7 Detail design

Statistical assessment of the probability of failure is presented as an alternative to the use of factor of safety. The emphasis is on a better understanding of the limiting factors associated with the design. Quality and reliability are presented together and distinction drawn between the two. The most common mode of failure for structural components is by cyclic loading or fatigue. The chapter includes a section covering avoidance of stress concentrations using pipe flow analogy. High cycle fatigue life prediction is the main subject matter.

7.1 Introduction

Following the embodiment stage the next stage is to consider individual components and ensure that the design or selection of these is optimized.

During the detail design process the design and selection of each component is verified and information prepared which will enable manufacture to commence. The input to the detail design stage is the scheme drawing and the design intent. As in all other stages, all decisions must be made within the constraints of the PDS. The output is a series of production drawings accompanied by documentation. Again solutions must be synthesized and decisions made in the design of one component will influence the design of others.

The detail design process is illustrated in Fig. 7.1. As with all other stages it is cyclical or iterative in nature following broadly the pattern indicated in the outer ring of the figure.

7.2 Factor of safety

All components carry a load of some kind, be this load electronic, chemical or structural. As a critical part of our design calculations a designer must ensure that all components can sustain the applied load for the working life of the product or process. In simple terms structural failure can occur due to breakage or significant deformation and the analysis process can be seen in terms of the three stage process outlined in Chapter 6 ‘Modelling’.

(1) load type and force analysis;
(2) stress analysis of critical sections;
(3) analysis of possible modes of component failure.

The vast majority of failures are caused by dynamic or cyclic loading leading to fatigue failures. We are only concerned with static loading here, fatigue being covered in the later robust design section of this chapter.

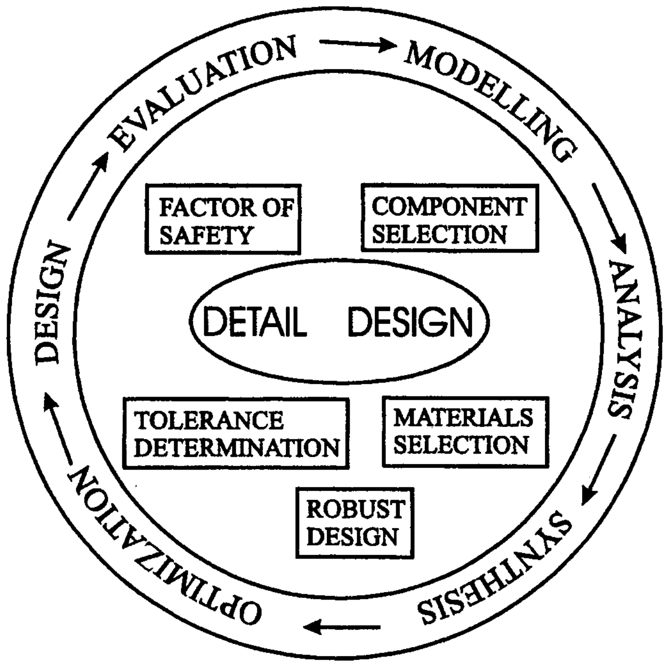


Figure 7.1 Cyclic nature of detail design

In the early stages of detail design the shape and form of a component are often ill defined so, with incomplete knowledge, we introduce a factor of safety to enable the analysis process to commence. Factor of safety is usually based upon the yield strength of the material and in exceptional circumstances upon the ultimate strength of the material.

The general equation for factor of safety is

$$\text{Factor of safety } (N) = \frac{\text{Load carrying capability}}{\text{Applied load}}$$

$$N = \sigma_Y/\sigma_L \qquad \text{or} \qquad N = \sigma_U/\sigma_L$$

Where, σ_L is the stress due to the applied load;
σ_Y is the yield strength of the material;
σ_U is the ultimate strength of the material.

In the first iteration we consider the nominal stress at the critical section to be the maximum applied stress. Subsequently, it is common practice to use the local maximum stress caused by such stress raisers as notches, shoulders, threads, holes, radii and undercuts. Such effects can be quantified using the many charts commonly available.

For ductile materials (< 1500 MN/m^2 tensile strength) typical factors of safety N are shown in Table 7.1.

Table 7.1

	Steady loads	*Occasional shock loads*
Tension and/or bending	3	6
Compression and/or contact pressure	3	6*
Torsion and /or shear	4.5	9

*This represents the crushing limit. You should also check buckling.

Since factors of safety compensate for uncertainties it is the designer's duty to gain further knowledge through investigation and research in order that doubts may be removed and lower factors of safety used.

Statistical assessment of factor of safety

The estimation of how reliable a product will be can be equated with a doctor taking a person's pulse rate. A high or low pulse rate indicates the patient is not well but does nothing to make the patient better. Similarly, estimation techniques will not improve the reliability of a product, merely indicate potential problem areas. Most texts concentrate on the very important area of tests for reliability estimation which are used in quality assurance and for scheduling preventative maintenance. However, the real issue confronting the designer is building in product reliability, in which case approximate reliability estimates will suffice and greater emphasis on reliability improvement techniques is required.

The situation facing design engineers is that most of the quantities upon which calculation procedures are based assume single figure values. However, it is not possible for materials, for example, to be produced economically without variation of properties. Variations also occur due to manufacturing methods such as heat treatment, weld quality, surface finish and dimensional accuracy. All other factors, such as fatigue life, fracture toughness, notch sensitivity, creep, abrasion and corrosion exhibit scatter.

The scatter exhibited does not always follow a normal distribution curve. However, for simplicity, in the later examples all data is assumed to follow the normal distribution. There are many different statistical distributions, some others are illustrated in the reliability section, which may in certain circumstances be more applicable. However, the principle here is that all values exhibit scatter and that quoting and using single figures is misleading. The reader must decide which distribution is most applicable to a particular situation. If a statistical distribution can be identified which fits the situation then a figure can be obtained, giving a more accurate representation of the situation.

For both applied load and load carrying capacity distributions it is normal to consider the area of the normal curve contained by six deviations. Indeed in engineering in general, including quality control, a tolerance specified is assumed to be three times the standard deviation. Taking three deviations a side covers 99.73% of all possibilities. The designer sets the allowable, acceptable chance of failure in the range of well-defined probabilities.

The concept of a factor of safety has been employed in engineering design for many years and although the necessity of its usage is well recognized the basis of its selection is often nebulous. Indeed, in some industries it is referred to as the factor of ignorance!

In calculating the factor of safety as $N = \sigma_Y/\sigma_L$ no account is taken of the shape of either the load distribution or the material strength distribution. Only mean values are used.

Clearly, the aim is to prevent overlap of the two distributions, therefore preventing failure. However, for a machine or structure it is impossible to predict precisely the external loads to which it will be subjected. Hence, in predicting the external load on an element, a tolerance band will accompany the specification of the mean load. Similarly, the load carrying capability of the element is affected by material strength variations and geometrical tolerances, both of which are susceptible to uncontrollable manufacturing flaws which can be significant. Thus the load carrying capability will be accompanied by a tolerance band.

The specific relationship which should exist between the two distribution curves for a satisfactory design will depend on the particular element under consideration and the significance attached to the occurrence of a particular failure mode. There are many options open to the designer once the likely distributions are known. To minimize the chance of failures occurring there are two main options, tighter manufacturing control and material selection reducing the spread of the capacity curve and increasing the separation of the mean values. Both of these will add cost to the design and an optimum or compromise solution must be arrived at.

Where failure would endanger human life, for example, the relationship shown in the figure would be desirable, since there is no realistic overlapping of the curves. However, there are many design situations where the occurrence of a particular failure mode could be tolerated occasionally with no significant consequence. For such cases the relationship between the distribution curves can be depicted as in Fig. 7.2.

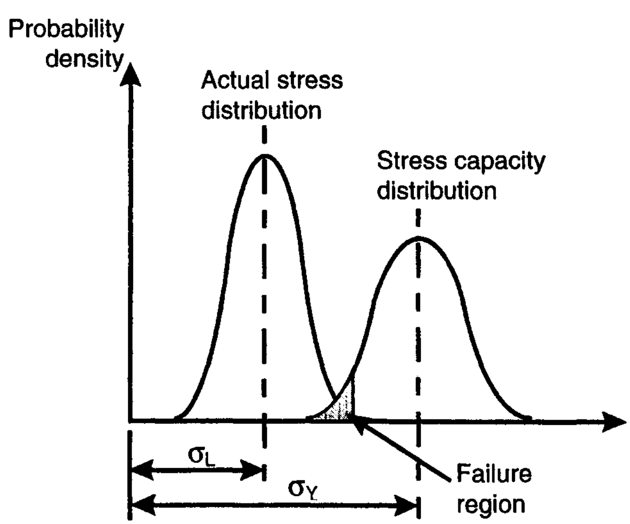


Figure 7.2 Load and strength distributions

For a component the particular failure mode will occur if the load capability based on that phenomenon is less than the corresponding actual load. Hence the distribution curve for the difference between load capability and actual load will be very significant since the probability of failure can be estimated from it. Depicted in Fig. 7.3 is the distribution curve for (σ_Y–σ_L) where all negative values correspond to the occurrence of failure.

For cases where failure cannot be tolerated, the distribution curve (σ_Y–σ_L) should be located so that all realistic values are +ve. When failure can occasionally be tolerated, some values can be –ve and the area under the curve represents the predicted percentage of

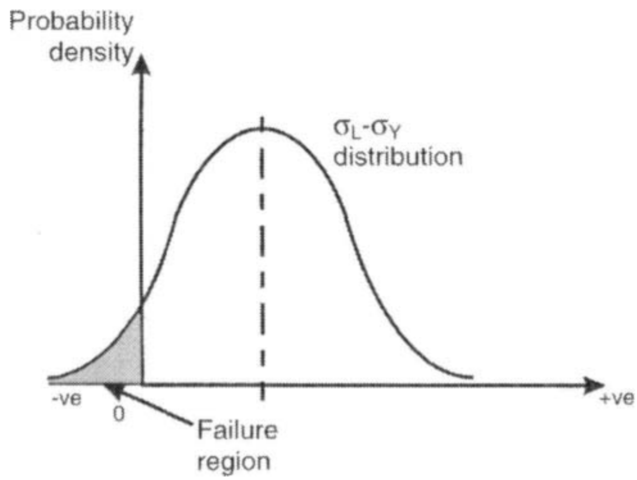


Figure 7.3 Combined curve

failures. For both loading and load capacity curves it is normal to consider six standard deviations, covering 99.73% of all cases. Practical distributions tend to exhibit a much more extended 'tail' than represented by the normal distribution. However, quality control of material or overload protection of a device are examples of methods employed to abruptly cut off the tail, making the use of the normal distribution valid.

In order to take advantage of available statistics tables a new unitless variable t is introduced.

$$t = \text{safety margin}$$

Any value of t can be converted to the proportion of the area under the normal curve between the mean ordinate and the ordinate at any standardized deviate from the mean using the normal table, reproduced in part in Table 7.2. The area shaded in the diagram is the area indicated from the tabular values. In order to obtain the full reliability figure we must add the more positive half of the area under the curve to the shaded area. Since we are only interested in high reliability figures only the higher values of t are quoted.

Since we are only interested in the point on the combined applied and capacity curve beyond which values go negative,

$$t = (\sigma_{\text{YMEAN}} - \sigma_{\text{LMEAN}})/((D_{\text{Y}})^2 + (D_{\text{L}})^2)^{1/2}$$

Consider the example of a component which has a load carrying capacity which is normally distributed, with a mean value of 5000 N and a standard deviation of 400 N. The load it has to withstand is also normally distributed, with a mean value of 3500 N and a standard deviation of 400 N.

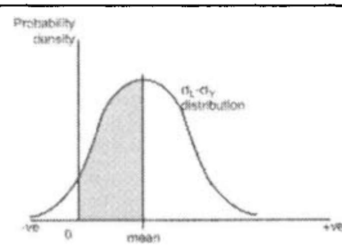
The reliability of the component per load application is estimated by calculating the safety margin.

$$t = (5000 - 3500)/((400)^2 + (400)^2)^{1/2}$$

$$= 2.65$$

Table 7.2

t	**0.00**	**0.01**	**0.02**	**0.03**	**0.04**	**0.05**	**0.06**	**0.07**	**0.08**	**0.09**
2.0										
2.1			*4840*							
2.2										
2.3										
2.4										
2.5	4938	4940	4941	4943	4945	4946	4948	4949	4951	4952
2.6	4953	4955	4956	4957	4959	4960	4961	4962	4963	4964
2.7	4965	4966	4967	4968	4969	4970	4971	4972	4973	4974
2.8	4974	4975	4976	4977	4977	4978	4979	4979	4980	4981
2.9	4981	4982	4982	4983	4984	4984	4985	4985	4986	4986
3.0	4987	4987	4987	4988	4988	4989	4989	4989	4990	4990
3.1	4990	4991	4991	4991	4992	4992	4992	4992	4993	4993
3.2	4993									
3.3	4995									
3.4	49966									
3.5	49977									
3.6	49984									
3.7	49989									
3.8	49993									
3.9	49995									



Referring to the section of normal table a value for *t* of 0.496 is obtained. This must be added to the remainder of the positive area under the curve, since 0.496 only represents the shaded area of the diagram. Thus we have a reliability figure of 0.996 which would normally be expressed as 99.6%.

Extending the example, it is illuminating to examine the effect of larger standard deviations (less control of parameters). Consider standard deviations of 500 N.

$$t = (5000 - 3500)/((500)^2 + (500)^2)^{1/2}$$
$$= 2.12$$

Reliability drops to 98.3%.

Note that reliance on the static factor of safety alone would be misleading since it is calculated on mean values and would be unchanged.

7.3 Selection procedure for bought out components

Many components and units which form part of a product or system can be bought from companies specializing in their manufacture. Generally in the design of any system the successful selection of suitable elements is the result of matching system requirements with the capabilities of one of a wide range of available alternatives. Thus information about the system and information about available hardware is necessary. Figure 7.4 shows a

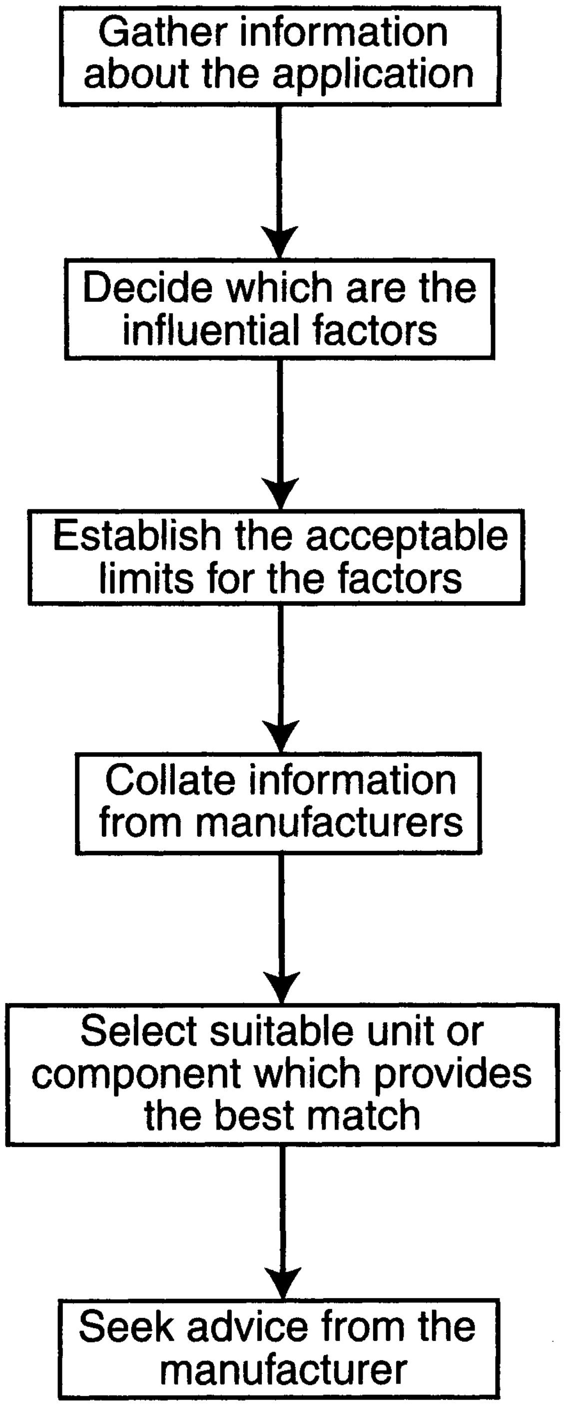


Figure 7.4 Unit or component selection procedure

generalized procedure for selecting the elements of a system. It shows the necessity for thorough information gathering regarding the application before selection can take place.

It is essential at the outset to define the boundaries within which the element must perform as comprehensively as possible. The information gathered relates to the purpose for which the element is required and to the criteria of life, performance, cost and operating environment with which it must comply. The information is needed in order to understand the total system so that the selection is consistent with the rest of the system. The temptation of the design engineer to consider only functionality should be resisted and time and effort expended at this stage is invariably rewarded when an optimum design is selected.

Factors which influence the selection of units and components must now be identified. The most important and common factors governing selection are often performance, application, geometry, environment, safety and commercial. Not all these factors are important in every design so careful study of the system is required to ensure that those considered are actually relevant. Reference should be made to the Product Design Specification for the system.

Each factor should be defined in terms which are as objective as possible. Thus, where appropriate and possible, numerical information should be given, terms must be explained and vagueness avoided. Following this the boundaries of satisfaction must be defined for each of the chosen factors. This assists the design engineer in making a selection which meets the stated requirements in every respect. Subjective judgements cannot always be avoided and in such cases a means of comparison must be established.

Manufacturers' data should be collated and arranged in a suitable format. There is a finite number of particular types of component available from manufacturers and the selection process is heavily constrained by the form and content of the information presented by them and the range of catalogues available to the design engineer when the need arises. There is a good case for maintaining a 'rolling' catalogue library, or data on microfilm/computer, since gathering such information can be very time consuming, particularly if a unique set of data is collected on each occasion. Data on size, cost and performance can often be noted in numerical form, giving a range where appropriate. In the case of less objective data a rating may be shown based on advice or opinion gathered.

Optimizing the choice is now a matter of selecting the best compromise, in the opinion of the design team, between the priorities of the system and the availability of hardware. In the initial stages some pruning of potential alternatives is called for. As far as the factors involving numerical data are concerned, some yield a go/no-go situation which will eliminate those which do not fit within the boundaries set.

Other requirements of a more subjective nature should be compared on the basis of the elements' ability to meet the criteria as laid down in the Product Design Specification. The evaluation technique used here is similar to that used elsewhere in the design activity, particularly for initial concept selection. References for further reading are included at the end of each subsequent chapter and many elaborate on the details of a variety of techniques.

Further advice on the detail of installation or specifying and ordering will be required from manufacturers' information. Normally this would be available in a catalogue but often it is necessary to communicate directly with a representative of the company.

As an example of a more detailed component selection procedure consider the decisions which influence the type of spring most suitable for a particular application. The flow chart in Fig. 7.5 illustrates this. The selection of a type of spring depends mainly on the space available and the magnitude and direction of the loading. One further important

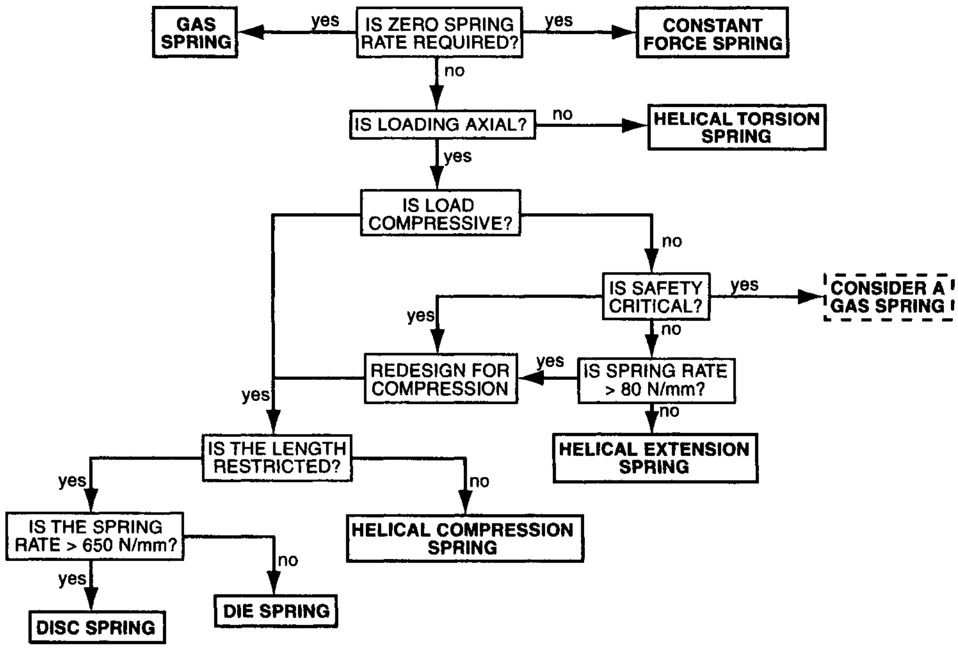


Figure 7.5 Selection chart for spring type

consideration which influences the selection of a particular type of spring is that for some applications safety codes require that a compression spring is used. This is because a failed compression spring can continue to provide a stop and hold components apart, in effect providing a fail safe design.

The flow chart in Fig. 7.5 illustrates the level of thought which must be applied to the smallest detail in any design. It only takes the failure of a relatively insignificant component to render the whole design worthless.

7.4 Robust design

Design for reliability

The reason for the steady increase in reliability engineering stems from the increasing awareness that the cost of ownership of a product or system comprises two components. The first is the capital cost and the second is the cost of operating, administering, maintaining and replacing the product or system. The second outlay, the running cost, can often exceed the capital cost and is a function of reliability. Indeed, because of the disastrous financial consequences of equipment failure most customers specify tightly reliability conditions.

One hundred per cent reliability testing is unthinkable since this implies that there would be no products for sale. The time required for reliability testing depends on the failure rate of the items under test. In general reliability adds cost to a product and although unreliability carries with it a cost penalty, the optimum level of reliability is always a

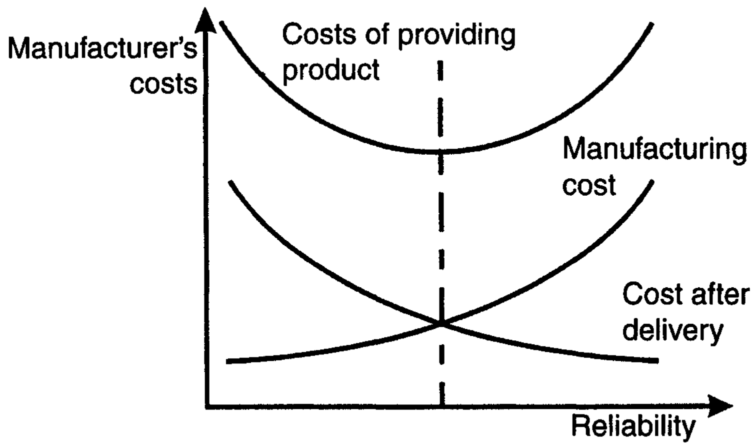


Figure 7.6 Cost of reliability

compromise between the two. Figure 7.6 shows the general relationship between reliability and cost.

Reliability is concerned with the causes, distribution and prediction of failure. Failure is defined as the termination of the ability of a component or system to perform its required function. The parameter, 'failure rate' is given the symbol $\lambda(t)$. Another method of describing the occurrence of failures is to state the mean time between successive failures. The two terms used, the mean time between failures (MTBF) and the mean time to fail (MTTF) are explained diagrammatically in Fig. 7.7. In many, but not all, cases MTTF and MTBF are the same. MTTF is the mean operating time between successive failures and the difference between the two terms is repair time. Hence,

$$\text{MTTF} + \text{mean time to repair} = \text{MTBF}$$

Components or systems which are not repaired do not extend beyond the point marked ▲ in which case MTTF and MTBF are the same.

The failure rate is not necessarily constant. If a reliability test were undertaken with a large sample and each product were tested until it failed and not replaced, the typical

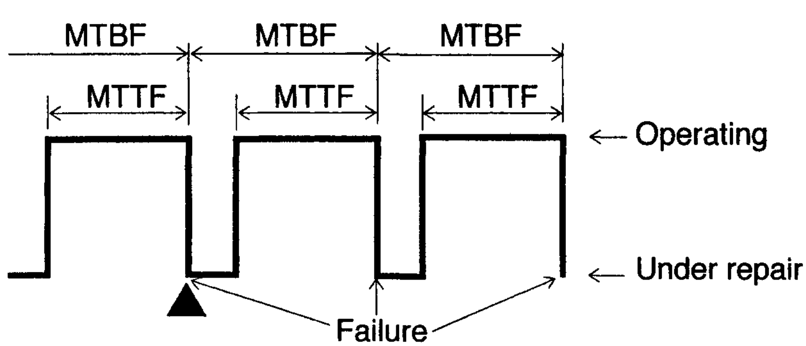


Figure 7.7 Difference between MTBF and MTFF

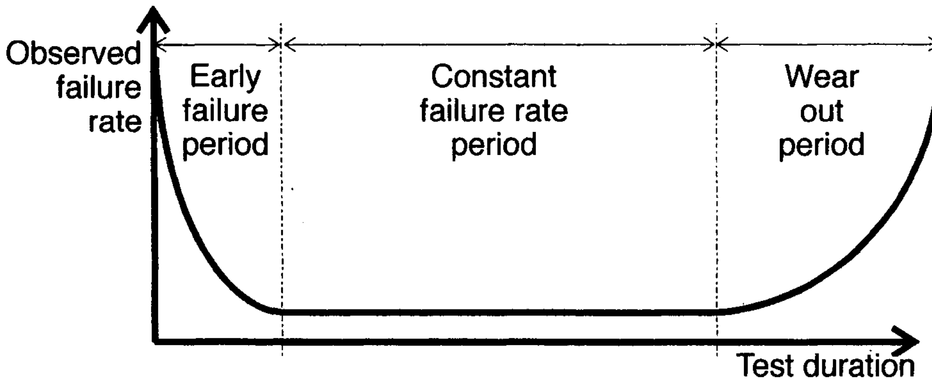


Figure 7.8 Failure rate against time curve

failure rate against time curve would be what is known as the 'bath-tub' curve shown in Fig. 7.8.

During the *early failure period*, within hours of commencement of operation it is quite likely that failures will occur due to various imperfections acquired during the manufacturing process, due to design faults or misuse. Gradually, these early failures will occur less frequently. This period is often covered by the manufacturer's guarantee.

The *constant failure rate period* is usually relatively long and the failure rate is normally approximately constant. During this period it is usual for failures to occur at a relatively low rate but from a wide variety of causes.

The beginning of the *wear out period* corresponds to the end of the useful working life. All products wear out and this occurs because of a variety of time-dependent mechanisms.

Reliability $R(t)$ is time related, and is defined as a probability and expressed as a value between 0 and 100%. Consider a number of components N_o being tested and allowed to fail without replacement so that at any time t there are N_s surviving. Then

$$R(t) = N_s(t)/N_o$$

If now the failure rate is expressed in terms of $N_s(t)$ the relationship between failure rate and $R(t)$ is

$$\lambda(t) = \mathrm{d}N_s(t)/N_s(t)\,\mathrm{d}t$$

Solving these two equations when the failure rate is constant gives

$$R(t) = \mathrm{e}^{-\lambda t}$$

For this condition, the MTBF (θ) is the reciprocal of the failure rate, hence

$$R(t) = \mathrm{e}^{-t/\theta}$$

As an example, records kept on 1000 engines of the same type show an average life to failure of 14 000 flying hours. What is the probability that one such engine will survive a transatlantic flight of 7 hours?

$$\text{MTBF} = 1/\lambda = 14\,000 \text{ hours}$$

$$\lambda = 1/14\,000$$

$$R(t) = e^{-\lambda t}$$

$$\begin{aligned} R(7) &= \text{Probability of surviving 7 or more hours} \\ &= e^{-7/14\,000} \\ &= 0.9995 \text{ or } \underline{99.95\%} \end{aligned}$$

This indicates that engine survival is relatively assured. The situation is obviously more complex since a modern passenger aircraft will always have more than one engine and could continue flying without one engine. Hence, 99.95% is only a measure of the reliability of the engine not the aircraft. Since this is only an introduction to design for reliability, the use of distributions, such as the binomial, which would indicate overall aircraft safety based on system reliability, is not covered.

Whilst it is failure rate, MTBF, and hence reliability of components that is measured, it is the reliability of complete systems that is the ultimate concern of the designer, salesman and customer. The reliability of a system can be obtained from the knowledge of the reliabilities of its components. Whatever the system, failure of one component may cause the whole system to fail. For example, a domestic television may have 500 components, and manned spacecraft several million. Thus, the problem facing designers is not so much that the parts are unreliable but that there are so many of them. Many types of system–component relationships exist. Series and parallel are two such relationships.

Consider a system comprising two components connected in series such that failure of either causes system failure. The reliability of the system is the product of the component reliabilities. If the components each have a reliability of 90% then

$$\text{Reliability of the system} = 0.9 \times 0.9 = 0.81 \text{ or } 81\%$$

Expanding to the case of N components connected in series;

$$R_{\text{an}} = R_a.R_b \ldots R_n$$

R_{an} = reliability of the system

If a series system comprises 100 components each with a reliability of 90% then

$$R_{100} = (0.9)^{100} = 0.000\,026$$

A reliability of this order implies that there is no hope of the system working satisfactorily over the required lifetime. If the individual reliability is increased to 0.999 then the system reliability becomes 0.906 or 90.6%. These examples are artificial since:

- It is unlikely to be essential for every component to function correctly in order to ensure system success.
- Critical components can be duplicated so that if failure occurs there is a spare to take over – this is called redundancy.

If the failure rates λ_a and λ_b apply to the two-component system such that

$$R_a = e^{-\lambda a t} \qquad \text{and} \qquad R_b = e^{-\lambda b t}$$

Applying the series rule gives

$$R_{ab} = e^{-(\lambda a + \lambda b)t}$$

This shows that the system is a constant failure-rate system of failure rate ($\lambda a + \lambda b$).

Consider now another type of system, again comprising two components but so connected that if one fails the system does not fail. This is a parallel system because the failure of a single component does not cause system failure. The probability of system success in this case is now the probability that either or both units succeed, hence

$$R_{ab} = R_a + R_b - R_a R_b$$

In general

$$R_{an} = 1 - (1 - R_a)(1 - R_b) \ldots (1 - R_n)$$

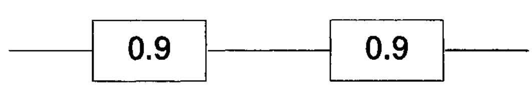
In order to achieve the degree of reliability required it may be necessary to duplicate components so that if one component fails there is another available to carry on working. The following are examples of this technique which is called *redundancy*:

- Altimeter in aircraft. One is insufficient in case of malfunction. Two would pose the problem of identifying which is giving the correct reading if they differed. Three are thus required. This is called active redundancy.
- In the operating theatres of hospitals. If the mains electricity fails provision is made for switching to an emergency supply. This is standby redundancy.
- The spokes of a traditional bicycle wheel illustrate another kind of redundancy. Several can break and the wheel will still function. This is called partial redundancy.

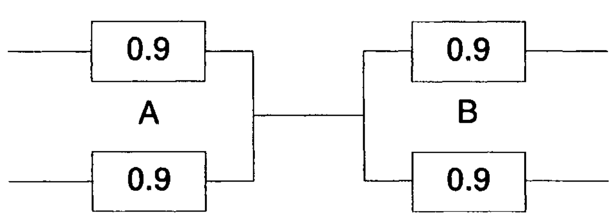
Consider the example of a gearbox containing six rolling element bearings. The L_{10} life of a bearing, which is quoted by a manufacturer, indicates the number of cycles after which 10% of the bearings would fail. That is, each bearing has a reliability of 0.9. The gearbox has six bearings and all must function for the gearbox to operate.

$$\text{The gearbox reliability} = (0.9)^6 = 0.54 \text{ or } 54\%$$

Considering a shaft requiring a bearing at each end. The reliability of the two bearings is 0.81 or 81%.



If this is unacceptable, two bearings could be used at each end of the shaft. The model now is as shown and the reliability is increased.



$$R_A = R_B = 1 - (1 - 0.9)^2$$

$$R_{shaft} = 0.99^2 = 0.98 \text{ or } 98\%$$

This gives an increase of 17% in reliability.

FMEA

Failure Modes and Effects Analysis (FMEA) was first used in the 1960s by the aerospace industry and is now a technique used by most industrial sectors. It is an objective method for evaluating system design. This is accomplished by establishing a multi-disciplinary team to consider all the potential failure modes of the components which make up a system or product and quantifying the influence such failures would have on overall reliability. It is one of the most powerful tools available for identifying reliability, safety, compliance, and product non-conformities during the design stages.

FMEA directs attention to those areas of a detail product design which may cause non-satisfaction of the reliability or safety criteria of the specification. Once critical components are identified then corrective action can be taken to improve the design. The technique can be used, for example, to identify critical parts within a system or product which benefit from the introduction of parallel or redundant components.

FMEA is an ongoing process that should start as a part of the first design review and continue throughout the life of the product. FMEA is a bottom up analysis technique.

The effect of a component failure depends upon the function of the component in the system. The severity of a potential failure is represented by the variable S and is assigned a value between 1 & 10, where 10 is the most severe. The occurrence of the failure (Relative Failure Rate) is represented by the variable O and is assigned a value between 1 & 10, where 10 is the highest failure rate. The ability to detect a failure is represented by the variable D which is assigned a value between 1 & 10 with 10 being the most difficult to detect. The relative importance of a failure mode is represented by its Risk Priority Number (RPN) calculated as

$$\text{RPN} = S \times O \times D$$

Every component has numerous potential failure modes and theoretically there is no limit to the depth one could go. Practically, there is a point of diminishing returns where the added cost exceeds the benefits derived. In practice a component with a RPN in excess of 100 is considered to be definitely worthy of attention. The FMEA process develops several very useful databases that provide manufacturers with the basic tools necessary to control the quality of their product.

Method

(1) A cross-functional team must be used to develop the FMEA.
(2) Identify the function of the component.
(3) List at least one potential failure mode for each function.
(4) Define the effects of failure in terms of what the customer might notice.

(5) Rate the severity (or seriousness) of the potential effect of the failure.
(6) Assign an occurrence rank to each of the potential causes/mechanisms of failure.
(7) Assign a detection ranking that assesses the ability of the design controls to detect a potential cause/mechanism or the ability of the design controls to detect the subsequent failure mode.
(8) Calculate the RPN numbers for each component.
(9) Identify and carry out remedial actions for potential significant and critical characteristics of components to lower the risk of the higher RPN failure modes.
(10) Calculate the new severity, occurrence, detection and RPN numbers.

The reasons for the FMEA being carried out by a multi-disciplinary team are explained by the subjective nature of arriving at the occurrence, severity and detectability scores. The following tables give general guidance for this process and the ratings quoted are those used in the automotive industry.

Table 7.3 Occurrence

Rating	*Failure Rates*	*Probability of Failure*
10	<1 in 2	Very High
8	1 in 8	High
6	1 in 80	Moderate
5	1 in 400	Occasional failures
3	1 in 15 000	Low
1	1 in 1 500 000	Remote: Failure is unlikely

Table 7.4 Severity

Rating	*Effect*	*Severity of Effect*
10	Hazardous – without warning	Potential failure mode affects safety or involves non-compliance with government regulation without warning.
9	Hazardous – with warning	Potential failure mode affects safety and/or involves non-compliance with government regulation with warning.
8	Very high	Item inoperable, with loss of primary function.
7	High	Item operable, but at reduced level of performance. Customer dissatisfied.
6	Moderate	Item operable, but customer experiences discomfort.
5	Low	Item operable, but at reduced level of performance. Customer experiences some dissatisfaction.
4	Very low	Poor fit or finish. Defect noticed by average customers.
3	minor	Poor fit or finish. Defect noticed by most customers.
2	Very minor	Poor fit or finish. Defect noticed by discerning customers.
1	None	No effect.

Table 7.5 Detectability

Rating	*Detection*	*Criteria*
10	Absolute uncertainty	Design Control will not detect a potential cause and subsequent failure mode.
9	Very remote	Very remote chance the Design Control will detect a potential cause and subsequent failure mode.
8	Remote	Remote chance Design Control will detect a potential cause and subsequent failure mode.
7	Very low	Very low chance Design Control will detect a potential cause and subsequent failure mode.
6	Low	Low chance Design Control will detect a potential cause and subsequent failure mode.
5	Moderate	Moderate chance Design Control will detect a potential cause and subsequent failure mode.
4	Moderately high	Moderately high chance Design Control will detect a potential cause and subsequent failure mode.
3	High	High chance Design Control will detect a potential cause and subsequent failure mode.
2	Very high	Very high chance Design Control will detect a potential cause and subsequent failure mode.
1	Almost certain	Design Controls will almost certainly detect a potential cause and subsequent failure mode.

Note: Zero (0) rankings for Severity, Occurrence or Detection are not allowed.

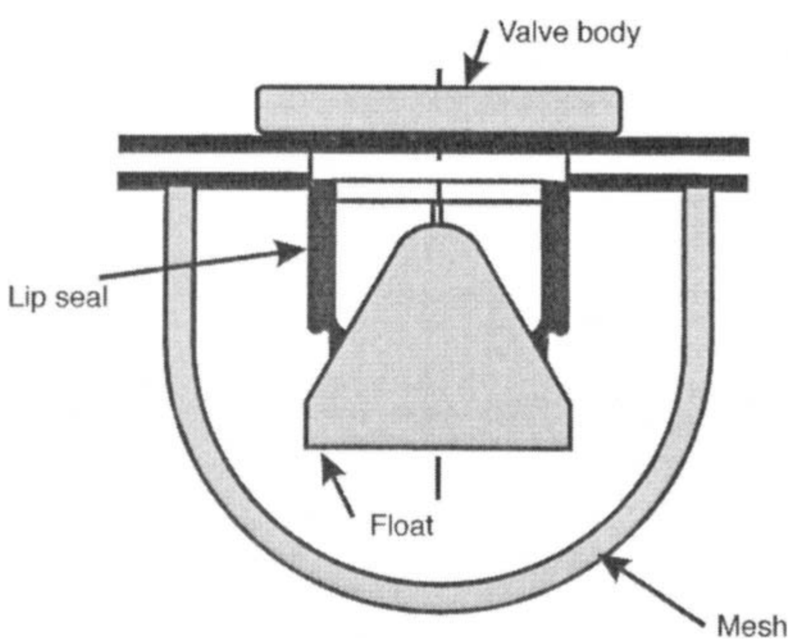


Figure 7.9 Pressure relief valve

Consider as an example the redesign of the pressure relief valve for a toilet cassette in a caravan. The redesign was required due to problems associated with 'blow-back' when the valve became blocked and therefore did not release pressure unless the toilet were operated. After consideration of various alternatives the concept shown in Fig. 7.9 was investigated using FMEA. The most important features of the new design aimed at increasing functionality are the rounded top of the float, preventing debris from settling, use of low friction coefficient material for the float and the mesh guard.

Table 7.6

Product function	*Potential failure mode*	*Potential effects of failure*	*Potential causes of failure*	*Current controls*	*O*	*S*	*D*	*RPN*
Equalize pressure between cassette and ambient	Valve fails to work	Cannot equalize pressure leakage	Mesh comes loose	Cleaning	4	4	3	48
			Valve blocks		5	5	2	50
			Float snaps off		2	5	2	20
			Seals perish		2	4	4	32
			Spring fails		1	5	2	10
			Plunger jams		1	5	3	15
			Valve works loose		1	4	2	8
			Chemical degradation		5	2	8	80
			Abuse in handling		6	4	1	24

Since there are no RPNs over 100 in Table 7.6 it is clear that there is no serious need for attention to any one element of the design. The two which are highlighted are the potential for chemical breakdown of the valve materials and possible clogging of the valve with solid waste. During detail design materials selection and surface finish were considered more carefully than might have been the case without carrying out an FMEA on the new design.

Design for quality

In the increasingly global competitive world in which companies must operate, persistent quality improvement, often allied with assembly and manufacture cost reduction, is essential if a company is to remain profitable. The quality of a product is a measure of the degree to which it meets the customer's requirements. Reliability is defined as the probability that a device or system will operate without failure for a given period of time. The difference between the two is the time element, reliability being concerned with how long quality exists.

A simple example which illustrates the difference between quality and reliability is a car tyre. The performance in cornering and braking deteriorates as the tyre wears. In other words, the quality deteriorates until it reaches an unacceptably low level and the tyre is replaced. By contrast, reliability is more a measure of the frequency with which tyre bursts occur.

The old concepts of improving product quality by controlling manufacturing processes and identifying poor quality items by inspection have given way to improving product and process design at the design stage. The main reasons for this dramatic switch of emphasis are illustrated in Fig. 4.2, which shows that the vast majority of engineering changes occur close to, and sometimes after, the product is released for production. The later in the design and development process these changes occur the greater the cost penalty. This is particularly true once manufacture has commenced and the effect is magnified by delaying the launch of the product. The aim is to design quality in by getting the product right first time. Expressed in terms of Fig. 4.2 the aim is to move the peak further left, as far before the release date as possible.

The adoption of the quality techniques which follow will result in shorter lead times, a reduced number of engineering changes, reduced costs and increased quality. Quality has no meaning unless it is related to cost, since, in general, the more expensive a product is the better its quality. The aim is to improve the quality of a product without increasing the cost of producing that product. This can be achieved by reducing the effects of variability of controlling parameters and the technique is called parameter design.

Within parameter design two different categories of factors are identified. The first category are control factors, which as the name suggests can be easily controlled. Examples of such factors include material selection, operating voltages and sizes. The second category are noise factors. These cannot be controlled easily and are often very costly when control is attempted. Examples include temperature and humidity. In an attempt at quantifying the effect of different parameters the signal to noise ratio has been introduced. This is so named because of the comparison in communications of the strength of a transmitted signal with the level of interference. There are three signal to noise ratios in general use: nominal, as in the case of colour density of a television; minimization, as for weight and noise; and maximization of such as strength or power.

The approach is to identify those control factors which are insensitive to noise factors and particularly those which exhibit non-linearity. The aim is to reduce the sensitivity of products to the source of variation. As an example consider the design of an electrical power supply circuit, the characteristics of which are illustrated in Fig. 7.10. Following conceptual (system) design the components making up the circuit are selected. There are many combinations of resistance, capacitance and transistor characteristics which will provide a functioning circuit.

If the desired output is 240 volts then parameter design seeks to reduce the variability about this mean for all circuits produced, without significant increase in associated costs. The output voltage can be determined by the gain on the transistor, which is non-linear. In order to obtain an output voltage of 240 V the designer selects gain A^1. If the actual gain deviation is as represented by the area under the normal curve then the voltage will also deviate about the target of 240 V. An alternative course of action is to use the effect of the non-linearity of the gain curve. If the flatter part of the curve were used, the target voltage would increase to around 270 V but the deviation about this value would be reduced.

Such adjustment of target value, and the introduction of an extra component, in this case a resistor to reduce 270 V to 240 V, must be considered carefully. It can often be a simpler and more cost effective way of controlling variance. Obviously the variance of the resistor must also be taken into account.

It is common practice to specify constraints in the PDS in terms of the limits of tolerance. If parameter design is to make its full impact then this must change and the constraints in the PDS must be specified in terms of target or ideal values with

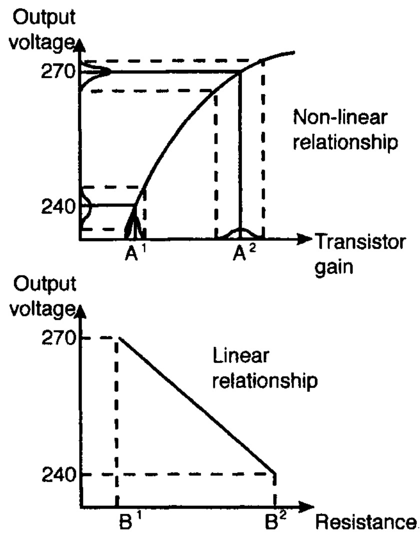


Figure 7.10 Transistor characteristics

accompanying limits of tolerance. Consider for example the fitting of a cab door. If the door is made to the smallest acceptable size and the frame to the largest acceptable size then the door will still operate satisfactorily but it may leak under certain conditions and allow more noise into the cab than a door and frame made to target. Target sizes guarantee optimum quality and should be the aim for the maximum amount of production possible.

Although it is difficult to quantify the quality of a product one measure is becoming accepted. This is called Loss to Society. The smaller the loss the greater the quality. A product causes losses when it deviates from target values, even when within the specified tolerance limits. An often quoted example, first published in a Japanese newspaper, involves two manufacturing plants making identical television sets for Sony, one in the USA and the other in Japan. The key characteristic involved was colour density, the target values for which were set after extensive customer trials.

The design philosophy in the Japanese plant was to aim for as many televisions as possible to be near to the target value for colour density and not to be too worried about a small number which were outside the acceptable limits. By contrast, the design philosophy of the American factory was to maximize the number within the acceptable limits. The results of the two philosophies are illustrated in Fig. 7.11. The Japanese philosophy resulted in a normal distribution whereas the American philosophy resulted in the flat profile. All of the televisions produced in the USA were within the specified limits of acceptability whilst a small percentage of the Japanese sets were outside these limits.

It is clear from the distributions in Fig. 7.11 that more of the Japanese sets were on and around the target value. Customer perception was that the Japanese sets were of much higher quality than the American sets. The conclusion that can be drawn is that quality is best achieved by minimizing variance rather than by strict adherence to the specification. The loss to society in this case is represented by a quadratic function with the loss to society increasing as the square of the deviation from target.

Of foremost importance is the definition of *'manufacturing' quality* as *product uniformity around a target* rather than conformance to specification limits.

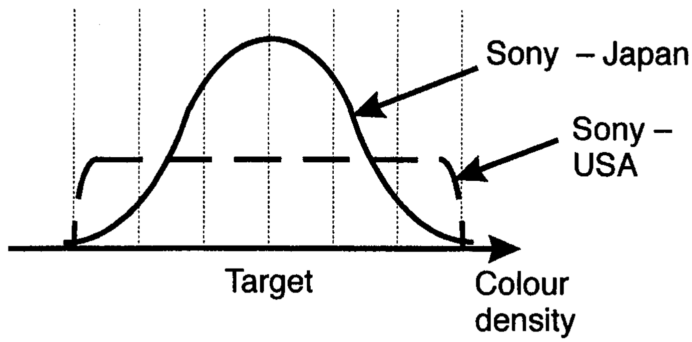


Figure 7.11 Colour density philosophies

Measures to counteract and reduce product variability can only be taken during design. Techniques used later, during manufacturing or process design can reduce variance introduced due to, for example, manufacturing tolerances and material imperfections. However, even these can be most effectively tackled during design.

Design against high cycle fatigue

Only rarely does the failure of a component or structure occur due to the application of a single static heavy load. Failure is more often caused by lighter repeated or cyclic loading and usually occurs at a point of stress concentration where the shape changes abruptly. The mechanisms by which cracks are propagated through the material are generally of more interest to the materials scientist than the designer. What is important for the designer is to recognize the features which influence fatigue life and to modify the design accordingly. Some very simple rules can be applied which will extend the life of a product greatly.

- Reduce stress raising effects.
- Provide best possible surface finish.
- Compress the surface.
- Stress relieve.
- Select materials which resist fatigue.

High cycle fatigue, above 1000 cycles during expected design life, is by far the most common cause of component failure in service accounting for an estimated 80% of all fractures. In the vast majority of cases the cause of the failure can be traced back to a lack of detail design consideration, particularly in component or joint shape. This is because most failures occur at changes in cross-section of a component or at the joints between components. Where this change in shape is accompanied by the welding of two components together the potential for disaster is significantly increased. This is due in the main to residual tensile stresses caused by the cooling and contracting of the weld pool. There are many factors which affect fatigue life.

Fortunately there are some general guiding principles and analogies which, if followed, will result in improved fatigue life. The main analogy used for predicting highly stressed regions of components and joints and therefore likely points at which fatigue failure will

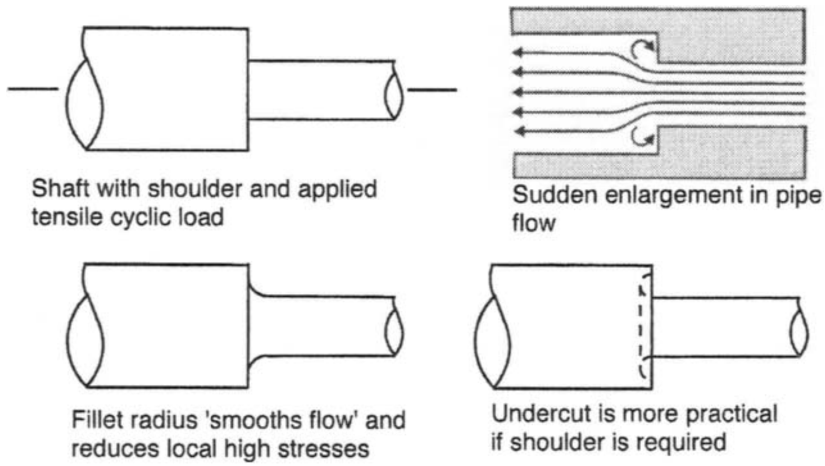


Figure 7.12 Stress and fluid flow analogy

occur is to compare the structure with fluid flow through pipes or channels. Consider the simple comparison, shown in Fig. 7.12, of a piece of round bar with a shoulder loaded in tension with pipe flow through a sudden expansion. The load is the same along the length of the bar, so where the cross-sectional area is large the stress will be low and the stress will be relatively large when the area is small. In the pipe flow analogy the amount of water flowing in each section is constant so the flow rate must increase in smaller sections. Two other features of pipe flow are helpful. Where eddy currents exist this indicates a region of very low stress and where the flow lines are locally closer there will be a local increase in stress.

Thus stress and velocity are related in the simple case of a steady flow. Perhaps surprisingly, the analogy also holds true where sudden changes in shape or flow restrictions cause turbulence. At the shoulder in the round bar we get a stress concentration whilst in the pipe the flow at the sharp corner is much faster. In the pipe flow there is also an area of slack water where the main flow cannot follow the walls into the corner. In the same position of the bar there will be a 'stress shadow'. This is a region of relatively low stress which could be used to advantage.

If the sharp corners were blended or radii then the local disturbance to the flow would be reduced and the highest levels of stress reduced significantly. In the pipe flow this would allow the streamlines to separate or come closer together gradually. The design of the shoulder on the shaft can be improved to 'smooth the flow' and reduce localized stress concentrations by the addition of a corner radius. However, if a shoulder is required for positioning of a bearing axially on the shaft then the undercut shown in the fourth sketch is probably the best option. Notice that the fatigue life is being improved by the removal of material.

Using the pipe flow analogy it becomes relatively easy to predict areas of high stress and therefore areas at high risk of fatigue failure. The technique, illustrated in the diagrams of Fig. 7.13 is to draw out a cross-section and then to rub out all internal outlines, leaving the shape of the metal. You now imagine that the section has water flowing through it along the direction of the applied load. In the first of the diagrams a structure to be loaded in tension has four fillet welds. In the second diagram the weld material is assumed to be

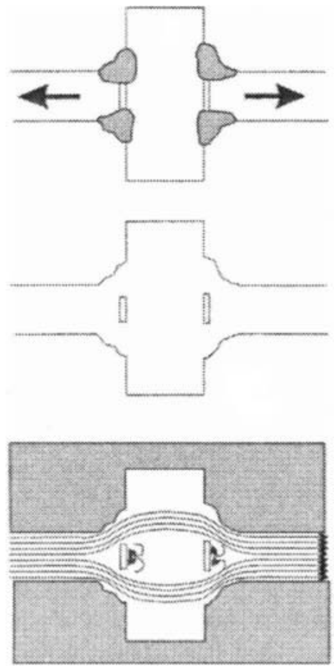


Figure 7.13 Determining high stress regions

homogeneous with the original metal. As illustrated in the final diagram the pipe flow analogy shows two potential areas of high stress. These are caused by the lack of penetration of the weld causing a blockage in the centre of the flow and the corners of the weld being sharp at the point of expansion and contraction of the flow.

It is already established that a sudden notch or change in section will cause local stresses to be significantly increased. Extending the flow analogy a little further, it is at first sight surprising that removal of material to give a smooth profile can also extend fatigue life. Figure 7.14 shows a few examples of poor weld shapes which could be filled out by extra weld or could be dealt with more effectively by grinding away the sudden change in section.

Many manufacturing processes leave residual or built in stresses in the component or structure. Most common of these are tensile stresses due to differential cooling following welding. These will relax with time but there are various techniques for speeding the process. Shot blasting, sand blasting or any similar process which compresses the surface reduces significantly any built in tensile stresses. In stress relieving the structure is heated and then allowed to relax.

Imagine a small crack in a component of width 4 μm. The fatigue damage from each cycle will be proportional to crack movement, so a tensile load which opens the crack to 8 μm before allowing the component to relax does 4 μm worth of damage. In the same way, a compressive load closing the crack from 4 μm will do 4 μm worth of damage. A tension followed by a compression would do 8 μm worth of damage. Heating the structure to around 650°C followed by a slow cooling allows local high stresses to even out and cracks to reduce in width as a result. Following stress relieving, tensile loads still have the same

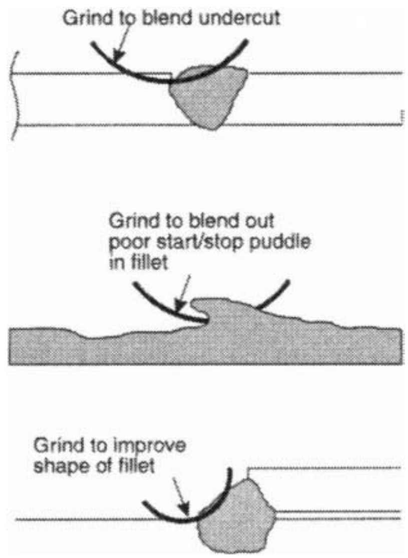


Figure 7.14 Removing material improves strength

damaging effect. However, since the crack is now closed up the damage caused by compression is reduced by approximately 75%.

There is a great deal of very advanced science in the most refined areas of fatigue life prediction, but a good deal of common sense and the application of these basic principles should ensure that many potential failures are avoided. Materials which exhibit a greater resistance to fatigue should be selected for critical areas of a design. For example, ceramic materials are not thought to be susceptible to fatigue at all. Although general awareness of the phenomenon of fatigue began with the Comet aircraft disasters some years ago, the underlying principles were appreciated more than a century ago by the British engineer Sir William Fairbairn, who carried out classic experiments on wrought-iron girders. He found that a girder which was statically loaded would support 12 tonf for an indefinite period but would fail if a load of only 3 tonf were raised and lowered on it more than a million times.

The quantification of the number of cycles that a product will last for, fatigue life, is based on testing data. There is a great wealth of this data which it is beyond the scope of this text to reproduce. However, the factors which must be taken into account in design are:

- mean stress
- alternating stress
- material ultimate tensile strength
- the type of loading: bending, axial or torsion
- the surface finish (the rougher the surface the lower the number of cycles the component will complete)
- the effect of any stress raiser
- the size of the component (since only relatively small components are normally tested to destruction).

7.5 Principles

Detail design principles

Optimization The search is for the best compromise between conflicting criteria.

Simplification Where possible rely on the expertise and knowledge of specialist manufacturers by using proprietary components.

Analysis Ensure that all components have appropriate factors of safety and are not over designed.

Robustness The designer should aim for a product which is fit for the purpose intended for the lifetime intended.

Synthesis A solution is often arrived at by a combination of techniques and elements.

Iteration Progress towards the production stage is made iteratively as knowledge of the important factors grows.

7.6 Exercises

1. Bolts installed on a production line are tightened with automatic wrenches. They are to be tightened sufficiently to yield the full cross-section in order to produce the highest possible initial tension. The limiting condition is twisting off the bolt head during assembly. The bolts have a mean twisting off torque of 20 N m with a standard deviation of 1 N m. The automatic wrenches have a standard deviation of 1.5 N m. What mean value of torque wrench setting would result in only 1 in 500 twisting off during assembly?

2. A shaft tolerance has a standard deviation of 0.01 mm. The hole tolerance has a standard deviation of 0.016 mm. The difference between the means is 0.045 mm. If the entire production is accepted for assembly, determine the proportion of assemblies with clearance less than the allowance of 0.01 mm and the proportion of assemblies expected to interfere.

3. A unit has a constant failure rate of 0.3% per 1000 hours. What is its MTBF? What are the probabilities of the unit successfully completing missions of 10 000, 100 000 and 1 000 000 hours?