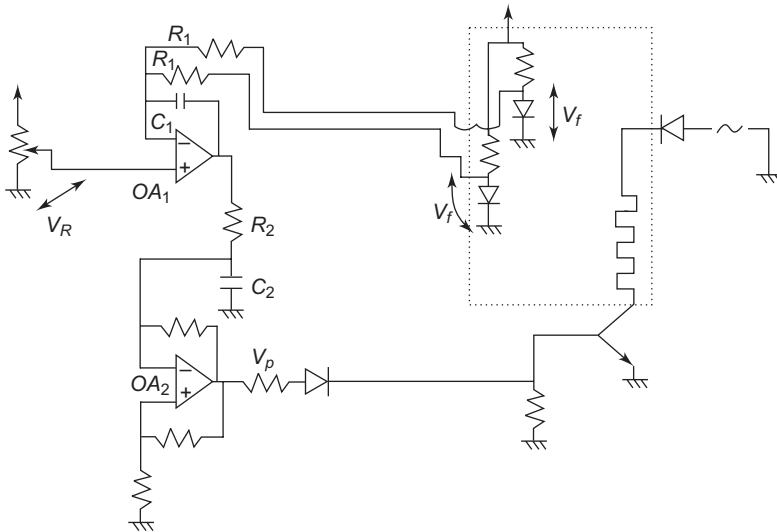


Fig. 5.5 (g) PWM Scheme for heater control

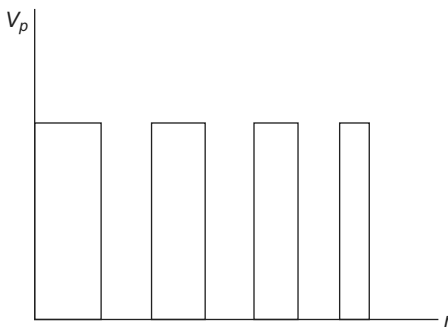


Fig. 5.5 (h) The output pulses from the scheme of Fig. 5.5(g)

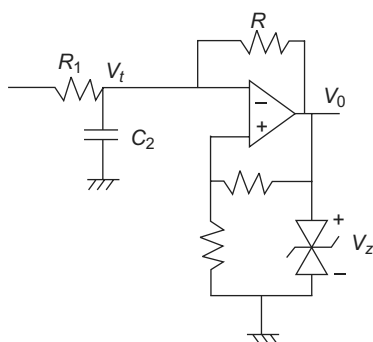


Fig. 5.5 (i) Circuit explaining the pulse production

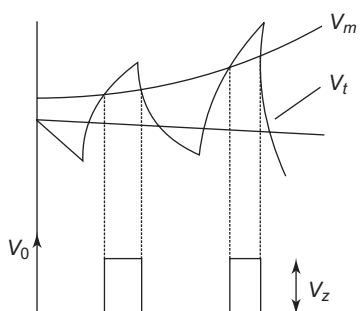


Fig. 5.5 (j) The graph showing generation of pulses

The charging and discharging rates of C_2 determine the output pulse widths Fig. 5.5(h) from OA_2 , this is explained in Fig. 5.5(i) and (j) the PWM circuit and the resulting duty cycles of heater current. With this, changes in pulse width error decreases. At equilibrium, the switching occurs for zero error. The duty cycle for Fig 5.5 (i) is given as

$$D/C = \frac{\ln\left(1 + 2 \frac{V_z}{V_m}\right)}{\ln(1 - 4 V_z^2 / V_m^2)}$$

5.4 TYPICAL PID CONTROLLER CHARACTERISTICS AND RELATED TERMINOLOGY

A generalized approach of proportional-integral-derivative (PID) controllers and control schemes are presented first and separate cases are then discussed. As reset and rate actions are normally associated with proportional action, these are accordingly discussed in relation to proportional action.

The PID controller transfer function that is often sought in the construction of most commercial controllers is given by

$$G_c(s) = K_c(1 + 1/(T_r s) + T_d s) \quad (5.14)$$

Proportional action associated with K_c is called proportional band, PB, and is inversely related to the concept of gain, K_c . Doubling gain K_c requires that the proportional band be halved.

Controller output is actually fed to the actuator. With proportional action only the controller output varies proportionally to the deviation from the set-point. When the input to the controller (with comparator) gradually increases above the set value, the actuator goes on closing and vice versa. If for a 50 per cent increase, the actuator completely closes and for a 50 per cent decrease, it completely opens, the proportional band is said to be 100 per cent. Likewise, if the actuator closes by 75 per cent for a 50 per cent increase and remains 25 per cent closed for a 50 per cent decrease, then the proportional band is 200 per cent and so on. The proportional band is, therefore, defined as the percentage of full scale change of the controlled variable required to operate the valve through full stroke. This is shown schematically in Fig. 5.6. If the proportional band becomes zero, an on-off controller results, and if it becomes infinity, control action ceases.

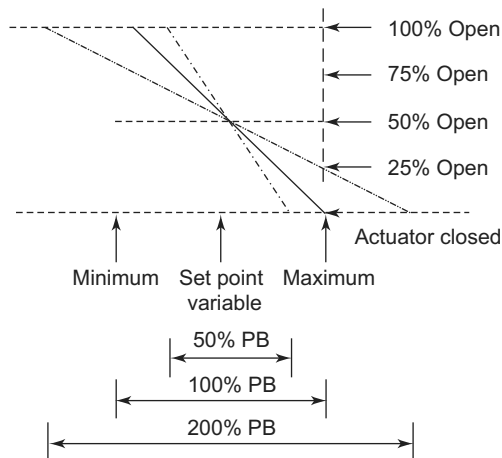


Fig. 5.6 Sketch showing the setting of proportional band

If the full scale change of the controlled variable c is V , controller output is y and K_s is the controller sensitivity, then

$$K_s = dy/dc \quad (5.15)$$

Again, if Δy_0 is the full scale change of controller output required to produce full actuator stroke, the required change in c is then

$$\Delta c_0 = \Delta y_0 / K_s \quad (5.16)$$

$\Delta y_0 = 1.0 - 0.2 = 0.8 \text{ kg/cm}^2$ for pneumatic controller and, $20 - 4 = 16 \text{ mA}$ for electronic controller.

Proportional band is, now, given by

$$PB = \Delta c_0 \times 100 / V = 100 \Delta y_0 / (VK_s) \text{ per cent} \quad (5.17)$$

Depending on the dynamics of the process to be controlled, simple combination of the ideal three actions is often used. One such combination is $P + I$ control in which

$$G_c(s) = K_c(1 + 1(sT_r)) \quad (5.18)$$

where T_r is the 'reset time' and $1/T_r$ is the reset rate.

The purpose of the reset action is to provide adequate control action on varying demands from and by the process. When a demand from the process is fully met by P -action alone, the I -action is unnecessary, but if a demand change is made from the process it is unlikely that the simple P -action will be able to cope with it. With proportional action alone, the actuator is set to open by some specified value even at zero error. This is the control point. This allows specific amount of input (energy) to the process. When demand increases (say), it means that at control point, more energy should be input to the system. But because of the control point setting this is not permitted, since, as proportional error becomes zero, valve opens by the specified amount only. As the additional demand is not met a discrepancy remains between the expected controlled variable and that obtained which, in this case, is less. This is the offset. With reset action, however, this error or deviation would send an additional signal to the actuator whose value is proportional to the time integral of the deviation allowing the actuator to open more, and more energy is input to the process at the control point. As long as the deviation is nonzero the corrective action would continue this way.

The demand by the process (and consequently from the process) can thus be met such that the deviation and the proportional output start falling towards the set point. The procedure will bring the offset to zero. By a proper choice of the P -action and I -action parameters, $1/K_c$ and T_r , it is possible to check the rise in the I -unit output and steady the P -unit output when the demand is fully met. Then outputs from the P -unit and I -unit together give the required output by the actuator. Over adjustment would give 'overshoot' and under adjustment retains a small offset. Figure 5.7 explains the situation. The statement $T_r \rightarrow 0$ means infinite I -action, the output increasing instantaneously tending to destabilize the system; and when $T_r \rightarrow \infty$, I -action is zero. Actually T_r is the interval for integral action to cause the same amount of output as that produced by proportional action due to a steady deviation from the set-point.

Another mode of control is $P + D$ control when

$$G_c(s) = K_c(1 + sT_d) \quad (5.19)$$

where T_d is the rate time.

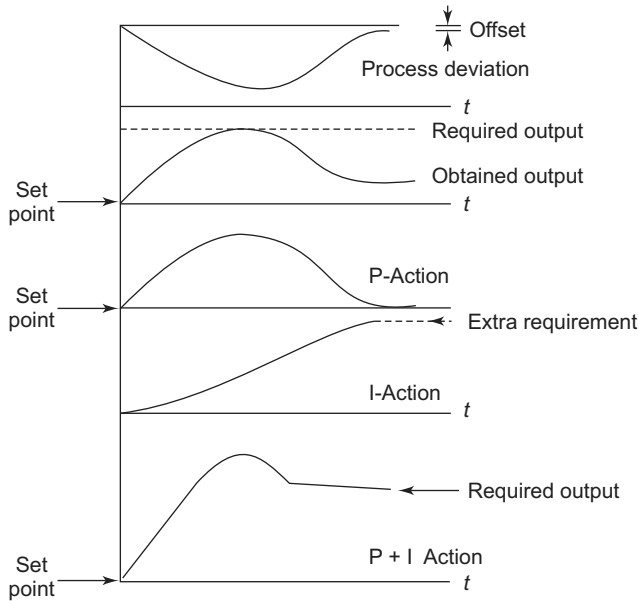


Fig. 5.7 Sketch showing the effects of P-action and PI-action, curve in the second figure from the top indicates settlement with P-action only, P: proportional, I: integral

This mode is complementary to the I -action. When a process runs with small disturbances, wide proportional band and low I -action (i.e., large T_r) are usually preferred. Because of these settings, the action of the controller is too slow, so to say, and less responsive. If now, suddenly, large disturbances occur at wide intervals, this PI action alone will not be sufficient and a D -action may be introduced to counteract such immediate deviations and ensure immediate corrective action. As soon as the deviation becomes zero, D -action stops. D unit thus detects a change in the increase or decrease of deviation and provides an output in proportion to the speed of the deviation. For a constant deviation, the output is zero. Figure 5.8 explains this situation. For $T_d \rightarrow 0$, no derivative action results. If $T_d \rightarrow \infty$, least speed change of deviation gives full output and control stability is obviously hampered. Hence, T_d should be chosen optimally in relation to K_c (and T_r where present).

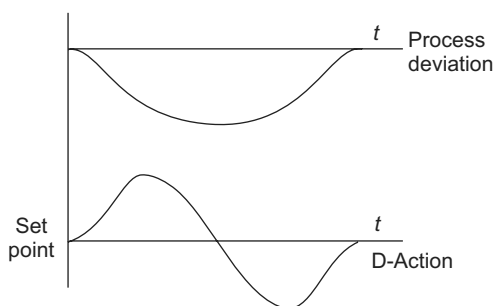


Fig. 5.8 Sketch showing the effect of derivative action, D: derivative

The third mode of control is represented by Eq. (5.14) and is known as $P + I + D$ control. Derivation of the form of Eq. (5.14) is simple. Refer to Fig. 5.9.

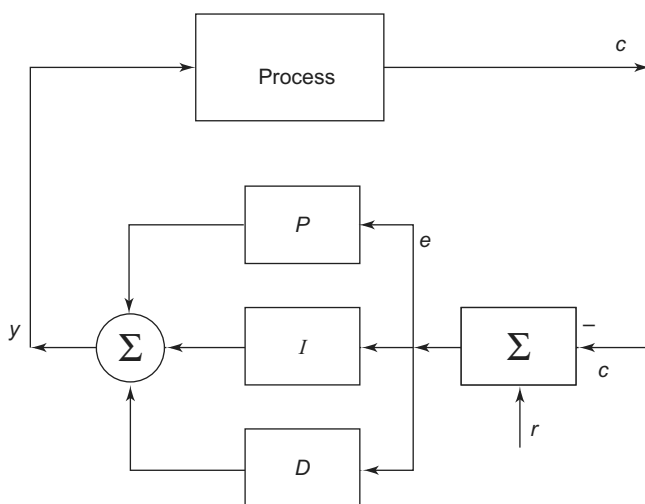


Fig. 5.9 General sketch of the controller with process for derivation of Eqn (5.14)

In this case one has

$$y = y_P + y_I + y_D = K_p(e + 1/T_r \int e \cdot dt + T_d dr/dt) + y_0 \quad (5.20)$$

where error $e = r - c$, $K_p/T_r = K_I =$ constant of proportionality for the I -action unit and $K_p T_d = K_D =$ constant of proportionality for the D -action unit with $K_p =$ the constant of proportionality for the P -action unit. The term y_0 represents the value of the manipulating variable at control point. Even when error is zero this value is kept, otherwise the control valve will be completely closed and the system cannot function properly.

Time response characteristics of *PID* controller with the mode of action used individually or in combination are obtained from the basic time response equations of the units. For *P* and *I* actions, let the error be a step, E , then for *P*-action

$$y_p = K_p E + y_0 \quad (5.21)$$

and for *I*-action

$$y_i = tE/T_i + y_0 \quad (5.22)$$

where K_I is replaced by $1/T_i$. Fig. 5.10(a) shows the response curves for $E = 1$. Combined action of *P* and *I* units is represented by

$$y_{pi} = (K_p + t/T_i)E + y_0 \quad (5.23a)$$

$$= K_p(1 + t/T_r)E + y_0 \quad (5.23b)$$

and is shown in Fig. 5.10(a) by the dotted for $E = 1$.

For the derivative action, a step error would yield no change in output. For time response of this type of action it is, therefore, assumed that the error is a ramp such that $e = Et$ and the guiding equation becomes

$$y_D = K_D E + y_0 \quad (5.24a)$$

As stated earlier derivative action is generally not used alone but associated with the proportional action so that the combined *P* and *D* actions produce a time response for ramp input as

$$\begin{aligned} y_{PD} &= K_p Et + T_D E + y_0 \\ &= K_p E(t + T_d) + y_0 \end{aligned} \quad (5.24b)$$

where T_D is used in place of K_D and $T_d = T_D/K_p$ is the rate time. The net effect of the derivative action is to shift the controller output ahead in time by T_d . For this reason, the derivative control is sometimes called the anticipatory control. Its response characteristics alongwith *P*-action is shown in Fig. 5.10(b) for $E = 1$.

Additionally, now, if *I* action is also included. The time response equation is given as

$$y_{pDi} = K_p E(t + T_d + t^2/(2T_r)) + y_0 \quad (5.25)$$

for a ramp input $e = Et$. The response curve for this case is also shown in Fig. 5.10(b).

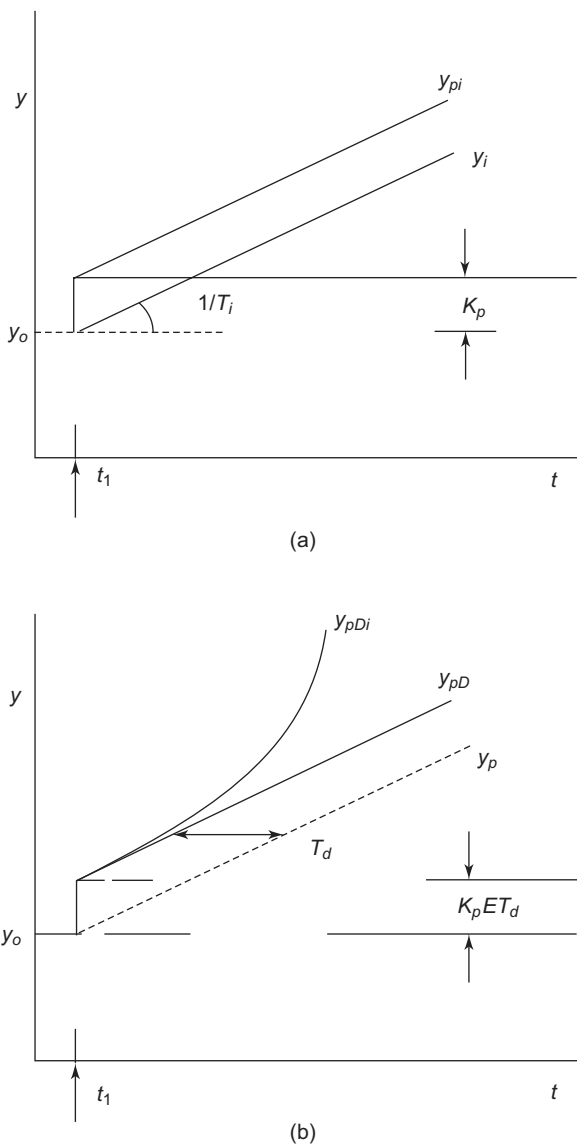


Fig. 5.10 PID controller characteristics test curves, (a) step input test, (b) ramp input test; t_1 : time of input

It should be remembered that as soon as specific control action or actions start to operate the error or deviation also starts to reduce and the response curves so drawn above are changed accordingly. However, these curves are still very useful for calibration of the control units with different actions when specific types of deviations are introduced.

5.4.1 Comments on Controller Transfer Functions and Controller Application Note

The operational equation of the PID controller with initial condition zero is given by

$$y(s)/e(s) = G_c(s) = K_c(1 + 1/sT_r + sT_d)$$

as given by Eq. (5.14). This is a quasi-ideal form of the transfer function with mild interaction between the modes of action, specifically when P -action parameter is adjusted. For optimal control of different types of processes, however, all the modes should be separately adjustable.

In most pneumatic controllers independent settings are impossible; however, electronic controllers can be designed in different ways to make settings independent to a large extent. But this is again rarely done, for it increases the cost.

The response function of a practical PI controller is usually obtained as

$$G_{pi}(s) = K_c A(T_r s + 1) / [(1 + A)T_r s + 1] \quad (5.26)$$

with $100 < A < 1000$ so that reasonable approximation is made when Eqs (5.26) and (5.18) are held identical.

For derivative action things are a little complicated. It is an action dependent on the speed of deviation; therefore, the band of frequencies over which this should be allowed to act is limited. This is because amplitude saturation may occur and cause the system to deteriorate. A practical PD controller is represented by

$$G_{PD(s)} = K_c \frac{T_d s + 1}{T_d s/B + 1} \quad (5.27)$$

where B is greater than 5 but less than 20 and is fixed for a single design. Thus a more realistic transfer function of a PID controller is given by

$$G_{CR(s)} = K_c \left(\frac{T_r s + 1}{T_r s} \right) \left(\frac{T_d s + 1}{T_d s/B + 1} \right) \quad (5.28)$$

The best way to set up an industrial control system is theoretically impossible because it is hardly possible to determine beforehand all the factors which are likely to influence the behaviour of the process. The sensitivity of the controller and of the sensing element, response time of the controllers and other instruments in the loop to a given variation of the input, response of the process to external disturbance, etc., are some of the items to be considered.

The shortest recovery time design is one of the best designs but in this the value of the controlled variable should not go beyond the specified

limits. The time lag of a control loop normally consists of the sum of the lags of individual blocks. However, the major contribution is from the process itself. The dynamic behaviour of the loop is usually evaluated from the transient response of the loop. Response to a unit step function is mainly evaluated for this. The transient response of the process actually determines the choice of the types of controllers. Table 5.1 below gives a brief account as to how to choose a controller for a specific process when the process response is known for a step input. The higher order response row in this table is a general case in which category 80 per cent of the industrial processes fall in.

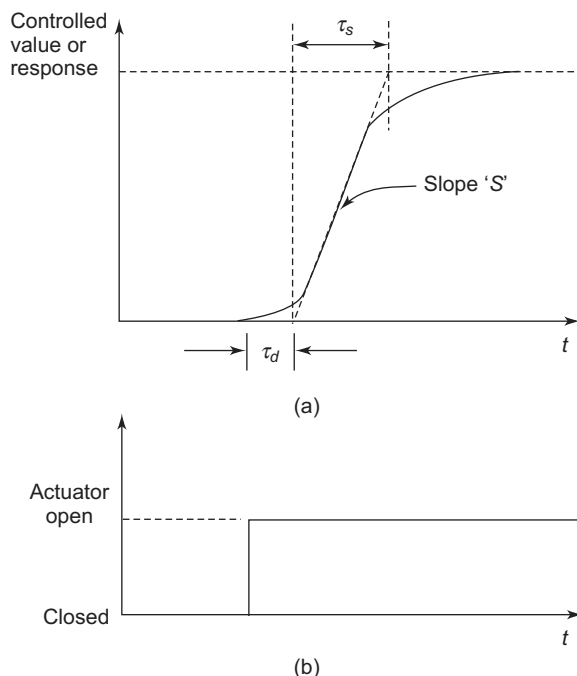


Fig. 5.11 Process reaction (a) along with the control valve positions (b)

Two characteristic parameters are important in the selection of controllers and control settings. Fig. 5.11 shows the response, in general, of a process to step input and the parameters. They are (i) the transfer lag or the time constant τ_s , and (ii) the dead time, τ_d .

If τ_s is relatively larger than τ_d , the process is easily controllable with a simple controller and a single control loop. A thumb-rule is laid out here for general reference:

$\tau_s / \tau_d \geq 10$: Easily controllable,

$\tau_s / \tau_d \approx 6$: Difficult to control with one controller,

$\tau_s / \tau_d \leq 3$: Controllable with only complex control loops.

Table 5.1 Choice of control mode depending on the transient response of the process

SERIAL	PROCESS RESPONSE TO INPUT	PRACTICAL EXAMPLE	CONTROL MODE
1	Ideal, without dead time	Short fluid pressure lines	$I, P + I$
2	Ideal, with a pure dead time	Conveyor belt, Gas pressure line	$P, P + I$
3	First order	High speed soldering iron	$P, P + I$
4	Non-stabilizing	Level in a tank with over head supply and bottom discharge	$P, P + D$
5	Higher order	Furnace and most others	$P + D, P + I,$ $P + I + D$

5.5 COMPARISON OF CONTROL ACTIONS: PID

It is now known that different control actions ‘improve’ or resist the change in the control variable. However, the optimum values of control parameters like proportional band, reset time and rate time can be obtained for different processes by actually obtaining the deviation time characteristics for different values of the parameters. This is all the more true as the process tends to become of a higher and higher order and exact solutions by computers are possible provided system functions are known sufficiently accurately. Rules for the adjustment of these parameters from the gain bandwidth product etc. have been elaborated later but it may be interesting to notice the characteristics for different types of processes and for different sets of parameter values. For lower order systems particularly, the results are analytically derivable and the idea for adjustment for higher order systems is, in fact, obtained from the characteristics of even the lower order ones.

Consider the unit feedback system shown in Fig. 5.12. Unit feedback is considered for brevity. The controlled variable is given as

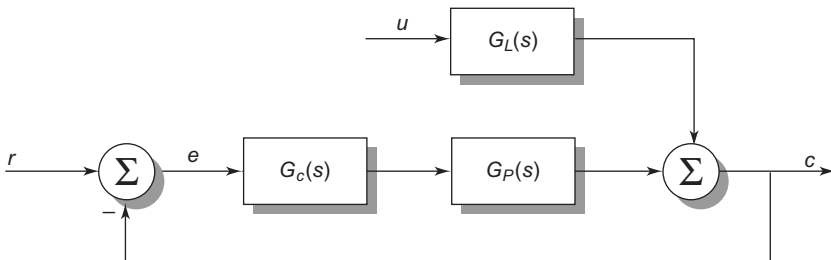


Fig. 5.12 Sketch of a simple feedback system with upset

$$c = \frac{G_c G_p r}{1 + G_c G_p} + \frac{G_L u}{1 + G_c G_p} \quad (5.29)$$

From Eq. (5.29), and knowing $r - c = e$, one obtains

$$e = \frac{r}{1 + G_c G_p} - \frac{G_L u}{1 + G_c G_p} \quad (5.30)$$

In this equation we can now choose different forms of G_c , and G_p can be chosen to be of the first order or of a higher order. For brevity G_a has been taken in the process or made unity. As a typical case of P and I action control of a first order process, assume

$$G_c = K_c(sT_r + 1)/sT_r, \quad G_p = K_p/(s\tau + 1) \text{ and } G_L = K_L/(s\tau + 1)$$

such that Eq. (5.30) changes to

$$e = \frac{sT_r(s\tau + 1)r}{s^2T_r\tau + sT_r(1 + K_cK_p) + K_cK_p} - \frac{sT_rK_Lu}{s^2T_r\tau + sT_r(1 + K_cK_p) + K_cK_p} \quad (5.31)$$

which is easily written in the form

$$K_cK_p e = \frac{sT_r(s\tau + 1)r}{s^2/\omega_N^2 + 2\zeta s/\omega_N + 1} - \frac{sT_rK_Lu}{s^2/\omega_N^2 + 2\zeta s/\omega_N + 1} \quad (5.32)$$

where

$$\omega_N = \sqrt{K_cK_p/(T_r\tau)} = \text{loop natural frequency,}$$

and

$$\zeta = (1 + K_cK_p)/2\sqrt{T_r/(K_cK_p\tau)} = \text{damping ratio}$$

Now, if the set point remains unchanged, response characteristics for load change/disturbance can be obtained from Eq. (5.32). If the disturbance is a unit step, the normalized deviation $c/\Delta u|_{\Delta u \rightarrow 1}$ is plotted with time for varying gain K_c with T_r constant and varying T_r with K_c constant as shown in Figs 5.13 (a) and (b) respectively.

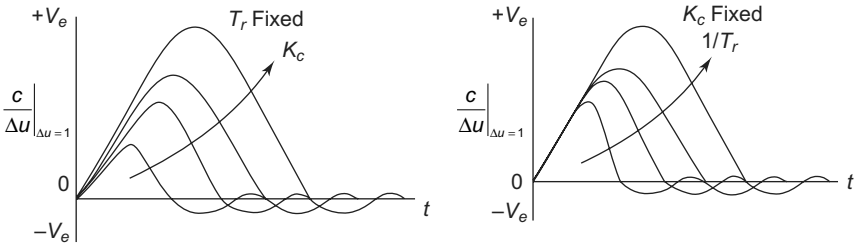


Fig. 5.13 Normalized deviation plotted against time

Likewise for a first order system, a derivative action along with a proportional action may be similarly treated. It must be borne in mind that derivative control action is useful in systems with a large number of storage elements and also when a dead time is present. Derivative action when added to proportional action, increases damping and hence proportional gain may be increased, without impairing the stability of the system, to a large extent.

When a large number of storage elements are present in a process, a still better control is PID. But the analysis of such a system is very complicated as far as the calculation of responses for different K_c , T_r and T_d are concerned. However, experimental trials have shown that whether or not the process contains dead time, addition of derivative action helps to reduce integral action time and increase K_c , simultaneously improving the performance in so far as maximum deviation and settling time are concerned.

Typical deviation-time response characteristics for multicapacity, that is, higher order systems with and without dead time can be obtained with Δu again equal to 1. However, for systems without dead time the response characteristics can be plotted to make a comparative study of *PID* control actions individually or in combination. Derivative action is not used alone although proportional and integral actions can be so used. The three counts on which they are compared, are: (i) maximum deviation, (ii) offset and (iii) stabilization (settling) time.

- (a) Proportional control, when used alone, produces a sufficiently large deviation and some offset, but has a comparatively low settling time, Fig 5.14(a).
- (b) Integral action, when used alone, produces a very large maximum deviation, no offset and a very long settling time, Fig. 5.14(b).
- (c) Proportional plus integral actions, when used at a time, produce maximum deviation that is quite large, larger than that produced when proportional action is used alone but smaller than that produced when integral action is used alone. Here offset is zero but settling time is large lying between the cases of (a) and (b) above, Fig. 5.14(c).
- (d) Proportional and derivative actions, when used together, produce the lowest maximum deviation. Offset is also not large, nearly half the value that occurs when proportional action is used alone and settling time is also the smallest, Fig. 5.14(d).
- (e) Proportional, integral and derivative actions, when used jointly, produce maximum deviation larger than that of case (d) only and no offset. But there is a substantial increase in the settling time, in fact, it is larger than those of cases (a) and (d) but smaller than those of (b) and (c), Fig. 5.14(d).

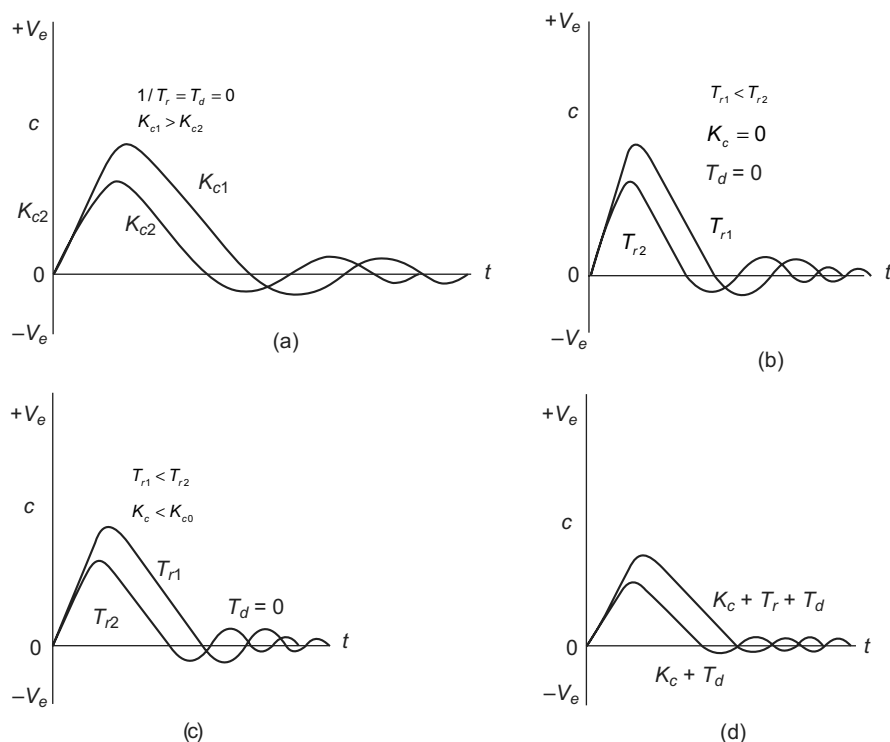


Fig. 5.14 (a) $c - t$ for different K_c (b) $c - t$ for different T_r , $K_c = 0$, $T_d = 0$, (c) $c - t$ for different T_r , K_c fixed, (d) With and without T_r for PID control action.

5.6 CONTROLLER TUNING OR CONTROLLER PARAMETER ADJUSTMENT

When the process characteristics are known approximately, the value at which K_c , T_r and T_d are to be adjusted are determinable. Final setting of these parameters will always be made by making a compromise between the steady state and dynamic performances. Different methods and schemes are available for setting—their comparison will yield an optimum method of doing this job.

There are three major approaches for adjusting these parameters. These use: (i) the stability limit of the control scheme, i.e., the gain bandwidth product or gain and bandwidth individually. This is also known as the ultimate method because it uses the ultimate values of parameters—gain and frequency. (ii) the process reaction curve, i.e., the transient response curve with a step input to the open loop process, and, (iii) the frequency response of the process.

(i) The closed loop method using the stability limit starts by putting off the integral and derivative actions of the controller and increasing its proportional gain till the system begins to have stable sinusoidal oscillations. Let this oscillation period be $T_0 (= 2\pi/\omega_0)$ and the gain often called the ultimate gain, be K_{c0} . The loop performance quality is then determined by the gain bandwidth product $2\pi K_{c0}/T_0 = K_{c0}\omega_0$. Both K_{c0} and T_0 are practically obtainable, and the parameter values for different control actions are now derived in terms of K_{c0} and T_0 . It is interesting to note that although K_{c0} and T_0 were initially considered from empirical tests for maximising their ratio, these can be correlated with the gain margin and phase margin of the frequency response studies, i.e., Bode-plots. Thus the rationale of choosing these quantities for adjusting controller parameters also becomes clear. Ziegler and Nichols were the first to give the adjustment rule in terms of K_{c0} and T_0 as quoted in Table 5.2.

For proportional action only, obviously, a gain margin of 2 is considered. As integral action is introduced a larger phase lag appears. Hence, from Bode-plot one knows that to have the same gain margin, K_c is to be reduced. The addition of derivative action, in contrast, induces phase lead such that a larger gain of $0.6K_{c0}$ may be accommodated. Choice of T_r and T_d (as also

Table 5.2 Parameter adjustment via gain-bandwidth

ACTION	K_c	T_r	T_d
<i>P</i>	$0.5 K_{c0}$	-	-
<i>PI</i>	$0.45 K_{c0}$	$0.825 T_0$	-
<i>PD</i>	$0.6 K_{c0}$	-	$0.125 T_0$
<i>PID</i>	$0.6 K_{c0}$	$0.5 T_0$	$0.125 T_0$

of K_c) is empirical, as already mentioned, although from considerations of phase margin the values of T_r and T_d suggested may be justified for *PI* and *PD* actions respectively. Also, T_d is adjusted assuming ideal derivative action and is acceptable if B is greater than 20 (Eqs (5.27) and (5.28)). For $B \leq 5$, modification of this parameter is necessary, e.g., *PD* action alone is given and T_d is adjusted till the critical frequency ω_0 is maximised. K_c and T_r are then chosen as for *PI* control only.

Harriott modified the procedure by considering the damped oscillation and not the sustained one. If T_δ is the period of the damped oscillation then for a 1/4 decay ratio with proportional control only, he suggested $T_d = T_\delta/6$ and $T_r = T_\delta/1.5$ for reset and rate time respectively and with these settings the proportional gain is to be further adjusted to get 1/4 decay ratio. The method is purely a trial and error method and in the process might affect other loops, not to mention the long time it takes for adjustment.

A procedure based on Bode-plot is quite common now. The Bode-plot of the open loop transfer function (without the controller) is made from which the frequency corresponding to -180° phase is found out. Using this frequency the amplitude ratio of the cascaded blocks is determined. If this is given as Y dB, then A , given as $20 \log_{10} A = Y$, is calculated and the value of K_{c0} , the ultimate gain, is $K_{c0} = 1/A$. For adjustment of K_c Table 5.2 would be used while for calculating T_r and T_d using the above table, the phase crossover frequency ω_0 , is considered so that $T_0 = 2\pi/\omega_0 = 2\pi/\omega|_{\phi=-180^\circ}$. The procedure is illustrated a little more elaborately in the next chapter.

(ii) The method using the process reaction curve is an open loop method and assumes that the process has a transfer function

$$G_p(s) = \frac{K_p \exp(-s\tau_d)}{(s\tau_s + 1)} \quad (5.33)$$

such that for a step input to the process a response is obtained as given in Fig. 5.11. For the given slope S in the linear region (at the point of inflexion) of the curve, the controller parameters are approximately given by

$$K_c = \tau_s/\tau_d = K_p/(S\tau_d) \text{ for proportional action only,}$$

$$K_c = 0.9\tau_s/\tau_d = 0.9 K_p/(S\tau_d)$$

$$T_r = 3.33\tau_d \text{ for proportional and integral action,}$$

$$K_c = 1.2 \tau_s/\tau_d = 1.2 K_p/(S\tau_d)$$

$$T_d = 0.5 \tau_d \text{ for proportional and derivative action and}$$

$$K_c = 1.2 \tau_s/\tau_d = 1.2 K_p/(S\tau_d)$$

$$T_i = 2 \tau_d$$

$$T_d = 0.5 \tau_d \text{ for proportional, integral and derivative action.}$$

These values are also proposed by Ziegler and Nichols. Interestingly, T_r and T_d of Table 5.2 are related to those of the method using process reaction curve by the equation $T_0 = 4\tau_d$. This correlation is not coincidental. The period of sustained oscillation as obtained from the stability limit in a system having a dead time is actually approximately given by the above equation.

Using the process reaction curve Cohen and Coon suggested a scheme based on controllability ratio, r , and process sensitivity K_{ps} defined as the incremental output Δc obtained from the process for an incremental input Δm from the controller, i.e., $K_{ps} = \Delta c/\Delta m$. Their suggested parameters are tabulated in Table 5.3.

Table 5.3 Values of K_c , T_r and T_d suggested by Cohen-Coon

ACTION	K_c	T_r	T_d
P	$(1/r + 0.33)/K_{ps}$	-	-
PI	$0.9(1/r + 0.092)/K_{ps}$	$3.33\tau_s \left(\frac{r + 0.1r^2}{1 + 2.2r} \right)$	-
PD	$1.35(1/r + 0.2)/K_{ps}$	-	$0.37\tau_s r / (1 + 0.2r)$
PID	$1.35(1/r + 0.2)/K_{ps}$	$2.5\tau_s \left(\frac{r + 0.2r^2}{1 + 0.6r} \right)$	$0.37\tau_s r / (1 + 0.2r)$

However using the above process transfer function and considering the 4:1 subsidence ratio and desirable features such as minimum error integral, negligible offset, etc., values of K_c , T_r and T_d for the different cases can be theoretically derived. Their values are then more rigorous.

With the advent of digital computers tuning of the controller parameters for the constraint of 1/4 decay ratio has become more convenient. Since there are three parameters, two more constraints would be needed for a unique solution. Integral square error criterion is one such constraint. Minimum offset is another as mentioned in the previous paragraph. But based on the suggestion of Cohen and Coon the third constraint used is given as $K_p K_{ps} \tau_d / \tau_s = 0.5$. With these constraints the tuning relations are obtained as given in Table 5.4. Because three constraints are involved in this tuning technique this method is also known as 3C method and is used in computer control schemes.

(iii) The method using the frequency response of the process may use the gain and phase margins or M and N circles directly. In process control systems the technique used for parameter adjustment, however, are through set criteria. One technique suggests the selection of (i) a given subsidence ratio (say 4) by controlling K_c ($T_r \rightarrow \infty$, $T_d \rightarrow 0$), (ii) the integral action time equal to the period of oscillation and (iii) the derivative action time fixed to have maximum controller gain with (i) and (ii) satisfied.

Table 5.4 Control parameters K_c , T_r and T_d by 3C method

ACTION	K_c	T_r	T_d
P	$1.208r^{-0.956}/K_{ps}$	-	-
PI	$0.928r^{-0.946}/K_{ps}$	$0.928\tau_s r^{0.583}$	-
PD	$1.37r^{-0.950}/K_{ps}$	-	$0.365\tau_s r^{0.950}$
PID	$1.37r^{-0.950}/K_{ps}$	$0.740\tau_s r^{0.738}$	$0.365\tau_s r^{0.950}$

As the damped sinusoid becomes the key factor since subsidence ratio of 4 is the starting point, a modification of the Bode diagram for the damped

sinusoid becomes essential. One of the frequency response techniques is thus adopted graphically following a modification of the Bode diagram as stated. The approximate relations between the damped and undamped waves are derived from the transformation:

$$j\omega_\delta = -\sigma + j\omega_0 = \lambda \exp(j\rho) \quad (5.34a)$$

where, suffix δ stands for damped condition. The relations are

$$\ln|G_p(j\omega_\delta)| - \ln|G_p(j\omega)| = -\rho[d\{\arg G_p(j\omega)\}/d(\ln\omega)] \quad (5.34b)$$

and

$$\arg G_p(j\omega_\delta) - \arg G_p(j\omega) = \rho[d\{\ln|G_p(j\omega)|\}/d(\ln\omega)] \quad (5.34c)$$

where $\arg$ means ‘polar angle of’.

The first equation has a right hand side given as $(-\rho \times \text{slope of the phase curve})$ and the right hand side of the second equation is $(\rho \times \text{slope of the gain curve})$. The slopes are, in general, negative such that the damped sinusoid has a higher gain and larger phase lag than the original sinusoid. The results are used in steps as follows. The Bode-plot is obtained for the system. Slopes of the phase and gain curves are known from these plots as also the critical or cross over frequency. Let this be ω_0 . When damping occurs, the damped frequency is given by

$$\omega_\delta = \omega_0 \sqrt{1 - \zeta^2}$$

For a subsidence ratio of 4, $\zeta = 0.22$. Using this ζ , ω_δ is calculated and using $|\omega_0| = \sqrt{\omega_\delta^2 + \sigma^2}$, σ is also evaluated. The parameter ρ is then found out as $\rho = \tan^{-1}(\omega_0/\sigma)$. The slopes obtained above are multiplied by ρ and the gain and phase curves are modified. The modified gain and phase curves are now used to obtain the controller parameters using (ii) and (iii) above.

Of the three methods suggested above, the stability limit method is probably the best as it takes less time and is quite suitable for normal conditions.

A very practical procedure but not necessarily less time consuming, is to adjust the parameters as indicated below. The procedure assumes no prior knowledge of the process.

- (i) T_r is set at maximum and T_d at zero,
- (ii) K_c is then gradually increased from zero till oscillation appears on the recorder with less than 4 mm peak to peak amplitude,
- (iii) T_r is gradually decreased till offset is eliminated and a low frequency low amplitude cycling appears about the set point.
- (iv) T_d is now increased step by step until the oscillation stops. By allowing a large value for T_d , K_c can also be allowed to be raised.

A few trials are made now to get the settings to optimum value when low amplitude low time period oscillation about the set point would be considered as the good adjustment.

Whichever of the three methods stated earlier is followed in practice, it should be remembered that disturbances have not been considered while suggesting these settings. Alterations to settings may be made by keeping in mind that one should use (i) for low frequency disturbance more integral action and increased proportional gain and (ii) for high frequency disturbance less integral action and small derivative action.

It has been shown in Chapter 4, Section 4.5, that compensators can help improve the system performance. From the transfer functions of the compensators given in Eqs (4.80) through (4.82), one would understand that *PID* controllers are basically compensators where compensator parameters are tunable. From the prescribed Ziegler–Nichols methods, one can write the transfer function of a *PID* (or *PI* and *PD* as well) controller as

$$G_{c1}(s) = K_c \left(1 + \frac{1}{sT_r} + sT_d \right) = 0.6K_{c0} \left[1 + \frac{2}{sT_0} + 0.125 sT_0 \right] \quad (5.34d)$$

when ultimate tuning rule is considered, and

$$G_{c2}(s) = 1.2 \frac{\tau_s}{\tau_d} \left[1 + \frac{1}{2\tau_d s} + 0.5 \tau_d s \right] \quad (5.34e)$$

when PRC method is followed.

Equations (5.34d) and (5.34e) can further be arranged to give

$$G_{c1}(s) = 0.075K_{c0}T_0 \frac{(s + 4/T_0)^2}{s} \quad (5.34f)$$

and

$$G_{c2}(s) = 0.6\tau_s \left[\frac{(s + 1/\tau_d)^2}{s} \right] \quad (5.34g)$$

Thus both the tuning techniques show that the controller has a pole at the origin and double zeros on the negative real axis. The parameters K_{c0} , τ_s , τ_d and T_0 are not adjustable for a given process, but if the compensator given by Eqs (5.34f) and (5.34g) shows its unacceptability because of, say, large peak deviation, there is scope of fine tuning. However, the Ziegler–Nichols tuning rules give the starting point.

Example 1

For a process transfer function $G(s) = 20/[s(s + 3)(s + 5)]$, design the *PID* controller following Ziegler–Nichols rule.

Solution If integral and derivative actions are zero and proportional controller has a gain K_c , the transfer function of the closed loop, assuming $G_v = G_m = 1$, is

$$T(s) = \frac{20K_c}{s^3 + 8s^2 + 15s + 20K_c}$$

Arranging the coefficients of the characteristic polynomial in the Routh array

$$\begin{array}{c|cc} s^3 & 1 & 15 \\ s^2 & 8 & 20K_c \\ s^1 & \frac{120 - 20K_c}{8} & 0 \\ s^0 & 20K_c & 0 \end{array}$$

From this array first element in the third row (s^1) would be zero for $K_{c0} = 6$ and for finding the ultimate gain subsidiary equation is formed as follows:

$$8s^2 + 20K_c = 0$$

giving $s^2 = -120/8$ so that

$$\pm j\omega_0 = s = \pm j\sqrt{15} \text{ from which}$$

$$\omega_0 = 3.88 \text{ r/s}$$

where ω_0 is the ultimate frequency. Thus ultimate time period is $T_0 = 6.28/3.88 = 1.618 \text{ s}$

Using Ziegler–Nichols tuning rule

$$K_c = 0.6 K_{c0} = 0.6 \times 6 = 3.6$$

$$T_r = 0.5 T_0 = 0.5 \times 1.618 = 0.809 \text{ s}$$

and

$$T_d = 0.125 T_0 = 0.125 \times 1.618 = 0.202 \text{ s}$$

When a transportation lag or delay appears to be large, the controller settings becomes quite complicated as is shown in Tables 5.3 and 5.4. Even there the delay is considered to be reasonably acceptable for controllability as charted out in Table 5.1. One method, for large delays, is to provide an input to the controller such that the process has no time delay. The scheme actually uses time delay and time constant units in the controller. This scheme also permits a higher controller gain and reset rate. Fig. 5.15(a) shows the extra part of the controller G_c . From the scheme one easily derives

$$q = c + f = mG_e + mG_p \exp(-s\tau) = mG_p \quad (5.35a)$$

giving

$$G_e = G_p[1 - \exp(-s\tau)] \quad (5.35b)$$

The scheme is commonly known as the Smith Predictor dead line compensator.

The Smith predictor is used to predict the process output that would be obtained after n times in the step where $n = \tau/\Delta t$, Δt being the sampling time in discretized operation (see Chapter 2). This predicted output is differenced from the actual process output and the resultant used in the controller is usually a PI one. The purpose is thus correcting in effect the set point for a mismatch between the process and the model. Obviously, compensation is a model-based approach. In Fig. 5.15(a), $G_e(s)$ is the transfer function of the model. Use of such a technique is adopted for the compensation, when the dead time is greater than the process lag—in fact, greater than twice the process lag. The model can be derived from the consideration of a disturbance-induced output situation and restructuring of the control architecture is then obtained. One such block diagram with models in Z-transform is given in Fig. 5.15 (b) where $K_l = e^{-\Delta t/\tau_1}$, τ_1 being the time

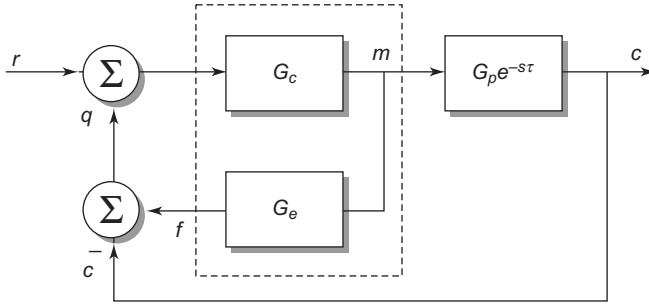


Fig. 5.15 (a) Block diagram of a controller for a process with dead time

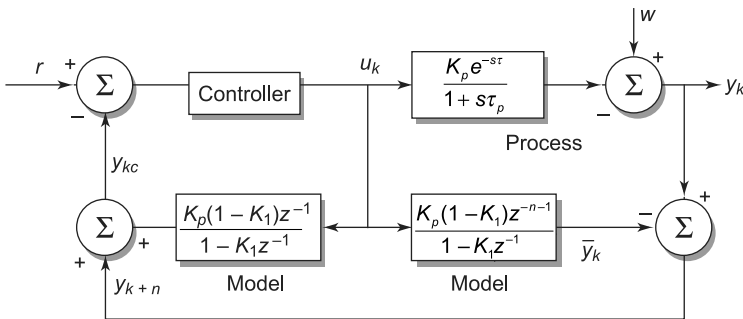


Fig. 5.15 (b) Alternative scheme of Smith predictor-based control

constant of the disturbance 'entry', the compensated process output y_{kc} can be shown to be given by

$$y_{kc} = y_k - [K_c (\bar{y}_{k+n-1} - \bar{y}_{k-1}) + K_p (1 - K_I) (u_{k-1} - u_{k-n-1})]$$

There can be other structures as well. Feed forward control often uses such compensation when $\tau \gg 2\tau_p$.

In many cases, another way of accepting the time delay in the process is by using a Pade' approximation, particularly a first order Pade' delay, and treating this as such in G_p . The expansion is

$$\exp(-s\tau) = \sum_0^n (-1)^k (s\tau)^k / k! \quad (5.36)$$

A few terms from this could also be used but are not easy to handle for frequency response or other studies. The first order Pade' delay is

$$\exp(-s\tau) = \frac{1 - s\tau/2}{1 + s\tau/2} \quad (5.37)$$

Higher order approximations are also acceptable but at the cost of computational disadvantages and complicated component designs.

5.7 PNEUMATIC CONTROLLERS

The basic principle of obtaining proportional action pneumatically is shown in Fig. 5.16. By a small movement of the flapper, the output pressure is obtained as directly proportional to the movement of the flapper. In this case valve V is closed so that the bellows element does not get any

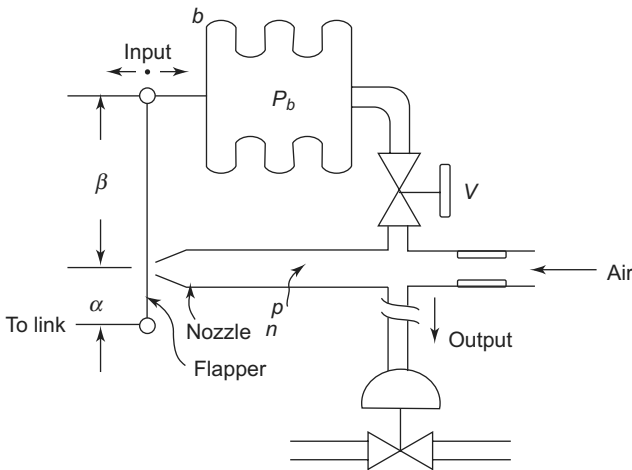


Fig. 5.16 Principle of obtaining proportional action in pneumatic controller

pressure. For wide band proportional action the valve is fully opened. Bellows element b is connected at the free end to the flapper pivot thereby providing the facility for giving extension. As the link moves the flapper to the right, via pressure increase in the bellows element, the pivot shifts to the left and, therefore, less sensitivity occurs (or less gain and hence more proportional band). The decrease in gain will be governed by the ratio α/β and the bellows stroke length per unit pressure. Length β can be changed by changing the position of attachment of the bellows link to the flapper. The feedback mechanism and action is very simple but give an accuracy of the order of 10^{-5} cm lengthwise. Such an arrangement can give a proportional band of the range of 0–600 per cent. However, in practice, 400 per cent is the limit. Usually, the less is the PB without affecting stability and providing the desired recovery for disturbances, the more is the advantage. Proportional band setting is usually done by a differential mechanism without affecting the set value and vice versa. When β is large feedback is small, the resultant output pressure change is large with proportional sensitivity large or PB small and vice versa. The relation can be deduced following the notations used in Fig. 3.18 and in the corresponding text. With error e fed through the link and assuming valve V to be fully open, $p_b = p$, so that

$$k_1 k_n e \beta / (\alpha + \beta) - k_3 k_n p \alpha / (\alpha + \beta) = p \quad (5.38)$$

which on rearrangement gives

$$p(s)/e(s) = [k_1 k_n \beta / (\alpha + \beta)] / [k_3 k_n \alpha / (\alpha + \beta) + 1] \quad (5.39)$$

As k_n is of very large magnitude, of the order of 10^5 , one can write the above transfer function as, with $k_3 k_n \alpha / (\alpha + \beta) \gg 1$,

$$p(s)/e(s) = k_1 \beta / (k_3 \alpha) = K_c = 100/PB \quad (5.40)$$

Figure 5.17 gives the method of generating integral action pneumatically. Bellows element C would oppose b in Fig. 5.16 and connected at ‘INPUT’ point. Thus a sort of positive feedback is operative with the integral action set up.

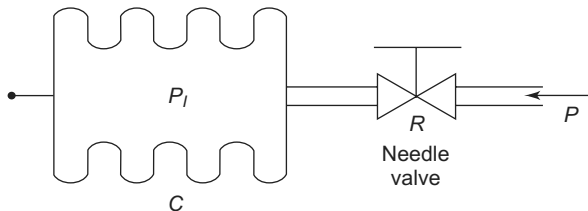


Fig. 5.17 Method of obtaining integral action in pneumatic system

The capacity C in the bellows element and resistance R in the needle valve give an exponential transfer of pressure such that for a step change Δp in the input pressure p , the pressure in the bellows element p_I starts changing proportional to Δp , as shown in Fig. 5.18.

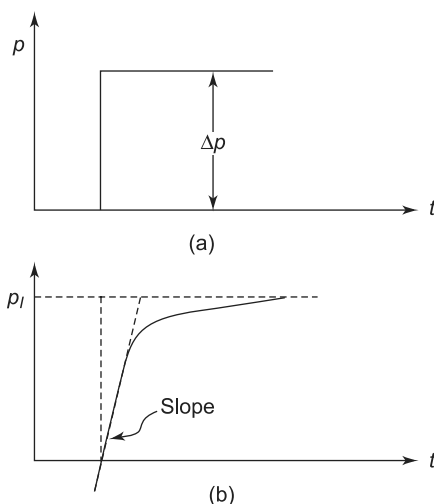


Fig. 5.18 (a) and (b) Explanation of integral action

Initially let p be equal to p_I . Then let there be an increase in p . Through resistance R a flow occurs which is proportional to the difference $p - p_I$ thereby increasing p_I slowly. As p_I increases, $p - p_I$ decreases and flow rate obviously decreases. The flow rate is evidently proportional to the rate of increase of pressure p_I since the resistance R is constant. This gives $\frac{dp_I}{dt} \propto (p - p_I)$ which is the error e and as $p - p_I$ is decreasing $\frac{dp_I}{dt}$ is also decreasing with time giving an exponential nature of the curve as shown in Fig. 5.18. Initially the increase of p_I being very small the rise of p_I with time is almost linear or ideal integration occurs as $p_I \propto \int e \cdot dt = \int (p - p_I) dt$.

A suitable spring in a bellows element can give an extension of the bellows which will be proportional to the change in p_I . But, as noted, the integral action is obtained only initially and subsequently there is a large departure from the actual integral action.

Obviously the PI action can be obtained from the set-up shown in Fig. 5.19. When a step deviation occurs so that the flapper comes closer to the nozzle, the output pressure increases. This is transmitted to the p -action bellows immediately repositioning the flapper but then across R a small flow occurs and p_I slowly increases, its rate of increase being proportional to $p - p_I$, this pressure slowly brings the flapper closer to the nozzle and the output

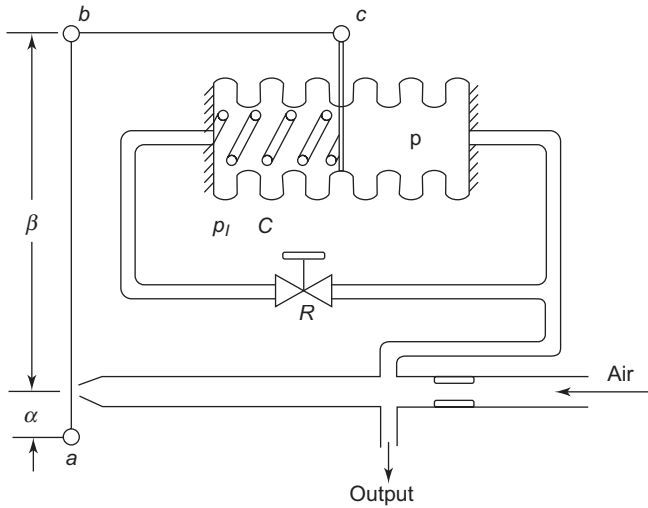


Fig. 5.19 Obtaining proportional and integral action in pneumatic controller

pressure gradually increases. This continues till $p_I \rightarrow p$. Thus as p_I builds up, c moves to the right and the gap between the flapper and nozzle decreases continuously and, therefore, output pressure p gets an additional continuously increasing component which is initially integrating in nature. Here again using the earlier notations, for an error e at point a ,

$$k_1 k_n e \beta / (\alpha + \beta) - (p - p_I) k_3 k_n \alpha / (\alpha + \beta) = p \quad (5.41)$$

Using $p_I = p / (sT_I + 1)$ and $sT_I k_3 K_n \alpha / [(sT_I + 1)(\alpha + \beta)] \gg 1$, Eq. (5.41) simplifies to

$$p(s)/e(s) = (k_1 \beta / (k_3 \alpha)) (1 + 1/sT_I) \quad (5.42)$$

which is the transfer function of a proportional plus integral controller.

The derivative action (along with proportional action) is obtained as shown in Fig. 5.16, when valve V acts as a restrictor and the bellows element as capacitor C . The change in e will produce a change in p but this will not be immediately transmitted to the bellows element because of the restrictor. As the deviation changes, the system sensitivity will increase and the output of the controller will also increase. This output is dependent on the rate of change of deviation. Let the deviation change linearly with time. The flapper now moves slowly closer to the nozzle, giving rise to slow increase in the output. The restriction before the feedback bellows element allows flow of air into this but delays feedback pressure to grow and thus reduces feedback. Thus, if the restriction would not have been there, output would have been less than when it is there. As error is linearly increasing, output pressure, due to this negative feedback would decrease in proportion to the rate of increase of error. With the valve in operation,

therefore, $p_b = p/(sT_b + 1)$ where T_b is the product of valve resistance R and bellows element capacitance C , then Eq. (5.39) changes to

$$k_1 k_n e \beta / (\alpha + \beta) - p_b k_3 k_n \alpha / (\alpha + \beta) = p$$

Putting the value of p_b as above and rearranging one obtains, for the condition $k_3 k_n \alpha / [(\alpha + \beta)(1 + sT_b)] \gg 1$

$$p(s)/e(s) = k_1 \beta (1 + sT_b) / (k_3 \alpha) = k_c (1 + sT_b) \quad (5.43)$$

Thus it becomes a proportional plus derivative controller with rate time T_b .

5.7.1 Three Action Pneumatic Controller

Combining the above three actions a controller is designed as shown in Fig. 5.20. There is a relay with unit pressure ratio. Usually the nozzle has a small diameter, because the force on the flapper from the nozzle output pressure should be negligible as compared to the working force, and air consumption should be less. For servo-operated cases the former restriction is not so binding. The other restriction that the diameters also should be small for linear operation is somewhat difficult to obtain. In fact, these are two to three times smaller. The usual size of the nozzle diameter is 0.5 to 1 mm. Such small diameters require pure clean air to prevent blocking. The main difficulty with these small diameters is slow response. A relay in the proper place compensates for this and provides rapid response and is thus a pressure amplifying device. It may be direct action or reverse-acting, as required.

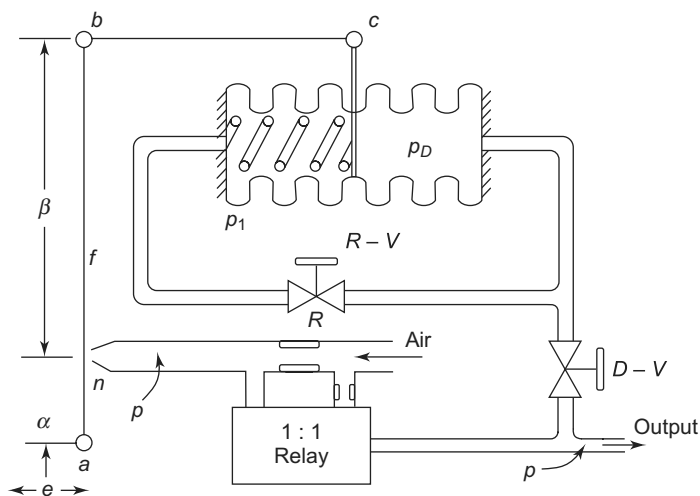


Fig. 5.20 Schematic of a three-term pneumatic controller;
f: flapper, n: nozzle, R-V: reset value D-V: rate value

Initially, let $p = p_I = p_D$, when $e = 0$. When deviation e has occurred, end a will move by $k_1 e$, where k_1 is a constant, a conversion factor of the link-lever system. Pressure p will now change and so will p_I and p_D , thus creating a movement of end b in opposition to a equal to $k_2(p_D - p_I)$, k_2 being given by the spring constant of the spring and the effective area of the bellows. The separation d between the flapper and the nozzle is, therefore, given at time t by

$$d = k_2 \alpha (p_D - p_I) / (\alpha + \beta) - k_1 \beta e / (\alpha + \beta) \quad (5.44)$$

For a gain k_n of the flapper nozzle system, the change in the output pressure is

$$p = -k_n d \quad (5.45)$$

If integral action time constant is T_I and derivative action time constant is T_D , then one easily obtains, from the orientation in the figure,

$$p_D = p / (sT_D + 1), p_I = p_D / (sT_I + 1) = p / [(sT_I + 1)(sT_D + 1)]$$

Using these in Eq.(5.44) and simplifying one obtains

$$k_1 k_n \beta e / (\alpha + \beta) = p [1 + k_2 k_n \alpha s T_I / \{(\alpha + \beta)\}(sT_D + 1)(sT_I + 1)] \quad (5.46)$$

Because k_n is very large, on the right hand side, 1 is negligible compared to the other part inside the bracket, so that,

$$\begin{aligned} p(s)/e(s) &= (k_1 \beta / k_2 \alpha) (1 + 1/sT_I)(1 + sT_D) \\ &= K_c (1 + 1/sT_I)(1 + sT_D) \end{aligned} \quad (5.47)$$

Thus a transfer function of a practical *PID* controller is obtained which is the same as Eq. (5.28) with $B \rightarrow \infty$. This is a design known as the series type, the parallel type being considered earlier in Chapter 3. In Eq. (5.47) an interaction factor $1 + T_D/T_I = \phi$ (Cf. Eq. (3.42)), exists so that it can be written as

$$p(s)/e(s) = K_c \phi (1 + 1/(sT_I \phi) + sT_D/\phi) \quad (5.48)$$

If k_n is not as is considered, the feedback in the controller design considerably complicates the relation. In such a situation controller parameters cannot be adjusted from the set criteria and more trial and error attempts should then be made.

The implementation of Fig. 5.20 is shown in Fig. 5.21 with the details sketched and marked. Another scheme without the link and lever system is shown in Fig. 5.22 in which the torque balance principle is utilized.

feedback actions, negative and positive. A third variation, a variation of that of Fig. 5.22 is shown in Fig. 5.23. Known as the beam balance type controller, it is drawn only schematically.

A controller, particularly the *PID* controller, is always used with an auto-manual-test-service station. The input to the controller is fed through this station instead of directly from the process. Similarly, the controller output is routed through this station to the control valve. A typical scheme is shown in Fig. 5.24. With reference to the connection piping numbers 1, 2,3,4 and 5 and switching positions *A*, *M*, *S* and *T* one can easily see that the connection schemes for the different positions are as given in Table 5.5.

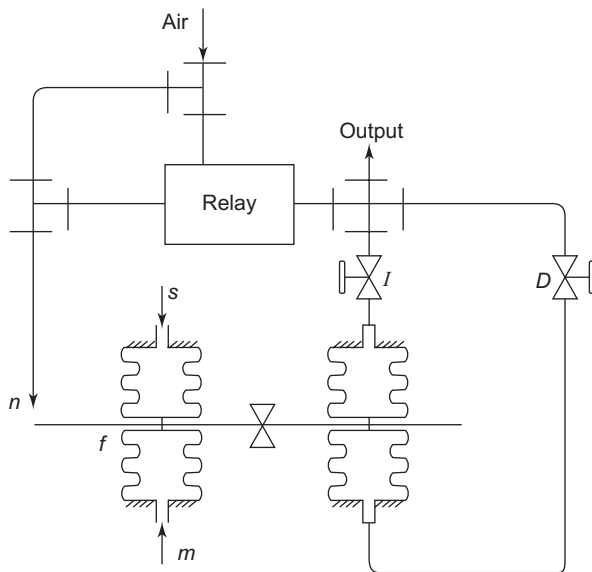


Fig. 5.23 Beam balance type three-term pneumatic controller

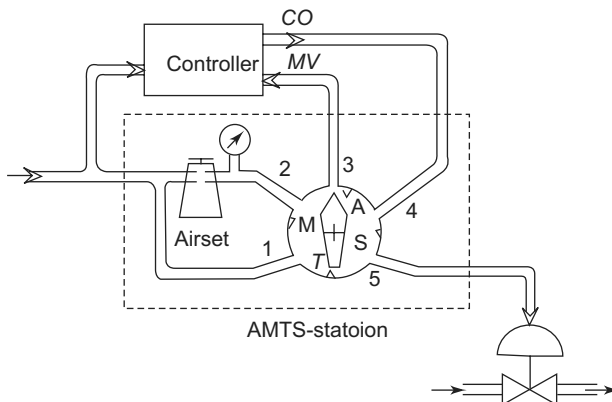


Fig. 5.24 Auto-Manual-Test-Service station

Table 5.5 Connection schemes of the AMTS station

SWITCH POSITION	CONNECTION OCCURS BETWEEN
A	1 and 3, 4 and 5
M	2 and 5, 2 and 4
T	2 and 5, 1 and 3
S	2 and 5

5.8 ELECTRONIC CONTROLLERS

Like the pneumatic controllers electronic controllers can also be of P , PI , PD and PID types. The controller part is made up of an amplifier, an integrator and/or a differentiator as required while the comparator part consists of a difference amplifier or a summer-subtractor. A typical comparator scheme is shown in Fig. 5.25. The error voltage is given in terms of the controlled output voltage and the reference voltage as

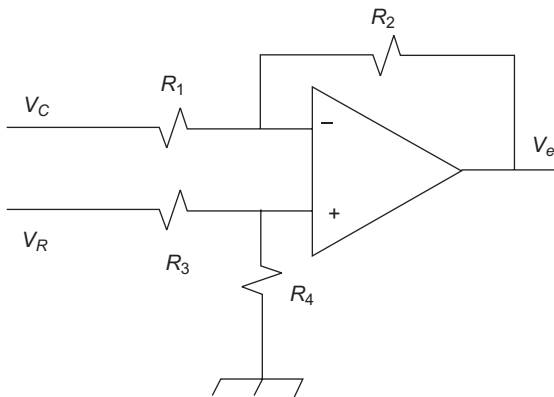


Fig. 5.25 Comparator circuit

$$V_e = V_R - V_c \quad (5.49)$$

for $R_1 = R_2 = R_3 = R_4$.

If, however, R 's are different, one easily derives

$$V_e = (V_R - V_c)R_2/R_1 \quad (5.50)$$

for $1 + R_1/R_2 = 1 + R_3/R_4$. Thus one can make $R_1 = R_3$ and gang equal valued potentiometers R_2 and R_4 to get both comparator and proportional action controller with a single active block, the proportional action gain being given by R_2/R_1 , and V_e becomes the output of the controller.

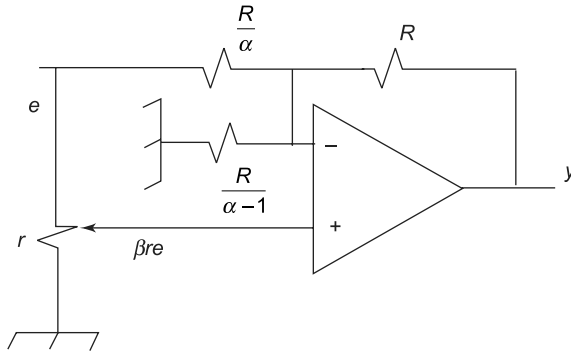


Fig. 5.26 (a) A versatile amplifier used as proportional action controller

A very wide range of proportional band, from negative to positive values, can be obtained using the scheme of Fig. 5.26. The controller output is given as

$$y = (2\beta - 1)\alpha e \quad (5.51a)$$

This gives negative output for $\beta < 0.5$.

Proportional controller with variable gain with more adaptable circuit is shown in Fig. 5.26b. In the figure, if $R_2 = \alpha R_q$ then $R_1 = (1 - \alpha) R_q$ and with simple analysis of the circuit one gets

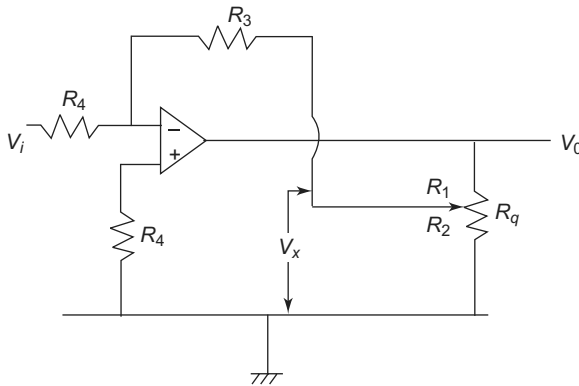


Fig. 5.26 (b) Variable gain proportional gain

$$\frac{V_0}{V_i} = -\frac{R_3}{\alpha R_4} \left[1 + \alpha(1 - \alpha) \frac{R_q}{R_3} \right] \quad (5.51b)$$

When $\alpha = 1$, $R_1 = 0$, the formal gain of the amplifier is $A_0 = -R_3/R_4$. Thus the variable gain factor is $1 + \alpha(1 - \alpha)R_q/R_3 = K_\alpha$. Obviously, α cannot be made 0, thus $1 > \alpha > 0$. For $R_q/R_3 = 1$, $\alpha = 0.5$, $V_0/V_i = -2 R_3/R_4 (1 + 0.25) = -2.5 R_3/R_4$

Integral action alone can be obtained by using a Miller integrator but it gives inverted output. A noninverting integral action is obtained cascading an inverter to this integrator or by bootstrapping a simple passive integrator as shown in Fig. 5.27. Simple analysis gives

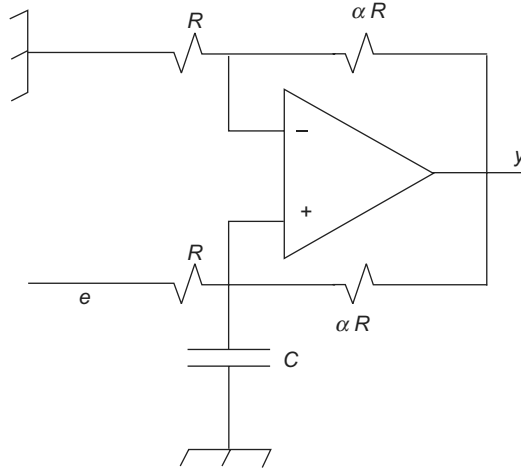


Fig. 5.27 Grounded capacitor integrator

$$y = (\alpha + 1) \int e dt / (RC) \quad (5.52)$$

which in transformed notation becomes

$$y = (\alpha + 1)e / (sCR) \quad (5.53)$$

Tuning of such an integrator is not easy because of 4 number of resistors involved in the process.

Proportional and integral action can be obtained with a single active block as shown in Fig. 5.28. Here

$$(\alpha y - e)sC = e/R \quad (5.54)$$

such that

$$y = e(1 + sCR) / (\alpha sCR)$$

giving

$$y = e(1/\alpha + 1/(\alpha sCR))$$

which in conventional time response form gives

$$y = K_c \left[e + (1/T_R) \int e dt \right] \quad (5.55)$$

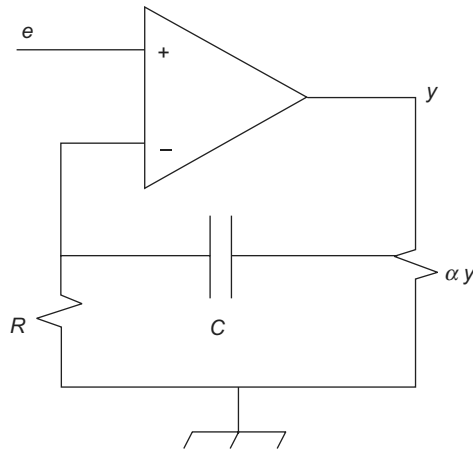


Fig. 5.28 (a) *PI-controller: Electronic type*

where $K_c = 1/\alpha$, and $T_R = CR$.

Thus α provides the proportional band adjustment and T_R , the integration time via R .

Interchanging C and R , the same scheme gives a proportional and derivative action controller.

The circuit of Fig. 5.26 (b) can be extended to get a *PI* controller with adjustable gain K_c and adjustable integrate time T_i . The scheme is shown in Fig. 5.28 (b). If $R_3 = \beta R_p$ and $R_4 = (1 - \beta)R_p$ and α is defined as in Fig. 5.26(b) one can derive the relations

$$K_c = \frac{R_5}{R_6} \cdot \frac{K_\alpha}{\alpha}$$

and
$$T_i = \frac{CR_5 R_p \beta}{R_5 + \beta R_p} K_\alpha$$

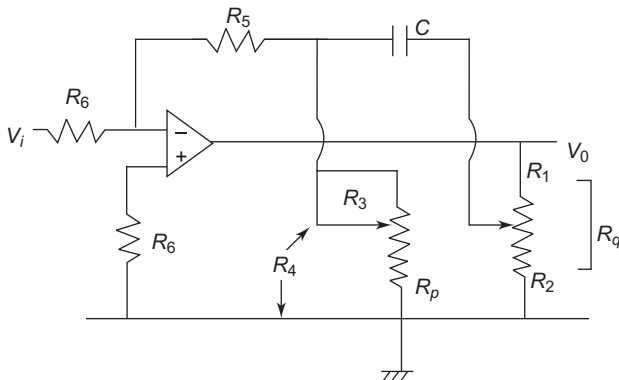


Fig. 5.28 (b) *Controller scheme with adjustable integral action*

Obviously K_c , and T_i are adjustable by R_6 and β respectively. If R_5 is made zero and R_p infinity, the transfer function is

$$\begin{aligned}\frac{V_0(s)}{V_i(s)} &= -\left[\frac{(1-\alpha)R_q}{R_6} + \frac{1}{s\alpha C_3 R_6}\right] \\ &= \frac{(1-\alpha)R_q}{R_6} \left[1 + \frac{1}{s\alpha(1-\alpha)C_3 R_q}\right]\end{aligned}$$

yielding yet a PI controller scheme but not independently controllable K_c or T_i .

The two-input operational amplifier can be used to design a PID controller as shown in Fig. 5.29. In this scheme P denoting a fraction of y and assuming value of the voltage at the junction of C_I and R_I , V , the circuit equations are

$$(Py - V)sC_I = V/R_I + (V - e)/R_D \quad (5.56a)$$

and

$$(V - e)/R_D = esC_D \quad (5.56b)$$

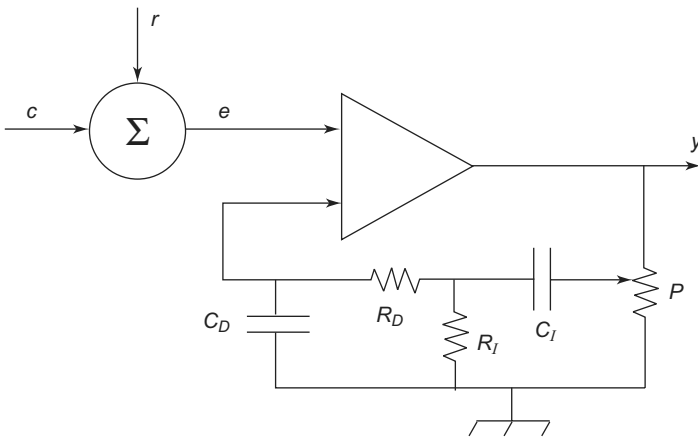


Fig. 5.29 Electronic PID controller using a single operational amplifier

Eliminating V from the above two equations, one easily obtains

$$P y sC_I R_I = e(1 + sC_I R_I + sC_D R_D + sC_D R_I + s^2 C_I C_D R_I R_D)$$

so that

$$\begin{aligned}
y(s)/e(s) &= \frac{1}{P} \left(1 + \frac{C_D}{C_I} + \frac{C_D R_D}{C_I R_I} \right) \\
&\quad \left[1 + \frac{1}{s C_I R_I \left(1 + \frac{C_D}{C_I} + \frac{C_D R_D}{C_I R_I} \right)} + \frac{s C_D R_D}{1 + \frac{C_D}{C_I} + \frac{C_D R_D}{C_I R_I}} \right] \\
&= \frac{1}{P} \left(1 + \frac{C_D}{C_I} \right) \left\{ 1 + \frac{C_D R_D (1 + C_D/C_I)}{C_I R_I \left(1 + \frac{C_D}{C_I} \right) \left(1 + \frac{C_D}{C_I} \right)} \right\} \\
&\quad \left[1 + \frac{1}{s C_I R_I \left(1 + \frac{C_D}{C_I} \right) \left(1 + \frac{C_D R_D (1 + C_D/C_I)}{C_I R_I \left(1 + \frac{C_D}{C_I} \right) \left(1 + \frac{C_D}{C_I} \right)} \right)} \right. \\
&\quad \left. + \frac{s C_D R_D}{\left(1 + \frac{C_D}{C_I} \right) \left(1 + \frac{C_D R_D (1 + C_D/C_I)}{C_I R_I (1 + C_D/C_I) (1 + C_D/C_I)} \right)} \right] \quad (5.57)
\end{aligned}$$

Writing $1 + C_D/C_I = A$, $C_D R_D/A = B$ and $C_I R_I A = C$, Eq. (5.57) is transformed to

$$\frac{y(s)}{e(s)} = \frac{A}{P} \left(1 + \frac{AB}{C} \right) \left[1 + \frac{1}{sC(1 + AB/C)} + \frac{sB}{1 + AB/C} \right] \quad (5.58)$$

Also one would note that $B/(1 + AB/C) = T_d$ and $C(1 + AB/C) = T_R$, the rate and reset times respectively and P is roughly proportional band controlling parameter.

The circuit of a commercial electronic controller using only one operational amplifier is shown in Fig. 5.30. The input current is passed through a fixed resistor R_1 and the voltage is subtracted from a set point voltage to form the input voltage e_1 . The proportional response is changed by stepwise control of αC , the feedback capacitor and the reset rate is controlled by adjusting the reset potentiometer R_f . Derivative action is obtained by providing R_D in the circuit and adding the capacitor $(1 - \alpha)C$. Such that T_d may be independent of gain change. Feedback capacitor C_D prevents excessive variation in controller output for a noisy signal. The diode limiter prevents the output from being controlled within limits, such as between 0.5 and 0.55 mA. The amplifier gain is about 1500, such that e_g is practically zero.

The following circuit equations are now easily derived.

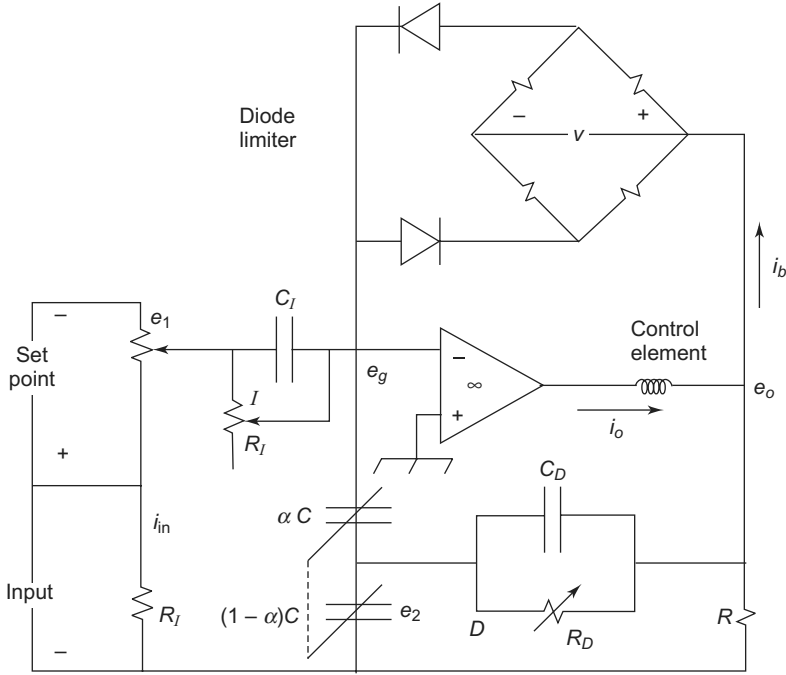


Fig. 5.30 A commercial version of an electronic PID controller

$$sC_I e_1 + e_1/R_I + s\alpha C e_2 = 0 \quad (5.59)$$

$$s(1-\alpha)C e_2 + s\alpha C e_2 + sC_D(e_2 - e_o) + (e_2 - e_o)/R_D = 0 \quad (5.60)$$

and

$$e_o/R = i_o - i_b - s(1-\alpha)C e_2 - s\alpha C e_2 \quad (5.61)$$

i_b is a small current flowing into the limiter biasing circuit. From Eqs (5.59) to (5.61) one easily obtains

$$\begin{aligned} -\frac{i_o - i_b}{e_1/R} &= \frac{C_I}{\alpha C} \left(1 + \frac{1}{sC_I R_I} \right) \frac{s(C + C_D)R_D + 1}{sC_D R_D + 1} \\ &\quad \left(1 + sCR \frac{sC_D R_D + 1}{s(C + C_D)R_D + 1} \right) \end{aligned} \quad (5.62)$$

If CR is very small and $i_b \ll i_o$, one simplifies this to

$$-i_o R/e_1 = K_C(1 + 1/sT_r)(sT_d + 1)/(sT'_d + 1) \quad (5.63)$$

Usually, $C = 20 C_D$, such that $T_d = 20 T'_d$ and this gives a maximum phase lead of 65° and a twenty times maximum gain increase for derivative action.

It must be remembered that as the cost of operational amplifiers are coming down sharply or as complete controllers are now available or may be made available in a single chip by the LSI process more number of operational amplifiers are not a deterrent in the design of electronic controller particularly when this fact enables independent control of the control parameters. This consideration has led many manufacturers to switch to multi-amplifier design.

A two *OA* scheme for capacitor isolation is known to provide a good PID controller with some independent adjustment facility. The scheme is shown in Fig. 5.31. Analysis gives its transfer function as, for $R_d = 0$

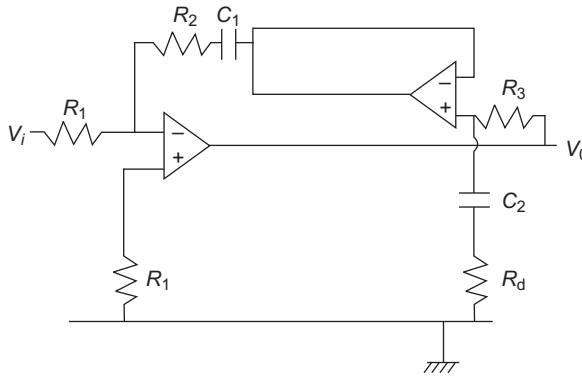


Fig. 5.31 PID controller scheme with independent adjustment facility

$$\begin{aligned} \frac{V_0(s)}{V_i(s)} &= \frac{R_2}{R_1} \left(1 + \frac{1}{sC_1R_2} \right) (1 + sC_2R_3) \\ &= K_c \left(1 + \frac{1}{sT_R} \right) (1 + sT_d) \end{aligned}$$

This makes the gain adjusting resistance as R_1 , T_R adjusting resistance R_2 and R_3 adjusts the T_d . However, R_2 also changes the gain. Hence, for a fixed T_R , K_c is independently adjusted by R_1 , for a fixed gain R_2 and R_1 are to be ganged for adjusting T_R .

Since the output current of the operational amplifier is limited, for the required output current, therefore, a power circuit is required—the simplest scheme of which is shown in Fig. 5.32(a). However, the load resistance is important in this case as well.

However, for a sudden change of the set value or at start up, when it usually takes a long time for the deviation to reach zero, this type of controllers tend to give a large overshoot because a very large deviation persists, maximum output acts for a long time till the set value is reached; after which the respective control actions start functioning normally. As this occurs because of the *I*-action acting too long, it may produce saturation. One way to check this is by shunting the *I*-action during start up. Additionally, the derivative unit is put to work with the input signal only, i.e., differentiated control signal is summed up with the deviation signal. This produces a smoother controller output at start up. The scheme of this is shown in Fig. 5.32(b). Most of the pneumatic controllers do not have this facility.

When the *P*, *I* and *D* controllers work in parallel receiving the same input from the comparator or error amplifier with their outputs summed up for the actuator input, it would not be advisable to make any step change in the set point because the derivative controller would then respond to this step change to give a very high output amounting to saturation. This makes the overall output also to saturate causing what is known as the 'lock-up' of the controller and causing the actuator to open full—a situation often referred to as 'hard-off'. The excess energy received by the process may cause large overshoot and large amplitude ringing before settling. This is loosely known as derivative overrun and is avoided by feeding the derivative controller with the process variable itself, as shown in Fig. 5.32(c). The error voltage is now

$$V_e = V_{sp} - V_{pv}$$

and the outputs from the *PID* controllers are

$$V_p = -K_p V_e$$

$$V_i = K_I \int V_e dt, \text{ and}$$

$$V_d = -K_d dV_{pv}/dt$$

respectively. For a constant setpoint, however, the last term becomes $K_d dV_e/dt$. The summing amplifier now should take into account the sign change as shown to give V_o as

$$V = V_{sp} - V_{pv}$$

$$V_o = K_p V_e + K_I \int V_e dt + K_d dV_e/dt$$

The zener is used for limiting the output to v_z .

Electronic controllers have so many variations in practice that it is very difficult to give even a brief idea of the types here. Only a very limited number of cases have been considered with analysis for giving a rough idea to the readers.

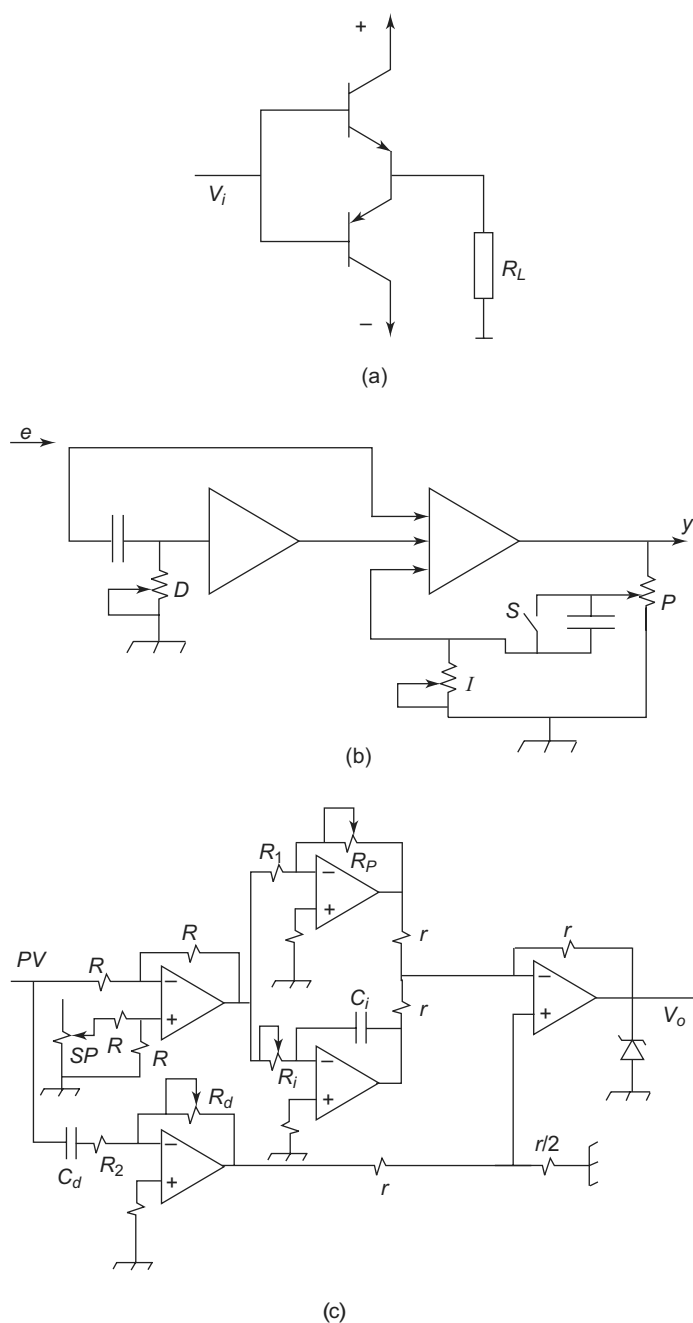


Fig. 5.32 (a) Simple scheme for power enhancement
 (b) An electronic controller for smooth start up; S : switch
 (c) Circuit to protect from derivative overrun

5.8.1 System Response with Controllers Having PID Actions

Earlier in this chapter we have demonstrated how change of proportional, integral and derivative actions individually or collectively change system response. Once the controller parameter adjustment rules are known and the transfer functions of the controller and the process are known, MATLAB can be used to find the response and this procedure would enable to further adjust the parameters.

By way of example, we write the *PID* controller transfer function as

$$Sys1 = K_c \frac{T_i T_d s^2 + T_i s + 1}{T_i s}$$

and the process transfer function as

$$Sys2 = \frac{64}{s^3 + 14s^2 + 56s + 64}$$

For simplicity's sake, measurement and actuator transfer functions are taken unity as has been considered earlier. We consider two sets of parameter values as tabulated below. Obviously, case I is not used in practice specially the T_d value.

Case	K_c	T_i	T_d
I	6.75	1	1
II	6.75	1	0.105

Case I, sys1 is given as $sys1 = (6.75 s^2 + 6.75 s + 6.75)/s$

Case II, sys1 is give as $sys1 = (0.7087 s^2 + 6.75 s + 6.75)/s$

The MATLAB program is written as

```
num 2 = [64]
den 2 = [1 14 56 64]
sys 2 = tf (num 2, den 2)
f1 = feedback (sys 2, 1)
t = 0 : 0.01 : 10
step (f1, t)
hold on;
Kc = input ('Enter the value of Kc')
Ti = input ('Enter the value of Ti')
Td = input ('Enter the value of Td')
num 1 = [Ti* Td* Kc Ti* Kc Kc]
den 1 = [Ti 0]
sys 1 = Tf (num 1, den 1)
sys 12 = Series (sys 1, sys 2)
```

```
sys 5 = feedback (sys 12, 1)
Step (sys 5, t)
hold on;
```

With step input the responses are tabulated and also plotted as shown.

Table 5.6 Response

SYSTEM	CASE I	CASE II
Peak amplitude	1.08	1.22
Peak time	3.26	1.47
Peak overshoot (0%)	8.46	21.9
Setting time	5.33	3.04
Rise time	1.08	0.572
Final	-	1

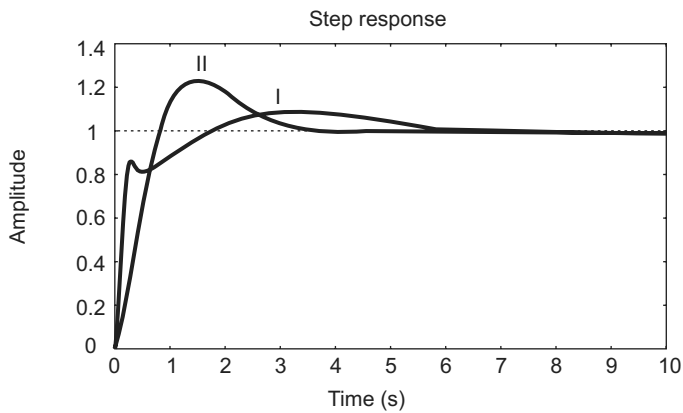


Fig. 5.32 (d) Step responses of the example obtained by Matlab

5.9 HYDRAULIC CONTROLLERS

Hydraulic control is very often used in valve positioning, dumper control, etc. Such systems may employ proportional position only but this often creates offset. Stabilization, is, therefore, obtained by integral action added to it. The schematic of a system employing a nozzle and orifice pair is shown in Fig. 5.33. With the change in pressure p , diaphragm D shifts, moving nozzle N against spring S_2 from zero condition, thereby creating an unbalanced supply of oil to orifices O_1 and O_2 . If N moves to the left, supply to O_1 is more than that to O_2 , the piston in C_2 moves left, applying more pressure to the bottom part of cylinder C_1 and pushing the piston up,

which, in turn, throttles the butterfly valve. Direct proportional action is obtained through the movement of the piston in cylinder C_2 , spring S_1 acting against the given setting; the lever against fulcrum F pushes the spring and thereby the nozzle for attaining the balance. However, unless the needle valve I is used for providing the integral action, proper positioning will not be obtained and offset will result. The action stops when the piston in cylinder C_2 assumes mid-position; this means that the pressure on both sides is equal, I ceases to act and spring S_1 helps to attain this position. Fulcrum F is moved to give more or less proportional gain and I is set for a repeat rate of the reset action.

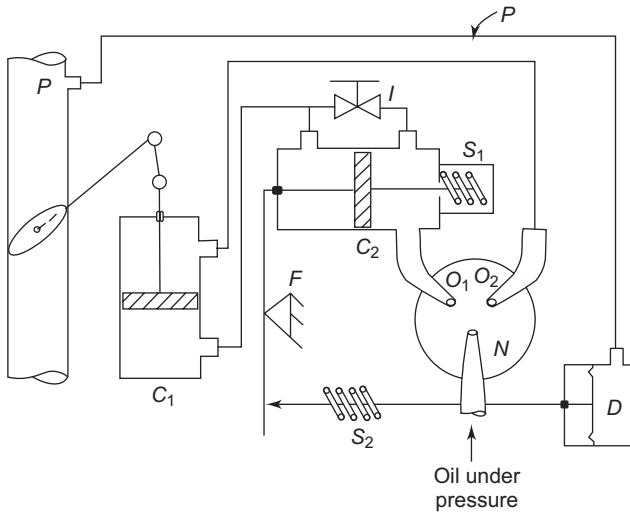


Fig. 5.33 Schematic diagram of a hydraulic controller

5.10 PROGRAMME CONTROLLERS

Programme controllers can be basically on-off, *PID* or any of *P*, *PI* and *PD* controllers, additionally provided with the facility of continuously adjusting the set point. A typical pneumatic scheme of a programme controller with *PID* action can be similar to that of Fig. 5.22 except for the set-point bellows arrangement which is now fed with a pressure adjuster in relation to the programmed quantity as shown by block representation in Fig. 5.34. Programmed quantity is either a pneumatic signal or a mechanical movement in relation to some desired value which is used to work any suitable valve. A typical arrangement is shown in Fig. 5.35. The basic arrangement of obtaining a different pressure for set point (s) adjustment is similar to that of proportional band setting which is usually done manually and in this it is done

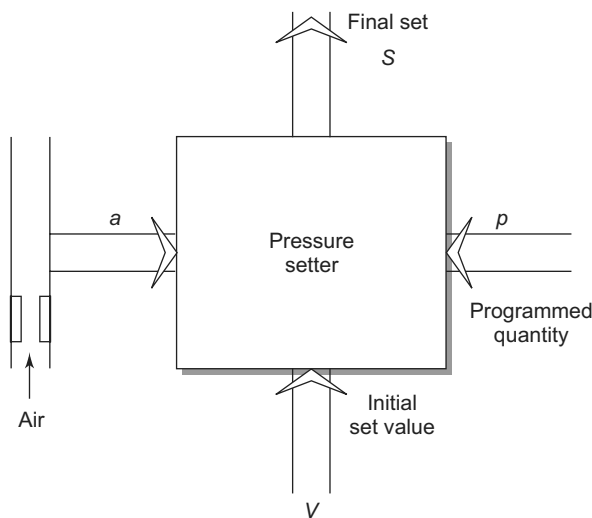


Fig. 5.34 Schematic representation of programmed control

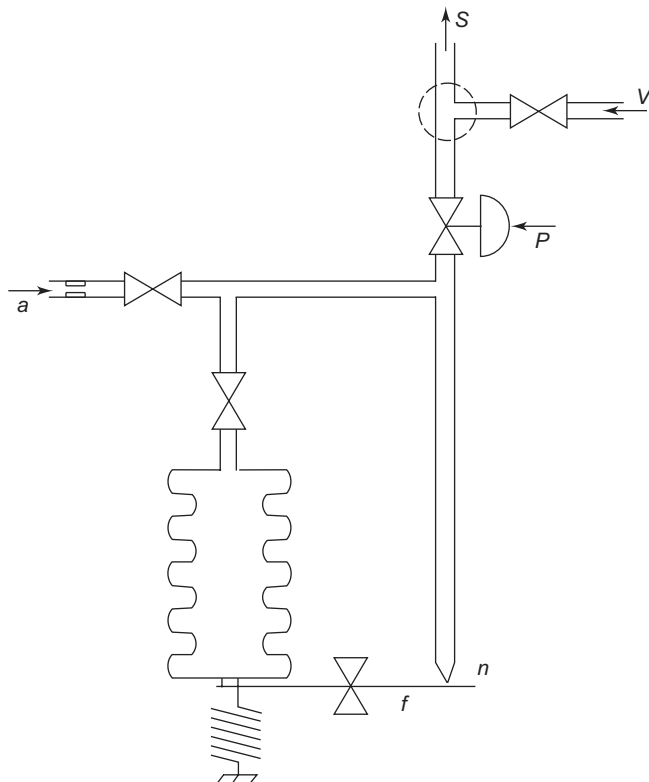


Fig. 5.35 Schematic arrangement of a pneumatic programme controller

automatically by P . Actually, the initial set value arrangement seems to some extent superfluous, but in case the programming action fails, the set value is fixed in accordance with this initial value. Also, the initial set value is 'modulated' by the programmed quantity. The scheme elaborated can provide a minimum value of 0.2 kg/cm^2 as is normally required.

The processes especially suited to programme control are mostly batch types like heat-treating metals, pulp digesting, food cooking, industrial waste treatment, plastic polymerizing, dyeing etc.

The programming of the index point or the set point adjustment is mostly time-bound. However, in exceptional cases time acts as an intermediate variable. Programme controlling can be effected by a few convenient means: by (i) control timers, (ii) cams, (iii) motor-driven index, (iv) curve followers, (v) card/tape programmers, and (vi) computers. These are briefly discussed as follows.

(i) Control timers

These are classified functionally as:

- (a) time switch,
- (b) time delay timer,
- (c) interval timer, and
- (d) time cycle controller.

Their state diagrams are shown in Fig. 5.36 for both direct action type (i) and reverse action type (ii). Electronic timers are now made with 555 timer chips. Realization of timers by electronic means has been discussed elsewhere (Patranabis, *Principles of Industrial Instrumentation*, TMH, 1976).

(ii) Cams

Cam-type programme controllers use programmes in cams to link time to the control of process variables. The cam is cut according to the process schedule or pattern in the controlled variable and the set point index is allowed to follow this. Obviously, in such cases, the rate of rise due to roller binding and fine adjustments pose limitations. Roller binding allows only 50° rotation for full scale variation. This is equivalent to a 70° slope in the developed diagram. For a return to the centre, there is no limitation and, therefore, for a still higher slope a reverse acting cam may be used. A typical programme of heating is shown in Fig. 5.37(a) with its cam design shown in Fig. 5.37(b). Roller diameter radius offset, however, should be compensated while actually constructing the cam. Cam speeds may vary from half an hour to 168 hours per revolution.

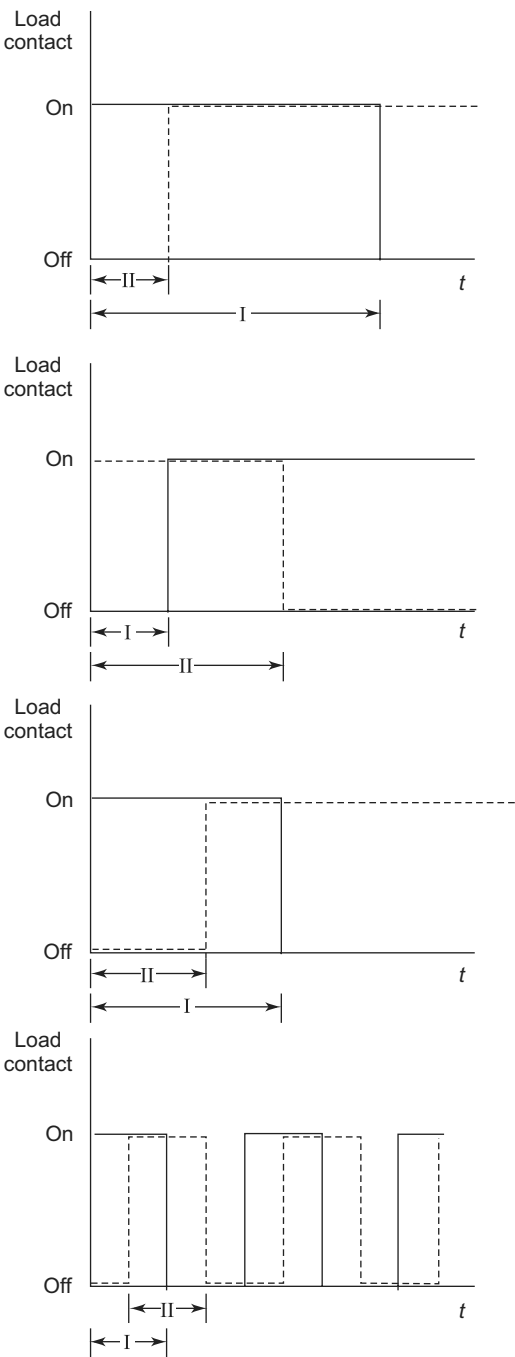


Fig. 5.36 State diagram of timers

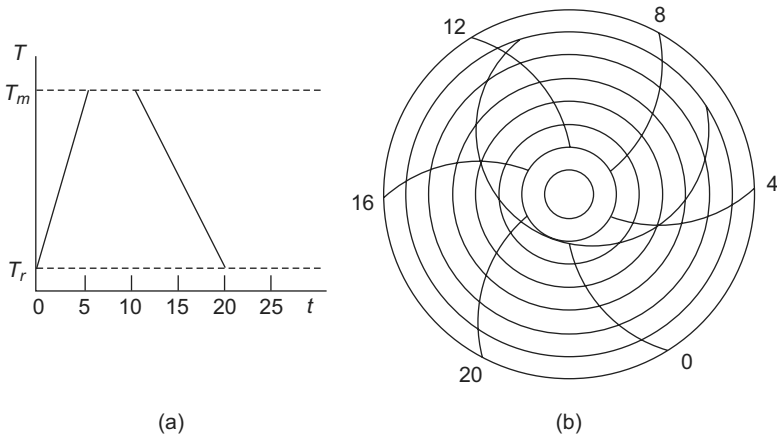


Fig. 5.37 (a) Programme of a typical process (b) Design of the cam

(iii) Motor-driven index

In some programme controllers where a variety of very slightly differing programmes are to be used, the set point index is driven upscale or downscale linearly with time by one or more integrally mounted motors. This system can, however, be made more versatile and complex, using timers and index operated switches.

(iv) Curve followers

Curve follower programme is used in process conditions where both widely varying and fast changing programmes are required. It usually utilizes a strip chart on which the programme pattern is predrawn by special ink or paint that establishes electrical contact with conductors. A slide wire spanning the entire chart is used to pick up the signal which is actually collected by a wiper, the signal itself being used to drive the set point index in a secondary controller.

(v) Card/tape programmers

Tape or card programming is now quite popular. In a tape different tracks pertaining to different programmes are made which are actually energized magnetic materials in response to current signals. Long tapes with a large number of tracks (0.22 million cm length, 20 tracks) provide enormous flexibility. Besides, digital recording of programming in the tapes is also possible. While cards are only for digital programme storage, magnetic tapes can be used both for analogue as well as digital programmes. In contrast to this, punched tapes are usable for the digital case only. These have large capacity and can be conveniently used directly with computers.

(vi) Computers

Even computers can be used as programme controllers. In accordance with a predetermined and preset programme the computer may be used to set the control points in the loops. CD's and or hard discs are used for storing the programmes and programming flexibility is infinite. On-line alteration of the programme is another facility offered here.

5.11 PROGRAMMABLE LOGIC CONTROLLER

Most industrial safety and interlock systems operate on inputs which are binary in nature as provided by elements like push buttons, limit switches, temperature/pressure/flow switches. They also set final control elements that are binary in nature, as for example, the elements like contactors to drive motors, solenoid valves, indicating lamps, etc. The operation of the system is logically stated as—is the limit switch closed? or, is the push button contact closed? Then, start the motor by closing its contactors.

A controller which handles binary inputs and outputs related by logic statements, which are stored in its memory in the form of a programme, is called a programmable logic controller (PLC). The controller checks the status of each individual input contact, performs the necessary logic and sets each of the outputs to the required status by energizing/deenergizing the contactor/solenoid or by giving supply to indicating lamp. Each interlock function is cyclically scanned at a fast rate and the outputs are set/reset depending on the current status or the inputs and/or outputs involved.

Not long ago, electromechanical relays were being used extensively for implementing control logic and interlocking systems for which requirement of the coils and contacts of the relays were specifically configured. A large space, extensive maintenance requirement, inconvenient means of alteration in the system logic were some of the drawbacks of such relay-based interlock system. In course of time solid state switches replaced the electromechanical relays to bring in more reliability and speed. However, system logic alteration and implementation was not made simpler by that. The hardwired system logic has subsequently been replaced by programmed memory in PLC's and as the contents of the memory can be easily altered, easy way of change in the control logic or interlock sequence has been made possible through such a controller.

The basic configuration of a typical PLC is shown in Fig. 5.38 which essentially consists of four distinct modules,

- (a) the input module,
- (b) the logic processor unit module, (LPU, often called CPU),
- (c) the memory module, and
- (d) the output module.

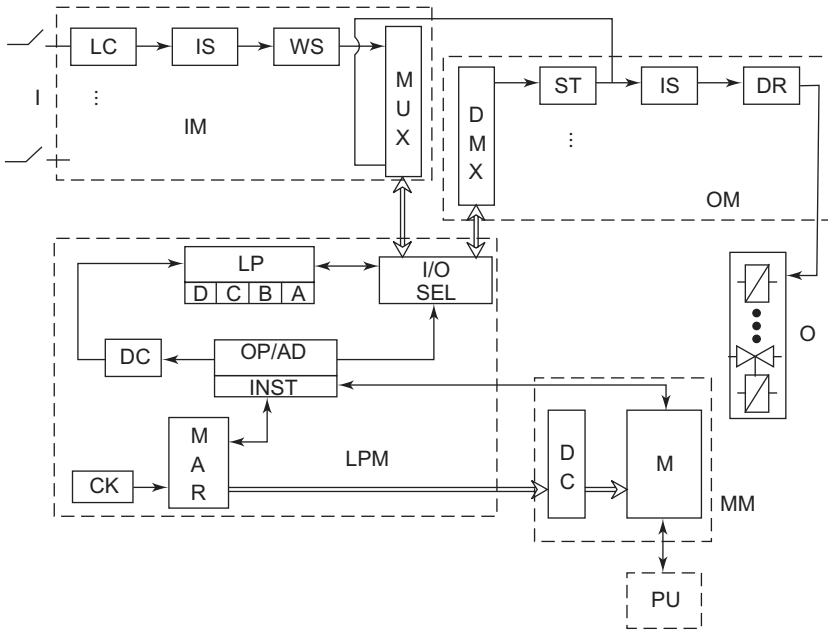


Fig. 5.38 The scheme of a programmable logic controller; IM: input module, OM: output module, LPM: logic processor module MM: memory module, I: inputs, LC: level change, IS: isolator, WS: wave shaper, MUX: multiplexer, DMX: demultiplexer, I/O SEL: input/output selector, ST: storage, DR: driver, LP: logic processor, DC: decoder, OP/AD: operation/address, INST: instruction, MAR: memory address register, M: memory PU: programming unit, O: outputs

5.11.1 PLC Operation

Every logic involving input and output is performed at a high speed via the programme statement stored in the memory. These statements are fetched sequentially as per instruction in the statement and necessary task is carried out through the logic processor. Clock pulses control fetching operation. An operation called memory scanning is the process of sequential execution of each statement. It is carried out cyclically at high speed.

The logic statements are written by the user in a simple, easy to understand and comprehensive programming language. The programme is stored in the memory. Each statement occupies one word of memory. The inputs and outputs are fed through field power supply units at a voltage 24 volts or 48 volts dc. Between the field and internal logic, isolation is provided by optical isolators for improving noise immunity of the system. Also, the internal protective circuitry is protected from the field environment.

The inputs require level changing and waveshaping for making them compatible to TTL/CMOS PLC before feeding to multiplexer unit in the

input module whose line address is the input address. On receiving the control input from the LPU (address part of the instruction) the multiplexer shows the status of the input line. The demultiplexer in the output module routes its input to the appropriate output line control element. The LPU receives a SET command and then directs the address part of the control input of the demultiplexer. There is storage in the output module as the status of the output needs be checked. The driver element actuates the control.

The memory module consists of a EPROM or EAPROM so that the user can reprogramme the memory as often as is necessary. Additions and modifications are also easily done. The memory capacity depends on the required use of the PLC. It may be 16 k bytes or even more.

As shown in the diagram of the LPU clock pulse increments the memory address register to enable sequential instruction fetch. The memory content is received by the instruction fetch where the operation is performed after decoding. The same unit performs to check input status or set output control element in which case the address part is directed to the control inputs of the multiplexer or the demultiplexer. The main processing unit consists of a shift register and the logic circuits which perform the logic operations on the bits of the shift register.

The usual instruction set in a simple case is READ, SET, AND, OR, NOT.

Others are there depending on the 'make' and design such as NOP (no operation), etc. Typically we can assume that instructions are each a byte (8 bits) long in which three bits are used to represent the operation (opcode), the remaining five bits are used as address for input/output. The processing unit performs the operation as follows:

For READ: it reads the status of the input from the input line address and pushes it into the shift register. If the shift register is of 4-bit ABCD type, *A* bit would hold the input status read now, its earlier content being pushed to bit *B*, and so on.

For AND: result of logical AND of bit *A* and bit *B* would be stored in bit *A* position with the content of *C* moving to *B*, that of *D* to *C*.

For OR: operation is logical OR and bit configuration becomes the same as that of AND-operation performance.

For NOT: the complement of bit *A* only is stored back in bit *A*, other bits (*B*, *C*, *D*) are not affected.

For SET: Corresponding to the value of the *A* bit output line address is directed to set the output element. A code is earlier given for this such as bit *A* value '1' may mean closing a solenoid circuit and '0' may mean opening it.

PLC units may provide indications of operations by LED's or monitoring facilities may be there which help to test and debug.

5.11.2 PLC Programming

A PLC can be programmed either by using (i) interlocking ladder diagrams which are basic relay logic diagrams or (ii) Boolean equations—control logic is represented by these equations. The Boolean equations can be obtained from the ladder diagrams straight away. A typical ladder diagram is shown in Fig. 5.39. This may, however, be considered as a single rung of a ladder

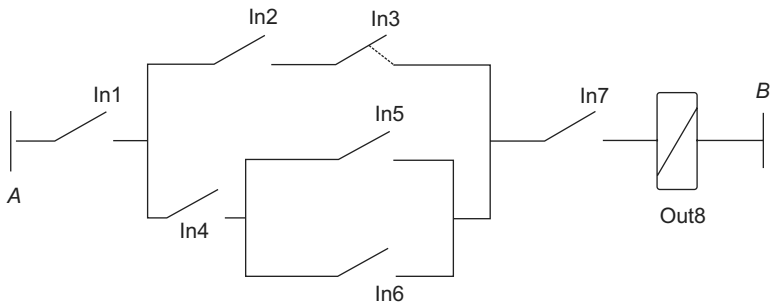


Fig. 5.39 A typical ladder structure

between *A* and *B*. In fact, a complete shopfloor interlock system may be represented by single ladder consisting of several rungs. From this diagram of Fig. 5.39 the Boolean equation can be written as

$$\text{out8} = \{(\text{In5} + \text{In6}) \cdot \text{In4} + (\text{In2} \cdot \text{In3})\} \cdot \text{In1} \cdot \text{In7}$$

For writing the programme corresponding to the above rung of the ladder or the Boolean equation, first, bit assignment for the operations are made as READ → 001, AND → 010, OR → 011, NOT → 100 and SET → 101. The programme is now illustrated in Table 5.6 below. Status of the shift register is easily understood from the comments column which also provides the execution ‘trace’ of the program.

Table 5.7 PLC Program

MEMORY LOCATION	INSTRUCTION OPERATION	I/O ADDRESS	ACTUAL MEMORY MAPPING	COMMENTS
001	READ	005	00100101	bit A = Stat of In5
002	READ	006	00100110	bit B ← bit A, bit A = In6
003	OR		01100000	bit A = bit A (OR)bit B
004	READ	004	00100100	bit B ← bit A, bit A = In4

005	AND		01000000	bit A = bit A (AND) bit B
006	READ	002	001 00010	bit B ← bit A, bit A = In2
007	READ	003	00100011	bit C ← bit B, bit B bit A
				bit A = In 3
008	NOT		100 00000	bit C ← bit B, bit B ← bitA,
				bit A = (Not) bit A
009	AND		010 00000	bit C ← bit B, Bit A = bit A
				(AND) bit B
010	OR		01100000	bit A = bit A (OR) bit B
011	READ	001	00100001	bit B ← bit A, bit A = In1
012	AND		01000000	bit A = bit A (AND) bit B
013	READ	007	00100111	bit B ← bit A, bit = In 7
014	AND		010 00000	bit A = bit A (AND) bit B
015	SET	008	101 01000	OUT8 = bit A

Any change to be done in the ladder because of operation of the plant, can be easily handled by the PLC as only the program has to be changed. For example, let us assume that across In2 another input In9 is to be provided. The steps required for program modification for this change are as follows:

Step 1 Copy the content of EPROM from location 001 to 006 into a RAM,

Step 2 Introduce the following instructions in location 007 to 008 into the RAM

READ 009

OR

Step 3 Copy the contents of EPROM from locations 007 to 015 into the RAM locations 009 to 017, i.e., rest of the program unaltered except location change.

Step 4 Copy the contents of RAM onto the EPROM

Step 5 The system is run as per modified program.

It should be borne in mind that the system allows on-line program modification as well in which case the program brought to RAM would be asked to fetch and execute cycles. During this stage the modification in RAM has been completed, operation is shifted to the normal operating memory, i.e., EPROM. The programming panel provided with the system allows such switch overs.

A word or two about the program entry to the controller must be said here. In such a case the steps are:

- (i) Programming panel is plugged into the controller (PLC);
- (ii) Program is loaded line by line by pressing push buttons;
- (iii) During Step 2, checking is carried out by looking at the CRT monitor provided, which displays the individual circuit, or by looking at the LED displays;
- (iv) The final program is ensured to be alright by making it operative;
- (v) Finally the program is transferred to the EPROM/EAPROM.

The controller is ready to provide the necessary service to the cause.

5.11.3 Further Development

With the overall basic principles, covering the operation, programming, etc., of a simple PLC, mentioned above, it is time to elaborate on the system with its functionalities as it has developed over the last twenty years. The 'power' of the PLC has increased manifold because of the use of bitslice processors when the execution operation also becomes faster. The operating system has also been accordingly updated which provides instructions to command the processor. Instructions used by the controller are different which are used for specified control problems.

Operation of a PLC can be charted as given below in Fig. 5.40.

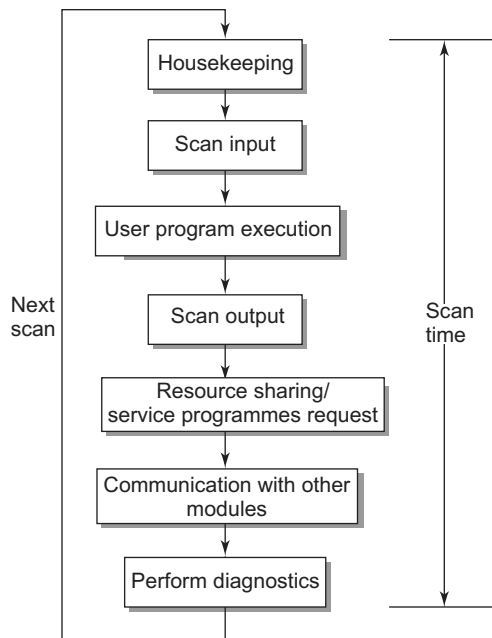


Fig. 5.40 PLC operation charted

Housekeeping is an operation that ensures the health of the system. Reading/scanning the inputs, programme execution, updating/scanning outputs, calling service programs/sharing of resources, interface handling/communication with intelligent modules and diagnostics are all in a scan. Diagnostics also means checking for normal operation, specially power supply and CPU/LPU. A 'watch dog' timer keeps check of the CPU—after every scan, a pulse sent by the processor indicates the normal operation. For nonfunctionality of the CPU, a 'fault output' is indicated. In fact, the housekeeping part generates a reset watchdog timer signal and the watchdog checks the normal execution time or scan time, 10 ms, perhaps. If the scan time exceeds 10 ms, a signal is sent to CPU and the system is shut down.

The CPU has some 'mark' number, higher this number larger is the memory capacity. Speed of the system is defined as the time required to execute 1 k Boolean binary instructions. A PLC is specified also by the numbers of I/P and O/P lines. CPU's are now designed for performing specialized function such as PID control, floating point arithmetic, etc. Recent times I/O modules can be analogue, discrete and special—the latter can accept thermocouples or RTD's directly with a built-in cold junction compensation unit, or can drive a stepper motor directly or can work with a high-speed counter—for this, communication with intelligent modules is necessary.

The system may be made HMI/MMI or HIS (human interface system) compatible by incorporating an extra card. PLC's communicate with the RS 485 port while PC with RS 232C serial link; a PLC is linked with PC's by conversion block RS 485 to RS 232C. Figure 5.41 shows the PLC architecture as it is available. There is always the possibility of an add-on, but provision should be made for that. Figure 5.41 shows expansion possibility with other racks that can be placed nearby or at a distance. For large separation, serial communication is generally used. It may be mentioned that PLC's are often used with redundancy with primary and secondary tags. In cases of emergency 'automatic' switching over to the secondary occurs without bump, i.e. bumpless transfer occurs. Another redundancy principle is the triple modular redundancy (TMR) where modules can be switched on/off or exchanged in cases of emergency or for expanded activities. The scheme is somewhat like the one shown in Fig. 5.42. The dotted connection lines show the switched state on demand.

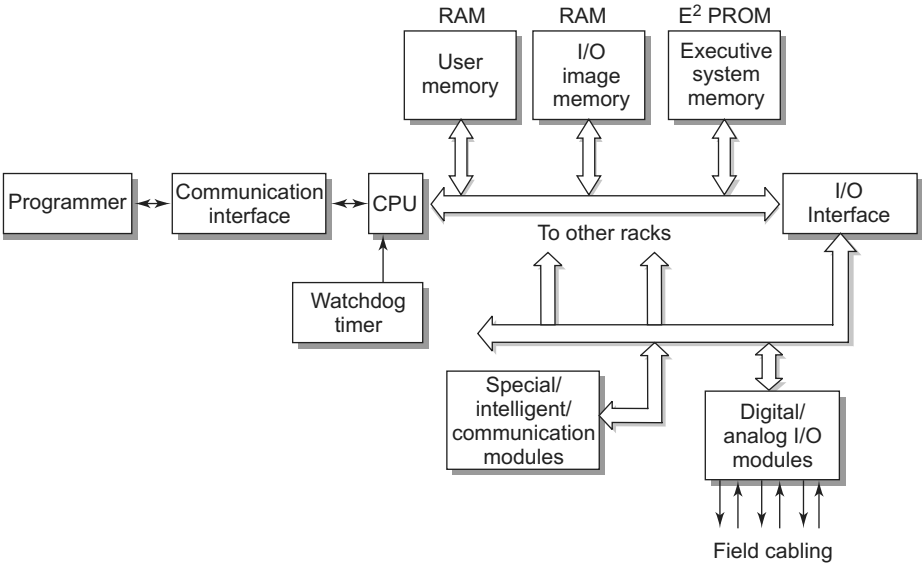


Fig. 5.41 Expansion possibility of PLC with other racks

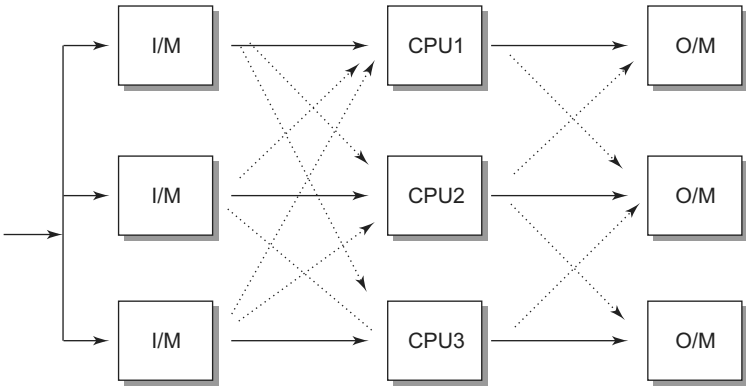


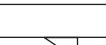
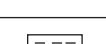
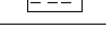
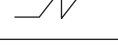
Fig. 5.42 Triple modular redundancy

5.11.4 Devices in Ladder Diagram

A PLC operates, as mentioned already, in logical sequence for which ladder ‘logic’ diagrams are prepared from a real-time system and accordingly a memory module is loaded with an instruction set which are scanned in clock pulse time and the user program is executed to produce the output as needed.

The ladder diagram uses symbols which are, although not fully standardized, representative of the operation functions. Ladder diagrams consist of input devices and output devices. For a simple system, a single rung of a ladder consists of the ‘input’ symbols which logically operate receiving commands/instructions and produce an output which also is symbolically presented in the same rung. However, more complex systems may contain a number of rungs for a single ‘stage’ of operation and scan numbers also increase then. Commonly used symbols are tabulated below:

Table 5.8 PLC symbols

<i>SYMBOL</i>	<i>INTERPRETATION</i>
	Normally open contact
	Normally closed contact
	Normally open push switch
	Normally closed push switch
	Normally open limit switch
	Normally closed limit switch
	Relay coil
	Timer/counter
	Driven unit like motor
	Solenoid
	Pilot light

The list is not exhaustive. As per requirement, newer symbols can be chosen for translating the real-world system into a ladder diagram. Each such symbol used in the ladder diagram is assigned an address which must be unique.

5.11.5 Ladder Commands and Programming

Once the real-world system is translated into a ladder, the vendor-prescribed codes are then used as commands and language to write the program to be executed. There is possibility, however, to simplify the diagram to make it compatible to be programmed by the codes available. For large PLC's, such commands are also many in number for programming and execution which are then divided into a few categories such as

- (1) Basic Instructions,
- (2) Timer-counter commands,
- (3) Comparison commands,
- (4) Data commands,
- (5) Mathematical commands,
- (6) Register sequence commands, and
- (7) Program control instructions.

In the following, we quote a few in each category. It is good to remember these are vendor-dictated and vary from vendor to vendor.

(1) Basic Instructions

<i>Mnemonics</i>	<i>Full Form</i>
OSR	One shot rising
OTE	Output energizer
OTL	Output latch
OTU	Output unlatch
XIC	Examine if closed
XIO	Examine if open

(2) Timer/Counter Commands

<i>Mnemonics</i>	<i>Full Form</i>
RES	Reset
RTO	Retentive timer on
TOF	Time off-delay
TON	Time on-delay
HSC	Highspeed counter
CTD	Count down
CTU	Count up
RTF	Retantive timer off

(3) Comparison Commands

<i>Mnemonics</i>	<i>Full Form</i>
EQU	Equal
GEQ	Greater than equal
GRT	Greater than
LEQ	Less than equal
MEQ	Mask comparison for equal
LIM	Limit test
NEQ	Not equal
LES	Less than

(4) Data Commands

<i>Mnemonics</i>	<i>Full Form</i>
AND	And
COP	Copy file
DEG	Convert from radian to degree
DCD	Decode (4 to 1 of 16)
FRD	Convert from BCD
OR	Or
RAD	Convert from degree to radian
TOD	Convert to BCD
XOR	Exclusive or
NOT	Not
NEG	Negate
MIM	Masked move
MOV	Move

(5) Mathematical Commands

<i>Mnemonics</i>	<i>Full Form</i>	<i>Mnemonics</i>	<i>Full Form</i>
ABS	Absolute	DDV	Double divide
ACS	Arc cosine	LN	Natural log
TAN	Tangent	LOG	Log to base 10
SIN	Sine	SCL	Scale data
COS	Cosine	SCP	Scale with parameter
ADD	Add	MUL	Multiply
ASN	Arc sine	SQR	Square root
ATN	Arc tangent	SUB	Subtract
CLR	Clear	SWP	Swap
CPT	Compute	XPY	X to the power Y
DIV	Divide		

(6) Register Sequence

<i>Mnemonics</i>	<i>Full Form</i>
BSR	Bit shift right
BSL	Bit shift left
SQL	Sequencer load
SQC	Sequence compare
SQO	Sequencer output

(7) Program Control

<i>Mnemonics</i>	<i>Full Form</i>
JMP	Jump
JSR	Jump to subroutine
LBL	Label
MCR	Master control reset
REF	Refresh
RET	Return to subroutine
SBR	Subroutine
SUS	Suspend
IIM	Immediate input with mask
IOM	Immediate output with mask
TND	Temporary end
LD	Load

Again the list is only representative. Besides, for simple cases all such instructions are not necessary. There are some commands which deserve some discussion like Latch (OTL) timer, retention timer RTO, TOF, TON.

Output latch is implemented to keep the output high for a required amount of time. Often in diagram latch circuits are used for this. A very simple case is shown in Fig. 5.43 with S1 closed or ‘start switch’ on, output coil SC is energized or turned on as well, this will energize S2 (SC) and keep it on even after S1 goes off. S2 (SC) would put off with the power cycle occurring, i.e., power is put off.

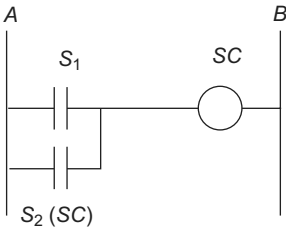


Fig. 5.43 A simple ladder with latch

Timers and counters are part and parcel of logic-based control particularly for providing delays and counting time, etc. Four types of timers are indicated in the tables above. In the ladder diagram, this is represented by a box with the function it performs. The normal function is to compare the current time with the preset time and to change the output at the end of time. On-delay timer (TON) is most common which delays the ‘turn-on’ situation by the preset time. The complementary one, the ‘delay-off’ or ‘off-delay’ (TOF), does the opposite. It turns on with the ‘supply’ on but puts off the line after a preset time. The RTO, retentive timer counts the time base intervals with ‘true’ instruction and also retains the count accumulated when the situation is ‘false’ or the power goes off. That is why it is also called an accumulating timer. With restarting the timer starts from the accumulated time value. Its accumulated ‘count’ requires to be reset and hence this is a timer with two inputs. Its counter part RTF is rare and is not discussed here.

Three types of counters have been listed. HSC, the high speed counter, has already been considered earlier. The others are CTU and CTD, up and down counters respectively. The former counts up from the current value to the preset value while the CTD counts down from the preset value. The two types may be combined to form the up-down counter. In these, HSC is a built-in hardware device while the CTU and CTD may be simulated time-set devices.

A typical timer (TON) put in ladder diagram is represented by the diagram shown in Fig. 5.44.

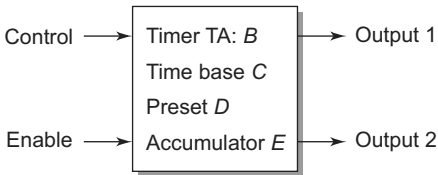


Fig. 5.44 Timber block

TA : B is the timer number, *TA* is the timer-memory indicator and *B* is the location. *C* is the time base and *D* is the preset value while *E* is the starting accumulated value. All these, *A* to *E*, are numerical values. There are two input and two output lines. Control is the selection input while *enable* is the run command.

In the table, a lot of ‘commands’ are indicated, specific vendors would provide the ‘operation manual’ and it is better to consult these manuals. We discuss one or two here which are typical of its kind. The last one in Table (4) MOV, it is used to move data from one location/register to another location/register and is represented in Fig. 5.45(a). It operates with a control signal and at the end output signal is received. Figure 5.45(b)

shows a comparator block which takes EQU, GRT and LES at a time with three outputs with comparison of values stored in two registers A and B when control signal is received. We take the last example here—the sequence compare SQC, which is shown in Fig. 5.45(c). It responds to a control pulse to provide output in the sequence from a sequence table in incremental manner. The sequence, however, has to be preset.

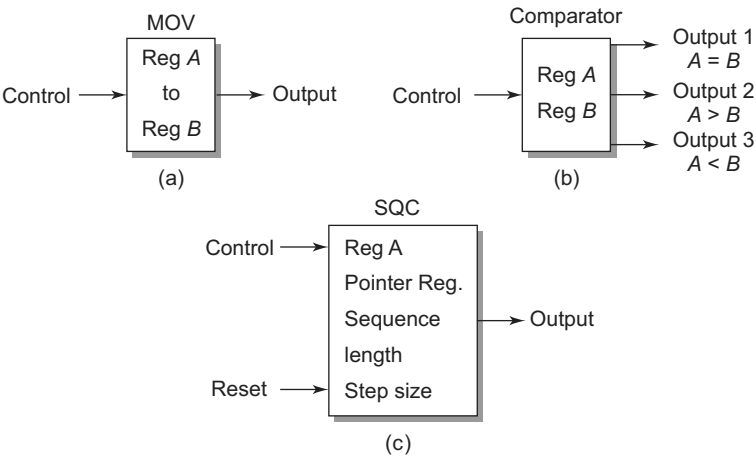


Fig. 5.45 Operator blocks (a) MOV (b) Compartor (c) SQC

Sequence length is the total number of steps contained and step size in bits is also indicated. The pointer register increments steps and content of the table is put at the output.

We are now in a position to demonstrate some simple programming with ladder diagrams derived from practical systems. It may be mentioned that functional blocks shown in Figs 5.45(a) to (c) are used in programming using higher level languages and such a language is block oriented like the one used in computers. Here we consider language using Boolean mnemonics. Here also vendors would provide the mnemonics.

We redraw the ladder of Fig. 5.39 in Fig 5.46 with the symbols as proposed now and write the program alongside.

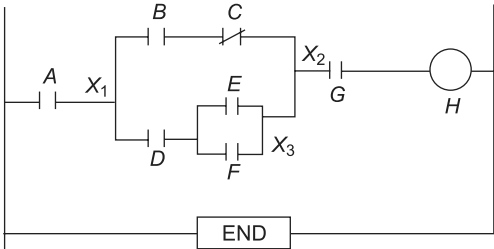


Fig. 5.46 Ladder diagram of Fig. 5.39 repeated

Program in Mnemonics

LD B	Here LD is for load
$\overline{AD}$ C	AD is for Add
	$\overline{AD}$ complement of Add
STO X2	STO is for store result
LD E	
OR F	
AD D	
OR X2	
LD A	
AD X2	
AD G	
STO H	
END	

We conclude with a ladder generation example of a pump which is operable both from control room and from field position. The installation/commissioning diagram is shown in Fig. 5.47(a). Left half shows the control room side and the right side the field/site side installations of control. Figure 5.47(b) shows the ladder diagram.

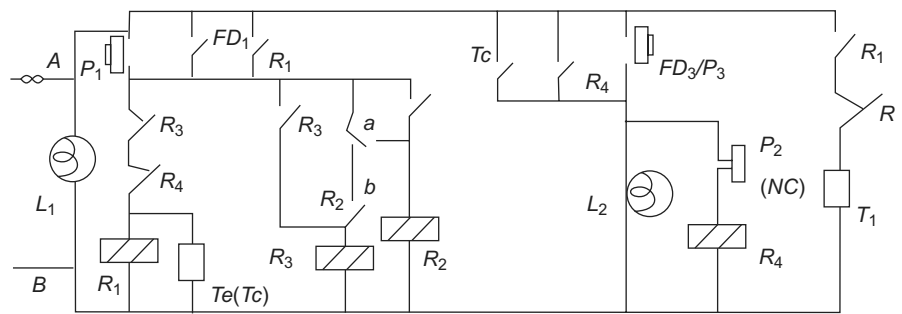


Fig. 5.47 (a) The installation scheme of a pump

The operation sequence is : Power on : P_1 (push button) closed, Relay R_1 energized, Contact R_1 closed, Power through contact R_1 (extreme right) to energize T_1 to put pump on. Timer Te starts simultaneously with preset value after which coil Tc closes contact Tc (right), L_2 glows, relay R_4 is energized, Contact R_4 closed and NC R_4 (left) opens, R_1 deenergized, contact R_1 open, Pump is put off. With FD switch (Field switches) (For emergency stop) : FD2 is as such connected to 'a' when it is connected to 'b' with R_2 closed, R_3 is energized, Contact R_3 is closed and normally closed R_3 is open, R_1 (right) is also open and pump stops.

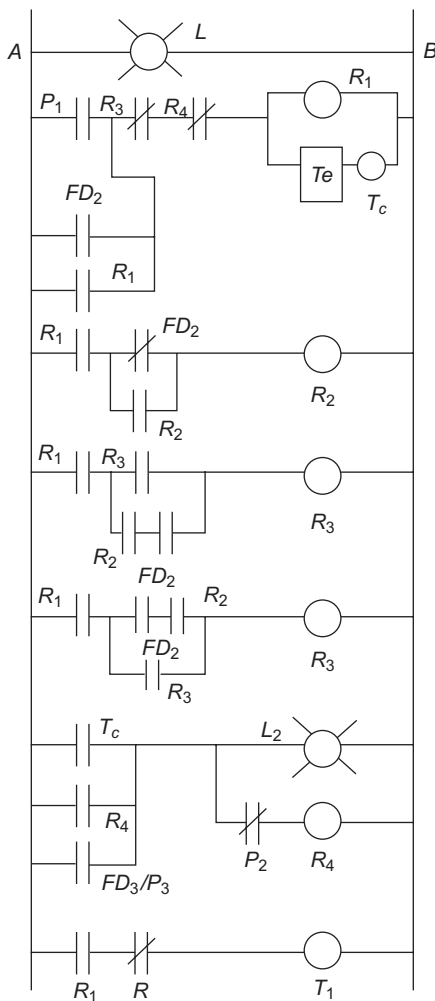


Fig. 5.47 (b) The ladder diagram

Example 2

Draw the scheme of a tank level control with PLC and draw its ladder diagram.

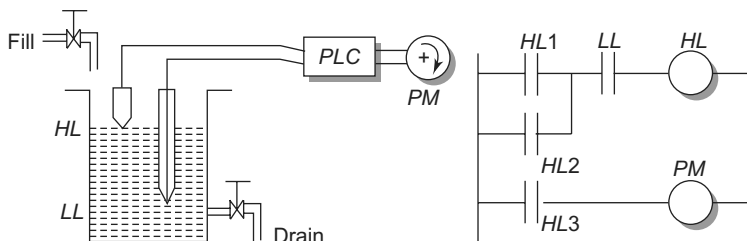


Fig. Ex. I

Solution Both HL and LL are to close to start relay HL which latches HL2 and HL3 to start motor. After some time, with tank liquid rising to LL, LL opens but the filling continues with pump till HL contacts open when the pump puts off.

5.11.6 PAC

PLC has been adapted to take analogue inputs and the letter *L* denoting 'logic' in it becomes superfluous although its operation is still in logic mode. However, the people dealing with devices of this type advocate use of the term PAC denoting Programmable Automation Controller which is a programmable microprocessor-based device and used for discrete automation/manufacturing, process control and also for remote monitoring applications. PAC is thus a hybrid of PLC and PC. Some of its features are listed below:

1. It operates with a single platform but operates on domains like logic, drives, motion, process control and others.
2. Its platform uses a single data base and common tagging to develop tasks in a wide range of disciplines.
3. The hardware and software parts of the system are integrated like embedded systems.
4. It can be programmed with tools that support the system on different units or machines.
5. Architecture is modular and is adapted to various applications from factory automation, project activities to unit operations in chemical process control.
6. It is to an extent an open-vendor system that allows data exchange with the protocols, languages provided with.
7. It can provide scanning and processing functions.

It must be remembered that fast progress in both hardware and software development devices is happening more frequently than we keep count of. However, stabilization of these products in use in industrial situation requires some time.

Example 3

A proportional controller has a sensitivity of $0.10 \text{ kg/cm}^2/^{\circ}\text{C}$; and the output range is from 0.2 to 1.0 kg/cm^2 . With the range of the controlled variable from 180 to 200°C , calculate the proportional band of the controller.

Solution

Using Eq. (5.17)

$$PB = 100 \times 0.8 \text{ kg/cm}^2 / (20^{\circ}\text{C} \times 0.10 \text{ kg/cm}^2/^{\circ}\text{C}) = 40 \text{ per cent}$$

Example 4

The loop transfer function with only proportional action of gain K_c of a control system is given by

$$L(s) = K_c / ((s + 1)(s + 2)(s + 3))$$

Obtain the optimum settings of controller parameters with PID action when no transportation lag is present in the system.

Solution We have $\arg 1/((s + 1)(s + 2)(s + 3)) = -180^\circ$ at $\omega_0 = 3.317$ r/sec

Now, the magnitude of the denominator of $L(s)$ at this frequency is 60. Hence, $K_{c_0} = 60$.

Using Ziegler–Nichols' rule from Table 5.2,

$$K_c = 0.6 \times 0.6 = 36$$

$$T_r = 0.5 \times 2\pi/3.317 = 0.946 \text{ s}$$

$$T_d = 0.125 \times 2\pi/3.317 = 0.236 \text{ s.}$$

Review Questions

- Where an on-off controller is recommended? How is its performance affected by the process dead time?
An idealized first order integrating process has a lag of 3 min and static gain 1.5. The on-off controller used has a dead zone of 2 volts and its operating voltage is 12 volts. Obtain the on-off periods for a step disturbance of equivalent 1v when the system has (i) a dead time of 2 min, (ii) no dead time.
(Ans: (i) 15.27 min, (ii) 2.18 min)
- How can you control on and off timings in a controller as per system requirement? Sketch a specific system and explain the operation.
- How are P and I actions are realized in a pneumatic controller? How are these actions varied in magnitude? Obtain the transfer function of such a controller.
- Sketch a one-amplifier electronic PID controller and show what would be the ranges of control of P , I and D actions in such a controller.
If the ratios of integral action to derivative action capacitances and resistances are 4 and 2 respectively, by what factors integral and derivative actions changed from the face values in the above controller?
(Ans: Derivative and integral action are both decreased by 37.5%)
- What are the different tuning schemes proposed for a PID controller? How have they been evolved?

A proportional controller is used in a process control system which is seen to oscillate for a proportional gain of 60 with an oscillation period of 3 min. If integral and derivative actions are also to be used, what should be the different tuned parameters of the controller?

6. How is process dead time compensated in a controller? What is the practicability of such a scheme?
7. Show by sketch how proportional and reset action rates are to be determined. What is the effect of the derivative action on maximum deviation and stabilizing time?
(Hint: See Figs 5.13 and 5.14)
8. List the types of process and operating situations when to select which mode of control actions. What control modes would you prefer if disturbance is small in magnitude and has low frequency content?
9. A *PID* controller has a steady output pressure of 0.4 kg/cm². The set point is now lowered at the rate of 1 cm/min. Obtain the nature of the output pressure. Assume $K_c = 0.1$ kg/cm²/cm, $T_R = 1.2$ min and $T_D = 0.5$ min.

(Hint: $y = 0.1(-t - (1/1.2) \int t \cdot dt - 0.5) + c'$ at $t = 0$, $y = 0.4$ kg/cm², hence $c' = 0.45$, and $y = -0.1t - t^2/24 + 0.4$; plot it).

10. In Problem 9, if the set point is raised suddenly by 1 cm, what would be the output pressure?
11. Briefly discuss the operational aspect of a programmable logic controller. The ladder diagram of Fig. 5.39 has its input switch In 4 permanently shorted, what would be the Boolean function and the program for the controller?
12. A unit feedback process control system with the process transfer function $0.5/(25s + 1)$ is controlled by a *PID* controller with $K_c = 8$ and varying T_r and T_d . Three different cases are considered: (1) $T_r = 1$ s, $T_d = 0$, (2) $T_r = 0.25$ s, $T_d = 0$, and (3) $T_r = 1$ s $T_d = 0.1$ s. Write the Matlab programme and obtain the transient response characteristics and plot.

[Hint. num 0 = [0.5];
den 0 = [25, 1];
sys 0 = tf (num 0, den 0)
 $K_c = \text{input}('value of Kc: =')$
 $t = \text{input}('value of t : =')$
 $td = \text{input}('value of td =')$
num1 = $K_c * [td \ t \ 1]$;
den 1 = [t 0];
sys 1 = tf (num1, den 1)
sys 01 = series (sys 0, sys 1)
sys 10 = feedback (sys 01, 1)

step (sys 10)
title ('Step Response of given system')

The programs are run with values of K_c , T_r , T_d given and the plots for sys 10, sys 20, sys 30 and transient response data are shown below.]

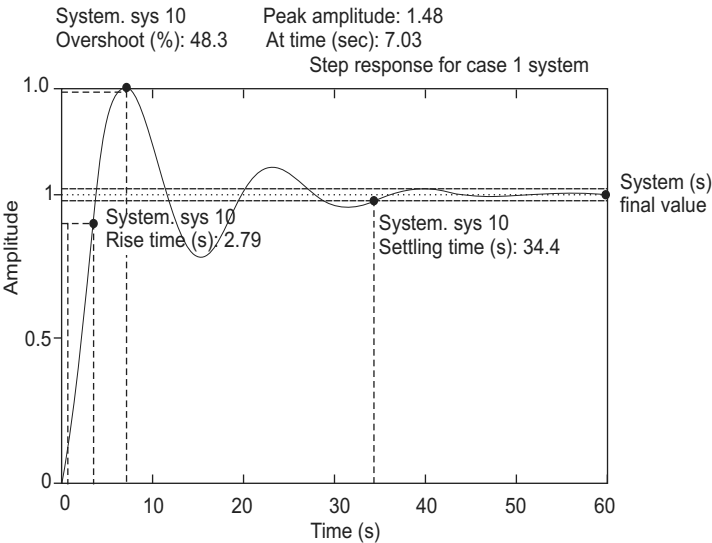


Fig. Q-5.1a

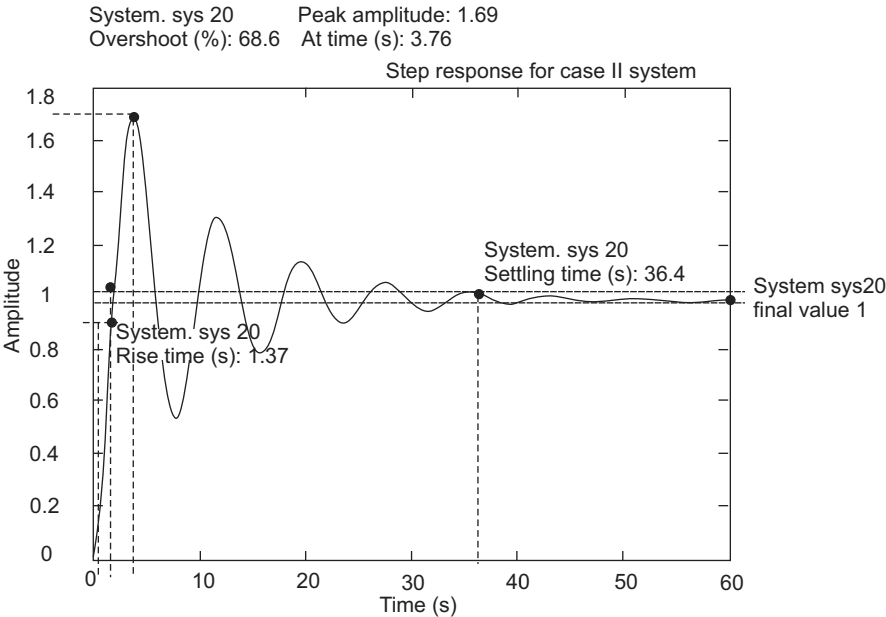


Fig. Q-5.1b

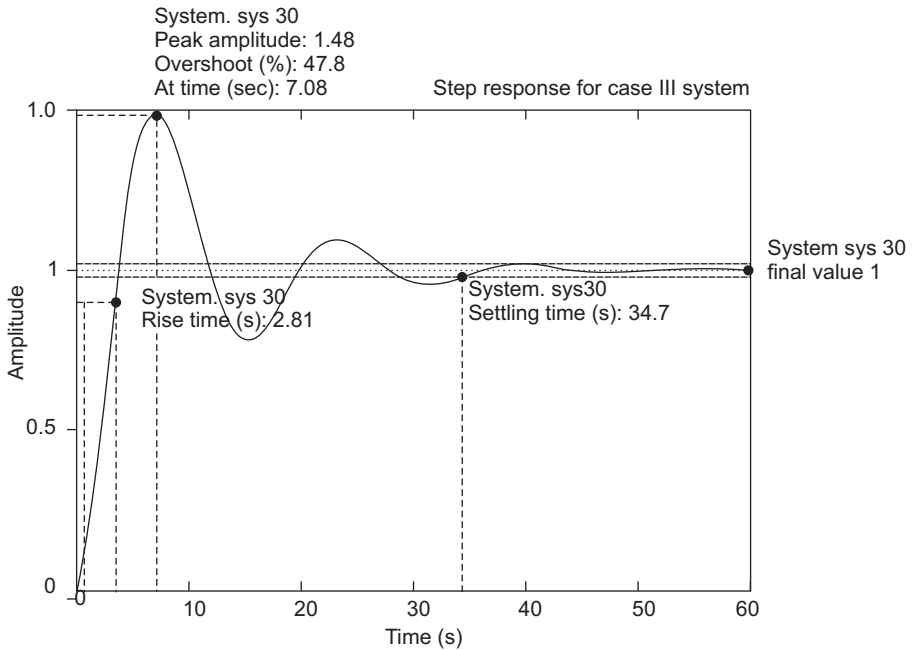


Fig. Q-5.1c

13. Draw the circuit of an on-off control system where the controller dead zone can be adjusted by a simple potentiometer. How are large overshoots in such systems reduced?
14. Chart the operational aspects of a PLC. Draw the set up of a PLC with CPU and interfacing units shown discreetly. Also indicate its expansion possibilities.
15. How do you classify the commands in ladder development and programming?
16. A pump motor has to be started with simple indication of level low in tank and stopped with level full. Before starting, the motor condition is to be checked such as bearing *ok*, and water line *ok*. Draw the system schematic and its ladder diagram.
17. An ideal integrating process of order one has a static gain of 0.4 and a time constant of 3.1 s. The response of the system with on-off controller of peak output of 2.2 units shows a differential gap of ± 1.3 units when the dead zone of the controller is ± 0.22 units,. What is the system dead time and what is its cycling period for unit step disturbance?

[Hint. Refer to Eq. (5.8):

$$K = 0.4, t = 3.1\text{ s}, M = 2.2, \pm \alpha = \pm 1.3, Z = \pm 0.22$$

$$\tau_d = 2\tau(\alpha - z)/KM = \{2 \times 3.1 \times (1.3 - 0.22)\}/(0.4 \times 2.2) = 7.61\text{ s}$$

Again, using Eq. (5.10),

$$T = (2 \times 1.3 \times 3.1 \times 2.2) / \{0.4 \times 1 \times (2.2 - 1)\} = 36.94 \text{ s}$$

18. A controller has a gain sensitivity K_s that sets the proportional band at 85% when the range of the controlled variable is 90 to 110 units for a desired value of 100 units. What should be the value of K_s ?

[Hint. See Eq. (5.17): $PB = 86\%$, $V = 20$

$$K_s = \Delta y_0 / (V \cdot PB) \times 100 = (0.8 \text{ kg/cm}^2 \times 100) / (20 \times 85) = 0.047 \text{ kg/cm}^2/\text{unit}]$$

19. The loop transfer function with P-action only in the controller is $K_c / \{(s + 0.4)(s + 2.2)(s + 3.1)\}$ what should the settings of K_c , T_R , T_d for a three-term controller?

[Hint. From the denominator, $\text{Re} = -5.7$, $\omega^2 = 2.728$, $\text{Im} = j\omega(8.94 - \omega^2)$
 $\tan^{-1} \phi = \omega(8.94 - \omega^2) / (2.728 - 5.7\omega^2)$ from which for $\phi = -180^\circ$, $\omega_0 = 2.99 \text{ r/s}$

Putting the value of ω in the expression for -1 , value $K_{c0} = 48.23$,
 Using table (Ziegler Nichols), $K_c = 0.6 \times 48.23 \approx 29$, $T_r = 0.5 \times 2\pi/2.99 = 1.05\text{s}$. $T_d = 0.125 \times 2\pi/2.99 = 0.2625\text{s}$]

20. A pneumatic *PI* controller has an output pressure of 0.65 kg/cm^2 when the set point and controlled variable (the pen point of the recorder) are together. The set point is suddenly displaced by 1.25 cm to obtain the following time-pressure values (table). Determine the actual gain $\text{kg/cm}^2/\text{cm}$ displacement and the integral time.

TIME (s)	PRESSURE kg/cm^2
0–	0.65
0+	0.53
20	0.47
60	0.33
90	0.23

[Hint. A sudden change in pressure of 0.12 kg/cm^2 for a (step) change of 1.25 cm in set point gives the gain as $K_{c \text{ actual}} = 0.12/1.25 = 0.096 \text{ kg/cm}^2/\text{cm}$. The integral time is obtained from the slope of the curve from 0+ time to 90 s time. Assuming it to be linear, between time 20 s to 60 s, pressure changes by 0.14 kg/cm^2 . Thus, for 40 s time change, K_c changes by 0.14 kg/cm^2 . Hence, $t_I = 40/0.14 \times 0.096 = 27.4 \text{ s}$]

21. Draw the scheme of a conveyor belt transport system where cans are filled for some given time with the belt stopped and then transported till the next can comes after a prescribed time. The belt is operated by a motor which is eventually started and stopped and restarted and so on.

Complex Control Schemes

6.1 INTRODUCTION

Control schemes, not usually made up of a single loop, deserve special attention as they are quite often used in practice for meeting the requirements of the processes. The more important ones in this category are:

- (1) Ratio control,
- (2) Split-range control,
- (3) Cascade control,
- (4) Feedforward control,
- (5) Selector control,
- (6) Inverse derivative control,
- (7) Antireset control, and
- (8) Multivariable control.

Most of these are multiloop schemes and are comparatively complex in nature. They are not easily analysed like single loop cases. Some attempt has been made here to give a brief analysis of the different schemes avoiding undue rigour. More stress is given on the application side.

6.2 RATIO CONTROL SYSTEMS

The ratio control scheme is a very convenient method of controlling variables maintaining a fixed relationship amongst them. For example, it can be used quite satisfactorily in proportioning fluid flows or for ensuring proper mixtures of solids or/and liquids in reactant processes.

The ratio control scheme is, in principle, shown in Fig. 6.1. In this scheme, one derives $c = x_1/x_2$, where x_1 and x_2 are two measured variables of the process. In fact, the ratio control system is only a part of the complete plant control system. The ratio indicates that either one of x_i s can be held uncontrolled while the other is varied as required for optimum control. Figure 6.1 does not, however, do proper justice to a ratio control system. The ratio element inside the loop presents certain difficulties. If x_2 is uncontrolled, then, since $c = x_1/x_2$,



Fig. 6.1 Block diagram of a ratio control scheme; CA: controller and actuator

$$dc/dx_1 = 1/x_2 \quad (6.1)$$

and the problem is not complicated. If, however, x_2 is controlled then, since

$$dc/dx_2 = -x_1/x_2^2 = -c/x_2 \quad (6.2)$$

there is a nonlinear relationship between c and x_2 and this is difficult to settle with linear controllers. Figure 6.2 shows that the ratio element is brought in the set point loop, making a separate loop and setting $r = x_1/K$, x_2 may be made controlled variable. In Fig. 6.2, K , in effect, is x_1/x_2 , the ratio of the two variables. For this reason, the block housing K is often known as the ratio station. The factor K is the ratio setter which usually lies between 0.3 and 3.0, the range being a decade with its midscale at 1. This is because the gain of the ratio element is the square of the control-ratio setting, since, flow squared is the transmitted variable: $q^2 \propto p$; hence $p_1 = Kp_2$ gives $q_1^2 = Kq_2^2$. This means if K varies from 0.3 to 3.0, $\sqrt{K}$ varies from 0.55 to 1.71. Since p is the process output, K should be considered as the actual ratio setter. By using two multipliers to set the rate of each one of the flowing streams, the range can, however, be increased.

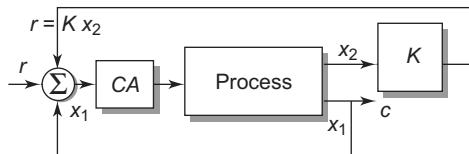


Fig. 6.2 Alternative arrangement of a ratio control system

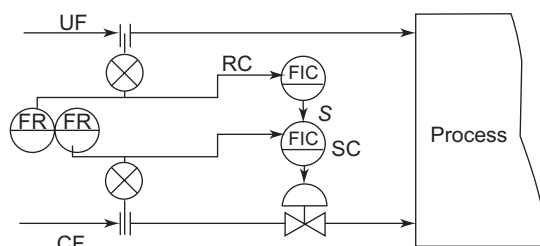
A practical example is the control of air flow in relation to controlled fuel (say oil) flow to an oil-fired furnace—a schematic representation of

which is given in Fig. 6.3(a) with flow recorders, ratio controller and a secondary controller which is in the line of the controlled flow. Besides such a combustion control scheme, ratio control is quite satisfactorily used in certain chemical processes where different variables like, flow, pressure, temperature and chemical compositions are interrelated.

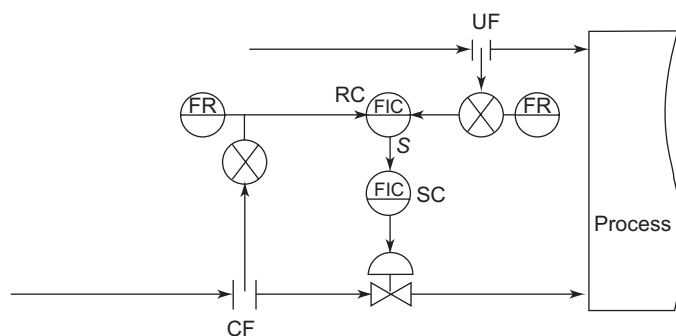
In Fig. 6.3 (a) the set point (S) for the flow controller is adjusted by the ratio control device RC-FIC, called the ratio station. The adjustment is done outside the control loop consisting of CF-FIC-SC and hence there is no interference between the adjustment and the loop response.

Sometimes an alternative arrangement has to be resorted to as shown in Fig. 6.3 (b). The ratio is calculated (RC) by CF and UF which is the input to the ratio controller (SC). This method has the ratio calculator (divider) within the control loop which would cause change in the gain of the loop because the divider is within the loop.

In Fig. 6.3(a) the setting of ratio in relation to the process has not been shown. It, therefore, assumes that this is perhaps a case of manual setting or setting without any relation to the process condition. It is often required that this setting be introduced as a control setting in relation to the process in terms of a specified parameter. The situation is explained in Fig. 6.3(b) of a blending process where it is clear that the primary variable is a function of the ratio of the two streams, that gives a condition that is designed in the



(a)



(b)

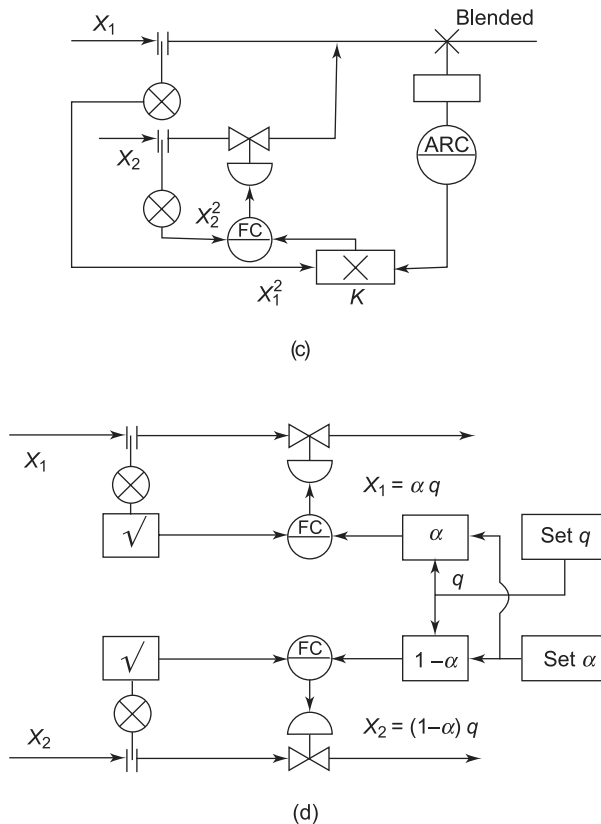


Fig. 6.3 (a) Practical example of a ratio control system; UF: uncontrolled flow, CF: controlled flow, RC: ratio controller, SC: setpoint controller, FR: flow recorder, FIC: flow indicator and controller, (b) Ratio control scheme with the divider (ratio set) inside the loop, (c) Ratio control in a blending process, (d) Ratio control with control over the flow rate; q : flow rate, α : blending factor

process. The output of the primary controller, thus, sets the ratio K , so that the secondary controller would obtain the control function as

$$x_2^2 = Kx_1^2$$

or

$$x_2 = \sqrt{K}x_1$$

Interestingly, the secondary loop gain would be $dx_2/dK = x_1^2/2x_2$.

The ratio setter here effectively is a multiplier which multiplies x_1^2 for setting the secondary controller.

The situation in Fig. 6.3(c), however, does not provide control over the total flow rate, say q , of the blended flow. Total flow rate, in fact, is the sum

of the two flow rates x_1 and x_2 . Hence, making $x_1 = \alpha q$ and $x_2 = (1 - \alpha)q$, one can have $x_1 + x_2 = q$. Thus the ratio x_1/x_2 can be varied from 0 to ∞ , varying α between 0 and 1 since $x_1/x_2 = \alpha/(1 - \alpha)$. The system will, however, work only with linear signal. Total flow and composition are independently selectable here. Fig. 6.3(d) explains the situation. Both the ratio setter or the multiplier here have their element factor (MEF) equal to 1 but the one in setting x_2 is reverse acting thus producing the desired result.

Some typical examples of ratio control systems are fuel blending, feed-water lines to boilers, gasoline product lines, etc. Ratio controllers are generally two mode (PI), although on special cases they may be three mode types.

6.3 SPLIT RANGE CONTROL

As the name indicates the control is effective in a very small range. This is made possible by the specific arrangement and 'pressure loading' of a pair of control valves in relation to a controller. A typical example is shown in Fig. 6.4 which is useful in a hydrogen absorber with nitrogen make-up. The nitrogen pressure build-up over a limited range depending on the overall system pressure is indicated. Control valve V_1 is closed on 0.2 to 0.6 kg/cm^2 signal and V_2 is open from 0.7 to 1 kg/cm^2 signal. The build-up range is only for 0.6 to 0.7 kg/cm^2 pressure signal obtained from the pressure controller.

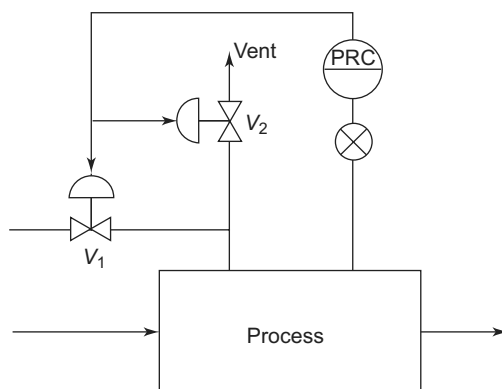


Fig. 6.4 Split range control system

6.4 CASCADE CONTROL

If a process consists of stages having different response characteristics, it is not possible to satisfactorily control such processes with a single controller. As an example, a two stage process, one stage of which has a much larger response time than the other, may be considered. This

occurs in heat exchangers as shown in Fig. 6.5, where C_1 is the cascade controller and C_2 is the set point controller for C_1 . The response time of the mainstream is evidently quite large compared to that of the jacketed steam because of less heat inertia and, therefore, controller C_2 would be inadequate to provide effective control in such a system. This is also true if there is any uncontrollable large disturbance in the steam line because then the performances with respect to these disturbances are also poor. As this scheme indicates, for disturbances in the steam line, the controller C_1 starts corrective action before the process output shows any deviation. The response may be 100 times faster than when a single controller is used. Thus an effective control scheme in such cases would be as shown and is known as cascade control. This implies an interaction between the mainstream temperature and pressure or flowrate of steam. Depending on the temperature of the main stream, C_2 will control the set value of controller C_1 . The pressure or flow rate input to C_1 from the steam jacket is differenced with the output of C_2 and an adequate control signal is sent to the control valve. Effectively, one secondary loop and one primary loop are formed in the system and the secondary loop may be designed such that the dynamic lag between the signal from the main controller (temperature) and the steam variable (pressure, or, flow rate) is reduced. The block diagram of the scheme of Fig. 6.5 is shown in Fig. 6.6. Upsets or disturbances are designated by u_1 and u_2 . In this case part of the process having transfer function G_p has the slowest speed of response, part of it having transfer function G_s responds more rapidly and control valve with transfer function G_v still more rapidly. The second loop around G_s speeds up the system as G_v actually determines the speed of response in this case. For a cascade control system to function properly, it is essentially necessary that the inner loop dynamics be, at least, as fast as those of the outer loop, if not faster. If these are faster, the inner loop controller would correct the effect of the disturbances in that loop before they get the chance to disrupt the controlled variable. If this condition is not satisfied, it would hardly be possible to tune the primary (outer loop) controller satisfactorily. The commonly encountered process variables with increasing response speed are temperature, pressure, liquid level and flow. Pressure and liquid level are often similar. Composition is slowest in response. The loop strategy is accordingly determined. Obviously, for the scheme of Fig. 6.5 to be effective, the secondary (inner) loop is to be much faster than the main loop. Assuming that the measurement block transfer functions are unity such that $G_{m1} = G_{m2} = 1$, the response to upset u_1 is obtained as

$$\frac{C}{u_1} = \frac{1}{G_{c2}G_{c1}G_v} \left[\frac{G_{c2}G_pQ}{1 + QG_{c2}G_p} \right] \quad (6.3)$$

where

$$Q = G_{c1}G_vG_s / (1 + G_{c1}G_vG_s) \quad (6.4)$$

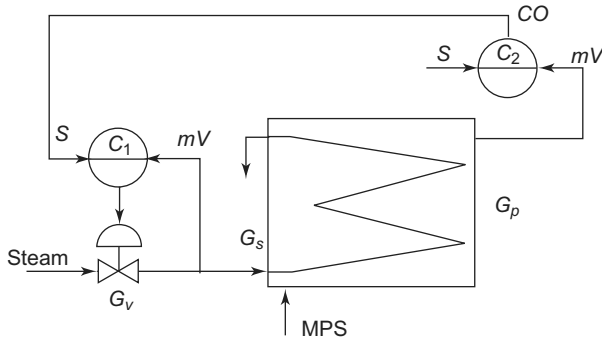


Fig. 6.5 Sketch of a heat exchanger process; mV : measured variable, CO : controlled output, S : set point, MPS : main process stream, G_s , G_p : process transfer functions, G_v : valve transfer function, C_1 , C_2 : controllers

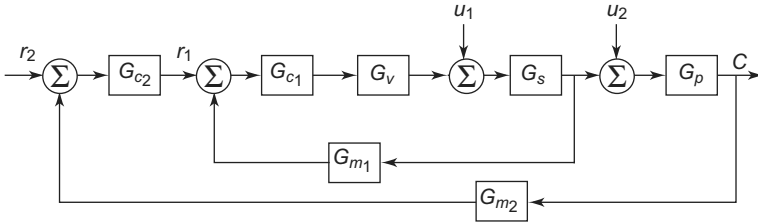


Fig. 6.6 Block diagram of the scheme of Fig. 6.5

In the absence of any cascade control this is given as

$$\frac{C}{u_1} = \frac{1}{G_{c2}G_v} \left[\frac{G_{c2}G_vG_sG_p}{1 + G_{c2}G_vG_sG_p} \right] \quad (6.5)$$

In terms of the frequency response analysis, one can show the use of the cascade control increases the critical frequency and also the allowable gain for the main controller C_2 . Both factors contribute to increase the controllability and, therefore, the error signal following a load change in the main loop (u_2) is reduced considerably in a short time. There is a large improvement in control because of the high gain in the secondary controller (C_1) and the load changes entering the secondary loop are well taken care of. In fact, the high gain in the secondary controller makes the phase curve of the inner loop quite flat and increases the critical frequency. This naturally requires that the inner loop should have a high gain control often termed as 'tight control'. For a proportional control with a proportional band of $1/K$ in a system with a time constant of $\tau/(1 + K)$, thereby reducing the time constant considerably. If this is the second largest time constant

it can be shown that there is a considerable increase in the controllability. For only one time constant, the secondary loop would reduce the delay to zero as $K \rightarrow \infty$, but then additional measurement and transmission lags may cancel the advantageous effects of the benefit so derived. Cascade control is actually prescribed where there is one dominant time constant in the secondary loop and two or three other lags in the main loop.

Considering Fig. 6.6, the inner loop has a transfer function with respect to reference as

$$T_i(s) = K_1/s^2 \tau_1 \tau_v + s(\tau_1 + \tau_v) + 1 + K_1 \quad (6.6)$$

where, $G_{c1} = K_1$, $G_v = 1/(s\tau_v + 1)$, $G_{m1} = 1$ and $G_s = 1/(s\tau_1 + 1)$.

Equation (6.6) is simplified to

$$T_i(s) = [K_1/(1 + K_1)]/[s^2 \tau_1 \tau_v/(1 + K_1) + s(\tau_1 + \tau_v)/(1 + K_1) + 1] \quad (6.7)$$

Response is very fast when damping is critical, i.e., $\zeta = 1$, which gives,

$$\frac{\tau_1 + \tau_v}{1 + K_1} = \frac{2\sqrt{\tau_1 \tau_v}}{\sqrt{(1 + K_1)}} \quad (6.8)$$

$$\text{or, } \sqrt{\tau_1/\tau_v} + \sqrt{\tau_v/\tau_1} = 2\sqrt{(K_1 + 1)}$$

$$\text{or, } \sqrt{\tau_1/\tau_v} = \sqrt{K_1 + 1} \pm \sqrt{K_1} \quad (6.9)$$

For very large K_1 , $K_1 \gg 1$,

$$\sqrt{\tau_1/\tau_v} = 2\sqrt{K_1}, \quad 0 \quad (6.10)$$

with K_1 very large it is thus obvious that τ_1/τ_v should also be very large as it is not zero, so that Eq. (6.10) gives

$$\tau_1/\tau_v = 4K_1 \quad (6.11)$$

Thus for fast response of the inner loop, i.e., for critical response condition τ_1/τ_v should be very large.

It is interesting to note that τ_v/τ_1 could also be made large as far as analysis is concerned but a high time constant valve would not produce the desired result and hence one should take τ_1/τ_v as the ratio in Eq. (6.9) and not the reciprocal of it.

Without the inner loop, it may similarly be shown that speed of recovery of the system is determined by the size of τ_2/τ_1 .

Supposing in the cascade control system of Fig. 6.6., $G_{m1} = G_{m2} = 1$, $G_{c1} = K_{c1}$, $G_{c2} = K_{c2}$, $G_v = 1/(s\tau_v + 1)$, $G_s = 1/(s\tau_1 + 1) = 1/(K_1 s\tau_v + 1)$ and

$G_p = 1/(s\tau_2 + 1) = 1/(K_2 s\tau_v + 1)$, where K_1 and K_2 are two multiplying factors such that $K_2 > K_1 > 1$. In the absence of the cascade control loop $c(s)/u_1(s)$ is

$$\begin{aligned} c(s)/u_1(s) &= \frac{\left(\frac{1}{K_1 s\tau_v + 1}\right)\left(\frac{1}{K_2 s\tau_v + 1}\right)}{1 + \frac{K_{c2}}{(K_1 s\tau_v + 1)(K_2 s\tau_v + 1)(s\tau_v + 1)}} \\ &= N_1(s)/(1 + F_1(s)) \end{aligned} \quad (6.12)$$

whereas, for the cascade control system

$$\begin{aligned} c(s)/u_1(s) &= \frac{\left(\frac{1}{K_1 s\tau_v + 1}\right)\left(\frac{1}{K_2 s\tau_v + 1}\right)}{1 + \frac{K_{c1}}{(s\tau_v + 1)(K_1 s\tau_v + 1)} + \frac{K_{c1}K_{c2}}{(s\tau_v + 1)(K_1 s\tau_v + 1)(K_2 s\tau_v + 1)}} \\ &= \frac{\left(\frac{1}{K_1 s\tau_v + 1}\right)\left(\frac{1}{K_2 s\tau_v + 1}\right)}{1 + \frac{K_{c1}(K_{c2} + 1)\left(1 + \frac{K_2}{1 + K_{c2}}s\tau_v\right)}{(s\tau_v + 1)(K_1 s\tau_v + 1)(K_2 s\tau_v + 1)}} = N_2(s)/(1 + F_2(s)) \end{aligned} \quad (6.13)$$

Both for $F_1(s)$ and $F_2(s)$, critical frequencies can be found so that $\angle F_1 = -180^\circ$ as also $\angle F_2 = -180^\circ$. With these frequencies used to evaluate $|F_1(s)|$ and $|F_2(s)|$, the amount of gain bandwidth product can be obtained and the improvement or otherwise with cascade control loop can be ascertained. Angle of $F_1(s)$ is

$$\angle F_1(s) = -\tan^{-1} \left[\frac{\omega_{c1}\tau_v(1 + K_1 + K_2) - \omega_{c1}^3\tau_v^3K_1K_2}{1 - \omega_{c1}^2\tau_v^2(K_1 + K_2 + K_1K_2)} \right] \quad (6.14a)$$

which must be equal to -180° . Then

$$\omega_{c1} = (1/\tau_v) (\sqrt{(1 + K_1 + K_2)/(K_1K_2)}) \quad (6.14b)$$

Similarly for $\angle F_2(s)$, assuming $K_2/(1 + K_{c2}) = K_3$.

$$\angle F_2(s) = -\tan^{-1} \frac{\omega_{c2}\tau_v(1 + K_1 + K_2) - \omega_{c2}^3\tau_v^3K_1K_2}{1 - \omega_{c2}^2\tau_v^2(K_1 + K_2 + K_1K_2)} + \tan^{-1} K_3\omega_{c2}\tau_v$$

which when equated to -180° , yields

$$\omega_{c2} = \frac{1}{\tau_v} \sqrt{\frac{1 + K_1 + K_2 - K_3}{K_1K_2 - (K_1 + K_2 + K_1K_2)K_3}} \quad (6.15)$$

Enhancement in ω_{c2} would obviously depend on K_3 which is a function of K_2 and K_{c2} . It would be seen from Eq. (6.15) that K_3 is constrained by two equations

$$K_3 \leq 1 + K_1 + K_2 \quad (6.16a)$$

and

$$K_3 \leq K_1 K_2 / (K_1 + K_2 + K_1 K_2) \quad (6.16b)$$

For K_1 and K_2 both greater than unity Eq.(6.16b) shows that K_3 is a fraction which automatically satisfies condition (6.16a) as well. From the condition $|F_1(s)| = 1$, K_{c2} is found out simultaneously satisfying conditions (6.16) for an appropriate K_3 , using which ω_{c2} is evaluated. Then from $|F_2(s)|$ at ω_{c2} K_{c1} is calculated. Gain bandwidth product can also be calculated by using the peak values of K_{c1} and K_{c2} and critical frequencies.

Example 1

Let $K_1 = 10$, $K_2 = 100$ and $\tau_v = 1$.

Solution Using Eq. (6.14b),

$$\omega_{c1} = \sqrt{111/1000} = 0.33 \text{ r/s}$$

Using this value of ω_{c1} , $|F_1(s)|$ is calculated as $K_{c2}/119.5$, so that maximum K_{c2} is 119.5. This gives a gain bandwidth product of about 40. Choose working K_{c2} for the second case from the inner loop consideration alone and use that for the cascade control performance evaluation. Let this K_{c2} be 115. Using Eq. (6.15), ω_{c2} is now found out as 1.77 r/sec. Using this value of frequency $|F_2(s)|$ is calculated as $K_{c1}/29$, so that $K_{c1\text{max}}$ is 29 giving the gain bandwidth product as 51.4 which is an improvement over 40. It must be noted that since the inner loop has only two dynamic elements, the controller gain can be given a still larger value than is assumed for further increasing the controllability. The above technique has been shown to be analytically used because of a limited number of dynamic elements. When the dynamic elements are more, such trigonometric procedure will fail because the critical frequency can no longer be computed by the simple procedure as shown. Bode-plot and Nichols' chart then become of considerable use for the controller design.

In such a case the open loop transfer function of the process is still used which, normally, is obtained through modelling. The step by step approach of the procedure is now described.

- (1) The transfer functions of the individual blocks are first considered.
- (2) The frequency response diagram in log frequency-log amplitude ratio and log frequency-phase of the individual blocks (presumably available in factored forms) are first obtained and are graphically added to obtain the composite Bode-plot of them in series for the inner loop.

- (3) From the -180° condition, the gain and frequency are found out which are the critical gain and the critical frequency. Using Ziegler-Nichols' tuning rule K_{c1} (and if necessary T_{r1} and T_{d1}) is determined.
- (4) With this controller the inner loop is now reduced to a single block using Nichols' chart.
- (5) Bode-plots of the elements of the outer loop including the reduced one in step (4) but excepting the controller are now obtained as before and the composite Bode-plots are drawn.
- (6) From -180° phase condition again the critical gain and frequency are determined and then using Ziegler and Nichols' tuning rule the controller parameters K_{c2} , T_{r2} and T_{d2} are obtained.

As indicated, often the inner controller is only a proportional one. The Bode-plots can also be used to check the improvements in the control using cascade system from that when not used. For this, from the overall plots including the controller as designed above for the inner loop with a proportional controller only, the critical frequency and gain are evaluated and the product of the two determined. In another case, controller G_{c1} is dispensed with and a single loop configuration is considered whose critical frequency and gain are obtained proceeding as above. This would give the gain bandwidth product for the noncascade condition. The ratio of the two would give a measure of the improvement. An example at this stage would illustrate the procedure clearly. Let us consider Fig. 6.6 again, and in that, let $G_{m1} = G_{m2} = 1$, $G_v(s) = 1/(s + 1)$, $G_s(s) = 1/((s + 1)(10s + 1))$, $G_p(s) = 1/(100s + 1)$ and $G_{c1} = K_{c1}$, indicating that the inner loop uses a proportional controller only. First step would be to prepare the Bode-plots of each of $G_s(s)$ and $G_v(s)$. Then corresponding to -180° phase in the composite phase plot of the above two transfer functions the frequency and gain are found out. For this particular case it can be trigonometrically computed as well, as discussed earlier. Thus

$$-180^\circ = \tan^{-1} \frac{10\omega_c + \omega_c + \omega_c - 10\omega_c^3}{1 - 10\omega_c^2 - 10\omega_c^2 - \omega_c^2}$$

giving $\omega_c = 1.09$ $r/sec.$ and the amplitude ratio at this frequency is obtained as 0.0413. Hence the value of K_{c1max} is $1/0.0413 = 24.2$. From Ziegler-Nichols' tuning law $K_{c1} = 24.2 \times 0.5 = 12.1$.

The inner loop is now replaced by a single block which has a transfer function $G_i(s)/(1 + G_i(s))$ where $G_i(s) = K_{c1}G_v(s)G_s(s)$. The Bode-plot $G_v(s)G_s(s)$ is already there (Fig. 6.7) on which that of K_{c1} is superposed to get the composite Bode-plot $G_i(s)$ which makes a change in the amplitude curve only by raising it by 12.1 units. The phase curve, however, remains unchanged. For a given $G_i(s)$, $G_i(s)/(1 + G_i(s))$ can be determined from the Nichols' chart which is a standard chart in open loop gains and phases

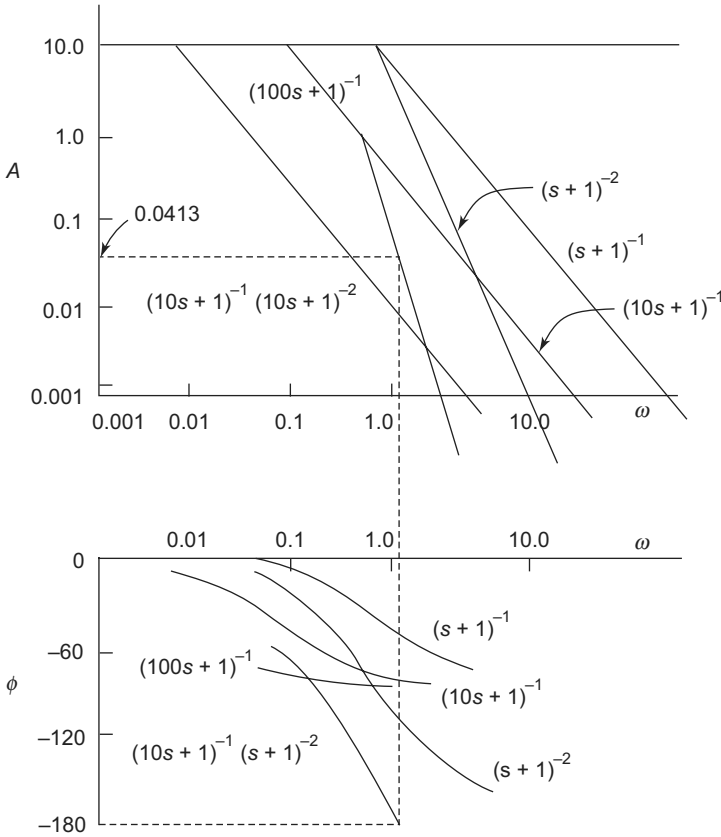


Fig. 6.7 Bode-plot, (a) Amplitude ratio frequency, (b) Phase frequency

with curves drawn for closed loop gains and phases as parameters. A table (Table 6.1) is now prepared with gain and phase entries of $G_i(s)$ from Bode-plot, $G_i(s)/(1 + G_i(s))$ from Nichols' chart, $G_p(s)$ from calculation and $G_p(s)G_i(s)/(1 + G_i(s))$ by multiplication for several values of ω . The table is shown below while the Bode-plot in Fig.6.7(a) and (b). Standard Nichols chart is consulted (see Appendix II, Fig. A-1).

Now, at $\phi_t = -180^\circ$, the amplitude ratio of the overall open loop transfer function is found by extrapolation to be equal to 0.0145 giving the ultimate gain as $K_{c2\max} = 68.9$. Again using Ziegler-Nichols' tuning rule one gets

$$K_{c2} = 0.45 \times 68.9 = 31$$

$$T_r = 0.825 \times 2\pi/\omega|_{\phi_t = -180^\circ} = 0.825 \times 6.28/1.05 = 4.9 \text{ min, for PI}$$

$$K_{c2} = 0.6 \times 68.9 = 41.34$$

$$T_r = 0.5 \times 2\pi/1.05 = 2.99 \text{ min}$$

$$T_d = 0.125 \times 2\pi/1.05 = 0.747 \text{ min, for PID}$$

Table 6.1 Frequency versus Open Loop Transfer Ratio

ω rad/sec	$\frac{G_i(s)}{ A_i } \quad \phi_i$		$\frac{G_i(s)/(1+G_i(s))}{ A_{ic} } \quad \phi_{ic}$		$\frac{G_p(s)}{ A_p } \quad \phi_p$		$\frac{G_p(s)G_i(s)/(1+G_i(s))}{ A_t } \quad \phi_t$	
0.1	8.34	-56°	0.90	-5°	0.1	-84°		
0.2	4.9	-85°	0.96	-12°	0.05	-87°		
0.4	2.6	-120°	1.10	-24°	0.025	-88.5°		
0.5	1.9	-133°	1.3	-32°	0.02	-88.8°	0.026	-120.8°
0.6	1.4	-145°	1.7	-48°	0.016	-89°		
0.8	1.3	-160°	1.6	-55°	0.011	-89.2°		
1.0	1.2	-170°	1.5	-60°	0.01	-89.4°	0.015	-149.4°
1.1	1.05	-175°	1.45	-100°	0.009	-89.47°	0.013	-189.47°
1.2	0.5	-180°	1.4	-180°	0.008	-89.5°	0.011	-269.5°

If cascade control is not used $G_{c1} = 1, G_{m1} = 0$ and from the combined Bode-plot of $G_s(s), G_p(s)$ and $G_v(s), K_{c2}$ is evaluated easily as also T_r and T_d . Nichols' chart is not necessary in this situation. Using trigonometric relation again critical frequency for single loop is obtained as 0.23 r/min and using this frequency the amplitude ratio is 0.0164 so that $K_{c2\max}$ in this case is 60.9. without cascade control, therefore, the gain-bandwidth product is $60.9 \times 0.23 = 14$, whereas with cascade control this is $68.9 \times 1.05 = 72.34$ showing an improvement of over 5 times.

A cascade control would be recommended:

- (i) where the overall process is slow to respond to process disturbances/corrections and large deviations result,
- (ii) where an intermediate process variable directly related to the controlled variable exists and is affected by the process disturbances and can be controlled by the main controlled variable.

The advantages of a cascade control:

- (i) disturbances in the secondary loop are taken care of by the secondary controller, before they can influence the primary variable,
- (ii) secondary loop reduces the phase lag, i.e., the response time in the secondary part considerably, thereby improving the response of the primary loop as well,
- (iii) secondary loop allows manipulation of the primary controller in such a way that exact mass/energy flow is stipulated by it, and
- (iv) process gain variations in the secondary part are taken care of in this loop itself.

The disadvantages:

- (i) cascade control cannot be employed indiscriminately; only when a suitable intermediate variable can be measured does this method of control fit in properly; and

- (ii) cascade action fails to yield the desired results if the inner loop is closed around the largest time constant of the part of the process. In fact, cascade control is effective only when the secondary time constant is smaller than the primary time constant. From the following approximate analysis this becomes clear.

In Fig. 6.6, let $G_{m1} = G_{m2} = G_v = 1$, $G_s = 1/(s\tau_1 + 1)$, $G_p = 1/(s\tau_2 + 1)$, $G_{c1} = K_{c1}$ and $G_{c2} = K_{c2}$, the overall transfer function is

$$T_0(s) = \frac{K_{c1}K_{c2}/(1 + K_{c1})}{\frac{s^2\tau_1\tau_2}{1 + K_{c1}} + s\left(\frac{\tau_1}{1 + K_{c1}} + \tau_2\right) + \left(1 + \frac{K_{c1}K_{c2}}{1 + K_{c1}}\right)}$$

The critical damping conditions give

$$\tau_1^2[1/(1 + K_{c1}) + \tau_2/\tau_1]^2 = 4\tau_1^2[\tau_2/\tau_1(1 + K_{c1})](1 + K_{c1}K_{c2}/(1 + K_{c1}))$$

and since K_{c1} is very large so that $1/(1 + K_{c1}) \rightarrow 0$, above equation transfers to

$$\tau_2/\tau_1 = 4K_{c1}K_{c2}(1 + K_{c1})^2 \approx 4K_{c2}/K_{c1}$$

Both K_{c1} and K_{c2} are large in cascade control system, although often $K_{c2} < K_{c1}$.

In any case, $\tau_2/\tau_1 > 1$. In fact for $K_{c2} = fK_{c1}$, where f is a fraction, $\tau_2 = 4f\tau_1$. Generally, $\tau_2 \geq 3\tau_1$ is considered as usual.

The speed of recovery with a cascade control or, more specifically, the performance improvement with cascade control is, however, governed by the speeds of responses of G_v and G_s as indicated in Eq. (6.11). The improvement factor is generally specified as F , and is given by

$$F = \frac{\text{Time constant of the second most-slow element in loop 2}}{\text{Time constant of the second most-slow element in loop 1}}$$

Some of the typical application fields of cascade control are

- (i) temperature control of a steam-jacketed kettle,
- (ii) temperature control of metallurgical furnaces,
- (iii) boiler control,
- (iv) reboiler temperature and flow control of distillation column, etc.

6.5 FEEDFORWARD CONTROL

An alternative for cascade control is the feedforward control. Specifically when cascade control cannot be used because of the incompatibilities mentioned above, feedforward may be tried, although certain constraints do exist in this case as well. The constraints are:

- (i) disturbances should be easily measurable, and
- (ii) frequency spectra of disturbances should not be too wide relative to the bandwidth of the regulatory system.

There are actually two different types of feedforward control schemes known as :

- (i) impulse type feedforward, and
- (ii) predictive feedforward.

Both are used with feedback.

Feedforward provides closer control of certain processes involving large lags including dead time. It also provides better stability with two or more closely interacting control loops. The impulse type feedforward control scheme is depicted in Fig. 6.8 for a fired heater which clearly shows the flow of

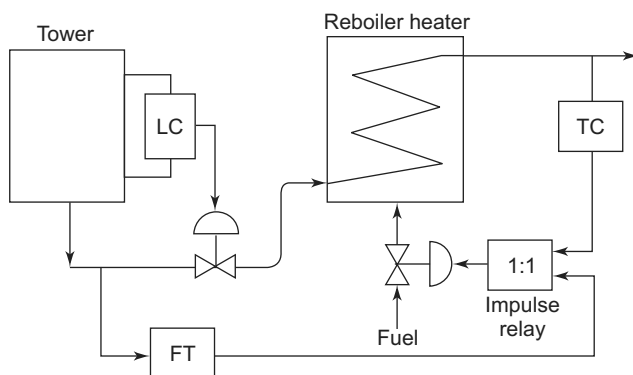


Fig. 6.8 Schematic diagram of an impulse type feedforward control system

signal in the forward direction, i.e., the direction of energy flow. The level change acts as a disturbance which can be sensed by TC only when this change reaches the output point in the absence of a feedforward control. The amount of fuel will not change and when TC starts acting, as the fuel inflow will remain the same during the travel of disturbance in the mean time, the output temperature will change. The feedforward line is added through FT to avoid this. Flow transmitter (FT) actually sends a signal proportional to the rate of the inlet flow and the expected change in the flow rate is reflected in the fuel control through the impulse relay which sends a signal proportional to the outlet temperature and proportional to the derivative of the flow inlet. This type of control is used in a two-element boiler drum level control. For a sudden change in the steam flow the impulse circuit compensates for the swell/shrink of the level. In some types of three-element control of a steam boiler drum level the predictive feedforward control is used (see Fig. 6.9).

In the feedforward control schemes, therefore, an attempt is primarily made to cancel the effect of the disturbances before they can appreciably affect the output. This means that as soon as the disturbance starts, corrective action should also start. Therefore, theoretically, the feedforward control is effective for providing perfect control to any type of process; practical limitations are, however, often there.

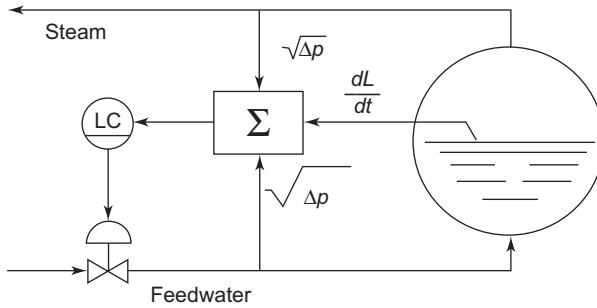


Fig. 6.9 Predictive type feedforward control of a boiler drum

In addition to the usual feedback loop, the feedforward control consists of the measuring elements for measurement of disturbances and the feedforward controllers, often called the load compensators, to form a separate loop. The sense of the outputs from these compensators should be in opposition to the respective disturbances. Essentially there is a forward flow of information in the auxiliary loops but the controlled variable is never involved. If the signal, which has the potential of upsetting the process if no action is taken, could be measured and transmitted to a controller which would, in turn, act on this signal and calculate a new value of the manipulated variable and send to the actuator which when acts under this, the controlled variable is not affected by the above signal, then there is a system where the error in the controlled variable is not fed back but changes in load is fed forward. A set point is, however, necessary.

Even a feedforward control cannot be a perfect control, mainly because of inadequate modelling and representation of the plant characteristics and the exclusion of some load components in this process of representation. This exclusion tends to induce uncertainties and offset which may not be negligible. By far the best way to at least partially annul such uncertainties is to use feedback in the reset mode to adjust the set point. The inclusion of the derivative mode is usually not necessary as the feedforward control takes care of the demanding processes; besides it tends to introduce oscillatory nature. A number of causes are known to produce the effect of offset and for this reason one initially decides whether or not the feedback is to come and if so where it should be introduced. These are really difficult questions. When the decision for the exact control strategy cannot be taken, set point readjustment for the feedforward case as shown in Fig. 6.10 is the best method. The similarity here with cascade control is obvious.

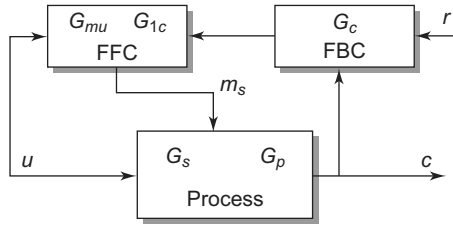


Fig. 6.10 Block diagram of a feedforward control scheme with provision for set point readjustment, FFC: feedforward controller, FBC: feedback controller

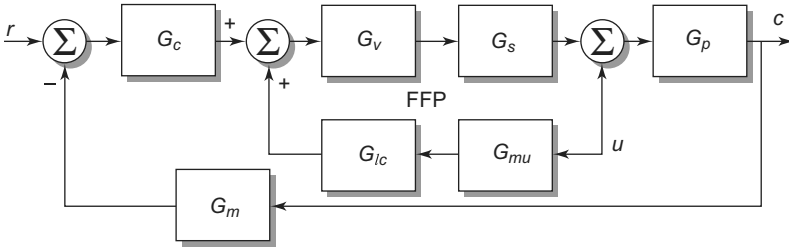


Fig. 6.11 Completed scheme of a feedforward control system; FFP: feedforward path

The feedforward control is often referred to as the disturbance feedback control as is evident in the scheme of Fig. 6.11 and then the output-disturbance response could be easily calculated as

$$\frac{c}{u} = \frac{G_p(1 + G_{lc}G_vG_s)}{1 + G_mG_cG_pG_vG_s} \quad (6.17)$$

with $G_{mu} = 1$.

Equation (6.17) shows that if

$$-G_{lc}G_vG_s = 1 \quad (6.18)$$

c/u can be made zero. Thus, when G_v and G_s are known, the load compensator or the feedforward controller is designed following

$$G_{lc} = -1/(G_vG_s) \quad (6.19)$$

If G_s is of the form $K_s/(1 + s\tau_s)$ and $G_v = 1$, the ideal load compensator would be

$$G_{lc} = -(1 + s\tau_s)/K_s \quad (6.20)$$

which is a proportional and derivative controller. Because of the nature of the load compensator, a ramp disturbance is rather easy to tackle. A step disturbance is very difficult to compensate because of infinite range of frequency content. If $G_v = 1/(1 + s\tau_v)$, we get

$$G_{lc} = -(1 + s\tau_v)(1 + s\tau_s)/K_s \quad (6.21)$$

and a second order derivative action would be needed for the compensation. It has, however, been suggested that the derivative time can be set at the sum of the two time constants and in such a situation the improvement depends on the ratio of total lag in the final process elements (G_p) to the sum of the lags in G_s .

In Fig. 6.11, load or upset transfer function has been considered unity. If, however, it has a transfer function $G_u(s)$, $G_u(s)$ would oppose $G_{lc}(s)$ $G_v(s)G_s(s)$ as the numerator of the right hand side of Eq. (6.17) would be $G_p(s)(G_u(s) + G_{lc}(s)G_v(s)G_s(s))$ and $G_{lc}(s)$ would be given as

$$G_{lc}(s) = -G_u(s)/(G_v(s)G_s(s)) \quad (6.22)$$

Now, with first order models of upset and process parts, $G_u(s) = K_u/(1 + s\tau_u)$, $G_v(s) = 1$, and $G_s(s) = K_s/(1 + s\tau_s)$, the compensator is then given by

$$G_{lc} = -(K_u/K_s)(1 + s\tau_s)/(1 + s\tau_u) \quad (6.23)$$

The forward controller has thus a steady state gain of K_u/K_s and the dynamic part of it represents a lead-lag element. For this to be physically realizable τ_u should be non-zero positive. For a very fast load variation it is assumed that τ_u is quite small but still not zero and $\tau_s \gg \tau_u$.

If the dynamic part $(1 + s\tau_s)/(1 + s\tau_u)$ is dropped from Eq. (6.23) the remaining part is referred to as the steady state feedforward controller. The effects of this controller only and that given by Eq. (6.23) are shown in Fig. 6.12 with pulse disturbance occurring as shown. The output curves are shown with feedforward control only—the curve of Fig. 6.12(b) is for the steady-state feedforward action only whereas that of Fig. 6.12(c) is with dynamic compensation as well. It is assumed that the dead times present in both the process and the disturbance are nearly the same and cancel out.

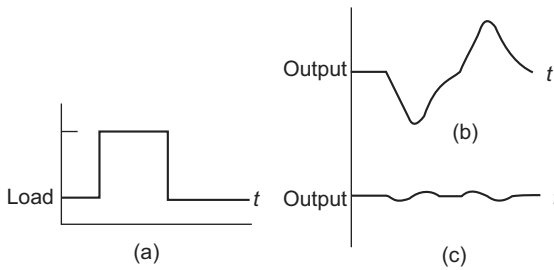


Fig. 6.12 Transient changes: (a) head change, (b) static feedforward control, (c) dynamic compensation

In most cases a simple lead-lag function would be quite adequate and reduce the dynamic area of the response curve by a factor of 10 or even more. It is easy to see that when a load response curve crosses the set point, a lag unit is necessary and vice versa. Usually, a proper matching of the

lead time τ_s and the lag time τ_u is necessary so that adequate compensation can be provided. If τ_p is the time for maxima or minima, it can be shown to be related to τ_s and τ_u as

$$\tau_p = \ln(\tau_s/\tau_u)/(1/\tau_u - 1/\tau_s) \quad (6.24)$$

It contains two variables τ_u and τ_s , and, therefore, the adjustment is not unique. However, τ_u is generally not alterable but τ_s is. If a depression occurs one should make τ_s greater than τ_u and vice versa. An initial setting starts with $\tau_s = 2\tau_u$ or $\tau_s = 0.5\tau_u$ in the two cases respectively. Perfect compensation is hardly attainable primarily because of the dead time which further complicates the process by its variation. When $\tau_u \rightarrow 0$, the compensator itself is provided with the lag term. It is good to remember that unless a sufficiently fast change in disturbance occurs, like that in flow rate, dynamic compensation with lead-lag unit is not essential.

It has been discussed above that a sudden change in upset has to be compensated for by a lead-lag network. This follows from the transfer function of the block schematic shown in Fig. 6.11. Arranging the scheme as shown in Fig. 6.13 and assuming ideal measurement ($G_{mu} = 1$), one easily derives

$$A(j\omega) = x(j\omega)/u_1(j\omega) = 1 - G_{lc}(j\omega)G_v(j\omega)G_s(j\omega) \quad (6.25)$$

where x is the upset modified by the compensating loop, $A(s)$ is the transmission function from u_1 to x and s is replaced by $j\omega$.

For the disturbance to contain an infinite number of frequencies, the minimization of $x(j\omega)$ would mean minimization of the 'square' integral

$$I = (1/2\pi) \int_{-\infty}^{\infty} |A(j\omega)|^2 d\omega \quad (6.26)$$

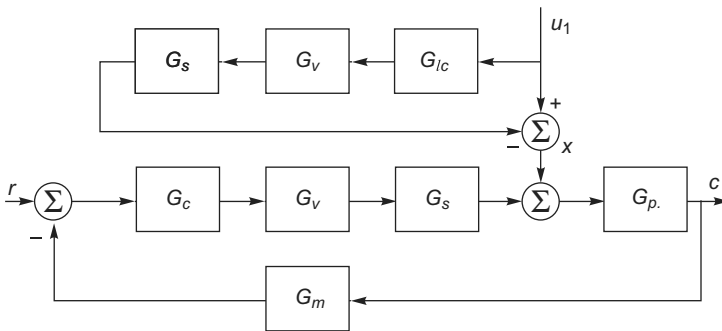


Fig. 6.13 Rearranged scheme of Fig. 6.11 for ease of calculation

Unfortunately I tends to be infinite. If u_1 is nonwhite, i.e., it contains only a finite number of frequencies, there is possibility of optimization by a proper choice of G_{lc} . If the spectrum of u_1 is unknown, a unit step function is assumed and then the filtering effect of the loop is assumed to act. With the step function of u_1 Parseval's theorem would give

$$\int_{-\infty}^{\infty} x^2(t)dt = (1/2\pi) \int_{-\infty}^{\infty} (|A(j\omega)|^2 / \omega^2) d\omega \quad (6.27)$$

Then

$$\frac{c(j\omega)}{u_1(j\omega)} = \frac{G_p(j\omega)A(j\omega)}{1 + G_p(j\omega)G_m(j\omega)G_c(j\omega)G_v(j\omega)G_s(j\omega)} \quad (6.28)$$

A minimization of c is more effective because of the loop filtering effect. A compensation network provided would perform filtering in addition to the system blocks.

The mean square value of x , $\bar{x}^2$ may be evaluated for specific G_{lc} for the given process loop components and for any assumed finite bandwidth of u_1 . A figure of merit is defined as

$$F = \frac{\bar{x}^2|_{\text{no feedforward}}}{\bar{x}^2|_{\text{feedforward optimum}}} \quad (6.29)$$

The wider the spectrum the less is the improvement. A sustained deviation is easily cancelled. A pulsed change in u_1 as shown in Fig. 6.12 would usually contain a wide spectrum and hence would be very difficult to compensate by this method, and as is pointed out the lead-lag filtering greatly improves the performance.

In conclusion it may be mentioned that if the main disturbance can be included in the inner loop, a cascade control is better for more reduction of error in a specified time. However, if the last process element has a very large process lag, much greater than the others and the disturbance occurs just before this lag, feedforward is more useful.

6.6 SELECTOR CONTROL

This is a specific type of control where the system has a normal state and the specified normal controller is in operation; however, whenever an abnormal condition arises a selector relay is used to select a suitable controller from a given lot to take over and as soon as the abnormal condition is over an automatic return to the normal state occurs. This type of control is also called override or limit control. This type of control is very widely used in industry. As has been mentioned already the operation is performed by selector relays and a continuous vigilance is maintained by a state change module. Although switching over to a new controller takes place in emergent conditions, actuator remains the same. A typical control scheme that for a compressor station discharge is shown in Fig. 6.14 where a single common actuator is shown connected to two control loops. Loops may be more depending on the requirement. Normally, discharge controller C_2 is in operation. If, however, the suction controller

goes below a preset limit, threatening to cause vapour lock and/or burn-out of the compressor, suction controller C_1 takes over control via the low pressure selector relay L . The reset feedback to both the controllers being the pressure of the valve, provides the equalizing connection for avoiding the reset wind-up in the stand-by controller.

In the absence of a selector control, in many cases such abnormalities are manually tackled, otherwise a shutdown would be imminent. The selector control thus can prevent shut down and boost up production economy. The type of selector control mentioned above is known as (i) limiting selector control. This is often used in combustion processes and heat exchangers for safe temperature limits or the same differential temperature limits.

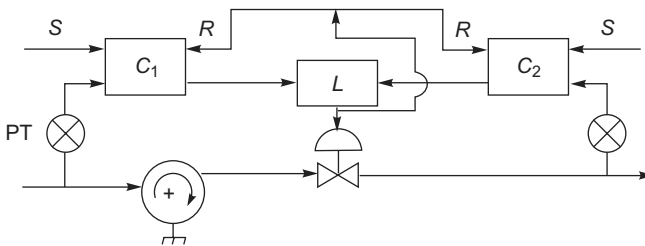


Fig. 6.14 Scheme of a compressor station discharge control;
S: set, R: reset, PT: pressure transmitter L: Load distributor

Another type is for the split load supply called (ii) selector split load supply control. This type is often used in air compressor control with N number of flow outputs. The selector relay receives signals from the outlet controllers and transmits a signal corresponding to the pressure of the outlet where the control valve is open widest. The compressor is controlled guided by this signal such that it can satisfy the needs of the most heavily loaded line, the needs of the other lines being automatically satisfied. Other applications are in fuel distribution at different points in a furnace. The schematic of such a system is known in Fig. 6.15.

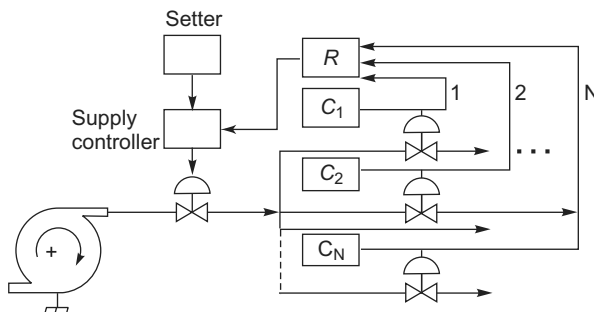


Fig. 6.15 Schematic diagram of a selector control for split load supply;
 C_i : i th controller, R: split relays

A third type of control known as (iii) flow distribution selector control is only a variation of type (ii) and is often used with a master controller. For N flow channels a selection of the control point of the individual line controllers is performed by a selector relay system which is fed from the line flow. Here also, the widest open control valve in a channel sends a signal which effects the throttling of the other channels but because of the variation in application this system is to be looked a bit differently. This type of selector control has been successfully used in the tuyere flow control in blast furnaces in steel plants for maximum economy as regards flow and protection to the furnace. When in a particular tuyere line, the butterfly valve opens widest (as it would like to send a maximum flow because a resistance to flow may have developed due to the furnace load coming down at that point), the signal is received by the selector relay and in correspondence to this other line, controllers are adjusted to have the same flow in the line. The scheme of this system is shown in Fig. 6.16(a).

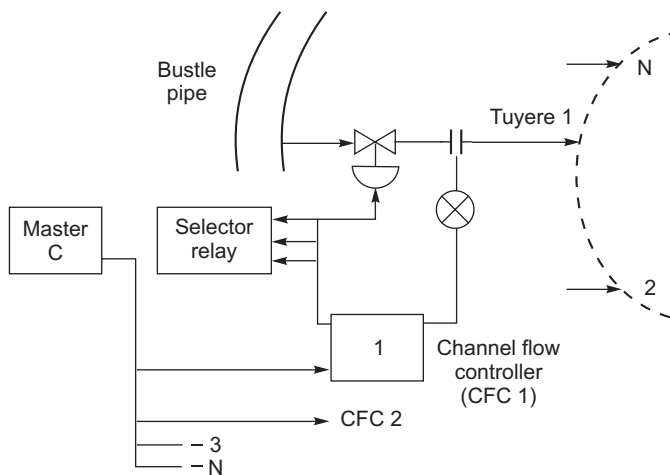


Fig. 6.16 (a) Schematic diagram of a selector control used in tuyeres of blast furnaces

A typical two-input autoselector pneumatic relay is shown in Fig. 6.16(b). Ports 1 and 2 are connected to the output of the controller C_1 and ports 4 and 5 to the output of controller C_2 . There are two diaphragms connected differentially but rigidly through a shaft. Two spring-opposed valves are connected at their stems to this shaft by brackets, as shown. When the output of C_1 is more, the differential movement of the shaft closes the valve connecting port 4 and opens the valve connecting port 2, thereby bringing C_1 in the control line. For the other case, C_2 comes in line. Slight adjustment facilities are also provided as shown. The output of the relay is drawn from port 3, which is connected to the input of the control valve.

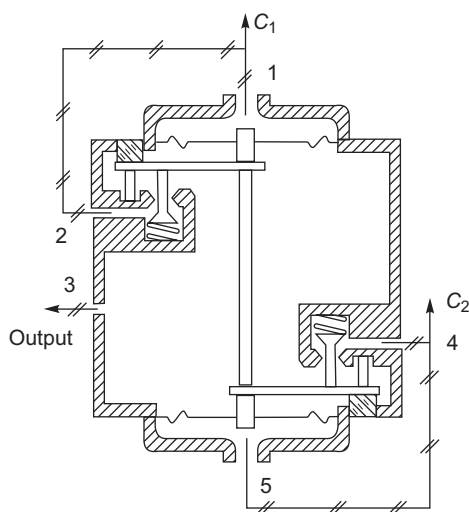


Fig. 6.16 (b) Schematic diagram of an autoselector pneumatic relay

A type of control scheme known as selective or auctioneering control is a little different from the selector control schemes described above. When a variable supposed to assume reasonably steady value along the process length at several places is seen to have different values having maximum and minimum as well, a control scheme is required to control the value at maximum or minimum points to a reasonably acceptable value. However the location of the maximum and minimum may also not remain fixed. A typical example of such a system is an exothermic cooled tubular reactor with constant coolant temperature. The maximum temperature spot is called the hot spot which changes position depending on the flowrates and compositions of the streams. Figure 6.17 shows the schematic diagram of such a reactor with only the auctioneering control loop. The reactor has its own control loop for throughput maximization. In this scheme temperature along the reactor length is measured at several points (5 shown here). The probes give their outputs to a logic block where the maximum is selected by 'greater than' (or 'smaller than') logic which is then used to control the coolant flowrate. At the bottom, three curves showing the location change for maximum are drawn.

One of the processes frequently met is the production of cumene via irreversible gas phase alkylation of benzene with propylene. Different configurations are suggested. Common 3-point configuration is tightness of hotspot temperature control, maximum process throughput subjected to the constraint of reactor cooling. The configurations manipulate the cooling process for tightmost hot spot temperature control, manipulate the reactor inlet temperature as there occurs large deviation in temperature

with increased throughput and manipulates fresh feed of propylene for maximizing throughput but hot spot temperature needs be regulated.

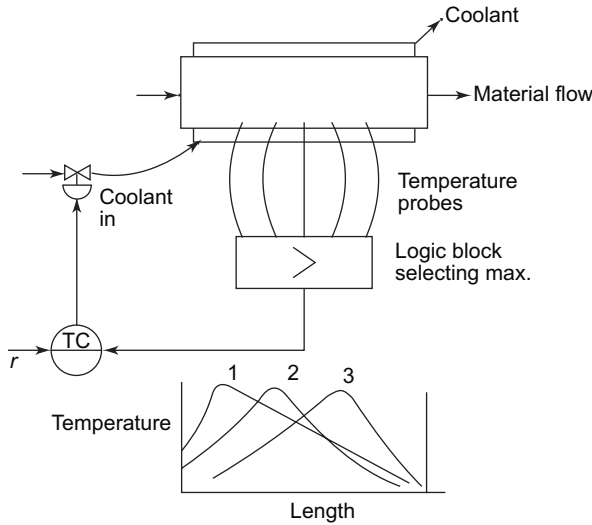


Fig. 6.17 Scheme of auctioneering control

6.7 INVERSE DERIVATIVE CONTROL

Direct derivative action is used for temporarily narrowing the throttling range, i.e., narrowing the ‘proportional band’—the amount of narrowing being dependent on the speed of the process variable and independent of the control point. Its inverse would, therefore, widen the throttling range temporarily and it lags the output of the controller unlike the derivative case. In flow control, i.e., in control of fast processes, oscillations tend to be generated unless proper control actions are given with requisite tuning; often these oscillations are avoided by widening the throttling range of a proportional action controller till stability is attained, which, in some cases may come only when the band reaches 500 per cent or even more. This, however, means that the control action is practically ineffective. When such a control process has stabilized, it is permissible to narrow the throttling range so that control action may return to its full swing but only if the process condition remains steady. Any slight change in the process may again initiate cycling, till band is widened again or some other form of damping is provided. However, the inverse derivative action may solve this problem of band-widening and narrowing automatically. As long as the process is steady the band may be narrowed to less than 20 per cent and on slightest change in the process this would widen up to the required value till stability is attained and then again automatically return to the

narrow range. Thus, this action combines the advantages of the good control by narrow throttling and good stability by wide throttling as and when necessary. This action is usually added to the proportional and reset units and used on processes of short time lags. A good application example is the boiler drum level control which basically is a flow control system with feed water inflow and steam outflow. Any sudden increase in the demand of steam would increase feed water inflow thereby tending to destabilize the process. Wide throttling would prevent this. In normal operation, however, narrow band operation continues. Flow and pressure being fast processes, inverse derivative action may be of convenience instead of direct derivative action.

6.8 ANTIRESET CONTROL

The reset/integral action in a *PI* controller causes its output to go on changing as long as the error is non-zero. For various reasons the error cannot be eliminated quickly in many situations and, therefore, if the time interval is long enough, larger and larger values of the manipulated variable develops due to reset action leading finally to saturation which means that the valve completely opens. This condition is known as the reset wind up or integral wind up and is found to occur during manual operational change like shut down, start-up, changeover, etc. On return to auto operation, the control action remains saturated which produces large overshoots. Special provisions are required to be made to cope with this saturation. One such is anti-reset control. The anti-reset wind-up relay can be used in such a case to throttle and operate at a specified pressure. It allows the control valve to remain fully open (1 kg/cm^2) as long as the controller output is less than full pressure with reset feedback connection as made in Fig. 6.14 and with the reset acting normally. With the output exceeding this pressure, the relay exhausts the reset feedback line and an output of 1 kg/cm^2 is maintained. Thus the *P*-action is only inducted at start-up, preventing overshoot. One other technique, used often in electronic controller, is suggested in Fig. 5.32(b) where the integral action is temporarily bypassed. A practical version of the same is shown in Fig. 6.18(a). To the controller a comparator *C* and a switch *S* are added as shown. When controller output remains below V_{omax} , the comparator output is such that the switch remains open and the integral action acts as usual. If the output exceeds V_{omax} , the comparator output is such as to close the switch *S* grounding the input point of the integral action controller so that integral action ceases. Resistance *R* is used to load the error amplifier in such a situation in the integration action instead of direct grounding of the comparator output.

The control law normally adopted in *P*, *PI*, *PID* controllers is the position control form, that is, the action of the controller is dependent on

the actual value of the controlled variable. With the change that the action should depend on the incremental change of the controlled variable, this type of problem can be avoided and is actually done in some situations

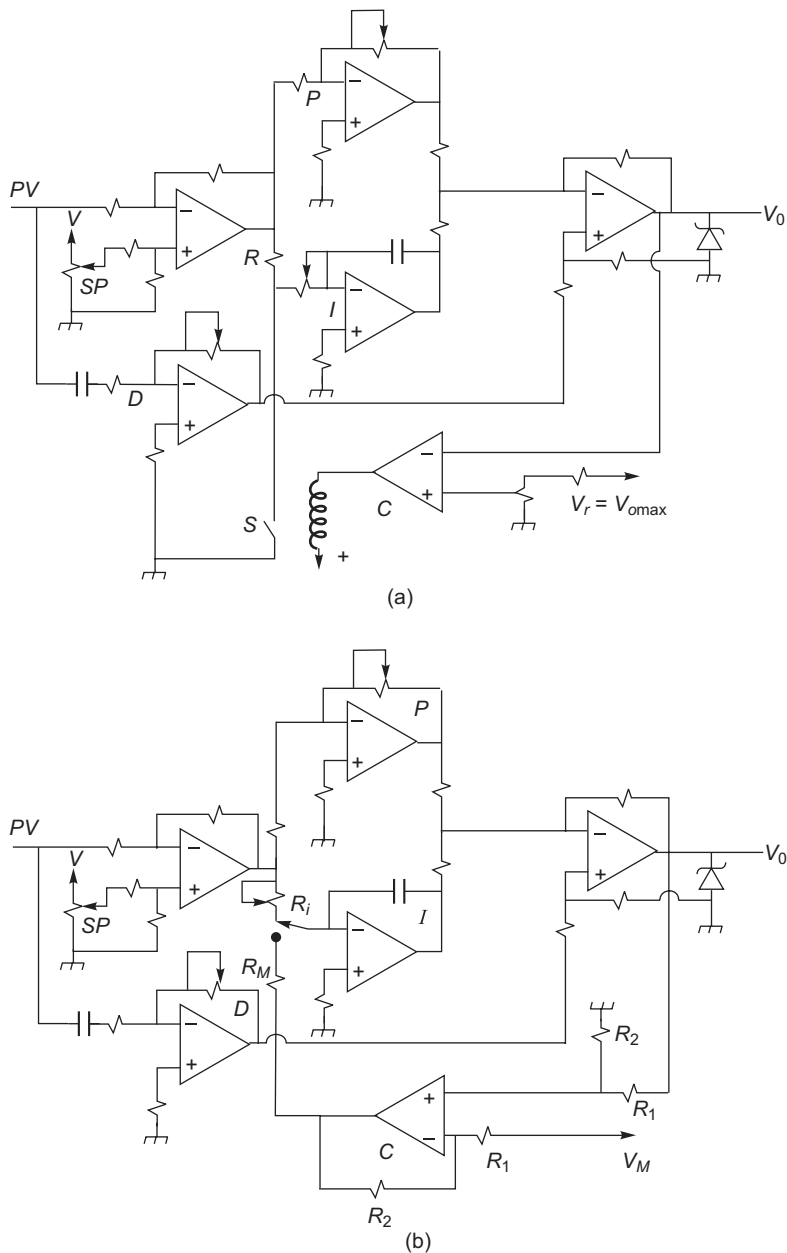


Fig. 6.18 (a) A practical electronic circuit for antireset control
(b) A scheme for bumpless transfer

specifically in digital control systems. One such situation is the transition of the system from manual to auto as specified already. In position algorithm the actual position of the valve must be read for effecting a smooth transfer by adjusting the manual and automatic condition manipulating variables equal. For the velocity or incremental algorithm this smooth transition is easily effected without having to read the variable. Such a smooth transition is known as bumpless transfer.

An analogue electronic scheme for bumpless transfer is shown in Fig. 6.18(b) where the integrator is shown to have two input possibilities — one for automatic operation via R_i and the other for manual via R_M — the latter is connected to the output of a high gain (R_2/R_1) differential amplifier. In the automatic mode, the high gain differential amplifier takes one input from a manual controller and always tracks the automatic mode output. Hence transfer from auto to manual is without any bump. In the manual mode, however, R_M is the integrating resistance and it is chosen as a small one compared to R_i allowing integration to occur much faster. In the manual mode control, the autocontroller tracks the output from the manual controller and the difference $V_M \sim V_0$ is amplified and fed to the integrator to produce an up or down ramp till V_0 equals V_M so that input to the integrator becomes zero. Contributions from P -action and D -action controllers are there but their effects are made negligible by keeping R_2/R_1 very high. At this zero input condition the change over is made so that transfer becomes bumpless.

6.9 MULTIVARIABLE CONTROL SYSTEMS

The multivariable control is the general nomenclature of a process control where the number of control variables is two or more. Such a control system differs from the cascade control system in the adjustment of the controller setting which is independently done.

As there are more than one controlled variables (c) the number of manipulated variables (m) are as important. In fact, it follows logically that there should be as many, or even more, m 's as there are c 's to have a non-interacting type control. This is proved analytically later in the section with the generalized matrix approach. Interaction is defined on the loop basis. Every c_i is paired with an m_i such that control of this m_i does not affect the other m 's and c 's. The system is equivalent to an independent multiloop control system. The general approach to achieve this requires extremely complex adjustment procedures for the controller parameters, and also the number of controllers should be increased $n \times n$ numbers instead of n as discussed subsequently. However, often the task is made easier by the instrument/process engineers by controlling the manipulating variable that has the maximum influence on a given process output. This

is made possible by following a semianalytical approach initially proposed by Bristol (1966) and is often termed as an approximated non-interacting control strategy via the relative static process gain. The method is given in brief here, the logic of which is apparent; the proof, being a little elaborate and involved (following matrix algebra), is omitted. The procedure consists in determining all the possible open-loop gains in terms of the controlled variables with respect to the manipulated variables and then normalizing them. These normalized gains are then arranged in a matrix form and from this matrix array the comparison is made. The greatest value of the relative gain in this matrix is then selected and its associated m and c are chosen for closing loop 1, then the next greatest value is selected and its associated m and c chosen for closing loop 2, and so on. The normalized relative static process gain is obtained by taking the ratio of $(\partial c_j / \partial m_i)$ for a specific m to a specific c with all other m s and c s taken as constants, respectively. Thus

$$\beta_{ji} = \frac{(\partial c_j / \partial m_i)_{m_k} \quad k \neq i}{(\partial c_j / \partial m_i)_{c_p} \quad p \neq j} = \text{constant} \quad (6.30a)$$

The matrix array is thus formed as

$$\begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} \beta_{11} & \beta_{12} & \cdots & \beta_{1w} \\ \beta_{21} & \beta_{22} & \cdots & \beta_{2w} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{n1} & \beta_{n2} & \cdots & \beta_{nw} \end{bmatrix} \begin{bmatrix} m_1 \\ m_2 \\ \vdots \\ m_w \end{bmatrix} \quad (6.30b)$$

Obviously, $w \geq n$ for w number of closed loops. For $w > n$, $(m_i)_i = w - n$ corresponding to the smallest β_{ji} s are kept fixed, i.e., uncontrolled. For example, for 2×3 system, if $\beta_{13} > \beta_{23} > \beta_{21} > \beta_{22} > \beta_{11} > \beta_{12}$, the closed loops for the least interaction would be via $c_1 - m_3$ and $c_2 - m_1$ and m_2 would be uncontrolled and fixed.

It is interesting to note that both $\sum_i^w \beta_{ji}, j = 1, 2, \dots, n$ and $\sum_{j=1}^n \beta_{ji}, i = 1,$

$2, \dots, w$ have a value unity. This property makes the calculation somewhat easier. Besides, the denominator elements of Eq. 6.30 need not be calculated separately, once numerator elements are found out. In fact, arranging the numerator elements in a $(j \times i)$ matrix its complementary is obtained, first taking its inverse and then transposing, to yield the reciprocals of the denominator elements. Thus

$$\mathbf{D}_r = (\mathbf{N} - \mathbf{1})^t$$

is obtained and element by element multiplication would yield the β_{ji} elements of the matrix equation of Eq. (6.30b).

In many situations, however, a single manipulated variable can significantly influence more than one controlled variable. Such coupling or interaction is the most general case.

Although we generalize by stating multiinput/multioutput systems, a number of simple interacting control systems are known in practice. In a particular process if two or more related variables are to be regulated by separate control loops, the situation obviously is an interacting one and unless a specific problem is known *a priori* and proper controller settings for the known ranges are made, the stability of the system as a whole may be in distress. A typical example is pressure and level/flow control in a continuous reactor. Another common example is found when both the top and bottom products of a distillation column are required to be controlled independently. A distillation column or such complicated control process present much more serious interaction problems and instead of raising individual independent issues regarding such control problems, a generalized approach for solving such problems is presented here. Although a rather analytical approach has been proposed for the general problem it must be remembered that often the designer's intuitive capacity and common sense act as better guides.

A typical 2×2 dimensional situation of a general multiinput-multioutput system is given in schematic form in Fig. 6.19(a). This may be generalized for m inputs and n outputs such that

$$r = [r_1, r_2, \dots, r_m]$$

$$\text{and } c = [c_1, c_2, \dots, c_n]$$

The method of treating this multivariable control system is to write the matrix equation

$$c = Tr \quad (6.31a)$$

T is the desired $n \times m$ transfer function matrix which is expressible in terms of G_v , G_f , G_m and the process transfer matrix G_p . Using the symbols of Fig. 6.19(a),

$$G_p = \begin{bmatrix} G_{11} & G_{12} & \cdots & G_{1k} \\ G_{21} & G_{22} & \cdots & G_{2k} \\ \vdots & \vdots & \cdots & \vdots \\ G_{n1} & G_{n2} & \cdots & G_{nk} \end{bmatrix} \quad (6.31b)$$

k being the order of the inputs to the process. If this input variable often referred to as the manipulating variable, is designated as m , and as seen in Fig. 6.19(a), if G_{ci} , G_{vi} can be replaced by a single transfer function, say G_{ai} , we obtain the following matrix equation for the $n \times m$ system

$$\mathbf{c} = \mathbf{G}_p \mathbf{m} \quad (6.32)$$

$$\mathbf{m} = \mathbf{G}_a (\mathbf{r} - \mathbf{G}_m \mathbf{c}) \quad (6.33)$$

For the generalized case considered here one should note that $\mathbf{G}_a$ is a $k \times m$ matrix and $\mathbf{G}_m$ is an $m \times n$ matrix. Combining Eqs (6.32) and (6.33), one gets

$$\mathbf{c} = \mathbf{G}_p \mathbf{G}_a (\mathbf{r} - \mathbf{G}_m \mathbf{c}) \quad (6.34)$$

yielding

$$\mathbf{c} = [\mathbf{1} + \mathbf{G}_p \mathbf{G}_a \mathbf{G}_m]^{-1} \mathbf{G}_p \mathbf{G}_a \mathbf{r} \quad (6.35)$$

Comparing Eqs (6.30) and (6.35) the desired transfer function matrix is obtained in terms of $\mathbf{G}_p$, $\mathbf{G}_a$ and $\mathbf{G}_m$ as

$$\mathbf{T} = [\mathbf{1} + \mathbf{G}_p \mathbf{G}_a \mathbf{G}_m]^{-1} \mathbf{G}_p \mathbf{G}_a \quad (6.36)$$

Given $\mathbf{G}_m$ and $\mathbf{G}_p$, only $\mathbf{G}_a$, i.e., $\mathbf{G}_c$ in $\mathbf{G}_a = \mathbf{G}_c \mathbf{G}_v$ is to be designed so that $\mathbf{T}$ is obtained. Considering now the order of $\mathbf{T}$ and $\mathbf{G}_c$, it will be seen that since for a unique solution to exist, the orders of $\mathbf{T}$ and $\mathbf{G}_c$, should be the same, it is necessary that there should be as many manipulated variables as controlled outputs. If the number of manipulated variables are less than those of controlled outputs then independent control of the outputs is in fact, not achievable. Obviously, the stress is on the separation of the controlled outputs from the inputs or on the conversion of the interacting system into a non-interacting one. Considering now the least number of manipulated variables for achieving independent control, i.e., $k = n$, one can easily achieve the controller function $\mathbf{G}_c$ or $\mathbf{G}_a = \mathbf{G}_c \mathbf{G}_v$ in general from the desired $\mathbf{T}$ and given $\mathbf{G}_p$ along with $\mathbf{G}_m$, by a very simple general procedure. From Eq.(6.36), one gets*

$$\mathbf{G}_a = \mathbf{G}_p^{-1} \mathbf{T} [\mathbf{1} - \mathbf{G}_m \mathbf{T}]^{-1} \quad (6.37)$$

i.e.,

$$\mathbf{G}_c = \mathbf{G}_p^{-1} \mathbf{T} [\mathbf{1} - \mathbf{G}_m \mathbf{T}]^{-1} \mathbf{G}_v^{-1} \quad (6.38)$$

Obviously, primarily one needs to see that $\mathbf{G}_p$, $\mathbf{G}_v$ and $\mathbf{1} - \mathbf{G}_m \mathbf{T}$ are nonsingular for such a $\mathbf{G}_c$ to exist. It is to be further seen that with the known $\mathbf{G}_m$, $\mathbf{G}_p$ and $\mathbf{G}_v$ the desired $\mathbf{T}$ should be such that $\mathbf{G}_c$ has components with

$$N^\circ(s) \leq D^\circ(s) \quad (6.39)$$

*From Eq. (6.36)

$$[\mathbf{1} + \mathbf{G}_p \mathbf{G}_a \mathbf{G}_m] \mathbf{T} = \mathbf{G}_p \mathbf{G}_a$$

$$\text{or } \mathbf{T} + \mathbf{G}_p \mathbf{G}_a \mathbf{G}_m \mathbf{T} = \mathbf{G}_p \mathbf{G}_a$$

$$\text{or } \mathbf{T} = \mathbf{G}_p \mathbf{G}_a [\mathbf{1} - \mathbf{G}_m \mathbf{T}]$$

$$\text{or } \mathbf{G}_a = \mathbf{G}_p^{-1} \mathbf{T} [\mathbf{1} - \mathbf{G}_m \mathbf{T}]^{-1}$$

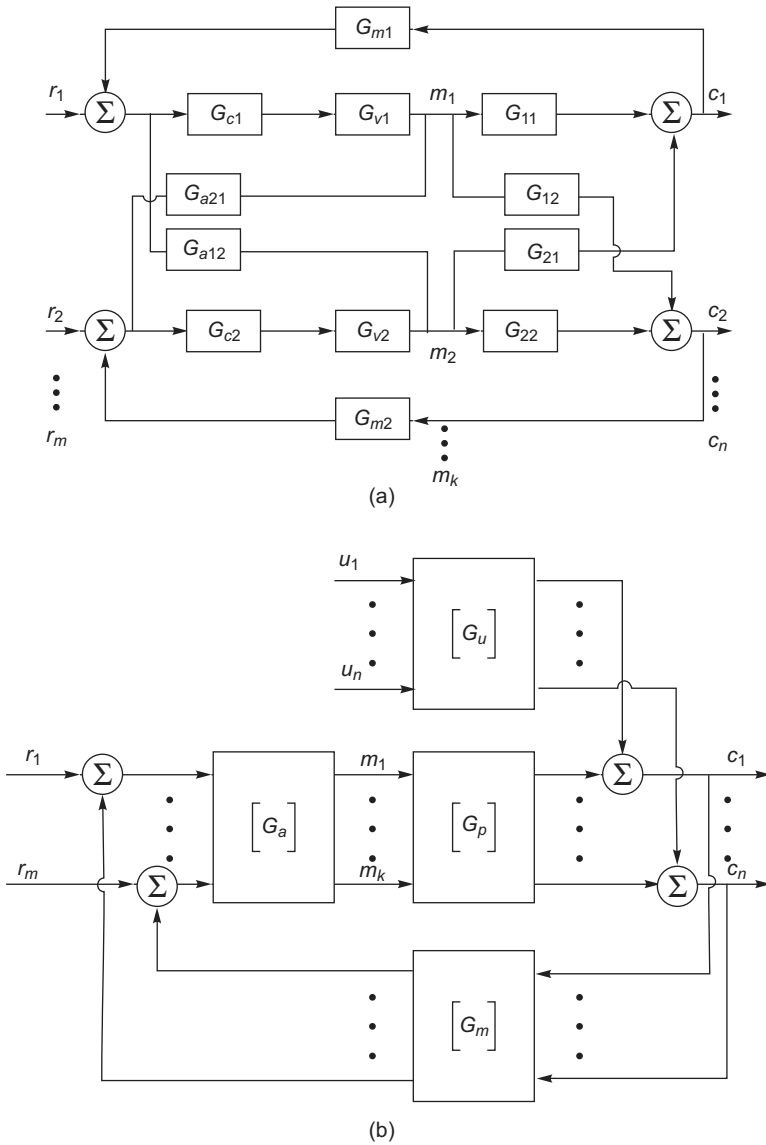


Fig. 6.19 (a) Block representation of a multivariable control system (G 's: block transfer functions; r 's: references; m 's: manipulated variables; c 's: controlled variables) (b) Generalized representation of a multivariable control system ($[G]$'s: block transfer function matrices; r 's: references; m 's: manipulated variables; u 's: upsets; c 's: controlled variables)

where N° and D° specify the degrees of the numerator and denominator polynomials, respectively. Additionally, upsets in the process also need to be considered for regulatory processes.

A generalized representation of such a scheme is shown in Fig. 6.19(b). For this case Eq. (6.32) is modified to

$$\mathbf{c} = \mathbf{G}_p \mathbf{m} + \mathbf{G}_u \mathbf{u} \quad (6.40)$$

$\mathbf{G}_u$ being $n \times q$ matrix, with $\mathbf{u}$ being q -dimensional. Equation (6.40) modifies Eq. (6.35) to

$$\mathbf{c} = [\mathbf{1} + \mathbf{G}_p \mathbf{G}_a \mathbf{G}_m]^{-1} \mathbf{G}_p \mathbf{G}_a \mathbf{r} + [\mathbf{1} + \mathbf{G}_p \mathbf{G}_a \mathbf{G}_m]^{-1} \mathbf{G}_u \mathbf{u} \quad (6.41)$$

$$= \mathbf{T} \mathbf{r} + \mathbf{D} \mathbf{u} \quad (6.42)$$

This equation points to the fact that, in general, response to upsets should also be specified and $\mathbf{T}$ and $\mathbf{D}$ cannot be specified independently. Solving Eqs (6.41) and (6.42) in parts for $\mathbf{G}_c$ with $\mathbf{T}$ and $\mathbf{D}$ respectively, one gets Eq. (6.38) and

$$\mathbf{G}_c = \mathbf{G}_p^{-1} [\mathbf{G}_u - \mathbf{D}] [\mathbf{G}_m \mathbf{D}]^{-1} \mathbf{G}_v^{-1} \quad (6.43)$$

Therefore, equating the right-hand sides of Eqs (6.38) and (6.43), the dependence relation between $\mathbf{T}$ and $\mathbf{D}$ is obtained as

$$\mathbf{D} = (\mathbf{1} - \mathbf{T} \mathbf{G}_m) \mathbf{G}_u \quad (6.44)$$

However, independence can be restored by introducing additional degrees of freedom in the controller schemes or, in other words, by increasing the system complexity.

In the above discussion no attempt has been made to show that the control system can be made to perform under optimum conditions with proper controller design, but on the contrary, a design issue has been raised where non-interacting control can be attempted and independent adjustment of controllers can be made. This, however makes the system more complex but allows the people operating the system to work with greater confidence.

By way of an example of a multivariable process control system we can recall the distillation column. In this, pressure control of the condenser coolant and temperature control of the reboiler steam supply may form a 2×2 interacting control system.

The idea of the non-interacting control presented above may be exemplified now for a 2×2 system. For this, from Eq. (6.36) one must have

$$\mathbf{T} = \begin{bmatrix} t_{11} & 0 \\ 0 & t_{22} \end{bmatrix} \quad (6.45)$$

Let

$$\mathbf{G}_p = \begin{bmatrix} \frac{6}{2s+1} & 2 \\ 4 & \frac{6}{3s+1} \end{bmatrix}, \mathbf{G}_{v1} = \mathbf{G}_{v2} = \frac{1}{s+1}, \mathbf{G}_{m1} = \mathbf{G}_{m2} = 1$$

and

$$\mathbf{G}_c = \begin{bmatrix} G_{c11} & G_{c12} \\ G_{c21} & G_{c22} \end{bmatrix}$$

Writing $\mathbf{G}_p \mathbf{G}_a \mathbf{G}_m = \mathbf{G}_o$, one obtains

$$G_{011} = \frac{6G_{c11}}{(2s+1)(s+1)} \left[1 + \frac{G_{c21}(2s+1)}{3G_{c11}} \right]$$

$$G_{012} = \frac{6G_{c12}}{(2s+1)(s+1)} + \frac{2G_{c22}}{s+1}$$

$$G_{021} = \frac{4G_{c11}}{s+1} + \frac{6G_{c21}}{(3s+1)(s+1)}$$

$$G_{022} = \frac{6G_{c22}}{(3s+1)(s+1)} \left[1 + \frac{2G_{c12}(3s+1)}{3G_{c22}} \right]$$

Thus, for non-interacting control, $G_{012} = G_{021} = 0$, giving

$$\begin{cases} G_{c12} = -\frac{1}{3}G_{c22}(2s+1) \\ G_{c21} = -\frac{2}{3}G_{c11}(3s+1) \end{cases} \quad (6.46)$$

G_{c11} and G_{c22} are, however, to be designed from the single-loop design procedure while the other loop is temporarily ignored. Thus, the loop transfer functions are given as

$$\begin{aligned} T_{11}(s) &= \left[\frac{6G_{c11}}{(s+1)(2s+1)} \right] \bigg/ \left[\frac{(2s+1)(s+1) + 6G_{c11}}{(2s+1)(s+1)} \right] \\ &= \frac{6G_{c11}}{2s^2 + 3s + 6G_{c11} + 1} \end{aligned}$$

$$\text{and } T_{22}(s) = \frac{6G_{c22}}{3s^2 + 4s + 6G_{c22} + 1}$$

The parameters G_{c11} and G_{c22} can be chosen following the loop design via the Bode-plot technique and then G_{c12} and G_{c21} obtained from Eq. (6.46). After completing the design, a stability check is, however, necessary. This may be done by verifying that $[\mathbf{1} + \mathbf{G}_o]$ is non-singular. It should be understood that any adjustment in G_{c11} and G_{c22} would involve adjustments in G_{c21} and G_{c12} also.

Example 2

A heat exchanger has a transfer function between the outlet temperature and process flow given by

$$G_s = \frac{T(s)}{q_p(s)} = \frac{\exp(-3s)}{(15s + 1)(5s + 1)}$$

The process condition also gives that percentage increase in temperature (T) per percentage increase in steam flow (q_s) as 0.8. With a process flow disturbance that can be measured, what is the compensating controller that may be necessary? Assume that $G_p = 1.5/(4s + 1)$.

Solution The system block diagram as drawn with the required FF control in Fig. 6.20. The relevant loop from Fig. 6.20 is redrawn in Fig. 6.21. Here $K_p = 0.8$, and using Eq. (6.19)

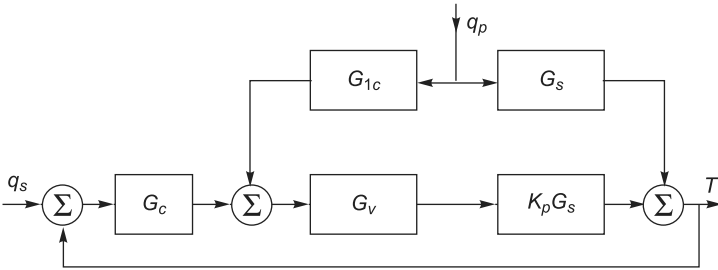


Fig. 6.20 Block diagram of the feedforward control scheme (G 's: block transfer functions; K_p : process gain; q_p : disturbance to process; q_s : reference flow rate; T : temperature output)

$$G_{lc} = K_p / G_v = 0.8(4s + 1) / 1.5$$

Hence

$$K_{lc} = 0.8 / 1.5 = 0.533$$

Assuming a proportional plus derivative action controller to be effective one can have

$$G_{lc} = 0.533(1 + sT_d)$$

such that the derivative time constant $T_d = 4$ s.

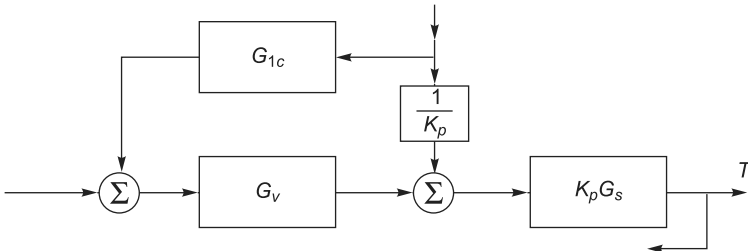


Fig. 6.21 Modified scheme of part of Fig. 6.20

Review Questions

1. What is a ratio control system? Discuss such a control system with a specific process. The final controlled variable may be taken as flow in both the cases.
2. Where would you use a split-range control? Mention some fields of application of such a control.
3. How would you determine the type of process that would require a cascade control and the type that would require feedforward control? What are the basic differences between them?
4. A cascade control system is shown in Fig. Q-6.4. Calculate the maximum gain and critical frequency of the primary controller. Eliminating the inner loop compare these values with the single loop system.
(Hint: Use Bode-plot technique)

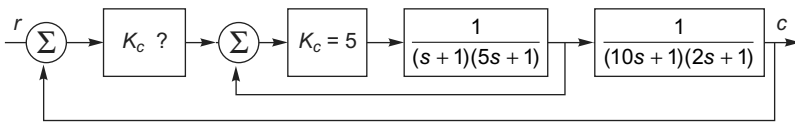


Fig. Q-6.4 Block diagram of a cascade control system

5. The block diagram of a cascade control system is shown in Fig. Q-6.5. The degree of stability of the inner loop is given by $M_p = 2$ (see Ch 4) and the outer loop is to be designed using the Ziegler-Nichols' criteria. Find the gain-bandwidth product of the system.

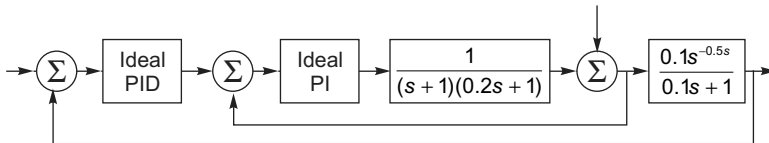


Fig. Q-6.5 Another block representation of a cascade control system

6. An oil-fired furnace is controlled by a cascade control system where the inner loop regulates the flow of oil. The inner process is approximated by a first order one having a lag of 2 sec in which loop measurement lag is 0.5 sec. Assuming the lag to be zero and the outer process lag to be 5 sec, obtain the controller parameters for effectively controlling the process. The outer loop measurement lag is zero. Compare your result with the case when the cascade control is not used.
7. A reactant stream is preheated by steam condensing on the tubes of an exchanger. Draw the block schematic if a very large change in output temperature is not allowed for a flow rate

variation of the reactant by 50 per cent. What steps would you follow to select the settings of the compensating controller? (Hint: Use FF control as shown in Fig. Q-6.7)

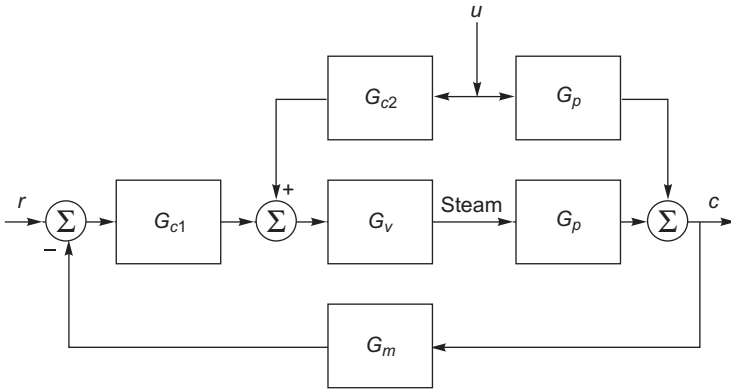


Fig. Q-6.7 Block diagram of feedforward control scheme

8. In the figure of problem 6.7, if the disturbance is at the demand side of the process of transfer function $K_p/(1 + s\tau_p)$ and the actuator transfer function is K_v , what should be the transfer function of the controller for this disturbance to be checked such that it does not affect the process operation effectively? How dynamic compensation can be introduced in such a system?
9. The open loop scheme of a blending system is shown in Fig. Q-6.9 where q is the flow rate and χ is the concentration. The flow rate in the streams are designated q_{m1} and q_{m2} respectively. Close the loop following the noninteracting control strategy suggested by Bristol.

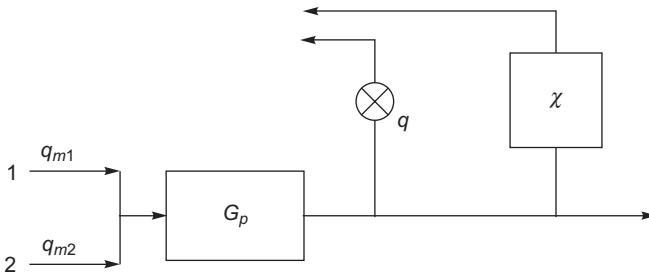


Fig. Q-6.9 Representation of a typical blending system

(Hint: The system equations are $q = q_{m1} + q_{m2}$, $\chi = q_{m1}/(q_{m1} + q_{m2})$. Hence

$$\partial q / \partial q_{m1}|_{(a)} = 1, \text{ and } \partial q / \partial q_{m1}|_{(b)} = 1/\chi \text{ giving the matrix}$$

$$\begin{bmatrix} q \\ \chi \end{bmatrix} = \begin{bmatrix} \chi & 1 - \chi \\ 1 - \chi & \chi \end{bmatrix} \begin{bmatrix} q_{m1} \\ q_{m2} \end{bmatrix}$$

Thus, if $\chi > 0.5$, $q - q_{m1}$ and $\chi - q_{m2}$ are the loops.)

10. What should be the gain and time constant of a load compensator used in a disturbance feedback system when the disturbance is fed back through a part of the process whose transfer function is

$$G_s(s) = 0.4/(1 + 2.2s)?$$

[Hint. Refer to Eq. (6.21)]

$$G_{lc} = -(1 + 2.2s)/0.4 = -2.5(1 + 2.2s)$$

so that gain is 2.5 and time constant 2.2s]

Final Control Elements

7.1 INTRODUCTION

Of the total control gears, the measurement system has been covered in a separate text (Patranabis, D., *Principles of Industrial Instrumentation*, 3e, TMH, New Delhi, (2010)), and controllers have been discussed in an earlier chapter. Final control elements, actuators and control valves remain to be discussed at certain length for a complete understanding of the systems. Final control element is a device that receives the output from the controller to perform a function that serves to take the process to its desired state usually by adjusting certain variable such as fluid flow rate. It is, therefore, pertinent and justified to introduce certain basic aspects of control valve characteristics so that the above function is properly carried out. Choice, selection and sizing of final control elements and their performance aspects are discussed here in a little detail.

There are different types of final control elements, the most popular being the common spring-loaded pneumatic actuator type. Others are electropneumatic actuators, hydraulic actuators, electrical actuators, etc. In the last variety, a digital type known as stepper motor is being used extensively now-a-days.

7.2 THE PNEUMATIC ACTUATOR

A typical pneumatically actuated control valve commonly known just as a control valve is sketched in Fig. 7.1 with proper labelling. It has been shown with a two-seat arrangement and V-port plugs which is quite common in

practice and would be discussed later in some details. A control valve should position its stem and plug in response to the signal received from the controller by striking balance with other active forces in the system such as (i) inertial forces because of the moving mass of the stem-diaphragm parts of the valve, (ii) static frictional forces between the impending motion of the stem with respect to packing, guide bushing, etc., and (iii) thrust forces due to fluid pressure and suspended weight. The signal received by the actuator is sufficient in some cases to deal with these forces, in some others extra power is needed which is availed off from a separate source. This extra power also helps to attain larger stroke length and linearity in operation. The common spring actuator without a separate power supply is often used

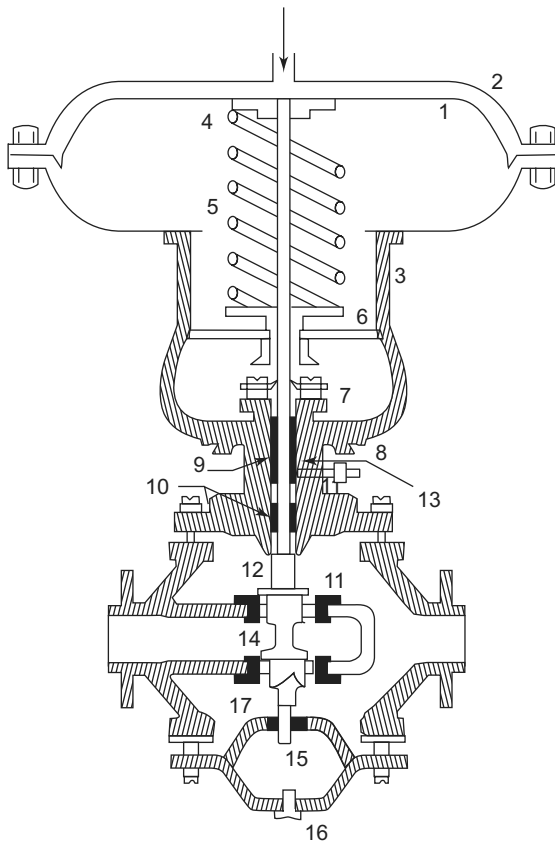


Fig. 7.1 A typical diaphragm operated control valve; 1: diaphragm, 2: diaphragm case, 3: yoke, 4: internal plate, 5: spring, 6: spring flange, 7: stuffing box, 8: valve bonnet, 9: packing, 10: guide bushing, 11: upper seating, 12: trim 13: stem, 14: plug, 15: lower guide bushing, 16: blind header stop, and 17: lower seat ring

in practice. Besides this, such an actuator is used with the support of a positioner, i.e., a separate supply. Other varieties are springless actuators, piston actuator and motor actuator all of which use power amplification systems for combating the additional forces mentioned above. The separate supply pressure may be as large as 7 kg/cm^2 starting from 1.4 kg/cm^2 .

As is shown in Fig. 7.1, the simple spring actuator uses a diaphragm which is moulded into its given shape from a fabric-base rubber with a backing plate (not shown in the Fig. 7.1).

Let m = controller output pressure which also is the input pressure to the actuator that actuates the diaphragm, kg/cm^2

m_0 = input pressure when stroke is zero, kg/cm^2

x = stroke/displacement of stem, cm

α = effective diaphragm area, cm^2

K = spring constant, kg/cm

Pressure m pushes the diaphragm downwards which is balanced against the spring and in equilibrium condition

$$(m - m_0)\alpha = Kx \quad (7.1)$$

The maximum value of m is 1 kg/cm^2 while $m_0 = 0.2 \text{ kg/cm}^2$. Stroke length varies from 0.5 to 7.5 cm in different designs depending on the design of the diaphragm. If mass of the moving parts is M ($= M_f/g$ in kg/cm/sec^2) the inertial force due to this may be obtained and is attempted to be limited by proper design. The natural frequency of the system should be made high so that $\omega_n = \sqrt{K/M}$ should be high. The lowest recommended value of ω_n is about 157 r/sec so that oscillation is prevented for low damping.

To keep the hysteresis down to a limited value the static frictional forces should also be low enough. Friction forces should be such as to keep the hysteresis to less than 1 per cent of the full stem travel. If the operating pressure range is m_{0p} , force for the total stem travel is $m_{0p}\alpha$. Hence, the frictional force should be given by the relation

$$F_f \leq m_{0p} \alpha / 100 \quad (7.2)$$

If the actuator is to make use of the full operating force, the thrust force must be smaller than the zero stroke force, that is,

$$F_f \leq m_0 \alpha \quad (7.3)$$

Also thrust force must be constant so as not to change the stem position with input pressure in a haphazard fashion. Stroke length becomes limited because, otherwise, the spring nonlinearity affects the stroke length as also the diaphragm area (effective) changes with larger stroke length.

Example 1

A diaphragm area (effective) of 600 sq.cm. of an actuator is operated between 0.2 to 1.0 kg/cm². Calculate the allowable friction and thrust forces.

Solution $F_f < (1.0 - 0.2) \times 600/100 \text{ kg} = 4.8 \text{ kg}$, and

$$F_t < 0.2 \times 600 \text{ kg} = 120 \text{ kg}.$$

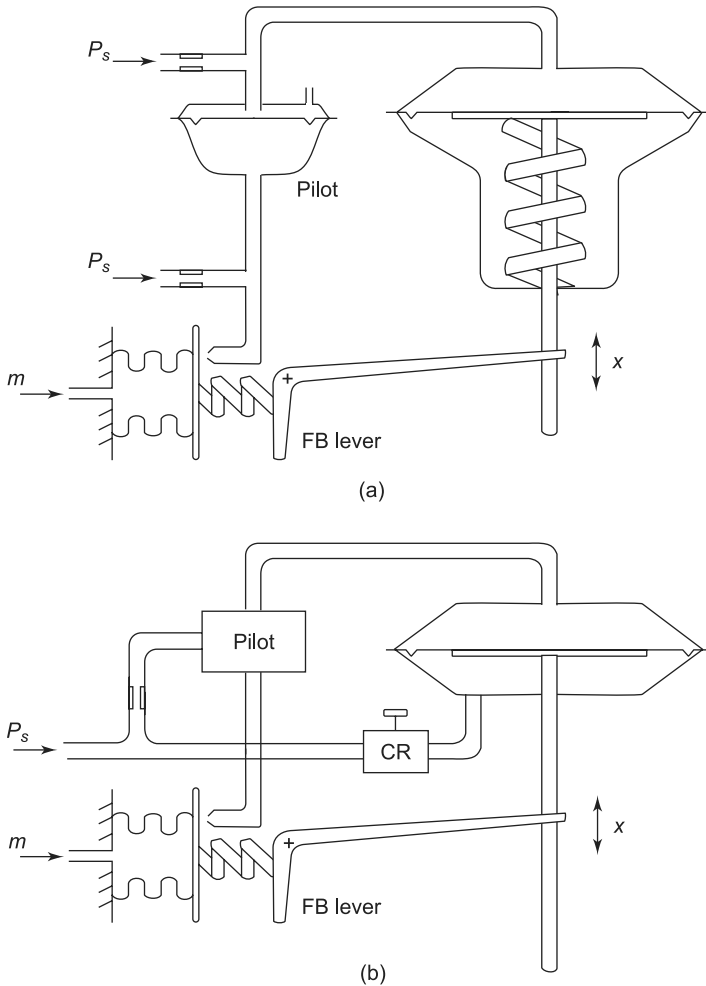


Fig. 7.2 Different types of actuators, (a) spring actuator with positioner, and (b) springless air cushion type.

Figure 7.2(a) to Fig. 7.2(d) show the other 4 types of control valve actuators, all with additional power supply facility for extra power and stroke lengths.

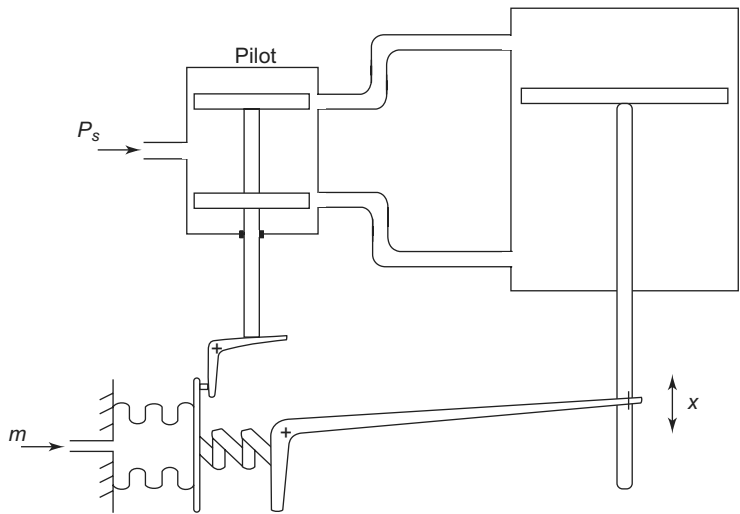
Obviously, such actuators can also accommodate additional friction and thrust forces. For large friction forces and sluggish movement of the stem, positioners have been incorporated with this extra power supply. In Fig. 7.2(a), the actuator operates with a positioner which consists of a pilot valve, a feedback system consisting of a lever, a spring, a flapper-nozzle system operated by the controller pressure through a bellows element. Input m , when increases, causes the flapper to cover the nozzle more, developing a back pressure in it which is amplified by the pilot and passes on to the diaphragm of the spring actuator. The diaphragm goes down compressing the spring. In the process, the feedback lever presses the spring of the feedback system which opposes the input bellows element and the flapper moves away from the nozzle and ultimately attains a balanced position. It is, finally the input pressure which controls the position of the diaphragm but the system is so designed with input to the actuator that the diaphragm and actuator spring characteristics turn out to be of little consequence in the system operation.

In Fig. 7.2(b) a cushion pressure derived from the main supply through a regulator replaces the opposing spring. This gives the advantage that a constant pressure can be maintained in the cushion and the variable characteristics of the spring no longer poses a problem; besides, the value of the pressure can be changed with the help of the regulator as necessary. Usual cushion pressure is 0.6 kg/cm^2 . At equilibrium condition, therefore, the upperside pressure should also be 0.6 kg/cm^2 . This assumes that there is no thrust forces on the actuator stem. If, now, the input pressure increases/decreases, the nozzle back pressure increases/decreases and correspondingly upper side pressure increases/decreases to a high/low value moving the actuator stem downward/upward, however, as the stem attains the new position, upper side pressure would also be 0.6 kg/cm^2 . If now thrust force acts on the actuator stem, upward or downward, the positioner would act to raise the upper side pressure or lower it above or below the cushion pressure which is fixed at 0.6 kg/cm^2 . Thus a thrust force equal to area times the cushion pressure can be counteracted by this type of actuator. This force is much larger than what is counteracted by the spring type design.

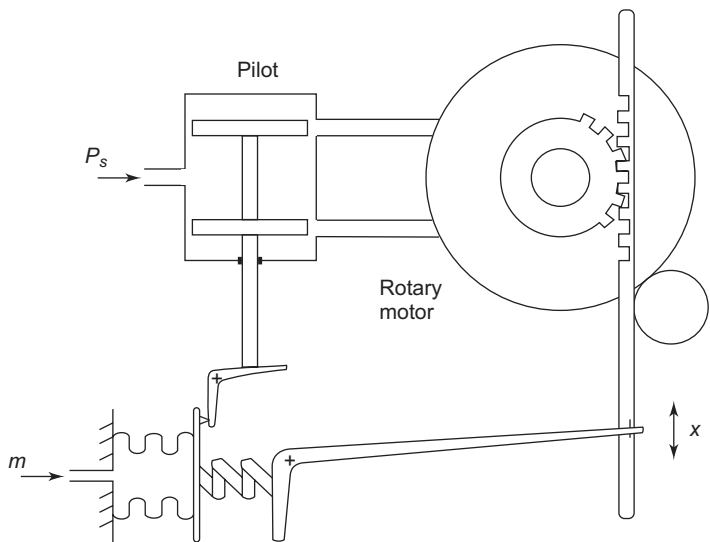
A double-acting piston actuator is shown in Fig. 7.2(c). The piston gives a longer stroke length. The pilot is a spool valve with both sides open to atmosphere. When input pressure m increases, the bellows element pushes the spool of the pilot valve through the crank lever opening the upper side of the piston-actuator to the air supply and exposing the lower side to atmosphere. This brings the piston to the neutral position. In fact, the position of the piston is proportional to the input pressure.

Figure 7.2(d) shows a high power rotary actuator which has its piston-cylinder actuator replaced by an air motor. When one side of the air motor

receives the pilot pressure the other side of it is exhausted. The motor is used to drive the rack through which the feedback mechanism also works to return the pilot piston to the neutral position.



(c)



(d)

Fig. 7.2 (c) piston type, (d) rotary actuator

In Table 7.1 a comparative study of the five actuators discussed above is made on counts of capacity in horse power, stroke length, capability of handling friction and thrust forces.

Table 7.1 Comparison of Actuators

ACTUATOR TYPE	Appr.HR	STROKE	FRICTION FORCE	THRUST FORCE
		cm	in kg (upto)	in kg (upto)
Common spring type	1/20	7.5	15	400
Spring type with positioner	1/7	7.5	150	400
Cushion type	1/7	7.5	150	1,000
Double-acting piston type	1	75	300	2,500
Air-motor rotary type	15	150		50,000

One interesting point to note with the positioner in a control valve is that it produces a subsidiary loop in the control system as is evident from Fig. 7.2. Positioner in the valve brings in another time constant and unless properly selected it is likely to produce dynamic instability. This loop inside the ‘valve-line’ is a sort of cascade loop and it is, therefore, imperative that a positioner would work best when the response of the positioner-operated control valve is much faster than the control valve itself. A well designed pneumatic positioner typically has an ‘open-loop’ gain of 40 and a dead zone of less than 0.2%. It will take not more than 40 sec to position the plug within 0.5% of the lift-span. These data are however for a 3” stroke valve and the time constant of the positioner itself is not more than 35 seconds.

7.2.1 Valve Port-Plug, etc., and Characteristics

There are three different types of port-plug arrangements useful for different situations. These are shown in Fig. 7.3(1, 2, 3) in which plug provision for two seat arrangement is also indicated by dotted lines. The characteristics for the plugs shown in Fig. 7.3 are

- (1) quick-opening,
- (2) linear, and
- (3) equal percentage.

These three types have other names as well. For example, quick-opening type also goes by the names (1a) bevelled disc and (1b) balanced popet; linear has names (2a) throttle plug, (2b) linear contoured, (2c) true linear

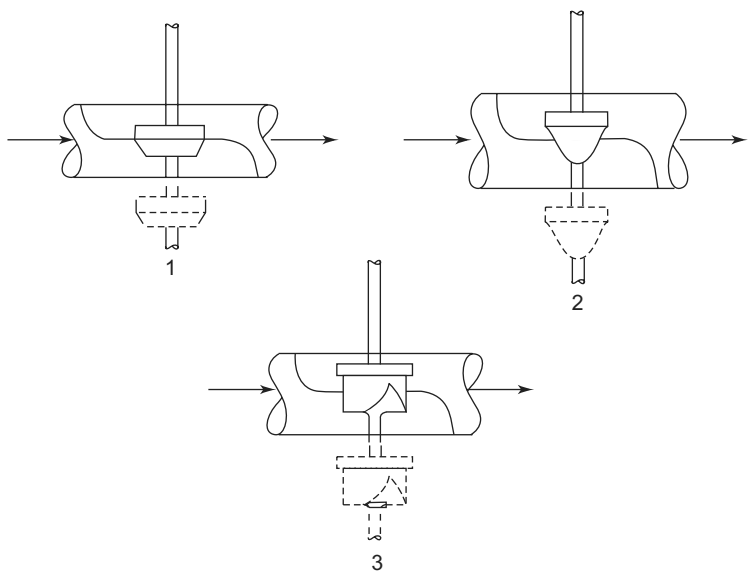


Fig. 7.3 Three different valve plugs; 1: quick opening type, 2: linear, and 3: equal percentage

contoured, and (2d) high flow, and finally, equal percentage has the names (3a) V-port, (3b) equal per cent-V-port and contoured, and, (3c) equal per cent parabolic. The basic characteristics of the three types are shown in Fig. 7.4 in per cent flow versus per cent lift coordinates when pressure drop

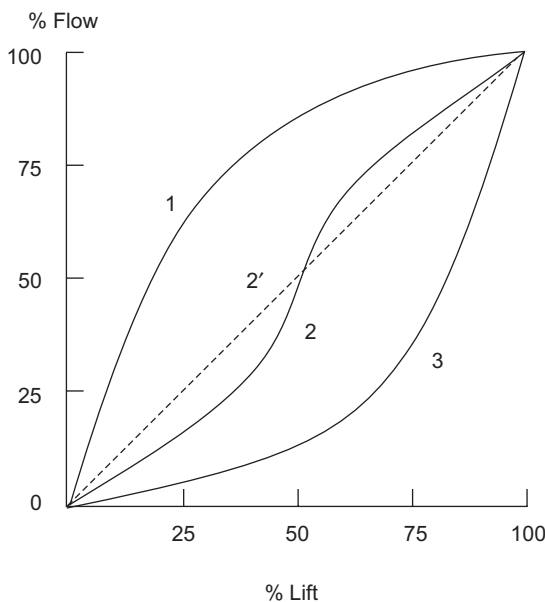


Fig. 7.4 Valve characteristics for the three types of Fig. 7.3

across the valve remains constant for all conditions of flow. This condition is somewhat idealistic and happens rarely.

Quick-opening type has limiting usage as its characteristic is a sort of a two-state one. The other two are used in different cases depending on the types of process and process conditions at different stages. Figure 7.5(a) and (b) show the characteristics of the equal percentage and linear valves for varying ratios of pressure drops across the valves with change in flow rates. Both the sets are curves obtained in per cent stroke versus per cent flow rate, however, in Fig. 7.5(a) abscissa is in log scale. It would be seen that (in Fig. 7.4 and Fig. 7.5(b)) for the linear valve a linear curve is obtained when the plot is in per cent stroke and per cent flow rate coordinates. In fact, equal percentage and linear characteristics are well defined in terms of what is known as valve capacity rating coefficient or simply C_v -factor, which is, therefore defined first.

The C_v -factor or the C_v -number, which is basically the capacity rating coefficient of the control valve, is defined as the number of US gallons of water at 60°F that flows through the valve per minute with a specified opening of the valve (usually wide open) and with a pressure drop of 1 psi across it.

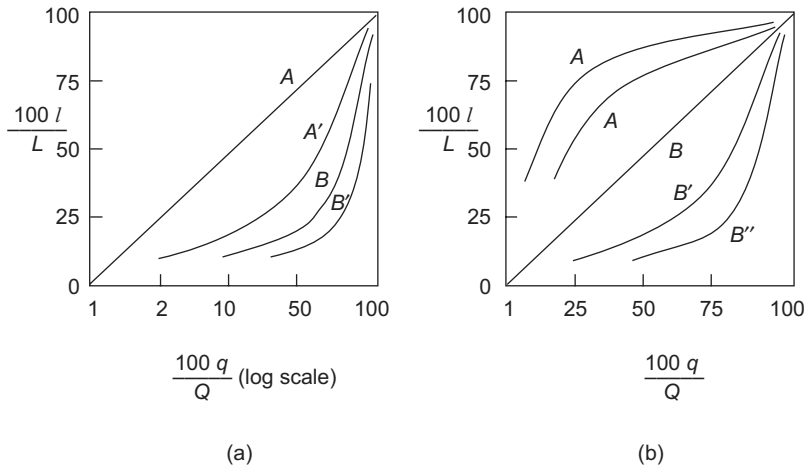


Fig. 7.5 Linear and equal percentage characteristics with variations in drops across the valves: (a) plots in linear-log scale. (b) plots in linear-scale;
 (for $p_r = \frac{\text{pr. drop across the valve at minimum flow rate}}{\text{pr. drop across the valve at maximum flow rate}}$)
 A: Eq. % with $p_r = 1$, A': Eq. % with $p_r = 2.5$, B: linear with $p_r = 1$, B': linear with $p_r = 2.5$, B'': linear with $p_r = 5$

Thus

$$C_v = q/\sqrt{\Delta p/G} \quad (7.4)$$

This value of C_v is generally used in USA and UK. The corresponding capacity rating coefficients in the continent in CGS units and SI units are denoted as K_v and A_v respectively. Their relations are as follows:

$$K_v = 0.856 C_v \quad \text{and} \quad A_v = 24 \times 10^{-6} C_v$$

The defining parameters of C_v , K_v and A_v are tabulated below in Table 7.2 in terms of their units.

Table 7.2 Valve Coefficients

PARAMETER	C_v	K_v	A_v
Fluid	water	water	fluid
Density	62.4 lb/ft ³	1 gm/cm ³	1 kg/m ³
Sp. Gravity	1	1	10 ⁻³
P , Pressure drop	1 psi	1 bar	1 Pa
Flow capacity	USGPM	m ³ /h	m ³ /sec

From Eq. (7.4) the valve area-flow relationship (sizing) is given as

$$q = K\alpha\sqrt{\Delta P/G} \quad (7.5)$$

such that $C_v = K\alpha$. In fact, relation (7.5) is a relation for orifice flow and from equality of C_v and $K\alpha$, the sizing can be done approximately. The relation is only a very approximate one, as although C_v should be proportional to the area of flow, all the practical data should be used for evaluating the area of flow. Usually valves are rated by the size of the pipes connecting the valve to the process. More of it would be discussed later. The C_v -flow relations for gases and steam are different from the one given above. For gas flow, for example the flow rate is

$$q(\text{cu} \cdot \text{ft} \cdot \text{hr}) = 60 C_v \sqrt{\Delta p(P_2/G)} \quad (7.6)$$

where P_2 is the downstream absolute pressure. For steam and vapour flow, flow rate

$$w(\text{lbs/hr}) = 63.3 C_v \sqrt{\Delta p(\rho)} \quad (7.7)$$

where ρ is the fluid density.

Equal percentage and linear valve characteristics are now definable in terms of the C_v -value.

For an equal percentage valve, an equal percentage change of valve capacity will occur for an equal increment of valve stem position (lift). This also means that same percentage change in flow would occur at any stage of operation of the valve when equal change in stem position has

occurred. For a linear valve, for equal change in valve stem position, there is a corresponding equal change in flow or C_v -value.

These would be clear from the following: Let the C_v values corresponding to the per cent valve openings (indicated in suffix numbers) be as, with per cent valve opening $\alpha\%$,

$\alpha\%$	α_{30}	α_{40}	α_{80}	α_{90}
$C_{v\%}$	C_{v30}	C_{v40}	C_{v80}	C_{v90}

then for equal percentage valve

$$\frac{C_{v40} - C_{v30}}{C_{v30}} = \frac{C_{v90} - C_{v80}}{C_{v80}} \quad (7.8)$$

and for linear valve

$$C_{v40} - C_{v30} = C_{v90} - C_{v80} \quad (7.9)$$

The term $(C_{vx} - C_{vy})/C_{vy}$ is, sometimes, referred to as the gain of the valve.

Example 2 *A 1½ in" control valve has a rated C_v of 20. It has equal per cent characteristics.*

Solution At 40 per cent valve opening its $C_v = 1.5$

per cent change in C_v is 66

at 30 per cent valve opening its $C_v = 0.9$

at 90 per cent opening its $C_v = 15.2$

per cent change in $C_v = 64.3$

and, at 80 per cent opening its $C_v = 9.25$

Equation (7.8) has been used to calculate the above per cent changes.

Example 3 *A 1½ in linear valve has a rated $C_v = 34$. At 30 per cent valve opening it has $C_v = 9.6$, at 40 per cent opening $C_v = 13.3$, at 90 per cent it is 29.6 and at 80 per cent it is 25.9.*

Solution Hence, using relation (7.9) it is seen that

$$C_{v(40-30)} = 3.7, \quad \text{and} \quad C_{v(90-80)} = 3.7$$

Another important factor which is often used in valve selection, etc. is the rangeability factor, R , which gives the usable range of the valve and is defined as the ratio of maximum to minimum controllable flow of the valve. It is often specified as, thus,

$$R = C_{v\max}/C_{v\min} \quad (7.10)$$

Using Eq. (7.4), however, it can be expressed as

$$R = q_{\max} \sqrt{\Delta p_1} / (q_{\min} \sqrt{\Delta p_2}) \quad (7.11)$$

where Δp_1 is the pressure drop across the valve corresponding to flow $q_{\min}$ and Δp_2 that corresponding to flow rate $q_{\max}$.

Minimum controllable flow is the flow below which the valve tends to close completely. For defining rangeability as the ratio of maximum to minimum flows one adds the clause that the flow characteristics are maintained within prescribed limits. Flow characteristics are of two types (1) Inherent and (2) Installed. Installed characteristic is largely different from inherent characteristic as the former is the actual operating characteristic and is dependent on the so called loading, that is, pressure drops across different installed equipment in the loop while inherent characteristics is the theoretically obtained one and serves as a guideline for selection and sizing. For the linear valve, the flow through the valve (percent of rated flow Q , q/Q 100) is proportional to the stem travel (percent of rated travel L , l/L 100) under constant pressure difference across the valve. For equal percent characteristics one has

$$\frac{d\left(\frac{q}{Q}\right)}{d\left(\frac{l}{L}\right)} = k \left(\frac{q}{Q}\right) \quad (7.12a)$$

where k is a constant; using instantaneous values q and l , one forms the equation

$$\frac{dq}{dl} = kq \quad (7.12b)$$

which has a solution of the form

$$q = \alpha e^{bl} \quad (7.12c)$$

where a and b are constants to be evaluated.

When l approaches 0, $q = q_n$ = minimum flowrate, then $q_n = a$, giving

$$q = q_n e^{bl} \quad (7.12d)$$

When lift is maximum, i.e. $l = l_x$, $q = q_x$, suffix x standing for maximum, then $q_x = q_n e^{bl_x}$, from which $b = \ln R / l_x$ where $R = q_x / q_n$ = rangeability

Thus the solution for q is

$$q = q_n e^{(\ln R) \frac{l}{l_x}} \quad (7.12e)$$

This can be rearranged by dividing throughout by q_x

$$\frac{q_f}{q_x} = \frac{1}{R} e^{(\ln R) \frac{l_f}{l_x}} \quad (7.12f)$$

where q_f and l_f are flowrate and lift normalized with q_x and l_x respectively. From Eq. (7.12f), one gets

$$\ln q_f + \ln R = l_f \ln R$$

$$\ln q_f + (\ln R) = [l_f - 1] \quad (7.12g)$$

For flow rates q_1 and q_2 and lift l_1 and l_2 Eq. (7.12g) can be modified as

$$\ln \left(\frac{q_{f1}}{q_{f2}} \right) = (l_{f1} - l_{f2}) \ln R \quad (7.12h)$$

Simpler relation has been proposed by some authors to account for installed characteristic as well. One such relation is

$$\frac{q}{Q} = R^{\frac{l}{L}-1} \quad (7.13a)$$

for equal percentage valve, and

$$\frac{q}{Q} = \frac{1}{R} \left[1 + (R - 1) \frac{l}{L} \right] \quad (7.13b)$$

for linear valve. However, these are also acceptable only for constant pressure drop across the valve.

Above relations are valid for constant pressure as mentioned. The relations change as the pressure drop across the valve changes with flowrate; for example, for linear control valve the relation changes to

$$\frac{q}{Q} = \frac{l/L}{[\rho + (1 - \rho)(l/L)^2]^{1/2}} \quad (7.14)$$

where $\rho = p_m/p_0$, p_m being the pressure across the valve at maximum flow and p_0 , that when flow is zero.

Rangeability for linear valve is usually 30 : 1 while for equal percentage type it is more, about 50 : 1 in most commercial designs.

Another term often used by control valve designers and users is turn down. It is the ratio of the normal maximum flow to minimum controllable flow. The ratio lies between 70 and 75 percent.

Before passing onto the topic of valve selection, some more discussion is made on the flow capacity and other limitations of valves used in vapour flow process. It should be remembered that for vapour and gas flows a critical flow condition reaches for

$$(p_u)_{\text{abs}} = 2(p_d)_{\text{abs}} \quad (7.15)$$

where suffixes u and d stand for upstream and downstream. After this there is no further increase in flow. Physically, the turbulence that starts

beyond this pressure drop opposes more flow lines. Hence, the maximum flow occurs for $\Delta p = (1/2)p_u$ in absolute units.

Valves are made in two different trims. Single-seated and double-seated double-plugs are these variations, as has already been mentioned in the beginning. In the double-seated arrangements the plugs move such that one plug moves with the stream, the other against it. This balances out the thrust forces considerably and such a design is, therefore, recommended when a high static pressure is likely to be present. If a single-seated valve is to be used for such a high static pressure a powerful design of the actuator may be necessary. As the plug moves closer to the seat, this thrust (force) increases, the overall action being like spring over the range of travel and thus providing a greater stability as compared to a double-seated design. Double-seated valves lead to cavities which can produce excessive erosion and, have different C_v values—usually lower. Single-seated valves allow tighter closing, particularly with a design using special synthetic inserts. But single seated valve require more power for working against the thrust, as already mentioned. This consideration has led to the designing of piston valves which provide the additional power.

A single-seated valve would be used when the

- (i) pressure drop across the valve is small,
- (ii) line pressure does not vary widely, and
- (iii) complete shut-off is required.

In the eventuality of the failure of the air supply, controller or even motor action of the valve should ensure the safety of the process. Depending on the nature of the process two designs of the trim are possible. If in failures, safety requires that material supply should be stopped, an air-to-open

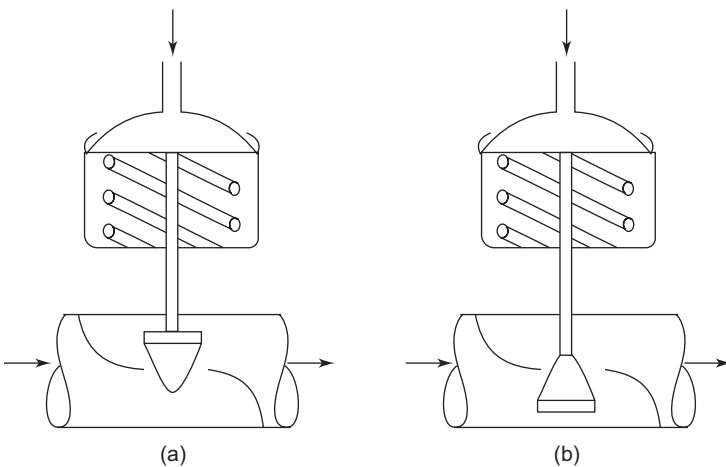


Fig. 7.6 Air-to-close and air-to-open valve designs

design, is preferred, while for the opposing requirement, an air-to-close design is recommended. The schematic arrangements of the two designs are shown in Fig. 7.6.

7.2.2 Materials and Services

Depending on the services, the materials of the valve bodies as well as trims differ. Tables 7.3 and 7.4 list different materials with services and other associated items.

Table 7.3 Materials for Valve Bodies

MATERIAL	SERVICE	PRESSURE	TEMPERATURE	END CONNECTIONS	REMARKS
		kg/cm ²	(max)°C		
Cast iron	Non-corrosive or slight corrosive fluids	8	180	Flange end	-
		15	220	Screw end	
Cast carbon steel, stainless steel	Steam, air, non-corrosive oils, corrosive fluids				-
		10	250	Flange end	
		40	400		
Bronze	Steam, air, water, non-corrosive gases, dilute acids, oils, etc.	20	250	Standard flanged	
Carbon-molybdenum steel	As for carbon steel	20	540	Flange end	
Chrome-molybdenum steel	For resistance to erosion	40	540	Flange end	-
			or more		
Nickel steel	For strong concentrates of reducing chemicals or neutral solutions	10	200	Flanged /screwed end	Not desirable at oxidising conditions

Table 7.4 *Materials for Valve Trims*

<i>MATERIAL</i>	<i>SERVICE</i>	<i>PRESSURE</i> <i>kg/cm²</i>	<i>TEMPERATURE</i> <i>°C</i>	<i>REMARKS</i>
Stainless steel	General	(standard) 20	(standard) 400	Not suitable for slurries or dust bearing gases, easily eroded
Hardened stainless steel (440°C)	General	20	400	Suitable where erosion is expected
Stalite (Co-Chro-W-Fe)	Where abrasive condition is to be resisted	20	800	Suitable for any chemicals with eroding and corroding properties for ordinary materials
Chrome carbide and tungsten carbide	-do-	medium	medium	-
Bronze	For low pressure	5	250	-

In recent years teflon is being used in different forms for packing. Teflon, in combination with other materials like asbestos, is also being used, particularly at different temperatures and pressures. Often the valve body is provided with cooling fins at high temperatures. This practice is very common for temperatures above 250°C.

There are wide variations in plug and port designs of valves as also in their body structures depending on the specialized services for which such designs are considered most suitable. In Table 7.5, eleven such valves are listed with notes on their services.

Table 7.5 *Valve Services*

<i>SL. NO.</i>	<i>TYPES</i>	<i>SERVICES</i>
1	Venturi flow angle valve	For flashing services, where high pressure drop occurs
2	Split-body valve with separable flange	For cases where easy disassembly and economic construction are needed
3	Long sweep-angle valve	For slurries and highly viscous materials
4	Needle valve (from bar-stock body)	For small flows
5	Ball valve with solid ball and cage	For tight shut-off and high range

6	Partial ball body (vee-ball) valve	For difficult-to-handle fluids like paper stock and polymer slurries
7	Butterfly valve with rubber lining	For tight shut-off characteristics
8	Butterfly valve with fish-tail disc design	For less operating torque and improved stability
9	Saunders valve	For slurries and highly viscous fluids
10	Pinch valve	For heavy slurry services including metallic ores, coal and paper stocks
11	Drag valve with multiple disc cage trim	For providing numerous fluid flow paths through the valve

Figure 7.7(a) shows the sketches of first nine of them for giving an idea of their construction.

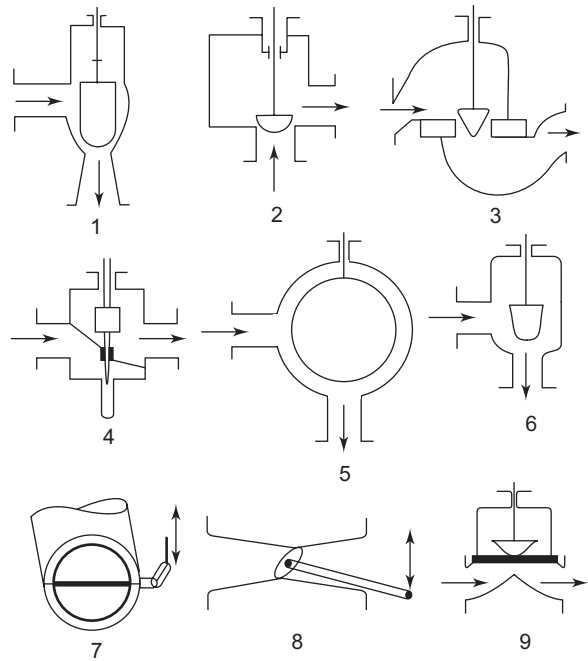


Fig. 7.7 (a) Nine types of valves referred to in Table 7.5 according to services

The commonly used port-plug design which has been considered till this section and beyond is the globe valve. It is a linear design, that is, stem motion is linear with actuating force. There are other designs both linear as well as rotary stem motion type. The chart below gives the commonly

used types in industrial practice which can be correlated with Table 7.5 and Fig. 7.7 (b)

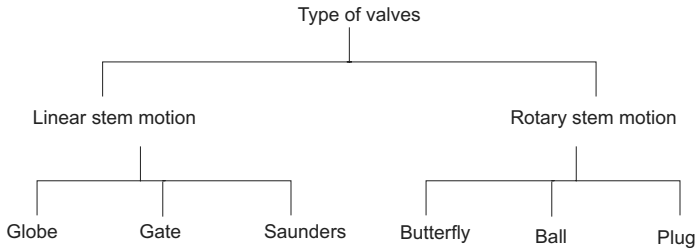


Fig. 7.7 (b) Chart of the types of valves

Each of these valves has separate flow chart symbol. These are given in Fig. 7.7 (c).

It may be noted that there are subclasses in some of these valves which are also given for reference.

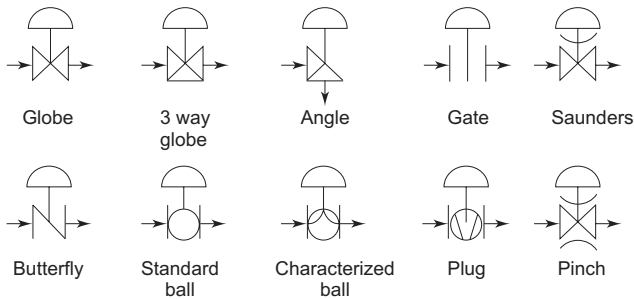


Fig. 7.7 (c) Symbols of the valves of chart of Fig. 7.7 (b)

The flow lift characteristics of some of the important valves of Fig. 7.7(c) are shown in Figs 7.7(d), (e), (f), (g), (h), (i) and (j)

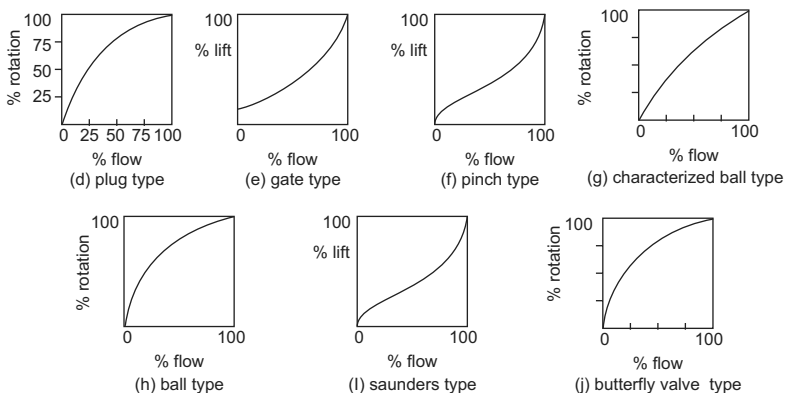


Fig. 7.7 (d), (e), (f), (g), (h), (i), (j)

7.2.3 Valve Sizing and Selection

There are two fundamental questions associated with control valves as used in process control, (1) Why valve-sizing is necessary at all? (2) Which type of control valve should go with which type of process? This section is devoted to answer these questions briefly.

Valve sizing is necessary because (i) for too small a valve, required flow will not be there; and, for too large a valve, it would be too expensive although it would allow more flow, (ii) an undersized valve will not deliver the full flow rate and thus there will be sharp narrowing down of the controllable flow range whereas for an oversized valve throttling will be near the closed position and full control range of the valve will not be utilized, (iii) for an oversized valve when the plug throttles very close to the seat, high fluid velocity will occur which causes erosive damage. From these considerations the ideal valve will be one that will function between 40 per cent and 70 per cent of the full operating range so that for maximum flow it is not wide open and for minimum flow it is not closing down too near to its seated position.

For liquids with low flash points valve sizing becomes important on other counts. When in the downstream side, pressure suddenly drops, such liquids may vaporize and expand. To accomodate this expansion, experience shows that, one valve size larger than that calculated would be a good choice and the downstream piping should be expanded as required.

Cavitation is another undesirable phenomenon which can occur in control valves in liquid service. When a portion of the liquid is transformed into its vapour during rapid acceleration of the fluids as it passes through the port of the valve and then sudden collapse of these vapour bubbles downstream can cause a very high localized pressure (10^4 kg/cm^2) which in turn, can cause wear of the valve trim, body and outlet piping. Associated with this high pressure, severe noise and vibration can also be generated.

If the static pressure at the vena contracta reaches the vapour pressure of the fluid, there is good chance of cavitation. At this point vapour bubbles will be formed which would collapse at the downstream. The critical flow factor C_f , defined as

$$C_f = \frac{C_v \text{ at the condition liquid vaporizes at vena contracta, i.e., critical condition}}{C_v \text{ at normal condition}}$$

has something to do with cavitation. The valves with low C_f values are likely to cavitate. The pressure drop condition $\Delta P_{\text{critical}}$ at which full cavitation occurs is defined as

$$\Delta P_{\text{critical}} = C_f^2 \Delta P \quad (7.16)$$

To avoid cavitation, pressure drop across the valve should be reduced to be made less than $\Delta P_{\text{critical}}$. This can be done by increasing the inlet pressure

or selecting a valve with larger C_f value. A V -port gives better result in comparison with a contoured plug. Another costly and cumbersome remedy is to cascade (install in series) two control valves whose combined C_f value is estimated as

$$C_{f(\text{total})} = \sqrt{C_{f(\text{individual})}} \quad (7.17)$$

From Fig. 7.5 it is clear that the valve characteristics change with the pressure drop across the valve with change in flow rate. In general, most of the drop occurs across the valve at low flow rates and at high flow rates the drop across the valve becomes least and the drop is distributed through the rest of the system. Selection of the valve is very much dependent on where the drop is maximum and at what flow conditions. This in turn is dependent on the type of the process. The two important considerations for analysis are

- (i) Dynamic response of the process;
- (ii) Combination of transmitter and primary device (such as linear or squared, etc.), indicating the signal state.

It is extremely difficult to have the dynamic analysis of individual control loops ready at hand all the time and hence a guideline is usually provided by the manufacturers basing upon their field experiences.

Table 7.6 is provided for the selection based on such experiences.

Table 7.6 Valve Selection

CONTROL SYSTEM	SIGNAL STATE	$\frac{\Delta p_{\max}}{\Delta p_{\min}}$	$\Delta p_{\max \text{flow}}$	TYPE OF VALVE
Pressure I	-	≈ 1	Nearly const. at all flow	Linear
Pressure II	-	2	50% of that at low flow	Equal per cent
Temperature	-	< 2	-	Equal per cent
Level	-	> 2.5	-	Equal per cent
Level	-	< 2.5	-	Linear
Flow I	Linear with flow rate	> 2.5	$< 40\%$ of System pressure	Equal per cent
Flow II	-do-	< 2.5	$> 40\%$ of system pressure	Linear
Flow III	Linear with Δp , i.e., square with flowrate	> 5	$< 20\%$ of system pressure	Equal per cent

Flow IV	-do-	< 5	> 20% of sys- tem pressure	Linear
Analysis, such as pH	-	> 2	-	Equal per cent
-do-	-	< 2	-	Linear

As the characteristics of the control valves show, with flow rate control valve gain changes. But for stability over the entire control range, control loops require that the flow be manipulated in uniform proportion to controller output. However, with increasing flow, frictional losses in the pipes and fittings increase and the available pressure drop across the valve decreases and hence, to maintain the uniform proportionality as desired, an equal percentage characteristic is often required. For a constant pressure drop, however, a linear characteristic is preferred. In a temperature control system, however, an equal percentage characteristic is always used, and, this is an exception in that sense.

Equal percentage valves have high gain at high flow rates and low gain at low flow rates which means that if ‘misused’ this can cause instability at high flow rates. This is tried to be avoided by adjusting the controller proportional band but then it leads to overdamping at low flow rates, causing sluggishness in the system. If a linear valve is ‘misused’ the effect is just opposite.

It is thus clear that the pressure drop across the valve or in the lines and fittings is very important for determining which valve to choose for a particular process. Different analytical or semianalytical approaches are there for estimating the line drop and hence the drop across the control valve in no load and full load conditions so that this estimation helps to design the control system and choose the appropriate valve. The effect of pressure drop on a control valve is not fully understood. This can be appreciated, however, if a graphical approach as shown in Fig. 7.8 is used from which it would be easier to do the design job. The technique is known as hydraulic gradient method and is normally used for sizing pumps and pipelines as well. In this method at maximum flow rate and minimum flow rate one requires to (i) plot the pressure and pressure drop from left to right (or otherwise) stopping at the valve, and (ii) plot the pressure and pressure rise from right to left (or as the case may be) stopping at the valve for a typical system. The difference of pressure at this point between these operations is the pressure drop which the valve must maintain at maximum/minimum flows. The technique is illustrated in Fig. 7.8 where the system scheme is shown at the top and the corresponding pressure and pressure drops (rises) are shown at the bottom diagram. In the plot of the diagram

the pump output pressure will be higher at low flow rate and low at high flow rate. At low flow rates, because of the low fluid velocities, pressure loss through the pipes and pipe fittings would be less than that at higher flow rates making pressure available at the inlet of the control valve high at low flow rate conditions. Also, for the same reason, pressure rise at the outlet of the control valve from that required at the final outlet is less at low flow rate conditions so that a large pressure drop needs be maintained across the valve at low flow rate. Reasoning similarly, it is obtained that the drop Δp across the control valve at high flow conditions is much less.

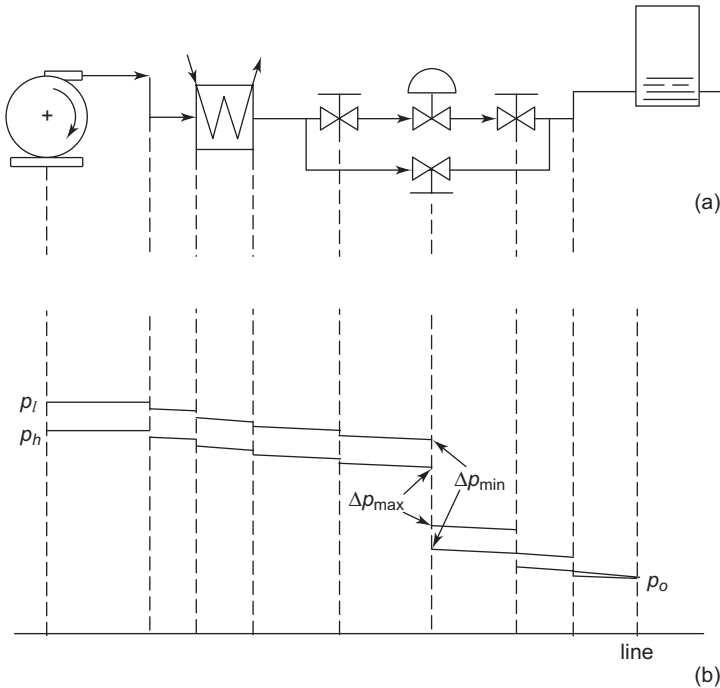


Fig. 7.8 Evaluation of pressure drop across the valve by the hydraulic gradient method; p_o : system output pressure, p_l : pump output pressure at low flow rate, p_h : pump output pressure at high flow rate

The pump characteristic is shown in Fig. 7.9. As the flow rate increases, output pressure (of the pump) decreases following a certain typical characteristics which is shown in the figure in terms of normalized parameters, flow q/Q and pressure p/P . For a centrifugal pump, as in this case, the relation between the parameters is given by $p/P = 1 - \beta(q/Q)^2$, where β is a capacity constant whose value is given by $0.2 < \beta < 0.4$.

Let the minimum and maximum flow rates be given by q_l and q_h respectively and the corresponding pump output pressures be p_l and p_h . Let the pipe size be d , total pipelength, l_p , pressure drop per unit length p_{ul} and p_{uh} at low and high flow rates respectively, fitting equivalent length of pipe, l_f , required outlet pressure both at maximum and minimum pressures be same as p_0 . Then, for maximum flow condition:

Total pressure loss at maximum flow rate in the pipe = $(l_p + l_f)p_{uh}$.

Hence drop across the valve at maximum flow rate $p_{\max}$ is given by

$$\Delta p_{\max} = p_h - p_0 - (l_p + l_f)p_{uh} \quad (7.18)$$

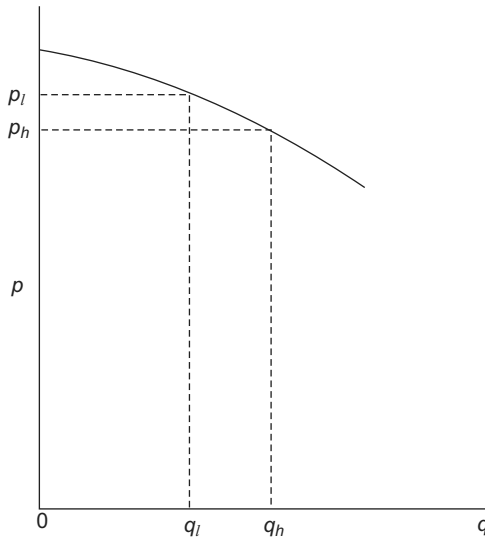


Fig. 7.9 The pump characteristics; q_l : low flow rate; q_h : high flow rate

Similarly for minimum flow rate the drop which must be absorbed by the control valve is

$$\Delta p_{\min} = p_l - p_0 - (l_p + l_f)p_{ul} \quad (7.19)$$

Eq. (7.18) and (7.19) allow one to calculate the C_v -factors of the valve following

$$C_{v_{\max\text{flow}}} = q_h / \sqrt{\Delta p_{\max} / G} \quad (7.20a)$$

and

$$C_{v_{\min\text{flow}}} = q_l / \sqrt{\Delta p_{\min} / G} \quad (7.20b)$$

Some typical numerical values are considered to demonstrate the order of change of pressure drop across the valve with change in flow rates.

Example 4

Following data are available for a process:

Solution

$q_h = 200$ gpm, $q_l = 50$ gpm, Pump $p = 100$ psi, $\beta = 0.3$,
with extrapolated

$Q = 500$ gpm, hence

$p_h = 100(1 - 0.3 \times 0.16) = 95$ psi, and

$p_l = 100(1 - 0.3 \times 0.01) = 99.7$ psi,

$l_p = 300$ ft and $l_f = 100$ ft, $p_{uh} = 0.05$ psi/ft and $p_{ul} = 0.002$ psi/ft

$p_0 = 40$ psi.

Then,

$\Delta p_{\max} = 95 - 40 - 400 \times 0.05 = 35$ psi, and

$\Delta p_{\min} = 99.7 - 40 - 400 \times 0.002 = 58.9$ psi

Correspondingly, with specific gravity $G = 1$,

$C_{v\max\text{flow}} = 200/\sqrt{35} = 35$, and

$C_{v\min\text{flow}} = 50/\sqrt{58.9} \approx 7$,

the rangeability is, therefore $R = 35/7 = 5$

We have assumed an empirical lumped line drop dependent on flow rate. For laminar flow the relationship is given by Darcy-Weisbach equation

$$\Delta p_{\text{line}} = k_1 \rho l v q / (\pi g d^4) = k_2 q \quad (7.21)$$

where ρ = density of the liquid, l = pipe length, v = kinetic viscosity, d = pipe diameter, g = gravitational acceleration and k_1 and k_2 are two constants. For turbulent flow with Reynold's number $Re > 2 \times 10^3$, this relation changes to

$$\Delta p_{\text{line}(t)} = 8 \zeta l \rho q^2 / (\pi^2 g d^5) = k_3 q^2 \quad (7.22)$$

where ζ is the dimensionless frictional factor assumed to be constant.

For selecting a valve certain idealized conditions are assumed initially in the overall system. For the valve these are

- (i) actuator diaphragm area remains constant over the range of operation,
- (ii) valve characteristics match the process characteristics,
- (iii) area of valve opening is proportional to flow through it,
- (iv) spring is linear and
- (v) thrust forces are negligible.

For low load changes and high valve gain as also for a large process tolerance, these conditions are either approximately true or, in effect, do not matter. With low valve gain and large load disturbances, the effects are definitely deteriorative.

We have considered valve sizing in terms of sizing formulae, i.e., C_v values. This is actually a preliminary selection and between C_v and actual size (diameter) a relation of the type

$$C_v = k'd^2 \tag{7.23}$$

exists particularly for single seated valves. However, it is better to consult manufacturers' slide rules or charts which provide relations between stroke and C_v with diameter as a parameter. The checkpoint is always followed while using this chart which states that a properly sized control valve will generally be one line size smaller than the line in which it is to be installed.

The calculated C_v factor is required to be corrected for viscosity. Since viscosity is the frictional resistance of the molecules of a fluid to internal as well as external motions, it seriously affects the C_v value. In fact a viscosity index is obtained in terms of the value of viscosity and the valve size (diameter) and is given by the relation

$$I_\mu = k''q/(d\mu) \tag{7.24}$$

where μ is the viscosity and k'' is any constant dependent on the unit of viscosity. For example; if μ is in Saybolt-sec $k'' = 14.7 \times 10^3$. Correction factors for I_μ are available in a chart or graph from the manufacturers and these can be applied for actual sizing. Fig. 7.10 shows the approximate curve between I_μ and the correction factor F_μ . The corrected size of the valve is obtained by multiplying the calculated size by the correction factor.

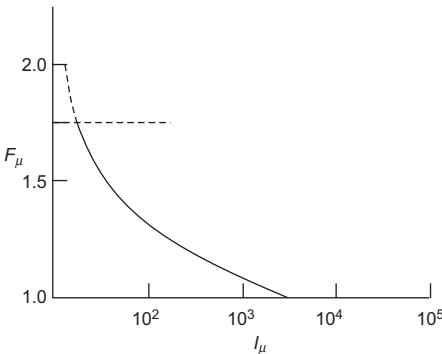


Fig. 7.10 Curve for obtaining the viscosity correction factor

It must be remembered that the index I_μ , if available in terms of C_v value, can be more conveniently adopted for calculation of corrected C_v

with the correction factor, also available in terms of this new index. In the following $C_v - I_\mu$ relations for some typical cases are given.

$$I_{\mu s1} = 94.8q/(\sqrt{C_v} \mu_{\text{stokes}}), \quad \text{for single-seated case} \quad (7.25a)$$

$$I_{\mu d1} = 67q/(\sqrt{C_v} \mu_{\text{stokes}}), \quad \text{for double-seated case} \quad (7.25b)$$

$$I_{\mu s2} = 44,100q/(\sqrt{C_v} \mu_{\text{saybolt-sec}}), \quad \text{for single-seated case} \quad (7.25c)$$

$$I_{\mu d2} = 31,200q/(\sqrt{C_v} \mu_{\text{saybolt-sec}}), \quad \text{for double-seated case} \quad (7.25d)$$

The nature of curve between the correction factor F_μ and this new I_μ , however, remains the same.

Usually, viscosity is obtained in poise, it is easily converted to Stokes by the relation

$$\text{poise}/G = \text{Stokes} \quad (7.26)$$

Example 5

A double-seated valve is used in a system for a liquid flowing at a maximum rate of 10 gpm, its specific gravity being 0.9 and viscosity 36,000 cp. The drop across the valve 1 psi, obtain the valve size.

Solution The value of C_v is calculated as

$$C_v = 10\sqrt{0.9/1} = 9.4$$

Viscosity in Stokes is $360/0.9 = 400$

$$I_\mu = 67q/(\sqrt{C_v} \mu_s) = 67 \times 10/(\sqrt{9.4} \times 400) \approx 0.54$$

From chart, $F_\mu = 20$, hence the corrected $C_v = 20 \times 9.4 = 188$

From manufacturers' slide rule or chart, a valve of 3" diameter is chosen.

Before taking up further problems, a typical sizing chart is shown in Fig. 7.11. It is a set of curves plotted in C_v (given in log scale) versus per cent

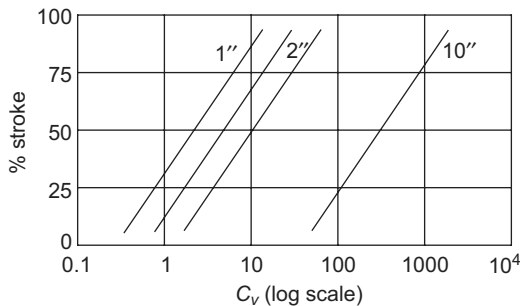


Fig. 7.11 The valve sizing curves

stroke with valve size as the parameter. The calculated C_v is corrected for viscosity and this corrected C_v is read at the abscissa. For ordinate the maximum stroke is considered which, as mentioned earlier, is usually taken as 75 per cent (slight tolerance may be allowed here). Corresponding size curve is thus obtained. But again, as specified earlier, a check point is followed and a size smaller by one line size is actually selected.

Example 6

A valve discharges from a tank with a head of 20 ft of water to a tank with a head of 10 ft and maximum flowrate is 100 gpm. What should be the size of the valve?

Solution The pressure difference is

$$\begin{aligned}\Delta p &= (20 - 10) \times 62.4/144 \text{ psi} \\ &= 4.33 \text{ psi}\end{aligned}$$

$$\begin{aligned}\text{Hence, } C_v &= 120/\sqrt{4.33} \text{ (since } G = 1) \\ &= 57.7\end{aligned}$$

From the chart, the smallest size of the valve is 2½ in.

It should be stressed that the C_v is always calculated for maximum flow rate and that is the reason that while reading the chart 75 per cent stroke point is considered.

Example 7

In an heat-exchanger steam is used at a normal pressure of 40 psig and water at 50°F enters it at a maximum flowrate of 50 gpm and pressure 65 psig and comes out at 200° F, $p_u/p_d \geq 2$ and $p_u = 100$ psig. Determine the size of the control valve needed for steam flow.

Solution Btu/min added to water

$$\begin{aligned}&= \Delta t \times q_{\max} \times 8.3 \text{ (where 8.3 lb/gal is the multiplier used)} \\ &= (200 - 50) \times 50 \times 8.3 \\ &= 6250 \text{ Btu/min.} \\ &= 3735000 \text{ Btu/hr}\end{aligned}$$

When condensed steam yields approximately 1000 Btu/lb, the required steam

$$w = 3735000/1000 = 3735 \text{ lb./hr}$$

The drop across the valve is large enough to make $p_u/p_d \geq 2$; therefore, the critical condition is reached and hence $\Delta p = p/2$ is used. Thus

$$\Delta p = (40 + 15)/2 \text{ psia} = 55/2 \text{ psi}$$

From steam chart/table density ρ is found out for downstream pressure. Its value is 0.065 (reciprocal of the specific volume),
Hence,

$$C_v = (w/63.3) \cdot 1/\sqrt{(55/2)(0.065))} = 3735/(63.3)\sqrt{55 \times 0.065/2}$$

$$\approx 45$$

From manufacturer's chart the valve size is 2".

7.2.4 Positioners and Power Cylinders

It has been mentioned earlier that often an actuator is supported by an extra supply pressure to position the valve stem as required. Such an arrangement is known as positioner. Generally, the actuator is required to supply sufficient force to position the valve accurately and overcome any opposition that flowing conditions apply to the plug. If controller output is small, the change in force available to accurately position the valve stem may also be too small. A positioner always helps in such situations. In fact, a positioner is used to overcome

- (i) stem friction,
- (ii) slow response of large capacity motors used with long transmission lines,
- (iii) line pressure changes that tend to reposition the plug, and
- (iv) plunger friction due to highly viscous fluids, gumming and/or sedimentation. Another important force is the spring force (when a spring is used).

The nonlinear spring force is often considered to be linearized for convenience but that is hardly ever justified. Also, spring force does not take into consideration the friction or thrust forces F . It will now be in context to list the specific situations where a positioner needs be used.

- (1) For valves which are expected to respond to very small control pressure changes, say 0.006 kg/cm^2 .
- (2) For single-seated valves to be used at high fluid velocities and low valve gain, controller output pressure is not sufficient to operate the valve.
- (3) For valves with large frictional forces due to tighter gland packing as required for certain types of fluids; self-lubricating teflon may ease the situation to a certain extent.
- (4) For split-range valves, where full opening to closing of one valve is to be made for one part of the control pressure, and similarly, closing to full opening of another valve is to be made by another part of the same control pressure, and so on.
- (5) For extra stroking speed and improved frequency response.

There has also been a suggestion as to where a positioner need be used in relation to flow rate and pressure drop ratios. If Q_x is the maximum flow rate the control valve may be required to handle, Q_y is the minimum flow rate required to sustain the variable under any conditions of operation, and the valve characteristic coefficient is defined as the ratio of the pressure drop across the valve when it is wide open to the pressure drop when it is closed and is denoted by α , then, if $\alpha < 1/2500(Q_x/Q_y)^2$ or $\alpha < 0.1$, the control valve may need positioner, specifically if, the operation has to be within 2 per cent of the total lift.

Unless absolutely necessary, a positioner should not be used because it increases the cost and complexity in control valve assembly, supply requirement and sensitivity to the extent of causing instability at times.

A positioner that may be used where a large power is required is often referred to as a power cylinder. Outputs from power cylinders are used for controlling the opening of the line where a large static pressure may exist. A typical arrangement for a power cylinder is shown in Fig. 7.12. Input pressure

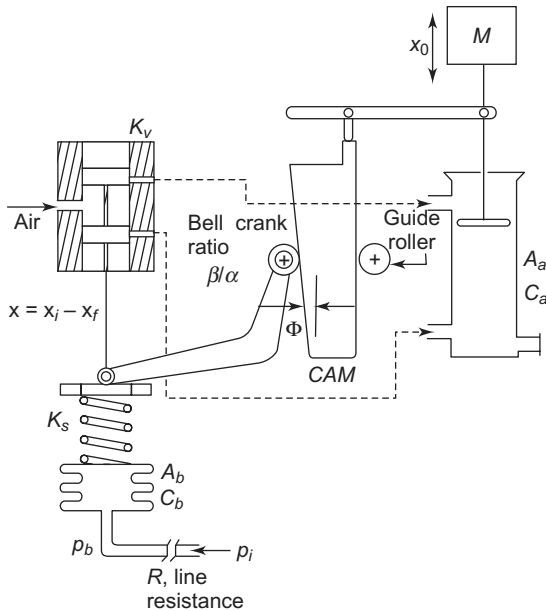


Fig. 7.12 Schematic of a power cylinder

acting through the bellows element and spring operates the spool valve. The spool valve actuates the power cylinder and the load. The load movement is, in turn, responsible for a feedback in the system via the command crank. This in turn attempts to restore the spool valve in a neutral position. The load position, x_0 , will then be proportional to the input pressure p_i . However, a dynamic change in p_i will not keep this proportionality

essentially linear. A brief analysis obtains the transfer function from which the dynamic characteristics may be studied.

The system equations are listed as follows:

(i) At bellows element: flow rate q_b is

$$q_b = (p_i - p_b)/R \quad (7.27)$$

and, pressure change p_b is

$$p_b = q_b/sC_b \quad (7.28)$$

(ii) At spool valve: initial displacement x_i is

$$x_i = p_b A_b / K_s \quad (7.29)$$

With a feedback displacement x_f , the effective displacement x of the spool is given by

$$x = x_i - x_f \quad (7.30)$$

(iii) At cylinder: air flow rate q_a is

$$q_a = x K_v \quad (7.31)$$

and pressure change p_a is given by

$$p_a = q_a / (s C_a) \quad (7.32)$$

The force balance equation is now written as

$$p_a A_a - F_f = M \ddot{x}_0 = M s^2 x_0 \quad (7.33)$$

where the opposing force F_f is given by

$$F_f = \dot{x}_0 B \quad (7.34)$$

B being the coefficient of damping. As the frictional effect is considered as lumped and with the geometry of the arrangement as shown, the displacement of the spool due to feedback is

$$x_f = x_0 \tan \phi \cdot (\beta/\alpha) \quad (7.35)$$

where β/α is the bellcrank ratio.

Hence the transfer function is easily derived as

$$\frac{x_0(s)}{p_i(s)} = \frac{\frac{\alpha A_b}{K_s \beta \tan \phi}}{(1 + s C_b R) \left(\frac{M s}{J} + \frac{B}{J} s^2 + 1 \right)} \quad (7.36)$$

where

$$J = \frac{A_a K_v \beta \tan \phi}{\alpha C_a} \quad (7.37)$$

The block representation of the power cylinder is shown in Fig. 7.13.

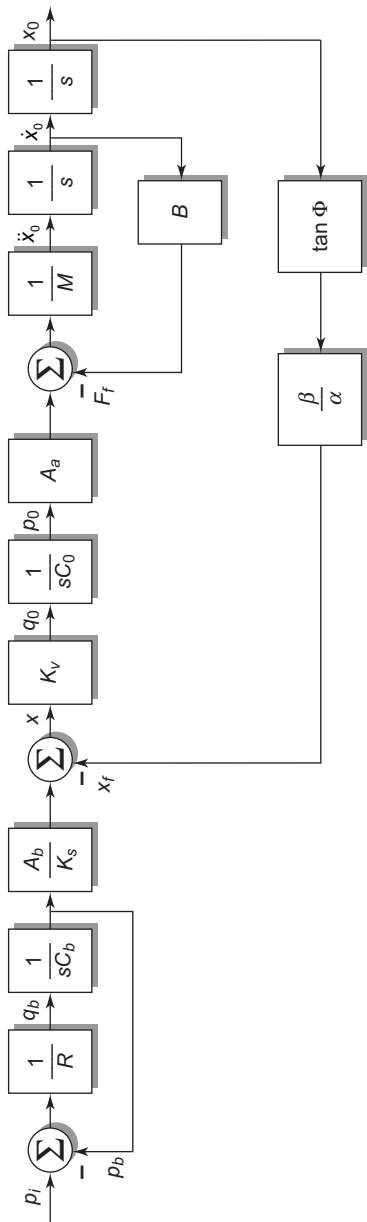


Fig. 7.13 Block diagrammatic representation of Fig. 7.12

7.3 ELECTRICAL ACTUATORS

Even when electrical/electronic control scheme is used the final control element can still be a pneumatic actuator. It is thus necessary to convert the electrical output of the controller (signal) into a pneumatic one. A typical scheme of such a converter is shown in Fig. 7.14. A beam which acts as a flapper to the flapper-nozzle assembly has mounted on it a voice coil which is fed with the 4–20 mA output of the controller. Depending on this current which passes through the coil that is supported in the field of a permanent magnet fixed in space, the coil itself is attracted more, or, less, towards the magnet. Consequently, the beam holding this coil is deflected more/or less, in turn. This results in an increase or decrease in output pressure p . The feedback bellows element provides a torque opposite to that provided by the voice coil motor. The output pressure is proportional to the d c current feeding the coil.

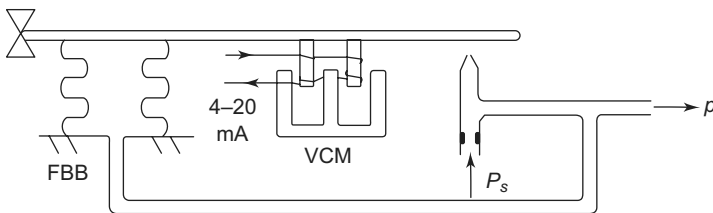


Fig. 7.14 Electro-pneumatic converter, FBB: feedback bellows, VCM: voice coil motor

Electrical motors are also used as final control elements. Servo motors with proper gearing arrangement are provided with separate windings one of which is supplied directly from the ac line. The other winding is supplied from the error through a modulator amplifier. When the latter has a phase leading the main line phase the servo-motor rotates in one direction, and when this phase is lagging, the motor rotates in the reverse direction. A feedback arrangement is also provided to make the position-balance to be fed back so that the controlling coil attains the phase of the main line and the rotation stops. The arrangement is schematically shown in Fig. 7.15.

7.3.1 Stepper Motors

A stepper motor is another electrical actuator used more often in digital control systems as it can take digital inputs. It is an electromagnetic device designed to convert a series of input pulses into discrete angular movements—one for each power pulse. These power pulses may be of same polarity or mixed polarity. They are sequentially delivered to the same winding of the motor or to different coils in the motor successively. Input pulse-rate need not remain fixed.

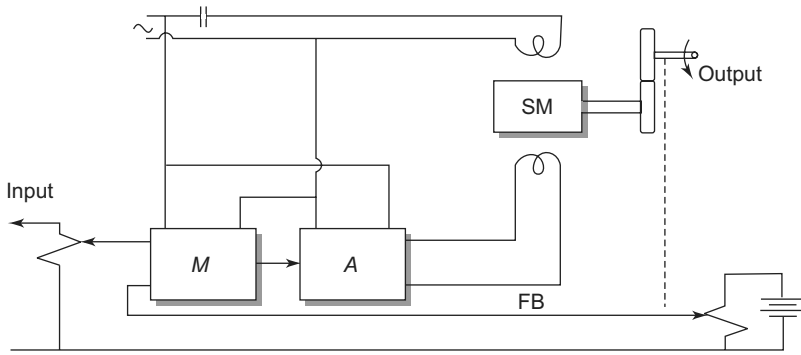


Fig. 7.15 Servomotor used as a final control element; *M*: modulator, *A*: amplifier, *FB*: feedback, *SM*: servomotor

A typical drive system of a stepper motor is shown in Fig. 7.16. Input controller is usually a microprocessor which generates a pulse train. The pulse train is fed to a logic sequencer which is a logic circuit that controls the winding excitation in a sequence. Logic sequencer output is passed through a power driver which, in turn, supplies the motor windings.

There are mainly two types of stepper motors:

- (a) variable reluctance (VR) type, and
- (b) permanent magnet (PM) type.

A hybrid type, combining the two, may also be quite useful and is sometimes used.

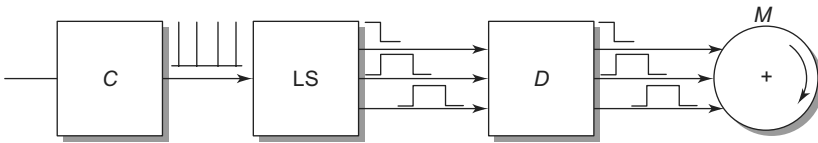


Fig. 7.16 Drive system of a stepper motor; *C*: controller, *LS*: logic sequencer, *D*: driver, *M*: stepper motor

In Fig. 7.17(a) and (b) are shown the cross-sectional model of a typical 3-phase VR type stepping motor and its winding arrangement. The switches and supply indicate the sequential pulse input model. In Fig. 7.17(a), the rotor position is shown with switch *S*₃ closed. If now switch *S*₂ is closed the rotor would rotate to align itself with pole-pairs 2-2' such that the marked tooth faces the 2'-pole tooth. In this, a sequential switching would initiate a counterclockwise movement. The variable reluctance between the pole pairs of the stator and the rotor has been obtained because of the difference in the number of pole-pairs in the stator and rotor. It may be noted that symmetry in the variable reluctance value may produce rotation in any direction. This may be countered by other means. For example, to assume

no loss or gain in position coil 2 is energized before the excitation in coil 1 is removed. Reversal in switching sequence reverses direction of rotation. Air gap should be small. Step angle is decreased by increasing number of phases and stator and rotor teeth.

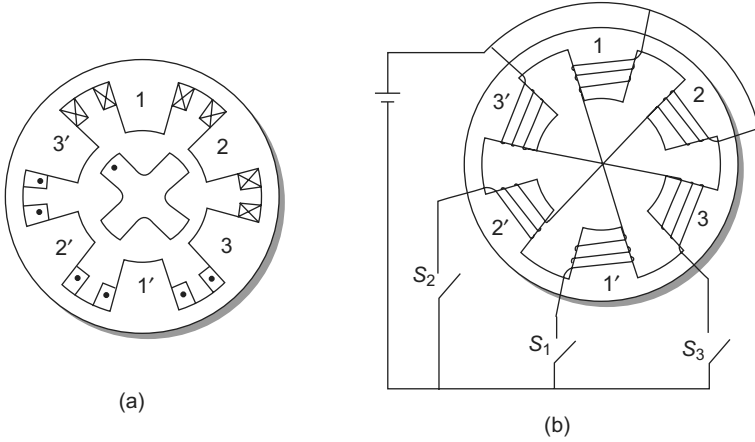


Fig. 7.17 Cross section models of a variable reluctance type stepper motor; (a) the constructional feature, and (b) the switching feature

Multistack VR type stepper motors are also available where successive stacks placed axially are excited cyclically leading to smaller step-angle due to the possibility of availing of small phase displacement of stator fields at each successive switching. If the number of rotor teeth is n_r and number of stacks is n_s , with the rotor teeth perfectly aligned, stator teeth of various stacks differ by an angular displacement of

$$\alpha_d = 360^\circ / (n_s / n_r) \quad (7.38)$$

This also is the resolution (angular) of the motor. In a multiple stack motor, the number of phases are the same as the number of stacks.

The permanent magnet type stepper motor uses a single tooth of the stator as a phase and the rotor is a permanent magnet. Figure 7.18 shows a typical three phase PM type stepper motor operation circuit. Again, switches with supply model the pulse input from sequencer-driver system. A 120° step shift occurs here in comparison to 30° shown in Fig. 7.17. PM stepper motors can withstand higher torque and come to a fixed position quickly even with the excitation off after starting. Magnet, however, is expensive and remanent flux limits the magnetic flux density. Usual construction of PM stepper motors is to have two stator phases effectively displaced by 90° electrical degrees. The motors are bidirectional and are often known as logic stepper motors.

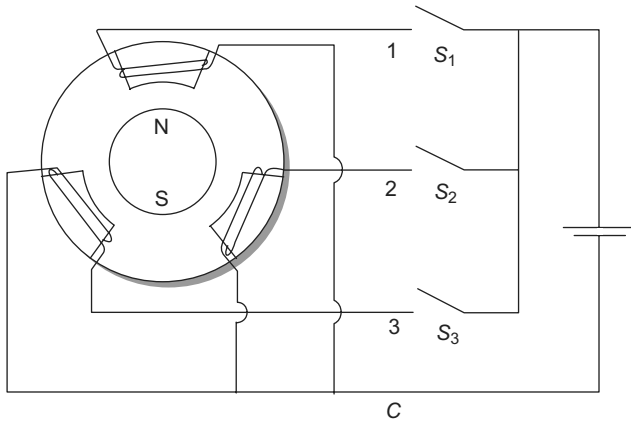


Fig. 7.18 Scheme of a permanent magnet type stepper motor with switching feature

Hybrid type combines the two principles as has already been mentioned. However, the coil connections are different and a cylindrical magnet lies in the core of the rotor which has a lengthwise magnetization to produce a unipolar field. Each pole is covered with uniformly toothed soft-steel disc. Teeth on the two sections are misaligned by a half-tooth pitch.

Cyclonome is a typical stepper motor and a trademark name of a commercial organization. Its principle of operation is shown in Fig. 7.19. It has a three-pole magnetic circuit. Pole *A* is the holding pole, known as the detent pole, and poles *B* and *C* are driver poles. Two large-area saturated alnico magnets M_1 and M_2 with their poles marked *N* and *S* in the figure are in the magnetic circuit of the stator. Stack soft iron laminations form the rest of the stator circuit including the bottom arm over which the pulse feeding coil is mounted. A solid ten-tooth rotor is mounted on a stainless steel shaft and centrally positioned in the air gap of the three poles. The rotor, as shown, is in one of the quiescent positions. One tooth of it is opposite one projection of *A* and the diametrically opposite tooth is positioned opposite to pole *B*. The flux ϕ_p , produced by the permanent magnet, as shown, provides a holding torque even when the coil does not provide any input. A similar quiescent state exists when the rotor is displaced by $1/2$ tooth pitch ($= 1/20$ th of revolution for 10-tooth wheel) from the previous position—in this case the other part of the detent pole becomes active with the drive pole changing to *C*. With the position shown, let a current pass through the coil with the polarity as shown. The mmf would destroy the south pole flux in *B* and make *C* a south pole and a switching would result with the flux now passing from pole *A* to pole *C*. This causes the rotor to turn by $1/2$ tooth-pitch clockwise and it takes the new quiescent position. If now another pulse of opposite polarity is applied, a similar switching occurs

causing the rotor to turn 1/2 tooth-pitch clockwise. Thus by applying pulses of alternating polarity of sufficient magnitude the motor is made to step continuously in one direction.

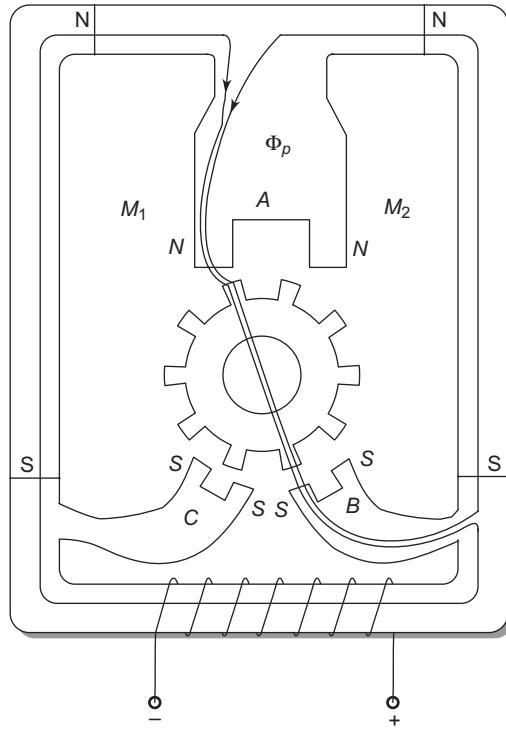


Fig. 7.19 A special type of a permanent magnet reluctance motor with switching feature

Stepper motors can be used in two distinct modes in control systems—the open loop mode and the closed loop mode. A stepper motor being a digital device itself, its angular shaft position is completely determined by the number of input pulses and a feedback device to determine its shaft position is not essential.

In the closed loop mode, which essentially is a position feedback mode it is used like a conventional servo-motor and a signal from the output is fed back to operate a gate controlling the pulses from a pulse generator.

7.3.2 Drive Circuit

It has been mentioned that drive to a stepper motor is through a train of pulses which is used to excite the winding in a sequence. For a higher power motor, direct excitation by these pulses is not sufficient and power

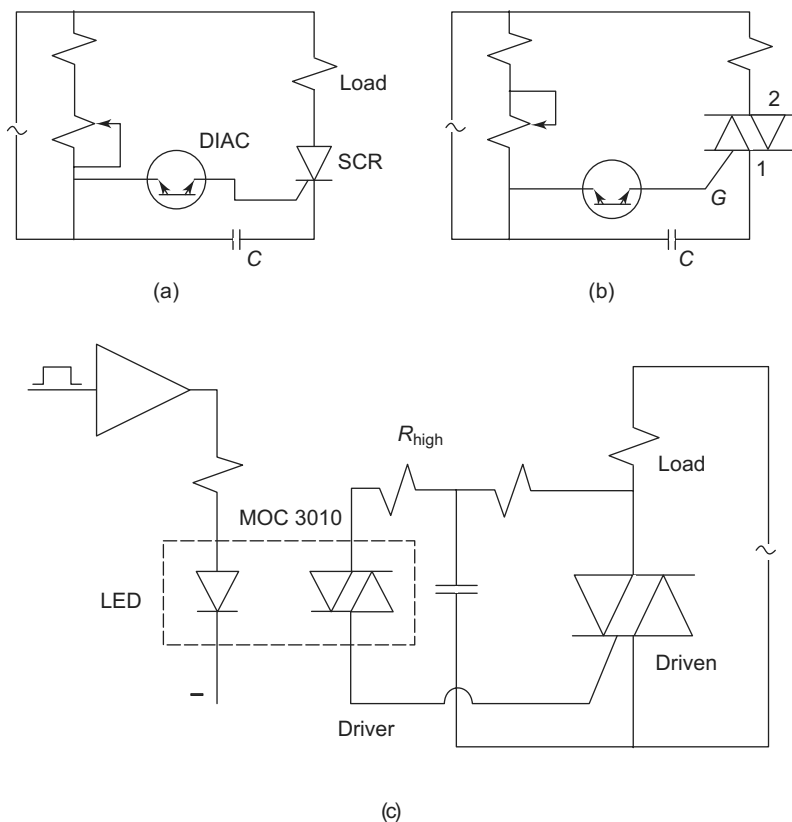


Fig. 7.20 (a) Drive for an SCR, (b) Drive for a triac.
(c) Drive with optoisolator chip

amplification is necessary. In fact, most of the modern day drives are SCR or triac controlled. Both SCR and triac are high power (high voltage and high current) semiconductor switches. They are termed as thyristors. While triacs can pass alternating current, SCR can pass current in one direction only.

To turn an SCR on, a voltage between the gate and the cathode is impressed to raise the gate current above a specified value. Once fired, the SCR remains on until cathode to anode voltage drops down below a value known as the end of the positive half cycle. The holding voltage is about 1.2 V because an SCR, in the on condition, is equivalent to two forward conducting diodes in series. The gate current required for turning on the SCR depends on the size, design and rating of the device—it may vary from 1 mA to well over 100 mA for a duration of 2 to 20 μ sec. The SCR driver itself is, therefore, a pulse generator that can deliver short pulses with high peak but low average current. A diac is often used when the control circuit is operated from the same high voltage source as the load.

A typical such scheme is shown in Fig. 7.20(a). The capacitor C charges up to build a voltage, typically 20 to 40V, when the diac turns on from its off condition and remains on till the voltage across it falls to less than about 1.2V due to discharge of the capacitor. As the diac is connected between the gate and the cathode through the capacitor, discharge occurs via the thyristor gate producing a high peak current. The SCR can be replaced by a triac as shown in Fig. 7.20(b) as the diac driver can deliver trigger pulses both during positive and negative cycles. When ac gate drive is used, it is necessary that G-I and 2-1 terminal voltages should be of same polarity.

It must be seen that neither side of the power line may be connected to the circuit ground, opto-isolators are now popularly used to counter this situation; although pulse transformers are also used for the purposes. This allows one to design ground isolated drive circuit. Nowadays IC chips are available that helps to design such isolated drive circuits. A typical circuit using an opto-isolator is shown in Fig. 7.20(c). The driving chip consists of a triac which is light activated and the light comes from a LED which is turned on for turning on the main triac also called the driven triac. In the first stage the driver triac is turned on by the LED which, in turn, turns on the main triac. When the main triac turns on, voltage to the driver triac becomes small and it gets off. For stepper motor drives, the main control would be on to the LED's.

Review Questions

1. Sketch a pneumatic spring type actuator, label its parts and explain its working principle. How may the power of such an actuator be increased?

The stroke length of a spring motor is 7.5 cm and the diameter of the diaphragm is 25 cm, what should be the spring constant? For a specified hysteresis of 1 per cent, what would be the thrust and frictional forces?

(Ans: 54.6 kg/cm, 98.75 kg, 3.95 kg.)

2. What is the C_v factor of a control valve? How is it useful in valve selection and sizing?
3. What are the factors that should be known for selecting a control valve? What different types of valves are commercially more used?
4. When are single-seated and double-seated valves used? List and compare their advantages and disadvantages.
5. How hydraulic gradient method is used in obtaining drops across the valve for different flow conditions?

At maximum flowrate of 210 gpm of water through a control valve in a system, the line losses are 0.01 kg/m while the equivalent

line length is 123 meter. The inlet pressure to the line is 7 kg/cm^2 and the system output pressure is 2.5 kg/cm^2 . Calculate the C_v value of the valve.

(Ans: 30)

6. How does viscosity affect the operation of a control valve? Indicate the method of correction for viscosity of the calculated C_v value.

A double-seated control valve is used for control of flow of water having a maximum flowrate of 200 gpm when the drop across the valve is 70% of that when the flow is minimum of 10 gpm. The pressure head across the system is 5 kg/cm^2 and line losses at minimum flow conditions are negligible. What type of valve should be selected and what should be its size? Assume viscosity of water as 400 stokes. Refer to Figs 7.10 and 7.11 for sizing.

(Ans: Linear, 6.25 cm)

7. What are the different types of stepper motors? Explain their operating principles with appropriate diagrams. How are they used in open loop and closed loop conditions?
8. In an equal percentage valve, the normalized flow rates at normalized lifts of 0.4 and 0.67 are 0.4 and 0.85 respectively. What is the rangeability of the valve?

[Hint. Use Eq. (7.12 h), where $q_{f1} = 0.4$, $q_{f2} = 0.85$, $l_{f1} = 0.4$ and $l_{f2} = 0.67$]

As, $\ln R = 1/(l_{f1} - l_{f2}) \ln(q_{f1}/q_{f2})$ putting values, one gets R]

Connecting Elements and Common Control Loops

8.1 INTRODUCTION

In a large majority of processes control parts consist of pneumatic systems. These system involve pneumatic transmission lines interconnecting the plant, the control valve, controller, measuring element, etc. In many instances the lengths of the transmission lines, and consequently, their effects on the signal moving in the closed loop are not negligible. Analogous to the electrical type, pneumatic lines can also be considered to consist of resistance, capacitance and inductance in a distributed fashion, calculable on a per unit length basis, but it should be remembered that the type of the velocity profile of the fluid in transmission is very important. For small signals ($< 0.04 \text{ kg/cm}^2$) and for laminar flow rigorous analysis at different frequencies assuming specific types of flow profiles has shown good correspondence with practical results. At large signals, the results tend to deviate because of turbulence in the fluid flow. Particularly at high flow rate the critical pressure gradient of $0.001 \text{ kg/cm}^2/\text{m}$ is always exceeded (for a 0.625 cm , i.e., $1/4''$ line) and the lag due to the transmission lines is significant enough for investigating the system behaviour and controller setting or loop rearrangement.

When a distributed model of the transmission line is considered, the analysis becomes involved and non-linear. At low frequencies such a complex analysis is unnecessary and lumped, linearized incremental models are good enough for the purpose. However, even at low frequencies the flow-pressure non-linearity of incremental models cannot be completely discarded as a prominent part is played by this relation in control problems.

The R-L-C (resistance-inductance-capacitance) elements are obtained in these systems by noting that pressure is an across variable, analogous to emf, and flow rate is a flow variable, analogous to current and the analogy is drawn on that basis.

8.2 RLC ELEMENTS

8.2.1 Resistance (R)

The resistance for pure lines results from viscous drags on the fluids by pipe surfaces and is usually defined on a per unit length basis as the pressure drop per unit flow rate

$$\frac{\Delta p}{q} \Big|_{l \rightarrow \text{unit}} = R_1 \quad (8.1)$$

Other resistances that come in fluid lines are due to head losses when the flowing fluids meet valves, orifices, tees, bends, etc.

The line resistance for capillary lines and incompressible flow is easily obtained from Poiseuille's equation as (Fig. 8.1)

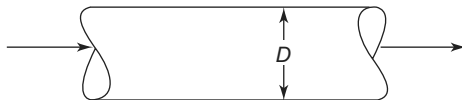


Fig. 8.1 Transmission or capillary line; D : diameter

$$R_1 = \frac{\Delta p}{q} \Big|_{l=1} = \frac{128 \mu \rho}{\pi D^4} \quad (8.2)$$

where μ = fluid viscosity, ρ = fluid density and D = line diameter.

For compressible fluids, if the pressure drop is within 10%, the above relation can, in practice be used. When the pressure drop is slightly larger the relation is complicated; an approximate mass-flow rate relation is

$$w = \frac{\pi D^4}{256 \mu R_g T} (p_1^2 - p_2^2) \quad (8.3)$$

where R_g = gas constant, T = absolute temperature and p_1 and p_2 indicate the pressures at sections 1 and 2 (Fig. 8.2a). It is not easy to use such a relation for calculating R_1 . For flows in lines other than capillary the flow-pressure relationship is non-linear, and for incompressible fluids, it is given by

$$R_1 = \frac{\Delta p}{q} \Big|_{l=1} = \frac{8 F \rho q}{\pi^2 D^4 \lambda} \quad (8.4)$$

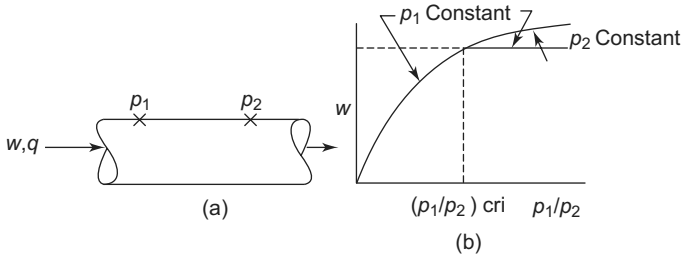


Fig. 8.2 (a) Diagram pertaining to the mass-flow rate calculation;
(b) $w - (p_1/p_2)$ plot

where F = dimensionless frictional coefficient and λ = a hydraulic parameter defined as equal to $\pi D^2/(4\phi)$, where $D/2$ = hydraulic radius of the pipe and ϕ = wetted perimeter. The relation for R_1 , as shown in Eq.(8.4) shows that the resistance varies directly with the flow rate. Even then the calculation of R_1 is not as simple as it is thought to be. The coefficient F is dependent on Reynold's number Re ; it is an approximate empirical relation which is given by

$$F = k_1 + k_2 Re^n \quad (8.5)$$

where k_1 and k_2 are constants and n is an index. Equation (8.5) is valid for commercial flows, i.e., for $Re > 2000$. Usually $k_1 \approx 3.5 \times 10^{-3}$, $k_2 \approx 264 \times 10^{-8}$ and $n \approx 0.4$. For laminar flow F is inversely proportional to Re , i.e.,

$$F = k_3/Re \quad (8.6)$$

usually $k_3 = 16$.

A square law relation is obtained from the lumped incremental model by considering a restriction of area $\alpha = \pi D^2/4$ actually present in the line, such that

$$q_l = C_d \alpha \cdot (\Delta p)^{\frac{1}{2}} \quad (8.7)$$

where q_l stands for the flow rate with a pressure drop Δp across a length l of the section of the line, therefore, Eq. (8.4) is to be expressed in terms of q_l for the elimination in conjunction with Eq. (8.7). Using the relations in Eqs (8.4) and (8.7) one derives the relation for the coefficient of discharge C_d for commercial pipings as

$$C_d = \sqrt{\frac{2\lambda}{\rho Fl}} \quad (8.8)$$

Another important consideration is the sonic flow. If the ratio of pressures at two different sections, from up to down, becomes larger than a ratio called critical pressure ratio (at which maximum flow occurs), then the flow rate becomes independent of the downstream pressure and also

the flow becomes sonic. This can be easily explained by a curve shown in Fig. 8.2(b). The q versus p_1 relation is then

$$q = \frac{k\alpha C_d p_1}{\sqrt{T}} \quad (8.9)$$

where k = a constant and T = absolute temperature.

Equation (8.9) bears enough similarity with Eq. (8.3). Equation (8.3) is actually a relation for a pressure ratio below $(p_1/p_2)_{cr}$ where the square law relation prevails and the flow also is then subsonic. As the relations are mostly based on the equivalence principle, the resistance calculation for the apparatus producing such resistance is similar.

8.2.2 Inductance (L)

Inductance arises because of the inertia effects in the flowing fluid and is obviously proportional to density. It is calculated on the basis of the consideration that a force, because of a differential pressure between two sections in a line, accelerates the fluid mass. Since it is difficult to know the acceleration, the average velocity in the line is approximated as its time integral. Thus referring to Fig. 8.3 one has

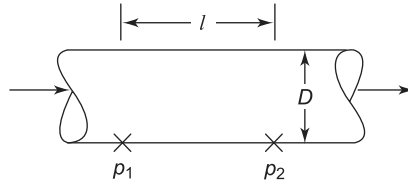


Fig. 8.3 Diagram pertaining to line-inductance calculation

$$(p_1 - p_2) \frac{\pi D^2}{4} = m \cdot \frac{du}{dt} \quad (8.10)$$

where m is the fluid mass to be accelerated and du/dt is the acceleration. It should be noted that the mass over length l is actually accelerated in the subsequent section and the velocity is approximately known from flow rate q as

$$q = \frac{\pi D^2}{4} u \quad (8.11)$$

Also, as in the previous case, l can be taken to be a unit length in which case

$$m_1 = \frac{\rho \pi D^2}{4} \quad (8.12)$$

Hence, combining Eqs (8.10), (8.11) and (8.12), with initial conditions zero

$$\frac{\rho\pi D^2}{4} \times \frac{4}{\pi D^2} \frac{dq}{dt} = \frac{\pi D^2}{4} (p_1 - p_2) \quad (8.13)$$

or

$$(p_1 - p_2) = \frac{4\rho}{\pi D^2} \frac{dq}{dt} \quad (8.14)$$

or

$$q = \frac{\pi D^2}{4\rho} \int (p_1 - p_2) dt \quad (8.15)$$

Thus the concept of inductance gives its per unit length values as

$$L_1 = \frac{4\rho}{\pi D^2} \quad (8.16)$$

It is well known that for the ranges of pneumatic pressure in signal lines, ρ is quite small and hence the effect of L_1 or L in general is also quite small. But in actual process fluid-flow channels where fluid density is sufficiently large, Eq. (8.16) will give an idea of the effect of the inertia of the fluid, particularly for frequency response studies.

8.2.3 Capacitance (C)

Capacitance arises in a fluid-flow system because of energy storage due to (i) compression of the fluid and (ii) flexure of certain component in the system such as a bellows (i.e., a receiver) element or a diaphragm motor where actually volume increases when pressure increases.

For incompressible fluids, only the flexure effect of any element is of consequence, whereas for compressible fluids, thermal conditions also have to be considered. When isothermal condition prevails (which means that changes are not appreciable or that the changes are slow), the ideal gas law of Eq. (8.17) holds good

$$pv = \eta RT \quad (8.17)$$

where

$$\eta = w/\mu = \frac{\text{mass of the gas}}{\text{molecular wt of the gas}} = \text{number of moles} \quad (8.18)$$

Defining the specific gas constant as

$$R_0 = R/\mu \quad (8.19)$$

one obtains from Eqs (8.17), (8.18) and (8.19)

$$q = \frac{1}{\rho} \frac{dw}{dt} = \frac{1}{\rho} \frac{v}{R_0 T} \frac{dp}{dt} \quad (8.20)$$

where ρ is assumed to be the average density of the gas for reasonable fluctuation in pressure. If the average pressure is p_a , then

$$\frac{p_a}{\rho} = R_0 T \quad (8.21)$$

From Eqs (8.20) and (8.21), therefore

$$q = \frac{v}{p_a} \frac{dp}{dt} \quad (8.22a)$$

or

$$p = \frac{1}{v/p_a} \int q \cdot dt \quad (8.22b)$$

giving the isothermal volumetric capacitance as

$$C_{iv} = \frac{v}{p_a} \quad (8.23)$$

For unit length this is obtained as

$$C_{iv1} = \frac{\pi D^2}{4 p_a} \quad (8.24)$$

where $\frac{\pi D^2}{4} \cdot l = v$. When the fluctuation rate is low (≤ 5 Hz), the assumption of an isothermal flow is justified. When rapid expansion and compression take place, the above description is required to be modified for the adiabatic case and then the volumetric capacitance is given as

$$C_{av} = \frac{v}{\gamma p_a} \quad (8.25)$$

where $\gamma = C_p/C_v$ = ratio of the specific heats of the fluid at constant pressure to constant volume. When the line terminates to big enough volume V , ($V \gg v$), the actual capacitance of the volume lies in a limit given as

$$\frac{V}{\gamma p_0} \leq C \leq \frac{V}{p_0} \quad (8.26)$$

In others, polytropic expansion is considered for flow rates which are neither very slow nor very fast and the capacitance is then

$$C_m = k \frac{V}{p_0} \quad (8.27)$$

where $1 < 1/k < \gamma$. In ordinary pneumatic control systems $1/k = 1.2$.

Now consider a receiver element like a bellows element at the terminating end of the line, as shown in Fig. 8.4. Assuming an incompressible fluid flowing in, one writes at force balance

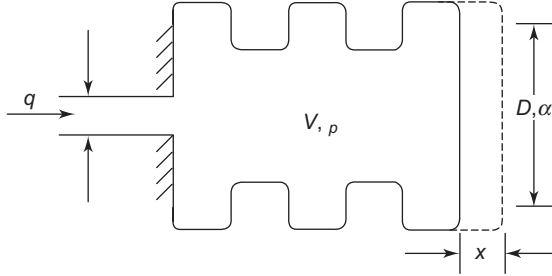


Fig. 8.4 Receiver element; D : diameter; α area; V : Volume; p : pressure; q : flow rate

$$p\alpha = K_b \cdot x \quad (8.28)$$

where K_b = stiffness of the bellows element. As the bellow element is axially flexible, extension is only in x -direction, so that

$$q = \frac{\dot{V}}{\rho} = \frac{1}{\rho} \frac{d}{dt}(\rho V) = \frac{dV}{dt} = \alpha \frac{dx}{dt} \quad (8.29)$$

Combining Eqs (8.28) and (8.29)

$$q = \frac{\alpha^2}{K_b} \cdot \frac{dp}{dt} \quad (8.30)$$

or

$$p = \frac{1}{\alpha^2/K_b} \int q \cdot dt \quad (8.31)$$

giving the volumetric capacitance as

$$C_f = \alpha^2/K_b \quad (8.32)$$

If instead of a bellows element, a non-flexure volume of a tank is considered, the bulk modulus, β , of the fluid would provide the energy storage parameter. Since

$$\beta = V \frac{dp}{dV} \quad (8.33)$$

Hence

$$q = dV/dt = (V/\beta)dp/dt \quad (8.34a)$$

or

$$p = \frac{1}{V/\beta} \int q \cdot dt \quad (8.34b)$$

giving the capacitance as

$$C_\beta = V/\beta \quad (8.35)$$

A compressible fluid ending in a receiver flexure element of volume ($V \gg v$) would produce a larger capacitance. As a matter of fact, this should be the sum total of C_m and C_f , thus

$$C_t = C_m + C_f = kV/p_0 + \alpha^2 K_b \quad (8.36)$$

However, one should note that the spring effect of the flexure element should be restricted such that V does not change by a large amount. In practice, there is only $\pm 10\%$ change which justifies this superposition.

While briefly discussing the theoretical means of calculating the RLC elements, one should simultaneously be alerted to the fact that the distributed nature of the RLC elements show a wide difference in practical performance when frequency response studies are made for different systems.

If the lags in the transmission only nominally affect the performance of the system, a lumped parameter modelling of the system is justified. In critical cases experimental test should provide the necessary data.

The lumped RLC elements calculated above are now arranged as shown in Fig. 8.5. The transmission line is terminated by a flexure element as a load whose capacitance is calculated as C_f . The parallel capacitance of the line is C_m , but for a distributed system it is taken as half this value as per the transmission line theory for high frequencies. If C_m is taken in its full value R should be halved. If $(1/2)C_m + C_f = C$, the transfer function given by Eq. (8.37) is easily obtained for the purpose of analysis.

$$\frac{p_2(s)}{p_1(s)} = \frac{1}{1 + sCR + s^2 LC} \quad (8.37)$$

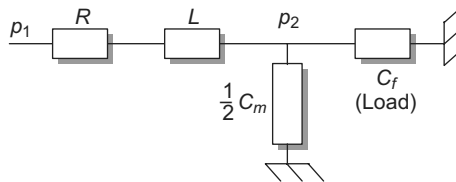


Fig. 8.5 Lumped RLC blocks in a transmission line

8.3 FLOW CONTROL

In most of the flow processes lag is negligible. The order of the lag varies from a fraction of a second to a few seconds except in the cases of long oil/gas pipe lines where a large transportation lag appears, but then such processes are to be treated in a different line than the industrial flow processes.

A flow process mainly consists of:

- (i) a process, having
 - (a) a source,
 - (b) a receiver and
 - (c) a flow channel,
- (ii) a transducer with a restrictor and a transmitter,
- (iii) a controller and
- (iv) a control valve.

The source may be a centrifugal, positive displacement or reciprocating type pump, a compressor, etc. A source of this type may introduce flow fluctuations in the systems having frequencies $\omega \geq 5$ rad/sec. With a restrictor in the line for measurement, pressure vortices on the two sides of the restrictor (like orifice) also introduce additional flow fluctuations of sufficiently high frequencies. In the control valve, restrictors or other irregularities such as bends, etc. change the flow pattern in a random fashion; these changes are fluctuations, due to which a sort of random noise is introduced in the system. These fluctuations at different frequencies and at different places should be considered for determining the control action needed as a transportation lag is likely to be introduced in such a situation. During analysis, the location of the disturbance centre should be chosen such that its maximum effect is taken care of, but at the same time, undue emphasis should not be given to this, complicating the choice of control gears unnecessarily.

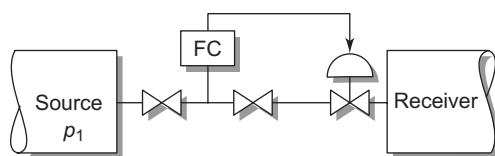


Fig. 8.6 A typical flow control scheme; FC: flow controller

From an equivalent model a rough calculation would show the amount of lag in the flow process. Figure 8.6 shows a typical control scheme. In Eq. (8.10), if one assumes that the source pressure is constant at p_1 and the receiver pressure is at p_2 , then

$$(\alpha/\rho)du/dt = \alpha(-\Delta p) \quad (8.38)$$

where α = pipe area, l = length, ρ = fluid density, u = fluid average velocity, Δp = total head which causes the acceleration. The head Δp has been assumed to be negative as this is the friction loss, p_1 and p_2 are constants, less of Δp would give more of du/dt . It would be a function of the drop across the control valve (a nonlinear function, because as the valve opens more, the drop across it is less) and of the line losses as also of drop across other restrictions. Actually this head is a nonlinear function of the stem-lift and of the stream velocity (Cf. Eq. (8.74) of Sec. 8.5). A linearization of this function for changes in velocity and valve position would give

$$\Delta p = \left\{ \frac{d(p_1 - p_2)}{du} \right\} \Delta u + \frac{dp_v}{dy} \Delta y \quad (8.39)$$

where p_v = drop across the valve and y = stem lift

Now if the line loss per unit length is p_l , then

$$\begin{aligned} p_1 - p_2 &= p_l \cdot l + p_v + p_r \\ &= p_l \cdot l \left[1 + \frac{p_v + p_r}{l \cdot p_l} \right] \\ &= p_l l [1 + \phi] \end{aligned} \quad (8.40)$$

where p_r = drop across the restrictors and other irregularities.

Combining Eqs (8.38), (8.39) and (8.40), and assuming velocity change Δu in the left-hand side term because of linearization

$$\rho l \left(\frac{du}{dt} \right) = - \left(\frac{dp_l}{du} \right) [l(1 + \phi)] \Delta u - \left(\frac{dp_v}{dy} \right) \Delta y \quad (8.41)$$

Therefore

$$\left\{ \rho l s + l(1 + \phi) \left(\frac{dp_l}{du} \right) \right\} \Delta u = - \left(\frac{dp_v}{dy} \right) \Delta y \quad (8.42)$$

Hence

$$\frac{\Delta q}{\Delta y} = \frac{\alpha \Delta u}{\Delta y} = \frac{\alpha \left(\frac{dp_v}{dy} \right) / \left[l(1 + \phi) \frac{dp_l}{du} \right]}{s \rho l / \left\{ (1 + \phi) \left(\frac{dp_l}{du} \right) \right\} + 1} \quad (8.43)$$

Thus the response of the process to the valve-stem position is with a first-order lag of value

$$\tau_p = \frac{\rho}{(1 + \phi)(dp_l/du)} \quad (8.44)$$

From Eq. (8.4), the value of dp_l/du is obtained as

$$\frac{dp_l}{du} = \frac{4F\rho u}{D} \quad (8.45)$$

Combining Eqs (8.44) and (8.45) one gets

$$\tau_d = \frac{D}{4(1+\phi)Fu} \quad (8.46)$$

From Eq. (8.46) the largest value of τ_p is.

$$\tau_{p \max} = \frac{D}{4Fu} \quad (8.47)$$

which occurs when drops across the valve and other restricters and irregularities are assumed negligible. In a low-size pipe (say 5cm dia.) with a low differential pressure (say 0.3 kg/cm²) so that u is small, $\tau_{p \max}$ can be calculated to be not more than 1.5 sec. Actually, the value of ϕ is quite large; in fact, more than 80% of pressure drop occurs across the valves and restrictions giving a process time constant of about 0.25 sec. In industrial processes F , ϕ , u and D determine τ_p and depending on the fluid and pressure differential this value changes but rarely exceeds 1 sec.

In the above discussion it has been assumed that the pressure of the source is constant at p_1 . If the source is a centrifugal pump or even a constant flow pump, it has its own regulation characteristics. As the flow increases, the pressure falls in a centrifugal pump and the approximate regulation curve is drawn from the relation

$$\frac{p_1}{p_{1m}} = 1 - r_c \left(\frac{q}{q_m} \right)^2 \quad (8.48)$$

where r_c is the regulation coefficient given by $0.2 < r_c < 0.4$. At a very large fraction of q/q_m , p requires to be corrected and thus also Δp in Eq. (8.38). As a matter of fact non-linearity comes in and the analysis becomes quite complex. The inference is, finally, that τ_p varies at different flow conditions. In a constant flow pump, the case is reversed. At high pressure conditions there is less flow as leakage increases. However, it has better regulation characteristics; regulation between 95 and 100% is, in fact, easily achievable. The process is affected as in the case of the centrifugal pump.

8.3.1 Transducer with Transmitter

Earlier, flow control loops had been using an enlarged lag mercury manometer in conjunction with a restriction like an orifice or a venturi. The manometer, in principle, is a second-order instrument the damping ratio of which could be controlled at will for a desired optimally flat response. When extra damping is not adjusted, i.e., the damping valve is kept fully open (Fig. 8.7), the value of ζ is about 0.1 and ω_n varies from 3 rad/sec to

6 rad/sec for such types of meters. The commercial meters have time lags $2 \text{ sec} \leq \tau_m < 20 \text{ sec}$. These instruments, therefore, have large time constants and would introduce a large phase lag. In recent years, a force-balance type of differential pressure transmitter is being used for convenience of operation because it does not require any manometric fluid, its range can easily be changed, its installation is simpler and electronic conversion is much easier. Besides, it has low response time (see Fig. 8.8). The transmitters are usually made with diaphragm and require a small volume of fluid to provide a signal. This makes the response time quite low, of the order of 0.2 to 0.3 sec. The pneumatic transmitter sends the signal through a pneumatic transmission line and is terminated at the other end in a volume (a receiver element) for operating the controller. The total response time and frequency is also determined by the length of this line and the terminating volume. Electronic/electrical transmission obviously has the advantage of reducing these values. With a commercial line diameter of 0.625 cm, the transmitter response is quite unaffected by line response up to about 30 m. Above this length, line response becomes the determining factor. Thus, if the length of the line $l > 30 \text{ m}$ the advantage of the d/p

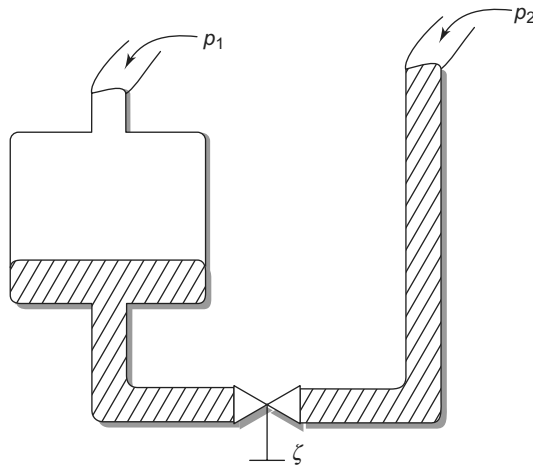


Fig. 8.7 Manometer with a damping valve; ζ : damping constant

transmitter is effectively marginal. In such cases, however, one can conveniently choose electronic transmission. Approximate line and transmitter frequency response curves are shown in Fig. 8.9 for different line lengths. The system schematic is shown in Fig. 8.8.

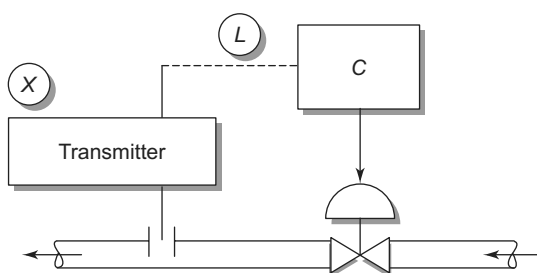


Fig. 8.8 Transducer-transmitter-line-controller system; C: controller; L: transmission line; X: transmitter

In commercial usage the transducer-transmitter time lag is kept between 0.1 and 0.5 sec for flow-control systems.

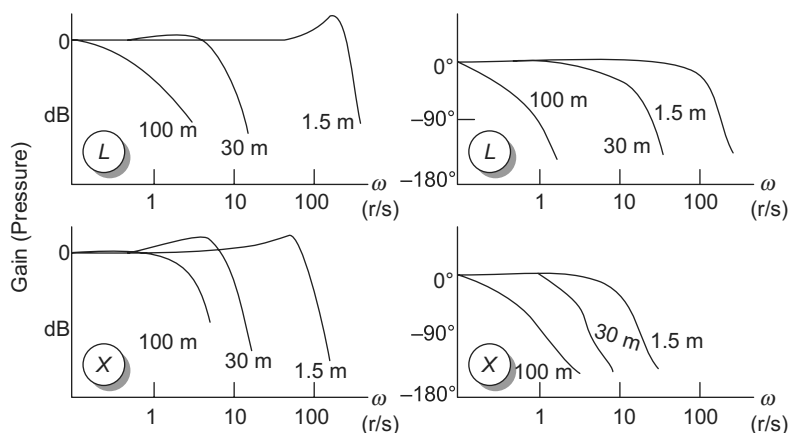


Fig. 8.9 (a) and (b) Frequency response curves of the transmitter (X) and transmission line L of Fig. 8.8; m: meters

8.3.2 Controller and Control Valve

The response characteristic of a controller has been discussed earlier. For proportional control alone, lag is negligible. *P-I* control may be used with a proper choice of the reset time. In any case the controller time constant is hardly allowed to be larger than 0.2 sec. As the transmitter output is fed to a receiver element for operating a line, its time constant is also small as has been pointed out in the previous section.

Control valve

Because of restriction and small storage associated with a control valve its transfer function is approximated by a first-order one and the function between the stem position and pressure is easily obtained as

$$\frac{y(s)}{p(s)} = \frac{K_v}{s\tau_v + 1} \quad (8.49)$$

The value of τ_v obviously depends on the design of the valve. This can be reduced by using a valve positioner which is primarily intended to reduce the stem friction and pressure unbalances about the valve plug (see Chapter 7). Depending on the choice of control valves the time lag may vary from 1 to 25 sec while with a positioner this can be lowered 10 to 20 times. A booster relay also improves the performance. A valve positioner, however, induces saturation in so far as at low signal it gives full air supply to the valve and a larger signal does not give a faster response.

From the different loop blocks that have been separately considered, it is apparent that in flow control systems, the largest contribution in the time lag is from the transmission lines and it can be reduced by: (i) increasing line diameter, (ii) eliminating the lines and (iii) providing electronic transmitters. In modern installations, where the last one is not permitted (because of cost and safety), the controller is actually mounted at the valve site and the transducer close to the valve. Where a control room monitoring is necessary, this type of installation, however, requires more pneumatic lines for monitoring flow and valve signals, for set point adjustments and for manual control. A reduction in the transmission line for purposes of control is shown in Fig. 8.10.

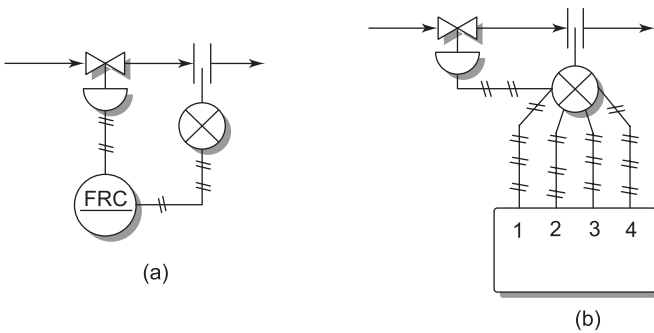


Fig. 8.10 (a) Scheme showing the reduction in transmission line length; FRC: flow recorder controller, (b) scheme of a general signal transmission method; 1: valve signal; 2: flow signal; 3: set point adjustment; 4: manual transfer

Once the transmission lags have been eliminated, the valve lag becomes predominant which with a booster can be reduced. A reasonable fast control can thus be made effective, but then controllability is also judged from the control valve characteristics. It has been discussed after Eq. (8.47) that the major drop occurs across the valve. This, in fact, is necessary for good control. Redrawing Fig. 8.6 as in Fig. 8.11 without loss in generality,

one obtains equations for the flow through the lumped restrictor and control valve as

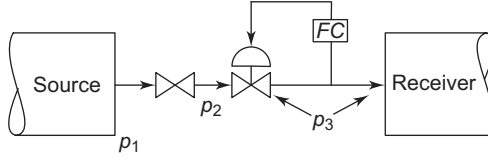


Fig. 8.11 Modified scheme of Fig.8.6; p_1 : source pressure; p_2 : control valve upstream pressure; p_3 : control valve downstream pressure; FC: flow controller

$$q = (p_1 - p_2)^{\frac{1}{2}} \cdot \alpha C_d \quad (8.50)$$

$$q = \alpha_v C_v (p_2 - p_3)^{\frac{1}{2}} \quad (8.51)$$

For a linear control valve with $\alpha_v = k_v y_1$ Eq. (8.51) changes to

$$\begin{aligned} q &= k_v C_v y_1 \sqrt{p_2 - p_3} \\ &= k_1 y_1 \sqrt{p_2 - p_3} \\ &= k_1 y_m y \sqrt{p_2 - p_3} \\ &= k y \sqrt{p_2 - p_3} \end{aligned} \quad (8.52)$$

where

$$y = y_1 / y_m = \frac{\text{instaneous lift}}{\text{maximum lift}} = \text{normalized lift}$$

When maximum flow is q_m and p_{2m} replaces p_2 for q_m , then defining

$$\begin{aligned} \beta &= \frac{\text{drop across valve}}{\text{total drop}} \text{ at maximum flow} \\ &= (p_{2m} - p_3) / (p_1 - p_3) \end{aligned} \quad (8.53)$$

Expressing Eqs (8.50) and (8.52) for q_m and dividing and then using Eq. (8.53) one easily derives

$$\beta = \frac{1}{1 + (k / \alpha C_d)^2} \quad (8.54a)$$

and

$$q_m = k \sqrt{\beta (p_1 - p_3)} \quad (8.54b)$$

From Eqs (8.50) and (8.52), also

$$p_1 - \left(\frac{q}{\alpha C_d} \right)^2 = \left(\frac{q}{ky} \right)^2 + p_3$$

or

$$q \left[\frac{1}{(ky)^2} + \frac{1}{(\alpha C_d)^2} \right]^{\frac{1}{2}} = (p_1 - p_3)^{\frac{1}{2}}$$

or

$$q = ky \left[\frac{p_1 - p_3}{1 + y^2 \left(\frac{k}{\alpha C_d} \right)^2} \right]^{\frac{1}{2}} \quad (8.55)$$

Using Eq.(8.54) this reduces to

$$q = \frac{k \cdot y \sqrt{(p_1 - p_3)\beta}}{\sqrt{\beta + (1 - \beta)y^2}} = \frac{q_m y}{\sqrt{\beta + (1 - \beta)y^2}} \quad (8.56)$$

showing a non-linear nature between q/q_m and y ; but as $\beta \rightarrow 1$, proportionality between q/q_m and y increases showing better controllability. This is shown also from

$$\frac{d(q/q_m)}{dy} = \frac{\beta}{[\beta + (1 - \beta)y^2]^{3/2}} \quad (8.57)$$

which approaches a constant value for $\beta \rightarrow 1$. As β is small, the drop across the valve is small and the control of the valve over the flow rate q is also small. For $\beta = 0.8$, as is mentioned, the control is quite good.

The flow loop performance with respect to disturbance in pressure is now considered. If proportional action alone is used and process is considered to have no lag, the disturbance at the start of the process gives the relation (from Fig. 8.12)

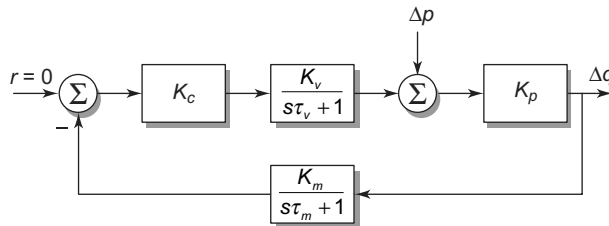


Fig. 8.12 Flow-control scheme with disturbance at the start of the process;
 K 's: gain parameter, τ 's: time constants; Δp : disturbance; Δq : output

$$\begin{aligned}
\frac{\Delta q}{\Delta p} &= \frac{K_p}{1 + \frac{K_p K_c K_v K_m}{(s\tau_m + 1)(s\tau_v + 1)}} = \frac{K_p (s\tau_m + 1)(s\tau_v + 1)}{K_p K_c K_v K_m + (s\tau_m + 1)(s\tau_v + 1)} \\
&= \frac{K_p [1 + s(\tau_m + \tau_v) + s^2(\tau_m \tau_v)]}{(1 + K_p K_c K_v K_m) \left[1 + \frac{s(\tau_v + \tau_m)}{1 + K_p K_c K_v K_m} + \frac{s^2 \tau_m \tau_v}{1 + K_p K_c K_v K_m} \right]} \quad (8.58)
\end{aligned}$$

Clearly, if the loop gain is very large, as is usually the case ($K_p K_c k_v K_m \gg 1$), the response is like a second-order high pass filter, and considering $\tau_v > \tau_m$, regulation is acceptable only up to a frequency $\omega_1 = 2\pi/\tau_v$. This is more clearly revealed when τ_m is negligible. Response is then that of a first-order high pass filter and regulatory control completely fails after a frequency $\omega_2 = 2\pi K_p K_c K_v K_m / \tau_v$. Now suppose integral plus proportional action is called upon to act and the process time constant is also considered. The last one is usually quite small, but for noise which is usually at high frequencies, this time constant cannot be totally ignored. Assume $K_p K_c K_v K_m = K$; a careful study of the system for values of τ_v , τ_p and τ_m is required for setting T_r and K_c . Usually K_c is kept low or the proportional band is kept large to minimize the effect of noise. This actually keeps the resonance peak smaller and error-free. However, for settings, the disturbance frequency range should be known. A loop gain of 2 and $T_r = 1$ sec is very often chosen. Actually, in a *PI* controller, regulation is better than in a *P* controller up to a frequency $\omega_3 = 2\pi/T_r$. Thus, the lower the T_r , the better is the controllability. A low gain choice should always be backed by *I*-action to eliminate offset. Derivative action (with *P*-action only) extends the bandwidth to $1/T_d$ and T_d needs to be small; but derivative action amplifies the high frequency noise and is often unsuitable. Also, the minimum derivative time available in a controller is quite high (5 sec). Inverse derivative action when added to the conventional controller in this loop widens the proportional band for high frequency signals but keeps it constant for low frequency ones; thus, requisite damping is provided for noise (Moore). A *PD* controller in the feedback path allows the gain to increase substantially without affecting stability, and the response with a disturbance improves considerably (Patranabis).

When measurement noise is large, because of turbulence and with an orifice type transducer, quite high frequency noise may be generated. This would change the block schematic as shown in Fig. 8.13. At frequencies near the gain cross-over frequency, the error response is amplified and correspondingly performance deteriorates by saturation and other undesirable exaggeration.

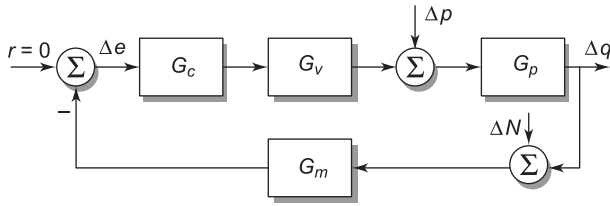


Fig. 8.13 Flow control scheme with large measurement noise
 ΔN : noise input; Δe : error; G 's: transfer functions

Thus

$$\frac{\Delta e}{\Delta N} = \frac{G_m}{1 + G_c G_v G_p G_m} = \frac{G_m \Delta q}{G_p \Delta p} \quad (8.59)$$

This effect is minimized by compromising with a more filtering effect in G_m but not in excess. This is often done in an on-line trial as optimum filtering is effective only when noise frequency bandwidth, valve response, control action and setting, saturation level, etc. are known.

8.4 PRESSURE CONTROL

Like level control, the flow equations in a pressure control system should also be linearized and the system transfer function derived about the nominal operating conditions. Because of gas being in a compressed state in the tank and the consequent energy storage effect, this process becomes integrating in nature. The storage capacity is volumetric and can be approximately calculated by the nominal operating pressure p_0 as V/p_0 . The gas tank may have a number of inlets and outlets and both inflow and outflow are dependent on tank pressure. This complicates the situation and the system time constant is rarely expressible as a simple function of the residence time as in the case of level control. Pressure differences may be such that at any stage sonic flow may occur and temperature change occurs during throttling, i.e., as the pressure decreases, gas expands adiabatically in the tank and the temperature falls. In small tanks the effect is isothermal at low fluctuations. These also are complications which are not very easily solved. For purposes of analysis we can consider a single input-single output tank, in the inlet a constant flow pump or any other type may be considered. Pressure must be controlled quite precisely. This means that conventional regulating systems should be used and the averaging control used in level control should be modified by using relief valves and surge tanks to account for flow pulsations.

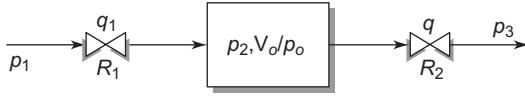


Fig. 8.14 Pressure process; R 's: restrictions; p 's: pressures; q 's: flow rates; V_0 : volume

A more generalized system is shown in Fig. 8.14. The balance equation, in terms of volume flow rates q ; and pressure p_0 at section j , is

$$q_1 - q = C \, dp_2/dt$$

or

$$\frac{p_1 - p_2}{R_1} - \frac{p_2 - p_3}{R_2} = C \frac{dp_2}{dt} \quad (8.60)$$

where C is the capacitance and R_1 and R_2 are lumped resistances taking into account all irregularities on the up and downstream sides. Depending on the types of resistors, R_1 and R_2 can be calculated from the physical dimensions and fluid parameters and $C = V_0/p_0$, suffix "0" being used for nominal conditions. From Eq. (8.60), one gets

$$C \frac{dp_2}{dt} + p_2 \left(\frac{1}{R_1} + \frac{1}{R_2} \right) = \frac{p_1}{R_1} + \frac{p_3}{R_2}$$

or

$$p_2 \left\{ sC + \left(\frac{1}{R_1} + \frac{1}{R_2} \right) \right\} = \frac{p_1}{R_1} + \frac{p_3}{R_2} \quad (8.61)$$

It is to be observed that the output pressure of the tank is the controlled variable. For n number of outlets and m number of inlets, this equation can be generalized for the controlled pressure p_0 , as (see Fig. 8.15)

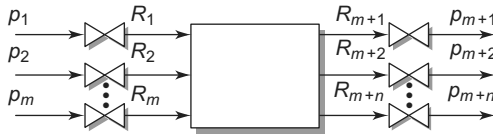


Fig. 8.15 More generalized scheme of a pressure process;
 R 's: restrictions; p 's: pressures

$$\left(sC + \sum_1^{m+n} \frac{1}{R_i} \right) p_0 = \sum_1^{m+n} \frac{p_i}{R_i} \quad (8.62)$$

or

$$(1 + s\tau_p)p_0 = \sum_1^{m+n} \left(\frac{p_i}{R_i} \right) / \sum_1^{m+n} \frac{1}{R_i} = R_t \sum_{m+n} \frac{p_i}{R_i} \quad (8.63)$$

where

$$\frac{1}{\sum_{m+n} \frac{1}{R_i}} = R_t = \text{the effective resistance acting in the process and } \tau_p = CR_t.$$

If there is sonic flow in k_u numbers of inlet resistors, p_0 is decoupled with flow through these resistances. If sonic flow occurs through k_d numbers of outlet resistances, p_0 couples the outlet flow but downstream pressures do not contribute. Or in short, in the inlet and outlet sides, resistances for sonic flow are given respectively as

$$\frac{p_i}{q_i}, i = 1, \dots, k_u$$

and

$$\frac{p_0}{q_j}, j = 1, \dots, k_d \quad (8.64)$$

meaning, from the outlet side, the contribution to the right-hand side of Eq.(8.63) is $n - k_d$ terms, whereas R_t is given as

$$R_t = \frac{1}{\sum_{n-kd+m} \frac{1}{R_i}} \quad (8.65)$$

The combined block schematic representation of Fig. 8.14 and Eq.(8.61) is shown in Fig. 8.16, where $k_{12} = R_2/(R_1 + R_2)$, $k_{32} = R_1/(R_1 + R_2)$ and

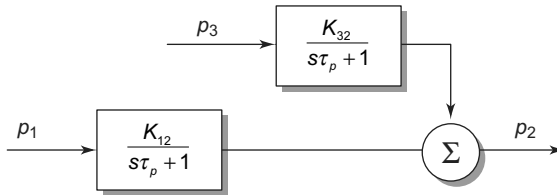


Fig. 8.16 Block representation of Fig. 8.14 (K 's: gain parameters; τ 's: time constant)

$\tau_p = C/(1/(R_1 + 1/R_2))$. The term $\frac{K_{32}p_3}{s\tau_p + 1}$ can be considered to represent the end side disturbance. An equivalent of this with starting end disturbance is shown in Fig. 8.17. It will be seen that the identity of p_1 and p_3 may be interchanged so that disturbance may be associated with p_1 .

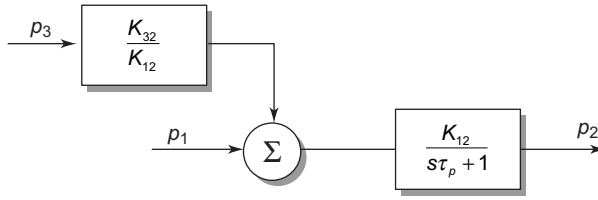


Fig. 8.17 Pressure process with starting end disturbance

We can have the pressure control system as shown in Fig. 8.18. Equation (8.60) is now written as

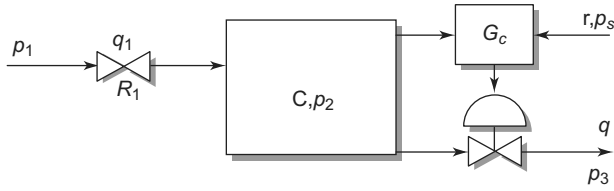


Fig. 8.18 Scheme of the pressure control system; G_c : controller transfer function; p_s : set pressure; C : capacity

$$C \frac{dp_2}{dt} = \frac{p_1 - p_2}{R_1} - k \cdot l \sqrt{p_2 - p_3} \quad (8.66)$$

where l is the valve lift.

For proportional action only in the controller

$$l = K_c(p_2 - p_3) \quad (8.67)$$

where p_3 is the 'set point' pressure.

Then, when Eqs (8.66) and (8.67) are combined, a non-linear equation as obtained below evolves.

$$C \frac{dp_2}{dt} = \frac{p_1}{R_1} - \frac{p_2}{R_2} - kK_c(p_2 - p_3)\sqrt{p_2 - p_3} \quad (8.68)$$

For operation around a nominal value p_0 , linearization, as shown in level control, should be followed and the results of controller settings and response studies be made with disturbance/load variation. Alternatively, Eq. (8.60) can be recast as

$$C \frac{dp_2}{dt} = q_1 - \left[q + \left(\frac{\partial q}{\partial l} \right)_{p_0} l \right] \quad (8.69)$$

with first-order linearization of the flow-lift relation. Thus, flow rates q_1 and q are at average conditions. The parameters q_1 and q can be replaced in terms of p_1, p_2, p_3, R_1, R_2 and l using Eq. (8.67), thus

$$C \frac{dp_2}{dt} = \frac{p_1 - p_2}{R_1} - \frac{p_2 - p_3}{R_2} - \left(\frac{\partial q}{\partial l} \right)_{p_0} \cdot K_c (p_2 - p_3) \quad (8.70)$$

or

$$C \frac{dp_2}{dt} + p_2 \left[\frac{1}{R_1} + \frac{1}{R_2} + K_c \left(\frac{\partial q}{\partial l} \right)_{p_0} \right] = \frac{p_1}{R_1} + \frac{p_3}{R_2} + K_c p_3 \left(\frac{\partial q}{\partial l} \right)_{p_0} \quad (8.71)$$

Then a similar procedure as stated earlier can be followed. Control requirements have already been discussed. These are verified from system Eqs (8.61) or (8.71).

8.5 LEVEL CONTROL

A typical level control system is shown in Fig. 8.19. This is an example of level control in a stirred tank reactor. There are in effect n number of inflows: $q_1, q_2, \dots, q_n$ and one outflow q . The level is to be maintained at h_s , the nominal level being h . The tank area is α and the volume of reactants is, therefore

$$V = \alpha \cdot h \quad (8.72)$$

The amount of liquid that is retained by the tank at any instant is

$$\sum_{i=1}^n q_i - q = \alpha \cdot dh/dt \quad (8.73)$$

If we choose a valve that has linear characteristic (lift-opening relationship linear), $\alpha_v = k \cdot l$, α_v = port area of the valve, l = valve stem lift and

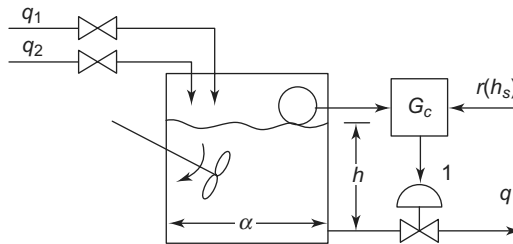


Fig. 8.19 A typical level-control scheme; h_s : set level; l : lift; G_c : controller transfer function

$k = a$ constant. Since the valve has a restriction, q - h relation is a squareroot one such that

$$q = k \cdot l(h)^{1/2} \quad (8.74)$$

Combining Eqs (8.73) and (8.74), one gets

$$\sum_i^n q_i - kl(h)^{\frac{1}{2}} = \alpha \frac{dh}{dt} \quad (8.75)$$

From Fig. 8.19, if $G_c(s) = K_c$ for proportional action only, then

$$l = K_c(h - h_s) \quad (8.76)$$

If h_s is the nominal value of the level, about this a set of linearization equations can be written. However, a different set of operating points with the suffix “0” can be chosen, in which case the incremental linearized approximations of Eqs (8.76) and (8.74) are written, respectively, as

$$\frac{\Delta l}{l_0} = K_c \left[\frac{\Delta h}{h_0} - \frac{\Delta h_s}{h_0} \right] \quad (8.77)$$

and

$$\frac{\Delta q}{q_0} = \frac{\Delta l}{l_0} + \frac{1}{2} \frac{\Delta h}{h_0} \quad (8.78)$$

Similarly the process equation is

$$\sum_1^n q_i - q = \frac{V_0 dh}{h_0 dt} \quad (8.79)$$

which is transformed to

$$\frac{\Delta \sum_i^n q_i}{q_0} - \frac{\Delta q}{q_0} = \frac{V_0}{q_0} \frac{d}{dt} \left(\frac{\Delta h}{h_0} \right) \quad (8.80)$$

with first-order approximation. Thus the incremental block diagram is obtained, as shown in Fig. 8.20, from Eqs (8.77), (8.78) and (8.80).

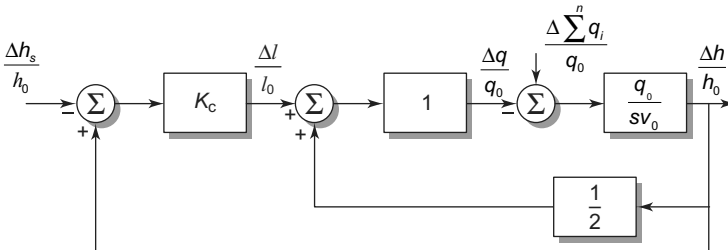


Fig. 8.20 Incremental block diagram of a level control scheme
(l : lift; h : level; suffix 0 for nominal value; suffix s for set value)

This can further be simplified following elimination; as from Eqs (8.77), (8.78) and (8.80)

$$\frac{d}{dt} \left(\frac{\Delta h}{h_0} \right) = \frac{q_0}{V_0} \left[\frac{\Delta \sum_n q_i}{q_0} + K_c \frac{\Delta h_s}{h_0} - \left(K_c + \frac{1}{2} \right) \frac{\Delta h}{h_0} \right]$$

or

$$\frac{\Delta h}{h_0} \left[s \frac{V_0}{q_0} + \left(K_c + \frac{1}{2} \right) \right] = \frac{\Delta \sum_n q_i}{q_0} + K_c \frac{\Delta h_s}{h_0}$$

or

$$\frac{\Delta h}{h_0} \left[\frac{1}{2} \left(2s \frac{V_0}{q_0} + 1 \right) K_c \right] = K_c \frac{\Delta h_s}{h_0} + \frac{\Delta \sum_n q_i}{q_0} \quad (8.81)$$

In Eq. (8.81), the first term on the right-hand side shows the incremental set point change and the second term shows the inflow disturbance. The simplified block diagram is now shown in Fig. 8.21. This shows that the static gain of the process is 2 and the linearized approach gives a process time constant $2V_0/q_0$. The quantity V_0/q_0 can be easily seen to be the residence time of the liquid at nominal conditions and is known as the nominal residence time τ_0 . The process itself has a filtering characteristic and filters high frequencies whether such changes occur in inflow or set point. Actual level control is now of two different types, namely that due to set point change and that due to disturbance in inflow. The transfer functions in the two cases are easily obtained from Eqs (8.81), respectively, as

$$\frac{\Delta h(s)}{\Delta h_s(s)} \Big|_{h \rightarrow h_0} = \frac{K_c}{K_c + \frac{1}{2}(1 + 2\tau_0 s)} = \frac{2K_c}{2K_c + 1 + 2\tau_0 s} \quad (8.82)$$

and

$$\frac{\Delta h(s)}{\Delta \sum_n q_i(s)} \Big|_{h \rightarrow h_0} = \frac{h_0}{q_0} \cdot \frac{2}{1 + 2K_c + 2\tau_0 s} \quad (8.83)$$

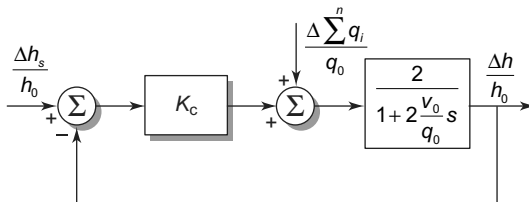


Fig. 8.21 Simplified block diagram of Fig. 8.20

When the set point is altered, frequency is unlikely to be large such that low K_c would be good enough to provide adequate control. If, however, inflow variations occur, it may have a high fluctuation rate. Rearrangement of the right-hand side of Eq. (8.83) gives

$$\frac{h_0}{q_0} \cdot \frac{2/K_c}{2 + (2\tau_0 s + 1)/K_c}$$

which shows that a high K_c would be required to provide the desired control action.

It will be seen that linearization of the system about its nominal conditions is essential as otherwise non-linear equation is derived even with a linear control valve as is obtained from Eqs (8.75) and (8.76)

$$\sum_n q_i - kK_c(h - h_s)h^{1/2} = \alpha \cdot dh/dt \quad (8.84)$$

and simple control loops cannot be established for a satisfactory wide range control of the level in a tank. A better way to control is by incorporating two controllers, i.e., with a multiloop system as in the case of the level control of a reboiler as shown in Fig. 8.22. Exact control as per Eq. (8.84) may not be carried out here because of independent control loops, but then it is not that critical.

When the level itself is not that critical and output is not allowed to change suddenly for a sudden change in the inflow, a different control known as averaging control is adopted which is primarily meant for smoothening flow fluctuations. This may occur if (any one or more of) the major inputs q_i suddenly stop because of stoppage of flow from the previous

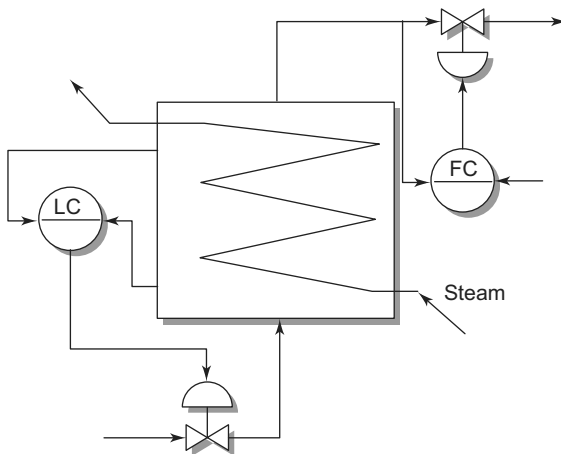


Fig. 8.22 Level-control scheme of a reboiler

reaction stages. Eqs (8.77), (8.78) and (8.80) may now be reorganized to obtain

$$\frac{\Delta q}{q_0} \left[1 + \frac{2V_0 s}{q_0(2K_c + 1)} \right] = \frac{\Delta \sum_n q_i}{q_0} - \frac{2K_c V_0 s}{q_0(2K_c + 1)} \frac{\Delta h_s}{h_0} \quad (8.85)$$

When the step point does not alter,

$$\frac{\Delta q}{\Delta \sum_n q_i} = \frac{1}{1 + \left(\frac{2\tau_0 s}{2K_c + 1} \right)} \quad (8.86)$$

This shows that as K_c is small, the variation of q with q_i is less, while a large K_c tends to increase this dependence. When outflow does not vary suddenly with a variation in inflow there may be a large change in level (h), but for the required tank this should not cause it either to overflow or empty out. The choice in this case depends on the size of the tank and then the value of K_c or $1/K_c$, i.e., the proportional band, and this is usually large for averaging control. In the conventional steady level control, as much variation in q as in q_i is desired, in which case K_c should become increasingly larger. This also increases the frequency response.

8.6 TEMPERATURE CONTROL

8.6.1 Introductory Remarks

Temperature control systems are not simple enough to be formulated by simple strokes. The sluggishness, exaggerated distributive effect and complicated heat transfer phenomena make it difficult to handle these systems by simple control methods. A study of the dynamic behaviour of thermal systems is all the more difficult because of the possible inaccuracy in the development of an adequate mathematical model of the system.

The majority of industrial processes in which temperature measurement and control are needed, are divided into two distinct types involving typical heat transfer problems. These are (i) single-phase systems and (ii) two-phase systems. In single-phase systems heat transfer takes place between single phase fluids, and in two-phase systems, heat transfer is between two phases. Often the situation becomes complex. In some cases the above conventional description does not apply and such systems are considered by the instrument engineers to be the most difficult processes. A typical example is the reheating of a furnace in a rolling mill.

The systems in which both the fluids between which heat transfer takes place, are in a single state, are more common industrial heat exchangers of

the double pipe or shell and tube type. Boilers and condensers are of the other type in which one of the fluids is in the two-phase systems (boiling or condensing). They are not only difficult to model but their control is also quite complicated and often the requirements are not fully met. In general, the temperature control scheme is complex. In Chapter 6 a few complex schemes have been discussed. These are very useful for the control of heat exchangers. The schemes for boilers are more rigorous and such control problems have been tried to be solved in Chapter 11. For ready reference, a few standard techniques of controlling heat exchangers are schematized by block diagrams and typical implementation schemes are suggested.

The dynamic behaviour of heat exchanger problems is additionally complicated by the simultaneous existence of mass transfer and heat transfer. The details of the complexity of different types of single-phase and two-phase systems and approximate models of them for the study of their dynamic behaviour are too extensive to be elaborated here. However, it is important to note that of the three possible ways of heat transfer, conduction, convection and radiation, “forced” convection plays the major or perhaps the only dominant role in heat exchangers. The heat transfer coefficient (h.t.c) that is to be considered for these problems will thus be theoretically and experimentally shown that this h.t.c. is directly proportional to the mass flow rate but inversely proportional to the viscosity (μ). With rising temperature, viscosity for liquids becomes less and h.t.c. increases, but usually this is small, as the decrease of μ with rising temperature is not large for most of the flowing fluids of industrial importance, such as water under pressure.

Whatever may be the situation, for dynamic behavioural study one has to formulate the mass balance and heat balance equations in a system and then from that estimate the various transients. Finally, the temperature to mass flow rate transfer functions are formulated and the control schemes are then considered. The difficulty in the process is, however, non-linearity and for simple control rules to be applied linearizing should be attempted first. The approximate linearized models are suitable for adaptation with instrumentation systems but even then single loop control is hardly ever suitable in most of the circumstances.

8.6.2 Elementary Control Schemes

The tube exchanger shown in Fig. 8.23 explains a simple situation of temperature regulation. Steam outlet temperature t_0 is to be controlled by the controller by regulating the mass flow rate w of what is known as “utility stream”. The inlet steam temperature t_i , steam flow rate w_s , utility stream temperature t_u and the pressure differential in the utility line in the exchanger $\Delta p = p_{ui} - p_{u0}$ are the disturbance factors. In these w_s and $p_{ui} - p_{u0}$ may vary sharply and pose more problems in control rule selection. The

block schematic representation is shown in Fig. 8.24. In this scheme the transfer function is easily calculated as

$$\frac{t_s(s)}{t_{\text{set}}(s)} = \frac{\{(w_s G_s + \Delta p G_p)/t_{\text{set}}\} + G_{ca} G_p}{1 + G_m G_{ca} G_p} \quad (8.87)$$

Of the various transfer functions, the G 's of the block G_s and G_p are important in exchanger problems. Others are typically known or are adjustable. In fact G_{ca} has to be adjusted in relation to G_p and G_s , specially G_p , the process transfer function $G_s(s)$ is given as

$$G_s(s) = t_0(s)/w_s(s) \quad (8.88)$$

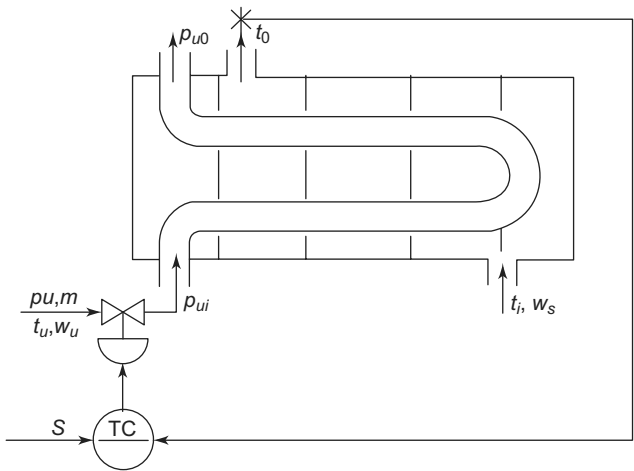


Fig. 8.23 Schematic of the tube exchanger; S : set; t_0 : temperature

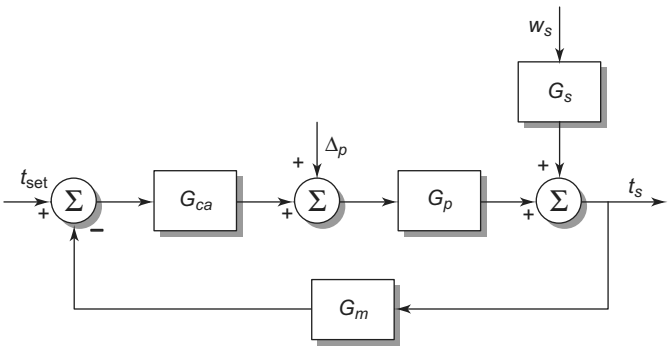


Fig. 8.24 Block schematic representation of Fig. 8.23

In the most common situations

$$G_p(s) = k_1 t_0(s)/w_u(s) \quad (8.89)$$

Also

$$G_{ca}(s) = G_c(s)G_a(s) = G_c(s)k_2/(1 + s\tau_a) \quad (8.90)$$

$$G_m(s) = k_3/(1 + s\tau_m) \quad (8.91)$$

If, now $G_p(s)$ is at least a first order system, which, in fact, is always true, a third order system results even if the controller chosen is only of the proportional action type. This requires that the stability be considered carefully and the response speed has now been considered by comparing the time constants of the individual elements particularly the second largest one. The basic difficulty faced in temperature control with this set-up is the shift in the performance characteristics with t_0/w_s . Usually, when t_0/w_s is the highest and the controller is set at that level for stable operation, the operation tends to be very sluggish whenever t_0/w_s falls.

Basically thermal process can be considered as a heat flow process for which transport equations are to be formed and their analysis made. But this direct approach fails to provide a good insight into the dynamic behaviour of the process and hence approximation techniques are resorted to. A very common technique of approximation is by representing the process in simpler partial differential equations via what is known as Taylor diffusion model. Representation can also be made by simpler transcendental equations. The other technique is the rational approximation. The results obtained through Taylor diffusion model is very instructive in so far as the overall process transfer function is approximated very closely by taking as many first order terms in series, as is necessary for the purpose. This model, however, is only a low frequency representation of the system behaviour and the results show that the heat or temperature distributions are expressed in terms of modes into which the decomposition of these distributions have been made.

Example 1

In a flow control system, for 60% of the maximum flow, the drop across the valve is one-third of the total drop for maximum flow in the system. If the percentage lift of the valve is to be the same but flow rate is to be increased to 75%, obtain the change in the percentage drop across the valve required. Make adequate comments on the controllability of the system.

Solution Using Eq. (8.56)

$$\left(\frac{q}{q_m}\right)^2 = y^2/[\beta + (1 - \beta)y^2]$$

one gets

$$0.6^2 = y^2/\left(\frac{1}{3} + \frac{2}{3}y^2\right)$$

yielding

$$y = 0.397$$

Hence the valve lift is 39.7%.

Using the same equation again

$$0.75^2 = (0.397)^2 / [\beta + (1 - \beta)(0.397)^2]$$

yielding

$$\beta = 0.14, \text{ i.e., } 14\%$$

Hence the percentage in pressure drop is

$$\left\{ \left(0.14 - \frac{1}{3} \right) / (0.14) \right\} \times 100, \text{ i.e. } -136\%$$

Such a large change in pressure drop evidently points to a large deviation from good controllability.

Example 2

A tank has a normal depth of 2 m and an area of 1 m². At the normal depth and normal flow rate the level is to be maintained within 1 % for a 20% change in inflow over the normal flow rate of 0.6 m³/min. Using proportional control estimate the controller gain when flow variation is (i) sudden and (ii) also varies at the rate of 2 times/min.

Solution From Eq. (8.33)

$$\left| \frac{\Delta h}{h_0} / \frac{\Delta q_i}{q_0} \right| = \left| \frac{2}{1 + 2K_c + 2\tau_{0s}} \right|$$

Now, normal volume of tank, $V_0 = 2 \times 1 = 2 \text{ m}^3$

$$q_0 = 0.6 \text{ m}^3/\text{min}$$

Residence time is, therefore, $\tau_0 = 2/0.6 = 3.33 \text{ min}$

Now,

$$(\Delta h/h_0) = 0.02/2 = 0.01 \text{ and } (\Delta q_i/q_0) = 0.04/0.6 = 0.067$$

For sudden change $s \rightarrow 0$, one gets

$$0.01/0.067 = 2/(1 + 2K_c), \text{ giving } K_c = 6.2$$

For a rate of 2 time/min, $f = 2 \text{ min}$, $\omega = 6.28 \text{ r/min}$

$$1/6.7 = 2/[(1 + 2K_c)^2 + (6.28)^2(3.33)^2]^{1/2}$$

This gives $K_c = [(153.76 - 430)^{1/2} - 1]/2$ which is imaginary. Thus with proportional gain setting alone this control cannot be made effective.

Review Questions

1. What are the equivalent elements in the transmission of signal and process materials and how are they found?
2. What are the basic differences in a flow control system and a temperature control system? Why and how are transmission lags cared for in flow control loops?
3. Establish the similarity between pressure control and level control. Why is the pressure control system a more general system?
4. What is an averaging control? How is it made effective? Where is such a type of control required?
5. Why is temperature control rather difficult in certain processes? Give example of such a process and suggest methods of meeting such stringent control requirement.
6. In an heat exchanger which one is more predominant—heat transfer or mass transfer? Can you suggest a system where both are equally important? How is such a system controlled effectively for temperature as a controlled variable?
7. In which processes are controllers and control valves mounted at the same place and why?
8. A flow control system has an actuator lag of 0.3 sec, process lag of 5 sec and a measurement lag of 0.2 sec. Calculate the response to load change of a sinusoidal nature for PI controller when $K_c = 4$ and $T_R = 1$ sec.
9. The level in a tank is maintained at a nominal value over which a tolerance value of 2 per cent may be accepted. If a proportional gain is at 10, by what percentage a sudden change in set point may be allowed?

(Ans: 2.1%, Hint: Use Eq. (8.82))

Computer Control of Processes

9.1 INTRODUCTION

Computer control of processes has been primarily motivated from the consideration of optimizing control operations of an entire plant or parts of it. From the users' viewpoint, optimisation means the optimisation of a performance function in relation to productivity and product quality, or in short, economy. Start-up and shut-down are additional functions to be performed by the computer in conformity with other terminal conditions called constraints. In a process, optimization would also consist of the efficient functioning of an enormous bulk of information (such as tapping, acquisition, assimilation, analysis and dissemination) relating to the conservation of material and energy with accuracy, speed and flexibility using an adequate and proper control strategy. This can only be made possible with the help of a suitable digital computer.

A modern computer has an enormous capacity and it is possible that a number of processes may be controlled with individually selected control constraints by such a computer following a predetermined sequence. This is known as sharing or time-sharing.

Simple details of the installation of a computer in a control centre are shown in Fig. 9.1. Blocks 4 to 10 are actually peripherals of the computer (3). Also, blocks 5 and 10 may often be combined in principle. Block 3 provides data-logging facilities.

9.2 CONTROL COMPUTERS

The effective and efficient control of a process with a computer requires that the computer communicates with the process equally efficiently. This

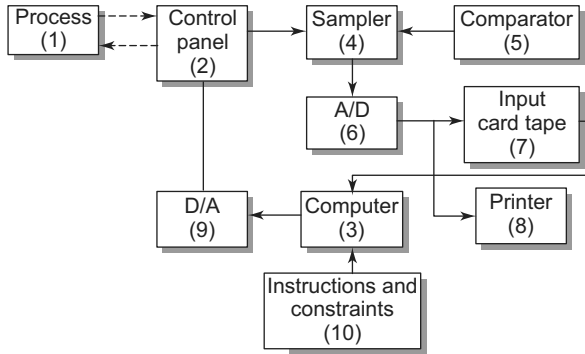


Fig. 9.1 Block diagrammatic representation of computer installation scheme in a control centre; A/D: analogue-to-digital converter; D/A: digital-to-analogue converter

is possible when: (i) the input and output channels of the computer are compatible with the plant outputs and inputs through the scanners provided with flexibility and a degree of control and (ii) the computer has large storage capacity, high speed, appropriate command and word structure, operational flexibility and versatility in software handling.

The features of a control computer are shown in Fig. 9.2. The heart of the computer is the central processing unit (CPU) which consists of the control

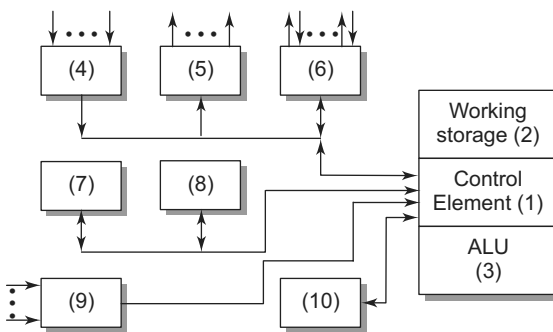


Fig. 9.2 Features of a control computer; 1: control element; 2: working storage; 3: arithmetic and logic unit; 4: inputs; 5: outputs; 6: telemetry channels; 7: printers; 8: bulk back-up storage; 9: interrupt; 10: console

element (1), working storage (2) and arithmetic or logic unit (ALU) (3). Additionally, the peripherals consist of the input (4), the output equipment (5), signal channels (6) and the print-out devices (7). For large works additional data storage facilities (8) are also provided. Then for checking the control action and special instruction, an interruption module (9) is also accommodated. Finally, the man-machine interface is the console (10) which, in general, consists of a set of input output devices. It follows the information and instructions to move to and from the CPU and blocks (4), (5) and (6). In between, registers are usually provided. Registers are of two kinds: (i) control registers and (ii) data registers. Control register is the control warehouse which receives instructions from the CPU regarding addressing (input calling), mode of operation, timing, initiation of converters, output sequencing, etc. Computer capability is enhanced by the registers.

9.2.1 Basic Functions of the Computer System

A computer system for process control which is shown by a very general block diagrammatic representation in Fig. 9.2, may be used to perform an infinite variety of functions. Some of the basic functions may be listed as:

- (i) data acquisition, both analogue and digital,
- (ii) data conversion with scaling and checking,
- (iii) data accumulation and formatting,
- (iv) visual displays,
- (v) comparing with limits and alarm raising,
- (vi) recording and monitoring of events, sequence and trend,
- (vii) data logging,
- (viii) computation, and
- (ix) control actions.

Data Acquisition

A knowledge of the computer functions in general would help to understand the functions specified above. By way of example, the data acquisition function may be considered. In case of analogue data, an infinite number of states is possible, whereas for digital data, only a finite number of states is known which may be as low as two in the binary form. The operations for the acquisition of analogue data are:

- (i) input point selection,
- (ii) setting of scale factors,
- (iii) storage location selection,
- (iv) sampling frequency selection and conversion of the data to the digital state.

For digital data, parts of steps (ii) and (iv) are superfluous. This processing is expected to bring in a certain amount of error in the system. Knowing

the accuracy and delay in each individual block the error can also be calculated. The time interval of sampling T , as also the delays in scaling and conversion, τ_1 , and in data shifting (from input/output register to storage location), τ_2 , are also important. A good acquisition system will obviously connote $T > \tau_1 + \tau_2$. The error can also be shown to decrease if the sampling interval is sufficiently close.

The input-point selection and sampling may be simultaneously performed by a scanner, the block schematic of which is shown in Fig. 9.3. The method of scanning or multiplexing shown here is for ready signals or data with appropriate scaling, matching and filtering. Often non-electrical data are present which require to be conditioned before they can be fitted into the termination block of the computer by what is known as signal conditioning when part-scaling is automatically effected. After the termination the rest is further conditioned by what is known as matching and filtering. Although this is not mentioned in the operation involved in the acquisition process, it can always be considered to be implied.

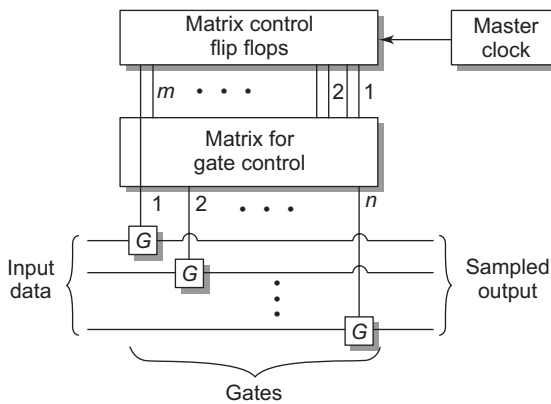


Fig. 9.3 Block schematic of a scanner

There are a few principles on which A/D converters are made. Of these the dual-slope integration principle is simple and popular but has comparatively less speed. The successive approximation converter has a faster speed of operation. In this a sequence of voltages having a binary coded decimal (8421 BCD usually) weighted code are successively compared with the input signal. Each successive voltage level is stored or rejected depending on the signal level in its most significant digit which is thus first approximated. Digit-by-digit comparison is thus effected to obtain the desired accuracy and precision. A typical analogue data acquisition channel is shown schematically in Fig. 9.4, starting from the termination to the registers. For the digital data, blocks A/D converters and limit comparators are not necessary as such. However, limit comparators are often not eliminated for level-selection purposes.

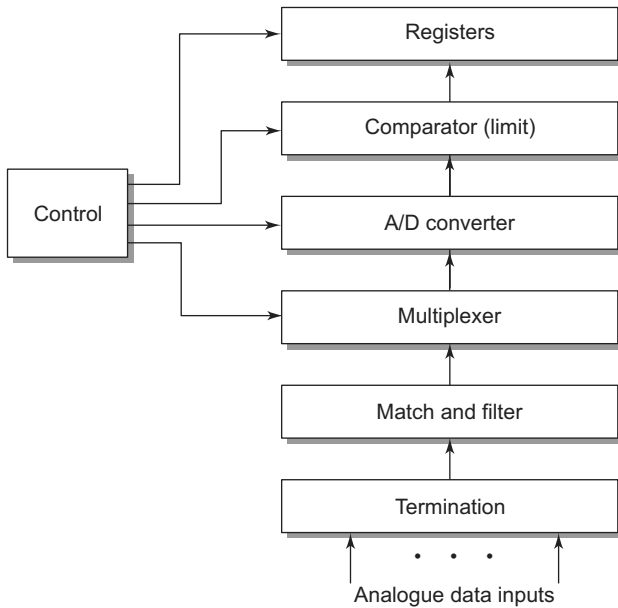


Fig. 9.4 Scheme of an analogue data acquisition channel

Comments on Other Important Functions

Alarm raising is one of the very important functions. This is a function valid only in exceptional conditions. It involves comparing the value of any point under consideration with high or low limit set values of the point, raising an alarm when the value exceeds the limits for the first time, printing out the details including time when it occurs, shutting off the alarm when the value has returned to within limits and reprinting the normalized conditions.

Computation is perhaps earmarked as the major function which a computer is required to perform. This is basically an arithmetical computation done by its arithmetic unit. The problem stated in mathematical form is first required to be transformed into a set of arithmetical operations through what is known as algorithm and the computation then follows. Besides the algorithm, computation also equires programming in its complete format.

Finally, computer control actions need to be considered. Here discrimination has to be made regarding the type of action. Two main basic actions are

- (i) set point control (SPC) and
- (ii) direct digital control (DDC).

Somewhat involved in these modes, but to a certain extent independently listed, is the optimal control. It will be shown in the subsequent chapter (Chapter 10) how optimal control is made effective.

Computer control action does not imply only deciding upon the required control through computation and initiating these actions through the output devices, but making provisions to check that the actions decided upon are safe.

Besides, when a computer is used in a process, the optimization of the process control should also be looked into. This can be done by making the mathematical formulation of a functional, from the dynamic process equations along with some constraints, whose optimum is sought through computation via the algorithm.

9.2.2 Direct-digital Control

In the set point control, the set points of the conventional controllers are adjusted by the computer from the process data and constraints following either a given programme (programme control) or on the basis of the required optimization of a given “cost” functional. In direct digital control, the control equipment reduces to the transducer and actuator (special types); controller, comparator, limiting and other safe-guarding operations/actions are provided by the digital computer itself. Control algorithm obviously has to be especially prepared for this. The schematic of a direct-digital control is shown in Fig. 9.5.

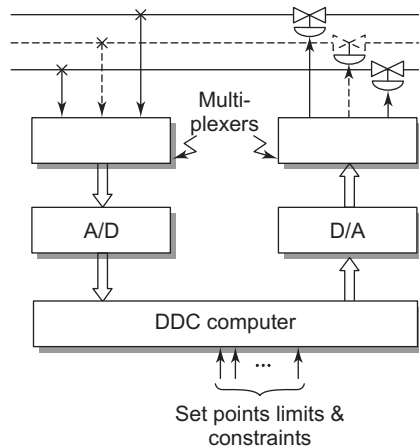


Fig. 9.5 Schematic of a direct-digital control (A/D: analogue to digital; D/A; digital-to-analogue; DDC: direct-digital control)

Since the control action is taken by the computer itself, the control equation can be chosen to suit the dynamic characteristics of the system (process) and thus it need not be limited to a 3-term control action only. The only constraint is that the chosen strategy must be programmable for the computer.

When the system complexity changes, the programme alone needs to be altered and this can be done without any change in the hardware. The programme may be designed for safe and consistent operation of the process as well. The rate of change of the controlled variable may be checked at stages and altered at will by a suitable programming schedule comparing with a given set of limiting values. This adaptability has made the direct-digital control quite attractive.

When a process gets more complex and elaborate one can easily check that a direct digital control may even be economically more viable considering initial equipment cost, operational cost and savings from performance improvement and modifications.

As already mentioned, many complex control functions can be implemented through the digital control algorithms. A typical example of the *PID* control function algorithm is

$$m_n = k_c e_n + \sum_n k_1 e_n (\Delta t) + \frac{k_d (e_n - e_{n-1})}{(\Delta t)} \quad (9.1)$$

where m is the manipulating variable, e is the error, Δt is the sampling interval and k_s are the constants, while the suffixes denote the order of sampling. In Section 9.5 this formulation has been elaborated in a more comprehensive manner. In general, an algorithm for a typical single loop linear control can be written as

$$m_n = \sum_{j=1}^{n-1} \alpha_j (m_{n-j}) + \sum_{j=1}^{n-2} \beta_j (e_{n-j}) \quad (9.2)$$

where α_j and β_j are constants.

9.2.3 Specifications

The optimum size of a computer in economic terms can be determined by the complexity and 'levels' of control. The levels are discussed briefly in the next section. Usually a computer is specified in terms of the characteristics which illustrate its capability. A few of the counts are:

- (i) Inputs/outputs (I/O): type—paper/magnetic tape/card/modules, etc., capacity of the I/O devices.
- (ii) Memory, storage: type and size, normal and back-up.
- (iii) Speed: cycling, operational.
- (iv) Arithmetic unit: size of words in binary digits including/excluding sign.
- (v) Instruction: address length in words, number of basic commands.
- (vi) Interface facility: existing or to be supplemented, capacity, range and conversion speed, etc.

These statements are explained here in detail.

- (i) Input/output (I/O) devices have both resource and handling requirements. A single programme may be required to serve different purposes needing different I/O devices. There may always be a number of programmes. The handling requirement is specified in terms of the number of signals/sec. Whereas the resource requirement may be in terms of different types and amount of storage as also peripherals like printers, typewriter, LCD/LED displays, other storage units, etc.
- (ii) Storage is again a resource for the programme. It is required for storing data, constants, instructions as well as for working storage. Storage is expressed in words. If for j th programme the required storage number is w_j words, for n programmes the total storage required will be

$$T = \sum_{j=1}^n w_j \quad (9.3)$$

- (iii) Speed, as mentioned, usually refers to the cycling speed (time) in the case of a general purpose computer. This, in modern days, is a very small figure, a typical value being 150 nsec. However, for the process control computer the speed is specified in terms of the time of execution. This involves the operational time-time for operations such as adding, subtracting, multiplying and dividing. In a usual specification schedule, these times are often mentioned. However, execution time does not exactly mean the time of some individual operations. It depends on factors like data bulk, interruption as well as the programme. In fact, the contribution of data to total execution time is variable. Total execution time is, therefore, broken down into a number of parts, namely, the fixed part, t_F , the variable part (a function of data) being given by $t_v(d_i, t)$, $i = 1, 2, \dots, k$, where d_i represents the data and the sum total of the duration of the interruptions which is represented by $= \sum_{j=1}^n t_{lj}$.

Thus, the total execution time of the p th programme is given as

$$T_p = t_{Fp} + t_v(d_{p1}, d_{p2}, \dots, d_{pk}, t) + \sum_{j=1}^n t_{lpj} \quad (9.4)$$

It is interesting to note that the execution time of a programme also depends on the storage W used by the programme as also on the input/output resources required (D) by the programme. With an increase in both D and W , T is shown to decrease. The implicit relationship is given by

$$T_p = F(W_p, D_p) \quad (9.5)$$

Fig. 9.6 gives the relationship for a typical programme. This background knowledge may be considered adequate when a specification is to be read. A typical specification sheet for an elementary process computer is given as follows.

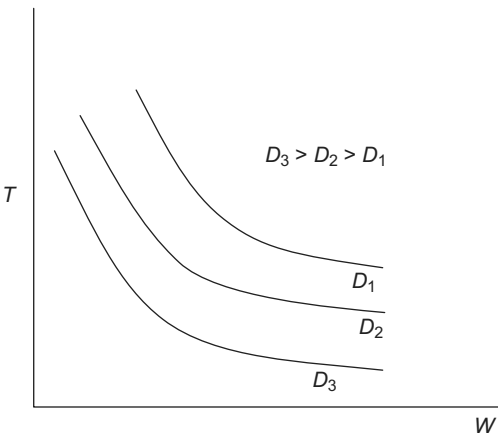


Fig. 9.6 Plot of the relation $T = F(W, D_p)$ for a typical programme (D_1, D_2, D_3 : resources)

Specifications

- (i) Input/output : More than 5000 variables/sec.
- (ii) Storage or memory : Core for 4096 to 8192 words with access time of 0.20 μ sec.
Drum for programme of 8192 to 16384 words with access time of 0.14 msec.
- (iii) Speed : Cycling, 1 μ sec. In view of the variable execution time of programmes, operational times are quoted
Add or subtract 0.10 msec
Multiply 0.70 msec
Divide 1.00 msec
Store 0.6 msec
- (iv) Arithmetic unit : Parallel operation type, clockrate of 50 kHz. For word size see next item.
- (v) Instructions : Single address, one instruction per word of 8-bit operation code 16-bit address; word length 24 binary digits.
Basic commands may be 60 instructions including 10 special.

- (vi) A/D and D/A : A/D-5000 conversions/sec with 1 Volt
conversion input 16-bit output
D/A-50 conversions/sec.

9.3 PROGRESS IN COMPUTER CONTROL IN PROCESS INDUSTRIES

From the time the digital computer was introduced in industrial process control about forty years back, there has been substantial progress and now the digital computer has almost become an integral part of and is nearly indispensable in a few plants like steel industries, refineries, etc. Earlier limitations on direct digital computer control of processes were:

- (i) small memory sizes,
- (ii) slow machine speeds, and
- (iii) poor reliability.

Then the applications were limited to supervisory control, where

- (i) data logging,
- (ii) monitoring, and
- (iii) alarm raising were the main functions performed by the computer.

Next came the closed-loop supervisory control—the digital computer was used to bring about economic optimization in steady state, the dynamic control being performed by a primary loop. Thus, static control by altering the dynamic controller's set-point was made use of for better economic returns and naturally the performance capability of the computer was not fully utilized (Fig. 9.7). As already mentioned such a control scheme is known as set point control or, in short, SPC.

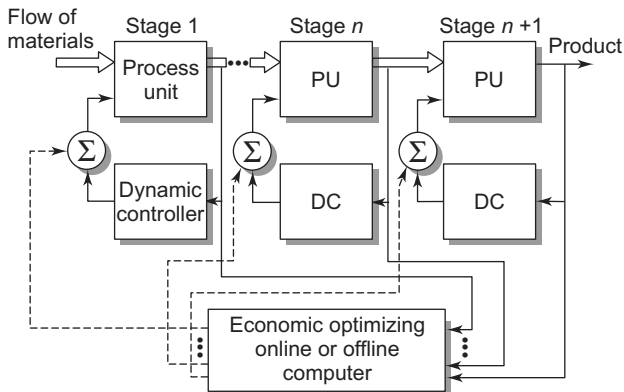


Fig. 9.7 Features of a set point control scheme with digital computer;
PU: process unit; DC: dynamic controller

With increasing memory sizes, speed and reliability, the small size and low cost being additional, the digital computer was required to perform the dynamic control, this hook-up being known as direct digital control, in short, DDC (Fig. 9.5).

Finally, there appeared the hierarchy control where several computers at different levels operate to control a large “complex” of an organization. These levels include “from management decision to valve lift” (Fig. 9.8). The decisions at all levels are important and are taken care of simultaneously. The three different levels that are truly representative of any organization using a hierarchy control are as follows.

- (i) **Company-level control (CLC):** Company production strategy, on the basis of time, is involved, i.e., (a) for short-term control, linear programmes are good enough and (b) for long-term control, dynamic optimization methods are adopted.
- (ii) **Plant-level control (PLC):** (a) Non-linear programming as a steady-state optimization medium is usually chosen—the exponential

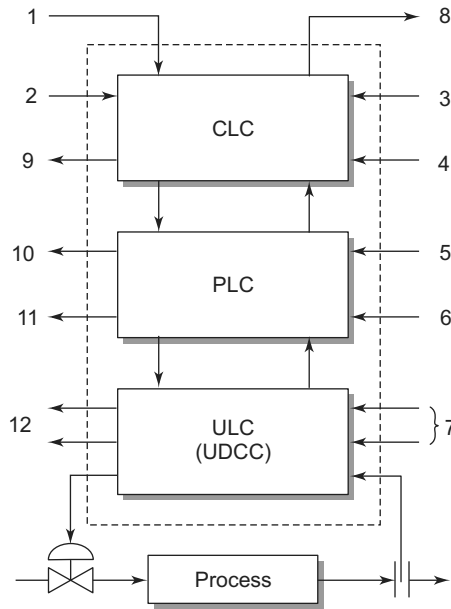


Fig. 9.8 Scheme of hierarchy control; CLC: company level control; PLC: plant-level control; ULC: unit-level control; UDCC: unit-direct-control computers; 1, management instructions; 2, sales order data; 3, related industrial economic data; 4, data links from other plants for production and inventory data; 5, data links from other plants; 6, unit-direct control computers; 7, input from sensors; 8, data to management; 9, production schedules, shipping instructions, etc.; 10, set-point data to plant; 11, unit-direct-control computers; 12, to control elements

nature of temperature, time, process reactions, etc. necessitate this. In the approximation, linear programming or even a gradient search method is also considered, (b) Dynamic optimization schemes of plant start-up, shut-down, etc. need to be made.

- (iii) **Unit-level control (ULC):** Usually dynamic optimization at the unit level is considered; some examples are feed forward or adaptive control particularly for non-linear process dynamics based on energy and material balance.

Although hierarchy control is the ulterior motive in computer control in which all the levels are simultaneously functional, the major stress is still on plant-level control where a direct digital control is adopted. In this situation the unit-level control merges with the plant-level control and a single computer is required to look after the entire process. It will be subsequently seen that with more advancement in digital soft and hardware technology plant-level control has acquired a new dimension where unit-level controls are performed by the microprocessors or microcomputers and a control centre installation of a digit computer works on the plant-level basis.

An idea of speed increase of computers during the first two decades is shown in Fig. 9.9 and that of reliability in Fig. 9.10. It may be mentioned that except for minor variations in the specifications the process control computer hardly varies in structure from a general purpose computer. The memory size has increased from 3K in the first ever process control computer (RW 3000) with drum type memory to at least ten times its original size in the seventies and the development of newer forms of memory modules have enormously reduced the size as well.

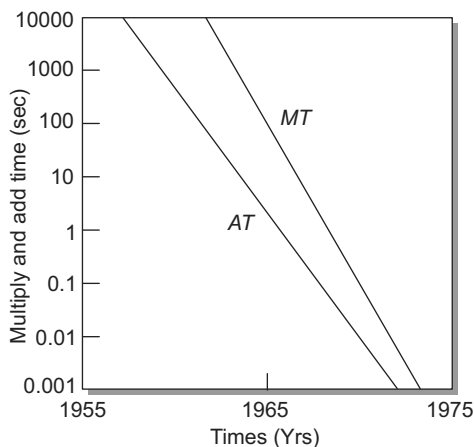


Fig. 9.9 Plot showing the speed increase of computers in first two decades; MT: multiply time; AT: add time

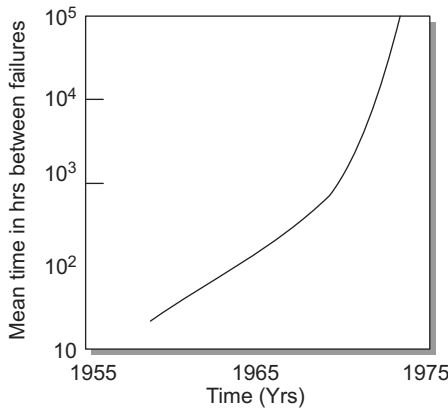


Fig. 9.10 Plot showing the increase in reliability

9.4 CONTROL ON LEVEL BASIS

9.4.1 Unit-level Control

Both unit-level control and plant-level control, as already pointed out, can be either set-point control (SPC) or direct-digital control (DDC). Company-level control, however, has to be treated on a different footing although DDC is more economical as well as convenient.

We can start a unit-level control to show how it can be adapted as either SPC or DDC. Consider the example of a reboiler. For an SPC the diagram is drawn as shown in Fig. 9.11. The arrangement is shown such that the computer in relation to the inputs:

- (i) feed-inflow rate (mass), q_1 ,
- (ii) inlet flow temperature, T_1 ,
- (iii) outlet flow to temperature, T_2 ,
- (iv) feed specific heat, C , and
- (v) fuel heat content, H would control the set point for the fuel controller FC.

The basic relation which the computer is required to solve for setting fuel controller can also be deduced from the material balance equation. Neglecting losses one has

$$\text{Feed-in} = \text{Feed-out}$$

so that

$$F \cdot H = q_1 C (T_2 - T_1) \quad (9.6a)$$

Hence,

$$F = q_1 C (T_2 - T_1) / H \quad (9.6b)$$

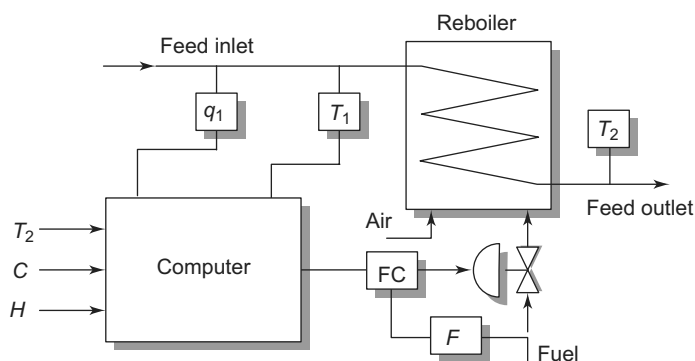


Fig. 9.11 Set point control scheme in a reboiler;
FC: flow controller, F: flow transducer

The main object of this control is to keep T_2 at its value. The computer receives T_2 from a setter and fuel conditions from the fuel analysis cell—these may have to be changed when fuel grade changes. Other input parameters q_1 and T_1 are continuously fed to the computer. Thus conditions of fuel and feed determine the flow of fuel which in turn is controlled through the fuel controller. For a direct digital control in the system, fuel flow rate appears both in the input and the output as shown in Fig. 9.12 and obviously, the analogue controller FC of Fig. 9.11 is dispensed with.

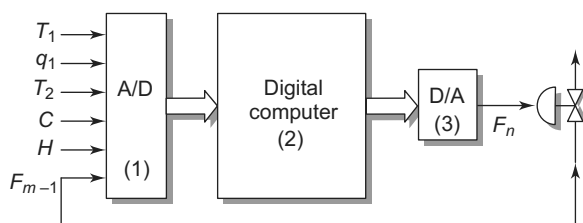


Fig. 9.12 Direct digital control scheme for the system of Fig. 9.11; F_{m-1} : fuel flow rate at $(m-1)$ th state, F_n : fuel flow rate at n th state.
 T 's : temperatures, H : heat, C : specific heat, q : mass flow rate

9.4.2 Plant-level Control

For a plant-level control a distillation plant is chosen. The main part of the plant is the fractionating tower. A case of modelling of a distillation column has been discussed in Chapter 2. There are variations in the process depending on the modes of operations such as

- (i) constant overhead product,
- (ii) constant bottom product rate, constant reflux rate,
- (iii) constant reflux rate, constant vapour rate,
- (iv) two point composition control, and so on.

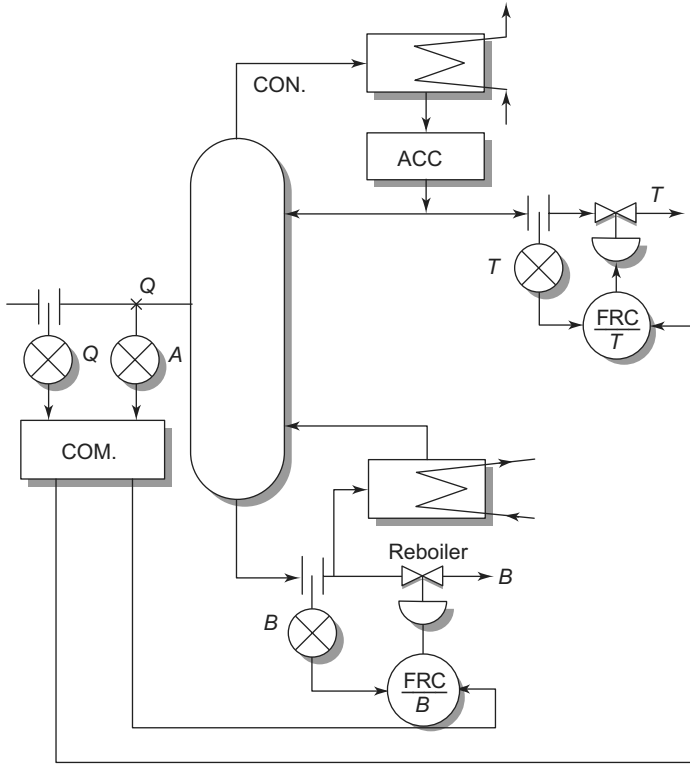


Fig. 9.13 The computer control scheme of a distillation process, B: bottom product, T: top product, ACC: accumulator, CON: condenser, COM: computer

A typical case here has been dealt with where top and bottom products of a distillation column are varied with variation of feed input quality and quantity. It is a case of feed forward control and the computer has been used to adjust the set points of the product flow controllers as a function of feed flow and composition. Fig. 9.13 shows the scheme of the system. Feed-forward control has been suggested for dead time and long time-lags. The set point control programme is given as

$$SPT = T \cdot Q / NU + SPT$$

$$SPB = B \cdot Q / NU + SPB$$

where $NU = (1 + \text{INDEX}) / \text{TIME}$, and $\text{INDEX} = \text{ITER} + 1$.

With this programme the feed forward control moves the set-points by one increment towards their final positions.

However, when a digital computer is to be brought into use, a solution of the above problem in such a straightforward manner will not justify the use of the computer itself. One will naturally be inclined to get the

process outputs optimized. This nomenclature of optimization is again very general, and the basic things which people want are already stated, namely (i) optimum (maximum) profit, and (ii) optimum product quality.

General Approach for SPC

When optimization is desired, it is necessary to decide the feasibility of computer control for which a decision forming study is to be made involving the

- (i) study of the process itself,
- (ii) study of the mathematical model for use in computers,
- (iii) study with different software techniques,
- (iv) computer itself for cost, etc. and
- (v) comparison with conventional control mainly in relation to cost and product quality.

Without reference to any specific process, the general approach of set point computer control (plant-level) can be represented by the schematic of Fig. 9.14 where the notations will be apparent from what follows next. The response function ϕ is required to be carefully selected which involves process study and its mathematical model. This function is to be optimised in terms of the desired variables of cost, product quality, etc. in relation to the constraints (χ_{ij}) which are to be known. For a process with uncontrollable (also called independent) variables (u_j), controlled or operating variables (also called controllable variables) (c_i), and product variables (x_k) and for operation conditions, the simple mathematical model should be of the form

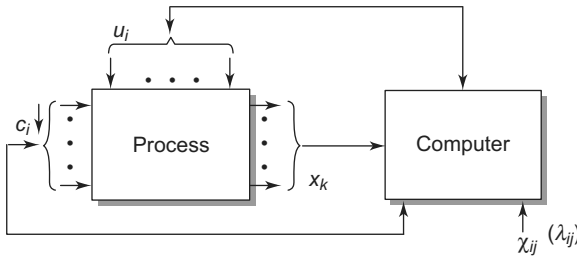


Fig. 9.14 Generalized set point computer control scheme

$$\phi = \phi_{ij}(c_i, u_j) \quad (9.7)$$

with the constraints

$$\chi_{ij}(c_i, u_j) \leq \lambda_{ij} \quad (9.8)$$

where λ_{ij} stands for the desired constraining limits. The optimum will be given by

$$d\phi_{ij}/dc_i = 0 \quad (9.9)$$

in one of the methods of optimization. From this the values of c_i in terms of an appropriate function ψ_j of u_j can be obtained, such that

$$c_i = \psi_j(u_j) \quad (9.10)$$

Eq. (9.9) with constraint (9.8) can be usually solved by a computer and is quite simply done when the u_j s are measurable. If not, the problem becomes sufficiently involved. If Eqs (9.7) and (9.8) can be approximated by linear equations, the well known method of dynamic programming can be adopted. In the method stipulated by Eq. (9.9) called the method of 'steepest ascent', the partial derivatives of ϕ with respect to c_i , and hence the gradient of ϕ , are calculated from a set of operating conditions of Eq.(9.9) as

$$\nabla\phi = \Sigma(\partial\phi / \partial c_i) \cdot \mathbf{I}_i \quad (9.11)$$

$\mathbf{I}_i$ being unit vectors along the coordinate axes. Then, $\nabla\phi$ are directed in the path along which ϕ increases most quickly and this path is called the path of steepest ascent. At every step of this calculation the constraints given by Eq. (9.8) should be checked.

In what has been discussed above, process modelling and constraints have been very briefly referred to and thus the procedure of formulating a suitable response function has not received adequate attention. In fact, formulation of the appropriate response function from the properly formulated process equations appears to be the most difficult part of the optimization procedure. When the theoretical approach from the physical principles fails, techniques with experimental studies, which are often more efficient, are relied upon to a greater degree.

Approach for DDC

From SPC to DDC the basic approach hardly differs. But now the controllers marked FRC in Fig. 9.13 are eliminated and the measured variable of each of these lines are additionally considered as input data to the computer so that after comparison with these values the final controlling action is directly taken by the computer through proper conversion and actuating mechanisms.

Obviously, the main difference is in making mathematical descriptions and programming or what is called software handling.

9.4.3 Company-level Control

In addition to the plant-level control which includes the unit-level control, data and instructions relating to customer side, sales office and warehouse are also processed by the computer following the prescribed strategies. A few of these may be mentioned, such as order processing, invoicing, payment record and scheduling, shipping instructions, sales study and analysis, inventory of raw materials, requisition, industrial economic data,

scheduling of product ranges, etc. It is obvious that it will be possible to take up such a wide range of processing when an overall optimization procedure is followed. In fact, as mentioned earlier, for formulating the adequate response functions, parameters representing these items must be incorporated and hence implementation of C-L-C will necessarily mean the implementation of both P-L-C and U-L-C also. Fig. 9.8 shows the block representation of different controls on a level basis.

9.5 ALGORITHMS FOR DIGITAL CONTROL

As digital control of the DDC type follows the SPC type, it is only natural to believe that the digital control programmes simulate the analog *PID* control laws. But the digital *PID* algorithm obviously has more flexibility and can thus be well adapted for interactions in loops as in ratio, cascade or feedforward controls. One advantage with the digital *PID* control programme is that the control gains for *P*, *I* and *D* can be made independent, but unfortunately they are affected by the sampling period. Also, low frequency fluctuation called noise aliasing may occur in the digital control. It is natural to think that the less the sampling rate the poorer will be the performance of the digital version of the *PID* algorithm. A very fast sampling rate is, however, uneconomical for the system. A typical performance function in terms of the integral square error (ISE) versus sampling time is illustrated in Fig. 9.15 when the analogue *PID* algorithm

$$m = K_c (e + (1/T_r) \int e \cdot dt + T_d de/dt) + m_0 \quad (9.12)$$

is approximated by the digital form

$$m_t = m_{t-T} + K_c [e_t - e_{t-T} + (T/T_r)e_t + (T_d/T)(e_t - 2e_{t-T} + e_{t-2T})] + m_0 \quad (9.13)$$

where m is the manipulating variable, e , the error and T , the sampling time with m_0 as the value of m for initial or control position of valve. The

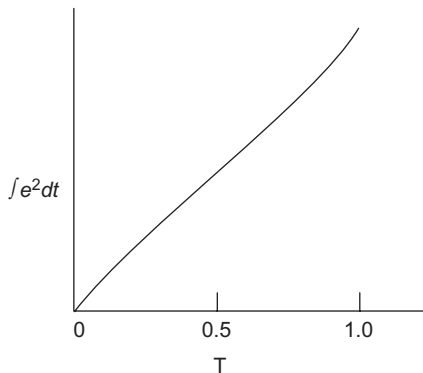


Fig. 9.15 Plot of the integral square error versus time

analogue-control equation such as the one given by Eq. (9.12) can be transformed to the digital form by using either Z -transform or difference equations. In simpler cases, the latter method offers certain advantages as regards understanding, implementation and error evaluation. PID algorithms developed here mostly use this latter technique.

The digital algorithm can be designed to be in the positional or incremental (also called velocity) form. Position form algorithm is quite common but the incremental form is used in some special cases. However, the choice depends on the use of the final actuating element. For the common spring actuated pneumatic control valve, position form is more suitable. Application examples are feedforward, cascade control, etc. The incremental form is more suitable where the valve acts as an integrator such as a stepping motor and cases where wind up is to be tackled (see Chapter 7). Wind up occurs when control valve is stopped at the limiting position but the integration of the error continues. In terms of sampling time T , the simplified position form and incremental form algorithms are written, respectively, as

$$m_n = \frac{100}{PB} \left[e_n + \frac{T}{T_r} \sum_{j=0}^n e_j + \frac{T_d}{T} \Delta e_n \right] + m_0 \quad (9.14)$$

and

$$\Delta m_n = \frac{100}{PB} \left[\Delta e_n + \frac{T}{T_r} e_n + \frac{T_d}{T} \Delta^2 e_n \right] \quad (9.15)$$

The flow chart for implementing the algorithm given by Eq. (9.14) is given in Fig. 9.16 (a).

It is always advisable to form the flowchart first and then go for programming. The simple flow charts for ON-OFF control and PROPORTIONAL CONTROL are shown in Figs. 9.16 (b) and (c) respectively as case studies. From the flow chart of Fig. 9.16 (a) flow charts for PI and PD controls can be easily prepared.

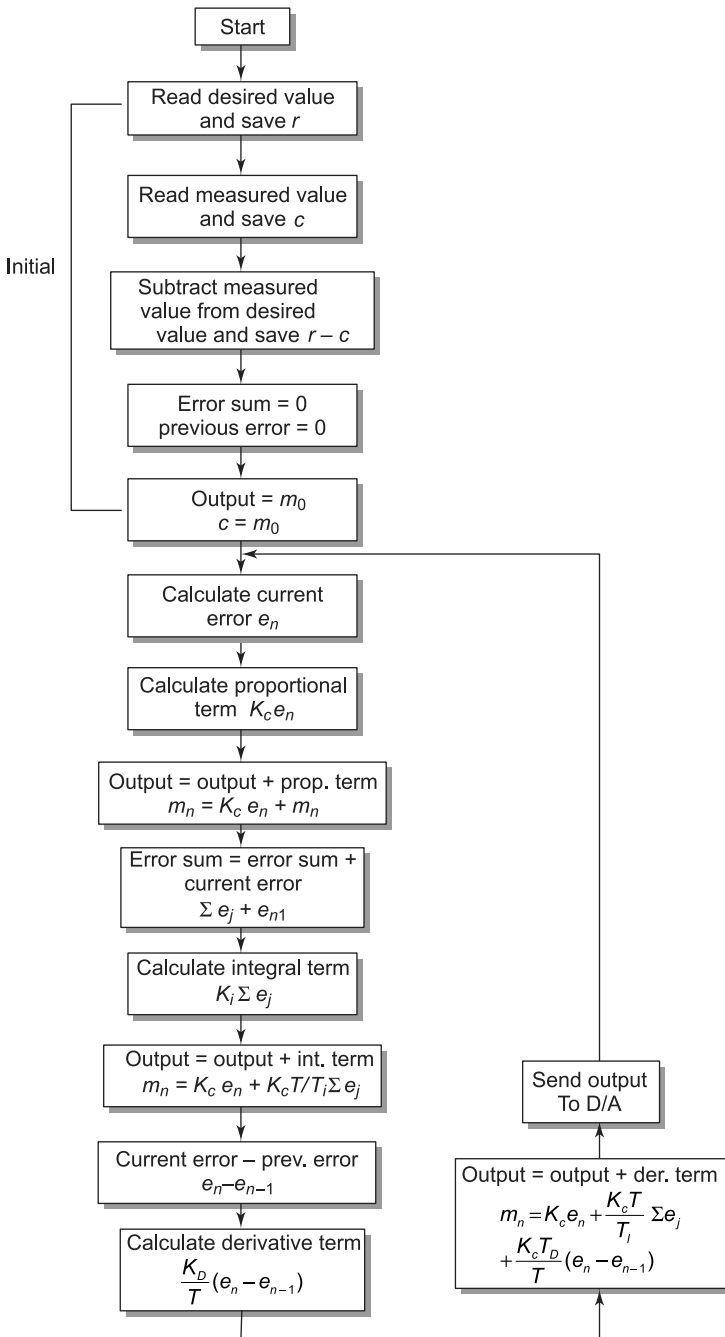


Fig. 9.16 (a) Flowchart for the algorithm of PID control

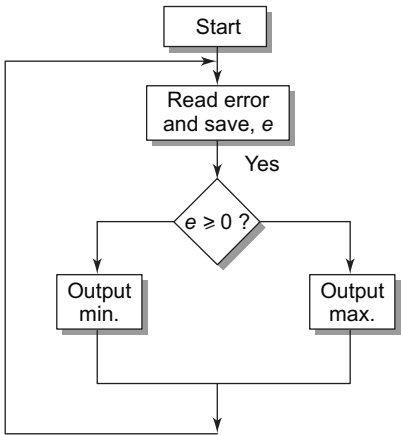


Fig. 9.16 (b) ON-OFF control

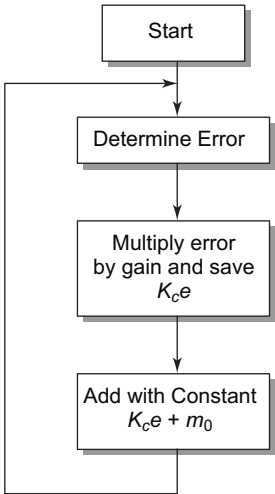


Fig. 9.16 (c) Proportional control

The program based on Eq. (9.14) is given by

DO 40 N = I, K

PRE(N) = ERR(N)

ERR(N) = MVL(N) – SPT(N)

PSI(N) = KP * ERR(N) + KI * ERR(N) + SUM(N)
 + KD * (ERR(N) – PRE(N))/TIME

SUM(N) = SUM(N – 1) + KI * ERR(N)

40 NCONTINUE

One important point to note is that when $T \gg T_r$ or T is very large, proportional action is not beneficial to the control; instead it tends to produce bump on achieving zero error at the n th sample, as, at that instant the term due to proportional action vanishes. One possible way to eliminate this is to sum up to the $(n - 1)$ th sample. One of the other methods is to design an algorithm based on lagging positive feedback. The primary advantage with such an algorithm is that an interruption in this feedback path can be made and a subsidiary input through this may be fed for bumpless transfer.

More generalized positional and incremental forms of algorithms are written as

$$m_n = K_c c_n + K_I \sum_{j=0}^n (r_j - c_j) + K_D (c_n - c_{n-1}) + m_0 \quad (9.16)$$

and

$$\Delta m_n = K_c (c_{n-1} - c_n) + K_I (r_n - c_n) + K_D (2c_{n-1} - c_{n-2} - c_n) \quad (9.17)$$

Now the gains K_c , K_I and K_D have been made independent of one another and to avoid a sudden change in the manipulated variable when the set point changes, a reference signal is included only in the integral action term. Fig. 9.16(d) shows the block-diagrammatic representation of Eq. (9.16) after appropriate transformation.

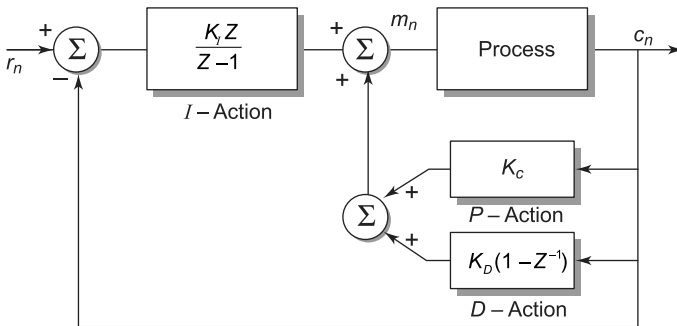


Fig. 9.16 (d) Block diagrammatic representation after transformation, Z : z-transform variable

From this the state-variable approach follows only too naturally and a control algorithm based on such an approach can easily be formulated. However, this aspect is too extensive to be further developed here. Instead, another example of writing a digital control algorithm from a practical *PID* control scheme is illustrated as follows.

9.5.1 Digital Algorithm from Practical PID-Law

Digital control algorithm for both computers and microprocessors are alike. The basic difference may be traced to the limited computing capability and memory of the microprocessor. Here an example of obtaining an algorithm from a practical analogue control law is presented.

Consider the *PID* controller transfer function

$$T_c(s) = \frac{K_c(1 + sT_r)(1 + sT_d)}{sT_r(1 + \alpha sT_d)} \quad (9.18)$$

where α is the rate amplitude constant. For formulation of the algorithm this is broken down into the proportional and, integral and derivative blocks as sketched in Fig. 9.17 and the following splitting up of the transfer function is allowed:

$$d/e = (1 + sT_d)/(1 + \alpha sT_d)$$

$$p/d = P = K_c$$

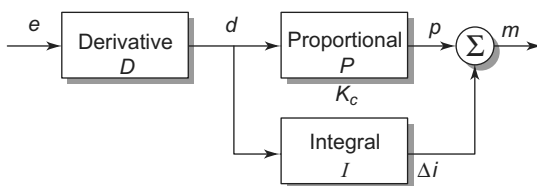


Fig. 9.17 Block representation to help formulate the algorithm

and

$$\Delta i/d = I = K_c/(sT_r) \quad (9.19)$$

The difference method procedure represents dy/dt as

$$dy/dt = (y_n - y_{n-1})/T = (y_n - y_{n-1})/(T_n - T_{n-1}) \quad (9.20)$$

so that the algorithm for the differentiator block is written as

$$d_n = d_{n-1} + (1/\alpha)(e_n - e_{n-1}) + \{T/(\alpha T_d + T)\}(e_n - d_{n-1}) \quad (9.21)$$

for $T \ll T_d$, similarly, the output Δi from the integral block is written as

$$\Delta i = K_c(T/T_r)d_n \quad (9.22)$$

Hence the output m_n at the n th instant is

$$\begin{aligned} m_n &= \Delta i + K_c d_n \\ &= K_c d_n (1 + T/T_r) \\ &= K_c (1 + T/T_r) [d_{n-1} + (1 + \alpha)(e_n - e_{n-1})] \end{aligned}$$

$$+ \{T/(\alpha T_d + T)\} (e_n - d_{n-1})] \quad (9.23)$$

This algorithm requires computations of d_{n-1} and storage of d_{n-1} and e_{n-1} for each e_n to get m_n .

Another algorithm very often used in dynamic compensation and feed forward control is the lead-lag algorithm as derived from the transfer function

$$\frac{f(s)}{x(s)} = T_f(s) = \frac{K(1 + sT_d)}{1 + sT_r} \quad (9.24)$$

where T_d = lead time, T_r = lag time and K is a gain constant. The transfer function can be represented by cascaded blocks whose individual transfer functions are, say, $(1 + sT_r)^{-1}$, $(1 + sT_d)$ and K respectively considered from the input side with outputs of the blocks as g , d and f respectively. From the first (lag) block, one easily obtains

$$g_n = g_{n-1} + T(x_n - g_{n-1})/(T_r + T) \quad (9.25)$$

The next block gives

$$d_n = T_d(g_n - g_{n-1})/T + g_n \quad (9.26)$$

and the last block gives

$$f_n = Kd_n \quad (9.27)$$

The complete algorithm is, therefore,

$$f_n = -KT_d g_{n-1}/T + (1 + T_d/T)\{g_{n-1} + T(x_n - g_{n-1})/(T_r + T)\} \quad (9.28)$$

Complexity arises when the process has a dead time. Considering a simple dead time process with a transfer function

$$T_p(s) = \alpha \exp(-s\tau_d)/(1 + s\tau) \quad (9.29)$$

it is easy to surmise that a large τ_d can hardly be taken care of and a stable process operation is almost impossible to attain unless, of course, the system is compensated as shown in Fig. 5.15. However, as mentioned there in Chapter 5, if τ_d is reasonably small a series expansion of $\exp(-s\tau_d)$ is made and let up to the second order terms only be considered for approximation. Equation (9.29) is then approximated as

$$T_p(s) = (\alpha/(1 + s\tau))\{1 - 1/(1 + s\tau_d/2)^2\} \quad (9.30)$$

One has to make blocks for α , $1/(1 + s\tau)$ and $(1 + s\tau_d/2)$ and arrange as shown in Fig. 9.18. The difference-equation formats for the individual blocks are

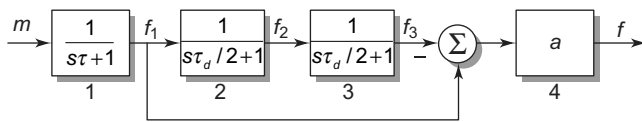


Fig. 9.18 Final block diagram with dead time

$$\begin{aligned}
 f_{1,n} &= f_{1,n-1} + (T/(\tau + T)) (m - f_{1,n-1}) \\
 &= f_{1,n-1}(T/\tau)(m - f_{1,n-1})
 \end{aligned} \tag{9.31a}$$

since $\tau \gg T$.

$$f_{2,n} = f_{2,n-1} + (2T/\tau_d)f_{1,n-1} - f_{2,n-1} \tag{9.31b}$$

$$f_{3,n} = f_{3,n-1} + (2T/\tau_d)f_{2,n-1} - f_{3,n-1} \tag{9.31c}$$

and finally,

$$f_n = a(f_{1,n} - f_{3,n}) \tag{9.31d}$$

Now, f_n will be added to c and the resultant subtracted from r for obtaining the input to the controller.

9.6 DIGITAL CONTROL VIA Z-TRANSFORM TECHNIQUE

It is, perhaps, clear by now that the development of a conventional *PID*-law based control algorithm to be used in a digital computer does not require the knowledge of the *Z*-transform. However, if the process transfer function is known algorithm may be designed by *Z*-transform which would enable the designer to obtain the desired response characteristics. In the sequel a simple case of designing a control algorithm via *Z*-transform would be discussed that achieves the desired performance of the system. For this, however, a knowledge of the *Z*-transform technique is essential which is in addition to what has been discussed in Ch. 1, briefly reviewed, although the course reader is expected to have the requisite knowledge of the same.

9.6.1 Brief Review of Z-transform

In digital control, the process variable is required to be sampled at regular intervals called sampling period by an A/D converter. Thus, the magnitude of the process variable at the sampling instant is only considered. The sampling is considered to be performed instantaneously, but because of the physical limitations, there is a finite sampling duration which for all practical purposes may be considered to be negligible compared to the time constants of the process. This consideration allows us to represent the sample function as

$$f^*(t) = f(t) \sum_{n=0}^{\infty} \delta(t - nT) \quad (9.32)$$

because the output of the sampler may in such a case be considered as impulses at sampling intervals T and the practical form of the impulse function is then used. Eq. (9.32) may be written as

$$f^*(t) = \sum_{n=0}^{\infty} f(nT) \delta(t - nT) \quad (9.33)$$

and thus the value of the k th sample is given by

$$f(kT) = \int_0^{\infty} f(t) \delta(t - kT) dt \quad (9.34)$$

Taking the Laplace transfer of $f^*(t)$ in Eq. (9.34), one gets

$$\begin{aligned} L[f^*(t)] &= f^*(s) = f(0) + f(T)\exp(-sT)L[\delta(t)] \\ &\quad + f(2T)\exp(-s2T)L[\delta(t)] + \dots + \\ &= \sum_{n=0}^{\infty} f(nT)\exp(-nsT) \end{aligned} \quad (9.35)$$

as $L[\delta(t)] = 1$. Now making the transformation called Z -transformation $z = \exp(sT)$ in Eq. (9.35)

$$Zf(t) = F(z) = f^*(s)|_{z=\exp(sT)} = \sum_{n=0}^{\infty} f(nT)z^{-n} \quad (9.36)$$

Table of Z -transforms is available (see Appendix I; A-1, A-2) wherein the Z -transforms of specific functions such as unit step, ramp, $\exp(-\alpha t)$, $\sin(\alpha t)$ etc. are given. Very important properties of Z -transforms are that:

- (i) It is linear.
- (ii) It follows the initial and final value theorems as

$$\lim_{t \rightarrow 0} f(t) = \lim_{z \rightarrow \infty} F(z) \quad (9.37a)$$

and

$$\lim_{t \rightarrow \infty} f(t) = \lim_{z \rightarrow 1} F(z)(1 - z^{-1}) \quad (9.37b)$$

- (iii) It follows the translation law the consequence of which is that if $f^*(t)$ is delayed by k -sampling periods, then

$$Z[f(t - kT)u(t - kT)] = z^{-k}F(z), f(t - kT) = 0 \quad \text{for } t < kT \quad (9.38)$$

where u here stands for unit step function.

The inverse of Z -transform is denoted as

$$z^{-1}\{F(z)\} = f^*(t) \quad (9.39)$$

clearly showing that it yields the sampled function and not the continuous form $f(t)$. To invert the Z -transform $F(z)$ to get $f^*(t)$ the steps are, therefore, as follows:

- (a) $F(z)$ is divided by z to get a new function, say, $F_1(z)$,
- (b) $F_1(z)$ is expanded in partial fractions,
- (c) $F_1(z)$ is then multiplied by z to get back $F(z)$,
- (d) Table of inverse transform is now consulted.

Alternatively, if $F(z)$ is in $N(z)/D(z)$ form, by long division the power series form of $F(z)$ is obtained and there the coefficients of z^{-n} gives the value of $f(t)$ at the n th sampling instant.

9.6.2 The Modified Z -transform

It would be seen that the ordinary Z -transform is used to determine the transient response only at the sampling instants. For obtaining responses in between, a modified Z -transform has been proposed. Such a transform is useful for analysis of systems having dead times. If a process transfer function is given as

$$P_0(s) = P(s)\exp(-s\tau_d) \quad (9.40)$$

where τ_d is the dead time and $P(s)$ is the transfer function without the dead time. For a sampling time T , let τ_d be written as

$$\tau_d(s) = k_1 T + \tau_0 \quad (9.41)$$

where k_1 is an integer denoting the largest number of times the sampling interval can go into τ_d , then the Z -transform of $P_0(s)$ is

$$\begin{aligned} Z[P_0(s)] &= Z[P(s)\exp(-(k_1 T + \tau_0)s)] \\ &= z^{-k} Z[P(s)\exp(-\tau_0 s)] \end{aligned} \quad (9.42)$$

In Eq. (9.42), the term $Z[P(s)\exp(-\tau_0 s)]$ is known as the modified Z -transform of $P(s)$ and is usually denoted as

$$P(z, m) = Z_m\{P(s)\} = Z[P(s)\exp(-\tau_0 s)] \quad (9.43)$$

Thus $L f(t) = \exp(-at)$ gives $f(s) = 1/(s+a)$, $Z[f(s)] = Z[1/(s+a)]$, but $Z_m[f(s)] = Z[\exp(-\tau_0 s)/(s+a)]$.

While evaluation of the former follows straight from the time function as

$$\begin{aligned} Z \exp(-at) &= \sum_{n=0}^{\infty} c^{-anT} z^{-n} = \sum_{n=0}^{\infty} \exp(-aT) z^{-1})^n \\ &= 1/(1 - z^{-1} \exp(-aT)), |z^{-1}| < \exp(aT) \end{aligned} \quad (9.44)$$

the latter requires modification as

$$\begin{aligned}
 Z_m[\exp(-at)] &= Z[u(t - \tau_0)\exp\{-a(t - \tau_0)\}] \\
 &= \sum_{n=0}^{\infty} u(nT - \tau_0)\exp(-a(nT - \tau_0))z^{-n}
 \end{aligned}$$

Putting $T - \tau_0 = mT$ so that $m = 1 - \tau_0/T$, one can write

$$\begin{aligned}
 Z_m[\exp(-at)] &= \exp(-amT)z^{-1} + \exp(-amT)\exp(-aT)z^{-2} + \dots \\
 &\quad \exp(-amT)\exp(-(k-1)aT)z^{-k} + \dots \\
 &= z^{-1}\exp(-amT)(1 + z^{-1}\exp(-aT) + \dots + z^{-k}\exp(-kaT) + \dots) \\
 &= \frac{z^{-1}\exp(-amT)}{1 - z^{-1}\exp(-aT)} \tag{9.45}
 \end{aligned}$$

9.6.3 Impulse Transfer Function

In sampled data system, the input and output are both pulsed and a relation between these is to be obtained. The pulsed input is conveniently expressed in Z -transform and likewise the pulsed output so that the pulse transfer function can be obtained by taking the ratio

$$P(z) = Y(z)/X(z) \tag{9.46}$$

It must be noted that if the input $X(t) = \delta(\tau)$, i.e., an impulse function,

$$X(z) = 1, \text{ so that}$$

$$P(z) = Y(z) = \sum_{n=0}^{\infty} Y(nT)z^{-n} \tag{9.47}$$

Also $X(s) = 1$ for such a case to give $P(s) = Y(s)$. Thus knowing $P(s)$ one has to get $P(t)$ by inversion, put $t = nT$ and then apply Eq. (9.36).

9.6.4 Hold Devices

The output measured by a measurement system in process control is sampled and compared with a sampled reference to obtain the error in discrete form. Since the process is continuous in nature, the discrete form control action from controller (computer) in response to the above error requires to be 'reconstructed' by means of a D/A converter which is simply known as a holding device. The holding device holds the value of the variable on its output side for a pulsed input till the next impulse is received by it and thus a smoothing occurs by extrapolation. The extrapolated value between two sampling instants nT and $(n+1)T$ would obviously depend on its values at the preceding instants nT , $(n-1)T$, $(n-2)T$, etc. and is thus given by a power series expansion of the variable between the intervals nT and $(n+1)T$. Thus, if the output variable is denoted by $y(t)$, then

$$y(t) = y_n(t), nT \leq t \leq (n+1)T \quad (9.48)$$

and

$$y_n(t) = \sum_{k=0}^n (d^k y(nT) / dt^k) (t - nT)^k / k! \quad (9.49)$$

where $d^k y(nT) / dt^k$ actually means $d^k y(t) / dt^k|_{t=nT}$ and $y(nT) = y(t)$ at $t = nT$.

If the value of k is zero, the right hand side of Eq. (9.49) becomes a zero order polynomial of a single term $y(nT)$ and the holding device is said to operate as a zero-order hold. In this way we have 1st order, 2nd order... k th order holds. Interestingly, as the device receives only at the sampling instants the derivatives are obtained from the sampled input data. Thus

$$d^k y(nT) / dt^k = (1/T^k) (-1)^j \sum_{j=0}^k k! y(n-j)T / ((k-j)!j!) \quad (9.50)$$

and one notes that to obtain an estimated value of a derivative of $y(t)$, the minimum number of data pulses should be the derivative order plus one. For zero order it is one, for 1st order it is two, and so on. Higher order holds introduce greater delays in the system and may cause instability in the system but it is likely to reproduce the desired function better. However, in process control systems zero order hold schemes are most common with first order hold used only occasionally.

The holding device appears in the control scheme as shown in Fig. 9.19, where $x^*(t)$ is a train of impulses of varying strengths. If the input strength at $t = 0$ is A_0 , $x_0^*(t)$ can be represented by

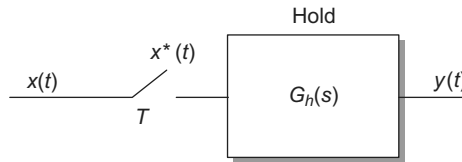


Fig. 9.19 The hold block

$$x_0^*(t) = A_0 \delta(t) \quad (9.51)$$

During the first interval $t = 0$ to $t = T$, therefore, for a zero order hold the output can be represented as

$$y(t) = A_0 [u(t) - u(t - T)] \quad (9.52)$$

Thus the input train of the pulses is represented by

$$x^*(t) = \sum_{n=0}^N A_n \delta(t - nT) \quad (9.53)$$

and its Laplace transform is

$$x^*(s) = \sum_{n=0}^N A_n \exp(-snT) \quad (9.54)$$

Similarly the output $y(t)$ is given by

$$y(t) = \sum_{n=0}^N A_n [u(t - nT) - u(t - (n+1)T)] \quad (9.55)$$

with its Laplace transform as

$$\begin{aligned} y(s) &= \sum_{n=0}^N A_n [\exp(-snT)/s - \exp(-s(n+1)T/s)] \\ &= \left[(1 - \exp(-sT)) \sum_{n=0}^N A_n \exp(-snT) \right] / s \end{aligned} \quad (9.56)$$

Thus the transfer function of the zero order hold is

$$y(s)/x(s) = G_{hz}(s) = (1 - \exp(-sT))/s \quad (9.57)$$

Proceeding similarly the transfer function of a first order hold can be obtained as

$$G_{hf}(s) = ((1 + sT)/T)((1 - \exp(-sT))/s)^2 \quad (9.58)$$

9.6.5 Loop Transfer Function

Before passing on to the Z-transform algorithm design of a controller for a control system, we discuss a little regarding the transfer function of the closed loop system in general as shown in Fig. 9.20. The derivation would normally be on pulse transfer function basis. But as shown, there are blocks which receive continuous signal and there are others which receive sampled signal. Obviously some may be combined to form a single block and some may not.

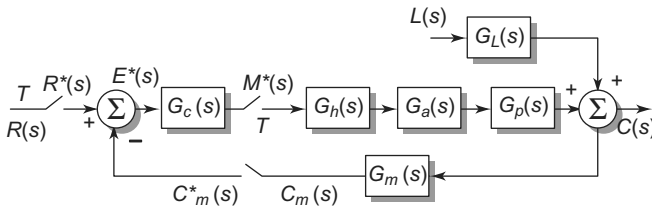


Fig. 9.20 The generalized sampled closed loop scheme

In the figure, blocks $G_h(s)$, $G_a(s)$ and $G_p(s)$ can be combined into a single block of transfer function $G(s)$ (say). For simplicity $G_m(s)$ may be

taken as a constant or even unity so that following equations are obtained from the diagram

$$C_m(s) = M^*(s)G(s) + L(s)G_L(s) \quad (9.59)$$

$$R^*(s) - C_m^*(s) = E^*(s) \quad (9.60)$$

and

$$M^*(s) = G_c(s)E^*(s) \quad (9.61)$$

Obviously $G_c(s)$ can be replaced by $G_c^*(s)$ and then using $M^*(s)$ in Eq. (9.59)

$$C_m(s) = G_c^*(s)E^*(s)G(s) + L(s)G_L(s) \quad (9.62)$$

Combining Eqs (9.60) and (9.62), further

$$C_m(s) = G_c^*(s)R^*(s) - C_m^*(s)G(s) + L(s)G_L(s) \quad (9.63)$$

For the pulse form, now,

$$C_m^*(s) = G_c^*(s)R^*(s)G^*(s) - G_c^*(s)C_m^*(s)G^*(s) + LG_L^*(s) \quad (9.64)$$

Here $LG_L^*(s) \neq L^*(s)G_L^*(s)$, and $G^*(s) \neq G_h^*(s)G_a^*(s)G_p^*(s)$

but $= G_h G_a G_p^*(s)$. From Eq. (9.64)

$$\frac{C_m(z)}{R(z)} = \frac{G_c(z)G(z)}{1 + G_c(z)G(z)} \quad (9.65)$$

and

$$C_m(z) = LG_L(z)/(1 + G_c(z)G(z)) \quad (9.66)$$

$C_m(z)/L(z)$ could not be written as $L^*(s)$ could not be obtained earlier or the load enters the process without being sampled.

9.6.6 The Response

If measurement is included in the process block $C_m = C$ so that from Eq. (9.65)

$$G_c(z) = (C(z)/R(z))/\{(G(z)(1 - C(z)/R(z))\} \quad (9.67)$$

To obtain the controller algorithm for the desired response of the process the following steps are to be taken:

- (a) Select suitable set point (step, ramp, etc),
- (b) Specify the desired response characteristic: one such specification is the dead beat algorithm which specifies finite settling time, minimum rise time and zero steady state error. One may say

that the controlled variable should reach the new value of the set point in one sampling period and thereafter remain there. This specification allows to compute $C(z)/R(z)$ using (a).

(c) For a given $G(z)$, i.e., $G(s)$ we are now able to find $G_c(z)$.

If the set-point is changed by a unit step so that $R(t) = u(t)$, then $R(z) = 1/(1 - z^{-1})$. For the dead beat algorithm as mentioned above, $C(z)$ would be

$$\begin{aligned} C(z) &= 0 + z^{-1} + z^{-2} + z^{-3} + \dots \\ &= z^{-1}/(1 - z^{-1}) \end{aligned} \quad (9.68a)$$

Thus

$$C(z)/R(z) = z^{-1} \quad (9.68b)$$

Assuming that $G_a(s) = 1$ and $G_p(s) = 1/(1 + s\tau)$, one has

$$\begin{aligned} G(z) &= G_{hz}G_p(z) = Z[(1 - \exp(-sT))/(s(1 + s\tau))] \\ &= Z[1/(s(1 + s\tau))] - Z[\exp(-sT)/(s(1 + s\tau))] \end{aligned} \quad (9.69)$$

Equation (9.69) changes to

$$\begin{aligned} G(z) &= Z[1/(s(1 + s\tau))] - z^{-1} \cdot Z[1/(s(1 + s\tau))] \\ &= (1 - z^{-1}) \cdot Z[1/(s(1 + s\tau))] \end{aligned}$$

Using the table now, this gives

$$\begin{aligned} G(z) &= [(1 - z^{-1})z(1 - \exp(-T/\tau))]/[(z - 1)(z - \exp(-T/\tau))] \\ &= (1 - \exp(-T/\tau))/(z - \exp(-T/\tau)) \end{aligned} \quad (9.70)$$

Using Eq. (9.67) now

$$G_c(z) = \frac{z - \exp(-T/\tau)}{1 - \exp(-T/\tau)} \cdot \frac{z^{-1}}{1 - z^{-1}} \quad (9.71)$$

Hence,

$$G_c(z) = \frac{M(z)}{E(z)} = \frac{1 - \exp(-T/\tau)z^{-1}}{1 - \exp(-T/\tau) - (1 - \exp(-T/\tau))z^{-1}} \quad (9.72)$$

From Eq. (9.72) now

$$\begin{aligned} (1 - e^{-T/\tau})M(z) - (1 - \exp(-T/\tau))z^{-1}M(z) \\ = E(z) - \exp(-T/\tau)z^{-1}E(z) \end{aligned} \quad (9.73)$$

As z^{-1} refers to a time delay by one sampling period, Eq. (9.73) is inverted to obtain the computation algorithm as

$$M_n = M_{n-1} + E_n / (1 - \exp(-T/\tau)) - E_{n-1} / \{\exp(-T/\tau) / (1 - \exp(-T/\tau))\} \quad (9.74)$$

For given T and τ , coefficients of E_n and E_{n-1} , the errors at n th and $(n-1)$ th instants are easily calculated. Also, for this case the coefficients of M_n and M_{n-1} , the controller output at n th and $(n-1)$ th sampling instants are both unity. Writing $1/(1 - \exp(-T/\tau)) = A$ and $\exp(-T/\tau)/(1 - \exp(-T/\tau)) = B$, the programme statements for this control are

$$M = M + A * E - B * E1$$

$$E1 = E$$

Before proceeding further, a note on the stability of the closed loop at this stage must be in order. For this the characteristic equation of Eqs (9.65) and (9.66) would be considered which, obviously, takes the form

$$1 + D(z) = 0 \quad (9.75)$$

and the nature of the roots of this equation would determine the system stability. As is well-known, in s -domain the stable region of having the roots is the left half plane and s and z are related by the equation $z = \exp(sT)$, mapping the left half of the s -plane in the z -plane would indicate that it actually is mapped inside an unit circle with, (i) the imaginary axis in the s -plane tracing the unit circle in the z -plane, (ii) the negative real axis of the s -plane mapping into the positive real axis in the z -plane inside the circle with the s -plane $(-\infty)$ - point mapping at z -plane - zero point, and (iii) any other point in the s -plane transforms or maps on to the z -plane inside the unit circle.

This means that any point outside this circle in z -plane would lead to system instability. Fig. 9.21 shows the region of stability.

For load changes, the controller design may be initiated from that angle. However, there are instances when algorithm designed on the basis of set point change performs as good as for load changes. When load change is of major interest the algorithm is designed on that basis. Again, considering $C_m = C$, $G_L = G_P$ and $G_a = 1$, Eq. (9.66) yields

$$G_c(z) = (LG_p(z) - C(z)) / (G(z)C(z)) \quad (9.76)$$

The procedure of design for $G_c(z)$ is, in steps, as follows:

Select appropriate $L(s)$ - step 1

Determine or specify the desired response $C(s)$ - step 2

Express in transformed notation (z -transform) - step 3

Solve for $G_c(z)$ - step 4

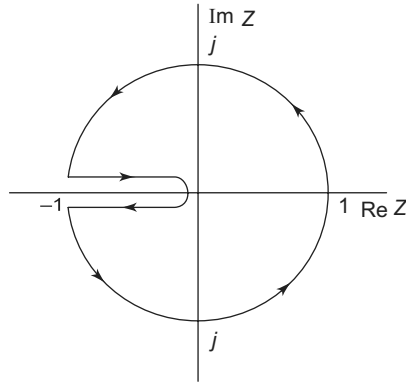


Fig. 9.21 The stability region

In this case, as mentioned earlier, load enters unsampled and hence the point of time load enters relative to sampling time is not in the control of the designer and hence a design is to be made for the worst possible case which is to admit the disturbance τ_d units of time before the sampling instant if the dead time of the process is τ_d . The controller is, therefore, not aware of this change until the sampling instant $2T$ as at T the disturbance is just to show its effect and is ignored. At instant $2T$ controller would cause any change in the manipulated variable for offsetting the disturbance but because of τ_d , the change would be reflected in the output only after this τ_d . Thus the system comes under control after $(\tau_d + T)$ time since disturbance enters the loop and correction appears only after $(2\tau_d + T)$ from its instant of entrance (see Fig. 9.22). Thus $C(z)$ cannot be arbitrarily chosen for a selected $L(s)$. If a process has a dead time equal to the sampling time, which is considered for convenience, and a first order lag such that $G_p(s) = \exp(-sT)/(1 + s\tau)$ and $G_a = 1$. With zero-order hold one has

$$G(z) = G_{hz}G_p(z) = Z\left[\frac{1 - \exp(-sT)}{s}, \frac{\exp(-sT)}{1 + s\tau}\right] \quad (9.77a)$$

$$= [(1 - \exp(-T/\tau))z^{-2}]/(1 - \exp(-T/\tau)z^{-1}) \quad (9.77b)$$

For a step disturbance $L(s) = 1/s$ (unit), so that

$$\begin{aligned} LG_p(z) &= Z[\exp(-sT)/(s(1 + s\tau))] \\ &= [(1 - \exp(-T/\tau))z^{-2}]/[(1 - \exp(-T/\tau)z^{-1})] \end{aligned} \quad (9.78)$$

In Eq.(9.76), only $C(z)$ remains to be specified for the design of $G_c(z)$. As the dead time and the sampling time are equal $C(0) = C(T) = 0$ and control

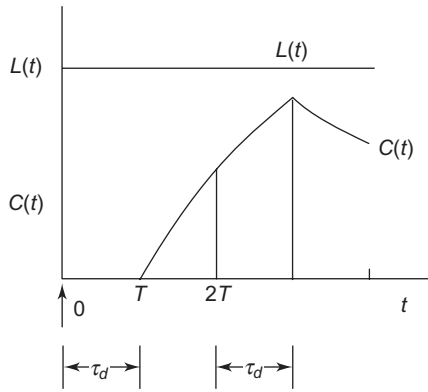


Fig. 9.22 Effect of dead time in digital control in load change

action does not start till $2T$, so between T and $2T$ the response is effectively open-loop and hence $C(2T) = 1 - \exp(-T/\tau)$, with control being effective at $2T$, it appears only at $3T$ (as $\tau_d = T$) (See Fig. 9.23), so that till $3T$ time response remains open-loop and $C(2T) = 1 - \exp(-2T/\tau)$. After this instant and onwards the response can be selected to be dead-beat, i.e., $C(kT) = 0, k > 3$.

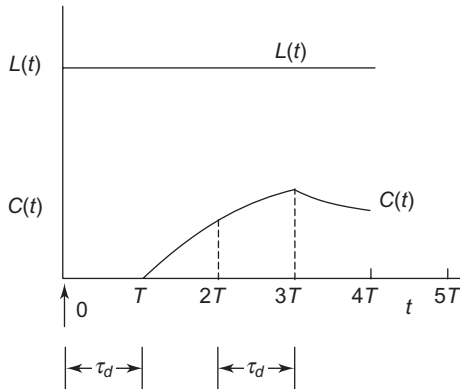


Fig. 9.23 Further elaboration of the effect of dead time

Thus

$$C(z) = (1 - \exp(-T/\tau))z^{-2} + (1 - \exp(-2T/\tau))z^{-3} \quad (9.79)$$

Using Eqs (9.77), (9.78) and (9.79) in Eq.(9.76)

$$G_c(z) = \frac{[1 + \exp(-T/\tau) + \exp(-2T/\tau)] \left[1 - \frac{\exp(-T/\tau)(1 + \exp(-T/\tau))z^{-1}}{1 - \exp(-T/\tau) + \exp(-2T/\tau)} \right]}{(1 - \exp(-T/\tau))(1 - z^{-1})\{1 + (1 + \exp(-T/\tau))z^{-1}\}} \quad (9.80)$$

Of the two poles the one at $-[1 + \exp(-T/\tau)]$ is outside the unit circle making the controller open-loop unstable producing considerable ringing or oscillations.

The dead-beat algorithm is hardly process-friendly as it is extremely difficult for a process to achieve the new set-point in one sampling time as suggested by $C(z)/R(z)$. Consequently other more ‘acceptable to process’ algorithm have been suggested. One such via Dahlin’s suggestion of a closed loop response for step input ($R(s) = 1/s$) is as given by

$$C(s)/R(s) = \exp(-s\tau_d)/(1 + s\lambda) \quad (9.81)$$

which in discrete form becomes

$$C(z)/R(z) = (1 - \exp(-T/\lambda))z^{-(N+1)}/(1 - \exp(-T/\lambda)z^{-1}) \quad (9.82)$$

where N = largest integral number of sampling times in τ_d .

From Eq. (9.67), therefore, the controller $G_c(z)$ is given as

$$G_c(z) = \frac{(1 - \exp(-T/\lambda))z^{-(N+1)}}{1 - \exp(-T/\lambda)z^{-1} - (1 - \exp(-T/\lambda))z^{-(N+1)}} \cdot \frac{1}{G(z)} \quad (9.83)$$

$G(z)$ using Eq. (9.77a) is calculated for dead time as 1.4s, process lag 4.22s and sampling interval 1s. The MATLAB programme is written for a process gain 0.5 as

num = [0.5];

den = [4.22, 1];

$g = t_f(\text{num}, \text{den}, \text{'input delay'}, 1.40)$

$h = e2d(g, 1)$

This gives

$$GH(z) = G(z) = [(0.006627 + 0.03922 z^{-1})/(1 - 0.789 z^{-1})] z^{-2}$$

This is used in Eq. (9.83) for the required $G_c(z)$.

For any given $G(z)$, $G_c(z)$ is calculated and then expressed in terms of E ’s and M ’s for writing programme statements. It is found that if λ is small control becomes better.

For load change, again, a design is made on the worst case consideration.

9.6.7 Sampling Frequency/Sampling Time

Since digital control requires a sampler or A/D converter, the sampling time T or the sampling frequency $\omega_s = 2\pi/T$ should be chosen on the basis of the frequency of the input signal. As per Shanon, the sampling frequency must be at least twice the maximum frequency to be recovered. This, then, means higher the sampling frequency, better the purpose is served. Often signals are not band-limited so that unless the sampling frequency is infinite the true content of the signal cannot be recovered from its counterpart.

Another consideration is the economic one. Plots have been made of the cost of decrease in loop performance and cost of computing effort each with sampling time (see Fig. 9.15), the former increases with T while the latter decreases and the intersection of the two curves gives the optimum sampling time.

Guidelines through Users Conference (1963) have also given recommendation for values of T depending on the types of processes. These are 1 sec for flow loops, 5 sec for level, and pressure loops and 20 sec for temperature and composition control. But study by individual designer for increasing T may always be on.

Also use of conventional *PI* and *PID* algorithm does not always permit reduction in the value of T because if T is too small, a reset dead-band may result when the implementation of the algorithm is being attempted through a fixed point calculation. Even with simpler proportional control type algorithm there is a limitation in lowering the value of the sampling time on consideration of stability. It is also argued that the algorithm should be designed to compensate for the dead time either through adequate choice of the sampling time or sampling time should be made independent of dead time but made related to process transfer lag.

The difference in the degrees of the denominator and numerator polynomials also is a consideration. If this degree is greater than two, T cannot be made arbitrarily small because of stability reasons.

9.7 DISTRIBUTED CONTROL SYSTEMS

There have been rapid advances in digital hardware and software, digital signal transducers, sophisticated data transmission/acquisition as also information display techniques. These have led to the design and implementation of what is presently known as distributed computer control or simply the distributed control systems (DCS). This basically is a scheme of realisation of control tasks on a multiple-computer system. Obviously, multiplicity of computers would not mean multiplicity of mainframe types but of local partially autonomous computing devices having input/output capability interconnected through digital communication link and coordinated by a mainframe computer. The resulting systems have the advantages of local control as well as centralized coordinated control. Failure of a part of the system would have a local effect or at the most would affect a part of the coordinated activity and there would never be a failure of the operation of the complete system.

The distributed control system consists of four sections:

- (a) Several microprocessor based controllers for first level control functions each of which is capable of handling several control simultaneously.

- (b) At least one coordinating controller—a large mainframe computer for higher level control functions. A number of minicomputers and a large mainframe are used for some systems.
- (c) Interconnecting digital datalinks and organising protocols needed for computer to computer communication and/or communication links between computers and other equipment.
- (d) A central information display unit for system status indication—this may additionally have centralized operator interface facility having access to all process and control variables.

A typical centralized distributed computer control network is shown in Fig. 9.24. Here the minicomputers are optional and may be housed in unit based control rooms as different from the coordinating control room. In such situations communication links are not that long from local controllers to minicomputers. Alternatively, the minicomputers may be dispensed with the mainframe computer taking care of the entire coordination task.

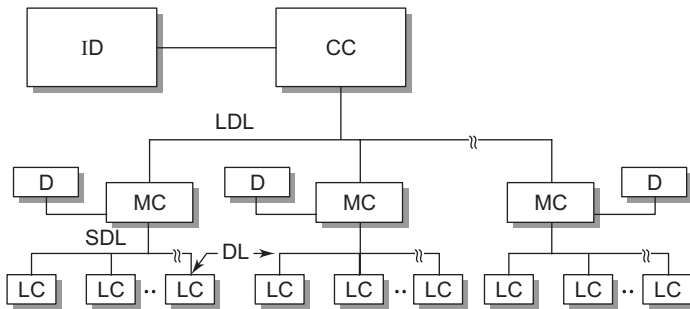


Fig. 9.24 The distributed computer control scheme; D: display, LC: local controller, MC: minicomputer, LDL: long distance data link, SDL: short distance data link, DL: local data link, ID: information display, CC: coordinating controller/computer

9.7.1 The Local Controllers

The local controllers are, as mentioned already, microprocessor-based types and can handle, in a typical case, 32 signals—16 of which are digital and the rest analogue. The signals are sequentially processed through suitable algorithms to produce the required outputs in closing the loops. The local controller has other associated functions such as local displays, putting up alarms, receiving local operator and supervising inputs, etc. A typical/functional diagram is shown in Fig. 9.25.

Number of loops controlled, highest sampling rate, types and levels of input/output signals, redundancy, method of local interaction and amount of local autonomy, modularity, etc. are some considerations that must be of consequence in the specification and design of such a local controller.

9.7.2 The Coordinating Controller

The coordinating controller in most usual situations, consists of a mainframe computer communicating with the local controllers, other man-machine interfaces, display modules via data links.

There are a large number of relatively simple control and signal processing functions which can be handled by microprocessor or analogue devices in an industrial process environment. Yet, there are many complex control tasks which such devices are unable to tackle. Execution of such tasks is left to one or a set of minicomputers or a large mainframe system. These control tasks include optimization, adaptation, scheduling, database management, data and event logging, management reporting, etc. These functions are often not bound by stringent response time requirements of the local controllers and there is consequently enough flexibility in terms of allocation of computer system resources among the computing tasks tolerating a large degree of operating costs.

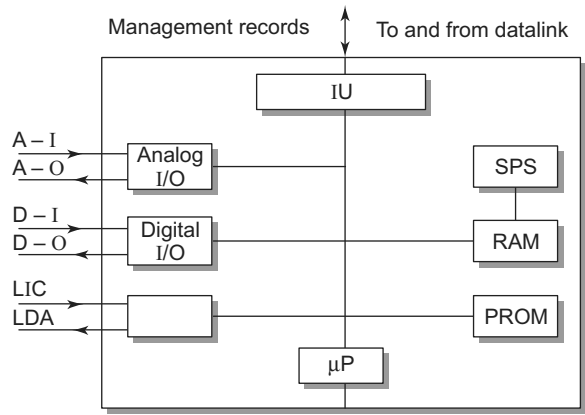


Fig. 9.25 The local controller, LIC: local operating inputs/commands, LDA: local display/alarms, IU: interfacing units, SPS: stand by power supply, μP: microprocessor

9.7.3 The Data Links

Data links are provided by coaxial cables, optical fibres, etc. They may be several kilometers long and are designed to carry high speed digital data serially transmitted.

There are various configurations of data links. The network is characterized by this configurations as also by the protocols for data transfer.

The communication system is responsible for the transfer of process data, programme, command and status information among the control computers and devices. The overall performance of a DCS is dependent on the management, parameters and the structure of the communication system or the data links. A good design of a DCS would connote a careful translation of the needs of the specific application into a data-link topology with all its communication aspects—level and location of redundancy, protocol, etc.

The most common configuration is the data highway configuration also known as the multidrop system. This belongs to the broadcast system in which every node in the network receives every message transmitted. It is shown in Fig. 9.26(a). The others, star, ring, hierarchical and mesh types, are shown in Figs 9.26 (b), (c), (d) and (e) respectively some of which belong to the routed block system or packed switching system in which informations are routed to specific receiving points. The disadvantage of configuration of Fig. 9.26(b) is that all communications must pass through the central node, hence its reliability is critically dependent on the reliability of the central switching unit. For configuration of Fig. 9.26 (c), one device has to transmit message to its neighbour which, in turn, retransmits to its own neighbours until the device defined by the address is reached. Failure of one device would disrupt the process. Such a structure has a delay much higher than a shared global bus (Fig. 9.26(a)).

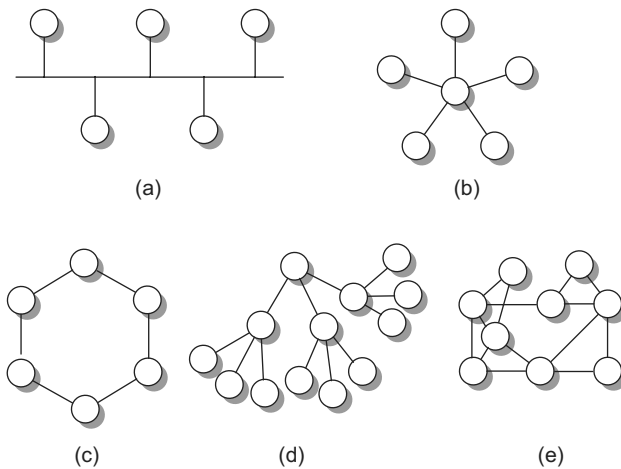


Fig. 9.26 The different types of data link, (a) data highway configuration (b) star type, (c) ring type, (d) hierarchical, (e) mesh type

When many devices intercommunicate along a shared data link, some system must be devised to organize the communication at a maximum efficiency keeping the priority and the conflicting requirements of speed and accuracy. For this purpose a set of rules are specified which is called protocol. Protocol is sometimes implemented by centralized explicit device and sometimes implicitly by ensuring that every device in the network is programmed to communicate as per the rules.

9.7.4 The Central Information Display Unit

It consists of a set of visual display units visible from a central location. The formats of display are, however, in accordance with the requirements of the users and are chosen as per their convenience.

9.8 THE NEWER TRENDS IN PROCESS AUTOMATION

In the present trend of wireless instrumentation which supports collection of signals—both analogue and digital by remote terminal unit (RTU) or PLC etc. from remote equipment and sensors where expensive hardwiring and associated constraints are dispensed with. Traditional hardwired systems are being replaced by wireless system with radio. Wireless network uses one radio system that communicates from a PC through RTU or PLC to the field instruments excepting, perhaps, the sensors. Except for the wireless radio network, there is no change in the system layout.

There had been a distinct difference between process automation and automation in manufacturing (which is more a batch process) but lately the difference in technological aspects between the two is drastically diminishing. These now use similar electronic systems for control via closed loop for human machine interface (HMI) and also for networking. Fieldbus is a network-driven set of field instruments which has been recommended for both type of control automation processes by the IEC.

Hierarchical structure of complex control processes has been in the reckoning for long where direct digital control (DDC) at the plant and unitary levels were considered and also supervisory systems were added to make the entire system function more efficiently when the DDC software could be multiplexed around several control loops. Gradually the sequential controllers were replaced by programmable logic controllers after the microelectronic technology appeared in the scene. But all these involved large amount of investment towards the cost of installing and configuring cables. This cost can be significantly reduced by eliminating the long lines between the controller, sensor and actuator. This is possible with the change in control scheme structure which is called distributed system described briefly above already. This architecture was first marketed by

Honeywell (TDC 2000 system, 1975) as also Yokogawa (Centum, 1975) almost simultaneously. The basic architecture as propounded by such vendors is shown in Fig. 9.27.

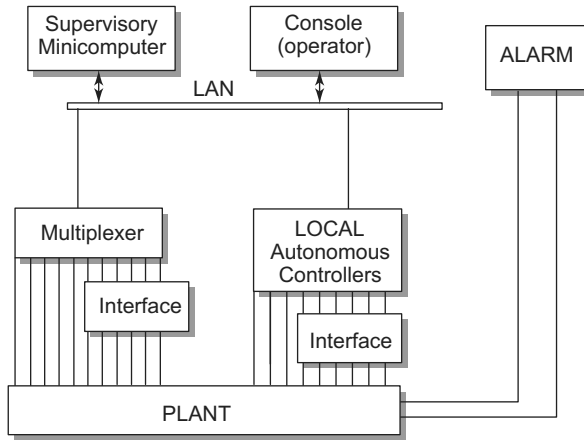


Fig. 9.27 The basic DCS architecture

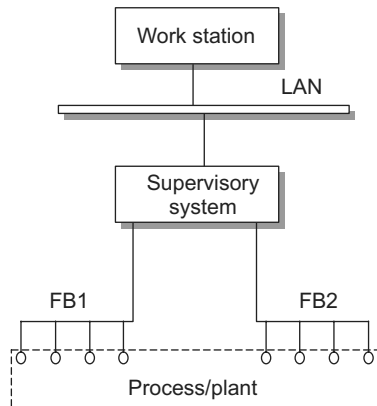


Fig. 9.28 Serial form Fieldbus system

Cabling cost could be reduced further by interconnecting devices located at fields with a serial bus also called fieldbus as shown in Fig. 9.28.

9.8.1 Field Instruments

Field instruments are there already and it is possible to connect these in serial bus system but only gradually, yet, star networks with 4.20 mA analogue signal are still to be considered to remain for sometime to come.

Field bus system is now under the purview of Fieldbus Foundation which is actually an agreed foundation of world FIP i.e. world Factory Instrumentation Protocol and ISP i.e. Interoperable System Project. Renowned companies like Honeywell, Bailey Instruments, Allen Bradley Cegelac, Elf etc. support the former while Siemens, ABB, Yokogawa, Rosemount, Fisher Porter, Foxboro etc are compliants of the latter. Standard bodies like IEC and ISA are collaborating with the support of all the manufacturers to standardize Fieldbus system—it is named SP-50.

Fieldbus is a serial bus and it is a two way (bidirectional) all digital communication system which serve as a LAN for instrumentation and control equipment/devices for plant/factories. In the hierarchy of plant digital networks, fieldbus environment is the base level group and is usable in both process control and manufacturing automation applications. It is capable of replacing a number of devices using 4-20 mA analogue standard and hence distributing the control application across the network.

The fieldbus segments (FB1, FB2, Fig.9.28) should be as large as possible reducing the number of segments to derive installation and wiring benefits. Also fieldbus protocol should support reliable and fast transfer of messages among different segments. Fieldbus is at the lower levels of automation/control/communication in the hierarchy but it can operate in self contained/efficient way and thus the higher levels would just be supervisory in nature.

For field serial networks as in Fieldbus, the time of information flow between field devices and between these devices and higher level is very critical, the constraint is due to the system dynamics. The total delay comes in (1) because of the time taken for message transfer through the levels of the protocol and (2) due to propagation delays introduced by the physical media.

The communication portion of the protocol is designed to support both periodic and aperiodic (event-driven) transactions. If the bandwidth of the bus is heavily used there may arise problems because of such transactions. This is solved by assigning priority tag to event driven transactions. Obviously, the aperiodic transactions are on-demand transactions. There should be synchronization commands as well for accurate time scheduling. The transactions should be guided by a protocol which should provide the facility of (1) detecting and reporting error (2) communication between other devices when one device is faulty, (3) maintaining time schedule, order and send correct data value, (4) redundancy of operational functions—often bus duplication provides network redundancy in full. However type of data and data length is open and not predefined by bus standard. These are set up at the configuration stage and are under the purview of network management function.

Network management is very important in field bus as this provides many of the features proposed. Information such as message and data between devices forms a group to be considered. Overall behaviour of the network used also plays a role. Bus length, number of devices per bus and the update rate in a system are determined by the system requirement such as process control and batch process/factory automation. In manufacturing processes i.e. factory automation, the update rates need be high since faster mechanical operations are required there. Typical cases of these values for process control application may be listed as

Bus length: 1000–2000 m

Devices/bus: 30–50

Update rate: 5 to 10/sec of 5 data bytes/device

For manufacturing processes:

Bus length serves PLC's

Devices: 32

Update rate: nearly 100/sec of 5 data bytes/device

Till current dates transaction is via twisted wire, coax cable, optical fibre etc. which are sooner or later going to be replaced by wireless *rf* radio. *FIP* in France and *PROFIBUS* in Germany have developed 'national' fieldbus standards—both, however, support *IEC* standard. *FIP* uses the broadcasting mode of transmission i.e., one station transmits a message to another or to all station. It is the prerogative of an individual station to receive it or not. The protocol has to be accordingly designed. This uses layers 1, 2 and 7 of the *OSI* communication model i.e., physical, data link and application layers of open system interconnection model. Profibus on the other hand, uses the command/response or the master/slave access system. This also uses 1, 2 and 7 layers with provision of sublayers of layer 7 which offer facilities for communication with field devices—something like *TCP/IP* model where layer 3 (transport) is included. This is a token passing system with the master holding the token is permitted to communicate with the slaves which are passive while the *FIP* has the bus arbitration technique which is responsible for organizing the broadcasting function. The *OSI* communication or reference model is shown in Fig. 9.29 (a) whereas the *TCP/IP* reference model is shown in Fig. 9.29 (b) in blocks. For protocol design process readers are referred to 'Telemetry Principles' by the same author published by McGraw Hill Education.

Quite a few proprietary standard serial networks are also used with specific advantages and disadvantages such as Echelon, *LON* works (Local operating network), Bitbus, Arcnet, CAN (Controller area network) etc.

IEC/ISA fieldbus standard has, however, been able to compromise the two modes adopting arbitration functionality in token passing bus. The coding in transmission is Manchester coding with half duplex communication in *IEC* standard field bus system where asynchronous

data transmission is possible. However, different standards use other codes to provide error control mechanism, *CRC* (Cyclic redundancy check) is one such used, for example in *CAN* standard.

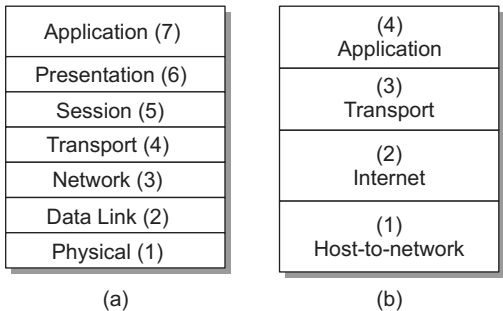


Fig. 9.29 (a) Standard OSI communication reference model
(b) TCP/IP model

It is interesting to note that for increasing bus traffic high-level system functions should be included in the field devices i.e. transducers and actuators etc. These functions are, for example, fault detection i.e. condition monitoring, data configuration configuring new devices—the latter functions are useful in distributed control systems. It thus appears that field devices should themselves be intelligent ones. A typical schematic block diagram of such a device is shown in Fig. 9.30 (a).

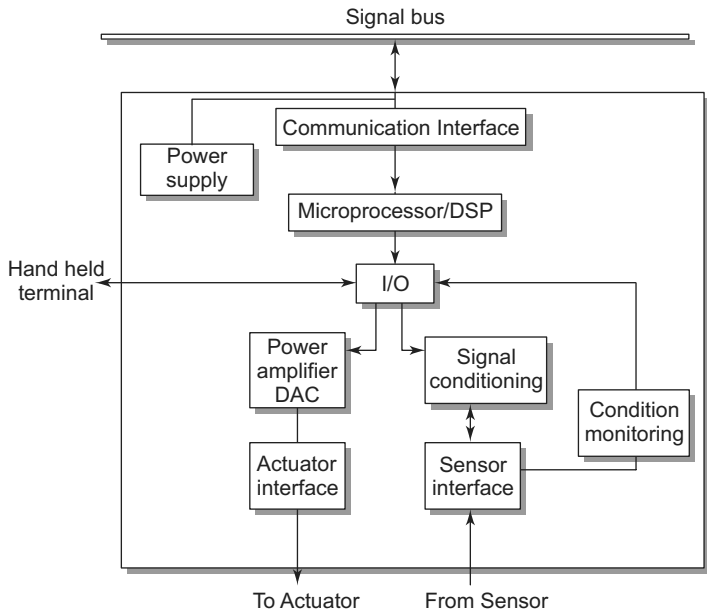


Fig. 9.30 (a) Intelligent field device

The smart or intelligent field device has to have any one or more of the features listed below:

- (1) Auto-calibration and autoranging facilities with calibrating constants acquisitioned and stored automatically, auto-configuration and auto-checking of the device,
- (2) Auto correction of drifts due to time and temperature, offsets and linearization facilities,
- (3) Fault-diagnosis monitoring which often is provided with additional sensors and signal processing and analysis software module,
- (4) Communication system for remote interfacing in the (serial) bus, a handheld communication unit is interfaced to this for checking, alternatively, special interfacing facility is provided using rf., optical or common inductive techniques,
- (5) Data acquired are to be sent or at least summary of the data in appropriate units. This requires processing memories,
- (6) Self-tuning algorithm,
- (7) Capability of dynamic re-configuration of control functions by downloading from host systems or from internal store,
- (8) Control implementation facility either directly via the bus or serial bus and a host system i.e. with special arrangement with the device,
- (9) Testing facility via the bus.

Field devices can be connected to *DCS* with multidrop configuration when cabled, reducing cabling cost and now using radio communication. Field device becomes the transmitter and the connection uses the master-slave protocol and communication is in multiplexed mode. HART (highway addressable remote transducer) protocol is one of the very useful early protocols with a speed of 1200 bauds which can allow 2 updates per second when standard *DCS* speed is around 10 million bits/sec. It allows simultaneous analogue and digital communications. It follows the reduced *OSI* model using only 1, 2 and 7 layers as many others do. It is a half-duplex protocol when communication is one way only. In the format, layer one (physical) uses the Bell 202 *FSK* technique with $1 \rightarrow 1200$ Hz and $0 \rightarrow 2200$ Hz; layer 2, data link layer specifies format and layer 7, the application layer specifies the HART commands. It allows upto four variables to be sent in a single message and allows two masters. It is an open protocol. A typical protocol structure is shown in Fig. 9.30 (b).

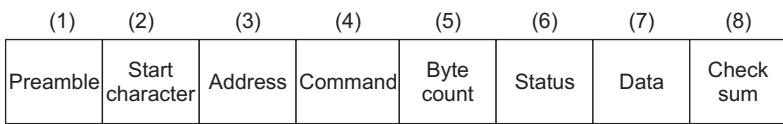


Fig. 9.30 (b) The protocol structure

Preamble block serves to synchronize the frequency detection circuit at the receiving end, has *FF* characters, and 8 numbers of 1's. Block (2) is for information to find which way to communicate, it consists of 1 byte. Address block (3) also is of 1 byte, can identify 1 to 15 transmitters, it is to choose the transmitter. Block (4) has 28 selection and is for different *OCS* commands to the transmitter. Byte count, block (5) counts how many bytes are in Status (block 6) for parity and data (block 7). Block (8) is the checksum block determining the longitudinal parity.

While talking of fieldbus, mention must be made of modbus which is also a serial communication protocol designed initially to be used with PLC's as early as 1979. It supports communication between devices more than two hundred in number and the supervisory computer in the network also can be a part of the data acquisition system other than RTU's. There are different protocols for use in this bus such as Modbus RTU, Modbus *ASCII*, Modbus *TCP/IP*, Modbus *PEMEX*, Modbus *UDP*, Modbus Plus etc. Mesh topology is typically followed in implementation. A typical Modbus *RTU* protocol format is as shown in Fig. 9.30(c)

Start	Address	Function	Data	CRC Check	End
3½ character silence (idle)	Station 8 bits	Command/ Instruction 8 bits	n*8 bits dependent on message	Error check 16 bits	3½ character idle

Fig. 9.30 (c) Another protocol format (Modbus RTU)

Data may be floating point (IEEE) or mixed type, 32 bit integer type, etc. The Modbus functions are coded, since designed for PLC's initially, functionality also was based on communication with components of a *PLC* and now with other RTU's. For example code 01 describes 'Read coil status', code 04 is for 'Read input registers', code 16 is for 'Read multiple registers' etc. Vendor manuals are the source documents for use.

Often a distinction/comparison is drawn between Modbus and Profibus. There are very little differences between them. However, it may be mentioned that Profibus has features that allow some versions to support multimaster mode on RS 485 while Modbus works only on one master. Also, profibus is supported by Profinet, while Modbus is supported by ethernet. Topological differences of the network are also mentioned for comparison.

9.8.2 System Hierarchy

In Fig. 9.8 a hierarchical control scheme is given where the levels are classified as unit level, plant level and company level. Unit level actually

corresponds to the field level. The present trend is to represent the hierarchy in four levels marking them as level 1 to level 4 as shown in Fig. 9.31.

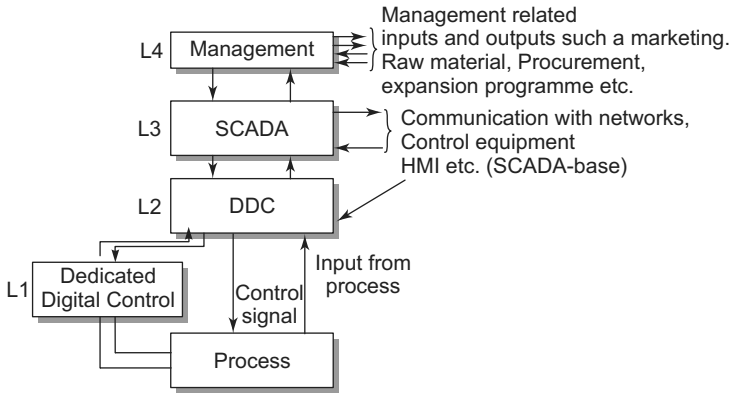


Fig. 9.31 The hierarchical control scheme

Level 1 is the field level, level 2 is the plant level which is essentially a DDC level, level 3 is the supervisory control level, while level 4 is the company management information level and control. Level 1 with 2, 2 with 3 and 3 with 4 are all connected by digital data link while level 1 is connected to process by both analogue and digital link. Levels 2 and 1 both receive process inputs in analogue and digital mode and send control signals to the process—the field units.

It would be noted that the diagram of Fig. 9.27 is only a variation in representation of the hierarchy. In Fig. 9.31 the dedicated digital control is to mean point-to-point connection bringing in individual sensors and actuators of the local field station. The fieldbus system introduced consists of interconnection of distributed data multiplexer (Fig. 9.27).

9.8.3 DCS Vendors

It may be mentioned that different manufactures design their DCS from their concept of better operation—functionally, however, there is no difference. Features like displays, availability of hardcopy output terminals, library of functions consisting of software packages may have small differences when overall packages are concerned. The trend in design is to develop systems with modular architecture. Various manufacturers have come in the market with their product (DCS), Honeywell (TDC 3000, TPS system), Brown Boveri (Procontrol), Fisher Porter (DCI 40000). Siemens (Teleperm-M), Kent (P.4000), Yokogawa (Centum system)—to name a few. Different manufacturers, however, have their own choice for the protocols viz; Honeywell, Bailey, Allan Bradley, Cegelac go for world

Factory Instrumentation protocol (world *FIP*) while Siemens, Yokogawa, *ABB*, Fisher, Foxboro, Rosemount use Interoperable System Project (*ISP*).

Example of the Honeywell systems may be considered as these are modular in design with large or small number of modules depending on the process size. The modules are, as can be seen in Fig. 9.27 : Process interface unit, Analogue unit (multiplexer), Highway traffic director (Network *LAN*), Operating station and Control files. The scope of the *TDC* series has been widened to provide the total plant solution (*TPS*) system, where the business and control information have been placed in an unified environment. Here the plant intranet i.e. ethernet has been merged with process networks which consist of fieldbus (Foundation) system, local control network and universal control network. Figure 9.32 shows a schematic diagram to give the idea.

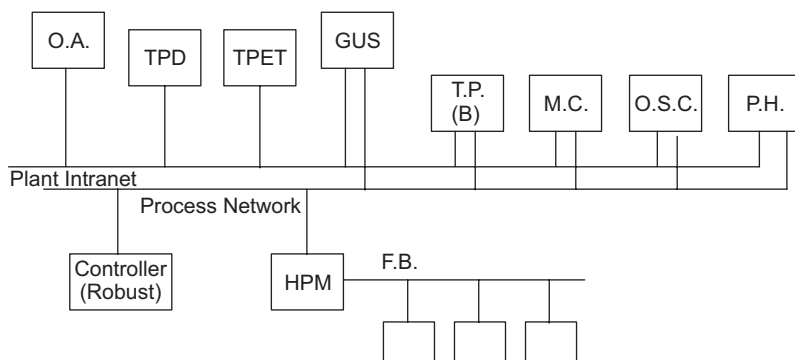


Fig 9.32 Total plant solution—the system illustrated

OA - Other Application OLE, OPC, etc.

TPD - Total plant Desktop

TPET - Total Plant Engineering Tools

TP (B) - Total (Batch) Plant

M.C - Multivariable Control

OSC - Optimization System Control

PH - Plant History

HPM - High Performance Manager (Process)

FB - Field Bus.

9.8.4 SCADA

Supervisory Control And Data Acquisition System is a large scale control system using acquired data from sensors at remote stations and sending

these to a computer that does the control as well. It is predominantly an 'openloop' control system since it does not use feedback to check/compare the result obtained with the inputs sent. SCADA is used in steelmaking, power generation, municipal works, oil pipe lines, making large scale control systems although it does not actually perform real-time control of the set of processes in the plant but does 'collect' the requisite data from it.

DCS is often compared with SCADA as SCADA incorporates a distributed data base usually referred to as tag data base which, in turn, is composed of tags or points i.e. data elements. A tag can be actual input/output within a system, or, it may be a soft point which evolves as a result of mathematical or logic operations applied to other points. Points can provide the history of the process as these are stored as 'value-time' format.

In general SCADA has a master terminal unit (*MTU*) which is the brain of the entire system. It would have one or more *RTU*'s which collect data locally and send these to *MTU* under its command. Collection, interpretation or management of all these data are done by customized or standard software. A SCADA can manage data elements (*I/O*) of the order of ten million or even more with the evolving technology. On the lower side it is a few thousands only. In comparison *DCS* is much smaller and communication there is carried out through *LANs* as described earlier. It operates in closed loop control, is faster and highly dependable.

SCADA has been discussed with a little detail in 'Telemetry Principles (TMH) by the same author. The basic structure of the SCADA is shown in Fig. 9.33

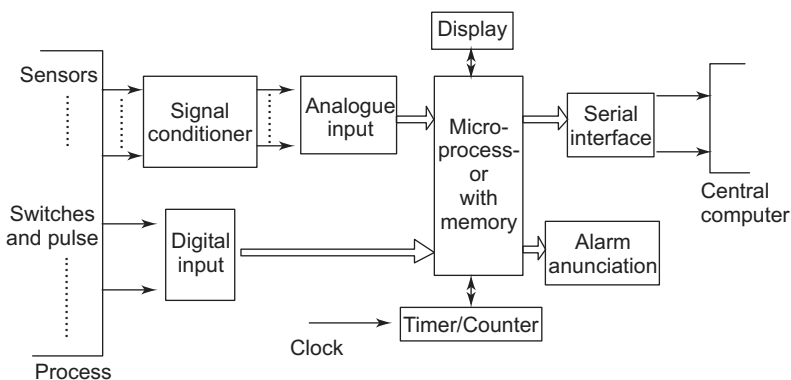


Fig. 9.33 The basic SCADA structure

For very large plants the conventional SCADA structure is extended by having more number of them. The scheme with a number of *DAS*'s

interfacing the central computer with 'star' connection is shown in Fig. 9.34. Other connections such as, daisy chain are also possible.

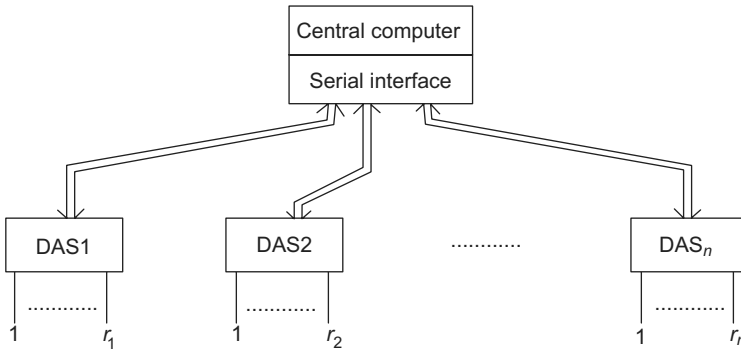


Fig. 9.34 Interfacing DAS with central scheme

9.8.5 Open Control Systems (OCS)

From the users point of view openness is defined as of having the capabilities to (1) integrate, (2) extend, and (3) reuse software modules in control systems (see Fig. 9.35). The required capabilities have to be supplied by the system platform of the control.

The architecture of such control has to have a system platform which is based on object oriented principle. It has three major components:

- (1) Operating systems,
- (2) Communication system, and
- (3) Configuration system.

The required capabilities of the modules are supplied by the system platform of the control as mentioned already. The capabilities are

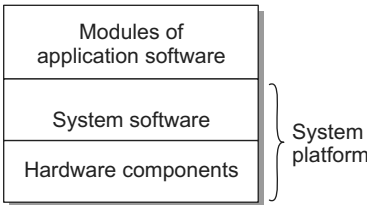


Fig. 9.35 Structure of modular control system

- (1) **Portability:** A module should be able to run in different control systems,
- (2) **Extendibility:** The module functionality can be extended,
- (3) **Exchangeability:** Replacement of the module should be possible with comparable functionality,

- (4) **Scalability:** Multiple adaptation of module is possible for enhanced performance,
- (5) **Interoperability:** Interchangeability or cooperative activity through data exchange in modules should be possible.

Above requirements lead to the derivation of the functionality of the system platform in so far as openness for modules (i.e. open control systems) are concerned. The platform is required to have the functionalities in conformity with the modular capabilities listed above. Thus, for

- (1) **Portability:** Application Program Interface (API) of the platform should be uniform.
- (2) **Extendability:** Necessary application of the module should be independent of the 'hosting' platform.
- (3) **Exchangeability:** Configuration that bind/hold the modules should be replaceable.
- (4) **Scalability:** Configuration system should be such that there can be multiple choice/accessibility of the modules.
- (5) **Interoperability:** Protocol on application layer should be standardized for the communication system adopted.

Obviously, for open control systems the platform must consist of the three basic elements.

- (1) The *operating system* which should provide the facility of parallel or quasi parallel execution of modules and there should not be any dependence on specific hardware.
- (2) The *communication system* ensures that the modules interact and cooperate in a 'standard' manner.
- (3) The *configuration system* which should help build a software topology with the modules available such that they can function both in time and space networking.

The system architecture of an *OCS* is shown in Fig. 9.36 the three elements mentioned above are integrated into the system platform such that these are amenable to be accessed through an Application Programme Interface (*API*). This interface can be designed to provide optimized solutions retaining the criteria for *OCS*. It is thus to be vendor-neutral allowing application modules to be used onto systems of different vendors. The application modules are of object-oriented nature and are required to be designed for simultaneous multiple access. These modules are also known as Architecture objects (*AO*) by some authors.

The Hardware

The platform hardware consists of processor boards, I/O boards, the necessary peripherals etc. The platform is required to be independent of a specific hardware. As and when necessary, suitable cost-effective hardware available at large may be integrated into it. The role of hardware is thus considered secondary in this platform although it is not avoidable.

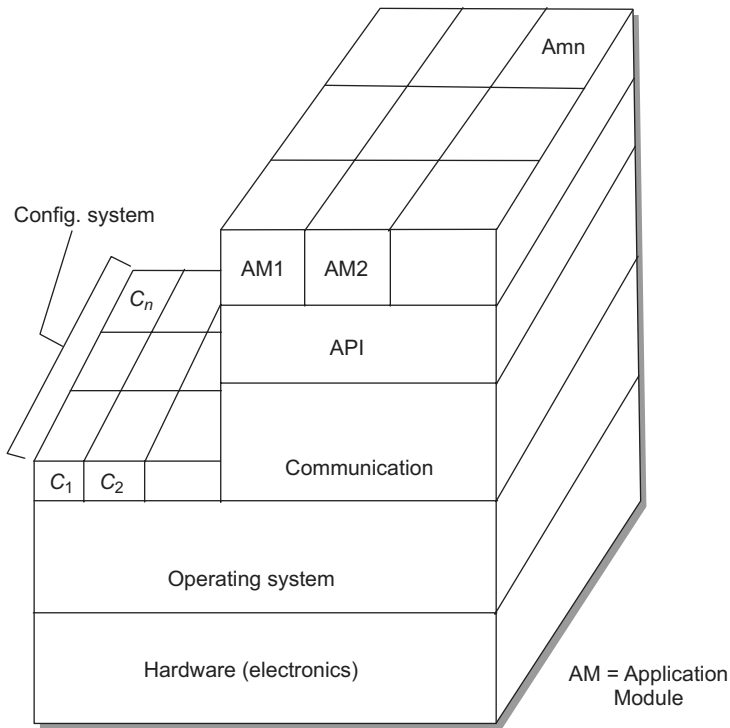


Fig. 9.36 System Architecture for OCS

Operating systems (OS)

Choice of operating systems may be kept open, instead the *API* is standardized. One *OS*, POSIX, is considered as an established standard for *OS* in the continent—it includes real time definitions making it a good choice for the purpose.

Communication system

This is the means to provide interchange of informations between AM's. It is required that the exchange of information should not only be between AM's on the same processor board but also between AM's on different locations (boards) connected, however, through a bus system. This requires a standardized protocol with uniform data formats having fixed set of instructions/messages.

The protocol architecture follows the *OSI* (open system interconnection) model but only selectively. The message transport system (*MTS*) takes help of layers 1 to 4 while the application service is supported by layers 5 to 7. The layers are represented in Fig. 9.29 (a).

MTS is adaptable to use any mechanism of information exchange such as LAN Protocols (*TCP/IP*), message queues etc. and is utilized for connection-oriented services for transport of message between AM's.

Application protocol is handled by Application Service Systems (*ASS*). Application protocol consists of

- (1) connection management,
- (2) message and disassemblage,
- (3) data conversion, etc.

The basis for application protocols which essentially is made of the communication objects like process, variables, events, is depicted in Fig. 9.37. The protocol is generated in a server/client basis using the object-oriented principle. The client views the server as a collection of a set of communication objects that can be accessed by the application services system for sending and receiving messages. In a server application module any information data or services which are assessable externally are mapped onto communication objects. AM's are therefore considered both as server and client.

The classes of objects shown are only there in Figure 9.37, there may be more. The variable class is for reading and writing data, the process class triggers actions (in the state machine), while the events class is for initiation of sending reports/events without being asked.

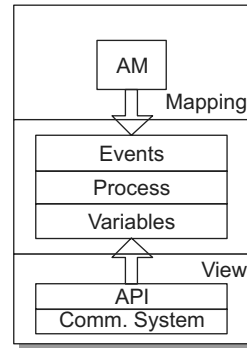


Fig. 9.37 Basis for application protocol

Specific application software is contained in the application modules (AM) for managing the communication objects (variable etc.) a specific software layer is introduced here which may be termed as communication object manager (*COM*). It holds the lists of the objects and renders the necessary services on request. The layout architecture is depicted in Fig. 9.38, where application software embedding (*ASE*) has been included. This is called the base architecture related to the object class.



Fig. 9.38 Base architecture object class

The base architecture is provided with the predefined links to specific functions such as

- (i) initializing,
- (ii) configuring,

(iii) resetting of the application modules.

The job of the programmer becomes easy—to test the required (to be selected) communication objects and provide the necessary links into the application software.

For achieving interoperability between AM's from different vendors, characteristic set of communication objects for every AM is to be defined, in the process, the functionality and external behaviour are also specified.

Configuration system

The actual topology of the system is required to be generated at the boot-up of the system making the system modular and in consequence the configuration becomes flexible/dynamic. The platform houses the configuration system which is designed to handle a library of AM's of different classes. At boot-up the AM's of different classes are chosen/selected and the communication between different AM's are established. Figure 9.39 shows the scheme. The actual topology is described externally via/by Configuration Editor and configuration order—the outputs of which are interpreted by the configuration system. The order is defined by the Editor graphically as is done in CAD system for layout design. The configuration order contains the list of all the AM's codes for a special control. For a class of AM, codes may be 'many' (diff.) for different control classes such as motion control, axis control (robotics) etc.

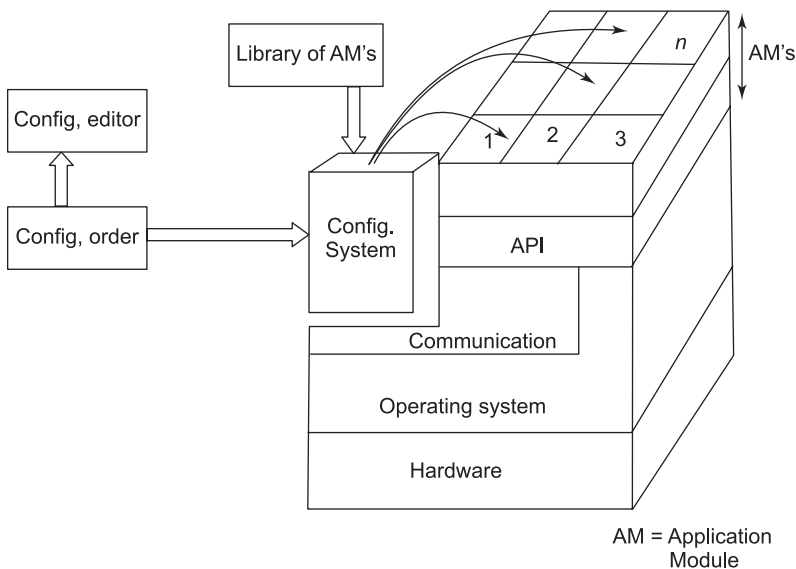


Fig. 9.39 Dynamic/Flexible generation of software topology

The configuration of control systems uses a PC-compatible computer for the operator and a special bus-based components system for control

of motion (for robotics, say) and *I/O*—the parts are connected by a serial bus—*TCP/IP* Ethernet is one such.

9.8.6 Open Connectivity

Open systems are yet to be standardized or even specified properly. Open systems are to have open standards and specification could be accordingly made. *OPC* has been in vogue for sometimes now meaning open connectivity via open standards. Interoperability is ensured through these standards. Based on fundamental standards and technology of the computation area the *OPC* foundation has been established which is creating specification for the industrial needs. A few such specifications have already been created. The *OPC* specification has been named recently as Data Access Specification which has defined a standard set of objects, methods and interfaces for use in process control and factory automations where interoperability has been facilitated. The initial specification (Data Access Spec.) was for printer-drivers. To support every printer separate software had to be written earlier, now, windows has incorporated the driver in the operating system. Microsoft's *OLE* technology accepted the *OPL* specification leading to standardization.

Some of the currently available *OPC* specification are :

- *OPC* Data Access
- *OPC* Alarms and Events including operator actions, messages etc.
- *OPC* Batch for specific needs of batch processes. It provides interfaces for the exchange of equipment capabilities besides operating conditions.
- *OPC* Data Exchange: Across Ethernet fieldbus networks communication is specified providing multivendor operability.
- *OPC* Security provides specification for controlling Client Access to the *OPC* servers for information relating to protection and guard against unauthorized modification of parameters.
- *OPC* unified Architecture basically provides a new set of specifications which will provide standards-based cross-platform capability.
- Basically, the approach is to have interoperability in multivendor systems and *OPC* standards are to facilitate this capability by conducting *OPC* certification programme, Interoperability workshops and Testing.

9.9 GENERAL COMMENTS

A well designed distributed computer control system would provide benefits such as

- (i) increased fault tolerance

- (ii) reliability,
- (iii) implementation possibility of hierarchical control algorithms,
- (iv) modularity,
- (v) flexibility,
- (vi) easy readability and maintenance, and
- (vii) reduced wiring costs.

A term 'survivability' is often used to convey the prime advantage a DCS can offer. It means the ability of a system to perform a set of functions satisfactorily over a particular time-span. Obviously this is manifest either by avoiding fault or tolerating fault. Fault avoidance and fault tolerance would require redundancy which may be made in the loop-level, i.e., within a local controller. Usually for rapid response and tight control redundancy is made in a set of local controllers or a backup in a higher level processor, when some degradation of performance with tolerance in response time can be accepted. In the former case a stand-by controller is run side by side with one operational controller. Diagnostic checks are carried out automatically at predetermined frequent intervals and on check data if it is found that the operational controller is not performing correctly changeover from main controller to stand-by controller takes place. On the other hand, in the latter case, for n number of operational stages, one spare stand-by is provided. This controller has to be larger than the usual one as it is required to compute concurrently with all the n controllers for their stand-by. Diagnostic checks are done from high level via this stand-by one. On fault of any one or more of the n operational ones the stand-by performs for it/them.

There is another technique of obtaining increased survivability which is by what is known as dynamic reconfiguration or reallocation. In this system all plant/process signals are available to all local controllers through data links and a coordinating controller which allocates the tasks to each such controller. When a fault occurs in any of these, the diagnostic checks continuously on in the coordinating system prompts this coordinator to reallocate the duties as per priority stored in the system. There is considerably increased complexity in the system which has been made for increased survivability. All these are possible, however, if communication itself has high reliability. Fig. 9.40 shows a simplified scheme of a system that can provide dynamic reconfiguration.

The programming and software part of the DCS is of crucial significance. It should have highly reliable real time programmes containing checking, diagnostic, fail-safe and redundancy features. The computer is connected usually to a wide variety of devices *I/O* interfaces and appropriate programme should be available to control the operation of data transfers through these interfaces. According to users' and designers' priorities, many subprogrammes of different importance and immediacy are to be

made available along with overall system software and an interrupt facility would provide the operator/designer to control over the priorities.

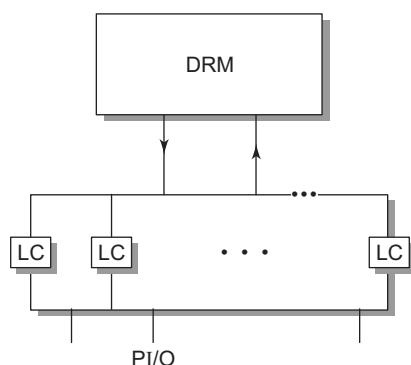


Fig. 9.40 System that provides dynamic reconfiguration; DRM: diagnosis and reallocation module, P I/O: process inputs/outputs

The design of a *DCS* is guided by many factors after its viability has been assured. Some of these are:

- (i) plant layout,
- (ii) process environment,
- (iii) sub-processes and their link/coupling with each other,
- (iv) sequence chain,
- (v) reliability requirement,
- (vi) nature of processes (batch, continuous etc.),
- (vii) response time and accuracy,
- (viii) types of sensors and actuators,
- (ix) complexity in supervisory control,
- (x) number of local loops per unit location, total loops and number of variables to be logged and displayed,
- (xi) operator needs,
- (xii) coordinated software,
- (xiii) management link,
- (xiv) possible extension in future of the overall system, its interfacing with old existing ones,
- (xv) possibility of system updating and its frequency, etc.

Only a broad outline of a *DCS* has been drawn and with increasing developments in different areas used by such a system continuous change in methodology is observed in the implementation of such systems as offered by different vendors. The presentation above has purposefully avoided a vendor-based description which, however, is necessary in actual field operation and is always available with the vendors themselves.

Review Questions

1. Discuss the development of *DDC* and show how the hierarchical type computer control has been adopted in the present-day process control system.
2. When a digital computer is to be used in a process control system, what are the initial studies to be made and constraints to be checked?
3. How are dead time taken care of in digital computer control? Make a software study of a simple integrating process for effective control.
4. Form the algorithm for the control programme of a microprocessor with a *PID* action for a process with a single dead time. What assumptions do you have to make for such an algorithm?
5. What are hold devices? What are their uses in a control loop? Obtain the transfer function of a zero order hold.
6. Why is *Z*-transform technique used in digital control systems? What are its disadvantages?

Is there any change in controller design when load change is effective in a system where originally set point change was considered? Explain with diagram.

7. The Kalman approach of controller design puts restrictions on $M(z)$ and $C(z)$ instead of $C(z)/R(z)$. If the response of the system is required to reach the final value in two sampling periods when a step input occurs find the controller function.

(Hint: Here $C(z) = \alpha z^{-1} + z^{-2} + z^{-3} + \dots$, with α arbitrary, and hence, $M(z)$ would take two intermediate values before assuming the final value β_f , say, thus

$$M(z) = \beta_0 + \beta_1 z^{-1} + \beta_f z^{-2} + \beta_f z^{-3} + \dots,$$

For an unit step input, $R(z) = 1/(1 - z^{-1})$, so that

$$C(z)/R(z) = (1 - z^{-1})(\alpha z^{-1} + z^{-2} + z^{-3} + \dots)$$

$$= \alpha z^{-1} + (1 - \alpha)z^{-2} = A(z), \text{ and}$$

$$M(z)/R(z) = (1 - z^{-1})(\beta_0 + \beta_1 z^{-1} + \beta_f z^{-2} + \dots)$$

$$= \beta_0 + (\beta_1 - \beta_0)z^{-1} + (\beta_f - \beta_1)z^{-2} = \beta(z),$$

But the process pulse transfer function $G(z)$ is the ratio $C(z)/M(z)$, $[G_a(s) \rightarrow 1]$, or. $A(z)/B(z)$. From Eq.(9.67), now $G_c(z) = B(z)/(1 - A(z))$.

8. The first order model of a process shows a process lag of 20 sec. Sampling frequency is fixed at 0.5/sec. What would be the controller function as per Dahlin's suggestion?

(Hint: Here $\tau = 20$ sec, so that with $G_a(s) = 1$, $G(z) = (1 - e^{-0.1})/(z - e^{-0.1}) = 0.1/(z - 0.9)$. As per Dahlin's suggestion, choice can be made for λ and τ_d for the desired response. Choose λ small for good response, let it be 2 sec and $\tau_d = 3$ sec, so that $N = 1$. giving

from Eq.(9.82), $C(z)/R(z) = (1 - e^{-1})z^{-2}/(1 - e^{-1}z^{-1})$ and $G_c(z) = (1 - e^{-1})z^{-2}/(1 - e^{-1}z^{-1}(1 - e^{-1})z^{-2}) \cdot 1/G(z)$. Using the value of $G(z)$ from above, now,

$$G_c(z) = (0.632 - 0.569z^{-2})/(0.1 - 0.0367z^{-1} - 0.06322z^{-2})$$

9. Write about the basic approach of the distributed computer control system. What are its advantages over the conventional digital control systems?
10. What forms of different data links are used in a distributed computer control system? What are their advantages and disadvantages?
11. Draw the flowchart for implementing the algorithm given by Eq. (9.14).
12. Foundation Fieldbus uses HSE and H1 Bus. Discuss these terms. (*Hint.* HSE stands for high-speed ethernet which runs at 100 MBit/s with random CSMA bus access (CSMA : carrier sense multiple access). Through a 'bridge', this connects the H1 bus which allows field devices powered over the bus and as per IEC 61158 - 2 spec. it operates at 31.25 KBit/s data transfer rate, coding is in Manchester coding.)
13. Discuss the features of a smart intelligent field device.
14. Briefly describe the structure of a SCADA and its utility in process control systems.
15. What specific importance is assigned to open control system (OCS)? How is it made open? Discuss the methodology adopted for making it open.
16. What is open connectivity? Comment on its adaptation in industrial area with special reference to automation.

Adaptive Control Systems

10.1 INTRODUCTION

If a control system is designed in such a way that when the process has variations in its characteristics and/or parameters, the control parameters automatically adapt/tune themselves to the needs of the process following certain preselected criteria, it is then called an adaptive control system. Obviously, depending on the types of variations in the process, different types of adaptation schemes may be designed. Generally, it is assumed that the variations and their types in the process are known *a priori* and further it is possible to design a programme to effect proper control of such variations by appropriate looping via the control variable and the controller. As long as the variations are expected, the control scheme self-adapts itself for restoration to the undisturbed condition and the system is then called a self-adaptive system. A simpler way to approximate this situation is to provide a second loop in the normal control system to regulate the damping. This presupposes that the first order control (or velocity control) is in order and enough for providing the desired dynamic control automatically taking care of acceleration and other higher order movements. Programmed setting of the controller parameters in relation to the expected process fluctuations resulting in process gain variations can produce the desired results of dynamic adaptation. For demonstrating the adaptive control system one has to first find out how the dynamic process gain varies with the relevant process variable. In the case where the former is inversely proportional to the latter, the controller gain should be varied linearly with the latter. In addition, if the loop period varies inversely with

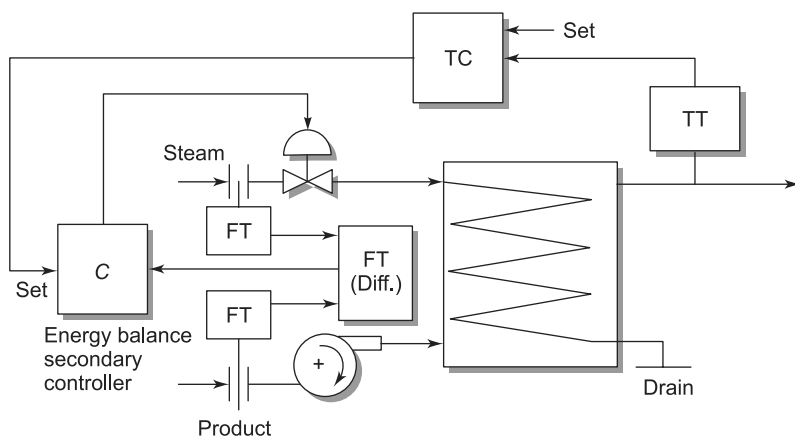


Fig. 10.2 Sketch of a programmed adaptive system as explained with a heat exchanger (C: controller; TT: temperature transmitter; FT: flow transmitter; TC: temperature controller; DIFF: differential)

and this response is achieved by simple feedback loops which are in turn obtained via approximate calculations. System identification approach is more complex and rarely justified for linear process dynamics. In this the process dynamical equations are to be accurately derived in terms of the system parameters which are obtained by on-line measurements. These equations are then used for the required controller settings. The approach based on the model reference technique is only an approximate one and under certain operating conditions or circumstances the optimum performance may not be achieved at all. But still, this approach is more akin to human behaviour. While it has been stated that this approach tries to obtain the desired response value, it is more appropriate, on the basis of approximating human behaviour, to design an adaptive controller which will follow a definite pattern like human beings.

The approach is simple. A prescribed response pattern (shape, with time) is continuously compared with the system output requiring that the set point be changed according to the response pattern. A continuous change in the controller gain is simultaneously necessary for the method to be effective. The required change in the set point is hardly possible by ordinary looping and obviously the adaptation of a digital computer in the loop ensures that the changes to be made are successfully carried out. This type of pattern-recognizing, adaptive-control system has application in both continuous and batch processes.

In the beginning it will be assumed that for a step disturbance the desired response curve is as shown in Fig. 10.3. The parameters τ_d and τ_l , represent the time for developing the response from 0 to 25% and from 25% to 75%, respectively. These are dead and lag time, respectively.

A controller is designed to identify these times. The parameters ϕ and ψ are the comparison

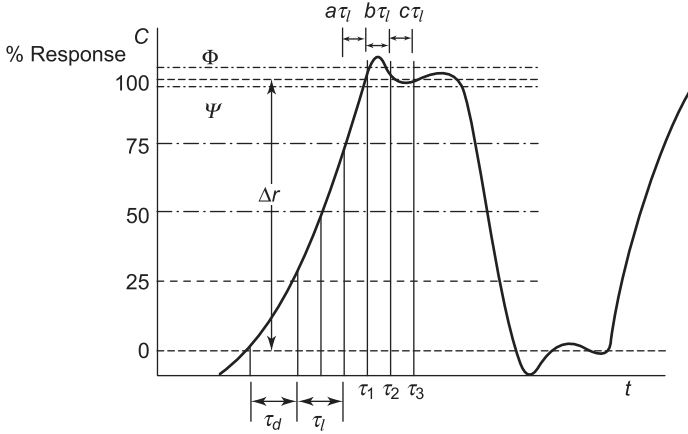


Fig. 10.3 Response with a step disturbance

levels, while a , b and c are multiplying constants such that a set of linear equations on the time-scale in the transient response part of the response curve can be established with τ_d and τ_l for proceeding with the adaptation programming in the computer. Now a , b and c require to be properly chosen such that a negative feedback is guaranteed with the comparison levels given by ϕ and ψ . The linear incremental time equations are

$$\tau_1 = \tau_d + \tau_l(1 + a) \quad (10.2a)$$

$$\tau_2 = \tau_1 + b\tau_l \quad (10.2b)$$

$$\tau_3 = \tau_2 + c\tau_l \quad (10.2c)$$

Now for an instantaneous error E_i , and initial proportional and reset gains K_c and K_n when a two-term controller is only used, the direct-digital control uses the algorithm

$$\Delta c = K_c \Delta E_i + K_R \cdot E_i \quad (10.3)$$

It is the controller gains K_c and K_R that require to be adjusted every time a set point change is effected and these gains are evaluated during the intervals τ_1 to τ_2 and τ_1 to τ_3 , the approximate relations for which are

$$K_{c(\text{new})} = K_c \left[1 + \frac{1}{b\tau_l} \sum_{T_x=\tau_1}^{\tau_2} f_p \left(\frac{E_i}{\Delta r} - \phi \right) \right] \quad (10.4)$$

and

$$K_{R(\text{new})} = \left[1 + \frac{1}{(b+c)\tau_l} \sum_{T_x=\tau_1}^{\tau_3} f_r \left(\frac{E_i}{\Delta r} - d\psi \right) \right] \quad (10.5)$$

where f_p and f_r , are preset scale factors adjustable manually or automatically for adjusting the speed of adaptation. New values of K_c and K_R are adopted at the end of τ_3 stopping the adaptation procedure. When the set point changes again and a deviation from ϕ and ψ is observed, the procedure is repeated with another adapted set of values for K_c and K_R . The adaptive terms in the above procedure finally are a , b and c .

Adaptive controlling is not a unique procedure in a physical system, mainly because it is extremely difficult, if not impossible, to choose an error criterion that would permit the controllers to be operative to optimize the system performance. Naturally, then, the solution of the problem lies in establishing the relevant error criterion for the individual system and performance requirement and a specially adapted controller may be employed for minimizing the error.

A judicious choice of a performance function that would guide the controller to economically control a process is possible when cost is considered as the prime factor. The cost function is due to: (i) failure of the system to perform the required duty and there is imperfection of the obtained product—this is reflected in the error and is called error cost, (ii) the requirement of power (or energy) for maintaining the desired condition in the process which is reflected in the manipulated variable for operating the actuator—this is called production cost, and (iii) the system construction including plant and controller—this is generally termed the plant or the overhead cost.

While the plant or the overhead cost is required to be considered only initially in a majority of cases and is decided by a fixed criterion computer solution (i) and (ii) may be considered as running cost and a performance function in terms of the weighted error (by ρ) and production cost (by λ) functions

$$p = \rho f_1(e) + \lambda f_2(m) \quad (10.6)$$

for minimizing cost may be adopted with the adapting controller being continuously manipulated by the performance function as shown in Fig. 10.4. Depending on the choice of functions f_1 and f_2 which are again guided by the requirement of “efficient” operation, adaptation becomes more complex. By way of example, we assume f_1 , and f_2 to be linear, such that

$$p = k_1 \rho e + k_2 \lambda m \quad (10.7)$$

Now, normally in the system described in Fig. 10.4, one easily gets

$$e(s) = \frac{r(s) + u(s)G(s)}{1 + G(s)J(s)} \quad (10.8)$$

where all functions are expressed as functions of s , $J(s)$ stands for the controller transfer function and it has been tentatively assumed that the measurement block has a unit-transfer function. Also

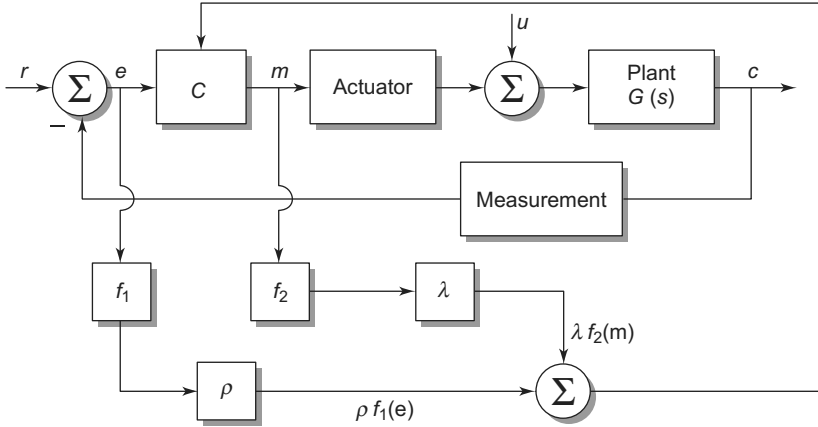


Fig. 10.4 Scheme for cost minimization with an adapting controller continuously manipulated by a performance function; ρ , λ : weighting parameters; $f_1(e)$: function of error variable; $f_2(m)$: function of manipulated variable; C : controller).

$$m(s) = J(s)e(s) \quad (10.9)$$

Combining Eqs (10.7), (10.8) and (10.9) and omitting $(s)s$, for brevity

$$\begin{aligned} p &= k_1 \rho \frac{r + uG}{1 + GJ} + k_2 \lambda J \frac{r + uG}{1 + GJ} \\ &= \frac{(k_1 \rho + k_2 \lambda J)(r + uG)}{1 + GJ} \end{aligned} \quad (10.10)$$

Usually it is the controller function that is to be adapted for minimizing function p . Hence, expressing Eq. (10.10) as

$$\frac{p}{k} = \frac{1 + \lambda' J}{1 + GJ} \quad (10.11)$$

where, for brevity, we have written

$$k_1 \rho (r + uG) = k = \text{a constant} \quad (10.12)$$

and we assume λ is the parameter that allows the controller to be adapted which is given in Eq. (10.11) by

$$\lambda = \frac{\lambda' k_1 \rho}{k_2} \quad (10.13)$$

For different values of λ' , J may be varied to obtain a minimum p/k which, as can be seen from Eq. (10.11), is not possible. This means that economic optimization is not possible in such a situation. However, if squared-cost functions are considered, such that

$$p = k_1 \rho e^2 + k_2 \lambda m^2 \quad (10.14)$$

for a range of values of λ' , proportional action alone may give rise to a minimum in the normalized performance-gain (proportional) curves. This can be easily shown by calculations.

10.3 SELF-ADAPTIVE SYSTEMS

In many instances it is also necessary to increase the number of errors or manipulating variables in appropriate functional form to cover the cost variations and consider their minimization. In such cases simple analyses with planar graphs are hardly satisfactory and a minimum-seeking loop has to be provided which adapts the controller in the system.

A typical block diagrammatic layout is shown in Fig. 10.5. A minimum seeking block actually performs the function $dp/dj = 0$ for $p = p_{\min}$.

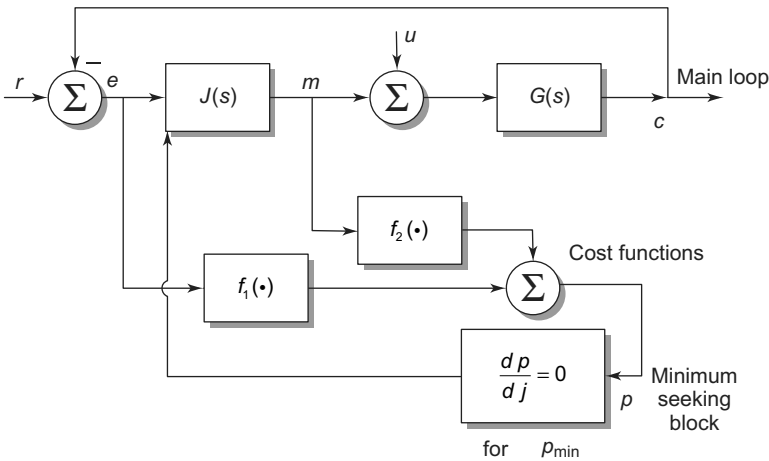


Fig. 10.5 Sketch of an adaptive system with increased number of manipulated variables; $J(s)$: controller transfer function; f_1, f_2 : functions

This is accomplished as

$$\frac{dp}{dJ} = \frac{dp}{dt} \bigg/ \frac{dJ}{dt} \quad (10.15)$$

There is basically, therefore, a division scheme which may only be implicit and this may cause a continuous system oscillation. For the system to be tolerant to change in the basic loop (process or/and controller) certain parameters are required in the adaptive loop which consists of the cost/performance functions and the minimum-seeking block in Fig. 10.5. These are required for increasing the adjusting facilities.

The type of the adaptive control system just described is often referred to as a self-adaptive system which is actually an optimization on the basis of the characteristic of the objective function at every instant and is accomplished by the feedback design. A simplified version of this is obtained when the manipulating variable is not considered to be of much importance for the feedback design as shown in Fig. 10.6 for a typical application. In this $J(s)$ is for a *PID* controller. Error is multiplied by a function $f(s)$ which is not exactly linear in the sense that a dead-band filter is used in the scheme for noise elimination. Dynamic gain is obtained by the ratio function of the low-pass and high-pass filter, while the integrator is used to retain the value in the steady-state gain. The parameter k has been shown in cascade with the low-pass filter (LPF) which is actually the adaptation parameter. This design obviously lowers the operational speed but that is not always a disadvantage. For example, a sudden change in the gain function in the main loop should not be and will not be immediately exactly reflected in the loop.

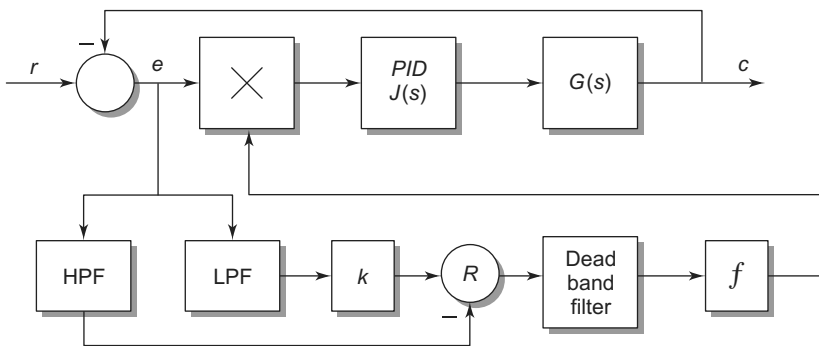


Fig. 10.6 A typical self-adaptive system; LPF: low-pass filter; HPF: high-pass filter; $J(s)$: controller transfer function; x : multiplier; J : integrator

Adaptive control systems naturally are intended to be self-adaptive. In continuous or periodic processes, when automatic adjustments for the process are to be made,¹ due consideration should be given to the changes (i) in process material (both qualitative and quantitative), (ii) in the process characteristics as also (iii) in the energy transport. For periodic processes, the programme needs to be changed befitting the operation for the particular stage and self-adaptation or adjustment may proceed according to this design. A still more involved problem of the self-adjusting control system is that the system itself is required to work out the best programme for controlling operations in specific circumstances that may arise during the run.

In continuous self-adjusting systems, in a generalized analysis, for the given index [such as product quality, target or raw material input, even economic (cost) function] or set of indices in relation to the possible disturbances (also called perturbations), a set of controlled variables belonging to an operational space is obtainable when the process is optimally controlled. In the relevant control terminology the index structure expressed as a function of controlled variable and upsets is known as “optimum entity structure”. Thus, from Fig. 10.7 the function $I(c, u)$ can be formulated from the possible us and cs .

Obviously, there will be a deviation in the index values from the given ones even in the best optimized states as the probability of the system to be ideal is almost zero. When the condition producing (i.e., causing) upset is too rapid a self-adjusting controller needs to be used and from the above discussion it will be apparent that unless a digital computer is requisitioned in the on-line basis it will be futile to obtain the self-adaptive operations for a completely general n -variable process.

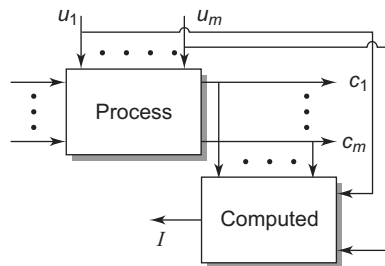


Fig. 10.7 Structural approach for generalized formulation of the system (u_1, u_m : upsets; c_1, c_m : controlled variables)

¹ In the strict terminology of adaptive systems in relation to formalisms adopted by theoreticians, controlled variables in space-time representation are termed “regime”, the process as “entity” and the desired values as “index” values.

A particular system using the above principle is shown in Fig. 10.8. This process is a gas-or oil-fired furnace in which a composite fuel mixture is used. The consumption of the fuel as also the supply of the oxygen-air mixture for a given concentration of oxygen-in-air needs to be optimized. The index could be the coefficient of “excess oxygen”, I , such that the deviation of this from a predicted optimum value I^* can be derived by computer analysis. If ε is the “oxygen-efficiency” of the flow material for a unit flow and q is the flow rate, the deviation is obtained as

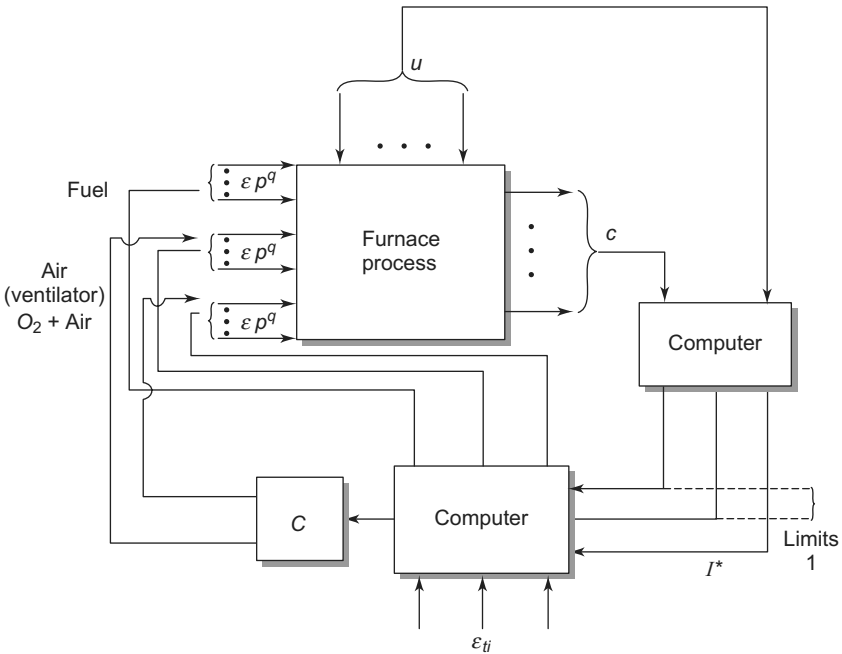


Fig. 10.8 Complete scheme using the approach of Fig. 10.7 for a furnace control; C : controller, ε_p : efficiency parameter; ε_n : theoretical efficiency; q : flow rate

$$\delta = I^* \sum_{i=1}^n \varepsilon_{ti} \cdot q_i - \sum_{j=1}^m \varepsilon_{pj} q_{pj} \quad (10.16)$$

where the second term represents the actual consumption and the first term the predicted one from the optimum coefficient as also from the theoretical “oxygen efficiency”, ε_{ti} . The parameter δ is now used to operate the controller for controlling the q s of the ventilating air and oxygen-air mixture, such that δ is minimum. In this problem, another interesting development is the optimization of the q s of the ventilating air and oxygen-air mixture for the desired minimum δ . The controller therefore, needs to be an optimizer. However, a complete optimizer includes the computers as

well. In fact, if the requirement is the maximum furnace temperature for the given fuel flow with a specified efficiency, the flow combination of the ventilating air with its oxygen concentration and the oxygen-air mixture with its oxygen concentration, has to be optimized with a search for a peak. Depending on the process, during its duration, there may be a number of peaks at different stages of the operation (as occurs in a steel-smelting furnace) which are shown in Fig. 10.9 for three different stages I, II and III. A simplified control scheme for a maximum c (here temperature) as a function of q_{air} and $q_{\text{air-oxygen}}$, i.e., for

$$c = c_{\text{max}}(q_{\text{air}}, q_{\text{oxygen-air}}) \tag{10.17}$$

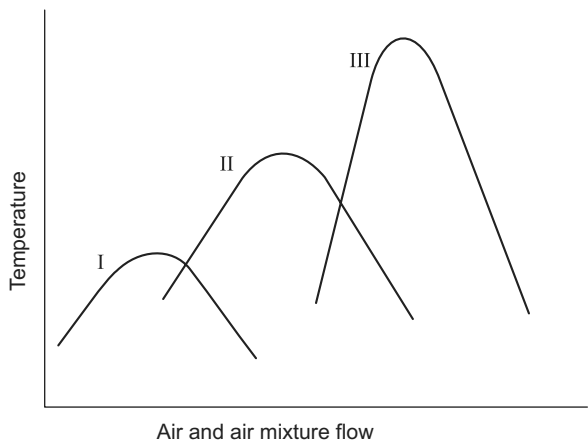


Fig. 10.9 Temperature peaks in a steel smelting furnace for different air mixtures

with a separate optimizer is shown in Fig. 10.10.

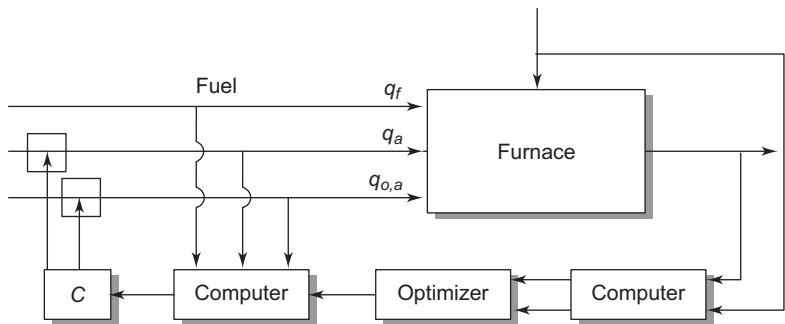


Fig. 10.10 Simplified control scheme for peak-value controlled variable with a separate optimizer (C: controller; q_f : fuel flow rate; q_a : air flow rate; $q_{o,a}$: oxygen-air ratio flow parameter)

10.4 PREDICTIVE APPROACH

It is time now to discuss about the three types of so-called predictive control systems. One can hardly imagine a predictive control system without the use of a computer. The one designed with the pattern recognition following a disturbance is often referred to as the 'exploratory control system'. This is an adaptive control system in which the controllable variable is perturbed to obtain certain effects which may be used to further adjust the controllable variables. Instead of using the process equations or measuring uncontrollable variables, the adaptive control first determines the current states of the process objectives and then uses a set of equations and/or rules to derive how the controlled variables should be manipulated for reaching the desired objectives. The predictive control system is by far the most rigorous of the lot that involves the measurement of uncontrollable variables such as the characteristics of the raw materials, ambient conditions, product markets, etc. The process equations in collusion with the uncontrollable variables determine the controllable variables, and then the desired control objective formulated which is usually complied with by certain adjustments made as already mentioned.

It should be remembered that the problem of optimization, self-adaptation or self-adjustment or self-tuning has been over-simplified here. Effective control involves the optimization of the working conditions in terms of mainly the upsets and the parameters of the control systems which are often approximately taken as lumped. When several such parameters vary simultaneously as well as in a distributed way by an appreciable amount, the optimum I^* (see above) is very difficult to formulate. The method of correlating upsets and controlled parameters and the use of a compensator for achieving the degree of compensation of the upsets in relation to the controlled quantities is also gaining ground because of marked improvement in the control of the process for the range and rate of disturbances that may normally appear. Figure 10.11 shows the scheme of compensation with the correlator.

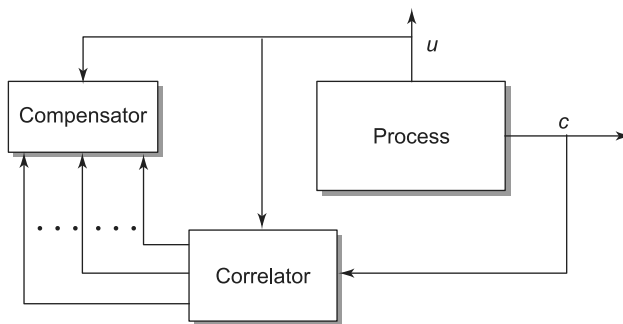


Fig. 10.11 Scheme of compensation with correlator

Generally stated, in predictive control, future occurrences and informations are predicted and are made use of in the current calculations and controller setting. The prediction period, often known as the 'horizon' is an important parameter. The calculation required in the predictive control becomes quite involved so that the services of a computer often turn out to be essential and the computer itself becomes the controller there. The computer is utilized to solve the complicated equations with nonlinearities and bounds and limits on variables for multiple-objective problems involving return, cost, energy flow, product quality, etc. The predictive control systems, as has already been said, try to obtain a desired objective without knowing certain variables by measurement and are only predicted in a prediction horizon. Such variables can be classified as unobservable ones. The success of the method obviously depends much on the 'accuracy' of prediction. The method is briefly explained as follows: If an objective ϕ is written as

$$\phi = \phi(u, c, v) \quad (10.18)$$

where,

u = upset, uncontrollable but measurable

c = parameter, controllable and measurable, and,

v = parameter, uncontrollable and unmeasurable

then the control equation for maximization will be given by

$$\partial\phi/\partial c = \phi_c(u, c, v) = 0 \quad (10.19)$$

which will solve for c as

$$c = c(u, v) \quad (10.20)$$

Depending on how the objective function in Eq.(10.18) has been laid out, the form of c will also appear as a function of the measurable parameter c and the completely unknown parameter v . A typical situation may be given as

$$c = \sum_{j=1}^n \gamma_j(v) u^j \quad (10.21)$$

As the γ_j s are dependent on v and are, therefore, varying; adaptation technique for solving c is to measure ϕ along with the corresponding c and u . Coefficients in Eqs (10.18) and (10.21) are continuously trimmed and Eq. (10.18) is kept current by an on-line adaptation algorithm. This procedure may be utilized to simultaneously correct the formulation of Eq. (10.18) and predict the system performance.

10.5 SELF-TUNING CONTROL

Self-tuning is an approach to automatic tuning of controllers for industrial processes. Like all other automatic systems discussed earlier this also requires computers for its implementation but often the self-tuning algorithms require predictable and modest computing power with its costs comparable to a *PID*-based system.

There is a sort of fuzziness as to the definition of a self-tuning control, specifically the control algorithm. A typical definition that commonly used, is in terms of known process dynamics. A feedback law is designed on this knowledge and a self-tuning algorithm would generate a control signal which would be same as that produced by such a feedback law when the number of input and output samples tend to infinity.

A self-tuning controller would normally have three basic elements arranged as follows:

- (a) the feedback controller modified by the
- (b) control design algorithm which is generated through changed 'parameters' available from a
- (c) recursive parameter estimator.

The recursive parameter estimator monitors the plant's output and input and makes an estimate of the dynamics of the plant in terms of a set of parameters within given constraints. These parameters then become the ingredients of the control design algorithm which, in turn, provides the new set of coefficients for the control law handled by the feedback controller. The control law usually is in the form of a difference equation. Fig. 10.12 shows the scheme of a self-tuning control system. The control design algorithm accepts the current estimates from the estimator for generating the coefficients. When the scheme consists of all the elements shown in the figure it is known as an explicit self-tuner. Alternately, the control design stage may be dispensed with and the estimator performs to reformulate the process model equations and produce the coefficients of the required control law. Such a system is known as an implicit one.

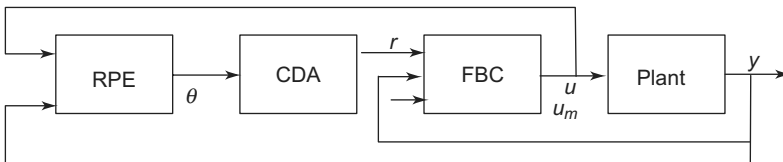


Fig. 10.12 A typical self-tuning controller scheme; RPE: recursive parameter estimator, CDA: control design algorithm, FBC: feedback controller

The predictive control theory which depends on knowledge of the system time delay τ_d , is the basis on which the simple self-tuners are designed. In

explicit methods, the time delay can be estimated as part of the process dynamics and hence this knowledge is not required. In the latter method, however, more computations are involved. This has other associated problems such as closed loop instability in some cases where no account is taken of the control efforts required. Consequently, simplest implicit algorithm is preferred for the general self-tuners.

The analysis of the self-tuning control system starts with the continuous time process model which is then converted into discrete time model, better written in predictive form. Then the recursive parameter estimation algorithm is obtained/derived and finally the prototype self-tuning controller is derived with the above to minimize an error criterion.

It must be remembered that all practical processes are to some extent nonlinear but good tuning can be obtained by using models or transfer functions of systems linearized around the current operating point.

As already mentioned a self-tuning algorithm usually generates a control signal following a feedback law designed on the basis of known process dynamics. This implies that the system is time-invariant but more general types cover time varying processes and algorithms are adjusted for effective control of such cases. It must be noted from Fig. 10.12 that variations occur in control design and in parameter estimation. The different control design approaches are summed up as:

- a. Stochastic minimum output variance control;
- b. Pole-placement technique;
- c. Amplitude and phase margin methods;
- d. Combination of minimization of output and control variances; and
- e. The linear quadratic Gaussian design

Likewise many different parameter estimation schemes are suggested. These are:

- a. Instrumental variables;
- b. Stochastic approximations;
- c. Least square and its extension;
- d. Extended Kalman filtering; and
- e. Maximum likelihood method.

In an implicit type design, the two are combined and, as mentioned, a predictor algorithm is formed for the objective to be achieved. This strategy involves the system delay (say k , for convenience) as the prediction horizon. This delay implies that the first output that can be influenced by the current control signal $u(t)$ is not $y(t)$ but $y(t + k)$. During this period, however, disturbance is acting on the system and if the required $u(t)$ is properly predicted, $y(t + c)$ would be 'optimised' with the neutralization of the disturbance because of the prediction. Prediction accuracy is thus very important which is dependent on the disturbance characteristics and interval of prediction k .

A typical process is given by the transfer function relation as

$$G_p(s) = \exp(-s\tau_d)B(s)/A(s) \quad (10.22)$$

For a disturbance $e(t)$, we can consider its contribution to the input as $C(s)/A(s) \cdot e(t)$, where in these representations of $A(s)$, $B(s)$ and $C(s)$, s stands for d/dt and hence the system model can be represented as

$$y(t) = B(s)u(t - \tau_d)/A(s) + C(s)e(t)/A(s) \quad (10.23)$$

However, the plant model is normally in the standard discrete time form so that $t = nT$ and the above equation can be written in terms of the forward shift operator z^{-1} as

$$y(t) = B(z^{-1})u(t - k)/A(z^{-1}) + C(z^{-1})e(t)/A(z^{-1}) \quad (10.24)$$

which, in the difference form, is written as

$$y(t) + \sum_{j=1}^n a_j y(t - j) = \sum_{j=0}^n b_j u(t - j - k) + \sum_{j=0}^n c_j e(t - j) \quad (10.25)$$

It is assumed that noise at some previous instants also affect the performance, otherwise $C(z^{-1}) = 1$ in Eq. (10.24). Now it can be shown that C/A can be resolved into an identity

$$C(z^{-1})/A(z^{-1}) = E(z^{-1}) + z^{-1}F(z^{-1})/A(z^{-1}) \quad (10.26)$$

where E and F are uniquely obtained by comparing coefficients of powers of z^{-1} with appropriate constraint in the degree of E and F .

Multiplying Eq. (10.24) by $E(z^{-1})$ and rearranging,

$$E(z^{-1})A(z^{-1})y(t) = E(z^{-1})B(z^{-1})u(t - k) + E(z^{-1})C(z^{-1})e(t) \quad (10.27)$$

Using Eq. (10.26), Eq. (10.27) changes to

$$\{C(z^{-1}) - z^{-k}F(z^{-1})\}y(t) = E(z^{-1})B(z^{-1})u(t - k) + E(z^{-1})C(z^{-1})e(t) \quad (10.28)$$

Replacing t by $t + k$, one obtains

$$\begin{aligned} y(t + k) &= F(z^{-1})y(t)/C(z^{-1}) + E(z^{-1})B(z^{-1})u(t)/C(z^{-1}) \\ &\quad + E(z^{-1})e(t + k) \end{aligned} \quad (10.29)$$

If the noise is not allowed to affect the system, noise may directly come to the output, then, we may also take $C(z^{-1}) = 1$, so that,

$$y(t + k) = E(z^{-1})e(t + k) \quad (10.30)$$

and

$$F(z^{-1})y(t) = -E(z^{-1})B(z^{-1})u(t) \quad (10.31a)$$

or

$$u(t) = -F(z^{-1})y(t)/[E(z^{-1})B(z^{-1})] \quad (10.31b)$$

Optimum prediction has been made with the error being made orthogonal to the prediction. The system diagram is now as shown in Fig. 10.13. Equation (10.31b) shows an implicit self-tuner, as the required feedback parameters are estimated directly rather than via a control design calculation, as already mentioned.

Self-tuning is a growing area and much interest is shown in it because of its possibility of practical implementation.

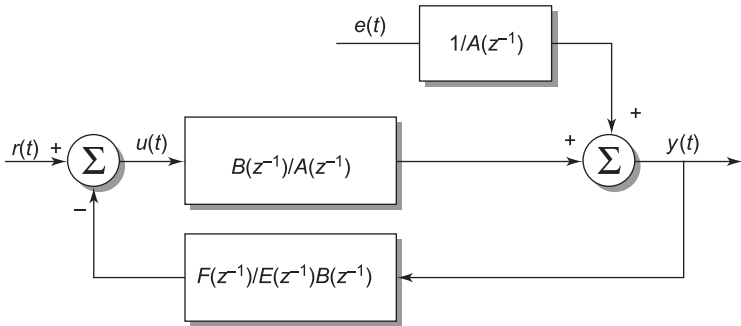


Fig. 10.13 A system in which error is made orthogonal to the prediction

10.5.1 A Practical Self-Tuner Via *PID* Algorithm

The *PID* algorithm provided with self-tuning feature has been introduced to obtain a user-friendly microprocessor based tool for controlling individual loops. It is the time-tested pattern recognition approach that is followed here. In this approach the closed loop is perturbed and the consequent response pattern is observed which is then compared with one that is desired. Knowledge of the process and experience enable the control engineer to adjust the control parameters for the purpose. The pattern, obviously, is the error-time curve which would have peaks or would not have depending on the amount of damping. Also, the features that are important in this pattern are (when peaks are obtained), time between peaks, i.e., the time period, deviation or overshoot, steady state error or offset and of course ratio of peak heights or damping. This approach of self-tuning monitors the process variable using direct performance feedback and determines the action required. The available tuning rules based on experience are used for self-tuning.

The algorithm monitors the closed loop recovery following a disturbance to set point or load, calculates the *P*, *I* and *D* parameters automatically to minimize process recovery time with the constraint of damping and overshoot specified by the user. The time period is also included for defining the shape. The integral and derivative times are actually normalized by the

period and the lead and lag angles of the controller are defined by these normalized quantities. Figs 10.14(a) and (b) show the error-time curves for disturbance to set point and load respectively. Overshoot is defined as $-E_{s2}/E_{s1}$, damping as $(E_{L3} - E_{L2})/(E_{L1} - E_{L2})$ and time period as T . The algorithm is designed to locate and verify peaks, count the time period length and the informations are to be stored for some time before the defined terms like overshoot damping, T_r/T and T_d/T are computed using Ziegler Nichols' or some other specified rules (for the latter two terms).

The computation of new *PID* values starts by using these stored informations to set directly T_r/T and T_d/T and proportional band PB which is computed on the basis of overshoot and damping, is subsequently adjusted to compensate for the changes in T_r/T and T_d/T . In final computation, the observed overshoot and damping are compared with the maximum allowed values set by the user. If the observed values are less, PB is decreased using a specified rule dependent on the difference in the values of the observed and specified ones. A properly tuned controller will not have its parameters changed for the same type of disturbances but would retune for the change in the type of disturbance or the change in the process.

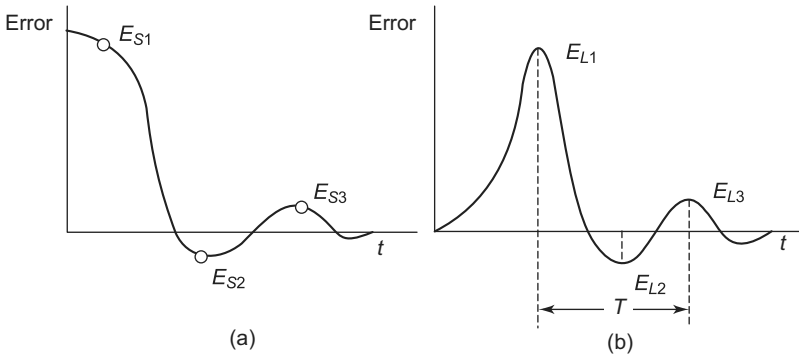


Fig. 10.14 Error time curves (a) set point change, and (b) disturbance change

For an overdamped system, pseudopeaks are considered which are assigned peaks based on the response curves that would give damping and overshoot (Cf. Ch. 4, c_m). When the process has comparatively large dead time, T_r/T and T_d/T need be given smaller values whereas for processes with dominant lag they must be larger values.

It would be of interest to note that the above self-tuner is based on what is known as expert system which is often defined as a computer programme that simulates the reasoning of a human expert in a certain domain. It thus uses what is called a knowledge base that contains facts and heuristics and some inference procedure for utilizing the knowledge.

10.5.2 A PI-action Based Auto-tuning Algorithm

The auto tuners that are commonly available to date are based on three specific methods: (i) that based on transient response as mentioned above, (ii) that based on frequency response, and (iii) that based on parametric models where tuning is done by recursive parameter estimation technique. Auto tuners can be characterized by their operating modes such as (i) tuning is performed on the demand by the operator, or (ii) it can be initiated automatically.

In what follows a new on-line auto-tuning algorithm based on the proportional and reset actions of a conventional *PI* controller is described. It uses the predictive approach with the prediction made through certain simple rules which are logically derived following a knowledge-base supported by existing expertise in the area. It is designed on the basis of the change in the manipulated variable as well as the error variable between two sampling instants. Utilizing their trends a new value of the manipulated variable is predicted and the desired algorithm is used to tune proportional gain K_c and the reset time T_r to obtain a dead beat control. Starting with the conventional *PI* control law.

$$m = K_c \{e + \int edt/T_r\} + m_0 \quad (10.32)$$

and using the simple trapezoidal technique for discretization, one gets the controller outputs at n th and $(n + 1)$ th instants respectively as

$$m_n = K_{cn-1}e_n + T \sum_{i=1}^n K_{ci-1}(e_i + e_{i-1})/2T_{ri-1} \quad (10.33)$$

and

$$m_{n+1} = K_{cn}e_{n+1} + T \sum_{i=1}^{n+1} K_{ci-1}(e_i + e_{i-1})/2T_{ri-1} \quad (10.34)$$

Subtracting Eq. (10.33) from Eq. (10.34)

$$m_{n+1} - m_n = -K_{cn-1}e_n + TK_{cn}e_n/2T_m \quad (10.35)$$

from which the tunable parameter ratio K_{cn}/T_m is obtained as

$$k_{cn}/T_m = (2/T)(m_{n+1} - m_n)/e_n + K_{cn-1} \quad (10.36)$$

All the terms on the right hand side are known except m_{n+1} which is now predicted through certain rules as already mentioned. The rules are

- (1) With increasing error, inverse extrapolation is used, i.e., $m_{n+1} \rightarrow m_{n-1}$,
- (2) With error decreasing but the magnitude remaining above 50 per cent of the set error range, double extrapolation is performed, i.e.,

$$m_{n+1} \rightarrow m_n + 2(m_n - m_{n-1})$$

- (3) With error decreasing but the magnitude remaining between 5 per cent and 50 per cent of the set error range, linear extrapolation is performed, i.e.,

$$m_{n+1} \rightarrow m_n + (m_n - m_{n-1})$$

- (4) Else, K_c and T_r are tuned to the lowest and highest values respectively with the set constraints on them.

With predicted m_{n+1} as above, K_{cn}/T_{rn} is now chosen following certain set constraints. It must be mentioned that if the error is attempted to be made zero in one sample as in the dead beat case, a very high value of K_c would, perhaps, be required which, in turn, may cause the response to become oscillatory at the initial stages. Also if $1/T_r$ is set at a high value a large maximum deviation occurs and stabilization time increases while with unduly low $1/T_r$ oscillation at the initial stage increases. Thus it is necessary that both K_c and T_r are varied with certain constraints to avoid the above situation. This knowledge base comparison of the analogue system suggests that at every sampling instant, these parameter values of the controller should be restricted to some per cent of the initial values which are set from the Ziegler and Nichols' criteria obtained from process reaction curves. These initial values are

$$K_c = 0.9/S\tau_d \quad (10.37)$$

and

$$T_r = 3.33/\tau_d \quad (10.38)$$

where S is the slope of the process reaction curve at its point of inflexion and τ_d is the process dead time.

Now, the process reaction curve may itself be either oscillatory and decaying or monotonically rising to saturation or expected value. The transfer functions in the two cases may be given respectively by

$$G_1(s) = \frac{\omega_n^2 \exp(-s\tau_d)}{s^2 + 2\zeta\omega_n s + \omega_n^2} \quad (10.39)$$

and

$$G_2(s) = \frac{\exp(-s\tau_d)}{(1 + s\tau_1)(1 + s\tau_2)} \quad (10.40)$$

where $\zeta < 1$, denoting undamped condition and ω_n is the natural frequency of oscillation.

The slopes in the two cases are given respectively by

$$S_1 = \omega_n \exp(-\zeta/\sqrt{1-\zeta^2}) \tan^{-1}(\sqrt{1-\zeta^2}/\zeta) \quad (10.41)$$

and

$$S_2 = (1/\tau_1) (\tau_2/\tau_1)^{\tau_2/(\tau_1 - \tau_2)} \quad (10.42)$$

The corresponding proportional gains are thus given by combining Eq. (10.37) with Eq. (10.41) and with Eq. (10.42) respectively. These are, therefore, the starting values in the two cases.

The value of K_c at the next instant is obtained by algebraically adding to the current value an increment

$$\Delta(K_c/T_r) = \text{initial value of } K_c/T_r - \text{value obtained from Eq. (10.36)}$$

with the new predicted value of m_{n+1} used in that equation. But this change must not be more than the set constraints given in per cent which the algorithm automatically takes care of by changing both K_c and T_r . Simulation and experimental results using this algorithm have been known to show improvement in performance in terms of integral absolute error reduction by about 10 per cent or even more. It should be remembered that adjustment or tuning is done by reducing reset action and keeping K_c as needed with decreasing error.

The technique is extendable to *PID*-based algorithm as well.

Review Questions

1. Where is adaptive control consideration important? In a single variable control how is it made effective?
2. In a generalized multivariable control system how can adaptive control action be adopted? Discuss with respect to a furnace process.
3. "Adaptive control systems are not fully adaptive"—comment.
4. What is a self-tuning control? Distinguish between explicit type and implicit type self-tuners.
5. What is parameter estimation as required in a self-tuning control system? What are its different approaches?
6. Discuss an implicit type self-tuner and show how a predictive algorithm may be designed for the purpose.

Process Control Systems

11.1 INTRODUCTION

After a generalized study has been made of the principles of process control in the foregoing chapters, it is time to introduce a few case studies of typical processes and plants with associated control schemes. These are believed to demonstrate that it is just not sufficient to know the principles of process control to be able to control any process/plant as per its demand. Each process has its own characteristics and the control of such a process would often call for specialized approaches. Yet the basic principles are no different. One major deviation in actual control principle from what has already been said is due to nonlinearity in process for which the strategy changes, but as has been demonstrated in Chapter 8 control at a nominal parameter value would allow linearization as well as application of the basic principles.

In the following a few typical processes are considered where the processes are first introduced briefly and then the control schemes that are usually adopted are discussed. It must be remembered that computers have changed the situation to a certain extent replacing the analogue controllers by the digital processors. However, the strategies otherwise have not changed substantially.

The processes considered here are (1) Boiler, (2) Part of a steel plant, (3) Part of a paper making industry, (4) Distillation column, (5) pH control, and (6) Batch process control. They are arranged in an arbitrary sequence and the presentation is done in more or less a piecemeal fashion. Each process has its own complexity and rigor and requires to be controlled on its own merit.

11.2 BOILER CONTROL

Steam used for generation of electrical power and also for other utilities in a plant is itself generated in a boiler. A boiler utilizes the energy latent in the fuel to convert water into high temperature steam. In a power plant this high temperature steam is converted into mechanical energy in a turbine which in turn drives an electrical generator. Both the power plant steam consumers and the utility series in plants consume steam at specified pressures and temperatures and the boiler is required to deliver steam at the desired conditions.

A boiler basically consists of a furnace process and a water/steam vessel which stands the steam pressure and temperature. Over the years the design of the boiler process has undergone evolution to suit the requirement. A typical scheme of the boiler process is shown in Fig. 11.1. There is a drum placed at a suitable position in the furnace from the bottom part of which

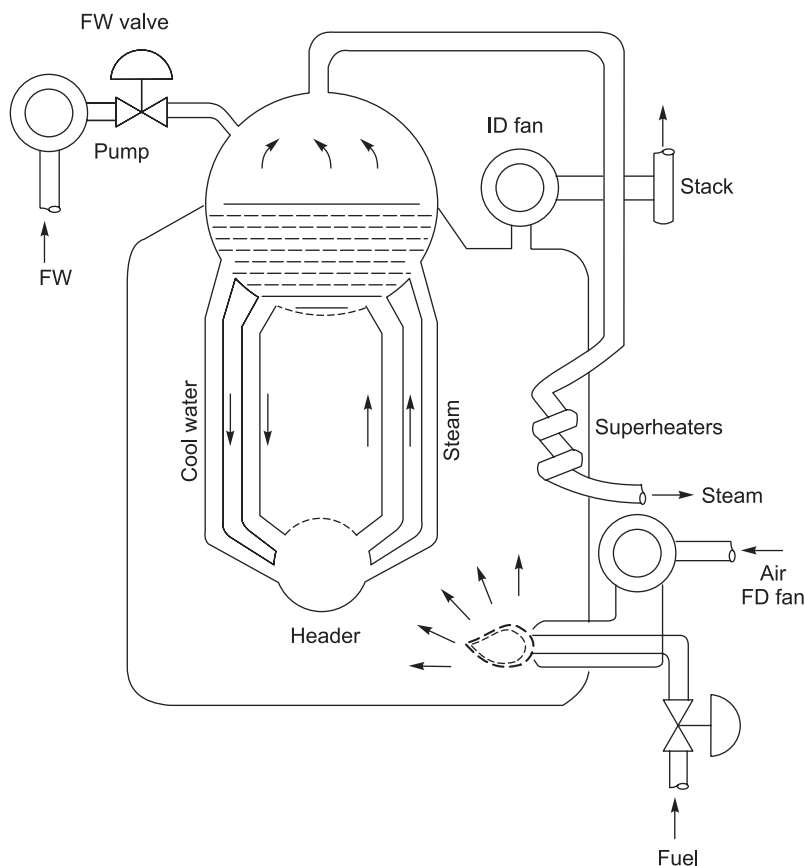


Fig. 11.1 *A typical boiler process*

a large number of pipes or tubes hang. These tubes carry the water down to distribution header at the base of the boiler through a relatively cool zone. From this header a similar array of tubes or pipes carry the water back to the drum through a zone where hot combustion gases flush from the burners. The tubes or pipes are small bore ones and the overall effect is to increase the surface area for the water to be exposed to hot gases. This exposure converts water to steam and a natural circulation induced by differing densities in the two halves of the thermal circuit forces the steam/steam-water mixture to rise to the drum. In large boilers assistance is provided by small pumps. Much higher rate of steam production is achieved here but steam being in contact with water cannot be raised above its saturation temperature without a corresponding increase in the steam pressure. Steam is, therefore, separately heated in superheater channels which are a system of heat exchangers put into the furnace as shown. Steam is heated here without being in contact with water and its temperature is raised much above the boiling point of water as is needed by a turbine. Its pressure also correspondingly increases. There are, in fact, a number of superheaters in succession each one picking up more heat to raise the temperature further. The superheaters are, therefore, arranged in the furnace in a specified way.

For assisting the passage of the combustion gases through the furnace in orderly way a fan assembly called Induced Draught (ID) fans or suction fans are used. This is a must for large boiler where the gases follow a complicated path through the heat exchanger to reach the chimney for being vented out or processed for pollution-level control. Similarly for combustion of the fuel, air is needed and large Forced Draught (FD) fans are used which discharge through ducts into the furnace. This air, blown into the furnace, is heated first by passing it through heat exchangers where the furnace waste gas is allowed to circulate (not shown in the figure), this is what is known as recuperation process.

The clean water required by the boiler is fed by what is known as feed water valve placed in the feedwater line supplied by feedwater pump. This supply is from de-mineralization plant as water must be very clean since dismantling the boiler for descaling is a very costly process. Water is a very critical commodity and requires to be preserved and hence it is used in a closed circuit. Steam after doing its work in the turbine, away from the superheater, is condensed in condenser/sump to start recirculation.

Although boilers are used in many plants for the purpose of indirect heating or direct steam injection, probably the largest use of boilers is in electricity generation plants or power stations. Now a days power stations are rarely of the unit type. This type means that one turbo generator is fed by one or more boilers without paralleling in the turbines. Unit type organizations should, however, not be entirely ignored. A typical unit type arrangement is shown in Fig. 11.2 where the turbine is pressure regulated

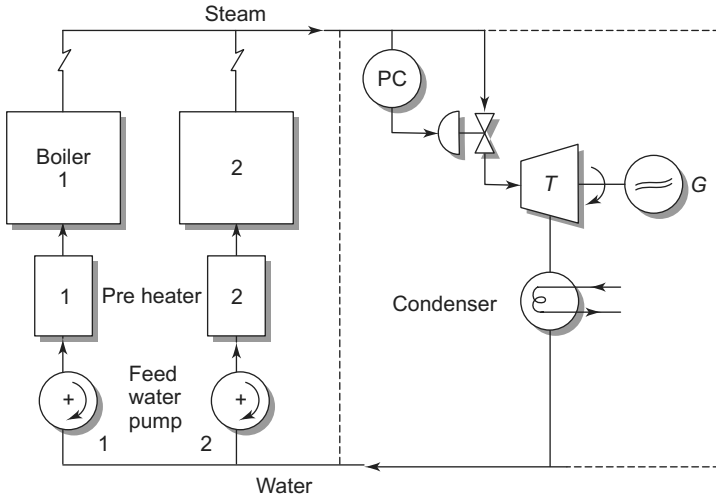


Fig. 11.2 Schematic of unit type organization in boiler control;
T: turbine, G: generator, PC: pressure controller

and is for base-load operation in which a fixed set point is adjusted as determined by the boiler output (in kg/hr or tonnes/hr). In this mode the boiler controls the turbine output and boiler output pressure is kept constant by a pressure controller.

11.2.1 Control Schemes

Given all the essential elements to a boiler-fuel supply, feedwater supply, air supply, one has to consider the proper functioning of the system for which following important aspects are to be checked: the level of water in the drum, the correct air-to-fuel or fuel-to-air flow rates with respect to the steam demand, the forced draught and the induced draught fans to prevent the furnace from becoming pressurized or sucking itself up the chimney. For modern-day large boiler systems efficient operation with perfect control requires high degree of interactions and complexity, and the vital areas where automatic control schemes are to be implemented are identified as

- (a) outlet steam pressure (the master steam pressure),
- (b) combustion control (air and fuel control),
- (c) feed-water,
- (d) furnace pressure, and
- (e) steam temperature.

As mentioned in Section 11.1 steam pressure control is the primary control in a boiler. Normally, for all boilers, the steam produced is collected at what is known as steam header or distribution header from where it is distributed to different users. Steam demand is thus reflected at this

header as steam pressure, with increase of demand, pressure falls and vice versa. Steam pressure is usually controlled by a PI controller with its set point manually adjusted for the desired boiler operation. The increase or decrease of steam generation on rise and fall of demand can be controlled by changing the rate of combustion. Hence the master steam pressure controller output sets the firing rate demand signal and is directly related to the combustion control scheme.

11.2.2 Combustion Control

Combustion control is, in turn, effected by control of fuel and air. There are three general types of combustion control schemes used now-a-days: (i) series, (ii) parallel, and (iii) series-parallel. In type (i) change in steam pressure in header causes a change in air-flow rate which in turn sequentially changes the fuel flow as shown in Fig. 11.3(a)

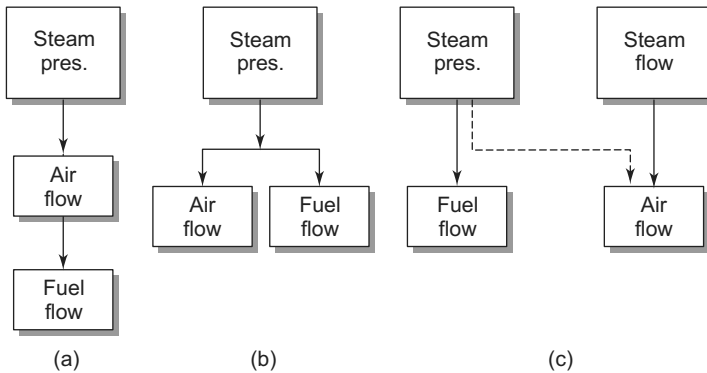


Fig. 11.3 Three types of combustion control schemes, (a) series, (b) parallel, and (c) series-parallel

The scheme is most suitable for boilers that are required to pick up load rapidly and shed it slowly, and here fuel becomes the secondary variable. It may be interchanged, with fuel becoming the main controlled variable from steam pressure signal with air as secondary one maintaining a series operation, for boilers having relatively constant steam load and with fuels having constant efficiency (Btu, calorific value). Also the boilers to which the latter is applied, are of small capacity ($<10^5$ lb/hr). Most commonly used scheme for boilers of any size using oils or gaseous fuels of fairly constant fuel calorific value, is the parallel scheme shown in Fig. 11.3(b) where steam pressure in header simultaneously adjusts both the fuel and air flow rates. Type (iii) is used where the fuel calorific value is likely to vary or the same is not easily monitored. A typical example is the pulverized coal firing scheme. The control strategy is shown in Fig. 11.3(c). However, in practice for solid fuels, a feed forward signal from steam pressure controller

is also given to the combustion air control. The principle of such a control (series-parallel) is as follows: Steam pressure set point variations are used to adjust fuel flow straightway. However, as steam flow is directly related to heat release of the fuel and since a relationship can also be established between heat release and air intake, steam flow rate is used as the index of this required air intake. The relationship, however, is valid only at steady load.

Actual hardwiring of the above is done through on/off, positioning (single point and parallel) and metering control systems. On-off is a preset adjustment and has single value control for both fuel and air. Positioning system simultaneously positions the forced draught and fuel valve to a set alignment. Depending on the mechanical means of aligning, the difference occurs in single-point (usually through a common shaft and cam-valve arrangement) or in the parallel positioning where adjustable cam positioned are used. The former is simple but used in single burner operated small boilers. The parallel positioning, in contrast, is able to independently adjust fuel and air through auto/manual stations. It is usable more for solid fuels like pulverized coal than liquid/gas fuels, i.e., for type (c) above with the difference already indicated. In metering control, combustion is controlled in accordance with the actually measured values of fuel and air inflows. These measured signals are used as feedback quantities for ensuring that flow is maintained in relation to the demand. This is the widely used method now. Because of metering, it can be effectively used to take account of fuel quantity change, barometric change, boiler performance, etc. to maintain combustion efficiency over wide load ranges and over longer periods.

A typical hardware scheme for type (b) of Fig. 11.3(b) is shown in Fig. 11.4 alongwith a modification made for varying demand during transients. The modification is done by incorporating high and low selectors in the set point control of the fuel and air flow controllers by the master steam pressure controller. During transients of demand, that is load change, this cross-limited (selector-based) metering control provides a positive interlock to prevent fuel-rich condition. The high selector ensures that the set point of the air-flow controller is either the master demand signal from master controller or the fuel flow signal whichever is greater and similarly the low selector ensures that the set point of the fuel flow controller is set either by the master steam pressure controller or the air flow signal whichever is lower. Thus the set point of the air-flow controller always equals or exceeds the set point of the fuel flow controller which makes air to lead fuel on load increase and to lag fuel on demand decrease. This is an extra measure of furnace safety but sacrifices the efficiency inherent in parallel control systems.

When load changes are rapid a steam flow rate feedforward circuit to the existing header pressure control scheme may be added so that a swifter corrective action to fuel and air controls can be imparted on extreme load changes. Figure 11.5 shows part of the control scheme with header pressure

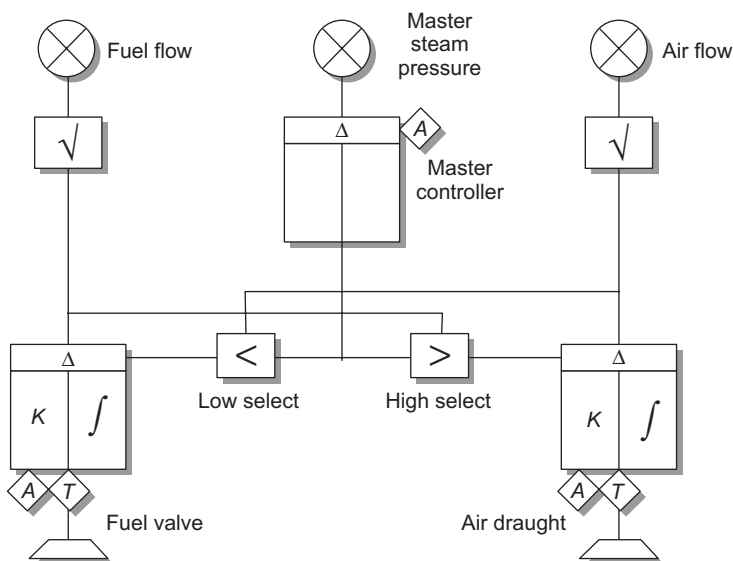


Fig. 11.4 A typical hardware scheme of Fig. 11.3(b) (Figure legends for this and subsequent figures: $f(x)$: specified/nonlinear function. ∇ : high limiting, Δ : low limiting, $>$: high selectivity, $<$: low selectivity, $+$ or $\pm$: bias. A: analogure signal generator. T: transfer)

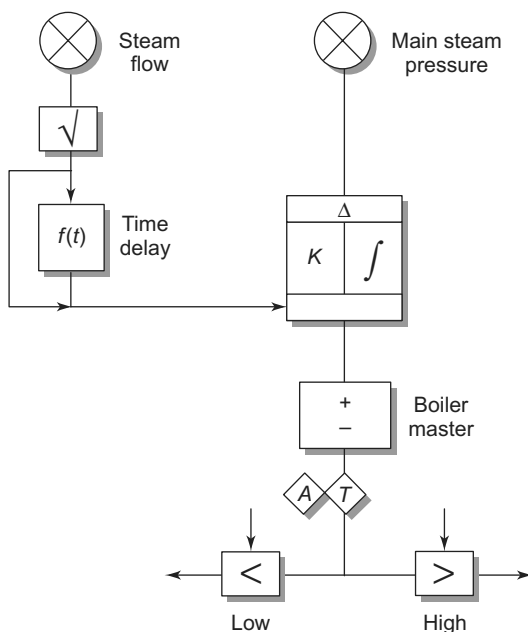


Fig. 11.5 Control scheme showing addition of steam flowrate signal to steam header pressure signal, (shown in part only)

controller output with a feedforward signal from steam flow summed with it and directed to the selectors via a boiler master. The linearized steam flow signal passes through an optional signal delay unit simultaneously with its usual course to the summer as shown. The summer output with this feedforward signal is passed on to the boiler master which passes the same signal to the air and fuel controls. The delay unit inverts the signal and holds it for a fixed period and then sends it to the summer to cancel the effects of the steam flow signal already sent to the summer and thus the main steam header pressure controller regains its full control.

In metering control often the fuel which is not directly meterable such as coal, an inferential signal is taken into consideration.

11.2.3 Optimizing Air-flow: Oxygen/CO Trimming

Even without oxygen or carbon-monoxide trimming air-flow control is often kept in check by modifying the air-flow measurement signal to the air-flow controller by a fuel-air ratio setter based on the calorific value of the fuel. The part of the scheme is shown in Fig. 11.6. The varying grade of fuel, however,

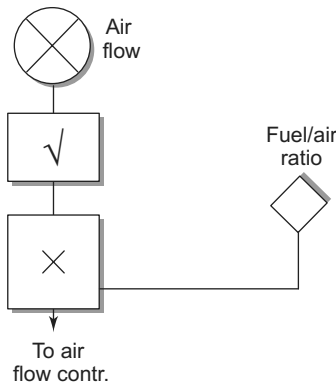


Fig. 11.6 Part of air-flow control scheme with fuel/air ratio setter

would demand varying fuel air ratio. Often for safety, excess air is permitted into the furnace which then becomes a necessary evil in boiler combustion processes as it adds significantly to heat loss. Without this there may be incomplete mixing, also insufficient retention time of the stoichiometric quantity of oxygen—these factors may lead to incomplete combustion and drop of oxygen below the calculated value at the burner during load changes may lead to a boiler explosion. The task, is, therefore, to keep excess air at the minimum level required for stable operation with effluent losses at minimum. The excess should be such as to maximize boiler efficiency by operating at a point (near theoretical value) where both combustible energy losses and effluent energy losses are minimized. This point is the

smoke point. Figure 11.7 shows plots of boiler heat losses versus excess oxygen (air). For trimming low-excess air the parameters that are of prime consideration are oxygen and carbon monoxide. The secondary parameters

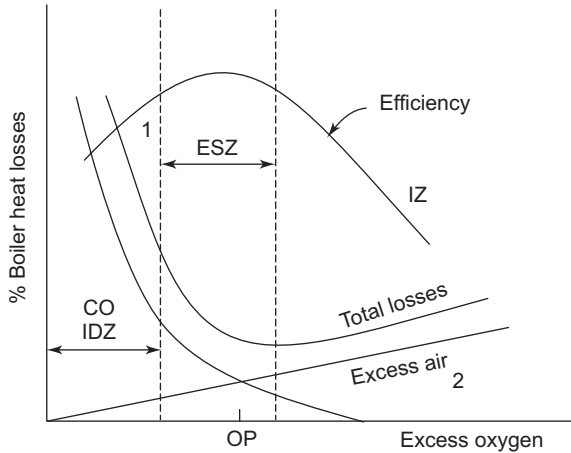


Fig. 11.7 Loss-efficiency curves in combustion control; IZ: inefficient zone, IDZ: inefficient and dangerous zone, ESZ: efficient safe zone, OP: operating point

are CO_2 , hydrocarbons, etc. Often in practice oxygen is used as the trim variable for control of excess air. However, both CO and O_2 are important to provide information relating to proper combustion and when used in combination they are likely to give the most accurate picture of the combustion condition inside the boiler.

Carbon monoxide is directly related to the amount of burnt fuel and its value is unaffected by infiltration of air from outside into the boiler; however, CO does not always provide the information regarding excess air. Carbon monoxide measurement is independent of fuel and boiler load. How much excess air should be there in a furnace? This, in turn, depends on the type of fuel used. In fact, for gaseous fuel it may be as low as 1%, for liquid fuels (oil) it may be 1.5 to 1.6 per cent whereas for solid fuels like coal it should be 3 to 4 per cent or sometimes even more, but the CO control range in all these cases remain around 150 ppm.

All in all, the base excess air trim is kept on oxygen monitoring and where economy of the entire system permits, CO trim is done to endorse the same as the CO analysers are at least 5 to 6 times more costly than oxygen analysers. Besides, addition of CO trim to O_2 trim provides a gain in efficiency by only 0.1 to 1 per cent.

As has already been mentioned metering system control has the facility to make sure that a ratio trim is incorporated as this allows the setting of a constant fuel/air ratio over the entire load range. An equal percentage trim

is usually preferred. A bias trim control is also recommended along with. Bias is often built mechanically in the positioning systems, for metering control systems an additional gain into it may also be incorporated.

Oxygen trim has been made easier for implementation after the zirconium oxide sensors have come to be used. Initially the trimming was manual, then it became semiautomatic when programming of oxygen set-point was being made which was then manually adjusted as per ratio of fuel to air requirement. In present day fully automatic case of oxygen trimming the steam flow signal is used to characterize the optimum air signal as a function of load of the boiler. This signal acts as the set point of the fuel to air ratio controller which is continuously updated with the actual value of oxygen in the flue gas. Figure 11.8 shows the part diagram. The air-flow controller also receives a signal from high selector fed by the fuel line and steam pressure line. In between the fuel line and the selector a bias trim with gain is also incorporated as already mentioned.

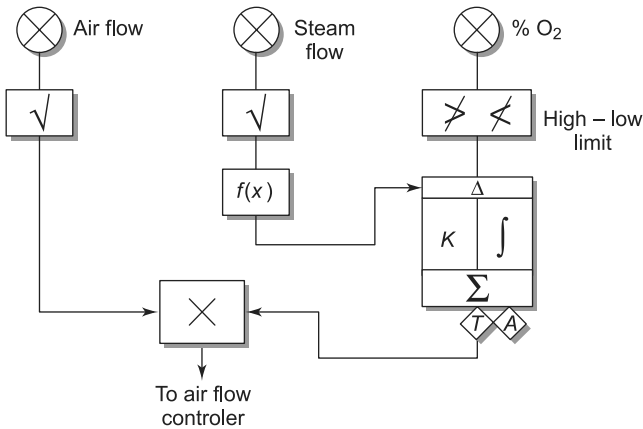


Fig. 11.8 Oxygen trimming for air flow control with steam flow acting as setter

Next CO measurement has been used in the above combination to improve the trim control to a great extent as it has the ability to control the combustion closely around the stoichiometric level. Carbon monoxide is used to fine-tune O_2 -trim when fuel has high carbon content such as pulverized coal or oil. There is an approach where this combined trim is used in a segregated manner—when load-change occurs trimming is kept on O_2 and in steady state condition it is on CO. It must be mentioned that manufacturers do differ in their choice of the combination which is based on the capacity of boilers, fuels used, likely load fluctuations, efficiency expected, etc. While economy remains the first priority. There are instances, particularly with grate-coal firing systems, where CO_2 also requires to be monitored and maximized.

11.2.4 Feedwater Control

Feedwater control system must be designed to maintain the mass balance with the expected load changes occurring in the boiler so that drum level remains within the limits of its safe and efficient operation. It is, thus, one of the most difficult controls as the types of boilers and their operating requirements are varying. If load changes frequently an adaptive approach is to be recommended. Complexity of control increases with the number of variables that can be measured for the purpose. A few are described in brief below:

(a) Single-element control

Such a scheme is dependent only on the drum level. Drum level deviation is taken into consideration for restoration of the same. When rapid changes do not occur in a boiler, such a scheme is good enough. It is usually recommended for small boilers ($\leq 10^5$ lb/hr). In such schemes response is very sluggish because drum level change occurs only after a sufficiently long time from the start of the change of steam or feedwater flow. Besides, when there is pressure fluctuation in the steam, water level in the drum shows volume change as transients resulting in what are known as swelling and shrinking. These are false changes. This will initiate control action which is not really needed. In fact, sometimes the indication becomes opposite and the control valve tends to close when it should open as in the case when water vapour is trapped within the liquid expanding the level instead of reducing it. The scheme is shown in Fig. 11.9(a).

(b) Two-element control

Steam flow changes as load changes and steam flow rate may be included as a second element and used as feedforward signal making the response much faster. The drum level remains the feedback signal. The two element control scheme can handle moderate load changes. Figure 11.9(b) shows the scheme.

(c) Three-element control

The third element based on the feedwater flow itself is introduced for boilers which have wide and rapid load changes. The three element control is most common in industrial and utility boilers. Steam flow signal is used as feedforward signal to compensate for swelling and shrinking during load transients. This signal and drum level signal are combined to act as set point control of the main feed water controller. Figure 11.9(c) shows the scheme.

To improve response accuracy a five-element control is also proposed, the scheme of which is shown in Fig. 11.10. The steam flow signal is compensated by the steam pressure (temperature) while drum level signal is compensated by the drum pressure signal.

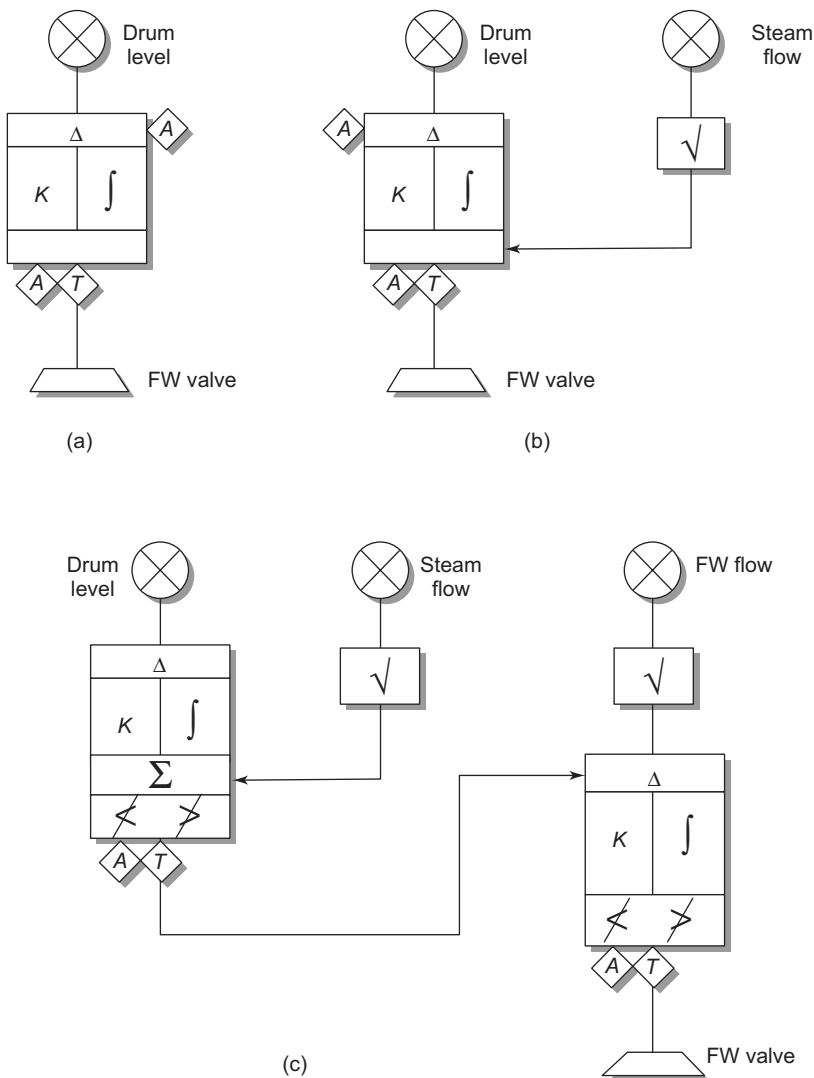


Fig 11.9 Feedwater control, (a) the simple single-element control, and (b) two-element control scheme (c) Three-element control scheme

11.2.5 Furnace Pressure Control

Slight low vacuum that is necessitated in the combustion chamber to ensure only low air infiltration in it is controlled by a draught controller. The draught is regulated by controlling the induced draught fan assembly or by controlling the flue gas damper against a manually adjusted set-point. The scheme is shown in Fig. 11.11.

The most popular method of steam temperature control is by using what is known as steam attemperation device. Very high temperature attained by the steam at the superheaters is controlled by this device by removing heat from the steam by spraying cold water into the flowing steam. There are three possible places where an attemperator may be located: (i) between the drum outlet and the first stage of superheater, (ii) between the successive stages of the superheaters and (iii) at the superheater outlet. Of these the most popular location is in between two stages of the superheaters. This arrangement is known to keep the average temperature of steam to the desired value at the outlet. It must be mentioned that a superheater is not a single piping structure but consists of a number of circuits for adequate heat transfer. Steam from all these circuits of the first stage is thoroughly mixed and cooled by the attemperator located at this position so that output steam from this stage enters the next stage at a uniform temperature. This, however, requires a bulk of piping work. Attemperator or spray-attemperator as it is often called, injects high purity cold water into the superheater steam through a spray nozzle in the throat of a venturi section made for the purpose. Interaction of the sprayed water through the nozzle and high velocity hot steam through the venturi neck makes the cold water to vaporize rapidly and to mix with the flushing steam thereby cooling it. The technique appears to be a very quick-acting and sensitive means of control for regulating the steam temperature.

The boiler superheater system, however, has a very large lag and steam temperature control often requires to be attained through cascade control loops. A typical scheme is shown in Fig. 11.12. Here the final superheater outlet temperature is measured and compared with the desired temperature, set manually or based on load, in the comparator of the controller C_1 whose output is used to set the controller C_2 which is used to regulate the inlet temperature of the final superheater. The spray-water valve often receives water from the feedwater main with a subsidiary line.

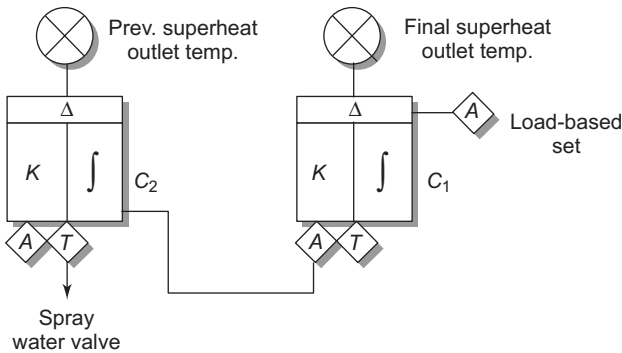


Fig. 11.12 Superheater control scheme

It must be mentioned here that besides injection coolers, surface or mixing coolers are also known. Such coolers serve two purposes—they maintain a fixed temperature at the turbine inlet as also prevent an abnormal rise in temperature in any single part of the superheater.

11.2.7 General Remarks

With the main control schemes indicated separately it appears that feedwater control is best obtained through a three-element control scheme where there is a feedforward loop, the superheater temperature control is best achieved by a cascade control scheme using attemperator. The schemes shown in figures earlier have indicated fuel inputs which can easily be recognized as gaseous or liquid fuels through burners though solid fuels may also be considered in schematic form. Figure 11.13 has been drawn from different angles to show the solid fuel charging through a feeder pulverizer and it is indicated that there can be any number of them. To avoid complexity in the drawing, feedwater control and superheater control have not been incorporated in this diagram.

The main control in boilers is the combustion/firing control and this should be taken care of by control functions such as:

- (a) adaptation of energy inputs to the boiler,
- (b) their optimum ratio, and
- (c) adaptation of furnace vacuum for the given conditions.

Other considerations in boiler control are the on-line efficiency calculations and the heat transfer calculations which are more easily done by microprocessor based controllers. When a number of boilers are in operation their efficiencies are likely to vary due to unequal ageing or design difference. It is necessary to distribute the load among the installed boilers to drive an overall optimum performance of the entire system. This distribution is best decided on cost consideration—which may also be handled by a microprocessor based system.

11.3 STEEL PLANT INSTRUMENTATION/CONTROL SYSTEM

To discuss the control of a steel plant it is necessary that the basic idea of steel making be kept in mind for which the example of an integrated steel plant would be in order. Such a plant starts its work from receipt of raw materials, their handling and preparation plants like

- (i) the coal washery,
- (ii) coal blending,
- (iii) sintering,
- (iv) lime and dolomite plants.

The requirement of coke for smelting iron in the blast furnaces makes coal carbonisation plant essential. The process of carbonisation provides the required coke for blast furnaces.

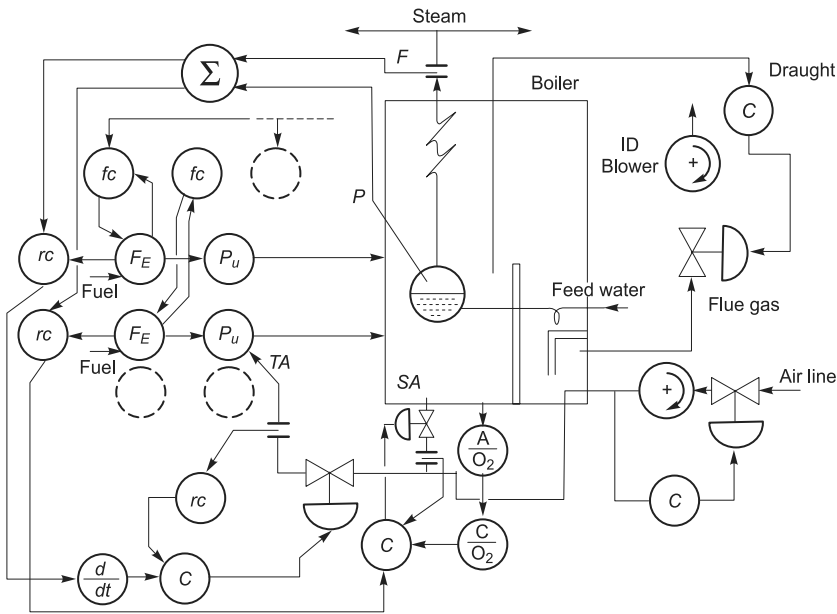


Fig. 11.13 Detailed control scheme of a drum type boiler, *fc*: fuel controller, *rc*: ratio controller, *F_E*: feeder, *P_v*: pulverizer, *A/O₂*: master oxygen controller, *C/O₂*: secondary oxygen controller, *ID*: induced draught, *TA*: transduced air (recuperated air), *SA*: secondary air, *F*: flow, *P*: pressure

It also generates gas as byproduct which is utilized as fuel for its own operation and the subsequent processes of smelting and shaping of steel. Other auxiliary units like the power plant, oxygen plant, foundry, engineering workshops are the integral part of a modern steel plant. The molten iron from the blast furnaces are transferred to steel making furnaces via its auxiliaries like the mixer bay and the desiliconising plant. The method of purification of steel varies from plant to plant; basic open hearth furnaces are one such technique. The casting bay of steel melting shop converts the molten steel produced from the furnaces into rollable forms by the process of teeming in moulds. These moulds are routed to the stripper bay where the hot ingots are stripped and finally despatched in ingot bogeys to soaking pits for reheating and further processing into finished product in the different rolling mills.

The elaborate and intricate activities as narrated above require measurement and control of process variables for successful contribution to the economy, productivity and quality of respective products. A moderate capacity steel plant would need around 4000 instrumentation and control accessories for its efficient operation. A block representation of the process mentioned above is given in Fig. 11.14.

11.3.1 Developing a Control Strategy

Making and shaping of steel requires massive furnace units working at high temperatures consuming high quantity of fuels—liquid, solid and/or gaseous, whose availability factor often overrides the technoeconomic justification for its selection. Efficient combustion control for regulating the heat input and hence final temperature at pre-determined desired values in the different plant units forms the bulk of control in a steel plant. For this the following arrangements are to be ensured:

- (a) Fuel at constant pressure (specifically for gaseous fuel);
- (b) Combustion air at controlled pressure and temperature;
- (c) Furnace pressure maintained at an optimum positive value to facilitate efficient heat transfer and for long life of the refractories;
- (d) Efficient recovery of heat from the outgoing products of combustion in regenerators, recuperators and/or waste heat boilers;
- (e) Provision for resetting the fuel/air ratio on appropriate analysis of the escaping flue gas;
- (f) Provision of safety shut-off valves as necessary.

Major sources of fuel for steel plant units are the coke-oven batteries (COB) and blast furnaces (BF). In addition to producing pig iron, BF produces one of the cheapest gas fuels. However, COB consumes 40 per cent of its own produced gas for underfiring and BF consumes 30 per cent of its own produced gas fuel for stove heating—the balance in either case is used for the rest of the plant units. The units which use only these gases for heating have a very simple strategy for combustion control utilizing a fixed ratio control with air. When mixed fuel firing is used, as in steel melting shops, where both gaseous fuel (from COB, for example) and liquid fuels (pitch and creosote mixture) are used, the control is complex particularly if the COB gas supply is poor (as in Durgapur plant) and then a second liquid fuel—usually furnace oil—is to be used as reimbursement. It then has an added problem of atomisation of fuel with steam and requisite fuel-air proportioning under changing fuel-mix condition and mutual fuel-air/steam ratio regulation. This problem is tackled by what is known as total heat input control (THIC). The scheme is briefly described here. A master controller, with variable thermal input scale and set-point, receives computed values of heat in calories of the three fuels (as the calorific value of the fuels often change) with due measurement and square root extraction—for final comparison against the set value. Output of the master controller will vary the dependent parameter which is either furnace oil or pitch-creosote. Summed up values of calories of the fuels will be utilized to control the air requirement. Additionally, a computing relay is used to proportionately deduce the combustion air demand when O₂ enrichment through burners is made. Another summation relay governs the demand of atomising steam with varied quantities of two liquid fuels.

There are cases where controlled condition is not directly measurable but it bears a definite given relationship with the measured condition. Soaking pit is such a case and is a very difficult to control. Soaking pit is used to uniformly heat the ingots from all sides to the desired level of temperature providing appropriate malleability/plasticity to the material without overheating. In the following, soaking pit control is taken up as an example of a very important steel plant process unit where there is a great scope of considerable energy savings.

11.3.2 Soaking Pit Control

Soaking pits are square, rectangular or circular shaped deep furnaces into which ingots are placed in upright position through a opening at the top. The oldest of the modern type is the regenerative type pit. The common types are the continuous fired design types also known as one-way fired pits. These pits are integral with high thermal efficiency recuperators designed to preheat the combustion air. In these pits combustion takes place above the level of the ingots where the space available is not affected by 'ingot coverage'. Ingot coverage is a term used to denote the tonnage or the number of ingots charged into a pit. Figure 11.15 shows a typical such arrangement. There are other designs such as bottom two-way fired pits, bottom centre-fired pits or vertically fired pits, etc. Choice of a specific design is guided by (i) the product mix, (ii) type of fuel used, (iii) ingot coverage and (iv) cost, etc. Since it takes several hours to heat an ingot to rolling temperature,

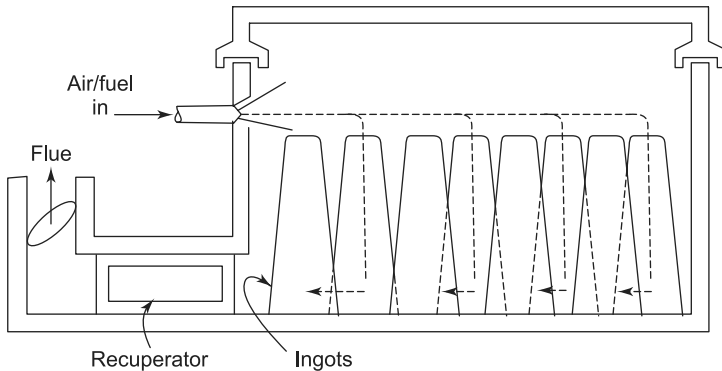


Fig. 11.15 Scheme of a soaking pit

a large number of soaking pits are required to meet the demand of a single rolling mill. Figure 11.16 shows a typical soaking pit installation in a steel mill. Hot metal, tapped from the steel making vessel, is brought to the ingot teeming area in a ladle where it is then poured into large moulds. These moulds are carried on ingot buggies, several such buggies are together called a drag. This drag is pulled to the ingot stripping area where the ingots

are allowed to cool down to be stripped off by the soaking pit crane. Ingots are transferred to an available soaking pit as soon as possible. If such a pit is not immediately available or at least within a reasonable length of time, the ingots are transferred to the cold ingot storage area for later use.

11.3.3 Control System Objectives

The control system for soaking pit furnaces should be designed for meeting many requirements a few of which are (i) to maintain safe furnace conditions in normal operation as well as in start up and shut down, (ii) to provide means to heat uniformly and consistently a wide variety of product mix, (iii) to increase productivity through optimum production scheduling, (iv) to protect material from excessive heating or scaling with abrupt change in demand, (v) to provide easy and flexible operator interfacing, (vi) to minimize the effects of changes in demand for enhancing furnace life, (vii) to provide monitoring and alarm functions, data logging and, energy consumption and production records.

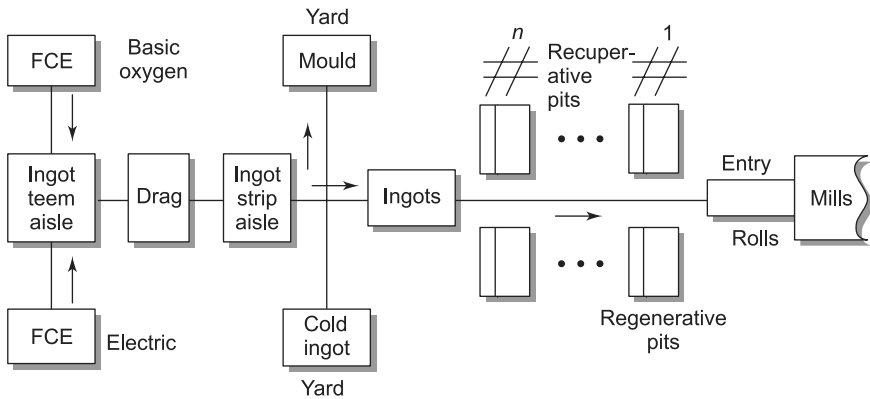


Fig. 11.16 Scheme of a system from steel making to rolling

Combustion control in a soaking pit performs two basic functions:

- (i) firing rate demand calculations, FRDC, and
- (ii) soaking pit temperature control (SPTC).

The function FRDC determines the best heat input rate required to be maintained at any time depending on the current soaking pit conditions and function SPTC is responsible for maintaining the calculated heat input rate. The latter function is thus made to be a slave sub-loop to the former function. The relationship between the two functions is diagrammatically represented in Fig. 11.17.

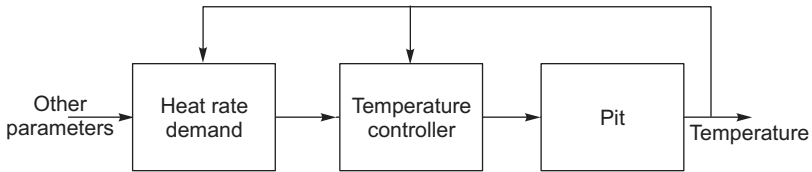


Fig. 11.17 Scheme of soaking pit control strategy

Earlier techniques were all for maximizing productivity of the mills and no care was taken of the pits as there is larger capital investment in the mills. In such cases fast heating of ingots and holding them in the pits was common. Increasing fuel costs have initiated ideas for minimization of fuel consumption in the soaking pits without sacrificing mill productivity. Models have been developed and firing strategies have been worked out based on such models that keep the rolling schedule, reducing the fuel cost.

FRDC is a function that provides for advanced control of a soaking pit and it can be performed easily with a computer as a part of the total control system.

A firing strategy is obtained that requires the initial thermal condition of the ingots charged in a soaking pit, the time to minimize gas flow, i.e., the time to switch the soaking cycle, time ready to draw a soaking pit, pit initial temperature, pit type, size and mixture of ingots in the pit, and pit efficiency. Besides, factors such as mill delays and pit conditions are also important.

The initial thermal condition is calculated from the particulars such as (i) elapsed time from teeming to stripping, (ii) elapsed time from stripping to charging, (iii) ambient air temperature, (iv) ingot dimensions and (v) ingot metallurgical data.

Firing strategy derived on such a basis is not an easy one but a linear firing strategy of the type shown in Fig. 11.18 is accepted because it can be easily implemented in a control system. Here the set-point is ramped at a fixed rate during the heating cycle. At the completion of this cycle, the set point is held constant until the ingots are completely soaked and withdrawn for rolling. To take account of any planned and unplanned mill delays, this strategy can easily be altered during the heating cycle itself by changing the ramp rate or by holding the set point constant for a specified time.

SPTC is a part of combustion control that performs the main functions of maintaining energy input to the furnace at a level demanded by the firing rate demand calculation (FRDC) function. Additionally it has to maintain a fuel to air ratio at the desired level, a safe pit condition and a slightly positive pressure in the soaking pit.

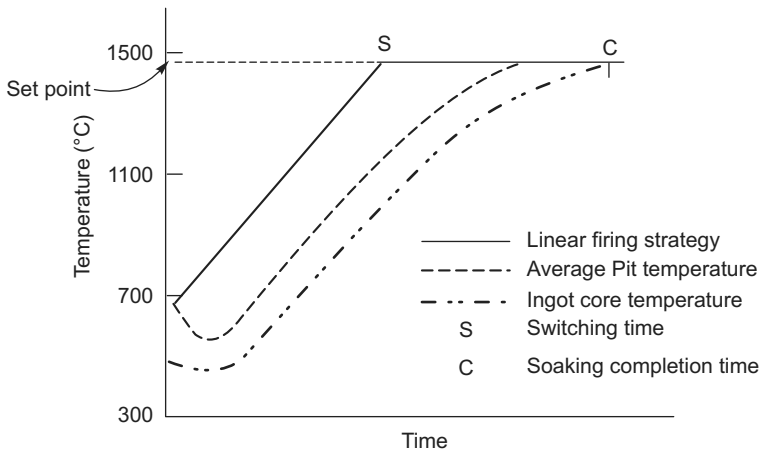


Fig. 11.18 Time-temperature curves in soaking pit indicating the control strategy

The most commonly employed technique for soaking pit control has been what is known as lead-lag parallel-series (LLPS) system as in boiler control. In this system, the firing rate signal is applied in parallel as the set-point to two slave flow control loops for quick response. One of these controls fuel and the other air. In addition, for pit safety, it provides interlocks between air and fuel flows with air leading on load increase and lagging on load decrease. The two commonly used modes in pit firing are (i) single fuel firing and (ii) duel fuel firing. The choice depends on the availability of fuels.

The block diagram of a typical single fuel firing LLPS control system is shown in Fig. 11.19. When fuel rate demand increases, the low selector rejects the demand signal and accepts the air-flow measurement signal, i.e., the lower one, with the fuel flow demand becoming equal to airflow measurement. At the same time the high selector rejects fuel flow measurement and accepts the increasing demand signal, i.e., the higher one with the air-flow demand signal (set point) becoming equal to firing rate demand. Thus the system acts as a series system with the fuel flow signal becoming equal to fuel set point in turn becoming equal to air flow measurement signal which equals to FRD signal, with fuel following air. However, when FRD decreases, the low selector accepts demand signal and fuel demand (set point) becomes equal to FRD signal. Now, as the fuel flow measurement signal becomes equal to fuel flow demand, and air flow demand (set point) also becomes equal to FRD for a short duration, a parallel state operation occurs for a short period. Subsequently high select operates and allows it to become series metering system again—this time air following fuel.

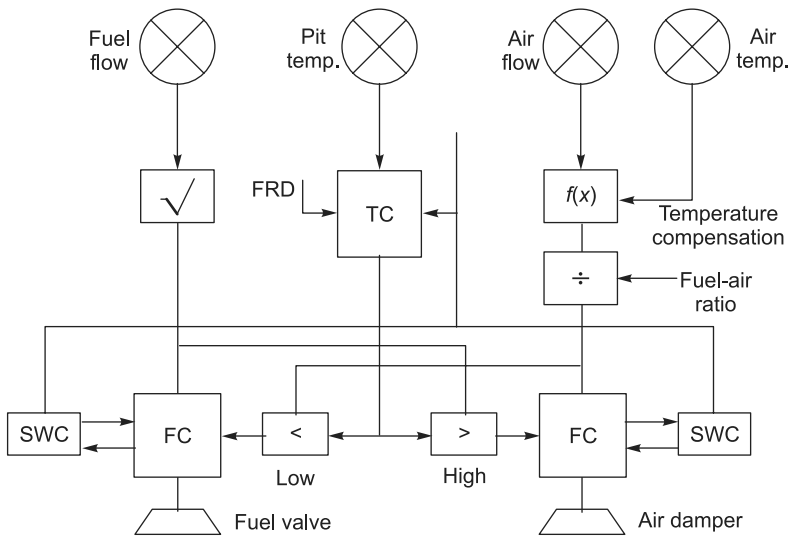


Fig 11.19 Lead-lag parallel system of soaking pit combustion control, SWC: switching controller

The soaking pit control must, however, have such provision that would enable the system to go from 'full fire' to 'off' when the pit cover is being lifted and again back to 'full fire' when the cover is replaced. In addition, in the case of the regenerative type pits it is also required that firing be periodically switched from one side of the soaking pit to the other. The light implementation of this feature is easy with microprocessor-based controllers. One particular way to implement this feature in a microprocessor-based system is by introducing switching controller as shown in the figure. After receiving the signal to switch 'firing off' from the pit reversing logic or pit cover logic, the switching controller performs the following functions:

- (i) switches the flow controller into manual mode, holds the output at the current value;
- (ii) saves the current value of the output and high limit of the flow controller;
- (iii) forces the high limit of the flow controller to zero to turn 'firing off';

Likewise, after receiving the signal to switch 'firing on' the controller takes the following actions:

- (i) restores the high limit of the flow controller (because it was saved, see step (2));
- (ii) switches the flow controller into tracking mode, if provided, so that it tracks the output;
- (iii) restores the output to full firing value (as it was saved previously);
- (iv) provides a short time delay for the controller to stabilize; and

- (v) finally switches the controller to auto-mode and the control occurs in accordance with the set point received from the temperature controller.

Dual Fuel Firing Pit Control System

It has been mentioned earlier that soaking pits and other furnace installations may use more than one fuel for reasons of economy and availability. In such situations it may be desired to fire a particular fuel on optional or preferential basis. Of the many different schemes proposed the most appropriate configuration of such firing types depends upon the operating and economic considerations of a particular installation.

As already pointed out, use of the blast furnace gas as fuel to the soaking pits and various other combustion processes in steel industry is quite common. This fuel contains only CO and is highly toxic. Besides, presence of nitrogen in it makes its heating value poor. It also carries fly-ash from the blast furnace which makes it dirty and abrasive.

Carbon monoxide as a fuel requires about 80 per cent of the quantity of oxygen required by other hydrocarbon fuels. This factor should be considered when the combustion control scheme is designed to use this gas as a fuel. Figure 11.20 shows a typical control scheme that uses blast furnace gas as a fuel (second), and is generally fired preferentially. Since in a steel mill, the gas may be available periodically and temporarily, the prime fuel burners are generally maintained in operation by base-loading the prime fuel which is accomplished by providing input (c) to the bias station to establish minimum set point to the prime fuel controller. The blast furnace gas controller set point (B) is calculated by subtracting the bias (C) from the firing rate demand (A).

The air-flow is usually calibrated in terms of the BF gas which is the preferential fuel. A part of the prime fuel measurement (approx 20 per cent) is subtracted from the air-flow measurement to provide for the additional air required for the prime fuel.

In addition, two summing units are used, one is for total BTU demand and is used to operate the high select interlock in the air-flow control loop. The other is used to provide the feedforward input to the pit pressure loop. Although the blast furnace gas has a lower calorific value, it has large volume and consequently loads the ID fans heavily for the same heat release. The second fuel summer can be used to account for this as well. The extra complexity added to the system is necessary because of the toxicity of the blast furnace gas.

Figure 11.20 also shows the automatic trimming of air based on the analysis of oxygen in the fuel gas. This trimming is accomplished by an oxygen trim controller in conjunction with the oxygen analyser. Set point to the oxygen trim controller is often received from a computer or set manually by an operator.

Fig. 11.21 A scheme of integrated 3-level system of steel plant control

11.4 CONTROL IN PAPER INDUSTRY

Pulp preparation from raw materials is the first stage in paper-making. This involves the use of 'digestors' where adequate control of steam inlet, pressure, temperature, etc. are necessary.

Prepared pulp is straightaway taken to paper-making plants but only after its water content, consistency and fibre-content are ensured to be up to the mark. When adequate supply of paper pulp is not there in the plant, air-dry pulp is required to be imported. In fact, the pulp is prepared at far-off places near to raw material resources and then sent to paper mills in air-dry condition. This air-dry pulp is 'prepared' to obtain a uniformly distributed suspension of fibres in water which is known as 'stock'. The stock is uniformly distributed on an endless wire mesh through which 'backwater' consisting of 'superfluous' water and some fibres, drains out. 'Surplus moisture' is then removed by suction, application of pressure and/or heat treatment. The latter two processing can be simultaneously done or done separately, by passing the stock through pressrolls and steam-heated dryer rolls.

Before the stock is worked with, the stock itself is prepared from air-dry pulp in what is known as 'hydrapulper' where it is mixed with measured quantity of brackwater and additives. Initially this mixing is done by pumping at 7 per cent concentration of fibres and then it is diluted to 6 per cent. The contents are then pumped to chest known as 'soak-chest'. Throughout the stock preparation close control of consistency is made which requires accurate control of brackwater supply to the hydrapulper. Large tanks of about 15,000 gallons are used as measuring tanks which are fed by diaphragm type positioner-operated valves and emptied by gate type cylinder operated valves. These tanks supply brackwater to hydrapulpers. There is on-off operation in the process and the valves respond sequentially to accurate fast-response on-off pneumatic level control signals. Miniature slack-type pneumatic level transmitters are used for the purpose with matching controllers.

There are three desired levels of the tanks that need be controlled, (i) the filled position—which is the reference position during batching, (ii) the low position—which is the level after the release of the pulping batch, and (iii) the medium position—which is the level after the release of the dilution batch.

These levels are set by pneumatically loading the controllers. The entire system requires interlocking controllers, pressure switches, relays, solenoid valves, auto-manual switches, etc., and are locally provided in usual cases. The operations are as described below.

(a) Filling of the measuring tank up to level (i) is automatically done, when (b) through interlocks and pressure switches pulping batch push is

operated to release the material and filled tank drops to its preset level (ii) at this stage the control equipment causes the gate type outlet valve to shut opening simultaneously the inlet diaphragm type valves which are allowed to remain open till level (i) is restored, (c) by time-controlled operation and indirect metering when 7 per cent pulping in the stock has been made complete, a 'dultion operation' interlock works to cause the level (i) to fall to position (ii) diluting the stock to 6 per cent and this operation also initiates another operation to return the tank level to position (i), (d) at this stage the hydrapulper transfer pump starts which also opens a 16" gate type cylinder operated valve through which the stock is delivered to the soak chests.

Hydrapulper and soak chest levels are indicated with the help of small transmitters. Level alarms are provided by the combination of pressure switches and electrical relay devices which are used to operate pumps and conveyor interlocks as well.

In the filled state of the hydrapulpers, interlocks are used to run a metering pump for a preset time to add china clay slurry and to keep a gravity-fed control valve open to add alum solution until an integrating transmitting flowmeter sends the batch-completion signal.

The next phase is for transferring the stuff, which is 6 per cent pulp, to hydrafiners after its consistency is reduced and controlled. In hydrafiners the fibre bundles are split into individual fibres. Stuff consistency is related to what is known as 'apparent viscosity' for the measurement and control of which an effective head is generated for the stuff to flow through a tube at constant rate. The head control is done by adding brackwater into the pump inlet as per demand initiated by a wideband (PB) proportional plus integral controller working via purged level transmitter. Stuff enters the hydrafiners at a controlled consistency of 5½ per cent.

The hydrafiners consist of conical plugs provided with surface bars, the plugs revolve in conical shells whose inner surfaces are also provided with bars (Fig. 11.22). In operation, the plug may be 'inched in', 'inched out' or 'run out' (removed to a large distance from the shell) by proper selection to regulate the degree of rubbing and cutting of the fibres which are fed in at the small end of the conical shell. Parameters monitored at this stage are input stuff pressure and flow and work done on the stuff. A two-term pneumatic level controller allows the hydrafiner stuff to enter storage chests which are known as day chests. The hydrafiners are stopped by withdrawing the plugs when it is detected that there is low flow, i.e., high level in day chests and low pressure or low level in soak chests. The supply pumps are stopped after a predetermined suitable time delay. The process described so far is schematically shown in Fig. 11.23.

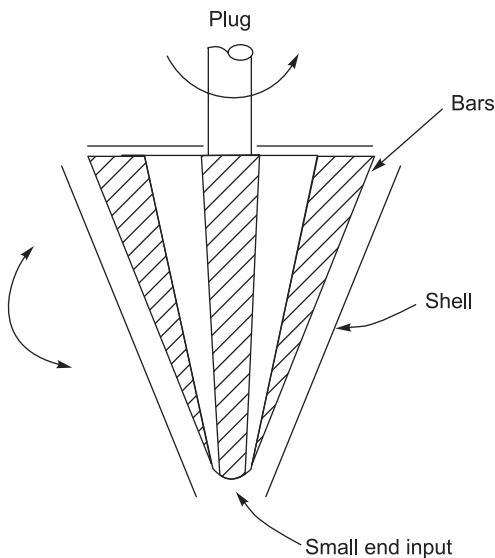


Fig. 11.22 Sketch of a hydrafiner

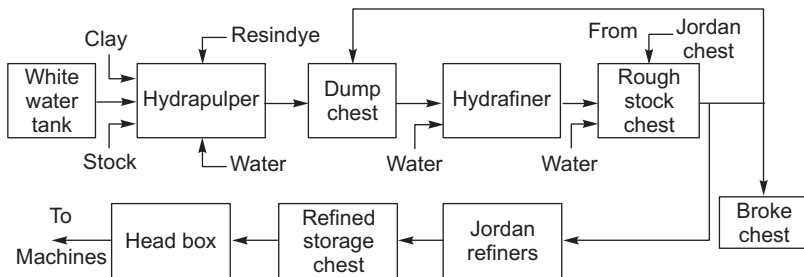


Fig. 11.23 Scheme of the basic process of paper pulp preparation

There are two types of pulp—groundwood and sulphite, each has its own technique of preparation. The two are mixed in appropriate proportion for paper-making, and pulp proportioning control, shown schematically in Fig. 11.24, is one of the major control systems used in paper industry. Each stuff is allowed to pass through area flow meters or magnetic flow meters at 4 per cent controlled consistency. A ratio control scheme is used for proportioning purpose. The sulphite flow is controlled through a level controller for the receiving Jordan chest which also receives the repulped broke (broke is a term used to denote trimming and spoiled paper from the machine) for use again. The high level of the Jordan chest is detected by an on-off level control installation which simultaneously operates the interlocks to stop the hydrafiners and stuff pumps and shut all input control valves.

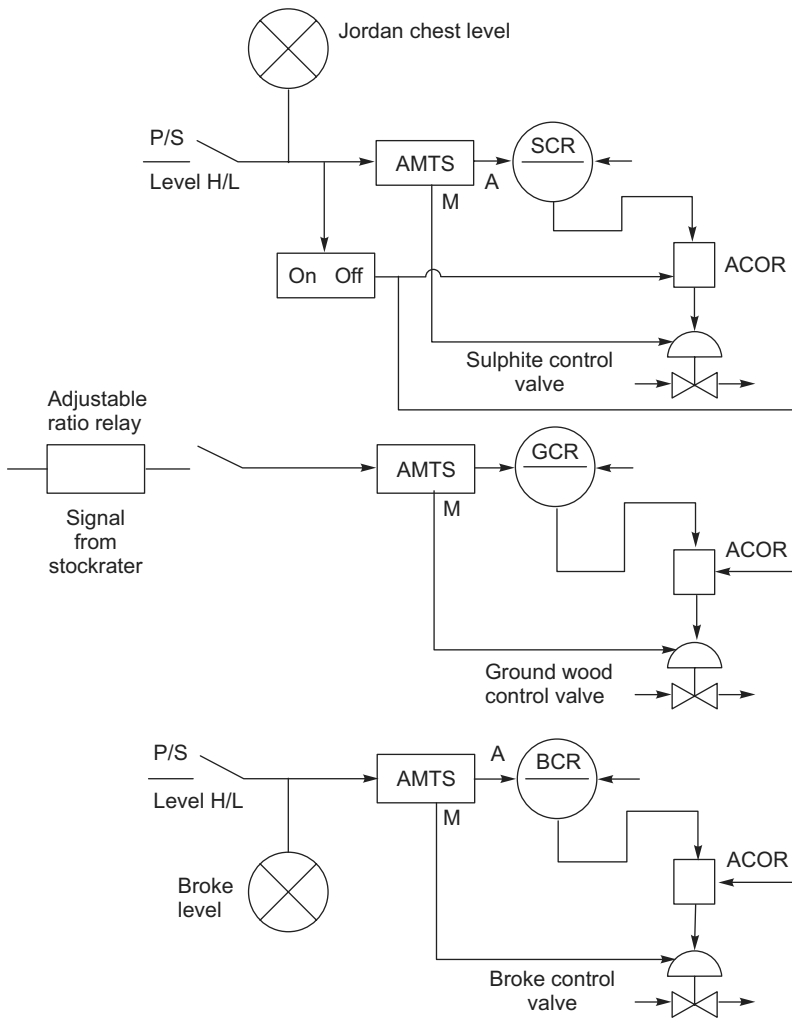


Fig. 11.24 Scheme of paper pulp preparation control, ACOR: alarm cutoff relay

The output passes to a constant head box (HB) through the refiners—the overflow from HB recirculating through the machine chest. The head box is fitted with a ‘stuff gate’ valve which is operated pneumatically. The desired weight and thickness of the paper is determined by the quantity of stuff and consistency leaving the headbox which obviously require metering. The regulated stuff is then further diluted, screened and cleaned and then passed into a pressurised flow box wherefrom it is ejected through a narrow opening, called slice, on to an endless moving bronze wire mesh. The diluted flow is also controlled. A comparison of flow rates in and from the headbox is used to control the consistency of the stuff

delivered in to the flowbox at the desired value of 1/2 per cent. At this stage the pressure of air over the stuff is controlled, the stuff temperature is recorded and pH controller regulates the flow of alum solution. The slice itself is pneumatically adjustable making possible to have control over the stuff speed on to the wire.

In the paper-making machines, a lot of interlocks and measurements are necessary, a few are

- (i) flow box pressure,
- (ii) photoelectric detection of paper break,
- (iii) pick-up suction (for dehumidification).

Paper from presses enters the steam-heated drying section at approximately 60 per cent moisture content and is to be maintained in moisture equilibrium with the atmosphere when it leaves the steam-heated cylinders. Control of steam supply is ensured at different sections.

11.4.1 Electric Drive and Control in Paper-making Industry

Paper-making industry requires extensive electric driving systems with proper control which may be different in different stages. These machines can be divided into three main groups:

- (a) machine for paper pulp preparation,
- (b) main paper-making machines, and
- (c) finishing machines.

The first group consists of individual driving systems with fixed speed for converting wood-pulp, waste paper, rags, straws, etc. into paper pulp which is subsequently processed further. Because of fixed speed requirement, induction motors are generally used for the driving systems of machines like pulp grinders, choppers, hydropulpers, etc. In a specific case of auxiliary driving system for a pulp-grinder angular velocity control is needed. This is because a proper pressure of wood blocks on the grinding wheel has to be maintained which is ensured by an auxiliary dc drive the angular velocity of which is regulated according to the power consumed by the main ac drive motor. The scheme is shown in Fig. 11.25. The induction motor M_1 drives the grinding wheel W in the toothed chain grinder T which is filled with wood blocks. The dc motor M_2 drives the chain system which presses the wood blocks to the grinding wheel. The control of this motor is from the amplifier A which receives the error from error block E that gives a signal which is proportional to the difference of the set signal S and a signal proportional to the power consumed by the ac motor which in turn is proportional to the pressure of the blocks on the grinding wheel. The power w is given by

$$w = \beta F \alpha v \cdot p \quad (11.1)$$

where, α = area of the grinding surface, F = frictional coefficient, v = linear velocity of the grinding wheel, p = pressure of wood blocks on the wheel, and β = a coefficient factor.

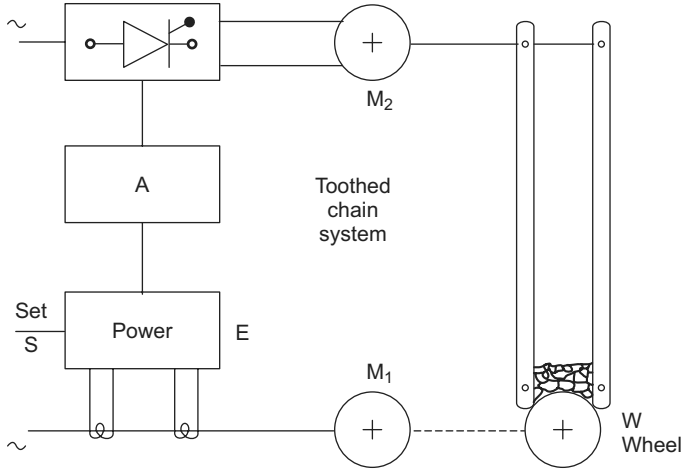


Fig. 11.25 Scheme of a pulp grinder control

Thus the error that controls the angular speed of the motor M_2 is proportional to this power or pressure as shown by Eq. (11.1).

The main paper, making part has a number of sections. Each section has its own set of operations. The sections are:

- (a) Sieve section, where preliminary dehydration of pulp occurs and web formations starts;
- (b) Press section, where further dehydration occurs and web structure starts to be consolidated;
- (c) Dry end section, where final dehydration occurs and the web becomes dried up to 95 per cent or so;
- (d) Calendering section, where smoothing, polishing, etc. are done and paper web of desired thickness is made;
- (e) Winder section, where the paper web is reeled on to a drum.

It must be remembered that thickness of the paper web passing through the above sections is gradually reduced, often to several hundred times and each section has to have its own angular velocity but must maintain a relation to the velocities of the adjacent sections to avoid rearing and folding the paper web. The angular velocity, may vary from a ratio 1:2 for coarse papers to 1:20 or more for fine papers (say for cigarette papers). The entire driving system comprising of all such sections must be continuously running, a proper web tension must be maintained and the angular velocity in each section must be regulable within the stipulated range.

Instead of going into the details of control in each section individually, a generalized scheme is described by block representation as in Fig. 11.26. Basically four different feedbacks are proposed in the scheme, (i) the current feedback for maintaining appropriate load current for the motor, (ii) angular speed feedback for appropriate speed control, (iii) position feedback for ensuring accuracy in the angular velocity in steady state condition, and (iv) web tension feedback for appropriate tension in the paper webs. Position pick-up gives pulses, the number per unit time of which being proportional to angular velocity; it is converted by D/A converter and fed back.

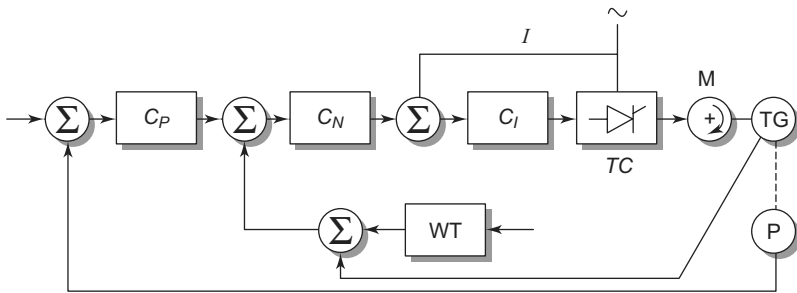


Fig. 11.26 Generalized scheme of control of sectional drives in a paper mill C_I : current regulator, C_N : speed regulator, C_P : power regulator. WT: web tension, I : current, M : motor, TG: tachogenerator, P : position pick up, TC: thyristor control

The finishing group of machines consists of (a) pre-winders for re-reeling the paper web, (b) calenders which are used for calendering and polishing, together called glazing, and (c) paper cutters.

Not all finishing side driving machines are continuously run, for example, re-reeler stops when the paper web is finished.

In finishing machines also, automatic paper web tension control is necessary. Besides, a controlled drive system for supercalender is required. Calendering is necessary for writing and printing papers. A supercalender consists of a number of cylinders placed one above another and the pressure of the top cylinder on the remaining ones is hydraulically controlled. The cylinders are alternately covered by metal and by rubber/fabric to make the paper glossy and smooth respectively.

Web tension can be measured by special transducers such as pressductors marketed by ASEA or inferred from an electromechanical system. The latter technique is briefly discussed here which is often used in reeling off the paper web from a drum.

The web tension has to remain constant and for that the power of the dc motor must be equal to the power needed to reel off the paper web. Thus, with armature emf and current denoted by e_a and i_a respectively and with

web tension force and linear web velocity denoted by F and v respectively, one gets

$$e_a i_a = k_1 F v \quad (11.2)$$

where k_1 is a constant. The torque to maintain the required tension, is

$$T_w = k_2 i_a \phi = Fd/2 \quad (11.3)$$

where ϕ = magnetic flux of the motor, d = drum diameter and k_2 = constant. If excitation ϕ is constant and so is paper web tension, F , the armature current should be proportional to drum diameter. Thus

$$i_a = k_3 d \quad (11.4)$$

where k_3 is a constant. Control through this equation is applicable to low power motors. With control of current directly in proportion to drum diameter (with paper web reeled on to it) the web tension can be kept constant.

There are alternative relations as well. For, if i_a is constant, then

$$\phi = k_4 d \quad (11.5)$$

where k_4 is another constant. Equation (11.5) means that motor flux has to be controlled. Also, for a dc motor emf e_a is constant and as

$$\phi = k_5 / \omega \quad (11.6)$$

where k_5 is a constant and ω is the angular speed, one can write, using Eq. (11.3),

$$i_a = k_6 \omega d \quad (11.7)$$

where k_6 is a constant. Such control is used in high speed systems.

The above simplified analysis alters greatly when losses are taken into consideration which have so far been ignored. The frictional torque T_f and the acceleration torque T_{ac} change the actual torque equation to

$$T_c = T_w - T_f \pm T_{ac} \quad (11.8)$$

Similarly the corrected armature current becomes

$$i_c = i_a - i_f \pm i_{ac} \quad (11.9)$$

Also, one has

$$T_f = k_7 \omega^\beta \quad (11.10)$$

and for constant tension

$$k_7 \omega^\beta = k_1 i_f \quad (11.11)$$

For excitation to remain constant, one has

$$i_f = k_8 \omega^\beta \quad (11.12)$$

The constant β has a value between 1 and 2. The acceleration torque T_{ac} is determined from

$$T_{ac} = J d\omega/dt = k_2 i_{ac} \phi \quad (11.13)$$

but $d\omega/dt = (1/d)dv/dt$, so that

$$T_{ac} = (J/d)dv/dt = K_2 i_{ac} \phi \quad (11.14)$$

The simplified form of J is

$$J = k_9 d^4 + k_{10} \quad (11.15)$$

Hence, for constant excitation, the armature current i_{ac} is

$$i_{ac} = (1/K_2)dv/dt(k_9 d^3 + k_{10}/d) \quad (11.16)$$

Thus Eqs (11.12) and (11.16) can be taken as correction terms for the earlier relation given by Eq. (11.4)

A typical scheme of drive system control, with paper web tension and position measurable, is shown in Fig. 11.27 where it is assumed that the motor field excitation is constant. Regulators C_1 , C_2 , and C_3 are for taking into account variations of the quantities as given by Eqs (11.5), (11.12) and (11.16). Controller C_N is the main controller.

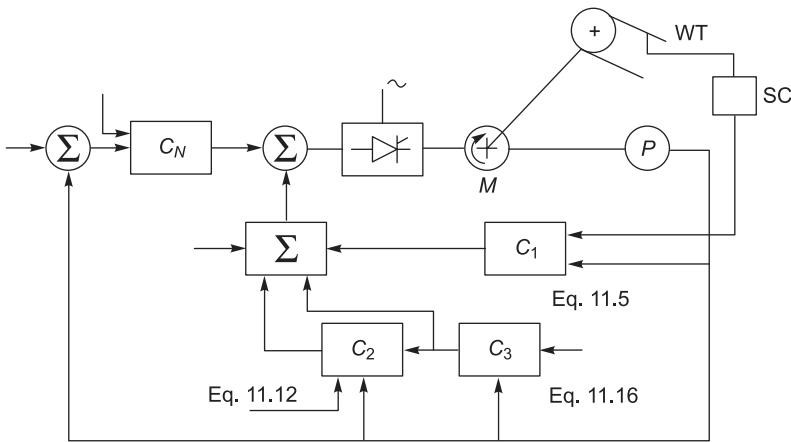


Fig. 11.27 A typical finishing side drive system control scheme, C_N : speed controller, WT: web tension, SC: Signal Conditioner. $C_1, 2, 3$: separate regulators, P: position pick up, M: motor

There are lots of variations in the design of the schemes and in recent times PC-based control schemes have appeared where all individual regulators have been dispensed with. Even the stock/pulp preparation schemes are designed around PLCs.

A very ingenious method of control of re-reeling finishing side machine is now described very briefly. A strip wound on a roller assumes the geometry of an equiangular spiral. The equation of such a spiral can thus be used for establishing a relation between roller rpm and roll diameter which can be used to control a dc motor by appropriate field change to keep the linear speed of the web constant, i.e., the web tension constant. If the nominal roll diameter is $2a$ and web thickness h , the instantaneous roll radius is

$$r = a \exp(m\theta) \quad (11.17)$$

where θ denotes the angular coverage by the strip over a certain small period of time and m is a parameter dependent on h and a . For one complete rotation, this become: $a \exp(2\pi m)$, so that the strip thickness is

$$h = a(\exp(2\pi m) - 1) \quad (11.18a)$$

and

$$m = (1/2\pi)\ln(1 + h/a) \quad (11.18b)$$

Also, since linear velocity $v = r\omega = r d\theta/dt$,

$$v = a \exp(m\theta) d\theta/dt \quad (11.19)$$

For constant v , $dv/dt = 0$, using which and writing $d\theta/dt = \omega$, one can obtain from Eq. (11.19) with appropriate mathematical manipulation and boundary conditions,

$$1/\omega = mt + 1/\omega_0 \quad (11.20)$$

which in terms of rpm N and nominal rpm N_0 becomes

$$N = N_0/(1 + kt) \quad (11.21)$$

where $k = mN_0/30 = AmN_0$.

For a separately excited dc motor the field current i_f is inversely proportional to rpm N , i.e.,

$$I_f = k_1/N \quad (11.22)$$

$$= (k_1/N_0)(1 + kt) = I_{f0} + \alpha t \quad (11.23)$$

From Eqs (11.19), (11.20) and (11.21) one obtains

$$v = AaN_0 \exp(m\theta)/(1 + AmN_0 t) \quad (11.24)$$

Also, from $d\theta = \omega dt$ and using Eq. (11.20), one gets

$$\theta = \ln(1 + \omega_0 mt)/m \quad (11.25)$$

Combining Eqs (11.24) and (11.25), now

$$v = AaN_0 \quad (11.26)$$

Thus with fixed base rpm and nominal roller diameter, the linear speed is constant. For this the field excitation is changed as per Eq. (11.23). The scheme is shown in Fig. 11.28. The armature current control is however to be done for making up frictional losses and acceleration torque compensation as indicated earlier.

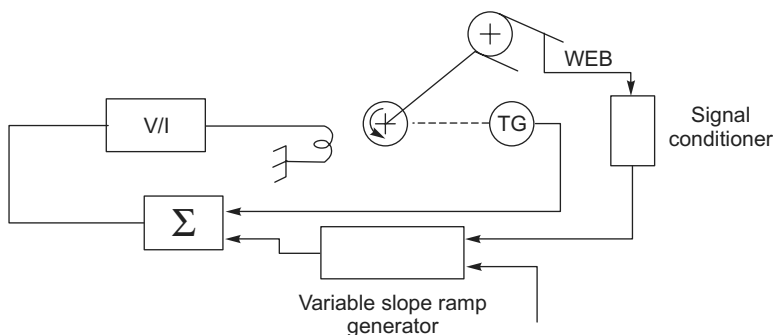


Fig. 11.28 A new simple scheme of constant linear speed control of strip like materials; V/I: Voltage/current converter, TG: tachogenerator

11.5 DISTILLATION COLUMN

A distillation column is a process which is both multivariable and nonlinear, sometimes distributed as well, making it exceedingly difficult to generalize about its dynamic behaviour. The control of such a system would consist of a number of interacting loops and is often realized through exhaustive simulation studies. For a given column, variety of control configuration exists. Determination of the structure of the control system, therefore, becomes the primary design task. However, for a given control system structure, determination of the control laws is also a major consideration.

Distillation process is widely used in chemical and petrochemical plants. It, basically, is a process of separating two or more mixed liquids by making use of different boiling points. When the liquid is heated and a vapour is produced, the vapour will show a higher ratio of molecules of the liquid with a lower boiling point to that of higher boiling point. When this vapour is condensed, it will produce a liquid which is richer in molecules of the low boiling point liquid. By repetition of this process, the enrichment can be increased obtaining finally a pure product. A three stage process is shown in Fig. 11.29 for understanding the principle of the basic process of distillation column.

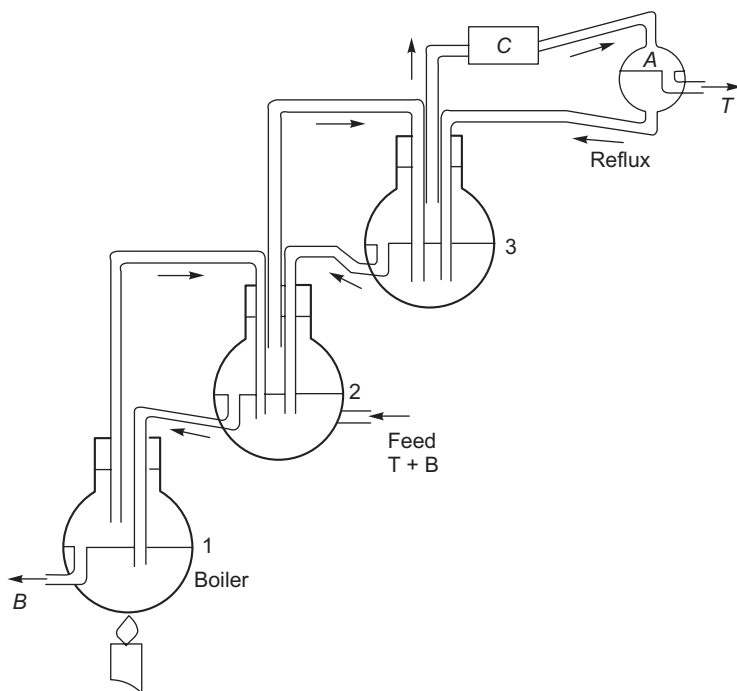


Fig. 11.29 Simplified scheme of a distillation column; C: condenser, A: accumulator, T: top product, B: bottom product

When the process starts, flasks 1 and 2 are filled to a certain level with mixed liquids, T of low boiling point and B of high boiling point. Flask 1 is heated, from which vapour enters flask 2 and condenses there as this flask is not heated. Excess liquid in this flask will return to flask 1 through an overflow pipe. If new feed is not given, the liquid in flask 2 will, in this way, grow richer in the low boiling component T , and in consequence its boiling point will also be lowered. Hot vapour from flask 1 heats up the contents of flask 2 and after a point of time this liquid would also start boiling. Its boiling point would be lower than the liquid in flask 1. The vapour produced by the boiling in flask 2 will be more enriched with component T . This vapour enters flask 3 where the same processing as in flask 2 occurs and the vapour from flask 3, which is highly enriched in component T , is condensed in a condenser, the output from which is almost the pure product T . The excess, here again, overflows back to flask 3. The control of reflux can optimize output quantity in relation to purity. Down below almost pure bottom product B is drawn out. Here, a boiler (called reboiler) is controlling the heating. If feed is made at a controlled constant stream of the mixed liquids, a continuity in the process can be maintained.

Separating action of distillation column follows from the equilibrium curve of Fig. 11.30. Starting with a liquid of composition x_0 , boiling it and collecting the vapour at equilibrium condition would give a vapour composition y_0 . If this vapour is condensed, the liquid that is formed has a composition $x_1 = y_0$. Boiling the liquid of this composition produces an equilibrium vapour with composition $y_1 > y_0$. The steps can be continued to any number of them to get a product of high degree of purity in the lighter component.

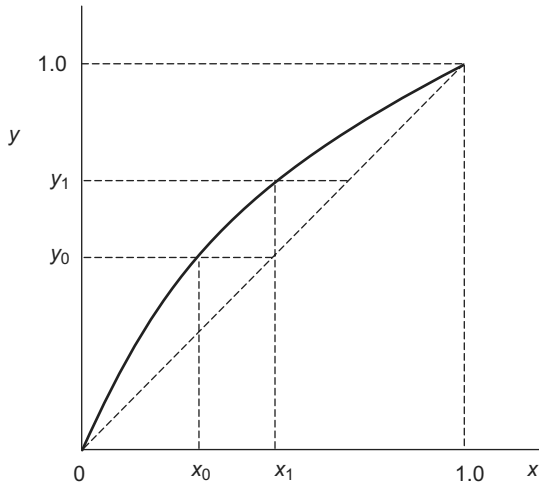


Fig. 11.30 The distillation principle explained

Distillation is a combined mass and heat transfer process. Typical mass transfer apparatus is designed specifically to effect the desired contact between the liquid and vapour phases during distillation. Two types are common, (i) packed towers which use a large number of small particles in the form of rings, spheres, saddles, etc. to pack the distillation tower in order to provide a very large surface area with the contact occurring on the surface of the particles and (iii) plate towers which use either bubble-cap trays or perforated plates so as to allow the vapour stream to bubble directly through the liquid.

In a binary distillation system the reflux rate at the top and the boil up rate at the bottom may be used to control compositions in the upper and lower halves of the column. The system is shown in Fig. 11.31 for analysis to show how separation is effected in such a system.

If now, F = feed rate, B = bottom product rate, T = top product rate, L = internal liquid rate, V = internal vapour rate, x_f = feed composition, x_t = top product composition, x_b = bottom product composition, and x, y are internal liquid and vapour compositions respectively, then L, V, x, y

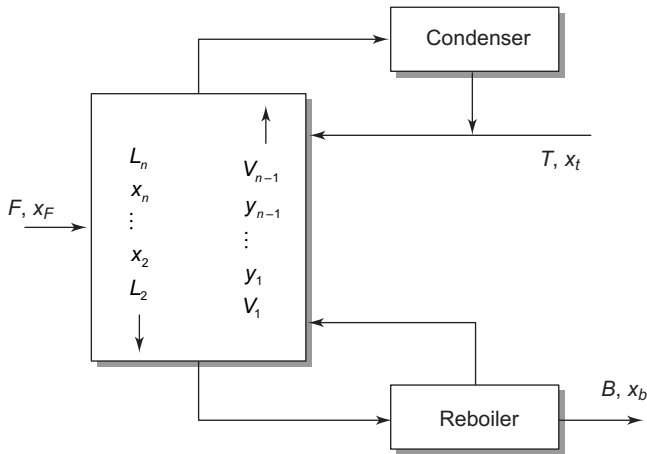


Fig. 11.31 Simplified block representation of the column; y, x : concentration, F : feed, T : top product, B : bottom product, L : liquid phase, V : vapour phase

are internal variables for which subscripts are used for plates increasing upwards in the column. The upper half of the column is known as the rectification section and the lower half, the stripping section.

In terms of rates and compositions the topmost and bottommost section mass balance equations are

$$V_{n-1}y_{n-1} = L_nx_n + Tx_t \quad (11.27)$$

and

$$L_2x_2 = V_1y_1 + Bx_b \quad (11.28)$$

Also,

$$V_{n-1} = L_n + T \quad (11.29)$$

and

$$L_2 = V_1 + B \quad (11.30)$$

Combining Eqs (11.27) and (11.29), for the topmost section

$$y_{n-1} = (L_n/V_{n-1})x_n + (1 - (L_n/V_{n-1}))x_t \quad (11.31)$$

and from Eqs (11.28) and (11.30) for the bottom section

$$y_1 = (L_2/V_1)x_2 + (1 - (L_2/V_1))x_b \quad (11.32)$$

For any section in the stripping side or the rectification side appropriate choice of suffix would yield the relations.

It is usually true that heat is conserved in column and there is no heat of mixing, the latent heat of vapourization is constant and the vapour leaving a tray is in equilibrium with the liquid on the tray. Then L_2/V_1 and L_n/V_{n-1} are constants, i.e., L_j/V_{j-1} is constant although $L_{j+1}/V_j > L_j/V_{j-1}$.

Now the overall system material balance yields

$$Fx_f = Tx_t + Bx_b \quad (11.33a)$$

and

$$F = T + B \quad (11.33b)$$

The control of distillation column, as already mentioned, is quite an intriguing problem and no specific unique set-out rules are known for it, particular strategy is to control a single product composition by conventional means for which regulation of secondary variables like pressure and temperature is proposed. The second one proposes to use a feedforward high speed loop for the inlet disturbances by which the reflux and vapour rates are adjusted. Also a composition check is made at the exit and a slow feedback loop is used to trim the final error. The third approach is to control compositions at both ends simultaneously, which is possible only with interactions between the loops. The interaction has been attempted to be avoided by designing a system through nodal analysis technique.

The main feature of the problem, however, remains and the alternatives galore through which control strategies are found out for specific cases. It is to be noted that the operation of a given binary distillation column is determined by many variables some of which are controllable/adjustable but some are not and are, therefore, considered as uncontrollable disturbances.

As top and bottom products control is also a strategy, it is worthy to note how these two products are related to feedrate and compositions of the above two products. From Eqs (11.33a) and (11.33b) one easily derives

$$T = ((x_f - x_b)/(x_t - x_b))F \quad (11.34)$$

and
$$B = ((x_t - x_f)/(x_t - x_b))F \quad (11.35)$$

11.5.1 Different Control Schemes

Without going into further elaboration of mathematical development, some important control strategies and schemes are now briefly presented.

(1) Constant Overhead Product Rate

The scheme is as shown in Fig. 11.32, where usually a kettle type reboiler is used. Product T is kept constant and any variation in F is absorbed in B , the change in which is accomplished by direct level control of the reboiler. In addition, if the steam-rate is fixed, the vapour rate V is approximately constant so that the liquids in the enriching and stripping sections, L_e and L_s , must increase. The material balance equations are

$$F = T + B$$

$$V = L_e + T, \text{ and } L_s = V + B = F + L_e$$

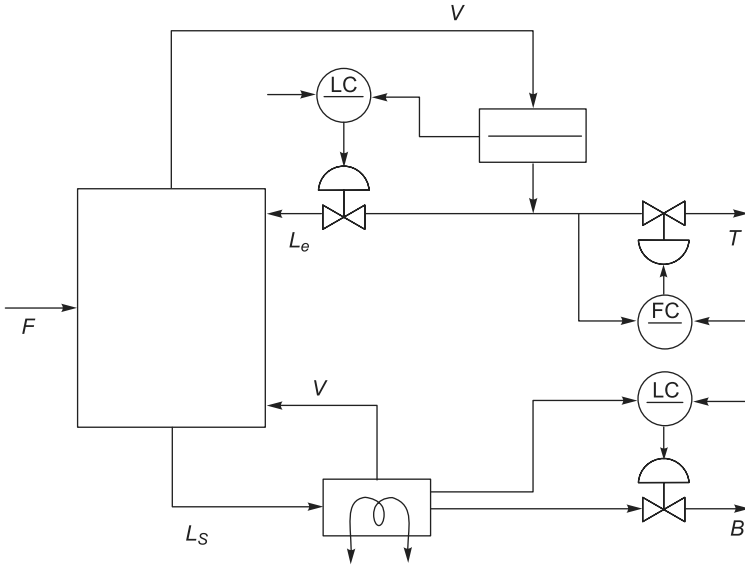


Fig. 11.32 Distillation column control scheme for constant top product

Since L_e increases for increase in F , V remaining nearly constant, top product quality, i.e., purity becomes better and since T also remains constant, the quantity of the lighter component in the top product becomes more. If the hold up in the accumulator is large, any attempt to adjust L_e by resetting T would respond slowly to bring L_e to the desired value.

(2) Constant Bottom Product Rate and Constant Reflux Rate

The scheme is shown in Fig. 11.33. As B is fixed here, feed rate fluctuation is absorbed in the change of T . Also L_e is fixed and L_s must change to accommodate feed rate change which implies a change in the vapour rate as well. Thus, if feed rate increases, vapour rate increases and as reflux rate is constant, the top product purity suffers and total production of lighter component in T becomes uncertain. This strategy is also poor in absorbing disturbances because of the lag in the steamside of the boiler. The dynamic lag of level control in reboiler would produce slow response to vapour rate adjustment via resetting of B .

(3) Constant Reflux Rate and Constant Vapour Rate

The scheme is shown in Fig. 11.34. It would be seen that as L_e is fixed and V is fixed, effectively T also is fixed and change in feed rate is absorbed in B . It should be stressed here that independent control of V and L_e is

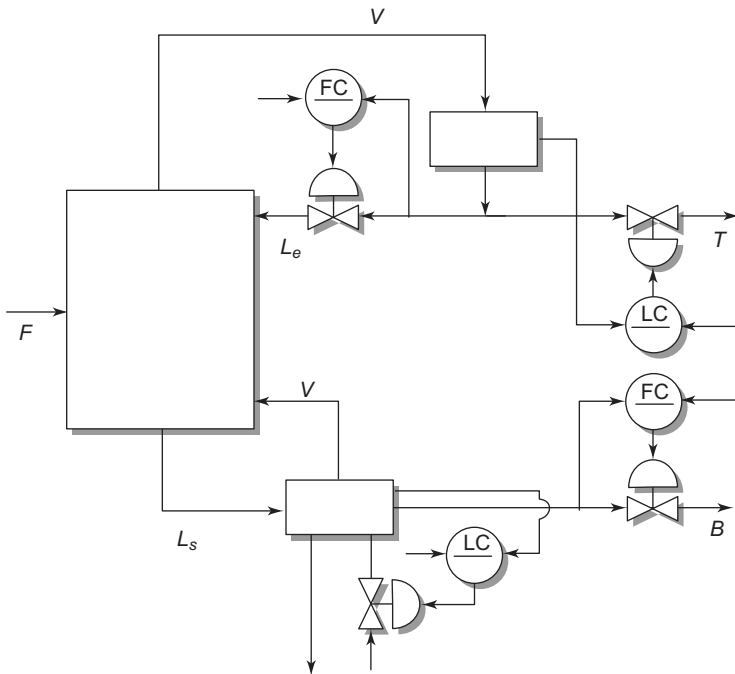


Fig 11.33 Distillation column control scheme for constant bottom product and reflux rate

made effective by cascade control of these two flow controllers with their set points being adjusted by appropriate composition sensors. Thus, this strategy leads to a scheme of composition control and has the ability to adjust the internal reflux in shortest possible time as the composition is a function of internal reflux ratios. In spite of large accumulator and reboiler hold-ups, the scheme works well because of composition-based controls.

Feed rate in all these schemes has been assumed to be varying within limits as the basic assumption in implementing the above strategies is that the feed rate, feed composition and feed enthalpy are relatively constant. If, however, there is a likelihood of a wide variation of feed rate, some form of feedforward control would be necessary and for wide change in feed composition, a scheme for switching the feed tray may have to be incorporated.

It is now largely clear that reflux feed loops as well as the vapour recycle loops are basically regenerative in nature and large hold-ups in the form of accumulator and reboiler in these loops tend to beget system instability and deteriorate the dynamic performance. Modification of the system has now been made by putting the accumulator hold-up outside the reflux loop and introducing vertical thermosyphon reboiler which, facilitates bottom hold-up to be isolated from the loop. This suggestion changes the schematic of Fig. 11.34 to that of Fig. 11.35.

11.6 BELT CONVEYOR CONTROL

Belt conveyor system is not a separate process like the ones already described but it has become a part and parcel of almost any industry and hence, a very brief discussion is included here for imparting a general idea to the readers about its operation and control.

Belt conveyors are used in various industries as (i) raw material carriers for process reactors such as coal carriers to feed boilers in thermal power plants, (ii) product carriers in the finishing side of process plants leading to stockyard or store, (iii) product carriers in assembly lines for facilitating production of assembled items in engineering industries, and, so on. A major use of belt conveyors in industry is in scales and weigh-feeders as included in item (i). Belt conveyor control would thus involve control of conveyor speed, weight-rate of flow of bulk materials it carries, counting and control of unit product items carried by the conveyor, control of filling cartons or cartels with process product transported by the conveyor, etc.

A typical conveyor weigh-feeder with flow rate control of materials is shown in Fig. 11.36(a). From the feeding hopper H through a fixed gate G , material is fed onto the belt conveyor. A weigh-platform bridge or weigh-bridge WB below the conveyor idlers is used to sense the weight of the materials passing over it and provide the weight signal WS . The speed sensor SS provides the speed signal and the two signals are multiplied in the multiplier M , the output of which is the rate output signal giving the total rate of flow per unit time. This signal is used for the controller C with a set point and the difference is sent to the SCR drive system to control the speed of the conveyor drive motors for belt conveyor speed control before the material leaves the end of the weigh platform to ensure maintenance of the flow rate as set. The fixed arrangement usually is manually set at a predetermined height for fixed depth of the feed material on the belt and this, to a large extent, keeps the material feed on the belt to be of constant weight per unit length as long as feed material density remains constant. Obviously an alternative arrangement, which is used only in very special cases, is to keep the belt conveyor speed constant and control the gate height through a weight controller WC and a motor M operating the gate. The measured variable for this controller is the weight signal WS from the weighbridge WB . The scheme is shown in Fig. 11.36(b).

Weigh-feeders and conveyors are now controlled using microprocessor technology. The power of the software as also the interfacing hardware system decide the functions to be allowed in such a scheme. The changing condition of weigh-feeding, scaling, solid flowmetering, fault diagnosis, speed proportioning, maximum and minimum set points, alarm, calibration, etc., are all taken care of by the software of such a system.

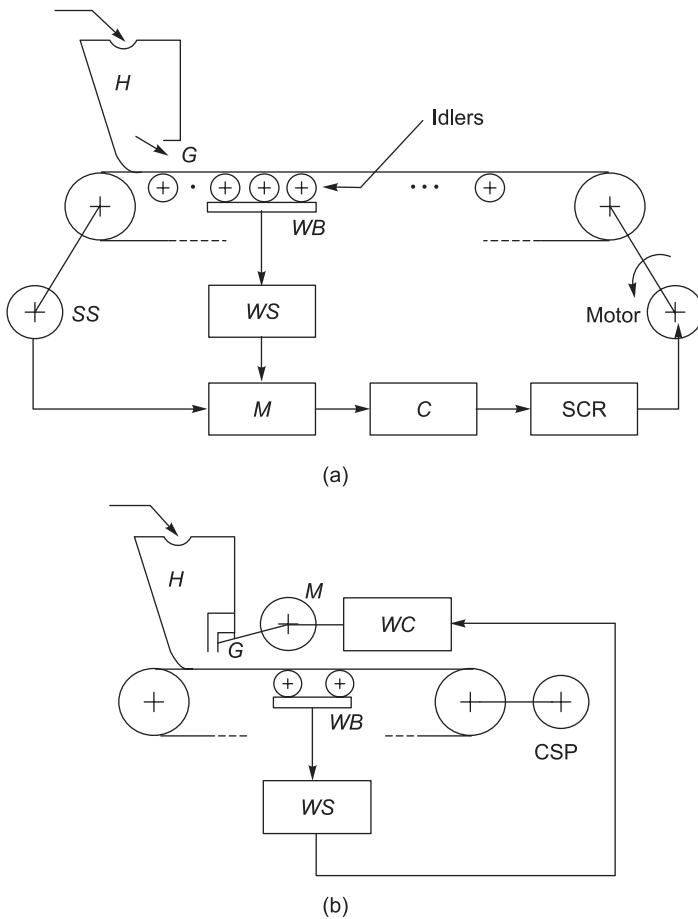


Fig. 11.36 (a) Belt conveyor speed control (b) Belt conveyor feed gate control, CSP: constant speed motor

As mentioned earlier, for the finishing side of a process plant, if it is bulk material, the conveyor belt carries the same and total weight to be deposited at a specified rate in a container for filling it is controlled almost similarly by the method already indicated. Additionally, container replacement, temporary stopping of the conveyor, etc., are all done automatically with microprocessor controlled system developed for the purpose.

In discrete product unit counting, filling, cutting to size, checking, etc., photoelectric sensors are used and through them a control system is developed where there arises the necessity of controlling the speed of the conveyor, its stopping and starting, etc. A typical case of such a system is illustrated schematically in Fig. 11.37(a) where carton filling is done with light beam interruption switching technique. The belt conveyor carries a number of identical cartons which allow light beam to pass through them

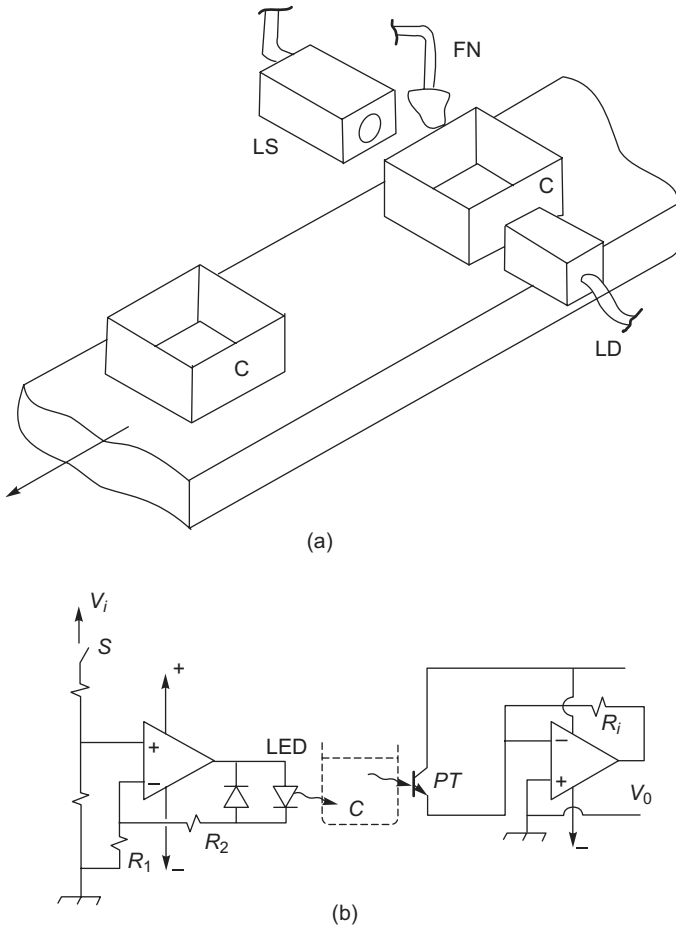


Fig. 11.37 (a) Carton filling control (b) The circuit scheme of (a)

when empty but not when full up to the desired level. The moving conveyor stops after a specified interval intermittently to bring an empty carton below the filling nozzle, FN , and across the same carton a photoelectric control system with a light source LS and a light detector LD installed. When the belt stops, the nozzle is positioned above the carton and opens to fill up the carton—this part of the control can be done electronically. After the carton is filled to the requisite level the light beam is interrupted and the light detector sends a converted electrical signal to stop the flow from the nozzle and remove it from position, simultaneously starting the conveyor which stops automatically via a timer control when the next carton in line is in position. Light source and light detector are traditionally incandescent lamp and photocell pair, but LED-phototransistor pair is also being used now. Figure 11.37(b) shows a scheme—the LED glows when switch S is closed receiving a signal from the conveyor movement condition. The

phototransistor remains irradiated after that till filling is complete—during this period only, an output V_0 is obtained, amplified and keeps the nozzle flowing via a solenoid. After that V_0 goes to zero, the solenoid closes and a gating circuit is used to start the conveyor.

There are wide variations in design and application in belt conveyor systems. Belt conveyors may be designed to carry loads of 0.5 kg/min at a speed of 0.3 m/min to 20,000 tonnes/hr at a speed of 300 m/min with its width varying from 0.3 m to 3 m.

11.7 pH CONTROL

Although pH control cannot be considered as a process control system in entirety, the control of pH assumes importance because of its widescale adoption in almost all chemical processes and because it has to have fundamentally nonlinear control strategy. Besides, the scale range of pH i.e., hydrogen-ion concentration from 100 to 10^{-14} moles/litre is tremendously large which would mean that reagent provider control valve had to have a rangeability $10^7:1$ for a set point of 7 pH. There are other stringent considerations which would be taken up later in the section.

Any component in a process control system may be nonlinear in nature—the process itself, or the actuator part, or the measurement part. Whichever it is, the superposition principle cannot be adopted and linear control strategy is also not acceptable. Often the method used is to linearize the system just around the control point. This is called perturbation principle. Other methods are sectional linearization and series linearization. If the system nonlinear equation is known with its parameters the equation may be solved either (1) by direct analysis which, however, is very difficult if not impossible, (2) by graphical procedure following phase-space technique—such techniques are considered in position control systems, (3) by numerical solution through digital computer. However, the analysis of nonlinear components like pH measuring electrodes pose problem. By definition,

$$\text{pH} = -\log [C] \quad (11.36)$$

where $[C]$ is hydrogen ion concentration in moles/litre.

11.7.1 The Nonlinearity Issue

The nonlinear nature of pH control is shown by the curve in Fig. 11.38 where the control is by adding a reagent after process stream pH is monitored which is to be maintained at a value. If this value is 5, control is simpler than if it is 8. In fact control above 5, becomes difficult because of the nonlinear nature and the gain associated with the process has to vary over a wide range. Analysis of the pH electrode reveals its nonlinear characteristic. There are two chemical processes involved in the electrode

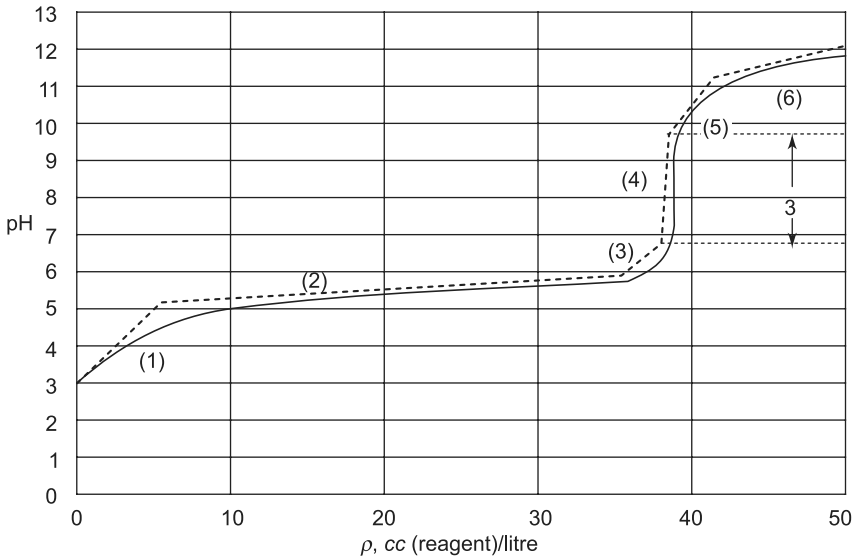


Fig. 11.38 pH-reagent curve

system as suggested by workers in the area—mixing in the cell with two electrodes, reference and measuring, and diffusion of ions in the measuring (glass) electrode. Figures 11.39 (a) and (b) represent the models.

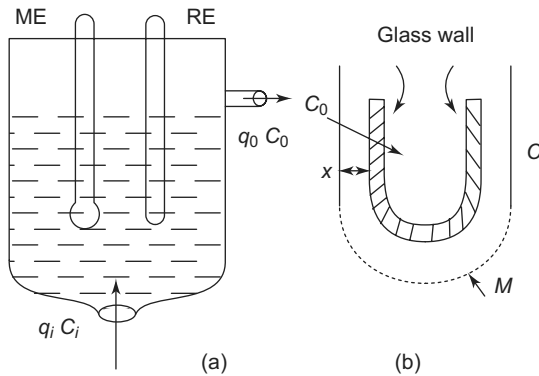


Fig. 11.39 (a) Mixing in the cell (b) Diffusion through film

In the mixing, if cell volume is V , then referring to the parameters marked in Fig. 11.39(a), for proper mixing, one gets

$$q_i C_i - q_0 C_0 = \frac{d}{dt} (V C_0) \quad (11.37)$$

where q = flowrate; usually $q = q_0 = q$ (say), C = hydrogen ion concentration. Equation (11.37) can be arranged as

$$C_0 = \frac{1}{1 + s\tau_1} C_i \quad (11.38)$$

where $\tau_1 = \frac{V}{q}$ which is the mixing time constant.

Diffusion of hydrogen ions through the boundary layer introduces a second time constant which may be calculated using the diffusion equation

$$\frac{dN}{dt} = K_d \alpha \frac{dC}{dx} \quad (11.39)$$

where N = Effective number of hydrogen ions diffusing in the process

α = Area through which diffusion occurs

K_d = Effective diffusion coefficient, and

C = Hydrogen ion concentration as indicated by the sensor.

If the concentration gradient across the film is linear and given by

$$\frac{dC}{dx} = \frac{C - C_0}{x} \quad (11.40a)$$

and if the capacitance of the electrode system for hydrogen ion is $\bar{C}$ given by

$$\bar{C} = \frac{dN}{dC} \quad (11.40b)$$

then one can write

$$\frac{dN}{dt} = \bar{C} \frac{dC}{dt} \quad (11.40c)$$

which combining with Eqs (11.39), (11.40a) gives

$$C = \frac{C_0}{1 + s\tau_2} \quad (11.40d)$$

where τ_2 is the diffusion time constant and is given by

$$\tau_2 = \frac{Cx}{K_d A} \quad (11.40e)$$

Combining Eqs (11.40d) and (11.38) one derives

$$C = \frac{1}{(1 + s\tau_1)(1 + s\tau_2)} C_0 \quad (11.41)$$

The mixing time constant τ_1 is small specially if volume is small and flowrate is high and it is predictable from the system consideration. However, diffusion time constant is relatively large and nonlinear as it depends on the cell configuration and flow conditions, direction of movement and amount of hydrogen ion concentration, the buffer material in type and amount and diffusion constant which, however, is relatively

constant for a given condition. Thus with $\tau_1 \ll \tau_2$ and τ_2 having nonlinear nature, Eqn. (11.41) is

$$C = \frac{1}{1 + s\tau_{2n}} C_i \quad (11.42)$$

Suffix n in τ_{2n} represents its nonlinear nature.

11.7.2 Linearized Control Strategy

The complex and nonlinear ρ -pH curve (Fig. 11.38) connotes apparent difficulty in controlling pH by addition of reagent which, however, is solved to an extent by piecewise approximate linearization of the curve and adapting the control process in ranges accordingly. The dotted linearized approximation is shown in Fig. 11.38 where six different segments are indicated. Over each linear segment, the linear control (ρ -pH relationship) is applicable. The control should be associated with appropriate stirring and mixing. The delay of pH sensing by the glass electrode or the electrode combination makes the situation still more complicated by increasing τ_{2n} and increasing the overall nonlinearity.

Looking at the segment 4 of the curve of Fig. 11.38 which is very steep, almost vertical, slight change in reagent flow or process flow and with simple control scheme large excursion is likely to occur. For such large slope change, the controller setting has to be carefully made and is usually kept considerably lower than the value obtained by linear analysis at the specified set points over this segment such that process gain can change. If pH is to be 8.5 and the pH range (segment length) is 3, for neutralization purpose a normal reagent flow is, say, 40 cc/litre and the reaction is in a small stirred tank, the slope at the control point would then correspond to a gain K_{cp} given by

$$K_{cp} = (8.5/3)/(1/40) = 113$$

Obviously such a high process gain would require a very low controller gain for reasons of stability.

It has been shown in Chapter 4 section 4.2.1, that Max error/Load change = $1.5K_l/(1 + K)$ indicating that peak error is 1.5 times the steady state error for proportional control. For an overall gain of the loop 10 (say) i.e. if the controller gain is 11.3, a 1% change in load would cause a peak error of

$$\begin{aligned} \text{Max error} &= [(1.5 \times 113)/(1 + 10)] \times 1\% \\ &= 15.4\% \end{aligned}$$

in the pH value.

Stability is improved to an extent by increasing tank size when large capacity of the tank dampens the pH fluctuation and hence higher controller

gain can be used. This technique, however, increases mixing time and then cascade control has to be called upon for improved performance. Linear control schemes adding dry chemical, gas feeding and liquid reagent are shown in Figs 11.40(a), (b) and (c) respectively.

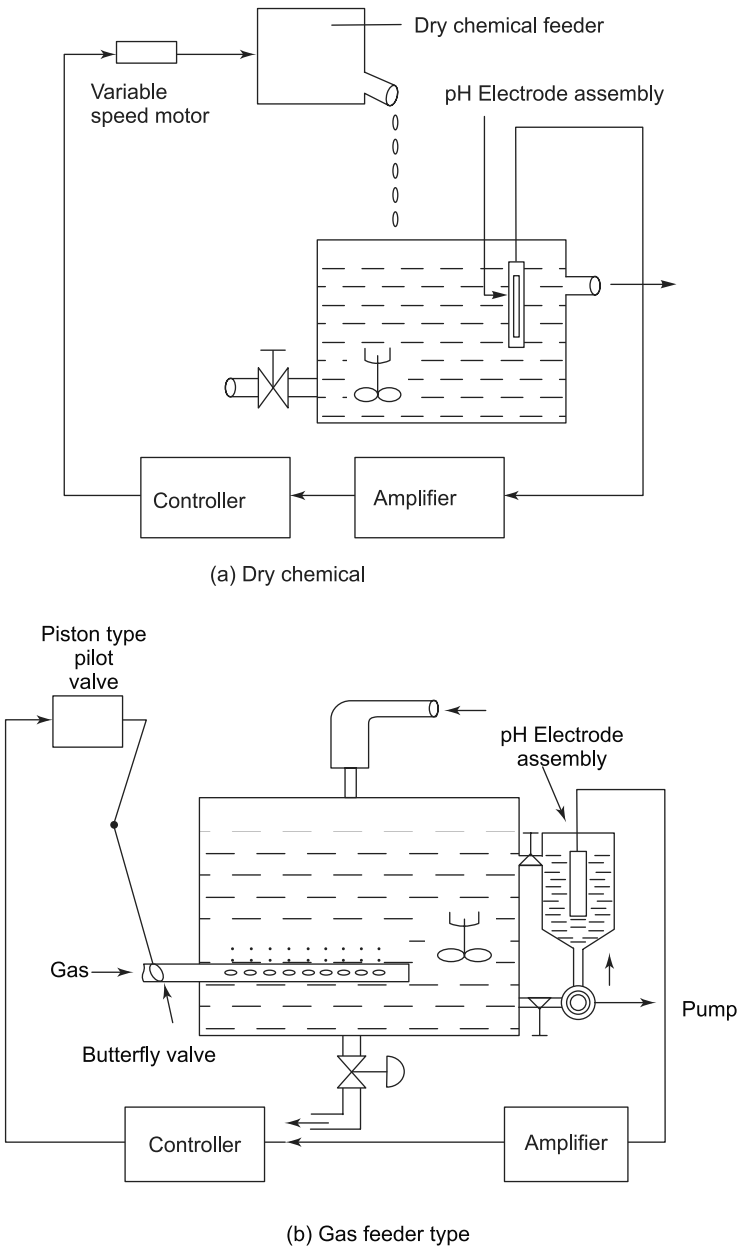


Fig. 11.40 pH control: (a) Addition of dry chemical (b) Gas feeder type

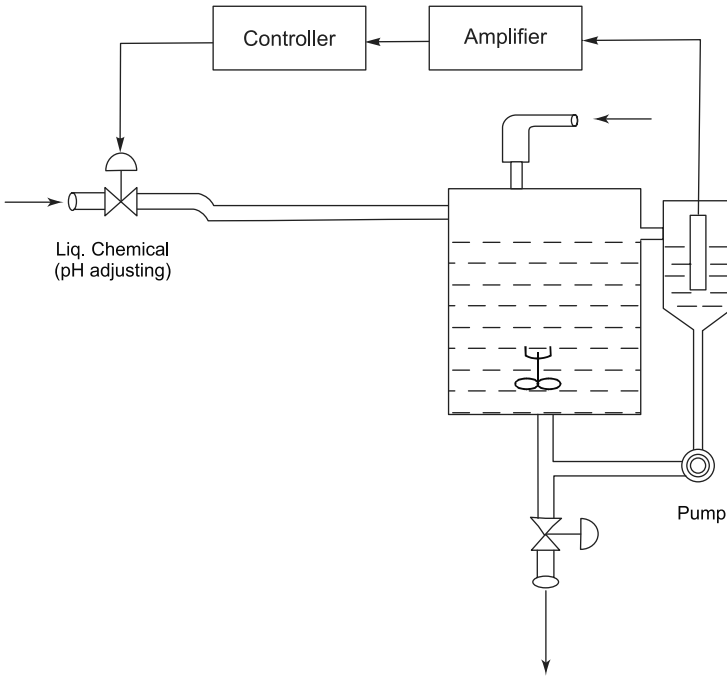


Fig. 11.40 (c) Liquid reagent

11.7.3 The Nonlinear Controller

The alternative approach is to use a nonlinear controller whose gain varies directly with deviation of the pH value and the controller function approximates a complementary function of the process. Even with approximation the improvement is reasonably good. The curve fitting (complementary) can be made by adjusting a nonlinearity parameter say η . The controller can be designed to have a nonlinear function between gain and deviation error amplitude. Also, I -action can be added for elimination of offset. The gain variation, however, is quite soft compared to dual mode (on-off) system which means accuracy is also less. With $f|e|$ as function of error magnitude with parametric representation in percent as

$$f|e| = \eta + \frac{(1-\eta)|e|}{100} \quad (11.43)$$

the continuous nonlinear controller equation is given as

$$m = \frac{100}{P} \left[\eta + \frac{(1-\eta)|e|}{100} \right] \left[e + \frac{1}{T_R} \int edt \right] \quad (11.44)$$

If $\eta = 1$, the controller becomes linear and $\eta \rightarrow 0$, the controller function tends to be square-law type (See Fig. 11.41). When $\eta = 0$, the controller becomes insensitive to small variation/error and hence an offset is likely to develop. For $\eta = 0.1$, the 'approximate' minimum gain is $10/P$. As is usual,

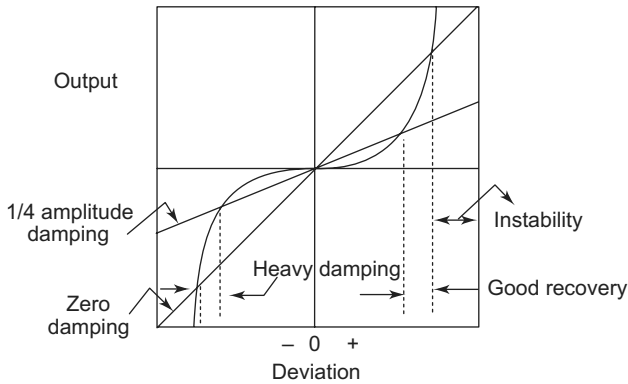


Fig. 11.41 The controller function indicated in output versus deviation

in linear control, the gain which produces constant amplitude oscillation — the state of the controller is said to have zero damping. Damping with decay ratio 1/4 is also linearly represented. The nonlinear functional curve superimposed on these straight lines intersect then and three distinct stability regions are produced as marked.

A nonlinear controller can be used in a linear process as shown in Fig. 11.42. It shows that small deviation persists only for one or two cycles and a large deviation provides more corrective action than is done by a linear controller.

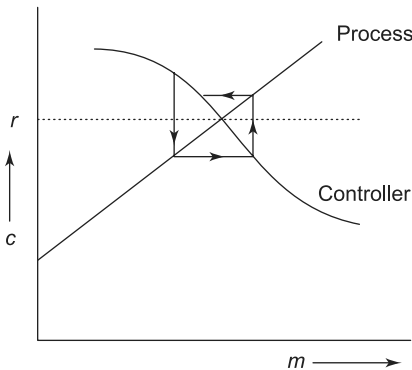


Fig. 11.42 Linearization approach

11.7.4 The pH-Control Loops

In pH control by neutralization of plant waste and sewage and sump outputs pose different problems such as

- (1) Effluent stream flowrate varies widely.
- (2) The stream may be acidic or basic at times requiring two different reagents.
- (3) The acidic and basic nature also vary widely covering several decades.
- (4) The acid or base may be weak or strong so that buffers may have to be provided.

If for reagent large flow-ranging is demanded linear control valve with a rangeability 25:1 is unsuitable. Equal percent valve characteristic would increase rangeability to 50:1. For flow rangeability of 200:1 two valves are used. Figure 11.43 (a) shows the single-valve set up while Fig. 11.43 (b) shows the two valve setup.

A small equal percent valve provides fine tuning while the linear valve takes care of the bulk load changes—the former is fed by a simple proportional controller while the latter linear valve takes care of offset changes with $P + I$ controllers. The characteristic of the equal percentage valve may be approximate as the complementary of the pH curve although not without large tolerance. In fact, this control action with the combination of two control valves is sort of compensatory and provides reasonably acceptable accurate control.

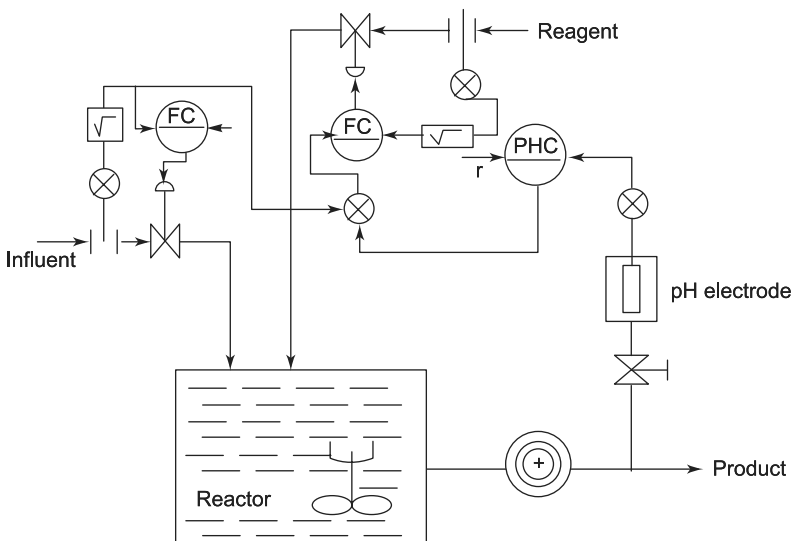


Fig. 11.43 (a) Single valve control

A feed-forward feed-back control can be used for pH control as shown in Fig. 11.44. Here

$$m = \log \beta F + \frac{100}{P}(r - pH) \quad (11.45)$$

where both feed-back (αF) and feed-forward $\left[\frac{100}{P}(r - pH) \right]$ actions are considered together. In Eq. (11.45), r = reference, feed-back factor is β and F is the measure of influent flow.

11.8 BATCH PROCESS CONTROL

Essentially there are two types of processes : Batch and Continuous. We have so far discussed the continuous processes and their control. A batch process produces products in batches unlike the continuous processes. Batch processes can have sequential operations in their completion. Often a term sequential process is used. Basically a batch process may be a subset of a sequential process.

Batch process control turns out to be programmable control which can provide stages of control operation such as in the case of firing process of thick film technology. Depending on the substrate, film and binder materials firing is done initially with slow ramping for about an hour say, then constant temperature firing for about one and a half hour and then again cooling by ramping down for 1 hr say. The operation is thus sequential as well as programmable. In recent times batch process controls are all automated using processor based programme controller or computer for single batch or distributed batches.

Batch processes are often parts of or units of complex sequential process or even short term continuous processes. Such processes are extensively used in pharmaceutical plants, food processing plants, wood pulping in paper plants and such other multiple sequence, multiple product chemical batching processes.

Like all process control systems batch process control is also guided by specification and design. Specification consists of functional details regarding batches, intervals, provision of check by alarm or such other monitoring equipment, constraints, implementation requirements. However, of these, functional details are of major interest. This specification may be stated in terms of number of levels which the batch process would have. This also considers the sequence of implementation guided by a set of rules.

Since batch process is also processor-controlled, the design would choose the logistics as best as it suits for a specific process. For safety one has to take care of the material being charged in quality and quantity, operation

of inputting devices like valve, rates of reaction—the mechanism of which needs be under surveillance.

A typical common batch process is annealing where the batch materials are pushed from behind the plant and final product are delivered at the front. The temperature is required to be raised up at a predetermined slope for a certain specified period till the soaking temperature is reached at which value it is maintained for a precalculated period of time. Often the cooling is also done at a precalculated slope. Depending on the quality and quantity of the material the programme may be re-done. The programmable controller does the job with the logic available to the operator. A typical centralized control system may be configured as given in Fig. 11.45(a) whereas a small batch process is represented in Fig. 11.45(b).

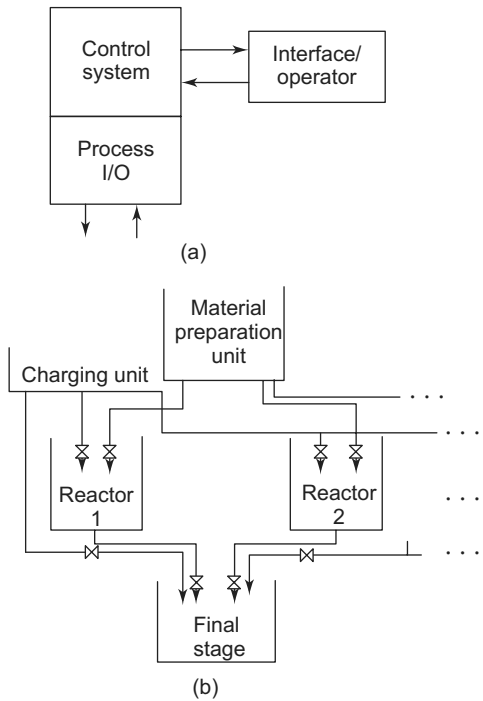


Fig. 11.45 (a) Centralized control scheme (b) Smaller distributed type

All the more modern control systems like hierarchical distributed control, fieldbus system in addition to PLC-aided support system are used in batch processes. For this adequate sequencing is necessary in the design stage. Another important aspect is redundancy. A typical example of redundancy is shown in Fig. 11.46.

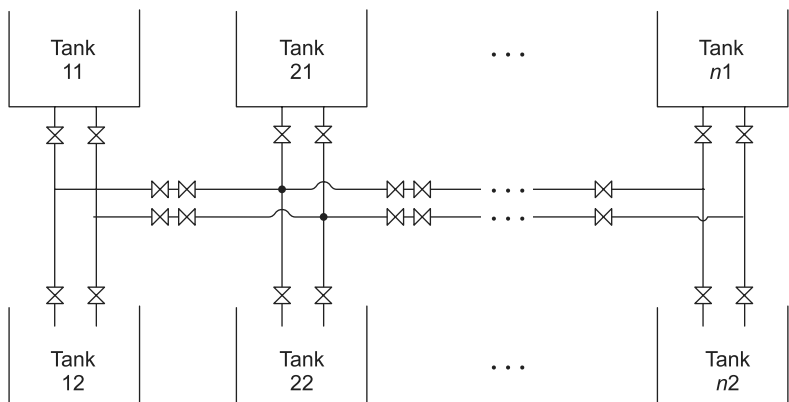


Fig. 11.46 Redundancy indicated

Since the control is not continuous type it must not be taken for granted that the system is not affected by dead time or, equivalently its dynamics is of some concern. A typical example is the air drying by spraying when inlet-outlet ‘dead time’ becomes of consequence in adequate control. For this feedforward approach is adopted to control the product to be dried as shown in Fig. 11.47.

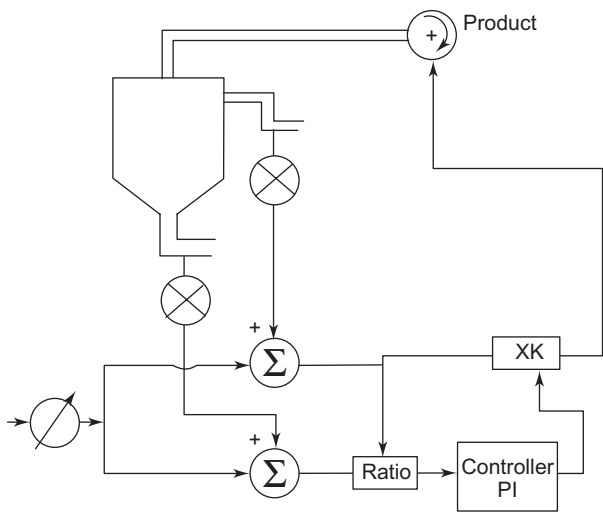


Fig. 11.47 Feedforward approach

All in all, the control schemes used in continuous control can be adopted in batch process control with the modification as demanded by the control strategy. A typical scheme of distributed batch controllers in a network is shown in Fig. 11.48.

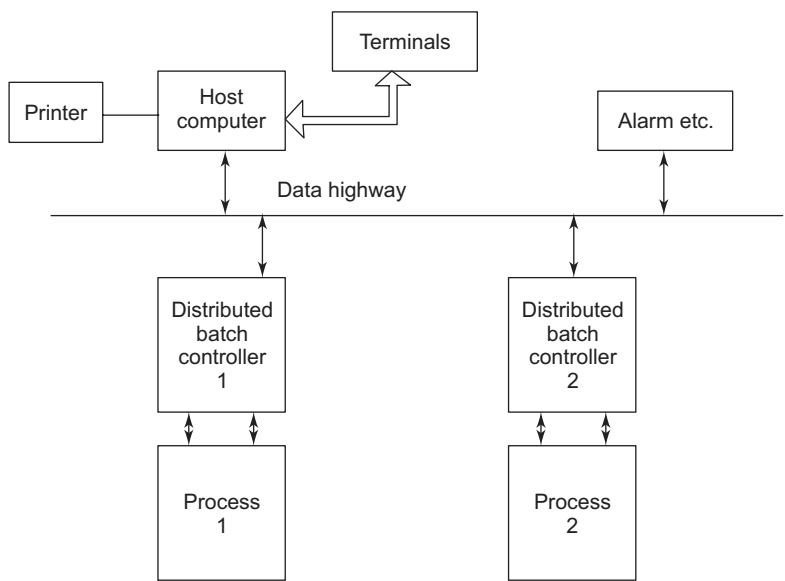


Fig. 11.48 Distributed type control with data highway

A batch reactor with direct digital control is shown in Fig. 11.49 where the computer controls the coolant by comparing the inlet and outlet temperature of the reactor.

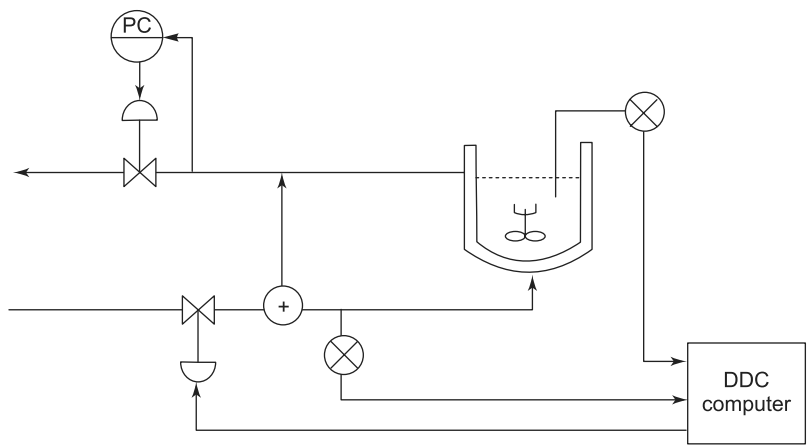


Fig. 11.49 DDC scheme in batch process

Review Questions

1. Explain with diagrams the different commonly adopted techniques of boiler control. In what situations are these control schemes used?
2. What are oxygen trimming and carbon monoxide trimming? Explain the efficacy of the two trimming techniques used separately as well as in combination.
3. What strategy is used for boiler furnace safety in combustion control systems? Explain with detailed logical analysis and diagrams.
4. How superheating of steam is done and how is it controlled? Describe a typical superheater control scheme discussing the constraints.
5. How is soaking pit furnace control different from boiler furnace control? Discuss a typical soaking pit control scheme where two types of fuels are used. How air-supply is controlled in such a case?
6. What are the parameters required to be controlled in paper pulp preparation? How are they controlled?
7. Why is paper mill drive system control so very intricate? How are the varying speeds in different mill sections controlled maintaining the requisite web tension?
8. Draw the control scheme of a distillation column for constant bottom product rate and constant reflux rate. What assumptions are made in developing such a scheme? What are its disadvantages?
9. Distillation column control faces some problems because of the accumulator and reboiler inertias appearing in the control loops. What are these problems and how are they solved?
10. Draw the scheme of a weigh-feeder speed control system. How is this used as a constant rate process feeder?
11. What different time constants play a role in the measurement and control of hydrogen ion concentration specially when the electrodes are considered?
How does pH of a solution change with addition of reagents to it? Draw the pH-reagent/litre curve.
12. How is control of pH done by piecewise linearization of the control curve, What may be the maximum error in such linearization?
13. How is nonlinear controller adopted in pH control schemes? Describe and discuss. Draw a pH-control loop with two-valve adoption. How is feedforward action provided for pH control?
14. Draw the scheme of feedforward approach as adapted in batch process control. Show with block diagram the DDC of batch reactors.

Appendix I

Table A-I Table of Z-Transforms

TIME FUNCTION	LAPLACE TRANSFORMS $L(s)$	Z-TRANSFORMS $Z(z)$
$\delta(t)$	1	1
$\delta(t - nT)$	$\exp(-nTs)$	z^{-n}
$u(t)$	$1/s$	$1/(1 - z^{-1})$
t	$1/s^2$	$Tz^{-1}/(1 - z^{-1})^2$
t^{n-1}	$(n-1)!/s^n$	$\lim_{a \rightarrow 0} (-1)^{n-1} \frac{\partial^{n-1}}{\partial a^{n-1}} \left(\frac{1}{1 - \exp(-aT)z^{-1}} \right)$
$\exp(-at)$	$1/(s + a)$	$1/(1 - \exp(-aT)z^{-1})$
$\frac{\exp(-at) - \exp(-bt)}{b - a}$	$1/((s + a)(s + b))$	$\frac{1}{b - a} (1/(1 - \exp(-aT)z^{-1}) - 1/(1 - \exp(-bT)z^{-1}))$
$(1/a)(u(t) - \exp(-at))$	$1/(s(s + a))$	$\frac{(1 - \exp(-aT))z^{-1}}{a(1 - z^{-1})(1 - \exp(-aT)z^{-1})}$
$t/a - (1 - \exp(-at))/a^2$	$1/(s^2(s + a))$	$\frac{1}{a} \left(\frac{Tz^{-1}}{(1 - z^{-1})^2} - \frac{1 - \exp(-aT)z^{-1}}{a(1 - z^{-1})(1 - \exp(-aT)z^{-1})} \right)$
$\frac{(a - b)u(t)}{a^2} + bt/a$ $+ (b - a)\exp(-at)/a^2$	$\frac{(s + b)/}{(s^2(s + a))}$	$\frac{bTz^{-1}}{a(1 - z^{-1})^2} + \frac{(a - b)(1 - \exp(-aT))z^{-1}}{a^2(1 - z^{-1})(1 - \exp(-aT)z^{-1})}$

(Contd)

Table A-1 (Contd)

$\frac{u(t)}{ab} + \frac{1}{ab(a-b)} \times$ $(b \exp(-at)/a \exp(-bt))$	$\frac{1/(s(s+a))}{(s+b))}$	$\frac{1}{ab} \left(\frac{1}{1-z^{-1}} + \frac{b}{(a-b)(1-\exp(-aT)z^{-1})} \right.$ $\left. - \frac{a}{(a-b)(1-\exp(-bT)z^{-1})} \right)$
$t \exp(-at)$	$1/(s+a)^2$	$T \exp(-aT)z^{-1}/(1-\exp(-aT)z^{-1})^2$
$\frac{u(t)}{a^2} - \frac{(1+at)\exp(-at)}{a^2}$	$1/(s(s+a)^2)$	$\frac{1}{a^2} \left[\frac{1}{1-z^{-1}} - \frac{1}{1-\exp(-aT)z^{-1}} - \frac{aT \exp(-aT)z^{-1}}{(1-\exp(-aT)z^{-1})^2} \right]$
$\frac{t^2}{2a} - \frac{t}{a^2} + \frac{u(t)}{a^3}$ $-\exp(-at)/a^3$	$1/(s^3(s+a))$	$\frac{T^2 z^{-2}}{(1-z^{-3})^3} + \frac{(aT-2)Tz^{-1}}{2a(1-z^{-2})^2} + \frac{1}{a^2(1-z^{-1})} - \frac{1}{a^2(1-\exp(-aT)z^{-1})}$
$\sin at$	$a/(s^2+a^2)$	$z^{-1} \sin aT/(1-2z^{-1} \cos aT + z^{-2})$
$\cos at$	$s/(s^2+a^2)$	$(1-z^{-1} \cos aT)/(1-2z^{-1} \cos aT + z^{-2})$
$\exp(-aT) \sin bt/b$	$1/((s+a)^2 + b^2)$	$(z^{-1} \exp(-aT) \sin bT)/(b(1-2z^{-1} \exp(-aT) \cos bT + \exp(-2aT)z^{-2}))$
$\exp(-aT) \cos bt$	$(s+a)/((s+a)^2 + b^2)$	$(1-z^{-1} \exp(-aT) \cos bT)/(1-2z^{-1} \exp(-aT) \cos bT + \exp(-2aT)z^{-2}))$

Table A-2 Table of Modified Z-Transforms

LAPLACE TRANSFORMS		MODIFIED Z-TRANSFORMS
$L(s)$		$Z_m(s)$
1	0	
$\exp(-nTs)$	z^{-n-1+m}	
$1/s$	$z^{-1}/(1-z^{-1})$	
$1/s^2$	$mTz^{-1}/(1-z^{-1}) + Tz^{-2}/(1-z^{-1})^2$	
$(n-1)!/s^n$	$\lim_{a \rightarrow 0} (-1)^{n-1} \frac{\partial^{n-1}}{\partial a^{n-1}} \left(\frac{\exp(-amT)z^{-1}}{1-\exp(-amT)z^{-1}} \right)$	
$1/(s+a)$	$\exp(-amT)z^{-1}/(1-\exp(-aT)z^{-1})$	
$1/((s+a)(s+b))$	$\frac{z^{-1}}{b-a} \left(\frac{\exp(-amT)}{1-\exp(-aT)z^{-1}} - \frac{\exp(-bmT)}{1-\exp(-bT)z^{-1}} \right)$	
$1/(s(s+a))$	$(z^{-1}/a) \left(\frac{1}{1-z^{-1}} - \frac{\exp(-amT)}{1-\exp(-aT)z^{-1}} \right)$	
$1/(s^2(s+a))$	$(z^{-1}/a) [T/(1-z^{-1})^2 + (amT-1)/(a(1-z^{-1})) + e^{-amT}/(a(1-\exp(-aT)z^{-1}))]$	

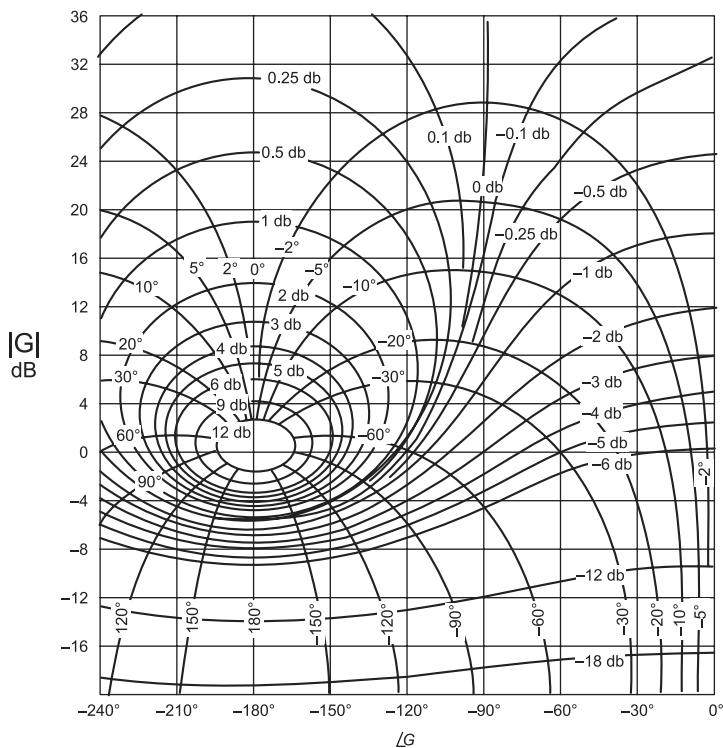
(Contd)

Table A-2 (Contd)

$(s + b)/(s^2(s + a))$	$\frac{z^{-1}}{a} \left[\frac{bTz^{-1}}{(1 - z^{-1})^2} + \left(bmT + 1 - \frac{b}{a} \right) / (1 - z^{-1}) + \frac{b - a}{a} \frac{\exp(-amT)}{1 - \exp(-aT)z^{-1}} \right]$
$1/(s(s + a)(s + b))$	$\frac{z^{-1}}{a} \left[\frac{1}{1 - z^{-1}} + \frac{b \exp(-amT)}{(a - b)(1 - \exp(-aT)z^{-1})} - \frac{a \exp(-bmT)}{(a - b)(1 - \exp(-bT)z^{-1})} \right]$
$1/(s + a)^2$	$(T \exp(-amT)z^{-1}(m + (1 - m)\exp(-aT)z^{-1})) / (1 - \exp(-aT)z^{-1})^2$
$1/(s(s + a)^2)$	$\frac{1}{a^2} \left[\frac{z^{-1}}{1 - z^{-1}} - \left(\frac{(1 + amT)z^{-1}}{1 - \exp(-aT)z^{-1}} + \frac{aT \exp(-aT)z^{-2}}{(1 - \exp(-aT)z^{-1})^2} \right) \exp(-amT) \right]$
$1/(s^3(s + a))$	$\frac{1}{a} \left[\frac{T^2 z^{-3}}{(1 - z^{-1})^3} + \frac{aT^2(2m + 1) - 2T}{2a(1 - z^{-2})^2} z^{-2} + \frac{(amT)^2 - 2amT + 2}{2a^2(1 - z^{-1})} z^{-1} - \frac{\exp(-amT)}{a^2(1 - \exp(-aT)z^{-1})} z^{-1} \right]$
$a/(s^2 + a^2)$	$\frac{z^{-1} \sin amT + z^{-2} \sin(1 - m)aT}{1 - 2z^{-1} \cos aT + z^{-2}}$
$s/(s^2 + a^2)$	$\frac{z^{-1} \cos amT - z^{-2} \cos(1 - m)aT}{1 - 2z^{-1} \cos aT + z^{-2}}$
$\frac{1}{(s + a^2) + b^2}$	$\frac{z^{-1} \exp(-amT)}{b} \frac{\sin bmT + z^{-1} e^{-aT} \sin(1 - m)bT}{1 - 2z^{-1} \exp(-aT) \cos bT + \exp(-2aT)z^{-2}}$
$\frac{s + a}{(s + a)^2 + b^2}$	$\frac{z^{-1} \exp(-amT)(\cos bmT + z^{-1} e^{-aT} \sin(1 - m)bT)}{1 - 2z^{-1} \exp(-aT) \cos bT + \exp(-2aT)z^{-2}}$

This is obviously not the complete table. Only the more common ones have been given. For other ones a standard book on sampled data control systems may be consulted.

Appendix II



NICHOLS CHART

Appendix III

REMEC'S METHOD

This method is another algorithm for the determination of the breakaway points on the root loci. The method closely resembles Routh's procedure. For the characteristic equation

$$P(s) = \sum_{j=0}^n b_j s^{n-j} = 0$$

a second equation is obtained such that

$$P'(s) = \sum_{j=0}^{n-1} a_j s^{n-j-1} = 0$$

where, $a_j = b_j(n-j)$

The coefficients of these two polynomials are now arranged as

$$\begin{array}{cccccc} b_0 & b_1 & b_2 & \dots & b_{n-1} & b_n \\ a_0 & a_1 & a_2 & \dots & a_{n-1} & \end{array}$$

Next, the third row is obtained as in the case of Routh's array by cross multiplication and division. The 4th row is the 2nd row repeated. The orders of the 2nd, 3rd and 4th rows are the same. The 5th row is obtained like Routh's array from the 3rd and the 4th rows. Similarly the 6th row is obtained from the 4th and the 5th rows. The 7th row is the 5th row repeated. The orders of the 5th, 6th and 7th rows are same and one less

than the order of the 2nd or 3rd or 4th row. This procedure is continued till the array is completed with the last row having only one element.

If a row, obtained in the process, has all its elements zero, then there are multiple roots of $P(s)$ indicating breakaway points on the root loci. The breakaway point or the pair of multiple roots is actually obtained by solving the equation formed by the coefficients of the rows just preceding this null row.

Bibliography

- ASEA. Pamphlet AP 41–103E, *Drive Systems for Winders, Supercalenders and Coating Machines*, 1st edn., Vasteras, 1983.
- ASEA, Pamphlet AP 33–103E, *Sectional Drives for Paper Machines*, 1st edn., Vasteras, 1980.
- Astrom, K.J. and Wittenmark, B., Self-tuning Controllers Based on Pole-zero Placement, *Proc. IEE*, 1980, **127**, pt. D, 120–123.
- Athans, M. and P.L. Falb, *Optimal Control*, McGraw-Hill, New York, 1966.
- Auricoste, J.G., “Application of Digital computer in Process Control”, *Proc. IFAC I, Moscow* (1960), Butterworths, London, 1961, **4**, 282.
- Auslander, D.M., Y. Takahashi and M. Tomizuka, “Direct Digital Process Control: Practice and Algorithms for Microprocessor Application”, *Proc. IEEE*, 1978, **66**, 199.
- Bailey, S.J., “Direct Digital Control: the On-line Scene Today”, *Control Engineering*, 1977, **24**, 20.
- Bandyopadhyay, S., B. Dam and D. Patranabis, “Linear Speed Control of Striplike Materials with a Variable Slope Current Ramp Excitation”, *J. Instn. Electr. Telecomm. Engrs.*, 1992, **38**, 328.
- Bandyopadhyay, R. and D. Patranabis, An autotuning algorithm for PID controllers using dead beat format, *ISA Transactions*, Vol 40, pp 255–266, 2001
- Bandyopadhyay, R., S. Bandyopadhyay and D. Patranabis, “A self-tuning Controller Based on PI Algorithm”, *J. Instn. Electr. Telecomm. Engrs.*, 1993, **39**, 247.

- Bandyopadhyay, R., U.K. Chakraborty and D. Patranabis, Autotuning a PID controller: A fuzzy-genetic approach, *Journal of System Architecture*, Vol 47, pp 663–673, 2001.
- Behera, L. and I. Kar, *Intelligent Systems and Control: Principles and Applications*. Oxford University Press, New Delhi 2009.
- Bibbero, R.J., *Microprocessors in Instrumentation and Control*, John Wiley, New York, 1977.
- Bristol, E.H., G.F. Inaloglu and J.F. Steadman, “Adaptive Process Control by Pattern Recognition”, *Instrument and Control Systems*, 1970, **43**, 101.
- Bristol, E.H., “A Simple Adaptive System for Industrial Control”, *Instrument and Control Systems*, 1967, **40**, 70.
- Bristol, E.H., “Designing and Programming Control Algorithms for DDC Systems”, *Control Engineering*; 1977, **24**, 24.
- Bristol, E.H., “On a New Measure of Interaction for Multivariable Process Control”, *IEEE Trans. Auto. Contr*, 1966, **AC-11**, 133.
- Buchting, M., “Evolution of Siemens Boiler Control”, *Siemens Review*, 1956, **23**, 123.
- Buckley, P.S., “Automatic Control of Process with Dead Time”, *Proc. IFAC I, Moscow* (1960), Butterworths, London, 1961, **1**, 33.
- Caldwell, W.I., G.A. Coon and L.M. Ross, *Frequency Response for Process Control*, McGraw-Hill, New York, 1959.
- Carr, J.J., *Designing Microprocessor-based Instrumentation*, Prentice Hall, 1982.
- Carslaw, H.S. and J.C. Jaeger, *Conduction of Heat in Solids*, Oxford University Press, New York, 1959.
- Ceaglaske, N.H., *Automatic Process Control for Chemical Engineers*, John Wiley, New York, 1956.
- Chakravorty, P.K., “Special Control Loops”, *Private Communication*, 1978.
- Chatterjee, H.K., “Multivariable Process Control”, *Proc. IFAC I, Moscow* (1960), Butterworths, London, 1961, **1**, 132.
- Chien, K.L., E.I. Ergin, C. Ling and A. Lee, “Dynamic Analysis of a Boiler”, *Trans-ASME*, 1958, **80**, 1809.
- Clarke, D.W. and P.J. Gawthrop, “Self-tuning Controller”, *Proc. IEE*, 1975, and 1979, **122** and **126**, 929 and 633.
- Cohen, G.H. and G.A. Coon, “Theoretical considerations of retarded control”, *Trans-ASME*, 1953, **75**, 827.

- Considine, D.M. (Ed.) *Process Instruments and Control Handbook*, McGraw-Hill, New York, 1987.
- Coughanour, D.R. and L.B. Koppel, *Process Systems Analysis and Control*, McGraw-Hill, New York, 1965.
- Cox, J.B., L.J. Heliums, T.J. Williams, R.S. Banks and O.J. Kirk (Jr.), "A Practical Spectrum of DDC Chemical Process Control Algorithms", *ISAJ*, 1966, **13**, 65.
- Coxon, W.F.(1) *Flow Measurement and Control*, (2) *Temperature Measurement and Control*, (3) *Automatic Process Control—An Introduction*, Hey wood, London, 1959, 1960, 1962.
- Dahlin, E.B., "Designing and Tuning Digital Controllers", *Instrum. Control Systems*, 1968, **41**, 77.
- Davies, O.L., *Design and Analysis of Industrial Experiments*, Oliver and Boyd, London, 1956.
- Davies, O.L., *Statistical Methods in Research and Production*, Oliver and Boyd, London, 1957.
- Davidov, N.I., "Development of Automatic Boiler Control System", *Proc IFAC I*, Moscow (1960), Butterworths, London, 1961, **4**, 373.
- Debolt, R.R. and R.E. Powell, "A Natural Three-mode Controller Algorithm for DDC". *ISAJ*, 1966, **13**, 43.
- Delancy, M.M., "A Chemical Company's Survey Gives Edge to Electronic Instrument", *Control Engineering*, 1961, **8**, 169.
- Derusso, P.M., R.J. Roy, and C.M. Close, *State Variable for Engineers*, John Wiley, New York, 1967.
- Deshpande, P.B. and R.H. Ash, *Elements of Computer Process Control*, ISA, Prentice-Hall. Englewood, N.J., 1983.
- Designing Control System Automatically, *European Report in Control Engineering*, 1962. **9**, 30
- Doebelin, E.O., *Measurement Systems: Application and Design*, McGraw-Hill, New York, 1966, 1990 (4th Edn).
- Dutta, B., A. Ray, S. Pal and D. Patranabis, A connectionist model for rainfall prediction, *Neural, Parallel and Scientific Computations*, Vol. 17, pp. 47–58, 2009.
- Eckman, D.P., "Frequency Response Analysis of a Flow Control System", *ISAJ*, 1954, **1**, 11.
- Eckman, D.P., *Automatic Process Control*, John Wiley, New York, 1958.
- Elonka, S.M. and A.R. Parsons, *Standard Instrumentation: Questions and Answers*, Vol. II, McGraw-Hill, New York, 1957.
- Endtz, J., M.L. Janssen and J.C. Vermeulen, *Plants and Process Dynamic Characteristics*, Butterworths, London, 1957.

- Fischer, A., "The Present State of Benson Boiler Control", *Siemens Review* 1959, **26**, 245.
- Ford, R.L., "Electrical analogues for heat exchangers", *Proc. JEE*, Pt. V., 1956, **B.103**, 65.
- Foxboro Company, *A Foxboro Boiler Control System—What it Does*, Pamphlets, IEP, CFS, FEPC, EPS
- Franks, R.G.F., *Modelling and Simulation for Chemical Engineers*, John Wiley, New York, 1972.
- Frensch, J., "Dynamic Behaviour of Steam Superheaters", *Proc. IFAC I*, Moscow, (1960), Butterworths. London, 1961, **4**, 334.
- Gallier, P.W. and R.E. Otto, Self tuning Computer Adapts DDC Algorithms *Instrum. Technology*, 1968, **15**, 65.
- Giusti, A.L., R.E. Otto and T.J. Williams, "Direct Digital Computer Control". *Control Engineering*, 1962, **9**, 104.
- Gollin, N.W.. "Cascade Control Systems", *Control Engineering*, 1956, **3**, 94.
- Gould, L.A., *Chemical Process Control—Theory and Applications*. Addison Wesley. Reading, 1969.
- Grabbe, E., S. Ramo and D.E. Wooldridge, *Handbook of Automation, Computation and Controls*, Vols. I to III, John Wiley, New York, 1958.
- Graeme, J.G., *Designing with Operational Amplifiers*, McGraw-Hill, New York, 1977.
- Hadley, W.A. and D.E. Longobardo, *Automatic Process Control*, Pitman, London, 1963.
- Hainsworth, B.D., V.V. Tivy and H.M. Paynter, "Dynamic Analysis of Heat Exchanger Control". *ISAJ*, 1957, **4**, 230.
- Hall, J., "Some Microprocessor Basics", *Instrum. Control System*, 1977, **50**, 25.
- Harriott, P., *Process Control*, McGraw-Hill, New York, 1964.
- Holzbock, W.G., *Instruments for Measurement and Control*, Van Nostrand Reinhold, New York, 1962.
- "Instrumentation flow-plan symbols", Tentative Recommended Practice, Instrument Society of America, 1950, ISA RP 51.
- Jury, E.I., Synthesis and Critical study of sampled data control systems, *Trans AIEE* 1956, **75**, Pt.2. 141.
- Jury, E.I., *Theory and Application of the Z-Transform Method*, John Wiley New York, 1964.

- Kaczmark, U., "Critical Comparison of Nuclear Reactors with Various Types of Controller", *Proc. IFAC I*, Moscow (1960), Butterworths, London, 1961, **4**, 398.
- Kavanagh, R.J., "The Application of Matrix Methods to Multivariable Control Systems", *J. Franklin Inst.* 1956, **262**, 349.
- Kestenbaum, A., R. Shinnar and F.E. Thau. "Design Concepts for Process Control", *Ind. Eng. Chem. Proc. Des. Dev.*, 1976, **15**, 2.
- Kuo, B.C., *Automatic Control Systems*, Prentice-Hall, New Delhi, 1967.
- Lacy, R. and B. McAdams, All in one network—a continuous system, Intech, pp. 12–16, Dec 2009.
- Landau, I.D., *Adaptive Control—the Model Reference Approach*, Marcel Dekker. New York, 1979.
- Langill, A.W. (Jr.) and D.A. Borses, "A Close Look at Two Microprocessor Control Applications", *Instrum. Control Systems*, 1977, **50**, 43.
- Ledgerwood, B.K., "Scale Model Tests Increase Odds of Full Scale Success". *Control Engineering*, 1959, **6**, 147.
- Lee, T.H., G.E. Adams and N.M. Gaines, *Computer Process Control: Modelling and Optimization*. John Wiley, New York, 1968.
- Leigh, J.R., *Applied Digital Control*. Prentice-Hall International, London, 1985.
- Lerner, A.Ya., "The Use of Self-adjusting Automatic Control Systems", *Proc. IFAC I*, Moscow (1960), Butterworths, London, 1961, **4**, 226.
- Liptak, B.G., *Instrument Engineers Handbook*, Vols I to III, Chilton Book Co., Philadelphia. 1969, 1970, 1995.
- Lupfer, D.E. and M.W. Oglesby, "Applying Dead Time Compensation for Linear Predictive Process Control", *ISAJ*. 1961, **8**, 53.
- Luyben, W.L., *Modelling, Simulation and Control for Chemical Engineers*, McGraw-Hill, New York, 1973.
- Mamzic, C.L., "Basic Multiloop Control Systems" *Bulletin, Moore Products*, Spring House, PA, 1960.
- Mamzic, C.L., "Improving the Dynamics of the Pneumatic Positioners". *ISAJ*, 1958, **5**, 38.
- Mamzic, C.L. "Basic Multiloop Control Systems", *ISAJ*, 1960, **7**, 63.
- Mandal, A.K., *Introduction to Control Engineering*, New Age International Publishers, New Delhi, 2006.
- Miller, J.G., "Introduction to Digital Control, Pt. I Characteristics, Pt. II Functional Considerations", *Instrum. Control Systems*, 1971, **44**, **97**, 138.

- Milsum, J.H., "Adaptive Control of Processes by an Economic Criterion", *Proc. IFAC I*, Moscow (1960), Butterworths. London 1961, **4**, 231.
- Mitchell, D.S. and C.R. Webb. "A Study of Interaction in a Multiloop Control System". *Proc. IFAC I*, Moscow (1960), Butterworths, London, 1961, **1**, 142.
- Mitra, A.K., "Notes on Boiler Control", *Private Communication*, 1978.
- Moise, J.C., "Pneumatic Transmission Lines", *ISAJ*, 1954, **1**, 35.
- Miller, J.A., A.M. Lopez, C.L. Smith and P.W. Murrill, "A Comparison of Controller Tuning Techniques", *Control Engineering*, 1967, **14**, 167.
- Moore, C.B., "The Inverse Derivative—A New Mode of Automatic Control", *Instruments*, 1949, **22**, 216.
- Moore, C.F., C.L. Smith and W. Murrill, "Improved Algorithms for Direct Digital Control", *Instrum. Control Systems*, 1970, **43**, 70.
- Murphy, G.J., *Basic Automatic Control Theory*, East West and Van Nostrand, 1969.
- Murrill, P.W., *Automatic Control of Processes*, International Text Book, Scranton, PA. 1967.
- Nagrath, I.J. and Gopal. M., *Control Systems Engineering*. Willey Eastern. New Delhi. 1982.
- Narendra, K.S. and R.V. Monopoli (Eds.). *Application of Adaptive Control*. Academic Press, New York, 1980.
- Newton, G.C., L.A. Gould and J. Kaiser, *Analytical Design of Linear Feedback Control*, John Willey, New York, 1957.
- Nichols, N.O., "The Linear Properties of Pneumatic Transmission Lines. *ISA Trans.*, 1962, **1**, 5.
- Oglesby, M.E. and D.E. Lupfer, "Feed Enthalpy Computer Control of Distillation Column" *Control Engineering*, 1962, **9**, 87.
- Oldenbourg, R.C. and H. Sartorius, *The Dynamics of Automatic Control*, ASME, New York, 1948.
- Parker, K.T., "Design of PID Controller by the Use of Optimal Linear Regulator Theory", *Proc. IEE*. 1972, **119**, 911.
- Patranabis, D., *Principles of Industrial Instrumentation*, Tata McGraw-Hill, New Delhi, 1976, 1996, 2010.
- Patranabis, D., "On Root Locus of the Second Order Systems", *App. Phys. Qrtrly.*, 1968, **9**, 101.
- Patranabis, D., "On the Optimum Gain Selection with First Derivative Action of Higher Order Linear Control Systems", *App. Phys. Qrtrly.*, 1969, **10**, 106.

- Patranabis, D. and K.B. Datta, "A note on the Limit Cycle Stability", *J. Instrn. Telecomm. Engrs.*, 1970, **16**, 277.
- Patranabis, D., "On the Application of the First Derivative Action to Increase the Critical Gain in a Linear Feedback Control System", *Int. J. Control*, 1972, **15**, 297.
- Patranabis, D and A.N. Dey, "Stability Condition of Boiler Type Process", *J. Instrn. Electron. Telecomm. Engrs.*, 1982, **28**, 293.
- Pearce, T.C., S.S. Schiffman, H.T. Nagle and J.W. Gardner, Handbook of M/C Olfaction, Wiley-VCH Weinheim, 2003.
- Phister, M, "A Comparison of Predictive and Exploratory Modes of Computer Control for Industrial Process", *Proc. IFAC I*, Moscow (1960), Butterworths, London, 1961, **4**, 241.
- Raffel, B. and B.H.L. Betlem, Advanced Practical Process Control, Springer-Verlag, Heidelberg, New York, 2006 (Indian Reprint)
- Reswick, J.B., "Disturbance Response Feedback—A New Control Concept", *Trans. ASME*. 1956, **78**, 153.
- Roberts, A.P., "The Optimum Principle", *Control*, 1963, **7**, 10.
- Rohman, C.P. and E.C. Grogan, "On the Dynamics of Transmission Lines", *Trans, ASME*, 1957, **79**, 853.
- Rosenbrock, H.H., "An Example of Optimal Adaptive Control", *J. Electron. Control*. 1964, **16**, 557.
- Resenbrock, H.H., "An Automatic Method for Finding the Greatest or Least Value of a function". *Computer J.*, 1960, **3**, 175.
- Rosensoff, H.P. and A. Ghosh, Batch Process Automation: Theory and Practice, Van Nostrand Reinhold, New York. 1987.
- Russo, P.M., C.C. Wang. P.K. Baltzer and J.A. Weisbecker, "Microprocessor in Consumer Products", *Proc. IEEE*, 1978, **66**, 131.
- Savas, E.A., *Computer Control of Industrial Processes*, McGraw-Hill, New York, 1965.
- Schick, K., "New Electrical Controller", *Siemens Review*, 1957, **24**, 5.
- Shah, M.A., "The Computer in a Control System", *Instrum. Control System*, 1971, **44**, 93.
- Shinsky, F.G., "Feedforward Control Applied", *ISAJ*, 1963, **10**, 61.
- Shinsky, F.G., *Process Control Systems: Application Design and Adjustment*, McGraw-Hill. New York, 1967.
- Singh, S., S. Lal and N. Kaistha, Case study of tubular reactor hot-spot temperature control for throughput maximization, *Ind. Eng. Chem. Res.* Vol. 47(19), pp. 7257–7263, 2008.

- Smith, C.L., *Digital Computer Process Control*, Intext Publishers, Scranton, PA, 1972.
- Sperling, W. and Peter Lutz, Enabling open control systems—An introduction to the OSACA system platform, *Robotics Manufacturing*, Vol 6—IS RAM '97, ASME Press, New York 1996.
- Stephanopoulos, G., *Chemical Process Control*, Prentice-Hall International, Englewood. Cliffs., N.J., 1984.
- Sundaresan, K.R. *et al.*, “Evaluating Parameters from Process Transients”, *Ind. Eng. Chem. Process. Des. Dev.*, 1978, **17**, 237.
- Szklarski, L., K. Jaraez and A. Horodecki, *Electric Drive Systems Dynamics, Selected Problems*. Elsevier. Amsterdam, 1990.
- Taylor Instruction Manual. No. 3E 211
- Tou, J.T., *Digital and Sampled Data Control Systems*, McGraw-Hill. New York. 1959.
- Tietze, U and Ch. Schenk. *Advanced Electric Circuits*. Springer Verlag. Berlin. 1978.
- Truxal, J.G.. *Automatic Feedback Control System Synthesis*, McGraw-Hill, New York, 1955.
- Välstar, J.E., “A Fresh Look at Selecting Control Valve Characteristics”. *Control Engineering*, 1959, **6**, 103.
- Webb, P.U., “Reducing Process Disturbances with Cascade Control”, *Control Engineering*, 1961, **8**, 63.
- Widemann, G., “New Pneumatic Controllers”, *Siemens Review*, 1956, **23**, 64.
- Weidmann, J.A. and W.J. Rowan, “Control Valve Plug Design”, *Trans. ASME*, 1956, **78**, 1367.
- Wilkinson, W.L. and W.D. Armstrong, *An Investigation of the Transient Response of a Distillation Column, in Plant and Process Dynamic Characteristics*, Butterworths, London. 1957, 56.
- Wills, D.M., “Tuning Maps for Three-mode Controllers”, *Control Engineering*, 1962, **9**, 104.
- Williams, T.J., “Computer Control in Continuous Processes”, *Proc. IEEE*, 1966, **54**, 1751.
- Woolverton, P.F. and P.W. Murrill, “An Evaluation of Four Ideas in Feedforward Control”, *Instrument Technology*, Jan. 1967, 35.
- Young, A.J., *An Introduction to Process Control System Design*, Longmans Green, New York, 1955.
- Ziegler, J.G. and N.B. Nichols, “Optimum Settings for Automatic controllers”, *Trans-ASME*, 1942, **64**, 759.
- Ziegler, J.G. “Electronic Flow Control”, *ISAJ* 1959, **6**, 58.

Index

A

Adaptive control 397
Anticipatory control 166
Antireset control 253
AO 387
API 387
ARMA model 52
ARMAX 52
ARX 52
ASS 389
Attenuation 430
Auctioneering control 251
Autotuning 414
Averaging control 329

B

Batch process 473
Bode plot 133, 238
Boolean equation 209
Box-Jenkins model 52
Bristol's method 256
Bumpless transfer 255

C

Canonical form 115
Capacity rating
 coefficient 274
Cascade control 233
Cavitation 284
Cohen and Coon 107, 175
Compensators 141
 Cascade 143
 Parallel 142
Computer control 337
Controllability 101
 ratio 106
Conversion factor 74

Conveyor 463
Corner plots 133
CO-trim 425
CRC 380
Critical flow factor 284
 C_v -factor 274
Cyclonome 300

D

Dahlin's algorithm 371
Damping factor 74
Darcy-Weisbach
 equation 289
DAS 386
Data acquisition 338
Data link 374
DC motor 35
DCS 372
DDC 341
Dead beat algorithm 371
Dead time 78, 105
Decay ratio 77, 109
Demand side capacity 149
Derivative action 12
Derivative overrun 197
Deviation reduction
 factor 103
Differential gap 151
Digital control 358
Discrete modelling 51
Distillation column 28, 455
Duration adjustable
 controller 154
Dynamic error 71

E

Enriching section 457
Expert system 413

F

FD fan 419
Feedforward control 242
Fieldbus 378, 474
Figure of merit 248
FIP 378
Flowchart 355
FRDC 436
Frequency Response 58, 73

G

Gain-bandwidth
 product 107
Gain margin 120, 133, 134
Glazing 449

H

HART 381
Heat Transfer 21
Hierarchy control 346
HIS 212
HMI 212, 376
Hold device 363
Hold time 25
Hydrafiner 444
Hydraulic gradient
 method 286
Hydrapulper 443

I

IAE 107
ID fan 419
Impulse lines 2
Impulse transfer
 function 363
Interaction factor 88
Integral action 12

Intrinsic stand-off
ratio 156

Inverse derivative
control 252

ISE 108

ISP 378

ITAE 108

J

Jordan canonical form 115

K

Kalman 116, 117

L

Ladder diagrams 209, 213

LAN 378

Laplace Transform 4

Limit control 248

LLPS 438

Load compensators 244,
245

Load reaction curve 105

Load static error 85

Logic stepper motors 299

M

Manchester coding 379

Masson rule 69

Material balance 23

MATLAB 199

Maximally flat 76

MEF 233

Miller integrator 191

MMI 212

Modbus 382

Modified Z-transform 362

MTS 389

MTU 385

Multivariable control 255

N

Nichols chart 238, 239

Nuclear reactor 31

O

Observability 114

OCS 386

Offset 83

On-off action 11

On-off control 149

OPC 391

Open connectivity 391

Optimum entity structure
404

OSI model 381

Oxygen trim 426

Override control 248

Overshoot 77

P

PAC 222

Pade approximation 181

Parseval's rule 247

pH 464

Phase margin 120, 134

Pneumatic actuator 266

PLC 206, 474

Poiseuille's equation 306

Positioners 270, 293

POSIX 388

Power cylinders 294

Process modelling 39, 41

Process reaction curve 11,
105

Process reaction rate 149

Profibus 379

Proportional action 12

band 12, 162

Proportional control

factor 103

PWM 161

R

Rangeability 277

Rate constant 26

Rate error 85

Ratio control 229

Rectification section 456

Remec's method 131

Remote terminal unit

(RTU) 376

Reset time 12

Reset wind-up 249, 253

Reynold's number 289, 307

Routh's algorithm 121

Routh-Hurwitz

criterion 121

Root locus 124

S

Shannon limit 371

SCADA 384

Scale modelling 55

Selector control 248

Self regulation 117

Self-tuning 409

Servomotor 298

Smith predictor 180

Soaking pit 435

SPC 345

Split load supply 249

Split range control 233

SPTC 436

State variables 111

Steepest ascent 352

Steel plant 433

Stepper motor 297

Stripping section 456

Structural approach 404

Subsidence ration 104

T

Taylor diffusion

model 333

TCP/IP 379

Thermal balance 24

Thermosyphon 459

Three element control 243

Time proportional

control 154

TPS 384

Transfer function 8

Transportation lag 11, 28,

105, 106

Triple modular redundancy

(TMR) 212

Turn down 278

V

Valve characteristics 266,
273, 275

Valve sizing and
selection 284

Variable reluctance
motor 298

Viscosity index 290

W

Warren-Ross method 142

Z

Ziegler Nichols 107, 174,
239, 415

Z-transform 13, 360