Engineering safety assessment

An introduction

J.R. Thomson,
BSc(Eng), PhD, CEng, MIMechE, MINucE
National Nuclear Corporation Ltd., Knutsford, Cheshire

Longman Scientific & Technical
Copublished in the United States with
John Wiley & Sons, Inc., New York

Engineering safety assessment

Longman Scientific & Technical,
Longman Group UK Limited,
Longman House, Burnt Mill, Harlow,
Essex CM20 2JE, England
and Associated Companies throughout the world.

Copublished in the United States with
John Wiley & Sons, Inc., 605 Third Avenue, New York, NY 10158

First published 1987

British Library Cataloguing in Publication Data
Thomson, J.R.
Engineering safety assessment: an
introduction.
1. Engineering – Safety measures
I. Title
363.1'1962 TA192

ISBN 0-582-41630-2

Library of Congress Cataloging-in-Publication Data
Thomson, J.R., 1955–
Engineering safety assessment.
Bibliography: p.
Includes index.
1. Reliability (Engineering) 2. System safety.
I. Title
TA169.T54 1987 620'.00452 86-20956

ISBN 0-470-20712-4 (Wiley)

Set in Compugraphic 10/12 Plantin

Produced by Longman Group (FE) Limited
Printed in Hong Kong

Contents

Preface *vii*

List of symbols *xi*

Chapter 1 Introduction *1*
1.1 A brief historical review 1
1.2 Some definitions 5
1.3 The structure of the text 6

Chapter 2 Basic probability *8*
2.1 Probabilities, permutations and combinations 8
2.2 Failure probability and system availability 10
2.3 Rules for combining probabilities 10
2.4 Probability distributions 15

Chapter 3 Systems reliability *39*
3.1 Reliability of time-independent systems 39
3.2 Time-dependent systems 62
3.3 Unrevealed faults and frequency of testing 68
3.4 Reliability data and the accuracy of reliability analysis 71
3.5 Common-mode failures 74
3.6 Human reliability 75
3.7 Software reliability 76
3.8 Conclusions 77

Chapter 4 Reliability of metal structures *80*
4.1 The origins of cracks 80
4.2 Fracture locus: some preliminary definitions 82
4.3 The determination of critical crack size 93
4.4 Subcritical crack growth 103
4.5 Non-destructive examination 107
4.6 Probabilistic fracture mechanics 109
4.7 Conclusions 117

Chapter 5 Major industrial hazards 120
5.1 Accident classification 120
5.2 Explosions 132
5.3 Dispersal of airborne material 149
5.4 Radiation and radioactivity 168
5.5 Conclusions 182

Chapter 6 Probabilistic Risk Assessment (PRA) 187
6.1 The practice of risk assessment 187
6.2 The implications of risk assessment 194
6.3 Conclusions 205

Recommended further reading 208
Appendix I : The behaviour of rising plumes 209
Appendix II : Factors for determining the effects of airborne radioactivity 212
Appendix III : Confidence limits for the expected value of a Poisson distribution 215
Appendix IV : Assessing the risk associated with iodine-131 contamination in milk – a sample calculation 216
Appendix V : Risks to members of the public from various means of electricity generation 218
Index 219

Preface

This book is intended to give an up-to-date introduction to the principal means of assessing the safety of hazardous plant. It is suitable for use as a textbook in senior undergraduate and postgraduate courses in chemical, mechanical and nuclear engineering, as well as providing an introduction to the subject for practising engineers and scientists who have become involved with risk assessment.

No attempt has been made to cover all aspects of what has become a very large subject. Comprehensive coverage would produce a very large book indeed, at a correspondingly high price. Since this book has been aimed at as big a readership as possible, the size of the text has been kept short. To achieve this, the discussion of some topics, in particular the reliability of repairable systems (Ch. 3), the methods of non-destructive examination (Ch. 4), and the effects of earthquakes and other external hazards (Ch. 5), has been minimised. There is no discussion of the effects of terrorist action. In addition, the text avoids wherever possible discussion of 'plant-specific' aspects of reliability or safety, and instead concentrates on the fundamental principles. Thus, for example, there is little discussion of the means employed to ensure the safety of any individual type of process plant. Similarly, a section of Chapter 5 is devoted to an introduction to radiation health physics, but there is little discussion of the likely magnitude of the 'source' term in nuclear accidents, since the size of the source term varies with differing reactor designs and accident sequences. Comprehensive lists of references are given so that readers may pursue these more detailed aspects if they so wish.

The author teaches a course in Safety and Reliability at Edinburgh University, which uses material mostly from Chapters 3, 5 and 6. In addition, Chapter 2 contains a brief review of the relevant areas of probability theory (including means of combining probabilities, probability distributions, confidence limits and the interpretation of statistical data) which may already be familiar to some readers, and Chapter 4 contains a review of those aspects of fracture mechanics which are relevant to safety assessment.

There is a growing need for engineers and scientists trained in the field of Safety and Reliability. The amount of research that has been carried out in this area in the last two decades is truly enormous, and it is hoped that this book will introduce the scope of this research to a wider audience. Also, new government regulations in the UK – the CIMAH (Control of Industrial Major Accident Hazards) Regulations 1984 – require the operators of hazardous installations to forward a safety case to the Health and Safety Executive. Similar regulations are in force in other European countries as a result of the EEC's 'Seveso Directive'. The preparation of such safety cases will require an understanding of the fundamental principles of risk assessment, as are presented in this text.

The author was introduced to this subject while training to be a shift manager at the Prototype Fast Reactor power station, at Dounreay in northern Scotland. I am greatly indebted to a large number of my former colleagues in the Atomic Energy Authority, in particular Messrs E.R. Adam, P.D. Humphreys, A.M. Judd, D. McCowan and G.M. Mitchell. In addition I would like to thank my colleagues at Edinburgh University who have assisted me, in particular Dr D. Drysdale of the Unit of Fire Safety Engineering, Dr D. Glass of the Department of Chemical Engineering, and Dr D. Mills of the Department of Mechanical Engineering, who have read and commented upon parts of the text. Mr A. Nightingale is also thanked for his assistance. Any mistakes or omissions, of course, remain entirely my responsibility. The manuscript was typed quickly and accurately by Miss M. McLeod. Finally, I must thank Professor J.A. McGeough for his guidance and encouragement over many years.

Department of Mechanical Engineering,
Edinburgh University.
February 1986.

Note added in proof

The Chernobyl accident occurred at a late stage in the writing of this book. At the time of going to press there have been 31 deaths due to radiation sickness from the accident. This figure is now unlikely to rise further.

The first reliable estimate for the number of delayed deaths arising from the accident (due to cancers and leukaemias) has been published (*Nature*, 9th September 1986). This suggests that 160 extra cancer deaths could be expected in years to come, amongst the 135 000 local residents who were evacuated at the time of the accident. This shows agreement with predictions of the consequences of major reactor accidents, published before the accident. (See for example, NRPB R–149, February 1983.)

Doses received by people outside the evacuation zone have been small, seldom exceeding a few milliSieverts. There is no evidence that anyone outside the evacuation zone will suffer any harm; however, if the linear dose–risk hypothesis is assumed valid for small doses, it is conceivable that an extra 4000 cancer deaths may occur in Eastern Europe and the Western Soviet Union over the next 50 to 70 years. During this timescale, there will be some tens of millions of cancer deaths in any case, due to the natural occurrence of the disease. Hence, any additional cancer mortality, if it occurs at all, will not be observable (except amongst the evacuees).

As Fig. 5.25 (p. 171) illustrates, it is possible that the effects of low-level radiation may even be beneficial. One can say with certainty that other environmental factors have a considerably greater effect on cancer mortality than low-level radiation; cancer mortality is measurably higher in cities than in rural areas, for example.

Claims that several hundred thousand fatalities may result from the accident are baseless, and should be ignored.

The accident's causes are attributable to poor design (as commented upon by a team of British engineers in 1974 – see *Nuclear Engineering International*, June 1986) compounded by maloperation. The main design flaws which contributed to the accident were a positive void coefficient (leading to the possibility of power instability) combined with a lack of fast shutdown capability, and a graphite moderator which operated at a temperature in excess of 700 °C (a temperature at which graphite spontaneously ignites when exposed to air).

The pressure-tube RBMK design is unique to the USSR. The design would appear to have been adopted because of the relative ease with which such a reactor can be built, since no large pressure vessels are required.

At the time of the accident, some 3800 GW(e)-years of nuclear generating experience had been accumulated worldwide. The relative risks of various means of electricity generation are listed in Appendix V.

October 1986

List of symbols

a	crack half-length
A	area
A	crack incidence coefficient
A	activity (i.e. quantity of radioactivity)
a, b	probit function coefficients
b	spreading length
c_p	specific heat at constant pressure
c_v	specific heat at constant volume
C	Paris' law coefficient
C	surface concentration
${}_nC_r$	number of combinations of r objects from a group of n objects
D	fireball diameter
D_1	dry deposition rate
D_2	diffusion coefficient
D^*	dosage
E	Young's modulus
E	energy
$E(x)$	expectation (or expected value or mean value)
f	number of failures
$f(x)$	probability density function (or failure density function)
f	test frequency (Ch. 3)
f	accident frequency (Ch. 6)
$F(x)$	cumulative probability function (or failure function or unreliability)
F	Froude number
g	gravitational acceleration
G	crack extension force/unit length
H	height of elevated source
k	reaction rate coefficient
k	thermal conductivity
K_c	fracture toughness or critical stress intensity factor
K_r	equilibrium constant
m	Paris' law exponent
m	mass
M	earthquake magnitude
M	molecular mass
M_I	mortality index

N	number of fatigue cycles (Ch. 4)
N	frequency of earthquake occurrence (para. 5.1.8)
N_i	number of kmols of i (para. 5.2)
N	Avogadro's number, 6.02×10^{26}/kmol (para. 5.4)
N	accident consequence (Ch. 6)
p	probability of success
p	pressure
Pr	probit function
${}_nP_r$	number of permutations of r objects from a group of n objects
q	probability of failure
Q	quality factor or Relative Biological Effectiveness
Q_1	quantity of airborne material released by instantaneous source
Q_2	rate of release of airborne material from a continuous source
R	gas constant
R	missile maximum range (para. 5.2.7)
R	stress intensity ratio (para. 4.4.1)
R	idealised cloud radius
$R(x)$	reliability function
s	number of successes
S	spreading length (Ch. 4)
S	burning velocity (Ch. 5)
t	time
T	temperature
T_d	adiabatic decomposition temperature
$T_{\frac{1}{2}}$	half-life
u	wind-speed
U	missile velocity
v	specific volume
v_d	deposition velocity
$\ddot{v}$	peak ground acceleration
V	volume
$V(x)$	variance
x	distance from centre of crack
x	crack length ($=2a$)
x	statistical variable
x, y, z	Cartesian coordinates
z	statistical variable ($=(x-\mu)/\sigma$)

Greek letters

α	Weibull parameter (Ch. 2)
α	coefficient of thermal expansion (Ch. 4)
α	frequency factor (Ch. 5)

α	risk aversion factor (Ch. 6)
β	Weibull parameter
γ	surface energy/unit area (Ch. 4)
γ	ratio of specific heats (Ch. 5)
Γ	gamma function (Ch. 2)
Γ	scaled range (Ch. 5)
δ	crack opening displacement COD
ΔG	Gibbs' Free Energy
ΔG^0	Standard Free Energy change
ΔG_f	Free Energy of formation
ΔH	change in enthalpy
ΔH_c	heat of combustion
ΔH_f	heat of formation
Δl	percentage elongation at fracture
Δp	overpressure
ΔS	change in entropy
ε	explosive yield (Ch. 5)
ε	normal strain (Ch. 4)
θ	crack incidence exponent (Ch. 4)
θ	angle of missile trajectory (Ch. 5)
λ	hazard rate
Λ	washout coefficient
μ	mean
μ	linear absorption coefficient (Ch. 5)
μ_m	mass absorption coefficient
ν	Poisson's ratio
ρ	reaction rate
$\sigma(x)$	standard deviation (Ch. 2)
σ	stress (Ch. 4)
$\sigma_x, \sigma_y, \sigma_z$	plume dispersal coefficients (Ch. 5)
ϕ	azimuthal angle (para. 5.2)
ϕ	radiation flux (para. 5.4)
χ	concentration

Subscripts

Chapter 3

A	pertaining to sub-system A
B	pertaining to sub-system B
c	common failure mode
u	unrevealed failure mode
S	pertaining to whole system

Pollution may be considered to be prolonged, low-level discharges of toxic material to the environment, with a possible long-term detrimental effect. On the other hand, a *toxic release* is a brief accidental high-level discharge of toxic material to the environment. This book is aimed primarily at the assessment of the probabilities and consequences of large, possibly multiple-fatality, accidents. The effects of pollution require consideration of ecological aspects which are beyond the scope of this book.

The effects of many accidents, e.g. explosions, are *prompt*. This means that all fatalities and injuries will occur at the time of the accident or very shortly thereafter. In some toxic release accidents, however, the effects can be *delayed*. This term usually (but not always) signifies a toxic release involving carcinogens, such as the Windscale fire in 1957 or the Seveso accident in 1976.

1.3 The structure of the text

Probabilistic safety assessment is now used widely in the nuclear and chemical industries, and reliability engineering (a subset of safety assessment) is used extensively in the aerospace industry. The probability of a given accident can be determined by consideration of the *systems reliability* (Ch. 3) and the *structural reliability* (Ch. 4). The *consequences* of the accident can then be determined by considering explosion thermodynamics, atmospheric plume dispersal and human toxicology (Ch. 5). This process may then be repeated for a number of different potential accident sequences in a given plant, e.g. different magnitudes of explosions or different sizes or rates of toxic release (Fig. 1.4).

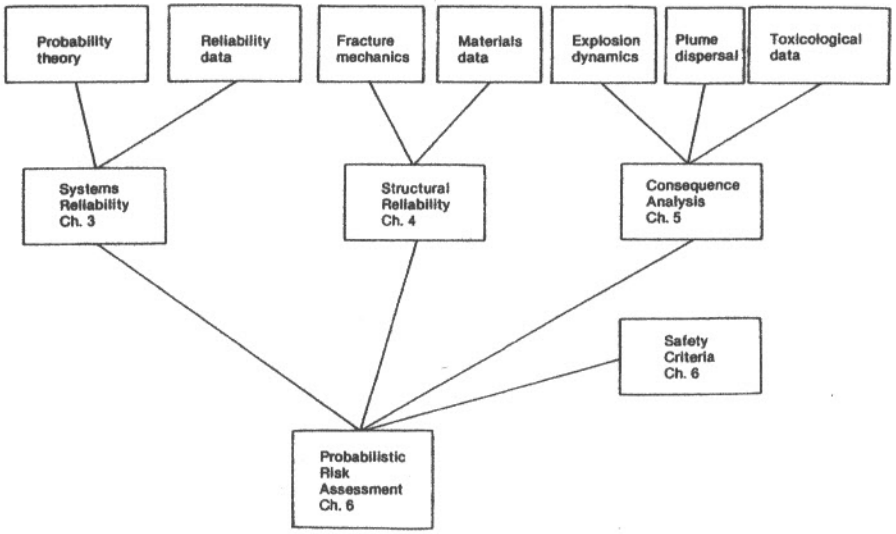


Fig. 1.4 The requirements for safety assessment, and the structure of the text

Chapter 2

Basic probability

One statement from Chapter 1 that bears repeating is this: in the field of safety assessment, it is axiomatic that there is no such thing as absolute safety. For example, no matter how many 'back-up' (redundant) systems are fitted to, say, a spacecraft for attitude control, it is always possible that all systems will fail simultaneously, or else another factor will cause the simultaneous failure of all systems ('common-mode' failure). By implication, this means that system failures and accidents are stochastic by nature. Hence, in order to assess risk, the mathematics of probability must be properly understood.

This chapter introduces elementary probability theory. The theory has been interspersed with worked examples in an attempt to make it more digestible, and also as a way of elaborating on some possible pitfalls.

2.1 Probabilities, permutations and combinations

Consider a system which either succeeds (s) or fails (f). The probability of the system succeeding may then be given as

$$P(\text{success}) = p = \frac{\text{the number of successes}}{\text{the total number of trials}} = \frac{s}{s+f} \qquad [2.1]$$

since there are only two possible outcomes, success or failure. Similarly, the probability of failure is

$$P(\text{failure}) = q = \frac{\text{the number of failures}}{\text{the total number of trials}} = \frac{f}{s+f} \qquad [2.2]$$

Furthermore, it is apparent that

$$p + q = 1 \qquad [2.3]$$

For the tossing of a coin, $p = q = 0.5$. For throwing a die, the probability of throwing a six (or any other number) in one throw is 1/6.

However, if we wish to calculate the probability of throwing, say, a total of eight with two dice, the problem becomes slightly more difficult. A total of eight could arise from a variety of outcomes of the dice, i.e.

(2+6), (3+5), (4+4), (5+3) or (6+2).

For the first possible outcome, (2+6), there will be a 1/6 probability of getting 2 with the first die. The probability of getting 6 with the second die, knowing we already have 2, will also be 1/6. Hence the possibility of getting (2+6) is $(1/6)^2$, i.e. 1/36. The same argument applies for the other outcomes which give a total of eight, so the probability of scoring eight in one throw of the dice is

$5 \times 1/36 = 5/36$

The probability of *not* scoring eight is therefore 31/36.

2.1.1 Permutations and combinations

The number of possible *permutations* ${}_nP_r$ of a group of n objects is the number of ways that a sub-group of r objects may be arranged when taken from the group. If three objects A, B and C are considered, and letting $r = n = 3$, the possible permutations of the three objects will be

ABC, ACB, CAB, CBA, BAC and BCA.

In other words, ${}_3P_3 = 6$. If $r = 2$, then the possible permutations are

AB, BA, AC, CA, BC and CB

so ${}_3P_2 = 6$ also. In general,

$$ {}_nP_r = \frac{n!}{(n-r)!} \qquad [2.4] $$

(Note: for $n = r$, $(n-r)! = 1$ and ${}_nP_n = n!$)

The number of permutations of a group of objects includes groupings of the same objects in a different order. However, the order of grouping of the objects is often unimportant. For this case the number of *combinations* of the objects is of interest.

The number of possible combinations ${}_nC_r$ of a group of n objects is the number of ways that a sub-group of r objects can be taken from the group, without regard to order. Hence the possible combinations of two from the three objects A, B and C will be

AB, AC and BC

so ${}_3C_2 = 3$. In general,

$$ {}_nC_r = \frac{n!}{r!(n-r)!} = \frac{n(n-1)\ldots((n-r)+1)}{r!} \qquad [2.5] $$

Also, ${}_nC_n = 1$, i.e. there is only one combination of n objects from a group of n objects. This is common sense, since combinations are defined to be without regard to order.

In the assessment of reliability, combinations are usually more important than permutations, since it is usually (but not always) more important to know which events can lead to system failure, than to know in which order they occur.

2.2 Failure probability and system availability

In general, the probability of the failure of a component or system may be defined to be

$$P(\text{failure}) = \lim_{n \to \infty} \left(\frac{f}{n} \right) \qquad [2.6]$$

where n is the number of times the component or system is used, and f is the number of failures. If, for example, the failure of a car starter motor is considered, the failure probability might be defined in a variety of ways, e.g.

1 failure per 2000 starts,
or 1 failure per 2 years,
or 1 failure per 40 000 kilometres.

For a car starter motor, the first definition is probably the most applicable, since the starter motor is only used once per journey and one might expect that the length of journey or the time between journeys was irrelevant. Hence, for this example, the 'failure rate per demand' is important. For other components, e.g. the fuel pump, the 'failure rate per annum' or the 'failure rate per kilometre' might be more significant.

Similarly, for many control systems in industrial plant, some systems or components are in continuous use whereas others are only used as required. The failure rate data must be in the appropriate form for each type of system. System *availability* may be defined for a given system as follows:

$$\text{Availability} = \frac{\text{(Operating time)}}{\text{(Operating time)} + \text{(Time under repair)}} \qquad [2.7]$$

The time that a system spends under maintenance or repair is usually called *outage time.*

2.3 Rules for combining probabilities

There are six basic rules for combining probabilities, as follows:

1. *Independent events* may be defined as events in which the occurrence or non-occurrence of one event does not affect the probability of occurrence of the other event. The events are not related in any way.

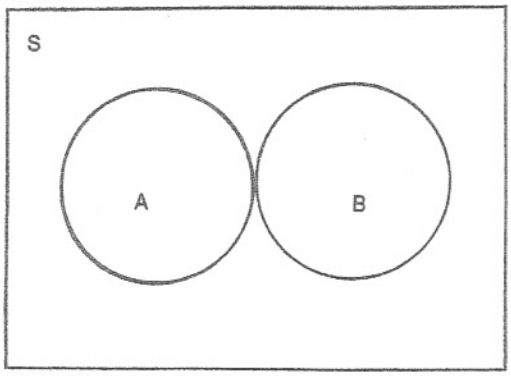


Fig. 2.1 Venn diagram for independent, mutually exclusive events

2. *Mutually exclusive events* cannot happen at the same time. The occurrence of one event prohibits the other event. For example a light bulb cannot be both on (A) and off (B). For mutually exclusive events, $P(A \cap B) = 0$. Such events are also called 'disjoint'.
3. *Complementary events* are events such that, if event A does not occur, event B must occur, and vice versa. Hence

$$P(A) + P(B) = 1 \quad [2.8]$$

or else $$P(B) = P(\bar{A}) \quad [2.9]$$

where $\bar{A}$ is 'NOT A'. See Fig. 2.2.

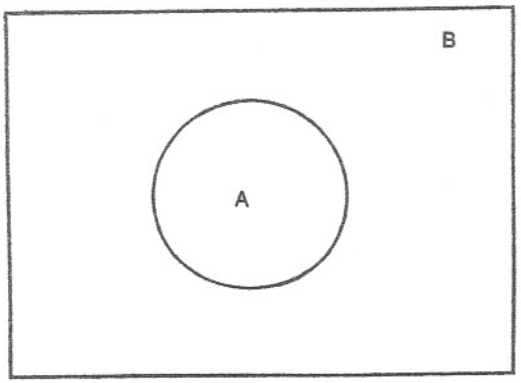


Fig. 2.2 Venn diagram for complementary events

4. *Conditional events* are events which occur conditionally on the occurrence of other events. We define a conditional probability $P(A \mid B)$ to be the probability of event A given that event B has occurred. From Fig. 2.3, it can be seen that

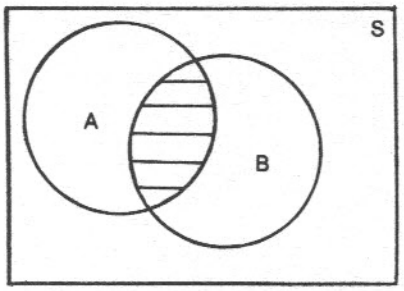


Fig. 2.3 Venn diagram for conditional events

$$P(\mathrm{B}) = \frac{\mathrm{B}}{S} \qquad [2.10]$$

and

$$P(\mathrm{A} \cap \mathrm{B}) = \frac{(\mathrm{A} \cap \mathrm{B})}{S} \qquad [2.11]$$

The conditional probability P(A | B) will be given by

$$P(\mathrm{A} \mid \mathrm{B}) = \frac{\mathrm{A} \cap \mathrm{B}}{\mathrm{B}} \qquad [2.12]$$

Equations [2.10], [2.11] and [2.12] together give that

$$P(\mathrm{A} \mid \mathrm{B}) = \frac{P(\mathrm{A} \cap \mathrm{B})}{P(\mathrm{B})} \qquad [2.13]$$

Similarly

$$P(\mathrm{B} \mid \mathrm{A}) = \frac{P(\mathrm{A} \cap \mathrm{B})}{P(\mathrm{A})} \qquad [2.14]$$

5. *Simultaneous events* (the occurrence of both A and B).
(a) If A and B are *independent* events,

$$P(\mathrm{A} \mid \mathrm{B}) = P(\mathrm{A}) \qquad [2.15]$$

and $$P(\mathrm{B} \mid \mathrm{A}) = P(\mathrm{B}) \qquad [2.16]$$

Hence, from eqn [2.13]

$$P(\mathrm{A} \cap \mathrm{B}) = P(\mathrm{A}).P(\mathrm{B}) \qquad [2.17]$$

Similarly

$$P(\mathrm{A}_1 \cap \mathrm{A}_2 \cap \mathrm{A}_3 \ldots \cap \mathrm{A}_n) = \prod_{i=1}^{n} P(\mathrm{A}_i) \qquad [2.18]$$

where A_i is the ith of n *independent* events.

(b) The probabilities of *non-independent* events can be determined from eqns [2.13] and [2.14], thus:

$$P(A \cap B) = P(A).P(B \mid A) = P(B).P(A \mid B) \quad [2.19]$$

6. *The occurrence of at least one of two events.*
In many cases the occurrence of *either* A or B *or* both A and B is of interest.
(a) For *independent, mutually exclusive* events (Fig. 2.1),

$$P(A \cap B) = 0 \quad [2.20]$$

Hence

$$P(A \cup B) = P(A) + P(B) \quad [2.21]$$

Similarly, for n independent, mutually exclusive events,

$$P(A_1 \cup A_2 \cup A_3 \ldots \cup A_n) = \sum_{i=1}^{n} P(A_i) \quad [2.22]$$

where A_i is the ith of n independent, mutually exclusive events.
(b) For *independent* events which are *not* mutually exclusive,

$$\begin{aligned} P(A \cup B) &= P(\text{A OR B OR BOTH A AND B}) \\ &= 1 - P(\bar{A} \cap \bar{B}) \\ &= 1 - P(\bar{A}).P(\bar{B}) \quad \text{(from [2.17])} \\ &= 1 - (1 - P(A))(1 - P(B)) \quad \text{(from [2.9])} \\ &= P(A) + P(B) - P(A).P(B) \\ &= P(A) + P(B) - P(A \cap B) \end{aligned} \quad [2.23]$$

(c) For *non-independent* events, eqn [2.19] can be substituted into eqn [2.23], to give

$$\begin{aligned} P(A \cup B) &= P(A) + P(B) - P(A).P(B \mid A) \\ &= P(A) + P(B) - P(B).P(A \mid B) \end{aligned} \quad [2.24]$$

2.3.1 The rare events approximation

In many cases in the reliability analysis of systems, a system failure can occur due to the simultaneous occurrence of events which are independent but not mutually exclusive (eqn [2.23]). However, it is often the case that the probabilities of these events are very small, i.e. the events are *rare*. In that case,

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \quad [2.23]$$

$$= P(A) + P(B) - P(A).P(B)$$

$$\simeq P(A) + P(B) \quad [2.25]$$

In other words, rare independent events may be treated as if they were also mutually

exclusive (eqn [2.21]) when calculating the probability of one or both of the events occurring. In a similar fashion, eqn [2.22] applies when the number of simultaneous rare events is greater than two.

2.3.2 Bayes' theorem

The probability for the simultaneous occurrence of non-independent events is given in eqn [2.19]

$$P(A_i \cap B) = P(A_i).P(B \mid A_i) = P(B).P(A_i \mid B) \qquad [2.19]$$

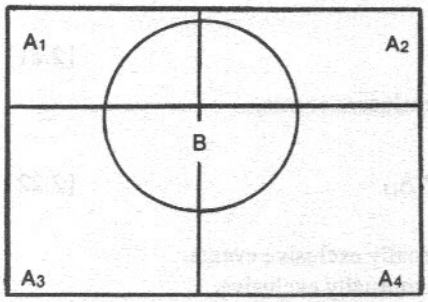


Fig. 2.4 Venn diagram illustrating Bayes' theorem

Here the subscript i denotes the ith of n mutually exclusive events, while B is some other event (Fig. 2.4). Hence,

$$P(A_i \mid B) = P(A_i).\frac{P(B \mid A_i)}{P(B)} \qquad [2.26]$$

Furthermore, Fig. 2.4 shows that for mutually exclusive events A_j,

$$\sum_{j=1}^{n} P(A_j \mid B) = 1 \qquad [2.27]$$

Multiplying by $P(B)$ gives

$$P(B) = \sum_{j=1}^{n} P(B)\,P(A_j \mid B)$$

$$= \sum_{j=1}^{n} P(A_j \cap B) \qquad [2.28]$$

Similarly,

$$P(B) = \sum_{j=1}^{n} P(A_j)\,P(B \mid A_j) \qquad [2.29]$$

Finally, substituting eqn [2.29] into [2.26] yields

$$P(A_i \mid B) = \frac{P(A_i)\,P(B \mid A_i)}{\sum_{j=1}^{n} P(A_j)\,P(B \mid A_j)} \qquad [2.30]$$

This is Bayes' theorem. This equation means that, if all the conditional probabilities for B, $P(B \mid A_i)$, are known, it is possible to determine the 'reverse' conditional probability for A_i given that B has occurred. This is a useful tool because in some cases it may be more convenient to determine $P(B \mid A_i)$ than $P(A_i \mid B)$. (This will be further discussed in Examples 2.5 and 2.6. Bayes' theorem can be a difficult concept to grasp; these practical examples may aid comprehension.)

2.4 Probability distributions

2.4.1 Discrete and continuous random variables

Whenever the magnitude of a quantity is determined to some extent by chance, the probability of obtaining a specific magnitude can be plotted against that magnitude to obtain a probability density function, $f(x)$ (Fig. 2.5). Some random variables can have

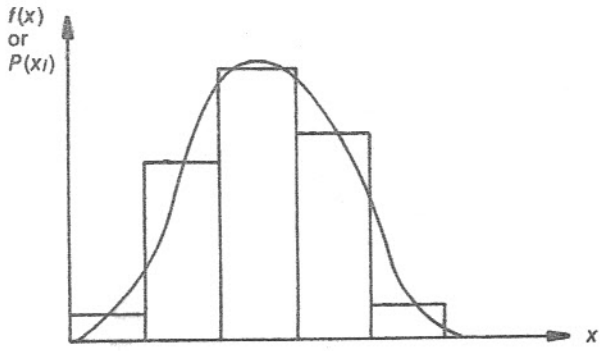


Fig. 2.5 Discrete and continuous random variable probability density functions

a continuous range of values, for example the lap times at a motor racing circuit. Other variables can only have *discrete* values, for example the number of racing cars passing a spectator in a given time interval. It is apparent that the summation over the whole range of probabilities must equal unity, i.e.

$$\sum_{i=1}^{n} P(x_i) = 1 \qquad [2.31]$$

for a discrete random variable and

$$\int_{-\infty}^{\infty} f(x)\,dx = 1 \qquad [2.32]$$

for a continuous random variable.

In a similar manner, we can define *cumulative* probability functions, i.e. the probability that the variable will be less than a specified size, thus:

$$F(x_j) = \sum_{i=1}^{j} P(x_i), \quad j < n \qquad [2.33]$$

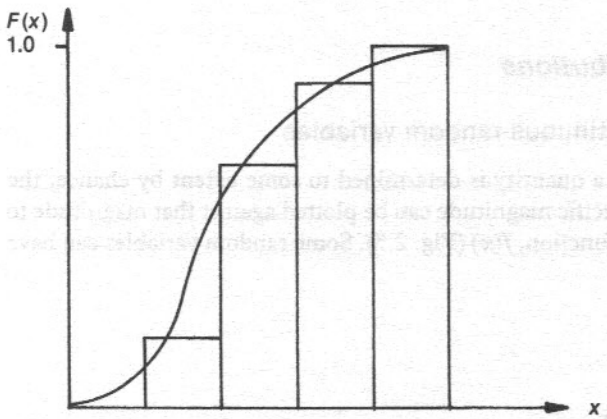


Fig. 2.6 Discrete and continuous random variable cumulative probability functions

for discrete random variables and

$$F(x_1) = \int_{-\infty}^{x_1} f(x)\,dx \qquad [2.34]$$

for continuous random variables.

Equation [2.34] may be differentiated to yield

$$f(x) = \frac{dF(x)}{dx} \qquad [2.35]$$

Furthermore, the probability that the variable x lies in an interval ($a \leqslant x \leqslant b$) will be given by

$$P(a \leqslant x \leqslant b) = \int_a^b f(x)\,dx \qquad [2.36]$$

Finally, the probability that the variable is greater than a value x_1 is called the *Reliability* function, $R(x_1)$.

$$R(x_1) = \int_{x_1}^{\infty} f(x)\,dx \qquad [2.37]$$

From eqns [2.32], [2.34] and [2.37] it is apparent that

$$R(x_1) + F(x_1) = 1 \qquad [2.38]$$

2.4.2 Expectation, variance, standard deviation and hazard rate

At this stage we must define some terms which will be used later

1. *Expectation* (or expected value) $E(x)$.
The expectation $E(x)$ is defined to be

$$E(x) = \sum_{i=1}^{n} x_i P(x_i) \qquad [2.39]$$

for a discrete random variable and

$$E(x) = \int_{-\infty}^{\infty} x f(x)\,dx \qquad [2.40]$$

for a continuous random variable. It is also called the *mean* value.

2. *Variance* $V(x)$
The variance is a measure of the degree of 'spread' that the variable has about its mean, or expected, value. For a discrete variable

$$V(x) = \sum_{i=1}^{n} (x_i^2 P(x_i)) - E^2(x) \qquad [2.41]$$

and for a continuous variable

$$V(x) = E(x^2) - E^2(x) \qquad [2.42]$$

Mathematically, this is the second central moment of the distribution.

3. *Standard deviation* $\sigma(x)$
The standard deviation is defined to be the positive square root of the variance:

$$\sigma(x) = +\sqrt{V(x)} \qquad [2.43]$$

4. *Hazard rate* $\lambda(x)$

This is particularly important in reliability engineering, where probability density functions of failures (from batches of similar components) as a function of time are of concern. The hazard rate is defined to be:

$$\lambda(x) = \frac{\text{failure density } f(x)}{\text{no. of components not yet failed}}$$

$$= \frac{f(x)}{1 - F(x)}$$

$$= \frac{f(x)}{R(x)} \qquad [2.44]$$

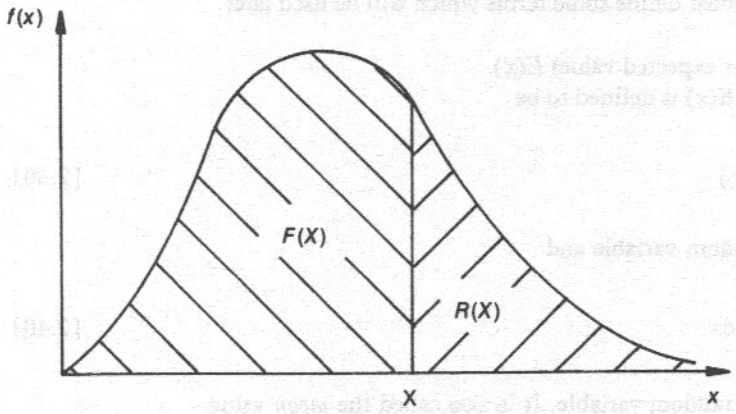


Fig. 2.7 Hypothetical failure (probability) density functions $f(x)$ illustrating the cumulative failure probability (failure function) $F(x)$ and the number of components not yet failed (reliability function) $R(x)$

With reference to Fig. 2.7 the following points on nomenclature should be clarified. In reliability analysis the probability density function $f(x)$ is usually called the *failure density function*. The cumulative probability function $F(x)$ is usually termed the *failure function*, or *unreliability*. The *reliability* function, $R(x)$, is the cumulative probability of *not* failing, i.e.

$$F(x) + R(x) = 1 \qquad [2.38]$$

2.4.3 Continuous random variables – the Gaussian and exponential distributions

The *Gaussian*, or normal, distribution (Fig. 2.8) has a probability density function given by

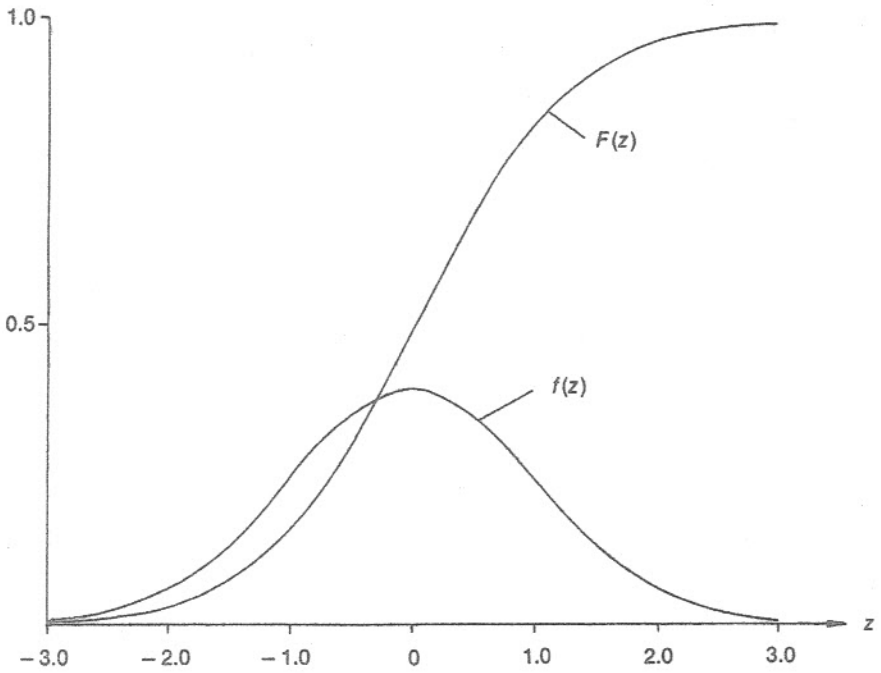


Fig. 2.8 The Gaussian distribution: probability density function $f(z)$ and cumulative probability function $F(z)$

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} .\exp\left[\frac{-(x-\mu)^2}{2\sigma^2}\right] \qquad [2.45]$$

where μ = mean (expected) value $E(x)$
and σ = standard deviation.

If we define a new variable z to be

$$z = \frac{x - \mu}{\sigma} \qquad [2.46]$$

then eqn [2.45] yields

$$f(z) = \frac{1}{\sqrt{2\pi}} .\exp\left[\frac{-z^2}{2}\right] \qquad [2.47]$$

The Gaussian distribution does not find much application in reliability analysis. However, it is of importance in analysing the consequences of accidents in which plumes of toxic material are released to the atmosphere (Section 5.3), since the concentration of such material follows a Gaussian distribution.

Table 2.1 The Gaussian distribution

$z\left(=\frac{x-\mu}{\sigma}\right)$	$f(z)$	$F(z)$
0	0.399	0.5
0.1	0.397	0.5398
0.2	0.391	0.5793
0.3	0.381	0.6179
0.4	0.368	0.6554
0.5	0.352	0.6915
0.6	0.333	0.7257
0.7	0.312	0.7580
0.8	0.290	0.7881
0.9	0.266	0.8159
1.0	0.242	0.8413
1.1	0.218	0.8643
1.2	0.194	0.8849
1.3	0.171	0.9032
1.4	0.150	0.9192
1.5	0.130	0.9332
1.6	0.111	0.9452
1.7	0.094	0.9554
1.8	0.079	0.9641
1.9	0.066	0.9713
2.0	0.054	0.9772
2.1	0.044	0.9821
2.2	0.035	0.9861
2.3	0.028	0.9893
2.4	0.022	0.9918
2.5	0.018	0.9938
2.6	0.014	0.9953
2.7	0.010	0.9965
2.8	0.008	0.9974
2.9	0.006	0.9981

Equation [2.47] cannot be integrated analytically to give the cumulative probability function $F(z)$, but values for this function are presented in Table 2.1.

The *exponential* distribution is of great importance in reliability analysis. In this distribution, *the hazard rate is constant*. (In studies of batches of similar manufactured articles, a constant hazard rate is often observed. This will be further discussed in Section 3.2.)

From eqn [2.44] we can therefore write

$$\lambda = \text{constant} = \frac{f(x)}{1-F(x)}$$

$$\Rightarrow \frac{dF(x)}{dx} = \lambda(1 - F(x)) \quad \text{from eqn [2.35]}$$

$$\Rightarrow \frac{-dR(x)}{dx} = \lambda R(x) \qquad \text{from eqn [2.38]}$$

Integrating gives

$$R(x) = e^{-\lambda x} \qquad [2.48]$$
$$F(x) = 1 - e^{-\lambda x} \qquad [2.49]$$
$$f(x) = \lambda e^{-\lambda x} \qquad [2.50]$$

Equations [2.48], [2.49] and [2.50] describe the exponential function. These equations are illustrated in Fig. 2.9.

The expected value or mean for the exponential distribution can be obtained from eqn [2.40]:

$$E(x) = \int_0^\infty x f(x)\,dx \qquad [2.40]$$

$$= \int_0^\infty x \lambda e^{-\lambda x}\,dx$$

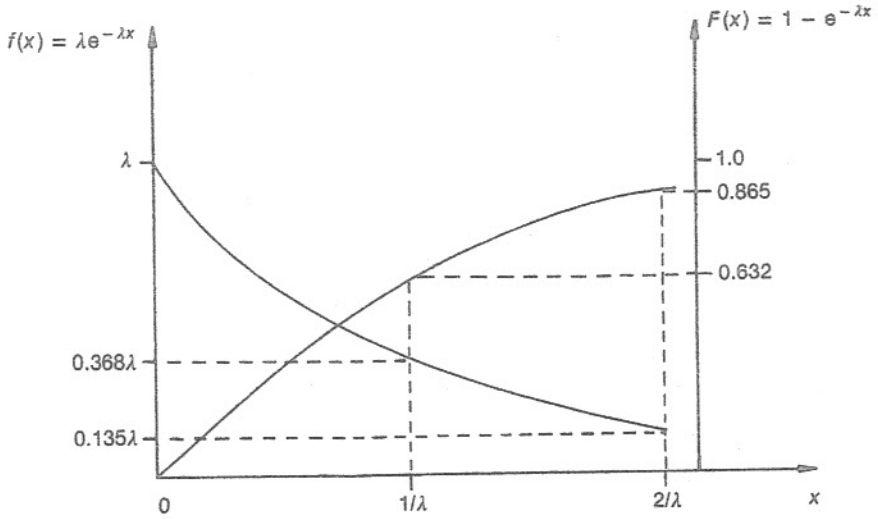


Fig. 2.9 The exponential distribution: probability (or failure) density function $f(x)$ and cumulative probability function (or failure function) $F(x)$

Integrating by parts gives

$$E(x) = 1/\lambda \qquad [2.51]$$

The expected value for the exponential distribution is often called the Mean-Time-To-Failure, or MTTF.

The standard deviation for this distribution can be obtained from eqns [2.42] and [2.43]:

$$\sigma^2 = E(x^2) - E^2(x)$$

$$= \int_0^\infty x^2 \lambda \, e^{-\lambda x} \, dx - 1/\lambda^2$$

Integrating by parts gives

$$\sigma = 1/\lambda \qquad [2.52]$$

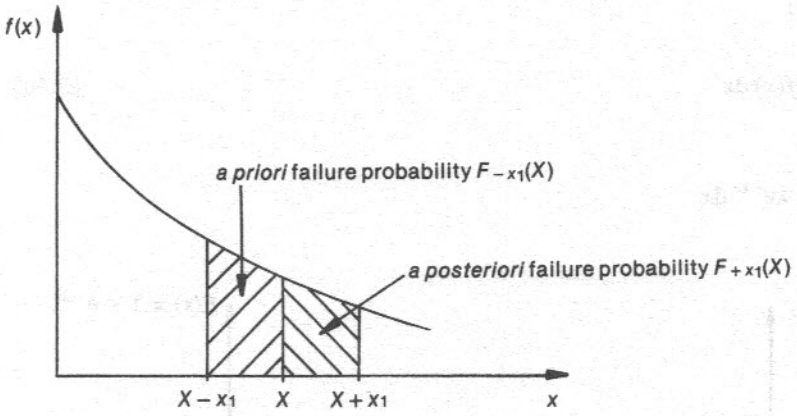


Fig. 2.10 A priori and *a posteriori* failure probabilities

We may also calculate failure probabilities over discrete intervals $(X, X+x_1)$ or $(X-x_1, X)$. The former is called the *a posteriori* failure probability $F_{+x_1}(X)$, and the latter is called the *a priori* failure probability $F_{-x_1}(X)$ (Fig. 2.10). The *a posteriori* failure probability will be given by

$$F_{+x_1}(X) = P \text{ [failure during } x_1 \text{ given the component has survived up to } X \text{]} \qquad [2.53]$$

Now the probability of surviving up to X will be given by

$$R(X) = e^{-\lambda X} \qquad [2.54]$$

and the probability of surviving up to X and failing during interval $(X, X+x_1)$ will be

$$\int_X^{X+x_1} f(x) \, dx$$

$$= e^{-\lambda X} - e^{-\lambda(X+x_1)} \qquad [2.55]$$

Hence, eqns [2.53], [2.54] and [2.55], together with eqn [2.19] give that

$$F_{+x_1}(X) = \frac{e^{-\lambda X} - e^{-\lambda(X+x_1)}}{e^{-\lambda X}}$$

$$= 1 - e^{-\lambda x_1} \qquad [2.56]$$

Thus we may, by comparing this result with eqn [2.49], conclude that for the exponential distribution the probability of failure over any interval $(X, X+x_1)$ or $(X-x_1, X)$ is independent of X. The reliability is constant for all equal intervals x_1. To put it another way, the probability of failure has no 'memory' of the past. This result applies only to the exponential distribution.

Finally, where the hazard rate λ is small, an important simplification can be made if the interval x_1 is also small such that

$$\lambda x_1 << 1$$

In this case

$$e^{-\lambda x_1} \simeq 1 - \lambda x_1 \qquad [2.57]$$

Hence, eqns [2.48] and [2.49] give that

$$F_{+x_1}(X) = \lambda x_1 \qquad [2.58]$$

$$R_{+x_1}(X) = 1 - \lambda x_1 \qquad [2.59]$$

These approximations are valid to two significant figures if $\lambda x_1 < 0.1$.

Example 2.1 An example using the Gaussian distribution

The concentration in a toxic cloud obeys a Gaussian distribution. The cloud is radially symmetrical and has a peak concentration of 500 ppm. Two hundred metres from the centre the concentration has dropped to 100 ppm. Determine the standard deviation ('dispersion coefficient'), and the radius at which the concentration has dropped to 25 ppm.

Method

Here the variable x is radial distance and $f(z)$ is a 'scale' concentration.

$$f(z) = \frac{1}{\sqrt{2\pi}} . \exp \frac{(-z^2)}{2}$$

where $z = \frac{x}{\sigma}$, a 'scaled' radial distance.

(The point of maximum concentration is the centre of the cloud. Hence $\mu = 0$.) Also

$$\frac{f(z)}{f(0)} = \frac{C(z)}{C(0)}$$

where $C(z)$ is the concentration at scaled radius z. Hence

$$C(z) = C(0) \exp\left[\frac{-z^2}{2}\right]$$

i.e. at 200 metres radius,

$$\exp\left(\frac{-x^2}{2\sigma^2}\right) = 0.2$$

i.e. σ = 111.5 metres

The radius x at which the concentration will have dropped to 25 ppm will be given by

$$\exp\left(\frac{-x^2}{2\sigma^2}\right) = 0.05$$

i.e. x = 273 metres for a concentration of 25 ppm.

Example 2.2 An example using the exponential distribution

A batch of 1000 cassette recorders are believed to have an exponential failure density function, $f(t) = \lambda e^{-\lambda t}$. Fifty cassette recorders fail in the first year. Determine

(a) the hazard rate λ,
(b) The time for half the recorders to fail, and
(c) how many fail in the tenth year.

Method

(a) The hazard rate can be obtained from eqn [2.49]

$$F(t) = 1 - e^{-\lambda t}$$
$$= 0.05$$
$$\lambda = 0.0513 \text{ per year}$$

(b) The time taken for half the recorders to fail is given by

$$R(t) = e^{-\lambda t}$$
$$= 0.5$$
$$\lambda t = 0.693$$
i.e. t = 13.5 years

(c) The number failing in the tenth year will be given by (failure probability × number remaining)

i.e. $1000 \times F_{+1}(t) \times R(t)$

where t equals nine years. Hence the number failing in the tenth year is

$$1000 \times (1 - e^{-\lambda.1}) \times e^{-\lambda t} = 32$$

2.4.4 Discrete random variables – the binomial and Poisson distributions

The *binomial* distribution applies in cases where there are a discrete number of independent trials n and only two possible outcomes, p and q. (e.g. success or failure). The probabilities for the two outcomes are fixed. The binomial distribution is obtained

from the polynomial expansion of $(p + q)^n$. Since $(p + q)$ must equal unity, $(p + q)^n$ must also equal unity.

$$(p + q)^n = 1$$

$$= p^n + n\,p^{n-1}q + n\frac{(n-1)}{2!}p^{n-2}q^2 + \ldots$$

$$\ldots + \frac{n(n-1)\ldots(n-r+1)}{r!}p^{(n-r)}q^r +$$

$$\ldots + npq^{n-1} + q^n \qquad [2.60]$$

Here, the coefficient for the $(r + 1)$th term is the combination ${}_nC_r$ (eqn [2.5]), and the probability of r successes in n trials is given by

$$P_r = \frac{n!}{r!(n-r)!}p^r q^{n-r}$$

$$= {}_nC_r\,p^r q^{n-r} \qquad [2.61]$$

The expected value is given by eqn [2.39]

$$E(x) = \sum_{i=1}^{n} x_i P(x_i) \qquad [2.39]$$

$$= \sum_{x=1}^{n} x\ {}_nC_x\,p^x\,q^{n-x}$$

$$= \sum_{x=1}^{n} \frac{n(n-1)!}{(x-1)!\,(n-x)!}\cdot p.p^{x-1}.q^{n-x}$$

$$= np\sum_{x=1}^{n} \frac{(n-1)!}{(x-1)!\,(n-x)!}\cdot p^{x-1}.q^{n-x}$$

Letting $s = n - 1$ and $a = x - 1$ gives

$$E(x) = np\sum_{a=0}^{s} \frac{s!}{a!\,(s-a)!}\,p^a q^{s-a} \qquad [2.62]$$

The summation here is equal to the whole binomial expression [2.60], so it equals unity. Hence,

$$E(x) = np \qquad [2.63]$$

The variance $V(x)$ may be calculated from eqn [2.41]

$$V(x) = \sum_{i=1}^{n} x_i^2 P(x_i) - E^2(x) \qquad [2.41]$$

where

$$\sum_{i=1}^{n} x_i^2 P(x_i) = \sum_{x=1}^{n} x(x-1)\,{}_nC_x\, p^x q^{n-x} + \sum_{x=1}^{n} x\,{}_nC_x\, p^x q^{n-x}$$

$$= np + \sum_{x=1}^{n} x(x-1)\,\frac{n!}{x!(n-x)!}\,p^x q^{n-x}$$

$$= np + p^2 n(n-1).\sum_{x=2}^{n} \frac{(n-2)!}{(x-2)!\,(n-x)!}\,p^{x-2}q^{n-x}$$

$$= np + n^2p^2 - np^2 \qquad [2.64]$$

Hence, eqns [2.41], [2.63] and [2.64] give

$$V(x) = np(1-p)$$
$$= npq \qquad [2.65]$$

and therefore

$$\sigma = +\sqrt{npq} \qquad [2.66]$$

The *Poisson* distribution represents the probability of an isolated event occurring a specified number of times in a given time interval when the rate of occurrence (the hazard rate λ) is fixed. Whereas in the binomial distribution both the occurrences and the non-occurrences of the event are counted, in the Poisson distribution only the occurrences are counted. This distribution might describe, for example, the number of car accidents at an accident 'black spot' in a given time interval, or the number of people struck by lightning in a given (large) population in a given period.

The Poisson distribution may be derived from the binomial distribution coefficients if the 'success' probability p is very much less than the number of trials n. In a small time period Δt, the expected number of occurrences will be $p = \lambda\Delta t$.
Hence, from eqn [2.61], the probability $P_r(t)$ of r 'successes' occurring in the interval $(0,t)$ will be given by:

$$P_r(t) = \frac{n!}{r!(n-r)!}\,(\lambda\Delta t)^r(1-\lambda\Delta t)^{n-r} \qquad [2.67]$$

If we let n tend to infinity, and noting that

$$1 - \lambda\Delta t \simeq e^{-\lambda\Delta t} \qquad [2.57]$$

for $\lambda\Delta t$ small, we obtain

$$P_r(t) \to \frac{n^r}{r!}\,(\lambda\Delta t)^r\, e^{-n\lambda\Delta t} \qquad [2.68]$$

Finally, $t = n\Delta t$ so

$$P_r(t) = \frac{n^r}{r!} \cdot \left(\frac{\lambda t}{n}\right)^r \cdot e^{-\lambda t}$$

$$= \frac{(\lambda t)^r e^{-\lambda t}}{r!} \qquad [2.69]$$

This is the Poisson distribution, which enables the probability of r events happening in time interval t to be determined given the hazard rate λ. Note also that if no events occur ($r = 0$), the expression reduces to the reliability function in the exponential distribution, $R(t) = e^{-\lambda t}$ since 0! is defined to equal unity.

Table 2.2 contains a summary of the main features of the Gaussian, exponential, binomial and Poisson distributions.

Table 2.2 Summary of Gaussian, exponential, binomial and Poisson distributions

Distribution	$f(x)$ or P_r	$F(x)$	$\lambda(x)$	$E(x)$	$\sigma(x)$
Gaussian	$\frac{1}{\sigma\sqrt{2\pi}} \cdot \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]$	$\int_{-\infty}^{x} f(x)\,dx$	$\frac{f(x)}{1-F(x)}$	μ	σ
Exponential	$\lambda e^{-\lambda x}$	$1-e^{-\lambda x}$	λ	$1/\lambda$	$1/\lambda$
Binomial	${}_nC_r\, p^r q^{n-r}$	$\sum_{i=1}^{j} P_i$	—	np	$+\sqrt{npq}$
Poisson	$\frac{(\lambda t)^r e^{-\lambda t}}{r!}$	$\sum_{i=1}^{j} P_i$	—	λt	$\sqrt{\lambda t}$

Example 2.3 An example using the binomial distribution

Use the binomial distribution to determine the probability of system success for a system of three emergency standby diesel generators in a hospital. Each generator has 90 per cent availability. Calculate the system success probability for meeting the demanded load if:

(a) each generator can supply 100 per cent of load demand,
(b) each generator can supply 50 per cent of load demand,
(c) each generator can supply 33⅓ per cent of load demand.

Method
The binomial distribution gives $(p + q)^r$, where $p = 0.90$, or $(p^3 + 3p^2q + 3pq^2 + q^3) = 1$

System state	*Probability*	
All 3 generators available	p^3	0.729
2 generators available	$3p^2q$	0.243
1 generator available	$3pq^2$	0.027
No generators available	q^3	0.001
		1.000

(a) If each supply 100 per cent of the load, then the success probability is

$$p^3 + 3p^2q + 3pq^2 = \underline{0.999}$$

(b) At least two generators must be available if each supplies 50 per cent of the load, so the success probability is

$$p^3 + 3p^2q = \underline{0.972}$$

(c) All three generators are required if each supplies $33\frac{1}{3}$ per cent of the load, so the success probability is

$$p^3 = \underline{0.729}$$

Example 2.4 An example using the Poisson distribution

The individual risk of death due to ischaemic heart disease (all ages) is 3.172×10^{-3} per year (Grist, 1978). In a random sample of 1000 people (with a standard age distribution) what is the probability that there will be

(a) zero deaths
(b) three deaths
(c) ten deaths

due to ischaemic heart disease in a one-year period?

Method
The Poisson distribution (eqn [2.69]) can be used directly to calculate the answers. The death rate λ is 3.172 per year.

(a) P_0 = 4.192 per cent
(b) P_3 = 22.298 per cent
(c) P_{10} = 0.119 per cent

If the sample size was larger, say 2000 people, then λ would be 6.344 per year. In this case, the corresponding probabilities would be

(a) P_0 = 0.176 per cent
(b) P_6 = 15.94 per cent
(c) P_{20} = 0.0008 per cent

It can be seen that the range of observed death rates (given the expected value of 3.172 per thousand per year) decreases as the sample size increases. This is illustrated in Fig. 2.11. We can say that the *confidence limits* narrow as the sample size increases. This is considered further in Examples 2.5 and 2.6.

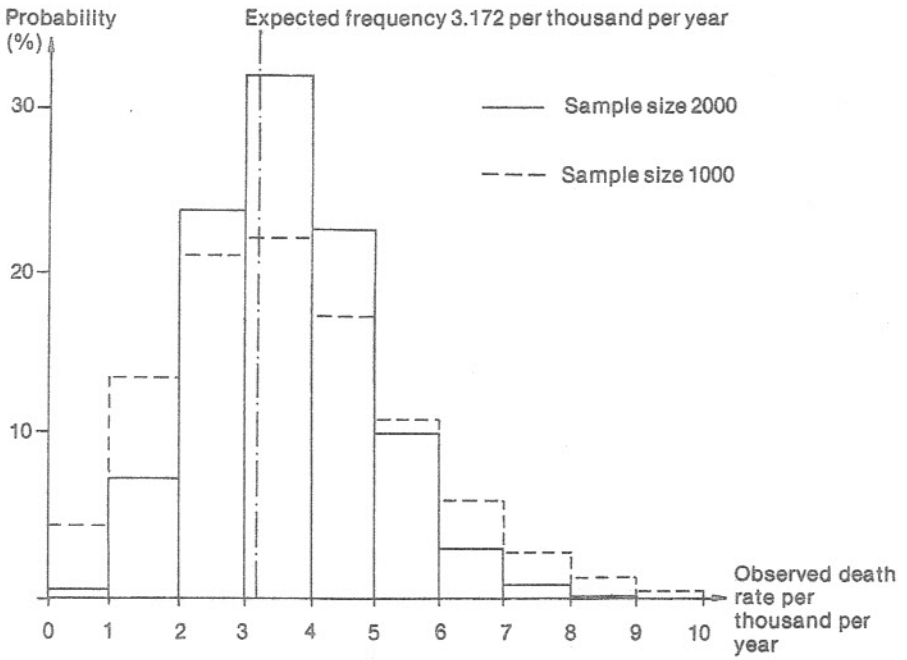


Fig. 2.11 The effect of sample size in statistical data: probabilities of observed death rates due to ischaemic heart disease for different population sample sizes

Example 2.5 An example of the use of Bayes' theorem

In an epidemiological survey of heart disease, 22 deaths in a sample of 1000 people are observed over five years. What is the probability that the 'true' rate of heart disease lies in the range 4–6 thousand per year?

Method

This is an example of the use of Bayes' theorem (see Section 2.3.2). A suitable form of eqn [2.30] for this problem is

$$P(\lambda \mid N)\mathrm{d}\lambda = \frac{P(\lambda)\mathrm{d}\lambda\, P(N \mid \lambda)}{\int_0^\infty P(\lambda)\, P(N \mid \lambda)\mathrm{d}\lambda} \qquad [2.70]$$

where

$P(\lambda \mid N)$ is the probability that (true rate = λ) given N observed deaths

$P(N \mid \lambda)$ is the probability of N observed deaths given λ, i.e.

$P(N \mid \lambda) = \dfrac{(\lambda t)^N \mathrm{e}^{-\lambda t}}{N!}$ (eqn [2.69]), where t is the period of observation, and

$P(\lambda)$ is the probability of λ having a particular value.

In other words, we are trying to 'reverse' the data to give us a probability that the true death rate due to heart disease lies within a specified range (λ, $\lambda + d\lambda$), given a single observation of death rate due to heart disease. (In Example 2.4, we calculated the probabilities of observing a specified death rate from the true rate.)

We have no prior knowledge of the probabilities of λ having a particular value; hence $P(\lambda)$ is a uniform distribution and eqn [2.70] simplifies to

$$P(\lambda \mid N)d\lambda = \frac{P(N \mid \lambda)d\lambda}{\int_0^\infty P(N \mid \lambda)d\lambda} \qquad [2.71]$$

Furthermore the denominator may be simplified since it may be integrated thus:

$$\int_0^\infty P(N \mid \lambda)d\lambda = \frac{\int_0^\infty (\lambda t)^N e^{-\lambda t} d\lambda}{N!} = 1/t \qquad [2.72]$$

Hence

$$P(\lambda \mid N)d\lambda = t\,P(N \mid \lambda)d\lambda \qquad [2.73]$$

where

$$P(N \mid \lambda) = \frac{(\lambda t)^N e^{-\lambda t}}{N!} \qquad [2.74]$$

Equations [2.73] and [2.74] can now be used to solve the problem. In this case $N = 22$, $t = 5$ and the 'expected value' of λ [$= N/t$] is 4.4 per thousand per year.

Range of λ	0–2	2–4	4–6	6–8	8–10
Central value of λ	1	3	5	7	9
$P(N \mid \lambda)$	1.43×10^{-8}	0.0204	0.0702	0.0052	5.98×10^{-5}
$t\,d\lambda$	10	10	10	10	10
$P(\lambda \mid N)d\lambda$	1.43×10^{-7}	0.204	0.702	0.0523	5.98×10^{-4}
			0.958		

Hence the probability that λ lies in the range 4–6 per thousand per year is 70.2 per cent. Alternatively, we may say with *better than 95 per cent confidence* that λ lies in the range 2–8 per thousand per year.

If epidemiological data are taken from smaller samples still, the likelihood of obtaining misleading data increases. If a study finds, say, 3 cases of a cancer which might be induced by environmental pollution instead of the 'expected' 2, then nothing is proved except that statistics can be abused.

Example 2.6 Confidence limits

Two valves, out of a batch of 132 that was manufactured five years ago, have failed in service. Determine the 95 per cent confidence limits for the hazard rate.

Method

This problem is similar to Example 2.5. Here $N = 2$, $t = 5$ and the 'expected value' of λ is 0.4 failures per year, i.e. the mean hazard rate is 3.03×10^{-3} per annum.

Range of λ	0–0.2	0.2–0.4	0.4–0.6	0.6–0.8	0.8–1.0	1.0–1.2	1.2–1.4
Central value of λ	0.1	0.3	0.5	0.7	0.9	1.1	1.3
$P(N \mid \lambda)$	0.0758	0.2510	0.2565	0.1850	0.1125	0.0618	0.0318
$t\,d\lambda$	1	1	1	1	1	1	1
$P(\lambda \mid N)d\lambda$	0.0758	0.2510	0.2565	0.1850	0.1125	0.0618	0.0318
	⌊					⌋	
				0.9426			

In this case, the paucity of data means that the true failure rate lies within the relatively wide range zero to 1.2 failures per year, with 94.26 per cent confidence. Alternatively, we can say that the *upper bound hazard rate* is (1.2 ÷ 132) or 9.1×10^{-3} per annum.

Appendix III gives 95 per cent and 99 per cent confidence limits for the expected value of a Poisson distribution.

2.4.5 Other probability distributions – the Weibull, gamma and lognormal distributions

Several other important probability distribution functions exist for continuous variables. Each can be used to describe differing aspects of reliability.

The *Weibull* distribution (Weibull, 1951) is a useful function with wide applicability. This function has two parameters α and β which can be adjusted to fit the experimental data.

$$R(x) = \exp\left(-\left[\frac{x}{\alpha}\right]^{\beta}\right) \qquad [2.75]$$

Weibull used this formula to describe data as diverse as yield strengths and fatigue lives of steels and the size distribution of particles of fly-ash. The parameters α and β can be determined by plotting the experimental data on special Weibull paper (Fig. 2.12). The cumulative failure probability, $F(x)$, is plotted against the variable x. The point Z has importance; it has coordinates $(e,(1 - e^{-1}))$. A line through Z parallel to the data yields a value for β where it intercepts $x = 1$, as shown in the diagram. The ordinate β can be read from the right-hand axis. Similarly, α is equal to the value of x for which $F(x)$ is equal to $(1 - e^{-1}) = 0.632$. Thus for the data shown in Fig. 2.12, β equals 1.35 and α equals 15 approximately. The failure density distribution for the Weibull distribution is given by

$$f(x) = \frac{\beta x^{\beta-1}}{\alpha^{\beta}} \cdot \exp\left(-\left[\frac{x}{\alpha}\right]^{\beta}\right) \qquad [2.76]$$

The distribution simplifies to the exponential distribution if $\beta = 1$. If $\beta = 2$ it is called the *Rayleigh* distribution (Fig. 2.13). See also Table 2.4 and Example 2.7.

The *gamma* distribution is a two-parameter function (like the Weibull distribution). The two parameters, α and β, can again be adjusted to fit experimental data. The failure density distribution for the gamma distribution is

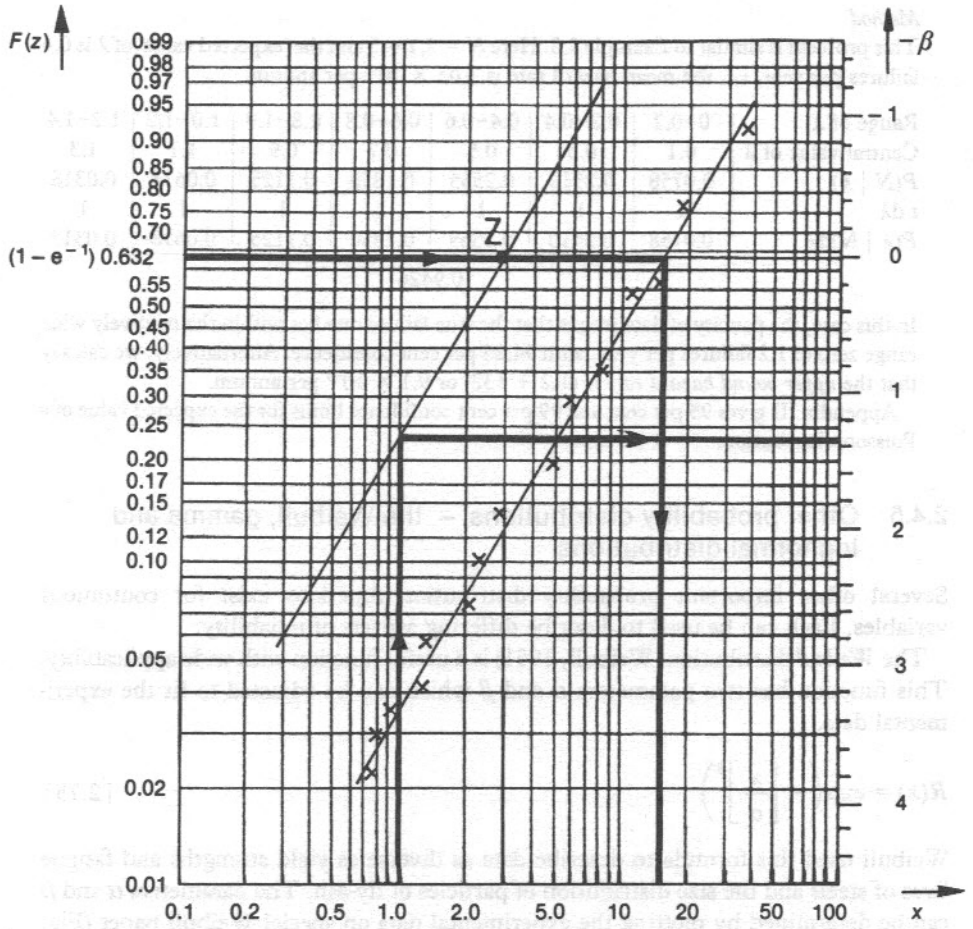


Fig. 2.12 Weibull graph paper: how to determine values of α and β from experimental data

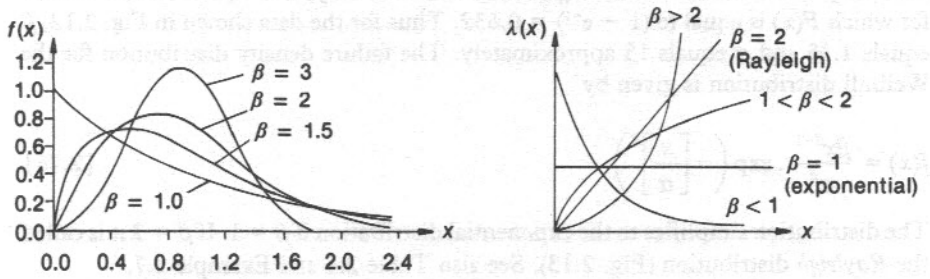


Fig. 2.13 The Weibull distribution: failure density function and hazard rates for different values of parameter β

Table 2.3 The gamma function $\Gamma(\beta)$

β	$\Gamma(\beta) = \int_0^\infty x^{\beta-1}e^{-x}\,dx$
1.00	1.0000
1.05	0.9735
1.10	0.9514
1.15	0.9330
1.20	0.9182
1.25	0.9064
1.30	0.8975
1.35	0.8912
1.40	0.8873
1.45	0.8857
1.50	0.8862
1.55	0.8889
1.60	0.8935
1.65	0.9001
1.70	0.9086
1.75	0.9191
1.80	0.9314
1.85	0.9456
1.90	0.9618
1.95	0.9799
2.00	1.0000

Notes: 1. For higher values of β, the following recursive formula applies:

$$\Gamma(\beta + 1) = \beta\,\Gamma(\beta)$$

2. For integer values of β

$$\Gamma(\beta) = (\beta - 1)!$$

$$f(x) = \frac{x^{\beta-1}}{\alpha^\beta \Gamma(\beta)} \exp\left[-\frac{x}{\alpha}\right] \qquad [2.77]$$

where

$$\Gamma(\beta) = \int_0^\infty x^{\beta-1}\,e^{-x}\,dx \qquad [2.78]$$

$\Gamma(\beta)$ is called a gamma function; values for this function are tabulated in books of mathematical tables, and a brief table of values is given in Table 2.3.

The gamma distribution simplifies to the exponential distribution for $\beta = 1$. If β is an integer the distribution is known as the *Erlangian* distribution. For this distribution, the failure density distribution is given by

$$f(x) = \frac{x^{\beta-1}}{\alpha^\beta\,(\beta-1)!} \exp\left[-\frac{x}{\alpha}\right] \qquad [2.79]$$

since, for integer values of β,

$$\Gamma(\beta) = (\beta - 1)! \qquad [2.80]$$

The gamma distribution is less frequently used for reliability evaluation than the Weibull distribution. Some features of the distribution are shown in Fig. 2.14; see also Table 2.4.

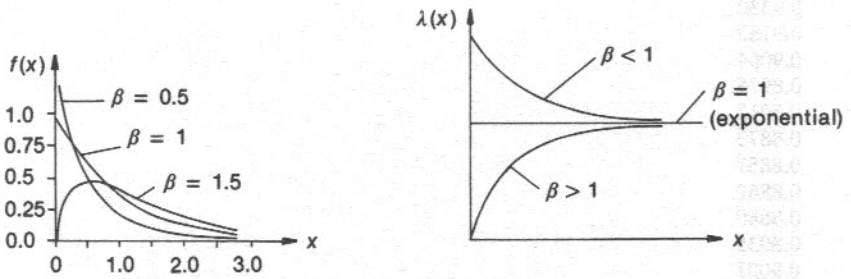


Fig. 2.14 The gamma distribution: failure density function and hazard rates for different values of parameter β

Table 2.4 The Weibull, gamma and lognormal distributions

Distribution	$f(x)$	$R(x)$	$\lambda(x)$	$E(x)$	$\sigma^2(x)$
Weibull	$\frac{\beta x^{\beta-1}}{\alpha^\beta} \cdot \exp\left[-\left(\frac{x}{\alpha}\right)^\beta\right]$	$\exp\left[-\left(\frac{x}{\alpha}\right)^\beta\right]$	$\frac{\beta x^{\beta-1}}{\alpha^\beta}$	$\alpha\Gamma\left(1+\frac{1}{\beta}\right)$	$\alpha^2\left[\Gamma\left(1+\frac{2}{\beta}\right) - \Gamma^2\left(1+\frac{1}{\beta}\right)\right]$
Gamma	$\frac{x^{\beta-1}}{\alpha^\beta\Gamma(\beta)} \cdot \exp\left[\frac{-x}{\alpha}\right]$	$\int_x^\infty f(x)\,dx$	$\frac{f(x)}{R(x)}$	$\alpha\beta$	$\alpha^2\beta$
Lognormal	$\frac{1}{x\sigma\sqrt{2\pi}} \exp\left[\frac{-(\ln x-\mu)^2}{2\sigma^2}\right]$	$\int_x^\infty f(x)\,dx$	$\frac{f(x)}{R(x)}$	$\exp(\mu + \frac{1}{2}\sigma^2)$	$\exp(2\mu + 2\sigma^2) - \exp(2\mu + \sigma^2)$

The *lognormal* distribution is similar to the Gaussian distribution (eqn [2.45]), except that it is the logarithm of the random variable x which follows a normal distribution in this case, i.e.

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left[-\frac{(\ln x - \mu)^2}{2\sigma^2}\right] \qquad [2.81]$$

where μ and σ are the mean and standard deviation, respectively (Fig. 2.15). The distribution can be used to describe the reliability of non-destructive examination (Section 4.6.2) and can also be used to fit the distribution of repair times in repairable systems.

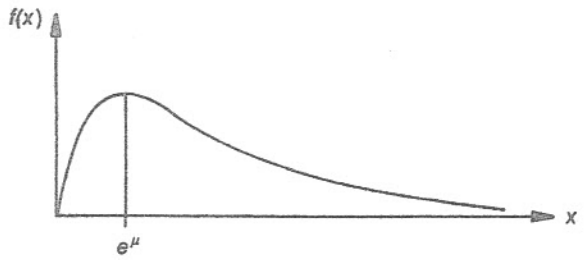


Fig. 2.15 The lognormal distribution: failure density function

Example 2.7 An example using the Weibull distribution

A follow-up study on a batch of 1000 automobiles reveals the following data (hypothetical)

Year in which useful life ended x	No. of autos. $f(x)$	Cumulative total $F(x)$	No. still in use $R(x)$	Hazard rate $\lambda = \frac{f(x)}{R(x)}$
1	3	3	997	0.003
2	7	10	990	0.007
3	5	15	985	0.005
4	4	19	981	0.004
5	8	27	973	0.008
6	3	30	970	0.003
7	5	35	965	0.005
8	15	50	950	0.016
9	35	85	915	0.038
10	55	140	860	0.064
11	29	169	831	0.035
12	35	204	796	0.044
13	46	250	750	0.061
14	110	360	640	0.172
15	121	481	519	0.233
16	119	600	400	0.298
17	141	741	259	0.544
18	109	850	150	0.727
19	62	912	88	0.705
20	40	952	48	0.833

If these figures are plotted on Weibull paper (Fig. 2.16) it becomes apparent that *two* separate processes are in action. When the automobiles are fairly new, there is only a small and approximately constant attrition rate, due to road accidents. Using the methods described in Section 2.4.5, in this early phase the distribution is exponential (i.e. $\beta = 1$) with a hazard rate $\lambda(=1/\alpha)$ of 0.004 per annum and a mean-time-to-failure (MTTF) $(=1/\lambda)$ of 250 years.

After seven years, the hazard rate increases due to the onset of 'wear-out'. The graphical method gives that $\beta = 4.5$ in this second stage, which corresponds to an accelerating hazard rate λ (see Fig. 2.13), and also that $\alpha = 16$. Using the formula in Table 2.4, we can calculate that $\lambda = (1.72 \times 10^{-5})x^{3.5}$ in the second stage.

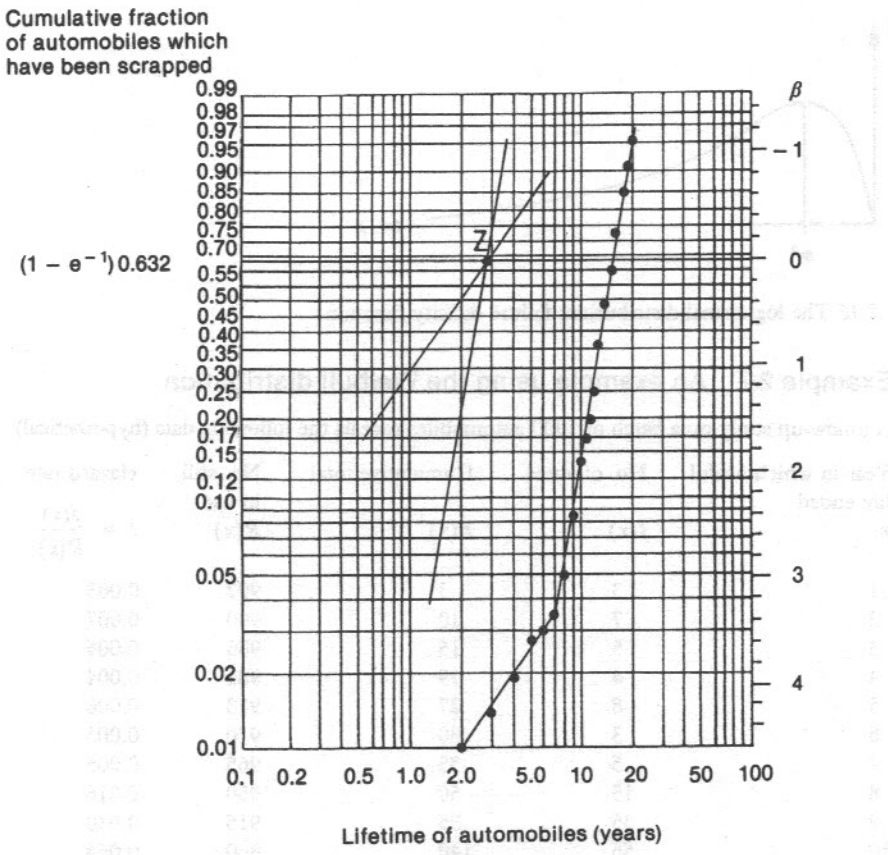


Fig. 2.16 Weibull plot of automobile lifetimes (Example 2.7)

A graph of hazard rate against time is shown in Fig. 2.17. The curves correspond to the λ values computed from the Weibull plot (Fig. 2.16). It is apparent that a hazard rate plot is 'noisier' than a Weibull plot, since a relatively small variation in the number of annual failures can have a large effect on the hazard rate. This illustrates how the Weibull plot, by using cumulative (integrated) data, is able to 'smooth out' results and enable data to be more readily interpreted, especially when sample size is relatively small.

The subject of time-dependent effects upon hazard rate will be discussed again in Section 3.2.

Questions

(See Appendix III for a table of confidence limits)

2.1 Ninety-six identical radar systems are fitted to a fleet of commercial aircraft. The manufacturers claim an in-service breakdown (hazard) rate of 2×10^{-5} per hour. If each plane flies 3000 hours per year on average, what is the most probable number of breakdowns per year?

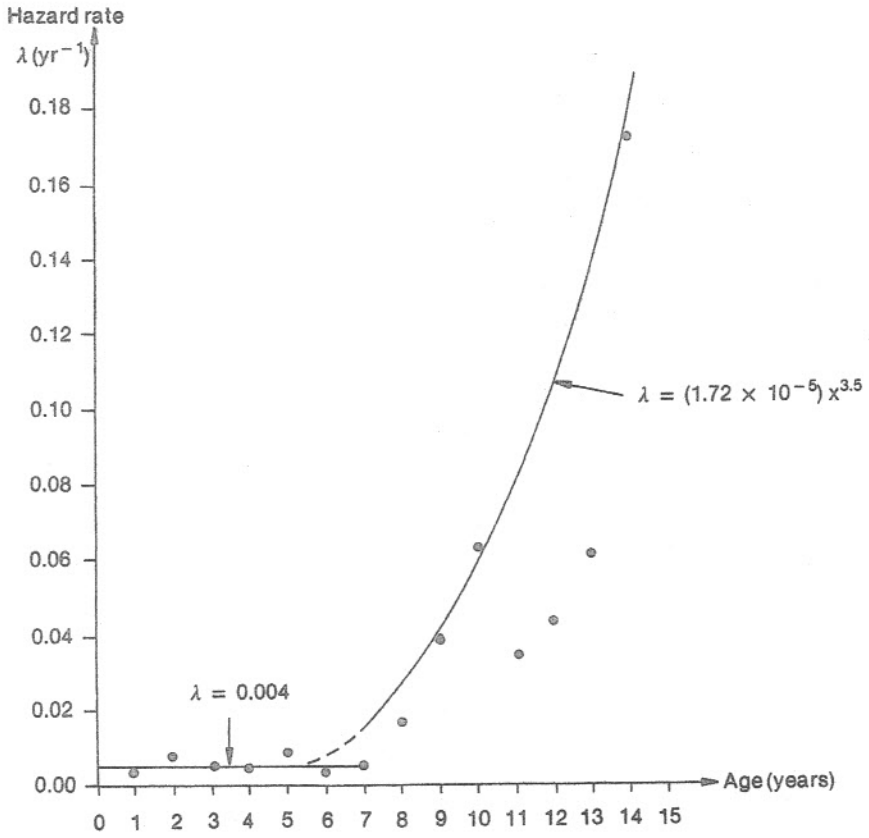


Fig. 2.17 Hazard rate versus age for automobiles (Example 2.7)

What is the probability of: (a) 8; (b) 10; (c) 12 breakdowns occurring in any one year? (5.76; 0.095; 0.035; 0.009)

2.2 A batch of 400 identical bearings have shown 12 in-service failures in their first three years of use. Determine the hazard rate confidence limits with: (a) 90%; (b) 95%; (c) 99% confidence.

	(lower bound	upper bound
(a)	5×10^{-3} per annum	15×10^{-3}
(b)	5×10^{-3}	17.5×10^{-3}
(c)	5×10^{-3}	20×10^{-3})

2.3 The individual risk per year of death due to cancers for children less than 4 years old is 79×10^{-6} (Grist, 1978).

A survey of child cancers reveals 10 deaths amongst children under 4 years old over a period of 5 years when the average number of children in that age group in the area is 16 200.

Does this result warrant further investigation of discharges from a chemical plant in that area?

('Expected' hazard rate is 1.3 per year. 95 per cent confidence limits for the 'measured' value are 1.0–3.5.)

2.4 Starting from the definition of the reliability function for the Weibull distribution,

$$R(x) = \exp\left(-\left(\frac{x}{\alpha}\right)^{\beta}\right)$$

determine the corresponding expressions for the probability density function $f(x)$, the cumulative probability function $F(x)$, the hazard rate $\lambda(x)$, the expectation $E(x)$, and the standard deviation $\sigma(x)$.

2.5 Supposing that commercial satellite launches show an average 80 per cent success rate, that an insurance company can call upon a maximum of £130 m. of reserves, and that the total cost per satellite launch is £60 m., determine a minimum premium to ensure that the probability of the insurance company exceeding its reserves within the first ten launches is less than: (a) 0.1%; (b) 1%; (c) 10%.

What is the most probable profit or loss in each case?

(£29 m.; £23 m.; £11 m.; £170 m.; £110 m.; −£10 m.)

2.6 The star signs of 992 female athletes are as follows: Aries 82; Taurus 90; Gemini 68; Cancer 51; Leo 60; Virgo 102; Libra 97; Scorpio 106; Sagittarius 92; Capricorn 75; Aquarius 87; Pisces 82 (*SHE* magazine. Sept. 1985). Do these results indicate anything more than random fluctuations?

(If each result is treated as being independent, then only the Cancer and Leo results lie outside the 99 per cent confidence limits.)

2.7 Tensile testing of 33 silicon carbide tensile specimens reveals the following tensile strengths (all data in MPa): 310; 330 (2 off); 380 (2 off); 390; 395; 400 (2 off); 405; 410; 430; 435; 470 (2 off); 480; 500; 510 (2 off); 520; 530; 535; 540; 545; 550 (2 off); 560; 570; 590; 600; 620; 700. One specimen had still not failed at 700 MPa.

Using Weibull graph paper, determine the coefficients α and β in the Weibull distribution for this data. Hence determine the expected value (mean) and the standard deviation.

(500 MPa; 8.2; 472 MPa; 67 MPa)

References and bibliography

Billinton R and Allan R N, *Reliability Evaluation of Engineering Systems*, Pitman, London 1983.

Grist D R, *Individual Risk – A Compilation of Recent British Data*, UKAEA Safety and Reliability Directorate Report No 125, HMSO 1978.

McCormick N J, *Reliability and Risk Analysis*, Academic Press, New York 1981.

Shooman M L, *Probabilistic Reliability: An Engineering Approach*, McGraw-Hill, New York 1968.

Weibull W, *J Appl. Mech.* **18** 1951, 293–7.

Woodcock E R and Eames A R, *Confidence Limits for Numbers from 0 to 1200 Based on the Poisson Distribution*, UKAEA Report AHSB(S)R179, HMSO 1970.

Systems reliability

In any engineering operation which is potentially hazardous, for example flying an aeroplane or operating a chemical plant which uses toxic materials, the operation will be controlled by means of control systems. Depending on the severity of the consequences of failure for a given control system, there may be a back-up system which operates in parallel with the operating control system. Thus cars generally have dual braking systems since brake failure may cause a crash, but seldom have dual fuel supply arrangements (petrol tank/fuel pump/carburettor), since failure of the fuel supply merely causes inconvenience, not tragedy.

We may therefore differentiate between a *series* system (such as the fuel supply) where only one component need fail for system failure, and a *parallel* system (such as dual brakes) where both (or all) systems must fail for a system failure. A parallel system is often called a *redundant* system.

3.1 Reliability of time-independent systems

We will first consider systems in which the reliability of the systems components is invariant with time. In other words, the reliability function $R(x)$ (Section 2.4.1) is a constant which will be called R, and

$$R + F = 1 \qquad [3.1]$$

as before, where the failure function $F(x)$ is now also constant with a value F.

The validity of these assumptions will be examined in Section 3.2.

3.1.1 Series and parallel systems

In a series system (Fig. 3.1(a)), every component must work for system success or, conversely, only one component need fail for system failure. The system is therefore non-redundant.

If we let R_A and R_B be the probabilities of the successful operation of sub-systems (or components) A and B respectively, and F_A and F_B be their failure probabilities, we can see that

$$R_A + F_A = 1$$

and

$$R_B + F_B = 1$$

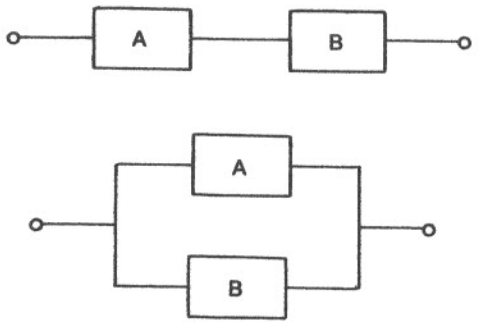


Fig. 3.1 Series and parallel systems

For the series system, both components must be working, so the reliability of the system R_S will be

$$R_S = R_A R_B \qquad [3.2]$$

In general, for an n-component series system, we may write

$$R_S = \prod_{i=1}^{n} R_i \qquad [3.3]$$

This is known as the Product Rule of reliability.
Similarly, the probability of system failure F_S will be

$$F_S = 1 - \prod_{i=1}^{n} R_i \qquad [3.4]$$

It can be seen that the probability of the successful operation of a series system decreases as the number of sub-systems increases.

For a *parallel* system the probabilities of system failure and success are calculated differently. In the system shown in Fig. 3.1(b), both sub-systems A and B must fail for system failure. Thus

$$F_S = F_A F_B \qquad [3.5]$$

For an n-component parallel system, we may write

$$F_S = \prod_{i=1}^{n} F_i \qquad [3.6]$$

and the success probability will be

$$R_S = 1 - \prod_{i=1}^{n} F_i \qquad [3.7]$$

R and F are thus interchanged between series and parallel systems. In general, the probability of the successful operation of a parallel system usually increases as the number of components or sub-systems increases. However some components may have failure modes which *reduce* the overall reliability in a parallel system. This will be discussed more fully in Section 3.1.8.

Using eqns [3.3] and [3.7] it is now possible to produce an overall reliability value for any general series – parallel system (Fig. 3.2) if the sub-system reliabilities are known.

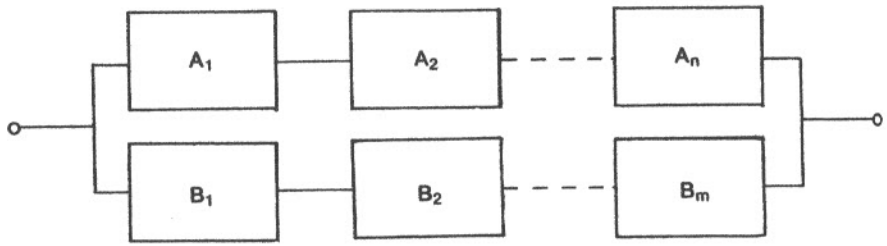


Fig. 3.2 A general series–parallel system

The system in Fig. 3.2 can be reduced to a system like that in Fig. 3.1(b), where

$$R_A = \prod_{i=1}^{n} R_{Ai}$$

and

$$R_B = \prod_{i=1}^{m} R_{Bi}$$

The overall system reliability then becomes, from eqn [3.7],

$$\begin{aligned} R_S &= 1 - F_A F_B \\ &= 1 - (1-R_A)(1-R_B) \\ &= R_A + R_B - R_A R_B \end{aligned} \qquad [3.8]$$

3.1.2 Majority voting and standby systems

Two types of system that are commonly used in aerospace, chemical and nuclear applications are the majority voting (or partially redundant) and standby systems. These are illustrated schematically in Fig. 3.3.

Majority voting systems (Fig. 3.3(a)) are used for high integrity shutdown systems in chemical or nuclear plant. Sub-systems A, B and C might represent temperature or pressure transducers and trip amplifiers. (A trip amplifier is a unit which generates an output signal if the input signal exceeds a preset value.)

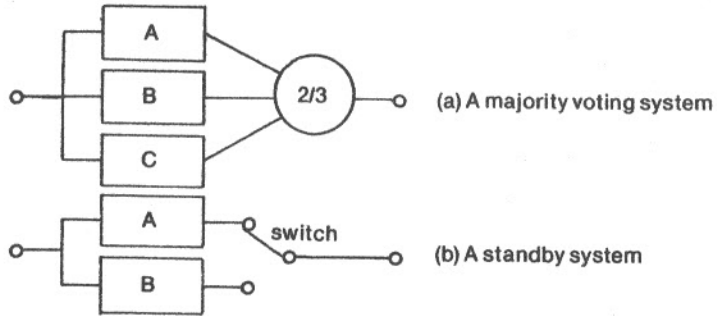


Fig. 3.3 Majority voting and standby systems

Two out of three 'voting' by the three sub-systems determines whether or not a plant shutdown should be initiated. (Depending on application, one out of two or two out of four voting is sometimes employed. Two out of three is the most common, however.) For system success we require at least two sub-systems to operate when required. Therefore the success probability is (from eqns [2.18] and [2.21])

R_S = (Probability of all three sub-systems operating correctly) + (Probability of two sub-systems operating correctly)
$= R_A R_B R_C + (R_A R_B F_C + R_A F_b R_C + F_A R_B R_C)$

If $R_A = R_B = R_C = R$, we get

$$R_S = R^3 + 3R^2F \qquad [3.9]$$

Furthermore, since $(R+F)^3 = 1$ from eqn [3.1], we can deduce that the system failure probability is

$$F_S = 1 - R_S = (R+F)^3 - R^3 - 3R^2F$$
$$= 3RF^2 + F^3 \qquad [3.10]$$

Standby systems (Fig. 3.3(b)) are used whenever sub-system failure might prove costly or dangerous but continuous running of both sub-systems is difficult or wasteful. Thus this arrangment might be used for pumps in process plant, or emergency diesels in nuclear plant, or fuel systems in aircraft. In this system the standby sub-system is only made operational when the operating sub-system fails. (See, for example, Stewart (1974) and Hensley (1968).)

To model such systems, it is necessary to consider the possibility of failure of the switching system which allows the standby sub-system to become operational. We thus replace Fig. 3.3(b) with Fig. 3.4, where sub-system C is the switch.

The probability of system failure F_S will be (from eqns [2.18] and [2.21])

F_S = (Probability of A and B both failing given successful switching) + (Probability of operating sub-system failing *and* a switching failure)
$= F_A F_B R_C + F_A F_C$
$= F_A F_B R_C + F_A(1 - R_C)$
$= F_A - F_A R_C (1 - F_B)$ [3.11]

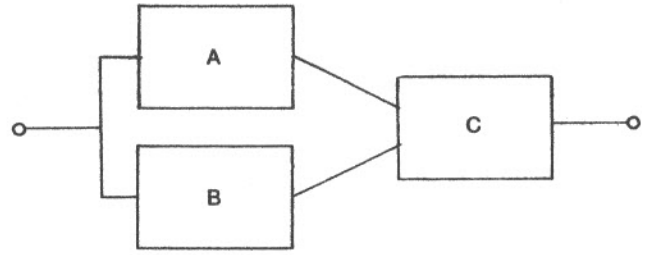


Fig. 3.4 A standby system

Hence the system reliability will be

$$R_S = 1 - F_S = 1 - [F_A - F_A R_C(1 - F_B)] \qquad [3.12]$$

There is also the possibility that the switch might fail in its normal operating mode. If we designate the reliability against such a failure as R^*, the overall system reliability becomes

$$R_S = R^* \{1 - [F_A - F_A R_C(1 - F_B)]\} \qquad [3.13]$$

It is worth pointing out that the failure probabilities of sub-systems A and B in Fig. 3.4 will not, in general, be the same, even if the components or sub-systems are similar. This is because their accumulated operating times may differ. (See Section 3.2).

3.1.3 Bridge networks

A bridge network, as shown in Fig. 3.5, cannot have its reliability assessed using the methods discussed in Sections 3.1.1 and 3.1.2. Assuming that this network must have at least one of the paths A_1A_2, A_1CB_2, B_1B_2 or B_1CA_2 unbroken, then we may write that the system reliability R_S will be, using rule (6) (eqn [2.22])

R_S = (Probability of system success if C is operating) + (Probability of system success if C is not operating)
$= (R_S(\text{C good}).\ R_C) + (R_S(\text{C failed}).\ F_C)$ [3.14]

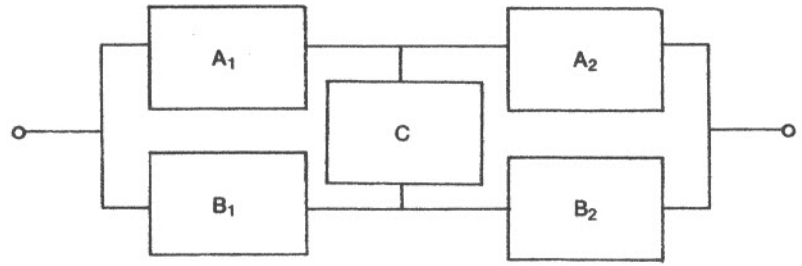
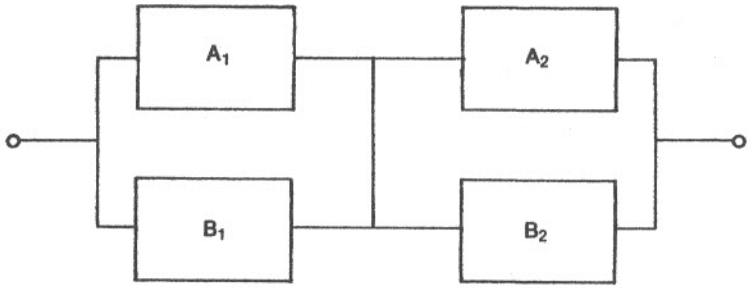
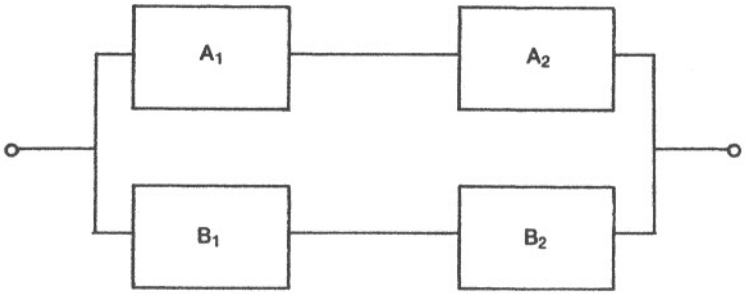


Fig. 3.5 A bridge network



(a) Subsystem C operable



(b) Subsystem C failed

Fig. 3.6 Simplified bridge networks

We must therefore devise expressions for the system reliability when C is operable and failed, respectively. For these two conditions, the network may be drawn as shown in Figs. 3.6(a) and 3.6(b).

For the case when sub-system C is operating, (Fig. 3.6(a)), failure of the system requires that either subsystems A_1 and B_1 should both fail or sub-systems A_2 and B_2 should both fail. Hence

$$R_S(\text{C good}) = (1-F_{A1}F_{B1})(1-F_{A2}F_{B2}) \qquad [3.15]$$

For the case where subsystem C has failed, (Fig. 3.6(b)), system reliability can be determined using the methods of Section 3.1.1, which yield,

$$R_S(\text{C failed}) = 1-(1-R_{A1}R_{A2})(1-R_{B1}R_{B2}) \qquad [3.16]$$

Hence the overall bridge network reliability may be determined from eqns [3.14], [3.15] and [3.16] to be

$$R_S = (1-F_{A1}F_{B1})(1-F_{A2}F_{B2})\,R_C + [1-(1-R_{A1}R_{A2}).(1-R_{B1}R_{B2})]\,F_C \qquad [3.17]$$

If we assume that all sub-systems have the same reliability R, we obtain

$$\begin{aligned} R_S &= (1-F^2)^2R + [1-(1-R^2)^2]F \\ &= (1-F)^2(1+F)^2R + [1-(1-R^2)]\,[1+(1-R^2)]F \\ &= 2R^5-5R^4+2R^3+2R^2 \end{aligned} \qquad [3.18]$$

Thus if $R = 0.98$, the system reliability R_S will be $= 0.99\,918\,484$, i.e. $F_S = 8.152 \times 10^{-4}$.

3.1.4 Cut sets

The foregoing analysis of a simple bridge network was straightforward enough, but cumbersome. Using the probability of conditional events does not lead to a means of readily solving a general system network, preferably by computer. A more generally applicable means of solution is the method of cut sets.

We must first define a cut set. A cut set is *a set of components or sub-systems which, when failed, causes system failure.* A *minimal* cut set is a cut set for which, if any one component or sub-system has *not* failed, then system failure does not occur.

Considering again the bridge network of Fig. 3.5 for which a path between the two terminals must be maintained for system success, we may identify the minimal cut sets to be

A_1B_1, A_1CB_2, B_1CA_2 and A_2B_2

since, if any one sub-system of each minimal cut set has not failed, the system will not fail. The bridge network in Fig. 3.5 may therefore be represented as a series system of minimal cut sets, in which each cut set is represented as a parallel arrangement of sub-systems (Fig. 3.7).

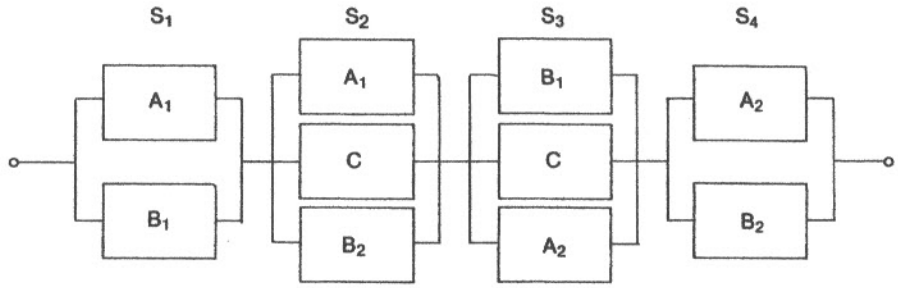


Fig. 3.7 Minimal cut set representation of bridge network

Cut sets S_1 and S_4 are called second-order cut sets, while S_2 and S_3 are called third-order cut sets.

We cannot reduce the system shown in Fig. 3.7 to a simpler system using the concept of series systems, since some sub-systems appear more than once. We can, however, use Venn Diagrams to yield an expression for system failure (Fig. 3.8).

Each circle represents the probability of failure of a minimal cut set. The system failure probability will therefore be the union of the four sets, since the failure of any cut set causes system failure.

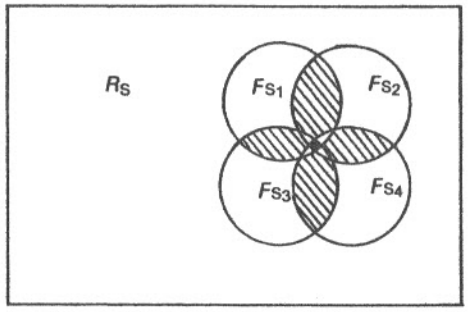


Fig. 3.8 Venn diagrams for cut sets for bridge network

Thus

$$1 - R_S = F_S = F_{S1} \cup F_{S2} \cup F_{S3} \cup F_{S4}$$

From Fig. 3.8, it may be seen that F_S will be given by

$$\begin{aligned} F_S = {} & (F_{S1}+F_{S2}+F_{S3}+F_{S4}) \\ & - (F_{S1} \cap F_{S2}+F_{S2} \cap F_{S3}+F_{S3} \cap F_{S4}+F_{S1} \cap F_{S3}+F_{S2} \cap F_{S4}+F_{S4} \cap F_{S1}) \\ & + (F_{S1} \cap F_{S2} \cap F_{S3}+F_{S1} \cap F_{S2} \cap F_{S4}+F_{S1} \cap F_{S3} \cap F_{S4}+F_{S2} \cap F_{S3} \cap F_{S4}) \\ & - F_{S1} \cap F_{S2} \cap F_{S3} \cap F_{S4} \end{aligned} \qquad [3.19]$$

(*Note:* Equation [3.19] may be checked by adding the number of times that the central shaded area in Fig. 3.8 is counted. This is $(4-6+4-1) = 1$, i.e. the central shaded area is counted once, as it should be.) Furthermore, we can readily show from Fig. 3.7 that

$$F_{S1} = F_{A1}F_{B1}$$

and

$$F_{S1} \cap F_{S2} = F_{S1}F_{S2} = F_{A1}F_{B1}F_{A1}F_CF_{B2}$$

and so on. If we follow this reasoning through, and substitute into eqn [3.19], and further assume that all failure probabilities of the sub-systems in the bridge network are equal, we obtain

$$F_S = 2F^5 - 5F^4 + 2F^3 + 2F^2 \qquad [3.20]$$

where F is the failure probability of each sub-system. Letting $F = 0.02$ (i.e. $R = 0.98$), we obtain $F_S = 8.152 \times 10^{-4}$, which is the same result as was obtained using conditional probabilities in Section 3.1.3.

The cut set method for determining the reliability of complex networks has a major advantage over any method based on the direct application of the laws of probability; namely, that the method can be programmed to calculate the reliability of any general network. One technique for doing this, using an algorithm developed by Allen *et al.* (1976), is as follows:

1. Deduce the *minimal paths* in the network.
   A minimal path is defined to be a path between input and output in which no intersection between branches is traversed more than once. For the bridge network in Fig. 3.5 the minimal paths will be A_1A_2, A_1CB_2, B_1CA_2 and B_1B_2.
2. Prepare an 'incidence matrix' showing the components in each minimal path. For the bridge network this will be:

path	*component*	A_1	A_2	B_1	B_2	C
1		1	1	0	0	0
2		1	0	0	1	1
3		0	1	1	0	1
4		0	0	1	1	0

   Here path 1 is A_1A_2, path 2 is A_1CB_2, path 3 is B_1CA_2 and path 4 is B_1B_2.
3. If none of the elements in a column is zero, then that component is a first-order cut set. There are no first-order cut sets in the simple bridge network.
4. Combine the columns of the matrix two at a time. If none of the elements of a combined column is zero, then that combination of components forms a second-order cut set. Thus A_1B_1 and A_2B_2 are second-order cut sets.
5. Repeat the process combining three columns together at a time, which yields third-order cut sets; in this case, these are A_1CB_2 and B_1CA_2.
6. Repeat for higher-order cut sets until the maximum order has been reached. This technique yields a programmable algorithm for calculating reliabilities of any general network configuration.

The technique also lends itself to approximate methods of reliability analysis of networks. For a complicated network we may neglect high-order cut sets. This greatly reduces calculation effort or CPU time. For the bridge network we may, for example, neglect the two third-order cuts A_1CB_2 and B_1CA_2.

The expression for system failure probability (eqn [3.19]) then reduces to (see Fig. 3.9)

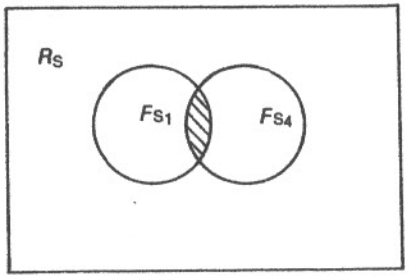


Fig. 3.9 Venn diagram for bridge network neglecting third-order cut sets

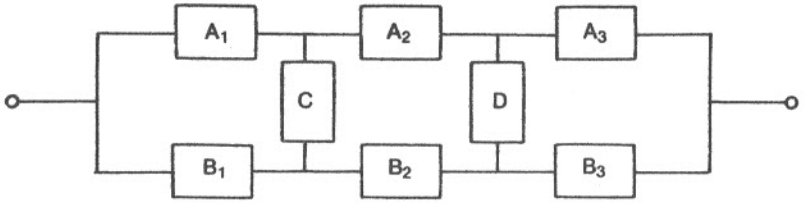


Fig. 3.10 Network of example 3.1

$$
\begin{aligned}
F_S &= F_{S1} \cup F_{S4} \\
&= F_{S1} + F_{S4} - F_{S1} \cap F_{S4} \\
&= F_{A1}F_{B1} + F_{A2}F_{B2} - F_{A1}F_{B2}F_{A1}F_{B2} \\
&= 2F^2 - F^4
\end{aligned}
\qquad [3.21]
$$

Letting $F = 0.02$, we obtain $F_S = 7.9984 \times 10^{-4}$.
This result is very close to the previous value of 8.152×10^{-4} (eqn [3.20]) but was obtained very much more simply.

Example 3.1

To calculate the reliability of the network shown in Fig. 3.10, using the method of cut sets.

The minimal paths are $A_1A_2A_3$, $A_1CB_2DA_3$, $A_1CB_2B_3$, $A_1A_2DB_3$, $B_1B_2B_3$, $B_1CA_2DB_3$, $B_1CA_2A_3$, and $B_1B_2DA_3$. The incidence matrix will be:

Path	*Component*	A_1	A_2	A_3	B_1	B_2	B_3	C	D
1		1	1	1	0	0	0	0	0
2		1	0	1	0	1	0	1	1
3		1	0	0	0	1	1	1	0
4		1	1	0	0	0	1	0	1
5		0	0	0	1	1	1	0	0
6		0	1	0	1	0	1	1	1
7		0	1	1	1	0	0	1	0
8		0	0	1	1	1	0	0	1

There are no first-order cut sets.

A_1B_1, A_2B_2 and A_3B_3 are second-order cut sets. A_1CB_2, B_1CA_2, A_2DB_3 are third-order cut sets. There are no higher-order sets.
Neglecting third-order minimal cut sets gives that

$$F_S = F_{S1} + F_{S2} + F_{S3} - F_{S1} \cap F_{S2} - F_{S1} \cap F_{S3} - F_{S2} \cap F_{S3} + F_{S1} \cap F_{S2} \cap F_{S3}$$

where F_{S1}, F_{S2} and F_{S3} are the failure probabilities of the second order minimal cut sets A_1B_1, A_2B_2 and A_3B_3, respectively. Therefore, assuming that all failure probabilities for the components are equal to F, we obtain

$$F_S = 3F^2 - 3F^4 + F^6$$

which for $F = 0.01$ yields $F_S \fallingdotseq 3 \times 10^{-4}$.

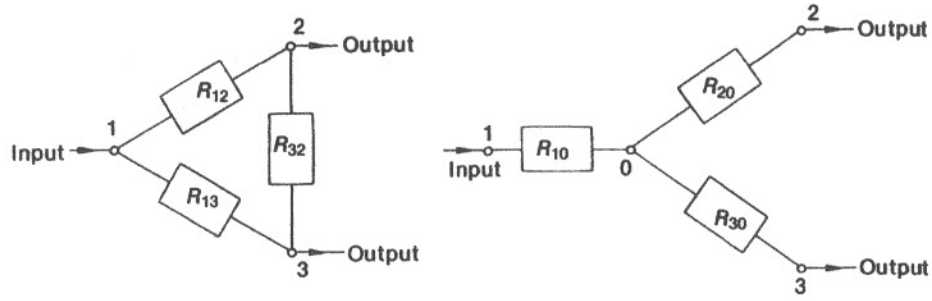


Fig. 3.11 The delta–star transformation

3.1.5 The delta–star transformation

Another way of solving the bridge network problem is to use the delta–star transformation (Fig. 3.11). Gupta and Sharma (1978) give a number of conditions before such a transformation is valid, including the following:

1. The reliability between input 1 and outputs 2 and 3 must be the same before and after the transformation when neither of the outputs are open-circuit (Fig. 3.12(a)).
2. The reliability between input 1 and output 2 must remain the same after the transformation when output 3 is open-circuit. (Fig. 3.12(b)).
3. The reliability between input 1 and output 3 must remain the same after the transformation when output 2 is open-circuit (Fig. 3.12(c)).

These conditions enable equations to be deduced for the reliabilities of the components in the star configuration in terms of the reliabilities of the equivalent components in the delta configuration. Applying the rules for series and parallel systems to Figs. 3.12(a), (b) and (c), we obtain the following equations:

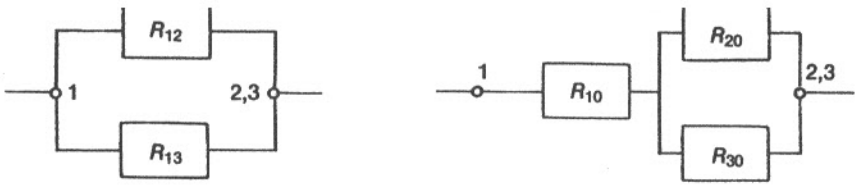
$$R_{10}[1-(1-R_{20})(1-R_{30})] = 1-(1-R_{12})(1-R_{13}) \qquad [3.22]$$
$$R_{10}R_{20} = 1-(1-R_{12})(1-R_{13}R_{32}) \qquad [3.23]$$
$$R_{10}R_{30} = 1-(1-R_{13})(1-R_{12}R_{32}) \qquad [3.24]$$

These equations may be solved thus:

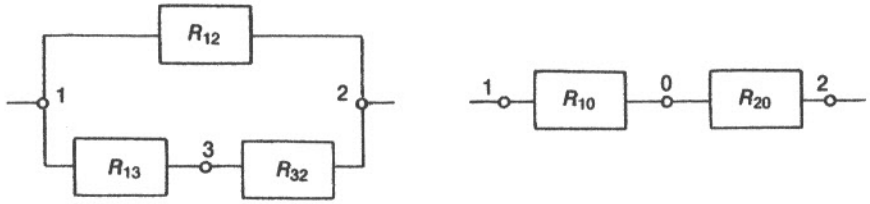
$$R_{10} = r_1r_2/r_3 \qquad [3.25]$$
$$R_{20} = r_3/r_2 \qquad [3.26]$$
$$R_{30} = r_3/r_1 \qquad [3.27]$$

where

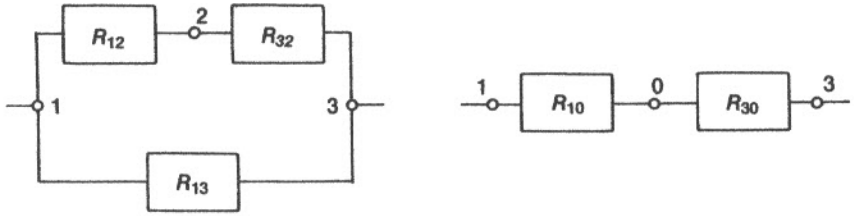
$$r_1 = R_{12} + R_{13}R_{32} - R_{12}R_{13}R_{32} \qquad [3.28]$$
$$r_2 = R_{13} + R_{12}R_{32} - R_{12}R_{13}R_{32} \qquad [3.29]$$
$$r_3 = R_{12}R_{32} + R_{12}R_{13} + R_{13}R_{32} - 2R_{12}R_{13}R_{32} \qquad [3.30]$$



(a) The equivalent transformation on condition (1)



(b) The equivalent transformation for condition (2)



(c) The equivalent transformation for condition (3)

Fig. 3.12 Necessary conditions for valid delta–star transformation

We may now apply these transformation equations to the bridge network shown in Fig. 3.13.

Equations [3.25] to [3.30] may then be applied to calculate R_{10}, R_{20} and R_{30}

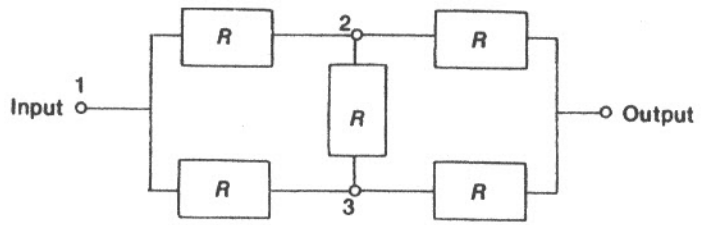
$$R_{10} = (1+R-R^2)^2/(3-2R) \qquad [3.31]$$

$$R_{20} = R_{30} = (3R-2R^2)/(1+R-R^2) \qquad [3.32]$$

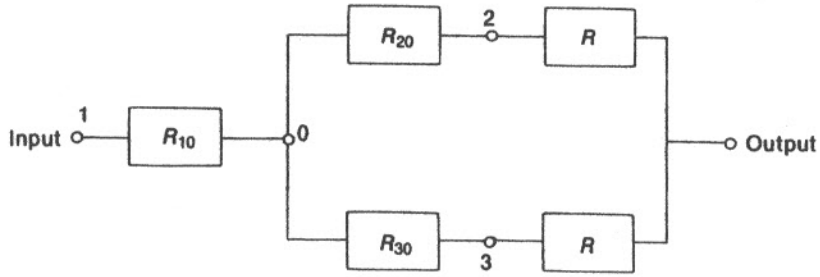
The overall system reliability R_S can now be determined directly from the transformed network using the principles of parallel and series systems, to give:

$$R_S = R_{10}[1-(1-R_{20}R)(1-R_{30}R)]$$
$$= R_{10}R_{20}R + R_{10}R_{30}R - R_{10}R_{20}R_{30}R^2$$

Substituting from eqns [3.31] and [3.32] yields



(a) Bridge network with all component reliabilities equal to R



(b) Transformed network with equivalent reliabilities

Fig. 3.13 Delta–star transformation applied to bridge network

$$R_S = 2R^5 - 5R^4 + 2R^3 + 2R^2 \qquad [3.18]$$

This is the same as the result that was obtained previously using conditional probability or cut sets (eqns [3.18] and [3.20]).

The inverse problem of a star–delta transformation is more difficult, as it requires the solution of three simultaneous cubic equations. Grosh (1983) has discussed this problem.

3.1.6 Review of network analysis techniques

Three methods for analysing any general system network have been presented. These are:

1. Conditional probability (Section 3.1.3), together with the rules for series–parallel systems (Section 3.1.1).
2. Cut sets (Section 3.1.4).

3. Delta–star transformations (Section 3.1.5), together with the rules for series–parallel systems (Section 3.1.1).

Of these three methods, only the method of cut sets is suitable for automated analysis of very complex systems. This method may also be used for approximate analyses of systems, by neglecting high-order cut sets. The methods of conditional probability and the delta–star transformation are really only useful for simple hand calculations.

3.1.7 Fault trees and event trees

Fault trees and event trees are two common ways of portraying failure conditions in engineering systems. In a *fault tree*, the conditions required for an unwanted fault or failure to occur are presented in a logical sequence, starting from the unwanted fault or 'top event'. This is called 'backward logic'. Fault trees are most often used at the early stages of engineering design as a means of qualitatively evaluating system failure modes. The method can also be used for quantitative evaluation of system reliability, however. An example of a fault tree is shown in Fig. 3.14. This diagram could be used as the basis for a quantitative assessment if the probabilities of the various failures were known. The overall system failure probability may then be calculated by combining the individual failure probabilities at the AND and OR gates as shown in Fig. 3.15.

In contrast, an *event tree* uses 'forward logic'. An event tree starts by assuming either a normally-operating system or a given initiating event. By means of binary logic (i.e. failed/not failed) the diagram then charts all possible system states.

There are two categories of event trees; those for continuously-operated systems and those for standby systems. For *continuously-operated systems* we will again consider the bridge network of Fig. 3.5. An event tree for component failure in this network is shown in Fig. 3.16. Of the 2^5 (=32) possible system states, sixteen states cause system failure. In addition, the sum of the probabilities for each of the 32 system states must be unity, i.e.

$$\sum_{i=1}^{n} P_i = 1 \qquad [3.33]$$

where n is the number of system states. The probability P_i of the ith system state will simply be the product of all the component reliabilities and failure probabilities which have led to that system state, e.g. in Fig. 3.16,

$$P_{19} = F_{A1}R_{B1}R_{A2}F_{B2}R_C \qquad [3.34]$$

We may therefore calculate the overall system reliability by adding together all the probabilities for those system states which lead to system success, i.e.

$$R_S = \sum_{i=1}^{6} P_i + \sum_{i=9}^{13} P_i + \sum_{i=17}^{19} P_i + P_{21} + P_{22} \qquad [3.35]$$

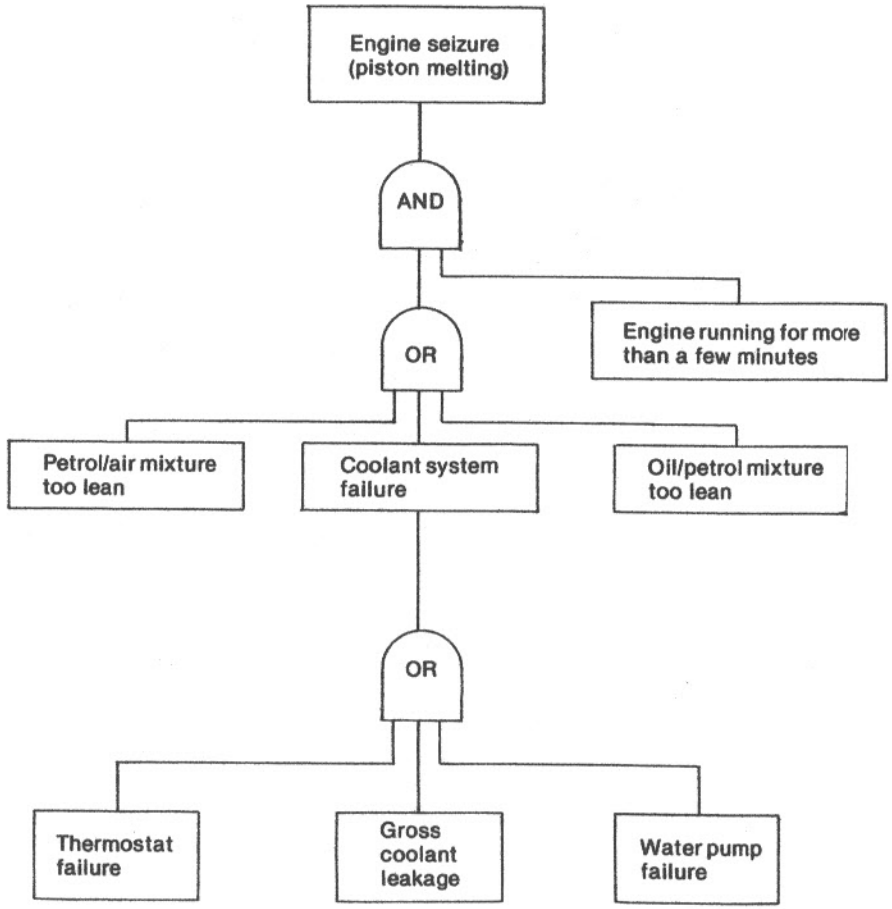


Fig. 3.14 Fault tree for seizure of a water-cooled two-stroke engine

If we assume, as before, that the reliabilities of all components are equal to R, then eqn [3.35] leads to the following expression for R_S

$$R_S = 2R^5 - 5R^4 + 2R^3 + 2R^2 \qquad [3.18]$$

This is the expression that was derived earlier using other methods (conditional probability, cut sets and the delta–star transformation).

For *standby systems* the chronology of events becomes important, since many standby systems are designed to operate only in the event of the failure of another system. A classic example of an event tree is shown in Fig. 3.17, which illustrates a simplified version of the major events in the first few minutes of the accident to the Three Mile Island-2 reactor at Harrisburg, Pennsylvania on the 28th March, 1979. (Kemeny, 1979; Bignell and Fortune, 1984).

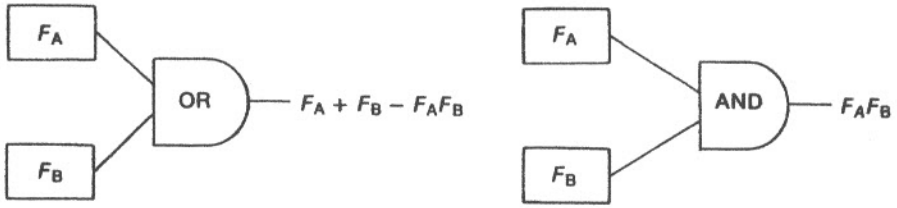


Fig. 3.15 Combining failure probabilities in fault trees

If reliability data are known, then the event tree shown in Fig. 3.17 can be used to determine the probability of the TMI-2 accident sequence occurring. For example, the WASH-1400 Reactor Safety Study (Rasmussen, 1975) (see Table 3.3) quotes the following data:

Pump, failure to run	3×10^{-5}/hour
Pump, failure to start	1×10^{-3}/demand
Relief valves, failure to open	1×10^{-5}/demand

This immediately illustrates a problem with this type of analysis. In the TMI-2 accident, the fault with the standby boiler feed pump was not that it failed to start but that its discharge valve had been accidentally left shut, i.e. human error. This means that the probability of the pump failing to do its job may have been greater than Rasmussen assumed. We shall return to the topic of human error in Section 3.6.

The TMI-2 accident caused a release of 592 GBq (16 Curies) of iodine-131. Rasmussen predicted that an accident of this magnitude would occur approximately every 400 reactor-years. At the time of the TMI-2 accident, approximately that amount of reactor operating experience had been accumulated around the world.

This single result proves little, of course, but it is at least *prima facie* verification of reliability analysis. The accuracy of reliability analysis is considered further in Section 3.4.

To summarise, event trees are commonly used for reliability assessment or operability analysis of engineering plant. Whereas a fault tree is used for qualitative safety or reliability assessment, an event tree gives quantitative reliability figures which may be used at the detailed-design stage. Furthermore, fault trees focus on hazardous outcomes which are anticipated. It is therefore possible in a fault tree to miss trains of events leading to unexpected hazards. An event tree, however, can lead, in principle, to all possible outcomes, although event trees are generally much more complex than fault trees.

3.1.8 Systems with more than one Failure Mode

Until now, we have tacitly assumed that any components or sub-systems can only be in one of two states, operating or failed. However, many systems can fail in more than one way. For example, a diode may fail short-circuit or it may fail open-circuit. A

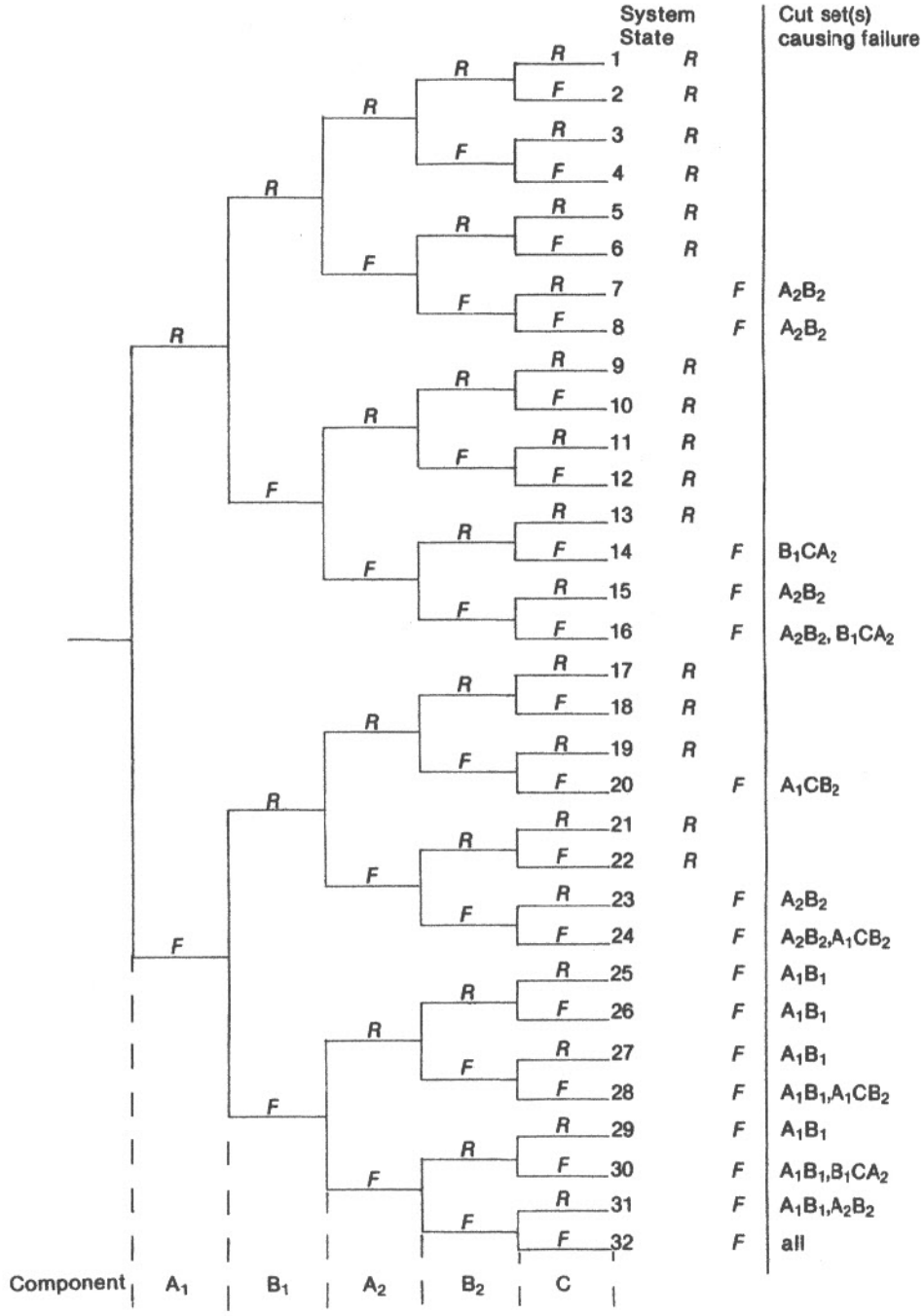


Fig. 3.16 Event tree for failure of the bridge network shown in Fig. 3.5

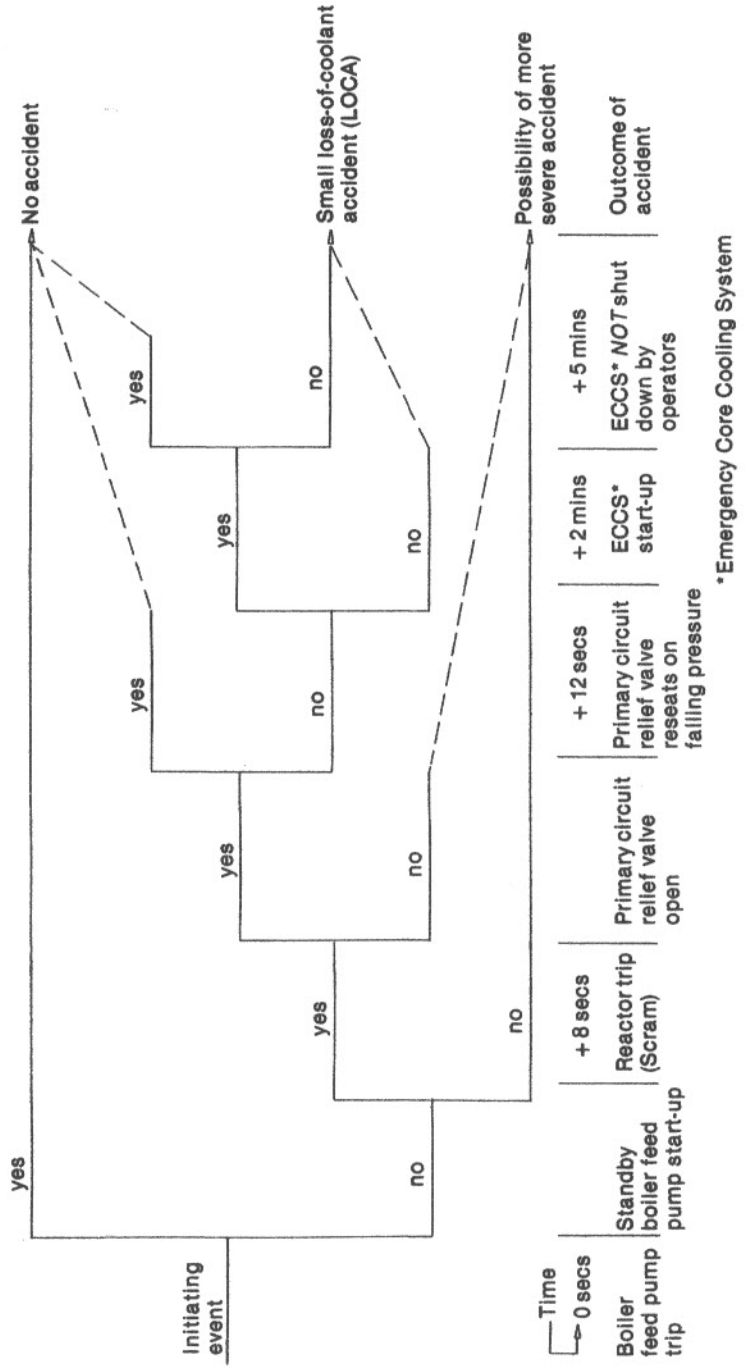


Fig. 3.17 Simplified, reduced event tree illustrating the major events at the start of the TMI-2 accident

designer might suggest using two diodes together to improve upon the reliability of a single diode. However, this raises a problem: should the diodes be connected in series or parallel? If they are connected in series, an open-circuit fault becomes more likely, but a short-circuit fault is less likely. If they are connected in parallel, a short-circuit fault is more likely but an open-circuit fault is less likely. The best configuration will depend upon the relative magnitudes of the failure probabilities in each *failure mode*, open-circuit and short-circuit.

Consider again the voting system introduced in Section 3.1.2. An event tree for such a system, assuming only one failure mode, is shown in Fig. 3.18. Letting the reliability of each detector be $R = 0.98$, then the overall system reliability will depend whether the voting is 1/3, 2/3 or 3/3, according to the binomial distribution $(R + F)^3$.

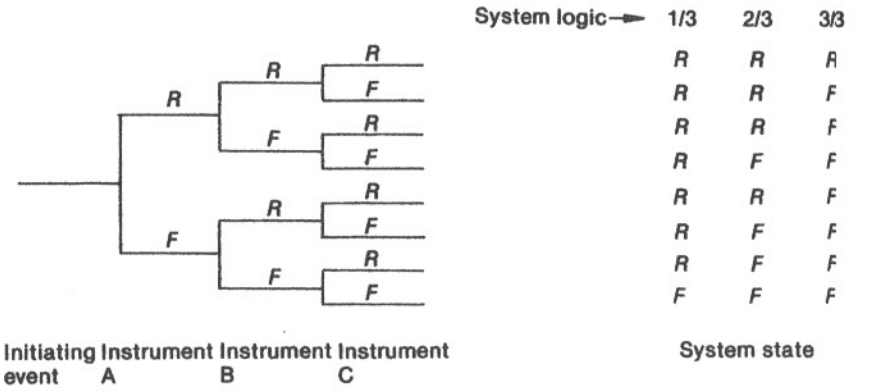


Fig. 3.18 Event tree for a majority voting system

The probability of successful operation of the voting system will therefore be:

Voting logic	*System reliability*	
1/3	$R^3+3R^2F+3RF^2$	= 0.99 999
2/3	R^3+3R^2F	= 0.9988
3/3	R^3	= 0.9412

At first sight this suggests that the 'best' voting logic is 1-out-of-3. However, each detector has *two* failure modes; *failure to operate on demand* and *spurious operation*. If 1-out-of-3 logic is used, the chances of spurious operation of the system may be unacceptably high. 2-out-of-3 voting reduces this probability of spurious operation while maintaining an acceptably high system reliability for operation on demand. 3-out-of-3 voting yields a low rate of spurious operation but a poor system reliability for operation on demand. For these reasons 2-out-of-3 voting logic is most commonly employed.

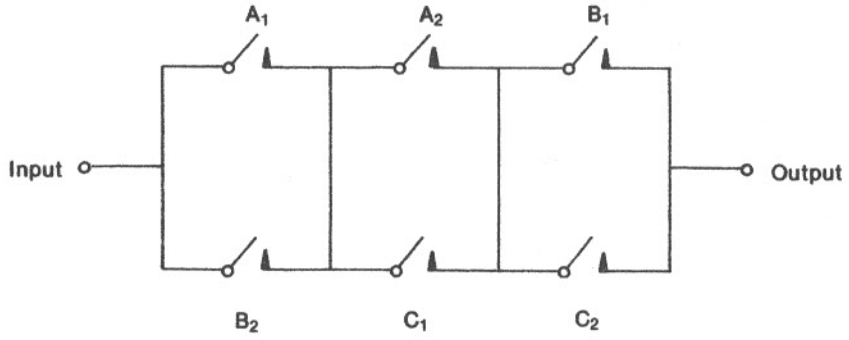


Fig. 3.19 Relay logic for 2-out-of-3 voting system

A typical arrangement for a 2-out-of-3 voting system employing relay logic is shown in Fig. 3.19 (cf. Fig. 3.3(a)). All relay contacts are normally closed.

The two failure modes for a system such as this will be (i) failure to operate caused by relay contacts sticking closed or an electrical or electronic fault, and (ii) spurious operation caused by an electrical or electronic fault. Example 3.2 shows one technique for calculating system reliabilities and failure probabilities when more than one failure mode exists.

Example 3.2

To calculate the probabilities of successful operation, spurious operation and failure to operate on demand for a majority voting system using 2-out-of-3 logic.

Let R be the probability of successful operation,
F_1 be the probability of failure to operate on demand, and
F_2 be the probability of spurious operation.

Since R, F_1 and F_2 are mutually exclusive and complementary their sum will equal unity, i.e.

$$R+F_1+F_2 = 1 \qquad [3.36]$$

In this example, we will let R, F_1 and F_2 be 0.98, 0.01 and 0.01 respectively.

The different possible system states for a three-element voting system will therefore be described by the following expression:

$$(R+F_1+F_2)^3 = R^3+F_1^3+F_2^3+3RF_1^2+3R^2F_1+3R^2F_2+3RF_2^2+3F_1^2F_2+3F_1F_2^2+6RF_1F_2 \qquad [3.37]$$

The terms on the right hand side of eqn [3.37] may be categorised into those terms which imply system success, those terms which imply system failure, and those terms which imply spurious operation, as follows:

Successful operation R_S	*Failure to operate* F_{S_1}	*Spurious operation* F_{S_2}
$R^3 = 0.941\,192$	$F_1^3 = 1 \times 10^{-6}$	$F_2^3 = 1 \times 10^{-6}$
$3R^2F_1 = 0.028\,812$	$3RF_1^2 = 294 \times 10^{-6}$	$3RF_2^2 = 294 \times 10^{-6}$
$3R^2F_2 = 0.028\,812$	$3F_1^2F_2 = 3 \times 10^{-6}$	$3F_1F_2^2 = 3 \times 10^{-6}$
$6RF_1F_2 = 0.000\,588$		
$R_S = 0.999\,404$	$F_{S_1} = 2.98 \times 10^{-4}$	$F_{S_2} = 2.98 \times 10^{-4}$
$0.000\,298$		
$0.000\,298$		
$1.000\,000$		

As one would expect, the system reliability R_S and the system failure probabilities F_{S_1} and F_{S_2} also add up to unity.

A summary table of the reliabilities and failure probabilities for 1-out-of-3, 2-out-of3 and 3-out-of-3 logic is presented in Table 3.1.

Table 3.1 System reliabilities and failure probabilities for 1/3, 2/3 and 3/3 logic voting systems

R	= reliability of logic element	= 0.98
F	= total failure probability of logic element	= 0.02
F_1	= probability of failure to operate *on demand* for individual logic element	= 0.01
F_2	= probability of *spurious* operation of logic element	= 0.01
R_S	= voting system reliability	
F_S	= voting system failure probability	

(a) *1 failure mode*

Logic	R_S		F_S
1/3	$R^3+3R^2F+3RF^2$	$= 0.999\,992$	$F^3 = 8 \times 10^{-6}$
2/3	R^3+3R^2F	$= 0.998\,824$	$3RF^2+F^3 = 1.176 \times 10^{-1}$
3/3	R^3	$= 0.9412$	$3R^2F+3RF^2+F^3 = 0.0588$

(b) *2 failure modes*

Logic	R_S	F_{S_1}	F_{S_2}
1/3	$R^3+3R^2F_1+3RF_1^2$	F_1^3	$F_2^3+3F_1F_2^2+3F_1^2F_2+3R^2F_2+$ $3RF_2^2+6RF_1F_2$
	$= 0.970\,298$	$= 1 \times 10^{-6}$	$= 0.029\,701$
2/3	$R^3+3R^2F_1+3R^2F_2$ $+6RF_1F_2$ $= 0.999\,404$	$F_1^3+3RF_1^2$ $+3F_1^2F_2$ $= 0.000\,298$	$F_2^3+3RF_2^2+3F_1F_2^2$ $= 0.000\,298$
3/3	$R^3+3R^2F_2+3RF_2^2$	$F_1^3+3R^2F_1+$ $3RF_1^2+3F_1^2F_2$ $+3F_1F_2^2+6RF_1F_2$	F_2^3
	$= 0.970\,298$	$= 0.029\,701$	$= 1 \times 10^{-6}$

3.1.9 Failure mode and effect analysis

Often the number of components in a system is just too large to make fault trees or event trees practicable. This may be the case in electronic systems, for example. In that event, a technique called Failure Mode and Effect Analysis may be employed. This technique is far less rigorous than fault tree or event tree methods, and relies for its efficacy upon the skill and experience of the reliability engineer.

In failure mode and effect analysis, the probabilities of each failure mode for each component in the system are listed. The probabilities of failure modes that might lead to the unwanted top event (i.e. dangerous failure modes) are then added and this figure is taken to be the probability of the top event. The method is illustrated in Example 3.3.

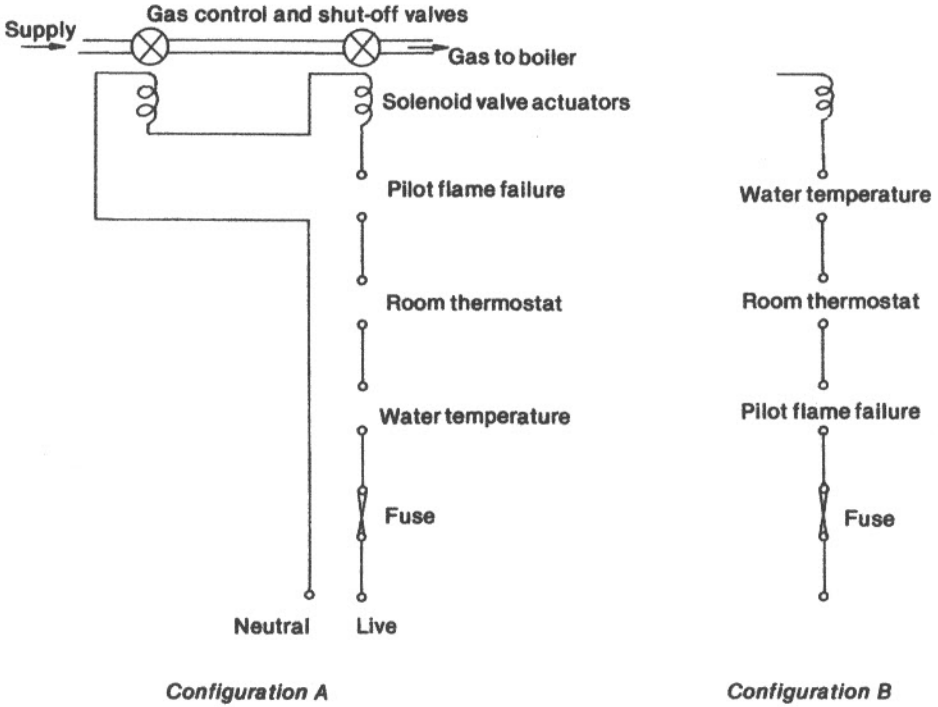


Fig. 3.20 Alternative arrangements for a domestic gas boiler control system

Example 3.3 Failure mode and effect analysis of a domestic gas boiler (Bowen, 1981)

A typical domestic gas boiler control system is shown in Fig. 3.20. What is the most reliable configuration, given that there is a 0.01 probability that the electrician will inadvertantly connect the live and neutral terminals the wrong way round?

The analysis is performed by means of Table 3.2, which shows reliability data for the different failure modes of the system components.

Assuming that there is a finite possibility that the live and neutral terminals will be crossed over, then the analysis suggests that configuration A is preferable. This is because the room thermostat has a high probability (10^{-2}/annum) of developing earth faults. If configuration B was used *and* the terminals had been crossed, and an earth fault subsequently occurred on the room thermostat, then there would be no protection if the pilot flame failed. (If the terminals were not crossed then the fuse would blow as soon as the earth fault developed.)

Table 3.2 Gas boiler failure mode and effect analysis

	Stuck closed	*Stuck open*	*Shorted*	*Open circuit*	*Earth fault*	*Through leakage*	*Probability of dangerous fault*
Gas valve 1	F/S	2×10^{-5}	—	—	—	3×10^{-4}	3.2×10^{-4}
Gas valve 2	F/S	2×10^{-5}	—	—	—	3×10^{-4}	3.2×10^{-4}
Flame failure relay	2×10^{-3}	F/S	2×10^{-3}	F/S	1×10^{-3}	—	5×10^{-3}
Solenoid coils	—	—	F/S	F/S	F/S	—	—
Water temperature switch	0	F/S	0	F/S	0	—	—
Room thermostat	F/S	F/S	F/S	F/S	1×10^{-2}	—	1×10^{-2}
						Total	1.56×10^{-2}

Notes: 1. All failure probabilities are 'per annum'.
2. F/S means 'fail-safe'.
3. A zero indicates insufficient data.

A fault such as this would be 'unrevealed', since there would be no indication that anything was wrong. (Unrevealed faults will be discussed more fully in Section 3.3.) A period of several months might then elapse before the pilot light was blown out by a draught, perhaps. A gas explosion would then very probably ensue.

3.1.10 Hazard and operability studies

A Hazard and Operability ('HazOp') Study is a qualitative, systematic technique used in the chemical industry to examine the operability of a chemical process. Using a full description of a proposed process plant, a design team will question every part of the process. In particular, they will consider the following:

1. What is the design *intention* of the plant?
2. What *deviations* from the design intention could occur?
3. What might *cause* such deviations from the design intention?
4. What would be the *consequences* of such deviations from the design intention? (e.g. toxic or fire hazards, or operating difficulties).

In each part of the plant, the design intention is considered and deviations from that intention are assessed using a set of 'Guide Words'. These are:

NO or NOT – meaning that the design intention is not achieved but nothing else happens.

MORE / LESS – temperatures, pressures or flow rates are changed from the design values.

AS WELL AS – an additional activity occurs together with the design intention.

PART OF – only some of the design intentions are achieved.

REVERSE – the logical opposite of the design intention occurs.

OTHER THAN – something completely different to the design intention happens.

The 'design intentions' might represent process conditions, operator or plant activities, substances, times or locations in the plant. Not all guide words will be applicable to every part of the plant.

To illustrate the HazOp technique, let us consider again the domestic gas boiler (Fig. 3.20). The design intention is that gas shall be supplied to the boiler at a sufficient rate to keep the room temperature at some predetermined level. The guide word 'NO' might therefore mean no gas flow, for which the consequences are obvious – the gas flame and the pilot light will go out. Subsequent restoration of the gas flow would then present a hazard, unless the pilot light is re-ignited.

Each guide word may then be applied in turn, and appropriate conclusions drawn.

The method has weaknesses and strengths. It does not yield quantitative results, but instead attempts to systematise a qualitative approach. It relies on the judgement of the engineer performing the assessment. The extent to which the guide word 'AS WELL AS' is applied will restrict the number of simultaneous faults which can be considered. Its strengths lie in its simplicity and ease of application, since no computer programs or reliability data are required. The method is readily applicable to any process plant.

The technique is, in many ways, complementary to other techniques such as Failure Mode and Effect Analysis and Fault or Event Trees. HazOp is an 'inductive' method, whereas the other methods are 'deductive'.

3.2 *Time-dependent systems*

In Section 3.1 it was assumed that the reliabilities (R) and failure probabilities (F) of components or systems could be treated as constants. The justification for this depends upon the assumption that the components or systems have exponential functions of reliability with time, i.e.

$$R(t) = e^{-\lambda t} \qquad [2.48]$$

but

$$R(t) + F(t) = 1$$

therefore

$$F(t) = 1 - e^{-\lambda t} \qquad [2.49]$$
$$\simeq \lambda t \quad \text{for} \quad \lambda t \text{ small.}$$

Now for an exponential reliability function, the hazard rate is by definition constant. Therefore, if we pick a fixed interval t (say 1 year), we may regard the failure probability F as equal to the hazard rate λ multiplied by the time interval t, provided $\lambda t \ll 1$. This is the assumption that has been made throughout Section 3.1.

3.2.1 Time-dependent effects on reliability

Another question now arises; how valid is the assumption that the reliability function is exponential? To answer this calls for the examination of experimental data for batches of similar articles, showing the variation of hazard rate λ with time. The conventional wisdom states that such experiments yield patterns like that shown in Fig. 3.21, which illustrates the well-known 'bathtub curve'.

This curve can be divided into three sections called Phases I, II and III, respectively. Phase I corresponds to a high rate of failure among recently-manufactured components, due to manufacturing defects. Such failures are called 'wear-in' or 'infant mortality' failures. This is followed by Phase II, in which the hazard rate is fairly constant. This phase represents the 'useful life' of the component. Thereafter, in Phase III, the hazard rate again increases until all components have failed. Failures in Phase III are termed 'wear-out' failures. See also Example 2.7.

The bathtub curve is frequently invoked in discussions on reliability. However it must be used with caution. Williams (1982) has pointed out that there is no evidence to support such a curve for the majority of electronic components, as there is little evidence for a rising hazard rate with old age. Carter (1979) has claimed that some mechanical components can be designed to be intrinsically reliable with a negligible hazard rate throughout their useful lives, and with the hazard rate only becoming measurable at the onset of wear-out. Figure 3.21 demonstrates that, as long as components of a system are within their periods of useful life (Phase II), then the assumption of constant hazard rate λ is valid. As has been shown (eqn [2.48]), a constant hazard rate corresponds to an exponential reliability function $R(t)$.

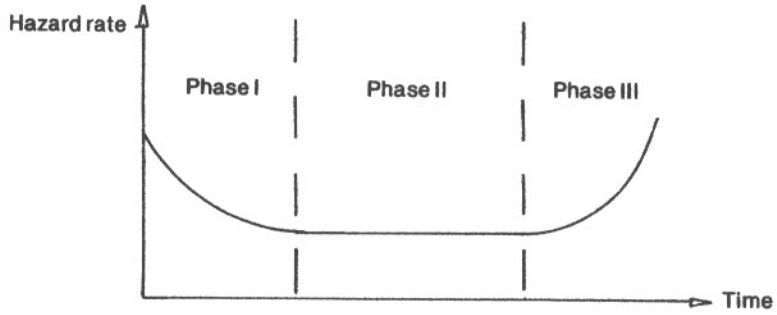


Fig. 3.21 The bathtub curve of reliability

In Section 3.2 we will deal exclusively with exponential reliability functions with constant hazard rate.

3.2.2 Series and parallel systems

Section 3.1.1 introduced these systems. Making the assumption of exponential reliability functions, i.e.

$$R_i(t) = \exp(-\lambda_i t) \qquad [2.48]$$

then eqn [3.3] yields, for an n-component series system, that the system reliability R_S will be

$$R_S = \exp\left(-\sum_{i=1}^{n} \lambda_i t\right) \qquad [3.38]$$

and the system hazard rate λ_S will be

$$\lambda_S = \sum_{i=1}^{n} \lambda_i \qquad [3.39]$$

Similarly, for an n-component *parallel* system, the system failure probability F_S will be, from eqn [3.6],

$$F_S = \prod_{i=1}^{n} (1\text{-}\exp(-\lambda_i t)) \qquad [3.40]$$

A single equivalent hazard rate λ_S cannot be deduced in this case, but if all λ_i's are small (such that $\lambda_i t \ll 1$), eqn [3.40] simplifies to yield

$$F_S \simeq \prod_{i=1}^{n} \lambda_i t \qquad [3.41]$$

3.2.3 Mean Time to Failure (MTTF)

This concept has already been introduced briefly in Chapter 2. The Mean Time to Failure (MTTF) is defined as the expectation $E(t)$ of a failure density function. Hence

$$\text{MTTF} = E(t) = \int_0^\infty t f(t)\,dt \qquad [2.40]$$

$$= -\int_0^\infty t\,dR(t)$$

$$\text{since } f(t) = \frac{-dR(t)}{dt}$$

Integrating by parts yields

$$\text{MTTF} = [-tR(t)]_0^\infty + \int_0^\infty R(t)\,dt \qquad [3.42]$$

However, $tR(t) \to 0$ as $t \to \infty$ (Shooman, 1968), so this expression reduces to

$$\text{MTTF} = \int_0^\infty R(t)\,dt \qquad [3.43]$$

For a *series* system with constant hazard rate, eqn [3.43] yields

$$\text{MTTF} = \int_0^\infty \exp\left(-\sum_{i=1}^{n} \lambda_i t\right) dt$$

$$= \frac{1}{\sum_{i=1}^{n} \lambda_i}$$

This result is consistent with eqn [3.39].
For a parallel system, the corresponding relationship will be

$$\text{MTTF} = \int_0^\infty (1 - F_S(t))dt$$

$$= \int_0^\infty \left[1 - \prod_{i=1}^{n} (1 - \exp(-\lambda_i t))\right] dt \qquad [3.44]$$

For *two* parallel components, eqn [3.44] yields the following expression for Mean Time To Failure:

$$\text{MTTF} = \left(\frac{1}{\lambda_1} + \frac{1}{\lambda_2}\right) - \left(\frac{1}{\lambda_1 + \lambda_2}\right) \qquad [3.45]$$

For *three* parallel components, eqn [3.44] yields

$$\text{MTTF} = \left(\frac{1}{\lambda_1} + \frac{1}{\lambda_2} + \frac{1}{\lambda_3}\right) - \left(\frac{1}{\lambda_1 + \lambda_2} + \frac{1}{\lambda_1 + \lambda_3} + \frac{1}{\lambda_2 + \lambda_3}\right) + \frac{1}{\lambda_1 \lambda_2 \lambda_3} \qquad [3.46]$$

Another commonly-used measure of reliability is the Mean Time Between Failure (MTBF). The MTBF is equal to the MTTF plus the repair time. In many cases the repair time is much less than the MTTF, so MTTF and MTBF are approximately equal. If the repair time is long, however, the two values may differ significantly.

3.2.4 Majority voting and standby systems

For a three-element *voting* system using 2/3 logic (Fig. 3.3(a)) and with one failure mode, the binomial distribution yields

$$[R(t) + F(t)]^3 = R^3(t) + 3R^2(t)F(t) + 3R(t)F^2(t) + F^3(t) \qquad [3.47]$$

System success is assured if two elements are operating. Hence, for constant hazard rate,

$$\begin{aligned} R_S(t) &= R^3(t) + 3R^2(t)F(t) \\ &= e^{-3\lambda t} + 3e^{-2\lambda t}(1-e^{-\lambda t}) \\ &= 3e^{-2\lambda t} - 2e^{-3\lambda t} \end{aligned} \qquad [3.48]$$

The corresponding MTTF will be

$$\text{MTTF} = \int_0^\infty R_S(t)\,dt \qquad [3.43]$$

$$= \int_0^\infty (3e^{-2\lambda t} - 2e^{-3\lambda t})\,dt$$

$$= \frac{5}{6\lambda} \qquad [3.49]$$

For a *standby* system as shown in Fig. 3.4, the reliability $R_S(t)$ will be given by

$R_S(t)$ = (Probability of neither component A nor component B failing) + (Probability of one component failing *and* the switch C operating successfully).

The terms in brackets may be calculated with the aid of the Poisson distribution,

$$P_m(t) = \frac{(\lambda t)^m e^{-\lambda t}}{m!} \qquad [2.69]$$

where m is the number of faults in the system.

Hence,

$$\begin{aligned} R_S(t) &= P_0(t) + P_1(t)\,R_C \\ &= e^{-\lambda t} + \lambda t\, e^{-\lambda t}\, R_C \end{aligned} \qquad [3.50]$$

where the hazard rates for both components A and B are given by λ, and R_C is the reliability *per demand* of the switch C.

Therefore,

$$R_S(t) = e^{-\lambda t}(1 + \lambda t\, R_C) \qquad [3.51]$$

From eqn [3.43], we may then calculate the Mean Time To Failure for this system:

$$\text{MTTF} = \int_0^\infty R_S(t)\,dt$$

$$= \frac{1 + R_C}{\lambda} \qquad [3.52]$$

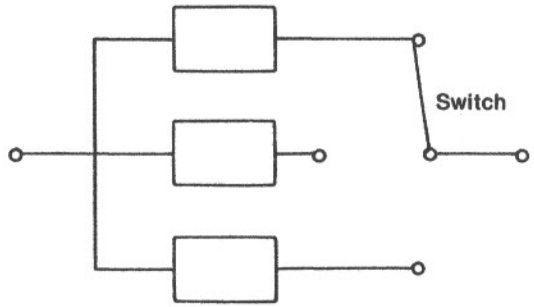


Fig. 3.22 A three-way standby system

Example 3.4 To calculate the MTTF for a 3 × 100 per cent standby diesel system

A standby power supply system employing three diesel generators is illustrated in Fig. 3.22. Each diesel is able to supply the full system load. Calculate the MTTF for the system given the hazard rate of each diesel unit is λ and the switch reliability is R_C (per demand).

In this case the system reliability will be given by

$R_S(t)$ = (probability of no failures) + (probability of one diesel failure *and* successful switching) + (probability of two diesel failures *and* two successful switching operations).

$$= P_0(t) + P_1(t).R_C + P_2(t).R_C^2$$

From the Poisson distribution, we get

$$P_0(t) = e^{-\lambda t} \qquad [3.54]$$

$$P_1(t) = \lambda t\, e^{-\lambda t} \qquad [3.55]$$

$$P_2(t) = \frac{(\lambda t)^2}{2} e^{-\lambda t} \qquad [3.56]$$

Hence,

$$R_S(t) = e^{-\lambda t}\left\{1 + R_C\left[\lambda t + R_C\frac{(\lambda t)^2}{2}\right]\right\} \qquad [3.57]$$

and

$$\text{MTTF} = \int_0^\infty R_S(t)\,dt$$

Integrating by parts yields

$$\text{MTTF} = \frac{1+R_C+R^2}{\lambda} \qquad [3.58]$$

3.3 *Unrevealed faults and frequency of testing*

A fault in a component or sub-system is termed *unrevealed* if that fault is not apparent unless the system is called upon to operate. For majority voting or standby systems this may be a dangerous failure mode.

If we consider again the voting system of Fig. 3.3(a), and assume there is an unrevealed failure mode which causes system failure and has a hazard rate λ_u. The failure probability in this mode will be given by

$$F_u(t) = 1 - e^{-\lambda_u t}$$

When the system is given a routine test and checked to be working correctly, $F_u(t)$ may be taken to have been restored to zero, i.e. $t = 0$. Hence, if there is a maximum permissible failure probability $F_{u\,max}$, we can specify the interval between tests T such that $F_u(t)$ does not exceed $F_{u\,max}$ (see Fig. 3.23).

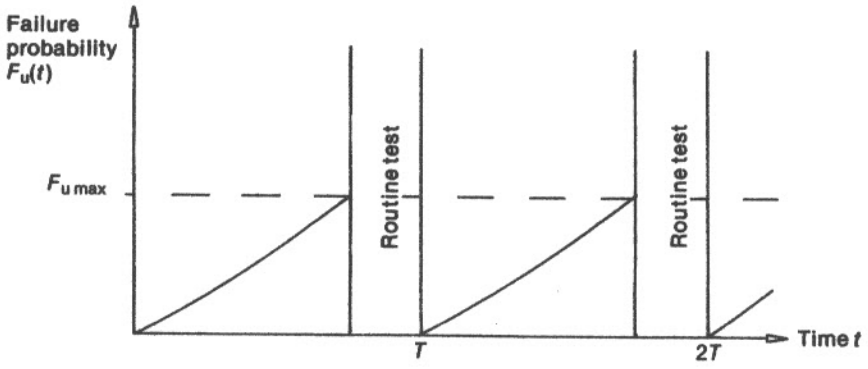


Fig. 3.23 System failure probability with routine testing

Say $F_{u\,max}$ is 10^{-3} and λ_u is 10^{-2} faults per year then

$$10^{-3} = 1 - e^{-10^{-2}T}$$

i.e. $T = 0.1$ year

i.e. test frequency = $1/T \fallingdotseq$ once per month.

For this analysis to be accurate, the time required to test the component must be much less than the test interval. If this is not so and, say, the time taken to test the system is 1 day, then the system unavailability due to maintenance (12 days per year or 0.33×10^{-2}) is larger than the probable unavailability due to breakdowns ($\sim 0.5F_{u\,max}$ or 0.5×10^{-3}). This difficulty is illustrated further in Example 3.5.

Example 3.5 The reliability of a single standby diesel generator: testing frequency and repair time

The block diagram for a three-way standby system (Fig. 3.22) is a gross oversimplification of the system required for a standby diesel-generator supply. A more realistic representation, showing a standby electrical supply system with only one diesel generator, is shown in Fig. 3.24.

A system like this remains idle for most of the time. To minimise the probability of unrevealed faults developing, the system requires regular testing. Most of the usage that standby diesel systems see is due to such testing. Excessive testing, however, will itself induce faults. The repair of these faults will make the machine unavailable. The object of this example, therefore, is to determine the optimum testing frequency to maximise the availability of the system.

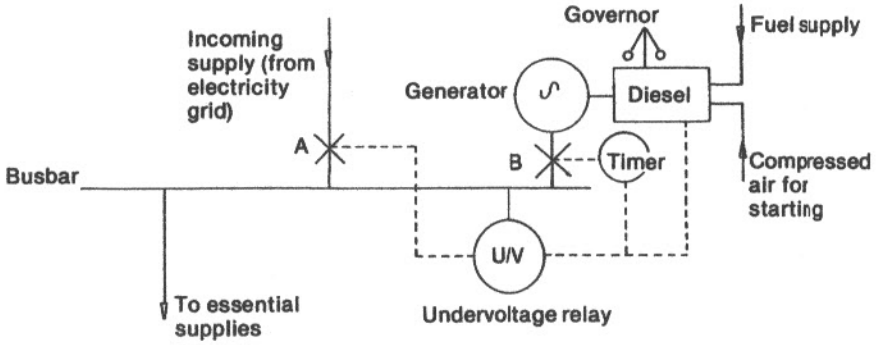


Fig. 3.24 Standby diesel-generator control arrangement

The system in Fig. 3.24 operates as follows:

1. On loss of supply, the undervoltage relay closes, sending a signal to open breaker A and to start up the diesel generator.
2. A timer operates (this allows time for the generator to reach the correct speed) before breaker B closes, thus re-establishing the electrical supply. There is also an interlock preventing breakers A and B both being closed. (If breakers A and B were both closed and the mains electricity supply was re-established, the possible asynchronicity between the generators could cause damage. This interlock may be overriden, and the generator manually synchronised to mains frequency, for test purposes.)

A block diagram for the reliability of this system is shown in Fig. 3.25. In this system, the failure of any component causes the system to fail. For such a series system, the system reliability is given by

$$R_S = \prod_{i=1}^{n} R_i \qquad [3.3]$$

Typical data for such a system might be as follows:

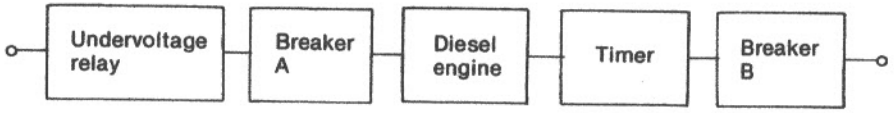


Fig. 3.25 Block diagram for a standby diesel generator

Component	*Failure probability* $F_{i(=1-R_i)}$
System demand frequency (i.e. the frequency of failure of mains supply)	10^{-1}/annum
Undervoltage relay	10^{-4}/demand
Timer	($\lambda_i = 3.5 \times 10^{-2}$/annum)
Breakers	10^{-3}/demand
Diesel engine	10^{-2}/demand

In this example, only one component (the timer) is shown to have a time-dependent failure probability. Other components are shown to have demand-dependent failure probabilities. This is a simplification for illustrative purposes.

The 'system demand frequency' in this example will be the frequency with which mains supplies are lost. An equivalent demand-dependent failure probability for the timer may be calculated by dividing its time-dependent failure rate by the demand frequency. This gives a value for F_i for the timer equal to 0.35/demand. This figure would, in many applications, be unacceptably high. This illustrates the need for regular testing which, in effect, increases the demand frequency.

In the absence of regular testing, the above data suggest that the system reliability R_S [$\Pi_{i=1}^{n}(1-F_i)$] would be approximately 0.64/demand. However, if the system is subjected to routine testing, eqn [3.3] gives that the system reliability will be given by

$$R_S = R_S^1\left(1 - \frac{\lambda}{f}\right)$$

where R_S^1 is the product of the demand-dependent reliabilities, λ is the hazard rate of the timer and f is the demand (test) frequency.

Hence, the more often the system is tested, the better the reliability. If the test frequency becomes too high, however, the number of failures per unit time ($f(1-R_S)$) increases accordingly. Each failure will necessitate a repair, and during the repair interval Δt the generator will not be available. The fraction of time that the system is unavailable will be given by $f(1-R_S)\Delta t$, the 'unavailability' or fractional outage time.

(We should note, in passing, that excessive testing increases the maintenance burden in two ways: the manpower required for carrying out the tests and the manpower required for repairing 'test-induced' faults.)

An optimum test frequency, from the viewpoint of overall failure probability per demand, can now be determined. If the unavailability, $f(1-R_S)\Delta t$, which represents a failure probability per demand due to test-induced repair, is added to the test failure probability per demand, $(1-R_S)$, an overall failure probability per demand, F is obtained. The minimum value for F (see Fig. 3.26) corresponds to the optimum test frequency f^*. The value for f^* will depend upon assumptions made about the repair time Δt, as the diagram shows. In this example, a repair time of 3.6 days corresponds to an optimum test interval of about 3 weeks. A repair time of 9 hours gives an optimum test interval of about 1 week.

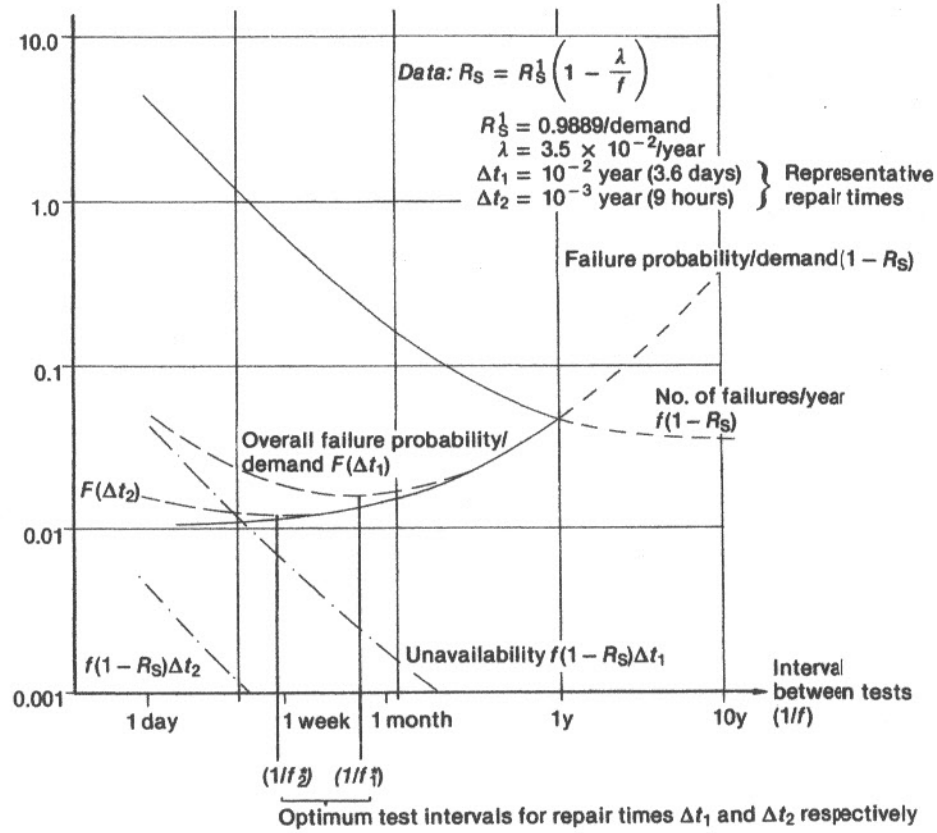


Fig. 3.26 (Example 3.5) Illustrating how the optimum test frequency f^* can be determined from the system reliability R_S and the repair time Δt

3.4 *Reliability data and the accuracy of reliability analysis*

It is readily seen that reliability analysis depends crucially upon experimental reliability data for its accuracy. A continuously updated databank of reliability data is a prerequisite for accurate reliability analysis. Such a databank is maintained by the United Kingdom Atomic Energy Authority's Systems Reliability Service (SRS). Associate companies to the SRS contribute reliability data and have access to the databank. These companies include (Green, 1975) the Central Electricity Generating Board, ICI, Shell, British Petroleum, the European Space Agency, the Civil Aviation Authority and the United States Atomic Energy Commission. A summary list of data has been published by Green and Bourne (1972), and the accuracy of reliability analysis is best judged by a graph published by Green (1970) and reproduced here (Fig. 3.27), which shows the correlation between observed and predicted system failure rates. In this diagram, one hundred per cent accuracy would yield a horizontal line through $Y=1$.

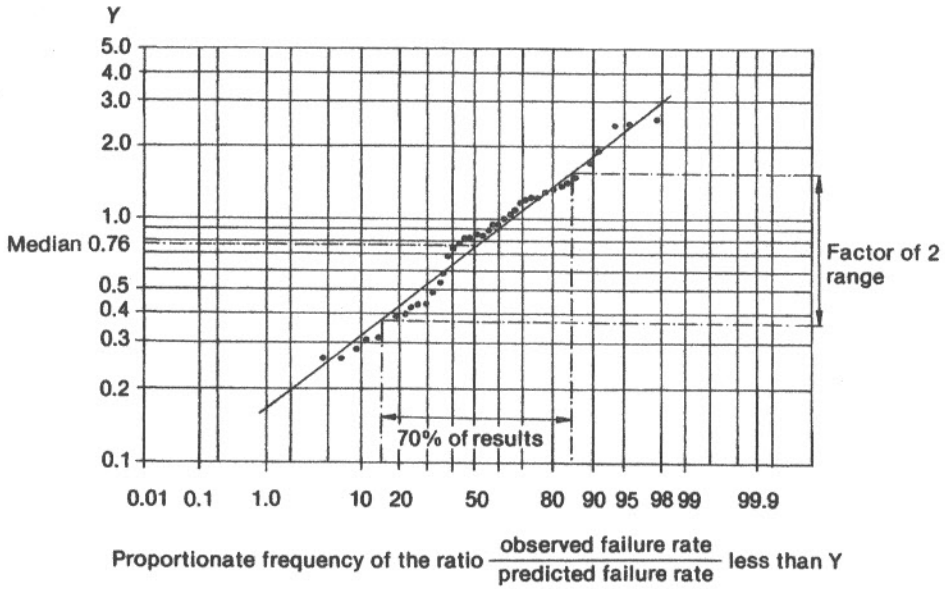


Fig. 3.27 The accuracy of reliability analysis. One hundred per cent accuracy would give a horizontal line through $Y = 1.0$ (Green, 1970; reprinted by permission of the Council of the Institution of Mechanical Engineers)

The results show that the median result was $Y=0.76$, and 70 per cent of results were within a factor of two of this.

Table 3.3 shows a list of typical failure data for mechanical and electrical hardware.

Table 3.3 Failure data for mechanical and electrical hardware (Rasmussen, 1975) (median results)

Components	*Failure mode*			
	Failure to operate on demand F (per demand)	Failure to remain operating $\lambda(hr^{-1})$	Internal leak/ short cct $\lambda(hr^{-1})$	Rupture/ earth-fault (hr^{-1})
Pump	1×10^{-3}	3×10^{-5}	—	—
Valve (motor driven)	1×10^{-3}	3×10^{-7}	—	1×10^{-8}
Valve (solenoid)	1×10^{-3}	—	—	1×10^{-8}
Valve (pneumatic)	3×10^{-4}	3×10^{-7}		1×10^{-8}
Stop valve	1×10^{-4}	—	3×10^{-7}	1×10^{-8}
Vacuum valve	3×10^{-5}	—	—	—
Manual valve	1×10^{-4}	—	—	1×10^{-8}
Relief valve	1×10^{-5}	1×10^{-5} (premature lift)	—	—
Test valves, flowmeters, orifices	3×10^{-4} (blockage)			1×10^{-8}
Pipes < 7.5 cm dia.	—	—	—	1×10^{-9}
Pipes > 7.5 cm dia.	—	—	—	1×10^{-10}
Clutch, mechanical	3×10^{-4}	—	—	—
Clutch, electrical	3×10^{-4}	1×10^{-6} (premature disengagement)	—	—
Relays	1×10^{-4}	1×10^{-7}	1×10^{-8}	1×10^{-8}
Circuit breakers	1×10^{-3}	1×10^{-6}	—	—
Limit switches	3×10^{-4}	1×10^{-7} (normally open) 3×10^{-8} (normally closed)	1×10^{-8}	
Torque switches	1×10^{-4}			
Pressure switches	1×10^{-4}			
Manual switches	1×10^{-5}			
Battery power supply	—	3×10^{-6}	—	—
Transformers	—	1×10^{-6}	1×10^{-6} (primary → secondary)	—
High-power solid-state devices	—	1×10^{-6}	1×10^{-7}	—
Diesel engines	—	3×10^{-4}	—	—
Diesel power supplies	3×10^{-2}	3×10^{-3}	—	—
General instrumentation	—	1×10^{-6}	—	—
Fuses	1×10^{-5}	1×10^{-6} (premature rupture)	—	—
Wires (typical cct.)	—	3×10^{-6}	1×10^{-8}	3×10^{-7}
Terminal boards	—	1×10^{-7}	1×10^{-8}	—

Note: A more up-to-date list of data is given in US MIL HDBK 217D, *Reliability Prediction of Electronic Equipment* (1982).

3.5 Common-mode failures

Edwards and Watson (1979) have defined common-mode failures thus:

> "A common-mode failure is the result of an event which, because of dependencies, causes a coincidence of failure states of components in two or more separate channels of a redundancy system, leading to the defined system failing to perform its intended function."

For example, a chemical plant shutdown system might be *diverse* in that the system monitors two independent parameters, such as temperature and pressure, and also *redundant* if the system relies on majority voting. However, it is conceivable that some other factor might cause all these systems to fail dangerously and simultaneously. The 'other factor' might be, say, an electrical failure or a fire.

Some examples of common-mode failures have been cited by Bignell and Fortune (1984). The capsize of the 'Alexander Kielland' rig in 1980 caused 123 deaths. The death toll was higher than it might have been because the initial tilt of the rig knocked out all normal and emergency electrical supplies, plunging the rig into darkness and making escape less easy.

The Turkish Airlines' DC-10 crash outside Paris in 1974 was caused by the door of the baggage hold, which was underneath the passenger cabin, opening at altitude. This caused the baggage hold to depressurise, which in turn caused the collapse of the cabin floor. The floor carried all control lines underneath it, so when it collapsed control of the aircraft was lost and a crash ensued in which all 346 occupants died.

Edwards and Watson (1979) have made recommendations for procedures to prevent common-mode failures.
These include:

1. Awareness of the possibility of common-mode failure at the design stage.
2. Functional diversity (e.g. monitoring different parameters in a shutdown system).
3. Equipment diversity (e.g. using different makes of transducer – preferably with different operating principles).
4. Fail-safe design wherever possible.
5. Physical protection and segregation of equipment.
6. The use of proven designs wherever possible.
7. The operation of equipment such that it is derated well below its maximum operating parameters.
8. Good quality control.
9. Proven operating and maintenance procedures.
10. Periodic proof testing of systems.

The same authors give data on common-mode failures in sub-systems of nuclear plant and aircraft. Nuclear plant sub-systems show an average common-mode failure rate of 3.3×10^{-2} per sub-system year. This includes failures which were not dangerous.

Aircraft accidents caused by common-mode failures were analysed, and separated into faults in individual sub-systems, as follows:

Sub-systems	*Percentage of total common-mode failure accidents*
Engines	41
Flight control system	21
Fuel system	14
Landing gear	12
Hydraulics	6

The overall accident rate due to common-mode failure is 6×10^{-4} per aircraft year.

The significance of common-mode failure lies in the fact that there is little point in having a high integrity, high reliability control system if common-mode routes with a relatively high probability exist. Overall system reliability with common-mode failure will be given by

$$R_S'(t) = e^{-\lambda_c t} R_S(t) \qquad [3.59]$$

where $R_S(t)$ is the system reliability without common-mode failure, and λ_c is the hazard rate for common-mode failure.

In order for the contribution from common-mode failures to be negligible, λ_c must be much less than the system hazard rate for 'normal' failure routes.

3.6 Human reliability

The subject of human reliability is a large one; a review of the subject by Embrey (1976) cites 146 references. Bignell and Fortune (1984) discuss some interesting case studies. From the viewpoint of the reliability engineer, however, the subject is clouded by an excess of data. Ideally, we would like to be able to represent a human being as a box in a block diagram with a given reliability. However the reliability of people in carrying out given operations varies with factors such as stress level, type of operation or time of day (e.g. night shift). These variables are difficult to quantify but qualitatively the relationship between human error rate and stress level is given in Fig. 3.28. The reliability with which a person carries out an operation is reduced if the person is bored or else over-excited.

Quantitatively, Embrey quotes data giving human failure probabilities in the range 1 to 5×10^{-3} per operation. Ireson (1966) gives human failure probabilities, for various operations, in the range 10^{-2} to 10^{-4} per operation, which is in general agreement with Embrey's data.

In emergency situations, Embrey quotes data suggesting that human reliability deteriorates markedly. Data from aircrashes indicates that human error rates (failure probabilities per operation) increase to at least 16 per cent and a figure of 25 per cent has been proposed for nuclear reactor safety studies.

The importance of human error in the maintenance and modification of industrial plant has been emphasised by the disasters at Flixborough and Bhopal. The Flixborough explosion in 1974 was caused by an inadequate temporary repair to plant containing cyclohexane at pressure (Department of Employment, 1975). The Bhopal tragedy in 1984 was caused by, *inter alia*, a temporary modification which allowed

water into a tank containing methyl isocyanate (MacKenzie, 1985)*. Many other industrial accidents are at least partly caused by human error, and human fallibility is, alas, likely to remain a major cause of such accidents.

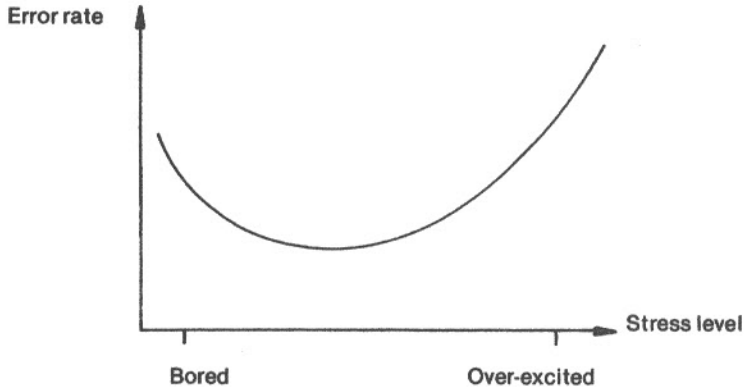


Fig. 3.28 Human error rates (or failure probabilities per operation) as a function of stress level

3.7 Software reliability

In modern engineering systems, real-time digital control systems and computer display or data retrieval systems are commonplace. Although great effort may be spent on trying to ensure that the computer software is free from errors, some residual errors ('bugs') often persist. If such software is involved in process control of potentially hazardous systems, the removal of these bugs assumes some considerable importance.

Shooman (1979) has proposed a model for software reliability which assumes that the average rate at which software bugs are found and removed from similar programs is approximately constant. The software failure rate will then be proportional to the number of remaining bugs.

Hence, if it is assumed that no new bugs are created while debugging is in progress, and that all detected errors are corrected, we may write

$$\varepsilon_r(\tau) = \left(\frac{E_T}{I_T}\right) - \varepsilon_c(\tau) \qquad [3.60]$$

where

τ is the debugging time,
ε_r is the fractional number of residual bugs,
ε_c is the fractional number of corrected bugs,
E_T is the total number of errors, and
I_T is the total number of instructions.

* Claims have been made that this 'temporary modification' was, in fact, sabotage.

Let us further assume that the hazard rate λ, after a period of debugging τ, will thereafter remain constant and proportional to the number of residual errors $\varepsilon_r(\tau)$. From eqn [2.56], the *a priori* failure probability for the program during the program operating time interval Δt may be writen as

$$F(\Delta t) = 1 - e^{-\lambda \Delta t} \qquad [2.56]$$
$$\simeq \lambda \Delta t$$
$$= K \varepsilon_r(\tau) \Delta t \qquad [3.61]$$

where K is a constant of proportionality.

Equations [3.60] and [3.61] are thus the basis of Shooman's model. His experimental findings suggest that the value for (E_T/I_T) is approximately constant and lies in the range 1.0 to 3.0×10^{-2}, and that the fractional number of corrected bugs ε_c may be described by

$$\varepsilon_c(\tau) = \rho\tau \qquad [3.62]$$

where ρ is the fractional rate at which errors are removed and the debugging time τ is in months. A value for ρ in the range 1.0 to 3.0×10^{-3} per month is suggested.

Shooman's model represents an attempt to quantify the reliability of software. However, this approach has its limitations; the constant of proportionality K must be calculated for each individual program by functional testing of the program, since K is dependent on the dynamic structure of the program and the degree to which faults are data-dependent

Musa (1982) has reviewed the various models for software reliability which have been proposed, of which Shooman's model is but one. The technique is still on its 'learning curve', but data and experience are accumulating.

3.8 Conclusions

In principle, the reliability of any engineering system can be determined if data on the reliabilities of the individual components are known. The accuracy of such reliability analysis is quite good; observed systems failure rates seldom differ from predicted failure rates by more than a factor of three (Fig. 3.27). For such analyses, the individual component hazard rates λ are usually taken to be invariant with time, although evidence suggests this is not the case near the beginning or end of a component's life (Fig. 3.21).

There are a number of other factors, however, which make life more difficult for the reliability engineer. These include unrevealed faults, common-mode faults, software faults (in engineering plant which is computer controlled or monitored) and human factors.

Questions

3.1 The failure probability of a jet engine on an airliner on a transatlantic journey is 0.01. Determine the probability of failure for

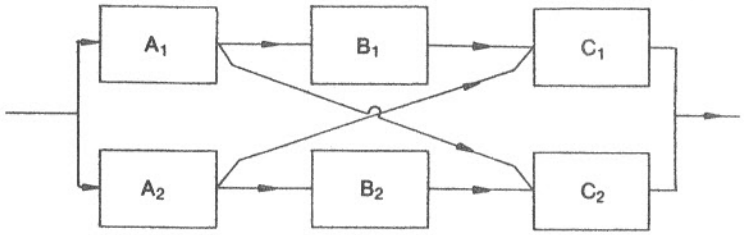
(a) a three-engine airliner which requires a minimum of two engines, and
(b) a four-engine airliner which requires a minimum of two engines, but both operating engines must not be on the same wing.

Draw the reliability block diagram for the latter system.

(3×10^{-4}; 2×10^{-4})

3.2 Use the delta–star transformation to determine the reliability of the network used in Example 3.1.

3.3 Use the method of cut sets to determine the reliability of the system show below.



The component reliabilities are R_A, R_B and R_C.

3.4 A hypothetical electrical grid system consists of 9 identical operating stations, one station on running standby (i.e. available at short notice) and one station in reserve.

If the average frequency of unplanned outage is 3 per station per year and the mean down-time per outage is 3 days, draw a fault tree for the grid control engineer to have to 'shed load', i.e. to cause localised black-out. The reserve station can be taken to 'running standby' state in eight hours. How often is load-shedding likely to be required? Why is this model too simplistic?

(1.21 per annum)

3.5 Using the guidelines given in Section 3.1.10, carry out a HazOp study on the gas boiler system shown in Fig. 3.20 (configuration A).

3.6 Draw fault trees for the condition 'gas valve open but pilot flame not lit' for both the gas boiler systems shown in Fig. 3.20. Which failure modes are unrevealed? Hence estimate the probability of a dangerous failure for each configuration. (Assume a frequency of pilot flame failure of 0.1 per annum.) Does this yield a similar conclusion to that given in Section 3.1.9?

Suggest simple changes to the design which might improve safety. Can you identify any common failure modes?

3.7 Starting from the definition of MTTF, derive an expression for the MTTF of a three-component parallel system (eqn [3.46]).

3.8 Calculate the system reliability and the MTTF for a 2 × 100 per cent standby diesel system.

3.9 For the single standby diesel system in Example 3.5, determine the optimum test frequency if the repair time is assumed to be five days. A graphical approach is recommended.

References and bibliography

Allan R N, Billinton R and De Oliveira M F, *IEEE Tans. Reliability* **R-25**, 1976, 226–33.

Bignell V and Fortune J, *Understanding System Failures*, Manchester University Press 1984.

Billinton R and Allan R N, *Reliability Evaluation of Engineering Systems*, Pitman, London 1983.

Bowen J H, Lecture notes on Risk Management, Reactor Safety Course, Harwell 1981.

Carter A D S, *Proc. I. Mech. Eng.* **193**, 1979, 81–92.

Department of Employment, *Flixborough Disaster – report of the Court of Inquiry*, HMSO 1975.

Edwards G T and Watson I A, *UKAEA Safety and Reliability Directorate Report R146*, 1979.

Embrey P E, *National Centre for Systems Reliability Report R10*, 1976.

Green A E, *Proc. I. Mech. Eng.* **184**(3B), 1970, 17–24.

Green A E, *Appendix to WASH-1400*, NRC, Maryland 1975.

Green A E and Bourne A J, *Reliability Technology*, Wiley Interscience 1972.

Grosh D L, *IEEE Trans. Reliability* **R-32**, 1983, 391–4.

Gupta H and Sharma J, *IEEE Trans. Reliability* **R-27**, 1978, 212–14.

Hensley G, *J. Inst. Meas. Cont.* **1**(4), 1968.

Ireson W G, *Reliability Handbook*, McGraw-Hill 1966.

Kemeny Commission, *Report of the President's Commission on the Accident at Three Mile Island*, Pergamon, 1979.

Lees F P, *Loss Prevention in the Process Industries*, Butterworths, London 1980.

MacKenzie D, *New Scientist* No. 1451, 11th April 1985, p. 4.

Musa J D, in *Reliability in Electrical and Electronic Components and Systems*, (Eds) E Langer and J Moltoft, North Holland Pub. Co. 1982.

Rasmussen N C, *Reactor Safety Study* WASH-1400, USNRC 1975.

Shooman M L, *Probabilistic Reliability: an Engineering Approach*, McGraw-Hill, New York 1968.

Shooman M L, in *Computing Systems Reliability*, (Eds) T Anderson and B Randell, Cambridge UP 1979.

Sinnott R K, *Chemical Engineering* vol. 6, Pergamon, Oxford 1983.

Stewart R M, *Chem. Eng.* October 1974, 622–6.

Williams V, in *Reliability in Electrical and Electronic Components and Systems*, (Eds) E Langer and J Moltoft, North Holland Pub. Co. 1982.

Reliability of metal structures

Metal failure has been the cause of many serious accidents, notably the 1954 Comet I air crashes, the Potchefstroom ammonia release accident in 1973 and the collapse of the 'Alexander Kielland' rig in the North Sea in 1980. Each of these accidents cost many lives. The conclusion to be drawn from accidents such as these is that a means of assessing the reliability of metal structures must be devised.

Surveys of pressure vessel failures yield interesting results. Ninety-four per cent of metal failures are due to cracking, mainly in weld affected material. Fifty per cent of failures are promoted by fatigue-assisted crack growth (Phillips and Warwick, 1968; Smith and Warwick, 1974).

The normal route in such cases follows three stages:

1. Crack initiation.
2. Crack growth to a 'critical' size.
3. Failure if the growth of the crack is not detected in time.

Reliable design of engineering structures demands that all three stages should be understood and managed. We need to know, therefore, (i) what causes cracks to form, (ii) what causes them to grow, (iii) what is the critical size of crack to cause failure and (iv) how can the size of a crack be determined.

We will also need to consider whether failure will be catastrophic (explosive) in the case of a pressure vessel, or whether the vessel will leak before failure. In the latter case, an opportunity would exist for the operator to depressurise the system, thereby preventing any danger of casualties.

The above may be described as the conventional, deterministic approach to fracture mechanics. However, for a proper risk assessment we need data on the probability of failure of metal structures. Probabilistic fracture mechanics will therefore be introduced.

The object of this chapter, therefore, is not to provide a comprehensive introduction to fracture mechanics, but merely to introduce those aspects of the subject which are of importance to engineers whose job it is to assess safety.

4.1 The origins of cracks

Cracks in steels may be caused by a number of mechanisms, of which these are among the most important:

1. *Porosity* or inclusions in the parent metal, as manufactured at the rolling mill.
2. *Lack of fusion* or inclusions in a weld joining different sections of a structure together.
3. *Stress corrosion cracking*. This is a form of intergranular corrosion which is more pronounced when a steel component is under stress. Environmental factors may also contribute to stress corrosion cracking. For example, stainless steels are subject to stress corrosion cracking in the presence of aqueous chloride ions.
4. *Weld decay*. This is another form of intergranular corrosion. Stainless steels may suffer from the depletion of chromium in the heat affected zone adjacent to a weld, due to the precipitation of chromium carbide. This makes the stainless steel more prone to corrosive attack.

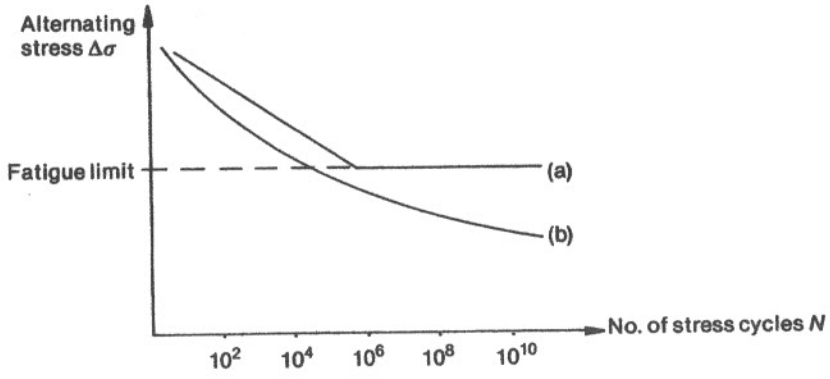


Fig. 4.1 Stress versus fatigue life for two different materials
(a) a typical alloy steel
(b) a typical aluminium alloy

Manufacturing defects (such as (1) or (2) above) are best handled by a careful regime of non-destructive examination (see Section 4.6). Stress corrosion cracking may be avoided by careful control of the environment. Weld decay can be avoided by heat treating all welds or else adding stabilising elements such as titanium or niobium to the steel, which react with carbon in preference to chromium.

A further cause of crack formation is *fatigue*. When a material is subjected to an alternating stress the material may fail after a number of stress cycles (see Fig. 4.1). For some materials (curve (a)) there is a threshold stress below which fatigue does not occur. This is called the fatigue limit. For other materials, for example aluminium alloys, no clearly defined threshold stress exists. For such materials it is possible to design for a limited life only.

Fatigue cracks form when about 85 per cent to 90 per cent of the material life has been reached. The detection of such cracks necessitates immediate remedial action.

4.2 Fracture locus: some preliminary definitions

It would be useful to be able to determine whether a component may fail suddenly and catastrophically by fast fracture. Sudden failures of pressure vessels may have explosive violence; if the failure is slower (so-called 'leak-before-break') it may be possible for the operator to depressurise the vessel before failure is complete.

The Comet I airframes failed suddenly. If 'leak-before-break' had occurred the aircrew would have been alerted and might have been able to descend to a safe altitude where the airframe could be depressurised (Fig. 4.2(a)).

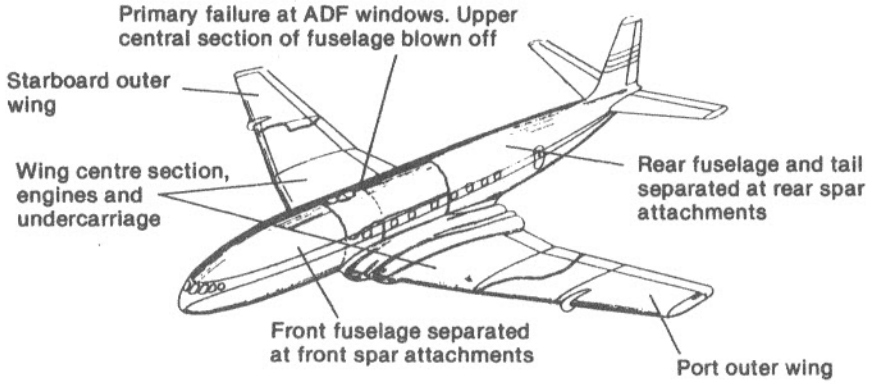


Fig. 4.2(a) Two Comet I airliners crashed in the Mediterranean in 1954. The cause was later identified to have been fast fracture of the fuselage, leading to explosive decompression. The fracture was caused by fatigue cracking due to higher-than-anticipated stress concentrations around two small windows used in conjunction with navigation (ADF) equipment. The first aircraft broke up into several pieces as shown (Courtesy HMSO)

It is a little difficult to believe, perhaps, that failures of massive steel pressure vessels do in fact occur. Figure 4.2(b) illustrates the failure of a steam drum at Cockenzie power station in Lothian region, Scotland, in 1966. This failure occurred after seven cycles of hydraulic pressure testing. The failure was due to a large crack caused by a manufacturing defect.

For the analysis of such failures, the conventional starting point is the Griffith equation (Griffith, 1921). From an argument based on energy considerations for brittle materials (i.e. materials which break without any plastic deformation), Griffith concluded that, for a material with an included crack $2a$ in length (Fig. 4.3), the critical stress for crack propagation (and hence failure) would be given by:

$$\sigma_c = \sqrt{\frac{2\gamma E}{\pi a}} \qquad [4.1]$$

where γ is the surface energy/unit area,
E is the Young's modulus,
a is the half-length of the crack, and
σ_c is the critical stress for failure.

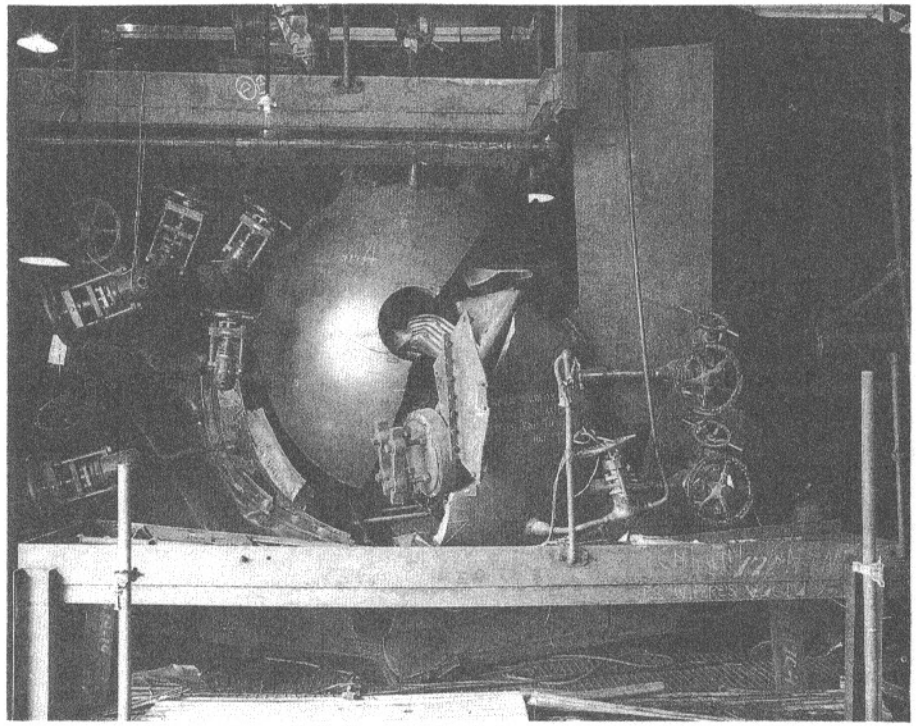


Fig. 4.2(b) A boiler drum at Cockenzie power station, Scotland, failed during its seventh hydraulic pressure test after fabrication, in 1966. The failure was caused by brittle fracture due to a 9 cm deep crack which had arisen during fabrication (Photo courtesy South of Scotland Electricity Board)

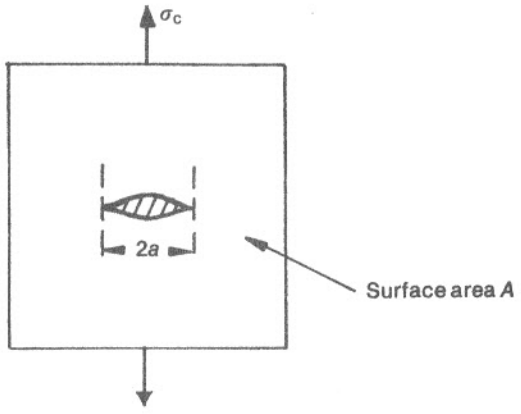


Fig. 4.3 Illustrating the Griffith equation

If we now define a quantity G_c called the *critical crack extension force/unit length of crack* where

$$G_c = 2\gamma \qquad [4.2]$$

then we obtain

$$\sigma_c^2 = \frac{EG_c}{\pi a} \qquad [4.3]$$

G_c is a material property since its magnitude is a function of interatomic bond strengths in the metallic crystal lattice.
(*Note:* The above relates to the *plane stress* condition, in which the third principal stress σ_3 is zero (see Fig. 4.4). This applies to thin plates. For thick sections we may consider the *plane strain* condition, for which condition the following relationship applies:

$$\varepsilon = (1 - \nu^2)\,\frac{\sigma_c}{E} \qquad [4.4]$$

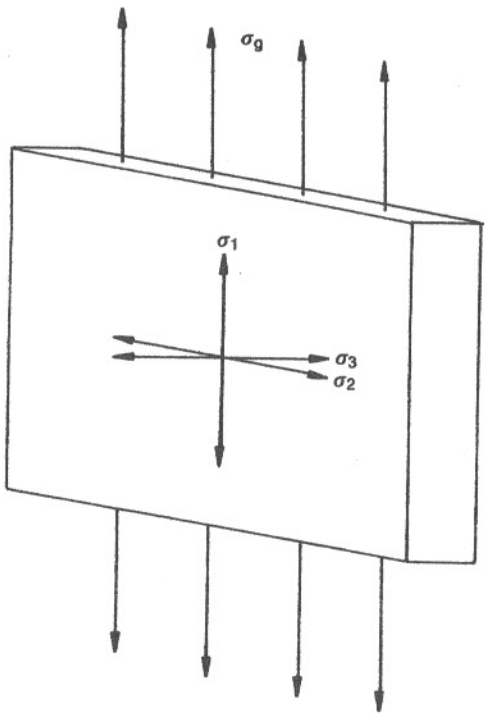


Fig. 4.4 The three principal stresses (σ_1, σ_2 and σ_3) in a plate under tension. σ_g is the average stress applied to the plate

Hence

$$\sigma_c^2 = \frac{EG_c}{\pi a(1-\nu^2)} \qquad [4.5]$$

for thick sections, where

ε = normal strain

and

ν = Poisson's ratio.)

We may also define a property called the *fracture toughness* or *critical stress intensity factor*, K_c

$$K_c = (\pi\sigma_c^2 a)^{1/2} \qquad [4.6]$$

i.e. from eqn [4.3]

$$K_c^2 = EG_c \qquad [4.7]$$

K_c must also be a material property, since E and G_c are material properties. Methods for determining K_c^* are described in ASTM E-399-72 and BS 5447.

4.2.1 Fracture locus: brittle and ductile fracture

Berry (1960) showed that the Young's modulus of a cracked plate, such as that shown in Fig. 4.3, was reduced from its uncracked value E as follows:

$$E_g = E.\frac{A}{A + 2\pi a^2} \qquad [4.8]$$

where E_g is the Young's modulus of the cracked plate,
A is the area of the plate, and
a is the half-length of the crack.

Hence the stress–strain relationship for the cracked plate will be, from Hooke's Law,

$$E_g = \frac{\sigma_g}{E_g} = \frac{\sigma_g}{E} + \frac{2\pi a^2\,\sigma_g}{AE} \qquad [4.9]$$

where σ_g is the stress applied to the cracked plate.

If the crack size a is eliminated between eqns [4.3] and [4.9], we obtain

$$E_g = \frac{\sigma_g}{E} + \frac{2E\,G_c^2}{\pi\sigma_g^3 A} \qquad [4.10]$$

* There are three different values for K_c for a material, corresponding to tensile fracture, forward shear failure and parallel shear failure. K_c for *tensile* failure is under consideration here. This is commonly given the symbol K_{IC}.

Equation [4.10] defines the fracture locus for the plate (Fig. 4.5).
Line (a) represents the stress–strain relationship for a crack-free plate.
Line (b) represents the reduced modulus of a cracked plate (eqn [4.9]).
Line (c) represents the fracture locus (eqn [4.10]).
The region to the right of line (c) represents an *unstable* region in which the crack will increase in size.
The region to the left represents a *stable* region where the crack size will remain fixed.

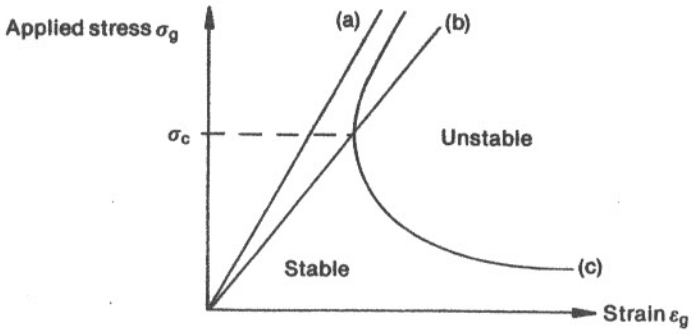


Fig. 4.5 Fracture locus of a cracked plate

Let us consider what happens when a cracked plate is loaded up. The stress will increase until the stress reaches the critical value σ_c. At this point fracture will occur and the system will *unload*. The way in which the system unloads will depend upon the way in which the plate is stressed, e.g.

1. If the loading system is rigid (Fig. 4.6) and the imposed strain is fixed, then *unstable crack growth* occurs, allowing the stress to fall until the crack re-enters the stable side of the locus at C. The growth of the crack then begins to slow until the crack stops at point D. This is termed *crack arrest*. The crack growth causes a further reduction in the modulus of the plate. The new modulus will be given by the gradient of the line ODE. The position of point D is defined, from considerations of the energy of crack arrest, to be such that area ABC is equal to area CDE.

   Further crack growth can only occur if the plate is again strained beyond point E. The crack will then again grow and reduce the stress, thereby re-entering the stable region. Repeated straining and unloading such as this is called *incremental tearing*. Unstable crack growth can not again occur with this loading system.
2. Alternatively, the plate may be loaded such that the stress remains constant no matter how large the crack becomes. This is termed a *fixed dead load*, and applies to structures such as bridges or aircraft wings. For fixed dead loading once crack growth has begun there is no mechanism for crack arrest. The plate will unload and fail by *fast fracture* (see Fig. 4.7).

At point X the stress–strain curve for the plate enters the unstable side of the locus at a point where the locus has a positive gradient, and *brittle fracture* ensues. For a plate

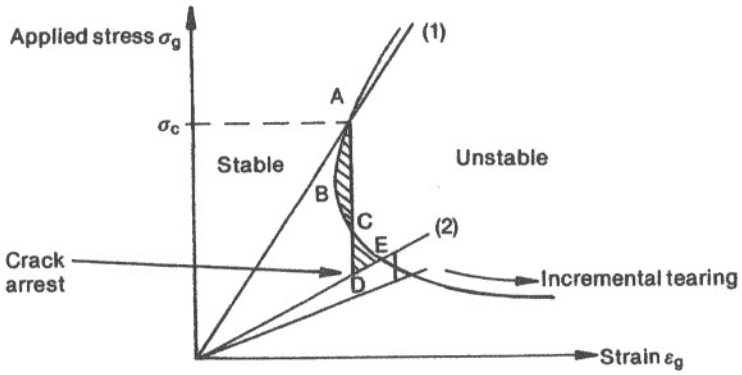


Fig. 4.6 Crack growth with rigid loading

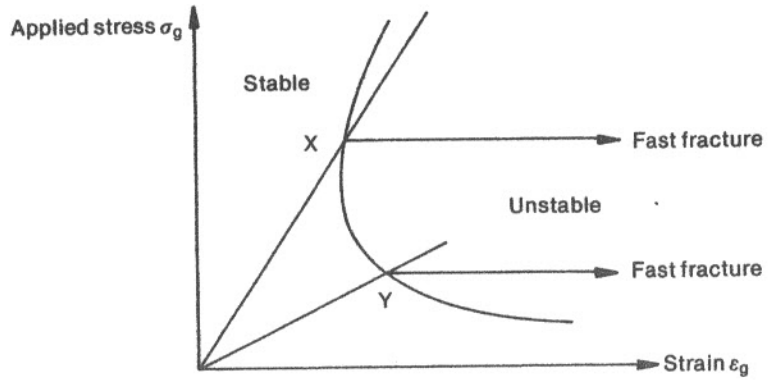


Fig. 4.7 Crack growth with fixed dead loading

with a larger crack (i.e. a lower modulus), the stress–strain curve enters the unstable side of the fracture locus at Y (where the locus has a negative gradient) and in this case *fast ductile fracture* will result. The consequences of brittle fracture and fast ductile fracture are the same; the plate fails very rapidly.

To summarise, then, the mode of fracture depends on (i) the type of loading and (ii) whether the stress–strain curve intersects the fracture locus at a point where the gradient of the locus is positive or negative.

4.2.2 Brittle and ductile fractures

The way in which a material has broken can be determined from the appearance of the fracture surface. For standard tensile test specimens, a point fracture indicates a ductile failure, whereas a flat fracture indicates a brittle failure (see Fig. 4.8). Similar

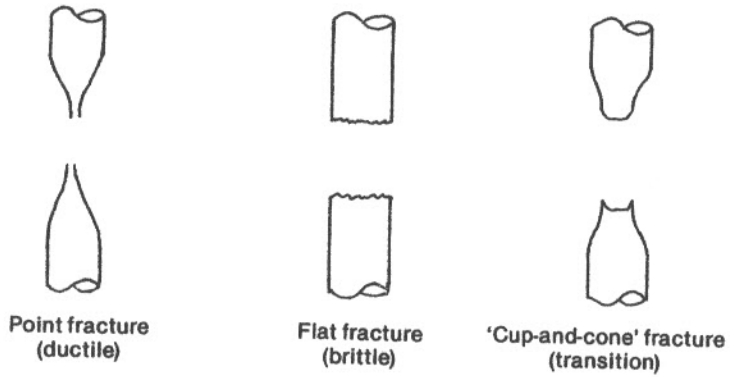


Fig. 4.8 The appearance of different types of fracture

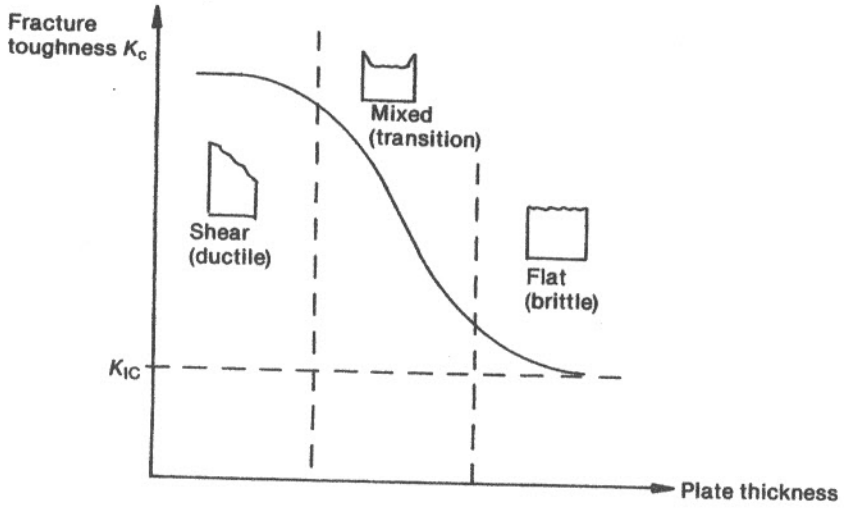


Fig. 4.9 The variation of fracture toughness with plate thickness

behaviour can be observed for the fracture of plates of differing thicknesses. Here, a ductile fracture is indicated by a sloping fracture surface, and a brittle fracture is, again, flat. A graph of the fracture toughness K_c against plate thickness yields interesting results (Fig. 4.9).

We can see that in brittle fracture the fracture toughness is greatly reduced. This is consistent with everyday experience. It is this value for fracture toughness which is defined to be K_{IC}, the tensile failure fracture toughness (see Section 4.2).

The cause of brittle fracture in thick plates may be described thus: in a thick plate, the principal stresses (Fig. 4.4) within the metal will be approximately equal. This means that an element of metal in the centre of the plate is being pulled equally in all directions. The metal cannot yield (i.e. deform plastically) since this would mean *contraction* perpendicular to the direction of yield – but since the principal stresses are equal, the material is being pulled equally strongly in those directions, and no such contraction can occur. This behaviour may be summarised by the Tresca criterion for yield:

$$|\sigma_1 - \sigma_3| \geqslant \sigma_{ys} \qquad [4.11]$$

where σ_1, σ_2 and σ_3 are the three principal stresses and σ_{ys} is the uniaxial yield stress.

Since the principal stresses are approximately equal, yielding will not be possible. Instead the material fails by a sudden cleavage or brittle failure.

For thin plates, however, σ_3 equals zero. The metal will therefore yield before failure, i.e. a ductile fracture.

4.2.3 Fracture resistance

Differentiating eqn [4.10] with respect to the stress on the plate σ_g gives an expression for the gradient of the fracture locus

$$\frac{d\varepsilon_g}{d\sigma_g} = \frac{1}{\text{gradient}} \qquad [4.12]$$

$$= \frac{1}{E} - \frac{6EG_c^2}{\pi\sigma_g^4 A}$$

It may be seen that the gradient of the fracture locus will always be negative if

$$\frac{G_c^2}{A} > \frac{\pi\sigma_g^4}{6E^2} \qquad [4.13]$$

G_c^2/A is called the *fracture resistance parameter.* The effect of the magnitude of this parameter is illustrated in Fig. 4.10.

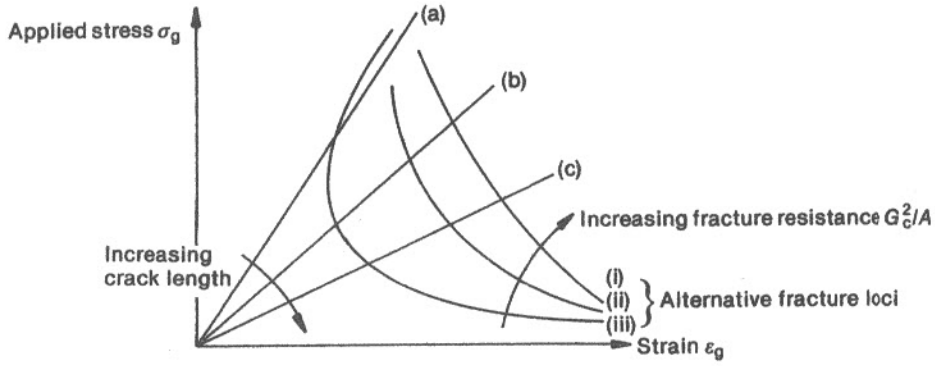


Fig. 4.10 The effect of the fracture resistance parameter

For fracture locus (i), the gradient of the locus is always negative. Hence (Section 4.2.1) ductile fracture will occur.

For fracture locus (iii), the point of intersection of the stress–strain curve will determine whether fracture is brittle or ductile, as previously discussed. Therefore brittle fracture will occur for curve (a) and ductile fracture for curves (b) and (c).

4.2.4 The effect of the loading system

So far, we have only considered plates which have either a fixed dead load or a fixed imposed strain. In many practical cases however, the unloading path will follow neither of these routes, but will instead follow an intermediate path. (Fig. 4.11).

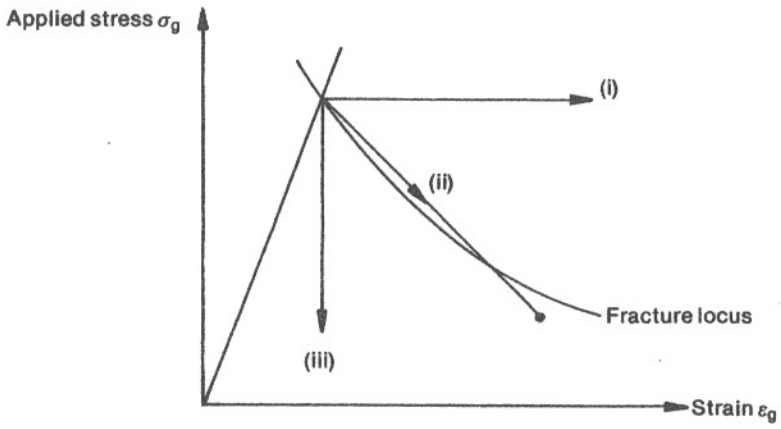


Fig. 4.11 The effect of unloading path

Case (i) (fixed dead load) will result in failure by fast ductile fracture, and case (iii) (fixed imposed strain) will not cause fracture unless further straining occurs, in which case incremental tearing will result. However, case (ii) corresponds to limited fast propagation of the crack followed by crack arrest. The loading mechanism in case (ii) might correspond to pneumatic pressure.

For pneumatic pressure, any straining of the pressure vessel caused by crack growth will cause the vessel volume to increase and, consequently, the pressure will fall. We may write

$$\frac{d\sigma_g}{d\varepsilon_g} = \frac{\partial \sigma_g}{\partial p} \cdot \frac{\partial p}{\partial v} \cdot \frac{\partial v}{\partial \varepsilon_g} \qquad [4.14]$$

Any increase in strain ε_g will cause a small increase in the specific volume of the fluid in the pressure vessel, v. Hence $\frac{\partial v}{\partial \varepsilon_g}$ is small and positive. An increase in vessel pressure p, will cause an increase in vessel stress σ_g, so $\frac{\partial \sigma_g}{\partial p}$ is positive. However, $\frac{\partial p}{\partial v}$, the rate of change of pressure with respect to specific volume, is large and negative for hydraulic (water) pressure, where $\frac{\partial p}{\partial v}$ is typically -2.5×10^{12} Nkg/m^5, but for pneumatic (gas) pressure,

$$\frac{\partial p}{\partial v} = \frac{-RT}{v^2} \qquad [4.15]$$

from $pv = RT$. Hence for pneumatic pressure $\frac{\partial p}{\partial v}$ is typically -1.07×10^5 Nkg/m^5. Thus $\frac{\partial p}{\partial v}$, and therefore $\frac{\partial \sigma_g}{\partial \varepsilon_g}$, is some 10^7 times smaller for pneumatic pressure than for hydraulic pressure. This means that, if a crack in a pressure vessel begins to grow, the pressure will fall very quickly if the pressure is hydraulic (i.e. if the contents are entirely liquid) but less quickly if the pressure is pneumatic. Therefore, curve (ii) in Fig. 4.11 corresponds to pneumatic pressure and curve (iii) corresponds to hydraulic pressure.

The significance of this result lies in its application to pressure vessel proof testing. If a pressure vessel fails during a pneumatic pressure test, the pressure vessel may explode and send debris over a large radius. If the test is carried out using hydraulic pressure, the rapid unloading path should ensure that crack arrest occurs before the vessel disintegrates. This effect can be appreciated by comparing Fig. 4.2(b) with Fig. 5.2.

4.2.5 Factors which may reduce fracture toughness K_{IC}

A number of environmental factors can act to reduce the fracture toughness of metals, i.e. to cause embrittlement. A reduction in fracture toughness K_{IC} would mean, via eqns [4.7], [4.2] and [4.1], that the critical stress for crack propagation would be reduced and that metal failure would be more likely.

The environmental factors that may lead to embrittlement include (a) low temperature, (b) the presence of hydrogen, (c) neutron irradiation and (d) the presence of impurities.

At low temperatures, the value of K_{IC} may be reduced by, typically, a factor of four. The temperature at which this fall in K_{IC} occurs is called the *transition temperature* or the *nil ductility temperature*. Figure 4.12 illustrates the typical variation of K_{IC} with temperature.

The presence of hydrogen may lead to *hydrogen embrittlement*, which arises where an inward diffusion of hydrogen into steel has occurred. The effect is normally reversible by low temperature annealing which drives out the absorbed hydrogen (Probert and

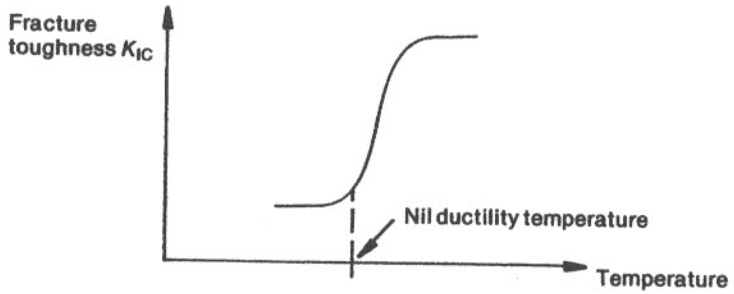


Fig. 4.12 The variation of fracture toughness with temperature

Rollinson, 1961). Figure 4.13(a) illustrates how such embrittlement reduces the tensile strength of iron foil.

Neutron irradiation of steels in nuclear reactors gives rise to *irradiation embrittlement*. The steel is damaged by repeated impact of high energy neutrons upon atoms in the crystal lattice, which increases the density of dislocations and vacancies. The effects are two-fold. Firstly, the nil ductility temperature is increased by, typically, 80°C, and secondly the maximum value for K_{IC} is reduced (Steele, 1975). (See Fig. 4.13(b).)

A number of impurities in metals can lead to embrittlement. In steel, these include nitrogen (*nitriding*) which raises the transition temperature, and *copper*. If an Inconel electrode, which contains copper, is inadvertently used to weld stainless steel, the copper may be drawn through the entire thickness of the weld. The weld will pass NDE, unless swabs are taken, and will pass a pressure test but will fail after a small number of loading cycles.

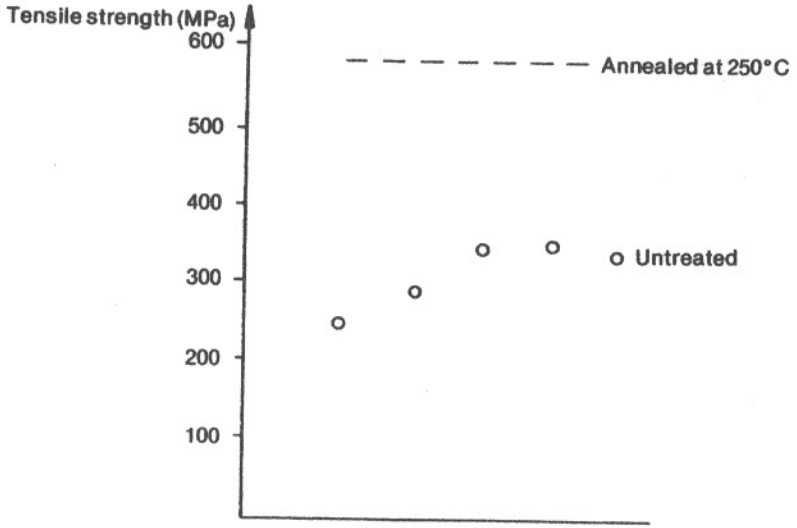


Fig. 4.13(a) Illustrating the effect of hydrogen embrittlement upon the tensile strength of iron foil (Thomson, 1979)

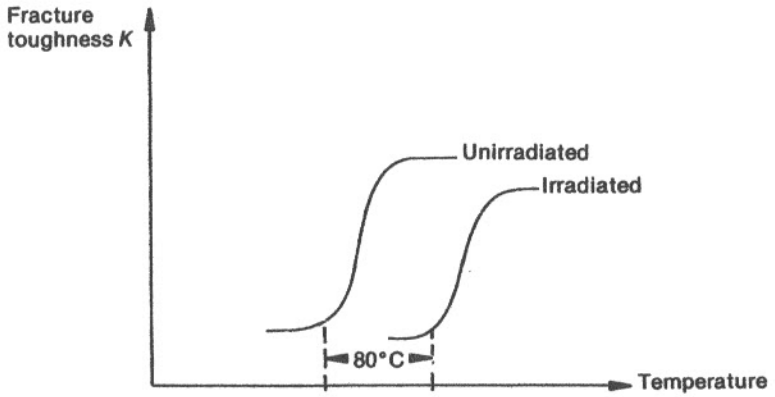


Fig. 4.13(b) The effect of irradiation embrittlement upon fracture toughness in steel

Finally, the existence of residual stresses in non-annealed welded steel structures should not be neglected. Such stresses may allow stress corrosion cracking to begin (Section 4.1), and the additional stress will also reduce the fracture toughness (Kihara *et al.*, 1971).

4.3 The determination of critical crack size

In Sections 4.1 and 4.2, different causes and types of metal failure have been discussed, but the underlying problem – how big a crack has to be to cause failure – has not yet been confronted. Westergaard (1939) considered this problem and showed that the stress distribution at the tip of a crack in a *perfectly elastic* material (Fig. 4.14) should be

$$\sigma_y = \sigma_g x (x^2 - a^2)^{-1/2} \qquad [4.16]$$

where σ_y is the stress in the metal normal to the plane of the crack,
σ_g is the uniform tensile stress field perpendicular to, and remote from, the crack,
x is the distance from the centre of the crack, and
a is the half-length of the crack.

The stress distribution near the crack for eqn [4.16] is illustrated in Fig. 4.15(a), which shows how this stress distribution implies that, even for very small values of σ_g, there will be an infinitely large stress at the crack tip (where $x = a$). This conclusion is obviously incorrect, since it would mean that materials would fail at low stresses even if vanishingly small cracks were present.

Hencky (1924) and Prandtl (1924) performed similar calculations for elasto-plastic materials (Fig. 4.14(b)). They showed that, for an elasto-plastic material with a sharp crack, there is a limiting value of stress, $\sigma_{y\,max}$, that can be attained at the crack tip (Fig. 4.15(b)) such that,

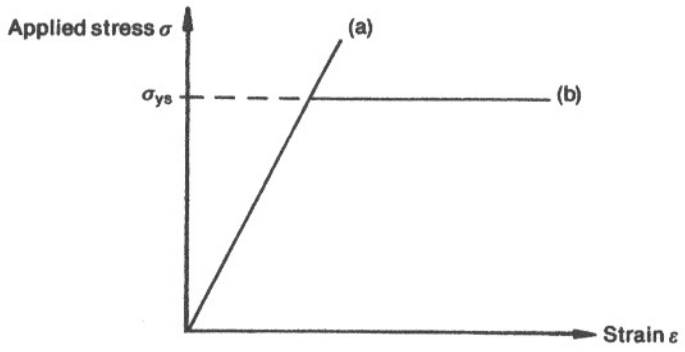


Fig. 4.14 The relationship between stress and strain for (a) perfectly elastic and (b) elasto-plastic materials

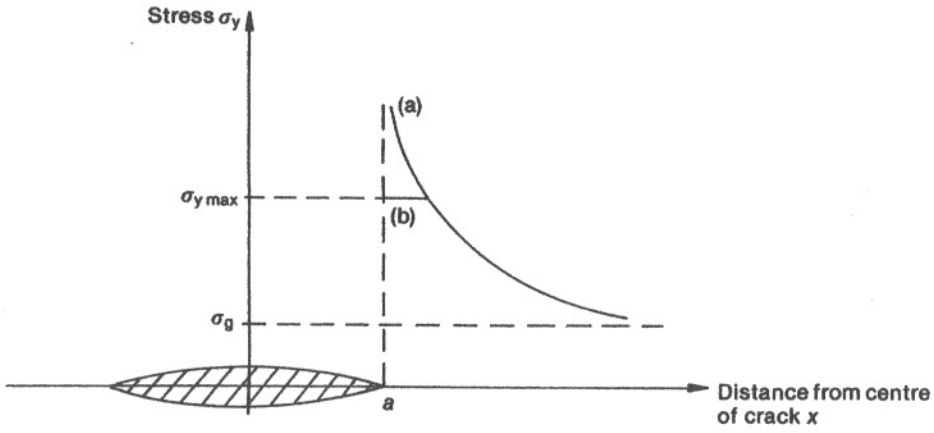


Fig. 4.15 The stress distribution at the tip of a crack for (a) a perfectly elastic material and (b) an elasto-plastic material

$$\sigma_{y\,max} = \sigma_{ys}(1 + \pi/2) \qquad [4.17]$$

σ_{ys} is the yield stress of the material which, for an elasto-plastic material (Fig. 4.14(b)), is equal to the tensile strength. However, eqn [4.17] implies that $\sigma_{y\,max}$ is greater than σ_{ys}. Hence this expression again implies that, for a vanishingly small crack and small loads, the material will fail since the stress at the crack tip exceeds the tensile strength. Again, this conclusion is obviously wrong.

Neuber (1937) concluded that failure of the material at the crack tip must only occur when the average stress over some characteristic distance from the crack tip b, attains a critical value, i.e. when

$$\bar{\sigma} = \frac{1}{b}\int_{a}^{a+b} \sigma_y \, dx \qquad [4.18]$$

exceeds a critical value. We will call b the 'spreading length'. (See Fig. 4.16.)

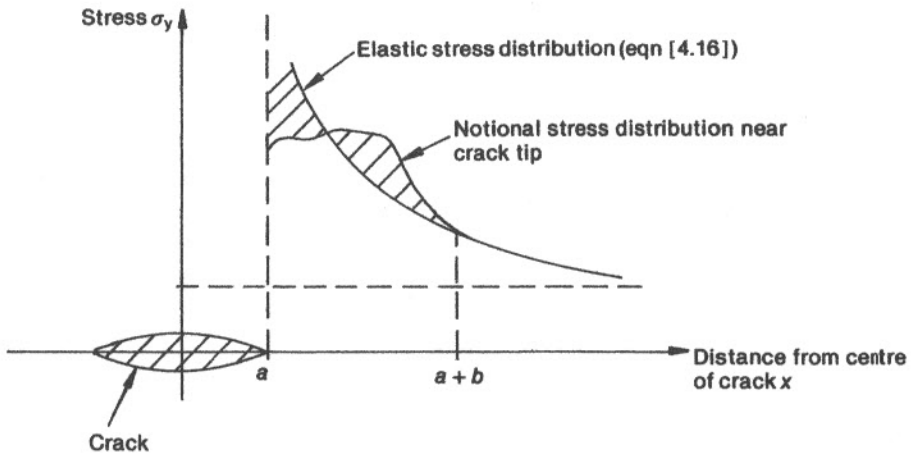


Fig. 4.16 Illustrating the significance of the spreading length b

This approach yields more promising results. Here we shall follow the analysis of Irvine (1976):

1. Assume that the two shaded areas in Fig. 4.16 are equal, and that the areas remain equal for any plastic relaxation that occurs if the spreading length b is large enough to contain the crack tip stress perturbation.
2. Substitute the Westergaard equation [4.16] into Neuber's equation [4.18] and integrate. This gives

$$\bar{\sigma} = \sigma_g\left(\frac{2a}{b} + 1\right)^{1/2} \qquad [4.19]$$

3. Equation [4.19] gives, for a vanishingly small crack ($a \to 0$), that the average stress near the crack equals the applied stress, i.e. $\bar{\sigma} = \sigma_g$.
This is the 'common sense' answer for very small cracks. However, for larger cracks at fracture, the average stress $\bar{\sigma}$ at the crack tip will equal the ultimate tensile strength σ_u, i.e.

$$\sigma_u = \sigma_c \cdot \left(\frac{2a}{b_{crit}} + 1\right)^{1/2} \qquad [4.20]$$

where σ_c is the applied stress at fracture, and where b_{crit} is the critical length of b for fracture. In other words, b_{crit} is the critical minimum length over which the

stress concentration at the crack tip must be 'spread' to prevent the stress at the crack tip exceeding σ_u.

It seems reasonable to propose that the critical spreading length b_{crit}, which represents the ability of a material to spread a stress concentration, will be a material property.

4.3.1 A relationship between spreading length, fracture toughness and critical crack size

Equation [4.20] is a formula which fits everyday observation. It does not predict that vanishingly small cracks may cause failure in the way that the Westergaard or Prandtl formulae do, but it does predict that the larger the crack size $2a$, the smaller the applied stress σ_c needed to cause failure.

However, a method must still be devised for determining the critical spreading length b_{crit}. If it is assumed that b_{crit} is very much less than the crack size, then eqn [4.20] simplifies to

$$\sigma_u \simeq \sigma_c \left(\frac{2a}{b_{crit}}\right)^{1/2} \qquad [4.21]$$

(This equation represents the *Linear Elastic Fracture Mechanics* model.)

The fracture toughness was defined (eqn [4.6]) thus:

$$K_c = (\pi \sigma_c^2 a)^{1/2} \qquad [4.6]$$

Eliminating $\sigma_c^2 a$ from these two equations yields

$$b_{crit} = \frac{2K_c^2}{\pi \sigma_u^2} \qquad [4.22]$$

The fracture toughness K_c and the ultimate tensile strength σ_u are properties which can be readily measured, so eqn [4.22] provides a means of determining the critical spreading length b_{crit}.

Equation [4.20] may be re-arranged to give an expression for the critical crack half-length a.

$$a = \left[\left(\frac{\sigma_u}{\sigma_g}\right)^2 - 1\right] \frac{b_{crit}}{2} \qquad [4.23]$$

This yields a value for the crack half-length a that will lead to crack propagation with an applied stress σ_g. Equations [4.22] and [4.23] together enable a value for a to be determined.

4.3.2 Practical determination of critical crack size

It would be more convenient to be able to describe the critical spreading length b_{crit} in

terms of a more readily measurable quantity than the fracture toughness K_c – in particular, tests such as the Charpy or Izod impact tests which are quick and simple to perform.

If we introduce another material constant S such that

$$b_{crit} = \frac{2}{3} S \qquad [4.24]$$

where S therefore becomes another definition for spreading length, then we get from eqn [4.22] that

$$S = \frac{3}{\pi} \cdot \left(\frac{K_c}{\sigma_u}\right)^2$$

$$\simeq \left(\frac{K_c}{\sigma_u}\right)^2 \qquad [4.25]$$

and eqn [4.23] becomes

$$\frac{a}{S} = \frac{1}{3} \cdot \left[\left(\frac{\sigma_u}{\sigma_g}\right)^2 - 1\right] \qquad [4.26]$$

(This is the *Stress Concentration Theory* of fracture mechanics.)

Irvine has determined an empirical correlation between S and Charpy energy

$$S = 0.133\, \Delta l^2 \left(\frac{\text{Charpy energy}}{\text{Yield stress}}\right)^{1/3} \qquad [4.27]$$

Here S is in mm, Δl is the percentage elongation of a $4\sqrt{A}$ gauge length tensile specimen, where A is the cross-sectional area of the specimen, Charpy energy is in Joules and yield stress is in MN/m².

Thus we can say that tensile test data (which will yield values for Δl, yield stress and ultimate tensile stress σ_u) and Charpy test data, together with eqns [4.27] and [4.26], enable a value for the critical half-length a to be determined, given the applied average stress σ_g.

It should be noted, however, that low temperature, which reduces the fracture toughness (Section 4.2.5), will also reduce the critical crack size for failure. Figure 4.17 illustrates the way in which the failure pressure in a pressurised water reactor (PWR) pressure vessel varies with temperature, with crack half-length as a parameter (Collier, Davies and Garne, 1982).

4.3.3 An alternative approach – BCS theory and the COD method

Bilby, Cottrell and Swinden (BCS) (1963) modelled a crack in a material as an array of dislocations. Using a rectilinear law for the interatomic forces across the crack faces, they were able to show that

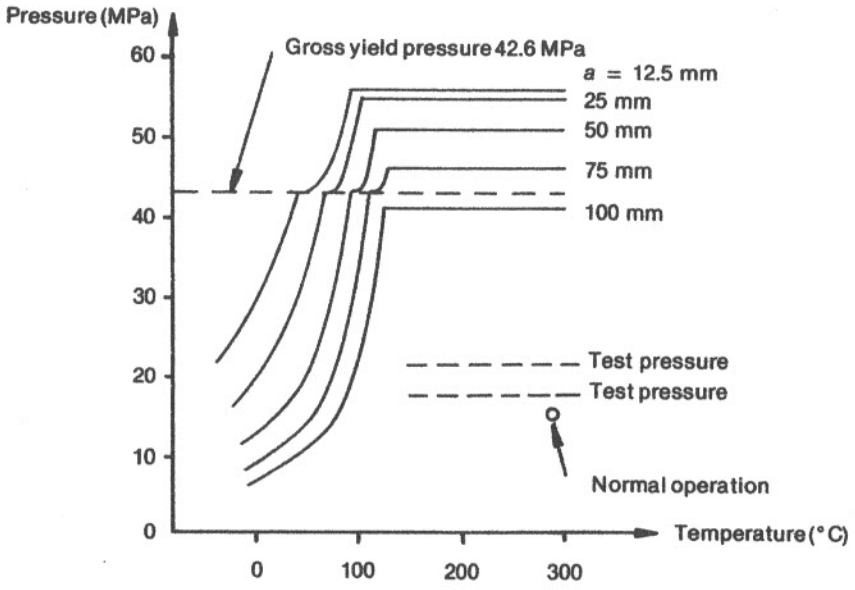


Fig. 4.17 Failure pressure for different flaw sizes (Collier *et al.*, 1982)

$$\frac{a}{S} = \frac{\pi}{8}\left[\ln \sec\left(\frac{\pi}{2} \cdot \frac{\sigma_g}{\sigma_{ys}}\right)\right]^{-1} \qquad [4.28]$$

We now introduce the *crack opening displacement* (COD) δ defined thus:

$$\delta = \frac{8a\,\sigma_u}{\pi E} \cdot \ln \sec\left(\frac{\pi\sigma_g}{2\sigma_u}\right)$$

$$= \frac{8a\,\sigma_u}{\pi E}\left[\frac{1}{2}\left(\frac{\pi\sigma_g}{2\sigma_u}\right)^2 + \frac{1}{12}\left(\frac{\pi\sigma_g}{2\sigma_u}\right)^4 + \ldots\right]$$

$$\simeq \frac{\pi\sigma_g^2\, a}{E\sigma_u} \qquad [4.29]$$

COD may be measured experimentally (BS 5762).

If we assume that the material is elasto-plastic as shown in Fig. 4.14(b), then σ_u is equal to σ_{ys}, and for failure $\sigma_g = \sigma_u$.

Hence

$$\delta = \frac{\pi\sigma_{ys}\, a}{E}$$

$$\Rightarrow a = \frac{E}{\pi}\frac{\delta}{\sigma_{ys}} \qquad [4.30]$$

The BCS theory therefore yields another means of determining the critical crack half-length a using two experimental measurements. This time the measurements require tensile testing (to obtain values for the yield stress σ_{ys} and Young's modulus E), and COD testing (to obtain a value for δ).

4.3.4 A comparison of the various methods for determining *a*

Three methods for determining the critical crack half-length a have been presented. Other available methods include the J integral method (Rice, 1968). This method has not been discussed here since it is difficult to apply in practice.

The methods presented here are summarised below:

1. *Linear elastic fracture mechanics* (LEFM)

Equations [4.21] and [4.24] together provide the LEFM model:

$$\frac{a}{S} = \frac{1}{3}\left(\frac{\sigma_u}{\sigma_g}\right)^2 \qquad [4.31]$$

2. *Stress Concentration Theory* (SCT)

Equation [4.26] describes the SCT model, a refinement of the LEFM model:

$$\frac{a}{S} = \frac{1}{3}\left[\left(\frac{\sigma_u}{\sigma_g}\right)^2 - 1\right] \qquad [4.26]$$

3. *BCS Theory*

This theory is described by eqn [4.28]

$$\frac{a}{S} = \frac{\pi}{8}\left[\ln \sec\left(\frac{\pi}{2} \cdot \frac{\sigma_g}{\sigma_u}\right)\right]^{-1} \qquad [4.28]$$

($\sigma_u = \sigma_{ys}$ for an elasto-plastic material.)

The methods are presented graphically in Fig. 4.18. It can be seen that all three methods are equally valid for low applied stresses σ_g (i.e. high σ_u/σ_g). For such low stresses the critical crack length is large, therefore such cracks are readily detectable using non-destructive examination (NDE) techniques.

Thus, the older pressure vessel design codes such as ASME section VIII or BS 1500, which specified that the primary stress σ_g in a pressure vessel should nowhere exceed $\sigma_u/4$ $\left(\text{i.e. } \frac{\sigma_u}{\sigma_g} > 4\right)$, ensured that the critical crack length was large and therefore readily detectable.

However, more recent design codes such as ASME section III, BS 3915 or BS 5500

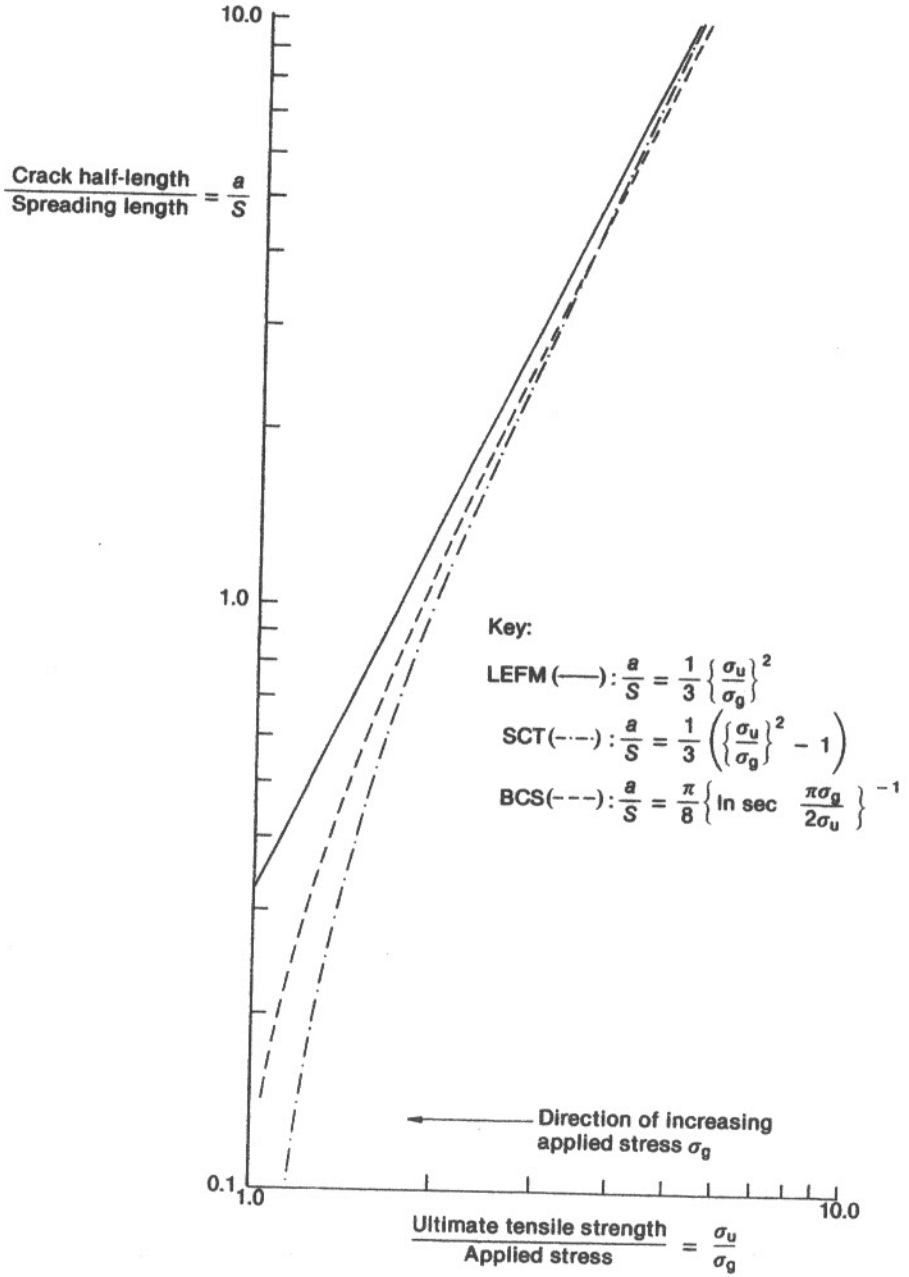


Fig. 4.18 Comparing three methods for determining the critical crack half-length a (Irvine 1976)

have shown a trend towards specifying the allowable stress σ_g in terms of the yield stress σ_{ys}. This may encourage designers to specify materials with a high σ_{ys}/σ_u ratio, thereby reducing thickness and cost. This, in turn, would mean that the σ_u/σ_g ratio may be reduced. From Fig. 4.18, a small σ_u/σ_g ratio will mean that the critical crack half-length a is much reduced.

Finally, for σ_u/σ_g much below 2.0, the LEFM theory is no longer reliable. For stresses in this range, either the SCT or the BCS formulae should be used.

To summarise, therefore, we may say that a high ratio of ultimate tensile stress to applied stress (σ_u/σ_g) ensures that a structure has a high tolerance to any cracks which may form. If structures are designed with σ_u/σ_g of much less than 3, the critical crack half-length a may be unacceptably small; so small, perhaps, that non-destructive examination (NDE) techniques may not be sufficiently sensitive to detect such a crack. NDE will be more fully discussed in Section 4.5.

Example 4.1

To determine the critical crack size in a steel pressure vessel with a design stress σ_t of 200 MN/m². The steel was subjected to tensile and Charpy testing with the following results:

Charpy Energy	120 J
σ_{ys}	250 MN/m²
σ_u	500 MN/m²
Δl	30 %

From eqn [4.27], we obtain

$$S = 0.133\,\Delta l^2 \left(\frac{\text{Charpy energy}}{\text{Yield stress}}\right)^{1/3}$$

$$= \underline{93.7 \text{ mm}}$$

Equation [4.26] gives, therefore

$$a = \frac{S}{3}\left[\left(\frac{\sigma_u}{\sigma_g}\right)^2 - 1\right]$$

$$= \underline{164 \text{ mm}}$$

Hence the critical crack size is $2a$ or 32.8 cm

4.3.5 The failure assessment diagram

The failure assessment diagram is the basis of the UK Central Electricity Generating Board's 'R6' method of defect assessment (Harrison and Milne 1981). The method has been recommended for use for defect assessment of British pressurised water reactors (Marshall 1982).

The failure of structures is bounded by two extremes of behaviour: linear elastic

fracture and fully plastic failure. If the two ratios K_r (stress intensity factor/fracture toughness) and σ_r (applied stress/stress required to cause plastic failure) are calculated, they form the coordinates for a point on the failure assessment diagram (Fig. 4.19). Failure will occur if the *assessment point* (σ_r, R_r) falls on or outside the *assessment line*. It is evident that the validity of the approach depends upon the failure assessment line being a reasonable approximation to the failure curve for any structural geometry.

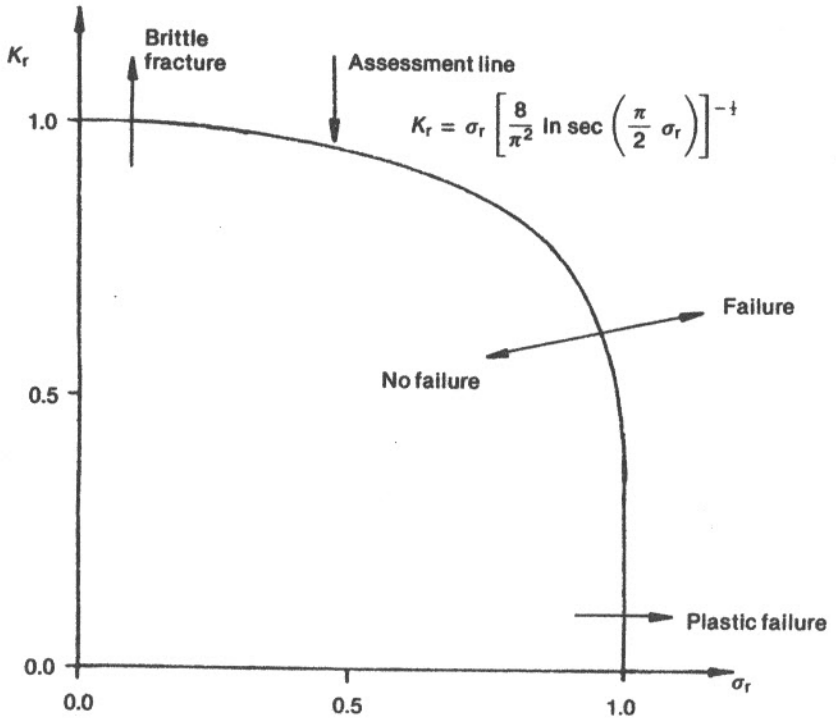


Fig. 4.19 The failure assessment diagram

The two parameters of the assessment point, K_r and σ_r, may be calculated as follows:

$$K_r = K_I(a)/K_{IC} \qquad [4.32]$$

$$\sigma_r = \sigma_g/\sigma_u \qquad [4.33]$$

where

$K_I(a) = \sigma_g\sqrt{\pi a}$, (eqn. [4.6])

K_{IC} is the tensile fracture toughness for the unflawed material
σ_g is the applied stress, and
σ_u is the tensile strength for plastic failure.

Equations [4.25], [4.28], [4.32] and [4.33] may be used to determine an expression for the assessment line, thus:

$$K_r = \sigma_r \left[\frac{8}{\pi^2} \ln \sec\left(\frac{\pi}{2}\sigma_r\right)\right]^{-1/2} \qquad [4.34]$$

Therefore, in order to use this method, information is needed about the following:

(i) The stresses within the metal (σ_g).
(ii) The size of any cracks present ($2a$).
(iii) The fracture toughness of the metal (K_{IC}).
(iv) The tensile strength for plastic failure.

These data enable the position of the assessment point (K_r, σ_r) to be determined.

4.4 Subcritical crack growth

In Section 4.3, various means of calculating the maximum crack size $2a$ in a pressure vessel with an allowable stress σ_g were presented. This provides a means of deciding whether a vessel is safe or not. However, a number of mechanisms exist which cause an existing subcritical crack to become larger. These mechanisms include:

1. Fatigue crack growth.
2. Creep crack growth.
3. Corrosion fatigue crack growth.

In many cases, all three of these crack growth mechanisms will be acting simultaneously, e.g. in chemical reactors where there is load cycling (leading to fatigue), high temperature (causing creep) and a corrosive environment.

4.4.1 Fatigue crack growth

If a series of cracked metal specimens are subjected to fatigue testing – repeated cyclic loading and unloading – with a stress amplitude $\Delta\sigma$, the length of the crack grows as the number of stress cycles increases. For a larger stress amplitude, the crack size will increase more rapidly. Figure 4.20 illustrates this.

These curves may be plotted as a single line if the logarithm of the gradient ($\ln da/dN$) is plotted against the logarithm of the amplitude of the stress intensity factor, ΔK, where

$$\Delta K = K_{max} - K_{min} = \Delta\sigma(\pi a)^{1/2} \qquad [4.35]$$

Equation [4.35] is similar to eqn [4.6]. (Note that the right hand side of eqn [4.35] must, in some circumstances, be multiplied by a geometrical or 'configuration correction' factor to allow for local stress concentrations.)

A typical plot of ($\ln da/dN$) vs ($\ln \Delta K$) is shown in Fig. 4.21. The curve shows three

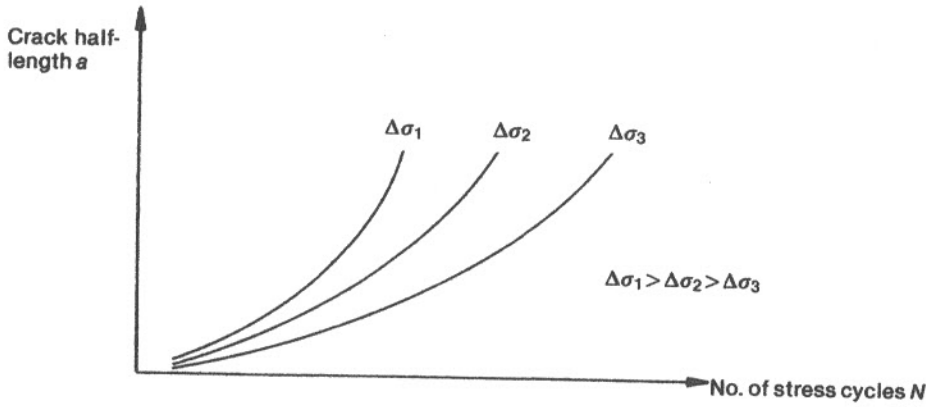


Fig. 4.20 Crack growth due to stress cycling

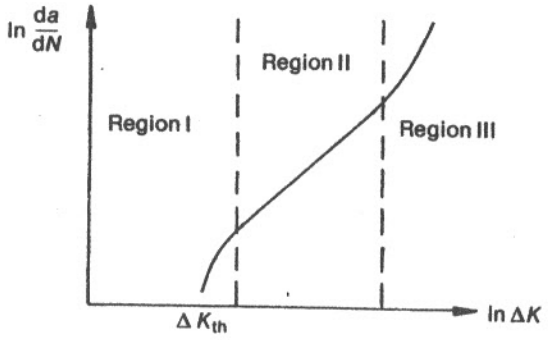


Fig. 4.21 A log–log plot of crack growth rate $\frac{da}{dN}$ against stress intensity factor ampliture ΔK, illustrating Paris' Law

regions, two of them curved. No crack growth occurs below a threshold amplitude of the stress intensity factor, ΔK_{th}. The central region of the curve (region II) is a straight line. This forms the basis of an empirical relationship called Paris' Law:

$$\frac{da}{dN} = C(\Delta K)^m \qquad [4.36]$$

where C and m are empirical constants.

Figure 4.22 shows typical experimental data for three types of steel, and the corresponding values for the constants C and m. The values for the constant C depend upon the units of ΔK; their numerical values are changed if ΔK is given in terms of $N/m^{3/2}$ instead of $MN/m^{3/2}$. Furthermore, ΔK is sometimes expressed in $MPa\sqrt{m}$ instead of $MN/m^{3/2}$; these units are, of course, the same.

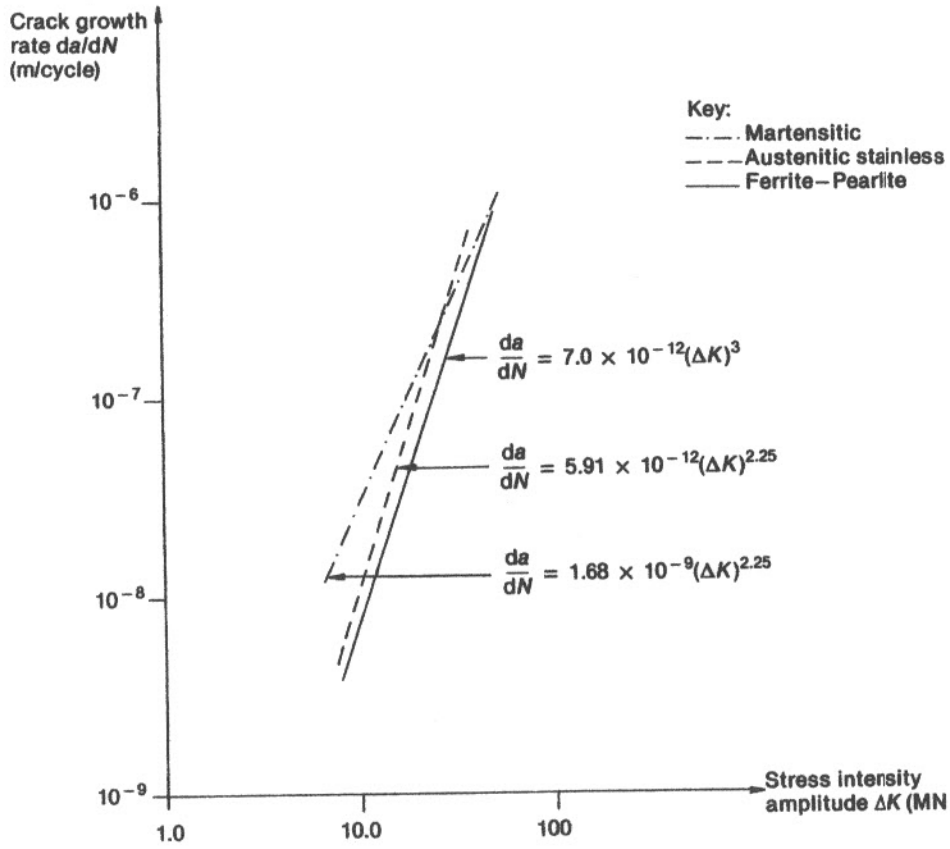


Fig. 4.22 Experimental data for fatigue crack growth in different types of steel

By integrating Paris' Law, we may obtain a value for the number of cycles required for a crack to grow from an initial half-length a_i to a final half-length a_f.

$$\Delta N = \int_{a_i}^{a_f} \frac{1}{C(\Delta K)m}\,da$$

$$= \frac{1}{C\pi^{m/2}(\Delta\sigma)^m}\left[\frac{a_f^{[1-(m/2)]} - a_i^{[1-(m/2)]}}{[1-(m/2)]}\right] \qquad [4.37]$$

From eqn [4.34], if a flaw is found in a pressure vessel by NDE to have size a_i and the maximum permissible size is calculated (Section 4.3), then the 'life' (the number of stress cycles allowable) may be determined. This has important application in all areas of industry.

Stress cycling due to temperature changes induces *thermal fatigue*. For this, the stress amplitude will be given by

$$\Delta\sigma = \frac{E \propto \Delta T}{1-\nu} \qquad [4.38]$$

where $\propto$ is the coefficient of thermal expansion, ΔT is the temperature difference within the metal, and ν is Poisson's ratio.

In some metals and alloys, notably aluminium-based alloys, variation of the *stress intensity ratio* R ($= K_{min}/K_{max}$) affects the shape of the crack growth rate/stress amplitude curve. This is illustrated in Fig. 4.23.

Variations in R do not have a major effect upon the fatigue life of steels.

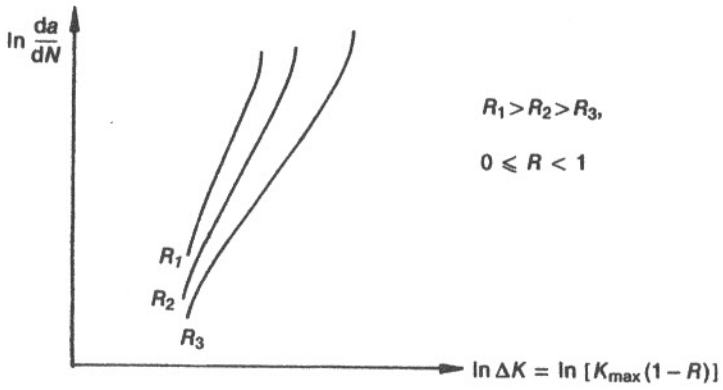


Fig. 4.23 Illustrating the effect of stress intensity ratio R upon crack growth rate $\frac{da}{dN}$

4.4.2 Other crack growth mechanisms

At elevated temperatures, cracks may grow at a steady load due to *creep crack growth*. Empirical correlations of the form

$$\frac{da}{dt} = f(K, T) \qquad [4.39]$$

can be determined, where $\frac{da}{dt}$ is the crack growth rate and T is the temperature (Tomkins, 1983; Ainsworth and Goodall, 1983).

The rate of fatigue crack growth may be enhanced if the metal is in a corrosive environment. This is termed *corrosion fatigue crack growth* (Scott and Truswell, 1983; Scott, Tomkins and Foreman, 1983). This may be characterised by an expression similar to eqn [4.36]:

$$\frac{da}{dN} = C_1(\Delta K)^m \qquad [4.40]$$

For corrosion fatigue the empirical constant C_1 is larger than for fatigue in an non-corrosive environment. In the case of A533-B steel plate, which is used in the fabrication of pressurised water reactor (PWR) pressure vessels, the value of C_1 is 4×10^{-11} $\left(\text{for } \Delta K \text{ in MPa}\sqrt{\text{m}} \text{ and } \frac{da}{dN} \text{ in m/cycle}\right)$ if the plate is immersed in PWR water. In an inert atmosphere, the constant C has a value of 1×10^{-11}. The constant m in both cases is 3.0. Hence fatigue crack growth occurs four times faster in water than in a non-corrosive medium in this case.

4.5 Non-destructive examination (NDE)

Sections 4.3 and 4.4 have shown how cracks greater than a given size may cause metal failure, and also how smaller cracks may grow due to load cycling or environmental factors. In order to be confident of the safety of pressure vessels, boilers and aircraft structures, methods of non-destructive examination have been devised which enable cracks to be located and their size measured.

We need to be able to tell at any time during the life of a component or structure whether it is safe to continue to use that component or structure. One approach is to *proof test* the article, which involves the application of stress above the design load, in order to demonstrate that the maximum crack size must be below a certain value. This is often done as a factory test on newly-manufactured articles. It is then possible to use the article with high confidence that failure will not occur at the lower, operating stresses. Strategies for proof testing of pressure vessels and components have been developed in many industries. For example, the technique used by British Gas for the proof testing of gas pipelines involves pressurising the pipes (using hydraulic pressure for reasons described in Section 4.2.4) to stress levels up to or exceeding the yield stress of the metal, with the objective of failing and eliminating all significant defects (Fearnehough and Jones, 1978). A disadvantage of this technique is that any defects thus detected cause irreparable damage to the section of pipework under test.

Non-destructive examination (NDE) techniques, as the name implies, aim to detect and measure cracks in a manner which does not damage the metal. Any defects thus detected may then be ground out and repaired, if so desired. NDE methods may be divided into two categories:

1. Methods for surface inspection (Section 4.5.1).
2. Methods for detecting internal flaws (Section 4.5.2).

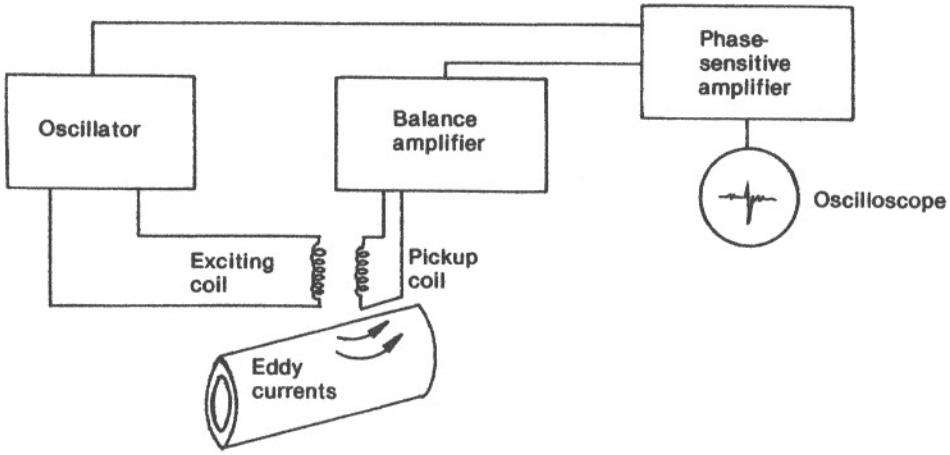


Fig. 4.24 Eddy current inspection

4.5.1 NDE methods for surface inspection

The most obvious NDE method is *visual inspection*. This can be made easier by spraying the surface with a dye or a suspension of fluorescent particles and wiping clean. Any surface cracks retain the liquid which are then more readily detected. This is called 'dye-penetration' testing.

Magnetic particle inspection is carried out by coating the surface to be inspected with a magnetic powder and then applying a local magnetic field using an electromagnet. The powder enables the field lines to be seen. Any defects on or just below the surface are revealed by distortion of the field lines. However, cracks which lie parallel to the field lines will not be revealed.

Eddy current inspection is illustrated in Fig. 4.24. The component under test is placed into a primary alternating current field, which induces eddy currents in the item. In turn the eddy currents produce a secondary alternating current field, which modifies the primary field. Any alteration in the primary field may be detected using an oscilloscope. Reference specimens with known defects are required for calibration purposes with this method. Defects up to 6 mm below the surface may also be detected using this technique.

4.5.2 NDE methods for internal inspection

The methods used for internal inspection are in two main categories, radiography and ultrasonics.

Radiography is performed using X-rays or a γ-ray source (cobalt-60), depending on the thickness of the material to be inspected. (γ-rays are more penetrating than X-rays.) The high energy rays are projected through the material onto a photographic plate. The contrast of the image is proportional to the degree of the change of density of the material at the defect site.

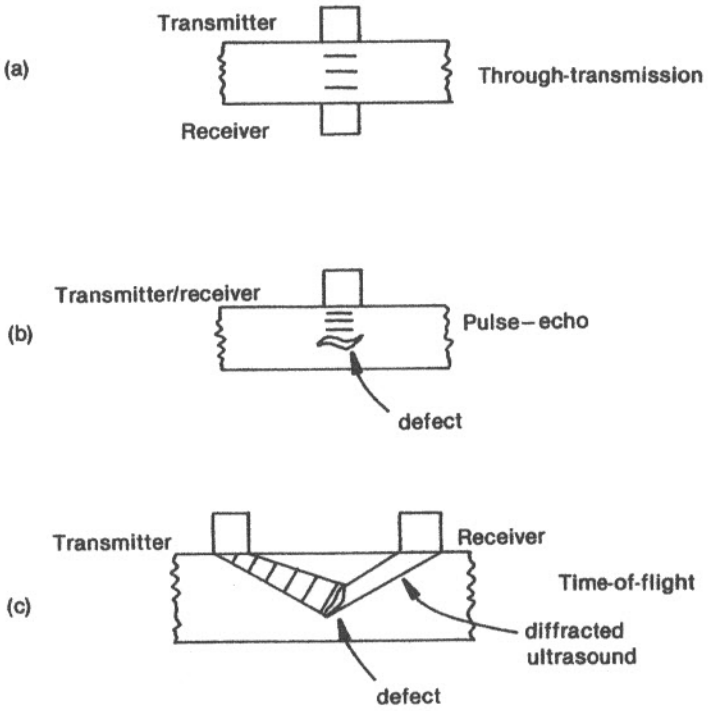


Fig. 4.25 Ultrasonic inspection techniques

Image quality indicators are placed in the field to verify that an adequate quality of radiograph has been achieved. The method has a disadvantage in that it requires access to both sides of the specimen, which in some applications may not be possible.

In *ultrasonic inspection*, a beam of ultrasound is passed through the material. Internal flaws may be detected by their interruption of the beam, or by their reflection of the beam. These methods are called 'through-transmission' and 'pulse–echo' respectively.

A further technique called 'time-of-flight' enables accurate sizing of deep cracks (Silk, 1977; Hawker, 1983; Temple, 1983). These methods are illustrated diagrammatically in Fig. 4.25. The time-of-flight technique is a fairly recent development which shows promise; the method requires access to one side of the specimen only, and can be automated.

The technique is not recommended when the crack penetrates less than 5 mm into the material. In that event, a pulse–echo technique should be adopted.

4.6 Probabilistic fracture mechanics

Sections 4.3, 4.4 and 4.5 have followed a conventional *deterministic* approach. A deterministic approach states that something is either good enough or it is not; thus, a crack

either exceeds the critical length or it does not. In practice, however, things are seldom so clearly definable. *Probabilistic* fracture mechanics attempts to replace this deterministic approach with a more realistic approach based on:

1. The probability that a crack of a given size will be present,
2. The probability that the crack will not be detected using NDE, and
3. The probability that the crack will grow to a critical size before the next inspection is due.

Alternative routes to failure, such as a mistake in the design procedure leading to overstressing, may also be relevant.

A statistical approach to the probability of component or structural failure may also be adopted in cases where sufficient data are available (Lucia, 1984). This can be done for certain categories of pressure vessels (O'Neil and Jordan, 1972).

4.6.1 The incidence of cracks in steel pressure vessels

Lidiard (1984) has suggested a function for describing the incidence of cracks in pressurised water reactor pressure vessels. For each pressure vessel, the average number of cracks in the size range x to $(x + dx)$ is given by $A(x)dx$, where

$$A(x) = Ae^{-\Theta x} \qquad [4.41]$$

and x is the crack size (depth) in *millimetres*,
A is a constant numerically equal to 0.59, and
Θ is a constant numerically equal to 0.16.

This result is based on surveys of PWR pressure vessels immediately after manufacture.

Hence we may write that the probability of a defect greater than x mm long being present after manufacture is given by

$$p_{>x}(x) = \int_x^\infty A(x)\,dx$$

$$= \frac{A}{\Theta} e^{-\Theta x} \qquad [4.42]$$

Thus there is a probability of 6.7 per cent that a crack greater than 25 mm long will be present in the pressure vessel after fabrication. Similarly, there is a probability of 4×10^{-7}/vessel that a crack greater than 100 mm long will be present.

For vessels other than PWR pressure vessels, the values of A and Θ will need to be reconsidered. It seems reasonable that the value of Θ will be genuinely constant; however, the value of A will be proportional to the mass of steel and the length of welding in any given vessel.

4.6.2 The reliability of non-destructive examination

The reliability of NDE has been subject to close study in recent years. An EEC subcommittee called the Plate Inspection Steering Committee (PISC) organised a set of international trials in the late 1970s and early 1980s, in which a steel plate with known defects was passed from one NDE inspection team to another, through several countries, before being destructively examined. Each team worked independently.

The final, destructive examination enabled accurate, direct measurements of defect sizes to be taken. This enabled the accuracy of the NDE results to be assessed. The initial results were reported by O'Neil (1980). A Defect Detection Probability (DDP) was defined thus:

$$\text{DDP} = \frac{\text{No. of teams reporting a given defect}}{\text{No. of teams carrying out inspections}} \qquad [4.43]$$

The inspections were carried out under laboratory conditions and the teams worked to a tightly-controlled version of the ASME XI inspection procedure, using ultrasonic inspection methods. The results for single, planar defects are shown in Fig. 4.26. In round figures this graph shows that the probability of detecting a 2.5 cm (1 in.) defect is about 50 per cent.

Haines (1983) has analysed the PISC results and has concluded that the Defect Detection Probability is dependent upon the following parameters:

1. The mean defect detection threshold.
2. The average received signal amplitude.
3. The standard deviation of signal amplitude
   (a) between scans, and
   (b) between different NDE teams.
4. The through-the-wall dimension of the defect.

A simplified approach will be used here. If it is assumed* that the DDP follows a *lognormal* cumulative distribution function $F(x)$, then the DDP may be characterised (Section 2.4.5) as follows:

$$\text{DDP}(x) = \text{Probability of detecting a defect with size } x$$

$$= \int_0^x f(x)\,dx$$

$$\text{where } f(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left[-\frac{(\ln x - \mu)^2}{2\sigma^2}\right]$$

Hence we may write that

$$\text{DDP}(x) = F(x)$$

* Marshall (1982) has used an exponential function for $(1 - \text{DDP}(x))$.

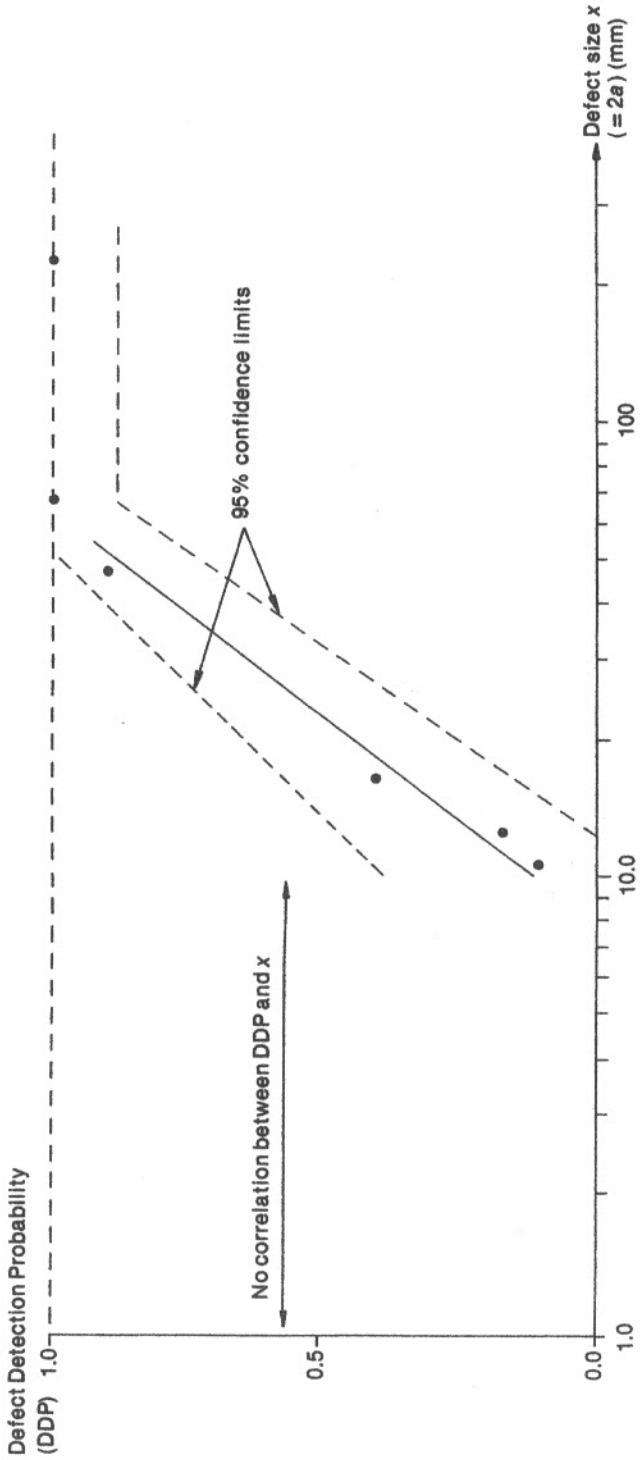


Fig. 4.26 The PISC results: defect detection probability as a function of defect size (O'Neil, 1980)

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} \exp\left(\frac{-z^2}{2}\right) dz \qquad [4.44]$$

where $z = \left(\dfrac{\ln x - \mu}{\sigma}\right)$

$e^{\mu} = 23$ mm, [4.45]

$\mu = 3.135$, [4.46]

and $\sigma = 0.667$.

Thus from the PISC data, we may calculate the probability that a given size of crack or defect will be detected using NDE.

4.6.3 A probabilistic approach to crack growth

The deterministic approach to crack growth (Section 4.4) enables predictions about crack growth rates to be made, if adequate information about material properties and environmental factors are known. Temple (1985) however, has proposed a probabilistic approach to crack growth. In this model, crack growth follows a Gaussian distribution. Given an initial crack size x_0 mm, a stress intensity factor amplitude ΔK and N load cycles, the probability that the crack will grow to size x mm is given by

$$p_c(x \mid x_0, \Delta K, N) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x - \mu}{\sigma}\right)^2\right] \qquad [4.47]$$

For PWR pressure vessel steel, Temple proposes the mean μ and standard deviation σ are described as follows:

$$\mu = x_0 + (6.56 \times 10^{-5}\, \Delta K.N) \qquad [4.48]$$

$$\sigma = 2.93 \times 10^{-5}\, \Delta K \qquad [4.49]$$

In these equations, ΔK has units of MPa$\sqrt{\text{m}}$ and μ and σ are in millimetres.

4.6.4 Reliability of metal structures

The results of the foregoing Sections 4.6.1, 4.6.2 and 4.6.3 mean that, in principle at least, probabilities for the failure of pressure vessels and other metal structures maybe calculated. The probability that a pressure vessel will fail in the time interval Δt between inspections will be given by:

Failure probability $F_{PV} = \sum_{x_0=0}^{x_c}$ (probability that a defect of size x_0 exists) × (probability that it is not detected by NDE) × (probability that the defect will grow to a critical size before the next NDE)

$$= \int_0^{x_c} A(x_0).(1 - \mathrm{DDP}(x_0)).p_c(x_c \mid x_0)\, dx \qquad [4.50]$$

In this expression, $A(x_0)$ may be calculated from eqn [4.41], $\mathrm{DDP}(x_0)$ may be calculated from eqn [4.44] and $p_c(x_c \mid x_0, \Delta K, \frac{dN}{dt} \Delta t)$ may be calculated from eqn [4.47]. Also

x_0 is the crack size at the time of inspection,
x_c is the critical crack size,
ΔK is the amplitude of the stress intensity factor,

$\frac{dN}{dt}$ is the load cycling frequency, and

Δt is the time interval between inspections.

In principle then, it should be possible to determine a failure probability for a pressure vessel by integrating eqn [4.50] numerically. The three functions $A(x_0)$, the crack incidence function, $\mathrm{DDP}(x_0)$, the defect detection probability, and $p_c(x_c \mid x_0)$, the crack growth probability, are illustrated schematically in Fig. 4.27.

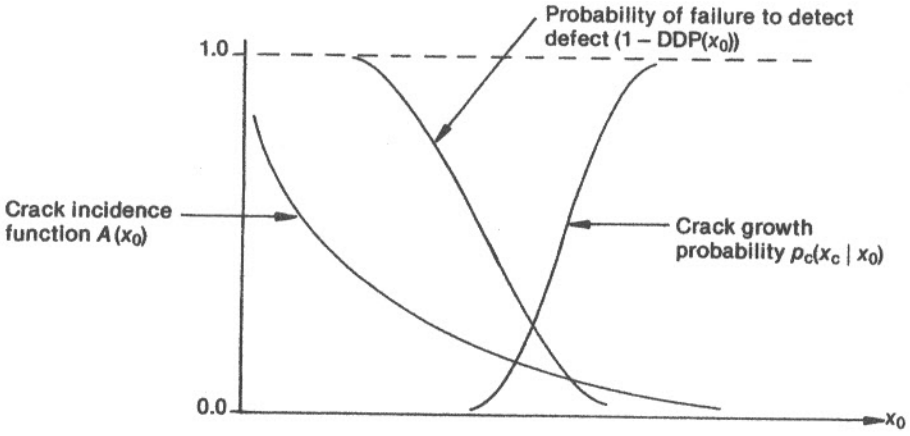


Fig. 4.27 Schematic illustration of the functions representing crack incidence, probability of failure to detect and crack growth probability

The numerical integration of the product of these functions is not easy. However, the calculation may be simplified if the deterministic model of crack growth (eqn [4.36]) is used in place of the probabilistic model. For deterministic crack growth,

failure will only occur if a crack grows by an amount $\left(\frac{dx}{dN} \times \text{number of cycles}\right)$ such that the critical crack size is exceeded. In this case eqn [4.50] simplifies to

$$F_{PV} = \int_{x_0}^{x_c} A(x).(1 - \mathrm{DDP}(x)) \frac{dx}{dN}.dN$$

$$= \int_{x_0}^{x_c} A(x).(1 - \mathrm{DDP}(x)).\frac{2da}{dN}.dN \qquad [4.51]$$

In this formula the lower bound of the integral x_0 will be given by

$$x_c - x_0 = \frac{2da}{dN}.\Delta N \qquad [4.52]$$

This approach is illustrated in Example 4.2.

Example 4.2

To estimate the failure probability for a steel pressure vessel.
A boiler drum with a two-year inspection period is subject to a stress cycle of $\Delta\sigma$ equal to 20 MPa with a frequency of 0.01 Hz. Estimate the probability of the vessel failing between inspections, assuming that Lidiard's crack incidence function is applicable (eqn [4.41]).

Materials data (illustrative):

UTS σ_u	: 450 MPa
Yield stress σ_{ys}	: 320 MPa
Charpy energy	: 110 J
Elongation Δl	: 20%
Design stress σ_g	: 250 MPa

Fatigue crack growth rate : $\frac{da}{dN} = 6.59 \times 10^{-9}\, \Delta K^{3.726}$ mm/cycle
(ΔK in MPa $\sqrt{\text{m}}$).

Method

Equation [4.27] yields a value for the spreading length S of 37.3 mm. Equation [4.26] therefore gives a critical crack length x_c (equals $2a$) of 55.7 mm.

Equation [4.35] gives the stress intensity factor amplitude ΔK, which equals 5.9 MPa$\sqrt{\text{m}}$.

The number of cycles in a two-year interval will be $(2 \times 365 \times 24 \times 3600 \times 0.01) = 6.3 \times 10^5$ cycles.

The lower bound of the integral in eqn [4.51] can be worked out using eqn [4.52] to yield $(x_c - x_0)$ equal to 6.19 mm. Over this small range we may assume that $A(x)$ and $\mathrm{DDP}(x)$ are constant, so eqn [4.51] simplifies to:

$$F_{PV} \fallingdotseq A(x_c).(1 - \mathrm{DDP}(x_c)).(x_c - x_0)$$
$$= (7.95 \times 10^{-5}).(0.09).(6.19)$$

Hence $F_{PV} = \underline{4.43 \times 10^{-5}}$; i.e. there is a probability of 4.43×10^{-5} that the vessel will fail between inspections.

4.6.5 A statistical approach to pressure vessel failure

The method used in Example 4.2 is alright as far as it goes; however, in many cases the materials data may not be known, or else the load cycling frequency or stress amplitude will not be known. Futhermore, all the possible causes of pressure vessel failure have not been considered. O'Neil and Jordan (1972) have described pressure vessel failure probabilities as follows:

$$p(\text{Failure}) = (P_D \times P_{FD1}) + (P_C \times P_{FD2}) + (P_M \times P_{FD3}) \qquad [4.53]$$

where P_D is the probability of failure due to bad design,
P_C is the probability of failure due to poor construction,
P_M is the probability of failure due to poor materials, and
P_{FD} is the probability of failure to detect the fault.

Obtaining reliable data for this type of approach becomes excessively difficult, however.

An alternative approach based on *measured* pressure vessel failure rates (Smith and Warwick, 1981) suggests that failure rates may be characterised by

$$F_{PV}(t) = 1 - e^{-\lambda t} \qquad [4.54]$$

where the hazard rate λ is estimated to be approximately 5×10^{-5} per pressure vessel year for catastrophic failure.

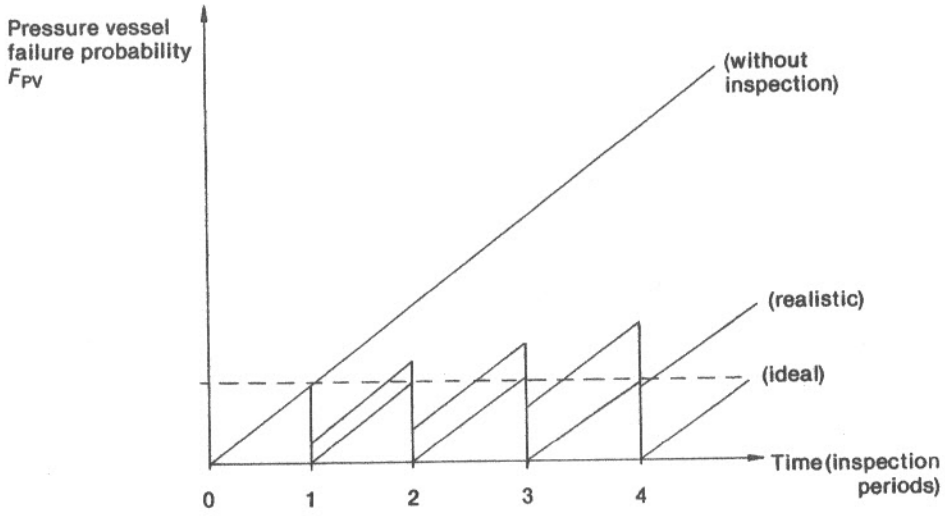


Fig. 4.28 The variation of pressure vessel failure probability with time

The function of periodic inspection of pressure vessels is, ideally, to restore t to zero in eqn [4.54]. Hence the failure probability will follow a saw-tooth pattern similar to that shown in Fig. 3.23, with steps at each in-service inspection. However, it is not realistic to assume that every inspection will be perfect and will restore the failure probability to zero. Due to imperfect inspection, it is likely that the failure probability will show a rising trend throughout the life of the pressure vessel (Fig. 4.28).

4.7 Conclusions

The main points arising from this chapter may be summarised as follows:

1. Embrittlement is not a prerequisite for fast fracture. Fast, ductile fracture can also occur, subject to loading conditions, crack size and material properties.
2. Once formed, cracks grow because of fatigue, creep or corrosion. Frequent inspection is necessary to ensure that the critical crack size is not exceeded.
3. Non-destructive examination can detect an internal 2.5 cm crack with a reliability of about 50 per cent. This crack size is close to the critical crack size in some materials and under certain loading conditions.

Questions

4.1 A sample of 0.36% C steel has a Charpy energy of 43 J. Tensile testing gives a yield stress of 226 MN/m², and an ultimate tensile stress of 453 MN/m² at 29 per cent elongation. If the metal is to be used in a pressure vessel designed in accordance with a code which allows applied stresses of up to two-thirds of the yield stress, determine, using linear elastic fracture mechanics, the critical crack size $2a$.
(386 mm)

4.2 Repeat Question 4.1 using stress concentration theory.
(343 mm)

4.3 Repeat Question 4.1 using BCS theory.
(351 mm)

4.4 A steel pressure vessel with a known defect 2 cm long has a design stress of 200 MN/m². It is subjected to thermal cycling with a period of 24 minutes and a maximum ΔT of 50 K. Estimate the time required for the crack to grow to 3 cm, given the following data:

Young's modulus	208 GN/m²
Coefficient of thermal expansion	11×10^{-6}/K
Poisson's ratio	0.3
$\frac{da}{dN}$	$7.0 \times 10^{-12}(\Delta K)^3$ (ΔK in MN/m$^{3/2}$)

(*c.* 1 year)

4.5 Estimate the probability of a large pressure vessel having a crack greater than 4 cm long after manufacture *and* of the crack not being detected. Assume that

$$A(x) = 0.59 \exp(-0.16x)$$

where x is the defect size in millimetres. A table of values for the Gaussian distribution is given in Table 2.1.
(less than 0.12%)

References and bibliography

Ainsworth R A and Goodall I W, *J. Press. Vess. Tech.* **105**, 1983, 263–8.
ASTM *Plane strain fracture toughness of metallic materials*, ASTM E-399-72, ASTM standards part 31, 960, 1973.
Berry J P, *J. Mech. Phys. Solids* **8**, 207–16.
Bilby B A, Cottrell A H and Swinden K H, *Proc. Roy. Soc* **A272**, 1963, p304.
British Standard BS 5447, *Plane strain fracture toughness of metallic materials*, British Standards Institution 1977.
British Standard BS 5762, *Crack opening displacement testing*, British Standards Institution 1979.
Cheng C G, *J. Nuc. Mat.* **56**, 1975, 11–33.
Collier J G, Davies L M and Garne L, *Nucl. Energy* **21**, 1982, 377–83.
Fearnehough G E and Jones D G, in *Tolerance of Flaws in Pressurised Components*, I.Mech.E., London 1978.
Griffith A A, *Phil. Trans. Roy. Soc.* **A221**, 1921, 163–98.
Haines N F, *Nucl. Energy* **22**, 1983, 349–56.
Harrison R P and Milne I, *Phil. Trans. Roy. Soc.* **A299**, 1981, 145–53.
Hawker B M, *Nucl. Energy* **22**, 1983, 309–18.
Hencky H Z, *Math. Mech.* **4**, 1924, p 323.
Irvine W H, *UKAEA Safety and Reliability Directorate Report R48*, 1976.
Kihara H, Kanazawa T, Ando Y and Machida S, *At. En. Review* **9**, 1971, 687–786.
Lidiard A B, *UKAEA Report AERE – M 3402*, 1984.
Lucia A C, in *Reliability of Engineering Materials*, (Ed.) A L Smith, Butterworths, London 1984.
Marshall W, *An Assessment of the Integrity of PWR Pressure Vessels*, Second Report, UKAEA 1982.
Neuber H, *Kerbspannungslehre*, Springer, Berlin 1937.
O'Neil R, *Atom* July 1980, 181–3.
O'Neil R and Jordan G, in *Periodic Inspection of Pressure Vessels*, I.Mech.E., London 1972.
Parker A P, *The Mechanics of Fracture and Fatigue*, Spon, London 1981.
Phillips C A G and Warwick R G, *UKAEA Report AHSB(S) R162*, 1968.
Prandtl L, *Proc. Congress App. Mech.* Delft, 1924, p 43.
Probert, L E and Rollinson, J J, *Electroplating and Metal Finishing*, September 1961, pp. 323–6.
Rice J R, *J. App. Mech.* June 1968, 379–86.
Scott P M, Tomkins B and Foreman A J E, *J. Press. Vess. Tech.* **105**, 1983, 255–61.
Scott P M and Truswell A E, *J. Press. Vess. Tech.* **105**, 1983, 245–54.
Silk M, in *Research Techniques in NDT*, (Ed.) R S Sharpe, Academic Press, London 1977.
Smith T A and Warwick R G, *UKAEA Safety and Reliability Directorate Report R30*, 1974.
Smith T A and Warwick R G, *UKAEA Safety and Reliability Directorate Report R203*, 1981.
Steele L E, *Neutron Irradiation Embrittlement of Reactor Pressure Vessel Steels*, IAEA, Vienna 1975.

Temple J A G, *Nucl. Energy* **22**, 1983, 335–48.
Temple J A G, *Nucl. Energy* **24**, 1985, 53–62.
Thomson J R, PhD thesis, Aberdeen University 1979.
Tomkins B, *J. Press. Vess. Tech.* **105**, 1983, 269–72.
Westergaard H M, *J. Appl. Mech.* June 1939, A49–A53.

Major industrial hazards

In order to be able to make an overall assessment of the safety of an engineering system, information is needed about the *probability* of various postulated accidents together with information about their *consequences*. Accident probabilities may be assessed using the methods for calculating reliabilities introduced in chapters three and four. The object of this chapter is to introduce means of calculating accident consequences. To put this more directly, we need to know how many people a given accident is likely to kill, or how large an area is likely to be destroyed or rendered unusable, in order that a true assessment of the safety of a plant can be made, and in order that the risks associated with different types of plant can be compared.

5.1 Accident classification

The accidents under consideration in this chapter will be restricted to those types of accidents at engineering installations or factories *which have the potential to cause harm to members of the public living outside the site boundary*. The categories of accident which have such potential are described below, with historical examples.

5.1.1 Chemical plant fires

5.1.1(i) *Pool fires*

A pool fire is where a flammable liquid is spilled and catches fire. The liquid may run through sewers and drains and emerge burning elsewhere if not contained by a low wall termed a *bund*.

In Cleveland, Ohio, in 1944 a spillage of 2000 tonnes of liquified natural gas (LNG) caused widespread fire – 130 people died.

The term 'running liquid fire' is often used to describe the event following leakage of flammable liquid from high level on chemical plant.

5.1.1(ii) *Flash fires*

A flash fire occurs when a flammable vapour cloud is ignited and burns without exploding (i.e. without causing a significant overpressure). In the Los Alfaques campsite disaster in Spain in 1978, a road tanker spilled 17 tonnes of liquified propene, which expanded to a 200 m vapour cloud before catching fire – 215 people were killed in the ensuing *deflagration*. The road tanker had been overfilled, and no relief valve was fitted.

The liquid contents had expanded due to solar heating, and thereby caused the tank to fracture.

5.1.2 Explosions due to flammable or unstable chemicals

An explosion may be defined as 'a sudden release of high pressure gas to the environment'.

5.1.2(i) *Boiling Liquid Expanding Vapour Explosions (BLEVEs)*

A BLEVE typically occurs as a result of a leak from a pressure or other storage vessel containing a flammable liquid. The leaking liquid catches fire underneath the vessel, which causes the vessel to heat up and the pressure to rise. After some time has elapsed (typically half an hour) the vessel fails due to creep. The contents are then discharged to the atmosphere and are ignited by the small fire, exploding in a classic mushroom fireball. The blast effects from such explosions are not usually too severe, but debris may be propelled over a wide range (Fig. 5.1); in some cases a 'domino' effect has occurred with missiles from one explosion initiating further failures and explosions (Advisory Committee on Major Hazards, 1984). In the USA alone, twelve BLEVEs occurred in the period 1970 to 1975.

To help prevent pool fires spreading, bunds may be built around storage vessels; however, the existence of a bund may encourage BLEVEs to occur, since the burning liquid under the vessel cannot escape. For this reason, the ground within the bund should slope, and drains should be installed to drain any leaking liquid away to a safe place.

Strehlow and Baker (1976) quote correlations which enable the size and duration of the fireball in a BLEVE to be determined. If the mass of combustible is m kg, then the fireball diameter D (m) and duration t (sec) are given by:

$$D = 3.86\, m^{0.32} \qquad [5.1]$$

$$\text{and } t = 0.299\, m^{0.32} \qquad [5.2]$$

5.1.2(ii) *Enclosed deflagrations*

If a deflagration occurs within a building, the degree of overpressure is greatly increased, and hence relatively small quantities of material can cause severe damage. Domestic gas explosions are typical examples of enclosed deflagrations. The Abbeystead (Lancashire) explosion in 1984, in which a water pumping station exploded killing 16 people following an accumulation of methane, is a further example.

Buildings in which there is a high risk of such accidents, such as grain stores where grain dust forms a hazard, should be designed to allow for explosion relief.

5.1.2(iii) *Unconfined Vapour Cloud Explosions (UVCEs)*

UVCEs are similar in origin to flash fires, but their consequences are radically different. In an UVCE, a cloud of flammable vapour is released and a period of time elapses, during which the vapour mixes with air, before ignition occurs. The mixture of vapour and air then *detonates* producing a shock wave with considerable overpressure.

Fig. 5.1 In Boiling Liquid Expanding Vapour Explosions (BLEVEs), the storage tank normally ruptures and a 'fireball' explosion ensues. In this case in New Jersey in 1949, however, the tank, containing 106 T of hydrocarbon, rocketed a distance of several hundred metres. There were four fatalities (photo courtesy NFPA)

UVCEs are the most devastating of chemical plant accidents, in terms of explosive violence. The best-known and most documented UVCE occurred at Flixborough, near Scunthorpe, in 1974, in which approximately 36 tonnes of cyclohexane exploded, destroying a chemical plant, killing 28 people and breaking some windows up to 12 kilometres away. Lampposts near the explosion were bent through a full right angle. Gugan (1979) has equated the blast effects to 18 tonnes of TNT. Blast damage from an UVCE can affect a much larger area than a flash fire involving a similar quantity of material; in the Los Alfaques flash fire, little damage occurred beyond the confines of the burning vapour cloud. The criteria for whether a vapour cloud burns or detonates will be examined in Section 5.2.3.

(Both the Flixborough and Los Alfaques accidents might be considered to have caused atypical death-tolls. The Flixborough accident would have killed many more if it had not occurred at the weekend, when relatively few people were in the vicinity. The Los Alfaques accident caused a freakishly high death-toll, because the vapour cloud burned directly over a campsite crowded with holidaymakers.)

5.1.2(iv) *Highly reactive materials*

Materials which may decompose suddenly and exothermically, such as high explosives, or mixtures of materials which may react together, such as solid-fuelled rocket propellants, present special handling problems. Such substances do not require atmospheric oxygen to support their reactions. TNT, ammonium nitrate and picric acid are typical examples; such materials are not restricted to military applications. Notable examples include an ammonium nitrate explosion at Oppau in Germany in 1921 which killed 430 people, and an explosion at an arsenal in Lanchow, China, in 1935 which killed 2000 people.

5.1.3 Physical explosions

In physical explosions, the source of the high pressure gas is either a pressurised vessel or sudden boiling of a liquid.

5.1.3(i) *Pressure vessel failure*

In the late nineteenth century, boiler explosions were fairly commonplace. Around the turn of the century, many industrialised countries began to pass laws making regular boiler inspections a legal requirement, and vessels began to be designed to standardised design codes. Since that time, the incidence of pressure vessel explosions has greatly reduced, but the potential for such explosions still exists.

The hazards resulting from such explosions are caused by (a) the shock wave of the explosion, (b) missiles ejected from the vessel during the explosion and (c) the possible toxic or flammable contents of pressure vessels in some process plant. In addition, the pressure vessel may be within a building. The explosive effect will then be magnified, as for an enclosed deflagration. Figure 5.2 illustrates the effect of a boiler explosion in a shoe factory in Brockton, Massachusetts, in 1905, in which 58 people died.

The mechanisms of fast fracture and crack growth have been discussed in chapter four. Another, more straightforward cause of failure is that of simple overpressurisa-

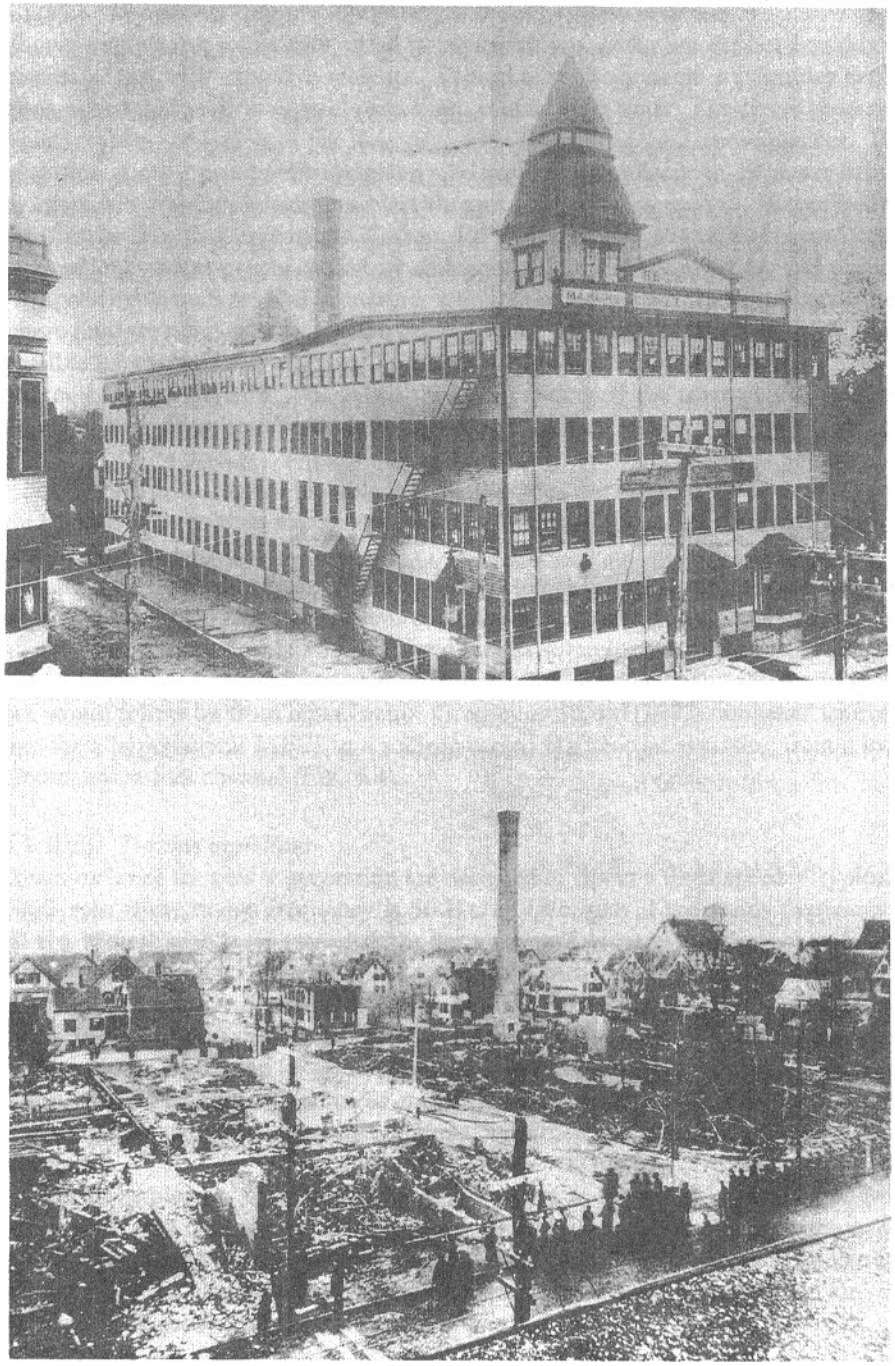
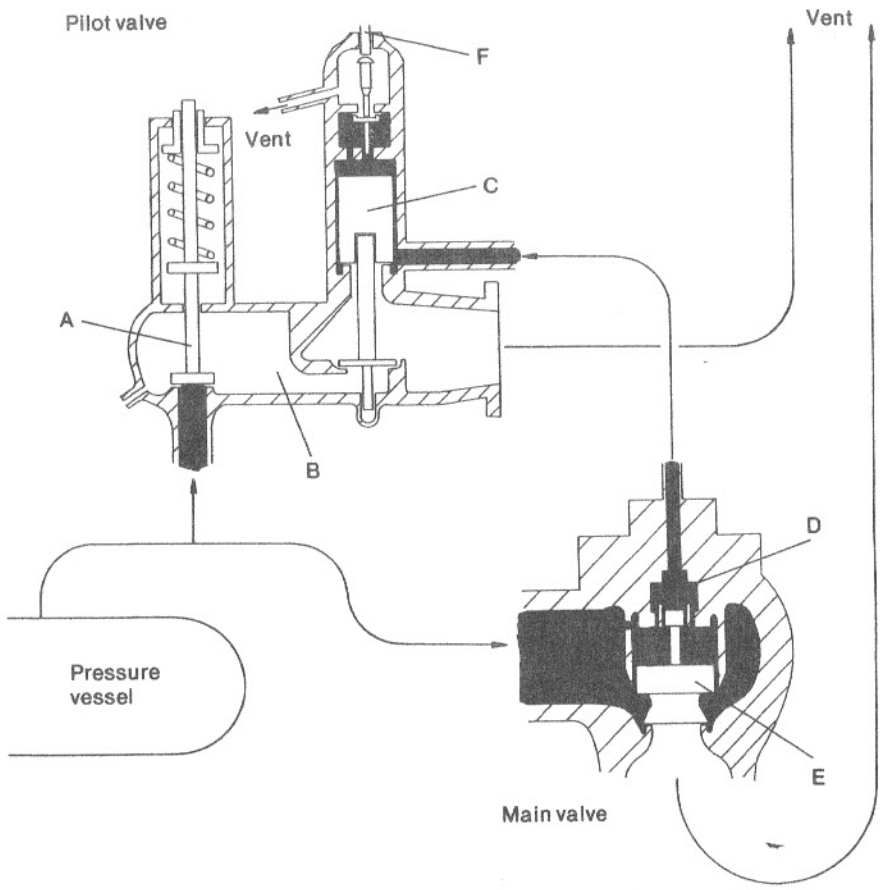


Fig. 5.2 A boiler pressure vessel exploded in a factory in Massachusetts in 1905. There were 58 fatalities (photos courtesy Hartford Steam Boiler Inspection & Insurance Company)

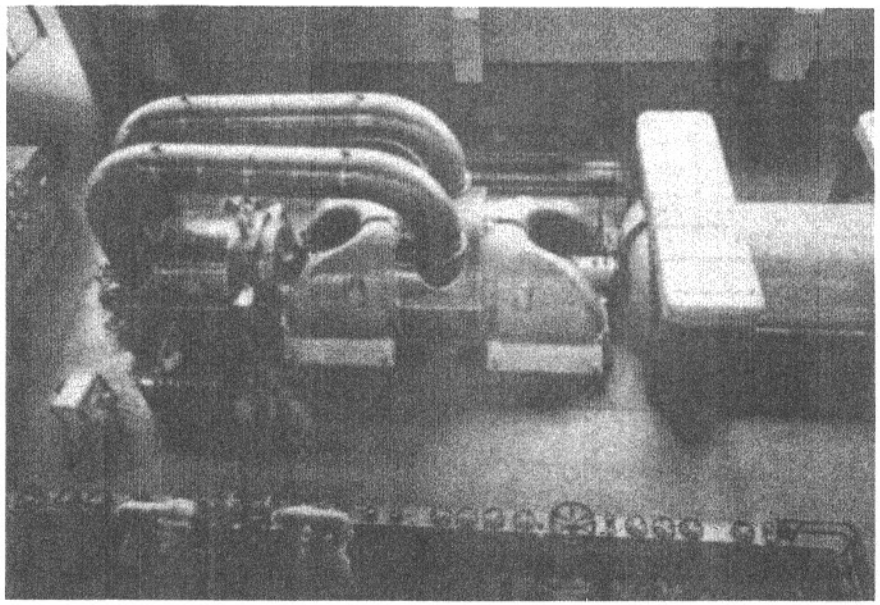
tion. This might occur due to maloperation, or a chemical reaction occurring when two substances are inadvertantly mixed in the vessel, or an autocatalytic chemical process 'running away'. To prevent overpressurisation, pressure vessels are fitted with safety valves which open to relieve the pressure if a preset value is exceeded. A typical safety valve is illustrated in Fig. 5.3.

The energy of pressure vessel explosions is normally divided about equally between shock wave (blast) energy and the kinetic energy of the vessel fragments (Lees, 1980).



Operation of a pilot-operated safety valve:
High pressure in the pressure vessel causes the pilot valve A to lift, thereby pressurising chamber B and causing piston C to lift. This vents chamber D, causing the main valve E to lift and depressurise the pressure vessel. The main valve will also lift if the plunger F is depressed

Fig. 5.3 A typical pilot-operated safety valve. Shaded volumes are normally pressurised



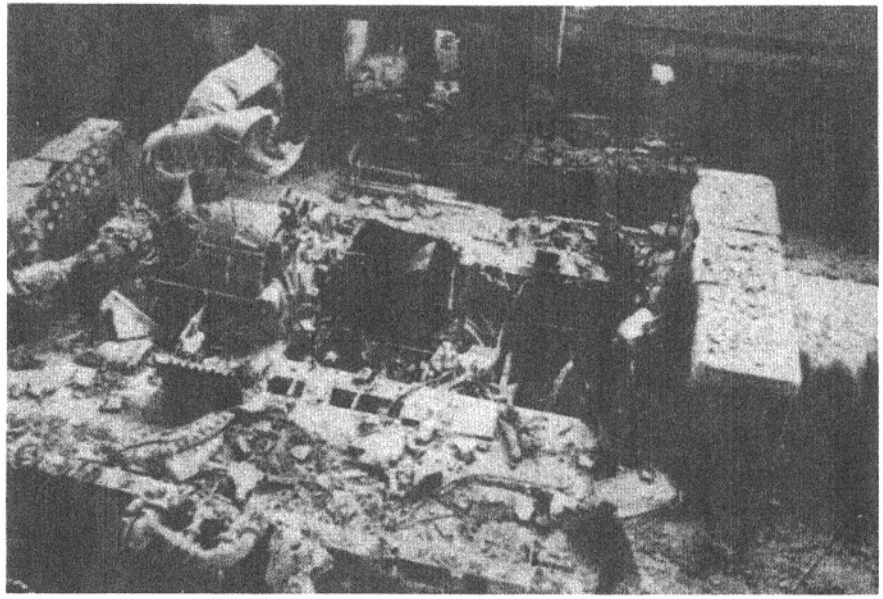


Fig. 5.5 The Uskmouth turbine accident (Reproduced by permission of the Council of the Institution of Mechanical Engineers. The author is grateful for the assistance of Professor Bernard Crossland and Sir Arnold Lindley)

Due to an extremely high state of awareness and preparation in the local authorities, prompt action was taken to evacuate 250 000 people up to 16 km downwind of the accident. Because of their prompt action, and a measure of good luck, there were no fatalities.

The Bhopal tragedy in Madhya Pradesh, India, in 1984 was the accident that Mississuaga might have been. Due to maloperation or poor maintenance*, water was allowed to pass via a temporary interconnecting line into a storage tank containing 30 tonnes of methyl isocyanate (MIC). MIC decomposes exothermically in the presence of water. The resulting pressure surge blew a bursting disc and overwhelmed a vent gas scrubber, which was not designed for such large discharges. The accident occurred at night and 2500 people living nearby were killed as they slept. Many thousands more were blinded or suffered lung damage. (MacKenzie, 1985; Union Carbide, 1985). Epidemiological studies for long-term effects are in progress.

5.1.4(ii) *Carcinogenic materials*

More than 1500 potentially carcinogenic materials have been identified (Schirripa, 1977). The safe handling of these materials presents difficulties, since it is generally held on theoretical grounds that there is no threshold dose for carcinogens to induce cancers (Royal Society, 1983). Hence zero risk can only be obtained by zero exposure, which may not be a practicable proposition. It is instead necessary to define an 'acceptable' dose, based on dose–risk data from animal studies and an assumption concerning 'acceptable' individual risk.

Such animal studies generally take the form of small-scale studies at high doses. Some functional relationship between dose and risk then has to be assumed so that the experimental data can be extrapolated down to low doses. The chosen functional relationship is often based on features of the biological mechanism of carcinogenesis; such relationships include the so-called 'one-hit' and 'multi-hit' models, where the individual risk equals

$$\{1 - \exp[-(q_0 + q_1 d + q_2 d^2 + \ldots)]\}$$

where d is the dose and the q_i's are coefficients. (In the one-hit model, all q_i's except q_0 and q_1 are zero.) These models are discussed by, e.g. Crump and Crockett (1985).

It follows, therefore, that for a one-hit model the dose–risk relationship is linear at low doses (cf. eqn [2.57]). For a two-hit model, i.e. with q_2 non-zero, an assumption of linear dose–response behaviour should not underestimate the consequences at low doses (Pochin, 1978).

Using reasoning such as this, the results of animal studies can be scaled up in proportion to body weight, and after the application of a factor of safety the results can be applied to man. Threshold Limit Values (TLVs) for carcinogens can then be used for guidance concerning safe exposure in the place of work (ACGIH, 1980).

The Seveso accident in Italy in 1976 is an example of an accident involving such a material. A runaway chemical reaction in a reactor producing trichlorophenol caused a bursting disc failure, and the contents of the reactor, which contained approximately

*See footnote on p. 76

2 kg of tetrachlorodibenzo-*p*-dioxin byproduct (TCDD or dioxin), were released to the environment. Dioxin has been proved to be fatal to experimental animals in doses of 10^{-9} times body weight, as well as being carcinogenic and mutagenic (Marshall, 1976, 1980). There were no prompt deaths in the Seveso accident, but the long-term effects are still unclear. (Dioxin impurity in 'Agent Orange', a defoliant used by US forces in the Vietnam war, is believed to have been responsible for birth deformities, a higher rate of spontaneous abortion and an increase in the incidence of cancer, although a recent report has suggested that the herbicide itself, rather than the dioxin impurity, was responsible (Newell, 1985).) Margerison and Wallace (1981), two journalists who investigated the Seveso accident, concluded that inadequate health records, as well as inadequate post-accident monitoring of nearby residents who were exposed to the release of dioxin, make it most likely that no firm conclusion about the effects of the accident will ever be drawn.

5.1.5 Nuclear accidents

5.1.5(i) *Nuclear reactor accidents*

Nuclear reactors cannot, under any circumstances, explode like nuclear bombs. Fetter and Tsipis (1981) state that "even in a reactor that is entirely out of control, the rate at which energy is released is more than 10^{12} times slower than it is in a nuclear weapon". The hazard of nuclear reactors lies instead in the possibility of accidental release of radioactivity to the environment. To prevent this, nuclear reactors must satisfy each of three criteria:

1. Reliable shutdown.
2. Reliable removal of decay heat. (This is the heat which continues to be generated by the radioactive decay of fission products after the neutron chain reaction has been stopped.)
3. Reliable containment of fission products and fuel.

Reliable shutdown is ensured by means of control rods, made of neutron absorbing materials, which are suspended above the reactor cone by means of electromagnets. The current to the electromagnets is controlled by a redundant, multiply-diverse Automatic Protective System (APS) using 2-from-3 majority voting logic. When the APS breaks the current to the electromagnets, the control rods drop into the core. Reliable decay heat removal is ensured by means of auxiliary boilers, dedicated decay heat rejection thermosyphons or emergency core cooling systems. Reliable containment is achieved by multiple containment barriers; the fuel pin cladding, the reactor pressure vessel and the containment building.

The Three Mile Island accident in Pennsylvania in 1979 occurred (*inter alia*) because of a failure to remove decay heat and a failure of containment. The accident at TMI caused an estimated release of 0.125 mg (16 Curies) of iodine-131. This may cause one extra cancer among the several hundred thousand nearby residents in years to come (Roberts, 1984). Hypothetically, Fremlin (1983), has calculated that an extremely severe nuclear reactor accident – a 1 GW(e) reactor volatilising one-third of its radioactivity into a 24 km/hr wind – might cause 100 to 200 prompt deaths and an

extra 200 cancers over the succeeding decades, if the population density around the plant was the same as the UK average population density. The Chernobyl accident in the Ukraine in April 1986 appears to have been of this order of magnitude. However, design features of this reactor (the RBMK light water-graphite design) would rule out its use in most countries, on safety grounds. Evidence given at the enquiry into the construction of a pressurised water reactor (PWR) at Sizewell, England, has suggested that the amount of radioactivity discharged in a hypothetical severe PWR accident might be less than this by a factor of ten or greater, with a correspondingly smaller death-toll resulting from the accident.

5.1.5(ii) *Accidental criticality*

In the processing and storage of nuclear fuel, the possibility of accidentally placing fuel in a critical configuration, such that a self-sustaining fission chain reaction can be supported, must not be excluded. Such 'criticality accidents' are effectively impossible with natural uranium, which contains only 0.7 per cent fissionable uranium-235. However, such accidents become possible with the handling and processing of enriched uranium or plutonium. There is no question of a nuclear explosion occurring, but the release of neutrons and gamma rays from a criticality will kill anyone standing nearby, unless they are behind shielding.

Nicholls, Woodcock and Gillieson (1961) give a useful introduction to the topic. About a dozen accidental criticalities have occurred in fuel processing plants in the western world to date, mostly in the United States. Two plant operators have been killed. No members of the public have been affected.

5.1.6 Dam failures

The worst industrial accident in history occurred in 1979 in Gujarat, India, when a dam failure caused between 5000 and 15 000 deaths. Although the failure probability of individual dams is difficult to estimate, statistics for dam failures suggest a hazard rate of about 10^{-4} per dam-year (Gast, 1973); in earthquake regions this figure may be expected to be greater. Ayyaswamy (1974) has estimated the effects of hypothetical instantaneous failures of some Californian dams. A failure of the Folsom dam might cause 260 000 deaths. An earthquake in the region could initiate such a failure.

In the UK, there have been no dam failures since the Reservoirs (Safety Provisions) Act was passed in 1930. Dams were, as a result of this Act, the first structures in the UK whose security was governed by Act of Parliament (Smith, 1972).

5.1.7 Aircraft crashes – risk on ground

The 'survivability' of aircraft crashes is low. Aircraft inherently fail dangerously, and for such crashes it may be assumed that reliability equals safety at least as far as the aircraft occupants are concerned.

However, a possibility that should often be considered in the safety assessment of chemical or nuclear plant is that of the installation being struck by a crashing aircraft.

The consequences of such an accident might be severe, even though the probability is very small.

The surface area of the United Kingdom is 2.38×10^5 square kilometres. Purely random civil aircraft crashes (as opposed to crashed near airports or on air force training routes, where it may be assumed that hazardous installations would not be built) occur with a frequency of about 3.5 per year (Hall, Philips and Peckover, 1985). If an area of 10^4 square metres is affected by each crash, then the probability of an installation being struck by a crashing aircraft is $(3.5 \times 10^4 \div 2.38 \times 10^{11})$, or 1.5×10^{-7}/annum, approximately.

5.1.8 Earthquakes

Aseismic design – the design of engineering installations to resist earthquake damage – is a large topic in its own right. (See, for example, Newmark and Rosenbleuth, 1971.) Typically, a structure is designed to resist the '10 000 year' earthquake for its locality; that is, the maximum magnitude of earthquake that can be expected in a 10 000 year interval.

The magnitude M of an earthquake is a measure of the maximum ground motion. An exact definition is 'the common logarithm of the peak trace amplitude (in microns) of a Wood–Anderson seismograph located on firm ground 100 km from the epicentre'.

Earthquake magnitude M and frequency N (yr^{-1}) may be related as follows (Alderson, 1979):

$$\log_{10} N = a - bM \qquad [5.3]$$

where a and b are constants for any given locality. In the UK, a and b are equal to 4.13 and 1.09, respectively.

For aseismic design, however, it is the peak ground acceleration that is of more interest than the earthquake magnitude. The peak ground acceleration associated with a 10 000 year earthquake varies widely between different parts of the world. In the UK, this acceleration is 2.5 m/s^2; in Japan the level is 6.4 m/s^2 (Hall, Philips and Peckover, 1985). From these values of acceleration, the additional loadings imposed by the earthquake may be determined, in order to arrive at a design which will survive the earthquake.

Peak ground acceleration $\dot{v}$ (m/s^2) has been correlated with earthquake magnitude M and the focal distance r (the distance from the earthquake epicentre, in kilometres). Newmark and Rosenbleuth give

$$\dot{v} = 12.3\, e^{0.8M} (r + 25)^{-2} \qquad [5.4]$$

5.2 *Explosions*

In Section 5.1.2, an explosion was defined as 'a sudden release of high pressure gas to the environment'. In this section, the thermodynamic and kinetic criteria* for the occurrence of unconfined vapour cloud explosions, the blast damage resulting from confined and unconfined explosions, and blast scaling and missile damage will be discussed.

* Only a brief introduction to chemical thermodynamics and kinetics will be given here. Those readers requiring a more detailed introduction are recommended to read, e.g. Warn (1969) or Stull (1977).

5.2.1 Basic combustion chemistry of flammable gas–air mixtures

5.2.1(i) *Chemical thermodynamics*

During any chemical reaction, there is a change in enthalpy ΔH (i.e. heat is either released to, or absorbed from, the environment) and a change in entropy ΔS. Entropy may be thought of as the 'degree of randomness' in a system, and, by the second law of thermodynamics, entropy increases in an irreversible reaction. Thus, we may write a general reaction

$$a\,\mathrm{A} + b\,\mathrm{B} \rightleftharpoons c\,\mathrm{C} + d\,\mathrm{D}, \pm \Delta H, + \Delta S$$

The change in enthalpy ΔH is called the heat of reaction. ΔH is positive for an endothermic reaction and negative for an exothermic reaction.

In some reactions, e.g. the dissolving of crystals in water, the enthalpy change may be endothermic (heat absorbed from surroundings) while the entropy change is positive. The criterion for whether a reaction will occur spontaneously or not is given by Gibb's Free Energy ΔG,

$$\Delta G = \Delta H - T\Delta S \qquad [5.5]$$

where the temperature T is in Kelvin.

If ΔG is negative, the reaction will occur spontaneously.

If ΔH is negative, the reaction releases heat. Enthalpy and entropy may thus be seen as two forces which sometimes compete to decide whether a reaction will proceed or not.

From knowledge of ΔH and the specific heat for the product c_p, it is possible to calculate an adiabatic reaction temperature change ΔT_r, by assuming that all heat generated goes towards heating up the products, and that no heat is lost to the environment.

$$\Delta T_r = \Delta H \left[\sum_{i=1}^{n} \frac{N_i}{c_{pi}} \right] \qquad [5.6]$$

where ΔH is the heat of reaction per kg of product mixture,

N_i is the number of kmol formed of the ith reaction product per kmol of reactant, and

c_{pi} is the specific heat of the ith reaction product.

For constant volume, the peak reaction pressure may now be calculated from the gas law:

$$p_r = \left[\frac{T_r}{T_o} \right] \cdot \left[\frac{N_r}{N_o} \right] p_o \qquad [5.7]$$

where N_o and N_r are the number of kmol of gas before and after the reaction, respectively, and p_o and p_r are the pressures before and after the reaction.
The heat of reaction is determined from

$$\Delta H = (\Sigma \Delta H_f(\text{products})) - (\Sigma \Delta H_f(\text{reactants})) \qquad [5.8]$$

where ΔH_f is the heat of formation of a substance.
Values for ΔH_f may be obtained from books of data, e.g. Aylward and Finlay (1971).

5.2.1(ii) *Chemical equilibrium*

Consider again the reaction

$$a\text{A} + b\text{B} \rightleftharpoons c\text{C} + d\text{D} \qquad [5.9]$$

The relative amounts of A, B, C and D present after the reaction has occurred define an equilibrium constant K_r, given by

$$K_r = \frac{(C_C)^c\,(C_D)^d}{(C_A)^a\,(C_B)^b} \qquad [5.10]$$

where C_A, C_B, etc. are the molar concentrations (or partial pressures) of the reactants and products at equilibrium. (*Note:* These should properly be replaced by their relevant activities.)

The Gibb's Free Energy change ΔG will be given by

$$\Delta G = (\Sigma \Delta G_f(\text{products})) - (\Sigma \Delta G_f(\text{reactants}))$$

where ΔG_f is the Free Energy of formation, which may be obtained from books of chemical data.
ΔG is related to the equilibrium constant K_r in the following manner:

$$\Delta G^\circ = -RT \ln K_r \qquad [5.11]$$

where ΔG° is the Standard Free Energy change (tabulated).

The Standard Free Energy change ΔG° is that which occurs at 298 K and 1 atm. pressure. Under other conditions, the overall change in Gibb's Free Energy ΔG_r can be determined thus:

$$\Delta G_r = \Delta G^\circ + RT \ln K_r \qquad [5.11a]$$

Hence, if ΔG_r has a large negative value, the reaction will proceed almost to completion. For explosions this will always be the case.

Thus, we may say that chemical thermodynamics can tell us *whether or not a reaction will proceed.* It also enables us to determine *to what extent the reaction will proceed,* i.e. the point of equilibrium. However, it tells us nothing about *the speed (or kinetics) of the reaction.*

5.2.1(iii) *Reaction kinetics*

The overall stoichiometric reaction [5.9] almost never represents the actual reaction mechanism. However, if the reaction is as represented, then the reaction is said to be 'of order $(a + b)$', i.e.

$$\rho = k(C_A)^a(C_B)^b \qquad [5.12]$$

where ρ is the reaction rate and
k is the reaction rate coefficient.

Most elementary reactions are either first or second order. Very few third-order reactions are known. A genuine third-order reaction implies a simultaneous meeting of three molecules in space, which is an unlikely event.

The reaction rate coefficient k normally varies with temperature according to the (experimental) Arrhenius equation:

$$k = \alpha \exp(-E/RT) \qquad [5.13]$$

where α is the frequency factor (an experimental coefficient), and
E is the activation energy.

The activation energy E is so called because it is associated with an energy barrier which the reactants must surmount before the reaction can occur (see Fig. 5.6).

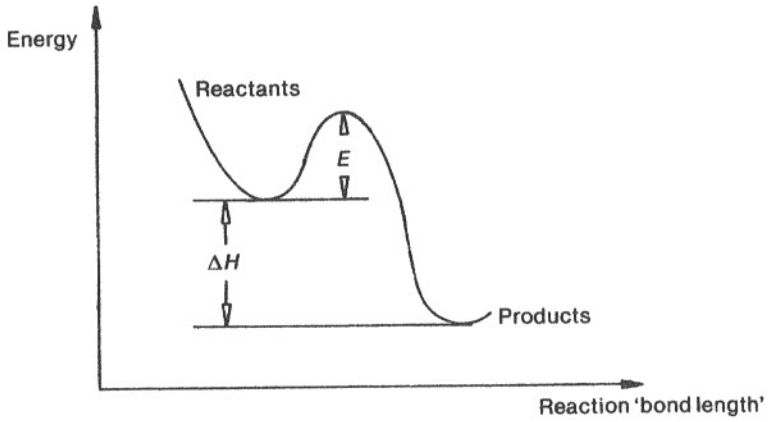


Fig. 5.6 Activation energy E and enthalpy change ΔH for an exothermic reaction

Figure 5.6 shows that if the reaction is exothermic then the products have a lower energy than the reactants. The energy release ΔH may then raise the temperature of the remaining reactants, thereby causing the reaction rate to accelerate (via the Arrhenius equation).
Other points to note are:

1. The magnitude of E is independent of ΔG or ΔH.

2. If E is suitable large, the reactants may be considered to be stable.
3. The effect of catalysts is to reduce the activation energy E.

To summarise so far, we may say that combustion is possible if ΔG_r is negative, ΔH is negative and E is small with respect to ΔH.

It should be noted that if ΔG_f for a compound is positive then the compound is thermodynamically unstable, i.e. it may decompose. If the products of decomposition are gaseous, an explosion is possible.

5.2.1(iv) *Explosion hazard potential*

Stull (1969) has proposed a means of classifying chemicals according to their potential fire or explosion hazard, using both thermodynamic and kinetic data. Stull used the adiabatic decomposition temperature T_d (a measure of the energy released in a reaction) and the activation energy E to propose a Reaction Hazard Index (RHI), defined thus:

$$\text{RHI} = \frac{10\ T_d}{T_d + 30\ E} \qquad [5.14]$$

where T_d is in Kelvin and the activation energy E is in kcal/mole. This index correlates well with the more arbitrary chemical hazard ratings given by the US National Fire Protection Association. High RHI values indicate that the substances or mixtures under consideration are particularly unstable and liable to detonate. Some values for RHI are included in Table 5.1.

Other indices for quantifying fire and explosion hazards have been proposed, most notably the Dow Fire and Explosion Index (Dow 1980). This uses thermodynamic data (the 'material factor'), as well as specific process hazards (e.g. high-pressure process plant), to produce an index between zero and one hundred depending on the degree of hazard. Preventive and protective measures are then prescribed according to the size of the index for a given process.

Other hazard indexing systems have been reviewed by Jones (1978). A handbook of reactive chemical hazards has been prepared by Bretherick (1975).

Table 5.1 Flammability data for some hydrocarbons in air at 1 bar and 25°C

Substance	AIT(°C)	$-\Delta H_c$(MJ/kg)	LFL(% by vol.)	UFL(% by vol.)	T_d(K)	E(kcal/mole)	RHI
Methane	537	55.6	5.0	15.0	298	103.0	0.88
Ethane	472	51.9	2.9	12.5	597	89.5	1.82
Propane	450	50.4	2.1	9.5	626	63.3	2.48
Butane (def.)	288	49.6	1.8	8.3	633	86.3	1.96
(det.)	288	49.6	2.9	5.2	—	—	—
Octane	206	47.9	0.95	—	—	—	—
Decane	201	47.7	0.75	5.6	—	—	—
Hexadecane	202	47.3	0.43	—	—	—	—
Cyclohexane	245	36.3	1.3	9.1	677	64.1	2.60
Hydrogen (def.)		142.5	4.1	74.0	—	—	—
(det.)		142.5	18.3	59.0	—	—	—
Kerosene	254		1.16	6.0	—	—	—
Benzene	538	32.0	1.4	8.0	—	—	—
Xylene	496	42.9	1.1	7.0	817	79.5	2.55
Ethanol	394	29.7	3.3	19.0	—	—	—
Acetone (Propanone)	538	30.8	2.6	12.8	774	68.0	2.75
Butanone	516	33.9	1.8	11.5	755	67.2	2.72

(def.) means 'deflagration', (det.) means 'detonation'

Refs: Zabetakis (1965), Stull (1977), Gugan (1979), Lees (1980)

5.2.2 Flammability of gas–air mixtures

All mixtures of combustible gases in air have upper and lower flammability limits (UFL and LFL) which vary according to pressure and flame temperature (Fig. 5.7). These limits also vary according to the initial bulk temperature of the gas mixture. However the observed LFL and UFL usually correspond to the same flame temperature, called the *flame threshold temperature* to produce a self-sustaining reaction. (The adiabatic flame temperature may be calculated from the heat of reaction and the specific heat of the products (at constant pressure), in a similar manner to eqn [5.6].)

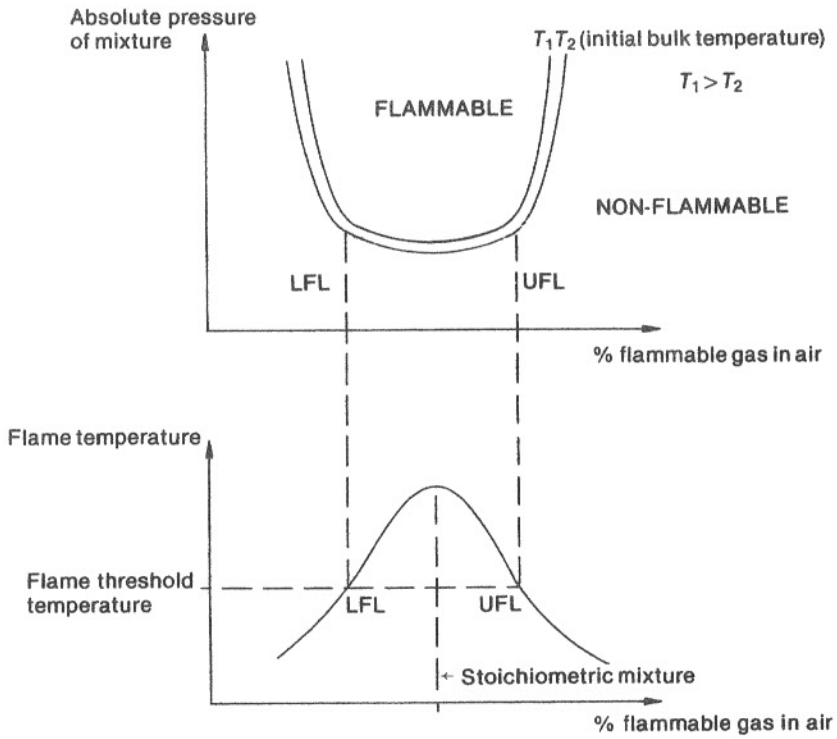


Fig. 5.7 Flammability and flame temperature of gas–air mixtures

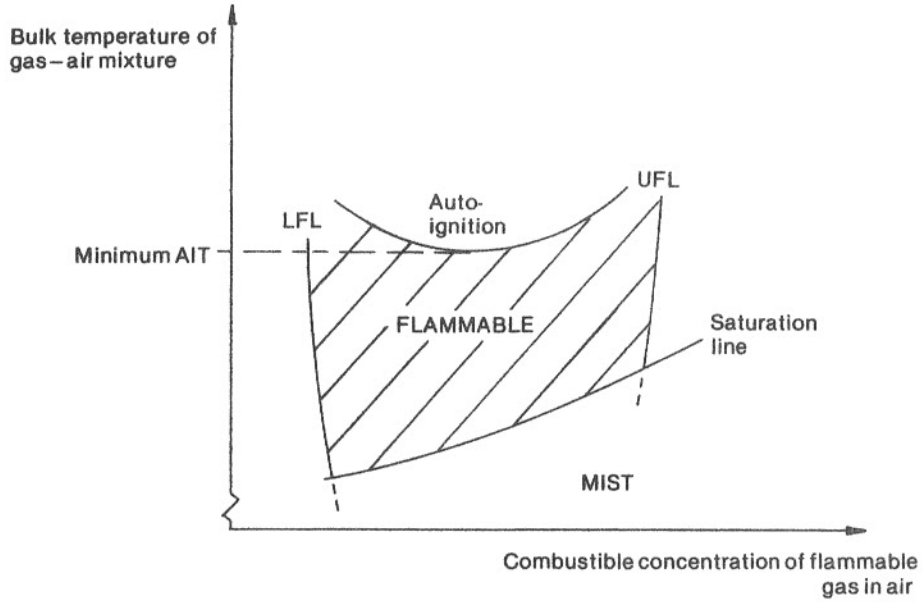


Fig. 5.8 The effects of temperature and concentration upon flammability

The flame threshold temperatures for hydrocarbon–air mixtures are in the range 1500 ± 150 K (Zabetakis, 1965). The maximum flame temperature occurs when the gas–air mixture ratio is stoichiometric.

Figure 5.8 shows how the combustion characteristics of a flammable gas–air mixture vary with concentration and bulk temperature. It can be seen that, at low temperatures, the flammable gas may condense and form a mist. The flammability of such a mist depends upon the size of the vapour droplets. For fine droplets, the combustible concentration at the lower flammability limit (LFL) is about the same as for uniform vapour–air mixtures. For larger droplets, though, the LFL may actually decrease under certain circumstances, since coarse droplets may fall towards the flame front, vaporise and locally increase the concentration.

The most commonly encountered mixtures are above the saturation line shown in Fig. 5.8. Here it can be seen that the range of concentrations (UFL–LFL) steadily widens as the temperature increases, until the Auto-Ignition Temperature (AIT) for that particular concentration is reached. At this temperature, the mixture spontaneously ignites. The AIT therefore represents an additional hazard in the handling of flammable materials. Table 5.1 presents data on AITs, LFLs and UFLs for some common flammable liquids and gases.

The burning velocity is the velocity of propagation of the flame front with respect to the unburnt gases. Ideal burning velocities can be calculated from mathematical models and reaction kinetics, and agree well with experiment. The velocities are typically in the range 0.10–2.0 m/s for air–fuel mixtures. The burning velocity

represents a minimum velocity for the flame front. The flame front accelerates to subsonic velocity in a deflagration, and supersonic velocity in a detonation.

5.2.3 Deflagration or detonation?

We are now ready to consider a problem that has already been introduced in Section 5.1.2(iii); what criteria determine whether a cloud of flammable gas will deflagrate (in a flash fire) or detonate (in an Unconfined Vapour Cloud Explosion)? The consequences of these different forms of combustion are very different. In a deflagration, the flame front moves subsonically through the gases, and the overpressure caused by the combustion is small. However, in a detonation, the flame front accelerates to supersonic velocities, thereby producing a shock wave. Overpressures of about 20 bar are typical for gas phase detonations in air.

The transition from deflagration to detonation has been discussed and considered by many investigators, e.g. Stull (1977), Strehlow and Baker (1976) and Zabetakis (1965). In principle, the combustion process generates hot expanding gas, which forces the flame front to move faster; the accelerated flame front then generates gaseous combustion products more quickly, and so the flame front is forced to accelerate further. For a detonation, the flame front continues accelerating until it becomes a shock wave, capable of causing auto-ignition of the gas ahead of the deflagration. The presence of turbulence may assist this transition. Table 5.2 gives some data on detonation velocities and overpressures resulting from detonations of hydrocarbons in air.

Table 5.2 Detonation velocities and overpressures for hydrocarbon–air detonations

Mixture	*Mixture Ratio*(%)	*Detonation Velocity*(km/sec)	*Overpressure*(×10⁵Pa)
H_2–air	29.53	1.96	14.6
CH_4–air	9.48	1.80	16.05
C_2H_2–air	7.73	1.87	18.0
C_2H_4–air	6.53	1.86	17.74
C_3H_8–air	4.0	1.80	17.27

Ref: Lee *et al.* (1977)

The problem of whether a given cloud of flammable gas will deflagrate or detonate when ignited remains unresolved. Kletz (1977) has suggested the following rough-and-ready criteria for a detonation to occur:

1. The vapour cloud must be large; at least five tonnes of hydrocarbon, and possibly as much as 15 or 20 tonnes, are required for detonation.
2. The release rate of vapour must be high, probably more than 1 tonne/minute.
3. There must be a delay before ignition to allow the vapour to mix with air so that the mixture ratio is between the lower and upper flammability limits.

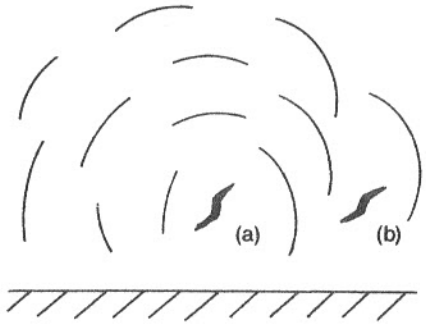


Fig. 5.9 Ignition sources in a vapour cloud

It is also probable that the position of the ignition source in relation to the cloud is important (Fig. 5.9). If the ignition source is in the centre of the cloud (a), the overpressure of the combustion products forces the flame front out into the cloud, and the flame front accelerates, possibly to yield a detonation. However, if the ignition source is near the edge of the cloud (b), the combustion products are free to disperse to the atmosphere without causing the flame front to accelerate.

Kletz also proposes a relationship between the probability that the vapour cloud will ignite, and the mass of vapour in the cloud (Fig. 5.10). For small leaks, the ignition probability is about 0.03, but for large leaks the ignition probability may rise to as much as 0.5.

Furthermore, Kletz suggests that only one in ten vapour cloud ignitions develop into a detonation. Hence, given that a large leak has occurred, the probability of a detonation (or UVCE) is about 0.05.

Finally, some experimental results suggest that the range of possible fuel–air mixture ratios (UFL–LFL) for detonation is narrower than the range of ratios for deflagration (see Table 5.1). However, reliable and repeatable data for this are not available. Thus, a conservative approach to safety would assume that a mixture which burns can also detonate.

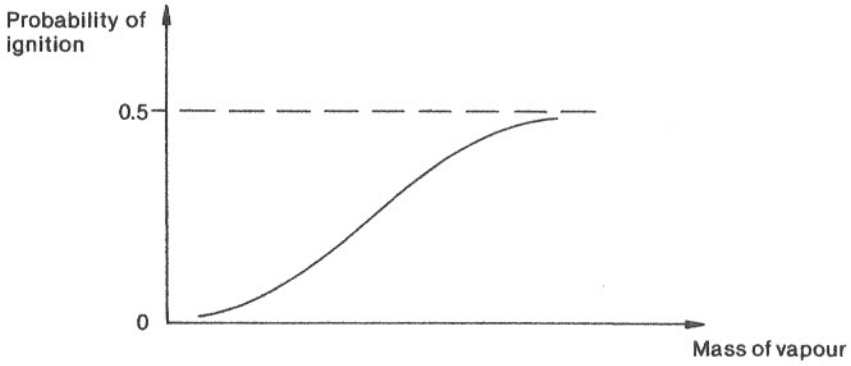


Fig. 5.10 Ignition probability versus vapour cloud

5.2.4 TNT equivalent and yield for detonations

The 'TNT equivalent' for a flammable gas–air detonation may be calculated as follows:

$$\text{TNT equivalent mass (kg)} = \varepsilon.m.\ \frac{\Delta H_c}{\Delta H_{TNT}} \qquad [5.15]$$

where ε is the *yield* of the explosion,
m is the mass of flammable gas,
ΔH_c is the heat of combustion for the flammable gas (MJ/kg), and
ΔH_{TNT} is the energy released per kg for a TNT explosion (4.773 MJ/kg).

Gugan (1979) has plotted yield against combustible mass for 22 UVCEs (Fig. 5.11). This shows no correlation but it can be seen that a yield of 10 per cent is seldom exceeded. Thus $\varepsilon = 0.10$ seems a reasonable value for calculating potential damage due to UVCEs. Brasie and Simpson (1968) suggest that lower values of yield should be used to calculate potential damage – as low as $\varepsilon = 0.03$. The more conservative value of $\varepsilon = 0.10$ is a better assumption for the purpose of safety assessment.

Values for ΔH_c for some common hydrocarbons are included in Table 5.1.

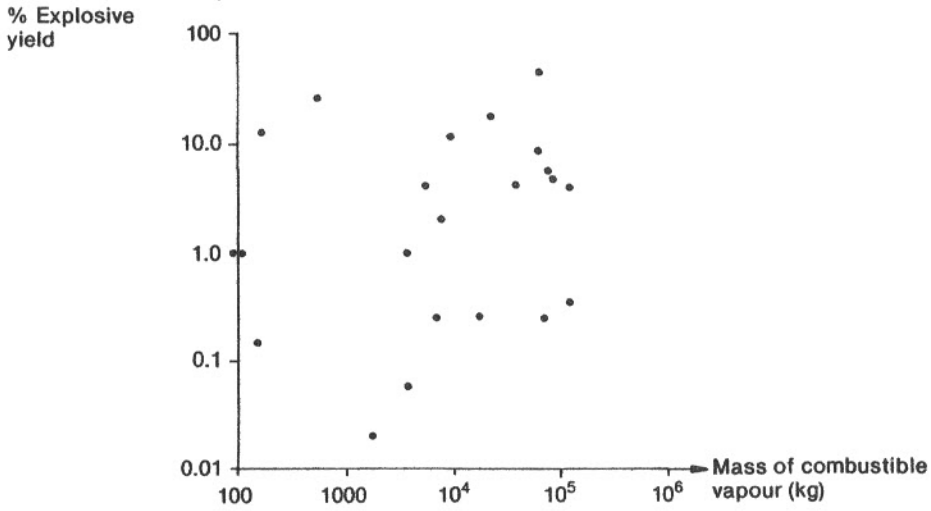


Fig. 5.11 Explosive yield versus combustible mass for 22 Unconfined Vapour Cloud Explosions (Gugan, 1979)

5.2.5 The effects of explosive blast

5.2.5(i) *Blast scaling*

The overpressure Δp caused by an explosion of m kg of explosive at a distance r metres from the explosion, obeys a fixed relationship such that

$$\Delta p = \Delta p(\Gamma) \qquad [5.16]$$

where

$$\Gamma = \frac{r}{m^{1/3}} \qquad [5.17]$$

Γ is called the 'scaled range' (m/kg$^{1/3}$). This is known as 'Hopkinson scaling' after Bertram Hopkinson, who developed the principle during the First World War (Baker, 1973).

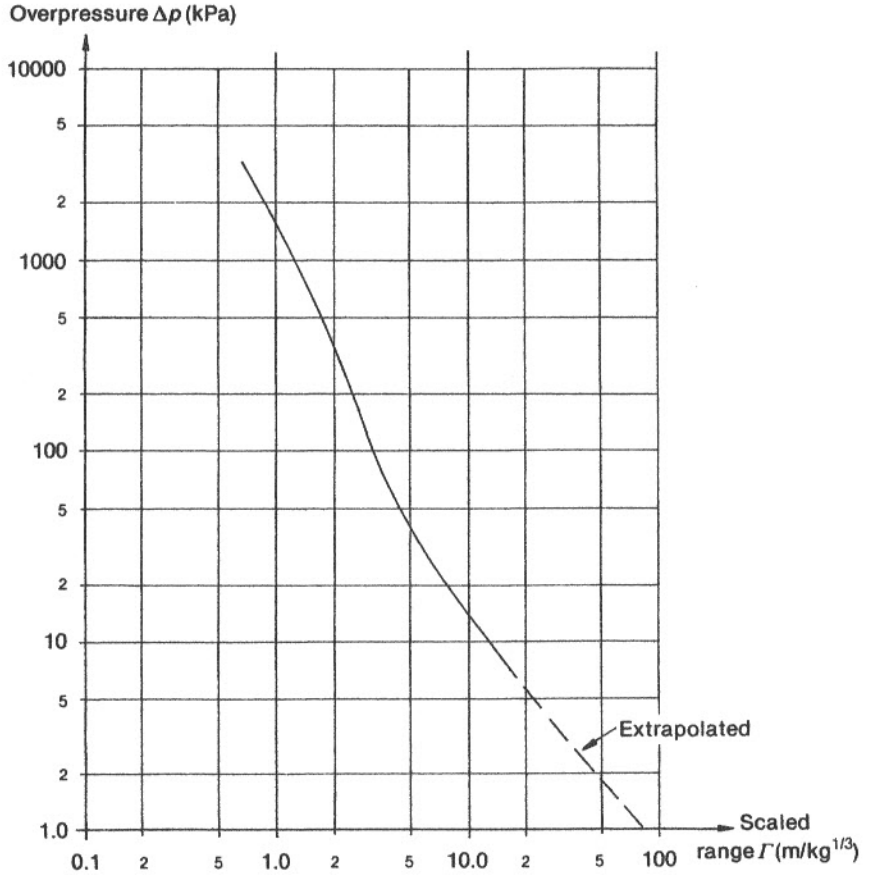


Fig. 5.12 The relationship between overpressure Δp and scaled range Γ for free surface blasts of TNT (Department of Defense, 1969)

Table 5.3 The effects of blast overpressure

Damage effect	*Approx. overpressure*
Crater lip	2000 kPa
Human LD_{50}	1000 kPa
Lung damage	400 kPa
Ear damage	100 kPa
Total building destruction	69 kPa
Loaded rail wagons overturned	48 kPa
Oil storage tanks rupture	27 kPa
50 per cent destruction of brickwork	17 kPa
Houses made uninhabitable	7 kPa
People knocked over	6 kPa
All windows shattered	5 kPa
Missile limit	1.5 kPa
Occasional window breakage	0.2 kPa

Refs: Gugan (1979), Brasie and Simpson (1968)

The side-on peak positive overpressure is a unique function of Γ for a given explosive in a free surface blast (Dept. of Defense, 1969). Figure 5.12 shows the relationship between overpressure Δp and scaled range Γ for TNT.

Figure 5.12 may thus be used in conjunction with eqns [5.15] and [5.17] to determine the peak overpressure resulting from the detonation of any given mass of flammable gas, by means of its 'TNT equivalent'.

5.2.5(ii) *The effects of blast overpressure*

The effects of blast are summarised in Table 5.3. It is notable that the LD_{50} overpressure (the overpressure that kills 50 per cent of people) is very high. The conclusion to be drawn from this is that direct blast effects kill relatively few people in explosions; flying debris or building collapse are a far more likely cause of death or injury. We are thus unable to say anything as categorical as 'nobody will be killed if they are more than x metres from a blast'.

Marshall (1977) examined past explosions in an attempt to determine a correlation between the size of an explosion and its lethality. Using data from 162 military and accidental explosions he proposed a 'mortality index', M_I thus:

$$M_I = 4m^{-1/2} \qquad [5.18]$$

where m is the mass, in tonnes, of combustible or explosive material (*not* the TNT equivalent mass). The mortality index M_I is then the number of fatalities per tonne of explosive. (See also Fig. 5.23).

Although this is obviously a rough-and-ready approach since it neglects personnel distribution at the time of the explosion, it probably represents the best means of estimating the cost (in lives) of accidental explosions. An approximate cost for damaged or destroyed property can be determined from overpressure calculations.

Example 5.1

As a result of a catastrophic storage vessel failure, 100 tonnes of ethane are released into the atmosphere. After a few minutes the ethane could explodes.

The available energy will be ($\Delta H_c \times$ mass), which equals 51.9×10^5MJ (from Table 5.1). Assuming a 10 per cent yield, eqn [5.15] gives a TNT equivalent mass of 107.7 tonnes.

The LD_{50} overpressure (from Table 5.3) occurs at a scaled range Γ equal to 1.2 m/kg$^{1/3}$. The actual radius for LD_{50} will therefore be 57 m in this case. The overpressure for building destruction occurs at a scaled range of 3.97 m/kg$^{1/3}$. The radius for building destruction will therefore be 189 m. Similarly, all windows will be shattered to a distance of 850 m from the blast.

Marshall's equation [5.18] suggests that the typical death toll from such an accident would be about 40.

It must be emphasised that the probability of such a disastrous explosion is low. If the probability of storage vessel failure is 5×10^{-5} per year (see Section 4.6.5) and the probability of a large leak detonating is 5×10^{-2} (Section 5.2.3) then the overall probability of such an accident would be about 2.5×10^{-6} per storage vessel-year.

5.2.5(iii) *Overpressure in enclosed deflagrations*

In an enclosed deflagration, the maximum pressure p_m is given by

$$p_m = p_1 \frac{M_1 T_2}{M_2 T_1} \qquad [5.19]$$

where p_1 is the initial pressure,

M_1 and M_2 are the initial and final average molecular weights, and

T_1 and T_2 are the initial and final (adiabatic) temperatures (Kelvin).

Zabetakis (1965) gives the following expression for the variation of pressure with time in the early stages of an enclosed deflagration:

$$\Delta p = K p_1 \frac{S^3 t^3}{V} \qquad [5.20]$$

where Δp is the pressure rise,

K is an experimental constant (typically equal to 70),

S is the burning velocity,

t is the time (consistent with Δp being less than $(p_m - p_1)$), and

V is the enclosed volume.

It is likely that a pressure vessel will burst if $p_m/p_1 > 4$. Hence any explosion relief device must act within a corresponding timescale, as given by t in eqn [5.20]. Typically, explosion relief should act within 0.1 second in a deflagration.

5.2.6 The energy of physical explosions

5.2.6(i) *The stored energy in pressure vessels*

Gas under pressure in a pressure vessel contains a large amount of stored energy. If the pressure vessel ruptures, the specific energy available for conversion to blast energy in an isothermal expansion will be:

$$E = \int_1^2 p \, dv$$

$$= RT \ln (p_1/p_2) \qquad [5.21]$$

for an ideal gas, where subscripts 1 and 2 refer to conditions before and after the explosion. The isothermal assumption leads to an overestimate of energy release. True energy release is nearer to

$$E = \frac{(p_1 - p_2) v_1}{(\gamma - 1)} \qquad [5.22]$$

where v_1 is the specific volume before the explosion and γ is the ratio of specific heats (Strehlow and Ricker, 1976). Hence the energy in a gas stored at 10 MPa is approximately 5500 kJ/kmol (for $\gamma = 1.4$), or 255 kJ/kg for compressed air. As already stated in Section 5.1.3(i), about half of this energy is converted to blast energy in the event of a catastrophic pressure vessel failure. The remaining energy is converted to missile kinetic energy.

5.2.6(ii) *Vapour explosions*

Vapour explosions (Section 5.1.3(ii)) are true 'explosions' in the sense that an accelerating shock wave passes through the exploding medium. The hot liquid at the interface fragments; this causes rapid vapour formation in the cold liquid, which induces more fragmentation in the hot liquid. The process therefore has 'positive feedback', and a shock wave develops and accelerates through the hot liquid. The thermal energy of the hot liquid is used to generate vapour, and the expanding vapour converts the thermal energy to blast energy (Lipsett, 1966).

The blast energy, neglecting any release of latent heat, will be given by

$$E = \varepsilon m c_p \Delta T \qquad [5.23]$$

where ε is the explosive yield,

m is the mass of hot liquid,

c_p is the specific heat of the hot liquid, and

ΔT is the temperature change in the hot liquid.

The yield ε usually lies in the range 0.01 to 0.16 (Briggs, 1983). Board and Caldarola (1977) selected a value of 0.10.

Fauske (1973) has proposed that vapour explosions can occur if the interface temperature between the hot and cold liquids, T_i, is greater than the spontaneous nucleation temperature T_n.

The interface temperature between the two liquids may be calculated from

$$\frac{T_i - T_C}{T_H - T_i} = \frac{k_H \rho_H c_H}{k_C \rho_C c_C} \qquad [5.24]$$

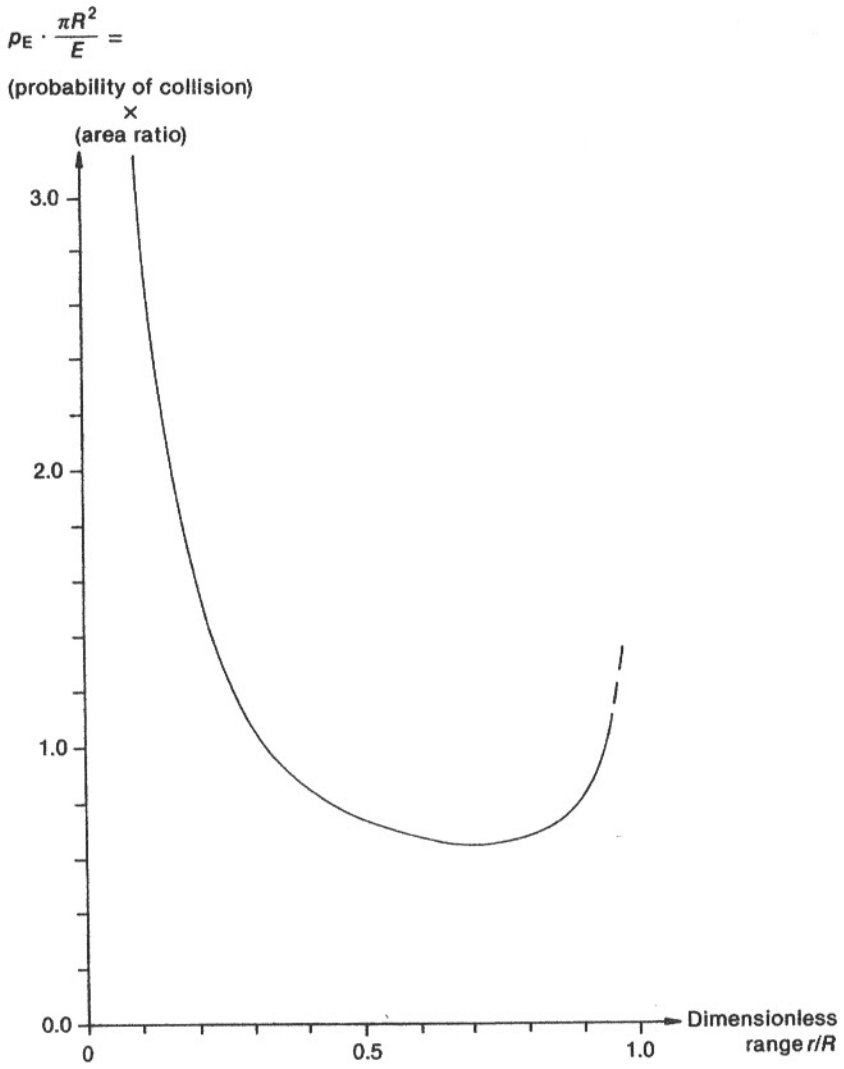


Fig. 5.14 The probability of a randomly fired missile with maximum range R colliding with an object of surface area $E (E \ll \pi R^2)$ at range r, neglecting air resistance

U, and hence the maximum range R, will be known. In some cases an estimate for U can be made, as the following example illustrates.

Example 5.3

To estimate the probability of a 'domino effect' fire or explosion in an LNG storage tank (surface area 100 m²) due to explosive failure of a three tonne pressure vessel containing 10 m³ of steam at 5 MPa. The pressure vessel is 500 metres from the storage tank.

The energy released by the vessel failure can be estimated from eqn [5.22] to be 122.5 MJ. Assuming that half this energy is missile energy, and that there are six equal fragments each of mass 0.5 T, then the initial velocity of each fragment will be about 200 m/s. The maximum range R of such a fragment will be about 4 km, neglecting air resistance. From eqn [5.29], the probability of one of the six missiles striking the storage tank is therefore 2.95×10^{-5}.

5.3 *Dispersal of airborne material*

The area affected by an accidental release of a toxic gas will depend on the quantity of gas released, the rate at which the cloud disperses to safe concentrations, and the toxicity of the gas. The dispersal of the cloud is dependent on several factors, including:

1. Atmospheric conditions,
2. The relative density of the toxic gas,
3. Local topography, e.g. the presence of hills or high buildings, and
4. The height at which the gas is released, e.g. at ground level or through a tall stack.

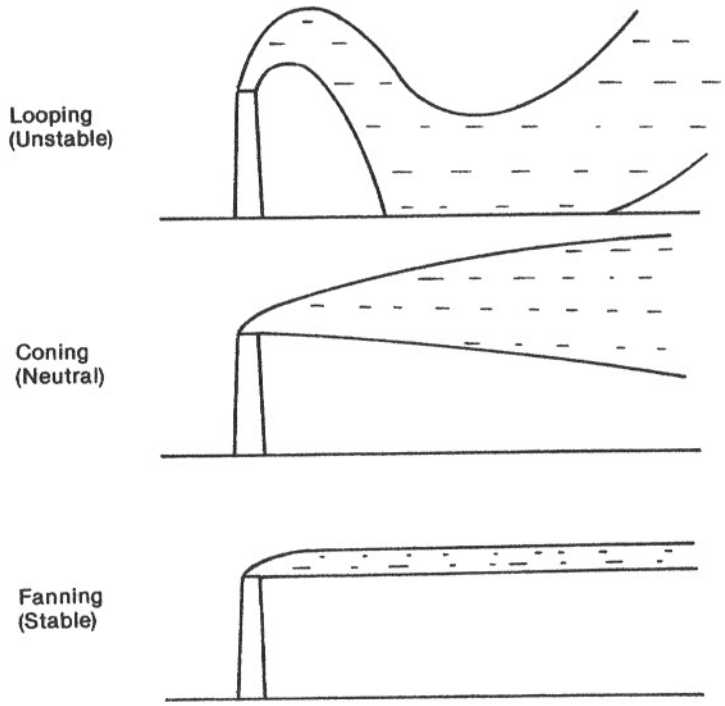


Fig. 5.15 Atmospheric stability

5.3.1 Atmospheric stability

Atmospheric stability can be qualitatively judged by observing the behaviour of smoke released from chimneys (Fig. 5.15). Conditions under which the smoke plume 'loops' are said to be *unstable*. Under different conditions, the smoke plume expands evenly in a vertical direction (*neutral* stability). Under *stable* conditions, little vertical dispersal of the smoke plume occurs.

The stability of the atmosphere is a function of the vertical temperature gradient dT/dz. This temperature gradient is, in turn, a function of the rate of solar heating (insolation), windspeed, and whether it is daytime or nighttime. Pasquill (1961) proposed categories of atmospheric stability as shown in Table 5.4. It can be seen that the most stable conditions occur when the vertical temperature gradient is positive. This is most likely on still, clear nights, and is commonly called a 'temperature inversion'.

Table 5.4 Pasquill's stability categories

Category	*Description*	*Vertical temperature gradient* (dT/dZ)
A	Extremely unstable	<-1.9 K per 100m
B	Moderately unstable	-1.8
C	Slightly unstable	-1.6
D	Neutral	-1.0
E	Slightly stable	$+1.0$
F	Moderately stable	$-1.5-+4.0$
G	Highly stable	$>+4.0$

Surface wind speed (m/s)	*Insolation*			*Night*	
	Strong	*Moderate*	*Slight*	*>4/8 cloud*	*<3/8 cloud*
<2	A	A–B	B	—	—
2–3	A–B	B	C	E	F
3–5	B	B–C	C	D	E
5–6	C	C–D	D	D	D
>6	C	D	D	D	D

Conversely, the least stable conditions arise on hot days with only a slight breeze. Strong winds promote neutral stability at all times.

The frequencies of occurrence of the various stability categories are shown in Table 5.5. It can be seen that the most stable categories (F and G) occur relatively infrequently. Stable atmospheric conditions are even less frequent at coastal sites (Van der Hoven, 1967).

In the event of, say, a toxic release from a chemical plant, the prevailing weather conditions may determine the magnitude of the consequences. A toxic release at ground level during a temperature inversion (stability category G) would take far longer to disperse than an equal magnitude of release under neutral (stability category D) conditions.

Table 5.5 Frequency of occurrence of Pasquill stability categories

Category	*Mean weighted frequency (UK)* (Jones, 1979)	*Typical US frequencies** (Rasmussen, 1975)
A	0.008	0.111
B	0.052	0.033
C	0.155	0.037
D	0.648	0.310
E	0.059	0.305
F	0.058	0.204
G	0.020	no data
Total	1.000	1.000

* US frequencies are the averages of data from seven sites

5.3.2 Plume dispersal

The equation governing the dispersal of clouds of gas, vapour, or small particles is the differential equation for diffusion, with the axes x, y and z referring to downwind, crosswind and vertical directions, respectively:

$$\frac{\partial \chi}{\partial t} + u\frac{\partial \chi}{\partial x} = D_z \nabla^2 \chi \qquad [5.30]$$

Here χ is the gas concentration, u is the windspeed and D_z is the diffusion coefficient. The diffusion process is, in general, anisotropic, so D_z has different values for each direction.

The boundary conditions are:

$$\chi \to 0 \text{ as } t \to 0, \quad x, y, z > 0$$
$$\text{and } \chi \to 0 \text{ as } t \to \infty \qquad [5.30a]$$

for an instantaneous point source, or, for a continuous source (see Fig. 5.16) where $\frac{\partial \chi}{\partial t} = 0$, the boundary conditions are:

$$\chi \to 0 \text{ as } x, y, z \to \infty$$
$$\text{and } \chi \to 0 \text{ as } x \to 0, \quad y, z \neq 0 \qquad [5.30b]$$

Also, the continuity equation gives

$$\int_{-\infty}^{+\infty}\int\int \chi \, dx\, dy\, dz = Q_1 \qquad [5.31]$$

for an instantaneous source, where Q_1 is the quantity of material released. If the source is continuous, then

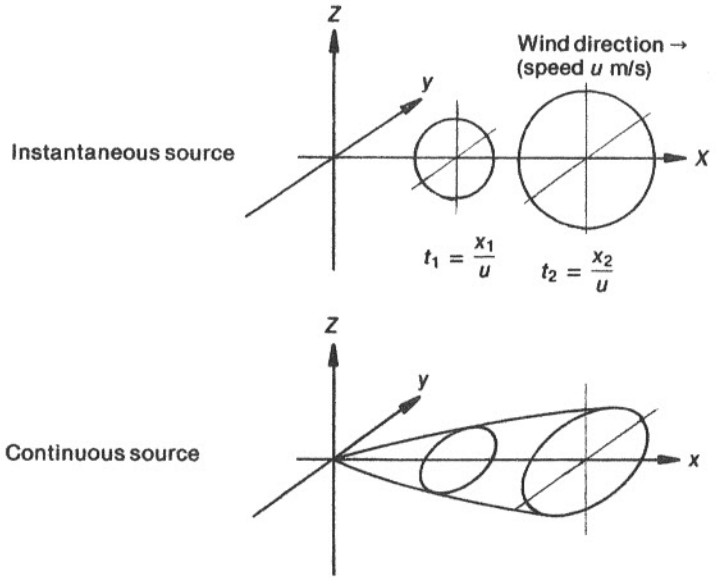


Fig. 5.16 Plume dispersal: idealised representations of instantaneous and continuous sources

$$\int_{-\infty}^{+\infty}\int u\chi\Big|_{x\ \text{constant}} \mathrm{d}y\,\mathrm{d}z = Q_2 \qquad [5.32]$$

where Q_2 is the rate of release of the material.

These equations yield solutions for χ in the form of Gaussian distributions (Section 2.4.3), as illustrated in Fig. 5.17. For an *instantaneous source at ground level*, the appropriate expression is

$$\chi(x,y,z,t) = \frac{Q_1}{(2\pi)^{3/2}\,\sigma_x\sigma_y\sigma_z} \cdot \exp\left\{-\frac{1}{2}\left[\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} + \frac{z^2}{\sigma_z^2}\right]\right\} \qquad [5.33]$$

where σ_x, σ_y and σ_z are the standard deviations, or *dispersion coefficients*, in the x, y and z directions respectively.

For *ground level continuous releases*, the solution of eqn [5.30] is

$$\chi(x,y,z) = \frac{Q_2}{2\pi u\,\sigma_y\sigma_z} \cdot \exp\left\{-\frac{1}{2}\left[\frac{y^2}{\sigma_y^2} + \frac{z^2}{\sigma_z^2}\right]\right\} \qquad [5.34]$$

Finally, for *elevated sources*, for example releases from a chimney stack, the concentration at ground level ($z = 0$) due to a continuous release from height H is given by

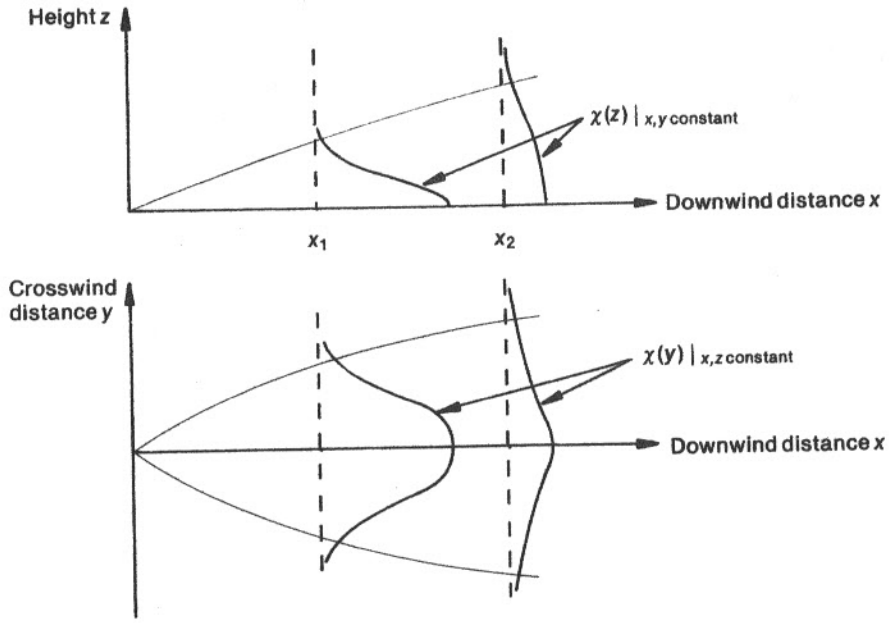


Fig. 5.17 Plume dispersal: the variation of concentration χ with position (x,y,z)

$$\chi(x,y,0) = \frac{Q_2}{\pi u\,\sigma_y\sigma_z} \cdot \exp\left\{-\frac{1}{2}\left[\frac{y^2}{\sigma_y^2} + \frac{H^2}{\sigma_z^2}\right]\right\} \qquad [5.35]$$

In this expression, the divisor 2 (eqn [5.34]) is no longer present. This is due to the effect of the ground 'reflecting' the cloud back upon itself (Fig. 5.18).

Before numerical results can be determined from these equations, values for the dispersion coefficients must be determined. Much work has been carried out to determine experimental correlations between the dispersion coefficients σ_x, σ_y and σ_z, the atmospheric stability categories A–F and the nature of the terrain, e.g. urban, pasture or forest. Sutton (1947) proposed the following expressions:

$$\frac{\sigma_y}{x} = \frac{1}{\sqrt{2}} \cdot C_y x^{-n/2} \qquad [5.36]$$

$$\frac{\sigma_z}{x} = \frac{1}{\sqrt{2}} \cdot C_z x^{-n/2} \qquad [5.37]$$

where n is a function of stability category, and C_y and C_z are functions of source height and atmospheric eddy velocity. For ground level release under neutral conditions and open terrain, Sutton suggests $C_y = 0.21$, $C_z = 0.12$ and $n = 0.25$. This yields, therefore, that

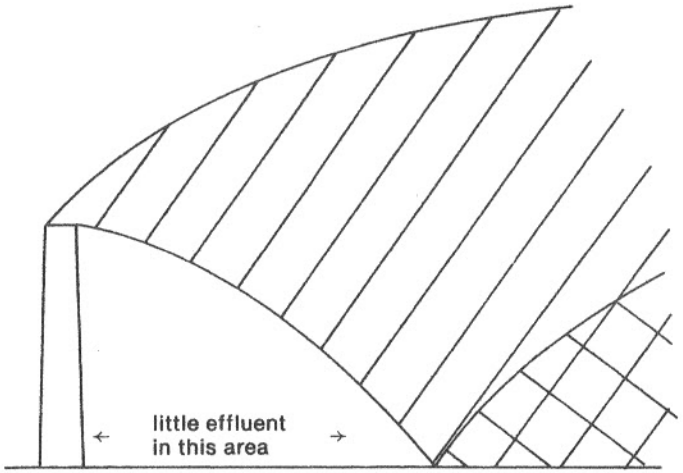


Fig. 5.18 Schematic illustration of continuous release from an elevated source

$$\sigma_y = 0.1485\, x^{0.875} \qquad [5.36a]$$
$$\sigma_z = 0.0848\, x^{0.875} \qquad [5.37a]$$

under the specified conditions. For a 50 metre elevated source, Sutton suggests $C_y = C_z = 0.10$. Hence, in this instance

$$\sigma_y = \sigma_z = 0.0707\, x^{0.875} \qquad [5.36b]$$

More recently, Hosker (1974) has published graphs illustrating the variation of crosswind and vertical dispersion coefficients, σ_y and σ_z, with downwind distance x under differing atmospheric conditions, and for rural as well as urban terrain. These are shown in Fig. 5.19. Sutton's formulae are shown for comparison.

It should be noted that the concentration in a fixed location (x,y,z) will vary with time. The expressions used above refer to *average* concentrations. It does not necessarily follow that the concentration of a toxic gas at a given location will be tolerable just because the average concentration is tolerable.

Finally, data on the frequency of wind directions and speeds in different geographical locations are available (e.g. Meteorological Office (1952), Bryson and Hare (1974)).

5.3.3 Chemical toxicity

The toxicity of chemicals to human beings is, of course, a difficult thing to assess. Toxicological studies are restricted to animal experiments. In some cases, these experiments can be scaled up to give a tentative dose–response curve for humans. In the case of chlorine, tentative human dose–response data are available from its use as a weapon in the First World War, as well as data from accidents.

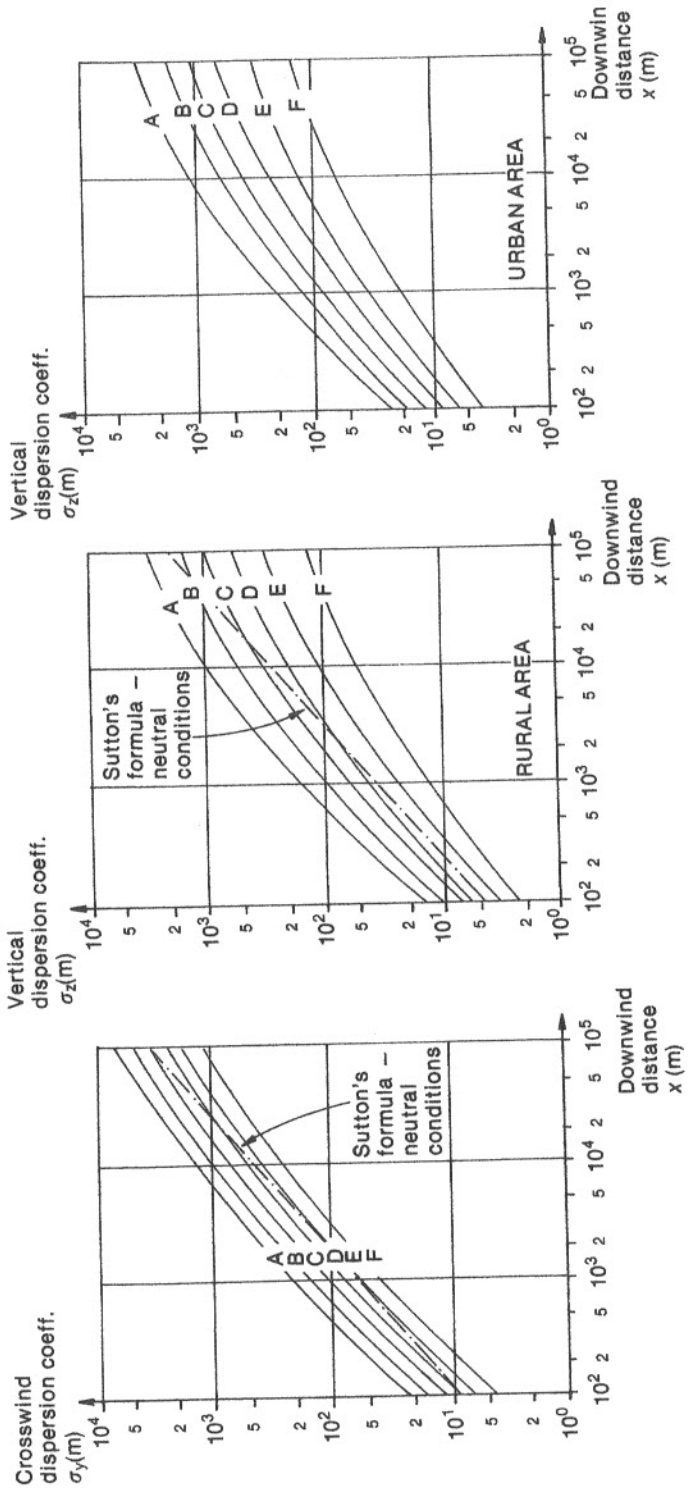


Fig. 5.19 Variation of the dispersion coefficients with downwind distance, for different atmospheric stabilities and terrain (Hosker, 1974). (Letters A to F refer to atmospheric stability categories)

The Canvey report (HSE, 1978) used the available data to reach the following conclusions:

1. The lethal exposure time for a conventional toxic material (i.e. non-carcinogenic) varies inversely with (concentration χ)^n, where n is taken to be 2.75.
2. The probability of fatality is a function of the dosage D^*, thus:

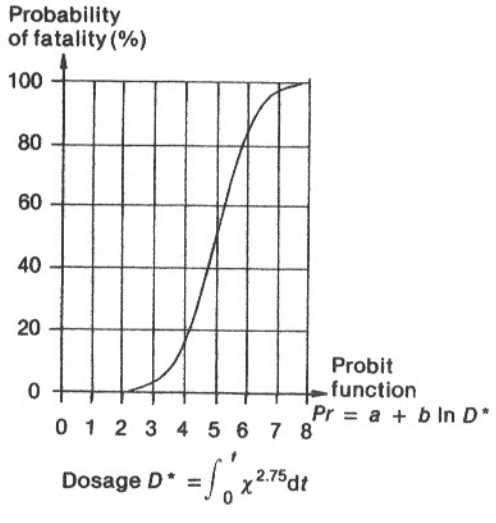
$$D^* = \int_0^t \chi^{2.75}\,dt \qquad [5.38]$$

$$\text{and } Pr = a + b \ln D^* \qquad [5.39]$$

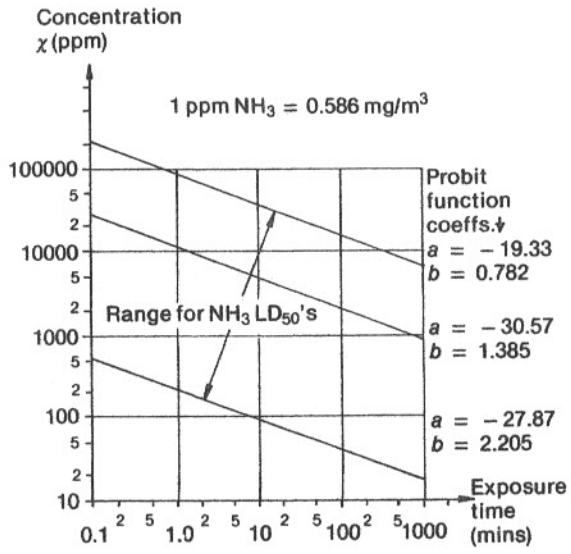
where Pr is called a *probit function*. The values of this function may be related to percentage probabilities of fatality by means of Fig. 5.20(a).

Although data are difficult to acquire, various authors have attempted to assess values for the coefficients a and b in eqn [5.39]. Griffiths and Megson (1984) have reviewed these efforts. Figures [5.20(b)] and [5.20(c)] illustrate the uncertainties in the available data. LD_{50} dosages (i.e. dosages which would kill 50 per cent of those exposed) for both ammonia and chlorine have been calculated using different estimates for coefficients a and b; the dosage has been converted into a line on a graph of concentration (in parts per million) versus exposure time (in minutes).

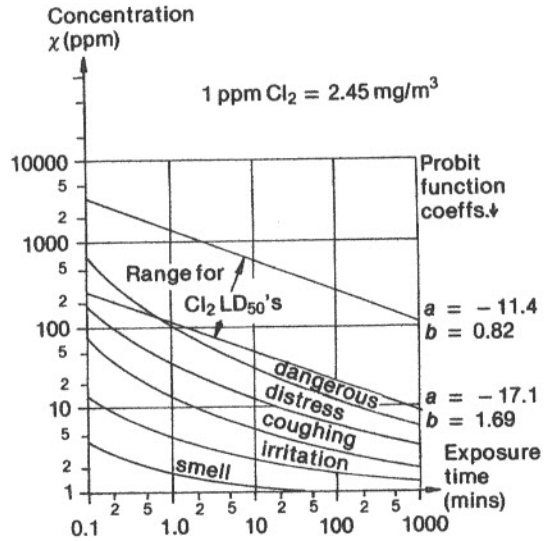
Other available data on the lethal effects of some common industrial gases are presented in Table 5.6. This table also shows data for TLVs (threshold limit values)



(a) The relationship between percentages and probits



(b) The toxicity of ammonia, illustrating the uncertainty of the available data



(c) The toxicity of chlorine, illustrating the uncertainty of the available data

Fig. 5.20 Probit function coefficients and LD_{50} data for ammonia and chlorine (Finney, 1971; Dicken, 1974; and Griffiths and Megson, 1984)

Note: Withers and Lees (1985) have published a review of historical data relating to the toxicity of chlorine.

Table 5.6 Exposure and storage limits for some acutely poisonous gases and vapours

	TLV[1]	IDLH[2]	*Lethal dose*	*Minimum UK Notifiable Inventory*
Ammonia NH_3	25 ppm	500 ppm	See Fig. 5.20(b)	60 tonnes
Arsine AsH_3	0.05 ppm	6 ppm	250 ppm for 30 mins; 10 ppm 'several hours'	10 kg
Bromine Br_2	0.1 ppm	10 ppm	1000 ppm 'rapidly fatal'; 50 ppm 'dangerous for short exposure'	500 tonnes
Chlorine Cl_2	1 ppm	25 ppm	See Fig. 5.20(c)	10 tonnes
Hydrogen Cyanide HCN	10 ppm	50 ppm	270 ppm 'immediately fatal'; 181 ppm for 10 mins; 135 ppm for 30 mins	20 tonnes
Hydrogen Fluoride HF	3 ppm	20 ppm	120 ppm for 1 min 'maximum tolerable'	50 tonnes
Hydrogen Sulphide H_2S	10 ppm	300 ppm	>500 ppm for short exposure	50 tonnes
Methyl Isocyanate CH_3CNO	0.02 ppm	20 ppm	21 ppm 'unbearable'	1 tonne
Phosgene $COCl_2$	0.1 ppm	2 ppm	800 ppm/min/m³ (400 ppm for 2 mins)	20 tonnes
Sulphur Dioxide SO_2	2 ppm	100 ppm	No data	1000 tonnes

Notes: 1. Threshold Limit Value.
2. Immediate Danger to Life or Health

Refs: ACGIH (1980), Sittig (1981), CIMAH Regulations (HSE, 1985)

and IDLH (immediate danger to life or health). The threshold limit value represents a concentration of gas which is believed not to cause any significant degree of harm no matter how long the exposure time. The IDLH level represents a concentration which will cause acute damage within a short timescale.

5.3.4 Assessing the area affected by toxic gas release

If the plume dispersal equations (eqns [5.33–5.35]) are used in combination with toxicity data, it is possible to estimate the area that is likely to be affected by a toxic release accident or else to estimate the area that should be evacuated in the event of an accident. This subject has been considered by a number of investigators, notably Howerton (1969), the Canvey report (HSE, 1978) and Griffiths and Megson (1984). The last authors calculated LD_{50} contours for chlorine and ammonia releases under different

Table 5.7 Estimates of areas affected by toxic gas release (Griffiths and Megson, 1984)

Toxic Gas Released	*Atmospheric Stability Category*	*Windspeed* (m/s)	LD_{50} *areas* (km²) (*min–max*)	*Downwind Distance* (km) (*min–max*)
50T NH_3	A	1.5	0.63–1.1	0.7–0.95
	D	2	0.64–1.2	0.75–1.1
	D	5	0.23–0.57	0.6–1.0
	D	10	0.09–0.30	0.45–0.85
	F	2	0.94–1.7	0.90–1.55
100T NH_3	D	5	0.4–1.0	0.9–1.4
10T Cl_2	D	5	0.2–0.9	0.7–2.6
50T Cl_2	A	1.5	1.3–2.6	1.1–1.9
	D	5	0.52–3.8	1.1–5.4
	F	2	1.1–24.0	1.8–26.9

atmospheric conditions. The contours are of the well-known 'cigar' shape, with the axis of the cigar pointing downwind from the source. Results are shown in Table 5.7. The results show a large degree of uncertainty, particularly when atmospheric conditions are stable (class F).

This subject is discussed further in the following worked examples (5.4, 5.5 and 5.6). The subject is again discussed in Example 5.10 (Section 5.4.3).

Example 5.4

To determine the minimum area which would have to be evacuated following the accidental release of 10 tonnes of chlorine at a constant rate over half an hour (i.e. 5.56 kg/s).

Data
Atmospheric stability class D, windspeed 5 m/s. Chlorine released at ground level in open country. Assume that evacuation should take place where concentration exceeds 5 ppm (equal to 12.25 mg/m³ for chlorine at standard temperature and pressure) and that Sutton's formulae apply.

Method
Substituting Sutton's formulae into eqn [5.34] and inserting numerical data yields

$$1.75 \ln \hat{x} + \frac{\hat{y}^2}{0.0441\hat{x}^{1.75}} = 13.95$$

where $\hat{x}$ and $\hat{y}$ are the downwind and crosswind positions for a concentration of 12.25 mg/m³.

The maximum downwind evacuation distance is therefore $\hat{x}$ = 2900 m. The same formula may be rearranged and differentiated to yield the maximum crosswind evacuation distance, 2y, which is 272 m. Thus an area of about 300 m × 3000 m in the direction of the wind would have to be evacuated. See also Fig. 5.21.

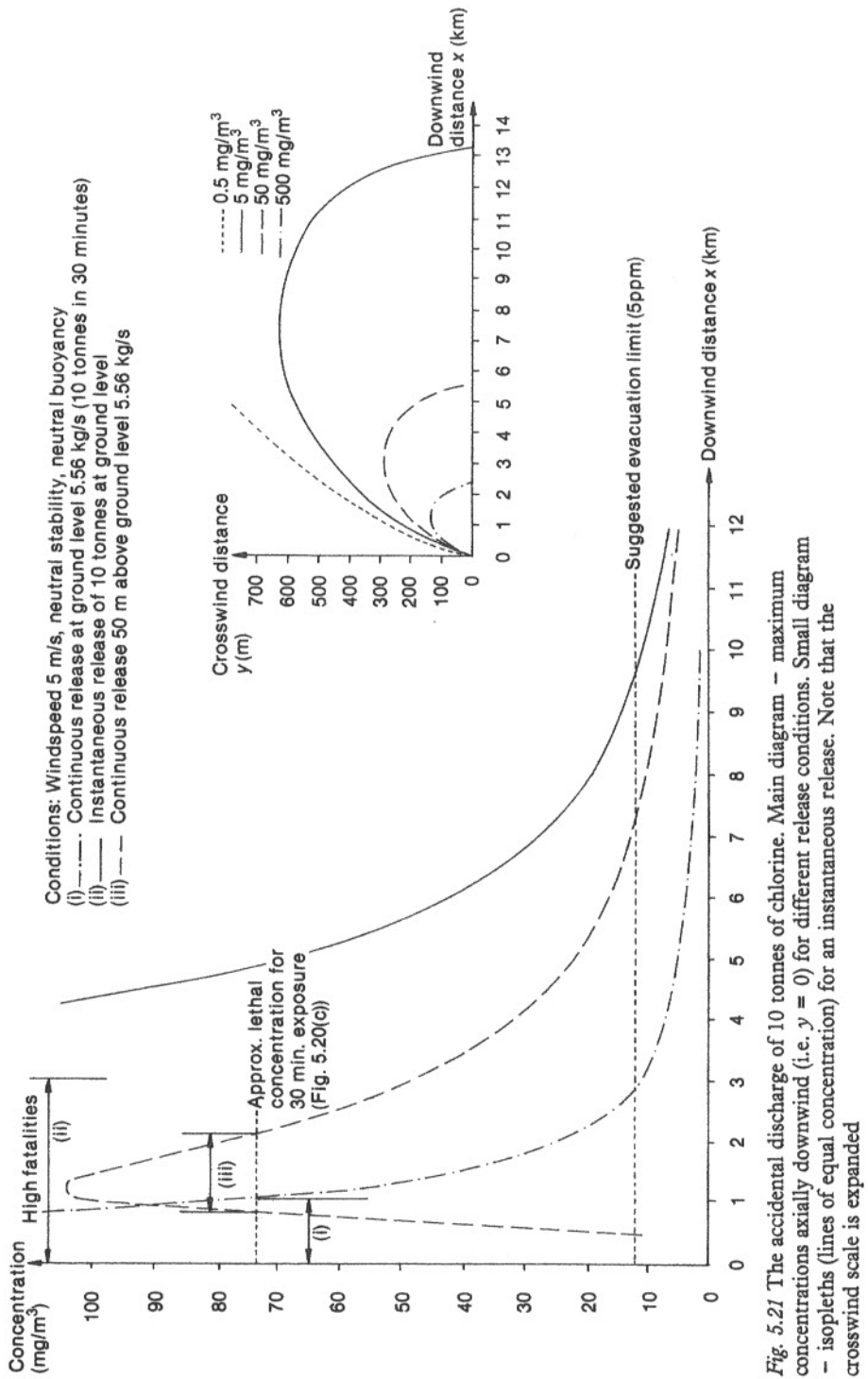


Fig. 5.21 The accidental discharge of 10 tonnes of chlorine. Main diagram – maximum concentrations axially downwind (i.e. $y = 0$) for different release conditions. Small diagram – isopleths (lines of equal concentration) for an instantaneous release. Note that the crosswind scale is expanded

Example 5.5

Repeat the calculation for an instanteous release of 10 tonnes of chlorine. (It is necessary to assume for this calculation that the downward dispersion coefficient σ_x equals the crosswind dispersion coefficient σ_y).

Method

The axial downwind concentrations can be determined by setting $x = y = z = 0$ in eqn [5.33]. (This corresponds to the centre of the cloud in a moving coordinate system.) The radius of the plume will be given by

$$r^2 = x^2 + y^2 = -2\sigma_x\sigma_y \ln\left[\frac{(2\pi)^{3/2}\,\sigma_x\sigma_y\sigma_z\,\chi}{Q_1}\right]$$

where $\sigma_x = \sigma_y$. σ_x, σ_y and σ_z may be determined by substituting $x = ut$ in eqns [5.36(a)] and [5.37a].

This yields an expression for plume radius as a function of time since release.

The above calculations give a downwind evacuation distance of 9.5 km, and a crosswind evacuation distance of $2r$ = 944 m maximum (at 5.37 km downwind). The maximum concentration at 5.37 km would be 54 mg/m³ (24 ppm) and the cloud would take $2r/u$ seconds to pass, i.e. 188 seconds.

This should be safe according to Fig. 5.20(c), but a conservative approach would be to evacuate out to 9.5 km (see Fig. 5.21).

Example 5.6

Repeat the calculation for continuous discharge at 5.56 kg/sec (10 tonnes in 30 minutes) from a 50 m stack.

Method

Using Sutton's formula for the dispersion coefficients for a 50 m high source (eqn [5.36(b)], eqn [5.35] can be employed to calculate the downwind evacuation distance of 7200 m, for a concentration of 5 ppm (12.25 mg/m³). The maximum ground level concentration can be determined by differentiating eqn [5.35] and calculating x for $d\chi/dx = 0$. This yields a maximum ground level concentration of 104 mg/m³ at a downwind distance of 1200 m. The '5 ppm' plume width at this point, $2y$, is equal to 144 m. Thus an area of about 7.5 km × 150 m should be evacuated in this case (see Fig. 5.21).

It is apparent from Fig. 5.21 that the area affected differs widely according to the circumstances of the release. An instantaneous release affects the largest area and would cause the greatest number of casualties. An elevated, prolonged release affects a much smaller area. However, an instantaneous release passes overhead much more quickly, so higher concentrations will be more tolerable without causing lasting harm.

The suggested 5 ppm evacuation limit is quite arbitrary. People living just inside this boundary might feel little more than some discomfort. If they remained indoors with windows and doors closed the effects would be further ameliorated.

5.3.5 The effect of vapour density on plume dispersal

All calculations and formulae so far in this section have as their basis an assumption that the plume is neutrally buoyant, i.e. that the plume has the same density as air.

Although in many cases this assumption is valid, it is important to be aware of the behaviour of lighter- or heavier-than-air plumes.

Plumes will rise upwards if (i) the gas is lighter than air, (ii) the gas is hot, or (iii) the gas possesses initial upward momentum, say due to a leak in a high pressure system. In cases (ii) and (iii), however, the plume will rise for some time (and downwind distance) until it cools or slows down. Thereafter the plume can be treated as an 'elevated source' (eqn [5.35]). The difficulty is to determine the height at which the plume ceases to rise. If this height remains indeterminate, it may be best, for the purposes of safety assessment, to adopt a conservative approach, i.e. to assume the least ameliorating circumstances. It can be seen from Fig. 5.21 that if plume rise is neglected (i.e. that a ground level release is assumed) then the consequences will not be underestimated in the vast majority of cases. However, some attempts have been made to calculate the height of plume rise, notable by Briggs (1984). (See Appendix I.)

Heavier-than-air plumes have attracted considerable attention, primarily for the potential toxicity of such plumes but also because in some chemical industry applications it is desirable to disperse flammable, heavier-than-air hydrocarbons without flaring. In the latter case the gas must be thoroughly mixed with air to ensure that the concentration falls below the lower flammability limit (LFL). The discharge velocity should normally exceed 150 m/s to ensure adequate mixing (Cude, 1974).

Large instantaneous releases of heavy gases were examined by Van Ulden (1974). He performed an experiment in which one tonne of freon (refrigerant-12) was almost instantly released. The density of freon is 4.2 times that of air. The freon cloud was observed to undergo *gravitational slumping*, spreading horizontally along the ground in the manner of a liquid. During this spreading, the freon mixed with air to produce a less-dense mixture with a density 1.25 times that of air.

From this liquid-spreading analogy, Van Ulden was able to propose that the rate of change of the radius of the dense cloud was given by:

$$\frac{dr}{dt} = \frac{F}{r}\sqrt{\frac{g(\rho_o - \rho_a)\,V_o}{\pi}} \qquad [5.40]$$

where r is the radius of the cloud,

ρ_o, ρ_a are the densities of the heavy phase and air, respectively,

V_o is the initial volume of the dense phase,

g is the acceleration due to gravity, and

F is a Froude number (dimensionless).

Figure 5.22 shows the results of Van Ulden's experiment, together with eqn [5.40] and the corresponding predictions assuming Gaussian dispersion. Van Ulden found good agreement between theory and experiment for a Froude number of 1.0. Subsequent work (Prince *et al.*, 1985) has refined this value to give a Froude number of 1.05 ± 0.12.

These results should be borne in mind when considering the consequences of accidents involving the release of heavier-than-air gases, since such gases linger near the ground instead of dispersing vertically. This means that the area for LD_{50} following an instantaneous toxic discharge may be considerably reduced below that predicted by an

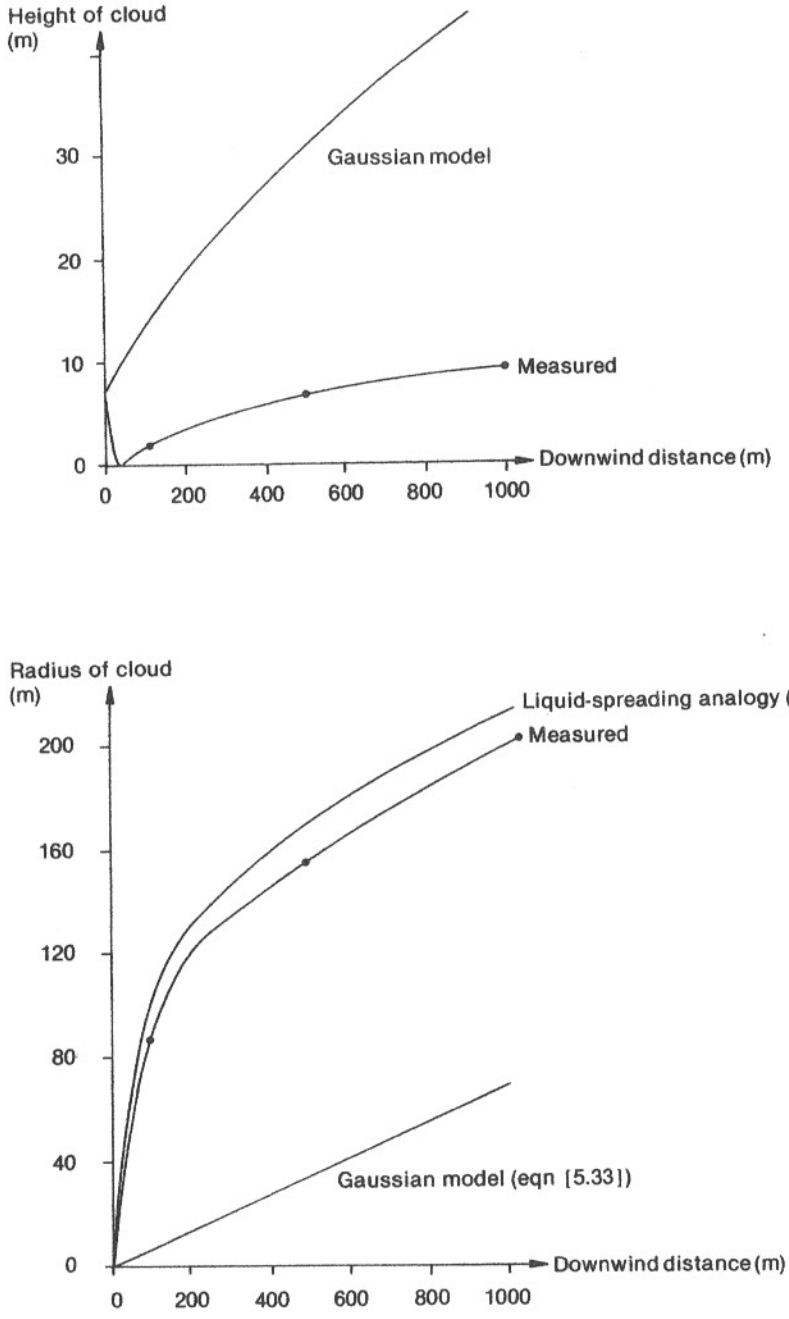


Fig. 5.22 The height and radius of a spreading cloud of heavier-than-air gas. (Van Ulden's experiment – 1 tonne of freon released instantly into a 3 m/s wind under class E conditions)

6 ppb. Stevens (1984) states that 1 ppb is a cause for concern. Evacuation at Seveso was enacted at concentrations greater than 0.15 ppb (WHO, 1981).

Example 5.8

Following an instantaneous release of 1 tonne of butane, the cloud disperses for 1 minute (without slumping) before encountering an ignition source. Determine for windspeeds of (a) 7 m/s or (b) 3 m/s whether the cloud will (i) not support combustion, (ii) deflagrate (LFL 1.8 per cent, UFL 8.3 per cent) or (iii) detonate (LFL 2.9 per cent, UFL 5.2 per cent).

Method

Equation [5.33] may be used together with Sutton's formulae for the dispersion coefficients. Downwind and crosswind dispersion may be assumed to be equal, and the downwind distance x equals ut. The maximum concentration will be for $x = y = z = 0$, i.e. at the centre of the cloud. Hence for a windspeed of 7 m/s the maximum concentration after one minute is 0.44 per cent (i.e. no combustion) but for a windspeed of 3 m/s the maximum concentration after one minute is 4.08 per cent. In the latter case the cloud *may* detonate.

5.3.7 Statistics of accidents involving gas release

5.3.7(i) *Safety of toxic gas transport*

It is self-evident that transport of toxic material by road or rail will, to some extent, be less safe than static storage of the same material. Westbrook (1974) compiled data on the rate of accidents relating to the transport of chlorine. His conclusions are summarised below.

1. The rate of significant accidents for heavy road transport vehicles in the UK is 1.27 per 10^6 vehicle kilometres.
2. The rate of significant accidents for rail vehicles in the UK is 2.37 per 10^6 vehicle kilometres.
3. Given that a rail accident has occurred, the probability of a chlorine spillage is 8.93 per cent (USA) or 4.08 per cent (UK). (The difference is largely due to the greater average speed of goods trains in the USA).
4. Given an accident has occurred, a rail tanker is about five times more likely to be punctured than a road tanker. However, a road accident may cause traffic congestion, thereby increasing the number of people put at risk.

Another review of accident data (WHO, 1981) gives the following USA data:

1. The probability of a chlorine release from a rail tank car is 1.9×10^{-4} per shipment.
2. The rate of occurrence of such accidents in the USA is 1.8 per annum.
3. The average annual number of fatalities from such accidents is 9.4.

5.3.7(ii) *Mortality in accidents involving toxic gas release*

Marshall, whose survey of mortality due to explosions was mentioned in Section 5.2.5(ii), also surveyed mortality in toxic gas releases in the same paper (Marshall, 1977). Again he includes military deaths (First World War gas attacks) in his survey.

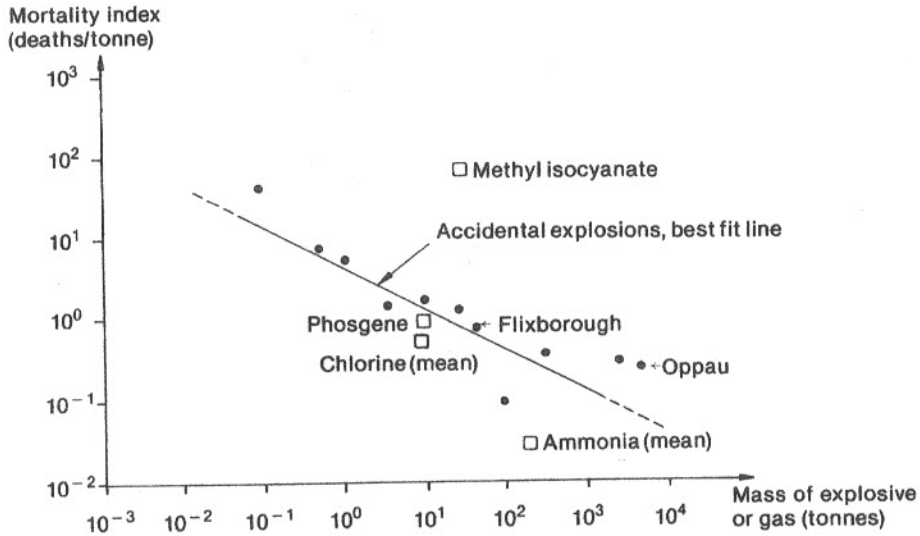


Fig. 5.23 Mortality index (deaths per tonne) as a function of the magnitude of the accident for explosions and gas release accidents (Marshall, 1977)

Unlike explosions, however, he concludes that there are insufficient data to define a relationship between 'mortality index' (deaths per tonne) and the quantity of toxic gas discharged. The lack of data is compounded by the widely differing toxicities of the gases involved in the accidents. His results, for both explosions and toxic gas release, are shown in Fig. 5.23, to which the data for the Bhopal accident in 1984 have been added.

A summary of gas-release accidents (i.e. non-military) for four different toxic agents is shown in Table 5.8. When the mean mortality index is multiplied by the threshold limit value (TLV) for the particular material (Table 5.6), a vague pattern begins to emerge. The product (TLV × mortality index) has a fairly narrow range of values

Table 5.8 Summary of gas release accident data

Gas	*Number of accidents*	*Mean mortality index* (deaths/tonne)	*Threshold limit value* (ppm)	$M_I \times$ TLV
Ammonia	6	0.027	25	0.675
Chlorine	30	0.52	1	0.52
Phosgene	1	1.0	0.1	0.1
Methyl isocyanate	1	83	0.01	0.83

Weighted mean ($M_I \times$ TLV) = 0.546 (ppm deaths/tonne)

with a weighted mean value of 0.546. This is tentatively proposed as a means for determining a central estimate of the likely death-toll of a given gas release accident. It should only be regarded as a rule-of-thumb, as the data used make no allowance for rate of release, population distribution or atmospheric conditions. Furthermore, TLV data should not properly be used in this way, since TLVs may reflect the effects of chronic (long-term) exposure. Proper assessments of the consequences of hypothetical gas release accidents should use dispersal calculations based on population and meteorological data, together with accurate IDLH data, where available (see Table 5.6).

Finally, it is worth noting that the mortality in accidents in which rail tankers are punctured is apparently lower than the above data might suggest. A typical tanker may carry ten tonnes of gas yet Westbrook (1974) concludes that, for chlorine, the rate of death is only 0.36 per punctured tanker. This may well be because many of these accidents occur following derailments in areas of low population density.

5.4 *Radiation and radioactivity*

As has already been stated in Section 5.1.5(i), nuclear reactors cannot under any circumstances explode like nuclear bombs. The hazard to the general public from nuclear power stations and nuclear fuel reprocessing plant comes from the possibility, however remote, that large quantities of radioactivity might be accidentally released to the environment.

Detailed consideration of nuclear safety has been given elsewhere, e.g. Farmer (1977), Lewis (1977) and Graham (1971). In this section only one aspect will be considered; that is, the way in which the consequences to the population of a (hypothetical) airborne discharge of radioactivity may be assessed. It is worth noting at this stage that the biological effects of radiation and radioactivity have been more carefully studied, and are better understood, than those of any 'conventional' toxic materials. Furthermore, radiation and radioactivity can be measured and detected very much more easily than conventional toxins, because of the ionising properties of radiation.

5.4.1 Units of radiation

Some preliminary definitions are required.* The basic unit of radioactivity is the Becquerel (Bq).

$$1\ \text{Bq} = 1\ \text{disintegration per second} \qquad [5.46]$$

The *activity A* in Becquerels of a quantity of radioactive material is given by:

* SI units have only recently replaced older units in common parlance in the nuclear industry. The appropriate conversion factors are as follows:

1 Becquerel (Bq) = 2.7×10^{-11} Curies (Ci)
1 Gray (Gy) = 100 Rads
1 Sievert (Sv) = 100 Rems

with a weighted mean value of 0.546. This is tentatively proposed as a means for determining a central estimate of the likely death-toll of a given gas release accident. It should only be regarded as a rule-of-thumb, as the data used make no allowance for rate of release, population distribution or atmospheric conditions. Furthermore, TLV data should not properly be used in this way, since TLVs may reflect the effects of chronic (long-term) exposure. Proper assessments of the consequences of hypothetical gas release accidents should use dispersal calculations based on population and meteorological data, together with accurate IDLH data, where available (see Table 5.6).

Finally, it is worth noting that the mortality in accidents in which rail tankers are punctured is apparently lower than the above data might suggest. A typical tanker may carry ten tonnes of gas yet Westbrook (1974) concludes that, for chlorine, the rate of death is only 0.36 per punctured tanker. This may well be because many of these accidents occur following derailments in areas of low population density.

5.4 *Radiation and radioactivity*

As has already been stated in Section 5.1.5(i), nuclear reactors cannot under any circumstances explode like nuclear bombs. The hazard to the general public from nuclear power stations and nuclear fuel reprocessing plant comes from the possibility, however remote, that large quantities of radioactivity might be accidentally released to the environment.

Detailed consideration of nuclear safety has been given elsewhere, e.g. Farmer (1977), Lewis (1977) and Graham (1971). In this section only one aspect will be considered; that is, the way in which the consequences to the population of a (hypothetical) airborne discharge of radioactivity may be assessed. It is worth noting at this stage that the biological effects of radiation and radioactivity have been more carefully studied, and are better understood, than those of any 'conventional' toxic materials. Furthermore, radiation and radioactivity can be measured and detected very much more easily than conventional toxins, because of the ionising properties of radiation.

5.4.1 Units of radiation

Some preliminary definitions are required.* The basic unit of radioactivity is the Becquerel (Bq).

1 Bq = 1 disintegration per second [5.46]

The *activity A* in Becquerels of a quantity of radioactive material is given by:

* SI units have only recently replaced older units in common parlance in the nuclear industry. The appropriate conversion factors are as follows:
1 Becquerel (Bq) = 2.7×10^{-11} Curies (Ci)
1 Gray (Gy) = 100 Rads
1 Sievert (Sv) = 100 Rems

$$A = \frac{\ln 2.m_o N}{T_{\frac{1}{2}} M} \quad [5.47]$$

where m_o is the mass of the particular radioactive isotope (kg),
$T_{\frac{1}{2}}$ is its half-life (seconds),
N is Avogadro's number, 6.02×10^{26} kmol^{-1}, and
M is the isotope's atomic mass (kg/kmol).

Because of radioactive decay, the mass of isotope remaining falls exponentially with time t.

$$m = m_o \exp\left(-\frac{t.\ln 2}{T_{\frac{1}{2}}}\right) \quad [5.48]$$

Thus half the initial quantity of isotope remains after one-half life, and only 0.1 per cent remains after 10 half-lives. In this section, however, we will not normally have to consider any reduction in the quantity of the isotope due to radioactive decay, since the major events in a nuclear accident would occur in a timescale of much less than one half-life of the more important isotopes.

Radioactive material disintegrates by emitting alpha particles (helium nuclei), beta particles (electrons), gamma rays (high energy photons) or neutrons. The *radiation field strength* or *flux* ϕ (m^{-2}s^{-1}) at a distance r from a small quantity (point source) of radioactive material of activity A will be

$$\phi = \frac{A}{4\pi r^2} \quad [5.49]$$

assuming that the material itself does not stop any of its own emitted radiation (self-shielding) and that the material is *in vacuo*. In air or in any other absorbing medium the flux falls exponentially with distance:

$$\phi = \frac{A\, e^{-\mu r}}{4\pi r^2} \quad [5.50]$$

where μ is the linear absorption coefficient (m^{-1}) for the medium.

A person standing in a gamma flux ϕ receives a *radiation dose D* proportional to that flux, and to the *gamma photon energy E* (Joules)* and exposure time t.

$$D = \mu_m \phi E t \quad [5.51]$$

where μ_m is the *mass absorption coefficient* for body tissue (m^2/kg).

The units of radiation dose are therefore J/kg. This unit is called the Gray (Gy).

1 Gray = 1 Joule of energy absorbed per kilogram of tissue [5.52]

At this juncture, we must differentiate between *external* and *internal* radiation. If the

* Photon energies are usually quoted in electron-volts (eV).
1 eV = 1.602×10^{-19} J.

radiation source is external, only gamma rays and neutrons need normally be considered, since alpha and beta particles can only travel short distances in air. However, if the radiation source is internal, due to material that has been ingested (swallowed or inhaled), then alpha and beta particles assume greater importance. In particular, alpha particles have a large mass, are highly charged and are stopped within a very short distance within the body. This means that they impart a greater amount of biological damage than, say, a gamma photon of similar energy.

Different types of radiation are accorded (empirically) different values for their *Relative Biological Effectiveness* (RBE), or *Quality Factor Q*. These values are given below.

Radiation	*RBE*
Gamma or Beta	1
Slow neutrons	3
Alpha or fast neutrons	20

These values are used to convert the absorbed radiation dose (measured in Grays) into a *dose-equivalent* measured in Sieverts (Sv), where

$$1 \text{ Sievert} = 1 \text{ Gray} \times \text{RBE} \qquad [5.53]$$

In hypothetical nuclear accidents involving the airborne discharge of radioactive material, we are primarily interested in the effects of alpha and beta particles and gamma rays. The main ways in which these would affect the population are illustrated in Fig. 5.24.

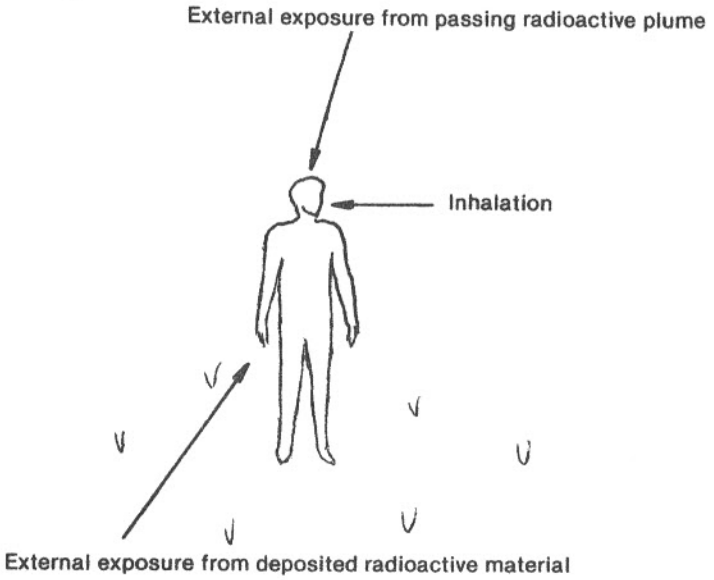


Fig. 5.24 Principle pathways for radiation exposure. (Other pathways exist, including the ingestion of contaminated food and the inhalation of resuspended deposited radioactivity. See also Appendix IV)

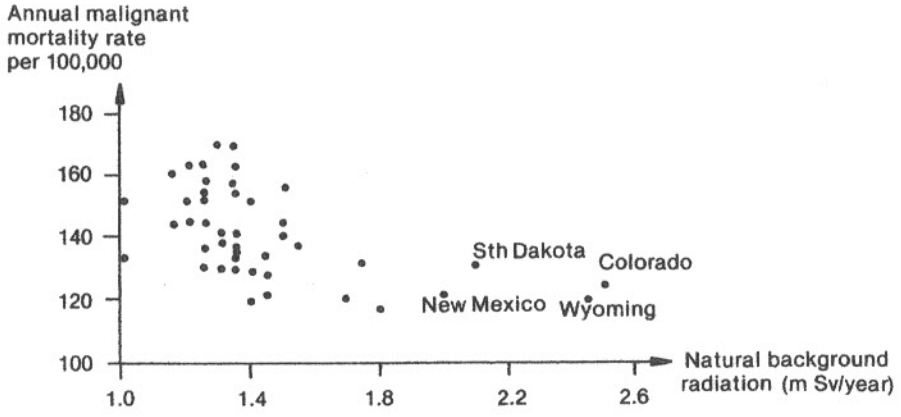
5.4.2 Effects of external radiation exposure

The effects of external radiation exposure may be divided into two categories – short term and long term.

For *short-term* effects, large doses of radiation can kill through damage to the central nervous system, the gastro-intestinal tract and the bone marrow. For an individual receiving a whole-body external dose-equivalent of less than 1.5 Sv, early death is unlikely although there may be some sickness and diarrhoea. Early deaths begin at about 2 Sv, and almost everyone subjected to doses in excess of 6 Sv would die within a few weeks. This therefore gives an LD_{50} (lethal dose to kill fifty per cent of the population) of about 4 Sv.

Genetic defect and cancer induction are *long-term* effects. For genetic defects, Kelly *et al.* (1977) gives a rate of induction to a 'standard' population of 15 per 10^4 man-Sv in the first generation and 57 per 10^4 man-Sv over all time. The 'natural' incidence of genetic abnormalities is 105 100 per million live births. The effect of giving a large population each 1 Sv (say) of gamma radiation will thus be an additional 5700 genetic abnormalities per million live births, or an increase of about six per cent. The difficulty of obtaining this statistical evidence is obvious; a large sample population must be subjected to radiation exposure to reveal such an effect. Indeed, these induction rates only represent upper bound values. No significant change in the incidence of genetic defects has been observed amongst the survivors of the Nagasaki and Hiroshima atomic bombs.

Data on cancer induction rates have been accumulated over the years from the Japanese bomb survivors, people subjected to medical radiotherapy and diagnosis, people who receive occupational exposures, and people who live in areas of high natural background radioactivity. The internationally agreed dose–risk relationship



(Frigerio *et al.*, 1973)

Fig. 5.25 Malignant mortality rate for USA white population 1950–67, by state and natural background radiation

for external whole-body gamma radiation is 1.25×10^{-2} cancer deaths per Sv. (This figure includes leukaemia.) This implies that the risk increases linearly with dose, although this has been questioned. It has been suggested that the risk is smaller at low doses and there are even some data (Fig. 5.25) which suggest that living in an area with high natural background radioactivity *reduces* the chance of dying from cancer. These statistics require very careful interpretation, however, since the expected effects of such low levels of radiation would be slight and about 1 person in 5 dies of cancer anyway. Other factors, such as population age distribution, may be significant. No biological mechanism for this apparent reduction in risk has yet been proposed.

The internationally-agreed limit for occupational exposure is 50 mSv/year, although few radiation workers receive this level of dose. If a radiation worker received this annual dose for every year of a 40-year working life, then, using the linear dose–risk relationship, his (or her) individual risk of dying from cancer becomes 22.5 per cent, instead of the 'natural' cancer death rate of about 20 per cent.* In fact mortality statistics for UK radiation workers show a *lower* than normal rate of death due to cancer. Again, though, other factors may be predominant (the so-called 'healthy worker' syndrome).

5.4.2(i) *External radiation dose from a radioactive plume*

A person standing at the centre of a uniform cloud of airborne radioactive material of concentration χ(Bq/m³) will receive from a volume element of $2\pi r^2\,dr$ (Fig. 5.26) a radiation flux contribution $d\phi$, which may be calculated from eqn [5.50] to be

$$d\phi = \frac{\chi\, e^{-\mu r}\, 2\pi r^2\, dr}{4\pi r^2} \qquad [5.54]$$

(This neglects 'build-up' due to Compton scattering.)

Therefore the total flux at the centre of a uniform hemispherical cloud of radius R will be

$$\phi = \frac{\chi}{2} \int_0^R e^{-\mu r}\, dr$$

$$= \frac{\chi}{2\mu} [1 - e^{-\mu R}] \qquad [5.55]$$

$$\simeq \frac{\chi}{2\mu} \quad \text{for an infinite cloud}$$

The 'infinite cloud' approximation is better than 95 per cent accurate for $R > \frac{3}{\mu}$, where μ is the linear absorption coefficient for air, which has a value of 3.24×10^{-3} m^{-1} at typical gamma energies. Hence this approximation is only valid for clouds with radii of 1000 m or greater.

* Grist (1978) gives a mean UK figure for deaths due to neoplasms of 21.2 per cent for the year 1975.

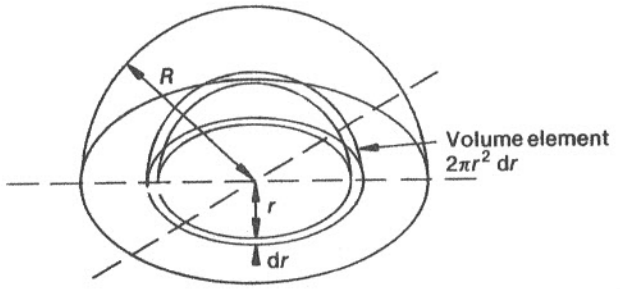


Fig. 5.26 Flux at the centre of a uniform hemispherical cloud of radius R

In all practical cases the cloud will not have a uniform distribution of activity. The concentration will instead follow a Gaussian distribution as given by eqns [5.33], [5.34] or [5.35], depending upon whether the cloud has been released instantaneously, or over a period of time, or from an elevated source. In general, therefore, we may write that for a point on the ground $(a,b,0)$, the radiation flux contribution $d\phi$ due to a volume element $dx\,dy\,dz$ at position (x, y, z) will be

$$d\phi = \frac{\chi(x,y,z)\, e^{-\mu[(x-a)^2 + (y-b)^2 + z^2]^{\frac{1}{2}}}}{4\pi[(x-a)^2 + (y-b)^2 + z^2]} \cdot dx\, dy\, dz \qquad [5.56]$$

and therefore

$$\phi(a,b,0) = \frac{1}{4\pi} \int_0^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \frac{\chi(x,y,z)\, e^{-\mu[(x-a)^2 + (y-b)^2 + z^2]^{\frac{1}{2}}}}{(x-a)^2 + (y-b)^2 + z^2} \cdot dx\, dy\, dz \qquad [5.57]$$

The radiation dose D received by a person standing at a fixed point $(a,b,0)$ may be determined by integrating

$$\frac{dD}{dt} = \mu_m E \phi(a,b,0) \qquad [5.58]$$

over the time t that the cloud takes to drift past the point $(a,b,0)$, where, in general, the flux ϕ will vary with time $\phi(t)$. The mass absorption coefficient μ_m for body tissue has a value of 0.0027 m²/kg at typical gamma energies.

It is evident that accurate solution of these equations, ([5.57] and [5.58]) can only be obtained numerically. Computer codes for these calculations have been written, e.g. TIRION. Approximate solutions are also possible. (See Example 5.9).

5.4.2(ii) *External radiation dose from deposited radioactivity*

Non-gaseous radioactive material released in an accident will be dispersed on the ground some way downwind, due to the processes of sedimentation, impaction and washout as discussed in Section 5.3.6. That section has a discussion of methods for calculating the amount of material deposited per unit area. Having determined this for

an airborne discharge of radioactive material, it is then desirable to calculate the dose-rate to a person due to that material.

Figure 5.27 shows how to proceed. The flux $d\phi$ at height h above the ground due to an element of surface $2\pi r\,dr$ with surface concentration C (Bq/m²) will be given by

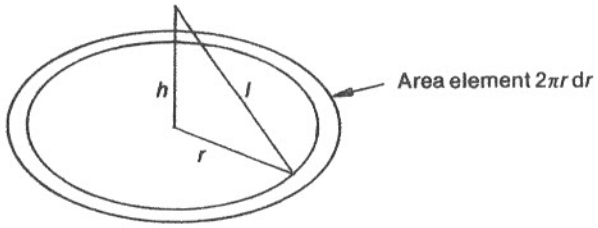


Fig. 5.27 Radiation flux above contaminated ground

$$d\phi = \frac{C\,e^{-\mu l}\,.\,2\pi r\,dr}{4\pi l^2} \qquad [5.59]$$

where

$$l = (r^2 + h^2)^{\frac{1}{2}} \qquad [5.60]$$

Therefore $l\,dl = r\,dr$,
so

$$d\phi = \frac{C\,e^{-\mu l}\,2\pi l\,dl}{4\pi l^2} \qquad [5.61]$$

Hence the flux ϕ at height h above the centre of an area of contaminated ground of radius R ($R \gg h$) will be

$$\phi = \frac{C}{2}\int_h^R l^{-1}\,e^{-\mu l}\,dl$$

$$\Rightarrow \phi = \frac{C}{2}\left[\ln \mu l - \frac{\mu l}{1!} + \frac{\mu^2 l^2}{2.2!} - \frac{\mu^3 l^3}{3.3!} + \ldots\right]_h^R \qquad [5.62]$$

Example 5.9

To estimate the external radiation dose to a person on the ground 500 m downwind from a severe nuclear accident in which 3700 TBq of Iodine-131 are released instantaneously. (This equals 100 000 Curies and corresponds to about 0.25 per cent of the core inventory of I-131 in a large power reactor. This compares with the Three Mile Island accident in 1979 which released 16 Curies or the Windscale fire in 1957 which released 20 000 Curies. The Chernobyl accident in 1986 released between 10 and 20 million Curies of I-131.)

Data

Linear absorption coefficient μ	3.24×10^{-3} m^{-1}
Gamma energy E	0.364 MeV
Mass absorption coefficient μ_m	0.0027 m²/kg
Windspeed u	5 m/s

Neutral weather conditions (class D)

Method

An approximate solution to this problem can be calculated if it is assumed that the effects of the cloud can be equated to the effects of an idealised hemispherical cloud of total activity Q Becquerels and uniform activity concentration $\bar{\chi}$, such that

$$\bar{\chi} = \frac{Q}{\frac{2}{3}\pi \bar{R}^3} \qquad [5.63]$$

Here $\bar{R}$ is the *idealised cloud radius*. In the CRAC2 code (Hemming *et al.*, 1983) this is taken to be equal to one and a half* times the crosswind dispersion coefficient σ_y, i.e.

$$\bar{R} = 1.5\sigma_y \qquad [5.64]$$

Substituting Sutton's formula (5.36a) for σ_y into this expression gives, for neutral (class D) conditions and open terrain, that

$$\bar{R} = 0.223\,(ut)^{0.875}$$

where (ut) is the downwind distance.

The flux at the *centre* of this idealised cloud will therefore be given by

$$\phi_c = \frac{\bar{\chi}}{2\mu}\,[1 - e^{-\mu \bar{R}}] \qquad [5.55]$$

In this example $\bar{R}$ = 51 m and $\bar{\chi}$ = 13.1 GBq/m³.

The integrated dose to a person on the ground can therefore be estimated as follows (from eqn 5.51):

$$D \simeq \phi_c \mu_m E \frac{2\bar{R}}{u} \qquad [5.65]$$

Two further assumptions have been made to obtain this expression. First, it is assumed that the flux is constant wherever a person is standing in the cloud and, secondly, it is assumed that the flux only affects those standing within the cloud. The validity of these approximations is not examined here; however the two approximations will tend to cancel one another out.

The various assumptions made in this analysis yield an integrated dose D from cloud gamma radiation equal to 0.99 mSv.

The dose rate due to deposited iodine can be estimated in a similar manner. The ground contamination C (Bq/m²) can be estimated from eqn [5.41], thus:

* The value of $\bar{R}/\sigma_y$ is dependent upon the amount of vertical dispersion, σ_y/σ_z. Hence the value of $\bar{R}/\sigma_y$ is a function of atmospheric stability and downwind distance. See Fig. 5.19. (Thomson and Nightingale, 1987)

$$C \simeq v_d \cdot \bar{\chi} \cdot \frac{2R}{\bar{u}} \qquad [5.66]$$

This expression gives a ground contamination level in the range 0.80 to 5.3 GBq/m² assuming that the value for the deposition velocity, v_d, lies in the range 0.3 cm/s (Kelly *et al.*, 1977) to 2 cm/s (Pasquill, 1968). The dose rate one metre above the ground can therefore be estimated from eqns [5.62] and [5.51] to be in range 0.27 to 1.79 μ Sv/s for a 100 metre contaminated radius.

If it is further estimated that people in the area will not be evacuated until (say) two hours after the accident, then their total external exposure will be (cloud gamma radiation + ground gamma radiation). This lies in the range 2.9 to 13.9 mSv. A typical annual exposure due to natural background radiation is about 1.5 mSv, so this does constitute a significant increase above what might be considered 'normal'. Nevertheless, the risk to an exposed individual – which may be determined by multiplying the dose by the whole body dose–risk factor of 1.25×10^{-2} cancer deaths per Sievert – only equals an additional risk of death from cancer of not more than 0.017 per cent. Therefore, a very large number of people would have to receive such a dose before the effects of the external radiation became apparent as increased mortality.

The additional risk of death (0.017 per cent) is called the *individual* risk. If (hypothetically) one hundred thousand people received such a dose, the expected additional number of cancer deaths (17) constitutes the *societal* risk. Twenty thousand of these people might be expected to die from cancer anyway, so this additional mortality would not be detectable with any degree of confidence.

Finally, it should be noted that for exposure to I-131 (and many other isotopes) a greater risk comes from internal exposure (due to inhalation) than from external exposure. This aspect is considered in the following section.

5.4.3 Effects of internal radiation exposure

Studies of hypothetical accidents often reveal that the most important exposure pathway is internal radiation due to inhalation of airborne material. This is because certain isotopes tend to concentrate in particular parts of the body after they have been inhaled. Those organs which concentrate the inhaled radioactive material then receive a higher than normal radiation dose. Furthermore, some of the isotopes are short-lived, so the organ only receives a radiation dose for a short time; other isotopes may have half-lives longer than a normal human lifespan. In the latter case, the organ may be continuously irradiated for the rest of the person's life.

In order to calculate the individual and societal risks from inhaled radioactivity, we need to know (i) in which organs different radioactive species concentrate, (ii) the dose that a given organ receives from a given quantity of inhaled isotope (the 'inhalation factor'), (iii) the dose–risk relationships for cancers to form in the various organs and, finally, (iv) the amount of isotope that is likely to be inhaled.

This information has been compiled, for a large number of different isotopes, by Kelly *et al.* (1977). Information for four isotopes of principal concern is given in Table 5.9. Of these, iodine, strontium and plutonium present special risk since each becomes concentrated in a particular part of the body. The integrated dose to that organ then rises steadily with time, levelling off only as the isotope decays or else is gradually excreted from the body. Further information is given in Appendix II.

Table 5.9 Inhalation factors and dose–risk relationships (Kelly *et al.*, 1977, ICRP, 1977 and Hemming *et al.*, 1983)

Isotope	*Time since Inhalation*	*Strontium-90*	*Iodine-131*	*Caesium-137*	*Plutonium-239*
Half-life		*28 y*	*8 d*	*30 y*	*24 300 y*
Radiation		*Beta*	*Beta/gamma*	*Gamma*	*Alpha*
Affected organ		*Red bone marrow*	*Thyroid*	*Whole body*	*Lungs*
Disease		*Leukaemia*	*Cancer*	*Cancer*	*Cancer*
Inhalation factor (Gy/GBq)	1 day	0.046	11.6	0.03	18.4
	1 year	76	270	6.5	3240
	50 years	760	270	7.3	16 200
Relative Biological Effectiveness (RBE)		1	1	1	20*
Dose–risk factor (deaths/man Sv)		15×10^{-4}	3.1×10^{-4}†	1.25×10^{-2}	13×10^{-4}
Derived working limit (Bq/m³)		300	700	2000	0.2

* For plutonium-239, the dose-equivalent (Sieverts) must be calculated by multiplying the inhalation factor by the RBE. (Section 5.4.1)

† The mortality from thyroid cancer is only about 5 per cent, so the incidence of thyroid cancer is twenty times greater than this figure. For thyroid doses greater than 15 Sv, the cancer incidence rate is approximately halved

Notes: 1. Further data are given in Appendix II

2. Dose–risk data are based on the conservative assumptions that there is a linear dose–risk relationship and that there is zero risk at zero dose

The normal inhalation rate varies from person to person, and also depends on rate of physical activity. Typical values lie in the range 2.2 to 3.5×10^{-4} m³/s.

Using this information, it is possible to calculate the permissible airborne concentrations (*derived working limits or DWL*) for occupational exposure of workers in the nuclear industry (commensurate with an 'acceptable' additional risk of cancer) by assuming that a radiation worker breathes in this concentration of radioactivity throughout a fifty-year working life, forty hours per week. The derived working limits given in Table 5.9 all, in principle, correspond to a small additional risk of cancer. As has already been stated, mortality statistics for radiation workers in fact show a lower than normal death rate due to cancer.

Because of its large inhalation factor and RBE, the DWL for plutonium-239 is very low – only 0.2 Bq/m³. This corresponds to only three micron-size dust particles of PuO_2 per cubic metre of air. Monitoring of such low levels in nuclear fuel processing plant requires continuous air filtration and measurement of the alpha-radioactivity of the filter. Such measurements may be confused by the alpha-radioactive daughter product of radon gas, polonium. Radon occurs naturally in masonry and brickwork, and

in poorly ventilated buildings its airborne concentration can rise to significant levels (Cliff *et al.*, 1984).

Example 5.10

To estimate the internal radiation dose to a person on the ground 500 m downwind from a severe nuclear accident in which 3700 TBq of iodine-131 are released instantaneously in neutral weather conditions with a windspeed of 5 m/s. (See also Example 5.9).

Method

If we assume, as before, that the Gaussian cloud can be treated as a uniform hemisphere of radius $\bar{R}$ equal to $0.223\ (\bar{u}t)^{0.875}$ and concentration $\bar{\chi}$ equal to $\frac{Q}{\frac{2}{3}\pi\bar{R}^3}$, then the inhaled activity can be equated to a dose to the thyroid thus:

Dose to thyroid (Sv) = (inhalation factor) × (activity concentration in cloud) × (inhalation rate) × (time for cloud to drift past) × (Relative Biological Effectiveness).

[5.67]

This gives an internal dose to the thyroid for a person 500 m downwind equal to 21.6 Sv over the next year (or over the next fifty years since I-131 has a short half-life) assuming an inhalation rate of 3×10^{-4} m³/s. This internal dose is very much larger than the external dose of 2.9 to 13.9 mSv (Example 5.9). From Table 5.9, the individual risk of dying from thyroid cancer will be 0.66 per cent and the individual risk of contracting thyroid cancer will be twenty times higher at about 13 per cent. For this reason members of the public living near nuclear power stations would be issued with potassium iodide tablets in the unlikely event of an accident. These tablets swamp the thyroid gland with non-radioactive iodine, thus limiting its uptake of radioactive iodine.

Because of the general paranoia surrounding nuclear power, it is probably wise to put this hypothetical accident, the instantaneous release of 3700 TBq of I-131, into some sort of perspective. With the single notable exception of the Chernobyl accident in April 1986 – which probably released about 500 000 TBq of I-131, although the reactor must be considered a poor design by Western safety standards – such an accident would be the largest ever nuclear accident by a large margin. A person 500 metres downwind would have an 87 per cent probability of suffering no after effects and a greater than 99 per cent probability of living a normal term of life.

If the accident released 10 tonnes of chlorine instead of 3700 TBq of iodine, that same person 500 m downwind would have zero chance of survival. (This calculation may be performed by assuming, as in Example 5.9, that the cloud is a uniform hemisphere instead of Gaussian. A person on the axis of the cloud would inhale 0.0746 kg/m³ of chlorine for twenty seconds. From Fig. 5.20(c), such an exposure would be fatal.) Nevertheless, the storage and transport of chlorine can be achieved with much less public concern than the construction of a nuclear power station or the transport of irradiated nuclear fuel.

The societal risk of a 3700 TBq release of iodine-131 is more difficult to calculate. The individual dose must be determined as a function of downwind distance and then the integrated value of (dose × population density) has to be determined. The results of this calculation are shown in Fig. 5.28, assuming a pessimistically high population density of 5000 per square kilometre. (Most nuclear plants are in rural or semi-rural districts). These results suggest that

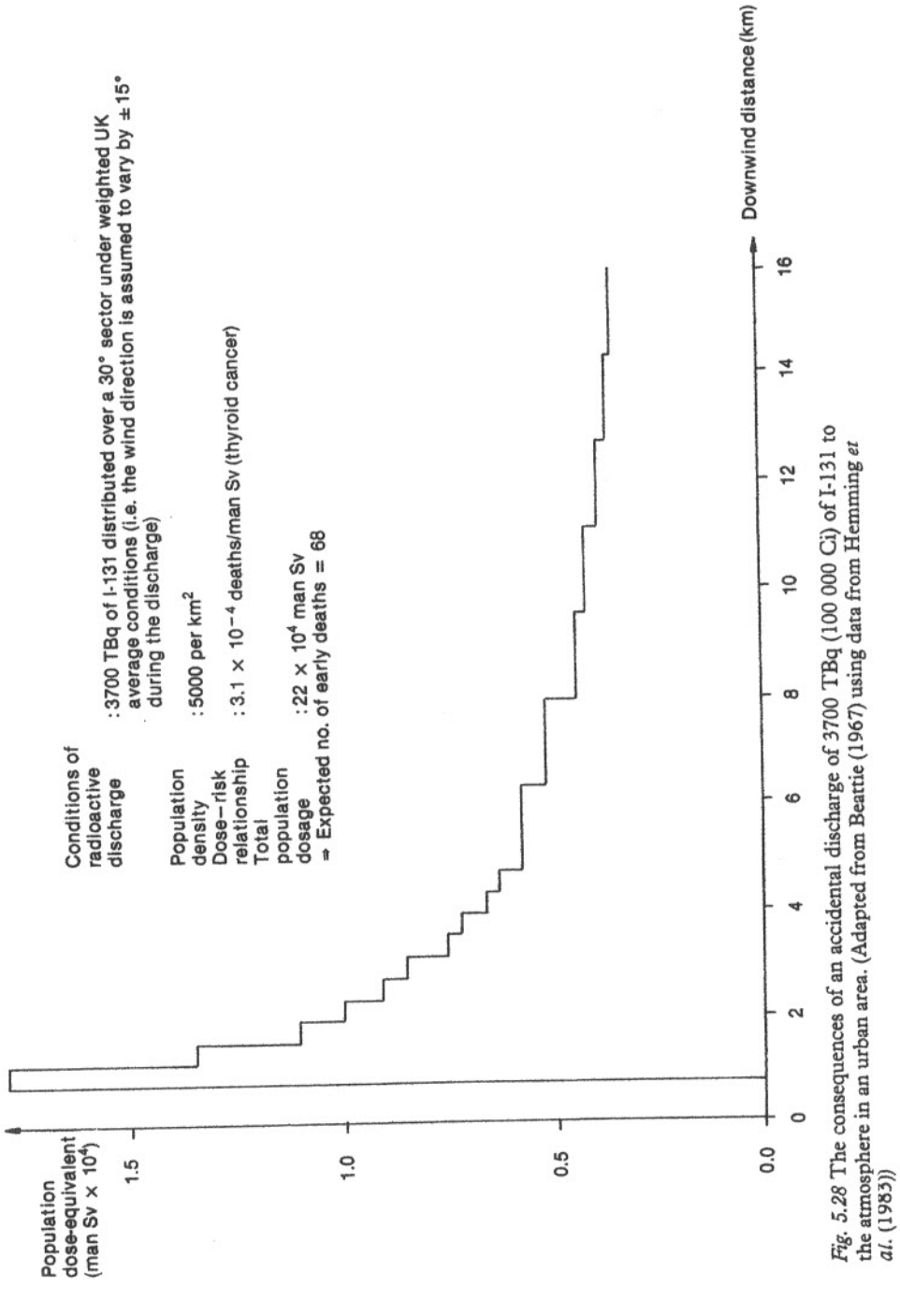


Fig. 5.28 The consequences of an accidental discharge of 3700 TBq (100 000 Ci) of I-131 to the atmosphere in an urban area. (Adapted from Beattie (1967) using data from Hemming *et al.* (1983))

about 68 extra early deaths due to cancer could be expected in this case. Four million people live within a ten-mile radius of this (hypothetical) reactor, and the 'normal' incidence rate of thyroid cancer is about twenty per million per annum with 5 per cent mortality (Campbell *et al.*, 1963). Hence the succeeding thirty years would normally produce 2400 cases and 120 deaths due to thyroid cancer, even if the reactor accident did not happen. Thus this accident has increased the death rate from one fairly rare disease by about 50 per cent.

There is, however, a lot of implicit pessimism in this figure of 68 extra early deaths. In addition to the unrealistically high population density, no allowance has been made for people escaping from the plume path, or for people remaining indoors with windows shut, or for the ameliorating effects of potassium iodide tablets. If similar, pessimistic assumptions were made for a calculation of the consequences of an accidental release of 10 tonnes of chlorine, then Table 5.7 (Section 5.3.4) suggests that a minimum of 1000 prompt deaths would result. (This is calculated by assuming that everyone within the LD_{50} area is killed, and that everyone outside the area survives.) However, no chlorine release accident has ever caused such a high death-toll. Of thirty such accidents reviewed by Marshall (1977), the highest death-toll has been 60, following a release of 25 tonnes of chlorine in Romania in 1939. From Table 5.7, the median death-toll for a release of 10 tonnes of chlorine would be only give. This illustrates the fallacious nature of the assumptions that were made. (Like nuclear reactors, storage tanks of toxic gas are seldom built in densely-populated areas.)

Plume dispersal modelling can only give accurate predictions of the consequences of accidents if reasonable assumptions are made at the beginning. ('Garbage in, garbage out'.) However, even rough calculations enable the relative magnitudes of the hazards associated with differing technologies to be assessed.

5.4.4 Accidents involving mixtures of isotopes

The accident considered in Examples 5.9 and 5.10 is a very idealised case; in any actual accident it is likely that a wide variety of different radioactive isotopes would be released. In a reactor accident, more short-lived isotopes would be released than in an accident in a fuel reprocessing plant, since a reprocessing plant handles fuel that has not been in a reactor for at least several months. Hence it is unlikely that a significant quantity of iodine-131, with a half-life of only 8 days, would be released in any reprocessing plant accident (e.g. a fire).

As has already been stated, iodine-131 is an isotope of great concern because of the body's ability to retain it in the thyroid. Furthermore, iodine in its elemental form is relatively volatile, so in the event of a reactor accident it is one of the fission products which is most likely to escape. Other fission products such as isotopes of xenon and krypton may be more volatile and more radioactive, but are also chemically inert, so they only present a hazard from external radiation.

The volatilities of fission products are therefore indicators of the relative probabilities of given isotopes being released in an accident. In order of decreasing volatility, the major fission products are:

The inert gases (krypton and xenon)
Iodine (b.p. 183°C)
Caesium (b.p. 690°C)
Strontium (b.p. 1380°C)

Table 5.10 Multiplying factors for equivalent dose (approximate maximum single organ dose)

Isotope	*Multiplying factor* (Bq/Bq I-131)	*Exposure pathway*
Sr-90	0.05	Inhalation, bone
I-131	1.0	Inhalation, thyroid
Cs-137	50	Inhalation, whole body
Inert gases (Kr and Xe)	5×10^4	External gamma radiation, whole body

For the non-inert fission products, much of each isotope will be present in the form of a chemical compound with, in all probability, a higher boiling point than the pure element. This, then, introduces the problem of determining the 'source term' or likely magnitude of radioactive release, in any given hypothetical nuclear accident. For further discussion of this problem, readers are referred to Farmer (1977).

Determining the effects on the population of a given release of radioactivity to the atmosphere necessitates summing the risks due to each exposure pathway (external radiation, inhalation and ingestion) for each isotope in turn; this requires a very lengthy calculation procedure. However, a simpler method can be used to yield approximate results. This approximate method consists of converting each isotope into an approximate equivalent quantity of iodine-131, such that the same maximum single organ dose results. These *Multiplying Factors for Equivalent Dose* are given in Table 5.10. These values indicate that the most toxic fission product is strontium-90, and that the least toxic are the inert gases. Using these data, an accidental release involving a mixture of fission products can be converted into an approximate equivalent quantity of iodine-131, thus simplifying the calculations of individual and societal risks.

Isotopes other than those shown in Table 5.10 are of lesser importance, either because they are very short-lived or else because they have very high melting and boiling points, and are therefore unlikely to become airborne following a reactor or process plant accident. Thus, although Appendix II indicates that plutonium and the other actinides are very much more toxic than, say, iodine-131, the melting point of plutonium oxide (the chemical form of plutonium in most reactors) is so high (*c.* 2700°C) that it is difficult to envisage how a significant quantity could become airborne following any accident.

Following any given accident, the minimum countermeasures to be taken are defined by Emergency Reference Levels (NRPB, 1981). These are the single organ dose levels at which specified countermeasures must be taken. See Table 5.11.

In each of the three major reactor accidents in the world to date, a cocktail of fission products has been released. The Three Mile Island accident released inert gases and a trace of iodine, totalling about 0.6 TBq of I-131 (equivalent). The Windscale fire released inert gases, iodine-131 and lesser quantities of other fission products, together

Table 5.11 Emergency reference levels for single organ dosage (NRPB, 1981)

Countermeasure	*Single organ dose* (mSv)
Evacuation	300
Sheltering	50
Distribution of stable iodine tablets	50

with some polonium-210,* totalling about 100 TBq of I-131 (equivalent). The Chernobyl reactor fire would appear to have released inert gases, iodine, caesium and other fission products totalling about 500 000 TBq of I-131 (equivalent). Given that, in the Chernobyl accident, there was a complete breach of reactor containment for a number of days while the reactor was on fire, it is quite difficult to conceive how a larger radioactive release from a reactor accident could occur.

5.5 Conclusions

This chapter has attempted to present to the reader a brief overview of some major industrial hazards, together with more detailed consideration of the effects of explosions, toxic gas release and the release of airborne radioactivity. The following general conclusions may be drawn:

1. Typically, the accidental release of a tonne of chlorine kills as many people as the unconfined detonation of a tonne of high explosive. All deaths will occur shortly after the event (so-called 'prompt' deaths).
2. Other airborne chemicals will cause deaths in proportion to their relative toxicity.
3. The quantification of delayed deaths (cancers), due to atmospheric releases of chemical carcinogens, is often made difficult by the lack of suitable dose–risk data.
4. Dose–risk data for radioactive toxins suggest that an instantaneous airborne release of 3700 TBq (10^5 Curies) of iodine-131 would most probably kill about as many people as the release of between one and ten tonnes of chlorine under the same conditions. Deaths due to the iodine-131 would occur after ten to thirty years, while deaths due to chlorine would be 'prompt'. Comparing radiotoxins with 'conventional' toxins in this way is difficult, however, because of the different nature of their dose–risk relationships.

Finally, it should be noted that death is not the only cost to society in accidents such as these. High explosive destroys property as well as people, and airborne releases of long-lived toxic or carcinogenic materials may require expensive clean-up procedures.

* The Windscale reactor was a military reactor. Polonium is not present in any civilian power reactor.

Questions

5.1 An airliner fuselage has a volume of 150 m³. Determine the energy released (in kg of TNT) if the pressurised fuselage failed at high altitude, when the pressure difference was 40 kPa. The internal pressure is maintained at 80 kPa, and the temperature is 25°C.
(3.1 kg)

5.2 Estimate the radii for blast damage causing (i) total building destruction, (ii) houses to be made uninhabitable, and (iii) window shattering following a UVCE in which 25 tonnes of propane has exploded. Assume 10 per cent yield.
(26.4 T TNT equivalent; 104 m; 447 m; 655 m)

5.3 Estimate the blast energy resulting from a steam explosion caused by 500 kg of molten steel at 1700°C entering water at 20°C and 0.1 MPa. The steel has a thermal conductivity of 30 W/mK and a specific heat of 0.45 kJ/kg K. The water has a thermal conductivity of 0.6 W/mK, a density of 100 kg/m³ and a specific heat of 4.2 kJ/kg K. The density of the steel is 7800 kg/m³. Assume 10 per cent yield.
(*c.* 2.8kg TNT)

5.4 A petrochemical complex houses storage vessels with an estimated projected area of 25 000 m² within a site radius of 500 m. If a ten tonne boiler drum at the centre of the complex were to suffer explosive failure, estimate the probability of a domino effect accident sequence occurring. The boiler drum contains 30 m³ of steam at 3 MPa ($\gamma = 1.4$). (For method, see Example 5.3.)
(*c.* 3%)

5.5 Determine whether an 'instantly released' cloud of 5 tonnes of propane will be able to burn after drifting for 500 metres under (i) class D conditions and (ii) class F conditions.
((i) no; (ii) yes)

5.6 A chemical plant stores 200 kg of hydrogen cyanide. Determine the downwind distance for fatal consequences from a hypothetical instant release, assuming neutral conditions (class D) and a 'fatal concentration' of 150 ppm.
(*c.* 500 m)

5.7 In an accident, 50 kg of particulate material are instantly released into a 3 m/s wind in neutral (class D) conditions. Estimate the ground concentration (kg/m²) 500 m downwind, assuming a deposition velocity of 2 cm/s. (Determine an idealised cloud radius $\bar{R}$ in the same way as for Example 5.9.)
(*c.* 200 mg/m²)

5.8 Estimate the dose rate (Gy/s) 1 metre above ground contaminated with 1.0 TBq/km² of Cs-137. Neglect any attenuation by air and the effects of β-particles. Hence determine the annual dose if someone were to live in such a contaminated area (as could conceivably happen after a major nuclear accident). Hence determine the individual risk from living in such an area.
(External dose 0.019 Gy/year; individual risk 2.4×10^{-4} per year of exposure)

5.9 Using data in Table II-3 (Appendix II), estimate the external dose (γ) to a person 1 km downwind from a 1000 MW(th) reactor following an extreme accident in which *all* radioactive gases were released 1000 seconds after reactor shutdown (about 17 minutes) and instantly dispersed to the atmosphere.
Use the same technique as was described in Example 5.9. Neglect β activity. Assume a wind speed of 5 m/s and neutral conditions. The mass absorption coefficient may be taken as 0.0027 m²/kg.
(*c.* 1.65 Sv)

5.10 Following an accident in a fuel reprocessing plant, a twenty-year old employee inhales an

amount of plutonium-239 estimated to be 10 000 Bq. Using the data in Table II-1 (Appendix II) estimate the total additional risk of cancer to the employee as a result of this incident. What mass of plutonium does this represent?
(0.61%; 4.34 μg)

References and bibliography

Advisory Committee on Major Hazards, 2nd report, HMSO 1979.
Advisory Committee on Major Hazards, 3rd report, HMSO 1984.
Alderson M, *UKAEA Safety and Reliability Directorate Report R135*, 1979.
ACGIH (American Conference of Governmental Industrial Hygienists), *Documentation of the Threshold Limit Values*, Cincinnati 1980.
Aylward G H and Finlay T J V, *SI Chemical Data*, 2nd edn, John Wiley and Sons, Sydney 1971.
Ayyaswamy P, *UCLA–ENG–7423*, 1974.
Baker W E, *Explosions in Air*, Univ. Texas Press, Austin 1973.
Beattie J R, Appendix to F R Farmer *Siting Criteria – a New Approach*, IAEA SM-89/34, Vienna 1967.
Blackmore D R, Eyre J A and Summers G G, *Trans.I.Mar.Eng* **94**, 1982, Paper 29.
Board S J and Caldarola L, ASME Symp. *Th. and Hyd. Aspects of Nucl. Reactor Safety*, Atlanta 1977.
Board S J, Hall R W and Hall R S, *Nature* **254**, 1975, 319–21.
Brasie W C and Simpson D W in *Loss Prevention*, vol. 2, AIChE 1968.
Bretherick L, *Handbook of Reactive Chemical Hazards*, Butterworths 1975.
Briggs A J, *UKAEA Report AEEW-R1692*, 1983.
Briggs G A, in *Atmospheric Science and Power Production*, (Ed.) D. Renderson, USDOE Tech. Info. Center, Oak Ridge, TN 1984.
Bryson R A and Hare F K (Eds), *Climates of North America*, Elsevier 1974.
Bush S H, *Nuc. Safety* **14**, 1977, p 3.
Campbell H, Doll W R S and Letchner J. *Brit. med J* 5369, Nov. 1963, p 1370.
Chuse R, *Pressure Vessels*, 5th edn, McGraw-Hill 1977.
Cliff K D, Miles J C H and Brown K, *National Radiological Protection Board Report R159*, 1984.
Cottrell W B, *Nuc. Safety* **3**, 1962, 64–74.
Crump K S and Crockett P W, *J.Haz.Mat.* **10**, 1985, 419–31.
Csanady G T, *Aust. J. Phys.* **10**, 1957, 558–64.
Cude A L, *Chem. Eng.* Oct 1974, p629.
Department of Defense, *Structures to Resist the Effects of Accidental Explosions*, TM5-1300/ NAVFAC P-397/AFM 88-22, 1969.
Dicken A N A, *Proc. Chlorine Institute Bicentennial Symposium*, San Francisco, 1974, 244–56.
Dow Chemical Co., The Dow Safety Guide, 1980, reprint from *Chem. Eng. Prog.* 5th edn, AIChE.
Farmer F R (Ed.), *Nuclear Reactor Safety*, Academic Press 1977.
Fauske H K, *Nuc. Sci. Eng.* **51**, 1973, 95–101.
Fetter S A and Tsipis K, *Scientific American*, **244**(4), 1981, 33–9.
Finney D J, *Probit Analysis*, 3rd edn, Cambridge UP, 1971.
Foster A R and Wright R L, *Basic Nuclear Engineering2*, 4th edn, Allyn and Bacon 1983.
Fremlin J H, *Nucl. Energy* **22**, 1983, 67–73.

Frigerio N A, Eckerman K F and Stowe R S, *Argonne Radiological Impact Program (ARIP), ANL/ES-26* (pt. 1), 1973.
Gast P F, *Trans. Am. Nuc. Soc.* **16**, 1973, p40.
Graham J, *Fast Reactor Safety*, Academic Press 1971.
Gregory C V and Lord D J, *Nucl. Energy* **13**, 1972, 251–60.
Griffiths R F and Megson L C, *Atmos. Environment* **18**, 1984, 1195–1206.
Grist D R, *UKAEA Safety and Reliability Directorate Report R-125*, HMSO, 1978.
Gugan K, *Unconfined Vapour Cloud Explosions*, I.Chem.Eng. 1979.
Hall S F, Philips D W and Peckover R S, *Nucl. Energy* **24**, 1985, 211–27.
Hemming C R, Charles D, Alpert D J and Ostmeyer R M, *National Radiological Protection Board Report R-149*, 1983.
HSE (Health and Safety Executive), *Canvey: an investigation of potential hazards*, HMSO 1978.
HSE (Health and Safety Executive), *Occupational Exposure Limits*, Guidance Note EH 40, HMSO 1984.
HSE (Health and Safety Executive), *A Guide to the Control of Major Accident Hazards Regulations 1984*, Health and Safety Series Booklet HS(R) 21, HMSO 1985.
Hosker R P, *IAEA-SM-181-19*, Vienna, 1974, 291–309.
Howerton A E, in *Loss Prevention*, vol. 3, AIChE 1969.
International Commission on Radiological Protection, *Recommendations of the ICRP* (Publication 9), Pergamon 1966.
International Commission on Radiological Protection, *Recommendations of the ICRP* (Publication 26), Pergamon 1977.
Jennings A J D, *Chem. Eng.* Oct. 1974, 637–41.
Jones C J, *J. Haz. Mat.* **2**, 1978, 363–89.
Jones J A, *National Radiological Protection Board Report R-88*, 1979.
Kelly G N, Jones J A and Hunt B W, *National Radiological Protection Board Report R-53*, 1977.
Kletz T A, in *Loss Prevention*, Vol. 11, AIChE 1977.
Knox J B, *Nuc. Safety* **21**, 1980, 569–71.
Lee J H, Guirao C M, Chui K W and Bach G G, in *Loss Prevention*, vol. 11, AIChE 1977.
Lees F P, *Loss Prevention in the Process Industries*, Butterworths 1980.
Lewis E E, *Nuclear Power Reactor Safety*, Wiley–Interscience 1977.
Lindley A L G and Brown F H S, *Proc. I. Mech. Eng.* **172**, 1958, 627–53.
Lipsett S G, *Fire Technology* **11**, 1966, 118–26.
MacKenzie D, *New Scientist* No. 1451, 11th April 1985, p4.
Margerison T and Wallace M, *The Superpoison*, Macmillan 1981.
Marshall V C, *Chem. Eng.* Oct. 1976, p697.
Marshall V C, *Chem. Eng.* Aug. 1977, 573–7.
Marshall V C, *Chem. Eng.* July 1980, 499–500.
Meteorological Office, *Climatological Atlas of the British Isles*, HMSO 1952.
Newell J, *New Scientist*, 7th Feb. 1985, p6.
Newmark N M and Rosenbleuth E, *Fundamentals of Earthquake Engineering*, Prentice-Hall 1971.
Nicholls C M, Woodcock E R and Gillieson A H, in *Chemical Processing of Reactor Fuels*, (Ed.) J F Flagg, Academic Press 1961.
NRPB, *Emergency Reference Levels ERL-2*, HMSO 1981.
Pasquill F. *Met. Mag.* **90**, 1961, 31–49.
Pasquill F, *Atmospheric Diffusion*, Van Nostrand, London 1968.

Pochin E E, *Env. Health Perspectives* **22**, 1978, 103–5.
Prince A J, Webber D M and Brighton P W M, *UKAEA Safety and Reliability Directorate Report R318*, 1985.
Rasmussen N C, *Reactor Safety Study WASH-1400*, USNRC 1975.
Ricci P F and Cirillo M C, *J. Haz. Mat.* **10**, 1985, 433–47.
Roberts L E J, *Nuclear Power and Public Responsibility*, Cambridge UP 1984.
Royal Society, *Risk Assessment, A Study Group Report*, Royal Society 1983.
Schirripa J T, in *Loss Prevention*, vol. 11, AIChE 1977.
Sittig M, *Handbook of Toxic and Hazardous Chemicals*, Noyes Publications, New Jersey 1981.
Smith N, *A History of Dams*, Citadel Press, New Jersey 1972.
Stevens C, *UKAEA Report AERE–M3409*, 1984.
Strehlow R A and Baker W E, *Prog. Energy Combust. Sci.* **2**, 1976, 27–60.
Strehlow R A and Ricker R E, in *Loss Prevention*, vol. 10, AIChE 1976.
Stull D R, in *Loss Prevention*, vol. 4, AIChE 1969.
Stull D R, *Fundamentals of Fire and Explosion*, AIChE 1977.
Stull D R, in *Loss Prevention*, vol. 7, AIChE 1973.
Sutton O G, *Quart. J. Roy. Met. Soc.* **73**, 1947, 426–36.
Tang Y S, Coffield R D and Markley R A, *Thermal Analysis of Liquid-Metal Fast Breeder Reactors*, ANS 1978.
Thomson J R and Nightingale A, to be published.
Union Carbide, *Bhopal Incident Investigation Team Report*, Danbury, Conn., March 1985.
Van der Hoven I, *Nucl. Safety* **8**, 1967, 490–9.
Van Ulden A P, *Loss Prevention Symposium*, (Ed.) C H Buschmann, Elsevier 1974.
Warn J R W, *Concise Chemical Thermodynamics*, Van Nostrand Reinhold 1969.
Wearne S H, *Proc. I. Mech. Eng.* **193**, 1979, 125–36.
Westbrook G W, *Loss Prevention Symposium*, (Ed.) C H Buschmann, Elsevier 1974.
Withers R M J and Lees F P, *J. Haz Mat.* **12**, 1985, 231–82.
World Health Organisation, *Planning Emergency Response Systems for Chemical Accidents*, WHO, Copenhagen 1981.
Zabetakis M G, *Flammability Characteristics of Combustible Gases and Vapours*, Bull. 627, US Bureau of Mines 1965.

Probabilistic Risk Assessment (PRA)

6.1 The practice of risk assessment

Risk was defined in Chapter 1 to be a function of accident frequency and accident consequences. Thus, to carry out a full probabilistic risk assessment, or PRA, on a given plant, the impact of possible process accidents must be assessed by determining both the likelihood and the consequences of the accidents. Furthermore, a full PRA on a plant will need to consider *all* of the possible process accidents within the plant.

The steps involved in probabilistic risk assessment are shown in the form of a flow diagram in Fig. 6.1. The results of such analysis may be portrayed in the form of a probability–consequence diagram (Fig. 1.2) or, in some cases, as lines of individual 'iso-risk' around a map of the plant. This latter approach has been used for the Rijnmond risk study (Openbaar Lichaam Rijnmond, 1980).

The number of PRA studies that have been undertaken has grown steadily since the Rasmussen report on light water reactors in 1975. By 1984, 22 PRA studies on individual nuclear power stations had been completed (Fussell, 1984) and major studies on chemical plant safety had been published in the UK and Holland. In addition, companies are increasingly using PRA as a means of auditing their own plant safety on an 'internal' basis.

Fig. 6.2 shows some accident data (actual and projected) together with some calculated risk curves from PRA studies.

6.1.1 Monte Carlo methods in PRA

The Monte Carlo method is a powerful mathematical tool for determining the approximate probability of a specific event which is the outcome of a series of stochastic processes. The method consists of determining the results of a (large) number of computational trials; the accuracy of the final answer increases as the number of trials increases.

The method was developed during the Second World War by von Neumann and Ulam as a means of analysing nuclear fission chain reactions. The method enabled calculations of critical mass to be performed by means of numerous computational trials following the histories of individual neutrons through the stochastic process of absorption, scattering and fission. The method only became readily practicable with the event of digital computers.

In PRA, the Monte Carlo method can be used in a number of ways, e.g.

1. The evaluation of the probabilities of system failure modes in large, complicated event trees.
2. The evaluation of pressure vessel failure probabilities by means of trials following the stochastic processes of crack formation, crack detection and crack growth (Section 4.6.4) (Pistone, Venzi and Re, 1984).
3. The evaluation of accident consequences by means of trials following the stochastic processes of population density variation (with the time of day, week or year), wind speed, wind direction and atmospheric stability.

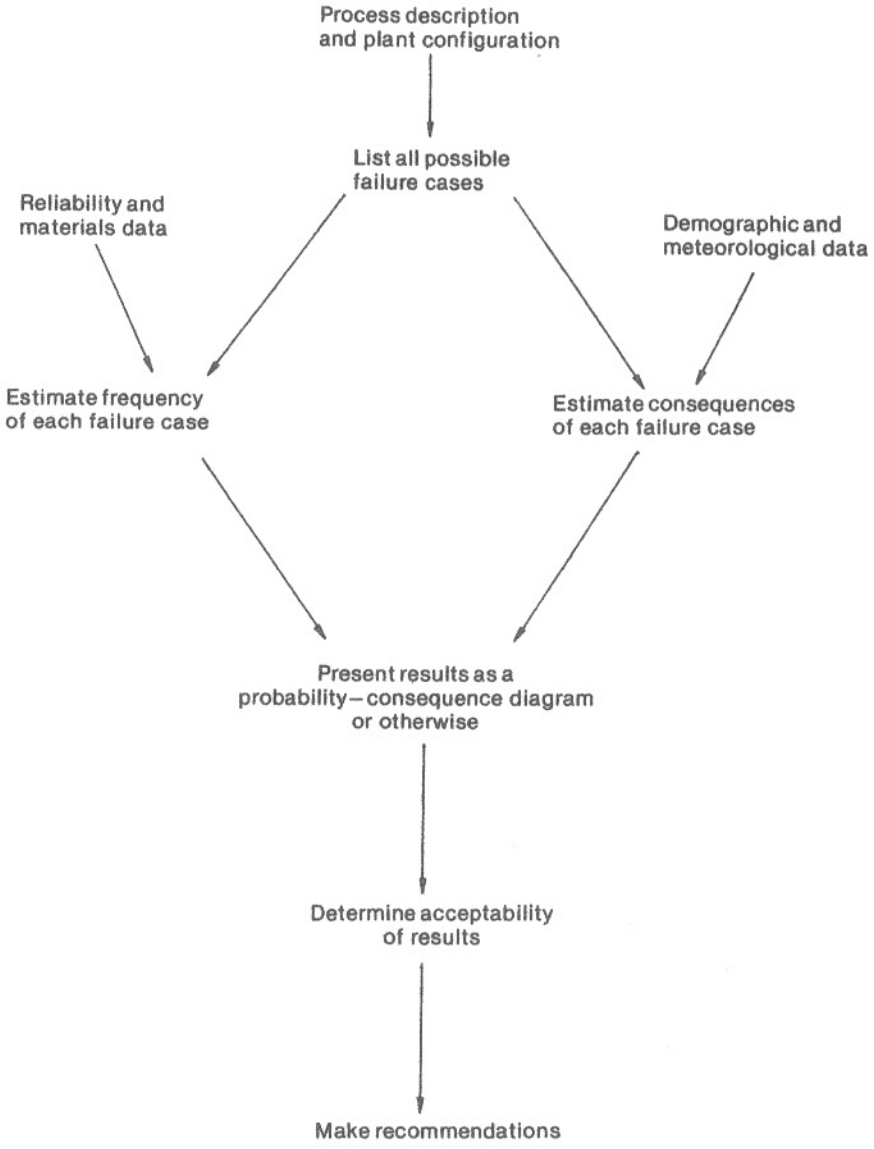


Fig. 6.1 Flow diagram for risk assessment

The use of the Monte Carlo method in PRA is illustrated in Example 6.1. In the method, continuous probability density functions are replaced with a discrete approximation; computer-generated random numbers then select a value for the variable, which in this case will be the position of a breach in a natural gas pipeline.

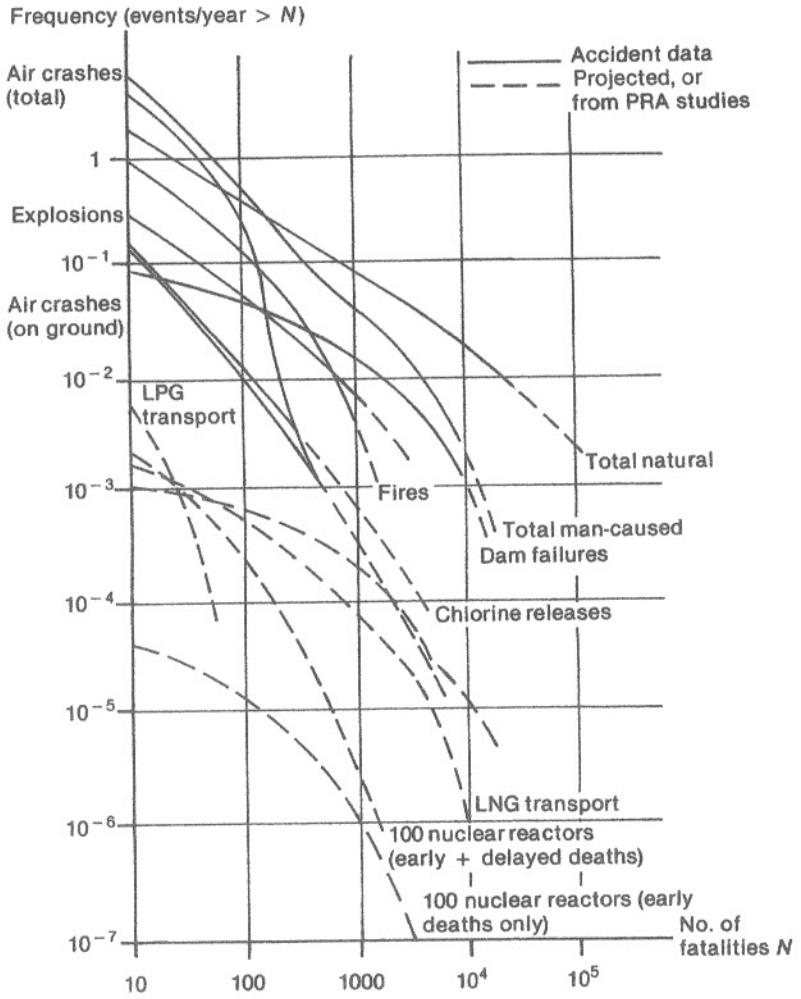


Fig. 6.2 Risk data for the United States (ref: Rasmussen, (1975); Roberts, 1984)). Note that auto accidents probably kill some 50 000 people annually)

Example 6.1

An example of the use of the Monte Carlo method in PRA: the risk assessment of natural ga pipelines (Abes, Salinas and Rogers, 1985).

A 40 km pipeline passes through an 8 km region of high population density. The pipe ir this region is replaced with thicker walled material; this lessens the chances of failure b reducing the hoop stress in the pipe. Provided that data on the frequency of pipe rupture (pe kilometre) as a function of hoop stress are available, then Monte Carlo trials can proceed a

Input data: R_1, R_2, population density distribution, ignition probability, probability of rupture/ probability of leakage, overpressure due to pipe rupture, time to depressurise line following failure.

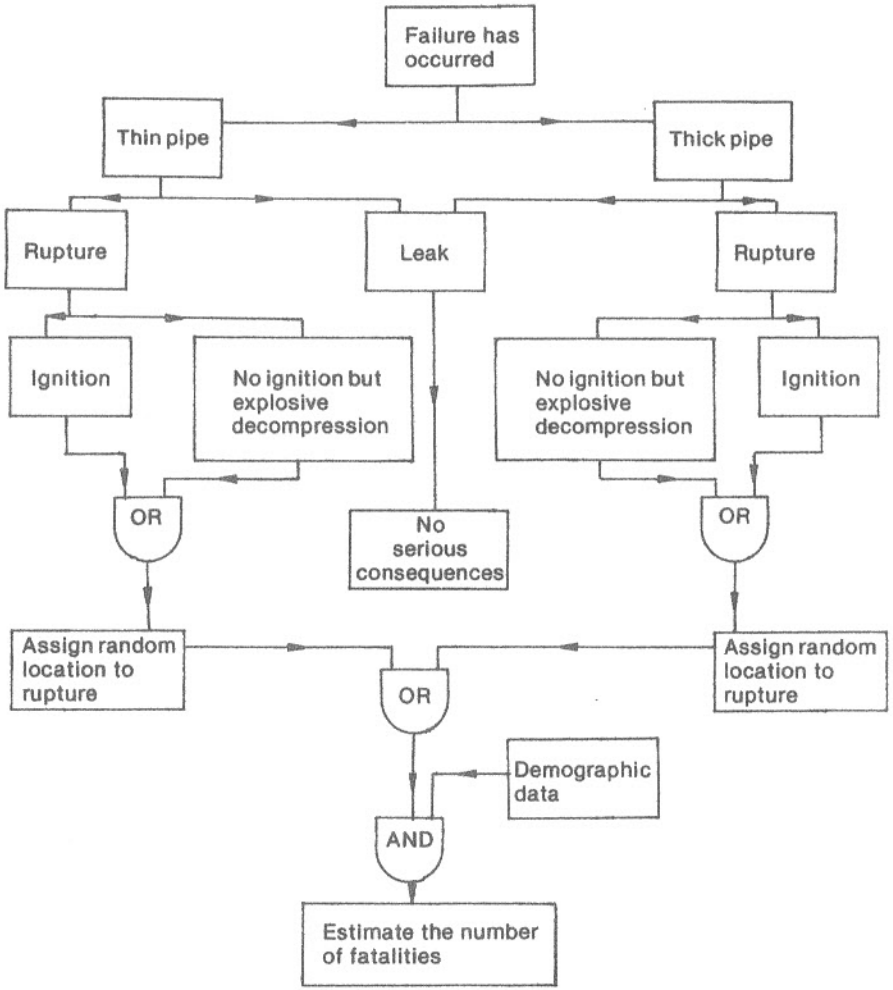


Fig. 6.3 Flow diagram for a single Monte Carlo trial for Example 6.1

shown in the flow diagram (Fig. 6.3). In this diagram the pipe failure probabilities are given by:

R_1 = frequency of failure of thick pipe/kilometre
R_2 = frequency of failure of thin pipe/kilometre

Hence the overall failure frequency is $(8R_1 + 32R_2)$, and the probabilities that a given failure has occurred in the high or low population density areas are

$$\frac{8R_1}{8R_1 + 32R_2} \text{ and } \frac{32R_2}{8R_1 + 32R_2} \text{ respectively.}$$

Abes gives the probability of a given failure being a leak as 66.4 per cent the the remaining 33.6 per cent being ruptures. 11.6 per cent of leaks and 24.9 per cent of ruptures ignite.

In the Monte Carlo computation, the pipe is divided up into a large number of lengths and the position of the failure in each trial is determined by a random number generator. The type of failure (leak/rupture, ignition/no ignition) is then combined with population (demographic) data (Fig. 6.4) and the number of 'fatalities' for each trial can be determined. Once several thousand trials have been completed, a frequency–consequence diagram could be proposed with reasonable confidence.

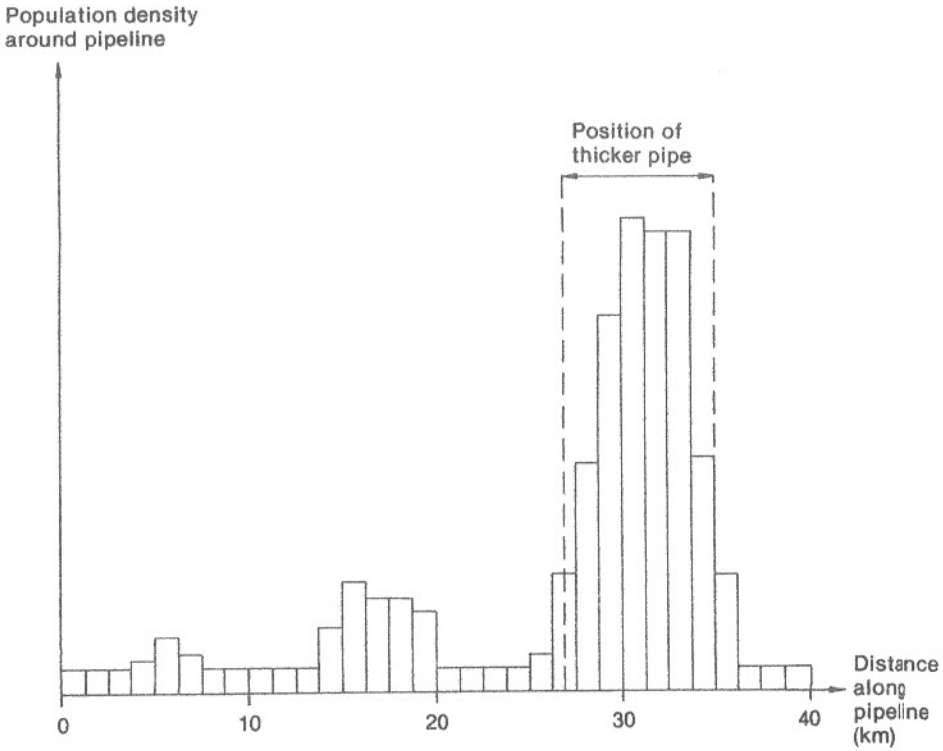


Fig. 6.4 Discretised population distribution near the (hypothetical) pipeline of Example 6.1

6.1.2 The applications for PRA

Cox (1982) has listed four main applications for PRA, as follows:

1. Planning studies.
2. Comparative studies.
3. Deciding priorities for plant improvements.
4. Insurance.

In *planning studies*, PRA may be used to determine the acceptability of proposals to construct new hazardous plant, or to construct dwellings adjacent to existing hazardous plant, or to determine whether existing hazardous plant should be improved. (Since it will often be the case that 'zero-risk' is not possible, it follows that deciding these issues requires the resolution of some difficult questions about the acceptability of risk. These will be examined further in Section 6.2.)

Comparative studies of the safety of competing designs of hazardous plant may assist at the design stage, in addition to other factors such as cost. The accuracy of PRA may not always be sufficient to discriminate clearly between alternative designs, however (Section 6.1.3).

Where the safety of a plant is to be improved, PRA may enable decisions to be made on the *priorities* of the improvements, i.e. which plant improvement yields the greatest improvement in plant safety? It may also be possible to introduce cost into the calculations, and attempt to answer the question "Which plant improvement saves the greatest number of (statistical) lives *and* is the expenditure worthwhile?" This, however, may imply some knowledge of the worth of a human life (Section 6.2.1(iii)).

PRA may also be used for determining *insurance* premiums for hazardous plant. Traditionally, premiums are determined by examining the accident record of similar plant, but this may not always be possible where the plant is of, say, a new design, or employs a new process. In addition, the recent occurrence of disastrous accidents such as Ixhuatepec and Bhopal (see Table 1.1) will probably lead to third-party damages settlements of many hundreds of millions of dollars. It would seem likely that a consequence of such accidents will be increased use of PRA as a means of assessing third-party liabilities.

6.1.3 The accuracy and cost of PRA

The accuracy of PRA is difficult to assess. The low frequency of multiple-fatality accidents mean than many hundred years of plant operation would have to be accumulated before any meaningful comparison between 'measured' and 'predicted' risk become possible. In addition, any PRA study will be subject to uncertainty because of the stochastic nature of the process involved. Thus, the dashed lines in Fig. 6.2 (representing the results of PRA studies) should properly be replaced by bands representing confidence limits (see Example 2.6). Confidence limits for the WASH-1400 reactor safety study (Rasmussen, 1975) are shown in Fig. 6.5. The merit of presenting confidence limits for such low probability events is questionable.

PRA studies are also subject to error for any or all of the following reasons:

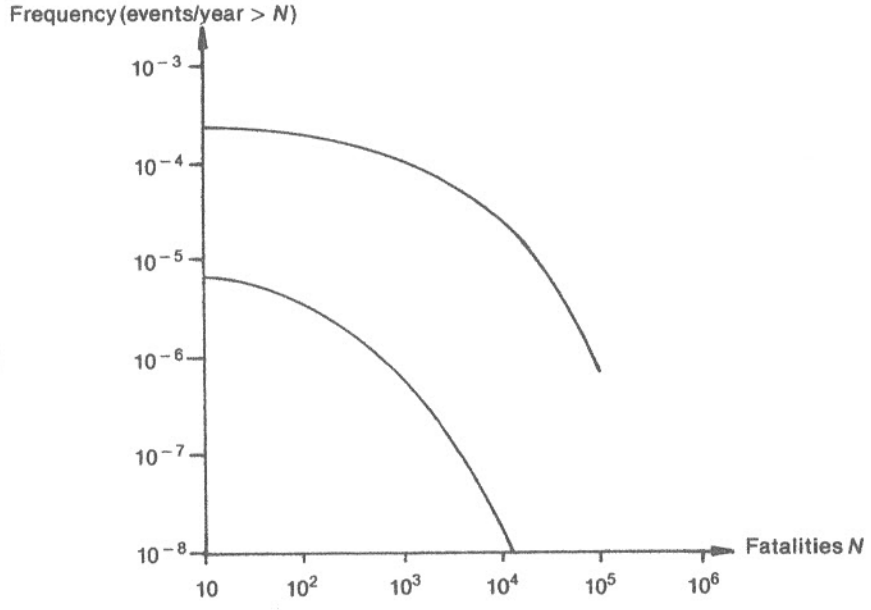


Fig. 6.5 Ninety per cent confidence limits for the WASH-1400 reactor safety study. (100 reactors, early deaths only) (Rasmussen, 1975)

1. The PRA does not consider all possible failure cases or failure routes. This omission will lead to underestimating accident probabilities, and possibly also accident consequences. There can be no guarantee against the omission of failure cases in PRA, as the assessment of possible failure cases relies on the judgement of the engineers doing the assessment.
2. Errors in the estimation of accident probability may occur if inaccurate data are used, or if the system is incorrectly modelled. The importance of a reliability databank is apparent. Cox (1982) states that the major cause of inaccuracy in PRA lies in this area.
3. Errors in the estimation of accident consequences may arise, for example, if incorrect toxicological data are used.

If PRA is to be used for comparative studies, it is important that the different types of consequences of accidents should be considered. These are:

1. Early or prompt deaths.
2. Delayed deaths and/or genetic effects.
3. Injuries.
4. Environmental effects.
5. Property damage.

In comparative studies, it is important to ensure that the basis for comparison is the

same.* Again, this may give rise to awkward questions concerning the value of human life; this will be further considered in Section 6.2.1(iii).

Finally, to carry out a full probabilistic risk analysis of a large complex plant can be an expensive proposition. The Canvey study reportedly cost £400 000, so a similar study at today's prices would probably cost in excess of £1 m.

6.2 *The implications of risk assessment*

6.2.1 Cost, risk and benefit

6.2.1(i) *The perception of risk*

Slovic *et al.* (1981) carried out an interesting study in which people were asked to judge the frequency of 41 cases of death. The participants were first told the annual death-toll from motor vehicle accidents (50 000 in the USA) and then asked to estimate frequencies for the other 40. The results are shown in Fig. 6.6. It would appear that people grossly overestimate death rates from rare causes such as botulism or tornadoes (by as much as two orders of magnitude) while underestimating death rates from more common causes such as cancer or heart disease by an order of magnitude or so.

The relevance of this result to PRA is open to question. However, it makes one point quite clear: the individual's perception of risk seldom bares any relation to the true magnitude of that risk. Therefore, it might be suggested that subjectivity should have no place in risk assessment. However, since the responses to accidents from the press and public are modulated by these same subjective risk assessments, it is very rash to ignore public perception of risks when assessing the acceptability of those risks.

Lee (1981) has pointed out that the public is much affected in its perception of risk by the potential size of a single catastrophe; public concern about nuclear power is thus due to the perception of the possibility of a devastating accident. This, in turn, is presumably due to a lingering suspicion (in spite of reassurances to the contrary) that a nuclear power station could blow up like a bomb. This idea that potential catastrophe size affects risk perception seems riddled with inconsistencies, however; for example, the public attitude to dam failures seems to be largely one of indifference. Who, outside India, remembers the Gujarat dam failure? (Bhopal, which killed fewer, will be remembered more vividly – presumably because of the television reports.) In a similar vein, the response of the UK public to the Canvey report seemed less than overwhelming, in spite of the discussion in the first report (HSE, 1978) of accidents which might cause 18 000 deaths.

This inconsistency in public attitude has been partially explained by McLean (1981) thus: injury following a truly 'accidental' event (being struck by lightning, say) is more readily acceptable to the public than the same injury inflicted, possibly

* The difficulties of comparing different types of harm have been discussed in ICRP Publication 45, Pergamon, Oxford 1985.

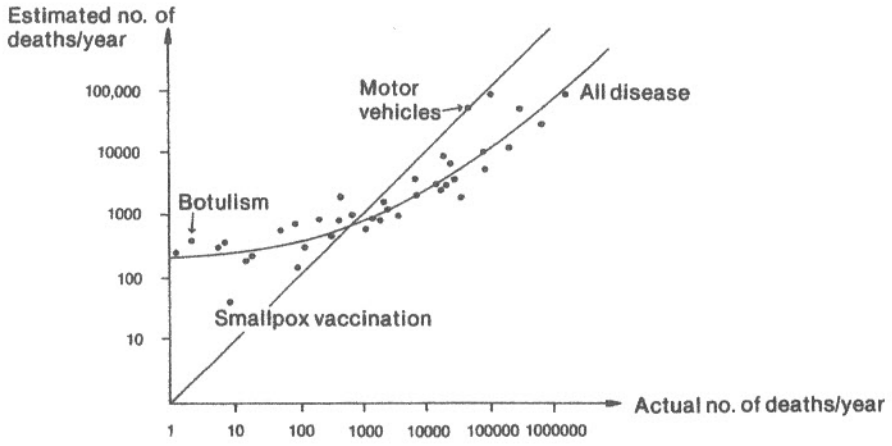


Fig. 6.6 The relationship between estimated and actual numbers of deaths per year in the USA from 41 causes of death (Slovic *et al.*, 1981)

indirectly, by a person. Hence an industrial accident in which malice or carelessness can be attributed will provoke a greater public response than an accident which is perceived (rightly or wrongly) as an 'act of God'. In the viewpoint of the ubiquitous man-in-the-street, dam failure may fall into the latter category, and chemical or nuclear plant accidents may fall into the former category.

A report by the Council for Science and Society (1977) had the following statement as its first conclusion: "The acceptability of risks cannot be simply derived from a scientific study of quantified probabilities, cost and benefits. The human factor influences the analysis at every point. But fairness in decisions and effectiveness in controls of risks can be approached by the use of scientific methods among others, provided that the diversity of human interests, values and perceptions of risks is always respected".

6.2.1(ii) *Comparative magnitudes of risk*

Three methods of measuring individual risk are in common use. These are:

1. Loss of expectation of life, in years.
2. Fatal accident rate (FAR) for industrial accidents, measured in deaths/10^8 working hours. (The expression 'Fatal Accident Frequency Rate', or FAFR, was formerly used. This, however, is tautological.)
3. Deaths per million per year.

For a 2000-hour working year, it is apparent that a FAR of 1 is equivalent to 20 deaths per million per year in the population exposed to the hazard.

Typical data for loss of expectation of life are shown in Table 6.1. This type of data is often used to describe the effects of chronic exposure to carcinogens over a long period: these effects will not normally be represented in FAR data. For a life expecta-

Table 6.1 Loss of expectation of life from various hazards

Population	*Cause of death*	*Loss of expectation of life*
All women	Complications of pregnancy	0.01 years
All men	Road accidents	0.30 years
Women handling blue asbestos in gas mask assembly Second World War	Lung diseases	1.51 years
Underground coal miners, male	Accidents, pneumoconiosis	2.40 years
Nickel-refiners, male	Respiratory diseases and cancers	3.97 years
Male doctors smoking 15–24 cigarettes per day	Diseases associated with smoking	5.45 years

Ref: Royal Society (1983)

tion of 72 years, and a working life of 40 years at 2000 working hours per year, the following approximate conversion factors may be used:

1 year loss of life expectation 347 deaths/10^6/year
1 year loss of life expectation 17 deaths/10^8 working hours

These conversion factors must be treated with great caution for a number of reasons. First, and most important, deaths per million per year may be defined in terms of *total* population, and not just exposed population (Grist, 1978). Furthermore, life expectation and working hours will not be the same for different groups of workers. Finally, these factors implicitly assume that 36 people each losing one year of life is equivalent to one person losing 36 years of life; this is questionable.

Some data showing FARs and deaths per million per year from various causes are given in Table 6.2. Columns (1) and (3) yield an interesting comparison – the trends are exactly reversed between them. The difference, of course, is due to Grist's death rate data being expressed in terms of whole population, and not just the exposed population. Since the first two columns of data neglect any effect of an industrial accident on the surrounding community, they are not really suitable for use in the assessment of major hazards. 'Whole population' death rate data is thus to be preferred in such assessments.

Data showing death rates due to all accidents and all causes are shown in Fig. 6.7. These data provide useful baselines against which other death rate data may be compared.

The presentation of data relating to transport accidents requires special consideration. At least six different methods of presenting such data are commonly employed, viz:

1. Fatal accidents per million journeys.
2. Fatal accidents per million hours.

Table 6.2 Fatal accident rates and death rates

(a) Classification by job or location

	FAR[1] (10^8 *working hours*)$^{-1}$ *Exposed population only*	*Death rate*[2] (10^6 *year*)$^{-1}$ *Exposed population only*	*Death rate*[3] (10^6*yr*)$^{-1}$ *Whole population*
All British industry	4	85	34
Chemical industry	4	85	—
Shipbuilding	7	105	—
Agriculture	10	110	2.5
Coal mining	12	210	2.1
Home	1	—	106

(b) Classification by type of accident[3]

	Death rate[3] (10^6 *year*)$^{-1}$ *Whole population*
Road accidents	122
Poisoning accidents (drugs)	11
Falls	111
Fire	16
Venomous animals and insects	0.085
Lightning	0.1
Drowning	12
Electrocution	2.5
Suicide	76
Homicide	10

Notes: 1. UK data, Carson and Mumford (1979). 2. UK data, Royal Society (1983). 3. England and Wales, Grist (1978).

3. Deaths per million passenger-journeys.
4. Deaths per million passenger-hours.
5. Deaths per hundred million passenger-miles.
6. Individual risk (death rate for whole population).

It is apparent that there is a great scope for confusion. Some relevant data are given in Table 6.3. Warren states that the target risk level for current aircraft design is 0.3 fatal accidents per million hours.

These data achieve little except to tell us what sort of risks are normally encountered in everyday life. The magnitude of these risks tell us little about their public acceptability; there are, doubtless, some people who are greatly concerned about the prospect of snakebite or being struck by lightning even though the magnitude of those risks is much smaller than, say, the risk of being murdered (Table 6.2).

6.2.1(iii) *The value of human life: risk versus cost*

Any potentially hazardous installation or operation can be made safer by the increased use of safeguards, diverse control systems, multiple containment barriers and redundant structures. These is no limit to the number of such safety-enhancing devices

which could be installed (except perhaps in aircraft, where there will be a weight limitation). However, the benefit in terms of marginal risk reduction will diminish as the amount spent increases (Fig. 6.8).

One method by which the amount that ought to be spent on risk reduction can be determined is given by point (i) on Fig. 6.8(b). This is the point of unity slope (after

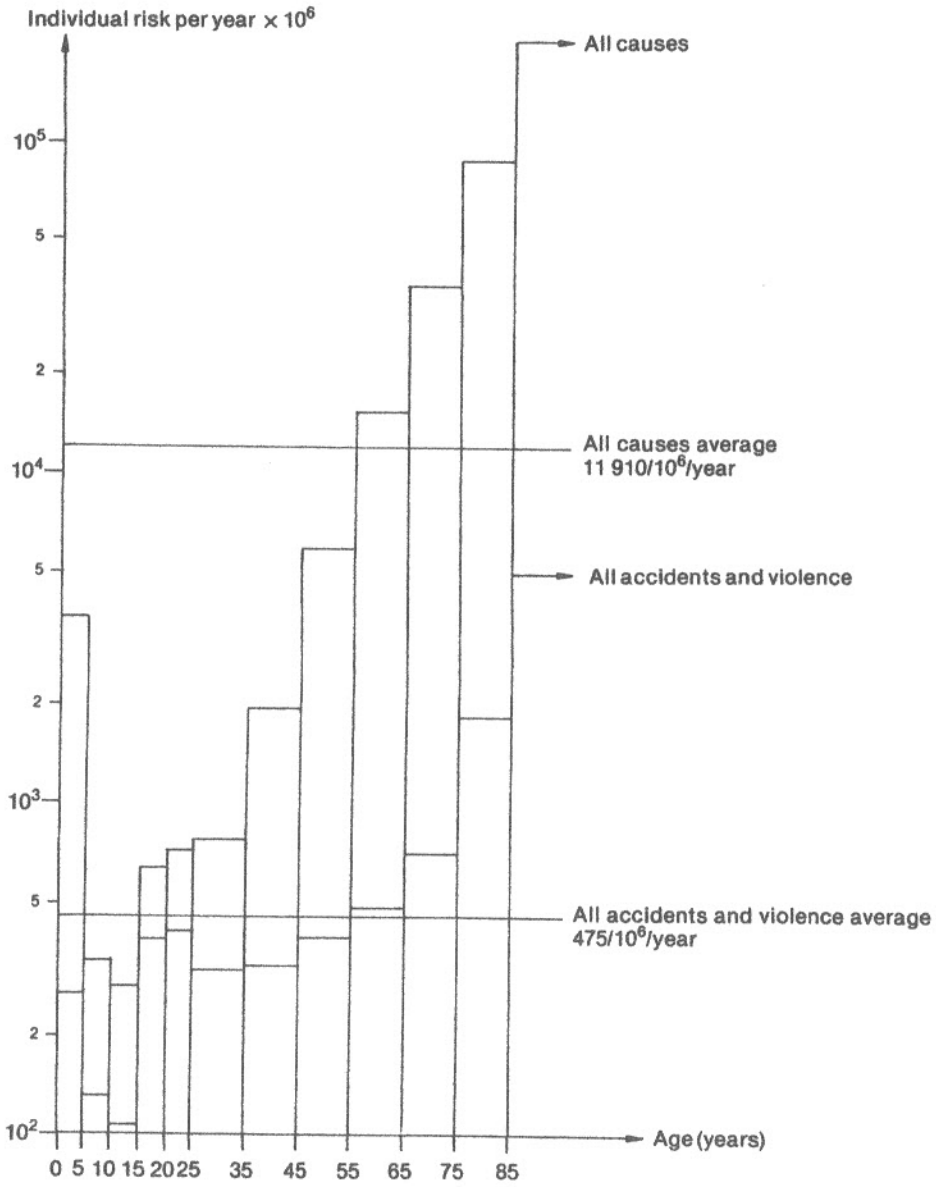


Fig. 6.7 Death rates as a function of age, England and Wales 1971–1975 (Grist, 1978)

Table 6.3 Risk data for transport accidents

	Deaths/10⁸ passenger-km[1]	*Deaths/10⁶ journeys, Exposed population*[1]	*Death rate (10⁶ yr)⁻¹ whole population*[2]
Aviation	0.24	1.8	1.4
Bicycles	8.3	—	—
Motor cycles	20	—	
Automobiles	1.2	0.027	122
Buses	0.22	—	
Trucks	0.5	—	
Rail passengers	0.14	0.059	3.3

Notes: 1. World data, Warren (1977). 2. England and Wales, Grist (1978).

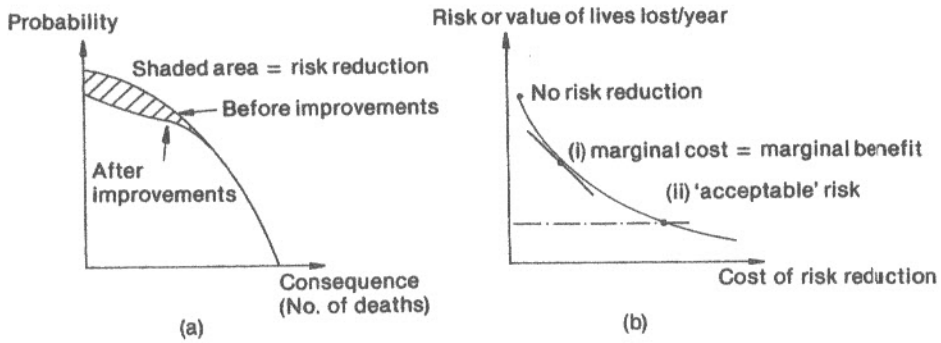


Fig. 6.8 The calculus of risk and benefit (Rowe, 1977)
(a) Risk reduction, where risk (statistical deaths per year) = Σ(probability × consequence)
(b) Cost effectiveness of risk reduction

normalising both axes in terms of cost) where the marginal cost of risk reduction equals the marginal benefit in terms of value of lives saved. This calculation, however, implies that a human life can be given an agreed monetary value. This is an extremely difficult value judgement; so difficult, in fact, that it may be unethical even to make the judgement. As Pearce (1981) has pointed out, "Some people see the attachment of money values to human life and suffering as in some way morally offensive". Nevertheless, various authors have attempted to answer the question by looking at the amount that is spent, in various areas of human activity, on either risk reduction or compensation for risk taken. These are summarised in Table 6.4. The range of values for a human life is thus $1500 to $282 million.

Other means of assessing how much should be spent on risk reduction, which avoid the human life value judgement, have been proposed. One criterion is that risk levels should be 'as low as practicable', defined thus (Rowe, 1977); "When the incremental cost per risk averted is such that a very large expenditure must be made for a relatively

Table 6.4 Some estimates of the value of a human life

Method	*Value of Life*
Mean discounted value of income[1]	$375 000
Amount spent on road safety[1]	$45 000
Amount spent on aviation safety[1]	$1 350 000
Amount spent on agricultural safety[1] (employees)	$15 000
Relationship between risky jobs and wage rates[1]	$300 000
Amount spent on high-rise apartment safety[1]	$30 000 000
Amount spent on nuclear safety[2]	$282 000 000
Third world famine relief[3]	$16 000
Cervical smears[3]	$9 000
UK Govt. decision not to introduce child-proof drug containers (1971)[4]	$1 500
Suggested expenditure on radiation dose reduction[5]	$150 000 to $64 000 000

All values in 1985 US dollars, approximately

Refs: [1]Okrent (1979); [2]Siddall (1980); [3]Roberts (1984); [4]Ennis (1985); [5]Calculated from NRPB (1980) using ICRP 26.

small decrease in risk as compared with previous risk reduction steps". This definition is obviously inexact and, furthermore, it implies that the risk–cost curve will not be smooth (as shown in Fig. 6.8(b)) but will have discrete steps or bumps; this does not seem an unreasonable supposition. Some authors differentiate between the 'as low as practicable' approach and the 'as low as reasonably achievable' approach (also called the ALARA principle), but there is no clear-cut distinction.

A further method of judging worthwhile safety investment was given in the Royal Society Study Group Report on Risk Assessment (1983), quoting Black, Niehaus and Simpson. In this method, the risk of construction and installation of the additional safety equipment is also considered in determining total risk (Fig. 6.9). This then enables the optimum safety investment, to yield minimum risk, to be assessed. This neatly avoids the difficulty of making any judgement of the value of a life. However it could be argued that this approach neglects other, less tangible, risks. Spending on safety implies the gainful employment of people who might otherwise be unemployed. Should not this factor also be considered for a truly humanitarian view of safety?

Finally, there is another criterion which avoids human life value judgement. If an absolute value of 'acceptable risk' can be agreed, then a universal safety standard could be adopted. This is shown as point (ii) in Fig. 6.8(b). This is discussed further in Section 6.2.3.

None of these criteria is unambiguous. A summary table is given as Table 6.5. Nonetheless, a major difficulty remains; if no criterion for investment on risk reduction can be agreed, then such spending may proliferate without limitation, particularly in those industries where the public perception of risk is high. Siddall (1980) has called this proliferation of safety costs 'ratchetting'. This phenomenon explains the particularly high value for spending on nuclear safety per (statistical) life saved, as given in Table 6.4.

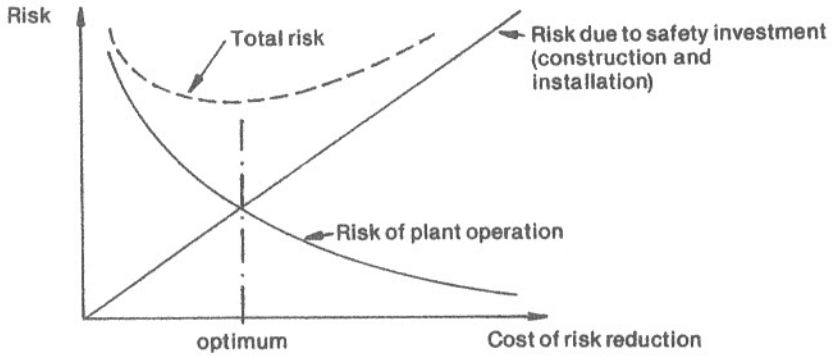


Fig. 6.9 Minimising risk (Royal Society, 1983)

Table 6.5 Some criteria for determining worthwhile investment in risk reduction

Criterion	*Difficulties*
1. Marginal cost = marginal benefit	(i) Value of human life (ii) Morality of criterion
2. As low as practicable, or ALARA	Imprecise criterion
3. 'Minimum' risk (Fig. 6.9)	(i) Neglect of less tangible risks (ii) Neglects any consideration of cost
4. 'Acceptable' risk	Definition of 'acceptable'

6.2.1(iv) *Cost/benefit analysis*

Until now, the definition of risk given in Chapter 1 has been retained:

$$\text{risk} = \Sigma\,(\text{frequency}) \times (\text{consequence})$$

This definition of risk implies that one fatal accident per year is as acceptable (or unacceptable) as an accident which could kill 10 000 people with a frequency of 10^{-4}/year; this implication is difficult to justify. Evidence suggests that societal aversion to risk varies as $(\text{consequence})^{\alpha}$, where the *aversion index* α is greater than unity (Okrent, 1981). Okrent has suggested that risk acceptability should be determined from the *equivalent social cost*, defined thus:

$$\text{equivalent social cost} = \Sigma\,(\text{frequency}) \times (\text{consequence})^{\alpha}$$

The proposed value for α is 1.2.

If equivalent social cost is used as a measure of risk, then those accidents with low frequencies which may kill large numbers are more likely to be deemed unacceptable.

The implications of this public aversion to large accidents should be considered when assessing the acceptability of industrial risks.

Using the idea of marginal costs and benefits introduced in Section 6.2.1(iii), a cost/benefit analysis for a proposed plant improvement might now proceed as follows:

1. Estimate the reduction (X) in the equivalent social cost due to the proposed modification:

$$X = \left[\int_0^\infty N^\alpha \, df\right]_{\substack{\text{before}\\\text{modification}}} - \left[\int_0^\infty N^\alpha \, df\right]_{\substack{\text{after}\\\text{modification}}}$$

   Here f is accident frequency, N is accident consequence and α is the aversion index.
2. Estimate the remaining operating lifetime of the plant Y.
3. Estimate the cost of the proposed modification Z.
4. The plant improvement is worthwhile if

$$XY > \frac{Z}{\text{(value of a human life)}}$$

   i.e. the benefits exceed the cost.
5. If the plant improvement is worthwhile but plant revenues are insufficient to pay for the modification, then there is a case for shutting down the plant.

6.2.1(v) *Cost/benefit analysis: some legal aspects*

In the United Kingdom, it is the duty of employers to ensure the health, safety and welfare of employees 'as far as reasonably practicable', under the terms of the Health and Safety at Work Act 1974. In the event of a court case following an accident, the onus is on the employer to demonstrate that all 'reasonably practicable' measures were taken to prevent such an accident occurring. If an employer had previously decided not to improve safety in a way that would have prevented the accident, then any argument based on cost/benefit analysis would have to be very carefully argued. The analysis would have to show clearly that the proposed improvements were not 'reasonably practicable' and that costs exceeded anticipated benefits by a sufficient margin to put the matter beyond reasonable doubt.

It seems not unreasonable to suggest that criteria for the acceptability of such cost/benefit analyses will eventually be determined in the courts. Such criteria might even go so far as to yield values for 'acceptable risk' (Section 6.2.2). In the meantime, however, there are very few legal precedents to give guidance.

An interesting and at least partly relevant case is that of the 1972 Ford Pinto, a 'compact' car produced by Ford in the United States (Grimshaw vs Ford Motor Company, Orange County Superior Court, 1981). This car had been shown, during manufacturer's tests, to have a high probability of catching fire in rear-end collisions above 20 miles per hour, in breach of a proposed new Federal regulation. Evidence produced during the trial suggested that Ford had nevertheless decided to proceed with production of the unmodified car, following an analysis which suggested that the costs of

design modifications, put at $10.9 m, would exceed the damages payable by Ford to the victims of such rear-end collisons.

Judge Leonard Goldstein disliked this reasoning and fined Ford $125 m punitive damages, in addition to $2.5 m compensatory damages.

This case, though interesting, is not directly applicable to cost/benefit analysis for two reasons. Firstly, Ford were in breach of a (proposed) new regulation in any case, and secondly, their cost/benefit analysis was incomplete in that it only considered the costs and benefits to the Ford Motor Company. No attempt was made to put a value on lives lost, except in so far as it affected Ford in terms of likely damages payable.

One tentative conclusion from this case might be as follows: for cost/benefit analysis of improvements in safety to stand up in court, the costs to society and the bereaved as well as the costs of the improvements in safety need to be considered. The equation of 'value of life' with 'likely damages payable' is not acceptable.

6.2.2 Emergency planning

Any factory or other industrial installation where the potential for severe accidents exists should have an emergency plan prepared; in some countries, certain categories of installation (e.g. nuclear plant) are required by law to have made such arrangements before plant operation can commence.

Any such plan should cover four main objectives:

1. Clear definitions of the lines of communication from middle to senior management, and on to the local authorities, for notification of any accident (actual or impending).
2. Procedures for the setting up of an accident control centre, together with clearly defined chains of command.
3. Procedures for the assessment of risks from data on types of potential accident, characteristics of materials involved, local topography and demography, weather conditions at time of accident, available resources, other industries nearby which may be affected, etc.
4. Actions to minimise risks by means of evacuation, calling up additional resources, cleaning up a contaminated area, shut down of any hazardous plant which may still be operating, confiscation of any contaminated foodstuffs, etc.

For an 'on-site' accident the plant operators should be able to *initiate* each of the above objectives. For an 'off-site' accident, e.g. a transport accident, the local authorities may have to deal with the problem entirely by themselves, at least in the early stages, until people with specialist knowledge of the hazard can be contacted. Following this line of reasoning, accidents may be categorised into four levels (WHO, 1981):

1. Operator level. A minor or site incident which does not exceed the site boundary.
2. District level. A minor incident which occurs off-site, e.g. a transport accident.
3. Regional or national level. An accident with major off-site implications, or a major transport accident.

4. International level. Such an accident could involve a hazardous installation close to an international boundary, or a maritime accident.

An accident may escalate from a lower to a higher category, and the geographical location of the emergency control centre may have to move accordingly. The chains of command and communication become more difficult the higher the accident level becomes. For international level accidents, the command chains multiply; in addition, the organisational frameworks differ from one country to another. The WHO booklet (WHO, 1981) discusses the different systems of emergency response in Europe.

It is often the case that a major incident calls into question the existing emergency planning arrangements. This occurred in Italy following the Seveso incident and in the USA following the Three Mile Island incident. The success of the emergency planning arrangements following the Mississauga incident in 1979 has been discussed in Section 5.1.4(i).

6.2.3 Criteria for engineering safety

The use of 'acceptable risk' as a criterion for engineering safety has already been mentioned in Section 6.2.1(iii) (Fig. 6.8(b)). Such a criterion could thus take the form of an acceptable number of statistical deaths per year. However, in view of the previously discussed public aversion to large-consequence accidents, a more suitable form of safety criterion takes the form of a line on a frequency–consequence diagram, as discussed in Chapter 1. Criteria of this sort (for single installations or complexes) have been proposed by Farmer (1967) and Provinciale Waterstaat Groningen (1979), and are shown in Fig. 6.10. The Groningen criterion is realistic in that it acknowledges the existence of large uncertainties in PRA studies, although the size of zone 'requiring further assessment' seems large, occupying no less than four orders of magnitude of accident frequency. Furthermore, the criterion stipulates that the possibility of any accident with a consequence greater than 10 deaths 'requires further assessment' even if the frequency of such an accident is only once per hundred million years. Due to the possibility of common-mode failures (Section 3.5) this part of the criterion is likely to be difficult to meet. For example, the frequency of an installation being struck by an aircraft is around once in fifteen million years (Section 5.1.7), and such a crash might well initiate a severe accident in say, a hazardous chemical plant. (Nuclear installations, however, are often designed to resist a light aircraft crash.) Thus, the Groningen criterion is severely restrictive for low probability, high consequence accidents.

The Farmer criterion (1967) was orginally based on 'quantity of iodine-131 released' instead of 'number of fatalities'. These values have been converted to fatalities for differing population densities (Bell, 1977) and are shown in Fig. 6.10. This criterion is simple to invoke in that it is clear cut; this, however, tends to suggest a degree of certainty in the results of PRA which is not warranted.

Finally, a somewhat simpler criterion has been given by Okrent (1979), quoting Bowen. This criterion simply suggests that the frequency of accidents causing off-site loss of life should not exceed 10^{-5} to 10^{-4} events per plant per year. The criterion does not therefore differentiate between low and high consequence accidents.

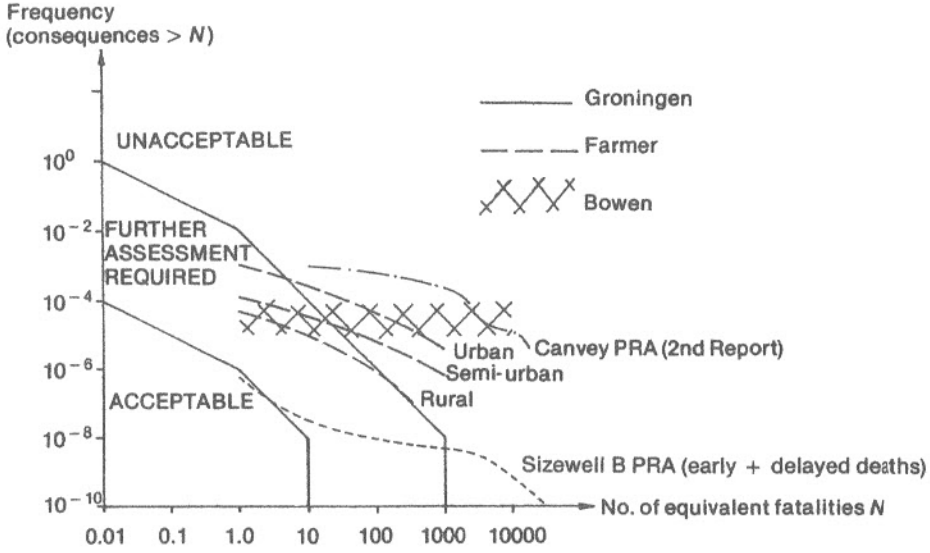


Fig. 6.10 Three proposed safety criteria (Groningen, Farmer and Bowen) compared with two PRA studies (Canvey petrochemical complex and Sizewell B pressurised water reactor)

6.3 Conclusions

Probabilistic Risk Assessment, or PRA, is a synthesis of the techniques of systems reliability assessment, structural reliability assessment and consequence analysis, together with consideration of the probabilities of external hazards such as earthquakes or aircraft crashes. It is applicable to all industrial activities which employ hazardous processes or materials.

Risk may (sometimes) be considered to be equal to the area under the curve on a frequency–consequence diagram, but human perception of risk seldom tallies with measured (or calculated) risk. However, this definition of risk can be used to give guidelines in making investment decisions relating to improvements in safety. If human life is accorded a monetary value, then those decisions can even be made on a basis of cost versus benefit. The morality of this type of decision-making is open to question.

An alternative approach is to attempt to define 'acceptable' risk in terms of a line on the frequency–consequence diagram. Some proposed criteria for acceptable risk have been presented in Fig. 6.10.

Questions

6.1 A modification to a plant is estimated to reduce risk by 0.01 statistical deaths per year. The cost of the modification will be \$1.6 m. and the plant has an expected remaining life of 15

years. Assuming an aversion index α of 1.0 (i.e. risk = equivalent social cost) try to decide whether the modification should be carried out, by using various values of human life from Table 6.4.

6.2 The modification in the above question will require an estimated 2000 man-days of site construction work. Using values for accident rates given in Table 6.2, suggest whether the 'benefit of construction' exceeds the 'risk of construction'.

6.3 (a) Estimate (graphically) the risk (in statistical deaths/year) attached to the Sizewell 'B' and Canvey plants, as shown in Fig. 6.10.
(b) Estimate the corresponding 'equivalent social costs', using an aversion index of 1.2.
(c) The first Canvey report, before improvements were made, gave the following frequency–consequence data for accidents:

No. of casualties	10	1500	3000	4500	6000	12 000	18 000
Frequency	31.4×10^{-4}	17×10^{-4}	10.8×10^{-4}	6.1×10^{-4}	3×10^{-4}	1.7×10^{-4}	1.0×10^{-4}

Suggest how much might have been worth spending to achieve such reductions in risk, assuming aversion indices of (i) 1.0 and (ii) 1.2.

References and bibliography

Abes A J, Salinas J J and Rogers J T, *Struct. Safety* **2**, 1985, 225–37.
Bell G P, in *Nuclear Reactor Safety*, (Ed.) F R Farmer, Academic Press, New York 1977.
Carson P A and Mumford C J, *J Haz. Mat.* **3**, 1979, 149–65.
Council for Science and Society, *The Acceptability of Risk*, Barry Rose Publishers Ltd, London 1977.
Cox R A, *J Haz. Mat.* **6**, 1982, 249–60.
Ennis J, *New Scientist* 2nd May 1985, 26–8.
Farmer F R, *Siting Criteria, a New Approach*, IAEA SM-89/34, Vienna 1967.
Fussell J B, *IEEE Trans. Reliability* **R33**, 1984, 41–7.
Grist D R, *UKAEA Safety and Reliability Directorate Report R125*, 1978.
HSE, *Canvey: an investigation of potential hazards*, HMSO 1978.
HSE, *Canvey: a second report*, HMSO 1981.
ICRP Publication 45, *Quantitative Bases for Developing a Unified Index of Harm*, Pergamon, Oxford, 1985.
Lee T R, *Proc. Roy. Soc.* **A376**, 1981, 5–17.
McLean A E M, *Proc. Roy. Soc.* **A376**, 1981, 51–65.
NRPB, *Application of Cost–Benefit Analysis to Radiological Protection*, HMSO 1980.
Okrent D, *Nuc. Safety* **20**, 1979, 148–64.
Okrent D, *Proc. Roy. Soc.* **A376**, 1981, 133–48.
Openbaar Lichaam Rijnmond, *Rapport van de COVO-Stuurgroep*, Rotterdam 1980.
Pearce D W, *Proc. Roy. Soc.* **A376**, 1981, 181–90.
Pistone V, Venzi S and Re G, in *Reliability of Engineering Materials*, (Ed.) A L Smith, Butterworths, London 1984.
Provinciale Waterstaat Groningen, *Pollution control and use of norms in Groningen – criteria for risks related to dangerous goods*, Groningen, Netherlands, April 1979, 43–57.
Rasmussen N C, *Reactor Safety Study WASH-1400*, USNRC 1975.
Roberts L E J, *Nuclear Power and Public Responsibility*, Cambridge UP 1984.
Rowe W D, *An Anatomy of Risk*, John Wiley and Sons, New York 1977.

Royal Society, *Risk Assessment, A Study Group Report*, Royal Society, London 1983.
Siddall E, *Nuc. Safety* **21**, 1980, 451–60.
Slovic P, Fischhoff B and Lichtenstein S, *Proc. Roy. Soc.* **A376**, 1981, 17–34.
Warren D V, Appendix to Council for Science and Society (1977).
World Health Organisation, *Planning Emergency Responses for Chemical Accidents*, WHO, Copenhagen 1981.

Recommended further reading

Although references have been cited in the text, a number of books deserve special mention for any reader who wishes to read in more detail about particular aspects of safety assessment.

American Nuclear Society, *Proc. Int. ANS/ENS Meeting on PRA*, New York 1981.

Bignell V and Fortune J, *Understanding System Failures*, Manchester UP, Manchester 1984.

Billinton R and Allan R A, *Reliability Evaluation of Engineering Systems*, Pitman, London 1983.

British Standards Institution, Handbook 22, *Quality Assurance*, BSI, London 1983.

Farmer F R (Ed.), *Nuclear Reactor Safety*, Academic Press, London 1977.

Gugan K, *Unconfined Vapour Cloud Explosions*, George Godwin, London 1979.

I.Chem.E., *The Assessment and Control of Major Hazards*, I.Chem.E. Symposium Series No. 93, 1985.

ICRP Publication 45, *Quantitative Bases for Developing a Unified Index of Harm*, Pergamon, 1985.

Kletz T A, *What Went Wrong?*, Gulf Publishing Co., Houston 1985.

Lees F P, *Loss Prevention in the Process Industries*, Butterworths, London 1980.

Pasquill F, *Atmospheric Diffusion*, van Nostrand, London 1962.

Shooman M L, *Probabilistic Reliability: An Engineering Approach*, McGraw-Hill, New York 1968.

Smith A L (Ed.), *Reliability of Engineering Materials*, Butterworths, 1984.

Stull D R, *Fundamentals of Fire and Explosion*, AIChE, New York 1977.

The Canvey reports and the WASH-1400 report are also recommended.

A useful source for publications covering the many applications of risk analysis is *An Annotated Bibliography on Risk Analysis*, published by the Institution of Mechanical Engineers, 1 Birdcage Walk, London SW1H 9JJ.

Storage and use of hazardous chemicals in the United Kingdom is governed by *The Control of Industrial Major Accident Hazards Regulations 1984* (known as CIMAH) published by Her Majesty's Stationery Office.

Appendix I The behaviour of rising plumes

As stated in Section 5.3.5, plumes of gas or aerosol will rise upwards if (i) the gas is buoyant, (ii) the gas or aerosol is hot (or generates its own heat, as might be the case for radioactive plumes), or (iii) the gas or aerosol possesses initial upward momentum in the manner of, say, a jet of material escaping via a leak in a high pressure system.

Plume rise modelling has been reviewed by Briggs (1984). Except in strong winds, plume rise increases the effective source height from 2 to 10 times the actual release height for typical elevated, buoyant sources. In this way the maximum ground concentration is reduced by a factor of 3 to 100 or even more.

Although theoretical solutions to the problem of plume final rise (or effective stack height) (see Fig. I.1) exist, adequate data for testing these models do not. Hence the accuracy of plume rise modelling must remain questionable. Also, the problem of plume rise under unstable conditions (Pasquill categories A, B and C) remains unresolved.

All analytical plume rise models are based on the conservation of mass, buoyancy and momentum. One further assumption is required, which is called the *closure assumption*. The usual closure assumption adopted is that the velocity of entrainment of air into the plume, v_e, is proportional to the velocity of rise, w, of the plume. The three conservation relations, together with the closure assumption, lead to the following expression for buoyant plumes. (A similar expression may be derived by straightforward dimensional analysis.)

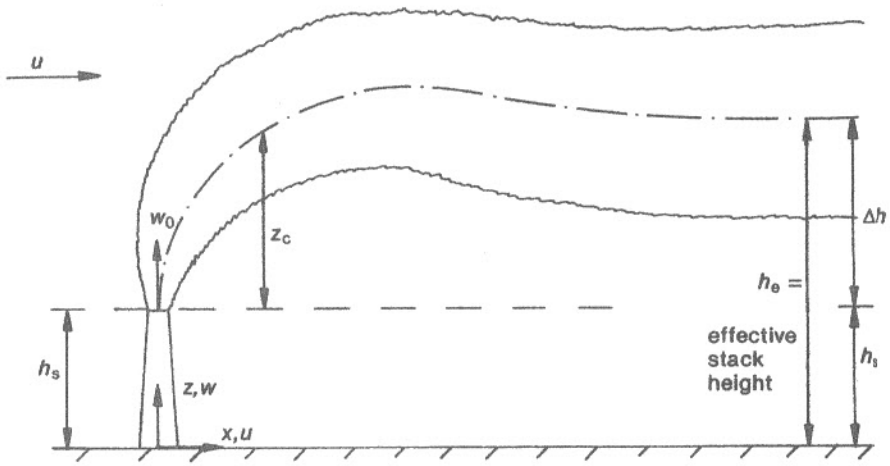


Fig. I.1 Schematic illustration of plume rise nomenclature

$$z_c = \frac{\alpha F_b^{\frac{1}{3}} x^{\frac{2}{3}}}{u} \qquad [I.1]$$

where z_c is the height of the plume centreline above the source (m),
x is the downwind distance (m),
u is the windspeed (m/s),
F_b is the buoyancy flux at the source, and
α is a coefficient equal to 1.6.

The buoyancy flux may be determined thus:

$$F_b = g\frac{\Delta\rho}{\rho_a}\dot{V} \qquad [I.2]$$

where $\Delta\rho$ is the difference in density between the plume and the surrounding air (kg/m³),
ρ_a is the density of air (kg/m³),
g is the acceleration due to gravity (m/s²), and
$\dot{V}$ is the volumetric flow rate of the plume at the source (m³/s).

(For a jet, the following relation applies:

$$z_c = \left[\frac{3 F_m}{\beta^2 u^2}\right]^{\frac{1}{3}} x^{\frac{2}{3}} \qquad [I.3]$$

where β is a coefficient equal to $\left(0.4 + \frac{1.2\,u}{w_o}\right)$,
F_m is equal to $(w_o \dot{V})$, i.e. momentum flux/ρ) (m⁴/s²), and
w_o is the initial upward velocity (m/s).)

The final rise, Δh, of even warm, buoyant plumes under stable conditions is given by

$$\Delta h = 2.6\left[\frac{F_b}{u\,s}\right]^{\frac{1}{3}} \qquad [I.4]$$

where s is the ambient stability parameter, which for gaseous plumes is given by

$$s = \left\{\frac{g}{\theta}\cdot\frac{\partial\theta}{\partial z}\right\} \qquad [I.5]$$

Here θ is the *ambient potential temperature*. The potential temperature of air does not change when it is adiabatically lifted, i.e.

$$\theta = T\left\{\frac{p_s}{p}\right\}^{\frac{\gamma-1}{\gamma}} \qquad [I.6]$$

where p_s is a standard pressure (usually 10^5 N/m²),
T is the absolute temperature (K), and
γ is the ratio of the specific heats (c_p/c_v) which is equal to 1.40 for air.

The potential temperature gradient is given by

$$\frac{\partial\theta}{\partial z} = \frac{\theta}{T}\left\{\frac{\partial T}{\partial z} + \frac{g}{c_p}\right\} \qquad [I.7]$$

where values for $\frac{\partial T}{\partial z}$ for different stability categories are given in Table 5.4. The term (g/c_p) represents the *adiabatic lapse rate* which for earth's lower atmosphere equals 0.98 K/100 m.

In conclusion, eqns [I.4], [I.5], and [I.7] illustrate how this analysis is not applicable to unstable conditions, since the equations would then predict a negative final plume rise. Although plumes may indeed sink in unstable conditions, these equations cannot accurately predict the degree to which this occurs. The equations also suggest that, under neutral conditions where $\frac{\partial T}{\partial z}$ equals $-(g/c_p)$, plume rise will continue indefinitely since $s = 0$.

Reference

Briggs G A, Plume rise and buoyancy effects, in *Atmospheric Science and Power Production*, (Ed.) D. Renderson, USDOE Tech. Info. Center, oak Ridge, TN 1984.

Appendix II Factors for determining the effects of airborne radioactivity

Table II.1 Inhalation and dose–risk factors (Kelly, Jones and Hunt, 1977; ICRP, 1977; Hemming, Charles, Alpert and Ostmeyer, 1983)

Organ	*Dose–risk factor (deaths/ man-Sv)*	*Principal isotopes affecting organ*	*Inhalation factor (Gy/GBq inhaled) after*			*Radiation and half-life*	*RBE*
			1 day	*1 year*	*50 years*		
Thyroid	3.1×10^{-4}	Te-131	0.297	32.4	32.4	β, γ, 30 h	1
		Te-132	18.9	54.0	54.0	β, γ, 77 h	1
		I-131	11.6	270	270	β, γ, 8.1 d	1
		I-133	15.7	43.2	43.2	β, γ, 22 h	1
		Pu-238	0.002	2.00	513	α, 89.6 y	20
		Pu-239	0.002	1.86	594	α, 2.4×10^4 y	20
		Pu-240	0.002	1.86	594	α, 6580 y	20
		Am-241	0.035	29.7	1620	α, 432 y	20
		Cm-242	0.038	16.5	29.7	α, 35 y	20
		Cm-244	0.038	32.4	783	α, 17.9 y	20
Red bone marrow (leukaemia)	1.5×10^{-3}	Sr-90	0.046	75.6	756	β, 27.7 y	1
		Pu-238	0.014	14.9	3240		
		Pu-239	0.013	14.0	3780		
		Pu-240	0.013	14.0	3780		
		Pu-241	—	0.013	81	α, 14 y	20
		Am-241	0.26	227	10 300		
		Cm-242	0.30	121	205		
		Cm-244	0.27	235	5130		
Endosteal cells (bone cancer)	3.1×10^{-4}	Sr-90	0.046	75.6	756		
		Pu-238	0.178	186	43 200		
		Pu-239	0.167	176	48 600		
		Pu-240	0.167	176	48 600		
		Pu-241	—	0.013	1026		
		Am-241	3.24	2970	130 000		
		Cm-242	3.78	1540	2570		
		Cm-244	3.51	2970	64 800		
Lower large intestine	6.3×10^{-4}	Ru-106	3.51	35.1	37.8	β, 1 y	1
		Ce-144	3.51	32.4	35.1	β, 285 d	1

Table II.1 continued

Organ	*Dose–risk factor (deaths/ man-Sv)*	*Principal isotopes affecting organ*	*Inhalation factor (Gy/GBq inhaled) after*			*Radiation and half-life*	*RBE*
			1 day	*1 year*	*50 years*		
Lungs	1.3×10^{-3}	Y-91	2.11	97.2	99.9	β, 61 d	1
		Zr-95	0.54	40.5	40.5	β, 63 d	1
		Ru-106	5.13	675	1053		
		Te-127	0.675	35.1	35.1	β, 9.3 h	1
		Te-129	1.97	40.5	40.5	β, 32 d	1
		Ce-144	4.59	567	783		
		Pm-147	0.22	35.1	78.3	β, 2.6 y	1
		Pu-238	19.7	3510	15 900		
		Pu-239	18.4	3240	16 200		
		Pu-240	18.4	3240	16 200		
		Pu-241	—	2.59	151		
		Am-241	18.9	918	918		
		Cm-242	21.0	783	783		
		Cm-244	20.0	945	972		
Liver	6.3×10^{-4}	Ce-144	0.019	11.3	29.7		
		Pu-238	0.049	48.6	9180		
		Pu-239	0.043	45.9	10 300		
		Pu-240	0.046	45.9	10 300		
		Pu-241	—	0.046	211		
		Am-241	0.864	756	27 000		
		Cm-242	0.972	405	648		
		Cm-244	0.918	783	14 850		
Ovaries or testes	1.9×10^{-3}	Pu-238	0.002	2.00	513		
		Pu-239	0.002	1.86	594		
		Pu-240	0.002	1.86	594		
		Am-241	0.035	29.7	1620		
		Cm-242	0.038	16.5	29.7		
		Cm-244	0.038	32.4	783		

Notes: 1. Radiation type, half-life and RBE (Relative Biological Effectiveness) are given only once for each isotope.

2. Inhalation factors are only given for those isotopes where the 50 year inhalation factor exceeds 27 Gy/GBq (10^5 Rads/Ci).

3. Dose–risk factors given take account of the time of appearance of cancer (assumed to have a log-normal distribution) and the population age distribution (assumed to be that of the UK population in 1977).

Table II.2 Some fission product activities in uranium fuel after 1000 days' irradiation

Isotope	*Radiation and half-life*	*Thermal fission yield in U-235 (% per fission)*	*Cooling time*	*Activity TBq/MW*
Kr-85	β(0.85 MeV) 10.6 y	0.293	0	15.0
			6 months	14.5
Sr-89	β(1.48 MeV) 51 d	4.79	0	895
			6 months	74.5
Sr-90	β(0.6 MeV) 28 y	5.77	0	89
			6 months	87.9
Ru-106	β(0.03 MeV) 1 y	0.38	0	588
			6 months	416
Te-132	β(0.35 MeV) γ(0.22 MeV) 77 h	4.7	0	1391
			6 months	—
I-131	β(0.60 MeV) 8.0 d	3.1	0	1125
			6 months	0.0001
I-133	β(1.4 MeV) 20.8 h	6.9	0	2054
			6 months	—
Xe-133	β(0.346 MeV) 5.3 d	6.62	0	2065
			6 months	—
Xe-135	β(0.9 MeV) 9.2 h	6.1	0	827
			6 months	—
Cs-137	β(1.17, 0.518 MeV) γ(0.662 MeV) 30 y	6.15	0	122
			6 months	121

Data from Abbey, F. in *Nuclear Reactor Safety*, (Ed.) F.R. Farmer, Academic Press 1977.

Table II.3 Activities of gaseous and volatile fission products in uranium fuel after 1000 days' irradiation

	Activity × *energy* (TBq–MeV/MW)				
Cooling time (secs)	0	10^2	10^3	10^4	10^5
Total gases (beta)	16 280	5960	2090	696	366
Total gases (gamma)	21 090	10 693	4366	2017	300
Total volatiles (beta)	38 295	17 575	9953	4625	2017
Total volatiles (gamma)	63 825	43 660	25 826	11 322	5439

Data from Abbey, F. in *Nuclear Reactor Safety*, (Ed.) F.R. Farmer, Academic Press 1977.

Appendix III Confidence limits for the expected value of a Poisson distribution

Total observed count	*99 per cent confidence limits*		*95 per cent confidence limits*		*Total observed count*	*99 per cent confidence limits*		*95 per cent confidence limits*	
	lower limit	*upper limit*	*lower limit*	*upper limit*		*lower limit*	*upper limit*	*lower limit*	*upper limit*
0	0.0	5.3	0.0	3.7					
1	0.0	7.4	0.1	5.6	26	14.7	43.2	17.0	38.0
2	0.1	9.3	0.2	7.2	27	15.4	43.5	17.8	39.2
3	0.3	11.0	0.6	8.8	28	16.2	44.8	18.6	40.4
4	0.6	12.6	1.0	10.2	29	17.0	46.0	19.4	41.6
5	1.0	14.1	1.6	11.7	30	17.7	47.2	20.2	42.8
6	1.5	15.6	2.2	13.1	31	18.5	48.4	21.0	44.0
7	2.0	17.1	2.8	14.4	32	19.3	49.6	21.8	45.1
8	2.5	18.5	3.4	15.8	33	20.0	50.8	22.7	46.3
9	3.1	20.0	4.0	17.1	34	20.8	52.1	23.5	47.5
10	3.7	21.3	4.7	18.4	35	21.6	53.3	24.3	48.7
11	4.3	22.6	5.4	19.7	36	22.4	54.5	25.1	49.8
12	4.9	24.0	6.2	21.0	37	23.2	55.7	26.0	51.0
13	5.5	25.4	6.9	22.3	38	24.0	56.9	26.8	52.2
14	6.2	26.7	7.7	23.5	39	24.8	58.1	27.7	53.3
15	6.8	28.1	8.4	24.8	40	25.6	59.3	28.6	54.5
16	7.5	29.4	9.4	26.0	41	26.4	60.5	29.4	55.6
17	8.2	30.7	9.9	27.2	42	27.2	61.7	30.3	56.8
18	8.9	32.0	10.7	28.4	43	28.0	62.9	31.1	57.9
19	9.6	33.3	11.5	29.6	44	28.8	64.1	32.0	59.0
20	10.3	34.6	12.2	30.8	45	29.6	65.3	32.8	60.2
21	11.0	35.9	13.0	32.0	46	30.4	66.5	33.6	61.3
22	11.8	37.2	13.8	33.2	47	31.2	67.7	34.5	62.5
23	12.5	38.4	14.6	34.4	48	32.0	68.9	35.3	63.6
24	13.2	39.7	15.4	35.6	49	32.8	70.1	36.1	64.8
25	14.0	41.0	16.2	36.8	50	33.6	71.3	37.0	65.9

Data from Ricker, W E, *J. Am. Stat. Assoc.* **32**, 1937, 349–86.

Appendix IV Assessing the risk associated with iodine-131 contamination in milk – a sample calculation

The Chernobyl accident in the Ukraine in 1986 caused some concern about food contamination, particularly contamination of fresh milk by iodine-131. Similar concern arose following the Windscale fire of 1957. This concern arises because the consumption of milk constitutes the major 'exposure pathway' to man for low-level radiation following reactor accidents. (It should be pointed out that iodine-131 cannot be released from reprocessing plants under any circumstances, since spent fuel despatched to such plants is always allowed to 'cool' for at least three months. Since the half-life of iodine-131 is only eight days, no significant quantity of the isotope ever reaches a reprocessing plant.)

Following the Chernobyl accident, iodine-131 levels in fresh milk in the UK rose to, typically, 50–100 Bq/litre. It is possible to determine approximate values for individual and societal risks in the UK by making the following (mostly conservative) assumptions:

1. Each person consumes half a litre of fresh milk per day.
2. Milk contamination remained at a constant level of 100 Bq per litre for two weeks. (Contamination level is a function of rainfall and weather patterns as well as time since release.) Contamination level was thereafter zero. This is most probably a conservative assumption.
3. Thirty per cent of ingested iodine-131 is transferred to the thyroid gland (Strather *et al.*, 1984).
4. All I-131 radioactive decay takes the form of 0.6 MeV beta particles. (This is conservative since a significant fraction of I-131 decays involve the release of a lower energy particle together with a gamma ray. Such a gamma ray is quite likely to escape from the body without doing biological damage, whereas beta particles will transfer all their energies to the body tissues.)
5. The thyroid dose–risk factor is taken to be 3.1×10^{-4} deaths/Sv. (Appendix II.)
6. The thyroid mass is typically 20 g for adults (Kelly *et al.*, 1977). The thyroid is smaller in children, which will lead to higher risk than that for adults.

Hence we can calculate that the total thyroid uptake is ($14 \times 50 \times 0.3$) or 210 Bq. The maximum possible thyroid dose from this uptake will be (total energies of radioactive decays/thyroid mass) or

$$\frac{210 \text{ Bq} \times (8 \times 24 \times 3600) \text{ secs} \times 0.6 \text{ MeV} \times 1.6 \times 10^{-13} \text{ J/MeV}}{\log 2 \times 20 \times 10^{-3} \text{ kg}}$$

which is equal to 1 mSv, assuming that all beta particles released within the thyroid gland dissipate their energies there.* From (5), the individual risk of death due to thyroid cancer will therefore be 3.1×10^{-7}.

(Note that the 'normal' individual risk of death from thyroid cancer is 1×10^{-6} per annum (Campbell *et al.*, 1963) so for a person with a remaining life expectancy of 40 years, the 'normal' individual risk due to thyroid cancer would be 4×10^{-5} in any case.)

The additional societal risk to the UK population can be estimated by multiplying 3.1×10^{-7} by the UK population (5.6×10^{7}) to yield a societal risk of *c.* 17 additional thyroid cancer deaths. There will be about 1500 deaths due to thyroid cancer over the next thirty years in any case, due to the normal incidence of the disease.

The above analysis has assumed that the linear dose–risk hypothesis is valid. As has been pointed out already (Fig. 5.25), some data suggest that this hypothesis overestimates the consequences of low level radiation exposure. The other assumptions made in this analysis suggest that this is an extreme upper bound value for the likely consequences in the UK from this exposure pathway. The true figure is likely to be considerably less.

References

Campbell H, Doll W R S and Letchner J, *Brit. Med. J.* 5369, Nov. 1963, p. 1370.

International Commission on Radiological Protection, *Recommendations of the ICRP* (Publication 30), Pergamon, Oxford, 1979.

Kelly G N, Jones J A and Hunt B W, *National Radiological Protection Board Report R-53*, 1977.

Strather J W, Wrixon A D and Simmonds J R, *National Radiological Protection Board Report R-171*, 1984.

* Doses to organs due to ingestion of food contaminated by other radioactive isotopes can be ascertained from data in ICRP 30.

Appendix V Risks to members of the public from various means of electricity generation

Coal or oil (cancers induced by benzo-a-pyrene)	*c.* 1–15 deaths/GW(e)-year
Hydro-electricity (dam failures)	*c.* 10 deaths/GW(e)-year
Nuclear (cancers due to routine discharges)	*c.* 0.05–0.2 deaths/GW(e)-year
(nuclear reactor accidents)	*c.* 1 death/GW(e)-year
Draught proofing of houses (i.e. energy saving) (cancers due to increased concentrations of radioactive radon from stone and brick)	*c.* 100 deaths/GW(e)-year

Ref: Fremlin, J H *Nuc. Energy* **22**, 1983, 67–73.

Notes: 1. Coal miner deaths, due to mine accidents and pneumoconiosis, have a rate of about 10/GW(e)–year in the United Kingdom.
2. Uranium miner deaths due to lung cancers have a rate of not more than 0.1/GW(e)–year.

Index

a posteriori failure probability, 22
a priori failure probability, 22
Abbeystead, 121
airborne material, dispersal of, 149
aircraft crashes
  Comet 1, 82
  DC-10, 74
  risk on ground, 131
ALARA principle, 200
Alexander Kielland, 74, 80
atmospheric stability, 150
availability, 10

bathtub curve, 63
Bayes' theorem, 14, 29
BCS theory, 97
Bhopal, 75, 129, 164, 167
blast
  overpressure, 143
  scaling, 142
  yield, 141
boiling liquid expanding vapour explosions (BLEVEs), 121
bridge networks, 43

Canvey, 4, 156, 194
carcinogens, 129
chemical toxicity, 154
Chernobyl, viii, 131, 178, 182
combinations, 9
combustion, 133
common mode failures, 74
complementary events, 11
conditional events, 11
confidence limits, 30, 193
confined explosions, 121
continuous variables, 18
cost/benefit analysis, 201
crack
  arrest, 86
  critical size, 93
  detection, 107
  deterministic growth, 103
  incidence in pressure vessels, 110
  opening displacement (COD), 97
  origins, 80
  probabilistic growth, 113
critical stress intensity factor, 85
criticality, accidental nuclear, 131
cumulative probability, 16
cut sets, 45

dam failures, 131
death rate, 198
deflagration, 120, 121, 139
delta–star transformation, 49
deposition, 164
derived working limits, 177
detonation, 121, 139
discrete variables, 15
dispersal of airborne material, 149
dispersion coefficients, 152, 155, 175
distributions
  binomial, 24
  Erlangian, 33
  exponential, 20, 64, 116
  gamma, 31
  Gaussian, 18, 152
  lognormal, 34, 111
  Poisson, 26, 66
  Rayleigh, 31
  Weibull, 31

earthquakes, 132
emergency planning, 203
emergency reference levels, 182
equivalent social cost, 201
event trees, 52
expectation, 17
explosions
  blast scaling, 142
  boiler, 91, 123
    energy of, 144
  boilding liquid expanding vapour explosions (BLEVEs), 121
  chemistry of, 133
  confined, 121
  effects of blast overpressure, 143
  hazard potential, 136
  physical, 123, 144
  probability of, 140
  TNT equivalent, 141
  turbine, 126
  unconfined vapour cloud explosions (UVCEs), 121, 141
  vapour (steam) explosions, 126, 145
  yield, 141
exponential distribution, 20, 64, 116

failure
  assessment diagram, 101
  function, 18, 39, 62
  mode and effect analysis, 60
  modes, 54
  probability
    a posteriori, 22
    a priori, 22
Farmer criterion, 4, 204
fatal accident rate (FAR), 195
fault trees, 52
flammability of gas–air mixtures, 137
flash fires, 120
Flixborough, 75, 123
fracture
  brittle, 87
  ductile, 87
  fast, 86
  locus, 85
  mechanics
    linear elastic, 96, 99
    probabilistic, 109
  resistance, 89
  toughness, 85

Gaussian distribution, 18, 152
Groningen risk criterion, 204

hazard rate, 18, 63, 116
HazOp, 61
highly reactive materials, 123
human
  reliability, 75
  life expectation, 195
  value of life (monetary), 200
hydrogen embrittlement, 91

incidence matrices, 47
incremental tearing, 86
independent events, 10
inhalation factors, 176, 212
inhalation rate, 177
irradiation embrittlement, 92

life expectation, human, 195
linear elastic fracture mechanics, 96, 99
loading system, effects of in fracture, 90
Los Alfaques, 120
loss prevention, 3

majority voting system, 41, 66
mean, 17
mean time between failures (MTBF), 65
mean time to failure (MTTF), 21, 64
missile damage, 146
Mississauga, 127, 204
Molten Fuel–Coolant Interactions (MFCIs), 126
Monte Carlo methods, 187
mortality
  in explosions, 143
  in nuclear accidents, 130
  in toxic release accidents, 166
mutually exclusive events, 11

natural gas systems (methane), 60, 190
nil ductility temperature, 91
non-destructive examination (NDE), 107
  reliability of, 111
nuclear accidents, 130, 168

Oppau, 123

parallel systems, 40, 64
permutations, 9
physical explosions, 123, 144
plume dispersal, 151, 161, 209
  effects of, 158
pool fires, 120
Potchefstroom, 80
pressure testing (pneumatic vs. hydraulic), 90
pressure vessel reliability, 113, 116
probabilistic fracture mechanics, 109
probabilistic risk assessment, 187
probit functions, 156

radiation
  activity, 168
  dose, 169
    equivalent, 170
    from cloud, 172
    from deposited activity, 173
  effects
    from external exposure, 171
    of internal exposure, 176
  field strength (flux), 169
  inhalation factors, 176
rare events approximation, 13
reaction hazard index, 136
relative biological effectiveness (RBE), 170
reliability
  accuracy of analysis, 71
  data, 71
  human, 75
  of metal structures, 113
  software, 76
  standby systems, 42, 66
  systems, 39
  time-dependent effects upon, 63
residual stresses, 93
risk, 5
  assessment, probabilistic, 187
  cost, and benefit, 194
  individual, 176
  magnitudes of, 195
  perception of, 194
  societal, 176
running liquid fire, 120

safety criteria, 7, 199, 204
safety valves, 125
series systems, 39, 64
Seveso, 129, 165, 204
simultaneous events, 12
social cost, equivalent, 201
software reliability, 76
spreading length, 95
stability, atmospheric, 150
standard deviation, 17
standby supplies, 67, 69
standby systems, 42, 66
steam explosions, 126, 145
stress concentration theory, 97, 99

test frequency, 68
Three Mile Island, 4, 54, 56, 130, 181
time-dependent effects upon reliability, 63
TNT equivalent, 141
toxic gas transport, 166
toxicity, chemical, 154
turbine explosions, 126

unconfined vapour cloud explosions (UVCEs), 121, 141
unrevealed faults, 68

value of human life (monetary), 200
vapour density, effect on plume dispersal, 161, 209
vapour explosions, 126, 145
variables
  continuous, 18
  discrete, 15
variance, 17

washout, 165
Windscale fire, 174, 181
working limits, derived, 177

Yield, blast, 141