

Advanced Series in Electrical and Computer Engineering – Vol. 8

AN INTRODUCTION TO
CONTROL SYSTEMS

(Second Edition)

KEVIN WARWICK

AN INTRODUCTION TO CONTROL SYSTEMS

ADVANCED SERIES IN ELECTRICAL AND COMPUTER ENGINEERING

Editor: W. K. Chen

Published

- Volume 1: Broadband Matching: Theory and Implementations (Second edition)
by W. K. Chen
- Volume 2: Active Network Analysis
by W. K. Chen
- Volume 3: Linear Networks and Systems: Computer-Aided Solutions and Implementations (Second edition)
by W. K. Chen
- Volume 4: Introduction to Engineering Electromagnetic Fields
by K. Umashankar
- Volume 6: Introductory Signal Processing
by R. Priemer
- Volume 7: Diagnostic Measurements in LSI/VLSI Integrated Circuits Production
by A. Jakubowski, W. Marciniak, and H. M. Przewlocki
- Volume 9: Orthogonal Functions in Systems and Controls
by K. B. Datta and B. M. Mohan
- Volume 10: Introduction to High Power Pulse Technology
by S. T. Pai and Q. Zhang
- Volume 11: Systems and Control: An Introduction to Linear, Sampled and Non-linear Systems
by T. Dougherty

Forthcoming

- Volume 5: Applied Graph Theory: Graphs and Electrical Networks (Third edition)
by W. K. Chen
- Volume 8: An Introduction to Control Systems (Second edition)
by K. Warwick
Infrared Characterization for Microelectronics
by W. S. Lau
Identification, Estimation and Statistical Signal Processing
by Y.-F. Huang

Advanced Series in Electrical and Computer Engineering – Vol. 8

AN INTRODUCTION TO CONTROL SYSTEMS

(Second Edition)

KEVIN WARWICK

Department of Cybernetics, University of Reading

Published by

World Scientific Publishing Co Pte Ltd

P O Box 128, Farrer Road, Singapore 912805

USA office: Suite 1B, 1060 Main Street, River Edge, NJ 07661

UK office: 57 Shelton Street, Covent Garden, London WC2H 9HE

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library.

First printed in 1989 by Prentice Hall International (UK) Ltd.

— A division of Simon and Schuster International Group

AN INTRODUCTION TO CONTROL SYSTEMS

Copyright © 1996 by World Scientific Publishing Co. Pte. Ltd.

All rights reserved. This book, or parts thereof, may not be reproduced in any form or by any means, electronic or mechanical, including photocopying, recording or any information storage and retrieval system now known or to be invented, without written permission from the Publisher.

For photocopying of material in this volume, please pay a copying fee through the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, Massachusetts 01923, USA.

ISBN 981-02-1563-0

This page is intentionally left blank

To the Staff and Students of the University of Reading,
Cybernetics Department

Contents

Preface	xi
Preface to the Second Edition	xiii
About the author	xiv
1 Introduction	1
1.1 General introduction	1
1.2 A concise history of control	3
1.3 Open-loop control	5
1.4 Closed-loop control	6
1.5 Some examples of control systems	8
1.6 Definitions of standard terminology	10
1.7 Summary	12
Problems	12
Further reading	14
2 System representations	15
2.1 Introduction	15
2.2 Transfer functions	16
2.3 Block diagrams	25
2.4 Physical system realizations	32
2.5 Electromechanical devices	42
2.6 Summary	47
Problems	48
Further reading	54
3 System response using differential equations	56
3.1 Introduction	56
3.2 Steady-state response	57
3.3 Transient response	65
3.4 Errors	73
3.5 Servomechanisms	84

viii	<i>Contents</i>	
	3.6 Summary	87
	Problems	88
	Further reading	92
4	System stability and performance	93
4.1	Introduction	93
4.2	The Routh–Hurwitz criterion	95
4.3	Root loci construction	102
4.4	Application of root loci	115
4.5	Compensation based on a root locus	125
4.6	Summary	131
	Problems	131
	Further reading	136
5	Frequency response analysis	137
5.1	Introduction	137
5.2	The Bode diagram	139
5.3	Compensation using the Bode plot	149
5.4	The Nyquist stability criterion	159
5.5	Compensation using the Nyquist plot	170
5.6	The Nichols chart	177
5.7	Compensation using the Nichols chart	184
5.8	Summary	191
	Problems	192
	Further reading	195
6	State equations	197
6.1	Introduction	197
6.2	State equations from transfer functions	198
6.3	Solution to the state equations	201
6.4	The characteristic equation	205
6.5	Canonical forms	209
6.6	Decomposition	213
6.7	Summary	216
	Problems	216
	Further reading	220
7	State variable feedback	222
7.1	Introduction	222
7.2	Controllability and observability	223
7.3	Controllability and observability in transfer functions	229
7.4	State variable feedback	232
7.5	State estimation	238

7.6	Summary	243
	Problems	243
	Further reading	248
8	Digital control	249
8.1	Introduction	249
8.2	Sampled data systems	249
8.3	Stability analysis	253
8.4	Digital compensation	256
8.5	State variable forms	261
8.6	Pole placement design	264
8.7	Summary	269
	Problems	269
	Further reading	273
9	Computing methods	274
9.1	Introduction	274
9.2	System simulation	275
9.3	Computer control	279
9.4	Computer-aided design	298
9.5	Summary	302
	Problems	303
	Further reading	305
10	Three-term controllers	306
10.1	Introduction	306
10.2	Proportional control	307
10.3	Proportional + integral control	312
10.4	PID control	319
10.5	Summary	329
	Problems	330
	Further reading	334
	Appendices	335
1	Matrix algebra	335
2	Tables of common transforms	343
3	Nonlinear systems	346
4	Unity feedback loops	355
	Index	359

This page is intentionally left blank

Preface

Automatic control encroaches upon many of the defined engineering and mathematics university courses in terms of both theory and practice. In common with most engineering disciplines it is a topic which is constantly changing in order to keep pace with modern day application requirements. In the 1960s control engineers were predominantly concerned with analog design procedures, the analog computer being a fundamental building block. Things have changed considerably since that period.

Many new techniques have risen to prominence in the area of control systems during the last decade. Significantly, the increasing influence of microcomputers has led to a much greater need for an emphasis to be placed on digital aspects of control. Many control ideologies that were once thought to be unrealizable, except in a few extreme cases, because of hardware limitations, have now become everyday realities in practical industrial applications due to the transference of complex algorithms to a software base. With the advent of VLSI design in association with special-purpose machines, higher speed and greater memory capacity computers are becoming available, allowing further control schemes to be considered.

The undergraduate students of engineering taking an introductory control course as part of their overall degree must necessarily become aware of the developments in digital control which have now become standard practice in many industrial applications and are fundamental to any further degree courses or research in control. This book treats digital control as one of the basic ingredients to an introductory control course, including the more regularly found topics with equal weight.

Organization of the book is as follows. A general overview of the field called control systems is introduced in Chapter 1. This includes a brief history of control and points out many of the standard definitions and terms used. In Chapter 2 it is shown how mathematical models can be formed in order that they may be representative of a system. Basic handling rules for these models are given and the model's relationship with everyday engineering tools and apparatus is stressed.

Differential equations are examined in Chapter 3 and in Chapter 4 system stability and performance, important aspects of control systems design, are considered. These two chapters, in effect, detail the necessary mathematical tools used in the analysis of a system and give an indication as to how the system responds to a particular set of

conditions. Chapter 4 concentrates mainly on the Routh–Hurwitz criterion in terms of stability and root-locus techniques for performance by investigating the effect of parameter variations on the roots of the system characteristic equation.

Chapter 5 looks into the frequency response methods due to Bode and Nyquist. Details are given with emphasis placed on how to design compensation networks using the methods described. The same approach is then taken in the discussion based on the Nichols chart which is used in obtaining the system closed-loop frequency response. It is considered that Chapters 3, 4 and 5 cover what is now termed classical control techniques; the rest of the book is therefore devoted to more recently devised methods.

Design philosophies and system descriptions based on the state space are brought together in Chapters 6 and 7. The use and variety of state equations and their relationship to system transfer functions is found in Chapter 6, whereas Chapter 7 is devoted more to system control using state feedback, subject to this being possible. Realization of the state variables when these are not measurable or physically obtainable is also considered.

Digital control uses the z operator and is closely linked with computer control via sampled-data systems. Computers, however, are also used for control design in terms of system simulations and modeling. Chapters 8 and 9 include the important topics of digital and computer control, computer-aided control system design being highlighted. The final chapter looks in depth at the most widely used control scheme, the PID controller, and includes details of its implementation in both continuous and discrete time.

The book is intended to be used in relation to an introductory course on control systems, for which students need a mathematical background such that Laplace transforms, solutions to differential equations and complex numbers have been taught.

This text has been put together with the help of several people. I would particularly like to thank John Westcott, John Finch, Mohammed Farsi and Keith Godfrey for their technical appraisal, and numerous former students who have provided valuable feedback on the material presented. I would also like to acknowledge the assistance of Sukie Clarke and Karen Smart in preparing the text and correcting my terrible English. Finally, I am immeasurably grateful for the support I have received from my family during the writing of this book, including practical help from my daughter Madeline: I hope the reader concludes that it has been worth it.

Preface to the Second Edition

One of the reviewer's comments on the first edition was that the section on Analog Computers was rather dated. In the second edition, apart from a brief mention, this section has been completely removed, to be replaced by a completely new section on Fuzzy Logic Controllers, which is very much of interest both from an academic and industrial viewpoint.

I am delighted that World Scientific Publishing Co. have decided to publish this second edition, following the first edition selling out very quickly. My thanks go to Liz Lucas for all her help in putting together this edition, her husband Rodney for helping with the figures and Barbara Griffin at World Scientific for her patience.

About the Author

Kevin Warwick has been Professor of Cybernetics and Head of the Department of Cybernetics at the University of Reading, UK, since 1988. He previously held appointments at Oxford University, Imperial College, London, Newcastle University and Warwick University, as well as being employed by British Telecom for six years. Kevin is a Chartered Engineer, a Fellow of the IEE and an Honorary Editor of the IEE Proceedings on Control Theory and Applications.

Professor Warwick has received Doctor of Science degrees for his research work on Computer Control both from Imperial College London and the Czech Academy of Sciences, Prague. He has also published well over 200 papers in the broad areas of his research interests, including intelligent systems, neural networks and robotics. He has written or edited, a number of books, including *Virtual Reality in Engineering* and *Applied Artificial Intelligence*.

Introduction

1.1 General introduction

When first encountering the subject of control systems, it should in no way be considered that the topic is a completely new field with which one has had no prior contact whatsoever. In everyday life we encounter control systems in operation and actually perform many controlling actions ourselves. Take, for example, picking up and opening this book, which involves the movement of our hands in a coordinated fashion. This movement is made in response to signals arising due to information obtained from our senses – in this case sight in particular. It is worth mentioning however that the selection of this book, as opposed to other books written by different authors, is an example of the human being behaving as an intelligent control system exhibiting, in this case, signs of good taste and common sense – both of which are difficult properties to quantify and study.

Certain properties are common between many different fields, examples being distance, height, speed, flow rate, voltage, etc., and it is the mere fact of this commonality that has given rise to the field of control systems. In its most general sense a system can be virtually any part of life one cares to consider, although it is more usual for a system to be regarded as something to which the concepts of cause and effect apply. A control system can then be thought of as a system for which we manipulate the cause element in order to arrive at a more desirable effect (if at all possible). In terms of engineering it therefore follows that the study of control systems is multi-disciplinary and is applicable equally well in the fields of chemical, mechanical, electrical, electronic, marine, nuclear, etc., engineering.

Although lying within distinctly different fields, possibly different branches of engineering, systems often exhibit characteristics which are of a similar, if not identical, nature. This is usually witnessed in terms of a system's response to certain stimuli, and although the physical properties of the stimuli themselves can take a different form, the response itself can be characterized by essentially the same information, irrespective of the field in which it lies. Consider, for example, heating a pot of water: this exhibits the same type of exponential response as that witnessed if either a capacitor is charged up or a

2 Introduction

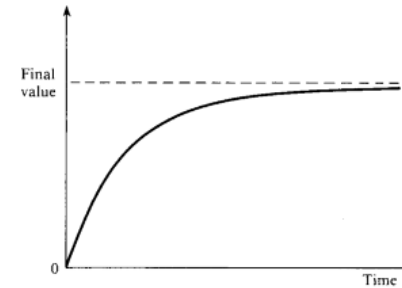


Fig. 1.1 Exponential response of a system

spring is compressed, see Fig. 1.1. The stimulus is, say, quantity of gas (if the water is gas heated), voltage and force respectively; however, the response, whether it be in terms of temperature, charge or stored energy, is still characterized as in Fig. 1.1. The study of control systems in terms of a multi-disciplinary basis therefore means that the results of performance and design tests in one discipline can readily be reposed and re-evaluated within an alternative discipline.

There are essentially two main features in the analysis of a control system. Firstly system modeling, which means expressing the physical system under examination in terms of a model, or models, which can be readily dealt with and understood. A model can often take the form of an appropriate mathematical description and must be satisfactory in the way in which it expresses the characteristics of the original system. Modeling a system is the means by which our picture of the system is taken from its own discipline into the common control systems arena. The second feature of control systems analysis is the design stage, in which a suitable control strategy is both selected and implemented in order to achieve a desired system performance. The system model, previously obtained, is therefore made great use of in the design stage.

The design of a control system only makes sense if we have some objective, in terms of system performance, which we are aiming to achieve. We are therefore trying to alter the present performance of a system in order to meet our objectives better, by means of an appropriate design, whether this is to be in terms of a modification to the system itself or in terms of a separate controller block. Actual performance objectives are very much dependent on the discipline in which the system exists, e.g. achieving a particular level of water or a speed of rotation, although the objectives can be stricter, e.g. requiring that a level of water does not vary by more than 2% from a nominal mean value.

An underlying theme in the study of automatic control systems is the assumption that any required controlling/corrective action is carried out automatically by means of actuators such as electromechanical devices. The concept of a human operative reading a value from a meter and applying what is deemed an appropriate response by pulling a lever or twisting a dial is therefore not really part of the subject matter covered in this

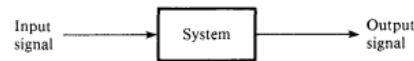


Fig. 1.2 Basic system schematic

book. The human being as an element in a control system is briefly considered later in this first chapter, merely to serve as an example. It is in fact becoming more often the case that a digital computer is employed to turn the measured (meter) value into an appropriate responsive action, the main advantage of such a technique being the speed with which a computer can evaluate a set of what could be complicated equations. It must be remembered however that interfacing is required, firstly to feed the measured value into the computer and secondly to convert the calculated response from the computer into a signal which is suitable for a physical actuator.

In terms of the system under control, an actuator is used to apply a signal as input to the system whereas a suitable measurement device is employed to witness the response in the form of an output from the system. Schematically, a system can therefore be depicted as a block which operates on an input signal in order to provide an output signal. The characteristics of that system are then contained within the block, as shown in Fig. 1.2.

1.2 A concise history of control

To put forward a chronologically ordered set of events in order to show the logical development of a technical subject is certain to be fraught with wrong conclusions and misleading evidence. The study of control systems is certainly no exception in this context, a particular problem being posed by much of the development of the mathematical tools which now form the basis of control systems analysis. It is also a difficult task to point the finger at someone way back in time and accuse them of starting the whole thing. Indeed some historical accounts refer to times well before the birth of Christ for relatively early examples.

As far as automatic control (no human element) is concerned, a historical example cited in many texts is James Watt's fly-ball governor. The exact date of the invention/development seems to vary, dependent on the particular text one is looking at, with dates ranging from 1767 to 1788 being quoted. Although the actual year itself is not particularly important, it serves a primary lesson in control systems analysis not to rely completely on any measured value. James Watt's fly-ball governor is shown in Fig. 1.3, where the control objective is to ensure that the speed of rotation is approximately constant. As the fly-balls rotate so they determine, via the valve, how much steam is supplied; the faster the rotation – the less steam is supplied. The rate of steam supplied then governs, via the piston and flywheel, the speed of rotation of the fly-balls. Although tight limits of operation, in terms of speed variation, can be obtained with such a device,

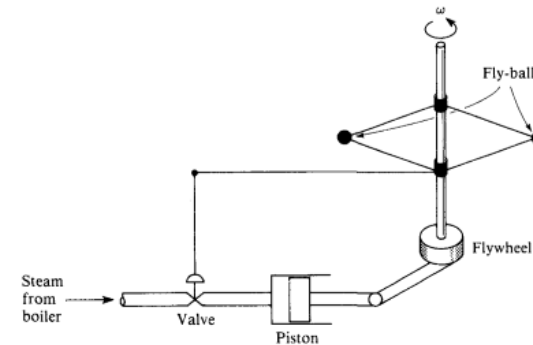


Fig. 1.3 James Watt's fly-ball governor

there are unfortunately several negative features, not the least of these being the tendency of the speed to oscillate about a mean (desired) speed value.

Around 1868, many years after Watt's governor, J. C. Maxwell developed a theoretical framework for such as governors by means of a differential equation analysis relating to performance of the overall system, thereby explaining in mathematical terms reasons for oscillations within the system. It was gradually found that Maxwell's governor equations were more widely applicable and could be used to describe phenomena in other systems, an example being piloting (steering) of ships. A common feature with the systems was the employment of *information feedback* in order to achieve a controlling action.

Rapid development in the field of automatic control took place in the 1920s and 30s when connections were drawn with the closely related areas of communications and electronics. In the first instance (1920s) this was due to analysts such as Heaviside who developed the use of mathematical tools, provided over a century before by Laplace and Fourier, in the study of communication systems, in particular forming links with the decibel as a logarithmic unit of measurement. In the early 1930s Harry Nyquist, a physicist who had studied noise extensively, turned his attention to the problem of stability in repeater amplifiers. He successfully tackled the problem by making use of standard function theory, thereby stressing the importance of the phase, as well as the gain, characteristics of the amplifier. In 1934 a paper appeared by Hazen entitled 'Theory of servomechanisms' and this appears to be the first use of the term 'servomechanism', which has become widely used as a name to describe many types of feedback control system.

The Second World War provided an ideal breeding ground for further developments in automatic control, particularly due to finance made available for improvements in the military domain. Examples of military projects of that time are radar tracking,

anti-aircraft gun control and autopilots for aircraft; each of which requires tight performance achievements in terms of both speed and accuracy of response.

Since that time mechanization in many industries, e.g. manufacturing, process and power generation, has provided a stimulus for more recent developments. Originally, frequency response techniques such as Bode's approach and Laplace transform methods were prominent, along with the root locus method proposed by Evans in the late 1950s. (Note: This was also known in the UK at one time as Westcott's method of π lines.) However, in the 1960s the influence of space flight was felt, with optimization and state-space techniques gaining in prominence. Digital control also became widespread due to computers, which were particularly relevant in the process control industries in which many variables must be measured and controlled, with a computer completing the feedback loop.

The 1970s saw further progress on the computer side with the introduction of microprocessors, thus allowing for the implementation of relatively complicated control techniques on one-off systems at low cost. The use of robot manipulators not simply for automating production lines but as intelligent workstations also considerably changed the requirements made of a control system in terms of speed and complexity. The need for high-speed control devices has in the 1980s been a contributing factor too and has made great use of hardware techniques, such as parallel processors, whereas at the same time ideas from the field of artificial intelligence have been employed in an attempt to cope with increased complexity needs. Finally the low cost and ease of availability of personal computers has meant that many control systems are designed and simulated from a software basis. Implementation, which may itself make use of a computer, is then only carried out when good control is assured.

1.3 Open-loop control

An open-loop control system is one in which the control input to the system is not affected in any way by the output of the system. It is also necessary however that the system itself is not varied in any way in response to the system output.

Such a definition indicates that open-loop systems are in general relatively simple and therefore often inexpensive. An excellent example is an automatic electric toaster in which the control is provided by means of a timer which dictates the length of time for which the bread will be toasted. The output from the toasting system is then the brownness or quality of the toast, the assumption being that once the timer has been set the operator can only wait to examine the end product.

Clearly the response of an open-loop system is dependent on the characteristics of the system itself in terms of the relationship between the system input and output signals. It is apparent therefore that if the system characteristics change at some time then both the response accuracy and repeatability can be severely impaired. In almost all cases however the open-loop system will present no problems insofar as stability is concerned,

6 Introduction

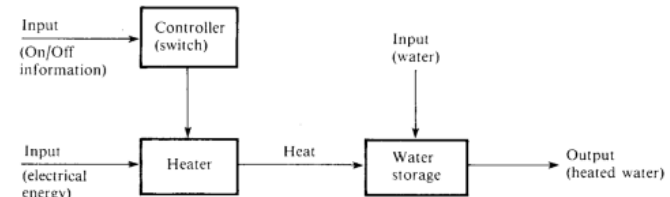


Fig. 1.4 Water heating device

i.e. if an input is applied the output will not shoot off to infinity – it is not much use as an open-loop system if this is the case.

Another example of an open-loop system is the water heater shown in Fig. 1.4, in which the controller is merely an on/off switch which determines when the heater is supplying heat in order to provide heated water at a certain specified temperature.

A problem with this open-loop system is that although today the water may be provided at a nice temperature at the output, tomorrow this might not be the case. Reasons for this could be a change in ambient conditions, a change in the temperature or amount of water input to the storage device or a drop in the voltage used to supply the heater. Merely leaving the system as it is and hoping for the best is clearly not satisfactory in the majority of cases. It would be much more sensible to measure the temperature of the heated water such that the on/off information can be varied appropriately in order to keep the output at approximately the temperature desired. We have now closed the loop, by providing feedback from the system output to an input, hence the system is no longer open-loop, but rather is closed-loop.

1.4 Closed-loop control

In a closed-loop system the control input is affected by the system output. By using output information to affect in some way the control input of the system, feedback is being applied to that system.

It is often the case that the signal fed back from the system output is compared with a reference input signal, the result of this comparison (the difference) then being used to obtain the control or actuating system input. Such a closed-loop system is shown in Fig. 1.5, where the error = reference input – system output.

Very often the reference input is directly related to the desired value of system output, and where this is a steady value with respect to time it is called a set point input.

By means of the negative feedback loop shown in Fig. 1.5 (negative because the system output is subtracted from the reference input) the accuracy of the system output in relation to a desired value can be much improved when compared to the response of an open-loop system. This is simply because the purpose of the controller will most likely be

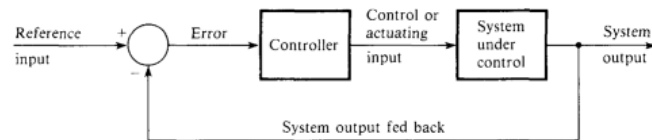


Fig. 1.5 Closed-loop system

to minimize the error between the actual system output and the desired (reference input) value.

A disadvantage of feedback is related to the fly-ball governor of Fig. 1.3, where oscillations can occur in the system output, the speed of rotation, which would not occur if the system were connected in open-loop mode. The oscillations are due to the attempt to get the error signal to as low a magnitude as possible, even if this means swinging the output first one way and then the other.

The open-loop heating system of Fig. 1.4 can be converted into a closed-loop system by measuring the temperature of the heated water (output) and feeding this measurement back such that it affects the controller by means of modifying the switch on/off information. It is then apparent that when variations in performance occur due to system modifications or a change in ambient conditions, the effect on the system output will be much reduced because of the feedback arrangement, i.e. with feedback the system is less sensitive to variations in conditions. Let us assume that we require the water temperature to be 50°C . With the system in open-loop, a variation in system characteristics would merely cause the temperature to drift away from the required value. With the system in closed-loop however, any variation from 50°C will be seen as an error when the actual measured temperature is compared with the required value, and the controller can then modify the control input in order to reduce the error to zero.

1.4.1 Effects of feedback

Open-loop systems rely entirely on calibration in order to perform with good accuracy; any system variations or effects caused by outside influences can seriously degrade this accuracy. Although only common in relatively simple systems, an important property is found in that an increase in system gain (amplification) does not affect the stability of an open-loop system (oscillations cannot be induced in this way).

Once feedback is applied, the system is in closed-loop. Closed-loop systems can achieve much greater accuracy than open-loop systems, although they rely entirely on the accuracy of the comparison between desired and actual output values and therefore on the accuracy of the measured output value. Effects of system variations or outside influences are much reduced, as are effects due to disturbances or nonlinear behavior. Unfortunately the advantages obtained when feedback is employed are at the expense of system stability (oscillations can often be induced by increasing system gain).

1.5 Some examples of control systems

Many simple control systems in everyday life include a human being, not only in the feedback loop but also to provide the actuating signals and even as the control system itself. Consider drinking a cup of coffee, which involves lifting a cup through a desired height. Feedback is provided via touch and sight, with a comparison being made between where the cup actually is at any time and where we want it to be, i.e. by our mouth – this is the error signal. An actuating signal is then provided in terms of lifting the cup with our hand at a rate which is dependent on how large the error signal is at any time. Performance is measured in terms of how quickly we can lift the cup – lifting it too quickly or going past the desired end position will result in coffee being spilt and possibly several broken teeth, conversely lifting it too slowly will result in severe boredom and even death through lack of liquid intake. This is really an example of a *biological control system*, everything within the system, except the cup and coffee, being part of the human being.

Note that outside influences such as windy weather, a lead weight in the bottom of the cup or the temperature of the coffee have not been considered. Although all of these would affect the performance in question, they were considered to be outside the scope of the system definition in the first instance. This is an important initial point because when modeling a control system the more characteristics that are included in the model, the more complex that model becomes. A trade-off therefore has to be made in terms of model complexity and the ability of that model to account for all eventualities. It would not be sensible to include the effect of high winds when modeling the act of drinking a cup of coffee because the vast majority of cups of coffee are not drunk in the presence of high winds.

Often the human being merely forms part of an overall control system, examples being driving an automobile and cooking, where the human can merely form the feedback loop – measuring the actual output, comparing this with the desired value and then modifying the control input accordingly.

Many control systems have been automated by replacing the human feedback element with more accurate equipment which also responds more rapidly. Examples of this are *ship steering* and *flight control*, for which pilots are replaced by autopilots whereby information from sensors, such as roll detectors and gyros respectively, is used to ensure constant velocity and constant heading. Any change in either of these requirements can be fed in merely as a change in set-point value. Taken further, both *ship docking* and *aircraft landing* can also be automatically carried out, by means of computational procedures either operating *in situ* or from a remote station. These latter examples serve as particular reminders that many different characteristics must be taken into account when modeling a system; neglecting the effect of rough weather or the presence of other craft could be fatal in this instance.

Other transit systems also employ automatic control techniques; certain automatic electric trains fall into this category along with autonomous guided vehicles (AGVs) which are not restricted to operation along certain lines, such as rails. With both visual

(camera) and distance (laser and sonar) information it is now possible for these vehicles to 'drive' themselves and it should not be long before greater use is made of such techniques in automobiles. The military interest in such vehicles is important for their development, as indeed it is for future developments in gun positioning, radar tracking and missile control, all of which involve rapid tracking of a target, with minimum error.

In manufacturing, *automated production lines* are not only supervised by an overall automatic controller, but each individual workstation can consist of a robot manipulator which either merely performs a set routine repeatedly with good accuracy and speed or which modifies its actions due to visual and/or tactile information feedback. On a similar basis the control of *machine tools* is gradually increasing in complexity from simple profile copying mechanisms, through numerical control for exact coordination, to machine tools which can modify their procedure in response to varying conditions such as differing workpiece requirements.

Power generation is also an important area in the application of automatic control systems. *Nuclear reactor control* is concerned with the rate at which fuel is fed into the reactor and it is in areas such as this that the effects of hazardous environment must also be taken into account. Certainly the large number of variables which need to be measured in order to provide a useful picture of the system presents many problems in terms of accuracy and reliability. Indeed other systems concerned with power generation and distribution share many instrumentation difficulties. Stricter controls imposed due to *environmental pollution* have meant not only increased efficiency requirements in the conversion schemes themselves but also the use of automatic control devices to control effluence.

Further down on the electrical power scale, *voltage stabilizers* are employed to retain the output voltage from a source as close as possible to a reference voltage value. This is done by comparing the actual voltage output with the reference, any difference (error) then being used to vary the resistance provided by a transistor which is connected in series with the source.

Uniformity in the thickness of such as paper, steel and glass is based on the control of the motor speed of rollers and drawing machines. Rollers are often not perfectly cylindrical and, combining this with possible large variations in the product quality (e.g. pulp obtained from a different factory), controlling output thickness within tight limits is not a simple task. Letting the thickness increase merely wastes money whereas letting it decrease produces an inferior product, hence tight limits are necessary. The control of *motors* both in terms of position and speed is however a much wider field than mentioned here, the complexity of control being dependent on the accuracy requirements made.

In terms of financial outlay, large amounts have been spent in providing accurate automatic systems for *process control*. In oil and chemical industries a small percentage improvement in reducing output variations (e.g. flow rate, concentration) in many cases results in millions of dollars being saved. Process plants are usually characterized as fairly slowly varying systems with only occasional changes in reference input. Controllers are required in order to minimize the effect of disturbances on the output signal. *Biochemical control* and *biomedical control* are in most cases closely related to process

control systems, with the former including such examples as drug production and fermentation, and the latter respiratory monitoring and pain relief.

Automatic control has become increasingly more important in *agricultural engineering*, with intelligent tractors now able to cope with wheel slip and to provide a constant ploughing level despite surface variations. Improved automobile *engine performance* and efficiency has become more and more dependent on complex control schemes, with power generated by means of the engine itself being used to provide energy for the electronic/microcomputer-based control circuitry.

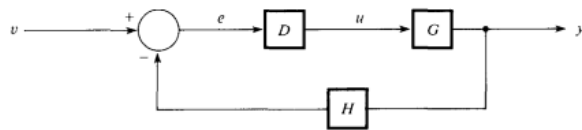
Obvious advantages can be obtained by successfully modeling *economic systems* in order to obtain suitable automatic control schemes. Apart from the most simple cases however it is extremely difficult to account for many of the spurious events that can occur in practice and which have considerable effect on the economic system under control, an example being the price of shares on the stock market. *Social systems* are similar to economic systems in terms of modeling difficulties and also suffer from the fact that very often there are both a large number of outputs and also different measures of performance which are often contradictory.

Hopefully a good idea of the wide variety of system types in which automatic control schemes operate has been given. When one is new to the subject, examples encountered are often related to relatively simple applications with which the reader may be familiar, e.g. electromechanical systems, electronic amplifiers, chemical plant, mechanical systems and electrical machines. It should be remembered that many of the simple techniques considered remain as simple techniques even when applied to the most complicated plant.

1.6 Definitions of standard terminology

As with most subject areas, terminology is used in the study of control systems which aptly describes phenomena within the field. It can, however, be seen as rather vague and in many cases confusing from the outsider's point of view. A list of definitions is therefore given here in order to help remove any barriers which do exist (see Fig. 1.6).

1. Lower case letters refer to signals, e.g. voltage, speed; and are functions of time, $u = u(t)$.
2. Capital letters denote signal magnitudes, as in the case of $u(t) = U \cos \omega t$, or otherwise Laplace transformed quantities, $U = U(s)$. Where $s = j\omega$, this is indicated by $U(j\omega)$.
Note: s is the Laplace operator and $\omega = 2\pi f$ where f is frequency.
3. The system under control is also known as the *plant* or *process*, G .
4. The *reference input*, v , also known as the set-point or desired output, is an external signal applied in order to indicate a desired steady value for the plant output.
5. The *system output*, y , also known as the controlled output, is the signal obtained from the plant which we wish to measure and control.

Fig. 1.6 Closed-loop system with feedback element H

6. The *error signal*, e , is the difference between the desired system output and the actual system output (when $H = 1$).
Note: See 8.
7. The *controller*, D , is the element which ensures that the appropriate control signal is applied to the plant. In many cases it takes the error signal as its input and provides an actuating signal as its output.
8. The *feedback element* H provides a multiplying factor on the output y before a comparison is made with the reference input v . When $H \neq 1$ the error e is the error between v and Hy , i.e. it is no longer the error between v and y .
Note: Often $H = 1$, although H can represent the characteristics of the measurement device in which case most likely $H \neq 1$.
9. The *feedback signal* is the signal produced by the operation of H on the output y .
10. The *control input*, u , also known as the actuating signal, control action or control signal, is applied to the plant G and is provided by the controller D operating on the error e .
11. The *forward path* is the path from the error signal e to the output y , and includes D and G .
12. The *feedback path* is the path from the output y , through H .
13. A *disturbance*, or noise (not shown in Fig. 1.6), is a signal which enters the system at a point other than the reference input and has the effect of undermining the normal system operation.
14. A *nonlinear* system is one in which the principles of superposition do not apply, e.g. amplifier saturation at the extremes, or hysteresis effects. Almost all except the most simple systems are nonlinear in practice, to an extent at least. The vast majority of systems can however be dealt with by approximating the system with a linear model, at least over a specific range.
15. A *time-invariant* system is one in which the characteristics of that system do not vary with respect to time. Most systems do vary slowly with respect to time, e.g. ageing, however over a short period they can be considered to be time-invariant.
16. A *continuous-time* system is one in which the signals are all functions of time t , in an analog sense.
17. A *discrete-time* system is a system such as a digital system or a sampled data system in which the signals, which consist of pulses, only have values at distinct time instants. The operator z is used to define a discrete-time signal such that $z^3 y(t) = y(t + 3)$ means the value of signal $y(t)$ at a point in time three periods in the future, where a period T (sample period) is defined separately for each system.

12 Introduction

18. A *transducer* converts one form of energy (signal) into another, e.g. pressure to voltage.
19. *Negative feedback* is obtained when $e = v - Hy$.
20. *Positive feedback* is obtained when $e = v + Hy$.
Note: This is not shown in Fig. 1.6.
21. A *regulator* is a control system in which the control objective is to minimize the variations in output signal, such variations being caused by disturbances, about a set-point mean value.
Note: A regulator differs from a servomechanism in which the main purpose is to track a changeable reference input.
22. A *multivariable* system is one which consists of several inputs and several outputs. In this text only single variable systems are considered.

1.7 Summary

In this first chapter the purpose has been to gently introduce the subject area of control systems, by providing a brief historical account of developments leading to its present state and also by putting forward the idea of feedback control in terms of some elementary forms. To do this it has been necessary to employ representative diagrams to show connections between signals and systems. These diagrams are described in much greater detail in the following chapter.

Many examples of the application of control systems were given in an attempt to show the multi-disciplinary nature of the subject. Once the fundamentals of a particular control technique have been studied in one field, they are generally applicable in many other fields. Control engineers therefore come from many different backgrounds in terms of subjects studied and an even more diverse set exists in terms of areas in which a control engineer can show a certain amount of understanding.

In the chapters which follow many techniques are shown for the modeling, control and study of system behavior. The ideas put forward should not be seen as limiting or restrictive in any way, in terms of applications, but rather as an insight to the wide variety of possibilities.

Problems

- 1.1 A thermostat is often used in order to control the temperature of a hot water tank. If the water is heated from cold, explain, with the aid of a sketch of water temperature versus time, the principle of operation of such a device. Why is a simple open-loop system not used instead?
- 1.2 Explain the open-loop operation of traffic signals at a road crossing. How can improved traffic control be achieved by means of a closed-loop scheme?

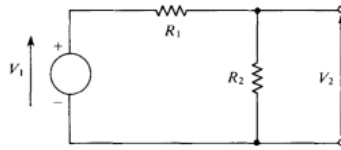


Fig. 1.7

- 1.3 Consider a missile guidance system which is based on an open-loop scheme, once the missile has been launched. How can a closed-loop technique be devised in order to track a target and what problems can this cause?
- 1.4 Consider the potential divider in Fig. 1.7. The output voltage is V_2 and the input voltage is V_1 .
How can the system be both an open- and closed-loop system at the same time?
- 1.5 The population of rabbits in a particular field is unaffected by events external to the field, including visiting rabbits. The population consists of adult males, adult females and young rabbits.
Each year 20% of adult males and females die, while the number of newborn young which survive to the following year is twice the number of adult females alive throughout the present year.
Each year 40% of the young become adult males the following year and 40% of the young become adult females the following year. Meanwhile, the remaining 20% of the young die.
Find equations to describe each section of the population from one year to the next.
- 1.6 Consider the simplified economic system describing the price of a toy. Demand for the toy will decrease if the price of the toy increases, whereas supply of the toy increases if its price increases. By finding the difference between supply and demand, show how this can be used as an error signal to effect a controlling action on the price.
- 1.7 Foxes move into the field inhabited by the rabbits of Problem 1.5, and use the rabbits as a food source. F is the number of foxes at any one time and R is the number of rabbits. The rate of change of foxes is equal to a constant K_1 multiplying the error between R and a number $\bar{R}$, whereas the rate of change of rabbits is equal to a constant K_2 multiplying the error between a number $\bar{F}$ and F .
How does the system operate; what happens to F and R with respect to time?
- 1.8 Consider the end effector of a robot manipulator. Explain how, in many cases, this is operated in open-loop mode. How could a closed-loop system be constructed for the manipulator?
- 1.9 In a liquid level system, what problems could occur if an open-loop scheme is employed? Can all of these problems be removed by employment of a

closed-loop technique? Are any problems brought about due to the closed-loop format itself?

- 1.10 Consider driving a vehicle without using any feedback, e.g. close your eyes and ears and drive (maybe you drive this way already!).

What advantages are obtained if a closed-loop vehicle driving scheme is used, which includes a human operator? (obvious answers). How could the vehicle driving scheme be configured without a human operator? Are there any particular advantages or disadvantages to such a technique?

Further reading

- Bennett, S., *A history of control engineering, 1800–1930*, Peter Peregrinus Ltd, 1986.
Black, H. S., 'Inventing the negative feedback amplifier', *IEEE Spectrum*, 1977, pp. 55–60.
Bode, H. W., 'Feedback – the history of an idea', in *Selected papers on mathematical trends in control theory*, Dover, New York, 1964, pp. 106–23.
Danai, K. and Malkin, S. (eds.), *Control of manufacturing processes*, ASME, 1991.
Dorf, R. C., *Modern control systems*, 6th ed., Addison-Wesley, 1992.
Enos, M. J. (ed.), *Dynamics and control of mechanical systems*, Mathematical Soc., 1993.
Hazen, H. L., 'Theory of servomechanisms', *J. Franklin Institute*, **218**, 1934, pp. 543–80.
Leigh, J. R., *Control theory: a guided tour*, Peter Peregrinus Ltd., 1992.
Lighthill, M. J., *Fourier analysis and generalised functions*, Cambridge University Press, 1959.
Maskrey, R. H. and Thayer, W. J., 'A brief history of electrohydraulic servo mechanisms', *ASME J. of Dynamic Systems, Measurement and Control*, 1978, pp. 110–16.
Maxwell, J. C., 'On governors', *Proceedings of the Royal Society (London)*, **16**, 1868, pp. 270–83.
Mayr, O., *Origins of feedback control*, MIT Press, 1971.
Minorsky, N., 'Directional stability of automatically steered bodies', *J. Am. Soc. Naval Eng.*, **34**, 1922, p. 280.
Minorsky, N., 'Control problems', *J. Franklin Inst.*, **232**, 1941, p. 451.
Nyquist, H., 'Regeneration theory', *Bell System Technical Journal*, **11**, 1932, pp. 126–47.
Popov, E. P., *The dynamics of automatic control systems*, Addison-Wesley, 1962.
Routh, E. J., 'Stability of a given state of motion', Adams Prize Essay, Macmillan, 1877.
Singh, M., *Encyclopedia of systems and control*, Pergamon Press, 1987.
Thaler, G. J., *Automatic control: classical linear theory*, Dowden, Hutchinson and Ross Inc., Stroudsburg, Pa., 1974.
Tischler, M. B. (ed.), 'Aircraft flight control', *Int. J. Control (Special Issue)*, **59**, No. 1, 1994.
Tzafestas, S. G. (ed.), *Applied control*, Marcel Dekker, 1993.

- MacFarlane, A. G. J., *Dynamical system models*, Harrap, 1980.
 Murphy, J., *Power control of AC Motors*, Pergamon Press, 1990.
 Nash, P., *Systems modelling and optimisation*, Peter Peregrinus Ltd, 1981.
 Nicholson, H. (ed.), *Modelling of dynamical systems – Vol. 1*, Peter Peregrinus Ltd, 1980.
 Nicholson, H. (ed.), *Modelling of dynamical systems – Vol. 2*, Peter Peregrinus Ltd, 1981.
 Raven, F. H., *Automatic control engineering*, 2nd ed., McGraw-Hill, 1990.
 Shearer, J., *Dynamic modelling and control of engineering systems*, Macmillan, 1990.
 Shinnars, S. M., *Modern control system theory and application*, Addison-Wesley, 1978.
 Truxal, J. G., *Automatic feedback control system synthesis*, McGraw-Hill, 1955.
 Van Dixhoorn, J. J. and Evans, F. J. (eds.), *Physical structure in systems theory*, Academic Press, 1974.
 Wellstead, P. E., *Introduction to physical system modelling*, Academic Press, 1979.
 Ziegler, B., *Theory of modelling and simulation*, John Wiley, 1976.

3

System response using differential equations

3.1 Introduction

In Chapter 2 various ways were considered in which a system to be controlled could be modeled in terms of a representative transfer function relating the system's input to its output. The Laplace operator was used to great effect, and it can be seen through the remainder of this text that this operator is a major mathematical tool in the formulation and design of continuous-time control systems.

The application of a controller to any system will result in an increase in overall complexity and, perhaps more obviously, will usually require some financial outlay. It is important therefore to decide, in the first instance, whether or not it is necessary to employ a controller in a particular case, the actual controller design is then of only secondary importance. So the initial question to be answered is, 'Does the performance of the present system satisfy our requirements?' But this raises a further problem in that we must have a clear definition of what we mean by performance. As an example, for a car, we could be interested in speed, reliability, comfort and/or fuel consumption in terms of one criterion in particular or more likely in some combination of all of them.

In this chapter some of the fundamental performance criteria used for system analysis are introduced by considering the output response of a system in terms of its dependence on one of a standard set of basic test input signals. Not only is it much simpler and cheaper to apply a fairly straightforward test input, the subsequent analysis is generally easier than that required for a complex test input and it is most likely the case that the input signals which are in practice applied to that particular system can be thought of as combinations of the basic forms of input.

In assessing the performance of a system, analysis can be concentrated on either the initial reaction of the system output, following a particular test input, or the reaction observed once things have settled down to a steady form of response. It is noticeable that the output response is viewed in terms of its time domain features, and although it may not be initially obvious, it is shown here that s -domain system representations are an ideal form and that no inverse transformation is necessary when investigating time domain specifications. The total output response witnessed is therefore considered to

consist of two parts (a) the initial reaction, and (b) the steady reaction, which is known as the steady-state response and is essentially that part of the response which is left when the transient response dies down to zero. In the detailed study of the steady-state response of a system, carried out in Section 3.2, it is shown how the observed output is dependent on the input signal applied and in Section 3.4 the possibility of an error existing in the steady-state between the actual output and the desired output/reference input is shown to be directly related to the application of specific input signals to certain types of systems.

All physical systems, when they are asked to respond, do not do so instantaneously by jumping to a final steady value. In some cases it might be that the response is rather sluggish and slowly reaches its final value, whereas in other cases it is possible for the output to very rapidly overshoot its final value and then to slowly home in on it by decreasing its overshoots as time goes on. These different transient response possibilities are considered in Section 3.3.

A further, but vital, contribution of this chapter is to introduce some of the positive aspects gained with the application of feedback to form a closed-loop control system, such as reduction of the effect of noise on the output signal and reduction in output sensitivity to plant variations. Some of the main points made in the chapter are then brought together in the extended servomechanism example of Section 3.5, in which the use of feedback to improve both transient and steady-state system performance is discussed and this results in a basic controller design problem.

3.2 Steady-state response

Because the total response of a system is made up of two parts, the transient and steady-state responses, it is possible to think of the steady-state response as being the response which remains when the transient response has died down, or more simply as the transient response subtracted from the total response. Indeed this latter definition is that employed to investigate response errors in Section 3.4.

It is wrong to think that the steady-state response of a system refers in every case to a steady system output value although, as is shown in this section this might be the case, it is not necessarily so. If a time invariant signal is applied at the input then a different steady output response would be expected when compared with the response obtained by applying a sinusoidal input to the same system. So the steady-state response of a system can only be considered by making reference to the type of input applied, and as the transient response will by definition tend to zero at time $t \rightarrow \infty$, so the steady-state response to an input will be that apparent as $t \rightarrow \infty$.

A variety of standard input signals are available for employment in order to determine the response characteristics of a system, a common element between these inputs being their relatively simple implementation. In most cases the system, when in normal operation, may well have to deal with a multitude of different input signals or conversely only have to face one specific type of input. Whatever the case it is usual to investigate

the system response to only one or two of the standard input forms, because in general this will provide sufficient information for further analysis.

In the section which follows directly, the most commonly encountered standard input test signals are described, with the more popular types discussed in greater detail. Subsequently the steady-state response of systems to certain of the inputs is considered with examples being given of some low order, fairly straightforward system transfer functions.

3.2.1 Control test inputs

The input to some systems can, during normal operation, be considered as random or at least pseudo-random, one example being systems whose input is affected by a large random disturbance. It may then be necessary to apply a pseudo-random input in order to obtain detailed information as to the characteristics of the system; the same is true for many complicated systems which are inherently of high order. However for essentially simple systems, and in order to obtain important fundamental information on more complicated systems, the test signals applied as inputs are generally of a fairly simple standard form.

Standardizing test input signals is important when assessing the relative merits of several systems by comparing the output responses to a common input signal. Perhaps the simplest of all inputs to employ is a *step input* which can be applied, for example, by switching a voltage in an electrical circuit or a sudden force which is retained in a mechanical network. If it is assumed that the step is of unit magnitude and is applied at $t = 0$, then this is shown in Fig. 3.1.

The unit step input is a primary test signal and can be described by

$$u(t) = \begin{cases} 1: & t > 0 \\ 0: & t < 0 \end{cases} \quad (3.2.1)$$

where the input magnitude is shown to be unity after time $t = 0$. In practice it will most likely be sensible to apply a step input of some magnitude other than unity, this being

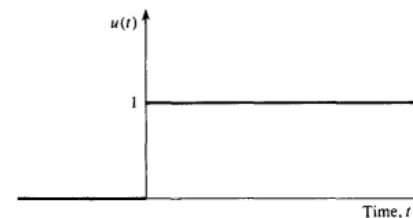


Fig. 3.1 Unit step input signal

particularly true for a low gain system when the presence of noise on the output could be problematic unless a high magnitude step is used. It can also be noticed that the step input is undefined for $t = 0$, the assumption being that at $t = 0$ the input changes from 0 to 1, an idealistic view, since a finite but very small time is actually necessary to carry out such a transition. In general though the response of systems considered in this text, governed by their time constants, will be much slower than the step transition time, which can therefore be neglected. For the study of very high speed systems, logic gates for example, a more complicated description of the step input would, however, most likely be necessary.

The Laplace transform of the ideal unit step input (3.2.1) is given by

$$U(s) = 1/s \quad (3.2.2)$$

and it is worth noting that a 4-unit step input, for example, has a Laplace transform $U(s) = 4/s$.

A second test signal is the *impulse function*. This was considered to an extent in the previous chapter, and is a simple derivative of the step input; it therefore has a Laplace transform of

$$U(s) = 1 \quad (3.2.3)$$

when applied as a control input signal.

In its strictest form an impulse function lasts for a time interval δt as $\delta t \rightarrow 0$, and is of infinite magnitude. As the practical application of such a signal poses obvious problems, it is usual to approximate the impulse with a signal of unit magnitude which lasts for a very short time period δt , and this is in fact discussed at length in Chapter 2.

If the integral is taken of the step input this results in a function which increases linearly with respect to time from a zero value at time $t = 0$. The *ramp function*, which can also be regarded as a velocity function, is shown in terms of a unit magnitude ramp in Fig. 3.2.

The unit ramp input can be written as

$$u(t) = \begin{cases} t: & t > 0 \\ 0: & t < 0 \end{cases} \quad (3.2.4)$$

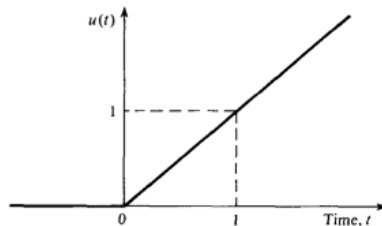


Fig. 3.2 Unit ramp input signal

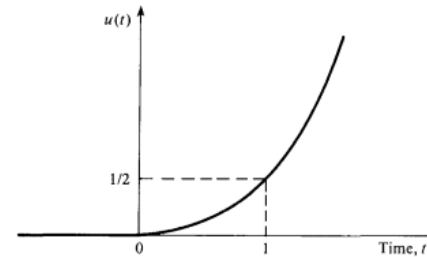


Fig. 3.3 Unit parabolic input signal

or in terms of the Laplace transform

$$U(s) = 1/s^2 \quad (3.2.5)$$

By integrating the ramp function, hence obtaining an acceleration function, a square law signal known as a *parabolic function*, is achieved. This is shown in terms of a unit parabolic signal in Fig. 3.3.

The unit parabolic input can be written as

$$u(t) = \begin{cases} t^2/2: & t > 0 \\ 0: & t \leq 0 \end{cases} \quad (3.2.6)$$

where the division by 2 in (3.2.6) is included in order to think of the function as a step acceleration term with respect to the original step input (3.2.1). This can also be considered as

$$U(s) = 1/s^3 \quad (3.2.7)$$

by means of the Laplace transform.

It is apparent that by combining the step, ramp and parabolic inputs to form one single input signal, for all $t > 0$, this can be written as

$$u(t) = 1 + t + \frac{t^2}{2}$$

It is quite possible, however, to consider further integrated terms in order to arrive at an overall input

$$u(t) = 1 + t + \frac{t^2}{2} + \dots + \frac{t^p}{p!}$$

although a more general expression is obtained by considering each separate term to have a magnitude equal to a constant d_i , resulting in a power series or *polynomial*

function

$$u(t) = d_0 + d_1 t + d_2 \frac{t^2}{2!} + \dots + d_p \frac{t^p}{p!} \quad (3.2.8)$$

for $t > 0$; $u(t) = 0$ for $t < 0$.

This polynomial function can though be written more concisely as

$$u(t) = \sum_{i=0}^p d_i t^i / i! \quad t > 0$$

where d_i are scalar constant values.

The final standard test signal introduced here is the *sinusoidal function* which can be written in unit magnitude form as

$$u(t) = \begin{cases} \cos \omega t: & t > 0 \\ 0: & t \leq 0 \end{cases} \quad (3.2.9)$$

This type of input signal is considered in the following section in terms of the steady-state response which results from its application.

3.2.2 Sinusoidal and polynomial inputs

Analyzing the steady-state response of a system to a sinusoidal input is not in general possible by means of the Laplace transform final value theorem. This theorem requires that there exists a limiting value to which a signal converges as time tends to infinity, and for a sine wave no such value exists.

For the sinusoidal input defined in (3.2.9) we have that for all $t > 0$, $u(t) = \cos \omega t$, remembering that if the input is of magnitude U , in terms of a phasor diagram this can be represented as in Fig. 3.4.

The output signal in the steady-state will then also be a sinusoid, as long as the system to which the input is applied does not contain any instabilities (the stability of a system is discussed in Chapter 4). So if the steady-state system output, $y(t)$, is obtained in

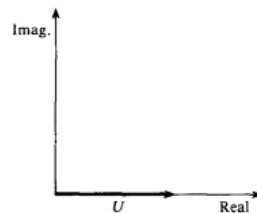


Fig. 3.4 Phasor diagram representation of $u(t) = U \cos \omega t$

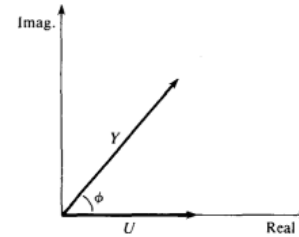


Fig. 3.5 Phasor diagram representation of $y(t) = Y \cos(\omega t + \phi)$, where $y(t)$ is the steady-state output

response to a sinusoidal input acted upon by a system whose transfer function may be described by means of a differential equation as defined in Chapter 2, then the output will be a sinusoid of magnitude Y , with a phase shift denoted by the angle ϕ , as shown on the phasor diagram in Fig. 3.5.

A general polynomial test input was defined in (3.2.8). It is assumed that the system to which the input is applied is defined by a differential equation, then we have that

$$(a_n D^n + a_{n-1} D^{n-1} + \dots + a_1 D + a_0) y(t) = (b_m D^m + b_{m-1} D^{m-1} + \dots + b_1 D + b_0) u(t)$$

in which the a_i and b_i coefficients determine the characteristics of the system transfer function, as described in Chapter 2. The application of a test input (3.2.8) will result in a steady-state output signal which will also contain enough terms in powers of time t to cause equality in the system differential equation. As an example consider $n = 2$ and $m = 1$ in the above equation with a control input defined by

$$u(t) = d_0 + d_1 t + \frac{d_2 t^2}{2}$$

We have then that:

$$\begin{aligned} (a_2 D^2 + a_1 D + a_0) y(t) &= (b_1 D + b_0) \left[d_0 + d_1 t + \frac{d_2 t^2}{2} \right] \\ &= b_1 d_1 + b_0 d_0 + (b_1 d_2 + b_0 d_1) t + \frac{b_0 d_2 t^2}{2} \end{aligned}$$

If in general the steady-state output is described by

$$y(t) = c_0 + c_1 t + \frac{c_2 t^2}{2!} + \dots + \frac{c_k t^k}{k!}$$

then in order to obtain equality in the example equation we must have that

$$y(t) = c_0 + c_1 t + \frac{c_2 t^2}{2!}, \quad \text{i.e. } k = 2$$

from which it follows that

$$(a_2 D^2 + a_1 D + a_0)y(t) = a_2 c_2 + a_1 c_1 + a_0 c_0 + (a_1 c_2 + a_0 c_1)t + \frac{a_0 c_2 t^2}{2}$$

The steady-state output coefficients c_0 , c_1 , c_2 can then be found by solving the three equations obtained, i.e.

$$a_0 c_2 = b_0 d_2$$

$$a_1 c_2 + a_0 c_1 = b_1 d_2 + b_0 d_1$$

and

$$a_2 c_2 + a_1 c_1 + a_0 c_0 = b_1 d_1 + b_0 d_0$$

Notice that the power index k is dependent solely on the complexity of the input signal, i.e. the index p , so

$$k = p \quad (3.2.10)$$

3.2.3 Step, ramp and parabolic inputs

The step input was defined in Section 3.2.1. If such an input is applied to a first-order system transfer function of the form

$$G(s) = \frac{K}{s + a_0}$$

where K is the system gain, then the system output is obtained from $Y(s) = G(s) \cdot U(s)$ to be

$$Y(s) = \frac{K}{s + a_0} \cdot \frac{1}{s}$$

in which $1/s$ is the unit step Laplace transform.

The output is then

$$Y(s) = \frac{K/a_0}{s} - \frac{K/a_0}{s + a_0}$$

or in terms of its inverse transform

$$y(t) = \frac{K}{a_0} - \frac{K}{a_0} \exp(-a_0 t)$$

the time response of which is shown in Fig. 3.6, where $a_0 > 0$.

It can be considered either that $1/a_0$ seconds is the time taken for the output to reach within $\exp(-1)$ of its final value, i.e. the time to reach 63.2% of the final value, or $1/a_0$ is the time taken to reach the final value if the initial rate of increase is continued.

The final value of the output signal, that is the steady-state output value, can be found

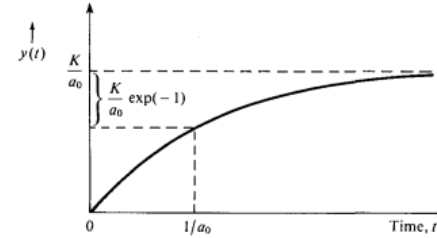


Fig. 3.6 First-order system response to a unit step input

by means of the final value theorem of the Laplace transform, which states that if the limit of a function $f(t)$ as $t \rightarrow \infty$ exists, then

$$\lim_{s \rightarrow 0} sF(s) = \lim_{t \rightarrow \infty} f(t) \quad (3.2.11)$$

So for the first-order example just discussed,

$$\lim_{s \rightarrow 0} sY(s) = s \cdot \frac{K}{s + a_0} \cdot \frac{1}{s} \bigg|_{s \rightarrow 0} = \frac{K}{a_0}$$

and this agrees with the result found from the inverse Laplace transform as $t \rightarrow \infty$; shown in Fig. 3.6.

By applying a unit ramp input to the same system transfer function, the output transform is now given by

$$Y(s) = \frac{K}{s + a_0} \cdot \frac{1}{s^2}$$

and on applying the final value theorem we find that

$$\lim_{s \rightarrow 0} sY(s) = s \cdot \frac{K}{s + a_0} \cdot \frac{1}{s^2} \bigg|_{s \rightarrow 0} = \infty$$

So, for a first-order system as shown, application of a unit ramp input will result in the system output tending to infinity as $t \rightarrow \infty$. This can also be seen by finding the inverse Laplace transform which is

$$y(t) = \frac{K}{a_0} \left[t - \frac{1}{a_0} + \frac{\exp(-a_0 t)}{a_0} \right]$$

such that the output tends to infinity as $t \rightarrow \infty$.

A similar result in which the output tends to infinity is witnessed as the first-order system response to a unit parabolic input, given from the final value theorem as

$$\lim_{s \rightarrow 0} sY(s) = s \cdot \frac{K}{s + a_0} \cdot \frac{1}{s^3} \bigg|_{s \rightarrow 0} = \infty$$

in which $1/s^3$ is the unit parabolic input.

In the following section the response of a second-order system, in particular to a unit step input, is considered at length in order to introduce the possibility of an output response overshooting its final value. For the first-order system considered in this section it is not possible to obtain any overshoot; the output value will always approach its final value from one direction only. The value of K merely part defines the actual final output value for a first-order system whereas the value of a_0 , as well as partly defining the actual final output value, characterizes how rapidly that final value is approached and is therefore an important measure of system performance, the figure $1/a_0$ being called the *time constant*.

A general second-order system transfer function can be written as

$$G(s) = \frac{K}{s^2 + a_1s + a_0} \quad (3.2.12)$$

Note: This is second order because the highest power of $s = 2$.

By analyzing the output of this system in response to a unit step input signal, the final value theorem reveals

$$\lim_{s \rightarrow 0} sY(s) = s \cdot \frac{K}{s^2 + a_1s + a_0} \cdot \frac{1}{s} \bigg|_{s \rightarrow 0} = \frac{K}{a_0}$$

which is an identical result to that found for a first-order system. It is worth noting that if $a_0 = 0$ the steady-state output signal tends to infinity, in fact with $a_0 = 0$ the system denominator reduces to $s(s + a_1)$, i.e. it contains a common s factor – it is a type 1 system, see Section 3.4.1, and this factor s , when set to zero, causes the infinite output.

The time response, to a step input, of a first-order system was shown in Fig. 3.6, and this diagram holds for all $a_0 > 0$, different values of a_0 merely affecting the time and magnitude axis scaling. The step input time response of a second-order system is rather more complicated, the type of response obtained being dependent on whether the system denominator roots (the system poles) are real or complex, and this is the subject of the following section.

3.3 Transient response

In the previous section (3.2), the steady-state response of a system was considered to be that part of the total response which remains as time $t \rightarrow \infty$. In a similar fashion the transient response of a system can be regarded as that part of the total response which tends to zero as time $t \rightarrow \infty$. The total response therefore consists of the steady-state response summed with the transient response. The concept that a system has a transient response which tends to zero implies that the system has certain properties which prevent the output from constantly increasing in magnitude towards infinity. This is the same as requiring that the system is 'stable', a topic discussed in Chapter 4.

All of the analysis in this section is based on the assumptions that (a) the control test input is a unit step function, and (b) the system is represented exactly by a second-order

transfer function. Ramp or parabolic inputs could be applied as well as or instead of the step input, however no further important aspects are revealed by such a procedure when compared with a step input response, and for the majority of systems it will be a step response which is most relevant in practice. Restricting the analysis to second-order systems only by no means limits the applicability of the results. Although certain systems are of order greater than two, they can in many cases be modeled adequately by a second-order system, the total response obtained being dominated by that due to a second-order model.

3.3.1 Overdamped response

If the system transfer function has two real denominator roots, then it follows that:

$$G(s) = \frac{K}{s^2 + a_1s + a_0} = \frac{K}{(s + \alpha)(s + \beta)} \quad (3.3.1)$$

By applying a unit step input to such a system, the output signal transform is found to be:

$$\begin{aligned} Y(s) &= \frac{K}{(s + \alpha)(s + \beta)} \cdot \frac{1}{s} \\ &= \frac{K}{\alpha\beta} \left[\frac{1}{s} + \frac{\beta}{(\alpha - \beta)} \cdot \frac{1}{(s + \alpha)} + \frac{\alpha}{(\beta - \alpha)} \cdot \frac{1}{(s + \beta)} \right] \end{aligned}$$

or in terms of its time-domain representation, this is

$$y(t) = K \left\{ \frac{1}{\alpha\beta} + \frac{1}{(\alpha - \beta)} \left[\frac{\exp(-\alpha t)}{\alpha} - \frac{\exp(-\beta t)}{\beta} \right] \right\} \quad (3.3.2)$$

It can be noted from the step response that by setting t to infinity in (3.3.2), and assuming $\alpha > 0$, $\beta > 0$,

$$\lim_{t \rightarrow \infty} y(t) = \frac{K}{\alpha\beta} = \frac{K}{a_0}$$

where, from (3.3.1), $a_0 = \alpha\beta$.

This result agrees with that obtained by application of the final value theorem in Section 3.2.3.

A further point to note is that at $t = 0$:

$$y(t) = K \left\{ \frac{1}{\alpha\beta} + \frac{1}{(\alpha - \beta)} \left[\frac{1}{\alpha} - \frac{1}{\beta} \right] \right\} = 0$$

and

$$\frac{dy(t)}{dt} = 0 \bigg|_{t=0}$$

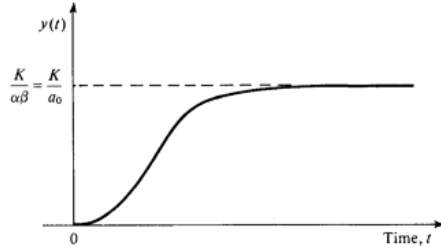


Fig. 3.7 Overdamped response to a step input

The overall time response for the overdamped case is shown in Fig. 3.7, where the initial rate of change of output signal is identically zero.

In common with the response of a first-order system to a step input, the output signal always approaches its final value from one side, as shown here it increases from zero to its final value with no overshoot.

The effect of having a numerator root in the system transfer function, e.g.

$$G(s) = \frac{K(s + \gamma)}{(s + \alpha)(s + \beta)}$$

is to cause the differential $dy(t)/dt$ at time $t = 0$ to be nonzero. It has very little effect on the time response other than that.

3.3.2 Critically damped response

A critically damped response can in a certain sense be regarded as a special or extreme case of the overdamped response. If the two denominator roots of the system transfer function are identical, then

$$G(s) = \frac{K}{(s + \alpha)^2} \quad (3.3.3)$$

When a unit step input is applied to this system the output signal transform is

$$\begin{aligned} Y(s) &= \frac{K}{(s + \alpha)^2} \cdot \frac{1}{s} \\ &= \frac{K}{\alpha^2} \left[\frac{1}{s} - \frac{1}{(s + \alpha)} - \frac{\alpha}{(s + \alpha)^2} \right] \end{aligned}$$

and in terms of its time-domain representation:

$$y(t) = \frac{K}{\alpha^2} \{ 1 - \exp(-\alpha t) - \alpha t \cdot \exp(-\alpha t) \}$$

or

$$y(t) = \frac{K}{\alpha^2} \{ 1 - \exp(-\alpha t)[1 + \alpha t] \} \quad (3.3.4)$$

Assuming that $\alpha > 0$, we have that

$$\lim_{t \rightarrow \infty} y(t) = \frac{K}{\alpha^2} = \frac{K}{a_0}$$

where

$$a_0 = \alpha^2$$

Initial conditions can be found by setting t to zero in (3.3.4) which gives

$$y(t) = 0$$

and also

$$\left. \frac{dy(t)}{dt} \right|_{t=0} = 0$$

The overall time response for the critically damped case is shown in Fig. 3.8, in which the initial rate of change of output signal is zero.

The critically damped response represents the most rapid type of response which does not result in any overshoot, i.e. the output signal in Fig. 3.8 never greater than its final value. This property is used to good effect in some measuring instruments and is particularly useful for example in certain machine tool applications or robot welders, in which it would be extremely dangerous to have any overshoot and yet taking a long time to reach the desired position, as would be the case for an overdamped response, would mean a waste of revenue due to less productivity. For many systems however it is desirable for an overshoot to occur before the output settles on its steady-state value, and such a response is described in the next section.

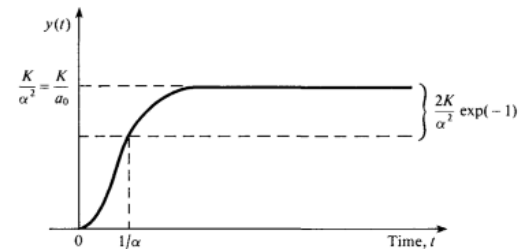


Fig. 3.8 Critically damped response to a step input

3.3.3 Underdamped response

The underdamped response is characterized principally by the output overshooting its final steady-state value following the application of a step input. This is by far the most widely studied type of transient response. More parameters are used to describe the response than in the previous two cases and hence in an academic environment it is a much more fruitful topic as far as exam questions are concerned.

If the two denominator roots of the system transfer function are complex (complex conjugates in fact) then the output signal will exhibit oscillations when a step input signal is applied. Just how quickly these oscillations die down is dependent on the degree of damping, such that if the response is not damped at all the oscillations will in theory carry on for ever. However, as the amount of damping is increased so the oscillations die down more quickly until we reach a critically damped response in which the degree of damping is so great that the oscillations do not even start up!

Assuming that the system transfer function contains complex roots, then we have

$$G(s) = \frac{K}{(s + \alpha + j\beta)(s + \alpha - j\beta)} \quad (3.3.5)$$

or

$$G(s) = \frac{K}{s^2 + 2\alpha s + (\alpha^2 + \beta^2)}$$

which is also

$$G(s) = \frac{K}{(s + \alpha)^2 + \beta^2}$$

such that when compared with the standard polynomial denominator form given in (3.3.1) it follows that in this case

$$a_0 = \alpha^2 + \beta^2 \quad \text{and} \quad a_1 = 2\alpha$$

By applying a unit step input to the system (3.3.5) the output signal transform is

$$\begin{aligned} Y(s) &= \frac{K}{(s + \alpha)^2 + \beta^2} \cdot \frac{1}{s} \\ &= \frac{K}{\alpha^2 + \beta^2} \left[\frac{1}{s} - \frac{(s + \alpha)}{(s + \alpha)^2 + \beta^2} - \frac{\alpha}{(s + \alpha)^2 + \beta^2} \right] \end{aligned}$$

In terms of its time-domain representation this is

$$y(t) = \frac{K}{\alpha^2 + \beta^2} \left\{ 1 - \exp(-\alpha t) \cdot \cos \beta t - \frac{\alpha}{\beta} \exp(-\alpha t) \sin \beta t \right\}$$

or

$$y(t) = \frac{K}{\alpha^2 + \beta^2} - \frac{K \exp(-\alpha t)}{\beta(\alpha^2 + \beta^2)^{1/2}} \left\{ \frac{\beta}{(\alpha^2 + \beta^2)^{1/2}} \cos \beta t + \frac{\alpha}{(\alpha^2 + \beta^2)^{1/2}} \sin \beta t \right\}$$

70 System response using differential equations

and if we define

$$\phi = \tan^{-1} \frac{\beta}{\alpha} \quad (3.3.6)$$

the time response can be written as

$$y(t) = \frac{K}{(\alpha^2 + \beta^2)} \left\{ 1 - \frac{(\alpha^2 + \beta^2)^{1/2}}{\beta} \cdot \exp(-\alpha t) \cdot \sin(\beta t + \phi) \right\} \quad (3.3.7)$$

So the response is a sine wave of frequency β with an exponential damping governed by the factor α and a steady-state value given by $K/(\alpha^2 + \beta^2)$, as shown in Fig. 3.9.

The most usual way to write the system denominator of (3.3.5) is in fact in the form $s^2 + 2\zeta\omega_n s + \omega_n^2$, where

$$s^2 + 2\zeta\omega_n s + \omega_n^2 = s^2 + 2\alpha s + (\alpha^2 + \beta^2) \quad (3.3.8)$$

in which ζ is the damping ratio and ω_n the natural frequency. These definitions arise because:

(a) We have from (3.3.8) that:

$$\zeta = \alpha/\omega_n \quad \text{and} \quad \omega_n^2 = \alpha^2 + \beta^2$$

such that if $\beta \rightarrow 0$, something which causes the roots of the denominator to tend to those of the critically damped system – see (3.3.5) and (3.3.3) – then the damping ratio $\zeta \rightarrow 1$. On the other hand if $\alpha \rightarrow 0$, so the exponential damping tends to zero and we approach the case of a purely oscillatory response – see (3.3.7), i.e. $\zeta \rightarrow 0$. Hence for an underdamped response $0 < \zeta \leq 1$, where the damping is heavier as ζ tends to unity. In fact for $\zeta > 1$ this corresponds to the overdamped case discussed in Section 3.3.1 in which the system denominator consists of two nonidentical real roots.

(b) If the damping ratio is zero then $\zeta = 0 = \alpha$ and $\omega_n = \beta$, such that ω_n is the frequency at which oscillations will occur when there is no damping and in such a case the oscillations would be sustained indefinitely. Note that for a particular response, β is the actual frequency of the oscillations, but this will only be equal to the natural frequency ω_n in the limit as the damping tends to zero.

The response shown in Fig. 3.9 indicates several basic characteristics of a system which are used to describe, in relatively simple terms, the performance of that system for comparison or design purposes. In the initial part of the response it must be remembered however that the phase shift ϕ will have a certain effect such that $\phi \rightarrow 90^\circ$ as the damping tends to zero, hence for an undamped system the response will be dictated from time $t = 0$ by a cosine waveform.

The initial part of the response is most often described by means of two terms, firstly the *delay time*, which is the time taken for the output step response to reach 50% of its final value, and secondly the *rise time*, which is the time taken for the output step response to rise from 10 to 90% of its final value. At a later stage in the time response, the *settling time*, shown as T_s in Fig. 3.9, is defined as the time taken for the step

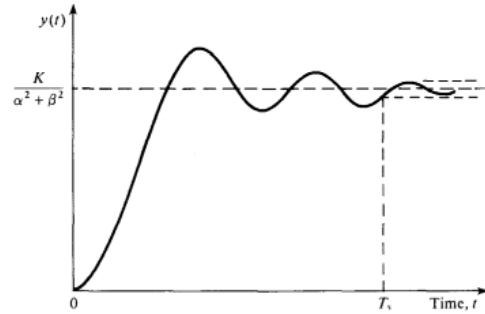


Fig. 3.9 Underdamped response to a step input

response to reach and stay within a specified percentage of its final value. A commonly encountered figure is 5%, and this was employed for the response shown in Fig. 3.9.

Another term used to describe the transient performance of a system is overshoot, or more usually *maximum overshoot*, which is the maximum difference between the transient and steady-state step response values that occurs after the response has first passed its final value. The maximum overshoot is frequently written as a percentage of the steady-state output response value, i.e.

$$\frac{\text{maximum overshoot} \times 100\%}{\text{steady-state value}}$$

The maximum and minimum values of the output response can be found by equating the differential form of (3.3.7) to zero, which results in

$$\frac{dy(t)}{dt} = \frac{K}{\beta} \cdot \exp(-\alpha t) \cdot \sin \beta t = 0 \quad (3.3.9)$$

It can be seen that the maximum and minimum values of the differential are dictated by the sine term only, such that they occur when $t = i\pi/\beta$, where $i = 0, 1, 2$, etc. This means that on substitution back into (3.3.7) the magnitude of each maximum and minimum value is given by

$$y(i) = \frac{K}{\alpha^2 + \beta^2} \{1 - \exp(-i\pi\alpha/\beta) \cdot \cos i\pi\} \quad (3.3.10)$$

in which $i = 0, 1, 2$, etc.

At the time instants denoted by the index i , the output response therefore has values of

$$y(i) = \frac{K}{\alpha^2 + \beta^2} \{1(-)^{i+1} \exp(-i\pi\alpha/\beta)\}$$

Relative to its final value the output response values for $i = 1, 2$, etc., are termed overshoots, the initial condition being $y(i) = 0$ for $i = 0$. The maximum overshoot is

then the first overshoot, i.e. when the exponential term is at its largest, and this is found from the actual magnitude at $i = 1$ to be

$$\text{maximum overshoot} = \frac{K}{\alpha^2 + \beta^2} \cdot \exp(-\pi\alpha/\beta)$$

where the steady-state value $K/(\alpha^2 + \beta^2)$ has been subtracted from the actual response magnitude in order to obtain the magnitude of overshoot at this time instant. In terms of a percentage this is

$$\text{maximum overshoot (\%)} = \exp(-\pi\alpha/\beta) \times 100\% \quad (3.3.11)$$

and the time taken to reach this maximum, first, overshoot is $t = \pi/\beta$.

The maximum overshoot value is in fact also the absolute value of the ratio of successive overshoots.

Consider for example the magnitude of the second overshoot divided by the magnitude of the first overshoot, which is

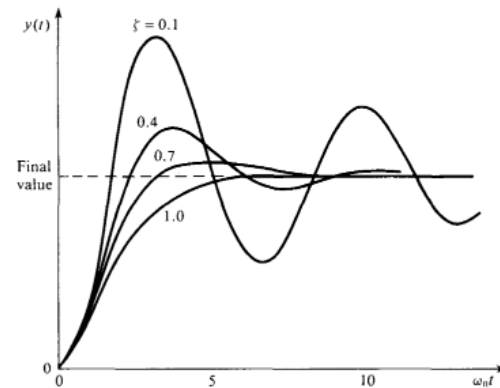
$$\text{successive overshoot ratio} = \frac{\exp(-2\pi\alpha/\beta)}{\exp(-\pi\alpha/\beta)} = \exp(-\pi\alpha/\beta)$$

This is then also an indication of how quickly the oscillations die away in magnitude. It is often written as a logarithmic value, that is

$$\text{logarithmic decrement} = 20 \log_{10} \exp(-\pi\alpha/\beta) \approx -27.29 (\alpha/\beta) \text{ dB/overshoot}$$

A special case of the logarithmic decrement is arrived at when $\alpha = \beta$, in which case we observe a loss of 27.29 dB per overshoot, and this particular relationship between α and β results in what is known as *ideal damping*, with a transfer function

$$G(s) = \frac{K}{(s + \alpha)^2 + \alpha^2} \quad (3.3.12)$$



However in the design of a system it might, in practice, not be possible to achieve an exact equality between α and β , and this leads to the idea of *acceptable damping*, for which

$$3^{-1/2} < \alpha/\beta < 3^{1/2}$$

To conclude this section, a set of step responses is given in Fig. 3.10 in order to show the difference in signal decay due to the damping factor in each case.

3.4 Errors

The performance of a system is usually assessed in terms of the error between the demanded or required output signal and the actual output signal that is obtained. Any error which does occur can be caused by one or more of a number of reasons, for instance the output measurement could be incorrect because of poor calibration or noise interference. The problem of noise is quite a serious one in many systems, and the objective of some controllers is simply to regulate the variation in output level, caused by disturbances, about a reference value. The effect of noise signals on a system is considered later in this section and the advantages gained by the employment of a feedback control scheme, in terms of disturbance error reduction, are discussed.

Another major source of errors is the employment of an incorrect system model for controller design purposes. It might be that a poor model was calculated in the first instance, although it is quite possible that even though the model was a good one when obtained, the system has aged or has been modified since then and the representation that we have of the system is no longer suitable. A control system design can be investigated in order to find out how sensitive it is to variations in the system description, i.e. if the actual system is not quite the same as we have pictured it to be, by how much is the control design affected? The sensitivity of a control system is also discussed later in this section.

The first classification of error to be considered here is that caused simply by the effect of certain system types on straightforward input signals, namely step, ramp and parabolic inputs. When applying a specific desired reference input to a system it is generally hoped that the system output will settle down to the desired level as steady-state conditions are approached. It is however shown here that this is by no means necessarily the case and is strongly dependent on the *system type*. In the first instance therefore a definition of system type is given.

3.4.1 System type

When discussing a system type it is important to note that reference is always made to the open-loop transfer function of a feedback system, on the assumption that unity feedback is applied as shown in Fig. 3.11.

74 System response using differential equations



Fig. 3.11 Unity feedback system

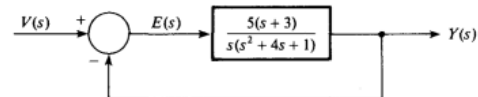
If the original block diagram representation of a system contains elements other than unity in its feedback path, the representation must be manipulated into a unity feedback system, by means of the techniques introduced in Chapter 2, in order to obtain $G_0(s)$, the open-loop transfer function – this is shown in detail in the following section.

The transfer function definition given in the previous chapter is now further generalized, by defining $G_0(s)$ as:

$$G_0(s) = \frac{K(s^m + b_{m-1}s^{m-1} + \dots + b_1s + b_0)}{s^q(s^n + a_{n-1}s^{n-1} + \dots + a_1s + a_0)} \quad (3.4.1)$$

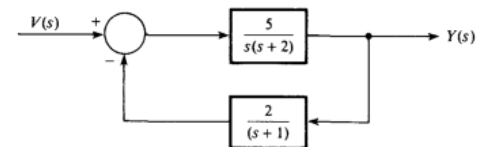
where q , m and n are integers, $a_0, b_0 \neq 0$ and $\{q, m, n \geq 0\}$. In many cases it is possible that the denominator and numerator polynomials of order n and m respectively can be factorized, these roots being termed poles and zeros of the transfer function, respectively. K is a gain constant and $q = 0, 1, 2, \dots$ denotes the type of the system, e.g. a type 1 system is achieved when $q = 1$. Note that the order of the system, usually given by the order of the denominator, is then $n + q$, e.g. if $n = 2$ and $q = 1$, this would signify a third-order system. It is worth noting that it is usual for $n \geq m$, and this is the case considered here.

Example 3.4.1



This is a type 1, third-order system.

Example 3.4.2



This is a type 0, third-order system.

The second of these examples, i.e. Example 3.4.2, is important in that it categorizes the error considered in this section always to be the difference between the desired and actual outputs $E(s) = V(s) - Y(s)$. In certain other texts, e.g. Kuo (1982), the error is considered to be the difference between the desired output $V(s)$ and the value $H(s)Y(s)$ – it is imperative that the reader is aware of this point. Only when $H(s)$ is unity are the two approaches identical.

Consider the closed-loop equation obtained from the unity feedback system shown in Fig. 3.11; this is given by

$$Y(s) = \frac{G_0(s)}{1 + G_0(s)} V(s) \quad (3.4.2)$$

For a general feedback system, as shown in Fig. 3.12, the closed-loop equation is, however:

$$Y(s) = \frac{G(s)}{1 + G(s)H(s)} V(s) \quad (3.4.3)$$

The equivalent open-loop transfer function $G_0(s)$ can therefore be found from $G(s)$ and $H(s)$ by means of the equation, obtained by connecting (3.4.2) and (3.4.3)

$$G_0(s) = \frac{G(s)}{1 + G(s)[H(s) - 1]} \quad (3.4.4)$$

where it can be noted that if $H(s) = 1$, the unity feedback case, the transfer function $G_0(s) = G(s)$.

3.4.2 Steady-state errors

Consider the unity feedback closed-loop system in Fig. 3.11, in which the error $E(s)$, is found to be

$$E(s) = \frac{1}{1 + G_0(s)} V(s) \quad (3.4.5)$$

such that the error between desired and actual output signals depends on the input signal $V(s)$, i.e. the desired reference output signal, and also on the open-loop transfer function $G_0(s)$ whether this is obtained directly or via equation (3.4.4).

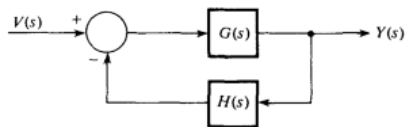


Fig. 3.12 General feedback system

The total error can be regarded as consisting of two parts; firstly the steady-state error can be written as

$$e(t) |_{t \rightarrow \infty} = sE(s) |_{s \rightarrow 0} \quad (3.4.6)$$

in which $sE(s)$ has no denominator root whose real part is positive, and where (3.4.6) is merely the Laplace transform final value theorem result. The remainder of the total error is then the transient error and this in fact links directly with the transient system response discussed in Section 3.3; it will therefore only be considered briefly here:

$$\text{Transient error} = e(t) - e(t) |_{t \rightarrow \infty} \quad (3.4.7)$$

Steady-state step error

If a unit step input is applied to a unity feedback system, we have that

$$V(s) = \frac{1}{s}$$

The steady-state error following a unit step input is then

$$sE(s) |_{s \rightarrow 0} = \left[\frac{1}{1 + G_0(s)} \right] |_{s \rightarrow 0} = \frac{1}{1 + [G_0(s)] |_{s \rightarrow 0}} \quad (3.4.8)$$

which is obtained from (3.4.5) and (3.4.6).

If $G_0(s)$ has any common s factors in its denominator the steady-state error to a step input will be zero, i.e.

$$e(t) |_{t \rightarrow \infty} = \frac{1}{1 + \infty} = 0$$

for $q > 0$, i.e. for a type > 0 system.

For a type 0 transfer function, however, the term

$$[G_0(s)] |_{s \rightarrow 0} = K_s, \text{ a constant}$$

Example 3.4.3

Consider the system

$$G_0(s) = \frac{3}{(s + 2)}$$

The steady-state step error is given as

$$sE(s) |_{s \rightarrow 0} = \frac{1}{1 + 1.5} = 0.4$$

where

$$G_0(s) \Big|_{s \rightarrow 0} = \left[\frac{3}{s+2} \right] \Big|_{s \rightarrow 0} = K_s = 1.5$$

It is important to note that if the step input is of a magnitude other than unity, the magnitude of the type 0 steady-state step error will be different from that calculated for a unit step input.

For a type 0 system, the steady-state step error can be made smaller by increasing K_s , which can be done by increasing the gain K , e.g. consider the system in Example 3.4.3, but with

$$G_0(s) = \frac{8}{(s+2)}$$

This would make

$$K_s = 4$$

and therefore

$$sE(s) \Big|_{s \rightarrow 0} = \frac{1}{1+4} = 0.2$$

i.e. a smaller steady-state error has been achieved.

Steady-state ramp error

If a unit ramp input is applied to a unity feedback system, we have that

$$V(s) = \frac{1}{s^2}$$

The steady-state error following a unit ramp input is therefore

$$sE(s) \Big|_{s \rightarrow 0} = \left[\frac{1}{1+G_0(s)} \cdot \frac{1}{s} \right] \Big|_{s \rightarrow 0} = \frac{1}{[sG_0(s)]} \Big|_{s \rightarrow 0}$$

For a type 0 system it is observed that

$$[sG_0(s)] \Big|_{s \rightarrow 0} = 0$$

such that the steady-state ramp error is ∞ .

For a type 1 system however, we have

$$[sG_0(s)] \Big|_{s \rightarrow 0} = K_r, \text{ a constant}$$

such that the steady-state ramp error is $1/K_r$.

For a type > 1 system though

$$[sG_0(s)] \Big|_{s \rightarrow 0} = \infty$$

such that the steady-state ramp error is 0.

It can be noted that if the ramp input is of a magnitude other than unity, the magnitude of the type 1 steady-state ramp error will be different from that calculated for a unit ramp input.

Steady-state parabolic (acceleration) function

If a unit parabolic input is applied to a unity feedback system, we have that

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

The steady-state error following a unit parabolic input is therefore

$$sE(s) \Big|_{s \rightarrow 0} = \left[\frac{1}{1+G_0(s)} \cdot \frac{1}{s^2} \right] \Big|_{s \rightarrow 0} = \frac{1}{[s^2 G_0(s)]} \Big|_{s \rightarrow 0}$$

For a type < 2 system it is observed that

$$[sG_0(s)] \Big|_{s \rightarrow 0} = 0$$

such that the steady-state parabolic error is ∞ .

For a type 2 system however, we have

$$[sG_0(s)] \Big|_{s \rightarrow 0} = K_a, \text{ a constant}$$

such that the steady-state parabolic error is $1/K_a$.

For a type > 2 system though

$$[sG_0(s)] \Big|_{s \rightarrow 0} = \infty$$

such that the steady-state parabolic error is 0.

If the parabolic input is of a magnitude other than unity, the magnitude of the type 2 steady-state parabolic error will be different from that calculated for a unit parabolic input.

The analysis carried out for the three classes of input can in fact be continued further if desired; however, a halt will be called here by summarizing the results obtained in Table 3.1.

In this section we have been concerned primarily with the steady-state error obtained from different system types in response to certain test inputs. In the following section the remaining part of the total error, namely the transient error, is considered.

Table 3.1 Steady-state errors

System type	Unit step input	Unit ramp input	Unit parabolic input
0	$\frac{1}{1+K_s}$	∞	∞
1	0	$\frac{1}{K_r}$	∞
2	0	0	$\frac{1}{K_a}$
3	0	0	0

3.4.3 Transient errors

The transient response of a system was considered at length in Section 3.3, and as the results in this section are closely related, the analysis is not carried out to such a great depth here.

The transient error is the error which remains when the steady-state error is subtracted from the total error. If the total error between desired and actual outputs is

$$e(t) = v(t) - y(t)$$

then the transient error can be written, in the s -domain, as

$$\text{transient error} = E(s) - [sE(s)]_{s \rightarrow 0} (s) \quad (3.4.9)$$

Transient step error

For a type 0 system, the steady-state error following a unit step input is given by

$$sE(s) \big|_{s \rightarrow 0} = \frac{1}{1+K_s} = e(t) \big|_{t \rightarrow \infty}$$

with the step input transform $V(s) = 1/s$ applied. The transient error of a type 0 system in response to a step input is therefore

$$\text{transient error} = \frac{1}{1+G_0(s)} \cdot \frac{1}{s} - \frac{1}{1+K_s} \cdot \frac{1}{s}$$

in which the transform of the steady-state error has been taken, as shown in (3.4.9). The steady-state error $1/(1+K_s)$ is simply a scalar value; this means that the denominator characteristics of the transient error transfer function are identical to those of the closed-loop transient response, see Problem 3.4. It follows then that the essential time

80 System response using differential equations

domain characteristics of the type 0 system transient error following a step input will also be those of the transient response to a step input.

For a type > 0 system, the steady-state error following a unit step input is zero. The transient error is therefore also the total error, i.e.

$$\text{transient error} = \frac{1}{1+G_0(s)} \cdot \frac{1}{s}$$

and as the closed-loop response to a step input is given by

$$Y(s) = \frac{G_0(s)}{1+G_0(s)} \cdot \frac{1}{s}$$

it follows that as the denominator of the transient error is also that of the output response, the characteristics of the two responses will be essentially the same.

Transient ramp error

For a type 0 system the steady-state error following a unit ramp input is ∞ . The transient error in such a case is therefore only of academic interest and is not really defined well in any useful sense, although it can be regarded as a special case of the following.

For a type 1 system the steady-state error following a unit ramp input is $1/K_r$. The transient error is then, by means of (3.4.9)

$$\text{transient error} = \frac{1}{1+G_0(s)} \cdot \frac{1}{s^2} - \frac{1}{K_r} \cdot \frac{1}{s}$$

where $K_r \rightarrow 0$ for a type 0 system.

For a type > 1 system the steady-state error following a unit ramp input is zero. The transient error is therefore also the total error, i.e.

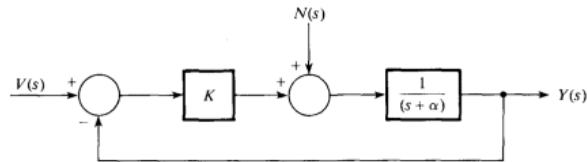
$$\text{transient error} = \frac{1}{1+G_0(s)} \cdot \frac{1}{s^2}$$

A similar analysis can be carried out for a parabolic input.

In conclusion, it must be remembered that the transient error behavior of a system depends on a denominator which is the same as that for a closed-loop transform of the same system, the same input signal being assumed for both cases.

3.4.4 Disturbances

No matter how well the system design and construction phases are carried out and how much protection is introduced, a certain amount of interference caused by disturbances will affect any control system. In some cases it is negligible with respect to the noise-free

Fig. 3.13 Unity feedback system with disturbance, $N(s)$

(deterministic) part of the system operation and need not be considered further. For a large number of systems, though, it presents a major problem and can result in the total objective of the control design being to reduce the effect of noise on the output signal to an acceptable level.

It is usual for disturbances to be regarded as either (a), a steady nonzero value – these are called load disturbances or offsets – or (b), a random noise signal, which may be corrupted. Disturbances are however considered in a fairly general way in this section and designs particular to either (a) or (b) are not discussed.

Consider a feedback control system, with a selectable gain K , which is affected by the disturbance term $N(s)$, as shown in Fig. 3.13.

The closed-loop transfer function for this system is

$$Y(s) = \frac{K}{(s + \alpha + K)} V(s) + \frac{1}{(s + \alpha + K)} \cdot N(s) \quad (3.4.10)$$

which in fact consists of two separate transfer functions, one which relates the signal $V(s)$ to the output $Y(s)$ and the other which relates the noise $N(s)$ to the output $Y(s)$. The total output value is therefore a summation of the desirable signal with the undesirable noise.

Assume the desired input $V(s)$ is a step of magnitude V and the disturbance $N(s)$ is a load disturbance of magnitude N . We would hope that in the steady state the output signal would be equal to the desired value V . By the Laplace transform final value theorem, the steady-state output value is in fact given by:

$$sY(s)|_{s \rightarrow 0} = \left[\frac{KV}{(s + \alpha + K)} + \frac{N}{(s + \alpha + K)} \right] \Big|_{s \rightarrow 0}$$

Hence, if we make K a very large value this will satisfy both the steady-state final value $y = V$ and noise reduction problems. However the $(s + \alpha + K)$ denominator term means that in the time domain the output signal will tend to its final value $\approx V$ by means of an exponential increase, the time constant of which is $\tau = 1/(\alpha + K)$. In practice such a rapid response may not be possible and the resulting large signal at the output of the gain K may not be achievable.

One of the main reasons for employing a feedback loop in a control system is to reduce the effects of noise on the output signal, and it has been shown here how this is possible merely through the selection of an appropriate gain value, which at the same time reduces the steady-state step error in the example given. Increasing the gain can

though cause problems, something which is easily seen if the first-order system example in Fig. 3.13 is replaced by a second-order system, for which an increase in K would result in oscillatory behavior such that for a very high value of K output signal oscillations would be wild and would take some time to settle down.

3.4.5 Sensitivity

When designing a controller for a system it is generally assumed that both the system model itself and the resulting controller description are accurate, exact representations of their real-world physical counterparts. This assumption ignores not only the errors inherent in the modeling and construction stages but also measurement problems, and the fact that either the system itself or its environment will have altered to some extent, since the system was modeled. One of the main reasons behind the employment of adaptive control schemes is the need to vary the control action in accordance with variations in the system characteristics. However, it is by no means always the case that an adaptive controller is necessary or even particularly desirable.

It is important to discover how sensitive system performance is to certain variations or inaccuracies in the parameters used to describe the system. For instance, if it is found that when one particular parameter varies by a few percent from its nominal value the output performance is degraded considerably, then it is sensible to redesign the controller in order to reduce the dependence of output performance on the accuracy of that parameter.

It is shown here how the use of feedback in the control of a system, reduces the sensitivity of that system to variations in the parameters which represent the system's characteristics. When investigating the sensitivity of a system though, the particular system performance that is being measured must be specified as must the parameter(s) which the sensitivity is considered in terms of. For simplicity, the following discussion restricts itself to the sensitivity of the output signal in respect of variations in the system transfer function. This can be summarized by means of the sensitivity function, S_G^Y , defined as

$$S_G^Y = \frac{\partial Y(s)/Y(s)}{\partial G(s)/G(s)} \quad (3.4.11)$$

which is simply the fractional change in output divided by the fractional change in system transfer function.

For an open-loop system, we have that the transfer function relating the input signal transform, $U(s)$, to the output signal transform, $Y(s)$, results in an equation

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

from which it follows that for an invariant input signal $U(s)$

$$\frac{\partial Y(s)}{\partial G(s)} = U(s) \quad (3.4.12)$$

such that sensitivity function becomes

$$S_G^Y = \frac{\partial Y(s)/Y(s)}{\partial G(s)/G(s)} = \frac{G(s)U(s)}{G(s)U(s)} = 1 \quad (3.4.13)$$

which means that the output varies in a directly proportional relationship to variations in the open-loop transfer function – or, looked at in another way, if the system transfer function is only accurate to $\pm 10\%$ then so is the system output.

Consider now a closed-loop system, as shown in Fig. 3.12, with the desired reference input transform $V(s)$ and feedback path $H(s)$ considered as invariant. The original closed-loop transfer function is given as

$$Y(s) = \frac{G(s)}{1 + G(s)H(s)} V(s)$$

such that

$$\frac{\partial Y(s)}{\partial G(s)} = \frac{V(s)}{1 + G(s)H(s)} - \frac{H(s)G(s)V(s)}{[1 + G(s)H(s)]^2} = \frac{V(s)}{[1 + G(s)H(s)]^2} \quad (3.4.14)$$

and the sensitivity function becomes

$$S_G^Y = \frac{\partial Y(s)}{\partial G(s)} \cdot \frac{G(s)}{Y(s)} = \frac{V(s)}{[1 + G(s)H(s)]^2} \cdot \frac{[1 + G(s)H(s)]}{V(s)}$$

or

$$S_G^Y = \frac{1}{[1 + G(s)H(s)]} \quad (3.4.15)$$

Hence (3.4.15) shows that the sensitivity of a system, when in closed-loop, is reduced from its open-loop value by a factor $1/(1 + G(s)H(s))$. The selection of feedback $H(s)$ is therefore very important in determining the sensitivity of a closed-loop system to variations in the system transfer function.

If it is the case that the system transfer function is invariant, but the feedback function $H(s)$ varies, we have that

$$\frac{\partial Y(s)}{\partial H(s)} = - \left[\frac{G(s)}{1 + G(s)H(s)} \right]^2 V(s) \quad (3.4.16)$$

with a sensitivity function, dependent on variations in $H(s)$ rather than $G(s)$,

$$S_H^Y = \frac{\partial Y(s)}{\partial H(s)} \cdot \frac{H(s)}{Y(s)} = \frac{-G(s)H(s)}{[1 + G(s)H(s)]} \quad (3.4.17)$$

This means that when $|G(s)H(s)| \gg 1$, the sensitivity function magnitude is equal to 1 and thus the accuracy of the output will be directly dependent on the accuracy of the feedback path, while for the same conditions, from (3.4.15), it will be only weakly linked with the accuracy of the system transfer function. Although this places much more importance on the accuracy of the feedback path, it is usually a desirable feature in that precision and modifications can be much more readily built in to $H(s)$ rather than $G(s)$.

3.5 Servomechanisms

In order to draw together the motor/generator analysis carried out in the previous chapter with the differential system equation properties of damping and steady-state error introduced in this chapter, a position control system, commonly referred to as a servomechanism, is considered here, rather in the form of an extended example.

The reference input signal $V(s)$ and sensed output position $Y(s)$ are provided by rotational potentiometers with gains K_v and K_y respectively. The latter signal is subtracted from the former at a summing junction in order to obtain the error, $E(s)$, between the amplified desired and actual output signals. This error is then itself amplified in order to provide the armature current, $I_a(s)$, for an armature controlled motor, the output of which is the output position $Y(s)$, as shown in Fig. 3.14. The motor is assumed to have a constant field current and also to contain a gearbox placed between the motor and the load. Before considering the overall input-output servomechanism transfer function, the effect of the gearbox, regarded as being part of the motor in Fig. 3.14, will be discussed in detail.

3.5.1 Gears

The block described in Fig. 3.14 as the 'motor' is expanded in Fig. 3.15 to introduce the idea of a gear ratio n , which means that n revolutions of the motor produce 1 revolution of the load, i.e. $\omega_m = n\omega_l$.

The motor impedance is $Z_m(s) = B_m + J_ms$, in which B_m is the resistance of the motor and J_m is the motor inertia. Also the load impedance is $Z_l(s) = B_l + J_ls$, in which B_l is the resistance of the load and J_l is the load inertia.

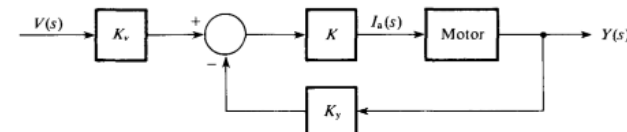


Fig. 3.14 Position control servomechanism

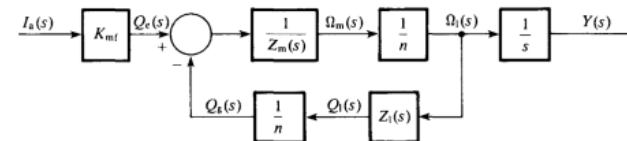


Fig. 3.15 Servomotor in open-loop

$\Omega_m(s)$ is the motor speed input to the gearbox and $\Omega_l(s)$ is the load speed output from the gearbox, also $Q_l(s)$ is the load torque output from the gearbox. Therefore the torque input to the gearbox $Q_g(s)$ must be such that for an ideal, lossless gearbox, power in = power out, that is

$$Q_g(s)\Omega_m(s) = Q_l(s)\Omega_l(s) \quad (3.5.1)$$

or

$$Q_g(s) = \frac{Q_l(s)}{n} \quad (3.5.2)$$

The total torque obtained from the motor, $Q_e(s)$ is in fact split up between that used on the motor and that input to the gearbox, i.e.

$$Q_e(s) = Z_m(s) \cdot \Omega_m(s) + Q_g(s) \quad (3.5.3)$$

or by substitution

$$Q_e(s) = \left[Z_m(s) + \frac{Z_l(s)}{n^2} \right] \Omega_m(s) \quad (3.5.4)$$

The effect of the gearbox is thus to reflect back the load impedance to combine with that of the motor, subject to the square of the gear ratio. Hence, the larger the gear ratio, the smaller the reflections.

This can also be regarded as

$$Q_e(s) = \left[nJ_m + \frac{J_l}{n} \right] \Omega_l(s) + \left[nB_m + \frac{B_l}{n} \right] \Omega_l(s) \quad (3.5.5)$$

or

$$Q_e(s) = Q_a(s) + Q_s(s)$$

in which

$$Q_a(s) = \left[nJ_m + \frac{J_l}{n} \right] s \Omega_l(s) \quad (3.5.6)$$

is the torque required in order to accelerate the load, whereas $Q_s(s)$ is that required in order to keep the speed of the load constant.

When selecting the gear ratio n for a particular servomechanism, a criterion often employed is that of obtaining n such that $Q_a(s)$ is a minimum, and this value of n is then referred to as the ideal gear ratio.

Now the torque required to accelerate the load, $Q_a(s)$, when differentiated gives

$$\frac{dQ_a(s)}{dn} = J_m - \frac{J_l}{n^2}$$

and this must be equal to zero for a minimum of $Q_a(s)$, hence

$$n = \sqrt{\frac{J_l}{J_m}} \quad (3.5.7)$$

It can be noted that:

- In the selection of n , to minimize $Q_a(s)$, the torque $Q_s(s)$ is disregarded completely;
- It is easy to show, by taking the second differential of $Q_a(s)$ with respect to n , that the value of n given in (3.5.7) results in a minimum value of $Q_a(s)$ rather than a maximum.

3.5.2 Servomechanism transient response

The relationship between torque $Q_e(s)$ and load speed $\Omega_l(s)$ can be viewed as that in the servomotor block diagram, Fig. 3.15, or in terms of the relationships (3.4.4) and (3.5.5) which reduce the servomotor block diagram to that in Fig. 3.16, where the impedance of the load is referred to that of the motor, such that

$$B_r = B_m + B_l/n^2 \quad (3.5.8)$$

and

$$J_r = J_m + J_l/n^2$$

By cascading the blocks shown in Fig. 3.16, we have that

$$Y(s) = \frac{K_{mf}}{ns(B_r + J_r s)} \cdot I_a(s) \quad (3.5.9)$$

and on substitution for $I_a(s)$, shown in Fig. 3.14 to be

$$I_a(s) = K[K_r V(s) - K_y Y(s)] \quad (3.5.10)$$

this results in an overall closed-loop equation relating desired output value (reference input) $V(s)$ to actual output $Y(s)$ of

$$Y(s) = \frac{K_{mf} K K_y}{ns(B_r + J_r s) + K_{mf} K K_y} \cdot V(s)$$

or

$$Y(s) = \frac{K_{mf} K K_y J_r n}{s^2 + \frac{B_r}{J_r} s + \frac{K_{mf} K K_y}{J_r n}} \cdot V(s) \quad (3.5.11)$$

By referring to the transient response forms considered in Section 3.3 it is possible to set the transient response of the servomechanism considered, by an appropriate choice of



Fig. 3.16 Servomotor with reflected load impedance

the potentiometer/amplifier gains. When $V(s)$ is a step input, i.e. when a desired output position value is asked for, the transient conditions which occur, will be those selected. Consider for example the case of ideal damping discussed in Section 3.3.3. The denominator of (3.3.12) is of the form

$$(s + \alpha)^2 + \alpha^2$$

and if the servomechanism control parameters are to be chosen with this in mind it is required that the denominator of (3.5.11) should also be of this form, i.e.

$$(s + \alpha)^2 + \alpha^2 = s^2 + 2\alpha s + 2\alpha^2$$

such that from (3.5.11)

$$2\alpha = \frac{B_t}{J_t}$$

and

$$2\alpha^2 = \frac{K_{mf} K K_y}{J_t n} = \frac{1}{2} \left[\frac{B_t}{J_t} \right]^2$$

Assuming B_t , J_t , n , K_{mf} and K to have been previously set, we therefore must set the feedback potentiometer gain K_y to

$$K_y = \frac{B_t^2 n}{2 J_t K_{mf} K}$$

Steady-state conditions following a unit step input, $V(s) = 1/s$, can be investigated by means of the final value theorem, whereby from (3.5.11)

$$\lim_{t \rightarrow \infty} sY(s) \Big|_{s=0} = \frac{K_v}{K_y} = \lim_{t \rightarrow \infty} y(t) \quad (3.5.12)$$

such that the output $y(t)$ will follow its required value, $v(t)$, in the steady-state, as long as the input potentiometer gain, K_v , is equal to that of the feedback potentiometer, K_y .

3.6 Summary

It was shown in this chapter how some fairly simple input signals can be used to test the response of a system and hence to investigate some of the fundamental properties of that system. Following the application of an input signal, the system response can be considered to consist of two separate parts, the steady-state response which remains after any initial response fluctuations have died down, i.e. as $t \rightarrow \infty$, and the transient response which is the name given to those initial response fluctuations and although these can persist for quite some time after the input signal is applied, it is considered that they die down to zero as $t \rightarrow \infty$.

The analysis of system steady-state and transient responses carried out made use of the

fact that once a system transfer function is obtained in terms of the Laplace operator s , on the application of an input signal, also considered in terms of its s -domain representation, it is not necessary to take inverse transforms in order to investigate the time domain properties of the response. Both the steady-state response and transient response can be discussed with reference to s -domain operators only, via the final value theorem and in terms of the degree of damping respectively.

Some of the main reasons for control system error were considered in Section 3.4. Firstly, the importance of system type and the input applied were shown in terms of their effect on the steady-state error between the desired and actual output signals. Secondly, some of the useful properties obtained from the employment of feedback in a control system were highlighted by showing how disturbance corruption and performance sensitivity can be reduced to an acceptable level simply by a suitable selection of feedback terms.

Many of the fundamental performance requirements in the design of a control system have been introduced in this chapter, and these are made use of in the chapters which follow.

Problems

- 3.1 A unity feedback closed-loop system contains an open-loop transfer function

$$G(s) = \frac{9}{s(s+5)}$$

A factor $D(s) = Ks + 1$ is cascaded with $G(s)$, where K is an adjustable gain, as shown in Fig. 3.17.



Fig. 3.17

- Find the value of K necessary to provide a critically damped closed-loop system.
 - Find the range of values of K which will provide acceptable damping for the closed-loop system.
 - What is the magnitude of overshoot which follows a unit step input, when $K = 0.1$?
- 3.2 A control system with unity feedback has an open-loop transfer function:

$$G(s) = \frac{K(s+12)}{s^2+25}$$

where K is a scalar gain.

Calculate the closed-loop denominator and find the range of values for K over which the closed-loop poles are real. With $K = 24$, find the step response of the closed-loop system and evaluate the steady-state error which follows a step response.

- 3.3 For the control system shown in Fig. 3.18, consider how derivative feedback, with gain K , can be employed to improve the transient performance.

Find the value of K for which the transient response is ideal and the steady-state and transient step errors in this case.

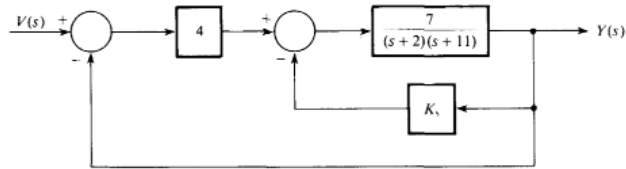


Fig. 3.18

- 3.4 Show that the denominator of a type 0 system's transient error following a unit step input is also the denominator of the system's closed-loop transient response. By using the open-loop transfer function:

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

show that (a) the steady-state error to a step input is $K_s = K/2$, and (b) a scalar multiple in the denominator does not affect the answer.

- 3.5 Given an open-loop, type 1 system:

$$G(s) = \frac{K}{s(s+\alpha)}$$

show that if $K = 100$ and $\alpha = 10$, this will result in a transient step error with damping ratio $\zeta = 0.5$ and natural frequency $\omega_n = 10$. Also, show that the steady-state ramp error which is produced by such a choice is 10%.

- 3.6 Consider the unity feedback closed-loop system with an open-loop transfer function:

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

Find the sensitivity function S_K^Y which relates the sensitivity of the output to variations in the gain K .

Show how this function is modified if a feedback gain of $H(s) = 200$ is employed, rather than $H(s) = 1$ in the unity feedback case.

90 System response using differential equations

- 3.7 For the feedback control system with open-loop transfer function

$$G(s) = \frac{K}{s(s+20)}$$

and feedback $H(s) = K_1s + 1$, find the gain value K and amount of derivative feedback K_1 for which the system has ideal transient performance and a ramp error of 5%.

- 3.8 For a unity feedback system with open-loop transfer function

$$G(s) = \frac{K}{s(\beta s + 1)}$$

show that for the closed-loop system, the damping ratio is

$$\zeta = \frac{1}{2\sqrt{K\beta}}$$

and the natural frequency

$$\omega_n = \sqrt{\frac{K}{\beta}}$$

- 3.9 A unity feedback closed-loop system contains an open-loop transfer function:

$$G(s) = \frac{K}{s(s+25)}$$

Find the value of K which will provide a steady-state ramp error of 5%, i.e. $K_r = 20$. For this value of K obtain (a) the steady-state error to a unit step input, and the transient errors produced by (b) a unit step input, and (c) a unit ramp input.

- 3.10 Find the steady-state parabolic error for the unity feedback closed-loop system whose open-loop transfer function is

$$G(s) = \frac{2.5}{s^2(2s+1)}$$

- 3.11 Find the steady-state step and ramp errors for the unity feedback closed-loop system whose open-loop transfer function is

$$G(s) = \frac{s+4}{s(s+5)}$$

- 3.12 Consider the system shown in Fig. 3.19, where $N(s)$ is an unwanted noise signal and

$$G(s) = \frac{1}{s(s+1)} \quad D(s) = \frac{K(s+1)}{(s+\alpha)}$$

Determine values for K and α which will achieve a critically damped output

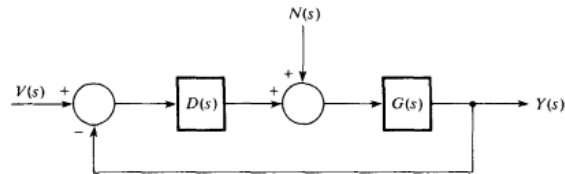


Fig. 3.19

response to a unit step input $V(s)$, and a disturbance effect on the output signal equal to $1/10$ that of a step signal $N(s)$.

- 3.13 A unity feedback closed-loop system contains an open-loop transfer function:

$$G(s) = \frac{0.03K}{s^2 + 0.35s + 0.03}$$

What value of K will result in:

- (a) Steady-state step error = 0.02?
 (b) Ideally damped transient step error?

For $K = 20$, what are the steady-state and transient errors if a unit step input signal is applied?

- 3.14 Consider the system diagram shown in Problem 3.12, in which $G(s) = 1/Js$ and $N(s)$ is a ramp signal of magnitude 0.5 units. It is required that the steady-state signal value is equal to 25 units $\pm 2\%$.

If

$$D(s) = \frac{K(s + \alpha)}{s}$$

find values for K , α and input signal, $V(s)$, step magnitude, in order to provide a damping ratio $\zeta = 0.5$.

- 3.15 For a particular armature current controlled servomechanism, the error between the reference position and sensed output position is amplified by a gain of K Nm/rad to achieve the total torque obtained from the motor. By reflecting back the load impedance to combine with that of the motor, the total inertia is J kgm² and the total resistance R Nm/rad/s with a gearbox ratio $n:1$.

Find the range of values for K over which acceptable damping is obtained.

- 3.16 Consider a unity feedback closed-loop system in which the open-loop transfer function is

$$G(s) = \frac{4(s + 3)K}{s(s^2 + 4s - 4)}$$

If $K = 10$, find the steady-state error to a ramp input of magnitude 0.3. If $K = 0.5$, find a compensator $D(s)$ which when connected in cascade with $G(s)$ will produce a damping ratio $\zeta = 1.75$ and a natural frequency $\omega = 1.25$ rad/s.

- 3.17 An armature current controlled servomechanism is subject to a total inertia of 4 kgm² and a total resistance of 3 Nm/rad/s with a gearbox ratio 1:1. The error between desired position and sensed output position is amplified by a gain K Nm/rad, where $0 < K < 10$. Find a value for K which results in a suitable damping.
- 3.18 A field current controlled servomechanism is subject to a total inertia of 15 kgm² and a total resistance of 2 Nm/rad/s (referred to the motor) with a gearbox ratio 10:1. Find a value for the amplifier gain K Nm/rad which results in ideal damping.
- 3.19 A field voltage controlled servomechanism is subject to a field resistance of 2 Ω and a field inductance of 4 H, with a motor resistance of 1 Nm/rad/s and a motor inertia of 2 kgm². The constant relating torque to field current is 25 Nm/A, and the load has a resistance of 10 Nm/rad/s with an inertia 15 kgm². Find the output speed following a 10 V step input.
- 3.20 An armature voltage controlled servomechanism is subject to a motor resistance of 0.2 Nm/rad/s and a motor inertia of 0.4 kgm², with a load resistance of 10 Nm/rad/s and a load inertia of 120 kgm². The motor armature has a resistance equal to 3 Ω and an inductance equal to 6 H, the field current is a constant 2.5 A and the electromechanical coupling constant is 25 Nm/A².

Find the transfer function which relates output load speed to input armature voltage, when the ratio of the gearbox is ideal.

Further reading

- Bower, J. L. and Schultheiss, P. M., *Introduction to the design of servomechanisms*, Chapman and Hall, 1958.
- Brown, G. S. and Campbell, D. P., *Principles of servomechanisms*, Chapman and Hall, 1950.
- Chestnut, H. and Mayer, R. W., *Servomechanisms and regulating system design*, Wiley (Vols 1 and 2), 1951.
- D'Azzo, J. J. and Houpis, C. H., *Linear control system analysis and design*, McGraw-Hill, 1983.
- DiStefano, J. J., Stubberud, A. R. and Williams, I. J., *Feedback and control systems*, McGraw-Hill (Schaum), 1976.
- Doebelin, E. O., *Control system principles and design*, John Wiley & Sons, 1985.
- Dorf, R. C., *Modern control systems*, 6th ed., Addison-Wesley, 1992.
- Elloy, J.-P. and Piasco, J.-M., *Classical and modern control*, Pergamon Press, 1981.
- Frank, P. M., *Introduction to system sensitivity theory*, Academic Press, 1978.
- Healey, M., *Principles of automatic control*, Hodder and Stoughton, 1981.
- Kuo, B. C., *Automatic control systems*, 3rd ed., Prentice-Hall, 1990.
- Ogata, K., *System dynamics*, Prentice-Hall, 1978.
- Oppenheim, A. V. and Willsky, A. S., *Signals and systems*, Prentice-Hall, 1983.
- Power, H. M. and Simpson, R., *Introduction to dynamics and control*, McGraw-Hill, 1978.
- Raven, F. H., *Automatic control engineering*, 2nd ed., McGraw-Hill, 1990.
- Richards, R. J., *An introduction to dynamics and control*, Longman, 1979.
- Sinha, N. K., *Control systems*, Wiley Eastern, 1994.