Principles of Measurement Systems

Principles of Measurement Systems

Fourth Edition

John P. Bentley
Emeritus Professor of Measurement Systems
University of Teesside

PEARSON
Prentice Hall

Harlow, England • London • New York • Boston • San Francisco • Toronto • Sydney • Singapore • Hong Kong
Tokyo • Seoul • Taipei • New Delhi • Cape Town • Madrid • Mexico City • Amsterdam • Munich • Paris • Milan

Pearson Education Limited
Edinburgh Gate
Harlow
Essex CM20 2JE
England

and Associated Companies throughout the world

Visit us on the World Wide Web at:
www.pearsoned.co.uk

First published 1983
Second Edition 1988
Third Edition 1995
Fourth Edition published 2005

ISBN 0 130 43028 5

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Cataloging-in-Publication Data
Bentley, John P., 1943–
Principles of measurement systems / John P. Bentley. – 4th ed.
p. cm.
Includes bibliographical references and index.
ISBN 0-13-043028-5
1. Physical instruments. 2. Physical measurements. 3. Engineering instruments. 4. Automatic control. I. Title.
QC53.B44 2005
530.8–dc22 2004044467

10 9 8 7 6 5 4 3 2 1
10 09 08 07 06 05

Typeset in 10/12pt Times by 35
Printed in Malaysia

The publisher's policy is to use paper manufactured from sustainable forests.

To Pauline, Sarah and Victoria

Contents

Preface to the fourth edition

Measurement is an essential activity in every branch of technology and science. We need to know the speed of a car, the temperature of our working environment, the flow rate of liquid in a pipe, the amount of oxygen dissolved in river water. It is important, therefore, that the study of measurement forms part of engineering and science courses in further and higher education. The aim of this book is to provide the fundamental principles of measurement which underlie these studies.

The book treats measurement as a coherent and integrated subject by presenting it as the study of measurement systems. A measurement system is an information system which presents an observer with a numerical value corresponding to the variable being measured. A given system may contain four types of element: sensing, signal conditioning, signal processing and data presentation elements.

The book is divided into three parts. *Part A (Chapters 1 to 7)* examines general systems principles. This part begins by discussing the static and dynamic characteristics that individual elements may possess and how they are used to calculate the overall system measurement error, under both steady and unsteady conditions. In later chapters, the principles of loading and two-port networks, the effects of interference and noise on system performance, reliability, maintainability and choice using economic criteria are explained. *Part B (Chapters 8 to 11)* examines the principles, characteristics and applications of typical sensing, signal conditioning, signal processing and data presentation elements in wide current use. *Part C (Chapters 12 to 19)* examines a number of specialised measurement systems which have important industrial applications. These are flow measurement systems, intrinsically safe systems, heat transfer, optical, ultrasonic, gas chromatography, data acquisition, communication and intelligent multivariable systems.

The fourth edition has been substantially extended and updated to reflect new developments in, and applications of, technology since the third edition was published in 1995. Chapter 1 has been extended to include a wider range of examples of basic measurement systems. New material on **solid state sensors** has been included in Chapter 8; this includes **resistive gas**, **electrochemical** and **Hall effect** sensors. In Chapter 9 there is now a full analysis of **operational amplifier circuits** which are used in measurement systems. The section on **frequency to digital conversion** in Chapter 10 has been expanded; there is also new material on **microcontroller** structure, software and applications. Chapter 11 has been extensively updated with new material on **digital displays**, **chart and paperless recorders** and **laser printers**. The section on **vortex flowmeters** in Chapter 12 has been extended and updated. Chapter 19 is a new chapter on **intelligent multivariable measurement systems** which concentrates on structure and modelling methods. There are around 35 additional problems in this new edition; many of these are at a basic, introductory level.

Each chapter in the book is clearly divided into sections. The topics to be covered are introduced at the beginning and reviewed in a conclusion at the end. Basic and important equations are highlighted, and a number of references are given at the end of each chapter; these should provide useful supplementary reading. The book contains about 300 line diagrams and tables and about 140 problems. At the end of the book there are answers to all the numerical problems and a comprehensive index.

This book is primarily aimed at students taking modules in measurement and instrumentation as part of degree courses in instrumentation/control, mechanical, manufacturing, electrical, electronic, chemical engineering and applied physics. Much of the material will also be helpful to lecturers and students involved in HNC/HND and foundation degree courses in technology. The book should also be useful to professional engineers and technicians engaged in solving practical measurement problems.

I would like to thank academic colleagues, industrial contacts and countless students for their helpful comments and criticism over many years. Thanks are again especially due to my wife Pauline for her constant support and help with the preparation of the manuscript.

John P. Bentley
Guisborough, December 2003

Acknowledgements

We are grateful to the following for permission to reproduce copyright material:

Figure 2.1(b) from *Repeatability and Accuracy*, Council of the Institution of Mechanical Engineers (Hayward, A.T.J., 1977); Figure 2.17(a) from Measurement of length in *Journal Institute Measurement & Control*, Vol. 12, July (Scarr, A., 1979), Table 5.1 from Systems analysis of instruments in *Journal Institute Measurement & Control*, Vol. 4, September (Finkelstein, L. and Watts, R.D., 1971), Table 7.3 from The application of reliability engineering to high integrity plant control systems in *Measurement and Control*, Vol. 18, June (Hellyer, F.G., 1985), and Figures 8.4(a) and (b) from Institute of Measurement and Control; Tables 2.3 and 2.4 from *Units of Measurement poster*, 8th edition, 1996, and Figures 15.22(a) and (b) from Wavelength encoded optical fibre sensors in *N.P.L. News*, No. 363 (Hutley, M.C., 1985), National Physical Laboratory; Figure 7.1 from The Institution of Chemical Engineers; Table 7.1 from Instrument reliability in *Instrument Science and Technology: Volume 1* (Wright, R.I., 1984), and Figure 16.14 from Medical and industrial applications of high resolution ultrasound in *Journal of Physics E: Scientific Instruments*, Vol. 18 (Payne, P.A., 1985), Institute of Physics Publishing Ltd.; Table 7.2 from The reliability of instrumentation in *Chemistry and Industry*, 6 March 1976, Professor F. Lees, Loughborough University; Table 8.2 from BS 4937: 1974 *International Thermocouple Reference Tables*, and Table 12.1 and Figure 12.7 from BS 1042: 1981 *Methods of measurement of fluid flow in closed conduits*, British Standards Institution; Figure 8.2(a) from *Instrument Transducers: An Introduction to their Performance and Design*, 2nd edition, Oxford University Press (Neubert, H.K.P., 1975); Figure 8.3(a) from *Technical Information on Two-point NTC Thermistors*, 1974, Mullard Ltd.; Table 8.4 from *Technical Data on Ion Selective Electrodes*, 1984, E.D.T. Research; Figures 8.4(b) and (c) from Thick film polymer sensors for physical variables in *Measurement and Control*, Vol. 33, No. 4, May, Institute of Measurement and Control and Professor N. White, University of Southampton (Papakostas, T.V. and White, N., 2000); Figures 8.8(a), (b) and (c) from Thick film chemical sensor array allows flexibility in specificity in *MTEC 1999, Sensor and Transducer Conference*, NEC Birmingham, Trident Exhibitions and Dr. A Cranny, University of Southampton (Jeffrey, P.D. *et al*., 1999); Figure 8.10 from Ceramics put pressure on conventional transducers in *Process Industry Journal*, June, Endress and Hauser Ltd. (Stokes, D., 1991); Figure 8.23(b) from Piezoelectric devices: a step nearer problem-free vibration measurement in *Transducer Technology*, Vol. 4, No. 1 (Purdy, D., 1981), and Figure 8.24 from IC sensors boost potential of measurement systems in *Transducer Technology*, Vol. 8, No. 4 (Noble, M., 1985), Transducer Technology; Figure 8.25(b) from *Analysis with Ion Selective Electrodes*, John Wiley and Sons Ltd. (Bailey, P.L., 1976); Figure 8.25(c) from pH facts – the glass electrode in *Kent Technical Review*, Kent Industrial Measurements Ltd., E.I.L Analytical Instruments (Thompson, W.); Figure 8.26(a) from *Electrical Engineering: Principles and Applications*, 2nd edition, reprinted by permission of Pearson Education Inc., Upper Saddle River, NJ, USA (Hambley, A.R.); Table 10.5 from Appendix A, *MCS BASIC-52 User's Manual*, Intel Corporation; Figure 11.10(c) from Instrumentation T 292 Block 6, part 2 Displays, 1986, The Open University Press; Figures 11.12(a) and (b) from Trident Displays technical literature on EL displays, Trident Microsystems Ltd. and M.S. Caddy and D. Weber; Figure 12.11(a) from Kent Process Control Ltd., Flow Products; Figure 15.10 from *Optical Fibre Communications* (Keiser, G., 1983), and Figure 15.12(b) from *Measurement Systems: Application and Design* (Doebelin, E.O., 1976), McGraw-Hill Book Co. (USA); Table 16.1 from Piezoelectric transducers in *Methods of Experimental Physics*, Vol. 19, Academic Press (O'Donnell, M., Busse, L.J. and Miller, J.G., 1981); Table 16.2 from *Ultrasonics: Methods and Applications*, Butterworth and Co. (Blitz, J., 1971); Figure 16.15 from ultrasonic image of Benjamin Stefan Morton, Nottingham City Hospital NHS Trust and Sarah Morton; Figure 17.5 from Process gas chromatography in *Talanta 1967*, Vol. 14, Pergamon Press Ltd. (Pine, C.S.F., 1967); Error Detection System in Section 18.5.2 from *Technical Information on Kent P4000 Telemetry Systems*, 1985, Kent Automation Systems Ltd.

In some instances we have been unable to trace the owners of copyright material, and we would appreciate any information that would enable us to do so.

Part A

General Principles

1 The General Measurement System

1.1 Purpose and performance of measurement systems

We begin by defining a **process** as a system which generates **information**. Examples are a chemical reactor, a jet fighter, a gas platform, a submarine, a car, a human heart, and a weather system.

Table 1.1 lists **information variables** which are commonly generated by processes: thus a car generates displacement, velocity and acceleration variables, and a chemical reactor generates temperature, pressure and composition variables.

Table 1.1 Common information/measured variables.

Acceleration	Density
Velocity	Viscosity
Displacement	Composition
Force–Weight	pH
Pressure	Humidity
Torque	Temperature
Volume	Heat/Light flux
Mass	Current
Flow rate	Voltage
Level	Power

We then define the **observer** as a person who needs this information from the process. This could be the car driver, the plant operator or the nurse.

The purpose of the **measurement system** is to link the observer to the process, as shown in Figure 1.1. Here the observer is presented with a number which is the current value of the information variable.

We can now refer to the information variable as a **measured variable**. The input to the measurement system is the **true value** of the variable; the system output is the **measured value** of the variable. In an ideal measurement system, the measured

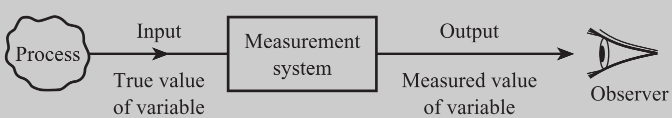


Figure 1.1 Purpose of measurement system.

value would be equal to the true value. The **accuracy** of the system can be defined as the closeness of the measured value to the true value. A perfectly accurate system is a theoretical ideal and the accuracy of a real system is quantified using **measurement system error** E, where

$$E = \text{measured value} - \text{true value}$$
$$E = \text{system output} - \text{system input}$$

Thus if the measured value of the flow rate of gas in a pipe is 11.0 m^3/h and the true value is 11.2 m^3/h, then the error $E = -0.2$ m^3/h. If the measured value of the rotational speed of an engine is 3140 rpm and the true value is 3133 rpm, then $E = +7$ rpm. Error is the main performance indicator for a measurement system. The procedures and equipment used to establish the true value of the measured variable will be explained in Chapter 2.

1.2 Structure of measurement systems

The measurement system consists of several elements or blocks. It is possible to identify four types of element, although in a given system one type of element may be missing or may occur more than once. The four types are shown in Figure 1.2 and can be defined as follows.

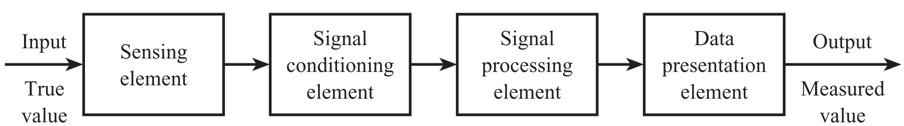


Figure 1.2 General structure of measurement system.

Sensing element

This is in contact with the process and gives an output which depends in some way on the variable to be measured. Examples are:

- Thermocouple where millivolt e.m.f. depends on temperature
- Strain gauge where resistance depends on mechanical strain
- Orifice plate where pressure drop depends on flow rate.

If there is more than one sensing element in a system, the element in contact with the process is termed the primary sensing element, the others secondary sensing elements.

Signal conditioning element

This takes the output of the sensing element and converts it into a form more suitable for further processing, usually a d.c. voltage, d.c. current or frequency signal. Examples are:

- Deflection bridge which converts an impedance change into a voltage change
- Amplifier which amplifies millivolts to volts
- Oscillator which converts an impedance change into a variable frequency voltage.

Signal processing element

This takes the output of the conditioning element and converts it into a form more suitable for presentation. Examples are:

- Analogue-to-digital converter (ADC) which converts a voltage into a digital form for input to a computer
- Computer which calculates the measured value of the variable from the incoming digital data.

Typical calculations are:

- Computation of total mass of product gas from flow rate and density data
- Integration of chromatograph peaks to give the composition of a gas stream
- Correction for sensing element non-linearity.

Data presentation element

This presents the measured value in a form which can be easily recognised by the observer. Examples are:

- Simple pointer–scale indicator
- Chart recorder
- Alphanumeric display
- Visual display unit (VDU).

1.3 Examples of measurement systems

Figure 1.3 shows some typical examples of measurement systems.

Figure 1.3(a) shows a temperature system with a thermocouple sensing element; this gives a millivolt output. Signal conditioning consists of a circuit to compensate for changes in reference junction temperature, and an amplifier. The voltage signal is converted into digital form using an analogue-to-digital converter, the computer corrects for sensor non-linearity, and the measured value is displayed on a VDU.

In Figure 1.3(b) the speed of rotation of an engine is sensed by an electromagnetic tachogenerator which gives an a.c. output signal with frequency proportional to speed. The Schmitt trigger converts the sine wave into sharp-edged pulses which are then counted over a fixed time interval. The digital count is transferred to a computer which calculates frequency and speed, and the speed is presented on a digital display.

The flow system of Figure 1.3(c) has an orifice plate sensing element; this gives a differential pressure output. The differential pressure transmitter converts this into a current signal and therefore combines both sensing and signal conditioning stages. The ADC converts the current into digital form and the computer calculates the flow rate, which is obtained as a permanent record on a chart recorder.

The weight system of Figure 1.3(d) has two sensing elements: the primary element is a cantilever which converts weight into strain; the strain gauge converts this into a change in electrical resistance and acts as a secondary sensor. There are two signal conditioning elements: the deflection bridge converts the resistance change into

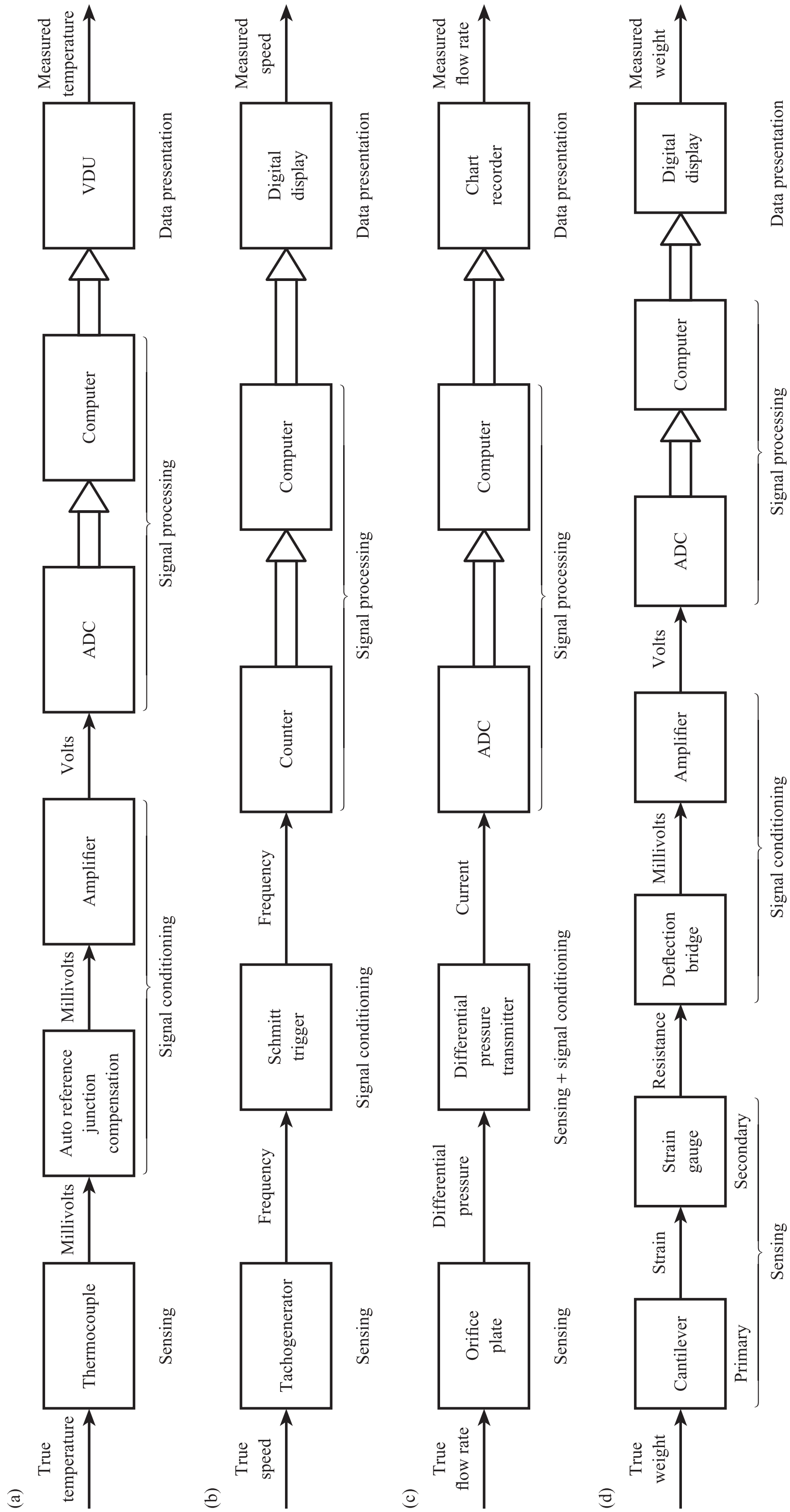


Figure 1.3 Examples of measurement systems.

millivolts and the amplifier converts millivolts into volts. The computer corrects for non-linearity in the cantilever and the weight is presented on a digital display.

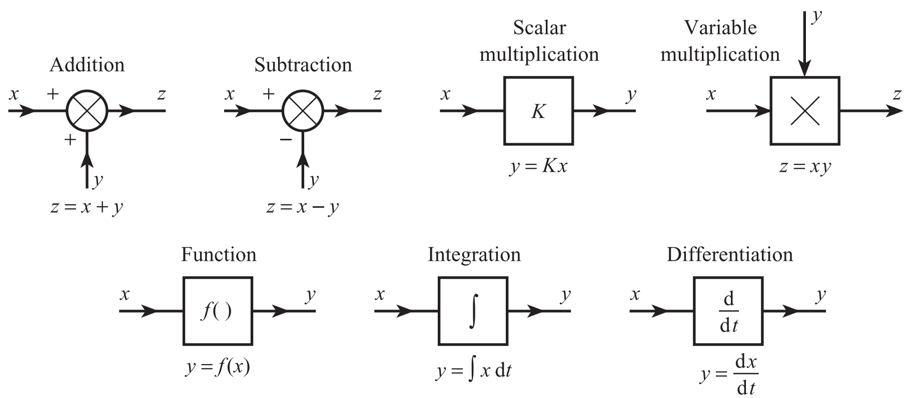
The word '**transducer**' is commonly used in connection with measurement and instrumentation. This is a manufactured package which gives an output voltage (usually) corresponding to an input variable such as pressure or acceleration. We see therefore that such a transducer may incorporate both sensing and signal conditioning elements; for example a weight transducer would incorporate the first four elements shown in Figure 1.3(d).

It is also important to note that each element in the measurement system may itself be a system made up of simpler components. Chapters 8 to 11 discuss typical examples of each type of element in common use.

1.4 Block diagram symbols

A block diagram approach is very useful in discussing the properties of elements and systems. Figure 1.4 shows the main block diagram symbols used in this book.

Figure 1.4 Block diagram symbols.



Conclusion

This chapter has defined the purpose of a measurement system and explained the importance of **system error**. It has shown that, in general, a system consists of four types of element: **sensing**, **signal conditioning**, **signal processing** and **data presentation** elements. Typical examples have been given.

2 Static Characteristics of Measurement System Elements

Figure 2.1 Meaning of element characteristics.

In the previous chapter we saw that a measurement system consists of different types of element. The following chapters discuss the characteristics that typical elements may possess and their effect on the overall performance of the system. This chapter is concerned with static or steady-state characteristics; these are the relationships which may occur between the output O and input I of an element when I is either at a constant value or changing slowly (Figure 2.1).

2.1 Systematic characteristics

Systematic characteristics are those that can be exactly quantified by mathematical or graphical means. These are distinct from statistical characteristics which cannot be exactly quantified and are discussed in Section 2.3.

Range

The input range of an element is specified by the minimum and maximum values of I, i.e. I_{MIN} to I_{MAX}. The output range is specified by the minimum and maximum values of O, i.e. O_{MIN} to O_{MAX}. Thus a pressure transducer may have an input range of 0 to 10^4 Pa and an output range of 4 to 20 mA; a thermocouple may have an input range of 100 to 250 °C and an output range of 4 to 10 mV.

Span

Span is the maximum variation in input or output, i.e. input span is $I_{MAX} - I_{MIN}$, and output span is $O_{MAX} - O_{MIN}$. Thus in the above examples the pressure transducer has an input span of 10^4 Pa and an output span of 16 mA; the thermocouple has an input span of 150 °C and an output span of 6 mV.

Ideal straight line

An element is said to be linear if corresponding values of I and O lie on a straight line. The **ideal straight line** connects the minimum point $A(I_{MIN}, O_{MIN})$ to maximum point $B(I_{MAX}, O_{MAX})$ (Figure 2.2) and therefore has the equation:

$$O - O_{\text{MIN}} = \left[\frac{O_{\text{MAX}} - O_{\text{MIN}}}{I_{\text{MAX}} - I_{\text{MIN}}}\right](I - I_{\text{MIN}}) \qquad [2.1]$$

Ideal straight line equation

$$O_{\text{IDEAL}} = KI + a \qquad [2.2]$$

where:

$$K = \text{ideal straight-line slope} = \frac{O_{\text{MAX}} - O_{\text{MIN}}}{I_{\text{MAX}} - I_{\text{MIN}}}$$

and

$$a = \text{ideal straight-line intercept} = O_{\text{MIN}} - KI_{\text{MIN}}$$

Thus the ideal straight line for the above pressure transducer is:

$$O = 1.6 \times 10^{-3} I + 4.0$$

The ideal straight line defines the ideal characteristics of an element. Non-ideal characteristics can then be quantified in terms of deviations from the ideal straight line.

Non-linearity

In many cases the straight-line relationship defined by eqn [2.2] is not obeyed and the element is said to be **non-linear**. Non-linearity can be defined (Figure 2.2) in terms of a function $N(I)$ which is the difference between actual and ideal straight-line behaviour, i.e.

$$N(I) = O(I) - (KI + a) \qquad [2.3]$$

or

$$O(I) = KI + a + N(I) \qquad [2.4]$$

Non-linearity is often quantified in terms of the maximum non-linearity $\hat{N}$ expressed as a percentage of full-scale deflection (f.s.d.), i.e. as a percentage of span. Thus:

$$\begin{array}{l}\text{Max. non-linearity as}\\ \text{a percentage of f.s.d.}\end{array} = \frac{\hat{N}}{O_{\text{MAX}} - O_{\text{MIN}}} \times 100\% \qquad [2.5]$$

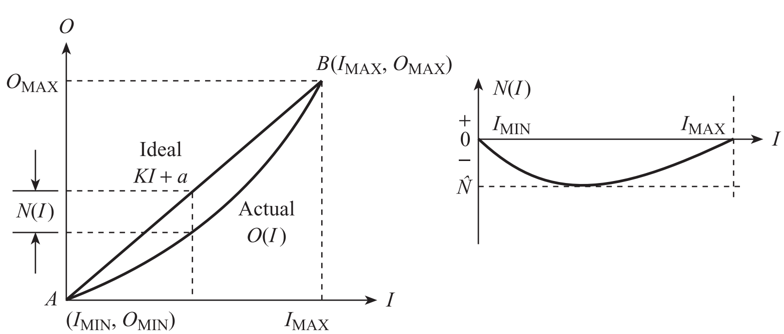


Figure 2.2 Definition of non-linearity.

As an example, consider a pressure sensor where the maximum difference between actual and ideal straight-line output values is 2 mV. If the output span is 100 mV, then the maximum percentage non-linearity is 2% of f.s.d.

In many cases $O(I)$ and therefore $N(I)$ can be expressed as a polynomial in I:

$$O(I) = a_0 + a_1 I + a_2 I^2 + \ldots + a_q I^q + \ldots + a_m I^m = \sum_{q=0}^{q=m} a_q I^q \qquad [2.6]$$

An example is the temperature variation of the thermoelectric e.m.f. at the junction of two dissimilar metals. For a copper–constantan (Type T) thermocouple junction, the first four terms in the polynomial relating e.m.f. $E(T)$, expressed in µV, and junction temperature T °C are:

$$E(T) = 38.74T + 3.319 \times 10^{-2}T^2 + 2.071 \times 10^{-4}T^3 - 2.195 \times 10^{-6}T^4 + \text{higher-order terms up to } T^8 \qquad [2.7a]$$

for the range 0 to 400 °C.[1] Since $E = 0$ µV at $T = 0$ °C and $E = 20\,869$ µV at $T = 400$ °C, the equation to the ideal straight line is:

$$E_{\text{IDEAL}} = 52.17T \qquad [2.7b]$$

and the non-linear correction function is:

$$\begin{aligned} N(T) &= E(T) - E_{\text{IDEAL}} \\ &= -13.43T + 3.319 \times 10^{-2}T^2 + 2.071 \times 10^{-4}T^3 \\ &\quad - 2.195 \times 10^{-6}T^4 + \text{higher-order terms} \end{aligned} \qquad [2.7c]$$

In some cases expressions other than polynomials are more appropriate: for example the resistance $R(T)$ ohms of a thermistor at T °C is given by:

$$R(T) = 0.04 \exp\left(\frac{3300}{T + 273}\right) \qquad [2.8]$$

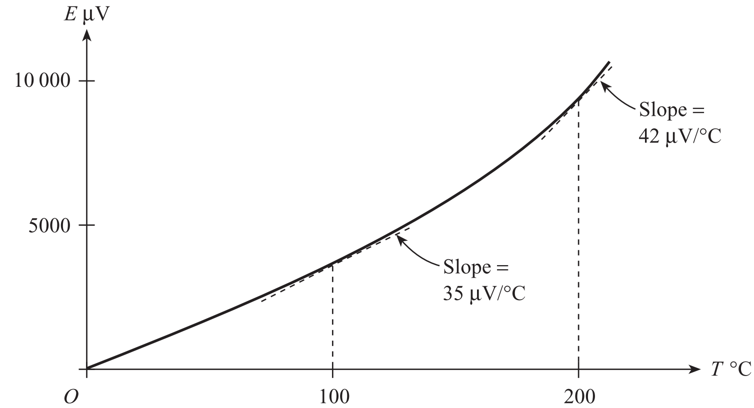
Sensitivity

This is the change ΔO in output O for unit change ΔI in input I, i.e. it is the ratio $\Delta O/\Delta I$. In the limit that ΔI tends to zero, the ratio $\Delta O/\Delta I$ tends to the **derivative** dO/dI, which is the rate of change of O with respect to I. For a linear element dO/dI is equal to the slope or gradient K of the straight line; for the above pressure transducer the sensitivity is 1.6×10^{-3} mA/Pa. For a non-linear element $dO/dI = K + dN/dI$, i.e. sensitivity is the **slope** or **gradient** of the output versus input characteristics $O(I)$. Figure 2.3 shows the e.m.f. versus temperature characteristics $E(T)$ for a Type T thermocouple (eqn [2.7a]). We see that the gradient and therefore the sensitivity vary with temperature: at 100 °C it is approximately 35 µV/°C and at 200 °C approximately 42 µV/°C.

Environmental effects

In general, the output O depends not only on the signal input I but on environmental inputs such as ambient temperature, atmospheric pressure, relative humidity, supply voltage, etc. Thus if eqn [2.4] adequately represents the behaviour of the element under 'standard' environmental conditions, e.g. 20 °C ambient temperature,

Figure 2.3
Thermocouple sensitivity.



1000 millibars atmospheric pressure, 50% RH and 10 V supply voltage, then the equation must be modified to take account of deviations in environmental conditions from 'standard'. There are two main types of environmental input.

A **modifying** input I_M causes the linear sensitivity of an element to change. K is the sensitivity at standard conditions when $I_M = 0$. If the input is changed from the standard value, then I_M is the **deviation** from standard conditions, i.e. (new value – standard value). The sensitivity changes from K to $K + K_M I_M$, where K_M is the change in sensitivity for unit change in I_M. Figure 2.4(a) shows the modifying effect of ambient temperature on a linear element.

An **interfering** input I_I causes the straight line intercept or zero bias to change. a is the zero bias at standard conditions when $I_I = 0$. If the input is changed from the standard value, then I_I is the **deviation** from standard conditions, i.e. (new value – standard value). The zero bias changes from a to $a + K_I I_I$, where K_I is the change in zero bias for unit change in I_I. Figure 2.4(b) shows the interfering effect of ambient temperature on a linear element.

K_M and K_I are referred to as environmental coupling constants or sensitivities. Thus we must now correct eqn [2.4], replacing KI with $(K + K_M I_M)I$ and replacing a with $a + K_I I_I$ to give:

$$O = KI + a + N(I) + K_M I_M I + K_I I_I \qquad [2.9]$$

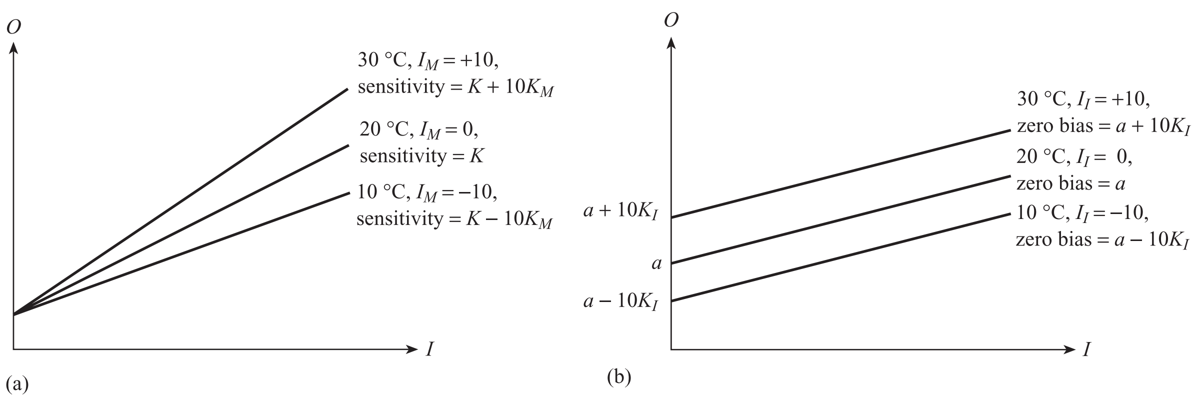
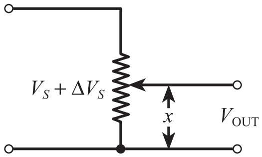


Figure 2.4 Modifying and interfering inputs.



If x is the fractional displacement, then
$V_{OUT} = (V_S + \Delta V_S)x$
$= V_S x + \Delta V_S x$

Figure 2.5

An example of a modifying input is the variation ΔV_S in the supply voltage V_S of the potentiometric displacement sensor shown in Figure 2.5. An example of an interfering input is provided by variations in the reference junction temperature T_2 of the thermocouple (see following section and Section 8.5).

Hysteresis

For a given value of I, the output O may be different depending on whether I is increasing or decreasing. Hysteresis is the difference between these two values of O (Figure 2.6), i.e.

$$\text{Hysteresis } H(I) = O(I)_{I\downarrow} - O(I)_{I\uparrow} \quad [2.10]$$

Again hysteresis is usually quantified in terms of the maximum hysteresis $\hat{H}$ expressed as a percentage of f.s.d., i.e. span. Thus:

$$\text{Maximum hysteresis as a percentage of f.s.d.} = \frac{\hat{H}}{O_{MAX} - O_{MIN}} \times 100\% \quad [2.11]$$

A simple gear system (Figure 2.7) for converting linear movement into angular rotation provides a good example of hysteresis. Due to the 'backlash' or 'play' in the gears the angular rotation θ, for a given value of x, is different depending on the direction of the linear movement.

Resolution

Some elements are characterised by the output increasing in a series of discrete steps or jumps in response to a continuous increase in input (Figure 2.8). Resolution is defined as the largest change in I that can occur without any corresponding change in O.

Figure 2.6 Hysteresis.

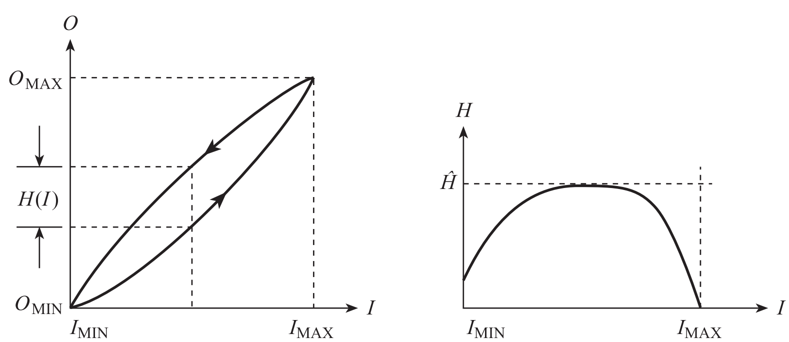
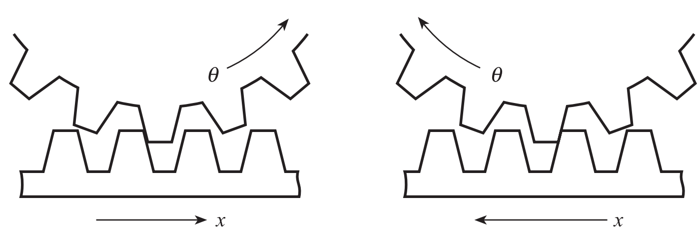


Figure 2.7 Backlash in gears.



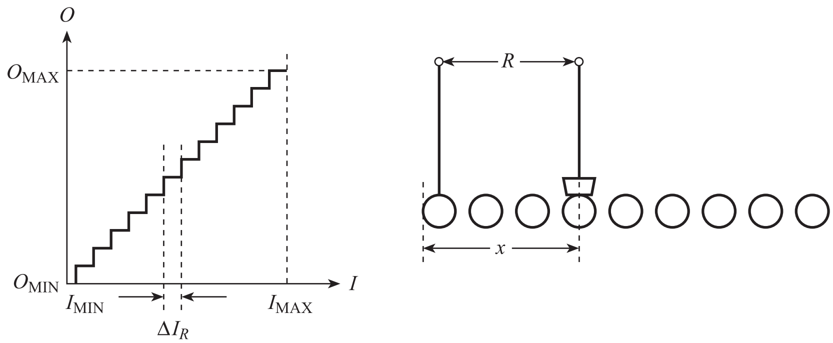


Figure 2.8 Resolution and potentiometer example.

Thus in Figure 2.8 resolution is defined in terms of the width ΔI_R of the widest step; resolution expressed as a percentage of f.s.d. is thus:

$$\frac{\Delta I_R}{I_{MAX} - I_{MIN}} \times 100\%$$

A common example is a wire-wound potentiometer (Figure 2.8); in response to a continuous increase in x the resistance R increases in a series of steps, the size of each step being equal to the resistance of a single turn. Thus the resolution of a 100 turn potentiometer is 1%. Another example is an analogue-to-digital converter (Chapter 10); here the output digital signal responds in discrete steps to a continuous increase in input voltage; the resolution is the change in voltage required to cause the output code to change by the least significant bit.

Wear and ageing

These effects can cause the characteristics of an element, e.g. K and a, to change slowly but systematically throughout its life. One example is the stiffness of a spring $k(t)$ decreasing slowly with time due to wear, i.e.

$$k(t) = k_0 - bt \qquad [2.12]$$

where k_0 is the initial stiffness and b is a constant. Another example is the constants a_1, a_2, etc. of a thermocouple, measuring the temperature of gas leaving a cracking furnace, changing systematically with time due to chemical changes in the thermocouple metals.

Error bands

Non-linearity, hysteresis and resolution effects in many modern sensors and transducers are so small that it is difficult and not worthwhile to exactly quantify each individual effect. In these cases the manufacturer defines the performance of the element in terms of error bands (Figure 2.9). Here the manufacturer states that for any value of I, the output O will be within $\pm h$ of the ideal straight-line value O_{IDEAL}. Here an exact or systematic statement of performance is replaced by a statistical statement in terms of a probability density function $p(O)$. In general a probability density function $p(x)$ is defined so that the integral $\int_{x_1}^{x_2} p(x)\,dx$ (equal to the area under the curve in Figure 2.10 between x_1 and x_2) is the probability P_{x_1,x_2} of x lying between x_1 and x_2 (Section 6.2). In this case the probability density function is rectangular (Figure 2.9), i.e.

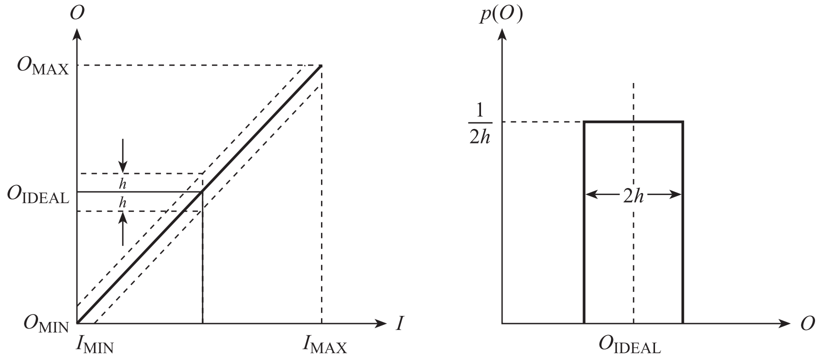


Figure 2.9 Error bands and rectangular probability density function.

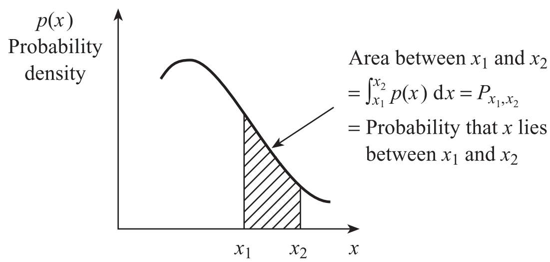


Figure 2.10 Probability density function.

$$p(O)\begin{cases} = \dfrac{1}{2h} & O_{IDEAL} - h \le O \le O_{IDEAL} + h \\ = 0 & O > O_{IDEAL} + h \\ = 0 & O_{IDEAL} - h > O \end{cases} \qquad [2.13]$$

We note that the area of the rectangle is equal to unity: this is the probability of O lying between $O_{IDEAL} - h$ and $O_{IDEAL} + h$.

2.2 Generalised model of a system element

If hysteresis and resolution effects are not present in an element but environmental and non-linear effects are, then the steady-state output O of the element is in general given by eqn [2.9], i.e.:

$$O = KI + a + N(I) + K_M I_M I + K_I I_I \qquad [2.9]$$

Figure 2.11 shows this equation in block diagram form to represent the **static** characteristics of an element. For completeness the diagram also shows the transfer function $G(s)$, which represents the **dynamic** characteristics of the element. The meaning of transfer function will be explained in Chapter 4 where the form of $G(s)$ for different elements will be derived.

Examples of this general model are shown in Figure 2.12(a), (b) and (c), which summarise the static and dynamic characteristics of a strain gauge, thermocouple and accelerometer respectively.

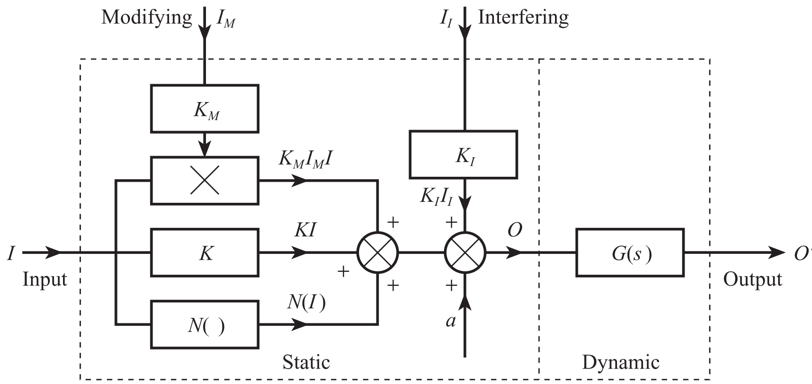


Figure 2.11 General model of element.

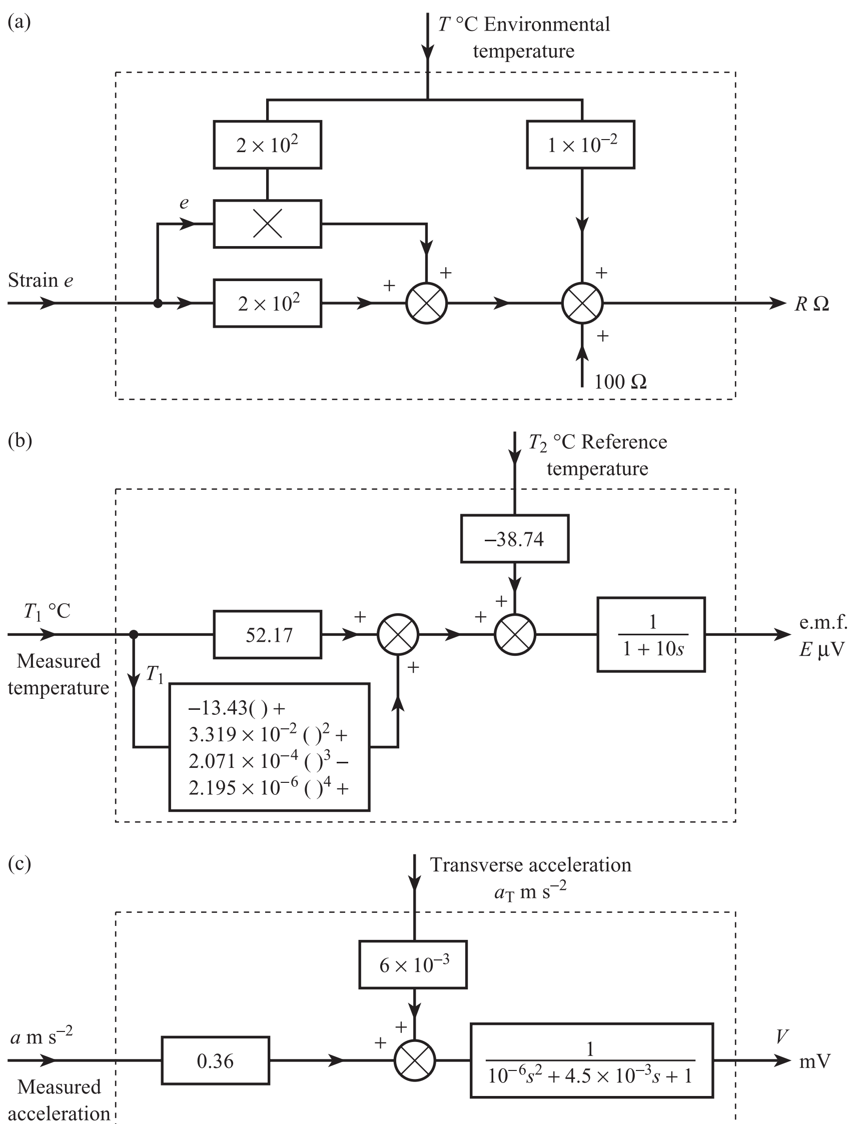


Figure 2.12 Examples of element characteristics:
(a) Strain gauge
(b) Copper–constantan thermocouple
(c) Accelerometer.

The strain gauge has an unstrained resistance of 100 Ω and gauge factor (Section 8.1) of 2.0. Non-linearity and dynamic effects can be neglected, but the resistance of the gauge is affected by ambient temperature as well as strain. Here temperature acts as both a modifying and an interfering input, i.e. it affects both gauge sensitivity and resistance at zero strain.

Figure 2.12(b) represents a copper–constantan thermocouple between 0 and 400 °C. The figure is drawn using eqns [2.7b] and [2.7c] for ideal straight-line and non-linear correction functions; these apply to a single junction. A thermocouple installation consists of two junctions (Section 8.5) – a measurement junction at T_1 °C and a reference junction at T_2 °C. The resultant e.m.f. is the difference of the two junction potentials and thus depends on both T_1 and T_2, i.e. $E(T_1, T_2) = E(T_1) - E(T_2)$; T_2 is thus an interfering input. The model applies to the situation where T_2 is small compared with T_1, so that $E(T_2)$ can be approximated by 38.74 T_2, the largest term in eqn [2.7a]. The dynamics are represented by a first-order transfer function of time constant 10 seconds (Chapters 4 and 14).

Figure 2.12(c) represents an accelerometer with a linear sensitivity of 0.35 mV m^{-1} s^2 and negligible non-linearity. Any transverse acceleration a_T, i.e. any acceleration perpendicular to that being measured, acts as an interfering input. The dynamics are represented by a second-order transfer function with a natural frequency of 250 Hz and damping coefficient of 0.7 (Chapters 4 and 8).

2.3 Statistical characteristics

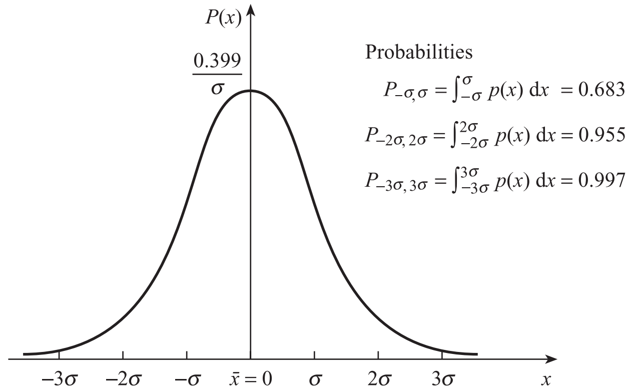
2.3.1 Statistical variations in the output of a single element with time – repeatability

Suppose that the input I of a single element, e.g. a pressure transducer, is held constant, say at 0.5 bar, for several days. If a large number of readings of the output O are taken, then the expected value of 1.0 volt is not obtained on every occasion; a range of values such as 0.99, 1.01, 1.00, 1.02, 0.98, etc., scattered about the expected value, is obtained. This effect is termed a lack of **repeatability** in the element. Repeatability is the ability of an element to give the same output for the same input, when repeatedly applied to it. Lack of repeatability is due to random effects in the element and its environment. An example is the vortex flowmeter (Section 12.2.4): for a fixed flow rate $Q = 1.4 \times 10^{-2}$ m^3 s^{-1}, we would expect a constant frequency output $f = 209$ Hz. Because the output signal is not a perfect sine wave, but is subject to random fluctuations, the measured frequency varies between 207 and 211 Hz.

The most common cause of lack of repeatability in the output O is random fluctuations with time in the environmental inputs I_M, I_I: if the coupling constants K_M, K_I are non-zero, then there will be corresponding time variations in O. Thus random fluctuations in ambient temperature cause corresponding time variations in the resistance of a strain gauge or the output voltage of an amplifier; random fluctuations in the supply voltage of a deflection bridge affect the bridge output voltage.

By making reasonable assumptions for the probability density functions of the inputs I, I_M and I_I (in a measurement system random variations in the input I to

Figure 2.13 Normal probability density function with $\bar{x} = 0$.



a given element can be caused by random effects in the previous element), the probability density function of the element output O can be found. The most likely probability density function for I, I_M and I_I is the normal or Gaussian distribution function (Figure 2.13):

Normal probability density function

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\bar{x})^2}{2\sigma^2}\right] \qquad [2.14]$$

where: $\bar{x}$ = mean or expected value (specifies centre of distribution)
σ = standard deviation (specifies spread of distribution).

Equation [2.9] expresses the independent variable O in terms of the independent variables I, I_M and I_I. Thus if ΔO is a small deviation in O from the mean value $\bar{O}$, caused by deviations ΔI, ΔI_M and ΔI_I from respective mean values $\bar{I}$, $\bar{I}_M$ and $\bar{I}_I$, then:

$$\Delta O = \left(\frac{\partial O}{\partial I}\right)\Delta I + \left(\frac{\partial O}{\partial I_M}\right)\Delta I_M + \left(\frac{\partial O}{\partial I_I}\right)\Delta I_I \qquad [2.15]$$

Thus ΔO is a linear combination of the variables ΔI, ΔI_M and ΔI_I; the partial derivatives can be evaluated using eqn [2.9]. It can be shown[2] that if a dependent variable y is a linear combination of independent variables x_1, x_2 and x_3, i.e.

$$y = a_1x_1 + a_2x_2 + a_3x_3 \qquad [2.16]$$

and if x_1, x_2 and x_3 have normal distributions with standard deviations σ_1, σ_2 and σ_3 respectively, then the probability distribution of y is also normal with standard deviation σ given by:

$$\sigma = \sqrt{a_1^2\sigma_1^2 + a_2^2\sigma_2^2 + a_3^2\sigma_3^2} \qquad [2.17]$$

From eqns [2.15] and [2.17] we see that the standard deviation of ΔO, i.e. of O about mean O, is given by:

Standard deviation of output for a single element

$$\sigma_0 = \sqrt{\left(\frac{\partial O}{\partial I}\right)^2 \sigma_I^2 + \left(\frac{\partial O}{\partial I_M}\right)^2 \sigma_{I_M}^2 + \left(\frac{\partial O}{\partial I_I}\right)^2 \sigma_{I_I}^2} \qquad [2.18]$$

where σ_I, σ_{I_M} and σ_{I_I} are the standard deviations of the inputs. Thus σ_0 can be calculated using eqn [2.18] if σ_I, σ_{I_M} and σ_{I_I} are known; alternatively if a calibration test (see following section) is being performed on the element then σ_0 can be estimated directly from the experimental results. The corresponding mean or expected value $\bar{O}$ of the element output is given by:

Mean value of output for a single element

$$\bar{O} = K\bar{I} + a + N(\bar{I}) + K_M \bar{I}_M \bar{I} + K_I \bar{I}_I \qquad [2.19]$$

and the corresponding probability density function is:

$$p(O) = \frac{1}{\sigma_0 \sqrt{2\pi}} \exp\left[\frac{-(O - \bar{O})^2}{2\sigma_0^2}\right] \qquad [2.20]$$

2.3.2 Statistical variations amongst a batch of similar elements – tolerance

Suppose that a user buys a batch of similar elements, e.g. a batch of 100 resistance temperature sensors, from a manufacturer. If he then measures the resistance R_0 of each sensor at 0 °C he finds that the resistance values are not all equal to the manufacturer's quoted value of 100.0 Ω. A range of values such as 99.8, 100.1, 99.9, 100.0 and 100.2 Ω, distributed statistically about the quoted value, is obtained. This effect is due to small random variations in manufacture and is often well represented by the normal probability density function given earlier. In this case we have:

$$p(R_0) = \frac{1}{\sigma_{R_0} \sqrt{2\pi}} \exp\left[\frac{-(R_0 - \bar{R}_0)^2}{2\sigma_{R_0}^2}\right] \qquad [2.21]$$

where $\bar{R}_0$ = mean value of distribution = 100 Ω and σ_{R_0} = standard deviation, typically 0.1 Ω. However, a manufacturer may state in his specification that R_0 lies within ±0.15 Ω of 100 Ω for all sensors, i.e. he is quoting **tolerance** limits of ±0.15 Ω. Thus in order to satisfy these limits he must reject for sale all sensors with $R_0 < 99.85$ Ω and $R_0 > 100.15$ Ω, so that the probability density function of the sensors bought by the user now has the form shown in Figure 2.14.

The user has two choices:

(a) He can design his measurement system using the manufacturer's value of $R_0 = 100.0$ Ω and accept that any individual system, with $R_0 = 100.1$ Ω say, will have a small measurement error. This is the usual practice.

(b) He can perform a calibration test to measure R_0 as accurately as possible for each element in the batch. This theoretically removes the error due to uncertainty in R_0 but is time-consuming and expensive. There is also a small remaining uncertainty in the value of R_0 due to the limited accuracy of the calibration equipment.

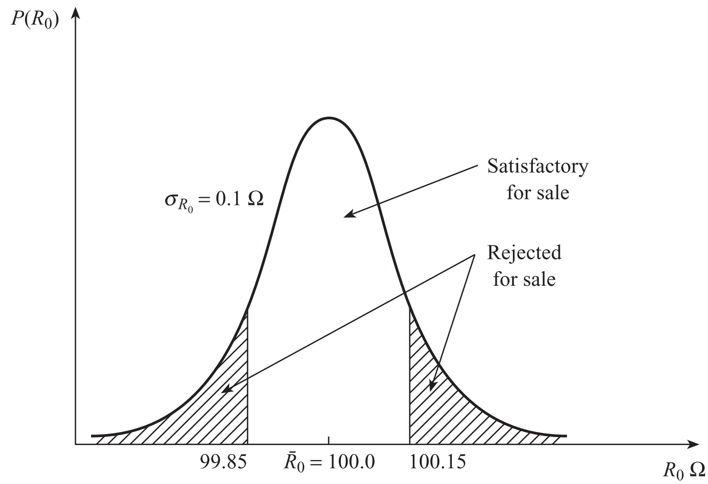


Figure 2.14
Tolerance limits.

This effect is found in any batch of 'identical' elements; significant variations are found in batches of thermocouples and thermistors, for example. In the general case we can say that the values of parameters, such as linear sensitivity K and zero bias a, for a batch of elements are distributed statistically about mean values $\bar{K}$ and $\bar{a}$.

2.3.3 Summary

In the general case of a batch of several 'identical' elements, where each element is subject to random variations in environmental conditions with time, both inputs I, I_M and I_I and parameters K, a, etc., are subject to statistical variations. If we assume that each statistical variation can be represented by a normal probability density function, then the probability density function of the element output O is also normal, i.e.:

$$p(O) = \frac{1}{\sigma_0\sqrt{2\pi}}\exp\left[\frac{-(O-\bar{O})^2}{2\sigma_0^2}\right] \qquad [2.20]$$

where the mean value $\bar{O}$ is given by:

Mean value of output for a batch of elements

$$\bar{O} = \bar{K}\bar{I} + \bar{N}(\bar{I}) + \bar{a} + \bar{K}_M\bar{I}_M\bar{I} + \bar{K}_I\bar{I}_I \qquad [2.22]$$

and the standard deviation σ_0 is given by:

Standard deviation of output for a batch of elements

$$\sigma_0 = \sqrt{\left(\frac{\partial O}{\partial I}\right)^2\sigma_I^2 + \left(\frac{\partial O}{\partial I_M}\right)^2\sigma_{I_M}^2 + \left(\frac{\partial O}{\partial I_I}\right)^2\sigma_{I_I}^2 + \left(\frac{\partial O}{\partial K}\right)^2\sigma_K^2 + \left(\frac{\partial O}{\partial a}\right)^2\sigma_a^2 + \ldots} \qquad [2.23]$$

Tables 2.1 and 2.2 summarise the static characteristics of a chromel–alumel thermocouple and a millivolt to current temperature transmitter. The thermocouple is

Table 2.1 Model for chromel–alumel thermocouple.

Model equation	$E_{T,T_a} = a_0 + a_1(T - T_a) + a_2(T^2 - T_a^2)$ (50 to 150 °C)
Mean values	$\bar{a}_0 = 0.00$, $\bar{a}_1 = 4.017 \times 10^{-2}$, $\bar{a}_2 = 4.66 \times 10^{-6}$, $\bar{T}_a = 10$
Standard deviations	$\sigma_{a_0} = 6.93 \times 10^{-2}$, $\sigma_{a_1} = 0.0$, $\sigma_{a_2} = 0.0$, $\sigma_{T_a} = 6.7$
Partial derivatives	$\dfrac{\partial E}{\partial a_0} = 1.0$, $\dfrac{\partial E}{\partial T_a} = -4.026 \times 10^{-2}$
Mean value of output	$\bar{E}_{T,T_a} = \bar{a}_0 + \bar{a}_1(\bar{T} - \bar{T}_a) + \bar{a}_2(\bar{T}^2 - \bar{T}_a^2)$
Standard deviation of output	$\sigma_E^2 = \left(\dfrac{\partial E}{\partial a_0}\right)^2 \sigma_{a_0}^2 + \left(\dfrac{\partial E}{\partial T_a}\right)^2 \sigma_{T_a}^2$

Table 2.2 Model for millivolt to current temperature transmitter.

Model equation	$i = KE + K_M E \Delta T_a + K_I \Delta T_a + a$ 4 to 20 mA output for 2.02 to 6.13 mV input ΔT_a = deviation in ambient temperature from 20 °C
Mean values	$\bar{K} = 3.893$, $\bar{a} = 3.864$, $\Delta\bar{T}_a = -10$ $\bar{K}_M = 1.95 \times 10^{-4}$, $\bar{K}_I = 2.00 \times 10^{-3}$
Standard deviations	$\sigma_a = 0.14$, $\sigma_{\Delta T_a} = 6.7$ $\sigma_K = 0.0$, $\sigma_{K_M} = 0.0$, $\sigma_{K_I} = 0.0$
Partial derivatives	$\dfrac{\partial i}{\partial E} = 3.891$, $\dfrac{\partial i}{\partial \Delta T_a} = 2.936 \times 10^{-3}$, $\dfrac{\partial i}{\partial a} = 1.0$
Mean value of output	$\bar{i} = \bar{K}\bar{E} + \bar{K}_M \bar{E} \Delta\bar{T}_a + \bar{K}_I \Delta\bar{T}_a + \bar{a}$
Standard deviation of output	$\sigma_i^2 = \left(\dfrac{\partial i}{\partial E}\right)^2 \sigma_E^2 + \left(\dfrac{\partial i}{\partial \Delta T_a}\right)^2 \sigma_{\Delta T}^2 + \left(\dfrac{\partial i}{\partial a}\right)^2 \sigma_a^2$

characterised by non-linearity, changes in reference junction (ambient temperature) T_a acting as an interfering input, and a spread of zero bias values a_0. The transmitter is linear but is affected by ambient temperature acting as both a modifying and an interfering input. The zero and sensitivity of this element are adjustable; we cannot be certain that the transmitter is set up exactly as required, and this is reflected in a non-zero value of the standard deviation of the zero bias a.

2.4 Identification of static characteristics – calibration

2.4.1 Standards

The static characteristics of an element can be found experimentally by measuring corresponding values of the input I, the output O and the environmental inputs I_M and I_I, when I is either at a constant value or changing slowly. This type of experiment is referred to as **calibration**, and the measurement of the variables I, O, I_M and I_I must be accurate if meaningful results are to be obtained. The instruments and techniques used to quantify these variables are referred to as **standards** (Figure 2.15).

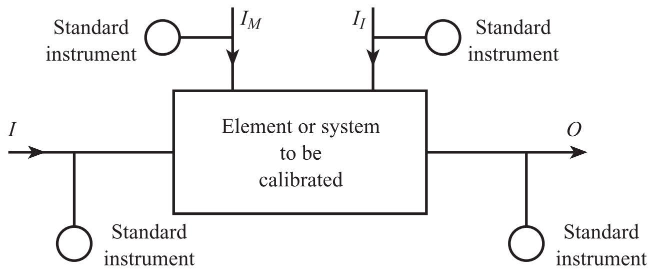


Figure 2.15 Calibration of an element.

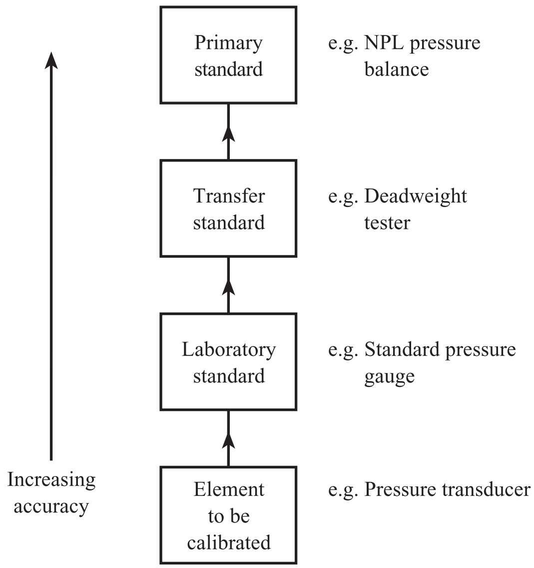
The **accuracy** of a measurement of a variable is the closeness of the measurement to the true value of the variable. It is quantified in terms of measurement error, i.e. the difference between the measured value and the true value (Chapter 3). Thus the accuracy of a laboratory standard pressure gauge is the closeness of the reading to the true value of pressure. This brings us back to the problem, mentioned in the previous chapter, of how to establish the true value of a variable. We define the true value of a variable as the measured value obtained with a standard of ultimate accuracy. Thus the accuracy of the above pressure gauge is quantified by the difference between the gauge reading, for a given pressure, and the reading given by the ultimate pressure standard. However, the manufacturer of the pressure gauge may not have access to the ultimate standard to measure the accuracy of his products.

In the United Kingdom the manufacturer is supported by the National Measurement System. **Ultimate** or **primary** measurement standards for key physical variables such as time, length, mass, current and temperature are maintained at the National Physical Laboratory (NPL). Primary measurement standards for other important industrial variables such as the density and flow rate of gases and liquids are maintained at the National Engineering Laboratory (NEL). In addition there is a network of laboratories and centres throughout the country which maintain **transfer** or **intermediate** standards. These centres are accredited by UKAS (United Kingdom Accreditation Service). Transfer standards held at accredited centres are calibrated against national primary and secondary standards, and a manufacturer can calibrate his products against the transfer standard at a local centre. Thus the manufacturer of pressure gauges can calibrate his products against a transfer standard, for example a deadweight tester. The transfer standard is in turn calibrated against a primary or secondary standard, for example a pressure balance at NPL. This introduces the concept of a **traceability ladder**, which is shown in simplified form in Figure 2.16.

The element is calibrated using the laboratory standard, which should itself be calibrated using the transfer standard, and this in turn should be calibrated using the primary standard. Each element in the ladder should be significantly more accurate than the one below it.

NPL are currently developing an **Internet calibration** service.[3] This will allow an element at a remote location (for example at a user's factory) to be calibrated directly against a national primary or secondary standard without having to be transported to NPL. The traceability ladder is thereby collapsed to a single link between element and national standard. The same input must be applied to element and standard. The measured value given by the standard instrument is then the true value of the input to the element, and this is communicated to the user via the Internet. If the user measures the output of the element for a number of true values of input, then the characteristics of the element can be determined to a known, high accuracy.

Figure 2.16 Simplified traceability ladder.



2.4.2 SI units

Having introduced the concepts of standards and traceability we can now discuss different types of standards in more detail. The International System of Units (SI) comprises seven base units, which are listed and defined in Table 2.3. The units of all physical quantities can be derived from these base units. Table 2.4 lists common physical quantities and shows the derivation of their units from the base units. In the United Kingdom the National Physical Laboratory (NPL) is responsible for the

Table 2.3 SI base units (after National Physical Laboratory 'Units of Measurement' poster, 1996[4]).

Time: second (s)	The second is the duration of 9 192 631 770 periods of the radiation corresponding to the transition between the two hyperfine levels of the ground state of the caesium-133 atom.
Length: metre (m)	The metre is the length of the path travelled by light in vacuum during a time interval of 1/299 792 458 of a second.
Mass: kilogram (kg)	The kilogram is the unit of mass; it is equal to the mass of the international prototype of the kilogram.
Electric current: ampere (A)	The ampere is that constant current which, if maintained in two straight parallel conductors of infinite length, of negligible circular cross-section, and placed 1 metre apart in vacuum, would produce between these conductors a force equal to 2×10^{-7} newton per metre of length.
Thermodynamic temperature: kelvin (K)	The kelvin, unit of thermodynamic temperature, is the fraction 1/273.16 of the thermodynamic temperature of the triple point of water.
Amount of substance: mole (mol)	The mole is the amount of substance of a system which contains as many elementary entities as there are atoms in 0.012 kilogram of carbon-12.
Luminous intensity: candela (cd)	The candela is the luminous intensity, in a given direction, of a source that emits monochromatic radiation of frequency 540×10^{12} hertz and that has a radiant intensity in that direction of (1/683) watt per steradian.

Table 2.4 SI derived units (after National Physical Laboratory Units of Measurement poster, 1996[4]).

Examples of SI derived units expressed in terms of base units

Quantity	SI unit	
	Name	Symbol
area	square metre	m^2
volume	cubic metre	m^3
speed, velocity	metre per second	m/s
acceleration	metre per second squared	m/s^2
wave number	1 per metre	m^{-1}
density, mass density	kilogram per cubic metre	kg/m^3
specific volume	cubic metre per kilogram	m^3/kg
current density	ampere per square metre	A/m^2
magnetic field strength	ampere per metre	A/m
concentration (of amount of substance)	mole per cubic metre	mol/m^3
luminance	candela per square metre	cd/m^2

SI derived units with special names

Quantity	SI unit			
	Name	Symbol	Expression in terms of other units	Expression[a] in terms of SI base units
plane angle[b]	radian	rad		$m \cdot m^{-1} = 1$
solid angle[b]	steradian	sr		$m^2 \cdot m^{-2} = 1$
frequency	hertz	Hz		s^{-1}
force	newton	N		$m\ kg\ s^{-2}$
pressure, stress	pascal	Pa	N/m^2	$m^{-1}\ kg\ s^{-2}$
energy, work quantity of heat	joule	J	N m	$m^2\ kg\ s^{-2}$
power, radiant flux	watt	W	J/s	$m^2\ kg\ s^{-3}$
electric charge, quantity of electricity	coulomb	C		s A
electric potential, potential difference, electromotive force	volt	V	W/A	$m^2\ kg\ s^{-3}\ A^{-1}$
capacitance	farad	F	C/V	$m^{-2}\ kg^{-1}\ s^4\ A^2$
electric resistance	ohm	Ω	V/A	$m^2\ kg\ s^{-3}\ A^{-2}$
electric conductance	siemens	S	A/V	$m^{-2}\ kg^{-1}\ s^3\ A^2$
magnetic flux	weber	Wb	V s	$m^2\ kg\ s^{-2}\ A^{-1}$
magnetic flux density	tesla	T	Wb/m^2	$kg\ s^{-2}\ A^{-1}$
inductance	henry	H	Wb/A	$m^2\ kg\ s^{-2}\ A^{-2}$
Celsius temperature	degree Celsius	°C		K
luminous flux	lumen	lm	cd sr	$cd \cdot m^2 \cdot m^{-2} = cd$
illuminance	lux	lx	lm/m^2	$cd \cdot m^2 \cdot m^{-4} = cd \cdot m^{-2}$
activity (of a radionuclide)	becquerel	Bq		s^{-1}
absorbed dose, specific energy imparted, kerma	gray	Gy	J/kg	$m^2\ s^{-2}$
dose equivalent	sievert	Sv	J/kg	$m^2\ s^{-2}$

Table 2.4 (*cont'd*)

Examples of SI derived units expressed by means of special names

Quantity	SI unit		
	Name	Symbol	Expression in terms of SI base units
dynamic viscosity	pascal second	Pa s	m^{-1} kg s^{-1}
moment of force	newton metre	N m	m^2 kg s^{-2}
surface tension	newton per metre	N/m	kg s^{-2}
heat flux density, irradiance	watt per square metre	W/m^2	kg s^{-3}
heat capacity, entropy	joule per kelvin	J/K	m^2 kg s^{-2} K^{-1}
specific heat capacity, specific entropy	joule per kilogram kelvin	J/(kg K)	m^2 s^{-2} K^{-1}
specific energy	joule per kilogram	J/kg	m^2 s^{-2}
thermal conductivity	watt per metre kelvin	W/(m K)	m kg s^{-3} K^{-1}
energy density	joule per cubic metre	J/m^3	m^{-1} kg s^{-2}
electric field strength	volt per metre	V/m	m kg s^{-3} A^{-1}
electric charge density	coulomb per cubic metre	C/m^3	m^{-3} s A
electric flux density	coulomb per square metre	C/m^2	m^{-2} s A
permittivity	farad per metre	F/m	m^{-3} kg^{-1} s^4 A^2
permeability	henry per metre	H/m	m kg s^{-2} A^{-2}
molar energy	joule per mole	J/mol	m^2 kg s^{-2} mol^{-1}
molar entropy, molar heat capacity	joule per mole kelvin	J/(mol K)	m^2 kg s^{-2} K^{-1} mol^{-1}
exposure (X and γ rays)	coulomb per kilogram	C/kg	kg^{-1} s A
absorbed dose rate	gray per second	Gy/s	m^2 s^{-3}

Examples of SI derived units formed by using the radian and steradian

Quantity	SI unit	
	Name	Symbol
angular velocity	radian per second	rad/s
angular acceleration	radian per second squared	rad/s^2
radiant intensity	watt per steradian	W/sr
radiance	watt per square metre steradian	W m^{-2} sr^{-1}

[a] Acceptable forms are, for example, $m \cdot kg \cdot s^{-2}$, m kg s^{-2}; m/s, $\frac{m}{s}$ or $m \cdot s^{-1}$.
[b] The CIPM (1995) decided that the radian and steradian should henceforth be designated as dimensionless derived units.

physical realisation of all of the base units and many of the derived units mentioned above. The NPL is therefore the custodian of ultimate or primary standards in the UK. There are secondary standards held at United Kingdom Accreditation Service (UKAS) centres. These have been calibrated against NPL standards and are available to calibrate transfer standards.

At NPL, the **metre** is realised using the wavelength of the 633 nm radiation from an iodine-stabilised helium–neon laser. The reproducibility of this primary standard

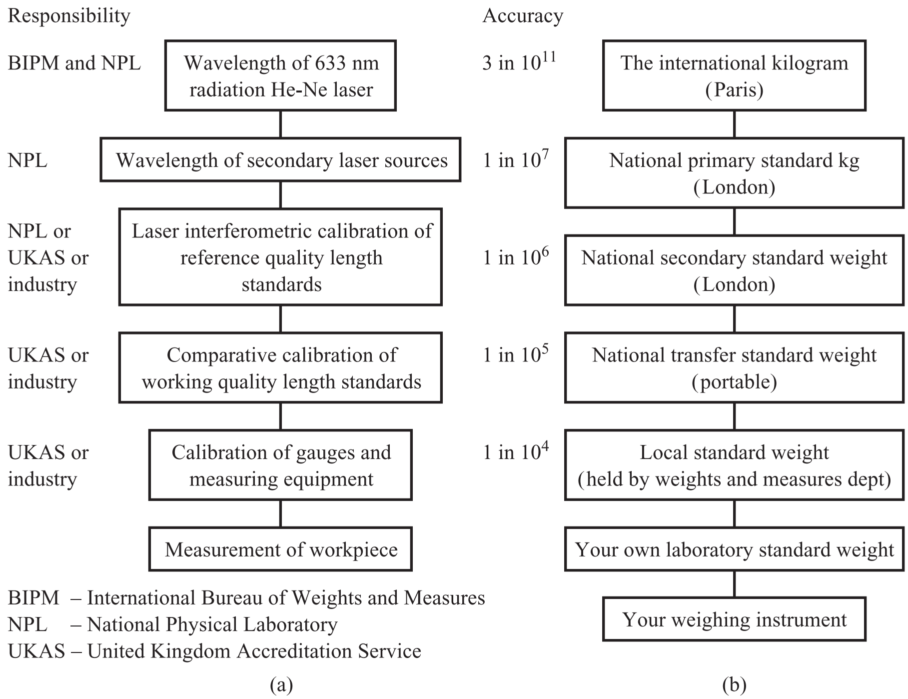


Figure 2.17
Traceability ladders:
(a) length (adapted from Scarr[5])
(b) mass (reprinted by permission of the Council of the Institution of Mechanical Engineers from Hayward[6]).

is about 3 parts in 10^{11}, and the wavelength of the radiation has been accurately related to the definition of the metre in terms of the velocity of light. The primary standard is used to calibrate secondary laser interferometers which are in turn used to calibrate precision length bars, gauges and tapes. A simplified traceability ladder for length[5] is shown in Figure 2.17(a).

The international prototype of the **kilogram** is made of platinum–iridium and is kept at the International Bureau of Weights and Measures (BIPM) in Paris. The British national copy is kept at NPL and is used in conjunction with a precision balance to calibrate secondary and transfer kilogram standards. Figure 2.17(b) shows a simplified traceability ladder for mass and weight.[6] The weight of a mass m is the force mg it experiences under the acceleration of gravity g; thus if the local value of g is known accurately, then a force standard can be derived from mass standards. At NPL deadweight machines covering a range of forces from 450 N to 30 MN are used to calibrate strain-gauge load cells and other weight transducers.

The **second** is realised by caesium beam standards to about 1 part in 10^{13}; this is equivalent to one second in 300 000 years! A uniform timescale, synchronised to 0.1 microsecond, is available worldwide by radio transmissions; this includes satellite broadcasts.

The **ampere** has traditionally been the electrical base unit and has been realised at NPL using the Ayrton-Jones current balance; here the force between two current-carrying coils is balanced by a known weight. The accuracy of this method is limited by the large deadweight of the coils and formers and the many length measurements necessary. For this reason the two electrical base units are now chosen to be the **farad** and the **volt** (or watt); the other units such as the ampere, ohm, henry and joule are derived from these two base units with time or frequency units, using

Ohm's law where necessary. The farad is realised using a calculable capacitor based on the Thompson–Lampard theorem. Using a.c. bridges, capacitance and frequency standards can then be used to calibrate standard resistors. The primary standard for the volt is based on the Josephson effect in superconductivity; this is used to calibrate secondary voltage standards, usually saturated Weston cadmium cells. The ampere can then be realised using a modified current balance. As before, the force due to a current I is balanced by a known weight mg, but also a separate measurement is made of the voltage e induced in the coil when moving with velocity u. Equating electrical and mechanical powers gives the simple equation:

$$eI = mgu$$

Accurate measurements of m, u and e can be made using secondary standards traceable back to the primary standards of the kilogram, metre, second and volt.

Ideally **temperature** should be defined using the thermodynamic scale, i.e. the relationship $PV = R\theta$ between the pressure P and temperature θ of a fixed volume V of an ideal gas. Because of the limited reproducibility of real gas thermometers the International Practical Temperature Scale (IPTS) was devised. This consists of:

(a) several highly reproducible fixed points corresponding to the freezing, boiling or triple points of pure substances under specified conditions;
(b) standard instruments with a known output versus temperature relationship obtained by calibration at fixed points.

The instruments interpolate between the fixed points.

The numbers assigned to the fixed points are such that there is exactly 100 K between the freezing point (273.15 K) and boiling point (373.15 K) of water. This means that a change of 1 K is equal to a change of 1 °C on the older Celsius scale. The exact relationship between the two scales is:

$$\theta \text{ K} = T \text{ °C} + 273.15$$

Table 2.5 shows the primary fixed points, and the temperatures assigned to them, for the 1990 version of the International Temperature Scale – ITS90.[7] In addition there are primary standard instruments which **interpolate** between these fixed points. Platinum resistance detectors (PRTD – Section 8.1) are used up to the freezing point of silver (961.78 °C) and radiation pyrometers (Section 15.6) are used at higher temperatures. The resistance R_T Ω of a PRTD at temperature T °C (above 0 °C) can be specified by the quadratic equation

$$R_T = R_0(1 + \alpha T + \beta T^2)$$

where R_0 is the resistance at 0 °C and α, β are constants. By measuring the resistance at three adjacent fixed points (e.g. water, gallium and indium), R_0, α and β can be calculated. An interpolating equation can then be found for the standard, which relates R_T to T and is valid between the water and indium fixed points. This primary standard PRTD, together with its equation, can then be used to calibrate secondary and transfer standards, usually PRTDs or thermocouples depending on temperature range.

The standards available for the base quantities, i.e. length, mass, time, current and temperature, enable standards for derived quantities to be realised. This is illustrated in the methods for calibrating liquid flowmeters.[8] The actual flow rate through the meter is found by weighing the amount of water collected in a given time, so that the accuracy of the flow rate standard depends on the accuracies of weight and time standards.

Table 2.5 The primary fixed points defining the International Temperature Scale 1990 – ITS90.

Equilibrium state	θ K	T °C
Triple point of hydrogen	13.8033	–259.3467
Boiling point of hydrogen at a pressure of 33 321.3 Pa	17.035	–256.115
Boiling point of hydrogen at a pressure of 101 292 Pa	20.27	–252.88
Triple point of neon	24.5561	–248.5939
Triple point of oxygen	54.3584	–218.7916
Triple point of argon	83.8058	–189.3442
Triple point of mercury	234.3156	–38.8344
Triple point of water	273.16	0.01
Melting point of gallium	302.9146	29.7646
Freezing point of indium	429.7485	156.5985
Freezing point of tin	505.078	231.928
Freezing point of zinc	692.677	419.527
Freezing point of aluminium	933.473	660.323
Freezing point of silver	1234.93	961.78
Freezing point of gold	1337.33	1064.18
Freezing point of copper	1357.77	1084.62

National primary and secondary standards for **pressure** above 110 kPa use pressure balances.[3] Here the force pA due to pressure p acting over an area A is balanced by the gravitational force mg acting on a mass m, i.e.

$$pA = mg$$

or

$$p = mg/A$$

Thus standards for pressure can be derived from those for mass and length, though the local value of gravitational acceleration g must also be accurately known.

2.4.3 Experimental measurements and evaluation of results

The calibration experiment is divided into three main parts.

O versus *I* with $I_M = I_I = 0$

Ideally this test should be held under 'standard' environmental conditions so that $I_M = I_I = 0$; if this is not possible all environmental inputs should be measured. I should be increased slowly from I_{MIN} to I_{MAX} and corresponding values of I and O recorded at intervals of 10% span (i.e. 11 readings), allowing sufficient time for the output to settle out before taking each reading. A further 11 pairs of readings should be taken with I decreasing slowly from I_{MAX} to I_{MIN}. The whole process should be repeated for two further '**ups**' and '**downs**' to yield two sets of data: an '**up**' set $(I_i, O_i)_{I\uparrow}$ and a '**down**' set $(I_j, O_j)_{I\downarrow}$ $i, j = 1, 2, \ldots, n$ $(n = 33)$.

Computer software regression packages are readily available which fit a polynomial, i.e. $O(I) = \sum_{q=0}^{q=m} a_q I^q$, to a set of n data points. These packages use a 'least squares' criterion. If d_i is the deviation of the polynomial value $O(I_i)$ from the data value O_i, then $d_i = O(I_i) - O_i$. The program finds a set of coefficients a_0, a_1, a_2, etc.,

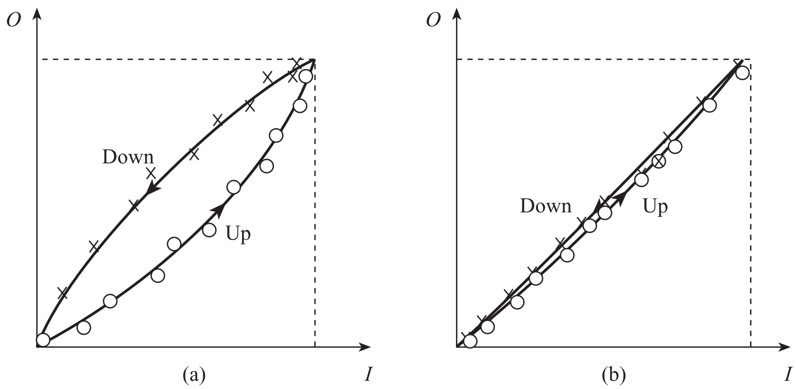


Figure 2.18
(a) Significant hysteresis
(b) Insignificant hysteresis.

such that the sum of the squares of the deviations, i.e. $\sum_{i=1}^{i=n} d_i^2$, is a minimum. This involves solving a set of linear equations.[9]

In order to detect any hysteresis, separate regressions should be performed on the two sets of data $(I_i, O_i)_{I\uparrow}$, $(I_j, O_j)_{I\downarrow}$, to yield two polynomials:

$$O(I)_{I\uparrow} = \sum_{q\;0}^{q\;m} a_q^{\uparrow} I^q \quad \text{and} \quad O(I)_{I\downarrow} = \sum_{q\;0}^{q\;m} a_q^{\downarrow} I^q$$

If the hysteresis is significant, then the separation of the two curves will be greater than the scatter of data points about each individual curve (Figure 2.18(a)). Hysteresis $H(I)$ is then given by eqn [2.10], i.e. $H(I) = O(I)_{I\downarrow} - O(I)_{I\uparrow}$. If, however, the scatter of points about each curve is greater than the separation of the curves (Figure 2.18(b)), then H is not significant and the two sets of data can then be combined to give a single polynomial $O(I)$. The slope K and zero bias a of the ideal straight line joining the minimum and maximum points (I_{MIN}, O_{MIN}) and (I_{MAX}, O_{MAX}) can be found from eqn [2.3]. The non-linear function $N(I)$ can then be found using eqn [2.4]:

$$N(I) = O(I) - (KI + a) \tag{2.24}$$

Temperature sensors are often calibrated using appropriate fixed points rather than a standard instrument. For example, a thermocouple may be calibrated between 0 and 500 °C by measuring e.m.f.'s at ice, steam and zinc points. If the e.m.f.–temperature relationship is represented by the cubic $E = a_1T + a_2T^2 + a_3T^3$, then the coefficients a_1, a_2, a_3 can be found by solving three simultaneous equations (see Problem 2.1).

O versus I_M, I_I at constant *I*

We first need to find which environmental inputs are interfering, i.e. which affect the zero bias a. The input I is held constant at $I = I_{MIN}$, and one environmental input is changed by a known amount, the rest being kept at standard values. If there is a resulting change ΔO in O, then the input I_I is interfering and the value of the corresponding coefficient K_I is given by $K_I = \Delta O/\Delta I_I$. If there is no change in O, then the input is not interfering; the process is repeated until all interfering inputs are identified and the corresponding K_I values found.

We now need to identify modifying inputs, i.e. those which affect the sensitivity of the element. The input I is held constant at the mid-range value $\frac{1}{2}(I_{MIN} + I_{MAX})$ and each environmental input is varied in turn by a known amount. If a change in input produces a change ΔO in O and is not an interfering input, then it must be a modifying input I_M and the value of the corresponding coefficient K_M is given by

$$K_M = \frac{1}{I}\frac{\Delta O}{\Delta I_M} = \frac{2}{I_{\text{MIN}} + I_{\text{MAX}}}\frac{\Delta O}{\Delta I_M} \qquad [2.25]$$

Suppose a change in input produces a change ΔO in O and it has already been identified as an interfering input with a known value of K_I. Then we must calculate a non-zero value of K_M before we can be sure that the input is also modifying. Since

$$\Delta O = K_I \Delta I_{I,M} + K_M \Delta I_{I,M} \frac{I_{\text{MIN}} + I_{\text{MAX}}}{2}$$

then

$$K_M = \frac{2}{I_{\text{MIN}} + I_{\text{MAX}}}\left[\frac{\Delta O}{\Delta I_{I,M}} - K_I\right] \qquad [2.26]$$

Repeatability test

This test should be carried out in the normal working environment of the element, e.g. out on the plant, or in a control room, where the environmental inputs I_M and I_I are subject to the random variations usually experienced. The signal input I should be held constant at mid-range value and the output O measured over an extended period, ideally many days, yielding a set of values O_k, $k = 1, 2, \ldots, N$. The mean value of the set can be found using

$$\bar{O} = \frac{1}{N}\sum_{k=1}^{k=N} O_k \qquad [2.27]$$

and the standard deviation (root mean square of deviations from the mean) is found using

$$\sigma_0 = \sqrt{\frac{1}{N}\sum_{k=1}^{N}(O_k - \bar{O})^2} \qquad [2.28]$$

A histogram of the values O_k should then be plotted in order to estimate the probability density function $p(O)$ and to compare it with the normal form (eqn [2.20]). A repeatability test on a pressure transducer can be used as an example of the construction of a histogram. Suppose that $N = 50$ and readings between 0.975 and 1.030 V are obtained corresponding to an expected value of 1.000 V. The readings are grouped into equal intervals of width 0.005 V and the number in each interval is found, e.g. 12, 10, 8, etc. This number is divided by the total number of readings, i.e. 50, to give the probabilities 0.24, 0.20, 0.16, etc. of a reading occurring in a given interval. The probabilities are in turn divided by the interval width 0.005 V to give the probability densities 48, 40 and 32 V^{-1} plotted in the histogram (Figure 2.19). We note that the area of each rectangle represents the probability that a reading lies within the interval, and that the total area of the histogram is equal to unity. The mean and standard deviation are found from eqns [2.27] and [2.28] to be $\bar{O} = 0.999$ V and $\sigma_0 = 0.010$ V respectively. Figure 2.19 also shows a normal probability density function with these values.

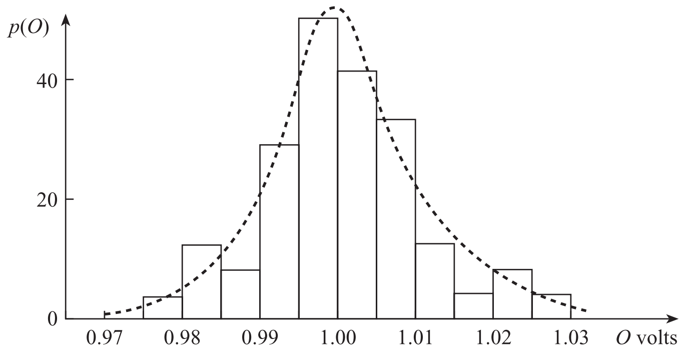


Figure 2.19 Comparison of histogram with normal probability density function.

Conclusion

The chapter began by discussing the **static** or **steady-state** characteristics of measurement system elements. **Systematic characteristics** such as non-linearity and environmental effects were first explained. This led to the **generalised model** of an element. **Statistical** characteristics, i.e. repeatability and tolerance, were then discussed. The last section explained how these characteristics can be measured experimentally, i.e. **calibration** and the use and types of **standards**.

References

[1] IEC 584.1:1995 *International Thermocouple Reference Tables*, International Electrotechnical Committee.

[2] KENNEDY J B and NEVILLE A 1986 *Basic Statistical Methods for Engineers and Scientists*, 3rd edn, pp. 345–60. Harper and Row, New York.

[3] STUART P R 1987 ‘Standards for the measurement of pressure’, *Measurement and Control*, vol. 20, no. 8, pp. 7–11.

[4] National Physical Laboratory 1996 *Units of Measurement*, poster, 8th edn.

[5] SCARR A 1979 ‘Measurement of length’, *Journal of the Institute of Measurement and Control*, vol. 12, July, pp. 265–9.

[6] HAYWARD A T J 1977 *Repeatability and Accuracy*, p. 34. Mechanical Engineering Publications, London.

[7] TC Ltd 2001 *Guide to Thermocouple and Resistance Thermometry*, issue 6.0.

[8] HAYWARD A T J 1977 ‘Methods of calibrating flowmeters with liquids – a comparative survey’, *Transactions of the Institute of Measurement and Control*, vol. 10, pp. 106–16.

[9] FREUND J E 1984 *Modern Elementary Statistics*, 6th edn, pp. 401–34. Prentice Hall International, Englewood Cliffs, NJ.

Problems

2.1 The e.m.f. at a thermocouple junction is 645 μV at the steam point, 3375 μV at the zinc point and 9149 μV at the silver point. Given that the e.m.f.–temperature relationship is of the form $E(T) = a_1T + a_2T^2 + a_3T^3$ (T in °C), find a_1, a_2 and a_3.

2.2 The resistance $R(\theta)$ of a thermistor at temperature θ K is given by $R(\theta) = \alpha \exp(\beta/\theta)$. Given that the resistance at the ice point ($\theta = 273.15$ K) is 9.00 kΩ and the resistance at the steam point is 0.50 kΩ, find the resistance at 25 °C.

2.3 A displacement sensor has an input range of 0.0 to 3.0 cm and a standard supply voltage $V_S = 0.5$ volts. Using the calibration results given in the table, estimate:

(a) The maximum non-linearity as a percentage of f.s.d.
(b) The constants K_I, K_M associated with supply voltage variations.
(c) The slope K of the ideal straight line.

Displacement x cm	0.0	0.5	1.0	1.5	2.0	2.5	3.0
Output voltage millivolts ($V_S = 0.5$)	0.0	16.5	32.0	44.0	51.5	55.5	58.0
Output voltage millivolts ($V_S = 0.6$)	0.0	21.0	41.5	56.0	65.0	70.5	74.0

2.4 A liquid level sensor has an input range of 0 to 15 cm. Use the calibration results given in the table to estimate the maximum hysteresis as a percentage of f.s.d.

Level h cm	0.0	1.5	3.0	4.5	6.0	7.5	9.0	10.5	12.0	13.5	15.0
Output volts h increasing	0.00	0.35	1.42	2.40	3.43	4.35	5.61	6.50	7.77	8.85	10.2
Output volts h decreasing	0.14	1.25	2.32	3.55	4.43	5.70	6.78	7.80	8.87	9.65	10.2

2.5 A repeatability test on a vortex flowmeter yielded the following 35 values of frequency corresponding to a constant flow rate of 1.4×10^{-2} m^3 s^{-1}: 208.6; 208.3; 208.7; 208.5; 208.8; 207.6; 208.9; 209.1; 208.2; 208.4; 208.1; 209.2; 209.6; 208.6; 208.5; 207.4; 210.2; 209.2; 208.7; 208.4; 207.7; 208.9; 208.7; 208.0; 209.0; 208.1; 209.3; 208.2; 208.6; 209.4; 207.6; 208.1; 208.8; 209.2; 209.7 Hz.

(a) Using equal intervals of width 0.5 Hz, plot a histogram of probability density values.
(b) Calculate the mean and standard deviation of the data.
(c) Sketch a normal probability density function with the mean and standard deviation calculated in (b) on the histogram drawn in (a).

2.6 A platinum resistance sensor is used to interpolate between the triple point of water (0 °C), the boiling point of water (100 °C) and the freezing point of zinc (419.6 °C). The corresponding resistance values are 100.0 Ω, 138.5 Ω and 253.7 Ω. The algebraic form of the interpolation equation is:

$$R_T = R_0(1 + \alpha T + \beta T^2)$$

where R_T Ω = resistance at T °C
R_0 Ω = resistance at 0 °C
α, β = constants.

Find the numerical form of the interpolation equation.

2.7 The following results were obtained when a pressure transducer was tested in a laboratory under the following conditions:

I Ambient temperature 20 °C, supply voltage 10 V (standard)
II Ambient temperature 20 °C, supply voltage 12 V
III Ambient temperature 25 °C, supply voltage 10 V

Input (barg)	0	2	4	6	8	10
Output (mA)						
I	4	7.2	10.4	13.6	16.8	20
II	4	8.4	12.8	17.2	21.6	28
III	6	9.2	12.4	15.6	18.8	22

(a) Determine the values of K_M, K_I, a and K associated with the generalised model equation $O = (K + K_M I_M)I + a + K_I I_I$.
(b) Predict an output value when the input is 5 barg, $V_S = 12$ V and ambient temperature is 25 °C.

Basic problems

2.8 A force sensor has an output range of 1 to 5 V corresponding to an input range of 0 to 2×10^5 N. Find the equation of the ideal straight line.

2.9 A differential pressure transmitter has an input range of 0 to 2×10^4 Pa and an output range of 4 to 20 mA. Find the equation to the ideal straight line.

2.10 A non-linear pressure sensor has an input range of 0 to 10 bar and an output range of 0 to 5 V. The output voltage at 4 bar is 2.20 V. Calculate the non-linearity in volts and as a percentage of span.

2.11 A non-linear temperature sensor has an input range of 0 to 400 °C and an output range of 0 to 20 mV. The output signal at 100 °C is 4.5 mV. Find the non-linearity at 100 °C in millivolts and as a percentage of span.

2.12 A thermocouple used between 0 and 500 °C has the following input–output characteristics:

Input T °C	0	100	200	300	500
Output E μV	0	5268	10 777	16 325	27 388

(a) Find the equation of the ideal straight line.
(b) Find the non-linearity at 100 °C and 300 °C in μV and as a percentage of f.s.d.

2.13 A force sensor has an input range of 0 to 10 kN and an output range of 0 to 5 V at a standard temperature of 20 °C. At 30 °C the output range is 0 to 5.5 V. Quantify this environmental effect.

2.14 A pressure transducer has an output range of 1.0 to 5.0 V at a standard temperature of 20 °C, and an output range of 1.2 to 5.2 V at 30 °C. Quantify this environmental effect.

2.15 A pressure transducer has an input range of 0 to 10^4 Pa and an output range of 4 to 20 mA at a standard ambient temperature of 20 °C. If the ambient temperature is increased to 30 °C, the range changes to 4.2 to 20.8 mA. Find the values of the environmental sensitivities K_I and K_M.

2.16 An analogue-to-digital converter has an input range of 0 to 5 V. Calculate the resolution error both as a voltage and as a percentage of f.s.d. if the output digital signal is:

(a) 8-bit binary
(b) 16-bit binary.

2.17 A level transducer has an output range of 0 to 10 V. For a 3 metre level, the output voltage for a falling level is 3.05 V and for a rising level 2.95 V. Find the hysteresis as a percentage of span.

3 The Accuracy of Measurement Systems in the Steady State

In Chapter 1 we saw that the input to a measurement system is the true value of the variable being measured. Also, if the measurement system is complete, the system output is the measured value of the variable. In Chapter 2 we defined accuracy in terms of measurement error, i.e. the difference between the measured and true values of a variable. It follows therefore that accuracy is a property of a complete measurement system rather than a single element. Accuracy is quantified using measurement error E where:

$$\begin{aligned} E &= \text{measured value} - \text{true value} \\ &= \text{system output} - \text{system input} \end{aligned} \quad [3.1]$$

In this chapter we use the static model of a single element, developed previously, to calculate the output and thus the measurement error for a complete system of several elements. The chapter concludes by examining methods of reducing system error.

3.1 Measurement error of a system of ideal elements

Consider the system shown in Figure 3.1 consisting of n elements in series. Suppose each element is ideal, i.e. perfectly linear and not subject to environmental inputs. If we also assume the intercept or bias is zero, i.e. $a = 0$, then:

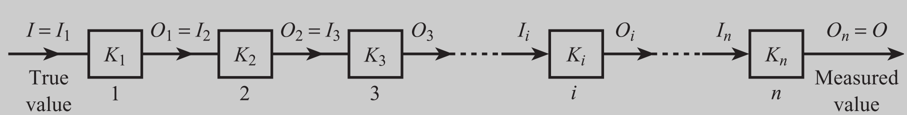
Input/output equation for ideal element with zero intercept

$$O_i = K_i I_i \quad [3.2]$$

for $i = 1, \ldots, n$, where K_i is the linear sensitivity or slope (eqn [2.3]). It follows that $O_2 = K_2 I_2 = K_2 K_1 I$, $O_3 = K_3 I_3 = K_3 K_2 K_1 I$, and for the whole system:

$$O = O_n = K_1 K_2 K_3 \ldots K_i \ldots K_n I \quad [3.3]$$

Figure 3.1.



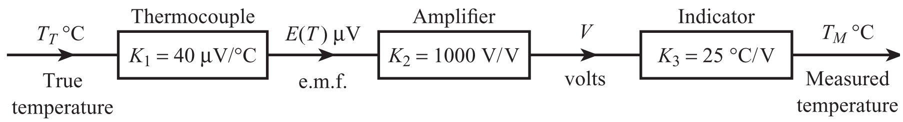


Figure 3.2 Simple temperature measurement system.

If the measurement system is complete, then $E = O - I$, giving:

$$E = (K_1K_2K_3 \dots K_n - 1)I \qquad [3.4]$$

Thus if

$$K_1K_2K_3 \dots K_n = 1 \qquad [3.5]$$

we have $E = 0$ and the system is perfectly accurate. The temperature measurement system shown in Figure 3.2 appears to satisfy the above condition. The indicator is simply a moving coil voltmeter (see Chapter 11) with a scale marked in degrees Celsius so that an input change of 1 V causes a change in deflection of 25 °C. This system has $K_1K_2K_3 = 40 \times 10^{-6} \times 10^3 \times 25 = 1$ and thus appears to be perfectly accurate. The system is not accurate, however, because none of the three elements present is ideal. The thermocouple is non-linear (Chapter 2), so that as the input temperature changes the sensitivity is no longer 40 μV °C^{-1}. Also changes in reference junction temperature (Figure 2.12(b)) cause the thermocouple e.m.f. to change. The output voltage of the amplifier is also affected by changes in ambient temperature (Chapter 9). The sensitivity K_3 of the indicator depends on the stiffness of the restoring spring in the moving coil assembly. This is affected by changes in environmental temperature and wear, causing K_3 to deviate from 25 °C V^{-1}. Thus the condition $K_1K_2K_3 = 1$ cannot be always satisfied and the system is in error.

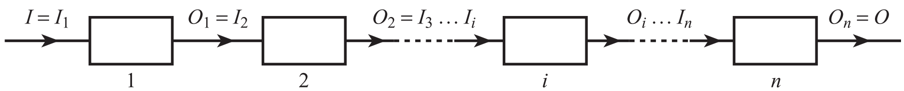
In general the error of any measurement system depends on the non-ideal characteristics – e.g. non-linearity, environmental and statistical effects – of every element in the system. Thus in order to quantify this error as precisely as possible we need to use the general model for a single element developed in Sections 2.2 and 2.3.

3.2 The error probability density function of a system of non-ideal elements

In Chapter 2 we saw that the probability density function of the output $p(O)$ of a single element can be represented by a normal distribution (eqn [2.20]). The mean value $\bar{O}$ of the distribution is given by eqn [2.22], which allows for non-linear and environmental effects. The standard deviation σ_0 is given by [2.23], which allows for statistical variations in inputs I, I_M and I_I with time, and statistical variations in parameters K, a, etc., amongst a batch of similar elements. These equations apply to each element in a measurement system of n elements and can be used to calculate the system error probability density function $p(E)$ as shown in Table 3.1.

Equations [3.6] (based on [2.22]) show how to calculate the mean value of the output of each element in turn, starting with $\bar{O}_1$ for the first and finishing with $\bar{O}_n = \bar{O}$ for the nth. The mean value $\bar{E}$ of the system error is simply the difference between the mean value of system output and mean value of system input (eqn [3.7]). Since the probability densities of the outputs of the individual elements are normal, then, using the result outlined in Section 2.3, the probability density function of the

Table 3.1 General calculation of system $p(E)$.



Mean values of element outputs

$$
\begin{aligned}
\bar{I}_1 &= \bar{I} \\
\bar{I}_2 = \bar{O}_1 &= \bar{K}_1 \bar{I}_1 + \bar{N}_1(\bar{I}_1) + \bar{a}_1 + \bar{K}_{M_1}\bar{I}_{M_1}\bar{I}_1 + \bar{K}_{I_1}\bar{I}_{I_1} \\
\bar{I}_3 = \bar{O}_2 &= \bar{K}_2 \bar{I}_2 + \bar{N}_2(\bar{I}_2) + \bar{a}_2 + \bar{K}_{M_2}\bar{I}_{M_2}\bar{I}_2 + \bar{K}_{I_2}\bar{I}_{I_2} \\
\vdots \quad & \vdots \\
\bar{I}_{i+1} = \bar{O}_i &= \bar{K}_i \bar{I}_i + \bar{N}_i(\bar{I}_i) + \bar{a}_i + \bar{K}_{M_i}\bar{I}_{M_i}\bar{I}_i + \bar{K}_{I_i}\bar{I}_{I_i} \\
\vdots \quad & \vdots \\
\bar{O} = \bar{O}_n &= \bar{K}_n \bar{I}_n + \bar{N}_n(\bar{I}_n) + \bar{a}_n + \bar{K}_{M_n}\bar{I}_{M_n}\bar{I}_n + \bar{K}_{I_n}\bar{I}_{I_n}
\end{aligned}
\qquad [3.6]
$$

Mean value of system error

$$\bar{E} = \bar{O} - \bar{I} \qquad [3.7]$$

Standard deviations of element outputs

$$
\begin{aligned}
\sigma_{I_1}^2 &= 0 \\
\sigma_{I_2}^2 = \sigma_{O_1}^2 &= \left(\frac{\partial O_1}{\partial I_1}\right)^2 \sigma_{I_1}^2 + \left(\frac{\partial O_1}{\partial I_{M_1}}\right)^2 \sigma_{I_{M_1}}^2 + \left(\frac{\partial O_1}{\partial I_{I_1}}\right)^2 \sigma_{I_{I_1}}^2 + \left(\frac{\partial O_1}{\partial K_1}\right)^2 \sigma_{K_1}^2 + \ldots \\
\sigma_{I_3}^2 = \sigma_{O_2}^2 &= \left(\frac{\partial O_2}{\partial I_2}\right)^2 \sigma_{I_2}^2 + \left(\frac{\partial O_2}{\partial I_{M_2}}\right)^2 \sigma_{I_{M_2}}^2 + \left(\frac{\partial O_2}{\partial I_{I_2}}\right)^2 \sigma_{I_{I_2}}^2 + \left(\frac{\partial O_2}{\partial K_2}\right)^2 \sigma_{K_2}^2 + \ldots \\
\vdots \quad & \vdots \\
\sigma_{I_{i+1}}^2 = \sigma_{O_i}^2 &= \left(\frac{\partial O_i}{\partial I_i}\right)^2 \sigma_{I_i}^2 + \left(\frac{\partial O_i}{\partial I_{M_i}}\right)^2 \sigma_{I_{M_i}}^2 + \left(\frac{\partial O_i}{\partial I_{I_i}}\right)^2 \sigma_{I_{I_i}}^2 + \left(\frac{\partial O_i}{\partial K_i}\right)^2 \sigma_{K_i}^2 + \ldots \\
\sigma_{O}^2 = \sigma_{O_n}^2 &= \left(\frac{\partial O_n}{\partial I_n}\right)^2 \sigma_{I_n}^2 + \left(\frac{\partial O_n}{\partial I_{M_n}}\right)^2 \sigma_{I_{M_n}}^2 + \left(\frac{\partial O_n}{\partial I_{I_n}}\right)^2 \sigma_{I_{I_n}}^2 + \left(\frac{\partial O_n}{\partial K_n}\right)^2 \sigma_{K_n}^2 + \ldots
\end{aligned}
\qquad [3.8]
$$

Standard deviation of system error

$$\sigma_E = \sigma_O \qquad [3.9]$$

Error probability density function

$$p(E) = \frac{1}{\sigma_E\sqrt{2\pi}} \exp\left[-\frac{1}{2\sigma_E^2}(E - \bar{E})^2\right] \qquad [3.10]$$

system output O and system error E is also normal (eqn [3.10]). Equations [3.8] (based on [2.23]) show how to calculate the standard deviation of the output of each element in turn, starting with σ_{O_1} for the first, and finishing with σ_{O_n} for the nth. We note that the standard deviation of the system input is zero and that the standard deviation of the error is equal to that of the system output (eqn [3.9]).

We now use the temperature measurement system shown in Figure 3.3 as an example of the calculation of $\bar{E}$ and σ_E. This consists of a platinum resistance temperature detector, current transmitter and recorder.

Table 3.2 gives the models for each of the elements in the temperature measurement system. In (a), the platinum resistance temperature detector is characterised by

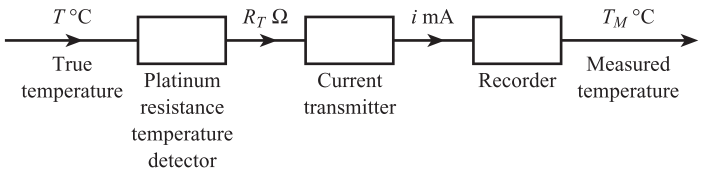


Figure 3.3 Temperature measurement system.

Table 3.2 Models for temperature measurement system elements.

(a) Platinum resistance temperature detector	
Model equation	$R_T = R_0(1 + \alpha T + \beta T^2)$
Individual mean values	$\bar{R}_0 = 100.0\ \Omega$, $\bar{\alpha} = 3.909 \times 10^{-3}$, $\bar{\beta} = -5.897 \times 10^{-7}$ (between 100 and 130 °C)
Individual standard deviations	$\sigma_{R_0} = 4.33 \times 10^{-2}$, $\sigma_\alpha = 0.0$, $\sigma_\beta = 0.0$
Partial derivatives	$\dfrac{\partial R_T}{\partial R_0} = 1.449$ at $T = 117$ °C
Overall mean value	$\bar{R}_T = \bar{R}_0(1 + \bar{\alpha}\bar{T} + \bar{\beta}\bar{T}^2)$
Overall standard deviation	$\sigma_{R_T}^2 = \left(\dfrac{\partial R_T}{\partial R_0}\right)^2 \sigma_{R_0}^2$
(b) Current transmitter	
Model equation	4 to 20 mA output for 138.5 to 149.8 Ω input (100 to 130 °C) ΔT_a = deviation of ambient temperature from 20 °C $i = KR_T + K_M R_T \Delta T_a + K_I \Delta T_a + a$
Individual mean values	$\bar{K} = 1.4134$, $\bar{K}_M = 1.4134 \times 10^{-4}$, $\bar{K}_I = -1.637 \times 10^{-2}$ $\bar{a} = -191.76$, $\Delta\bar{T}_a = -10$
Individual standard deviations	$\sigma_K = 0.0$, $\sigma_{K_M} = 0.0$, $\sigma_{K_I} = 0.0$ $\sigma_a = 0.24$, $\sigma_{\Delta T_a} = 6.7$
Partial derivatives	$\dfrac{\partial i}{\partial R_T} = 1.413$, $\dfrac{\partial i}{\partial \Delta T_a} = 4.11 \times 10^{-3}$, $\dfrac{\partial i}{\partial a} = 1.00$
Overall mean value	$\bar{i} = \bar{K}\bar{R}_T + \bar{K}_M \bar{R}_T \Delta\bar{T}_a + \bar{K}_I \Delta\bar{T}_a + \bar{a}$
Overall standard deviation	$\sigma_i^2 = \left(\dfrac{\partial i}{\partial R_T}\right)^2 \sigma_{R_T}^2 + \left(\dfrac{\partial i}{\partial \Delta T_a}\right)^2 \sigma_{\Delta T_a}^2 + \left(\dfrac{\partial i}{\partial a}\right)^2 \sigma_a^2$
(c) Recorder	
Model equation	$T_M = Ki + a$
Individual mean values	$\bar{K} = 1.875$, $\bar{a} = 92.50$ (100 to 130 °C record for 4 to 20 mA input)
Individual standard deviations	$\sigma_k = 0.0$, $\sigma_a = 0.10$
Partial derivatives	$\dfrac{\partial T_M}{\partial i} = 1.875$, $\dfrac{\partial T_M}{\partial a} = 1.00$
Overall mean value	$\bar{T}_M = \bar{K}\bar{i} + \bar{a}$
Overall standard deviation	$\sigma_{T_M}^2 = \left(\dfrac{\partial T_M}{\partial i}\right)^2 \sigma_i^2 + \left(\dfrac{\partial T_M}{\partial a}\right)^2 \sigma_a^2$

Table 3.3 Summary of calculation of $\bar{E}$ and σ_E for temperature system.

Mean $\bar{E}$

$$\bar{T} = 117\ °\text{C} \quad \bar{R}_T = 144.93\ \Omega$$

$$\bar{i} = 13.04\ \text{mA} \quad \bar{T}_M = 116.95\ °\text{C}$$

$$\bar{E} = \bar{T}_M - \bar{T} = -0.005\ °\text{C}$$

Standard deviation σ_E

$$\sigma_T^2 = 0$$

$$\sigma_{R_T}^2 = \left(\frac{\partial R_T}{\partial R_0}\right)^2 \sigma_{R_0}^2 = 39.4 \times 10^{-4}$$

$$\sigma_i^2 = \left(\frac{\partial i}{\partial R_T}\right)^2 \sigma_{R_T}^2 + \left(\frac{\partial i}{\partial \Delta T_a}\right)^2 \sigma_{\Delta T_a}^2 + \left(\frac{\partial i}{\partial a}\right)^2 \sigma_a^2$$

$$= 78.7 \times 10^{-4} + 8.18 \times 10^{-4} + 5.76 \times 10^{-2}$$

$$= 6.62 \times 10^{-2}$$

$$\sigma_{T_M}^2 = \left(\frac{\partial T_M}{\partial i}\right)^2 \sigma_i^2 + \left(\frac{\partial T_M}{\partial a}\right)^2 \sigma_a^2 = 24.3 \times 10^{-2}$$

$$\sigma_E = \sigma_{T_M} = 0.49\ °\text{C}$$

a small amount of non-linearity and a spread of values of R_0 (resistance at 0 °C). The current transmitter (b) is linear but temperature acts as both a modifying and an interfering input. The zero bias and sensitivity are adjustable: we cannot be certain that the transmitter will be set up exactly as stated in the table, and this is reflected in the non-zero value of σ_a. In (c) the recorder is linear but again calibration uncertainties are modelled by a non-zero value of σ_a.

Table 3.3 summarises the calculation of $\bar{E}$ and σ_E for the system when the mean value $\bar{T}$ of the true temperature is 117 °C. The corresponding mean values for the system are resistance $\bar{R}_T$ = 144.93 Ω (Table 3.2(a)), current $\bar{i}$ = 13.04 mA (Table 3.2(b)) and measured temperature $\bar{T}_M$ = 116.95 °C (Table 3.2(c)). The mean error is therefore $\bar{E}$ = −0.05 °C. The standard deviations σ_{R_T}, σ_i and σ_{T_M} are calculated using Table 3.2 to give σ_E = 0.49 °C.

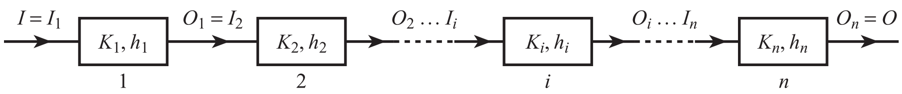
Modelling using error bands

In Section 2.1 we saw that in situations where element non-linearity, hysteresis and environmental effects are small, their overall effect is quantified using error bands. Here a systematic statement of the exact element input/output relationship (e.g. eqn [2.9]) is replaced by a statistical statement. The element output is described by a rectangular probability density function, of width $2h$, centred on the ideal straight line value $O_{\text{IDEAL}} = KI + a$. If every element in the system is described in this way, then the mean output value $\bar{O}_i$ will have the ideal value $\bar{O}_i = K_i\bar{I}_i + a_i$ for each element in the system. In the special case $a_i = 0$ for all i, $\bar{O}_i = K_i\bar{I}_i$, the mean value of system output is given by

$$\bar{O} = K_1K_2 \ldots K_i \ldots K_n\bar{I}$$

and the mean value of error $\bar{E} = \bar{O} - \bar{I}$. In the special case $K_1K_2 \ldots K_i \ldots K_n = 1$, $\bar{E} = 0$ (Table 3.4). The error probability density function $p(E)$ is the result of

Table 3.4 $\bar{E}$ and σ_E for system of elements described by error bands.



$$\bar{O}_i = K_i \bar{I}_i$$
$$\bar{O} = K_1 K_2 \ldots K_i \ldots K_n \bar{I}$$
$$\bar{E} = \bar{O} - \bar{I} \quad [= 0, \text{ if } K_1 K_2 \ldots K_n = 1]$$

$$\sigma^2_{I_1} = 0$$
$$\sigma^2_{I_2} = \sigma^2_{O_1} = \frac{h_1^2}{3}$$
$$\sigma^2_{I_3} = \sigma^2_{O_2} = K_2^2 \sigma^2_{I_2} + \frac{h_2^2}{3}$$
$$\vdots \qquad \vdots$$
$$\sigma^2_{I_{i+1}} = \sigma^2_{O_i} = K_i^2 \sigma^2_{I_i} + \frac{h_i^2}{3}$$
$$\vdots \qquad \vdots$$
$$\sigma^2_{O} = \sigma^2_{O_n} = K_n^2 \sigma^2_{I_n} + \frac{h_n^2}{3}$$
$$\sigma_E = \sigma_0$$

combining n rectangular distributions, each of width $2h_i$, $i = 1, 2, \ldots, n$. If $n > 3$, then the resultant distribution $p(E)$ approximates to a normal distribution; the larger the value of n the closer the distribution is to normal.

In order to calculate the standard deviation σ_E of $p(E)$ we return to eqns [2.16] and [2.17]. If a dependent variable y is a linear combination of independent random variables x_1, x_2, x_3, i.e.

$$y = a_1 x_1 + a_2 x_2 + a_3 x_3 \qquad [2.16]$$

then the standard deviation σ of the probability distribution of y is given by

$$\sigma^2 = a_1^2 \sigma_1^2 + a_2^2 \sigma_2^2 + a_3^2 \sigma_3^2 \qquad [2.17]$$

Equation [2.17] applies whatever the probability distributions of the individual x_1, x_2, x_3 are, provided that the individual standard deviations σ_1, σ_2, σ_3 are small.[1] Consider the ith element in Table 3.4 with sensitivity K_i and error band width $2h_i$. The standard deviation σ_{O_i} of the output O_i is equal to

$$\sigma^2_{O_i} = \sigma^2 \text{ due to input} + \sigma^2 \text{ due to element}$$

If σ_{I_i} is the standard deviation of the input I_i, then this contributes $K_i \sigma_{I_i}$ to the output standard deviation. The standard deviation of the rectangular distribution of width $2h_i$, due to the element, is $h_i/\sqrt{3}$. We therefore have:

$$\sigma^2_{O_i} = K_i^2 \sigma^2_{I_i} + h_i^2/3$$

Table 3.4 gives the calculation procedure for a complete system of n elements described by error bands.

3.3 Error reduction techniques

In the two previous sections we saw that the error of a measurement system depends on the non-ideal characteristics of every element in the system. Using the calibration techniques of Section 2.4, we can identify which elements in the system have the most dominant non-ideal behaviour. We can then devise compensation strategies for these elements which should produce significant reductions in the overall system error. This section outlines compensation methods for non-linear and environmental effects.

One of the most common methods of correcting a non-linear element is to introduce a **compensating non-linear element** into the system. This method is illustrated in Figure 3.4. Given a non-linear element, described by $U(I)$, we need a compensating element $C(U)$, such that the overall characteristics $C[U(I)]$ of the elements together are as close to the ideal straight line as possible. The method is illustrated in Figure 3.4 by the use of a deflection bridge to compensate for the non-linear characteristics of a thermistor. A detailed procedure for the design of the bridge is given in Section 9.1.

The most obvious method of reducing the effects of environmental inputs is that of **isolation**, i.e. to isolate the transducer from environmental changes so that effectively $I_M = I_I = 0$. Examples of this are the placing of the reference junction of a thermocouple in a temperature-controlled enclosure, and the use of spring mountings to isolate a transducer from the vibrations of the structure to which it is attached. Another obvious method is that of **zero environmental sensitivity**, where the element is completely insensitive to environmental inputs, i.e. $K_M = K_I = 0$. An example of this is the use of a metal alloy with zero temperature coefficients of expansion and resistance as a strain gauge element. Such an ideal material is difficult to find, and in practice the resistance of a metal strain gauge is affected slightly by changes in ambient temperature.

A more successful method of coping with environmental inputs is that of **opposing environmental inputs**. Suppose that an element is affected by an environmental

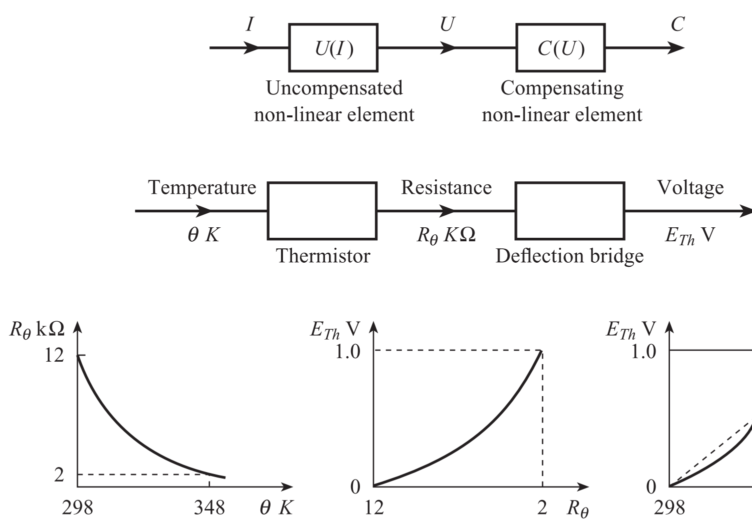


Figure 3.4 Compensating non-linear element.

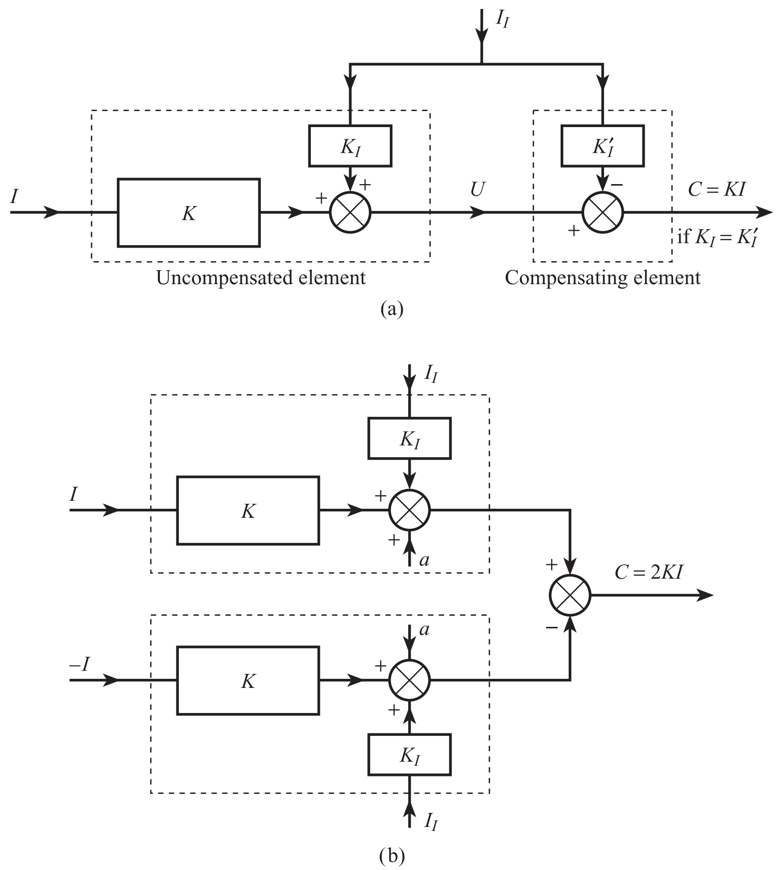


Figure 3.5 Compensation for interfering inputs: (a) Using opposing environmental inputs (b) Using a differential system.

input; then a second element, subject to the same environmental input, is deliberately introduced into the system so that the two effects tend to cancel out. This method is illustrated for interfering inputs in Figure 3.5 and can be easily extended to modifying inputs.

An example is compensation for variations in the temperature T_2 of the reference junction of a thermocouple. For a copper–constantan thermocouple (Figure 2.12(b)), we have $K_I I_I$ equal to $-38.74T_2$ μV so that a compensating element giving an output equal to $+38.74T_2$ μV is required. The design of the reference junction compensation element is discussed in Sections 8.5 and 9.1 and Problem 9.2.

An example of a **differential system** (Figure 3.5(b)) is the use of two matched strain gauges in adjacent arms of a bridge to provide compensation for ambient temperature changes. One gauge is measuring a tensile strain $+e$ and the other an equal compressive strain $-e$. The bridge effectively subtracts the two resistances so that the strain effect is doubled and the environmental effects cancel out.

The use of **high-gain negative feedback** is an important method of compensating for modifying inputs and non-linearity. Figure 3.6 illustrates the technique for a force transducer. The voltage output of a force-sensing element, subject to a modifying input, is amplified by a high-gain amplifier. The amplifier output is fed back to an element (e.g. a coil and permanent magnet) which provides a balancing force to oppose the input force.

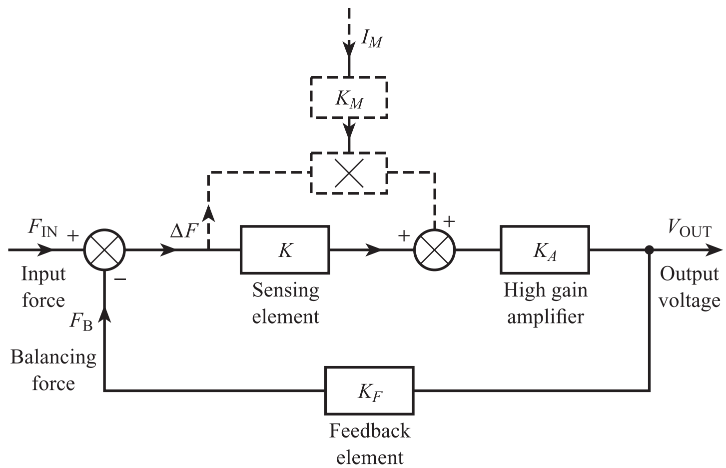


Figure 3.6 Closed loop force transducer.

Ignoring the effects of the modifying input for the moment we have:

$$\Delta F = F_{IN} - F_B \quad [3.11]$$

$$V_{OUT} = KK_A \Delta F \quad [3.12]$$

$$F_B = K_F V_{OUT} \quad [3.13]$$

i.e.

$$\frac{V_{OUT}}{KK_A} = F_{IN} - K_F V_{OUT}$$

giving

Equation for force transducer with negative feedback

$$V_{OUT} = \frac{KK_A}{1 + KK_A K_F} F_{IN} \quad [3.14]$$

If the amplifier gain K_A is made large such that the condition

$$KK_A K_F \gg 1 \quad [3.15]$$

is satisfied, then $V_{OUT} \approx \frac{1}{K_F} F_{IN}$. This means that the system output depends only on the gain K_F of the feedback element and is independent of the gains K and K_A in the forward path. This means that, provided the above condition is obeyed, changes in K and K_A due to modifying inputs and/or non-linear effects have negligible effect on V_{OUT}. This can be confirmed by repeating the above analysis with K replaced by $K + K_M I_M$, giving

$$V_{OUT} = \frac{(K + K_M I_M) K_A}{1 + (K + K_M I_M) K_A K_F} F_{IN} \quad [3.16]$$

which again reduces to $V_{OUT} \approx \frac{F_{IN}}{K_F}$ if $(K + K_M I_M) K_A K_F \gg 1$. We now, of course, have to ensure that the gain K_F of the feedback element does not change due to non-linear or environmental effects. Since the amplifier provides most of the required power, the feedback element can be designed for low power-handling capacity,

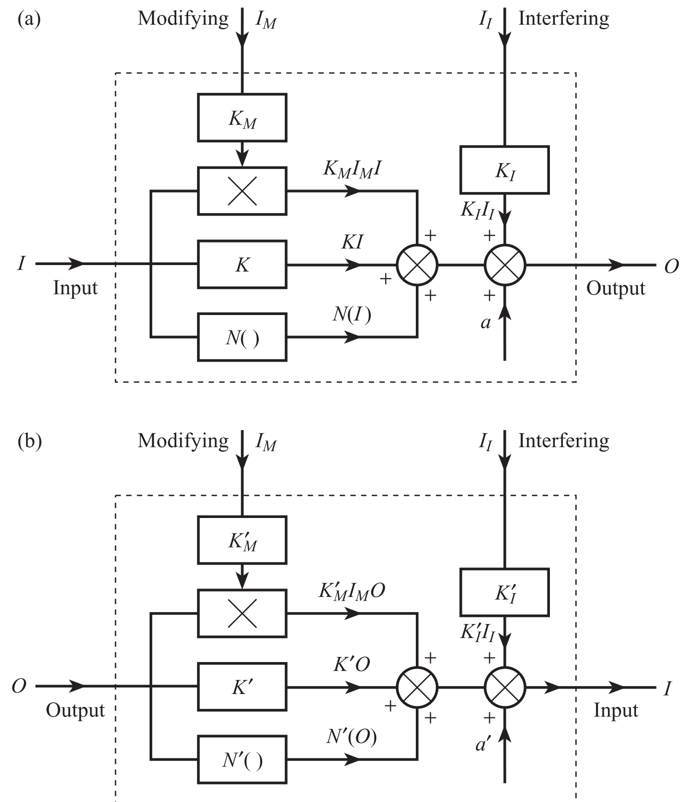


Figure 3.7 Models for system element: (a) Direct (b) Inverse.

giving greater linearity and less susceptibility to environmental inputs. A commonly used device which employs this principle is discussed in Section 9.4.

The rapid fall in the cost of digital integrated circuits in recent years has meant that microcontrollers are now widely used as signal-processing elements in measurement systems (Chapter 10). This means that the powerful techniques of **computer estimation of measured value** can now be used. For this method, a good model of the elements in the system is required. In Sections 2.1 and 2.2 we saw that the steady-state output O of an element is in general given by an equation of the form:

Direct equation

$$O = KI + a + N(I) + K_M I_M I + K_I I_I \qquad [2.9]$$

This is the **direct equation** (Figure 3.7(a)); here O is the dependent variable, which is expressed in terms of the independent variables I, I_M and I_I. In Section 2.4.2 we saw how the direct equation could be derived from sets of data obtained in a calibration experiment. In Section 3.2 eqn [2.9] was used to derive the error probability density function for a complete measurement system.

The steady-state characteristics of an element can also be represented by an alternative equation. This is the **inverse equation** (Figure 3.7(b)); here the signal input I is the dependent variable and the output O and environmental inputs I_I and I_M are the independent variables. The general form of this equation is:

Inverse equation

$$I = K'O + N'(O) + a' + K'_M I_M O + K'_I I \qquad [3.17]$$

where the values of K', $N'(\ )$, a', etc., are quite different from those for the direct equation. For example, the direct and inverse equations for a copper–constantan (type T) thermocouple with reference junction at 0 °C are:

DIRECT

$$E = 3.845 \times 10^{-2}T + 4.682 \times 10^{-5}T^2 - 3.789 \times 10^{-8}T^3 + 1.652 \times 10^{-11}T^4 \text{ mV}$$

INVERSE

$$T = 25.55E - 0.5973E^2 + 2.064 \times 10^{-2}E^3 - 3.205 \times 10^{-4}E^4 \text{ °C} \qquad [3.18]$$

where E is the thermocouple e.m.f. and T the measured junction temperature between 0 and 400 °C. Both equations were derived using a least squares polynomial fit to IEC 584.1 data; for the direct equation E is the dependent variable and T the independent variable; for the inverse equation T is the dependent variable and E the independent variable. While the direct equation is more useful for **error estimation**, the inverse equation is more useful for **error reduction**.

The use of the inverse equation in computer estimation of measured value is best implemented in a number of stages; with reference to Figure 3.8(a), these are:

1. Treat the uncompensated system as a single element. Using the calibration procedure of Section 2.4.2 (or any other method of generating data) the parameters K', a', etc. in the inverse model equation:

$$I = K'U + N'(U) + a' + K'_M I_M U + K'_I I_I \qquad [3.19]$$

   representing the overall behaviour of the uncompensated system can be found. This procedure will enable major environmental inputs I_M and I_I to be identified (there may be more than one of each type).
2. The uncompensated system should be connected to the estimator. This consists firstly of a computer which stores the model parameters K', a', $N'(\ )$, etc. If errors due to environmental inputs are considered significant, then environmental sensors to provide the computer with estimates I'_M, I'_I, of these inputs are also necessary. The output U of the uncompensated system is also fed to the computer.
3. The computer then calculates an initial estimate I' of I using the inverse equation:

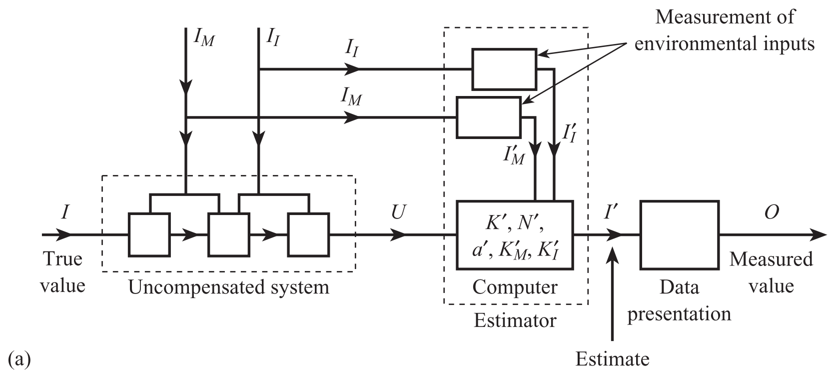
$$I' = K'U + N'(U) + a' + K'_M I'_M U + K'_I I'_I \qquad [3.20]$$

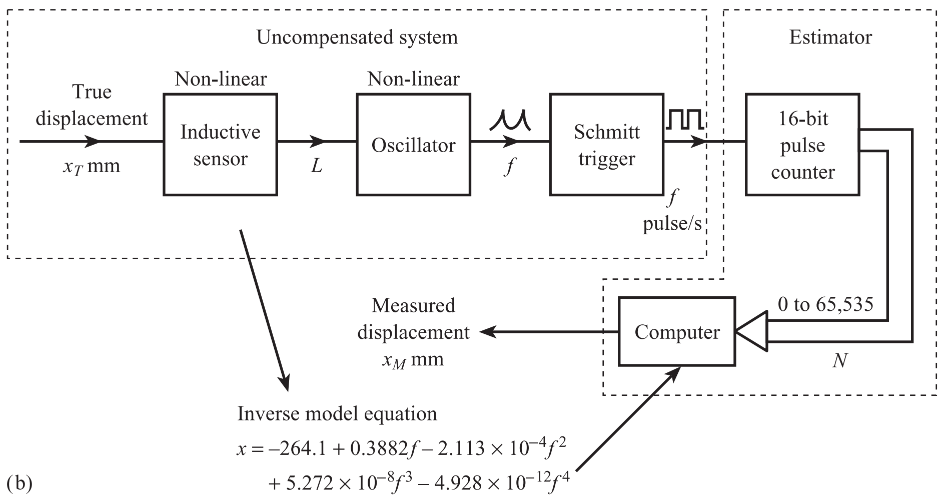
4. The data presentation element then displays the measured value O, which should be close to I'. In applications not requiring the highest accuracy the procedure can be terminated at this stage.
5. If high accuracy is required, then it may be possible to further improve the estimator by calibrating the complete system. Values of system output O are measured for a range of known standard inputs I and the corresponding values of system error $E = O - I$ are calculated. These error values will be mainly due to random effects but may also contain a small systematic component which can be corrected for.
6. An attempt should now be made to fit the data set (O_i, E_i), $i = 1, 2, \ldots n$, by a least squares straight line of the form:

$$E = kO + b \qquad [3.21]$$

   where b is any residual zero error and k specifies any residual scale error. There is little point in attempting a polynomial fit at this stage.

Figure 3.8 Computer estimation of measured value using inverse model equation:
(a) Principles
(b) Example of displacement measurement system.





7. The correlation coefficient:

$$r = \frac{\sum_{i=1}^{i=n} O_i E_i}{\sqrt{\sum_{i=1}^{i=n} O_i^2 \times \sum_{i=1}^{i=n} E_i^2}} \qquad [3.22]$$

between E and O data should now be evaluated. If the magnitude of r is greater than 0.5, then there is reasonable correlation between the E and O data; this means the systematic error of eqn [3.7] is present and we can proceed to stage 8 to correct for it. If the magnitude of r is less than 0.5, then there is no correlation between the E and O data; this means that the errors E are purely random and no correction can be made.

8. If appropriate, eqn [3.21] can be used to calculate an improved measured value:

$$O' = O - E = O - (kO + b) \qquad [3.23]$$

The displacement measurement system of Figure 3.8(b) shows this method. The uncompensated system consists of an inductive displacement sensor, an oscillator (Section 9.5) and a Schmitt trigger (Section 10.1.4). The sensor has a non-linear relation between inductance L and displacement x, the oscillator a non-linear relation between frequency f and inductance L. This means that the inverse model equation, relating displacement x and frequency f of the Schmitt trigger output signal, has the non-linear form shown. The estimator consists of a 16-bit pulse counter and a computer. The computer reads the state of the counter at the beginning and end of a fixed time interval and thus measures the frequency f of the pulse signal. The computer then calculates x from the inverse model equation using model coefficients stored in memory.

Conclusion

This chapter has shown how to find the **error** of a complete measurement system under **steady-state** conditions. **Measurement error** was first defined and then the **error probability density function** was derived, firstly for a general system of non-ideal elements and then for the typical example of a temperature measurement system. The last section discussed a range of methods for **error reduction**.

Reference

[1] KENNEDY J B and NEVILLE A 1986 *Basic Statistical Methods for Engineers and Scientists*, 3rd edn, pp. 345–60. Harper and Row, New York.

Problems

3.1 A measurement system consists of a chromel–alumel thermocouple (with cold junction compensation), a millivolt-to-current converter and a recorder. Table Prob. 1 gives the model equations, and parameters for each element. Assuming that all probability distributions are normal, calculate the mean and standard deviation of the error probability distribution, when the input temperature is 117 °C.

Table Prob. 1.

	Chromel–alumel thermocouple	e.m.f-to-current converter	Recorder
Model equation	$E = C_0 + C_1T + C_2T^2$	$i = K_1E + K_ME\Delta T_a + K_I\Delta T_a + a_1$	$T_M = K_2 i + a_2$
Mean values	$\bar{C}_0 = 0.00$ $\bar{C}_1 = 4.017 \times 10^{-2}$ $\bar{C}_2 = 4.66 \times 10^{-6}$	$\bar{K}_1 = 3.893$ $\Delta T_a = -10$ $\bar{a}_1 = -3.864$ $\bar{K}_M = 1.95 \times 10^{-4}$ $\bar{K}_I = 2.00 \times 10^{-3}$	$\bar{K}_2 = 6.25$ $\bar{a}_2 = 25.0$
Standard deviations	$\sigma_{C_0} = 6.93 \times 10^{-2}$ $\sigma_{C_1} = \sigma_{C_2} = 0$	$\sigma_{a_1} = 0.14$, $\sigma_{\Delta T_a} = 10$ $\sigma_{K_1} = \sigma_{K_M} = \sigma_{K_I} = 0$	$\sigma_{a_2} = 0.30$ $\sigma_{K_2} = 0.0$

Table Prob. 2.

Element	Linear sensitivity K	Error bandwidth $\pm h$
Pressure sensor	$10^{-4}\ \Omega\ \text{Pa}^{-1}$	$\pm 0.005\ \Omega$
Deflection bridge	$4 \times 10^{-2}\ \text{mV}\ \Omega^{-1}$	$\pm 5 \times 10^{-4}\ \text{mV}$
Amplifier	$10^3\ \text{mV mV}^{-1}$	$\pm 0.5\ \text{mV}$
Recorder	$250\ \text{Pa mV}^{-1}$	$\pm 100\ \text{Pa}$

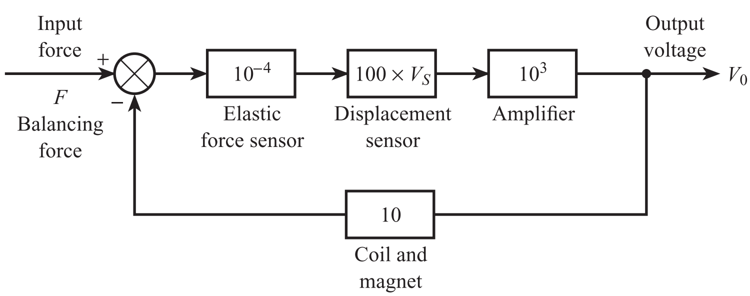
3.2 A pressure measurement system consists of a pressure sensor, deflection bridge, amplifier and recorder. Table Prob. 2 gives the linear sensitivities and error bandwidths for each element in the system.

(a) Calculate the standard deviation σ_E of the error distribution function.
(b) Given that the recorder is incorrectly adjusted so that its sensitivity is 225 Pa mV^{-1}, calculate the mean error $\bar{E}$ for an input pressure of 5×10^3 Pa.

3.3 Figure Prob. 3 shows a block diagram of a force transducer using negative feedback. The elastic sensor gives a displacement output for a force input; the displacement sensor gives a voltage output for a displacement input. V_S is the supply voltage for the displacement sensor.

(a) Calculate the output voltage V_0 when
 (i) $V_S = 1.0$ V, $F = 50$ N
 (ii) $V_S = 1.5$ V, $F = 50$ N.
(b) Comment on the practical significance of the variation of the supply voltage V_S.

Figure Prob. 3.

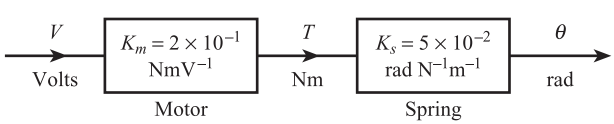


3.4 Figure Prob. 4 is a block diagram of a voltmeter. The motor produces a torque T proportional to voltage V, and the output angular displacement θ of the spring is proportional to T. The stiffness K_s of the spring can vary by ±10% about the nominal value of 5×10^{-2} rad N^{-1}m^{-1}. Given that the following are available:

 (i) a d.c. voltage amplifier of gain 1000
 (ii) a voltage subtraction unit
 (iii) a stable angular displacement transducer of sensitivity 100 V rad^{-1},

(a) draw a block diagram of a modified system using these components, which reduces the effect of changes in K_s;
(b) calculate the effect of a 10% increase in K_s on the sensitivity of the **modified** system.

Figure Prob. 4.



3.5 A temperature measurement system consists of a thermocouple, an amplifier, an 8-bit analogue-to-digital converter and a microcontroller with display facilities. Table Prob. 5 gives the model equations and parameters for each element in the system. The temperature of the thermocouple measurement junction is T_1 °C and the temperature of the reference junction is T_2 °C. The microcontroller corrects for T_2 having a non-zero mean value.

(a) Estimate the mean and standard deviation of the error probability density function when the input temperature T_1 is 100 °C. Treat rectangular distribution as normal with $\sigma = h/\sqrt{3}$.

(b) Explain briefly what modifications should be made to the system to reduce the quantities calculated in (a).

Table Prob. 5.

	Thermocouple	Amplifier	Analogue-to-digital converter	Microcontroller with display
Model equations	$E_{T_1,T_2} = a_0 + a_1(T_1 - T_2) + a_2(T_1^2 - T_2^2)$	$V = K_1 E_{T_1,T_2} + b_1$	$n = K_2 V + b_2$	$T_m = K_3 n + b_3$
Mean values	$\bar{a}_0 = 0$ $\bar{a}_1 = 4.3796 \times 10^{-2}$ $\bar{a}_2 = -1.7963 \times 10^{-5}$ $\bar{T}_2 = 20$	$\bar{K}_1 = 255$ $\bar{b}_1 = 0.0$	$\bar{K}_2 = 0.1$ $\bar{b}_2 = 0.0$ n rounded to nearest integer	$\bar{K}_3 = 1.0$ $\bar{b}_3 = 20$
Statistical distributions	Normal with $\sigma_{a_0} = 0.05$ $\sigma_{T_2} = 2.0$ $\sigma_{a_1} = \sigma_{a_2} = 0.0$	Normal with $\sigma_{b_1} = 5.0$ $\sigma_{K_1} = 0.0$	b_2 has a **rectangular distribution** of width ±0.5 $\sigma_{K_2} = 0.0$	$\sigma_{K_3} = 0.0$ $\sigma_{b_3} = 0.0$

3.6 A fluid velocity measurement system consists of a pitot tube, a differential pressure transmitter, an 8-bit analogue-to-digital converter and a microcontroller with display facilities. Table Prob. 6 gives the model equations and parameters for each element in the system. The microcontroller calculates the measured value of velocity assuming a constant density.

(a) Estimate the mean and standard deviation of the error probability density function assuming the true value of velocity v_T is 14.0 m s^{-1}. Use the procedure of Table 3.1, treating the rectangular distributions as normal with $\sigma = h/\sqrt{3}$.

(b) Explain briefly what modifications could be made to the system to reduce the quantities calculated in (a).

Table Prob. 6.

	Pitot tube	Differential pressure transmitter	Analogue-to-digital converter	Microcontroller with display
Model equations	$\Delta P = \frac{1}{2}\rho v_T^2$	$i = K_1 \Delta P + a_1$	$n = K_2 i + a_2$	$v_M = K_3\sqrt{n - 51}$
Mean values	$\bar{\rho} = 1.2$	$\bar{K}_1 = 0.064$ $\bar{a}_1 = 4.0$	$\bar{K}_2 = 12.80$ $\bar{a}_2 = 0.0$ n rounded off to nearest integer	$\bar{K}_3 = 1.430$
Half-widths of rectangular distribution	$h_\rho = 0.1$	$h_{a_1} = 0.04$	$h_{a_2} = 0.5$	$h_{K_3} = 0.0$

3.7 A temperature measurement system consists of a thermistor, bridge and recorder. Table Prob. 7 gives the model equations, mean values and standard deviations for each element in the system. Assuming all probability distributions are normal, calculate the mean and standard deviation of the error probability density function for a true input temperature of 320 K.

Table Prob. 7.

	Thermistor	Bridge	Recorder
Model equations	$R_\theta = K_1 \exp\left(\frac{\beta}{\theta}\right)$	$V_0 = V_S \left\{ \frac{1}{1 + \frac{3.3}{R_\theta}} - a_1 \right\}$	$\theta_M = K_2 V_0 + a_2$
Mean values	$\bar{K}_1 = 5 \times 10^{-4}$ kΩ $\bar{\beta} = 3 \times 10^3$ K	$\bar{V}_S = -3.00$ V $\bar{a}_1 = 0.77$	$\bar{K}_2 = 50.0$ K/V $\bar{a}_2 = 300$ K
Standard deviations	$\sigma_{K_1} = 0.5 \times 10^{-4}$ $\sigma_\beta = 0$	$\sigma_{V_S} = 0.03$ $\sigma_{a_1} = 0.01$	$\sigma_{K_2} = 0.0$ $\sigma_{a_2} = 3.0$

Basic problems

3.8 A force measurement system consists of four elements with sensitivities 10^{-2}, 5×10^{-2}, 10^3 and 1.9. Find the system error for a true value input of 10 kN.

3.9 A level measurement system consists of three linear elements with sensitivities of 0.050, 21.5 and 0.99. Find the system error for a true value input of 5.0 metres.

4 Dynamic Characteristics of Measurement Systems

If the input signal I to an element is changed suddenly, from one value to another, then the output signal O will not instantaneously change to its new value. For example, if the temperature input to a thermocouple is suddenly changed from 25 °C to 100 °C, some time will elapse before the e.m.f. output completes the change from 1 mV to 4 mV. The ways in which an element responds to sudden input changes are termed its **dynamic characteristics**, and these are most conveniently summarised using a **transfer function** $G(s)$. The first section of this chapter examines the dynamics of typical elements and derives the corresponding transfer function. The next section examines how standard test signals can be used to identify $G(s)$ for an element. If the input signal to a multi-element measurement system is changing rapidly, then the waveform of the system output signal is in general different from that of the input signal. Section 4.3 explains how this **dynamic error** can be found, and the final section outlines **dynamic compensation** methods that can be used to minimise errors.

4.1 Transfer function *G*(*s*) for typical system elements

4.1.1 First-order elements

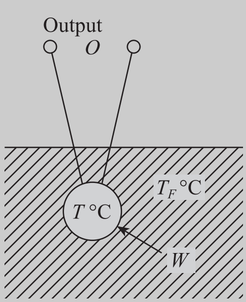


Figure 4.1 Temperature sensor in fluid.

A good example of a first-order element is provided by a temperature sensor with an electrical output signal, e.g. a thermocouple or thermistor. The bare element (not enclosed in a sheath) is placed inside a fluid (Figure 4.1). Initially at time $t = 0-$ (just before $t = 0$), the sensor temperature is equal to the fluid temperature, i.e. $T(0-) = T_F(0-)$. If the fluid temperature is suddenly raised at $t = 0$, the sensor is no longer in a steady state, and its dynamic behaviour is described by the **heat balance equation**:

$$\textbf{rate of heat inflow} - \textbf{rate of heat outflow} = \begin{array}{l}\textbf{rate of change of}\\ \textbf{sensor heat content}\end{array} \qquad [4.1]$$

Assuming that $T_F > T$, then the rate of heat outflow will be zero, and the rate of heat inflow W will be proportional to the temperature difference $(T_F - T)$. From Chapter 14 we have

$$W = UA(T_F - T) \text{ watts} \tag{4.2}$$

where U W m^{-2} °C^{-1} is the overall heat transfer coefficient between fluid and sensor and A m^2 is the effective heat transfer area. The increase of heat content of the sensor is $MC[T - T(0-)]$ joules, where M kg is the sensor mass and C J kg^{-1} °C^{-1} is the specific heat of the sensor material. Thus, assuming M and C are constants:

$$\text{rate of increase of sensor heat content} = MC\frac{\mathrm{d}}{\mathrm{d}t}[T - T(0-)] \tag{4.3}$$

Defining $\Delta T = T - T(0-)$ and $\Delta T_F = T_F - T_F(0-)$ to be the deviations in temperatures from initial steady-state conditions, the differential equation describing the sensor temperature changes is

$$UA(\Delta T_F - \Delta T) = MC\frac{\mathrm{d}\Delta T}{\mathrm{d}t}$$

i.e.

$$\frac{MC}{UA}\frac{\mathrm{d}\Delta T}{\mathrm{d}t} + \Delta T = \Delta T_F \tag{4.4}$$

This is a **linear differential equation** in which $\mathrm{d}\Delta T/\mathrm{d}t$ and ΔT are multiplied by constant coefficients; the equation is **first order** because $\mathrm{d}\Delta T/\mathrm{d}t$ is the highest derivative present. The quantity MC/UA has the dimensions of time:

$$\frac{\text{kg} \times \text{J} \times \text{kg}^{-1} \times {}^\circ\text{C}^{-1}}{\text{W} \times \text{m}^{-2} \times {}^\circ\text{C}^{-1} \times \text{m}^2} = \frac{\text{J}}{\text{W}} = \text{seconds}$$

and is referred to as the **time constant** τ for the system. The differential equation is now:

Linear first-order differential equation

$$\tau\frac{\mathrm{d}\Delta T}{\mathrm{d}t} + \Delta T = \Delta T_F \tag{4.5}$$

While the above differential equation is a perfectly adequate description of the dynamics of the sensor, it is not the most useful representation. The transfer function based on the Laplace transform of the differential equation provides a convenient framework for studying the dynamics of multi-element systems. The Laplace transform $\bar{f}(s)$ of a time-varying function is defined by:

Definition of Laplace transform

$$\bar{f}(s) = \int_0^\infty \mathrm{e}^{-st} f(t)\,\mathrm{d}t \tag{4.6}$$

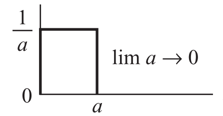
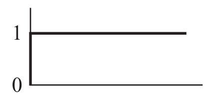
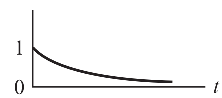
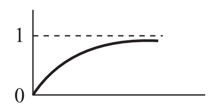
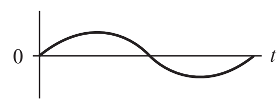
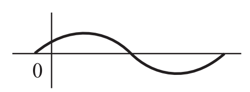
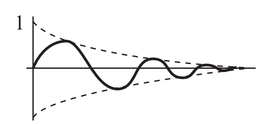
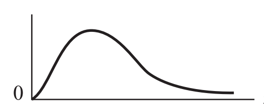
where s is a complex variable of the form $\sigma + j\omega$ where $j = \sqrt{-1}$.

Table 4.1 gives Laplace transforms for some common standard functions $f(t)$. In order to find the transfer function for the sensor we must find the Laplace transform of eqn [4.5]. Using Table 4.1 we have:

$$\tau[s\Delta\bar{T}(s) - \Delta T(0-)] + \Delta\bar{T}(s) = \Delta\bar{T}_F(s) \tag{4.7}$$

Table 4.1 Laplace transforms of common time functions $f(t)$.[a]

$$\mathcal{L}[f(t)] = \bar{f}(s) = \int_0^\infty e^{-st} f(t)\,dt$$

Function	Symbol	Graph	Transform
1st derivative	$\frac{d}{dt}f(t)$		$s\bar{f}(s) - f(0-)$
2nd derivative	$\frac{d^2}{dt^2}f(t)$		$s^2\bar{f}(s) - sf(0-) - \dot{f}(0-)$
Unit impulse	$\delta(t)$		1
Unit step	$\mu(t)$		$\frac{1}{s}$
Exponential decay	$\exp(-\alpha t)$		$\frac{1}{s+\alpha}$
Exponential growth	$1 - \exp(-\alpha t)$		$\frac{\alpha}{s(s+\alpha)}$
Sine wave	$\sin \omega t$		$\frac{\omega}{s^2+\omega^2}$
Phase-shifted sine wave	$\sin(\omega t + \phi)$		$\frac{\omega\cos\phi + s\sin\phi}{s^2+\omega^2}$
Exponentially damped sine wave	$\exp(-\alpha t)\sin \omega t$		$\frac{\omega}{(s+\alpha)^2+\omega^2}$
Ramp with exponential decay	$t\exp(-\alpha t)$		$\frac{1}{(s+\alpha)^2}$

[a] Initial conditions are at $t = 0-$, just prior to $t = 0$.

where $\Delta T(0-)$ is the temperature deviation at initial conditions prior to $t = 0$. By definition, $\Delta T(0-) = 0$, giving:

$$\tau s \Delta\bar{T}(s) + \Delta\bar{T}(s) = \Delta\bar{T}_F(s)$$

i.e.

$$(\tau s + 1)\Delta\bar{T}(s) = \Delta\bar{T}_F(s) \qquad [4.8]$$

The transfer function $G(s)$ of an element is defined as the ratio of the Laplace transform of the output to the Laplace transform of the input, provided the initial conditions are zero. Thus:

Definition of element transfer function

$$G(s) = \frac{\bar{f}_0(s)}{\bar{f}_i(s)} \qquad [4.9]$$

and $\bar{f}_0(s) = G(s)\bar{f}_i(s)$; this means the transfer function of the output signal is simply the product of the element transfer function and the transfer function of the input signal. Because of this simple relationship the transfer function technique lends itself to the study of the dynamics of multi-element systems and block diagram representation (Figure 4.2).

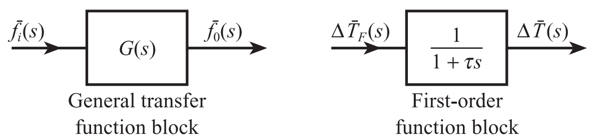


Figure 4.2 Transfer function representation.

From eqns [4.8] and [4.9] the transfer function for a first-order element is:

Transfer function for a first-order element

$$G(s) = \frac{\Delta\bar{T}(s)}{\Delta\bar{T}_F(s)} = \frac{1}{1+\tau s} \qquad [4.10]$$

The above transfer function only relates changes in sensor temperature to changes in fluid temperature. The overall relationship between changes in sensor output signal O and fluid temperature is:

$$\frac{\Delta\bar{O}(s)}{\Delta\bar{T}_F(s)} = \frac{\Delta O}{\Delta T}\frac{\Delta\bar{T}(s)}{\Delta\bar{T}_F(s)} \qquad [4.11]$$

where $\Delta O/\Delta T$ is the **steady-state sensitivity** of the temperature sensor. For an ideal element $\Delta O/\Delta T$ will be equal to the slope K of the **ideal straight line**. For non-linear elements, subject to small temperature fluctuations, we can take $\Delta O/\Delta T = \mathrm{d}O/\mathrm{d}T$, the derivative being evaluated at the steady-state temperature $T(0-)$ around which the fluctuations are taking place. Thus for a copper–constantan thermocouple measuring small fluctuations in temperature around 100 °C, $\Delta E/\Delta T$ is found by evaluating $\mathrm{d}E/\mathrm{d}T$ at 100 °C (see Section 2.1) to give $\Delta E/\Delta T = 35\ \mu\mathrm{V}\ {}^\circ\mathrm{C}^{-1}$. Thus if the time constant of the thermocouple is 10 s the overall dynamic relationship between changes in e.m.f. and fluid temperature is:

$$\frac{\Delta\bar{E}(s)}{\Delta\bar{T}_F(s)} = 35 \times \frac{1}{1+10s} \qquad [4.12]$$

In the general case of an element with static characteristics given by eqn [2.9] and dynamic characteristics defined by $G(s)$, the effect of small, rapid changes in ΔI is evaluated using Figure 4.3, in which steady-state sensitivity $(\partial O/\partial I)_{I_0} = K + K_M I_M + (\mathrm{d}N/\mathrm{d}I)_{I_0}$, and I_0 is the steady-state value of I around which the fluctuations are

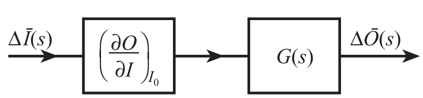
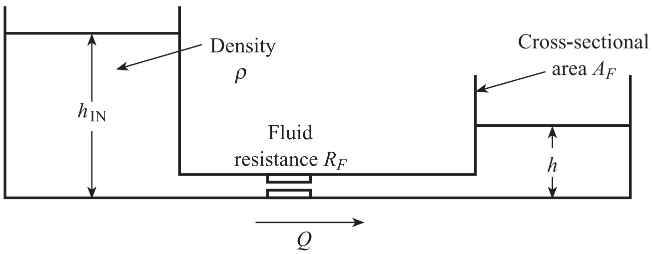


Figure 4.3 Element model for dynamic calculations.

Table 4.2 Analogous first-order elements.

Fluidic



Volume flow rate $Q = \frac{1}{R_F}(P_{IN} - P)$

Pressures $P_{IN} = h_{IN}\rho g$, $P = h\rho g$

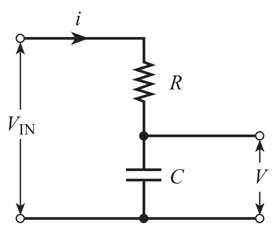
$$A_F \frac{dh}{dt} = Q = \frac{\rho g}{R_F}(h_{IN} - h)$$

$$\frac{A_F R_F}{\rho g}\frac{dh}{dt} + h = h_{IN}$$

i.e.

$$\tau_F \frac{dh}{dt} + h = h_{IN}, \ \tau_F = \frac{A_F R_F}{\rho g}$$

Electrical



$V_{IN} - V = iR$

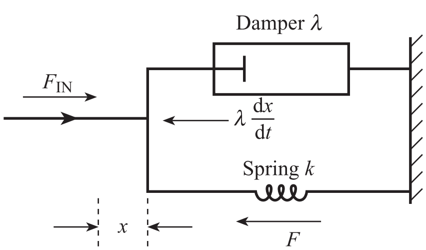
Charge $q = CV$, current $i = \frac{dq}{dt} = \frac{CdV}{dt}$

$$RC\frac{dV}{dt} + V = V_{IN}$$

i.e.

$$\tau_E \frac{dV}{dt} + V = V_{IN}, \ \tau_E = RC$$

Mechanical



$F_{IN} - F = \lambda \frac{dx}{dt}$, λ N s m^{-1} = damping constant

Displacement $x = \frac{F}{k}$, k N m^{-1} = spring stiffness

$$\frac{\lambda}{k}\frac{dF}{dt} + F = F_{IN}$$

$$\tau_M \frac{dF}{dt} + F = F_{IN}, \ \tau_M = \frac{\lambda}{k}$$

Thermal	$\tau_{Th} = \frac{MC}{UA} = R_{Th}C_{Th}$; $R_{Th} = \frac{1}{UA}$, $C_{Th} = MC$
Fluidic	$\tau_F = \frac{A_F R_F}{\rho g} = R_F C_F$; $R_F = R_F$, $C_F = \frac{A_F}{\rho g}$
Electrical	$\tau_E = RC = R_E C_E$; $R_E = R$, $C_E = C$
Mechanical	$\tau_M = \frac{\lambda}{k} = R_M C_M$; $R_M = \lambda$, $C_M = \frac{1}{k}$

taking place. Table 4.2 shows analogous fluidic, electrical and mechanical elements, which are also described by a first-order transfer function $G(s) = 1/(1 + \tau s)$. All four elements are characterised by 'resistance' and 'capacitance' as illustrated in the table. Temperature, pressure, voltage and force are analogous 'driving' or effort variables; heat flow rate, volume flow rate, current and velocity are analogous 'driven' or flow variables. These analogies are discussed further in Section 5.2.

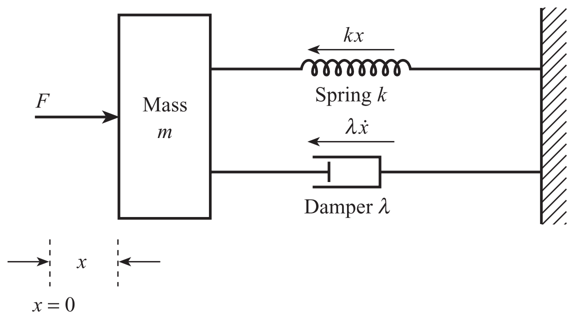


Figure 4.4 Mass–spring–damper model of elastic force sensor.

4.1.2 Second-order elements

The elastic sensor shown in Figure 4.4, which converts a force input F into a displacement output x, is a good example of a second-order element. The diagram is a conceptual model of the element, which incorporates a mass m kg, a spring of stiffness k N m^{-1}, and a damper of constant λ N s m^{-1}. The system is initially at rest at time $t = 0-$ so that the initial velocity $\dot{x}(0-) = 0$ and the initial acceleration $\ddot{x}(0-) = 0$. The initial input force $F(0-)$ is balanced by the spring force at the initial displacement $x(0-)$, i.e.

$$F(0-) = kx(0-) \qquad [4.13]$$

If the input force is suddenly increased at $t = 0$, then the element is no longer in a steady state and its dynamic behaviour is described by Newton's second law, i.e.

resultant force = mass × acceleration

i.e.

$$F - kx - \lambda\dot{x} = m\ddot{x} \qquad [4.14]$$

and

$$m\ddot{x} + \lambda\dot{x} + kx = F$$

Defining ΔF and Δx to be the deviations in F and x from initial steady-state conditions:

$$\Delta F = F - F(0-), \qquad \Delta x = x - x(0-) \qquad [4.15]$$
$$\Delta\dot{x} = \dot{x}, \qquad \Delta\ddot{x} = \ddot{x}$$

the differential equation now becomes:

$$m\Delta\ddot{x} + \lambda\Delta\dot{x} + kx(0-) + k\Delta x = F(0-) + \Delta F$$

which, using [4.13], reduces to:

$$m\Delta\ddot{x} + \lambda\Delta\dot{x} + k\Delta x = \Delta F$$

i.e.

$$\frac{m}{k}\frac{d^2\Delta x}{dt^2} + \frac{\lambda}{k}\frac{d\Delta x}{dt} + \Delta x = \frac{1}{k}\Delta F \qquad [4.16]$$

This is a **second-order linear differential equation** in which Δx and its derivatives are multiplied by constant coefficients and the highest derivative present is $d^2\Delta x/dt^2$. If we define

undamped natural frequency $\omega_n = \sqrt{\dfrac{k}{m}}$ rad/s and

$$\textbf{damping ratio } \xi = \frac{\lambda}{2\sqrt{km}} \qquad [4.17]$$

then $m/k = 1/\omega_n^2$, $\lambda/k = 2\xi/\omega_n$, and [4.16] can be expressed in the standard form:

Linear second-order differential equation

$$\frac{1}{\omega_n^2}\frac{d^2\Delta x}{dt^2} + \frac{2\xi}{\omega_n}\frac{d\Delta x}{dt} + \Delta x = \frac{1}{k}\Delta F \qquad [4.18]$$

In order to find the transfer function for the element we require the Laplace transform of eqn [4.18]. Using Table 4.1 we have:

$$\frac{1}{\omega_n^2}[s^2\Delta\bar{x}(s) - s\Delta x(0-) - \Delta\dot{x}(0-)] + \frac{2\xi}{\omega_n}[s\Delta\bar{x}(s) - \Delta x(0-)] + \Delta\bar{x}(s)$$
$$= \frac{1}{k}\Delta\bar{F}(s) \qquad [4.19]$$

Since $\Delta\dot{x}(0-) = \dot{x}(0-) = 0$ and $\Delta x(0-) = 0$ by definition, [4.19] reduces to:

$$\left[\frac{1}{\omega_n^2}s^2 + \frac{2\xi}{\omega_n}s + 1\right]\Delta\bar{x}(s) = \frac{1}{k}\Delta\bar{F}(s) \qquad [4.20]$$

Thus

$$\frac{\Delta\bar{x}(s)}{\Delta\bar{F}(s)} = \frac{1}{k}G(s)$$

where $1/k$ = steady-state sensitivity K, and

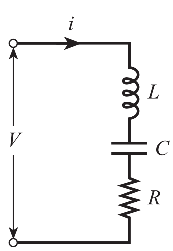
Transfer function for a second-order element

$$G(s) = \frac{1}{\frac{1}{\omega_n^2}s^2 + \frac{2\xi}{\omega_n}s + 1} \qquad [4.21]$$

Figure 4.5 shows an analogous electrical element, a series L–C–R circuit.

Comparing eqns [4.14] and [4.22] we see that q is analogous to x, V is analogous to F, and L, R and $1/C$ are analogous to m, λ and k respectively (see Table 5.1). The electrical circuit is also described by the above second-order transfer function with $\omega_n = 1/\sqrt{LC}$ and $\xi = (R/2)\sqrt{C/L}$.

Figure 4.5
Series L–C–R circuit.



$$V = iR + \frac{q}{C} + L\frac{di}{dt}$$

where $i = \frac{dq}{dt}$ (q = charge on the capacitance)

$$\text{thus } L\frac{d^2q}{dt^2} + R\frac{dq}{dt} + \frac{1}{C}q = V \qquad [4.22]$$

4.2 Identification of the dynamics of an element

In order to identify the transfer function $G(s)$ of an element, standard input signals should be used. The two most commonly used standard signals are step and sine wave. This section examines the response of first- and second-order elements to step and sine wave inputs.

4.2.1 Step response of first- and second-order elements

From Table 4.1 we see that the Laplace transform of a step of unit height $u(t)$ is $\bar{f}(s) = 1/s$. Thus if a first-order element with $G(s) = 1/(1 + \tau s)$ is subject to a unit step input signal, the Laplace transform of the element output signal is:

$$\bar{f}_o(s) = G(s)\bar{f}_i(s) = \frac{1}{(1 + \tau s)s} \qquad [4.23]$$

Expressing [4.23] in partial fractions, we have:

$$\bar{f}_o(s) = \frac{1}{(1 + \tau s)s} = \frac{A}{(1 + \tau s)} + \frac{B}{s}$$

Equating coefficients of constants gives $B = 1$, and equating coefficients of s gives $0 = A + B\tau$, i.e. $A = -\tau$. Thus:

$$\bar{f}_o(s) = \frac{1}{s} - \frac{\tau}{(1 + \tau s)} = \frac{1}{s} - \frac{1}{(s + 1/\tau)} \qquad [4.24]$$

Using Table 4.1 in reverse, i.e. finding a time signal $f(t)$ corresponding to a transform $\bar{f}(s)$, we have:

$$f_o(t) = u(t) - \exp\left(\frac{-t}{\tau}\right)$$

and since $u(t) = 1$ for $t > 0$:

Response of first-order element to unit step

$$f_o(t) = 1 - \exp\left(\frac{-t}{\tau}\right) \qquad [4.25]$$

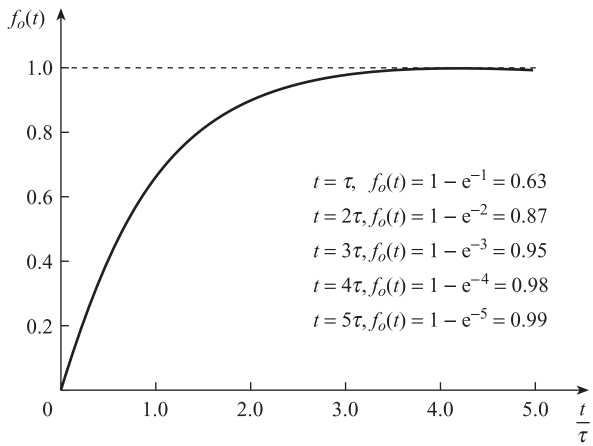
The form of the response is shown in Figure 4.6.

As an example of the use of eqn [4.25], consider the temperature sensor of Section 4.1.1. Initially the temperature of the sensor is equal to that of the fluid, i.e. $T(0-) = T_F(0-) = 25$ °C, say. If T_F is suddenly raised to 100 °C, then this represents a step change ΔT_F of height 75 °C. The corresponding *change* in sensor temperature is given by $\Delta T = 75(1 - e^{-t/\tau})$ and the actual temperature T of the sensor at time t is given by:

$$T(t) = 25 + 75(1 - e^{-t/\tau}) \qquad [4.26]$$

Thus at time $t = \tau$, $T = 25 + (75 \times 0.63) = 72.3$ °C. By measuring the time taken for T to rise to 72.3 °C we can find the time constant τ of the element.

Figure 4.6 Response of a first-order element to a unit step.



If the second-order element with transfer function

$$G(s) = \frac{1}{\dfrac{1}{\omega_n^2}s^2 + \dfrac{2\xi}{\omega_n}s + 1}$$

is subject to a unit step input signal, then the Laplace transform of the element output signal is:

$$\bar{f}_o(s) = \frac{1}{\left(\dfrac{1}{\omega_n^2}s^2 + \dfrac{2\xi}{\omega_n}s + 1\right)s} \qquad [4.27]$$

Expressing [4.27] in partial fractions we have:

$$\bar{f}_o(s) = \frac{As + B}{\left(\dfrac{1}{\omega_n^2}s^2 + \dfrac{2\xi}{\omega_n}s + 1\right)} + \frac{C}{s} \qquad [4.27]$$

where $A = -1/\omega_n^2$, $B = -2\xi/\omega_n$, $C = 1$. This gives:

$$\begin{aligned}\bar{f}_o(s) &= \frac{1}{s} - \frac{(s + 2\xi\omega_n)}{s^2 + 2\xi\omega_n s + \omega_n^2} \\ &= \frac{1}{s} - \frac{(s + 2\xi\omega_n)}{(s + \xi\omega_n)^2 + \omega_n^2(1 - \xi^2)} \\ &= \frac{1}{s} - \frac{(s + \xi\omega_n)}{(s + \xi\omega_n)^2 + \omega_n^2(1 - \xi^2)} - \frac{\xi\omega_n}{(s + \xi\omega_n)^2 + \omega_n^2(1 - \xi^2)} \qquad [4.28]\end{aligned}$$

There are three cases to consider, depending on whether ξ is greater than 1, equal to 1, or less than 1. For example, if $\xi = 1$ (**critical damping**) then:

$$\bar{f}_o(s) = \frac{1}{s} - \frac{1}{s + \omega_n} - \frac{\omega_n}{(s + \omega_n)^2} \qquad [4.29]$$

Using Table 4.1 we have:

Response of second-order element to a unit step, critical damping $\xi = 1$

$$f_o(t) = 1 - e^{-\omega_n t}(1 + \omega_n t) \quad [4.30]$$

Using standard tables it can be shown that if $\xi < 1$ **(underdamping)** then:

Second-order step response, underdamping $\xi < 1$

$$f_o(t) = 1 - e^{-\xi\omega_n t}\left[\cos \omega_n \sqrt{(1-\xi^2)}\,t + \frac{\xi}{\sqrt{1-\xi^2}} \sin \omega_n \sqrt{(1-\xi^2)}\,t\right] \quad [4.31]$$

and if $\xi > 1$ (**overdamping**) then:

Second-order step response, overdamping $\xi > 1$

$$f_o(t) = 1 - e^{-\xi\omega_n t}\left[\cosh \omega_n \sqrt{(\xi^2-1)}\,t + \frac{\xi}{\sqrt{\xi^2-1}} \sinh \omega_n \sqrt{(\xi^2-1)}\,t\right] \quad [4.32]$$

We consider the case of the underdamped response with $\xi < 1$, given by eqn [4.31]. Here the **damped angular frequency** of the oscillations is given by:

$$\omega_d = \omega_n\sqrt{1-\xi^2}$$

where ω_n is the natural or undamped angular frequency. The corresponding period T_d of the damped oscillations is given by:

$$T_d = \frac{2\pi}{\omega_d} = \frac{2\pi}{\omega_n\sqrt{1-\xi^2}}$$

The time T_p at which the first oscillation peak occurs is correspondingly given by:

$$T_p = T_d/2 = \frac{\pi}{\omega_n\sqrt{1-\xi^2}}$$

The **settling time** T_s taken for the response to settle out at the steady-state value is governed by the exponential decay term $e^{-\xi\omega_n t}$. When $t = 5/\xi\omega_n$, i.e. $\xi\omega_n t = 5$, then $e^{-5} = 0.0067$. The time for the response to settle out approximately within 1% is therefore:

$$T_s = 5/\xi\omega_n$$

From Figure 4.7, we see that T_s is minimum when $\xi = 0.7$. The response $f_o(t)$ has a maximum value f_o^p at the peak of the first oscillation. The difference $(f_o^p - 1)$ between maximum and steady-state values of $f_o(t)$, for $\xi < 1$, is termed the **maximum overshoot** and is given by:

$$\exp\left[\frac{-\pi\xi}{\sqrt{1-\xi^2}}\right]$$

and depends only on ξ. Thus for $\xi < 1$, an estimate of ξ can be found from measurement of maximum overshoot and then, knowing ξ, ω_n can be estimated from measurements of T_d, T_p and/or T_s.

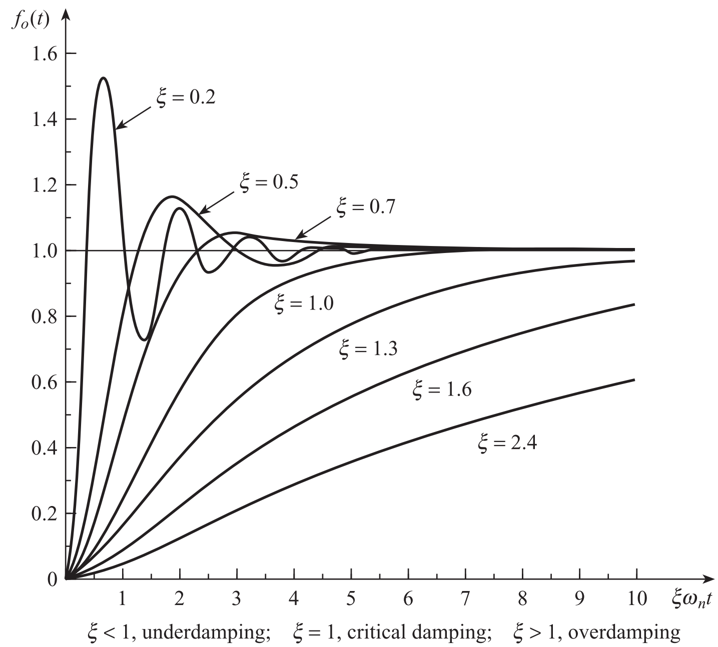


Figure 4.7 Response of a second-order element to a unit step.

The form of the responses is shown in Figure 4.7. As an example consider the step response of a force sensor with stiffness $k = 10^3$ N m^{-1}, mass $m = 10^{-1}$ kg and damping constant $\lambda = 10$ N s m^{-1}. The steady-state sensitivity $K = 1/k = 10^{-3}$ m N^{-1}, natural frequency $\omega_n = \sqrt{k/m} = 10^2$ rad s^{-1} and damping coefficient $\xi = \lambda/2\sqrt{km} = 0.5$. Initially at time $t = 0-$, a steady force $F(0-) = 10$ N causes a steady displacement of $(1/10^3) \times 10$ metre, i.e. 10 mm. Suppose that at $t = 0$ the force is suddenly increased from 10 to 12 N, i.e. there is a step change ΔF of 2 N. The resulting change $\Delta x(t)$ in displacement is found using

$$\Delta x(t) = \text{steady-state sensitivity} \times \text{step height} \times \text{unit step response } f_o(t) \quad [4.33]$$

i.e.

$$\Delta x(t) = \frac{1}{10^3} \times 2 \times [1 - e^{-50t}(\cos 86.6t + 0.58 \sin 86.6t)] \text{ metre}$$

$$= 2[1 - e^{-50t}(\cos 86.6t + 0.58 \sin 86.6t)] \text{ mm} \quad [4.34]$$

Here the damped angular frequency $\omega_d = 86.6$ rad/s and the period of the damped oscillations $T_d = 72.6$ ms. The maximum overshoot is 0.16, giving $f_o^p = 1.16$ so that the maximum value of Δx is 2.32 mm. This occurs at time $T_p = 36.3$ ms. Eventually as t becomes large Δx tends to 2 mm, i.e. x settles out at a new steady-state value of 12 mm with a settling time T_s of 100 ms.

4.2.2 Sinusoidal response of first- and second-order elements

From Table 4.1 we see that the Laplace transform of sine wave $f(t) = \sin \omega t$, with unit amplitude and angular frequency ω, is $\bar{f}(s) = \omega/(s^2 + \omega^2)$. Thus if a sine wave of amplitude $\hat{I}$ is input to a first-order element, then the Laplace transform of the output signal is

$$\bar{f}_o(s) = \frac{1}{(1+\tau s)}\frac{\hat{I}\cdot\omega}{(s^2+\omega^2)} \qquad [4.35]$$

Expressing [4.35] in partial fractions we have

$$\bar{f}_o(s) = \frac{A}{(1+\tau s)} + \frac{Bs+C}{s^2+\omega^2} \qquad [4.36]$$

where:

$$A = \frac{\omega\tau^2\hat{I}}{(1+\tau^2\omega^2)}, \quad B = \frac{-\omega\tau\hat{I}}{(1+\tau^2\omega^2)}, \quad C = \frac{\omega\hat{I}}{(1+\tau^2\omega^2)}$$

so that:

$$\bar{f}_o(s) = \frac{\omega\tau^2\hat{I}}{(1+\tau^2\omega^2)}\frac{1}{(1+\tau s)} + \frac{\hat{I}}{(1+\tau^2\omega^2)}\left\{\frac{-\omega\tau s+\omega}{s^2+\omega^2}\right\}$$

$$= \frac{\omega\tau^2\hat{I}}{(1+\tau^2\omega^2)}\frac{1}{(1+\tau s)} + \frac{\hat{I}}{\sqrt{1+\tau^2\omega^2}}\left\{\frac{\omega\dfrac{1}{\sqrt{1+\tau^2\omega^2}} + s\dfrac{-\omega\tau}{\sqrt{1+\tau^2\omega^2}}}{s^2+\omega^2}\right\}$$

$$= \frac{\omega\tau^2\hat{I}}{(1+\tau^2\omega^2)}\frac{1}{(1+\tau s)} + \frac{\hat{I}}{\sqrt{1+\tau^2\omega^2}}\left\{\frac{\omega\cos\phi + s\sin\phi}{s^2+\omega^2}\right\} \qquad [4.37]$$

where

$$\cos\phi = \frac{1}{\sqrt{1+\tau^2\omega^2}}, \quad \sin\phi = \frac{-\omega\tau}{\sqrt{1+\tau^2\omega^2}}$$

Using Table 4.1 we have:

$$f_o(t) = \underbrace{\frac{\omega\tau\hat{I}}{1+\tau^2\omega^2}e^{-t/\tau}}_{\text{Transient term}} + \underbrace{\frac{\hat{I}}{\sqrt{1+\tau^2\omega^2}}\sin(\omega t+\phi)}_{\text{Sinusoidal term}} \qquad [24.38]$$

In a sine wave test experiment, we wait until the transient term has decayed to zero and measure the sinusoidal signal:

$$f_o(t) = \frac{\hat{I}}{\sqrt{1+\tau^2\omega^2}}\sin(\omega t+\phi) \qquad [4.39]$$

which remains. We see therefore that the output signal is also a sine wave of frequency ω but with amplitude $\hat{I}/\sqrt{1+\tau^2\omega^2}$, and shifted in phase by $\phi = -\tan^{-1}(\omega\tau)$ relative to the input sine wave. These amplitude and phase results can be found directly from the transfer function $G(s) = 1/(1+\tau s)$ without having to use the table of transforms. If we replace s by $j\omega$ ($j = \sqrt{-1}$) in $G(s)$ we form the complex number $G(j\omega) = 1/(1+j\tau\omega)$. The **magnitude** $|G(j\omega)| = 1/\sqrt{1+\tau^2\omega^2}$ of this complex number is equal to the ratio of output amplitude to input amplitude, and the **angle** or **argument** $\arg G(j\omega) = -\tan^{-1}(\omega\tau)$ is equal to the phase difference ϕ between output and input sine waves. Figure 4.8 shows amplitude ratio versus frequency and phase versus frequency graphs for a first-order element; these are known as the **frequency response** characteristics of the element. From the above equations we see that when $\omega\tau = 1$,

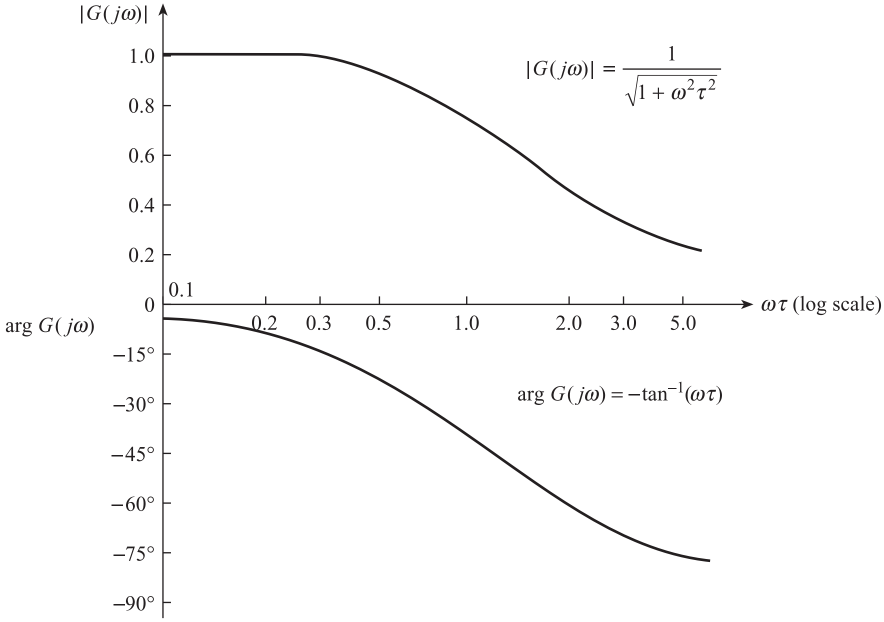


Figure 4.8 Frequency response characteristics of first-order element with $G(s) = \frac{1}{1+\tau s}$.

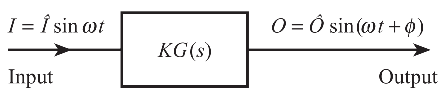


Figure 4.9 Frequency response of an element with linear dynamics.

i.e. $\omega = 1/\tau$, the amplitude ratio $= 1/\sqrt{2}$ and phase difference $\phi = -45°$. These results enable the value of τ to be found from experimental frequency response data.

The above results can be generalised to an element with steady-state sensitivity K (or $\partial O/\partial I$) and transfer function $G(s)$, subject to a sinusoidal input signal $I = \hat{I}\sin\omega t$ as in Figure 4.9. In the steady state we can make four statements about the output signal:

(a) O is also a sine wave.
(b) The frequency of O is also ω.
(c) The amplitude of O is $\hat{O} = K|G(j\omega)|\hat{I}$.
(d) The phase difference between O and I is $\phi = \arg G(j\omega)$.

Using the above rules we can quickly find the amplitude ratio and phase relations for a second-order element with:

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

Here we have:

$$G(j\omega) = \frac{1}{\frac{1}{\omega_n^2}(j\omega)^2 + \frac{2\xi}{\omega_n}(j\omega) + 1}$$

so that

Frequency response characteristics of second-order element

$$\textbf{Amplitude ratio} = |G(j\omega)| = \frac{1}{\sqrt{\left(1 - \frac{\omega^2}{\omega_n^2}\right)^2 + 4\xi^2\frac{\omega^2}{\omega_n^2}}}$$

$$\textbf{Phase difference} = \arg G(j\omega) = -\tan^{-1}\left[\frac{2\xi\omega/\omega_n}{1 - \omega^2/\omega_n^2}\right] \qquad [4.40]$$

These characteristics are shown graphically in Figure 4.10; both amplitude ratio and phase characteristics are critically dependent on the value of ξ.

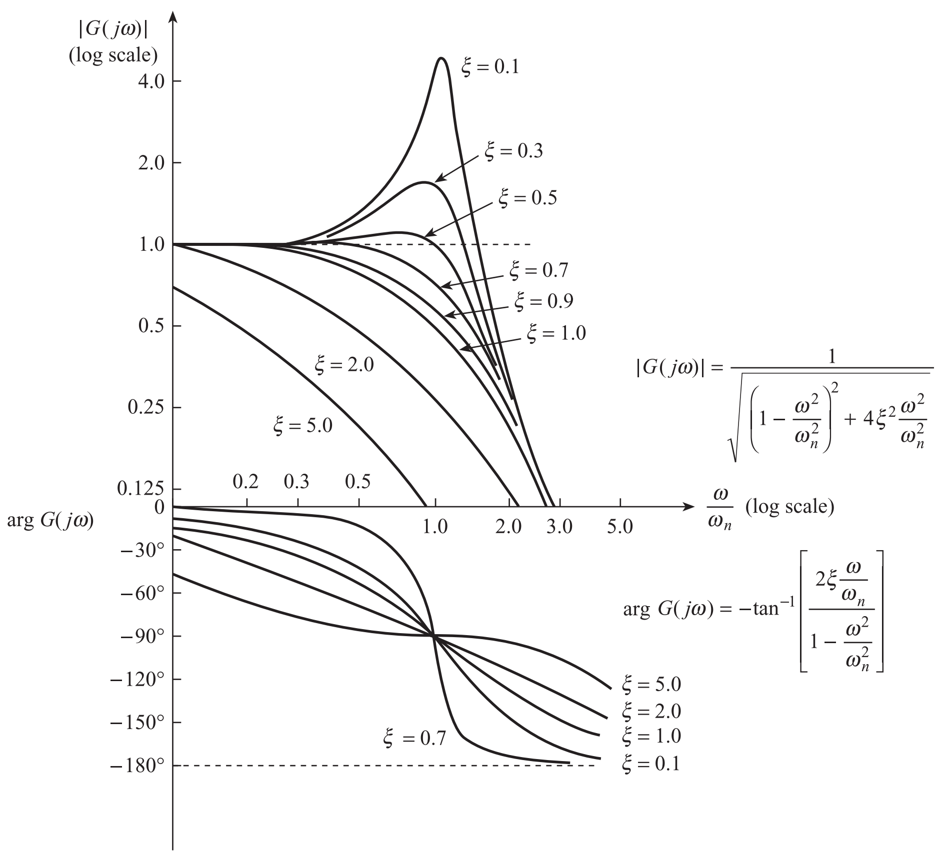
We note that for $\xi < 0.7$, $|G(j\omega)|$ has a maximum value which is greater than unity. This maximum value is given by:

$$|G(j\omega)|_{\text{MAX}} = \frac{1}{2\xi\sqrt{1-\xi^2}}$$

and occurs at the **resonant frequency** ω_R where:

$$\omega_R = \omega_n\sqrt{1-2\xi^2}, \quad \xi < 1/\sqrt{2}$$

Figure 4.10 Frequency response characteristics of second-order element with $G(s) = \dfrac{1}{\dfrac{1}{\omega_n^2}s^2 + \dfrac{2\xi}{\omega_n}s + 1}$



Thus by measuring $|G(j\omega)|_{MAX}$ and ω_R, ξ and ω_n can be found. An alternative to plotting $|G(j\omega)|$ versus ω is a graph of the number of decibels N dB versus ω, where $N = 20\log_{10}|G(j\omega)|$. Thus if $|G(j\omega)| = 1$, $N = 0$ dB; if $|G(j\omega)| = 10$, $N = +20$ dB; and if $|G(j\omega)| = 0.1$, $N = -20$ dB.

4.3 Dynamic errors in measurement systems

Figure 4.11 shows a complete measurement system consisting of n elements. Each element i has ideal steady-state and linear dynamic characteristics and can therefore be represented by a constant steady-state sensitivity K_i and a transfer function $G_i(s)$.

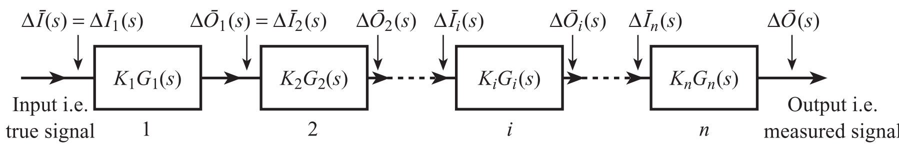


Figure 4.11 Complete measurement system with dynamics.

We begin by assuming that the steady-state sensitivity $K_1 K_2 \ldots K_i \ldots K_n$ for the overall system is equal to 1, i.e. the system has no steady-state error (Section 3.1). The system transfer function $G(s)$ is the product of the individual element transfer functions, i.e.

Transfer function for complete measurement system

$$\frac{\Delta\bar{O}(s)}{\Delta\bar{I}(s)} = G(s) = G_1(s)G_2(s)\ldots G_i(s)\ldots G_n(s) \qquad [4.41]$$

In principle we can use eqn [4.41] to find the system output signal $\Delta O(t)$ corresponding to a time-varying input signal $\Delta I(t)$. We first find the Laplace transform $\Delta\bar{I}(s)$ of $\Delta I(t)$, then using [4.9] the Laplace transform of the output signal is $\Delta\bar{O}(s) = G(s)\Delta\bar{I}(s)$. By expressing $\Delta\bar{O}(s)$ in partial fractions, and using standard tables of Laplace transforms, we can find the corresponding time signal $\Delta O(t)$. Expressing this mathematically:

$$\Delta O(t) = \mathscr{L}^{-1}[G(s)\Delta\bar{I}(s)] \qquad [4.42]$$

where $\mathscr{L}^{-1}$ denotes the inverse Laplace transform. The dynamic error $E(t)$ of the measurement system is the difference between the measured signal and the true signal, i.e. the difference between $\Delta O(t)$ and $\Delta I(t)$:

Dynamic error of a measurement system

$$E(t) = \Delta O(t) - \Delta I(t) \qquad [4.43]$$

Using [4.42] we have:

$$E(t) = \mathscr{L}^{-1}[G(s)\Delta\bar{I}(s)] - \Delta I(t) \qquad [4.44]$$

The simple temperature measurement system (Figure 4.12), first introduced in Section 3.1, provides a good example of dynamic errors. The thermocouple has a time constant of 10 s, the amplifier has a time constant of 10^{-4} s (Chapter 9), and the recorder (Chapter 11) is a second-order element with $\omega_n = 200$ rad/s and $\xi = 1.0$. The overall steady-state sensitivity of the system is unity.

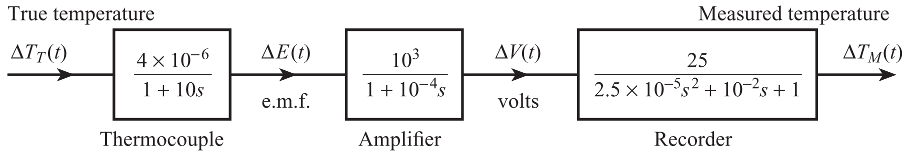


Figure 4.12 Simple temperature measurement system with dynamics.

We can now calculate the dynamic error of the system for a step input of +20 °C, i.e. $\Delta T_T(t) = 20u(t)$ and $\Delta \bar{T}_T(s) = 20(1/s)$. Thus the Laplace transform of the output signal is:

$$\Delta \bar{T}_M(s) = 20\frac{1}{s}\frac{1}{(1+10s)}\frac{1}{(1+10^{-4}s)}\frac{1}{(1+1/200s)^2}$$

$$= 20\left\{\frac{1}{s} - \frac{A}{(s+0.1)} - \frac{B}{(s+10^4)} - \frac{Cs+D}{(s+200)^2}\right\} \qquad [4.45]$$

Using Table 4.1 and eqn [4.30],

$$\Delta T_M(t) = 20\{u(t) - A\,e^{-0.1t} - B\,e^{-10^4 t} - E\,e^{-200t}(1+200t)\}$$

and the dynamic error:

$$E(t) = \Delta T_M(t) - \Delta T_T(t)$$
$$= -20\{A\,e^{-0.1t} + B\,e^{-10^4 t} - E\,e^{-200t}(1+200t)\} \qquad [4.46]$$

where the negative sign indicates too low a reading. The $B\,e^{-10^4 t}$ term decays to zero after about 5×10^{-4} s, and the $E\,e^{-200t}(1+200t)$ term decays to zero after about 25 ms. The $A\,e^{-0.1t}$ term, which corresponds to the 10 s time constant of the thermocouple, takes about 50 s to decay to zero and so has the greatest effect on the dynamic error.

We can use the rules developed in Section 4.2.2 to find the dynamic error of a system, with transfer function $G(s)$ subject to a sinusoidal input $\Delta I(t) = \hat{I}\sin\omega t$. From Figure 4.9 we have:

$$\Delta O(t) = |G(j\omega)|\hat{I}\sin(\omega t + \phi)$$

giving

$$E(t) = \hat{I}\{|G(j\omega)|\sin(\omega t + \phi) - \sin\omega t\} \qquad [4.47]$$

where $\phi = \arg G(j\omega)$.

Suppose that the above temperature measurement system is measuring a sinusoidal temperature variation of amplitude $\hat{T}_T = 20$ °C and period $T = 6$ s, i.e. angular frequency $\omega = 2\pi/T \approx 1.0$ rad s^{-1}. The frequency response function $G(j\omega)$ is:

$$G(j\omega) = \frac{1}{(1+10j\omega)(1+10^{-4}j\omega)[1+10^{-2}j\omega+2.5\times10^{-5}(j\omega)^2]} \qquad [4.48]$$

so that at $\omega = 1$

$$|G(j\omega)|_{\omega=1} = \frac{1}{\sqrt{(1+100)(1+10^{-8})[(1-2.5\times10^{-5})^2+10^{-4}]}} \approx 0.10$$

and [4.49]

$$\arg G(j\omega)_{\omega=1} \approx 0 - \tan^{-1}(10) - \tan^{-1}(10^{-4}) - \tan^{-1}(10^{-2}) \approx -85°$$

We note from the above equations that the values of $|G(j\omega)|$ and arg $G(j\omega)$ at $\omega = 1$ are determined mainly by the 10 s time constant; the dynamic characteristics of the other elements will only begin to affect the system performance at much higher frequencies. Since $T_T(t) = 20 \sin t$ and $T_M(t) = 0.1 \times 20 \sin(t - 85°)$, the error is:

$$E(t) = 20\{0.1 \sin(t - 85°) - \sin t\} \qquad [4.50]$$

We note that in the case of a sine wave input, the output recording is also a sine wave, i.e. the **waveform** of the signal is unaltered even though there is a reduction in amplitude and a phase shift.

In practice the input signal to a measurement system is more likely to be **periodic** rather than a simple sine wave. A periodic signal is one that repeats itself at equal intervals of time T, i.e. $f(t) = f(t + T) = f(t + 2T)$ etc. where T is the period. One example of a periodic measurement signal is the time variation of the temperature inside a diesel engine; another is the vibration of the casing of a centrifugal compressor. In order to calculate dynamic errors for periodic signals, we need to use **Fourier analysis**. Any periodic signal $f(t)$ with period T can be expressed as a series of sine and cosine waves; these have frequencies which are **harmonics** of the fundamental frequency $\omega_1 = 2\pi/T$ rad s^{-1}, i.e.

Fourier series for periodic signal

$$f(t) = a_0 + \sum_{n\;1}^{n\;\infty} a_n \cos n\omega_1 t + \sum_{n\;1}^{n\;\infty} b_n \sin n\omega_1 t \qquad [4.51]$$

where

$$a_n = \frac{2}{T}\int_{-T/2}^{+T/2} f(t) \cos n\omega_1 t \, dt$$

$$b_n = \frac{2}{T}\int_{-T/2}^{+T/2} f(t) \sin n\omega_1 t \, dt \qquad [4.52]$$

and

$$a_0 = \frac{1}{T}\int_{-T/2}^{+T/2} f(t)\, dt = \text{average value of } f(t) \text{ over } T$$

If $f(t) = \Delta I(t)$, where $\Delta I(t)$ is the deviation of measurement input signal $I(t)$ from steady-state or d.c. value I_0, then $a_0 = 0$. If we also assume that $f(t)$ is odd, i.e. $f(t) = -f(-t)$, then $a_n = 0$ for all n, i.e. there are only sine terms present in the series. This simplifying assumption does not affect the general conclusions drawn in the following section. The system input signal is thus given by

$$\Delta I(t) = \sum_{n\;1}^{n\;\infty} \hat{I}_n \sin n\omega_1 t \qquad [4.53]$$

where $\hat{I}_n = b_n$ is the amplitude of the nth harmonic at frequency $n\omega_1$. In order to find $\Delta O(t)$, let us first suppose that only the nth harmonic $\hat{I}_n \sin n\omega_1 t$ is input to the system. From Figure 4.9 the corresponding output signal is

$$\hat{I}_n |G(jn\omega_1)| \sin(n\omega_1 t + \phi_n)$$

where $\phi_n = \arg G(jn\omega_1)$.

We now require to use the principle of superposition, which is a basic property of linear systems (i.e. systems described by linear differential equations). This can be stated as follows:

> If an input $I_1(t)$ causes an output $O_1(t)$ and an input $I_2(t)$ causes an output $O_2(t)$, then an input $I_1(t) + I_2(t)$ causes an output $O_1(t) + O_2(t)$, provided the system is linear.

This means that if the total input signal is the sum of many sine waves (equation [4.53]), then the total output signal is the sum of the responses to each sine wave, i.e.

$$\Delta O(t) = \sum_{n\ 1}^{n\ \infty} \hat{I}_n |G(jn\omega_1)| \sin(n\omega_1 t + \phi_n) \qquad [4.54]$$

The dynamic error is thus

Dynamic error of system with periodic input signal

$$\Delta E(t) = \sum_{n\ 1}^{n\ \infty} \hat{I}_n \{|G(jn\omega_1)| \sin(n\omega_1 t + \phi_n) - \sin n\omega_1 t\} \qquad [4.55]$$

As an example, suppose that the input to the temperature measurement system is a square wave of amplitude 20 °C and period $T = 6$ s (i.e. $\omega_1 = 2\pi/T \approx 1$ rad s^{-1}). The Fourier series for the input signal is:

$$\Delta T_T(t) = \frac{80}{\pi}\left[\sin t + \tfrac{1}{3}\sin 3t + \tfrac{1}{5}\sin 5t + \tfrac{1}{7}\sin 7t + \ldots\right] \qquad [4.56]$$

Figure 4.13 shows the waveforms of the input square wave and the first four Fourier components with frequencies 1, 3, 5 and 7 rad s^{-1}.

Figure 4.14 shows the amplitude–frequency and phase–frequency relationships for the input temperature; these define the frequency spectrum of the signal. The spectrum consists of a number of lines at frequencies ω_1, $3\omega_1$, $5\omega_1$, etc., of decreasing length to represent the smaller amplitudes of the higher harmonics. In practical cases we can terminate or truncate the series at a harmonic where the amplitude is negligible; in this case we choose $n = 7$. In order to find the output signal, i.e. the recorded waveform, we need to evaluate the magnitude and argument of $G(j\omega)$ at $\omega = 1$, 3, 5 and 7 rad s^{-1}. Thus

$$|G(j)| \approx 0.100, \quad |G(3j)| \approx 0.033, \quad |G(5j)| \approx 0.020, \quad |G(7j)| \approx 0.014$$

and

$$\begin{aligned} &\arg G(j) \approx -85°, \quad \arg G(3j) \approx -90°, \\ &\arg G(5j) \approx -92°, \quad \arg G(7j) \approx -93° \end{aligned} \qquad [4.57]$$

Again the above values are determined mainly by the 10 s thermocouple time constant; the highest signal frequency $\omega = 7$ is still well below the natural frequency of the recorder $\omega_n = 200$. The system output signal is:

$$\begin{aligned} \Delta T_M(t) = \frac{80}{\pi}[&0.100 \sin(t - 85°) + 0.011 \sin(3t - 90°) \\ &+ 0.004 \sin(5t - 92°) + 0.002 \sin(7t - 93°)] \end{aligned} \qquad [4.58]$$

Figure 4.13 Waveforms for input square wave and Fourier components.

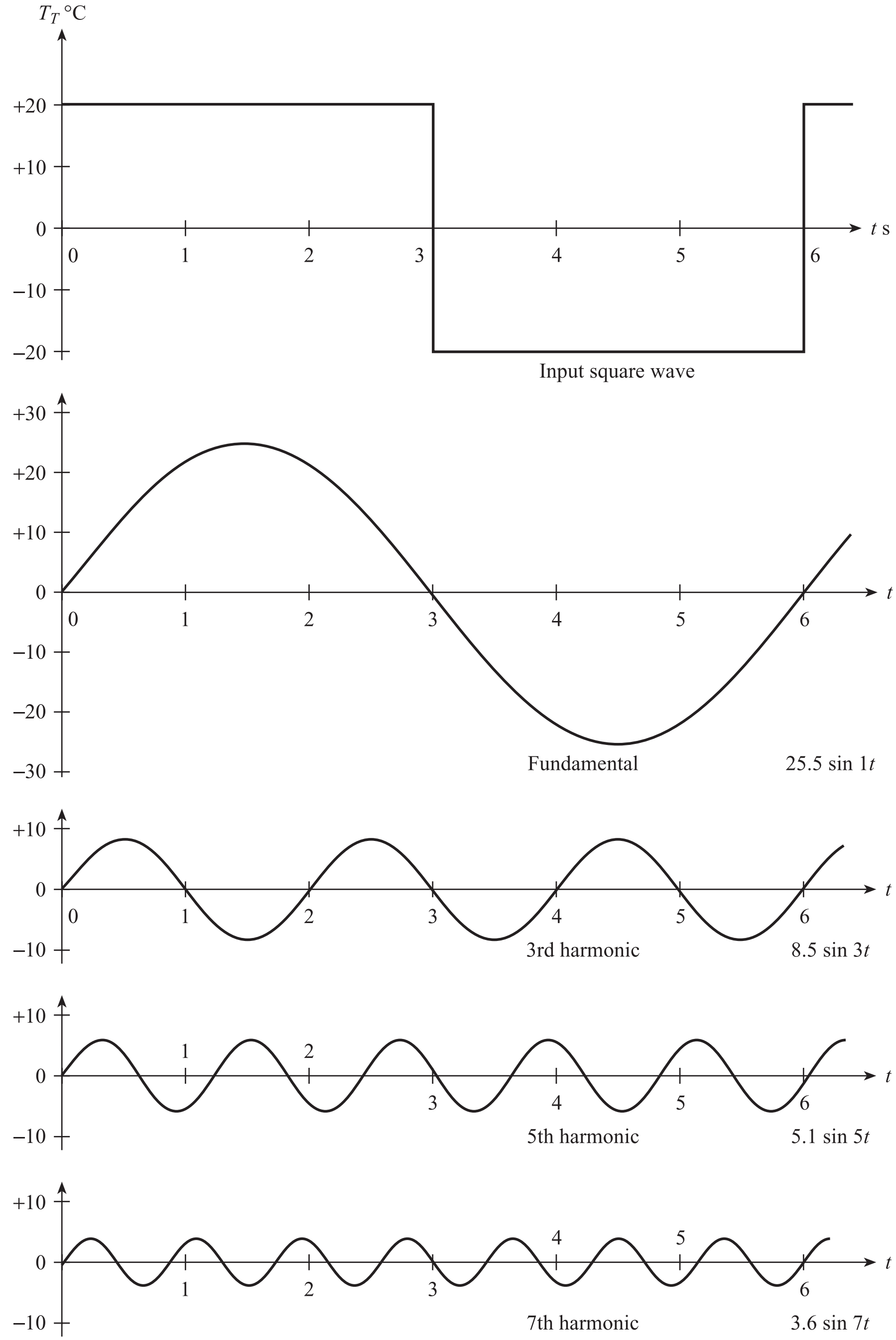


Figure 4.14 shows the system frequency response characteristics, the output signal frequency spectrum and the output waveform. We note that, in the output signal, the amplitudes of the 3rd, 5th and 7th harmonics have been reduced relative to the amplitude of the fundamental. The recorded waveform has therefore a different **shape** from the input signal as well as being reduced in amplitude and changed in phase.

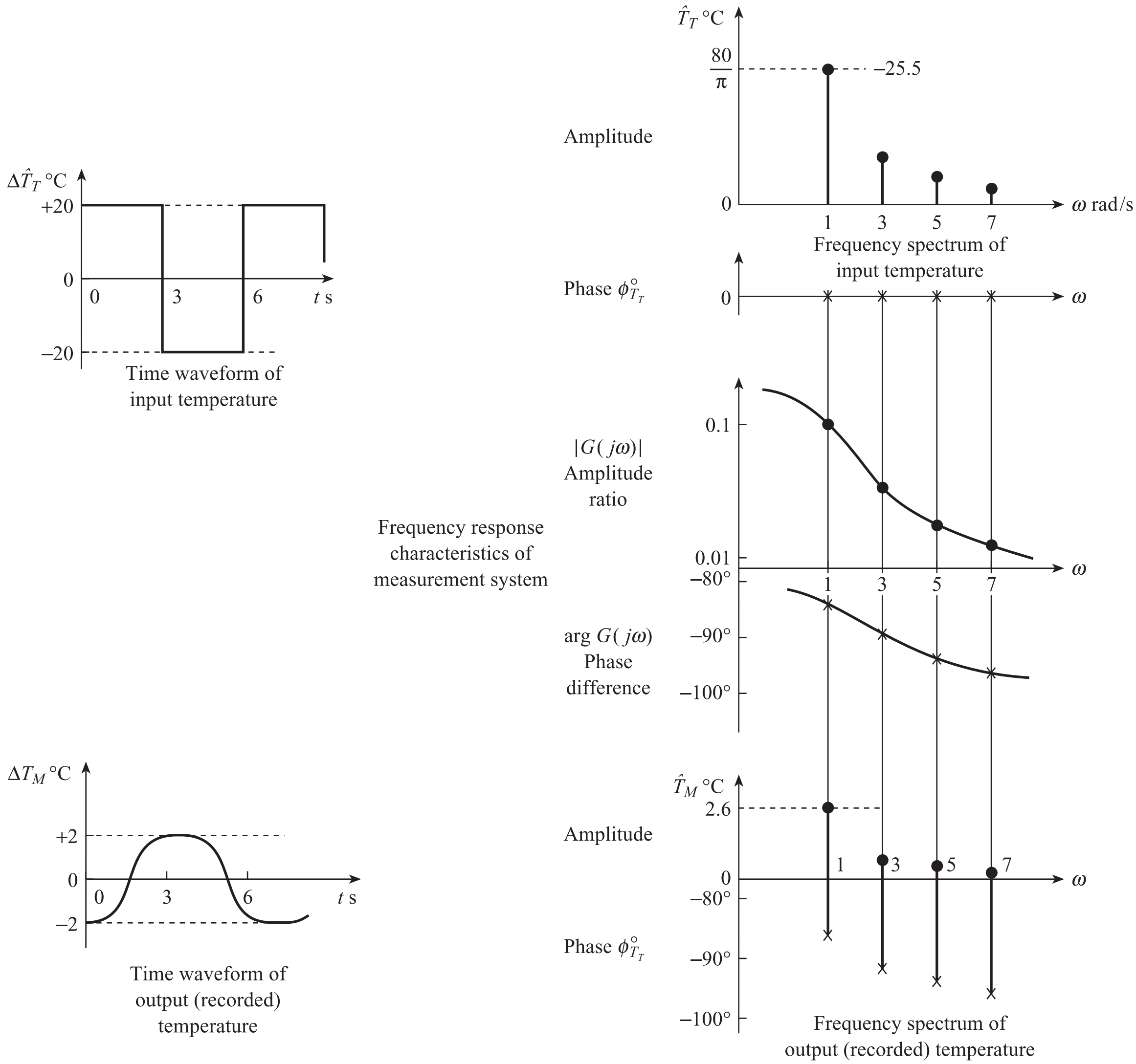


Figure 4.14 Calculation of dynamic errors with periodic input signal.

The above ideas can be extended to calculating the dynamic error for **random** input signals. Random signals can be represented by continuous frequency spectra (Chapter 6).

4.4 Techniques for dynamic compensation

From eqn [4.55] we see that in order to have $E(t) = 0$ for a periodic signal, the following conditions must be obeyed:

$$|G(j\omega_1)| = |G(j2\omega_1)| = \ldots = |G(jn\omega_1)| = \ldots = |G(jm\omega_1)| = 1$$
$$\arg G(j\omega_1) = \arg G(j2\omega_1) = \ldots = \arg G(jn\omega_1) = \ldots = \arg G(jm\omega_1) = 0 \quad [4.59]$$

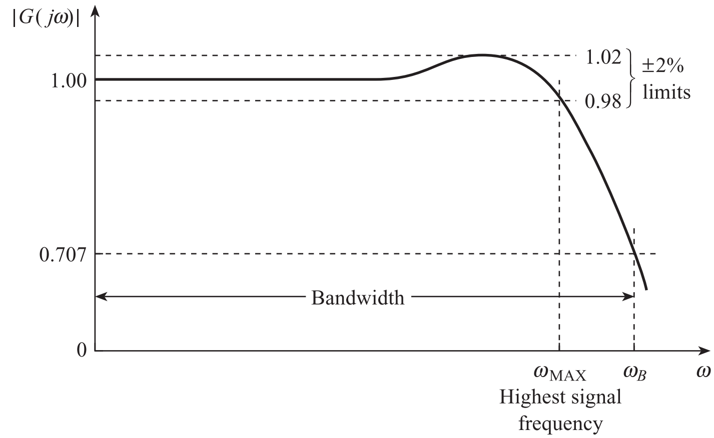


Figure 4.15 Percentage limits and bandwidth.

where m is the order of the **highest *significant* harmonic**. For a random signal (Chapter 6) with a continuous frequency spectrum containing frequencies between 0 and ω_{MAX}, we require:

$$|G(j\omega)| = 1 \quad \text{and} \quad \arg G(j\omega) = 0 \quad \text{for} \quad 0 < \omega \leq \omega_{MAX} \qquad [4.60]$$

The above conditions represent a theoretical ideal which will be difficult to realise in practice. A more practical criterion is one which limits the variation in $|G(j\omega)|$ to a few per cent for the frequencies present in the signal. For example the condition

$$0.98 < |G(j\omega)| < 1.02 \quad \text{for} \quad 0 < \omega \leq \omega_{MAX} \qquad [4.61]$$

will ensure that the dynamic error is limited to $\approx \pm 2$ per cent for a signal containing frequencies up to $\omega_{MAX}/2\pi$ Hz (Figure 4.15).

Another commonly used criterion is that of **bandwidth**. The bandwidth of an element or a system is the range of frequencies for which $|G(j\omega)|$ is greater than $1/\sqrt{2}$. Thus the bandwidth of the system, with frequency response shown in Figure 4.15, is 0 to ω_B rad s^{-1}. The highest signal frequency ω_{MAX} must be considerably less than ω_B. Since, however, there is a 30% reduction in $|G(j\omega)|$ at ω_B, bandwidth is not a particularly useful criterion for complete measurement systems.

Bandwidth is commonly used in specifying the frequency response of amplifiers (Chapter 9); a reduction in $|G(j\omega)|$ from 1 to $1/\sqrt{2}$ is equivalent to a decibel change of $N = 20\log(1/\sqrt{2}) = -3.0$ dB. A first-order element has a bandwidth between 0 and $1/\tau$ rad s^{-1}.

If a system fails to meet the specified limits on dynamic error $E(t)$, i.e. the system transfer function $G(s)$ does not satisfy a condition such as [4.61], then the first step is to identify which elements in the system dominate the dynamic behaviour. In the temperature measurement system of the previous section, the dynamic error is almost entirely due to the 10 s time constant of the thermocouple.

Having identified the dominant elements in the system, the most obvious method of improving dynamic response is that of **inherent design**. In the case of a first-order temperature sensor with $\tau = MC/UA$, τ can be minimised by minimising the mass/area ratio M/A – for example by using a thermistor in the form of a thin flake. In the case of a second-order force sensor with $\omega_n = \sqrt{k/m}$, ω_n can be maximised by maximising k/m, i.e. by using high stiffness k and low mass m. Increasing k, however, reduces the steady-state sensitivity $K = 1/k$.

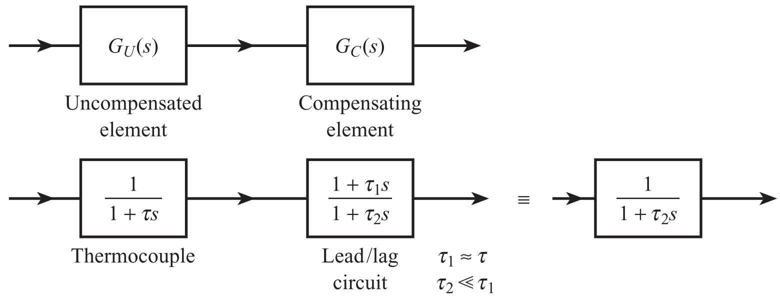


Figure 4.16 Open-loop dynamic compensation.

From second-order step and frequency response graphs we see that the optimum value of damping ratio ξ is around 0.7. This value ensures minimum settling time for the step response and $|G(j\omega)|$ closest to unity for the frequency response.

Another possible method is that of **open-loop dynamic compensation** (Figure 4.16). Given an uncompensated element or system $G_U(s)$, a compensating element $G_C(s)$ is introduced into the system, such that the overall transfer function $G(s) = G_U(s)G_C(s)$ satisfies the required condition (for example eqn [4.61]). Thus if a lead/lag circuit (Figure 9.12) is used with a thermocouple (Figure 4.16), the overall time constant is reduced to τ_2 so that $|G(j\omega)|$ is close to unity over a wider range of frequencies. The main problem with this method is that τ can change with heat transfer coefficient U, thus reducing the effectiveness of the compensation (Chapter 14).

Another method is to incorporate the element to be compensated into a closed-loop system with **high-gain negative feedback**. An example of this is the constant temperature anemometer system for measuring fluid velocity fluctuations (Section 14.3). Another example is the closed-loop accelerometer shown in schematic and block diagram form in Figure 4.17.

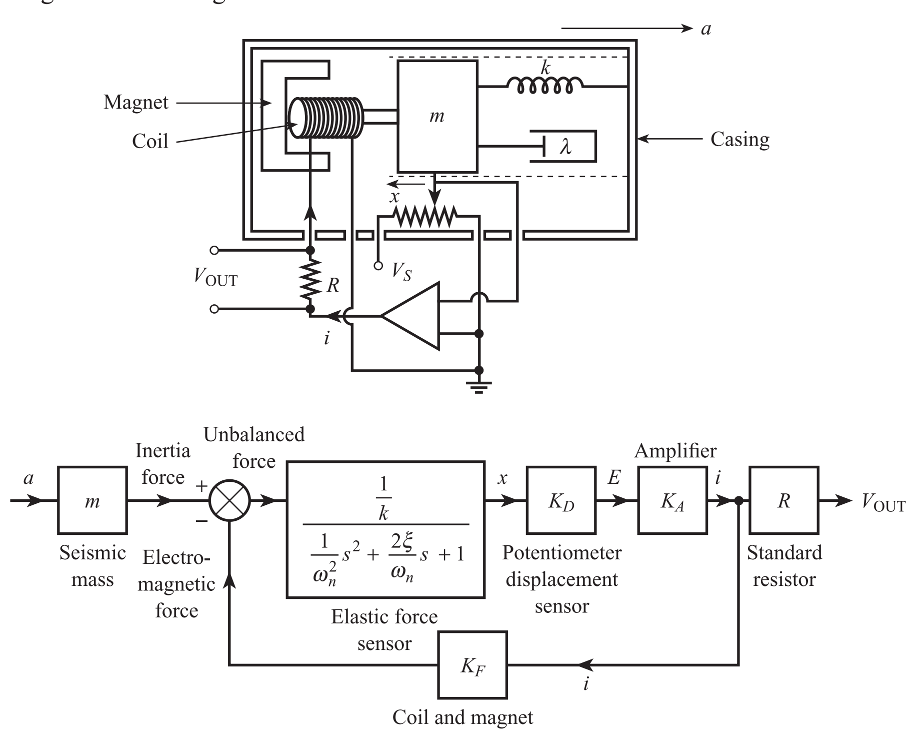


Figure 4.17 Schematic and block diagram of closed-loop accelerometer.

The applied acceleration a produces an inertia force ma on the seismic mass m (Chapter 8). This is balanced by the force of the permanent magnet on the current feedback coil. Any imbalance of forces is detected by the elastic force element to produce a displacement which is detected by a potentiometric displacement sensor (Chapter 5). The potentiometer output voltage is amplified, giving a current output which is fed to the feedback coil through a standard resistor to give the output voltage.

Analysis of the block diagram shows that the overall system transfer function is:

$$\frac{\Delta\bar{V}(s)}{\Delta\bar{a}(s)} = \frac{mR}{K_F}\left\{\frac{1}{\dfrac{k}{K_A K_D K_F}\dfrac{1}{\omega_n^2}s^2 + \dfrac{2\xi}{\omega_n}\dfrac{ks}{K_A K_D K_F} + \left(1 + \dfrac{1}{K_A K_D K_F}\right)}\right\} \quad [4.62]$$

If K_A is made large so that $K_A K_D K_F/k \gg 1$ (Chapter 3), then the system transfer function can be expressed in the form:

$$\frac{\Delta\bar{V}(s)}{\Delta\bar{a}(s)} = \frac{K_s}{\dfrac{1}{\omega_{ns}^2}s^2 + \dfrac{2\xi_s}{\omega_{ns}}s + 1}$$

where

system steady-state sensitivity $\quad K_s = \dfrac{mR}{K_F}$

system natural frequency $\quad \omega_{ns} = \omega_n\sqrt{\dfrac{K_A K_D K_F}{k}}$

system damping ratio $\quad \xi_s = \xi\sqrt{\dfrac{k}{K_A K_D K_F}} \quad [4.63]$

We see that the system natural frequency ω_{ns} is now much greater than that of the elastic force element itself. The system damping ratio ξ_s is much less than ξ, but by making ξ large a value of $\xi_s \approx 0.7$ can be obtained. Furthermore the system steady-state sensitivity depends only on m, K_F and R, which can be made constant to a high degree.

Conclusion

The **dynamic characteristics** of typical measurement system elements were initially discussed; in particular the **transfer functions** of **first-** and **second-order** elements were derived. The response of both first- and second-order elements to **step** and **sine wave inputs** was then studied. A general description of the **dynamic error** of a complete measurement system was then developed and applied to a temperature measurement system subject to step, sine wave and **periodic** input signals. Finally methods of **dynamic compensation**, which reduce dynamic error, were explained.

Problems

4.1 A temperature measurement system consists of linear elements and has an overall steady-state sensitivity of unity. The dynamics of the system are determined by the first-order transfer function of the sensing element. At time $t = 0$, the sensing element is suddenly transferred from air at 20 °C to boiling water. One minute later the element is suddenly transferred back to air. Using the data given below, calculate the system dynamic error at the following times: $t = 10$, 20, 50, 120 and 300 s.

Sensor data
Mass = 5×10^{-2} kg
Surface area = 10^{-3} m^2
Specific heat = 0.2 J kg^{-1} °C^{-1}
Heat transfer coefficient for air = 0.2 W m^{-2} °C^{-1}
Heat transfer coefficient for water = 1.0 W m^{-2} °C^{-1}

4.2 A force sensor has a mass of 0.5 kg, stiffness of 2×10^2 N m^{-1} and a damping constant of 6.0 N s m^{-1}.

(a) Calculate the steady-state sensitivity, natural frequency and damping ratio for the sensor.
(b) Calculate the displacement of the sensor for a steady input force of 2 N.
(c) If the input force is suddenly increased from 2 to 3 N, derive an expression for the resulting displacement of the sensor.

4.3 A force measurement system consists of linear elements and has an overall steady-state sensitivity of unity. The dynamics of the system are determined by the second-order transfer function of the sensing element, which has a natural frequency $\omega_n = 40$ rad s^{-1} and a damping ratio $\xi = 0.1$. Calculate the system dynamic error corresponding to the periodic input force signal:

$$F(t) = 50(\sin 10t + \tfrac{1}{3}\sin 30t + \tfrac{1}{5}\sin 50t)$$

4.4 An uncompensated thermocouple has a time constant of 10 s in a fast-moving liquid.

(a) Calculate the bandwidth of the thermocouple frequency response.
(b) Find the range of frequencies for which the amplitude ratio of the uncompensated thermocouple is flat within ±5%.
(c) A lead/lag circuit with transfer function $G(s) = (1 + 10s)/(1 + s)$ is used to compensate for thermocouple dynamics. Calculate the range of frequencies for which the amplitude ratio of the compensated system is flat within ±5%.
(d) The velocity of the liquid is reduced, causing the thermocouple time constant to increase to 20 s. By sketching $|G(j\omega)|$ explain why the effectiveness of the above compensation is reduced.

4.5 An elastic force sensor has an effective seismic mass of 0.1 kg, a spring stiffness of 10 N m^{-1} and a damping constant of 14 N s m^{-1}.

(a) Calculate the following quantities:
  (i) sensor natural frequency
  (ii) sensor damping ratio
  (iii) transfer function relating displacement and force.
(b) The above sensor is incorporated into a closed-loop, force balance accelerometer. The following components are also present:

Potentiometer displacement sensor: sensitivity 1.0 V m^{-1}
Amplifier: voltage input, current output, sensitivity 40 A V^{-1}
Coil and magnet: current input, force output, sensitivity 25 N A^{-1}
Resistor: 250 Ω.

(i) Draw a block diagram of the accelerometer.
(ii) Calculate the overall accelerometer transfer function.
(iii) Explain why the dynamic performance of the accelerometer is superior to that of the elastic sensor.

4.6 A load cell consists of an elastic cantilever and a displacement transducer. The cantilever has a stiffness of 10^2 N m^{-1}, a mass of 0.5 kg and a damping constant of 2 N s m^{-1}. The displacement transducer has a steady-state sensitivity of 10 V m^{-1}.

(a) A package of mass 0.5 kg is suddenly dropped onto the load cell. Use eqn [4.31] to derive a numerical equation describing the corresponding time variation of the output voltage ($g = 9.81$ m s^{-2}).
(b) The load cell is used to weigh packages moving along a conveyor belt at the rate of 60 per minute. Use the equation derived in (a) to explain why the load cell is unsuitable for this application. Explain what modifications to the load cell are necessary.

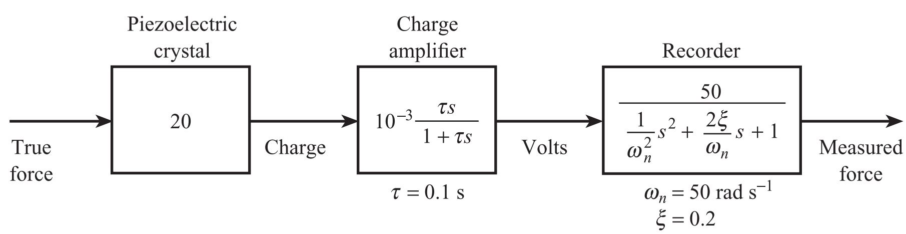
4.7 A force measurement system consisting of a piezoelectric crystal, charge amplifier and recorder is shown in Figure Prob. 7.

(a) Calculate the system dynamic error corresponding to the force input signal:

$$F(t) = 50(\sin 10t + \tfrac{1}{3}\sin 30t + \tfrac{1}{5}\sin 50t)$$

(b) Explain briefly the system modifications necessary to reduce the error in (a) (hint: see Figure 8.22).

Figure Prob. 7.



4.8 A thermocouple is used to measure the temperature inside a vessel, which is part of a high-speed batch process. At time $t = 0$, with the vessel at an initial temperature of 50 °C, the vessel is instantaneously filled with gas at 150 °C. One minute later, instantaneously the gas is removed and the vessel is filled with liquid at 50 °C. The thermocouple can be regarded as having linear steady-state characteristics and first-order dynamics.

(a) Use the data given below to sketch a graph of how the thermocouple e.m.f. changes with time. The axes of the graph should have suitable scales and the answer should include supporting numerical calculations.
(b) Comment on whether the thermocouple is suitable for this application.
(c) What modifications should be made?

Data

Thermocouple sensitivity	= 40 µV °C^{-1}
Thermocouple mass	= 5×10^{-2} kg
Thermocouple specific heat	= 0.2 J kg^{-1} °C^{-1}
Thermocouple surface area	= 10^{-3} m^2
Heat transfer coefficient for gas	= 0.2 W m^{-2} °C^{-1}
Heat transfer coefficient for liquid	= 1.0 W m^{-2} °C^{-1}

Response of first-order element, with unit sensitivity, to unit step:

$$F^0(t) = 1 - \exp\left(\frac{-t}{\tau}\right)$$

4.9 A temperature measurement system for a gas reactor consists of linear elements and has an overall steady-state sensitivity of unity. The temperature sensor has a time constant of 5.0 s; an *ideal* low-pass filter with a cut-off frequency of 0.05 Hz is also present. The input temperature signal is periodic with period 63 s and can be approximated by the Fourier series:

$$T(t) = 10(\sin \omega_0 t + \tfrac{1}{2}\sin 2\omega_0 t + \tfrac{1}{3}\sin 3\omega_0 t + \tfrac{1}{4}\sin 4\omega_0 t)$$

where ω_0 is the angular frequency of the fundamental component.

(a) Calculate expressions for the time response of:

(i) the system output signal
(ii) the system dynamic error.

(b) Explain what modifications are necessary to the system to minimise the dynamic error in (a).

Note
An ideal low-pass filter has a gain of one and zero phase shift up to the cut-off frequency. The gain is zero above the cut-off frequency.

5 Loading Effects and Two-port Networks

In our discussion of measurement systems no consideration has yet been given to the effects of **loading**. One important effect is that of **inter-element loading** where a given element in the system may modify the characteristics of the previous element (for example by drawing current). In turn, the characteristics of this element may be modified by the following element in the system. Inter-element loading is normally an electrical loading effect which is described in the first section of this chapter using **Thévenin** and **Norton equivalent circuits**. The second section begins by discussing the analogies between electrical and non-electrical variables. This means that mechanical and thermal systems can be described by equivalent circuits and sensing elements by **two-port networks**. Two-port networks are then used to describe **process loading**; here the introduction of the sensing element into the process or system being measured causes the value of the measured variable to change. Finally two-port networks are used to describe **bilateral transducers** which use reversible physical effects.

5.1 Electrical loading

We have so far represented measurement systems as blocks connected by single lines where the transfer of information and energy is in terms of one variable only. Thus in the temperature measurement system of Figure 3.2 the information transfer between elements is in terms of voltage only. No allowance can therefore be made for the amplifier drawing current from the thermocouple and the indicator drawing current from the amplifier. In order to describe both voltage and current behaviour at the connection of two elements, we need to represent each element by equivalent circuits characterised by two terminals. The connection is then shown by two lines.

5.1.1 Thévenin equivalent circuit

Thévenin's theorem states that any network consisting of linear impedances and voltage sources can be replaced by an equivalent circuit consisting of a voltage source E_{Th} and a series impedance Z_{Th} (Figure 5.1). The source E_{Th} is equal to the open circuit voltage of the network across the output terminals, and Z_{Th} is the impedance looking back into these terminals with all voltage sources reduced to zero and replaced by their internal impedances. Thus connecting a load Z_L across the output

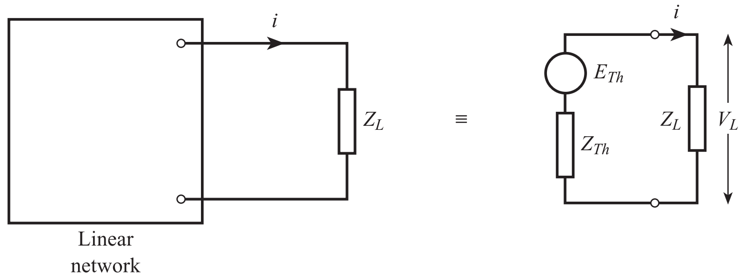


Figure 5.1 Thévenin equivalent circuit.

terminals of the network is equivalent to connecting Z_L across the Thévenin circuit. The current i in Z_L is then simply given by:

$$i = \frac{E_{Th}}{Z_{Th} + Z_L} \qquad [5.1]$$

and the voltage V_L across the load by:

Loading of Thévenin equivalent circuit

$$V_L = iZ_L = E_{Th}\frac{Z_L}{Z_{Th} + Z_L} \qquad [5.2]$$

From eqn [5.2] we see that if $Z_L \gg Z_{Th}$, then $V_L \to E_{Th}$; i.e. in order to get **maximum voltage transfer** from the network to the load, the load impedance should be *far greater* than the Thévenin impedance for the network. In order to get **maximum power transfer** from network to load, the load impedance should be equal to the network impedance; i.e. $Z_L = Z_{Th}$.[1] (An example of the calculation of E_{Th} and Z_{Th} for a potentiometer displacement transducer is given in the following section and for a deflection bridge in Section 9.1.)

We can now discuss the Thévenin equivalent circuit for the temperature measurement system of Figure 3.2. The thermocouple may be represented by $Z_{Th} = 20\ \Omega$ (resistive) and $E_{Th} = 40T\ \mu\text{V}$, where T is the measurement junction temperature, if non-linear and reference junction temperature effects are ignored. The amplifier acts both as a load for the thermocouple and as a voltage source for the indicator. Figure 5.2 shows a general equivalent circuit for an amplifier with two pairs of terminals. Using typical amplifier data (Chapter 9), we have input impedance $Z_{IN} = R_{IN} = 2 \times 10^6\ \Omega$, closed-loop voltage gain $A = 10^3$, and output impedance $Z_{OUT} = R_{OUT} = 75\ \Omega$. The indicator is a resistive load of $10^4\ \Omega$. The complete equivalent circuit for the system is shown in Figure 5.3, and using eqn [5.2] we have:

$$V_{IN} = 40 \times 10^{-6}T\left(\frac{2 \times 10^6}{2 \times 10^6 + 20}\right)$$
$$V_L = 1000\ V_{IN}\left(\frac{10^4}{10^4 + 75}\right) \qquad [5.3]$$

If the indicator scale is drawn so that a change of 1 V in V_L causes a change in deflection of 25 °C, then the measured temperature $T_M = 25V_L$. This gives:

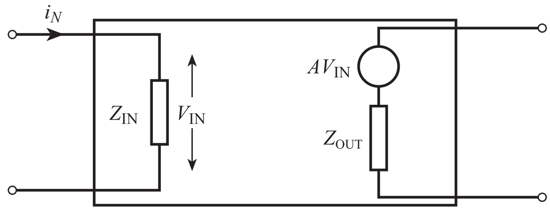


Figure 5.2 Equivalent circuit for amplifier.

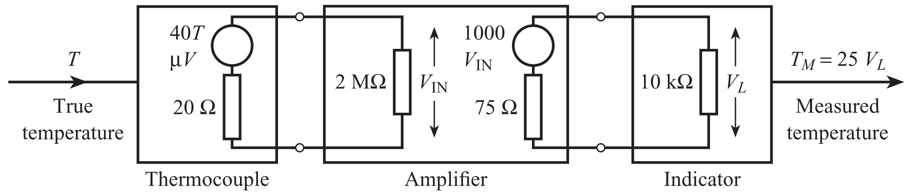


Figure 5.3 Thévenin equivalent to temperature measurement system.

$$T_M = \left(\frac{2 \times 10^6}{2 \times 10^6 + 20}\right)\left(\frac{10^4}{10^4 + 75}\right)T = 0.9925T \qquad [5.4]$$

i.e. we have to introduce the factor $Z_L/(Z_{Th} + Z_L)$ at every interconnection of two elements to allow for loading. The loading error $= -0.0075T$; this is in addition to the steady-state error due to element inperfections calculated in Chapter 3.

The loading error in the above example is small, but if care is not taken it can be very large. Suppose a pH glass electrode (Chapter 8), with sensitivity 59 mV per pH, i.e. $E_{Th} = 59\text{pH mV}$ and $Z_{Th} = R_{Th} = 10^9\ \Omega$, is directly connected to an indicator with $Z_L = R_L = 10^4\ \Omega$ and a scale of sensitivity $\frac{1}{59}$pH/mV. The measured pH is:

$$\text{pH}_M = 59\text{pH}\left(\frac{10^4}{10^4 + 10^9}\right)\frac{1}{59} \approx 10^{-5}\text{pH} \qquad [5.5]$$

i.e. there will be effectively a zero indication for any non-zero value.

This problem is solved by connecting the electrode to the indicator via a **buffer amplifier**. This is characterised by large Z_{IN}, small Z_{OUT} and unity gain $A = 1$. For example, an operational amplifier with a field effect transistor (FET) input stage connected as a voltage follower (Figure 9.9) would have $Z_{IN} = 10^{12}\ \Omega$, $Z_{OUT} = 10\ \Omega$. The indicated pH value for the modified system (Figure 5.4) is:

$$\text{pH}_M = \frac{10^{12}}{10^{12} + 10^9} \times \frac{10^4}{10^4 + 10}\text{pH} \qquad [5.6]$$

and the loading error is now -0.002pH, which is negligible.

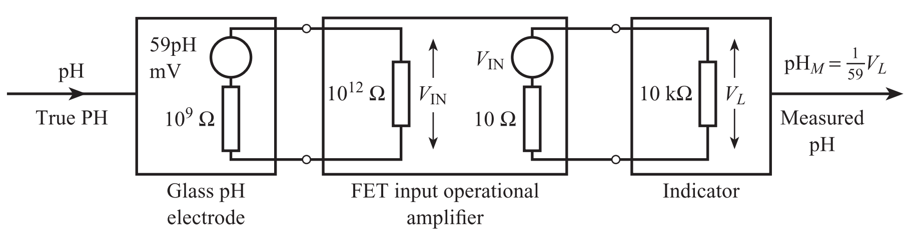


Figure 5.4 Equivalent circuit for pH measurement system.

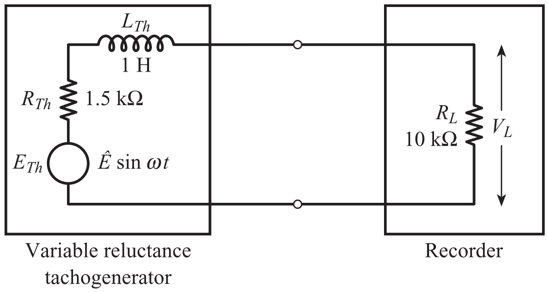


Figure 5.5 A.C. loading of tachogenerator.

An example of a.c. loading effects is given in Figure 5.5, which shows the equivalent circuit for a variable reluctance tachogenerator connected to a recorder. The Thévenin voltage E_{Th} for the tachogenerator is a.c. with an amplitude $\hat{E}$ and angular frequency ω, both proportional to the mechanical angular velocity ω_r (Section 8.4). In this example, $\hat{E} = (5.0 \times 10^{-3})\ \omega_r$ volts and $\omega = 6\omega_r$ rad s^{-1}. The Thévenin impedance Z_{Th} for the tachogenerator is an inductance and resistance in series (coil surrounding magnet), i.e. $Z_{Th} = R_{Th} + j\omega L_{Th}$. Thus if $\omega_r = 10^3$ rad s^{-1}:

$$\hat{E} = 5 \text{ V}, \ \omega = 6 \times 10^3 \text{ rad s}^{-1}$$

and

$$Z_{Th} = 1.5 + 6.0j \text{ k}\Omega$$

so that the amplitude of the recorded voltage is

$$\hat{V}_L = \hat{E}\frac{R_L}{|Z_{Th} + R_L|} = 5\frac{10}{\sqrt{(11.5)^2 + (6.0)^2}} = 3.85 \text{ V} \qquad [5.7]$$

If the recorder scale sensitivity is set at $1/(5 \times 10^{-3})$ rad s^{-1} V^{-1}, then the recorded angular velocity is 770 rad s^{-1}. This error can be removed either by increasing the recorder impedance, or by changing the recorder sensitivity to allow for loading effects. A better alternative is to replace the recorder by a counter which measures the frequency rather than the amplitude of the tachogenerator signal (Section 10.3).

5.1.2 Example of Thévenin equivalent circuit calculation: potentiometric displacement sensor

Figure 5.6 shows a schematic diagram of a potentiometric sensor for measuring displacements d. The resistance of the potentiometer varies linearly with displacement.

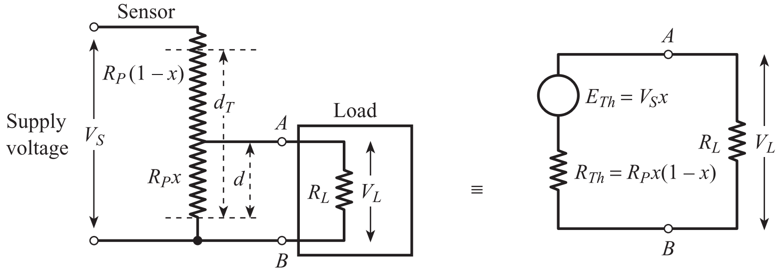


Figure 5.6 Potentiometer displacement sensor and Thévenin equivalent circuit.

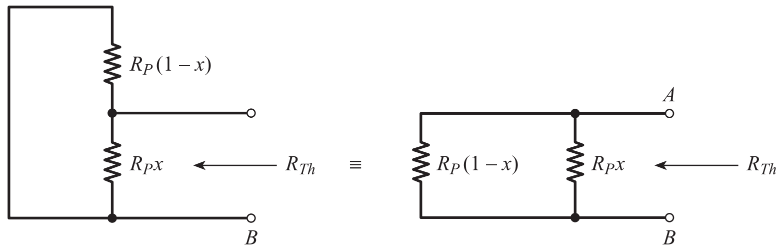


Figure 5.7 Calculation of R_{Th} for potentiometer.

Thus if $x = d/d_T$ is the fractional displacement, the corresponding resistance is $R_P x$, where R_P Ω is the total resistance of the potentiometer. The Thévinin voltage E_{Th} is the open circuit voltage across the output terminals AB. The ratio between E_{Th} and supply voltage V_S is equal to the ratio of fractional resistance $R_P x$ to total resistance R_P; that is

$$\frac{E_{Th}}{V_S} = \frac{R_P x}{R_P}, \quad \text{giving} \quad E_{Th} = V_S x \qquad [5.8]$$

The Thévenin impedance Z_{Th} is found by setting supply voltage $V_S = 0$, replacing the supply by its internal impedance (assumed to be zero), and calculating the impedance looking back into the terminals AB as shown in Figure 5.7. Thus:

$$\frac{1}{R_{Th}} = \frac{1}{R_P x} + \frac{1}{R_P(1-x)}$$

giving

$$R_{Th} = R_P x(1-x) \qquad [5.9]$$

Thus the effect of connecting a resistive load R_L (recorder or indicator) across the terminals AB is equivalent to connecting R_L across the Thévenin circuit. The load voltage is thus:

$$V_L = E_{Th}\frac{R_L}{R_{Th} + R_L} = V_S x \frac{R_L}{R_P x(1-x) + R_L}$$

i.e.

Voltage–displacement relationship for a loaded potentiometer

$$V_L = V_S x \frac{1}{(R_P/R_L)x(1-x) + 1} \qquad [5.10]$$

The relationship between V_L and x is non-linear, the amount of non-linearity depending on the ratio R_P/R_L (Figure 5.8). Thus the effect of loading a linear potentiometric sensor is to introduce a non-linear error into the system given by:

$$N(x) = E_{Th} - V_L = V_S x \left\{1 - \frac{1}{(R_P/R_L)x(1-x) + 1}\right\}$$

i.e.

$$N(x) = V_S \left\{\frac{x^2(1-x)(R_P/R_L)}{1 + (R_P/R_L)x(1-x)}\right\} \qquad [5.11]$$

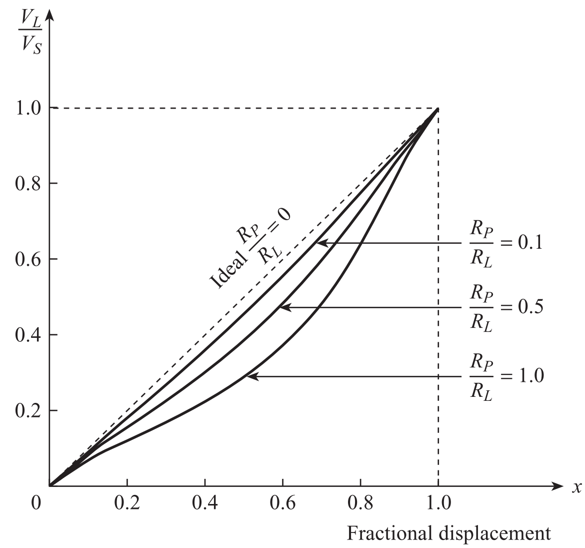


Figure 5.8 Non-linear characteristics of loaded potentiometer.

which reduces to:

$$N(x) \approx V_S(R_P/R_L)(x^2 - x^3)$$

if $R_P/R_L \ll 1$ (the usual situation). $N(x)$ has a maximum value of $\hat{N} = \frac{4}{27}V_S(R_P/R_L)$ when $x = \frac{2}{3}$, corresponding to $dN/dx = 0$ and negative d^2N/dx^2. Expressing $\hat{N}$ as a percentage of full-scale deflection or span V_S volts gives:

$$\hat{N} = \frac{400}{27}\frac{R_P}{R_L}\ \% \approx 15\frac{R_P}{R_L}\ \% \qquad [5.12]$$

Non-linearity, sensitivity and maximum power requirements are used to specify the values of R_P and V_S for a given application. Suppose that a 10 cm range potentiometer is to be connected to a 10 kΩ recorder. If the maximum non-linearity must not exceed 2%, then we require $15R_P/R_L \le 2$, i.e. $R_P \le \frac{20}{15} \times 10^3\ \Omega$; thus a 1 kΩ potentiometer would be suitable.

Since sensitivity $dV_L/dx \approx V_S$, the greater V_S the higher the sensitivity, but we must satisfy the requirement that power dissipation V_S^2/R_P should not exceed maximum value $\hat{W}$ watts. If $\hat{W} = 0.1$ W we require $V_S \le \sqrt{0.1 \times 10^3}$, i.e. $V_S \le 10$ V; if $V_S = 10$ V, then corresponding sensitivity = 1.0 V cm^{-1}.

5.1.3 Norton equivalent circuit

Norton's theorem states that any network consisting of linear impedances and voltage sources can be replaced by an equivalent circuit consisting of a current source i_N in parallel with an impedance Z_N (Figure 5.9). Z_N is the impedance looking back into the output terminals with all voltage sources reduced to zero and replaced by their internal impedances, and i_N is the current which flows when the terminals are short circuited. Connecting a load Z_L across the output terminals of the network is equivalent to connecting Z_L across the Norton circuit. The voltage V_L across the load is given by $V_L = i_N Z$, where $1/Z = 1/Z_N + 1/Z_L$, giving:

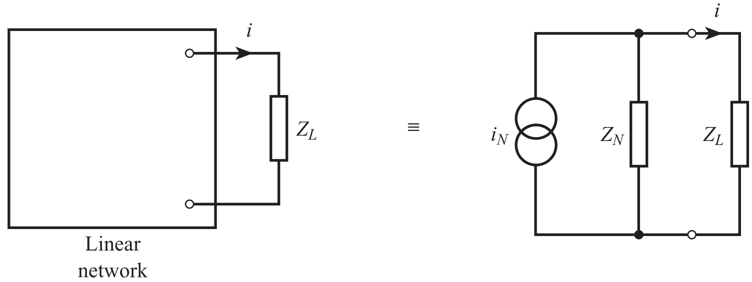


Figure 5.9 Norton equivalent circuit.

Loading of Norton equivalent circuit

$$V_L = i_N \frac{Z_N \cdot Z_L}{Z_N + Z_L} \quad [5.13]$$

From eqn [5.13] we note that if $Z_L \ll Z_N$, then $V_L \to i_N Z_L$; i.e. in order to develop the maximum current through the load, the load impedance should be *far smaller* than the Norton impedance for the network.

A common example of a current source is an electronic differential pressure transmitter giving an output current signal, range 4 to 20 mA, proportional to an input differential pressure, typical range 0 to 2×10^4 Pa (Section 9.4). Figure 5.10 shows a typical equivalent circuit for the transmitter connected to a recorder via a cable. Using eqn [5.13], the voltage across the total load $R_C + R_R$ of recorder and cable is

$$V_L = i_N \frac{R_N(R_C + R_R)}{R_N + R_C + R_R} \quad [5.14]$$

and the ratio $V_R/V_L = R_R/(R_C + R_R)$, giving the recorder voltage:

$$V_R = i_N R_R \frac{R_N}{R_N + R_C + R_R} \quad [5.15]$$

Using the data given, we have $V_R = 0.9995 i_N R_R$, so that the recorded voltage deviates from the desired range of 1 to 5 V by only 0.05%.

A second example of a current generator is provided by a piezoelectric crystal acting as a force sensor. If a force F is applied to any crystal, then the atoms of the crystal undergo a small displacement x proportional to F. For a piezoelectric material the crystal acquires a charge q proportional to x, i.e. $q = Kx$. The crystal can therefore be regarded as a Norton current source of magnitude $i_N = \mathrm{d}q/\mathrm{d}t = K(\mathrm{d}x/\mathrm{d}t)$,

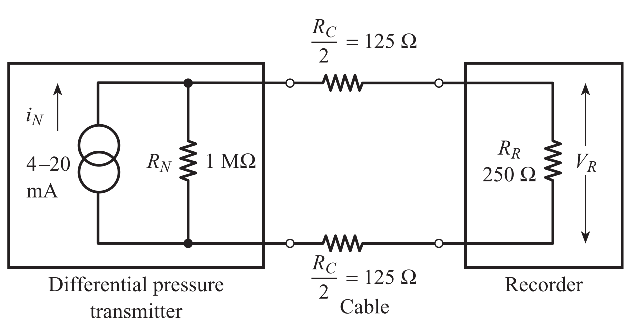


Figure 5.10 Typical current source and load.

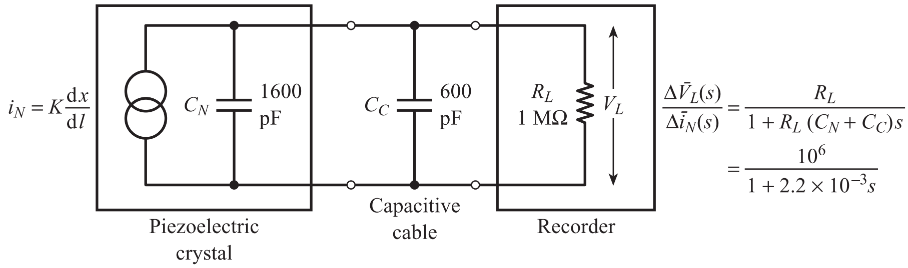


Figure 5.11 Piezoelectric force measurement system.

where $\mathrm{d}x/\mathrm{d}t$ is the velocity of the atomic deformations. This effect is discussed more fully in Section 8.7, where we see that the crystal acts as a capacitance C_N in parallel with the current source i_N. Figure 5.11 shows the equivalent circuit and typical component values for a crystal connected via a capacitive cable C_C to a recorder acting as a resistive load R_L. The voltage V_L across the load is given by $i_N Z$, where Z is the impedance of C_N, C_C and R_L in parallel. Since

$$\frac{1}{Z} = C_N s + C_C s + \frac{1}{R_L}$$

$$Z = \frac{R_L}{1 + R_L(C_N + C_C)s}$$

where s denotes the Laplace operator. The transfer function relating dynamic changes in source current and recorder voltage is thus:

$$\frac{\Delta\bar{V}_L(s)}{\Delta\bar{i}_N(s)} = \frac{R_L}{1 + R_L(C_N + C_C)s} \qquad [5.16]$$

Thus the effect of electrical loading in this example is to introduce a first-order transfer function into the force measurement system; this will affect dynamic accuracy.

5.2 Two-port networks

5.2.1 Generalised effort and flow variables

We have seen in the previous section how electrical loading effects can be described using a pair of variables, voltage and current. Voltage is an example of an **across** or **effort variable** y; current is an example of a **through** or **flow variable** $\dot{x}$. An effort variable drives a flow variable through an impedance. Other examples of effort/flow pairs are force/velocity, torque/angular velocity, pressure difference/volume flow rate and temperature difference/heat flow rate.[2] Each y–$\dot{x}$ pair has the following properties:

(a) The product $y\dot{x}$ represents power in watts.
(b) The ratio $y/\dot{x}$ represents impedance.

The only exception is the thermal variables where the product has the dimensions of watts and temperature. Table 5.1 lists the effort/flow pairs for different forms of energy and for each pair defines the related quantities of impedance, stiffness, compliance

Table 5.1 Flow/effort variables and related quantities.

Variables	$\frac{d}{dt}\dot{x}$	$\int \dot{x}\,dt$	Flow $\dot{x}$	Effort y	Impedance $\frac{y}{\dot{x}}$	Stiffness $\frac{y}{\int \dot{x}\,dt}$	Compliance $\frac{\int \dot{x}\,dt}{y}$	Inertance $\frac{y}{d\dot{x}/dt}$
Mechanical-translation	acceleration	displacement	velocity	force	$\frac{\text{force}}{\text{velocity}} =$ damping constant	$\frac{\text{force}}{\text{displacement}} =$ mechanical stiffness	$\frac{\text{displacement}}{\text{force}} = \frac{1}{\text{mech. stiffness}}$	$\frac{\text{force}}{\text{acceleration}} =$ mass
Mechanical-rotation	angular acceleration	angular displacement	angular velocity	torque	$\frac{\text{torque}}{\text{ang. velocity}} =$ damping constant	$\frac{\text{torque}}{\text{angular disp.}} =$ mechancial stiffness	$\frac{\text{angular disp.}}{\text{torque}} = \frac{1}{\text{mech. stiffness}}$	$\frac{\text{torque}}{\text{angular accn}} =$ moment of inertia
Electrical	$\frac{d}{dt}$ (current)	charge	current	voltage	$\frac{\text{voltage}}{\text{current}} =$ electrical resistance	$\frac{\text{voltage}}{\text{charge}} = \frac{1}{\text{elect. capacitance}}$	$\frac{\text{charge}}{\text{voltage}} =$ electrical capacitance	$\frac{\text{voltage}}{\frac{d}{dt}\text{(current)}} =$ inductance
Fluidic		volume	volume flow rate	pressure	$\frac{\text{pressure}}{\text{vol. flow rate}} =$ fluidic resistance	$\frac{\text{pressure}}{\text{volume}} = \frac{1}{\text{fluid capacitance}}$	$\frac{\text{volume}}{\text{pressure}} =$ fluidic capacitance	$\frac{\text{pressure}}{\frac{d}{dt}\text{(flow rate)}} =$ fluidic inertance
Thermal		heat	heat flow rate	temperature	$\frac{\text{temperature}}{\text{heat flow rate}} =$ thermal resistance	$\frac{\text{temperature}}{\text{heat}} = \frac{1}{\text{therm. capacitance}}$	$\frac{\text{heat}}{\text{temperature}} =$ thermal capacitance	

Source: adapted from Finkelstein and Watts, 1971[2].

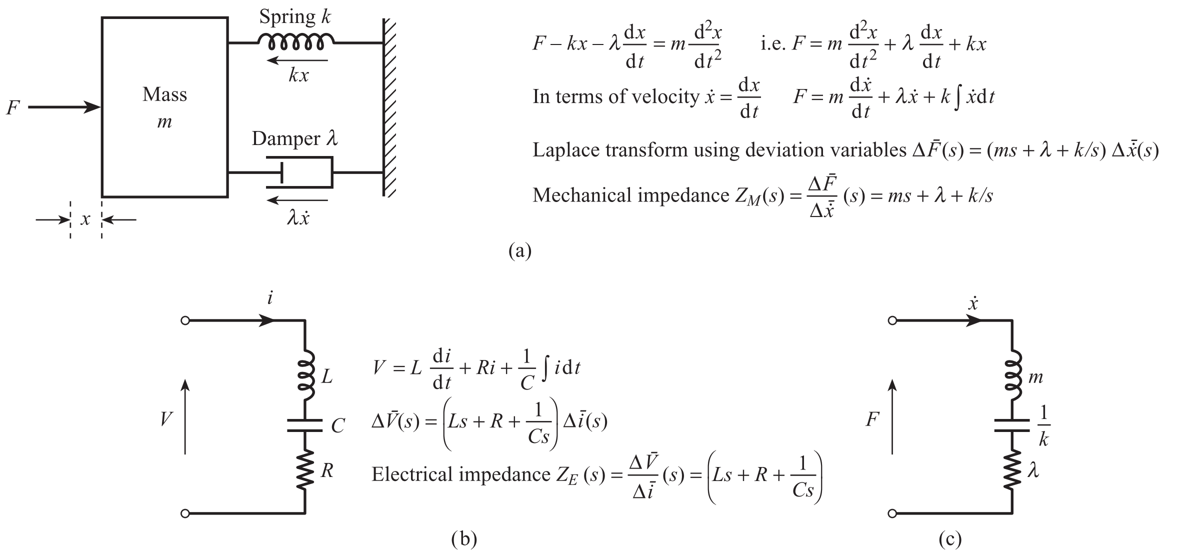


Figure 5.12 Equivalent circuit for a mechanical system:
(a) Parallel mechanical system
(b) Series electrical circuit
(c) Equivalent mechanical circuit.

and inertance. Thus we see that the concept of impedance is applicable to mechanical, fluidic and thermal systems as well as electrical. For a mechanical system, mass is analogous to electrical inductance, damping constant is analogous to electrical resistance, and 1/stiffness is analogous to electrical capacitance. For a thermal system, thermal resistance is analogous to electrical resistance, and thermal capacitance is analogous to electrical capacitance. This means we can generalise the electrical equivalent circuits of Thévenin and Norton to non-electrical systems.

Figure 5.12(a) shows a parallel mechanical system consisting of a mass m, spring stiffness k and damper constant λ. Figure 5.12(b) shows a series electrical circuit consisting of an inductance L, a capacitance C and a resistance R. Since mechanical impedance is the ratio of force/velocity, the mechanical impedance transfer function is

$$Z_M(s) = \frac{\Delta\bar{F}}{\Delta\bar{\dot{x}}} = ms + \lambda + \frac{k}{s} \qquad [5.17]$$

The impedance transfer function for the electrical circuit is

$$Z_E(s) = \frac{\Delta\bar{V}}{\Delta\bar{i}} + Ls + R + \frac{1}{Cs} \qquad [5.18]$$

We see that these have a similar form, with m corresponding to L, λ corresponding to R, and k corresponding to $1/C$. Thus the parallel mechanical system can be represented by an equivalent circuit consisting of an inductive element m, a resistive element λ and a capacitive element $1/k$ in series (Figure 5.12(c)).

Figure 5.13(a) shows a thermal system consisting of a body at temperature T immersed in a fluid at temperature T_F. The body has a mass M, specific heat C_H and surface area A; U is the heat transfer coefficient between the body and the fluid. Figure 5.13(b) shows a series electrical circuit with resistance R, capacitance C, input voltage V_{IN} and output voltage V_{OUT}. The differential equation for the thermal system is:

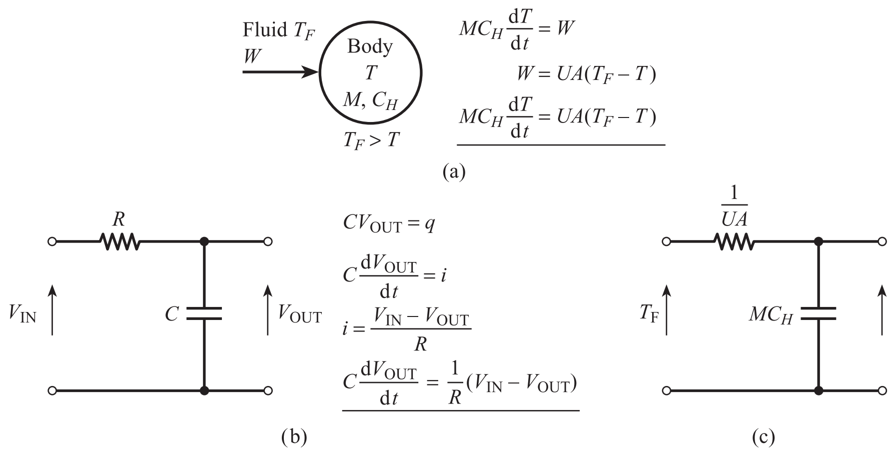


Figure 5.13 Equivalent circuit for a thermal system:
(a) Thermal system
(b) Electrical circuit
(c) Equivalent thermal circuit.

$$MC_H \frac{dT}{dt} = UA(T_F - T) \qquad [5.19]$$

and the differential equation for the electrical circuit is:

$$C\frac{dV_{OUT}}{dt} = \frac{1}{R}(V_{IN} - V_{OUT}) \qquad [5.20]$$

We see that the equations have a similar form and the temperature effort variables T_F and T are analogous to the voltage effort variables V_{IN} and V_{OUT}. Heat flow W is analogous to current flow i, UA is analogous to $1/R$, i.e. the reciprocal of electrical resistance, and MC_H is analogous to electrical capacitance C. Thus the thermal system can be represented by an equivalent circuit consisting of a resistive element $1/UA$ in series with a capacitive element MC_H as in Figure 5.13(c).

5.2.2 Two-port networks

We saw in Section 5.1 that the electrical output of a sensing element such as a thermocouple or piezoelectric crystal can be represented by a Thévenin or Norton equivalent circuit. The sensor has therefore two output terminals which allow both voltage and current flow to be specified; this is referred to as an electrical output port. The sensing element will have a mechnical, thermal or fluidic input; we saw in the previous section that mechanical and thermal systems can be represented by equivalent circuits which show the relation between the corresponding effort and flow variables. Thus the input to a mechanical or thermal sensor can be represented by two input terminals which allow both the effort and flow variables to be specified; this is either a mechanical or a thermal input port. Thus a sensing element can be represented by a **two-port** or four-terminal network. Figure 5.14(a) shows a two-port representation of a mechanical sensor with input mechanical port and output electrical port; Figure 5.15(a) shows the two-port representation of a thermal sensor.

Figures 5.14(b) and (c) show the detailed two-port networks for a range of mechanical sensing elements. Figure 5.14(b) shows the equivalent circuit for sensing elements with a Thévenin equivalent circuit at the electrical output port. Z_M is the input mechanical impedance; E_{Th} and Z_{Th} are the Thévenin voltage and impedance. For a displacement sensor E_{Th} is proportional to displacement x, i.e.

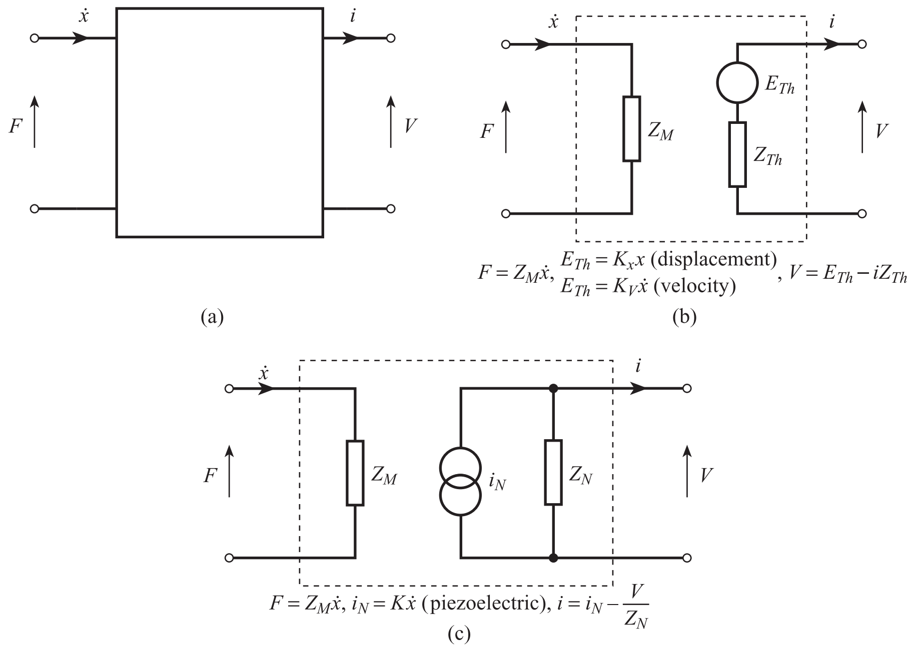


Figure 5.14 Mechanical sensing elements as two-port networks:
(a) Overall two-port representation
(b) Equivalent circuit with Thévenin output
(c) Equivalent circuit with Norton output.

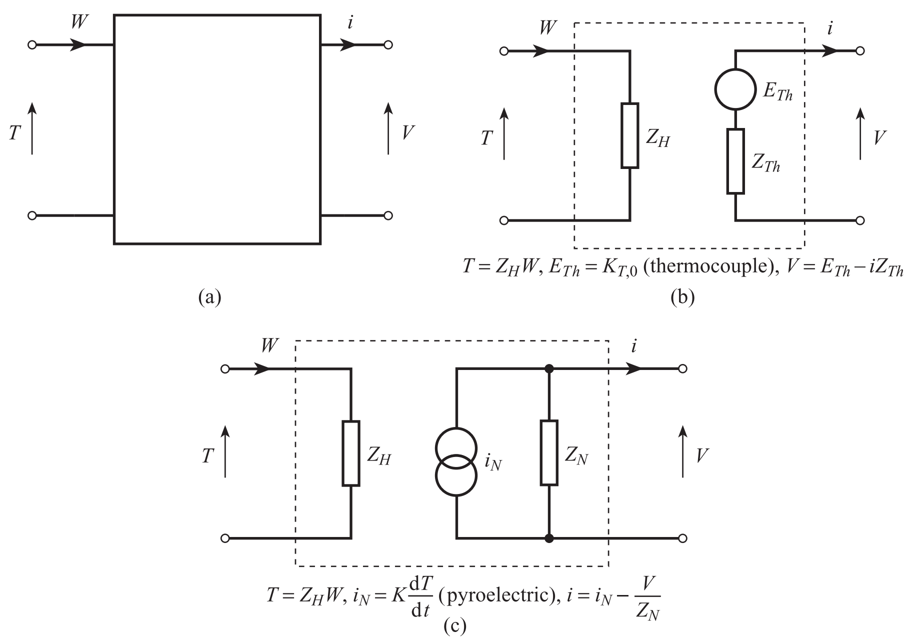


Figure 5.15 Thermal sensing elements as two-port networks:
(a) Overall two-port representation
(b) Equivalent circuit with Thévenin output
(c) Equivalent circuit with Norton output.

$$E_{Th} = K_x x \qquad [5.21]$$

where K_x is the sensitivity and $x = \int \dot{x}\, dt$.

For a potentiometer displacement sensor (Section 5.1.2) $K_x = V_S$ (supply voltage), and for a linear variable differential transformer (LVDT, Section 8.3.2) K_x is the slope of the linear portion of the a.c. voltage versus displacement characteristics. For a velocity sensor, E_{Th} is proportional to velocity $\dot{x}$, i.e.

$$E_{Th} = K_V \dot{x} \qquad [5.22]$$

For an electromagnetic linear velocity sensor (Section 12.5.1) $K_V = Bl$ where B is the applied magnetic field and l the length of the conductor. In the case of an electromagnetic angular velocity sensor or tachogenerator (Section 8.4) we have $E_{Th} = K_V \omega_r$, where ω_r is the angular velocity and $K_V = \mathrm{d}N/\mathrm{d}\theta$ the rate of change of flux N with angle θ. Figure 5.14(c) shows the equivalent circuit for sensing elements with a Norton equivalent circuit i_N, Z_N at the electrical output port. For a piezoelectric sensor i_N is proportional to velocity $\dot{x}$ (Section 8.7), i.e.

$$i_N = K\dot{x} \qquad [5.23]$$

where $K = dk$, d is the charge sensitivity to force and k is the stiffness of the crystal.

Figures 5.15(b) and (c) show the detailed two-port networks for two examples of thermal sensing elements. Figure 5.15(b) shows the equivalent circuit for a sensing element with a Thévenin equivalent circuit E_{Th}, Z_{Th} at the electrical output port; Z_H is the thermal input impedance. For a thermocouple temperature sensor (Section 8.5) with the reference junction at 0 °C we have $E_{Th} = E_{T,0}$ where $E_{T,0}$ is the e.m.f. at the measured junction at T °C and is given by the power series

$$E_{T,0} = a_1 T + a_2 T^2 + a_3 T^3 + \ldots \qquad [5.24]$$

Figure 5.15(c) shows the equivalent circuit for a sensing element with a Norton equivalent circuit i_N, Z_N at the electrical output port. For a pyroelectric detector (Section 15.5.1) i_N is proportional to rate of change of temperature $\mathrm{d}T/\mathrm{d}t$, i.e.

$$i_N = K\frac{\mathrm{d}T}{\mathrm{d}t} \qquad [5.25]$$

where $K = A(\mathrm{d}P/\mathrm{d}T)$, A is the area of the electrodes and $\mathrm{d}P/\mathrm{d}T$ is the slope of the polarisation temperature characteristics.

5.2.3 Process loading

Having introduced the concepts of equivalent circuits and two-port networks for mechanical and thermal systems, we can now use these concepts to study examples of how a primary sensing element can 'load' the process or element being measured.

Figure 5.16 shows a mechanical system or 'process' represented by a mass, spring and damper. The force F applied to the 'process' is being measured by a force sensor, consisting of an elastic element in conjunction with a potentiometric displacement sensor. The elastic force sensor can also be represented by a mass spring and damper (Section 4.1.2). Under steady-state conditions when both velocity $\dot{x} = 0$ and acceleration $\ddot{x} = 0$, we have the following force balance equations:

process $F = k_P x + F_S$ [5.26]

sensor $F_S = k_S x$

showing that the relationship between the measured force F_S and the true force F is:

Steady-state loading of mechanical system

$$F_S = \frac{k_S}{k_S + k_P} F = \frac{1}{1 + k_P/k_S} F \qquad [5.27]$$

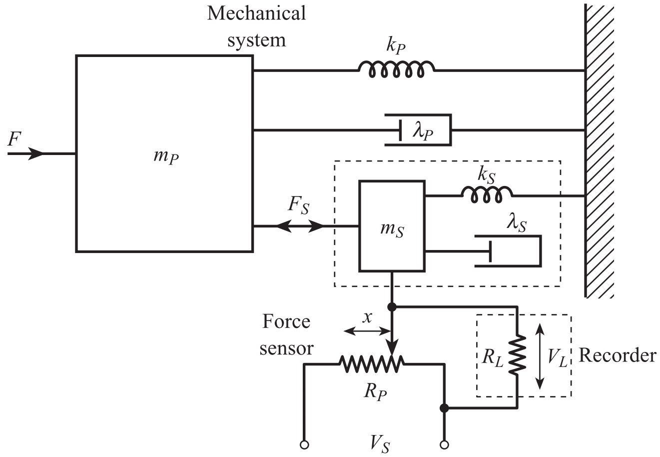


Figure 5.16 Loading of mechanical system by force sensor.

We see that in order to minimise the loading error in the steady state the sensor stiffness k_S should be very much greater than the process stiffness k_P.

Under unsteady conditions when $\dot{x}$ and $\ddot{x}$ are non-zero, Newton's second law gives the following differential equations:

$$\textbf{process} \quad F - k_P x - \lambda_P \dot{x} - F_S = m_P \ddot{x} \qquad [5.28]$$

$$\textbf{sensor} \quad F_S - k_S x - \lambda_S \dot{x} = m_S \ddot{x}$$

i.e.

$$m_P \frac{\mathrm{d}\dot{x}}{\mathrm{d}t} + \lambda_P \dot{x} + k_P \int \dot{x}\,\mathrm{d}t = F - F_S$$

$$m_S \frac{\mathrm{d}\dot{x}}{\mathrm{d}t} + \lambda_S \dot{x} + k_S \int \dot{x}\,\mathrm{d}t = F_S \qquad [5.29]$$

Using the analogues given earlier, the sensor can be represented by F_S driving $\dot{x}$ through the mechanical L, C, R circuit m_S, $1/k_S$, λ_S; and the process can be represented by $F - F_S$ driving $\dot{x}$ through the mechanical L, C, R circuit m_P, $1/k_P$, λ_P. If Δx, ΔF and ΔF_S are deviations from initial steady conditions, then the Laplace transforms of eqns [5.29] are:

$$\left(m_P s + \lambda_P + \frac{k_P}{s}\right)\overline{\Delta \dot{x}} = \Delta \bar{F} - \Delta \bar{F}_S$$

$$\left(m_S s + \lambda_S + \frac{k_S}{s}\right)\overline{\Delta \dot{x}} = \Delta \bar{F}_S \qquad [5.30]$$

Using Table 5.1 we can define mechanical impedance transfer functions by $Z_M(s) = \Delta \bar{F}(s)/\overline{\Delta \dot{x}}(s)$, so that:

$$\textbf{process impedance} \quad Z_{MP}(s) = m_P s + \lambda_P + \frac{k_P}{s}$$

$$\textbf{sensor impedance} \quad Z_{MS}(s) = m_S s + \lambda_S + \frac{k_S}{s} \qquad [5.31]$$

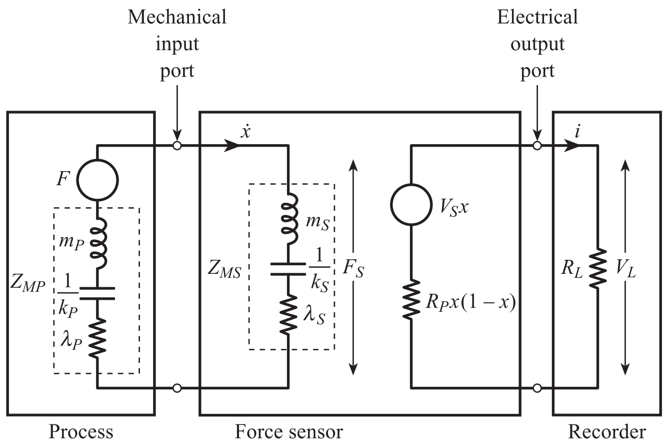


Figure 5.17 Equivalent circuit for complete system showing a force sensor as a two-port network.

From [5.30] and [5.31] the relationship between measured and actual dynamic changes in force is:

Dynamic loading of mechanical system

$$\Delta \bar{F}_S(s) = \frac{Z_{MS}}{Z_{MS} + Z_{MP}} \Delta \bar{F}(s) \qquad [5.32]$$

Thus in order to minimise dynamic loading effects, sensor impedance Z_{MS} should be very much greater than process impedance Z_{MP}. Figure 5.17 shows the equivalent circuit for the system: process, force sensor and recorder.

Figure 5.18 shows a hot body, i.e. a thermal 'process' whose temperature T_P is being measured by a thermocouple sensor. Under unsteady conditions, heat flow rate considerations give the following differential equations (Section 4.1):

$$\begin{aligned} &\textbf{process} \quad M_P C_P \frac{dT_P}{dt} = W_P - W_S, \quad W_P = U_P A_P (T_F - T_P) \\ &\textbf{sensor} \quad M_S C_S \frac{dT_S}{dt} = W_S, \qquad\quad W_S = U_S A_S (T_P - T_S) \end{aligned} \qquad [5.33]$$

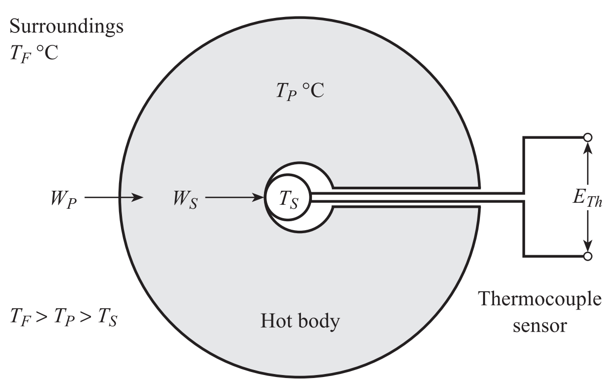


Figure 5.18 Loading of thermal 'process' by thermocouple temperature sensor.

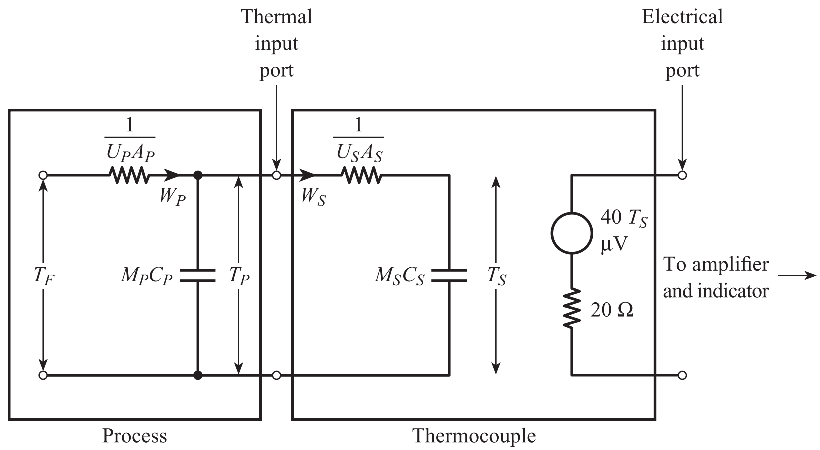


Figure 5.19 Equivalent circuit for thermal system showing thermocouple as a two-port network.

where M denotes masses
C denotes specific heats
U denotes heat transfer coefficients
A denotes heat transfer areas.

The quantities $M_P C_P$ and $M_S C_S$ have the dimensions of heat/temperature and are analogous to electrical capacitance. The quantities $U_P A_P$ and $U_S A_S$ have the dimensions of heat flow rate/temperature and are analogous to 1/(electrical resistance). The equivalent circuit for the process and thermocouple is shown in Figure 5.19. We see that the relationship between T_F and T_P depends on the potential divider $1/U_P A_P$, $M_P C_P$, and the relationship between T_P and T_S depends on the potential divider $1/(U_S A_S)$, $M_S C_S$. Again the thermocouple can be represented as a two-port network with a thermal input port and an electrical output port.

In conclusion we see that the representation of measurement system elements by two-port networks enables both **process** and **inter-element** loading effects to be quantified.

5.2.4 Bilateral transducers

Bilateral transducers are associated with reversible physical effects. In a reversible effect the same device can, for example, convert mechanical energy into electrical energy and also convert electrical energy into mechanical energy. When the device converts electrical energy into mechanical energy it acts as a **transmitter** or **sender**.

This can be represented by a two-port network with an input electrical port and an output mechanical port as in Figure 5.20(a), When the device converts mechanical energy into electrical energy it acts as a **receiver** or **sensor**, and can be represented by an input mechanical port and an output electrical port as in Figure 5.20(b).

The piezoelectric effect is a common example of a reversible effect (Section 8.7). In the **direct effect** a force F applied to the crystal produces a charge q, proportional to F, according to:

$$q = \mathrm{d}F \qquad [5.34]$$

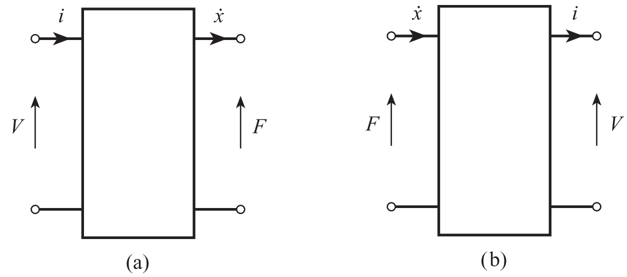


Figure 5.20 Bilateral transducers:
(a) Transmitter/sender
(b) Receiver/sensor.

i.e. this is a conversion of mechanical energy to electrical energy and the device acts as a receiver or sensor. In the **inverse effect** a voltage V applied to the crystal produces a mechanical deformation x, proportional to V, according to:

$$x = \mathrm{d}V \qquad [5.35]$$

This is a conversion of electrical energy to mechanical energy and the device acts as a transmitter or sender. The detailed equivalent circuits for a piezoelectric transmitter and receiver are given in Section 16.2.1.

Another reversible physical effect is the **electromagnetic effect**. In the direct effect, a conductor of length l moving with velocity $\dot{x}$ perpendicular to a magnetic field B has a voltage

$$E = Bl\dot{x} \qquad [5.36]$$

induced across the ends of the conductor. This is a conversion of mechanical to electrical energy and the device acts as a receiver or sensor. In the inverse effect a conductor of length l carrying a current i in a transverse magnetic field B experiences a force

$$F = Bli \qquad [5.37]$$

This is a conversion of electrical to mechanical energy and the device acts as a transmitter or sender. Figure 5.21 gives the detailed equivalent circuits for an electromagnetic transmitter and receiver. At the electrical ports the applied voltage drives a current through the electrical impedance L, R; at the mechanical ports the applied force drives a velocity through the mechanical impedance m, $1/k$, λ. The transmitter can be used as a voltage indicator (Section 11.2) and the receiver as a velocity sensor (Sections 5.1.1 and 8.4).

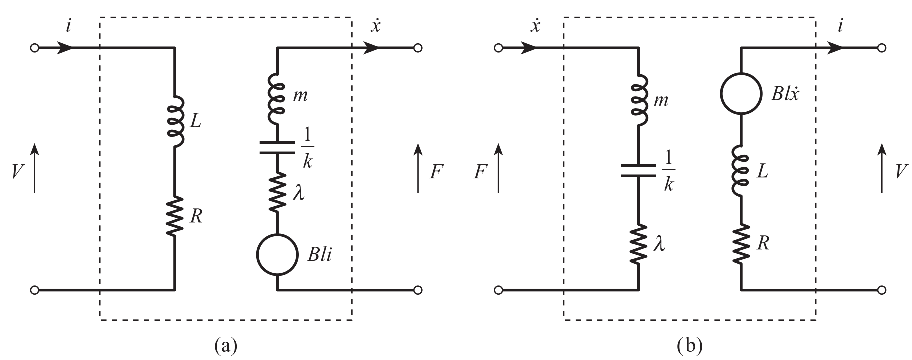


Figure 5.21 Equivalent circuits for bilateral electromagnetic transducers:
(a) Transmitter/sender
(b) Receiver/sensor.

Conclusion

We have seen how the use of equivalent circuits and two-port networks has enabled both inter-element and process loading effects to be described.

References

[1] BELL E C and WHITEHEAD R W 1979 *Basic Electrical Engineering and Instrumentation for Engineers*, Granada, London, pp. 46–7.
[2] FINKELSTEIN L and WATTS R D 1971 'Systems analysis of instruments', *Journal of the Institute of Measurement and Control*, vol. 4, Sept., pp. 236–7.

Problems

5.1 A glass pH electrode with a sensitivity of 59 mV pH^{-1} and a resistance of 10^9 Ω is used to measure pH in the range 0 to 15. The electrode is to be connected to a recorder of input range 0 to 100 mV and resistance 100 Ω using a buffer amplifier of unity gain and output resistance 100 Ω.

(a) Calculate the input impedance of the amplifier, and the sensitivity of the recorder scale necessary to obtain an accurate recording of pH.
(b) The resistance of the electrode increases to 2×10^9 Ω due to chemical action. Calculate the resulting measurement error in the above system, as a percentage of full scale, for a true pH of 7.

5.2 The motion of a hydraulic ram is to be recorded using a potentiometer displacement sensor connected to a recorder. The potentiometer is 25 cm long and has linear resistance displacement characteristics. A set of potentiometers with maximum power rating of 5 W and resistance values ranging from 250 to 2500 Ω in 250 Ω steps is available. The recorder has a resistance of 5000 Ω and the non-linear error of the system must not exceed 2% of full scale. Find:

(a) the maximum potentiometer sensitivity that can be obtained;
(b) the required potentiometer resistance and supply voltage in order to achieve maximum sensitivity.

5.3 An electronic differential transmitter gives a current output of 4 to 20 mA linearly related to a differential pressure input of 0 to 10^4 Pa. The Norton impedance of the transmitter is 10^5 Ω. The transmitter is connected to an indicator of impedance 250 Ω via a cable of total resistance 500 Ω. The indicator gives a reading between 0 and 10^4 Pa for an input voltage between 1 and 5 V. Calculate the system measurement error, due to loading, for an input pressure of 5×10^3 Pa.

5.4 A sensor with mass 0.1 kg, stiffness 10^3 N m^{-1} and damping constant 10 N s m^{-1} is used to measure the force on a mechanical structure of mass 5 kg, stiffness 10^2 N m^{-1} and damping constant 20 N s m^{-1}. Find the transfer function relating measured and actual changes in force.

Basic problems

5.5 A linear thermocouple with a sensitivity of 0.04 mV/°C and resistance of 100 Ω is connected to a load with a resistance of 1 kΩ. Find the voltage across the load for a temperature of 250 °C.

5.6 A potentiometer displacement sensor has a supply voltage of 15 V and a resistance of 50 kΩ. The fractional displacement of the wiper is 0.3. Find the Thévenin voltage and resistance for the circuit.

5.7 A potentiometer has a supply voltage of 10 V, a resistance of 10 kΩ and a length of 10 cm. A recorder of resistance 10 kΩ is connected across the potentiometer. Calculate the Thévenin equivalent circuit for the sensor and the recorder voltage for each of the following displacements:

(a) 2 cm
(b) 5 cm
(c) 8 cm.

5.8 A potentiometer has a total length of 10 cm and a resistance of 100 Ω.

(a) Calculate the supply voltage so that power dissipation = 1 W.
(b) Draw the Thévenin equivalent circuit for 7 cm displacement.
(c) The potentiometer is connected to a recorder with a resistance R_L. Find R_L such that the recorder voltage is 5% less than the open circuit voltage at 7 cm displacement.

5.9 A pressure transducer consists of a Bourdon tube elastic element connected to a potentiometer displacement sensor. The input range of the Bourdon tube is 0 to 10^4 Pa and the output range is 0 to 1 cm. The potentiometer has a length of 1 cm, a resistance of 10 kΩ and a supply voltage of 10 V. If the input pressure is 5×10^3 Pa, calculate:

(a) the displacement of the potentiometer wiper (assume a linear Bourdon tube)
(b) the open circuit transducer output voltage
(c) the voltage indicated by a voltmeter of resistance 10 kΩ connected across the potentiometer.

6 Signals and Noise in Measurement Systems

6.1 Introduction

In Chapter 4 we studied the dynamic response of measurement systems to step, sine wave and square wave input signals. These signals are examples of **deterministic signals**: a deterministic signal is one whose value at any future time can be exactly predicted. Thus if we record these signals for an observation period T_O (Figure 6.1), the future behaviour of the signal, once the observation period is over, is known exactly. The future behaviour of real processes, such as chemical reactors, blast furnaces and aircraft, will depend on unknown factors such as the type of feedstock, reliability of equipment, changes in throughput and atmospheric conditions, and cannot be known in advance. This means that the future value of measured variables, such as reactor temperature, flow in a pipe and aircraft speed, *cannot* be exactly predicted. Thus in real measurement applications the input signal to the measurement system is not deterministic but **random**. If a random signal is recorded for an observation period T_O (Figure 6.1) the behaviour of the signal, once the observation period is over, is *not* known exactly. However, five statistical quantities – mean, standard deviation, probability density function, power spectral density and autocorrelation function – are used to estimate the behaviour of random signals. These are defined and explained in Section 6.2.

Random variations in the measured input variable produce corresponding random variations in the electrical output signals of elements in the system. More precisely the Thévenin voltage and Norton current signals defined in Chapter 5 vary randomly with time and will be referred to as the **measurement signal**. Examples are random fluctuations in the millivolt output of a thermocouple, random fluctuations in the current output of a differential pressure transmitter, and random fluctuations in the amplitude and frequency of the a.c. output voltage of a variable reluctance tacho-generator. However, unwanted electrical signals may also be present in the measurement circuit. These may be due to sources inside the measurement circuit or caused by coupling to sources outside the circuit. The magnitude of the unwanted signal may be comparable to or larger than that of the measurement signal itself, resulting in a measurement error for the overall system. This is an example of the **interfering inputs** discussed in Chapter 2. The unwanted signal then may be either random, e.g. signals caused by the random motion of electrons, or deterministic, e.g. sinusoidal signals at 50 Hz caused by power cables. Unwanted random signals are usually referred to as **noise signals** and unwanted deterministic signals as **interference signals**.

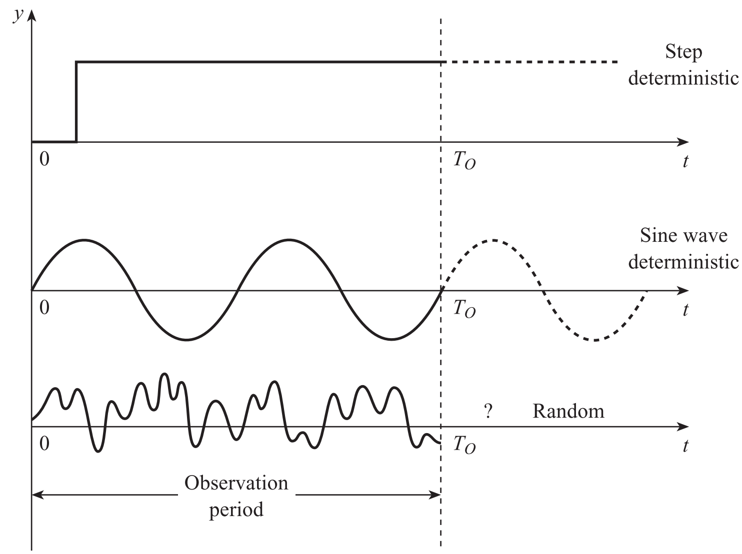


Figure 6.1 Deterministic and random signals.

Sections 6.3 and 6.4 discuss the sources of noise and interference signals and how they affect the measurement circuit. Section 6.5 examines ways of reducing the effects of noise and interference.

6.2 Statistical representation of random signals

Figure 6.2 shows a recording of a section of a random signal obtained during an observation period T_O. Since the signal is random we cannot write down a continuous algebraic equation $y(t)$ for the signal voltage y at time t. We can, however, write down the values y_1 to y_N of N samples taken at equal intervals ΔT during T_O. The first sample y_1 is taken at $t = \Delta T$, the second y_2 is taken at $t = 2\Delta T$, and the ith y_i is taken at $t = i\Delta T$, where $i = 1, \ldots, N$. The sampling interval $\Delta T = T_O/N$ must satisfy the Nyquist sampling theorem, which is explained in Section 10.1. We can now use these samples to calculate statistical quantities for the observed section of the signal. These observed statistical quantities will provide a good estimate of the future behaviour of the signal, once the observation period is over, provided:

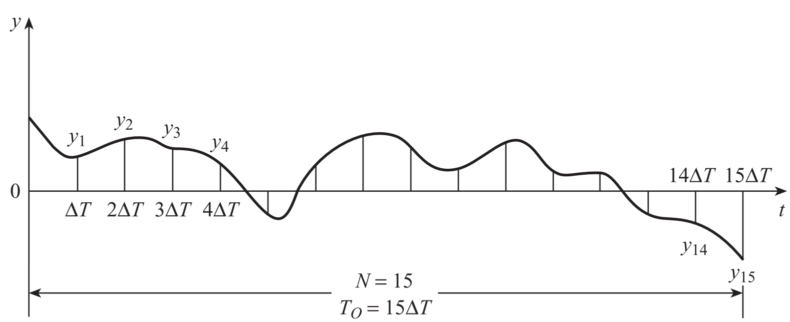


Figure 6.2 Sampling of a random signal.

(a) T_O is sufficiently long, i.e. N is sufficiently large;
(b) the signal is **stationary**, i.e. long-term statistical quantities do not change with time.

6.2.1 Mean $\bar{y}$

For a signal defined in terms of a continuous function $y(t)$ in the interval 0 to T_O, the mean is given by:

Mean for continuous signal

$$\bar{y} = \frac{1}{T_O}\int_0^{T_O} y(t)\,\mathrm{d}t \qquad [6.1]$$

If, however, the signal is represented by the set of sampled values y_i we have

Mean for sampled signal

$$\bar{y} = \frac{1}{N}\sum_{i=1}^{i=N} y_i \qquad [6.2]$$

6.2.2 Standard deviation σ

This is a measure of the average spread or deviation of the signal from the mean value $\bar{y}$. In the continuous case:

Standard deviation for continuous signal

$$\sigma^2 = \frac{1}{T_O}\int_0^{T_O} [y(t) - \bar{y}]^2\,\mathrm{d}t \qquad [6.3]$$

and in the sampled case:

Standard deviation for sampled signal

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{i=N} (y_i - \bar{y})^2 \qquad [6.4]$$

In the special case that $\bar{y} = 0$, the standard deviation σ is equal to the **root mean square** (r.m.s.) value y_{rms}, where:

$$y_{\mathrm{rms}} = \sqrt{\frac{1}{T_O}\int_0^{T_O} y^2\,\mathrm{d}t} \qquad [6.5]$$

or

$$y_{\mathrm{rms}} = \sqrt{\frac{1}{N}\sum_{i=1}^{i=N} y_i^2} \qquad [6.6]$$

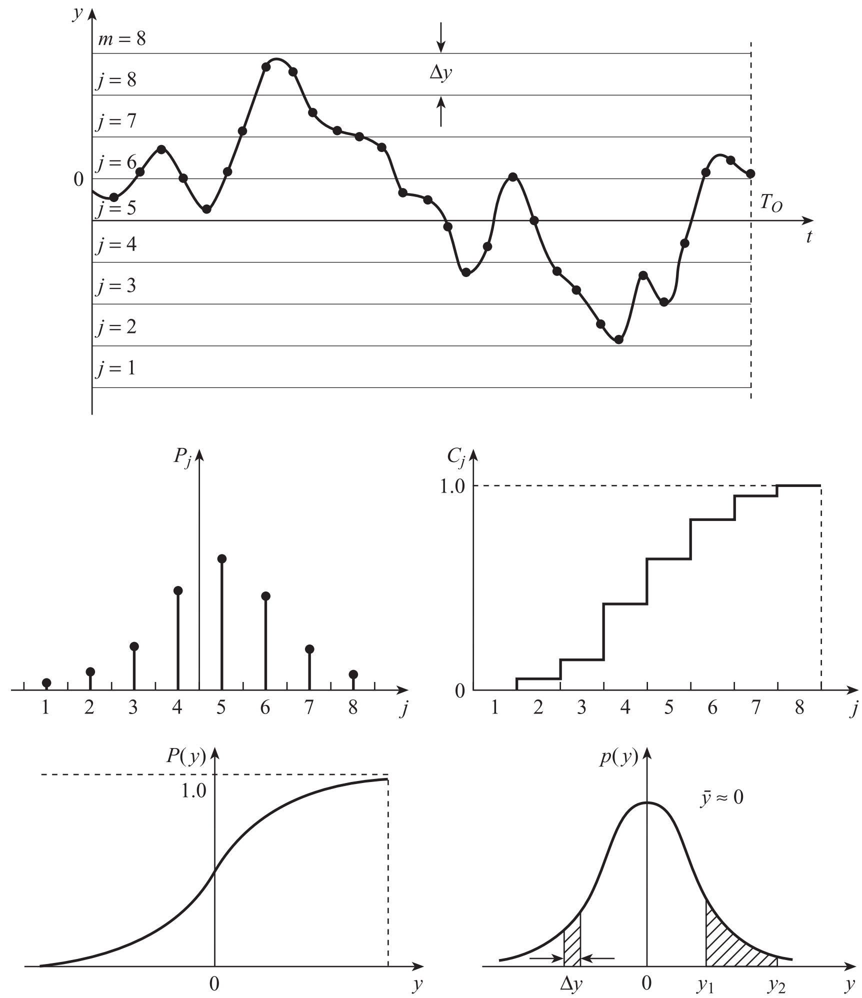


Figure 6.3 Probability and probability density.

6.2.3 Probability density function *p*(*y*)

This is a function of signal value y and is a measure of the probability that the signal will have a certain range of values. Figure 6.3 shows the set of sample values y_i and the y axis divided into m sections each of width Δy. We can then count the number of samples occurring within each section, i.e. n_1 in section 1, n_2 in section 2, n_j in section j, etc., where $j = 1, \ldots, m$. The **probability** P_j of the signal occurring in the jth section is thus:

$$P_j = \frac{\text{number of times sample occurs in the } j\text{th section}}{\text{total number of samples}} \qquad [6.7]$$

$$= \frac{n_j}{N}, \quad j = 1, \ldots, m$$

The **cumulative probability** C_j is the total probability that the signal will occur in the first j sections and is given by:

Cumulative probability

$$C_j = P_1 + P_2 + \ldots + P_j = \frac{1}{N}(n_1 + n_2 + \ldots + n_j) \quad [6.8]$$

Figure 6.3 shows the corresponding forms of P_j and C_j. The final value of C_j when $j = m$ is:

$$C_m = \frac{1}{N}(n_1 + n_2 + \ldots + n_m)$$

$$= \frac{1}{N} \times N = 1 \quad (\text{since } n_1 + n_2 + \ldots + n_m = N) \quad [6.9]$$

i.e. the total probability of finding the signal in all m sections is unity. In the limit as the interval Δy tends to zero the discrete cumulative probability C_j tends to a continuous function (Figure 6.3). This is the **cumulative probability distribution function** (c.d.f.) $P(y)$, which is defined by:

Cumulative probability distribution function

$$P(y) = \lim_{\Delta y \to 0} C_j \quad [6.10]$$

The **probability density function** (p.d.f.) $p(y)$ is more commonly used and is the derivative of $P(y)$, i.e.

Probability density function

$$p(y) = \frac{dP}{dy} \quad [6.11]$$

Thus the probability $P_{y,y+\Delta y}$ that the signal will lie between y and $y + \Delta y$ is given by:

$$P_{y,y+\Delta y} = \Delta P = p(y)\Delta y \quad [6.12]$$

i.e. by the area of a strip of height $p(y)$ and width Δy. Similarly the probability P_{y_1,y_2} that the signal will lie between y_1 and y_2 is given by:

$$P_{y_1,y_2} = \int_{y_1}^{y_2} p(y)\, dy \quad [6.13]$$

i.e. the area under the probability density curve between y_1 and y_2. The total area under the probability density curve is equal to unity, corresponding to the total probability of the signal having any value of y. The normal probability density function (Section 2.3), i.e.

$$p(y) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(y - \bar{y})^2}{2\sigma^2}\right] \quad [6.14]$$

usually provides an adequate description of the amplitude distribution of random noise signals.

6.2.4 Power spectral density $\phi(\omega)$

If we record a random signal for several observation periods, each of length T_O (Figure 6.4), the waveform will be different for each period. However, the average signal power will be approximately the same for each observation period. This means that signal power is a **stationary** quantity which can be used to quantify random signals. In Section 4.3 we saw that a **periodic** signal can be expressed as a **Fourier series**, i.e. a sum of sine and cosine waves with frequencies which are harmonics of the fundamental frequency. The power in a periodic signal is therefore distributed amongst these harmonic frequencies. A random signal is not periodic and cannot be represented by Fourier series but does contain a large number of closely spaced frequencies. Power spectral density is a stationary quantity which is a measure of how the power in a random signal is distributed amongst these different frequencies.

In order to explain the meaning of $\phi(\omega)$ we approximate the random signal by a periodic signal (Figure 6.4) in which the waveform recorded during the first observation period T_O is exactly repeated during each subsequent observation period. This periodic approximation is valid:

(a) in the limit that the observation period, i.e. the signal period T_O, becomes large,
(b) provided we use it to calculate the power distribution in the signal.

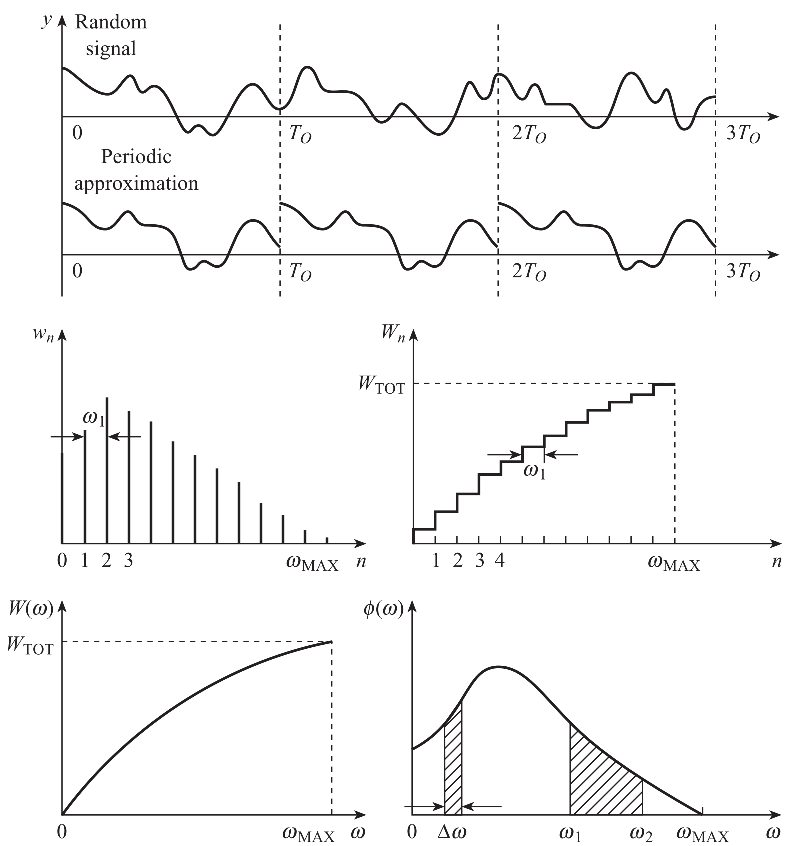


Figure 6.4 Power spectrum and power spectral density.

The Fourier series for a voltage signal with period T_O is

$$y(t) = a_0 + \sum_{n\;1}^{n\;\infty} a_n \cos n\omega_1 t + \sum_{n\;1}^{n\;\infty} b_n \sin n\omega_1 t \qquad [4.51a]$$

where the fundamental frequency $\omega_1 = 2\pi/T_O$, a_0 = mean $\bar{y}$, and

Fourier series coefficients for sampled signal

$$a_n = \frac{2}{N}\sum_{i=1}^{i=N} y_i \cos(n\omega_1 i\Delta T) = \frac{2}{N}\sum_{i=1}^{i=N} y_i \cos\left(2\pi n\frac{i}{N}\right)$$

$$b_n = \frac{2}{N}\sum_{i=1}^{i=N} y_i \sin(n\omega_1 i\Delta T) = \frac{2}{N}\sum_{i=1}^{i=N} y_i \sin\left(2\pi n\frac{i}{N}\right) \qquad [6.15]$$

Equations [6.15] use the set of sample values y_1 and are equivalent to eqns [4.52] for the Fourier coefficients of a continuous signal. If the nth harmonic $a_n \cos n\omega_1 t$ is applied across a 1 Ω resistor, the instantaneous power in the resistor at time t is $a_n^2 \cos^2 n\omega_1 t$ watts, and the average power over period T_O is:

$$a_n^2 \frac{1}{T_O}\int_0^{T_O} \cos^2 n\omega_1 t \, dt = \frac{a_n^2}{2} \qquad [6.16]$$

Similarly the average power in a 1 Ω resistor due to $b_n \sin n\omega_1 t$ is $b_n^2/2$, so that the power due to the nth harmonic at frequency $n\omega_1$ is:

Power due to nth harmonic

$$w_n = \tfrac{1}{2}(a_n^2 + b_n^2) \qquad [6.17]$$

Figure 6.4 shows that the relationship between w_n and ω is a series of lines at the harmonic frequencies $n\omega_1$. This is referred to as the **power spectrum** of the signal and is terminated at ω_{MAX}, the harmonic frequency beyond which w_n becomes negligible. The **cumulative power** W_n is the total power in a 1 Ω resistor due to the first n harmonics and the d.c. component a_0, i.e.

Cumulative power

$$W_n = w_0 + w_1 + w_2 + \ldots + w_n \qquad [6.18]$$

The diagram shows the relation between W_n and ω; it is in the form of a staircase, each step having width ω_1. At $\omega = \omega_{MAX}$, $W_n = W_{TOT}$, the total power in the signal. However, in the limit that signal period $T_O \to \infty$, $\omega_1 \to 0$ and W_n becomes a continuous function of ω. This is the **cumulative power function** (c.p.f.) $W(\omega)$, which is defined by:

Cumulative power function

$$W(\omega) = \lim_{\omega_1 \to 0} W_n \qquad [6.19]$$

The **power spectral density** (PSD) $\phi(\omega)$ is more commonly used and is the derivative of $W(\omega)$ (Figure 6.4), i.e.

Power spectral density

$$\phi(\omega) = \frac{dW}{d\omega} \qquad [6.20]$$

The PSD ϕ has units watt s rad^{-1} or W/Hz. Thus the power produced in a 1 Ω resistor due to frequencies between ω and $\omega + \Delta\omega$ is:

$$W_{\omega,\omega+\Delta\omega} = \Delta W = \phi(\omega)\Delta\omega \text{ watts} \qquad [6.21]$$

i.e. the area of a strip of height $\phi(\omega)$ and width $\Delta\omega$. Similarly the power due to frequencies between ω_1 and ω_2 is:

$$W_{\omega_1,\omega_2} = \int_{\omega_1}^{\omega_2} \phi(\omega)\, d\omega \text{ watts} \qquad [6.22]$$

i.e. the area under the power spectral density curve between ω_1 and ω_2. The total area under the PSD curve is the total power in the signal, i.e.

$$W_{\text{TOT}} = \int_{0}^{\omega_{\text{MAX}}} \phi(\omega)\, d\omega \text{ watts} \qquad [6.23]$$

Internal noise sources in electrical circuits can often be regarded as **white noise**, which has a uniform PSD over an infinite range of frequencies, i.e.

$$\phi(\omega) = A, \quad 0 \le \omega \le \infty \qquad [6.24]$$

Another useful representation for both noise and measurement signals is a power spectral density which is constant up to a cut-off frequency ω_C and zero for higher frequencies (**band limited white noise**):

$$\phi(\omega) = \begin{cases} A, & 0 \le \omega \le \omega_C \\ 0, & \omega > \omega_C \end{cases} \qquad [6.25]$$

i.e. $W_{\text{TOT}} = A\omega_C$.

The maximum frequency ω_{MAX} for a measurement signal depends on the nature of the measured variable. Thus the vibration displacement of part of a machine may contain frequencies up to many kHz, whereas the temperature variations in a chemical reactor may only contain frequencies up to 0.01 Hz. In order to accurately measure random fluctuations in a measured variable, the transfer function $G(s)$ for the measurement system must satisfy the conditions of Section 4.4, i.e. $|G(j\omega)| = 1$ and arg $G(j\omega) = 0$ for all ω up to ω_{MAX}.

6.2.5 Autocorrelation function

Figure 6.5 shows a block diagram of a simple correlator. The input signal $y(t)$ is passed through a variable time delay unit to give a delayed signal $y(t - \beta)$. The signals $y(t)$ and $y(t - \beta)$ are multiplied together to give the product waveform $y(t)y(t - \beta)$. This product waveform is passed through an averager and the average value $\overline{y(t)y(t - \beta)}$ is displayed on a meter: this is the **autocorrelation coefficient** R_{yy}. If the time delay β is altered, the shape of the product waveform changes, causing the meter reading R_{yy} to change. The relationship between R_{yy} and time delay β is the **autocorrelation function** $R_{yy}(\beta)$ of the signal. We note that $R_{yy}(\beta)$ has a maximum value

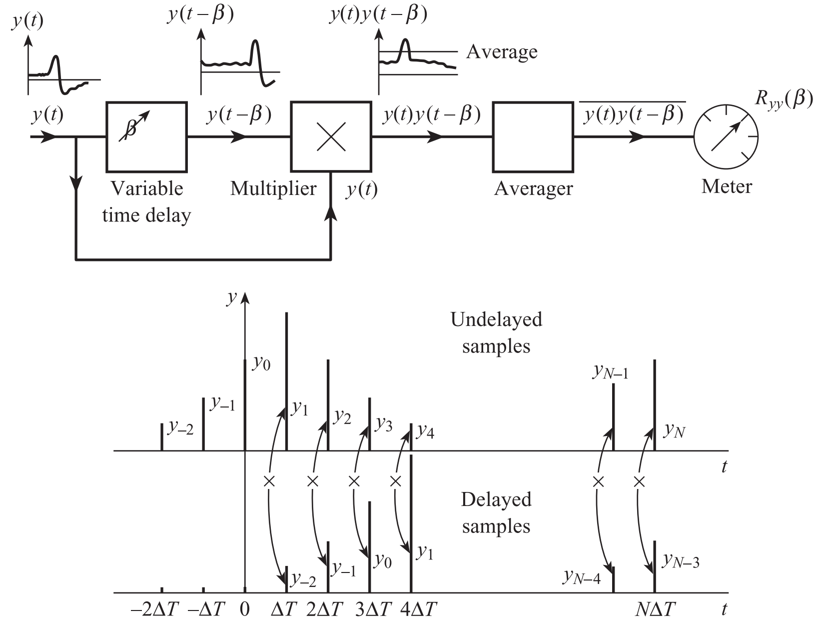


Figure 6.5 Autocorrelation and evaluation of autocorrelation coefficient $R_{yy}(3\Delta T)$.

$R_{yy}(0)$ when $\beta = 0$; this is because the corresponding product waveform is $y^2(t)$, which is always positive and has a maximum average value.

If the signal is defined by a continuous function $y(t)$ in the interval 0 to T_O, then $R_{yy}(\beta)$ can be evaluated using:

Autocorrelation function of a continuous signal

$$R_{yy}(\beta) = \lim_{T_O \to \infty} \frac{1}{T_O} \int_0^{T_O} y(t)y(t - \beta)\,\mathrm{d}t \qquad [6.26]$$

Thus, if $y(t) = b \sin(\omega_1 t + \phi)$,

$$R_{yy}(\beta) = b^2 \lim_{T_O \to \infty} \frac{1}{T_O} \int_0^{T_O} \sin(\omega_1 t + \phi)\sin[\omega_1(t - \beta) + \phi]\,\mathrm{d}t$$

$$= b^2 \lim_{T_O \to \infty} \frac{1}{T_O} \int_0^{T_O} \tfrac{1}{2}\{\cos\,\omega_1\beta - \cos[\omega_1(2t - \beta) + 2\phi]\}\,\mathrm{d}t$$

$$= \frac{b^2}{2}\cos\,\omega_1\beta \lim_{T_O \to \infty} \frac{1}{T_O} \int_0^{T_O} \mathrm{d}t - \frac{b^2}{2} \lim_{T_O \to \infty} \frac{1}{T_O} \int_0^{T_O} \cos[\omega_1(2t - \beta) + 2\phi]\,\mathrm{d}t$$

$$= \frac{b^2}{2}\cos\,\omega_1\beta \qquad [6.27]$$

since the average value of $\cos[\omega_1(2t - \beta) + 2\phi]$ is zero. Thus the autocorrelation function of a sinusoidal signal is a cosine function of the same frequency, but the phase information ϕ in the sine wave is lost. The autocorrelation function of any periodic signal has the same period as the signal itself (see Problem 6.2).

A random signal is usually characterised by a set of N sample values y_i. Since information is only available at discrete time intervals, the time delay β is normally an integer multiple of the sampling interval ΔT, i.e.

$$\beta = m\Delta T, \quad m = 0, 1, 2 \ldots \qquad [6.28]$$

In this case the autocorrelation function of the signal is found by evaluating the set of autocorrelation coefficients $R_{yy}(m\Delta T)$, which are given by:

Autocorrelation coefficient for sampled signal

$$R_{yy}(m\Delta T) = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{i=N} y_i y_{i-m} \qquad [6.29]$$

where y_i is the sample value at time $i\Delta T$ and y_{i-m} the value at time $(i-m)\Delta T$ (m sampling intervals earlier). Figure 6.5 shows the evaluation of $R_{yy}(3\Delta T)$ for a sampled waveform; this involves calculating the products $y_1 y_{-2}$, $y_2 y_{-1}$, $y_3 y_0$, $\ldots$ $y_N y_{N-3}$, summing, and dividing by N.

The autocorrelation function of a random signal can also be found from the power spectral density $\phi(\omega)$. To illustrate this we first consider a periodic signal which is the sum of three harmonics:

$$y(t) = b_1 \sin \omega_1 t + b_2 \sin 2\omega_1 t + b_3 \sin 3\omega_1 t$$

Using eqn [6.27] the autocorrelation function is:

$$R_{yy}(\beta) = \frac{b_1^2}{2} \cos \omega_1 \beta + \frac{b_2^2}{2} \cos 2\omega_1 \beta + \frac{b_3^2}{2} \cos 3\omega_1 \beta$$

which has period $2\pi/\omega_1$ (Figure 6.6). Using eqn [6.17], the power spectrum of the signal consists of three lines at frequencies ω_1, $2\omega_1$ and $3\omega_1$ with heights $b_1^2/2$, $b_2^2/2$ and $b_3^2/2$ respectively. Thus the power spectrum can be obtained from the autocorrelation function by Fourier analysis. Similarly the autocorrelation function can be obtained from the power spectrum by adding the harmonics together as in Fourier synthesis. For random signals, $R_{yy}(\beta)$ and $\phi(\omega)$ are related by the Fourier transform or Wiener–Khinchin relations:

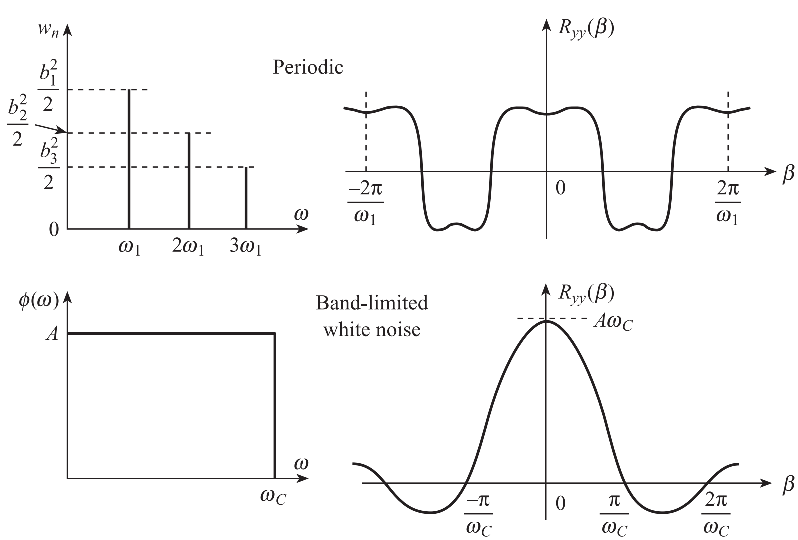


Figure 6.6 Relationships between power spectrum and autocorrelation function for periodic and random signals.

$$R_{yy}(\beta) = \int_0^\infty \phi(\omega) \cos \omega\beta \, d\omega$$
$$\phi(\omega) = \frac{2}{\pi}\int_0^\infty R_{yy}(\beta) \cos \omega\beta \, d\beta \qquad [6.30]$$

Thus for a signal with $\phi(\omega)$ constant up to ω_C and zero for higher frequencies we have:

$$R_{yy}(\beta) = \int_0^{\omega_C} A \cos \omega\beta \, d\omega = A\left[\frac{\sin \omega\beta}{\beta}\right]_0^{\omega_C} = A\frac{\sin \omega_C\beta}{\beta} \qquad [6.31]$$

The form of both functions is shown in Figure 6.6; we see that $R_{yy}(\beta)$ has its first zero crossings at $\beta = \pm\pi/\omega_C$, i.e. the width of the central 'spike' is $2\pi/\omega_C$. Thus a rapidly varying random signal has a high value of ω_C, i.e. a broad power spectrum but a narrow autocorrelation function, that falls off sharply as β is increased. A slowly varying random signal, however, has a low value of ω_C, i.e. a narrow power spectrum but a broad autocorrelation function that falls slowly as β is increased.

6.2.6 Summary

In order to specify a random signal we need to know:

either **probability density function** / or **mean and standard deviation** — to specify amplitude behaviour

and

either **power spectral density** / or **autocorrelation function** — to specify frequency/time behaviour

Important relations between these different quantities can be derived by considering $R_{yy}(0)$, the autocorrelation coefficient at zero time delay. From eqns [6.26] and [6.30] we have

$$R_{yy}(0) = \lim_{T_O\to\infty} \frac{1}{T_O}\int_0^{T_O} y^2 \, dt = \int_0^\infty \phi(\omega) \, d\omega \qquad [6.32]$$

From eqn [6.5], the first expression is equal to y^2_{RMS} in the limit of infinitely long observation time T_O. From equation [6.23] the second expression is equal to W_{TOT}, the total power produced by the signal in a 1 Ω resistor. Thus:

$$R_{yy}(0) = y^2_{\text{RMS}} = W_{\text{TOT}} \qquad [6.33]$$

6.3 Effects of noise and interference on measurement circuits

In Section 5.1 we saw that the interconnection of two measurement system elements, e.g. a thermocouple and an amplifier, or a differential pressure transmitter and a recorder, could be represented by an equivalent circuit in which either a Thévenin voltage

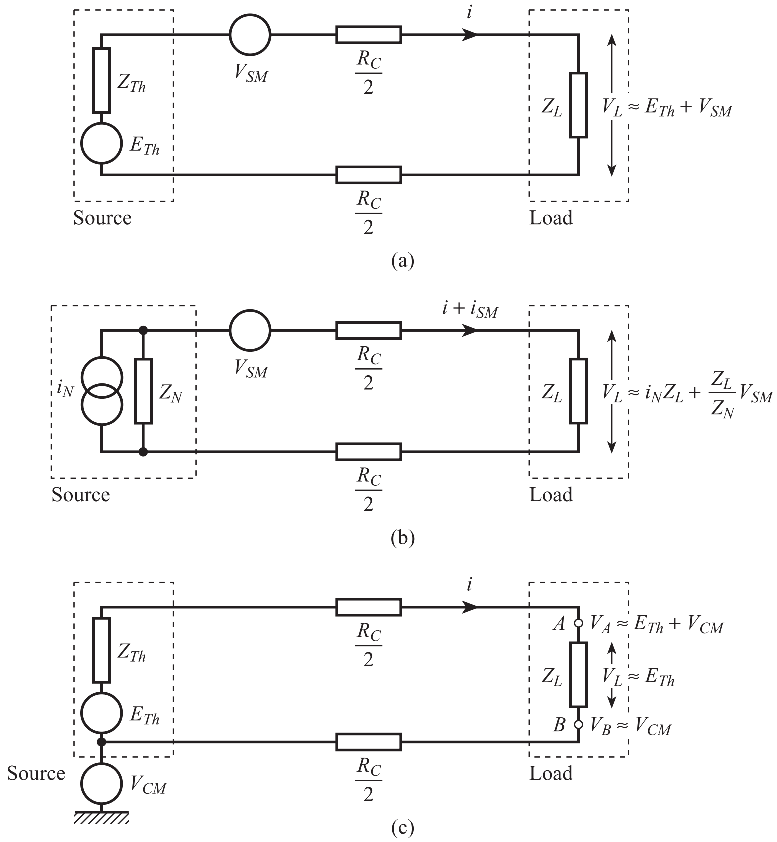


Figure 6.7 Effects of interference on measurement circuit:
(a) Voltage transmission – series mode interference
(b) Current transmission – series mode interference
(c) Voltage transmission – common mode interference.

source or a Norton current source is connected to a load. In industrial installations, source and load may be typically 100 metres apart and noise and/or interference voltages may also be present.

Figure 6.7(a) shows a **voltage transmission system** subject to **series mode interference**; here a noise or interference voltage V_{SM} is in series with the measurement signal voltage E_{Th}. The current i through the load is:

$$i = \frac{E_{Th} + V_{SM}}{Z_{Th} + R_C + Z_L}$$

and the corresponding voltage across the load is:

$$V_L = \frac{Z_L}{Z_{Th} + R_C + Z_L}(E_{Th} + V_{SM}) \qquad [6.34]$$

Normally we make $Z_L \gg R_C + Z_{Th}$ to obtain maximum voltage transfer to the load (Section 5.1.1); under these conditions eqn [6.34] becomes:

$$V_L \approx E_{Th} + V_{SM} \qquad [6.35]$$

This means that with a voltage transmission system all of V_{SM} is across the load; this affects the next element in the system and possibly results in a system measurement error. We define **signal-to-noise** or **signal to interference ratio** S/N in decibels by:

Signal-to-noise ratio

$$\frac{S}{N} = 20\log_{10}\left(\frac{E_{Th}}{V_{SM}}\right) = 10\log_{10}\left(\frac{W_S}{W_N}\right)\text{dB} \qquad [6.36]$$

where E_{Th} and V_{SM} are the r.m.s. values of the voltages, and W_S and W_N are the corresponding total signal and noise powers. Thus if E_{Th} = 1 V and V_{SM} = 0.1 V, $S/N = +20$ dB.

Figure 6.7(b) shows a **current transmission system** subject to the same series mode interference voltage V_{SM}. The Norton source current i_N divides into two parts, one part through the source impedance Z_N, the other part through Z_L. Using the current divider rule, the current through the load due to the source is:

$$i = i_N \frac{Z_N}{Z_N + R_C + Z_L}$$

In addition there is an interference current

$$i_{SM} = \frac{V_{SM}}{Z_N + R_C + Z_L}$$

through the load due to the interference voltage. The total voltage across the load is therefore:

$$\begin{aligned} V_L &= iZ_L + i_{SM}Z_L \\ &= i_N Z_L \cdot \frac{Z_N}{Z_N + R_C + Z_L} + V_{SM} \cdot \frac{Z_L}{Z_N + R_C + Z_L} \end{aligned} \qquad [6.37]$$

Normally we make $R_C + Z_L \ll Z_N$ to obtain maximum current transfer to the load (Section 5.1.3); under these conditions eqn [6.37] becomes:

$$V_L \approx i_N Z_L + \frac{Z_L}{Z_N} V_{SM} \qquad [6.38]$$

Since $Z_L/Z_N \ll 1$, this means that with a current transmission system only a small fraction of V_{SM} is across the load. Thus a current transmission system has far greater inherent immunity to series mode interference than a voltage transmission system. In a thermocouple temperature measurement system, therefore, it may be better to convert the thermocouple millivolt e.m.f. into a current signal (Section 9.4.1) prior to transmission, rather than transmit the e.m.f. directly.

Figure 6.7(c) shows a voltage transmission system subject to **common mode interference** in which the potentials of both sides of the signal circuit are raised by V_{CM} relative to a common earth plane. If, as above, $Z_L \gg R_C + Z_{Th}$, then current $i \to 0$ so that the potential drops $iR_C/2$, etc., can be neglected. Under these conditions:

Potential at $B = V_{CM}$
Potential at $A = V_{CM} + E_{Th}$

and

$$V_L = V_B - V_A = E_{Th}$$

This means that the voltage across the load is unaffected by V_{CM}; there is, however, the possibility of conversion of a common mode voltage to series mode.

6.4 Noise sources and coupling mechanisms

6.4.1 Internal noise sources

The random, temperature-induced motion of electrons and other charge carriers in resistors and semiconductors gives rise to a corresponding random voltage which is called thermal or Johnson noise. This has a power spectral density which is uniform over an infinite range of frequencies (white noise) but proportional to the absolute temperature θ K of the conductor, i.e.

$$\phi = 4Rk\theta \text{ W/Hz} \qquad [6.39]$$

where $R\Omega$ is the resistance of the conductor and k is the Boltzmann constant = 1.4×10^{-23} J K^{-1}. From eqn [6.22] the total thermal noise power between frequencies f_1 and f_2 is:

$$W = \int_{f_1}^{f_2} 4Rk\theta \, \mathrm{d}f = 4Rk\theta(f_2 - f_1) \text{ W} \qquad [6.40]$$

and from [6.33] the corresponding r.m.s. voltage is:

$$V_{RMS} = \sqrt{W} = \sqrt{4Rk\theta(f_2 - f_1)} \text{ V} \qquad [6.41]$$

Thus if $R = 10^6$ Ω, $f_2 - f_1 = 10^6$ Hz and $\theta = 300$ K, $V_{RMS} = 130$ μV and is therefore comparable with low-level measurement signals such as the output from a strain gauge bridge.

A similar type of noise is called **shot noise**; this occurs in transistors and is due to random fluctuations in the rate at which carriers diffuse across a junction. This is again characterised by a uniform power spectral density over a wide range of frequencies.

6.4.2 External noise and interference sources[1]

The most common sources of external interference are nearby **a.c. power circuits** which usually operate at 240 V, 50 Hz. These can produce corresponding sinusoidal interference signals in the measurement circuit, referred to as **mains pick-up** or **hum**. Power distribution lines and heavy rotating machines such as turbines and generators can cause serious interference.

D.C. power circuits are less likely to cause interference because d.c. voltages are not coupled capacitively and inductively to the measurement circuit.

However, **switching** often occurs in both a.c. and d.c. power circuits when equipment such as motors and turbines is being taken off line or brought back on line. This causes sudden large changes in power, i.e. steps and pulses, which can produce corresponding **transients** in the measurement circuit.

The air in the vicinity of high voltage power circuits can become ionised and a **corona discharge** results. Corona discharge from d.c. circuits can result in random noise in the measurement circuit and that from a.c. circuits results in sinusoidal interference at the power frequency or its second harmonic.

Fluorescent lighting is another common interference source; arcing occurs twice per cycle so that most of the interference is at twice the power frequency.

Radio-frequency transmitters, welding equipment and electric arc furnaces can produce **r.f. interference** at frequencies of several MHz.

6.4.3 Coupling mechanisms to external sources

Inductive coupling[2]

Figure 6.8(a) shows inductive or **electromagnetic coupling** between the measurement circuit and a nearby power circuit. If the circuits are sufficiently close together, then there may be a significant mutual inductance M between them. This means that an alternating current i in the power circuit induces a series mode interference voltage in the measurement circuit of magnitude

$$V_{SM} = M\frac{di}{dt} \qquad [6.42]$$

Thus if $M \approx 1\ \mu H$ and $di/dt \approx 10^3\ A\ s^{-1}$ (possible in a 1 horsepower motor) then $V_{SM} \approx 1$ mV, which can be comparable with the measurement signal. The mutual inductance M depends on the geometry of the two circuits, namely on the overlapping length and separation, but is distributed over the entire length of the circuits rather than the 'lumped' equivalent value shown in Figure 6.8. Inductive coupling will occur even if the measurement circuit is completely isolated from earth.

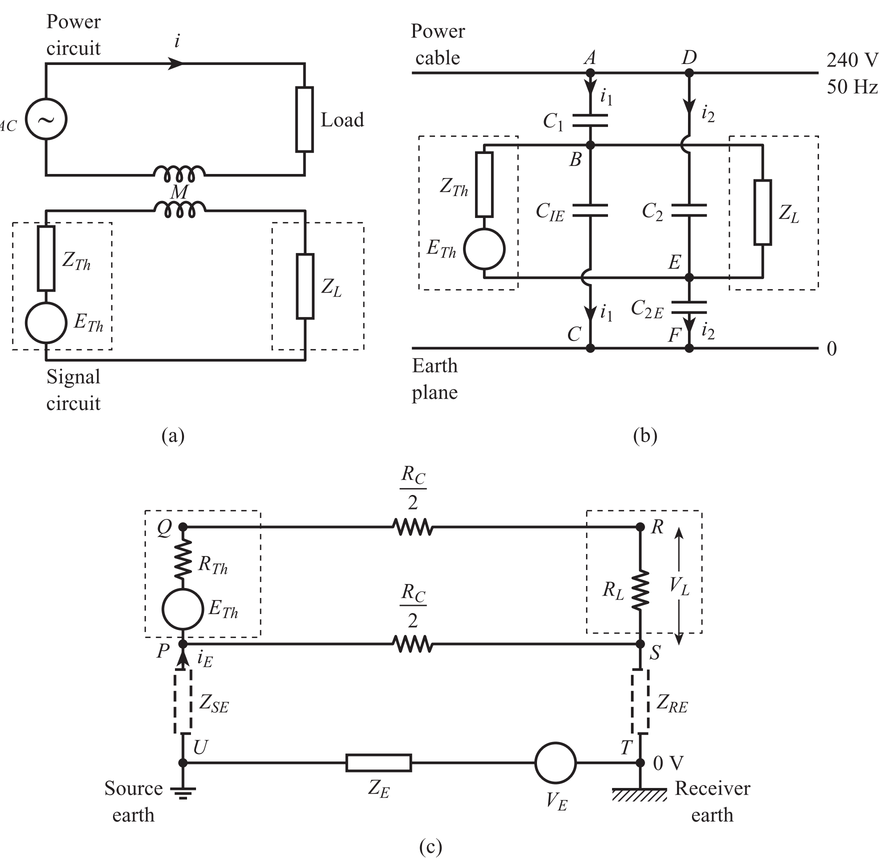


Figure 6.8 Coupling mechanisms to external sources:
(a) Electromagnetic coupling
(b) Electrostatic coupling
(c) Multiple earths.

Capacitive coupling[1,2]

Another important coupling mechanism is **capacitive** or **electrostatic coupling**, which is illustrated in Figure 6.8(b). The diagram shows the measurement circuit close to an a.c. power cable which is at a potential of 240 V (r.m.s.) relative to the earth plane. The power cable, earth plane and signal leads are all conductors, so that there may be some capacitance between the power cable and the signal leads and between the signal leads and the earth plane. These capacitances will be distributed over the entire length of the measurement circuit, but are represented by 'lumped' equivalents. C_1 and C_2 are the capacitances between the power cable and signal leads, and C_{1E} and C_{2E} the capacitances between the signal lead and the earth plane; all four capacitances will be proportional to the length of the measurement circuit, which could be tens of metres in an industrial installation. Ignoring the measurement signal voltage E_{Th} for the moment, the potentials at B and E are determined by the potential dividers ABC and DEF:

$$V_B = 240\left[\frac{1/(j\omega C_{1E})}{1/(j\omega C_{1E}) + 1/(j\omega C_1)}\right] = 240\frac{C_1}{C_1 + C_{1E}}$$

$$V_E = 240\left[\frac{1/(j\omega C_{2E})}{1/(j\omega C_{2E}) + 1/(j\omega C_2)}\right] = 240\frac{C_2}{C_2 + C_{2E}} \qquad [6.43]$$

Thus we have a common mode interference voltage $V_{CM} = V_E$ and a series mode interference voltage:

$$V_{SM} = V_B - V_E = 240\left(\frac{C_1}{C_1 + C_{1E}} - \frac{C_2}{C_2 + C_{2E}}\right) \qquad [6.44]$$

Thus series mode interference is zero only if there is perfect balance between the coupling capacitances, i.e. $C_1 = C_2$ and $C_{1E} = C_{2E}$; in practice small imbalances are present due to slightly different distances between each signal lead and the power cable/earth plane.

Multiple earths[1]

The above explanation assumes an earth plane having a potential of 0 volts at every point on its surface. Heavy electrical equipment can, however, cause currents to flow through the earth, causing different potentials at different points. If the measurement circuit is completely isolated from the earth plane there is no problem. In practice, however, there may be a leakage path connecting the signal source to one earth point and another leakage path connecting the recorder or indicator to a different earth point, some distance away. If the two earth points are at different potentials, then common and series mode interference voltages are produced in the measurement circuit.

Figure 6.8(c) illustrates the general problem of **multiple earths**. The measurement signal source E_{Th} is connected via a resistive cable to a receiver represented by a resistive load R_L. Provided $R_L \gg R_C + R_{Th}$, the current flow in $PQRS$ is negligible and $V_L \approx E_{Th}$, provided the circuit is completely isolated from earth. However, leakage paths Z_{SE} and Z_{RE} exist between source/source earth and receiver/receiver earth. If

V_E is the difference in potential between source earth and receiver earth, then a current i_E flows in the circuit *UPST*, given by:

$$i_E = \frac{V_E}{Z_E + Z_{SE} + (R_C/2) + Z_{RE}} \qquad [6.45]$$

Thus

Potential at $P = V_E - i_E(Z_E + Z_{SE})$
Potential at $Q, R = V_E - i_E(Z_E + Z_{SE}) + E_{Th}$
Potential at $S = i_E Z_{RE}$

Thus there is a common mode interference voltage:

$$V_{CM} = V_S = V_E \frac{Z_{RE}}{Z_E + Z_{SE} + (R_C/2) + Z_{RE}} \qquad [6.46]$$

To find the series mode interference voltage we need to calculate the voltage across R_L:

$$V_L = V_R - V_S = V_E - i_E(Z_E + Z_{SE} + Z_{RE}) + E_{Th} = E_{Th} + i_E R_C/2$$

i.e. there is a series mode interference voltage:

$$V_{SM} = V_E \frac{R_C/2}{Z_E + Z_{SE} + (R_C/2) + Z_{RE}} \qquad [6.47]$$

Ideally we require both Z_{SE} and Z_{RE} to be as large as possible in order to minimise i_E and V_{SM}; this, however, is not always possible in an industrial application. A common example is a thermocouple installation where in order to achieve as good a speed of response as possible the tip of the thermocouple touches the thermowell or sheath (Section 14.2). The thermowell is itself bolted to a metal vessel or pipe which is in turn connected to one point in the earth plane. Thus Z_{SE} will be very small, say $Z_{SE} = 10\ \Omega$ (resistive), so that the receiver must be isolated from earth to minimise V_{SM}. Taking $Z_E = 1\ \Omega$, $R_C/2 = 10\ \Omega$, $V_E = 1$ V and $Z_{RE} = 10^6\ \Omega$, we have:

$$V_{SM} = 1 \times \frac{10}{1 + 10 + 10 + 10^6}\ \text{V} \approx 10\ \mu\text{V}$$

The worst case is when the receiver is also directly connected to earth, i.e. $Z_{RE} \approx 0$, giving $V_{SM} = 0.48$ V. Thus if the measurement circuit must be connected to earth, the connection must be made at one point only.

6.5 Methods of reducing effects of noise and interference

6.5.1 Physical separation

Since mutual inductances and coupling capacitances between measurement and power circuits are inversely proportional to the distance between them, this distance should be as large as possible.

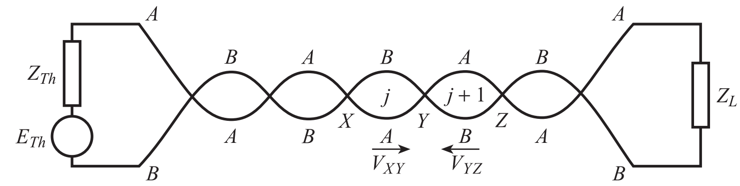


Figure 6.9 Reduction of electromagnetic coupling by twisted pairs.

6.5.2 Electromagnetic shielding

The simplest way of reducing the effects of inductive coupling to an external interference source is shown in Figure 6.9. The two conductors A and B of the measurement circuit are twisted into loops of approximately equal area. This arrangement is commonly known as **twisted pairs** and is explained in ref. [2]. The magnitude of the interference voltage induced in a given loop is proportional to the area of the loop and the rate of change of the external magnetic field. The sign of the induced voltage depends on the orientation of conductors A and B. Thus if a voltage V_{XY} is induced in the jth loop between points X and Y, then an opposing voltage V_{YZ} is induced in the $(j+1)$th loop between Y and Z. In the ideal case of both loops having the same area and experiencing the same magnetic fields, $|V_{XY}| = |V_{YZ}|$, i.e. there is a zero resultant induced voltage between X and Z. This process is repeated for the whole length of the twisted pair, giving a reduced overall interference voltage.

6.5.3 Electrostatic screening and shielding

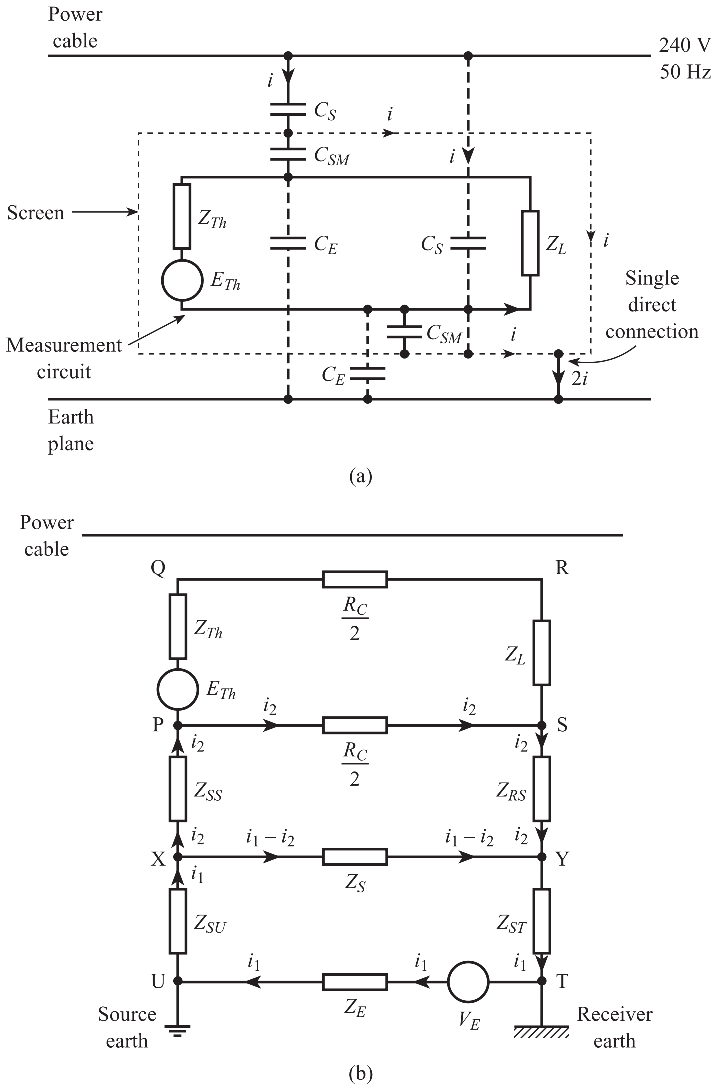
The best method of avoiding the problem of capacitive coupling to a power circuit (Section 6.4.3) is to enclose the entire measurement circuit in an earthed metal screen or shield. Figure 6.10(a) shows the ideal arrangement; the screen is connected directly to earth at a single point, at either the source or the receiver. There is no direct connection between the screen and the measurement circuit, only high impedance leakage paths via the small (screen/measurement circuit) capacitances C_{SM}. The screen provides a low impedance path to earth for the interfering currents i; the currents through C_{SM} and C_E are small, thus reducing series and common mode interference.

The above ideal of the measurement circuit completely insulated from the screen and the screen earthed at one point only may be difficult to achieve in practice for the following reasons:

(a) The signal source may be directly connected to a local earth point via the structure on which it is mounted; an example is the thermocouple installation mentioned in Section 6.4.3.
(b) The receiver may be directly connected to a local earth; an example is in a computer-based system where the receiver must be directly connected to the computer earth.
(c) There may be indirect connections via leakage impedances.

Figure 6.10(b) illustrates the general problem. The measurement circuit PQRS is connected to the screen (impedance Z_S) via source/screen impedance Z_{SS} and receiver/screen impedance Z_{RS}. The screen is connected to earth point U at the source end via Z_{SU} and to earth point T at the receiver end via Z_{ST}. The measurement circuit can be affected by interference voltages from both V_E (potential difference between U and T) and nearby power circuits.

Figure 6.10 Reduction of electrostatic coupling using screening:
(a) Ideal arrangement
(b) Practical equivalent circuit.



Analysis of circuits UXYT and XPSY using Kirchhoff's laws gives:

$$\text{UXYT} \quad V_E = i_1 Z_E + i_1 Z_{SU} + (i_1 - i_2)Z_S + i_1 Z_{ST} \qquad [6.48]$$

$$\text{XPSY} \quad 0 = -(i_1 - i_2)Z_S + i_2 Z_{SS} + i_2 R_C/2 + i_2 Z_{RS} \qquad [6.49]$$

Solution of these equations gives:

$$i_1 = \frac{(Z_S + Z_{SS} + R_C/2 + Z_{RS})V_E}{(Z_E + Z_{SU} + Z_S + Z_{ST})(Z_S + Z_{SS} + R_C/2 + Z_{RS}) - Z_S^2} \qquad [6.50]$$

$$i_2 = \frac{Z_S V_E}{(Z_E + Z_{SU} + Z_S + Z_{ST})(Z_S + Z_{SS} + R_C/2 + Z_{RS}) - Z_S^2} \qquad [6.51]$$

The series mode interference voltage in the measurement circuit PQRS is the voltage drop across PS, i.e.

$$V_{SM} = i_2 R_C/2 = \frac{Z_S(R_C/2)V_E}{(Z_E + Z_{SU} + Z_S + Z_{ST})(Z_S + Z_{SS} + R_C/2 + Z_{RS}) - Z_S^2} \quad [6.52]$$

The common mode interference voltage is the voltage drop across ST, i.e.

$$V_{CM} = V_{ST} = V_{SY} + V_{YT} = i_2 Z_{RS} + i_1 Z_{ST} \quad [6.53]$$

$$= \frac{[Z_{RS}Z_S + Z_{ST}(Z_S + Z_{SS} + R_C/2 + Z_{RS})]V_E}{(Z_E + Z_{SU} + Z_S + Z_{ST})(Z_S + Z_{SS} + R_C/2 + Z_{RS}) - Z_S^2}$$

To minimise V_{SM} and V_{CM}, we want the product term:

$$(Z_E + Z_{SU} + Z_S + Z_{ST})(Z_S + Z_{SS} + R_C/2 + Z_{RS})$$

to be large. Since in practice the earth impedance Z_E, screen impedance Z_S and cable resistance $R_C/2$ are all small, the above condition reduces to:

$$(Z_{SU} + Z_{ST})\ (Z_{SS} + Z_{RS}) \text{ to be large} \quad [6.54]$$

However, we cannot have *both* Z_{SU} *and* Z_{ST} large; either Z_{ST} or Z_{SU} must be small, otherwise the screen will not be earthed and there is therefore no low impedance path to earth for the capacitively coupled interference currents. Condition (6.54) therefore reduces to the two conditions:

$Z_{SU}(Z_{SS} + Z_{RS}) = \text{HIGH};\ Z_{ST} = \text{LOW}$

OR [6.55]

$Z_{ST}(Z_{SS} + Z_{RS}) = \text{HIGH};\ Z_{SU} = \text{LOW}$

which are satisfied if:

$Z_{SU} = \text{HIGH AND } (Z_{SS} \text{ OR } Z_{RS} \text{ OR both} = \text{HIGH}),\ Z_{ST} = \text{LOW}$

OR [6.56]

$Z_{ST} = \text{HIGH AND } (Z_{SS} \text{ OR } Z_{RS} \text{ OR both} = \text{HIGH}),\ Z_{SU} = \text{LOW}$

As mentioned above, in many practical situations it may be impossible to have the measurement circuit completely isolated from the screen, i.e. *both* Z_{SS} *and* Z_{RS} high. In this situation, possible confusion is avoided if Z_{SU} *and* Z_{SS} are *both* high, *or* Z_{ST} *and* Z_{RS} are *both* high.

In this situation the conditions become:

$Z_{SU} = \text{HIGH AND } Z_{SS} = \text{HIGH},\ Z_{RS} = \text{LOW AND } Z_{ST} = \text{LOW}$
(Isolated source)

OR [6.57]

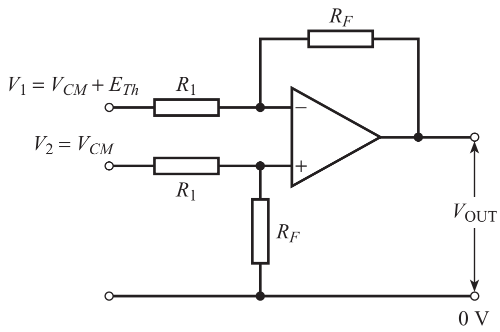
$Z_{ST} = \text{HIGH AND } Z_{RS} = \text{HIGH},\ Z_{SS} = \text{LOW AND } Z_{SU} = \text{LOW}$
(Isolated receiver)

6.5.4 Use of differential amplifiers

Common mode interference voltages can be successfully rejected by the use of a differential amplifier (Figure 6.11 and Section 9.2). An ideal differential amplifier has an output:

$$V_{OUT} = \frac{R_F}{R_1}(V_2 - V_1) = -\frac{R_F}{R_1}E_{Th} \quad [6.58]$$

Figure 6.11 Use of differential amplifier.



i.e. only the sensor voltage E_{Th} is amplified. The output of a practical amplifier (Section 9.2.2) contains a contribution proportional to V_{CM}; from eqn [9.39] we have:

$$V_{OUT} = -\frac{R_F}{R_1}E_{Th} + \left(1 + \frac{R_F}{R_1}\right)\frac{V_{CM}}{\text{CMRR}} \qquad [6.59]$$

The common mode rejection ratio (CMRR) of the amplifier is the ratio of differential voltage gain to common mode voltage gain and should be as large as possible to minimise this effect. Thus if we have $E_{Th} = 1$ mV, $R_1 = 1$ kΩ, $R_F = 1$ MΩ, $V_{CM} = 1$ V and CMRR = 10^5 (100 dB) then:

$$V_{OUT} \approx -1.0 + 0.01 \text{ V}$$

i.e. the resultant series mode interference is only 1%.

6.5.5 Filtering

A filter is an element which transmits a certain range (or ranges) of frequencies and rejects all other frequencies. An **analogue filter** is an electrical network, consisting usually of resistors, capacitors and operational amplifiers, which conditions continuous signals. A **digital filter** is usually a digital computer programmed to process sampled values of a signal (Chapter 10). Provided that the power spectrum of the measurement signal occupies a *different* frequency range from that of the noise or interference signal, then filtering improves the signal-to-noise ratio.

Figures 6.12(a)–(d) show the use of **low pass**, **high pass**, **band pass** and **band stop** filters in rejecting noise. In all cases the filter transmits the measurement signal but rejects the noise signal, which occupies a different frequency range. The diagrams show the amplitude ratio $|G(j\omega)|$ for each filter and the power spectral densities $\phi(\omega)$ for signal and noise. In order to transmit the measurement signal without distortion the transfer function $G(s)$ of the filter must, ideally, satisfy the conditions of Section 4.4, i.e. that $|G(j\omega)| = 1$ and $\arg G(j\omega) = 0$ for all the frequencies present in the measurement signal spectrum. Analogue filtering can be implemented at the signal conditioning stage; the a.c. amplifier of Figure 9.12 is an example of a band pass filter. Digital filtering can be implemented at the signal processing stage (Chapter 10).

If, however, measurement signal and noise spectra overlap, filtering has limited value. Figure 6.13(a) shows a measurement signal affected by wide band noise; a low-pass filter with bandwidth matched to the signal spectrum removes as much of the noise as possible, but the noise inside the filter bandwidth still remains. Figure 6.13(b)

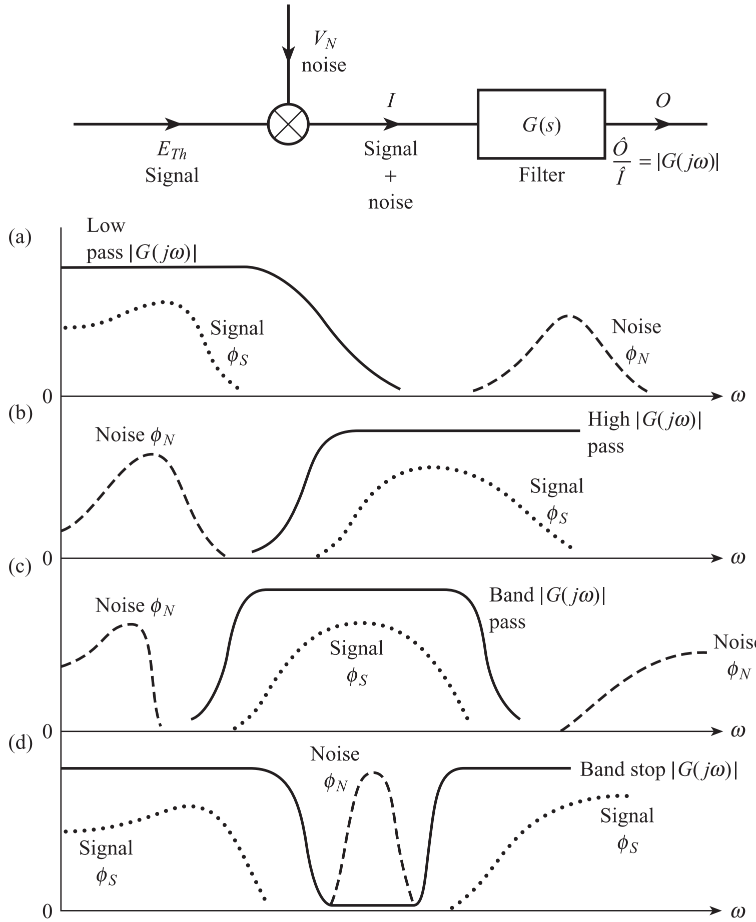


Figure 6.12 Use of filtering to reject noise.

shows the measurement signal affected by 50 Hz interference; a narrow band-stop filter centred on 50 Hz rejects the interference but also rejects measurement signal frequencies around 50 Hz and causes amplitude and phase distortion over a wider range of frequencies.

6.5.6 Modulation

The problem of Figure 6.13(b) can be solved by modulating the measurement signal onto a higher frequency carrier signal, e.g. a 5 kHz sine wave as shown in Figure 6.13(c). The simplest form of modulation is **amplitude modulation**; this involves the multiplication of measurement and carrier signals and is discussed in detail in Section 9.3.

Modulation causes the spectrum of the measurement signal to be shifted to around 5 kHz (Figure 6.13(d)). If the 50 Hz interference is added *after* modulation, i.e. during transmission from sensor/modulator to a remote amplifier/demodulator, the interference spectrum is not shifted. The interference can then be easily rejected by an a.c. amplifier, i.e. a band pass filter with bandwidth matched to the spectrum of the amplitude modulated signal. Modulation, however, does not help the problem

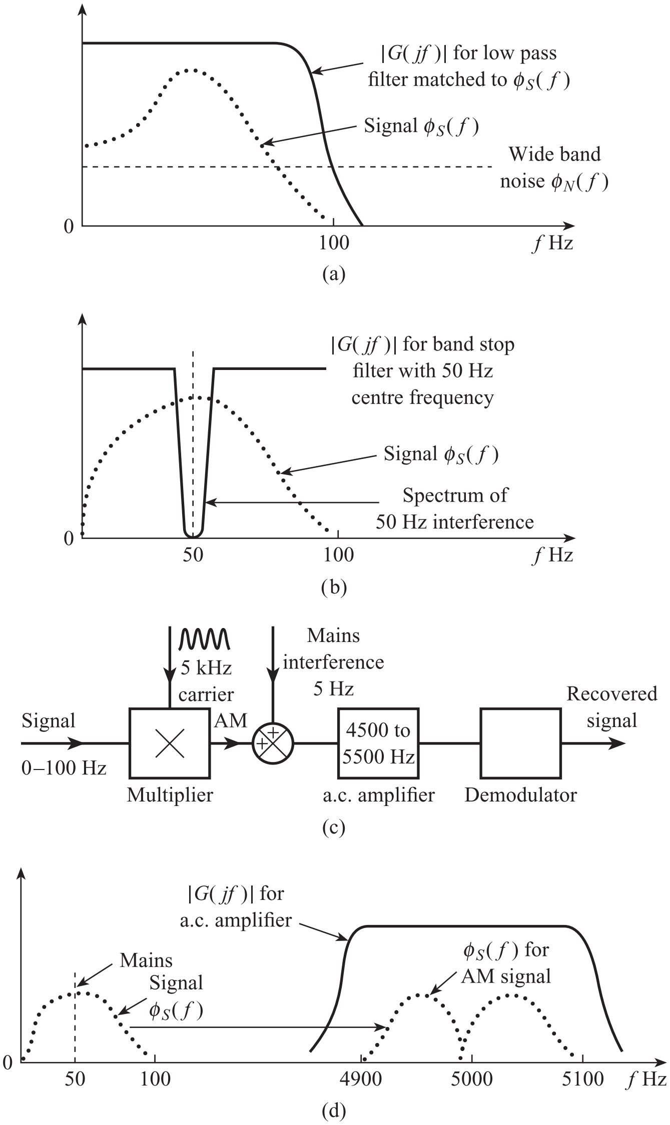


Figure 6.13 Limitations of filtering and use of modulation.

of Figure 6.13(a); the noise has a uniform power spectral density over a wide band of frequencies, so that moving the measurement signal to a different frequency range does not improve the signal-to-noise ratio.

6.5.7 Averaging

Signal averaging can be used to recover a **repetitive** measurement signal affected by random noise, even if the signal r.m.s. value is much less than that of the noise.[3] The process is shown in Figure 6.14.

Suppose that T is the time for each complete cycle of the repetitive signal; p sections of the noise-affected signal, each of duration T, are fed into the averager. N samples from each section are taken and stored, giving pN samples in total; typically

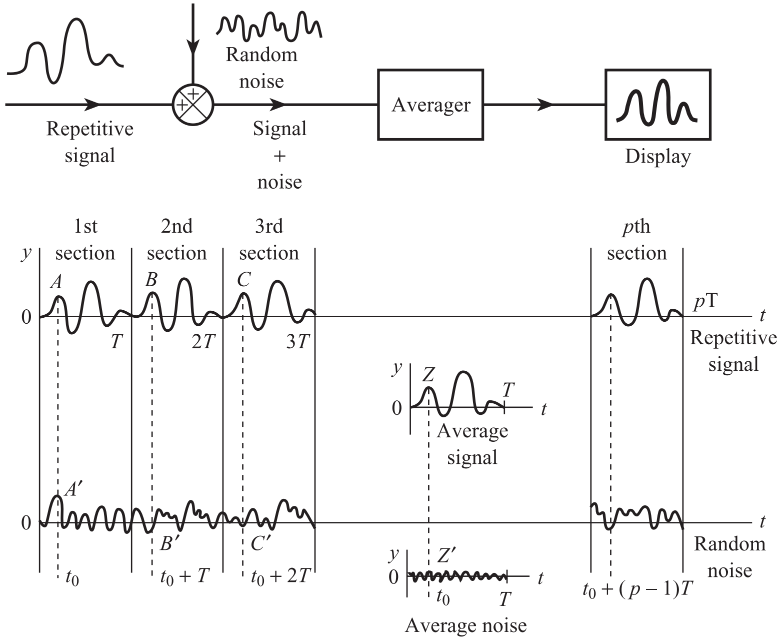


Figure 6.14 Signal averaging.

we may have $p = 50$, $N = 100$. The sampling is exactly synchronised: i.e. if the ith sample of the 1st section is taken at time t_0, the ith sample of the 2nd section is taken at $t_0 + T$, the ith sample of the 3rd at $t_0 + 2T$, and so on. Corresponding samples from each section are then averaged: e.g. the first sampled values from each of the p sections are added together and divided by p. Thus the average value of the ith sample is:

$$y_i^{AV} = \frac{1}{p}(y_{i1} + y_{i2} + \ldots + y_{ip}), \quad i = 1, \ldots, N \tag{6.60}$$

Each of these N average sample values are then displayed at the appropriate time to give the averaged signal. Corresponding sample values A, B, C, . . . of the signal component are approximately equal, so that the average value Z has a similar magnitude. Corresponding sample values A', B', C', . . . of the noise component are very different, some positive and some negative, so that the average value Z' is reduced in magnitude. Averaging therefore maintains the r.m.s. value of the measurement signal while reducing the r.m.s. value of the random noise.

This improvement in signal-to-noise ratio can be readily calculated for random noise with a normal probability density function. Suppose we have p normal signals $y_1(t)$ to $y_p(t)$, with standard deviations σ_1 to σ_p respectively; then from Section 2.3 the average signal

$$y_{AV}(t) = \frac{1}{p}[y_1(t) + y_2(t) + \ldots + y_p(t)] \tag{6.61}$$

is also normal with standard deviation

$$\sigma_{AV} = \sqrt{\frac{\sigma_1^2}{p^2} + \frac{\sigma_2^2}{p^2} + \ldots + \frac{\sigma_p^2}{p^2}} \tag{6.62}$$

If $\sigma_1 = \sigma_2 = \sigma_p = \sigma$, then

Reduction in noise standard deviation due to averaging

$$\sigma_{AV} = \sqrt{\frac{p\sigma^2}{p^2}} = \frac{\sigma}{\sqrt{p}} \qquad [6.63]$$

Thus if we average 50 sections $\sigma_{AV} = \sigma/\sqrt{50} \approx \sigma/7$, i.e. the noise r.m.s. value is reduced by a factor of 7, giving an increase in signal-to-noise ratio of 17dB.

6.5.8 Autocorrelation

Autocorrelation can be used to *detect* the presence of a sinusoidal or any periodic signal buried in random noise. The actual waveform of the measurement signal cannot be recovered because phase information is lost in correlation, but we can measure the amplitude and period of the signal from the autocorrelation function (ACF) of the noise-affected signal. The ACF for the (signal + noise) is the sum of the signal ACF and noise ACF, i.e.

$$R_{yy}^{S+N}(\beta) = R_{yy}^{S}(\beta) + R_{yy}^{N}(\beta) \qquad [6.64]$$

Thus using eqns [6.27] and [6.31] the ACF for a sinusoidal signal affected by band-limited white noise is:

$$R_{yy}^{S+N}(\beta) = \frac{b^2}{2}\cos\omega_1\beta + A\frac{\sin\omega_C\beta}{\beta} \qquad [6.65]$$

The form of $R_{yy}^{S+N}(\beta)$ is shown in Figure 6.15; at large values of β the $A(\sin\omega_c\beta)/\beta$ term due to the noise decays to zero, leaving the $(b^2/2)\cos\omega_1\beta$ term due to the signal. Thus the amplitude b and period $2\pi/\omega_1$ of the original signal can be found from the amplitude and period of the autocorrelation function at large values of time delay. This method can be used in the vortex flowmeter (Chapter 12) to measure the vortex frequency in the presence of random flow turbulence.

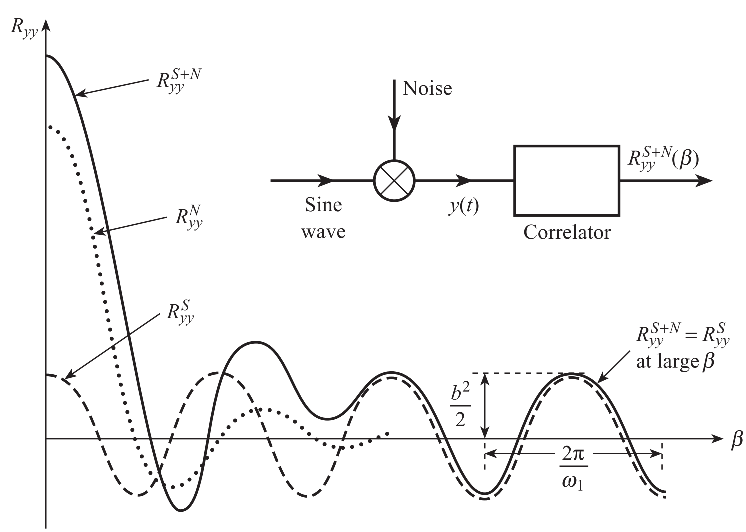


Figure 6.15
Autocorrelation detection of periodic signal buried in noise.

Conclusion

The chapter began by defining **random signals** and **deterministic signals** and explained that in many practical situations the wanted signal may be random. Unwanted signals may also be present in the measurement circuit; these can be classified as either **interference** (deterministic) or **noise** (random).

The chapter then explained how random signals can be quantified using the following statistical functions: **mean**, **standard deviation**, **probability density function**, **power spectral density** and **autocorrelation function**.

The effects of noise and interference voltage on measurement circuits using both voltage and current transmission were then discussed. In the following section internal noise and external interference sources were discussed and the mechanisms whereby external sources are coupled to the measurement circuit were explained. The chapter concluded by explaining methods of reducing the effects of noise and interference, including **electromagnetic shielding**, **electrostatic screening**, **filtering**, **modulation**, **averaging** and **autocorrelation**.

References

[1] OLIVER F J 1972 *Practical Instrumentation Transducers*. Pitman, London, pp. 290–333.

[2] COOK B E 1979 'Electronic noise and instrumentation', *Journal of Instrumentation Measurement and Control*, vol. 12, no. 8.

[3] Hewlett Packard, Technical Literature on Model 3721A Correlator.

Problems

6.1 Table Prob. 1 gives 50 sample values of a random signal.

(a) Estimate the mean and standard deviation of the signal.

(b) Using an interval $\Delta y = 0.5$ V, draw the P_j and C_j discrete probability distributions for the signal.

Table Prob. 1.

–0.59	1.02	–0.25	–0.34	0.95	1.24	–0.30	0.21	–0.89
–1.00	1.36	0.03	0.04	–0.13	–0.71	–1.23	0.03	–1.00
0.65	0.11	0.99	0.17	0.39	2.61	–0.08	–0.33	0.99
2.15	0.91	0.89	1.43	–1.69	–0.25	2.47	–1.97	–2.26
0.42	0.05	0.26	0.33	–0.42	0.79	–0.07	–0.32	–0.66
–0.63	–0.06	–0.61	0.77	1.90				

6.2 Two complete periods of a square wave can be represented by the following 20 sample values:

+1+1+1+1+1 –1–1–1–1–1 +1+1+1+1+1 –1–1–1–1–1

Find the autocorrelation function of the signal over one complete period by evaluating the coefficients $R_{yy}(m\Delta T)$ for $m = 0, 1, 2, \ldots, 10$.

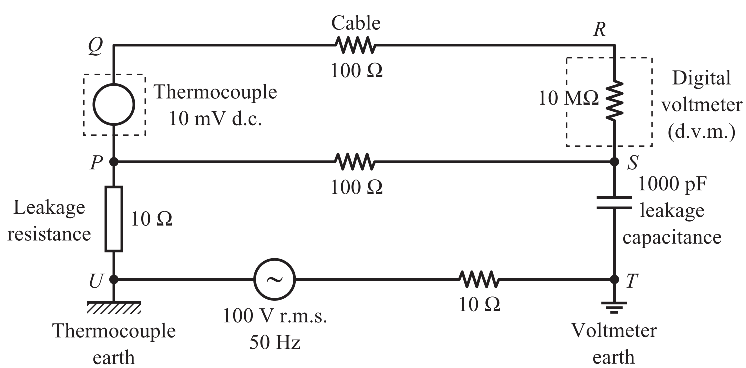
6.3 A sinusoidal signal of amplitude 1.4 mV and frequency 5 kHz is 'buried' in Gaussian noise with zero mean value. The noise has a uniform power spectral density of 100 pW Hz^{-1} up to a cut-off frequency of 1 MHz.

(a) Find the total power, r.m.s. value and standard deviation for the noise signal.
(b) What is the signal-to-noise ratio in dB?
(c) Sketch the autocorrelation function for the combined signal and noise.
(d) The combined signal is passed through a band-pass filter with centre frequency 5 kHz and bandwidth 1 kHz. What improvement in signal-to-noise ratio is obtained?
(e) The filtered signal is then passed through a signal averager which averages corresponding samples of 100 sections of signal. What further improvement in signal-to-noise ratio is obtained?

6.4 A thermocouple giving a 10 mV d.c. output voltage is connected to a high impedance digital voltmeter some distance away. A difference in potential exists between earth at the thermocouple and earth at the voltmeter. Using the equivalent circuit given in Figure Prob. 4.

(a) Calculate the r.m.s. values of the series mode and common mode interference voltages at the voltmeter input.
(b) If the digital voltmeter has a common mode rejection ratio of 100 dB, find the minimum and maximum possible measured voltages.

Figure Prob. 4.

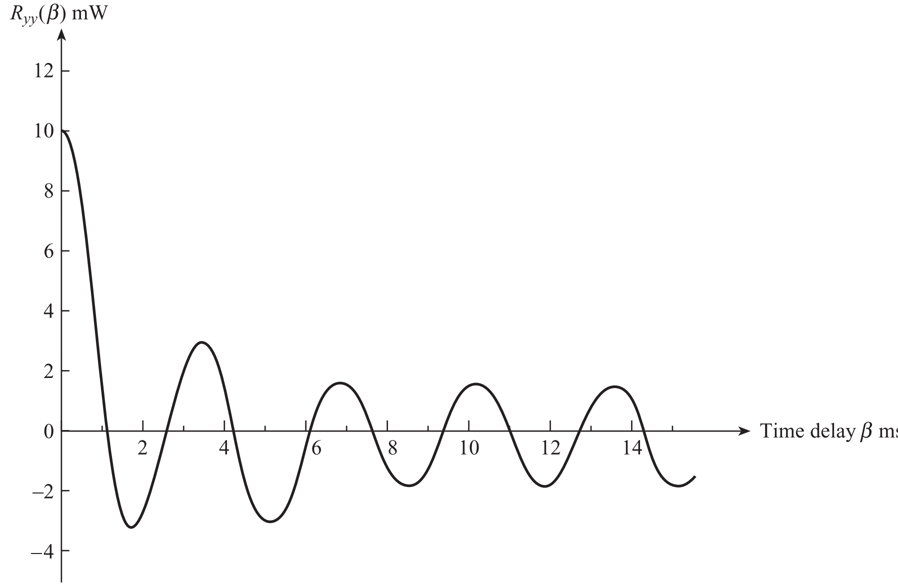


6.5 A sinusoidal signal is transmitted over a noisy transmission link to a remote correlator acting as a receiver. Figure Prob. 5 shows a typical autocorrelation function. Use the figure to estimate the following quantities:

(a) Signal power
(b) Noise power
(c) Signal-to-noise ratio in decibels
(d) Signal amplitude
(e) Signal frequency
(f) Noise standard deviation (assume zero mean).

Hint: use eqns [6.27] and [6.33].

Figure Prob. 5.



7 Reliability, Choice and Economics of Measurement Systems

In Chapters 3 and 4 we defined the accuracy of a measurement system and explained how measurement error can be calculated, under both steady-state and dynamic conditions. **Reliability** is another important characteristic of a measurement system; it is no good having an accurate measurement system which is constantly failing and requiring repair. The first section of this chapter deals with the reliability of measurement systems, first explaining the fundamental principles of reliability and the reliability of practical systems, then failure rate data, and finally examining ways of improving reliability. The following section examines the problems of how to choose the most appropriate measurement system, for a given application, from several competing possibilities. Initially a specification for the required application can be drawn up: this will be a list of important parameters such as accuracy, reliability, etc., each with a desired value. This can then be compared with the manufacturer's specification for each of the competing measurement systems and the system with the closest specification is chosen. Even if all the required information is available this procedure is far from satisfactory because it takes no account of the relative importance of each parameter. A better method, explained in the final section, is to choose the system with minimum total lifetime operating cost.

7.1 Reliability of measurement systems

7.1.1 Fundamental principles of reliability

Probability *P*

If a large number of random, independent trials are made, then the probability of a particular event occurring is given by the ratio:

$$P = \frac{\text{number of occurrences of the event}}{\text{total number of trials}} \quad [7.1]$$

in the limit that the total number of trials tends to infinity. Thus the probability of a tossed coin showing heads tends to the theoretical value of $\frac{1}{2}$ over a large number of trials.

Reliability $R(t)$

The reliability of a measurement element or system can be defined as: 'the probability that the element or system will operate to an agreed level of performance, for a specified period, subject to specified environmental conditions'. In the case of a measurement system 'agreed level of performance' could mean an accuracy of ±1.5%. If the system is giving a measurement error outside these limits, then it is considered to have failed, even though it is otherwise working normally. The importance of environmental conditions on the reliability of measurement systems will be discussed more fully later. Reliability decreases with time; a measurement system that has just been checked and calibrated should have a reliability of 1 when first placed in service. Six months later, the reliability may be only 0.5 as the probability of survival decreases.

Unreliability $F(t)$

This is 'the probability that the element or system will *fail* to operate to an agreed level of performance, for a specified period, subject to specified environmental conditions'. Since the equipment has either failed or not failed the sum of reliability and unreliability must be unity, i.e.

$$R(t) + F(t) = 1 \quad [7.2]$$

Unreliability also depends on time; a system that has just been checked and calibrated should have an unreliability of zero when first placed in service, increasing to, say, 0.5 after six months.

7.1.2 Practical reliability definitions

Since $R(t)$ and $F(t)$ are dependent on time, it is useful to have measures of reliability which are independent of time. We will consider two cases: in the first the items are non-repairable and in the second the items are repairable.

Non-repairable items

Suppose that N individual items of a given non-repairable component are placed in service and the times at which failures occur are recorded during a test interval T. We further assume that all the N items fail during T and that the ith failure occurs at time T_i, i.e. T_i is the survival time or **up time** for the ith failure. The total up time for N failures is therefore $\sum_{i=1}^{i=N} T_i$ and the **mean time to failure** is given by

$$\text{Mean time to fail} = \frac{\text{total up time}}{\text{number of failures}}$$

i.e.

$$\text{MTTF} = \frac{1}{N}\sum_{i=1}^{i=N} T_i \qquad [7.3]$$

The **mean failure rate** $\bar{\lambda}$ is correspondingly given by:

$$\text{Mean failure rate} = \frac{\text{number of failures}}{\text{total up time}}$$

i.e.

$$\bar{\lambda} = \frac{N}{\sum_{i=1}^{i=N} T_i} \qquad [7.4]$$

i.e. mean failure rate is the reciprocal of MTTF.

There are N survivors at time $t = 0$, $N - i$ at time $t = T_i$, decreasing to zero at time $t = T$; Figure 7.1(a) shows how the probability of survival, i.e. reliability, $R_i = (N - i)/N$ decreases from $R_i = 1$ at $t = 0$, to $R_i = 0$ at $t = T$. The ith rectangle has height $1/N$, length T_i and area T_i/N. Therefore from eqn [7.3] we have:

$$\text{MTTF} = \text{total area under the graph}$$

In the limit that $N \to \infty$, the discrete reliability function R_i becomes the continuous function $R(t)$. The area under $R(t)$ is $\int_0^T R(t)\,dt$ so that we have in general:

$$\text{MTTF} = \int_0^\infty R(t)\,dt \qquad [7.5]$$

The upper limit of $t = \infty$ corresponds to N being infinite.

Repairable items

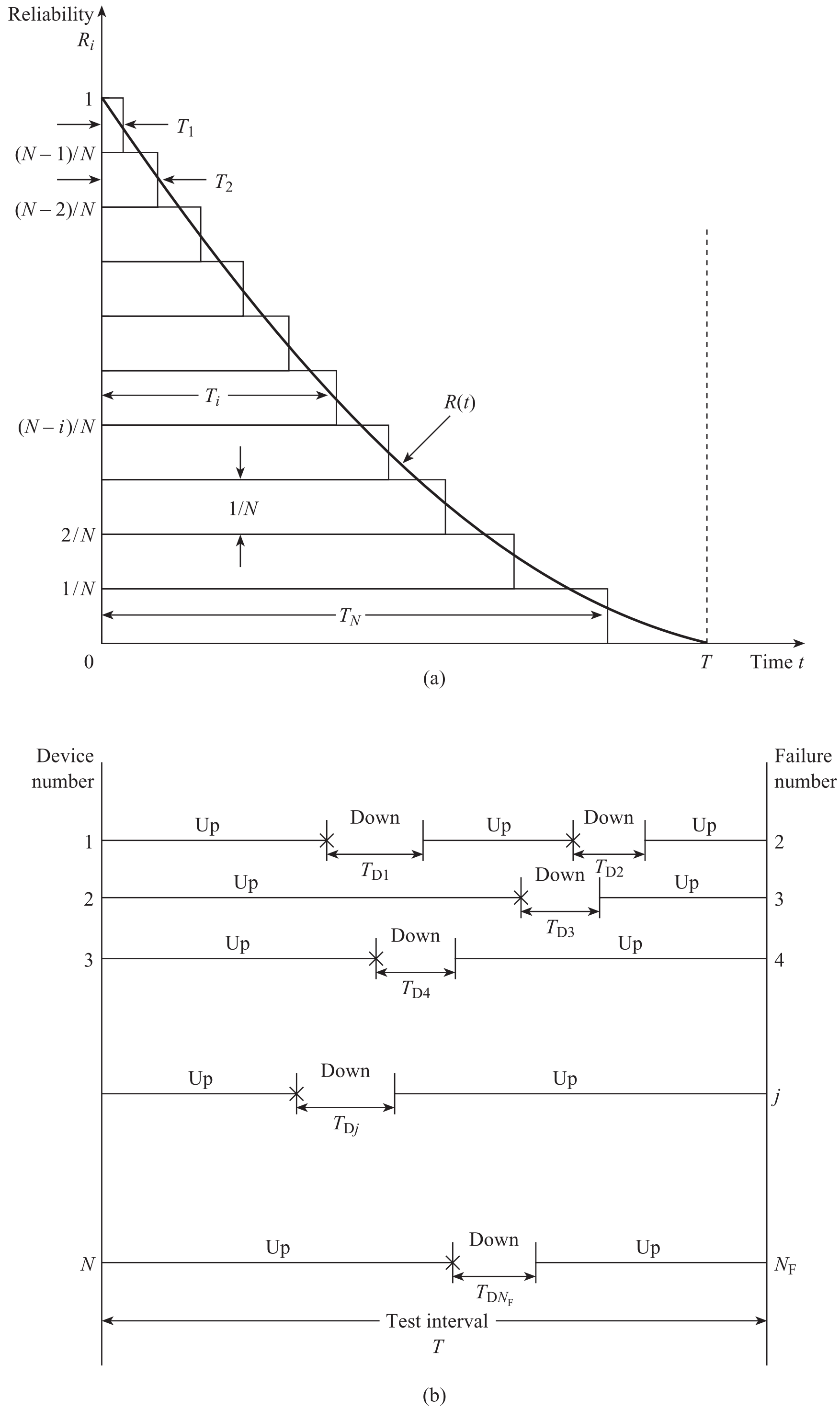
Figure 7.1(b) shows the failure pattern for N items of a repairable element observed over a test interval T. The **down time** T_{Dj} associated with the jth failure is the total time that elapses between the occurrence of the failure and the repaired item being put back into normal operation. The total down time for N_F failures is therefore $\sum_{j=1}^{j=N_F} T_{Dj}$ and the **mean down time** is given by

$$\text{Mean down time} = \frac{\text{total down time}}{\text{number of failures}}$$

i.e.

$$\text{MDT} = \frac{1}{N_F}\sum_{j=1}^{j=N_F} T_{Dj} \qquad [7.6]$$

Figure 7.1
Failure patterns:
(a) Non-repairable items
(b) Repairable items.



The **total up time** can be found by subtracting the total down time from NT, i.e.:

$$\text{Total up time} = NT - \sum_{j=1}^{j=N_F} T_{Dj}$$
$$= NT - N_F\,\text{MDT}$$

The mean up time or the **mean time between failures** (MTBF) is therefore given by:

$$\text{Mean time between failures} = \frac{\text{total up time}}{\text{number of failures}}$$

i.e.

$$\text{MTBF} = \frac{NT - N_F\,\text{MDT}}{N_F} \qquad [7.7]$$

The **mean failure rate** $\bar{\lambda}$ is correspondingly given by:

$$\text{Mean failure rate} = \frac{\text{number of failures}}{\text{total up time}}$$

i.e.

$$\bar{\lambda} = \frac{N_F}{NT - N_F\,\text{MDT}} \qquad [7.8]$$

Again mean failure rate is the reciprocal of MTBF.

Thus if 150 faults are recorded for 200 transducers over 1.5 years with a mean down time of 0.002 years, then the observed MTBF is 1.998 years and the mean failure rate is 0.5005 yr^{-1}.

The **availability** A of the element is the fraction of the total test interval over which it is performing within specification, i.e. up; thus we have:

$$\text{Availability} = \frac{\text{total up time}}{\text{test interval}}$$
$$= \frac{\text{total up time}}{\text{total up time} + \text{total down time}}$$
$$= \frac{N_F \times \text{MTBF}}{N_F \times \text{MTBF} + N_F \times \text{MDT}}$$

i.e.

$$A = \frac{\text{MTBF}}{\text{MTBF} + \text{MDT}} \qquad [7.9]$$

Using the above data of MTBF = 1.998 years and MDT = 0.002 years gives $A = 0.999$.

Unavailability U is similarly defined as the fraction of the total test interval over which it is not performing to specification, i.e. failed or down; thus we have:

$$\text{Unavailability} = \frac{\text{total down time}}{\text{test interval}}$$

giving:

$$U = \frac{\text{MDT}}{\text{MTBF} + \text{MDT}} \qquad [7.10]$$

It follows from eqns [7.9] and [7.10] that:

$$A + U = 1 \qquad [7.11]$$

7.1.3 Instantaneous failure rate and its relation to reliability

We assume to begin with that n items of an element survive up to time $t = \xi$ and that Δn items fail during the small time interval $\Delta\xi$ between ξ and $\xi + \Delta\xi$. The probability of failure during interval $\Delta\xi$ (given survival to time ξ) is therefore equal to $\Delta n/n$. Assuming no repair during $\Delta\xi$ the corresponding **instantaneous failure rate** or **hazard rate** at time ξ is, from eqn [7.8], given by:

$$\lambda(\xi) = \frac{\Delta n}{n\Delta\xi} = \frac{\text{failure probability}}{\Delta\xi} \qquad [7.12]$$

The unconditional probability ΔF that an item fails during the interval $\Delta\xi$ is:

ΔF = probability that item survives up to time ξ

AND

probability that item fails between ξ and $\xi + \Delta\xi$ (given survival to ξ).

The first probability is given by $R(\xi)$ and from eqn [7.12] the second probability is $\lambda(\xi)\Delta\xi$. The combined probability ΔF is the product of these probabilities:

$$\Delta F = R(\xi)\lambda(\xi)\Delta\xi$$

i.e.

$$\frac{\Delta F}{\Delta\xi} = R(\xi)\lambda(\xi)$$

Thus in the limit that $\Delta\xi \to 0$, we have:

$$\frac{dF}{d\xi} = R(\xi)\lambda(\xi) \qquad [7.13]$$

Also since $F(\xi) = 1 - R(\xi)$, $dF/d\xi = -(dR/d\xi)$, giving:

$$-\frac{dR}{d\xi} = R(\xi)\lambda(\xi)$$

i.e.

$$\int_{R(0)}^{R(t)} \frac{dR}{R} = -\int_0^t \lambda(\xi)\, d\xi \qquad [7.14]$$

In eqn [7.14], the left-hand integral is with respect to R and the right-hand integral with respect to ξ. Since at $t = 0$, $R(0) = 1$, we have:

$$[\log_e R]_1^{R(t)} = -\int_0^t \lambda(\xi)\, d\xi$$

$$\log_e R(t) = -\int_0^t \lambda(\xi)\, d\xi$$

i.e.

Relation between reliability and instantaneous failure rate

$$R(t) = \exp\left[-\int_0^t \lambda(\xi)\, d\xi\right] \qquad [7.15]$$

7.1.4 Typical forms of failure rate function

In the previous section instantaneous failure rate or hazard rate $\lambda(t)$ was defined. Figure 7.2 shows the most general form of $\lambda(t)$ throughout the lifetime of an element. This is the so-called **bathtub curve** and consists of three distinct phases: early failure, useful life and wear-out failure. The **early failure region** is characterised by $\lambda(t)$ decreasing with time. When items are new, especially if the element is of a new design, early failures can occur due to design faults, poor quality components, manufacturing faults, installation errors, operator and maintenance errors; the latter may be due to unfamiliarity with the product. The hazard rate falls as design faults are rectified, weak components are removed, and the user becomes familiar with installing, operating and maintaining the element. The **useful life region** is characterised by a low,

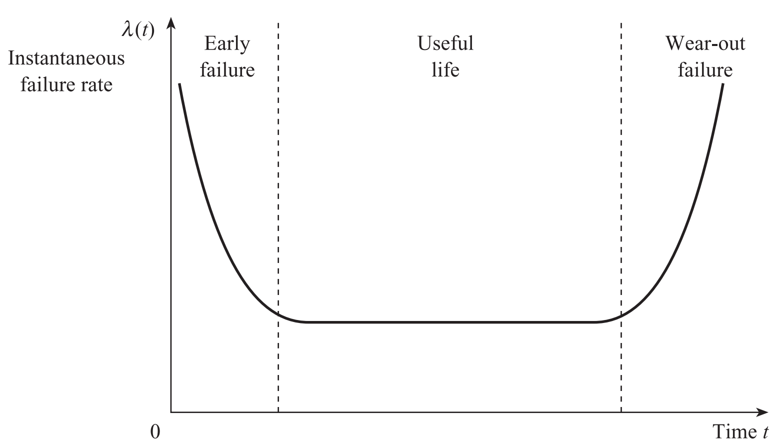


Figure 7.2 Typical variation in instantaneous failure rate (hazard rate) during the lifetime of element – ‘bathtub curve’.

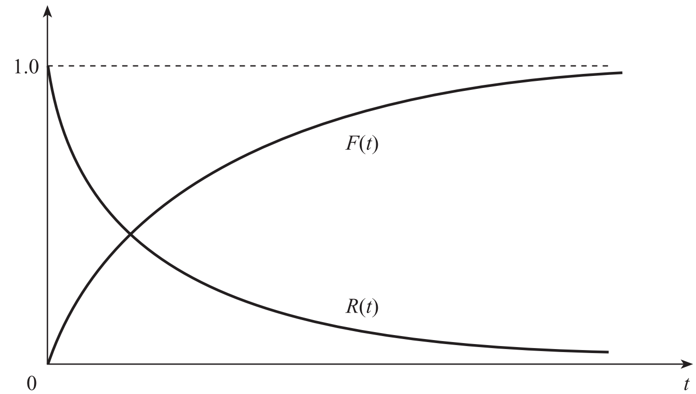


Figure 7.3 Reliability and unreliability with constant failure rate model.

constant failure rate. Here all weak components have been removed: design, manufacture, installation, operating and maintenance errors have been rectified so that failure is due to a variety of unpredictable lower-level causes. The **wear-out region** is characterised by $\lambda(t)$ increasing with time as individual items approach the end of the design life for the product; long-life components which make up the element are now wearing out.

Many measurement elements have a useful life region lasting many years, so that a **constant failure rate** model is often a good approximation. Here we have:

$$\lambda(t) = \lambda(\xi) = \lambda = \text{constant} \tag{7.16}$$

so that:

$$R(t) = \exp\left[-\lambda \int_0^t \xi\right] = \exp(-\lambda t) \tag{7.17}$$

$$F(t) = 1 - \exp(-\lambda t)$$

Thus a constant failure or hazard rate gives rise to an **exponential reliability** time variation or distribution as shown in Figure 7.3.

7.1.5 Reliability of systems

Series systems

We saw in Chapter 1 that a complete measurement system consists of several elements usually in series or cascade. Figure 7.4 shows a series system of m elements with individual reliabilities $R_1, R_2, \ldots, R_i, \ldots, R_m$ respectively. The system will only survive if every element survives; if one element fails then the system fails.

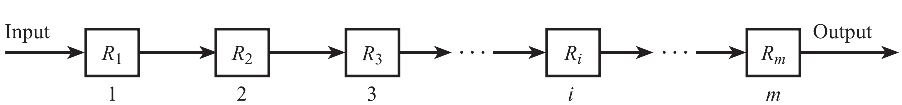


Figure 7.4 Reliability of series system.

Assuming that the reliability of each element is independent of the reliability of the other elements, then the probability that the system survives is the probability that element 1 survives *and* the probability that 2 survives *and* the probability that 3 survives, etc. The system reliability R_{SYST} is therefore the product of the individual element reliabilities, i.e.

Reliability of series system

$$R_{\mathrm{SYST}} = R_1 R_2 \ldots R_i \ldots R_m \qquad [7.18]$$

If we further assume that each of the elements can be described by a constant failure rate λ (Section 7.1.4), and if λ_i is the failure rate of the *i*th element, then R_i is given by the exponential relation (eqn [7.17]):

$$R_i = \mathrm{e}^{-\lambda_i t} \qquad [7.19]$$

Thus

$$R_{\mathrm{SYST}} = \mathrm{e}^{-\lambda_1 t}\, \mathrm{e}^{-\lambda_2 t} \ldots \mathrm{e}^{-\lambda_i t} \ldots \mathrm{e}^{-\lambda_m t} \qquad [7.20]$$

so that if λ_{SYST} is the overall system failure rate:

$$R_{\mathrm{SYST}} = \mathrm{e}^{-\lambda_{\mathrm{SYST}} t} = \mathrm{e}^{-(\lambda_1+\lambda_2+\ldots+\lambda_i+\ldots+\lambda_m)t} \qquad [7.21]$$

and

Failure rate of system of m *elements in series*

$$\lambda_{\mathrm{SYST}} = \lambda_1 + \lambda_2 + \ldots + \lambda_i + \ldots + \lambda_m \qquad [7.22]$$

This means that the overall failure rate for a series system is the sum of the individual element or component failure rates. Equations [7.18] and [7.22] show the importance of keeping the number of elements in a series system to a minimum; if this is done the system failure rate will be minimum and the reliability maximum. A measurement system consisting of a thermocouple ($\lambda_1 = 1.1$), a millivolt to current converter ($\lambda_2 = 0.1$) and a recorder ($\lambda_3 = 0.1$) in series will therefore have a failure rate $\lambda_{\mathrm{SYST}} = 1.3$.

Parallel systems

Figure 7.5 shows an overall system consisting of *n* individual elements or systems in parallel with individual unreliabilities $F_1, F_2, \ldots, F_j, \ldots, F_n$ respectively. Only one individual element or system is necessary to meet the functional requirements placed on the overall system. The remaining elements or systems increase the reliability of the overall system; this is termed **redundancy**. The overall system will only fail if every element/system fails; if one element/system survives the overall system survives. Assuming that the reliability of each element/system is independent of the reliability of the other elements, then the probability that the overall system fails is the probability that element/system 1 fails *and* the probability that 2 fails *and* the probability that 3 fails, etc. The overall system unreliability F_{SYST} is therefore the *product* of the individual element system unreliabilities, i.e.

Unreliability of parallel system

$$F_{\mathrm{SYST}} = F_1 F_2 \ldots F_j \ldots F_n \qquad [7.23]$$

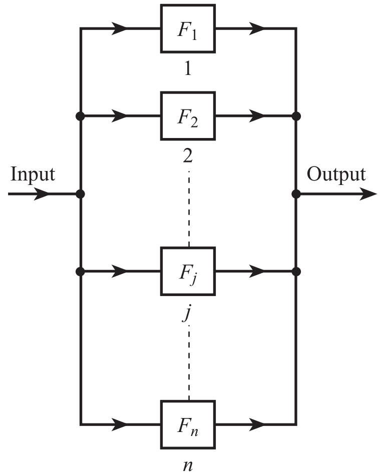


Figure 7.5 Reliability of parallel system.

Comparing eqns [7.18] and [7.23] we see that, for series systems, system reliability is the product of element reliabilities, whereas for parallel systems system unreliability is the product of element unreliabilities. Often the individual elements/systems are identical, so that $F_1 = F_2 = \ldots = F_i = \ldots = F_n = F$, giving:

$$F_{\mathrm{SYST}} = F^n \tag{7.24}$$

The temperature measurement system of the previous section has a failure rate $\lambda_{\mathrm{SYST}} = 1.3\ \mathrm{yr}^{-1}$; the corresponding unreliability F is given by

$$F = 1 - \mathrm{e}^{-\lambda_{\mathrm{SYST}} t}$$

Thus if $t = 0.5$ year then $F = 0.478$. Figure 7.6 shows a redundant system consisting of three single temperature measurement systems in parallel. The overall system unreliability is therefore:

$$F_{\mathrm{OVERALL}} = F^3 = 0.109$$

so that the probability of a failure with the overall system is less than a quarter of that of a single system.

The above parallel system, while reliable, is expensive. Since the thermocouple failure rate is 11 times greater than the converter and recorder failure rates, a more

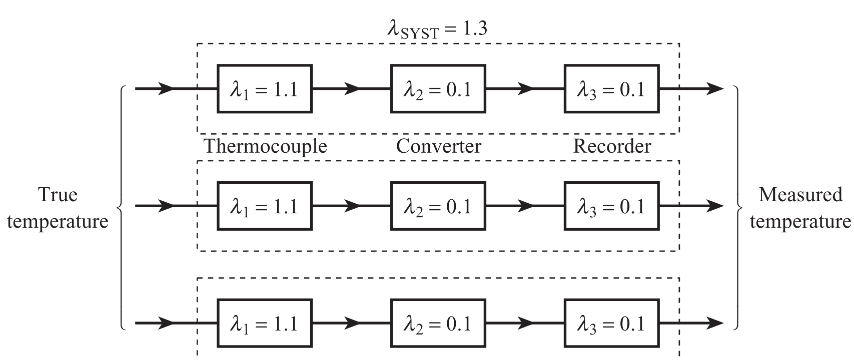


Figure 7.6 Reliability of three thermocouple temperature measurement systems in parallel.

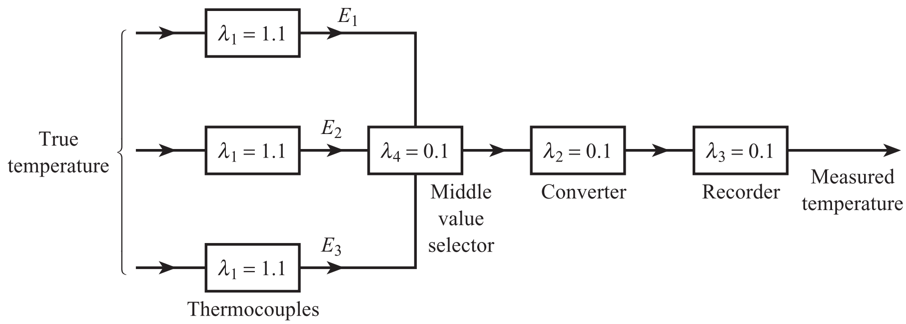


Figure 7.7 Reliability of system with three thermocouples and middle value selector.

cost-effective redundant system would have three thermocouples in parallel and only one converter and recorder. One possible system is shown in Figure 7.7. The three thermocouple e.m.f.'s E_1, E_2 and E_3 are input into a middle value selector element. The selector output signal is that input e.m.f. which is neither the lowest nor the highest; thus if $E_1 = 5.0$ mV, $E_2 = 5.2$ mV and $E_3 = 5.1$ mV, the output signal is E_3. If, however, thermocouple 3 fails so that $E_3 = 0$ mV, the selector output signal is E_1. The reliability of this system can be analysed by replacing the three thermocouples by a single element of unreliability

$$F_1 = (1 - e^{-\lambda_1 t})^3 = (1 - e^{-0.55})^3 = 0.076$$

or reliability $R_1 = 1 - 0.076 = 0.924$. The reliability of the other elements with $\lambda = 0.1$ is:

$$R_2 = R_3 = R_4 = e^{-\lambda t} = e^{-0.05} = 0.951$$

Using [7.18], the overall system reliability is:

$$R_{\text{OVERALL}} = R_1 R_2 R_3 R_4 = 0.924(0.951)^3 = 0.795$$

i.e. $F_{\text{OVERALL}} = 0.205$. This is almost twice the unreliability of the parallel system but less than half that of a single system.

7.1.6 Failure rate data and models

A distinction must now be made between **components** and **elements**. A component is defined as a 'non-repairable' device, i.e. when it fails it is removed and thrown away. Examples are a resistor or an integrated circuit. An element, however, is a repairable part of a system, which is usually made up of several components. Examples are a pressure sensor, a temperature transmitter and a recorder.

Failure rate data for both components and elements can be found experimentally by direct measurements of the frequency of failure of a number of items of a given type. Equation [7.4] can be used to find the observed failure rate of non-repairable components and eqn [7.8] to find $\bar{\lambda}$ for repairable elements.

Table 7.1 gives observed average failure rates for typical instruments. These data have been taken from the UK data bank operated by the Systems Reliability Service (SRS).[1] The table specifies the environment in which each type of instrument is located. For an element located in the process fluid, the environment is determined by the nature of the fluid, e.g. temperature, corrosion properties, presence of dirt or solid particles.

Table 7.1 Observed average failure rates for instruments.

Instrument	Environment	Experience (item-years)	No. of failures	Failure rate (failures/yr)
Chemical analyser, Oxygen	Poor, chemical/ship	4.34	30	6.92
pH meter	Poor, chemical/ship	28.08	302	10.75
Conductivity indicator	Average, industrial	7.53	18	2.39
Fire detector head	Average, industrial	1 470	128	0.09
Flow transmitter, pneumatic	Average, industrial	125	126	1.00
Level indicator, pneumatic	Average, industrial	898	201	0.22
Pressure controller	Average, industrial	40	63	1.58
Pressure indicator, dial, mechanical	Average, industrial	575	178	0.31
Pressure sensor, differential, electronic	Poor, chemical/ship	225	419	1.86
Pressure transmitter	Average, industrial	85 045	806	0.01
Recorder, pen	Average, industrial	26.02	7	0.27
Temperature indicator and alarm	Fair, laboratory	47.2	101	2.14
Temperature indicator, resistance thermometer	Fair, laboratory	212.3	68	0.32
Temperature indicator, bimetal	Average, industrial	165	215	1.30
Temperature trip unit	Average, industrial	120	70	0.58
Thermocouple	Poor, chemical/ship	317	127	0.40
Valve, gate	Poor, chemical/ship	11 564	841	0.07
non-return	Poor, chemical/ship	1 530	101	0.07
solenoid	Poor, chemical/ship	1 804	66	0.04

Source: after Wright[1].

For an element located in the atmosphere, the environment is determined by atmospheric conditions, e.g. temperature, humidity, salinity, presence of dust. The failure rate of a given type of element will depend on the environment in which it is located: an iron–copper thermocouple will have a higher failure rate in a corrosive acid than in water.

Table 7.2 shows the observed failure rates for given types of elements at three works A, B, C which process different materials and fluids and have different background environments.[2] The observed failure rates can be regarded as the product of a **base failure rate** λ_B and an **environmental correction factor** π_E:

Element failure rate model

$$\lambda_{OBS} = \pi_E \times \lambda_B \qquad [7.25]$$

Here λ_B corresponds to the best environmental conditions and π_E has values 1, 2, 3 or 4, the highest figure corresponding to the worst environment.

The failure rate of elements can alternatively be calculated from the failure rate data/models for the basic components which make up the element.

Table 7.3 shows the calculation of the overall failure rate, from basic component data, for an electronic square root extractor module.[3] The module gives an output voltage signal in the range 1–5 V, proportional to the square root of the input signal with range 4–20 mA; this type of module is commonly used in fluid flow rate control systems. The module is made up from basic electronic components of various types, all connected in series. Several components of each type are present. From

Table 7.2 Observed failure rates for instruments in different chemical plant environments.

Instrument (p = pneumatic)	Observed failure rate, faults/year	Environmental correction factor	Base failure rate, faults/year
Control valve (p)			
–	0.25	1	0.25
Works A	0.57	2	0.29
Works B	2.27	4	0.57
Works C	0.127	2	0.064
Differential pressure transmitter (p)			
–	0.76	1	0.76
Works A (flow)	1.86	3	0.62
Works A (level)	1.71	4	0.43
Works B (flow)	2.66	4	0.67
Works C (flow)	1.22	2	0.61
Variable area flowmeter transmitter (p)			
–	0.68	1	0.68
Works A	1.01	3	0.34
Thermocouple			
Works A	0.40	3	0.13
Works B	1.34	4	0.34
Works C	1.00	4	0.25
Controller			
–	0.38	1	0.38
Works A	0.26	1	0.26
Works B	1.80	1	1.80
Works C	0.32	1	0.32
Pressure switch			
–	0.14	1	0.14
Works A	0.30	2	0.15
Works B	1.00	4	0.25

Source: after Lees[2].

eqn [7.22] the failure rate λ of a module containing m different component types in series with failure rates $\lambda_1, \lambda_2, \ldots, \lambda_i, \ldots, \lambda_m$, and one of each type present is:

$$\lambda = \lambda_1 + \lambda_2 + \ldots + \lambda_i + \ldots + \lambda_m \qquad [7.26]$$

If there are multiple components of each type, all connected in series, then the module failure rate is given by:

Module failure rate – multiple components of each type

$$\lambda = N_1\lambda_1 + N_2\lambda_2 + \ldots + N_i\lambda_i + \ldots + N_m\lambda_m \qquad [7.27]$$

where $N_1, N_2, \ldots, N_i, \ldots, N_m$ are the quantities of each component type. The failure rate of each component type is calculated using the model equation:

$$\lambda_i = (F_{1i} + F_{2i} + F_{3i}) \times K_{1i} \times K_{2i} \times K_{3i} \times K_{4i} \qquad [7.28]$$

Table 7.3 gives values of F_1, F_2, F_3, K_1, K_2, K_3, K_4 and failure rate λ_i for each component. Each failure rate is then multiplied by the appropriate quantity N_i and the $N_i\lambda_i$ values are added together to give a total module failure rate of 3.01×10^{-6} per hour.

Table 7.3 Calculation of overall failure rate for electronic square root module.[a]

Component	Failure rates per 10^{10} hours								
	F_1	F_2	F_3	K_1	K_2	K_3	K_4	Rate[b]	Qty	Value
RESISTORS										
Carbon film										
$0 < R \leq 100$ K	100	0	0	1	1	1	1	100	17	1700
100 K $< R \leq 1$ M	100	0	0	1	1.1	1	1	110	2	220
1 M $< R \leq 10$ M	100	0	0	1	1.6	1	1	160	3	480
Metal film										
$0 < R \leq 100$ K	150	0	0	1	1	1	1	150	17	2550
100 K $< R \leq 1$ M	150	0	0	1	1.1	1	1	165	3	495
POTENTIOMETERS										
$0 < R \leq 50$ K	700	0	0	1	1	1	1	700	8	5600
50 K $< R \leq 100$ K	700	0	0	1	1	1.1	1	770	1	770
CAPACITORS										
Metal film										
$0 < C \leq 33$ nF	200	0	0	1	1	1	1	200	3	600
33 nF $< C \leq 1$ μF	200	0	0	1	1	1.3	1	260	2	520
1 μF $< C \leq 10$ μF	200	0	0	1	1	1.5	1	300	1	300
Ceramic										
$0 < C \leq 3.3$ nF	150	0	0	1	1	1	1	150	1	150
Electrolytic										
$3.2 < C \leq 62$ μF	500	0	0	0.29	1	0.7	1	102	1	102
DIODES										
Silicon LP	200	0	0	0.55	1	1	1	110	2	220
Zener	1000	0	0	1.3	1	1	1	1300	1	1300
TRANSISTORS										
NPN LP	400	0	0	1.4	1	1	1	560	2	1120
INTEGRATED CIRCUITS										
OP AMP	160	50	600	1	1	1	1	810	9	7290
Quad switch	38	320	560	1	1	1	1	918	1	918
OTHERS										
Edge connectors	300	0	0	1	1	1	1	300	8	2400
Soldered joints	20	0	0	1	1	1	1	20	167	3340
PCB	60	0	0	1	1	1	1	60	1	60
	Total rate: 3.01×10^{-6} per hour									
	MTBF: 3.32×10^{5} hours									

[a] Data sources:
1: *Electronic Reliability Data*, National Centre of Systems Reliability, Application Code 2, 25C
2: *Reliability Prediction Manual for Guided Weapon Systems*, MOD
3: Component supplier's information.

[b] Rate $= (F_1 + F_2 + F_3) \times K_1 \times K_2 \times K_3 \times K_4$.

Source: after Hellyer[3].

7.1.7 Design and maintenance for reliability

Design for reliability

The following general principles should be observed.

Element selection. Only elements with well-established failure rate data/models should be used. Furthermore some technologies are inherently more reliable than others. Thus an inductive LVDT displacement sensor (Chapter 8) is inherently more reliable than a resistive potentiometer; the latter involves a contact sliding over a wire track, which will eventually become worn. A vortex flowmeter (Chapter 12) involves no moving parts and is therefore likely to be more reliable than a turbine flowmeter which incorporates a rotor assembly.

Environment. The environment in which the element is to be located should first be defined and the element should consist of components and elements which are capable of withstanding that environment. Thus the diaphragm of a differential pressure transmitter on a sulphuric acid duty should be made from a special alloy, e.g. Hastelloy C, which is resistant to corrosion.

Minimum complexity. We saw above that, for a series system, the system failure rate is the sum of the individual component/element failure rates. Thus the number of components/elements in the system should be the minimum required for the system to perform its function.

Redundancy. We also saw that the use of several identical elements/systems connected in parallel increases the reliability of the overall system. Redundancy should be considered in situations where either the complete system or certain elements of the system have too high a failure rate.

Diversity. In practice faults can occur which cause either more than one element in a given system, or a given element in each of several identical systems, to fail simultaneously. These are referred to as **common mode failures** and can be caused by incorrect design, defective materials and components, faults in the manufacturing process, or incorrect installation. One common example is an electronic system where several of the constituent circuits share a common electrical power supply; failure of the power supply causes all of the circuits to fail. This problem can be solved using **diversity**; here a given function is carried out by two systems in parallel, but each system is made up of different elements with different operating principles. One example is a temperature measurement system made up of two subsystems in parallel, one electronic and one pneumatic.

Maintenance

The **mean down time**, MDT, for a number of items of a repairable element has been defined as the mean time between the occurrence of the failure and the repaired element being put back into normal operation. It is important that MDT is as small as possible in order to minimise the financial loss caused by the element being out of action.

There are two main types of maintenance strategy used with measurement system elements. **Breakdown maintenance** simply involves repairing or replacing the element when it fails. Here MDT or **mean repair time**, T_R, is the sum of the times taken for a number of different activities. These include **realisation** that a fault has occurred, **access** to the equipment, **fault diagnosis**, **assembly** of repair equipment, components and personnel, **active repair/replacement** and finally **checkout**. **Preventive maintenance** is the servicing of equipment and/or replacement of components at regular fixed intervals; the corresponding **maintenance frequency** is m times per year. Here MDT or **mean maintenance time**, T_M, is the sum of times for access, **service/replacement** and checkout activities and therefore should be significantly less than mean repair time with breakdown maintenance.

7.2 Choice of measurement systems

The methods to be used and problems involved in choosing the most appropriate measurement system for a given application can be illustrated by a specific example. The example used will be the choice of the best system to measure the volume flow rate of a clean liquid hydrocarbon, range 0 to 100 $m^3 h^{-1}$, in a 0.15 m (6 inch) diameter pipe. The measured value of flow rate must be presented to the observer in the form of a continuous trend on a chart recorder. The first step is to draw up a specification for the required flow measurement system. This will be a list of all important parameters for the complete system such as measurement error, reliability and cost, each with a desired value or range of values. The first two columns of Table 7.4 are an example of such a 'job specification'. As explained in Chapter 3, system measurement error in the steady state can be quantified in terms of the mean $\bar{E}$ and standard deviation σ_E of the error probability distribution $p(E)$. These quantities depend on the imperfections, e.g. non-linearity and repeatability, of every element in the system. System failure rate λ and repair time T_R were defined in Section 7.1. Initial cost C_I is the cost of purchase, delivery, installation and commissioning of the complete system. C_R is the average cost of materials for each repair.

Table 7.4 Comparison table for selection of flow measurement system.

Parameter	Job specification	System 1 Orifice plate	System 2 Vortex	System 3 Turbine	System 4 Electromagnetic
Measurement error (at 50 $m^3 h^{-1}$)	$\bar{E} \le 0.25$ $\sigma_E \le 0.8$ $m^3 h^{-1}$	0.2 0.7	0.1 0.3	0.03 0.1	Not technically feasible
Initial cost	$C_I \le$ £4000	3500	3000	4200	
Annual failure rate	$\lambda \le 2.0$ failures yr^{-1}	1.8	1.0	2.0	
Average repair time	$T_R \le 8$ h	6	5	7	
Material repair cost	$C_R \le$ £200	100	100	300	

In this example the choice could be between four competing systems based on the orifice plate, vortex, turbine and electromagnetic primary sensing elements. The principles and characteristics of all four elements are discussed in Chapter 12. The configuration of the four systems could be:

1. Orifice plate – electronic D/P transmitter – square rooter – recorder.
2. Vortex element – frequency to voltage converter – recorder.
3. Turbine element – frequency to voltage converter – recorder.
4. Electromagnetic element – recorder.

The next step is to decide whether all the competing systems are technically feasible. The electromagnetic device will not work with electrically non-conducting fluids such as hydrocarbons, so that system 4 is technically unsuitable. Systems 1, 2 and 3 are feasible and the specification for each competing system must then be written down to see whether it satisfies the job specification. Table 7.4 gives possible specifications for the orifice plate, vortex and turbine systems. This data is entirely fictitious and is given only to illustrate the problems of choice. In practice a prospective user may not have all the information in Table 7.4 at his disposal. The manufacturer will be able to give estimates of initial cost C_I and the limits of measurement error, e.g. ±2% of full scale, for the orifice plate system at 50 $m^3 h^{-1}$. He will not, however, be able to give values of mean error $\bar{E}$, failure rates and repair times. The last two quantities will depend on the environment of the user's plants and the maintenance strategy used. This information may be available within the user's company if adequate maintenance records have been kept. From Table 7.4 we see that the turbine flowmeter system 3 does not satisfy the job specification; both initial cost and material repair cost are outside the limits set. This would appear to rule out system 3, leaving 1 and 2. Both these systems satisfy the job specification but the vortex system 2 is cheaper, more accurate and more reliable than the orifice plate system 1. Thus, based on a straightforward comparison of job and system specification, the vortex measurement system 2 would appear to be the best choice for this application.

Under certain circumstances, however, the above conclusion could be entirely wrong. The turbine system is more expensive and less reliable than the vortex system, but is three times more accurate. We must now ask how much this increased accuracy is worth. Suppose the market value of the hydrocarbon is £100 m^{-3}. A measurement error of one standard deviation in the turbine system, where $\sigma_E = 0.1$ $m^3 h^{-1}$, represents a potential cash loss of £10 per hour or approximately £80 000 per annum. A corresponding error in the vortex system, where $\sigma_E = 0.3$ $m^3 h^{-1}$, represents an approximate potential loss of £240 000 per annum. The difference between these two figures far outweighs the extra initial and maintenance costs, so that the turbine system 3 is the best choice in this case. We can conclude, therefore, that in order to choose the correct system for a given application, the financial value of each parameter in the job specification must be taken into account. In a costing application of this type, a digital printout of flow rate is more suitable than an analogue trend record.

7.3 Total lifetime operating cost

The total lifetime operating costing (TLOC) of a measurement system is the total cost penalty, incurred by the user, during the lifetime of the system. The TLOC is given by

TLOC = initial cost of system (purchase, delivery, installation and commissioning)
+ cost of failures and maintenance over lifetime of the system
+ cost of measurement error over lifetime of the system [7.29]

and therefore takes account of the financial value of each parameter in the job specification. The best system for a given application is then the one with minimum TLOC. This method also enables the user to decide whether a measurement system is necessary at all. If no system is installed, TLOC may still be very large because no measurement implies a large measurement error. A measurement system should be purchased if it produces a significant reduction in TLOC.

Using eqn [7.29], we can derive an algebraic expression for TLOC using the parameters listed in Table 7.4. The initial cost of the system is £C_I. If the system lifetime is T years and average failure rate λ faults yr^{-1}, then the total number of faults is λT. Since the average repair time is T_R hours then the total 'downtime' due to repair is λTT_R hours. The total lifetime cost of failures is the sum of the repair cost (materials and labour) and the process cost, i.e. the cost of lost production and efficiency while the measurement system is withdrawn for repair. If we define

£C_R = average materials cost per repair
£C_L = repair labour cost per hour
£C_P = process cost per hour

then the total repair cost is $(C_R\lambda T + C_L T_R \lambda T)$ and the total process cost is $C_P T_R \lambda T$, giving:

$$\text{Total lifetime cost of failures} = [C_R + (C_L + C_P)T_R]\lambda T \qquad [7.30]$$

The above costs only apply to breakdown maintenance; many users also practise preventive maintenance in order to reduce failure rates. Suppose preventive maintenance is carried out on a measurement system m times yr^{-1}, the average maintenance time is T_M hours and the materials cost per service is £C_M. The total number of services is mT and the total time taken for preventive maintenance is mTT_M hours. Usually preventive maintenance of measurement systems is carried out at a time when the process or plant itself is shut down for repair and maintenance. This means that no process costs are incurred during preventive maintenance, giving:

$$\text{Total lifetime maintenance cost} = (C_M + C_L T_M)mT \qquad [7.31]$$

The last term in eqn [7.29] involves the total lifetime cost of measurement error. In order to evaluate this we first need to evaluate the cost penalty function $C(E)$ (£ yr^{-1}) associated with a given steady-state measurement system error E. The form of $C(E)$ depends on the economics of the process on which the measurement is being made. A good example is temperature measurement in a chemical reactor where a degradation reaction is taking place.[4] Here two reactions occur simultaneously: the feedstock A is converted into a desired product B but B is also degraded to an undesired product C. The rates of both reactions increase sharply with temperature, the rate of B to C being more temperature sensitive than the rate of A to B.

There is an optimum temperature at which the yield of B is maximum. If the reactor is operated at either above or below this optimum temperature, then the yield of B is sharply reduced and a cost penalty is incurred. This situation will occur if there is a measurement system error E between measured and true values of reactor temperature: the system tells the operator that the reactor is at optimum temperature

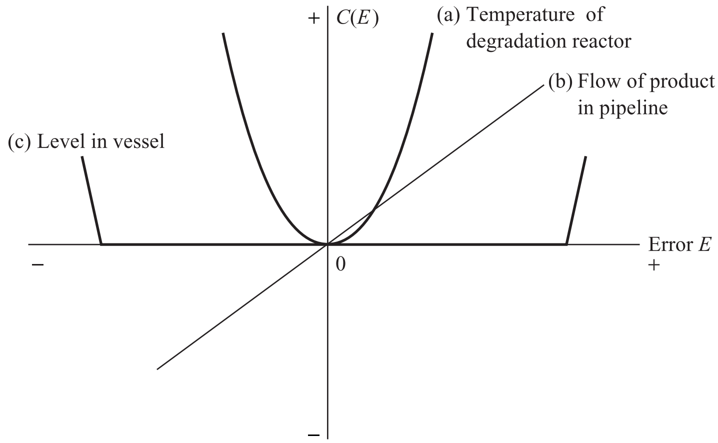


Figure 7.8 Error cost penalty function $C(E)$ for different processes.

when it is not. Figure 7.8(a) shows the form of the cost penalty function $C(E)$ in this case. We see that when $E = 0$, $C(E) = 0$ corresponding to optimum temperature; but $C(E)$ increases rapidly with positive and negative values of E as the yield of B decreases.

In Figure 7.8(b) $C(E)$ represents the cost penalty, due to imperfect flow measurement, incurred by a customer receiving a fluid by pipeline from a manufacturer. If E is positive, the customer is charged for more fluid than he actually receives and so is penalised, i.e. $C(E)$ is positive. If E is negative, the customer is charged for less fluid than was actually received and $C(E)$ is negative. Figure 7.8(c) refers to a non-critical liquid level measurement in a vessel. The vessel should be about half full, but plant problems will occur if the vessel is emptied or completely filled. Thus a cost penalty is incurred only if there is gross measurement error, i.e. the measurement system shows the vessel to be half full when almost empty.

We saw in Chapter 3 that the exact value of measurement error E, for a given measurement system at a given time, cannot be found. We can, however, find the probability that the system will have a certain error. This is quantified using a normal probability density function $p(E)$, with mean $\bar{E}$ and standard deviation σ_E (Table 3.1). The probability of getting a measurement error between E and $E + \mathrm{d}E$ is $p(E)\,\mathrm{d}E$; the corresponding cost penalty is $C(E)p(E)\,\mathrm{d}E$ per year, or $TC(E)p(E)\,\mathrm{d}E$ throughout the system lifetime. The total lifetime cost of measurement error is then found by integrating the above expression over all possible values of E, i.e.

$$\text{Total lifetime cost of measurement error} = T\int_{-\infty}^{\infty} C(E)p(E)\,\mathrm{d}E \qquad [7.32]$$

The integral has a finite value, because the value of $p(E)$ becomes negligible for $|E|$ greater than 3 or 4 standard deviations; it can be evaluated numerically[4] using values of the normalised Gaussian distribution.[5]

From eqns [7.29]–[7.32] we have

$$\text{TLOC} = C_I + [C_R + (C_L + C_P)T_R]\lambda T + (C_M + C_M T_M)mT + T\int_{-\infty}^{\infty} C(E)p(E)\,\mathrm{d}E \qquad [7.33]$$

The relative importance of the terms in the above equation will depend on the application. In the chemical reactor $C(E)$ and C_P will be the major factors so that accuracy and reliability will be far more important than initial cost. In the tank level application, measurement error is unimportant and minimum TLOC will be obtained by the best trade-off between initial cost, reliability, and maintainability.

Conclusion

The first section of this chapter discussed the reliability of measurement systems. The fundamental principles and practical definitions of reliability were first explained and the relationship between reliability and instantaneous failure rate was derived. The typical variation in instantaneous failure rate throughout the lifetime of an element was then discussed and the reliability of series and parallel systems examined. The section concluded by looking at failure rate data and models and general strategies in design and maintenance for reliability.

The second section dealt with the problem of choice of measurement systems approached by comparing the job specification with those of the competing systems. This method does not take account of the financial value of each item in the specification. A better method, discussed in the final section, is to choose the systems with minimum lifetime operating cost.

References

[1] WRIGHT R I (SRD Warrington) 1984 'Instrument reliability', *Instrument Science and Technology*, Institute of Physics, Bristol, vol. 1, pp. 82–92.

[2] LEES F P 1976 'The reliability of instrumentation', *Chemistry and Industry*, March, pp. 195–205.

[3] HELLYER F G (Protech Instruments and Systems) 1985 'The application of reliability engineering to high integrity plant control systems', *Measurement and Control*, 18 June, pp. 172–6.

[4] BENTLEY J P 1979 'Errors in industrial temperature measurement systems and their effect on the yield of a chemical degradation reaction', *8th IMEKO Congress of the International Measurement Confederation*, Moscow, May.

[5] WHITE J, YEATS A and SKIPWORTH G 1974 *Tables for Statisticians*, Stanley Thornes, London, pp. 18–19.

Problems

7.1 A batch of 100 identical thermocouples were tested over a 12-week period. Twenty failures were recorded and the corresponding down times in hours were as follows:

5, 6, 7, 8, 4, 7, 8, 10, 5, 4, 8, 5, 4, 5, 6, 5, 4, 9, 8, 6

Calculate:

(a) Mean down time
(b) Mean time between failures
(e) Mean failure rate
(d) Availability.

7.2 A flow measurement system consists of an orifice plate (λ = 0.75), differential pressure transmitter (λ = 1.0), square root extractor (λ = 0.1) and recorder (λ = 0.1). Calculate the probability of losing the flow measurement after 0.5 year for the following:

(a) A single flow measurement system
(b) Three identical flow measurement systems in parallel
(c) A system with three orifice plates, three differential pressure transmitters and a middle value selector relay (λ = 0.1). The selected transmitter output is passed to a single square root extractor and recorder.

Annual failure rate data are given in brackets; assume that all systems were initially checked and found to be working correctly.

7.3 Use the data given in Table Prob. 3 to decide which level measurement system should be purchased. Assume a breakdown maintenance only strategy is practised, each system has the same measurement error, and there is a 10-year total lifetime.

Table Prob. 3.

	Parameter	System 1	System 2
Initial cost	£	1000	2000
Materials cost per repair	£	20	15
Labour cost per hour	£	10	10
Process cost per hour	£	100	100
Repair time	h	8	12
Annual failure rate	yr^{-1}	2.0	1.0