CHAPTER 10

CONTACT STRESS, RESIDUAL STRESS AND STRESS CONCENTRATIONS

Summary

The maximum pressure p_0 or compressive stress σ_c at the centre of contact between two curved surfaces is:

$$p_0 = -\sigma_c = \frac{3P}{2\pi ab}$$

where a and b are the major and minor axes of the Hertzian contact ellipse and P is the total load.

For *contacting parallel cylinders* of length L and radii R_1 and R_2,

maximum compressive stress, $\sigma_c = -0.591\sqrt{\dfrac{P\left(\dfrac{1}{R_1}+\dfrac{1}{R_2}\right)}{L\Delta}} = -p_0$

with $\Delta = \dfrac{1}{E_1}[1-\nu_1^2] + \dfrac{1}{E_2}[1-\nu_2^2]$

and the maximum shear stress, $\tau_{max} = 0.295 p_0$ at a depth of $0.786b$ beneath the surface, with:

contact width, $b = 1.076\sqrt{\dfrac{P\Delta}{L\left(\dfrac{1}{R_1}+\dfrac{1}{R_2}\right)}}$

For *contacting spheres* of radii R_1 and R_2

maximum compressive stress, $\sigma_c = -0.62\sqrt[3]{\dfrac{P}{\Delta^2}\left[\dfrac{1}{R_1}+\dfrac{1}{R_2}\right]^2} = -p_0$

maximum shear stress, $\tau_{max} = 0.31 p_0$ at a depth of $0.5b$ beneath the surface with:

contact width (circular) $b = 0.88\sqrt[3]{\dfrac{P\Delta}{\left(\dfrac{1}{R_1}+\dfrac{1}{R_2}\right)}}$

For a *sphere on a flat surface* of the same material

maximum compressive stress, $\sigma_c = -0.62\sqrt[3]{\dfrac{PE^2}{4R^2}}$

For a *sphere in a spherical seat* of the same material

$$\text{maximum compressive stress, } \sigma_c = -0.62\sqrt[3]{PE^2\left[\frac{R_2 - R_1}{R_1R_2}\right]^2}$$

For *spur gears*

$$\text{maximum contact stress, } \sigma_c = -0.475\sqrt{K}$$

with
$$K = \frac{W}{F_w d}\left[\frac{m+1}{m}\right]$$

with W = tangential driving load; F_w = face width; d = pinion pitch diameter; m = ratio of gear teeth to pinion teeth.

For *helical gears*

$$\text{maximum contact stress, } \sigma_c = -C\sqrt{\frac{K}{m_p}}$$

where m_p is the profile contact ratio and C a constant, both given in Table 10.2.

$$\text{Elastic stress concentration factor } K_t = \frac{\text{maximum stress, } \sigma_{\max}}{\text{nominal stress, } \sigma_{\text{nom}}}$$

$$\text{Fatigue stress concentration factor } K_f = \frac{S_n \text{ for the unnotched material}}{S_n \text{ for notched material}}$$

with S_n the endurance limit for n cycles of load.

$$\text{Notch sensitivity factor } q = \frac{K_f - 1}{K_t - 1}$$

or, in terms of a significant linear dimension (e.g. fillet radius) R and a material constant a

$$q = \frac{1}{(1 + a/R)}$$

$$\text{Strain concentration factor } K_\varepsilon = \frac{\text{max. strain at notch}}{\text{nominal strain at notch}}$$

Stress concentration factor K_p in presence of plastic flow is related to K_ε by Neuber's rule

$$K_p K_\varepsilon = K_t^2$$

10.1. Contact Stresses

Introduction

The design of components subjected to contact, i.e. local compressive stress, is extremely important in such engineering applications as bearings, gears, railway wheels and rails, cams, pin-jointed links, etc. Whilst in most other types of stress calculation it is usual to neglect local deflection at the loading point when deriving equations for stress distribution in general bodies, in contact situations, e.g. the case of a circular wheel on a flat rail, such an assumption would lead to infinite values of compressive stress (load ÷ "zero" area = infinity). This can only be avoided by local deflection, even yielding, of the material under the load to increase the bearing area and reduce the value of the compressive stress to some finite value.

Contact stresses between curved bodies in compression are often termed "Hertzian" contact stresses after the work on the subject by Hertz[(1)] in Germany in 1881. This work was concerned primarily with the evaluation of the maximum compressive stresses set up at the mating surfaces for various geometries of contacting body but it formed the basis for subsequent extension of consideration by other workers of stress conditions within the whole contact zone both at the surface and beneath it. It has now been shown that the strength and load-carrying capacity of engineering components subjected to contact conditions is not completely explained by the Hertz equations by themselves, but that further consideration of the following factors is an essential additional requirement:

(a) Local yielding and associated residual stresses

Yield has been shown to initiate sub-surface when the contact stress approaches 1.2 σ_y (σ_y being the yield stress of the contacting materials) with so-called "*uncontained plastic flow*" commencing when the stress reaches 2.8 σ_y. Only at this point will material "escape" at the sides of the contact region. The ratio of loads to produce these two states is of the order of 350 although tangential (sliding) forces will reduce this figure significantly.

Unloading from any point between these two states produces a thin layer of residual tension at the surface and a sub-surface region of residual compression parallel to the surface. The residual stresses set up during an initial pass or passes of load can inhibit plastic flow in subsequent passes and a so-called "*shakedown*" situation is reached where additional plastic flow is totally prevented. Maximum contact pressure for shakedown is given by Johnson[(14)] as 1.6 σ_y.

(b) Surface shear loading caused by mutual sliding of the mating surfaces

Pure rolling of parallel cylinders has been considered by Radzimovsky[(5)] whilst the effect of tangential shear loading has been studied by Deresiewicz[(15)], Johnson[(16)], Lubkin[(17)], Mindlin[(18)], Tomlinson[(19)] and Smith and Liu[(20)].

(c) Thermal stresses and associated material property changes resulting from the heat set up by sliding friction. (Local temperatures can rise to some 500°F above ambient).

A useful summary of the work carried out in this area is given by Lipson & Juvinal[(21)].

(d) The presence of lubrication – particularly hydrodynamic lubrication – which can greatly modify the loading and resulting stress distribution

The effects of hydrodynamic lubrication on the pressure distribution at contact (see Fig. 10.1) and resulting stresses have been considered by a number of investigators including Meldahl[(22)], M'Ewen[(4)], Dowson, Higginson and Whitaker[(23)], Crook[(24)], Dawson[(25)] and

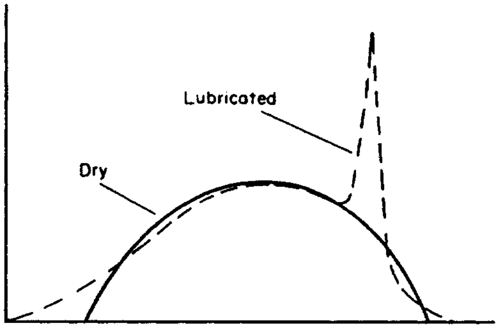


Fig. 10.1. Comparison of pressure distributions under dry and lubricated contact conditions.

Scott[26]. One important conclusion drawn by Dowson *et al.* is that at high load and not excessive speeds hydrodynamic pressure distribution can be taken to be basically Hertzian except for a high spike at the exit side.

(e) The presence of residual stresses at the surface of e.g. hardened components and their distribution with depth

In discussion of the effect of residually stressed layers on contact conditions, Sherratt[27] notes that whilst the magnitude of the residual stress is clearly important, the depth of the residually stressed layer is probably even more significant and the biaxiality of the residual stress pattern also has a pronounced effect. Considerable dispute exists even today about the origin of contact stress failures, particularly of surface hardened gearing, and the aspect is discussed further in §10.1.6 on gear contact stresses.

Muro[28], in X-ray studies of the residual stresses present in hardened steels due to rolling contact, identified a compressive residual stress peak at a depth corresponding to the depth of the maximum shear stress – a value related directly to the applied load. He therefore concluded that residual stress measurement could form a useful load-monitoring tool in the analysis of bearing failures.

Detailed consideration of these factors and even of the Hertzian stresses themselves is beyond the scope of this text. An attempt will therefore be made to summarise the essential formulae and behaviour mechanisms in order to provide an overall view of the problem without recourse to proof of the various equations which can be found in more advanced treatments such as those referred to below:-

The following special cases attracted special consideration:

(i) *Contact of two parallel cylinders* – principally because of its application to roller bearings and similar components. Here the Hertzian contact area tends towards a long narrow rectangle and complete solutions of the stress distribution are available from Belajef[2], Foppl[3], M'Ewen[4] and Radzimovsky[5].

(ii) *Spur and helical gears* – Buckingham[6] shows that the above case of contacting parallel cylinders can be used to fair accuracy for the contact of spur gears and whilst Walker[7] and Wellaver[8] show that helical gears are more accurately represented by contacting conical frustra, the parallel cylinder case is again fairly representative.

(iii) *Circular contact* – as arising in the case of contacting spheres or crossed cylinders. Full solutions are available by Foppl[3], Huber[9] Morton and Close[10] and Thomas and Hoersch[11].

(iv) *General elliptical contact.* Work on this more general case has been extensive and complete solutions exist for certain selected axes, e.g. the axes of the normal load. Authors include Belajef[2], Fessler and Ollerton[12], Thomas and Heorsch[11] and Ollerton[13].

Let us now consider the principal cases of contact loading:-

10.1.1. General case of contact between two curved surfaces

In his study of this general contact loading case, assuming elastic and isotropic material behaviour, Hertz showed that the intensity of pressure between the contacting surfaces could be represented by the elliptical (or, rather, semi-ellipsoid) construction shown in Fig. 10.2.

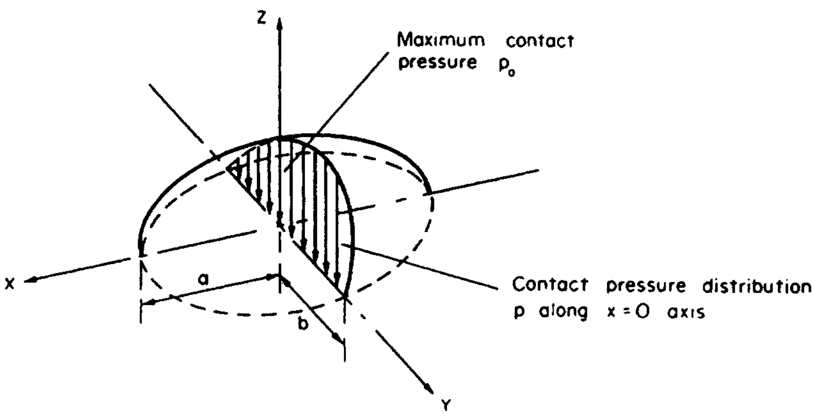


Fig. 10.2. Hertzian representation of pressure distribution between two curved bodies in contact.

If the maximum pressure at the centre of contact is denoted by p_0 then the **pressure at any other point within the contact region** was shown to be given by

$$p = p_0\sqrt{1 - \frac{x^2}{a^2} - \frac{x^2}{b^2}} \tag{10.1}$$

where a and b are the major and minor semi-axes, respectively.
The **total contact load** is then given by the volume of the semi-ellipsoid,

i.e.
$$P = \frac{2}{3}\pi a b p_0 \tag{10.2}$$

with the **maximum pressure** p_0 therefore given in terms of the applied load as

$$p_0 = \frac{3P}{2\pi ab} = \textbf{maximum compressive stress } \sigma_c \tag{10.3}$$

For any given contact load P it is necessary to determine the value of a and b before the maximum contact stress can be evaluated. These are found analytically from equations suggested by Timoshenko and Goodier[29] and adapted by Lipson and Juvinal[21].

i.e. $$a = m\left[\frac{3P\Delta}{4A}\right]^{1/3} \quad \text{and} \quad b = n\left[\frac{3P\Delta}{4A}\right]^{1/3}$$

with $$\Delta = \frac{1}{E_1}[1-\nu_1^2] + \frac{1}{E_2}[1-\nu_2^2]$$

a function of the elastic constants E and ν of the contacting bodies and

$$A = \frac{1}{2}\left[\frac{1}{R_1} + \frac{1}{R_1'} + \frac{1}{R_2} + \frac{1}{R_2'}\right]$$

with R and R' the maximum and minimum radii of curvature of the unloaded contact surfaces in two perpendicular planes.

For flat-sided wheels R_1 will be the wheel radius and R_1' will be infinite. Similarly for railway lines with head radius R_2 the value of R_2' will be infinite to produce the flat length of rail.

$$B = \frac{1}{2}\left[\left(\frac{1}{R_1} - \frac{1}{R_1'}\right)^2 + \left(\frac{1}{R_2} - \frac{1}{R_2'}\right)^2 + 2\left(\frac{1}{R_1} - \frac{1}{R_1'}\right)\left(\frac{1}{R_2} - \frac{1}{R_2'}\right)\cos 2\psi\right]^{1/2}$$

with ψ the angle between the planes containing curvatures $1/R_1$ and $1/R_2$.

Convex surfaces such as a sphere or roller are taken to be positive curvatures whilst internal surfaces of ball races are considered to be negative.

m and n are also functions of the geometry of the contact surfaces and their values are shown in Table 10.1 for various values of the term $\alpha = \cos^{-1}(B/A)$.

Table 10.1.

α degrees	20	30	35	40	45	50	55	60	65	70	75	80	85	90
m	3.778	2.731	2.397	2.136	1.926	1.754	1.611	1.486	1.378	1.284	1.202	1.128	1.061	1.000
n	0.408	0.493	0.530	0.567	0.604	0.641	0.678	0.717	0.759	0.802	0.846	0.893	0.944	1.000

10.1.2. Special case 1 – Contact of parallel cylinders

Consider the two parallel cylinders shown in Fig. 10.3(a) subjected to a contact load P producing a rectangular contact area of width $2b$ and length L. The contact stress distribution is indicated in Fig. 10.3(b).

The elliptical pressure distribution is given by the two-dimensional version of eqn (10.1)

i.e. $$p = p_0\sqrt{1 - \frac{y^2}{b^2}} \tag{10.5}$$

The **total load** P is then the volume of the prism

i.e. $$\boldsymbol{P = \tfrac{1}{2}\pi b L p_0} \tag{10.6}$$

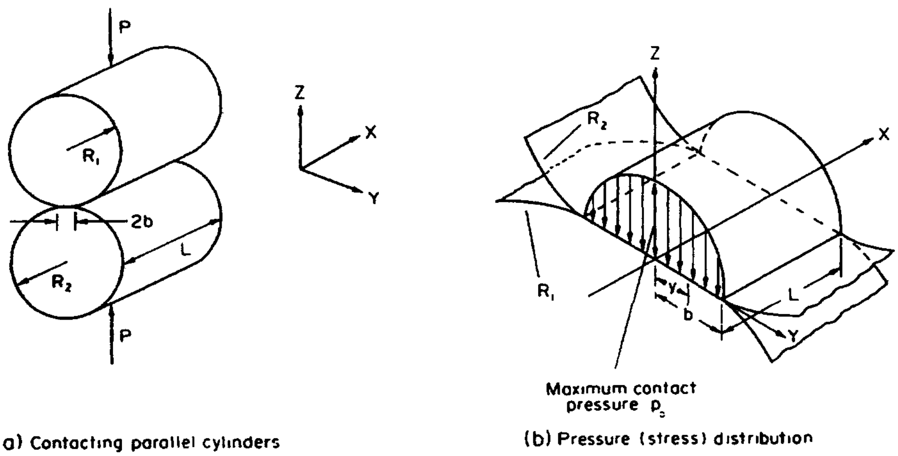


Fig. 10.3. (a) Contact of two parallel cylinders; (b) stress distribution for contacting parallel cylinders.

and the **maximum pressure or maximum compressive stress**

$$p_0 = \sigma_c = \frac{2p}{\pi b L} \tag{10.7}$$

The **contact width** can be related to the geometry of the contacting surfaces as follows:-

$$b = 1.076\sqrt{\frac{P\Delta}{L\left(\frac{1}{R_1}+\frac{1}{R_2}\right)}} \tag{10.8}$$

giving the **maximum compressive stress** as:

$$\sigma_c = -p_0 = -0.591\sqrt{\frac{P\left(\frac{1}{R_1}+\frac{1}{R_2}\right)}{L\Delta}} \tag{10.9}$$

(For a flat plate R_2 is infinite, for a cylinder in a cylindrical bearing R_2 is negative). **Stress conditions at the surface** on the load axis are then:

$$\sigma_z = \sigma_c = -p_0$$

$$\sigma_y = -p_0$$

$$\sigma_x = -2\nu p_0 \quad \text{(along cylinder length)}$$

The **maximum shear stress** is:

$$\tau_{max} = 0.295 p_0 \simeq 0.3 p_0$$

occurring at a depth beneath the surface of 0.786 b and on planes at 45° to the load axis.

In cases such as gears, bearings, cams, etc. which (as will be discussed later) can be likened to the contact of parallel cylinders, this shear stress will reduce gradually to zero as the rolling load passes the point in question and rise again to its maximum value as the next load contact is made. However, this will not be the greatest reversal of shear stress since there is another shear stress on planes parallel and perpendicular to the load axes known as the *"alternating" or "reversing"* shear stress, at a depth of 0.5 b and offset from the load axis by 0.85 b, which has a maximum value of 0.256 p_0 which changes from positive to negative as the load moves across contact.

The maximum shear stress on 45° planes thus varies between zero and 0.3 p_0 (approx) with an alternating component of 0.15 p_0 about a mean of 0.15 p_0. **The maximum alternating shear stress,** however, **has an alternating component of 0.256 p_0 about a mean of zero** – see Fig. 10.4. The latter is therefore considerably more significant from a fatigue viewpoint.

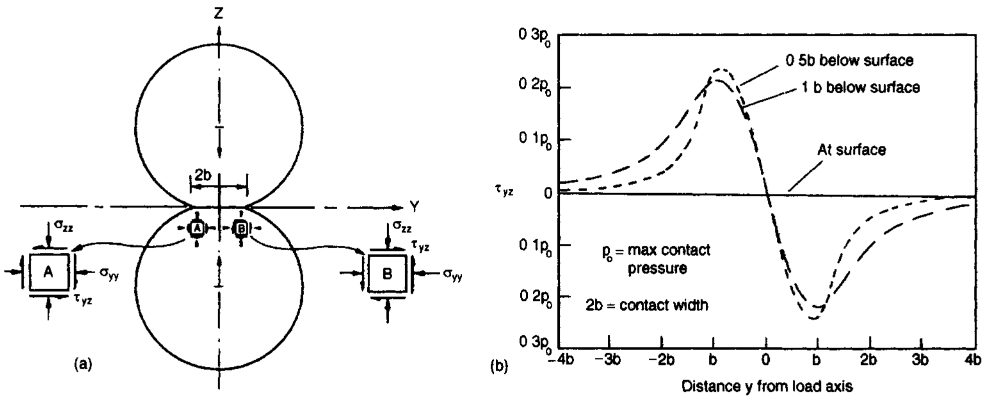


Fig. 10.4. Maximum alternating stress variation beneath contact surfaces.

N.B.: The above formulae assume the length of the cylinders to be very large in comparison with their radii. For short cylinders and/or cylinder/plate contacts with widths less than six times the contact area (or plate thickness less than six times the depth of the maximum shear stress) actual stresses can be significantly greater than those estimated by the given equations.

10.1.3. Combined normal and tangential loading

In normal contact conditions between contacting cylinders, gears, cams, etc. friction will be present reacting the sliding (or tendency to slide) of the mating surfaces. This will affect the stresses which are set up and it is usual in such cases to take the usual relationship between normal and tangential forces in the presence of friction

viz. $$F = \mu R \quad \text{or} \quad q = \mu p_0$$

where q is the tangential pressure distribution, assumed to be of the same form as that of the normal pressure. Smith and Liu[20] have shown that with such an assumption:

(a) A shear stress now exists on the surface at the contact point introducing principal stresses which are different from σ_x, σ_y and σ_z of the normal loading case.

(b) The maximum shear stress may exist either at the surface or beneath it depending on whether μ is greater than or less than 1/9 respectively.

(c) The stress range in the y direction is increased by almost 90% on the normal loading value and there is also a reversal of sign. A useful summary of stress distributions in graphical form is given by Lipson and Juvinal[21].

10.1.4. Special case 2 – Contacting spheres

For contacting spheres, eqns. (10.9) and (10.8) become

Maximum compressive stress (normal to surface)

$$\sigma_c = -p_0 = -0.62\sqrt[3]{\frac{P}{\Delta^2}\left[\frac{1}{R_1}+\frac{1}{R_2}\right]^2} \tag{10.10}$$

with a maximum value of

$$\sigma_c = -1.5P/\pi a^2 \tag{10.11}$$

Contact dimensions (circular)

$$a = b = 0.88\sqrt[3]{\frac{P\Delta}{\left[\frac{1}{R_1}+\frac{1}{R_2}\right]}} \tag{10.12}$$

As for the cylinder, if contact occurs between one sphere and a flat surface then R_2 is infinite, and if the sphere contacts inside a spherical seating then R_2 is negative.

The other two **principal stresses in the surface** plane are given by:

$$\sigma_x = \sigma_y = -\frac{(1+2\nu)}{2}p_0 \tag{10.13}$$

For steels with Poisson's ratio $\nu = 0.3$ the **maximum shear stress** is then:

$$\tau_{\max} \simeq 0.31p_0 \tag{10.14}$$

at a depth of half the radius of the contact surface.

The **maximum tensile stress** set up within the contact zone occurs at the edge of the contact zone in a radial direction with a value of:

$$\sigma_{t_{\max}} = \frac{(1-2\nu)}{3}p_0 \tag{10.15}$$

The circumferential stress at the same point is equal in value, but compressive, whilst the stress normal to the surface is effectively zero since contact has ended. With equal and opposite principal stresses in the plane of the surface, therefore, the **material is effectively in a state of pure shear.**

The **maximum octahedral shear stress** which is also an important value in consideration of elastic failure, occurs at approximately the same depth below the surface as the maximum shear stress. Its value may be obtained from eqn (8.24) by substituting the appropriate values of σ_x, σ_y and σ_z found from Fig. 10.5 which shows their variation with depth beneath the surface.

The **relative displacement, e, of the centres of the two spheres** is given by:

$$e = 0.77\sqrt[3]{P^2\left(\frac{1}{E_1}+\frac{1}{E_2}\right)^2\left(\frac{1}{R_1}+\frac{1}{R_2}\right)} \tag{10.16}$$

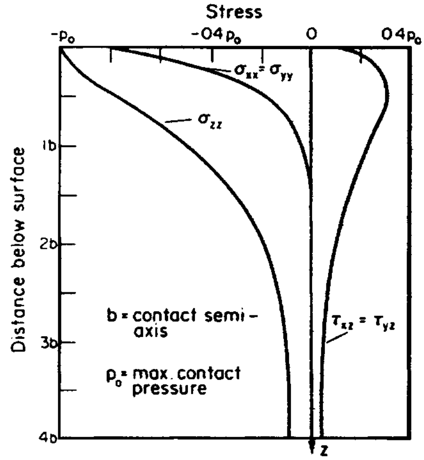


Fig. 10.5. Variation of stresses beneath the surface of contacting spheres.

For a *sphere contacting a flat surface* of the same material $R_2 = \infty$ and $E_1 = E_2 = E$. Substitution in eqns. (10.10) and (10.16) then yields

maximum compressive stress

$$\sigma_c = -0.62\sqrt[3]{\frac{PE^2}{4R_1^2}} \tag{10.17}$$

and relative displacement of centres

$$e = 1.54\sqrt[3]{\frac{P^2}{2E^2R_1}} \tag{10.18}$$

For a *sphere on a spherical seat* of the same material

$$\sigma_c = -0.62\sqrt[3]{PE^2\left[\frac{R_2-R_1}{R_1R_2}\right]^2} \tag{10.19}$$

with

$$e = 1.54\sqrt[3]{\frac{P^2}{2E^2}\left[\frac{R_2-R_1}{R_1R_2}\right]} \tag{10.20}$$

For *other, more general, loading cases* the reader is referred to a list of formulae presented by Roark and Young[33].

10.1.5. Design considerations

It should be evident from the preceding sections that the maximum Hertzian compressive stress is not, in itself, a valid criteria of failure for contacting members although it can be used as a valid design guide provided that more critical stress states which have a more direct influence on failure can be related directly to it. It has been shown, for example, that alternating shear stresses exist beneath the surface which are probably critical to fatigue life but these can be expressed as a simple proportion of the Hertzian pressure p_0 so that p_0 can be used as a simple index of contact load severity.

The contact situation is complicated under real service loading conditions by the presence of e.g. residual stresses in hardened surfaces, local yielding and associated additional residual stresses, friction forces and lubrication, thermal stresses and dynamic (including shock) load effects.

The failure of brittle materials under contact conditions correlates more closely with the maximum tensile stress at the surface rather than sub-surface shear stresses, whilst for static or very slow rolling operations failure normally arises as a result of excessive plastic flow producing indentation (*"brinelling"*) of the surface. In both cases, however, the Hertzian pressure remains a valuable design guide or reference.

By far the greatest number of failures of contacting components remains the surface or sub-surface fatigue initiated type variably known as *"pitting"*, *"spalling"*, *"onion-peel spalling"* or *"flaking"*. The principal service areas in which this type of failure occurs are gears and bearings.

10.1.6. Contact loading of gear teeth

Figure 10.6 shows the stress conditions which prevail in the region of a typical gear tooth contact. Immediately at the contact point, or centre of contact, there is the usual position of maximum compressive stress (p_0). Directly beneath this, and at a depth of approximately one-third of the contact width, is the maximum shear stress τ_{max} acting on planes at 45° to the load axis. Between these two positions lies the maximum alternating or reversed shear stress τ_{alt} acting on planes perpendicular and parallel to the surface. Whilst τ_{alt} is numerically smaller than τ_{max} it alternates between positive and negative values as the tooth proceeds

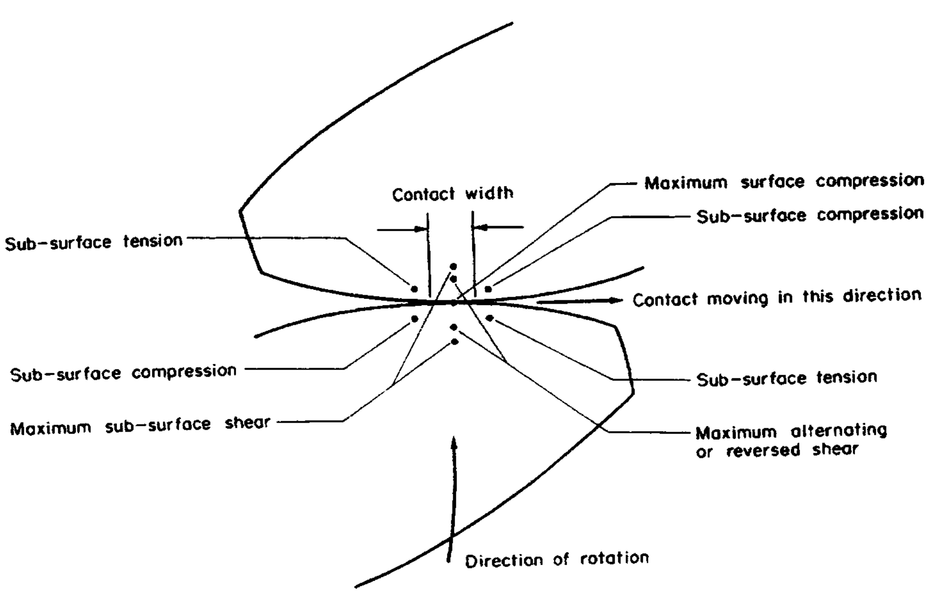


Fig. 10.6. Stress conditions in the region of gear tooth contact.

through mesh giving a stress range greater than that of τ_{max} which ranges between a single value and zero. It is argued by many that, for this reason, τ_{alt} is probably more significant to fatigue life than τ_{max} – particularly if its depth relates closely to that of peak residual stresses or case–core junctions of hardened gears.

As the gears rotate there is a combination of rolling and sliding motions, the latter causing additional surface stresses not shown in Fig. 8.6. Ahead of the contact area there is a narrow band of compression and behind the contact area a narrow band of tension. A single point on the surface of a gear tooth therefore passes through a complex variety of stress conditions as it goes through its meshing cycle. Both the surface and alternating stress change sign and other sub-surface stresses change from zero to their maximum value. Add to these fatigue situations the effects of residual stress, lubrication, thermal stresses and dynamic loading and it is not surprising that gears may fail in one of a number of ways either at the surface or sub-surface.

The majority of gear tooth failures are surface failures due to "pitting", "spalling", "flaking", "wear", etc. the three former modes referring to the fracture and shedding of pieces of various size from the surface. Considerable speculation and diverse views exist even among leading workers as to the true point of origin of some of these failures and considerable evidence has been produced of, apparently, both surface and sub-surface crack initiation. The logical conclusion would therefore seem to be that both types of initiation are possible depending on precisely the type of loading and contact conditions.

A strong body of opinion supports the suggestion of Johnson[(14,16)] and Almen[(30)] who attribute contact stress failures to local plastic flow at inclusions or flaws in the material, particularly in situations where a known overload has occurred at some time prior to failure. The overload is sufficient to produce the initial plastic flow and successive cycles then extend the region of plasticity and crack propagation commences. Dawson[(25)] and Akaoka[(31)] found evidence of sub-surface cracks running parallel to the surface, some breaking through to the surface, others completely unconnected with it. These were attributed to the fatigue action of the maximum alternating (reversed) shear stress. Undoubtedly, from the evidence presented by other authors, cracks can also initiate at the surface probably producing a "pitting" type of failure, i.e. smaller depth of damage. These cracks are suggested to initiate at positions of maximum tensile stress in the contact surface and subsequent propagation is then influenced by the presence (or otherwise) of lubricant.

In the case of helical gears, three-dimensional photoelastic tests undertaken by the author[(32)] indicate that maximum sub-surface stresses are considerably greater than those predicted by standard design procedures based on Hertzian contact and uniform loading along the contact line. Considerable non-uniformity of load was demonstrated which, together with dynamic effects, can cause maximum loads and stresses many times above the predicted nominal values. The tests showed the considerable benefit to be gained on the load distribution and resulting maximum stress values by the use of tip and end relief of the helical gear tooth profile.

10.1.7. Contact stresses in spur and helical gearing

Whilst the radius of an involute gear tooth will change slightly across the width of contact with a mating tooth it is normal to ignore this and take the contact of spur gear teeth as equivalent to the contact of parallel cylinders with the same radius of curvature at the point of contact. The Hertzian eqns. (10.8) and (10.9) can thus be applied to **spur gears** and, for typical steel elastic constant values of $\nu = 0.3$ and $E = 206.8\ \mathrm{GN/m^2}$, the **maximum contact** stress becomes

$$\sigma_c = -p_o = -0.475\sqrt{K}\ \mathrm{MN/m^2} \tag{10.21}$$

where

$$K = \frac{W}{F_w d}\left[\frac{m+1}{m}\right]$$

with W = tangential driving load = pinion torque ÷ pinion pitch radius
F_w = face width
d = pinion pitch diameter
m = ratio of gear teeth to pinion teeth; the pinion taken to be the smaller of the two mating teeth.

For **helical gears, the maximum contact stress** is given by

$$\sigma_c = -p_o = -C\sqrt{\frac{K}{m_p}} \tag{10.22}$$

where K is the same factor as for spur gears
m_p is the profile contact ratio
C is a constant

the values of m_p and C being found in Table 10.2, for various helix angles and pressure angles.

Table 10.2. Typical values of C and m_p for helical gears.

Pressure angle	Spur		15° Helix		30° Helix		45° Helix	
	C	m_p	C	m_p	C	m_p	C	m_p
$14\frac{1}{2}°$	0.546	2.10	0.528	2.01	0.473	1.71	0.386	1.26
$17\frac{1}{2}°$	0.502	1.88	0.485	1.79	0.435	1.53	0.355	1.13
20°	0.474	1.73	0.458	1.65	0.410	1.41	0.335	1.05
25°	0.434	1.52	0.420	1.45	0.376	1.25	0.307	0.949

10.1.8. Bearing failures

Considerable care is necessary in the design of bearings when selecting appropriate ball and bearing race radii. If the radii are too similar the area of contact is large and excessive wear and thermal stress (from frictional heating) results. If the radii are too dissimilar then the contact area is very small, local compressive stresses become very high and the load capacity of the bearing is reduced. As a compromise between these extremes the radius of the race is normally taken to be between 1.03 and 1.08 times the ball radius.

Fatigue life tests and service history then indicate that the life of ball bearings varies approximately as the cube of the applied load whereas, for roller bearings, a 10/3 power relationship is more appropriate. These relationships can only be used as a rough "rule of thumb", however, since commercially produced bearings, even under nominally similar and controlled production conditions, are notorious for the wide scatter of fatigue life results.

As noted previously, the majority of bearing failures are by spalling of the surface and most of the comments given in §10.1.6 relating to gear failures are equally relevant to bearing failures.

10.2. Residual Stresses

Introduction

It is probably true to say that all engineering components contain stresses (of variable magnitude and sign) before being subjected to service loading conditions owing to the history of the material prior to such service. These stresses, produced as a result of mechanical working of the material, heat treatment, chemical treatment, joining procedure, etc., are termed *residual stresses* and they can have a very significant effect on the fatigue life of components. These residual stresses are "locked into" the component in the absence of external loading and represent a datum stress over which the service load stresses are subsequently superimposed. If, by fortune or design, the residual stresses are of opposite sign to the service stresses then part of the service load goes to reduce the residual stress to zero before the combined stress can again rise towards any likely failure value; such residual stresses are thus extremely beneficial to the strength of the component and significantly higher fatigue strengths can result. If, however, the residual stresses are of the same sign as the applied stress, e.g. both tensile, then a smaller service load is required to produce failure than would have been the case for a component with a zero stress level initially; the strength and fatigue life in this case is thus reduced. Thus, both the magnitude and sign of residual stresses are important to fatigue life considerations, and methods for determining these quantities are introduced below.

It should be noted that whilst preceding chapters have been concerned with situations where it has been assumed that stresses are zero at zero load this is not often the case in practice, and great care must be exercised to either fully evaluate the levels of residual stress present and establish their effect on the strength of the design, or steps must be taken to reduce them to a minimum.

Bearing in mind that most loading applications in engineering practice involve fatigue to a greater or less degree it is relevant to note that surface residual stresses are the most critical as far as fatigue life is concerned since, almost invariably, fatigue cracks form at the surface. The work of §11.1.3 indicates that whilst tensile mean stresses promote fatigue crack initiation and propagation, compressive mean stresses are beneficial in that they impede fatigue failure. Compressive residual stresses are thus generally to be preferred (and there is not always a choice of course) if fatigue lives of components are to be enhanced. Indeed, compressive stresses are often deliberately introduced into the surface of components, e.g. by chemical methods which will be introduced below, in order to increase fatigue lives. There are situations, however, where compressive residual stress can be most undesirable; these include potential buckling situations where compressive surface stresses could lead to premature buckling failure, and operating conditions where the service loading stresses are also compressive. In the latter case the combined service and residual stresses may reach a sufficiently high value to exceed yield in compression and produce local plasticity on the first cycle of loading. On unloading, tensile residual stress "pockets" will be formed and these can act as local stress concentrations and potential fatigue crack initiation positions. Such a situation arises in high-temperature applications such as steam turbines and nuclear plant, and in contact load applications.

Whilst it has been indicated above that tensile residual stresses are generally deleterious to fatigue life there are again exceptions to this "rule", and very significant ones at that! It is now quite common to deliberately overload structures and components during proof testing to produce plastic flow at discontinuities and other stress concentrations to reduce their stress concentration effect on subsequent loading cycles. Other important techniques which involve the deliberate overloading of components in order to produce residual stress distribution favourable to subsequent loading cycles include "autofrettage" of thick cylinders (see §3.20(a)), "overspeeding" of rotating discs (see §3.20(b)) and pre-stressing of springs (see §3.8).

Whilst engineers have been aware of residual stresses for many years it is only recently that substantial efforts have been made to investigate their magnitudes and distributions with depth in components and hence their influence on performance and service life. This is probably due to the conservatism of old design procedures which generally incorporated sufficiently large safety factors to mask the effects of residual stresses on component integrity. However, with current drives for economy of manufacture coupled with enhanced product safety and reliability, design procedures have become far more stringent and residual stress effects can no longer be ignored. Principally, the designer needs to consider the effect of residual stress on structural or component failure but there is also need for detailed consideration of distortion and stability factors which are also closely related to residual stress levels.

10.2.1. Reasons for residual stresses

Residual stresses generally arise when conditions in the outer layer of a material differ from those internally. This can arise by one of three principal mechanisms: (a) mechanical processes, (b) chemical treatment, (c) heat treatment, although other mechanism are also discussed in the subsequent text.

(a) Mechanical processes

The most significant mechanical processes which induce surface residual stresses are those which involve plastic yielding and hence "cold-working" of the material such as rolling, shot-peening and forging. Practically all other standard machining procedures such as grinding, turning, polishing, etc., also involve local yielding (to a lesser extent perhaps) and also induce residual stresses. Reference should also be made to §3.9 and §3.10 which indicate how residual stresses can be introduced due to bending or torsion beyond the elastic limit.

Cold working

Shot peening is a very popular method for the introduction of favourable compressive residual stresses in the surface of components in order to increase their fatigue life. It is a process whereby small balls of iron or steel shot are bombarded at the component surface at high velocity from a rotating nozzle or wheel. It is applicable virtually to all metals and all component geometries and so is probably the most versatile of all the mechanical working processes. The bombardment tends to compress the surface layer and thus laterally try to expand it. This lateral expansion at the surface is resisted by the core material and residual compression results, its magnitude depending on the size of shot used and the peening velocity. Typically, residual stresses of the order of half the yield strength of the material are readily obtained, with peak values slightly sub-surface. However, special procedures such as "strain peening" which bombard the surface whilst applying external tensile loads can produce residuals approaching the full yield strength.

The major benefit of shot peening arises in areas of small fillet radii, notches or other high stress gradient situations and on poor surface finishes such as those obtained after rough machining or decarburisation. It is widely used in machine parts produced from high-strength steels and on gears, springs, structural components, engine con-rods and other motor vehicle components when fatigue lives have been shown to have been increased by factors in excess of 100%.

A number of different peening procedures exist in addition to standard shot peening with spherical shot, e.g. needle peening (bombardment by long needles with rounded ends), hammer peening (surface indented with radiused tool), roller-burnishing (rolling of undersized hole to required diameter), roto peening (impact of shot-coated flexible flaps). Figure 10.7 shows a typical residual stress distribution produced by shot peening, the maximum residual stress attainable being given by the following "rule of thumb" estimate

$$\sigma_m \simeq 500 + (0.2 \times \text{tensile strength})$$

for steels with a tensile strength between 650 MN/m^2 and 2 GN/m^2.

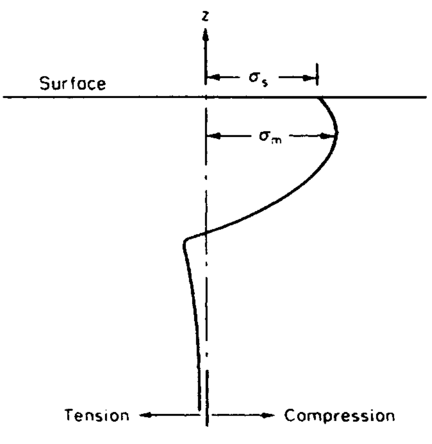


Fig. 10.7. Typical residual stress distribution with depth for the shot-peening process.

For lower-strength steels and alloys σ_m can initially reach the yield stress or 0.1% proof stress but this will fade under cyclic loading.

Cold rolling of threads, crankpins and axles relies on similar principles to those outlined above with, in this case, continuous pressure of the rollers producing controlled amounts of cold working. Further examples of cold working are the bending of pipes and conduits, cold shaping of brackets and clips and cold drawing of bars and tubes – sometimes of complex cross-section.

In some of the above applications the stress gradient into the material can be quite severe and a measurement technique which can produce results over reasonable depth is essential if residual stress–fatigue life relationships and mechanisms are to be fully understood.

Machining

It has been mentioned above that plastic deformation is almost invariably present in any machining process and the extent of the plastically deformed layer, and hence of the residually stressed region, will depend on the depth of cut, sharpness of tool, rates of speed and feed and the machineability of the material. With sharp tools, the heat generated at the tip of the tool will not have great influence and the residually stressed layer is likely to be compressive and relatively highly localised near the surface. With blunt tools or multi-tipped tools, particularly grinding, much more heat will be generated and if cooling is not sufficient this will produce thermally induced compressive stresses which can easily exceed the tensile stresses applied by the mechanical action of the tool. If they are large enough to exceed yield then tensile residual stresses may arise on cooling and care may need to be exercised in the type and level of service stress to which the component is then subjected. The depth of the residually stressed layer will depend upon the maximum temperatures reached during the machining operation and upon the thermal expansion coefficient of the material but it is likely that it will exceed that due to machining plastic deformation alone.

Residual stresses in manufactured components can often be very high; in grinding, for example, it is quite possible for the tensile residual stresses to produce cracking, particularly sub-surface, and etching techniques are sometimes employed after the grinding of e.g. bearings to remove a small layer on the surface in order to check for grinding damage. Distortion is another product of high residual stresses, produced particularly in welding and other heat treatment processes.

(b) Chemical treatment

The principal chemical treatments which are used to provide components with surface residual stress layers favourable to subsequent service fatigue loading conditions are nitriding, tufftriding and carburising.

Nitriding

Nitriding is a process whereby certain alloy steels are heated to about 550°C in an ammonia atmosphere for periods between 10 and 100 hours. Nitrides form in the surface of the steel with an associated volume increase. The core material resists this expansion and, as a result, residual compressive stresses are set up which can be very high (see Fig. 10.8). The surface layer, which typically is of the order of 0.5 mm thick, is extremely hard and the combination of this with the high surface residual compressive stresses make nitrided components exceptionally resistant to stress concentrations such as surface notches; fatigue lives of nitrided components are thus considerably enhanced over those of the parent material.

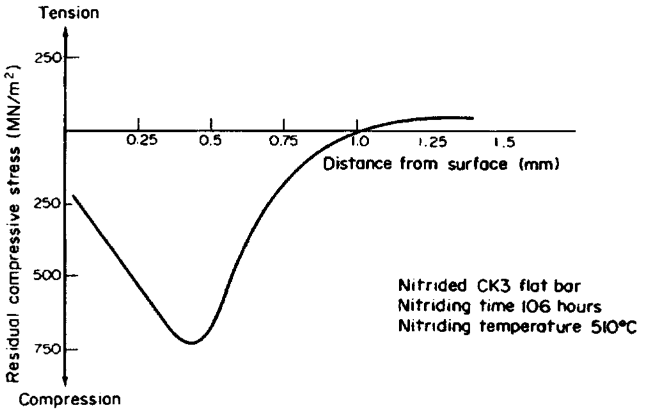


Fig. 10.8. Typical residual stress results for nitrided steel bar using the X-ray technique.

Minimal distortion or warping is produced by the nitriding process and no quenching is required.

Tufftriding

A special version of nitriding known as "tough nitriding" or, simply, "Tufftriding" consists of the heating of steel in a molten cyanic salt bath for approximately 90 minutes to allow nitrogen to diffuse into the steel surface and combine with the iron carbide formed in the outer skin when carbon is also released from the cyanic bath. The product of this combination is carbon-bearing epsilon-iron-nitride which forms a very tough but thin, wear-resistant layer, typically 0.1 mm thick. The process is found to be particularly appropriate for plain medium-carbon steels with little advantage over normal nitriding for the higher-strength alloy steels.

Carburising

Introduction of carbon into surface layers to produce so-called carburising may be carried out by solid, liquid or gaseous media. In each case the parent material contained in the selected medium such as charcoal, liquid sodium cyanide plus soda ash, or neutral gas enriched with propane, is heated to produce diffusion of the carbon into the surface. The depth of hardened case resulting varies from, typically, 0.25 mm on small articles to 0.37 mm on bearings (see Fig. 10.9).

(c) Heat treatment

Unlike chemical treatments, heat treatment procedures do not alter the chemical composition at the surface but simply modify the metallurgical structure of the parent material. Principal heat treatment procedures which induce favourable residual stress layers are induction hardening and flame hardening, although many other processes can also be considered within this category such as flame cutting, welding, quenching and even hot rolling or

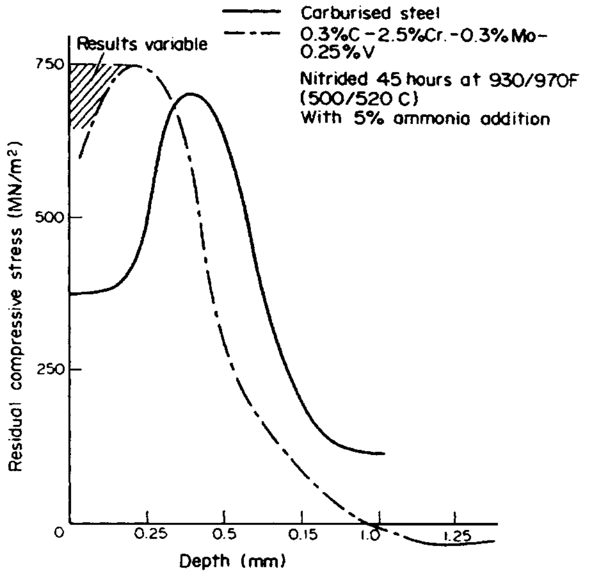


Fig. 10.9. Comparison of residual stress pattern present in nitrided chromium–molybdenum–vanadium steel and in a carburised steel – results obtained using the X-ray technique.

forging. In the two latter cases, however, chemical composition effects are included since carbon is removed from the surface by oxidation. This "decarburising" process produces surface layers with physical properties generally lower than those of the core and it is thus considered as a weakening process.

Returning to the more conventional heat treatment processes of flame and induction hardening, these again have a major effect at the surface where temperature gradients are the most severe. They produce both surface hardening and high compressive residual stresses with associated fatigue life improvements of up to 100%. There is some evidence of weakening at the case to core transition region but the process remains valuable for components with sharp stress gradients around their profile or in the presence of surface notches.

In both cases the surface is heated above some critical temperature and rapidly cooled and it is essential that the parent material has sufficient carbon or alloys to produce the required hardening by quenching. Heating either takes place under a gas flame or by electric induction heating caused by eddy currents generated in the surface layers. Typically, flame hardening is used for such components as gears and cams whilst induction hardening is applied to crankshaft journals and universal joints.

Many other components ranging from small shafts and bearings up to large forgings, fabricated structures and castings are also subjected to some form of heat treatment. Occasionally this may take the form of simple stress-relief operations aiming to reduce the level of residual stresses produced by prior manufacturing processes. Often, however, the treatment may be applied in order to effect some metallurgical improvement such as the normalising of large castings and forgings to improve their high-temperature creep characteristics or the surface hardening of gears, shafts and bearings. The required phase change of such processes usually entails the rapid cooling of components from some elevated temperature and it is this cooling which induces thermal gradients and, if these are sufficiently large (i.e. above yield), residual stresses. The component surfaces tend to cool more rapidly, introducing tensile stresses in the outer layers which are resisted by the greater bulk of the core material and result in residual compressive stress. As stated earlier, the stress gradient with depth into the material will depend upon the temperatures involved, the coefficient of thermal expansion of the material and the method of cooling. Particularly severe stress gradients can be produced by rapid quenching in water or oil.

Differential thermal expansion is another area in which residual stresses–or stress systems which can be regarded as residual stresses–arise. In cases where components constructed from materials with different coefficients of linear expansion are subjected to uniform temperature rise, or in situations such as heat exchangers or turbine casings where one material is subjected to different temperatures in different areas, free expansions do not take place. One part of the component attempts to expand at a faster rate and is constrained from doing so by an adjacent part which is either cooler or has a lower coefficient of expansion. Residual stresses will most definitely occur on cooling if the differential expansion stresses at elevated temperatures exceed yield.

When dealing with the quenching of heated parts, as mentioned above, a simple rule is useful to remember: *"What cools last is in tension"*. Thus the surface which generally cools first ends up in residual biaxial compression whilst the inner core is left in a state of triaxial tension. An exception to this is the quenching of normal through-hardened components when residual tensile stresses are produced at the surface unless a special process introduced by the General Motors Corporation of the U.S.A. termed "Marstressing" is used. This probably explains why surface-hardened parts generally have a much greater fatigue life than corresponding through-hardened items.

Should residual tensile stresses be achieved in a surface and be considered inappropriate then they can be relieved by tempering, although care must be taken to achieve the correct balance of ductility and strength after completion of the tempering process.

(d) Welds

One of the most common locations of fatigue failures resulting from residual stresses is at welded joints. Any weld junction can be considered to have three different regions; (a) the parent metal, (b) the weld metal, and (c) the heat-affected-zone (H.A.Z.), each with their own different physical properties including expansion coefficients. Residual stresses are then produced by the restraint of the parent metal on the shrinkage of the hot weld metal when it cools, and by differences in phase transformation behaviour of the three regions.

The magnitude and distribution of the residual stresses will depend upon the degree of preheat of the surfaces prior to welding, the heat input during welding, the number of weld passes, the match of the parent and weld metal and the skill of the operator. Even though the residuals can often be reduced by subsequent heat treatment this is not always effective owing to the different thermal expansions of the three zones. Differences between other physical properties in the three regions can also mean that failures need not always be associated with the region or part which is most highly stressed. Generally it is the heat-affected zones which contain sharp peaks of residual stress.

In welded structures, longitudinal shrinkages causes a weld and some parent material on either side to be in a state of residual tension often as high as the yield stress. This is balanced in the remainder of the cross-section by a residual compression which, typically, varies between 20 and 100 MN/m^2. When service load compressive stresses are applied to the members, premature yielding occurs in the regions of residual compressive stress, the stiffness of the member is reduced and there is an increased tendency for the component to buckle. In addition to this longitudinal "tendon force" effect there are also transverse effects in welds known as "pull-in" and "wrap-up" effects (see Fig. 10.10) again dictated by the level of residual stress set up.

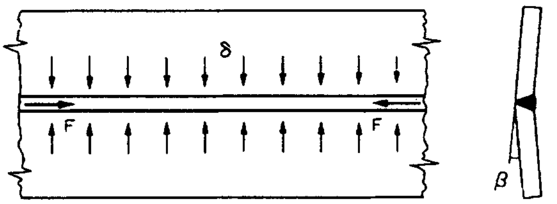


Fig. 10.10. The three basic parameters used to describe global weld shrinkage: F = tendon force; δ = pull-in; β-wrap-up.

The control of distortion is a major problem associated with large-scale welding. This can sometimes be minimised by clamping parts during welding to some pre-form curvatures or templates so that on release, after welding, they spring back to the required shape. Alternatively, components can be stretched or subjected to heat in order to redistribute the residual stress pattern and remove the distortion. In both cases, care needs to be exercised that unfavourable compressive stresses are not set up in regions which are critical to buckling failure.

(e) Castings

Another common problem area involving thermal effects and associated residual stresses is that of large castings. Whilst the full explanation for the source of residuals remains unclear (even after 70 years of research) it is clear that at least two mechanisms exist. Firstly, there are the physical restrictions placed on contraction of the casting, as it cools, by the mould itself and the differential thermal effects produced by different rates of cooling in different sections of e.g. different thickness. Secondly, there are metallurgical effects which arise largely as a result of differential cooling rates. Metal phase transformations and associated volume changes therefore occur at different positions at different times and rates. It is also suspected that different rates of cooling through the transformation range may create different material structures with different thermal coefficients. It is likely of course that the residual stress distribution produced will be as a result of a combination of these, and perhaps other, effects. It is certainly true, however, that whatever the cause there is frequently a need to subject large castings to some form of stress-relieving operation and any additional process such as this implies additional cost. It is therefore to be hoped that recent advances in measurement techniques, notably in the hole drilling method, will lead to a substantially enhanced understanding of residual stress mechanisms and to the development, for example, of suitable casting procedures which may avoid the need for additional stress-relieving operations.

10.2.2. The influence of residual stress on failure

It has been shown that residual stresses can be accommodated within the elastic failure theories quite simply by combining the residual and service load stresses (taking due account of sign) and inserting the combined stress value into the appropriate yield criterion. This is particularly true for ductile materials when both the Von Mises distortion energy theory and the Tresca maximum shear stress theory produce good correlation with experimental results. Should yielding in fact occur, there will normally be a change in the residual stress magnitude (usually a reduction) and distribution. The reduction of residual stress in this way is known as *"fading"*.

It is appropriate to mention here another type of failure phenomenon which is related directly to residual stress termed *"stress corrosion cracking"*. This occurs in metals which are subjected to corrosive environments whilst stressed, the cracks appearing in the surface layers.

Another source of potential failure is that of residual stress systems induced by the assembly of components with an initial lack of fit. This includes situations where the lack of fit is by deliberate design, e.g. shrinking or force-fit of compound cylinders or hubs on shafts and those where insufficient clearance or tolerances have been specified on mating components which, therefore, have to be forced together on assembly. This is, of course, a totally different situation to most of the cases listed above where residual stresses arise within a single member; it can nevertheless represent a potentially severe situation.

In contact loading situations such as in gearing or bearings, consideration should be given to the relationship between the distribution of residual stress with depth in the, typically, hardened surface and the depth at which the peak alternating shear stress occurs under the contact load. It is possible that the coincidence of the peak alternating value with the peak residual stress could explain the sub-surface initiation of cracks in spalling failures of such components. The hardness distribution with depth should also be considered in a similar way to monitor the strength/stress ratio, the lowest value of which can also initiate failure.

10.2.3. Measurement of residual stresses

The following methods have been used for residual stress investigations:

(1) Progressive turning or boring – Sach's method[34]
(2) Sectioning
(3) Layer removal – Rosenthal and Newton[35]
– Waisman and Phillips[36]
(4) Hole-drilling – Mather[37]
– Bathgate[38]
– Procter and Beaney[39, 40, 41]
(5) Trepanning or ring method – Milbradt[42]
(6) Chemical etch – Waisman and Phillips[36]
(7) Stresscoat brittle lacquer drilling – Durelli and Tsao[43]
(8) X-ray – French and Macdonald[44]
– Kirk[45, 46]
– Andrews et al[47]
(9) Magnetic method – Abuki and Cullety[48]

(10) Hardness studies – Sines and Carlson[49]
(11) Ultrasonics – Noranha and West[50]
– Kino[51]
(12) Modified layer removal – Hearn and Golsby[52]
– Spark machining–Denton[53]
(13) Photoelasticity – Hearn and Golsby[52]

Of these techniques, the most frequently applied are the layer removal (either mechanically or chemically), the hole-drilling and the X-ray measurement procedures. Occasionally the larger scale sectioning of a component after, e.g. initially coating the surface with a photoelastic coating, a brittle lacquer or marking a grid, is useful for the semi-quantitative assessment of the type and level of residual stresses present. In each case the relaxed stresses are transferred to the coating or grid and are capable of interpretation. In the case of the brittle lacquer method the surface is coated with a layer of a brittle lacquer such as "Tenslac" or "Stress coat" and, after drying, is then drilled with a small hole at the point of interest. The relieved residual stresses, if of sufficient magnitude, will then produce a crack pattern in the lacquer which can be readily evaluated in terms of the stress magnitude and type.

The layer removal, progressive turning or boring, trepanning, chemical etch, modified layer removal and hole-drilling methods all rely on basically the same principle. The component is either machined, etched or drilled in stages so that the residual stresses are released producing relaxation deformations or strains which can be measured by mechanical methods or electrical resistance strain gauges and, after certain corrections, related to the initial residual stresses. Apart from the hole-drilling technique which is discussed in detail below, the other techniques of metal removal type are classed as destructive since the component cannot generally be used after the measurement procedure has been completed.

Most layer removal techniques rely on procedures for metal removal which themselves introduce or affect the residual stress distribution and associated measurement by the generation of heat or as a result of mechanical working of the surface – or both. Conventional machining procedures including grinding, milling and polishing all produce significant effects. Of the 'mechanical' processes, spark erosion has been shown to be the least damaging process and the only one to have an acceptably low effect on the measured stresses. Regrettably, however, it is not always available and it may prove impractical in certain situations, e.g. site measurement. In such cases, either chemical etching procedures are used or, if these too are impractical, then standard machining techniques have to be employed with suitable corrections applied to the results.

X-ray techniques are well established and will also be covered in detail below; they are, however, generally limited to the measurement of strains at, or very near to, the surface and require very sophisticated equipment if reasonably accurate results are to be achieved.

Ultrasonic and magneto-elastic methods until recently have not received much attention despite the promise which they show. Grain orientation and other metallurgical inhomogeneities affect the velocity and attenuation of ultrasonic waves and further development of the technique is required in order to effectively separate these effects from the changes due to residual stress. A sample of stress-free material is also required for calibration of the method for quantitative results. Considerable further development is also required in the case of the magneto-elastic procedure which relies on the changes which occur in magnetic flux densities in ferromagnetic materials with changing stress.

The attempts to relate residual stress levels to the hardness of surfaces again appear to indicate considerable promise since they would give an alternative non-destructive technique which is simple to apply and relatively inexpensive. Unfortunately, however, the proposals do not seem to have achieved acceptance to date and do not therefore represent any significant challenge to the three "popular" methods.

Let us now consider in greater detail the two most popular procedures, namely hole-drilling and X-ray methods.

The hole-drilling technique

The hole-drilling method of measurement of residual stresses was initially proposed by Mathar[(37)] in 1933 and involves the drilling of a small hole (i.e. small diameter and depth) normal to the surface at the point of interest and measurement of the resulting local surface deformations or strains. The radial stress at the edge of the hole must be zero from simple equilibrium conditions so that local redistribution of stress or "relaxation" must occur. At the time the technique was first proposed, the method of measurement of the relaxations was by mechanical extensometers and the accuracy of the technique was limited. Subsequent workers, and particularly those in recent years,[(38–41)] have used electric resistance strain gauges and much more refined procedures of hole drilling metal removal as described below.

The particular advantages of the hole drilling technique are that it is accurate, can be made portable and is the least "destructive" of the metal removal techniques, the small holes involved generally not preventing further use of the component under test–although care should be exercised in any such decision and may depend upon the level of stress present. Stress values are obtained at a point and their variation with depth can also be established. This is important with surface-hardening chemical treatments such as nitriding or carburising where substantial stress variation and stress reversals can take place beneath the surface – see Fig. 10.9.

Whichever method of hole drilling is proposed, the procedure now normally adopted is the bonding of a three-element strain gauge rosette at the point under investigation and the drilling of a hole at the gauge centre in order to release the residual stresses and allowing the recording of the three strains ε_1, ε_2 and ε_3 in the three gauge element directions. Beaney[(39)] then quotes the formula which may be used for evaluation of the principal residual stresses σ_1 and σ_2 in the following form:

$$\left.\begin{matrix}\sigma_1\\ \sigma_2\end{matrix}\right\} = -\frac{1}{K_1}\cdot\frac{E}{2}\left\{\frac{(\varepsilon_1+\varepsilon_3)}{1-\nu(K_2/K_1)} \pm \frac{1}{1+\nu(K_2/K_1)}\sqrt{(\varepsilon_3-\varepsilon_1)^2+[(\varepsilon_1+\varepsilon_3)^2-2\varepsilon_2]^2}\right\}$$

Values of K_1 and K_2 are found by calibration, the value of K_1 and hence the "sensitivity" depending on the geometry of the hole and the position of the gauges relative to the hole-close control of these parameters are therefore important. For hole depths of approx. one hole diameter little error is introduced for steels by assuming the "modified" Poissons ratio term $\nu(K_2/K_1)$ to be constant at 0.3.

It should be noted that the drilled hole will act as a stress raiser with a stress concentration factor of at least 2. Thus, if residual stress levels are over half the yield stress of the material in question then some local plasticity will arise at the edge of the hole and the above formula will over-estimate the level of stress. However, the over-estimation is predictable and can be calibrated and in any case is negligible for residuals up to 70% of the yield stress.

Scaramangas *et al.*[55] show how simple correction factors can be applied to allow for variations of stress with depth, for the effects of surface preparation when mounting the gauges and for plastic yielding at the hole edge.

Methods of hole-drilling: (a) high-speed drill or router

Until recently, the 'standard' method of hole-drilling has been the utilisation of a small diameter, high-speed, tipped drill (similar to that used by dentists) fitted into a centring device which can be accurately located over the gauge centre using a removable eyepiece and fixed rigidly to the surface (see Fig. 10.11). Having located the fixture accurately over the required drilling position using cross-hairs the eye-piece is then removed and replaced by the drilling head. Since flat-bottomed holes were assumed in the derivation of the theoretical expressions it is common to use end-milling cutters of between $\frac{1}{8}$ in and $\frac{1}{4}$ in (3 mm to 6 mm) diameter. Unfortunately, the drilling operation itself introduces machining stresses into the component, of variable magnitude depending on the speed and condition of the tool,

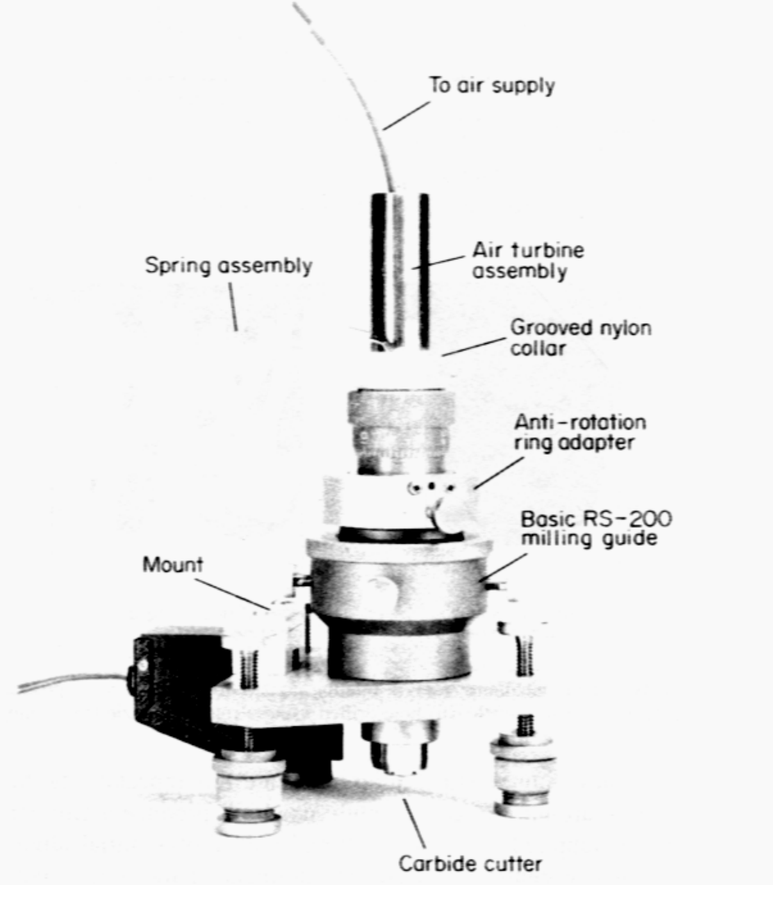


Fig. 10.11. Equipment used for the hole-drilling technique of residual stress measurement.

and these cannot readily be separated. Unless drilling is very carefully controlled, therefore, errors can arise in the measured strain values and the alternative "stress-free" machining technique outlined below is recommended.

(b) Air-abrasive machining

In this process the conventional drilling head is replaced by a device which directs a stream of air containing fine abrasive particles onto the surface causing controlled erosion of the material – see Fig. 10.12(a). The type of hole produced – see Fig. 10.12(b) – does not have a rectangular axial section but can be trepanned to produce axisymmetric, parallel-sided, holes

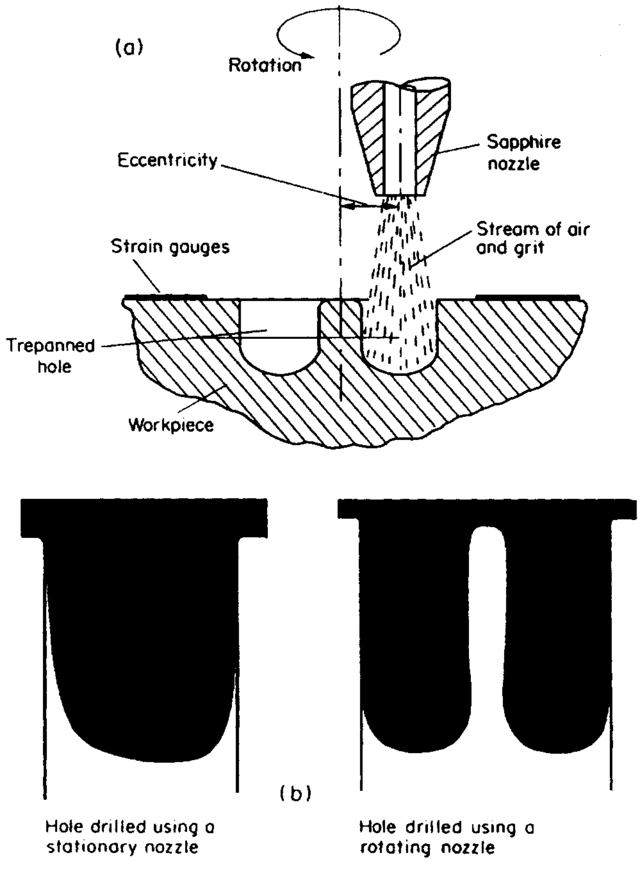


Fig. 10.12. (a) Air abrasive hole machining using a rotating nozzle: (b) types of holes produced: (i) hole drilled using a stationary nozzle, (ii) hole drilled using a rotating nozzle (trepanning)

with repeatable accuracy. Square-sided holes often aid this repeatability, but rotation of the nozzle on an eccentric around the gauge centre axis will produce a circular trepanned hole.

The optical device supplied with the "drilling" unit allows initial alignment of the unit with respect to the gauge centre and also measurement of the hole depth to enable stress distributions with depth to be plotted. It is also important to ascertain that the drilled depth is at least 1.65 mm since this has been shown to be the minimum depth necessary to produce full stress relaxation in most applications.

Whilst the strain gauge rosettes originally designed for this work were for hole sizes of 1.59 mm, larger holes have now been found to give greater sensitivity and accuracy. Optimum hole size is now stated to be between 2 and 2.2 mm.

X-ray diffraction

The X-ray technique is probably the most highly developed non-destructive measurement technique available today. Unfortunately, however, although semi-portable units do exist, it is still essentially a laboratory tool and the high precision equipment involved is rather expensive. Because the technique is essentially concerned with the measurement of stresses in the surface it is important that the very thin layer which is examined is totally representative of the conditions required.

Two principal X-ray procedures are in general use: the *diffractometer approach* and the *film technique*. In the first of these a diffractometer is used to measure the relative shift of X-ray diffraction lines produced on the irradiated surface. The individual crystals within any polycrystalline material are made up of families of identical planes of atoms, with a fairly uniform interplanar spacing d. The so-called lattice strain normal to the crystal planes is then $\Delta d/d$ and at certain angles of incidence (known as Bragg angles) X-ray beams will be diffracted from a given family of planes as if they were being reflected. The diffraction is governed by the Bragg equation:

$$n\lambda = 2d \sin\theta \tag{10.23}$$

where n is an integer corresponding to the order of diffraction
λ is the wavelength of the X-ray radiation
θ is the angle of incidence of the crystal plane.

Any change in applied or residual stress caused, e.g. by removal of a layer of material from the rear face of the specimen, will produce a change in the angle of reflection obtained by differentiating the above equation

i.e. $$\Delta\theta = -\frac{\Delta d}{d} \cdot \tan\theta \tag{10.24}$$

This equation relates the change in angle of incidence to the lattice strain at the surface. Typically, $\Delta\theta$ ranges between 0.3° and 0.02° depending on the initial value of θ used.

Two experimental procedures can be used to evaluate the stresses in the surface: (a) the $\sin^2\psi$ technique and (b) the two-exposure technique, and full details of these procedures are given by Kirk[45]. Certain problems exist in the interpretation of the results, such as the elastic anisotropy which is exhibited by metals with respect to their crystallographic directions. The appropriate values of E and ν have thus to be established by separate X-ray experiments on tensile test bars or four-point bending beams.

In applications where successive layers of the metal are removed in order to determine the values of sub-surface stresses (to a very limited depth), the material removal produces a redistribution of the stresses to be measured and corrections have to be applied. Standard formulae[54] exist for this process.

An alternative method for determination of the Bragg angle θ is to use the film technique with a so-called *"back reflection"* procedure. Here the surface is coated with a thin layer of powder from a standard substance such as silver which gives a diffraction ring near to that from the material under investigation. Measurements of the ring diameter and film–specimen distance are then used in either a single-exposure or double-exposure procedure to establish the required stress values.[47]

The X-ray technique is valid only for measurement in materials which are elastic, homogeneous and isotropic. Fortunately most polycrystaline metals satisfy this requirement to a fair degree of accuracy but, nevertheless, it does represent a constraint on wider application of the technique.

10.2.4. Summary of the principal effects of residual stress

(1) Considerable improvement in the fatigue life of components can be obtained by processes which introduce residual stress of appropriate sign in the surface layer. Compressive residual stresses are particularly beneficial in areas of potential fatigue (tensile stress) failure.
(2) Pre-loading of components beyond yield in the same direction as future service loading will produce residual stress systems which strengthen the components.
(3) Surface-hardened components have a greater fatigue resistance than through-hardened parts.
(4) Residual stresses have their greatest influence on parts which are expected to undergo high numbers of loading cycles (i.e. low strain–high cycle fatigue); they are not so effective under high strain–low cycle fatigue conditions.
(5) Considerable benefit can be obtained by local strengthening procedures at, e.g. stress concentrations, using shot peening or other localised procedures.
(6) Failures always occur at positions where the ratio of strength to stress is least favourable. This is particularly important in welding applications.
(7) Consideration of the influence of residual stresses must be part of the design process for all structures and components.

10.3. Stress concentrations

Introduction

In practically all the other chapters of this text loading conditions and components have been analysed in which stresses have been assumed, or shown, to be uniform or smoothly varying. In practice, however, this rarely happens owing to the presence of grooves, fillets, threads, holes, keyways, points of concentrated loading, material flaws, etc. In each of these cases, and many others too numerous to mention, the stress at the "discontinuity" is likely to be significantly greater than the assumed or nominally calculated figure and such discontinuities are therefore termed *stress raisers or stress concentrations.*

Most failures of structural members or engineering components occur at stress concentrations so that it is important that designers understand their significance and the magnitude of their effect since it is practically impossible to design any component without some form of stress raiser. In fatigue loading conditions, for example, virtually all failures occur at stress concentrations and it is therefore necessary to be able to develop a procedure which will take them into account during design strength calculations.

Geometric discontinuities such as holes, sharp fillet radii, keyways, etc. are probably the most prevalent causes of failure and typical examples of failure are shown in Figs. 10.13 and 10.14.

Fig. 10.13. Combined bending and tension fatigue load failure at a sharp fillet radius stress concentration position on a large retaining bolt of a heavy-duty extrusion press.

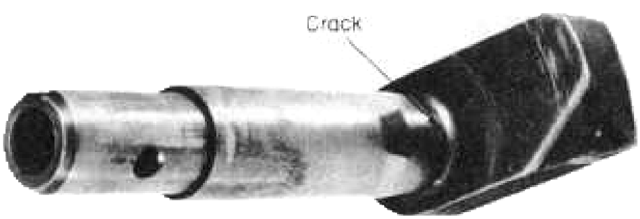


Fig. 10.14. Another failure emanating from a sharp fillet stress concentration position.

In order to be able to understand the stress concentration mechanism consider the simple example of the tensile bar shown in Figs. 10.15(a) and 10.15(b). In Fig. 10.15(a) the bar is solid and the tensile stress is nominally uniform at $\sigma = P/A = (P/bt)$ across the section.

In Fig. 10.15(b), however, the bar is drilled with a transverse hole of diameter d. Away from the hole the stress remains uniform across the section at $\sigma = P/bt$ and, using a similar calculation, the stress at the section through the centre of the hole should be $\sigma_{\text{nom}} = P/(b-d)t$ and uniform.

The sketch shows, however, that the stress at the edge of the hole is, in fact, much greater than this, indeed it is nearly 3 times as great (depending on the diameter d).

The ratio of the actual maximum stress $\sigma_{\max}$ and the nominal value σ_{nom} is then termed the *stress concentration factor* for the hole.

$$\textbf{Stress concentration factor } K_t = \frac{\textbf{maximum stress}}{\textbf{nominal stress}} = \frac{\sigma_{\max}}{} \tag{10.25}$$

For a small hole $K_t \simeq 3$.

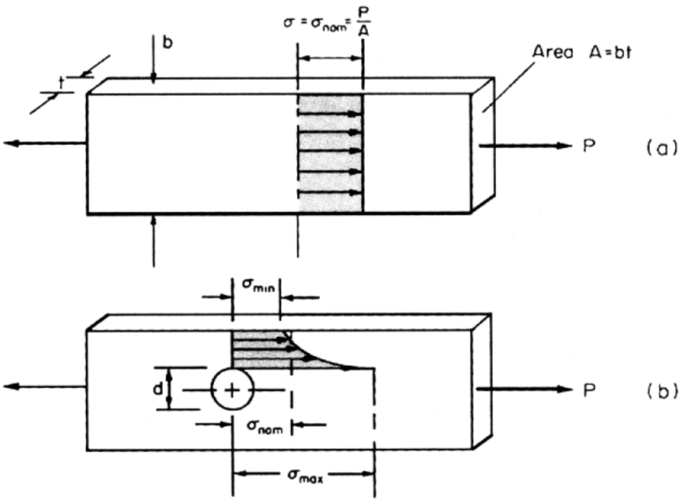


Fig. 10.15. Stress concentration effect of a hole in a tension bar.

It should also be observed that whilst the stress local to the hole is greater than the nominal stress, at distances greater than about one hole diameter away from the edge of the hole the stresses are less than the nominal value. This must be true from simple equilibrium condition since the sum of (stress × area) across the section must balance the applied force; if the stress is greater than the nominal or average stress at one point it must therefore be less in another.

It should be evident that even had a safety factor of, say, 2.8 been used in the stress calculations for the tensile bar in question the bar would have failed since the stress concentration factor exceeds this and it is important not to rely on safety factors to cover stress concentration effects which can generally be estimated quite well, as will be discussed later. Safety factors should be reserved for allowing for uncertainties in service load conditions which cannot be estimated or anticipated with any confidence.

The cause of the stress concentration phenomenon is perhaps best understood by the use of a few analogies; firstly, that of the flow of liquid through a channel. It can be shown that the distribution of stress through a material is analogous to that of fluid flow through a channel, the cross-section of which varies in the same way as that of the material cross-section. Thus Fig. 10.16 shows the experimentally obtained flow lines for a fluid flowing round a pin of diameter d in a channel of width b, i.e. the same geometry as that of the tensile bar. It will be observed that the flow lines crowd together as the fluid passes the pin and the velocity of flow increases significantly in order that the same quantity of fluid can pass per second through the reduced gap. This is directly analogous to the increased stress at the hole in the tensile bar. Any other geometrical discontinuity will have a similar effect see Fig. 10.17.

An alternative analogy is to consider the bar without the hole as a series of stretched rubber bands parallel to each other as are the flow lines of Fig. 10.16(a). Again inserting a pin to represent the hole in the bar produces a distortion of the bands and pressure on the pin at its top and bottom diameter extremities – again directly analogous to the increased "pressure" or stress felt by the bar at the edge of the hole.

It is appropriate to mention here that the stress concentration factor calculation of eqn. (10.25) only applies while stresses remain in the elastic range. If stresses are increased

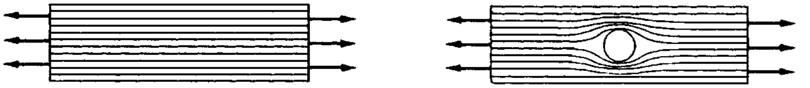


Fig. 10.16. Flow lines (a) without and (b) with discontinuity.

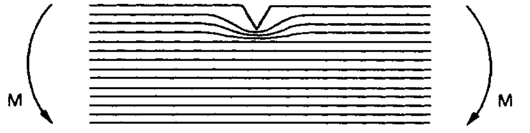


Fig. 10.17. Flow lines around a notch in a beam subjected to bending.

beyond the elastic region then local yielding takes place at the stress concentration and stresses will be redistributed as a result. In most cases this can reduce the level of the maximum stress which would be estimated by the stress concentration factor calculation. In the case of a notch or sharp-tipped crack, for example, the local plastic region forms to blunt the crack tip and reduce the stress-concentration effect for subsequent load increases. This local yielding represents a limiting factor on the maximum realistic value of stress concentration factor which can be obtained for most structural engineering materials. For very brittle materials such as glass, however, the high stress concentrations associated with very sharp notches or scratches can readily produce fracture in the absence of any significant plasticity. This, after all, is the principle of glass cutting!

The ductile flow or local yielding at stress concentrations is termed a *notch-strengthening effect* and stress concentration factors, although defined in the same way, become *plastic stress concentration factors* K_p. For most ductile materials, as the maximum stress in the component is increased up to the maximum tensile strength of the material, the value of K_p tends towards unity thus indicating that the *static* strength of the component has not been reduced significantly by the presence of the stress concentration. This is not the case for impact, fatigue or brittle fracture conditions where stress concentrations play a very significant part.

In complete contrast, stress concentrations of the types mentioned above are relatively inconsequential to the strength of heterogeneous brittle materials such as cast iron because of the high incidence of "natural" internal stress raisers within even the un-notched material, e.g. internal material flaws or impurities.

It has been shown above that the magnitude of the local increase in stress in the tensile bar caused by the stress concentration, i.e. the value of the stress concentration factor, is related to the geometry of both the bar and the hole since both b and d appear in the calculation of eqn (10.25).

Figure 10.18 shows the way in which the stress concentration value changes with different hole/bar geometries. It will be noted that the most severe effect (when related to the nominal area left after drilling the hole) is obtained when the hole diameter is smallest, producing a stress concentration factor (s.c.f.) of approximately 3. Whilst this is the largest s.c.f. value it does not mean, of course, that the bar is weaker the smaller the hole. Clearly a very large hole leaves very little material to carry the tensile load and the nominal stress will increase to produce failure. It is the combination of the nominal stress and the stress concentration factor which gives the value of the maximum stress that eventually produces

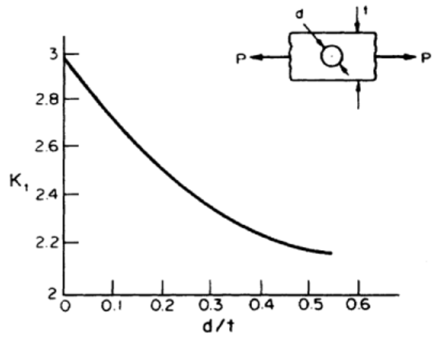


Fig. 10.18. Variation of elastic stress concentration factor K_t for a hole in a tensile bar with varying d/t ratios.

failure – both must therefore be considered.

i.e. Maximum stress = nominal stress × stress concentration factor.

If load on the bar is increased sufficiently then failure will occur, the crack emanating from the peak stress position at the edge of the hole across the section to the outside (see Fig. 10.19).

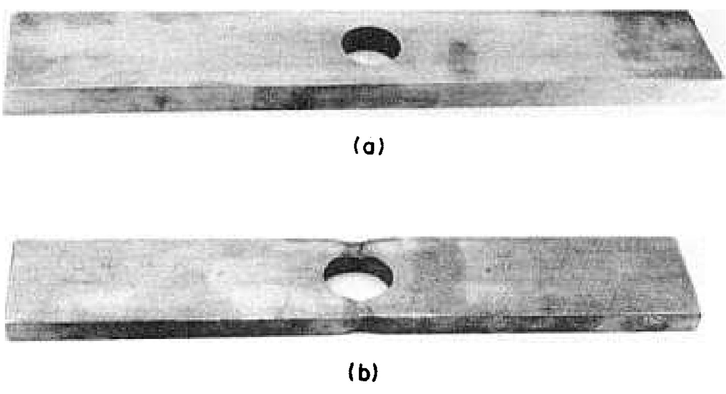


Fig. 10.19. Tensile bar loaded to destruction – crack initiates at peak stress concentration position at the hole edge.

Other geometric factors will affect the stress-concentration effect of discontinuities such as the hole, e.g. its shape. Figure 10.20 shows the effect of various hole shapes on the s.c.f. achieved in the tensile plate for which it can be shown that, approximately, $K_t = 1 + 2(A/B)$ where A and B are the major and minor axis dimensions of the elliptical holes perpendicular and parallel to the axis of the applied stress respectively. When $A = B$, the ellipse becomes the circular hole considered previously and $K_t \simeq 3$.

For large values of B, i.e. long elliptical slots parallel to the applied stress axis, stress concentration effects are reduced below 3 but for large A values, i.e. long elliptical slots perpendicular to the stress axis, s.c.f.'s rise dramatically and the potentially severe effect of slender slots or cracks such as this can readily be seen.

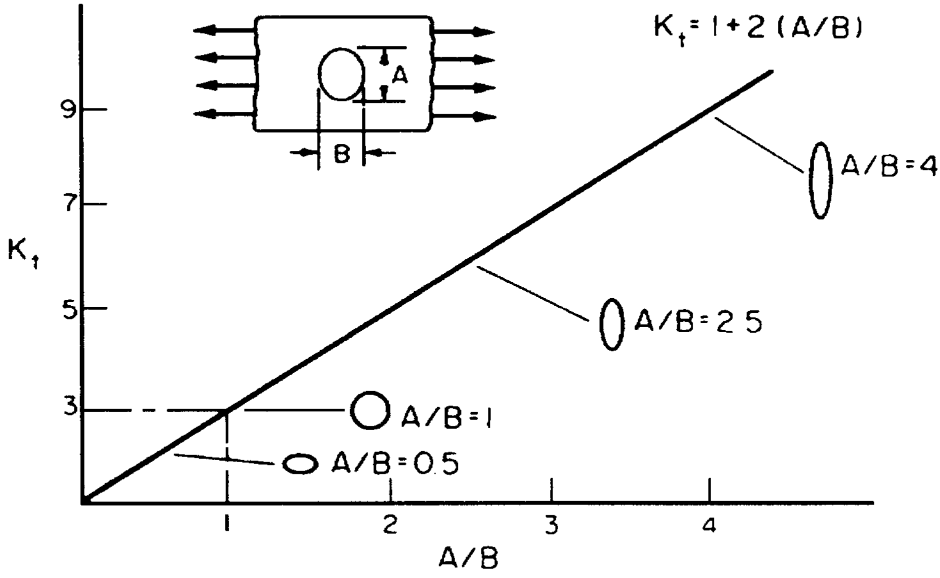


Fig. 10.20. Effect of shape of hole on the stress concentration factor for a bar with a transverse hole.

This is, of course, the theory of the perforated toilet paper roll which should tear at the perforation every time–which only goes to prove that theory very rarely applies perfectly in every situation!! (Closer consideration of the mode of loading and material used in this case helps to defend the theory, however.)

10.3.1. Evaluation of stress concentration factors

As stated earlier, the majority of the work in this text is devoted to consideration of stress situations where stress concentration effects are not present, i.e. to the calculation of nominal stresses. Before resulting stress levels can be applied to design situations, therefore, it is necessary for the designer to be able to estimate or predict the stress concentration factors associated with his particular design geometry and nominal stresses. In some cases these have been obtained analytically but in most cases graphs have been produced for standard geometric discontinuity configurations using experimental test procedures such as photoelasticity, or more recently, using finite element computer analysis.

Figures 10.21 to 10.30 give stress concentration factors for fillets, grooves and holes under various types of loading based upon a highly recommended reference volume(57). Many other geometrical forms and loading conditions are considered in this and other reference texts(60) but for non-standard cases the application of the photoelastic technique is also highly recommended (see §6.12).

The reference texts give stress concentration factors not only for two-dimensional plane stress situations such as the tensile plate but also for triaxial stress systems such as the common case of a shaft with a transverse hole or circumferential groove subjected to tension, bending or torsion.

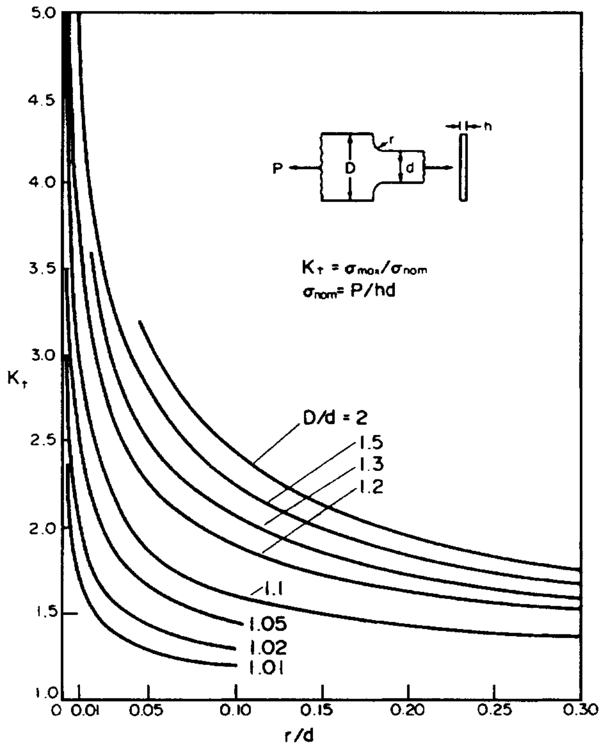


Fig. 10.21. Stress concentration factor K_t for a stepped flat tension bar with shoulder fillets.

Figures 10.31, 10.32 and 10.34 indicate the ease with which stress concentration positions can be identified within photoelastic models as the points at which the fringes are greatest in number and closest together. It should be noted that:

(1) Stress concentration factors are different for a single geometry subjected to different types of loading. Appropriate K_t values must therefore be obtained for each type of loading. Figure 10.33 shows the way in which the stress concentration factors associated with a groove in a circular bar change with the type of applied load.

(2) Care must be taken that stress concentration factors are applied to nominal stresses calculated on the same basis as that of the s.c.f. calculation itself, i.e. the same cross-sectional area must be used–usually the net section left after the concentration has been removed. In the case of the tensile bar of Fig. 10.15 for example, σ_{nom} has been taken as $P/(b-d)t$. An alternative system would have been to base the nominal stress σ_{nom} upon the full 'un-notched' cross-sectional area i.e. $\sigma_{nom} = P/t$. Clearly, the stress concentration factors resulting from this approach would be very different, particularly as the size of the hole increases.

(3) In the case of combined loading, the stress calculated under each type of load must be multiplied by its own stress concentration factor. In combined bending and axial load, for example, the bending stress ($\sigma_b = My/I$) should be multiplied by the bending s.c.f. and the axial stress ($\sigma_d = P/A$) multiplied by the s.c.f. in tension.

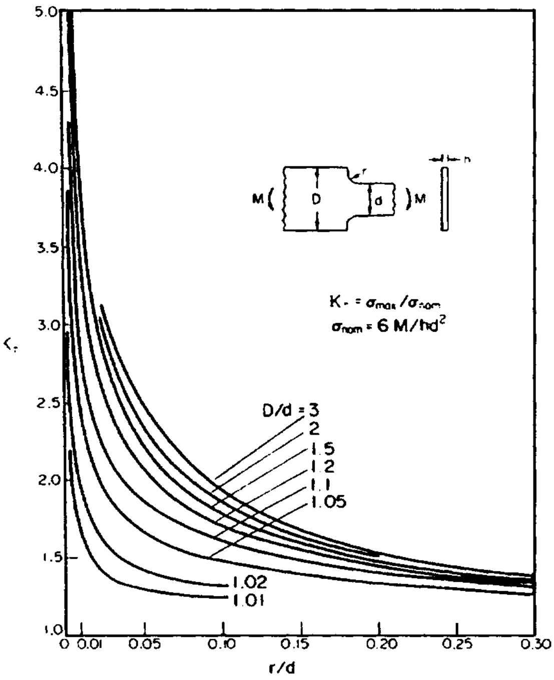


Fig. 10.22. Stress concentration factor K_t for a stepped flat tension bar with shoulder fillets subjected to bending.

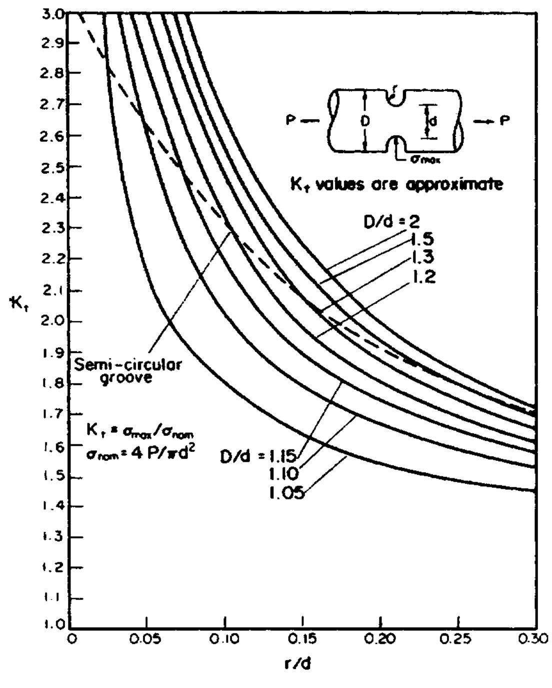


Fig. 10.23. Stress concentration factor K_t for a round tension bar with a U groove.

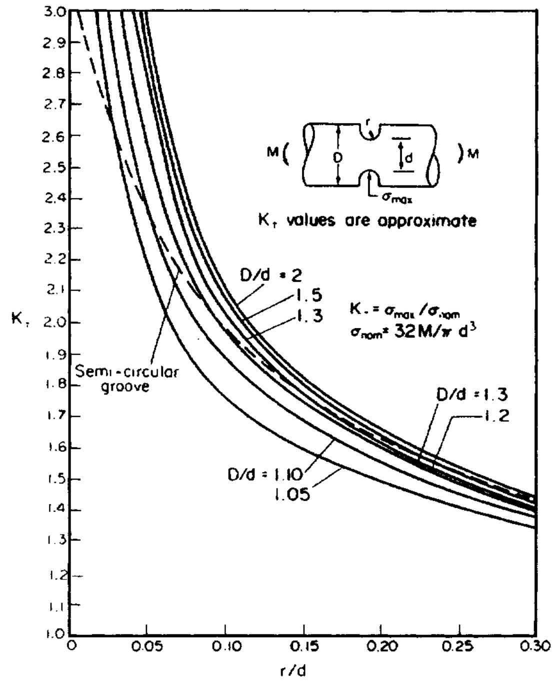


Fig. 10.24. Stress concentration factor K_t for a round bar with a U groove subjected to bending.

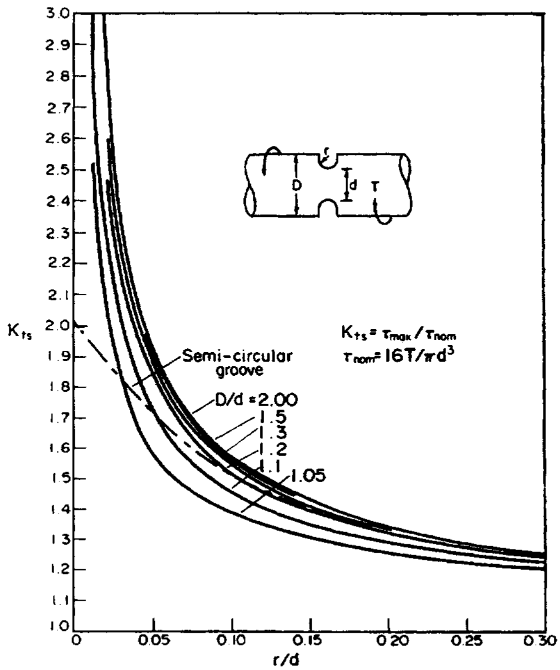


Fig. 10.25. Stress concentration factor K_t for a round bar with a U groove subjected to torsion.

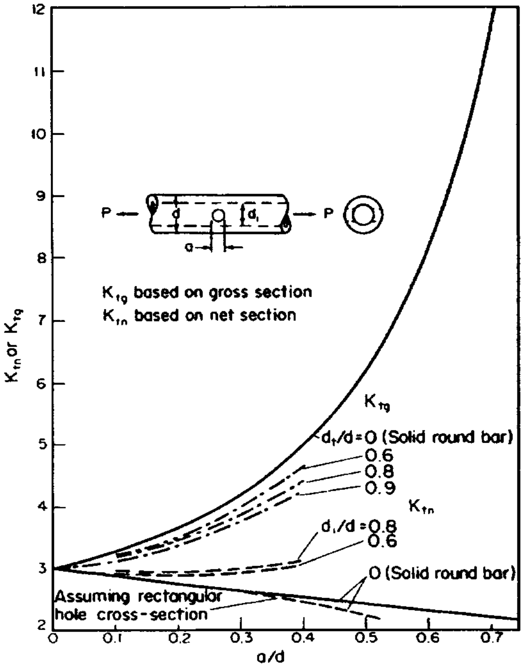


Fig. 10.26. Stress concentration factor K_t for a round bar or tube with a transverse hole subjected to tension.

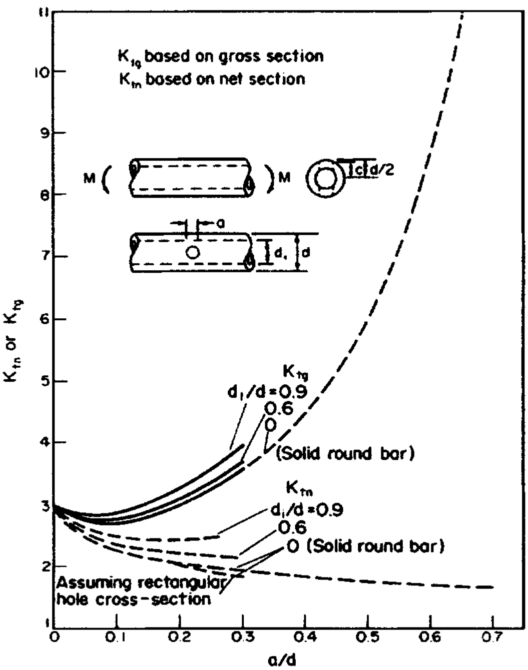


Fig. 10.27. Stress concentration factor K_t for a round bar or tube with a transverse hole subjected to bending.

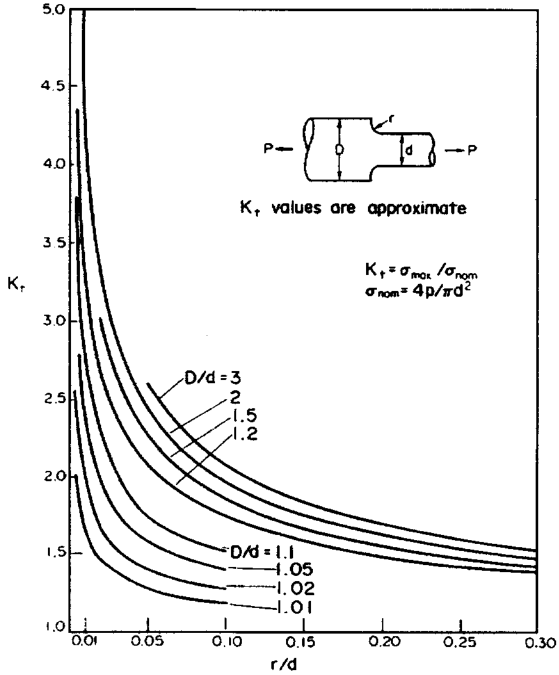


Fig. 10.28. Stress concentration factor K_t for a round bar with shoulder fillet subjected to tension.

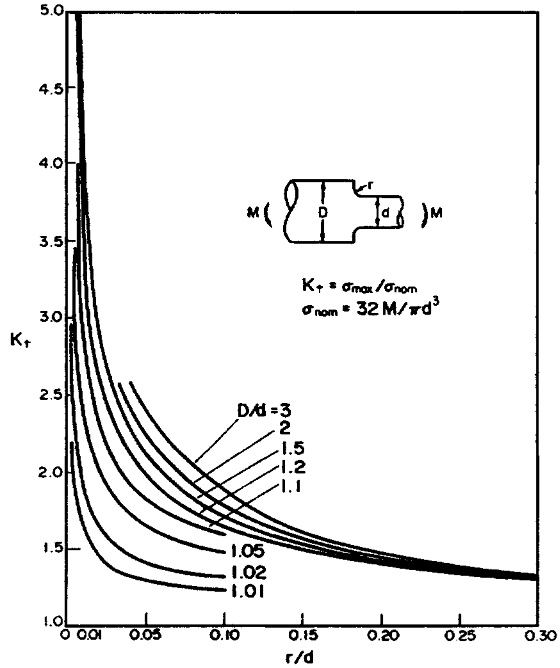


Fig. 10.29. Stress concentration factor K_t for a stepped round bar with shoulder fillet subjected to bending.

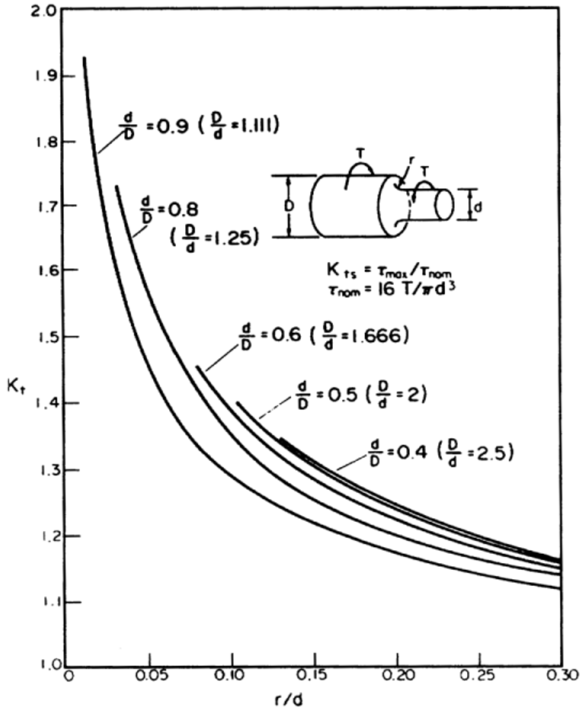


Fig. 10.30. Stress concentration factor K_t for a stepped round bar with shoulder fillet subjected to torsion.

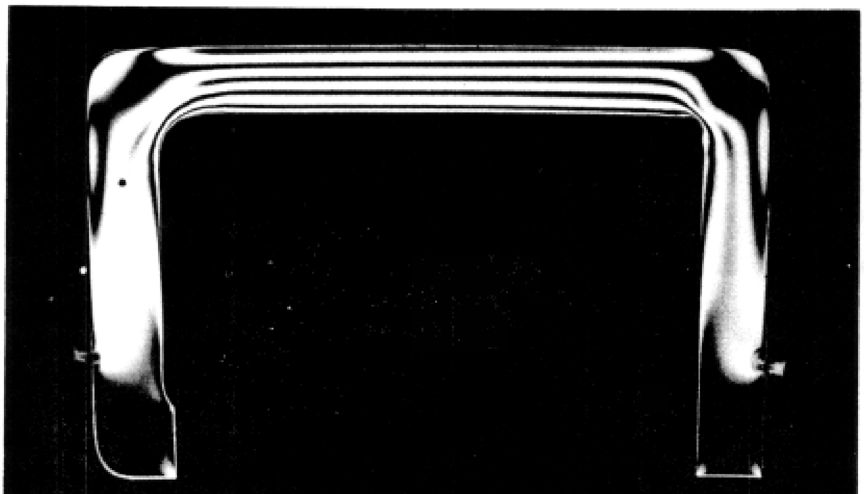
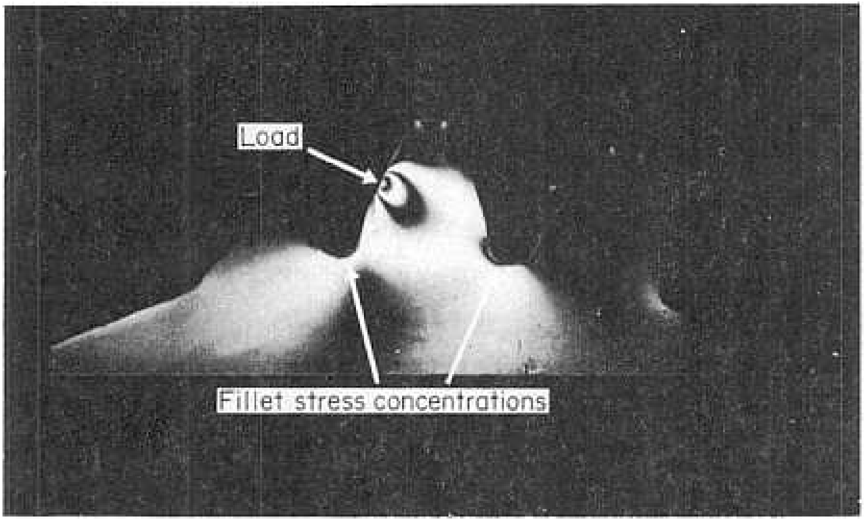


Fig. 10.31. Photoelastic fringe pattern of a portal frame showing stress concentration at the corner blend radii (different blend radii produce different stress concentration factors)



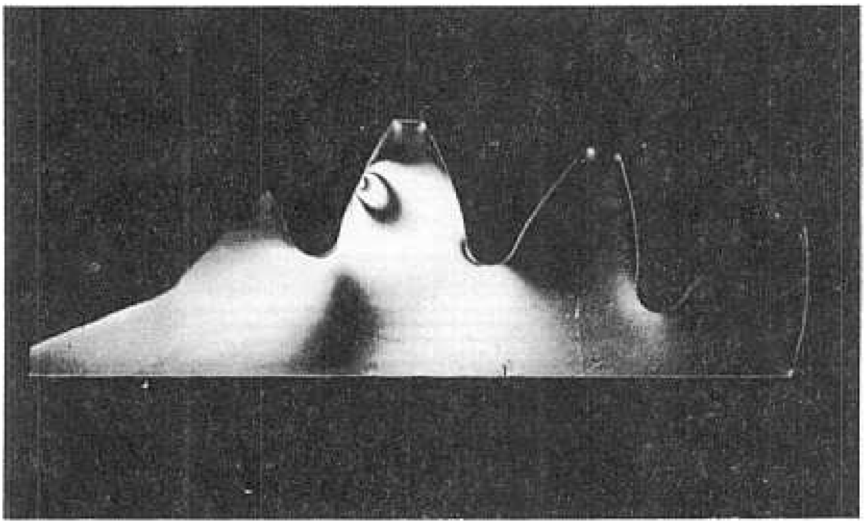


Fig. 10.32. Photoelastic fringe pattern of stress distribution in a gear tooth showing stress concentration at the loading point on the tooth flank and at the root fillet radii (higher concentration on the compressive fillet). Refer also to Fig. 10.45.

10.3.2. Saint-Venant's principle

The general problem of stress concentration was studied analytically by Saint-Venant who produced the following statement of principle: "If the forces acting on a small area of a body are replaced by a statically equivalent system of forces acting on the same area, there will be considerable changes in the local stress distribution but the effect at distances large compared with the area on which the forces act will be negligible". The effect of this principle is best demonstrated with reference to the photoelastic fringe pattern obtained in a model of a beam subjected to four-point bending, i.e. bending into a circular arc between the central

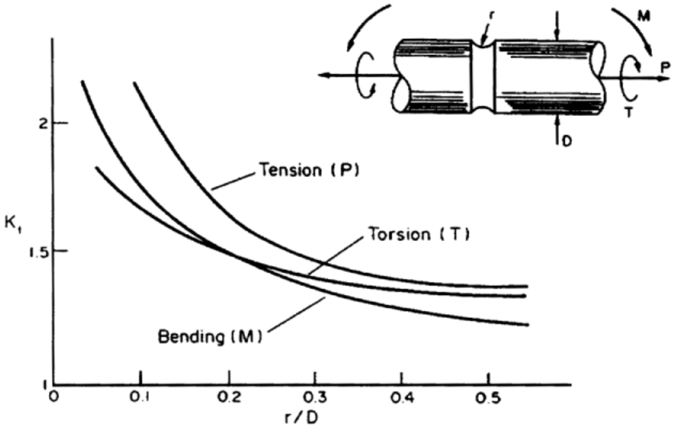


Fig. 10.33. Variation of stress concentration factors for a grooved shaft depending on the type of loading.

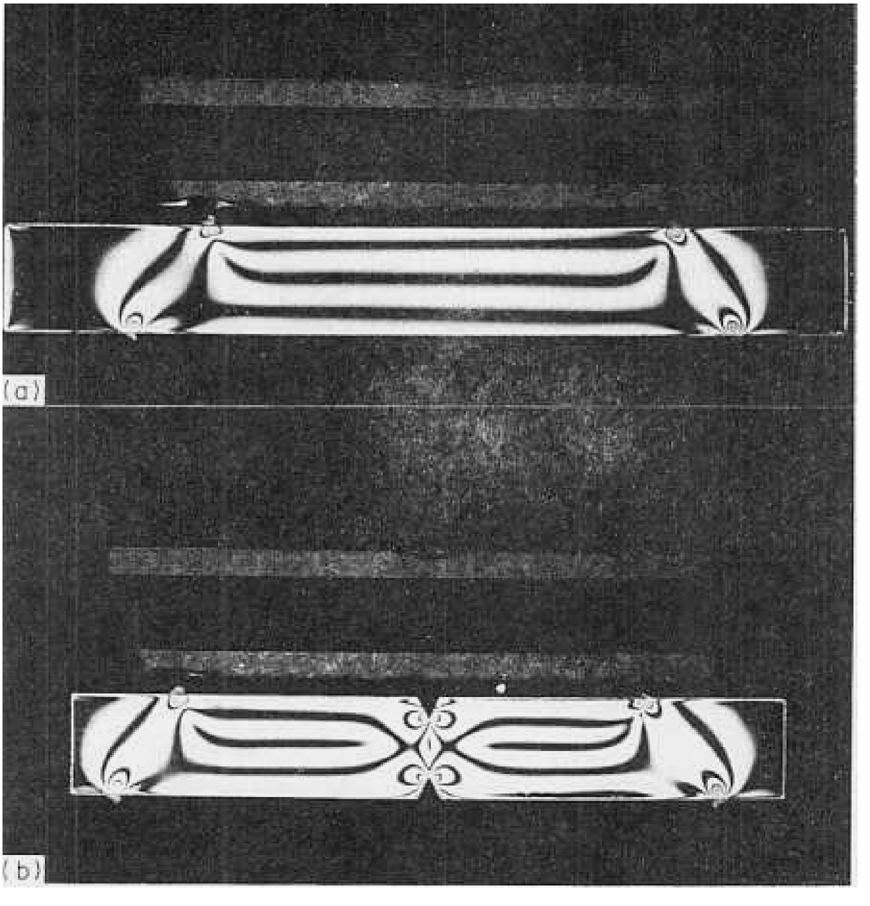


Fig. 10.34. (a) Photoelastic fringe pattern in a model of a beam subjected to four-point bending (i.e. circular arc bending between central supports): (b) as above but with a central notch.

supports – see Fig. 10.34(a). If the moment could have been applied by some other means so as to avoid the contact at the loading points then the fringe pattern would have been a series of parallel fringes, the centre one being the neutral axis. The stress concentrations due to the loading points are clearly visible as is the effect of these on the distribution of the fringes and hence stress. In particular, note the curvature of the neutral axis towards the inner loading points and the absence of the expected parallel fringe distribution both near to and outside the loading points. However, for points at least one depth of beam away from the stress concentrations (St. Venant) the fringe pattern is unaffected, the parallel fringes remain undisturbed and simple bending theory applies. If either the beam length is reduced or further stress concentrations (such as the notch of Fig. 10.34(b)) are introduced so that every part of the beam is within "one depth" of a stress concentration then at no point will simple theory apply and analysis of the fringe pattern is required for stress evaluation–there is no simple analytical procedure.

Similarly, in a round tension bar the stresses at the ends will be dependent upon the method of gripping or load application but within the main part of the bar, at least one diameter away from the loading point, stresses can again be obtained from simple theory. To the other extreme comes the case of a screw thread. The maximum s.c.f. arises at the first contacting thread at the plane of the bearing face of the head or nut and up to 70% of the load is carried by the first two or three threads. In such a case, simple theory cannot be applied anywhere within the component and the reader is referred to the appropriate B.S. Code of Practice and/or the work of Brown and Hickson[(59)].

10.3.3. Theoretical considerations of stress concentrations due to concentrated loads

A full treatment of the local stress distribution at points of application of concentrated load is beyond the scope of this text. Two particular cases will be introduced briefly, however, in order that the relevant useful equations can be presented.

(a) Concentrated load on the edge of an infinite plate

Work by St. Venant, Boussinesq and Flamant (see §8.7.9) has led to the development of a theory based upon the replacement of the concentrated load by a radial distribution of loads around a semi-circular groove (which replaces the local area of yielding beneath the concentrated load) (see Fig. 10.35). Elements in the material are then, according to Flamant, subjected to a radial compression of

$$\sigma_r = \frac{2P\cos\theta}{\pi br} \quad \text{with } b = \text{ width of plate} \tag{10.26}$$

This produces element cartesian stresses of:

$$\sigma_{xx} = \sigma_r \sin^2\theta = -\frac{2P\cos\theta\sin^2\theta}{\pi br} = -\frac{2Px^2y}{\pi b(x^2+y^2)^2} \tag{10.27}$$

$$\sigma_{yy} = \sigma_r \cos^2\theta = \frac{-2P\cos^3\theta}{\pi br} = \frac{-2Py^3}{\pi b(x^2+y^2)^2} \tag{10.28}$$

$$\tau_{xy} = \sigma_r \sin\theta\cos\theta = -\frac{2P\sin\theta\cos^2\theta}{\pi br} = -\frac{2Pxy^2}{b(x^2+y^2)^2} \tag{10.29}$$

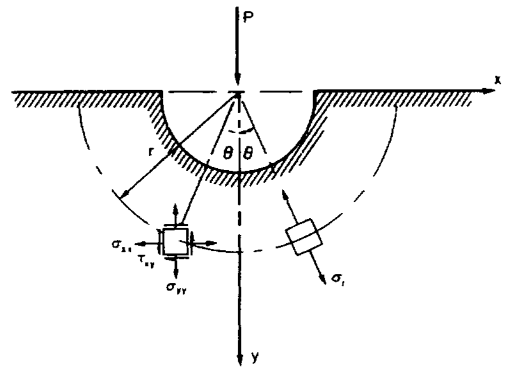


Fig. 10.35. Elemental stresses due to concentrated load P on the edge of an infinite plate.

(b) Concentrated load on the edge of a beam in bending

In this case a similar procedure is applied but, with a finite beam, consideration must be given to the horizontal forces set up within the groove which result in longitudinal stresses additional to the bending effects.

The total stress across the vertical section through the loading point (or groove) is then given by the so-called "*Wilson–Stokes equation*".

$$\sigma_{xx} = \frac{P}{\pi bd} \pm \left[\frac{L}{4} - \frac{d}{2\pi}\right]\frac{2Py}{bd^3} \tag{10.30}$$

where d is the depth of the beam, b the breadth and L the span.

This form of expression can be shown to indicate that the maximum longitudinal stresses set up are, in fact, less than those obtained from the simple bending theory alone (in the absence of the stress concentration).

10.3.4. Fatigue stress concentration factor

As noted above, the plastic flow which develops at positions of high stress concentration in ductile materials has a stress-relieving effect which significantly nullifies the effect of the stress raiser under static load conditions. Even under cyclic or fatigue loading there is a marked reduction in stress concentration effect and this is recognised by the use of a fatigue *stress concentration factor* K_f.

In the absence of any stress concentration (i.e. for $K_t = 1$) materials exhibit an "*endurance limit*" or "*fatigue limit*" – a defined stress amplitude below which the material can withstand an indefinitely large (sometimes infinite) number of repeated load cycles. This is often referred to as the un-notched fatigue limit – see Fig. 10.36.

For a totally brittle material in which the elastic stress concentration factor K_t might be assumed to have its full effect, e.g. $K_t = 2$, the fatigue life or notched endurance limit would be reduced accordingly. For materials with varying plastic flow capabilities, the effect of stress-raisers produces notched endurance limits somewhere between the un-notched value and that of the 'theoretical' value given by the full K_t – see Fig. 10.36, i.e. the fatigue stress concentration factor lies somewhere between the full K_t value and unity.

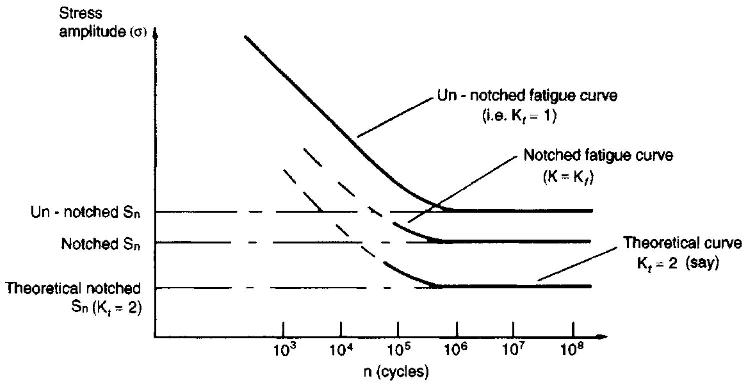


Fig. 10.36. Notched and un-notched fatigue curves.

If the endurance limit for a given number of cycles, n, is denoted by S_n then the fatigue stress concentration factor is defined as:

$$K_f = \frac{S_n \text{ for unnotched material}}{S_n \text{ for notched material}} \tag{10.31}$$

K_f is sometimes referred to by the alternative titles of "*fatigue strength reduction factor*" or, simply, the "*fatigue notch factor*".

The value of K_f is normally obtained from fatigue tests on identical specimens both with and without the notch or stress-raiser for which the stress concentration effect is required.

It is well known (and discussed in detail in Chapter 11) that the fatigue life of components is affected by a great number of variables such as mean stress, stress range, environment, size effect, surface condition, etc., and many different approaches have been proposed to allow realistic estimations of life under real working conditions as opposed to the controlled laboratory conditions under which most fatigue tests are carried out. One approach which is relevant to the present discussion is that proposed by Lipson & Juvinal[60] which utilises fatigue stress concentration factors, K_f, suitably modified by various coefficients to take account of the above-mentioned variables.

10.3.5. Notch sensitivity

A useful relationship between the elastic stress concentration factor K_t and the fatigue notch factor K_f introduces a *notch sensitivity* q defined as follows:

$$q = \frac{K_f - 1}{K_t - 1} \text{ or, in shear, } q = \frac{K_{fs} - 1}{K_{ts} - 1}$$

which may be re-written in terms of the fatigue notch factor as:

$$K_f = 1 + q(K_t - 1) \quad \text{with } 0 \leqslant q \leqslant 1 \tag{10.32}$$

It will be seen that, at the extreme values of q, eqn. (10.32) is valid since when $q = 1$ the full effect of the elastic stress concentration factor K_t applies and $K_f = K_t$; similarly when $q = 0$ and full ductility applies there is, in effect, no stress concentration and $K_f = 1$ with the material behaving in an unnotched fashion.

The value of the notch sensitivity for stress raisers with a significant linear dimension (e.g. fillet radius) R and a material constant "a" is given by:

$$q = \frac{1}{\left(1 + \frac{a}{R}\right)} \tag{10.33}$$

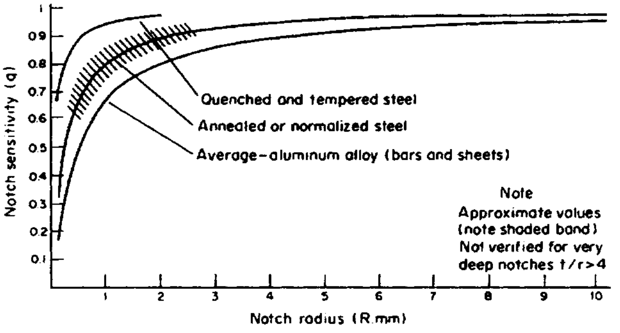


Fig. 10.37. Average fatigue notch sensitivity q for various notch radii and materials.

Typically, $a = 0.01$ for annealed or normalised steel, 0.0025 for quenched and tempered steel and 0.02 for aluminium alloy. However, values of "a" are not readily available for a wide range of materials and reference should be made to graphs of q versus R given by both Peterson(57) and Lipson and Juvinal(60).

The stress and strain distribution in a tensile bar containing a "through-hole" concentration are shown in Fig. 10.38 where the elastic stress concentration factor predictions are compared with those taking into account local yielding and associated stress redistribution.

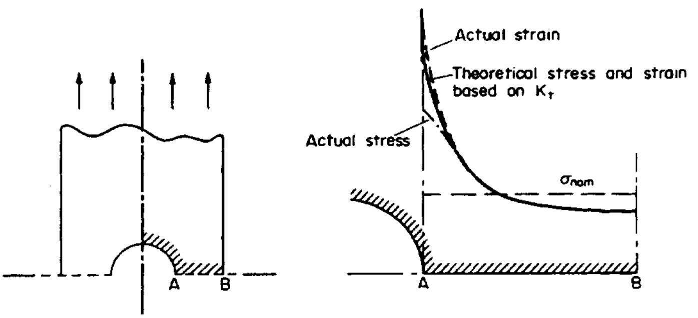


Fig. 10.38. Effect of a local yielding and associated stress re-distribution on the stress and strain concentration at the edge of a hole in a tensile bar.

10.3.6. Strain concentration – Neuber's rule

Within the elastic range, the concentration factor expressed in terms of strain rather than stress is equal to the stress concentration factor K_t. In the presence of plastic flow, however, the elastic stress concentration factor is reduced to the plastic factor K_p but local strains clearly exceed those predicted by elastic considerations – see Fig. 10.39.

A strain concentration factor can thus be defined as:

$$K_\varepsilon = \frac{\text{maximum strain at the notch}}{\text{nominal strain at the notch}}$$

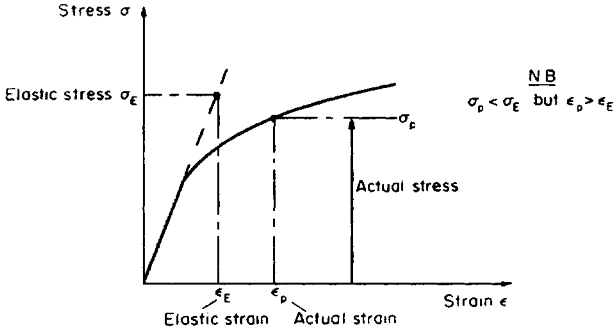


Fig. 10.39. Comparison of elastic and plastic stresses and strains.

the value of K_ε increasing as the value of K_p decreases. One attempt to relate the two factors is known as "*Neuber's Rule*", viz.

$$K_pK_\varepsilon = K_t^2 \tag{10.34}$$

It is appropriate here to observe that recent research in the fatigue behaviour of materials indicates that the strain range of fatigue loading may be more readily related to fatigue life than the stress range which formed the basis of much early fatigue study. This is said to be particularly true of low-cycle fatigue where, in particular, the plastic strain range is shown to be critical.

10.3.7. Designing to reduce stress concentrations

From the foregoing discussion it should now be evident that stress concentrations are critical to the life of engineering components and that fatigue failures, for example, almost invariably originate at such positions. It is essential, therefore, for any design to be successful that detailed consideration is given to the reduction of stress concentration effects to an absolute minimum.

One important rule in this respect is concerned with the initial placement of the stress concentration. Assuming that some freedom exists as to the position of e.g. oil-holes, keyways, grooves, etc., then it is essential that these be located at positions where the nominal stress is as low as possible. The resultant magnitude of stress concentration factor × nominal stress is then also a minimum for a particular geometry of stress raiser.

In situations where no flexibility exists as to the position of the stress raiser then one of the procedures outlined below should be considered. In many cases a qualitative assessment of the benefits, or otherwise, of design changes is readily obtained by sketching the lines of stress flow through the component as in Fig. 10.17. Sharp changes in flow direction indicate high stress concentration factors, smooth changes in flow direction are the optimum solution.

The following standard stress concentration situations are common in engineering applications and procedures for reduction of the associated stress concentration factors are introduced for each case. The procedures, either individually or in combination, can then often be applied to produce beneficial stress reduction in other non-standard design situations.

(a) Fillet radius

Probably the most common form of stress concentration is that arising at the junction of two parts of a component of different shape, diameter, or other dimension. In almost every shaft, spindle, or axle design, for example, the component consists of a number of different diameter sections connected by shoulders and associated fillets.

If Fig. 10.40(a) is taken to be either the longitudinal section of a shaft or simply a flat plate, then the transition from one dimension to another via the right-angle junction is exceptionally bad design since the stress concentration associated with the sharp corner is exceedingly high. In practice, however, either naturally due to the fact that the machining tool has a finite radius, or by design, the junction is formed via a fillet radius and the wise designer employs the highest possible radius of fillet consistent with the function of the component in order to keep the s.c.f. as low as possible. Whilst, historically, circular arcs have generally been used for fillets, other types of blend geometry have been shown to produce even further reduction of s.c.f. notably elliptical and streamline fillets[61], the latter following similar contours to those of a fluid when it flows out of a hole in the bottom of a tank. Fig. 10.41 shows the effect of elliptical fillets on the s.c.f. values.

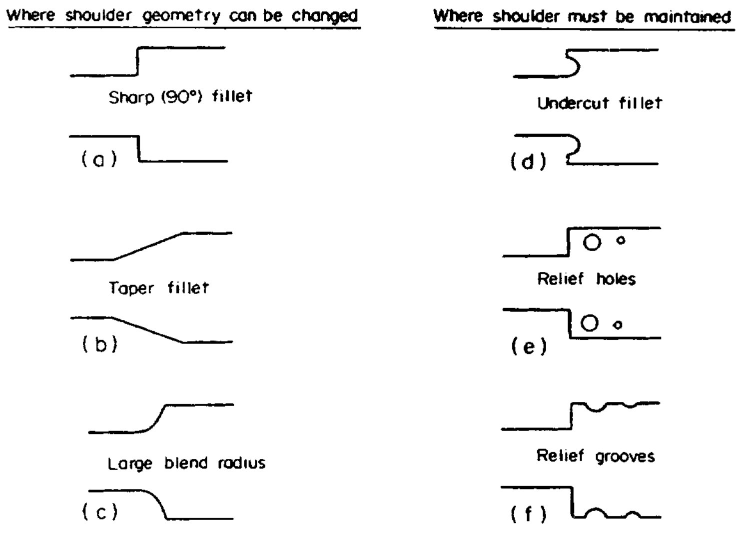


Fig. 10.40. Various methods for reduction of stress concentration factor at the junction of two parts of a component of different depth/diameter.

There are occasions, however, where the perpendicular faces at the junction need to be maintained and only a relatively small fillet radius can be allowed e.g. for retention of bearings or wheel hubs. A number of alternative solutions for reduction of the s.c.f's are shown in Fig. 10.40(d) to (f) and Fig. 10.42.

(b) Keyways or splines

It is common to use keyways or splines in shaft applications to provide transfer of torque between components. Gears or pulleys are commonly keyed to shafts, for example, by square

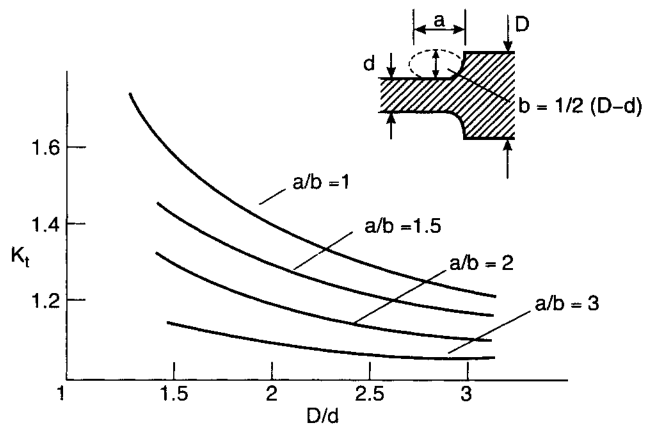


Fig. 10.41. Variation of elliptical fillet stress concentration factor with ellipse geometry.

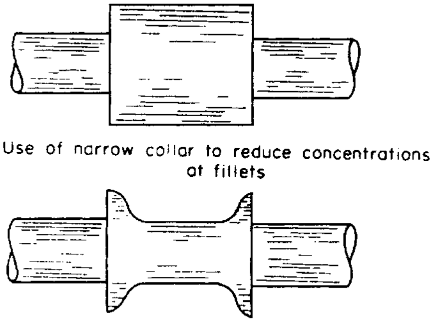


Fig. 10.42. Use of narrow collar to reduce stress concentration at fillet radii in shafts.

keys with side dimensions approximately equal to one-quarter of the shaft diameter with the depth of the keyway, therefore, one-eighth of the shaft diameter.

Analytical solutions for such a case have been carried out by both Leven[63] and Neuber[65] each considering the keyway without a key present. Neuber gives the following formula for stress concentration factor (based on shear stresses):

$$K_{t_s} = 1 + \sqrt{\frac{h}{r}} \tag{10.35}$$

where h = keyway depth and r = radius at the base of the groove or keyway (see Fig. 10.43). For a semi-circular groove $K_{t_s} = 2$.

Leven, considering the square keyway specifically, observes that the s.c.f. is a function of the keyway corner radius and the shaft diameter. For a practical corner radius of about one-tenth the keyway depth $K_{t_s} \simeq 3$.

If fillet radii cannot be reduced then s.c.f.'s can be reduced by drilling holes adjacent to the keyway as shown in Fig. 10.43(b).

The presence of a key and its associated fit (or lack of) has a significant effect on the stress distribution and no general solution exists. Each situation strictly requires its own solution via practical testing such as photoelasticity.

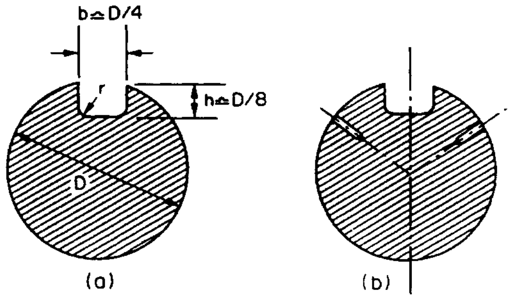


Fig. 10.43. Key-way dimensions and stress reduction procedure.

(c) Grooves and notches

Circumferential grooves or notches (particularly U-shaped notches) occur frequently in engineering design in such applications as C-ring retainer grooves, oil grooves, shoulder or grinding relief grooves, seal retainers, etc; even threads may be considered as multi-groove applications.

Most of the available s.c.f. data available for grooves or notches refers to U-shaped grooves and circular fillet radii and covers both plane stress and three-dimensional situations such as shafts with circumferential grooves. In general, the higher the blend radius, the lower the s.c.f; the optimum value being $K_t = 2$ for a semi-circular groove as calculated by Neuber's equation (10.35) above.

Some data exists for other forms of groove such as V notches and hyperbolic fillets but, particularly in bending and tension, the latter have little advantage over circular arcs and V notches only show significant advantage for included angles greater than 120°. In cases where s.c.f. data for a particular geometry of notch are not readily available recourse can be made to standard factor data for plates with a central hole.

Stress concentrations at notches and grooves can be reduced by the "metal removal – stiffness reduction" technique utilising any procedure which improves the stress flow, e.g. multiple notches of U grooves or selected hole drilling as shown in Fig. 10.44. Reductions of the order of 30% can be obtained.

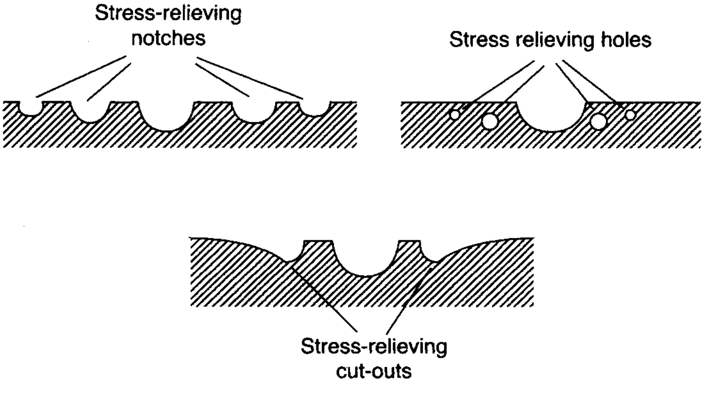


Fig. 10.44. Various procedures for the reduction of stress concentrations at notches or grooves.

This procedure of introducing secondary stress concentrations deliberately to reduce the local stiffness of the material adjacent to a stress concentration is a very powerful stress reduction technique. In effect, it causes more of the stiffer central region of the component to carry the load and persuades the stress lines to follow a path removed from the effect of the single, sharp concentration. Figures 10.40(d) to (f), 10.42 and 10.43 are all examples of the application of this technique, sometimes referred to as an "interference effect" the individual concentrations interfering with each other to mutual advantage.

(d) Gear teeth

The full analysis of the stress distribution in gear teeth is a highly complex problem. The reader is only referred in this section to the stress concentrations associated with the fillet radii at the base of the teeth – see Fig. 10.45.

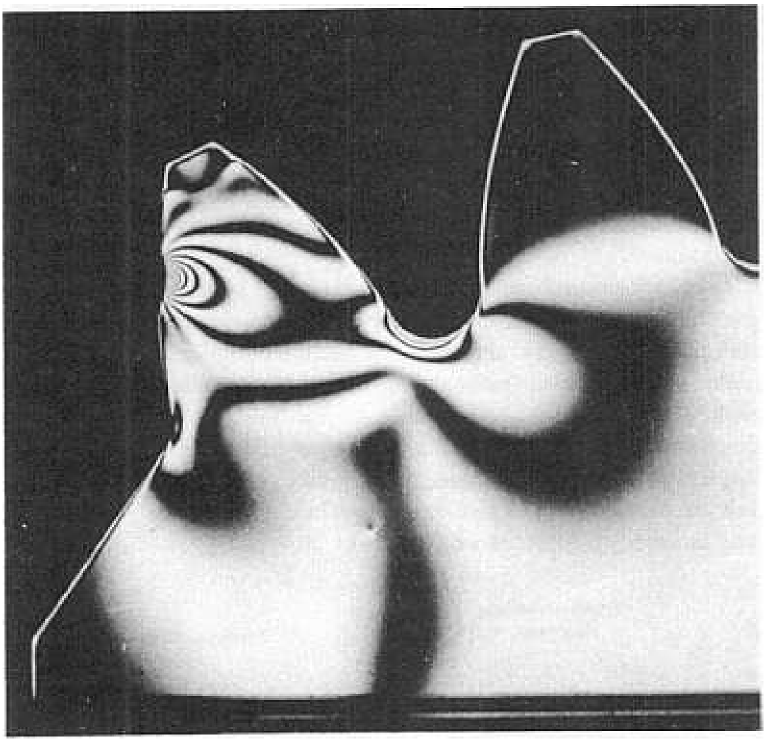


Fig. 10.45. Stress concentration at root fillet of gear tooth.

The loading on the tooth produces both direct stress and bending components on the root section and Dolan and Broghammer[68] in early studies of the problem gave the following formula for the combined stress concentration effect (for 20° pressure angle gears)

$$K_t = 0.18 + \frac{1}{\left(\dfrac{r}{t}\right)^{0.15}\left(\dfrac{h}{t}\right)^{0.45}}$$

Later work by Jacobson[69], again for 20° pressure angle gears, produced a series of charts of strength factors and more recently Hearn[66,67] has carried out photoelastic studies of both two-dimensional involute tooth forms and three-dimensional helical gears which introduce new considerations of stress concentration factors, notably their variation in both magnitude and position as the load moves up and down the tooth flank.

(e) Holes

From much of the previous discussion it should now be evident that holes represent very significant stress raisers, be they in two-dimensional plates or three-dimensional bars. Fortunately, a correspondingly high amount of information and data is available, e.g. Peterson[57], covering almost every foreseeable geometry and loading situation. This includes not only individual holes but rows and groups of holes, pin-joints, internally pressurised holes and intersecting holes.

(f) Oil holes

The use of transverse and longitudinal holes as passages for lubricating oil is common in shafting, gearing, gear couplings and other dynamic mechanisms. Occasionally similar holes are also used for the passage of cooling fluids.

In the case of circular shafts, no problem arises when longitudinal holes are bored through the centre of the shaft since the nominal torsional stress at this location is very small and the effect on the overall strength of the shaft is minimal. A transverse hole, however, is a significant source of stress concentration in any mode of loading, i.e. bending, torsion or axial load, and the relevant s.c.f. values must be evaluated from standard reference texts[57, 60]. Whatever the type of loading, the value of K_t increases as the size of the hole increases for a given shaft diameter, with minimum values for very small holes of 2 for torsion and 3 for bending and tension.

In cases of combined loading, a conservative estimate[58] of the stress concentration may be obtained from values of K_t given by either Peterson[57] or Lipson[60] for an infinite plate containing a transverse hole and subjected to an equivalent biaxial stress condition.

One procedure for the reduction of the stress concentration at the point where transverse holes cut the surface of shafts is shown in Fig. 10.46.

(g) Screw threads

Again the stress distribution in screw threads is extremely complex, values of the stress concentration factors associated with each thread being dependent upon the tooth form, the fit between the nut and the bolt, the nut geometry, the presence or not of a bolt shank and the load system applied. Pre-tensioning also has a considerable effect. However, from numerous photoelastic studies carried out by the author and others[59, 61, 62] it is clear that the greatest stress most often occurs at the first mating thread, generally at the mating face of the head of the nut with the bearing surface, with practically all the load shared between the first few threads. (One estimate of the source of bolt failures shows 65% in the thread at the nut face compared with 20% at the end of the thread and 15% directly under the bolt

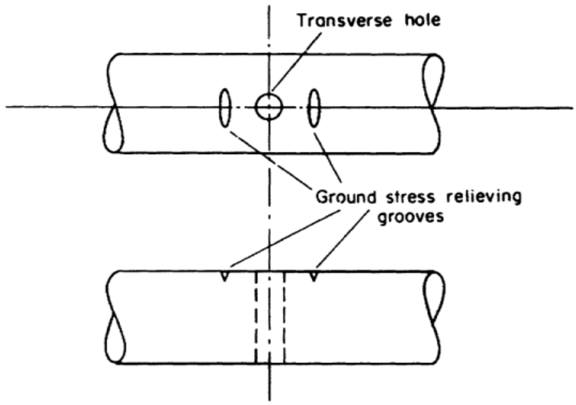


Fig. 10.46. Procedure for reduction of stress concentration at exit points of transverse holes in shafts.

head). Alternative designs of nut geometry can be introduced to spread the load distribution a little more evenly as shown in Figs. 10.47 and tapering of the thread is a very effective load-distribution mechanism.

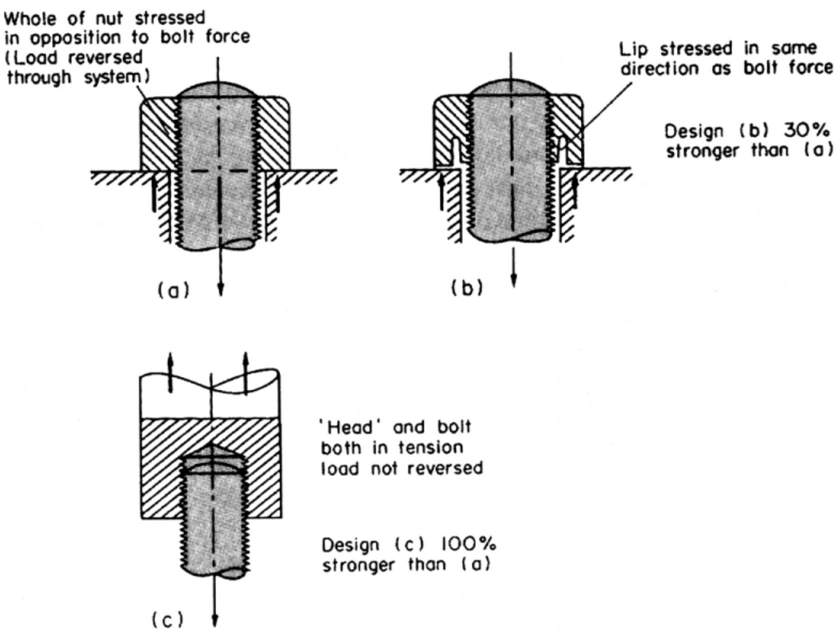


Fig. 10.47. Alternative bolt/nut designs for reduction of stress concentrations.

Reduction in diameter of the bolt shank and a correspondingly larger fillet radius under the bolt head also produces a substantial improvement as does the use of a material with a lower modulus of elasticity for the nut compared with the bolt; fatigue tests have shown strength improvements of between 35 and 60% for this technique.

Stress concentration data for various nut and bolt configurations are given by Hetenyi[62], again based on photoelastic studies. As an example of the severity of loading at the first thread, stress concentration factors of the order of 13 are readily obtained in conventional nut designs and even using the modified designs noted above s.c.f.'s of up to 9 are quite common. It is not perhaps surprising, therefore, that one of the most common causes of machinery or plant failure is that of stud or bolt fracture.

(h) Press or shrink fit members

There are some applications where discontinuity of component profile caused by two contacting members represents a substantial stress raiser effectively as great as a right-angle fillet. These include shrink or press-fit applications such as collars, gears, wheels, pulleys, etc., mounted on their drive shafts and even simple compressive loading of rectangular faces on wider support plates – see Fig. 10.48(a).

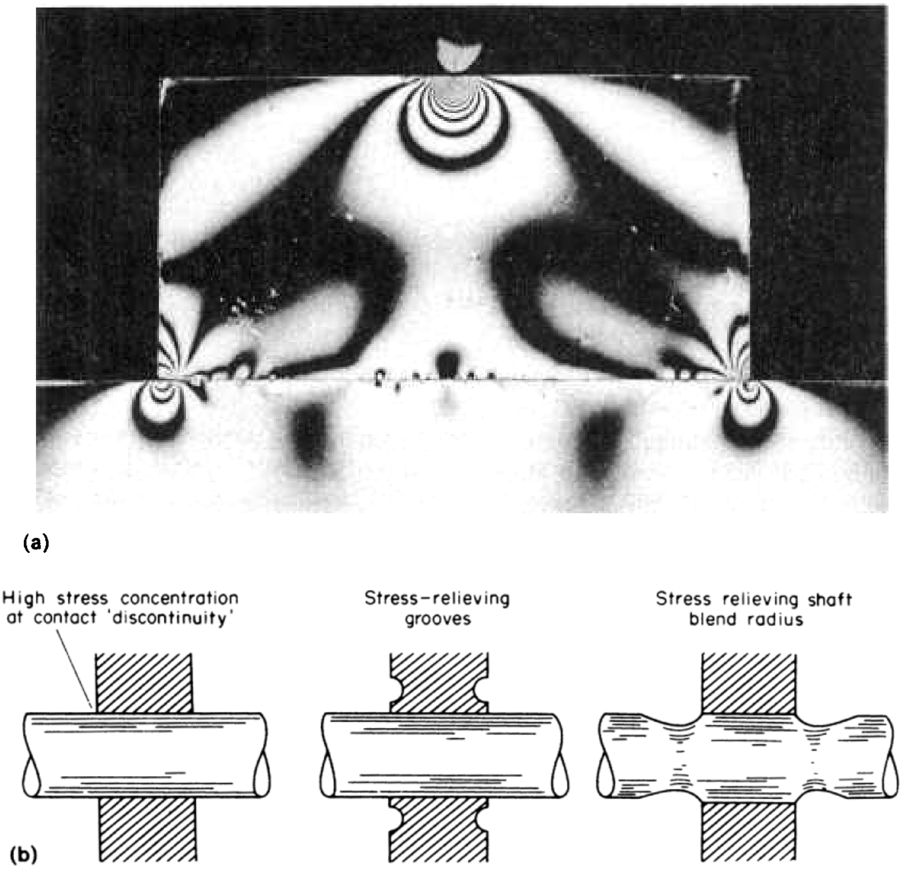


Fig. 10.48. (a) Photoelastic fringe pattern showing stress concentrations produced at contact discontinuities such as the loading of rectangular plates on a flat surface (equivalent to cross-section of cylindrical roller bearing on its support surface); (b) the reduction of stress concentration at press and shrink fits.

Significant stress concentration reductions can be obtained by introducing stress-relieving grooves or a blending fillet (or taper) in the press-fit member or the shaft – see Fig. 10.48.

10.3.8. Use of stress concentration factors with yield criteria

Whilst stress concentration factors are defined in terms of the maximum individual stress at the stress raiser it could be argued that, since stress conditions there are normally biaxial, it would be more appropriate to express them in terms of some "equivalent stress" employing one of the yield criteria introduced in Chapter 15.†

Since the maximum shear strain energy (distortion energy) theory of Von Mises is usually considered to be the most applicable to both static and dynamic conditions in ductile materials then, for a biaxial state the Von Mises equivalent stress can be defined as:

$$\sigma_e = \sqrt{\sigma_1^2 - \sigma_1\sigma_2 + \sigma_2^2} \tag{10.36}$$

and, since there is always a direct relationship between σ_1 and σ_2 within the elastic range for biaxial states (i.e. $\sigma_1 = k\sigma_2$) then

$$\sigma_e = \left[\sigma_1^2 - \frac{\sigma_1^2}{k} - \frac{\sigma_1^2}{k^2}\right]^{1/2}$$

$$= \sigma_1\left[1 - \left(\frac{1}{k}\right) - \left(\frac{1}{k^2}\right)\right]^{1/2}$$

Then the stress concentration factor expressed in terms of this equivalent stress will be

$$K_e = \frac{\sigma_e}{\sigma_{nom}} = \frac{\sigma_1}{\sigma_{nom}}\left[1 - \left(\frac{1}{k}\right) - \left(\frac{1}{k^2}\right)\right]^{1/2} \tag{10.37}$$

Except for the special case of equal bi-axial stress conditions when $\sigma_1 = \sigma_2$ and $K = 1$ the value of K_e is always less than K_t.

A full treatment of the design procedures to be adopted for both ductile and brittle materials incorporating both yield criteria (Von Mises and Mohr) and stress concentration factors is carried out by Peterson(57) with consideration of static, alternating and combined static and alternating stress conditions.

10.3.9. Design procedure

The following procedure should be adopted for the design of components in order that the effect of stress concentration is minimised and for the component to operate safely and reliably throughout its intended service life.

(1) Prepare a draft design incorporating the principal features and requirements of the component. The dimensions at this stage will be obtained with reference to the nominal stresses calculated on the basis of known or estimated service loads.

(2) Identify the potential stress concentration locations.

† E.J. Hearn, *Mechanics of Materials 1*, Butterworth-Heinemann, 1997.

(3) Undertake the procedures outlined in §10.3.7 to reduce the stress concentration factors at these locations by:
  (a) streamlining the design where possible to avoid sharp changes in geometry and producing gradual fillet transitions between adjacent parts of different shape and size.
  (b) If fillet changes cannot be effected owing to design constraints, of e.g. bearing surfaces, undertake other modifications to the design to produce smoother "flow" of the stresses through the component.
  (c) Where appropriate, reduce the stiffness of the material adjacent to the stress concentration positions to allow greater flexibility and a reduction in the associated stress concentration factor. This is probably best achieved by removal of material as discussed earlier.

(4) Evaluate the stress concentration factors for the modified design using standard tables[57, 60] or experimental test procedures such as photoelasticity. Depending on the material and the loading conditions either K_t or K_f may be appropriate.

(5) Ensure that the maximum stress in the component taking into account both the stress concentration factors and an additional safety factor to account for service uncertainties, does not exceed the safe working stress for the material concerned.

References

1. Hertz, H. "On the contact of elastic solids", *J. reine angew Math.* (1881), **92**, 56.
2. Belajef, N. M. "On the problem of contact stresses", *Bull. Eng. Ways of Communication"*, St. Petersburg, 1917.
3. Foppl, L. "Zertschrift fur Angew", *Math und Mech.* (1936), **16**, 165.
4. M'Ewen, E. "The load-carrying capacity of the oil film between gear teeth". *The Engineer*, (1948), **186**, 234–235.
5. Radzimovsky, E. I. "Stress distribution of two rolling cylinders pressed together," *Bulletin* 408, Eng. Exp. Sta. Univ. Ill., 1953.
6. Buckingham, E. "Dynamic loads on gear teeth", ASME Special Committee on Strength of Gear Teeth. New York, 1931.
7. Walker, H. "A laboratory testing machine for helical gear tooth action", *The Engineer*, (6 June 1947), **183**, 486–488.
8. Wellauer, E. J. "Surface durability of helical and herringbone gears", *Machine Design*, May 1964.
9. Huber, M. T. *Ann. Phys. Paris* (1904), **14**, 153.
10. Morton W. B. and Close, L. J. "Notes on Hertz's theory of the contact of elastic bodies" *Phil. Mag.* (1922), 6th series, **43**, 320.
11. Thomas H. R. and Hoersch, V. A. "Stresses due to the pressure of one elastic solid on another" *Univ. Illinois Eng. Exp. Sta. Bull.* (1930), **212**.
12. Fessler, H. and Ollerton, E. "Contact stresses in toroids under radial loads", *Brit. J. Appl. Phys.* (1957), **8**, 387.
13. Ollerton, C. "Stresses in the contact zone". Convention on Adhesion. I. Mech. E., London, 1964.
14. Johnson, K. L. "A review of the theory of rolling contact stresses", O.E.C.D. Sub-group on Rolling Wear. Delft, April, 1965. *Wear* (1966), **9**, 4–19.
15. Deresiewicz, H. "Contact of elastic spheres under an oscillating torsional couple". *Ibid.* (1954), **76**, 52.
16. Johnson, K. L. "Plastic contact stresses" BSSM Conference. Sub-surface stresses." Nov. 1970, unpublished.
17. Lubkin, J. L. "The torsion of elastic spheres in contact" *J. Appl. Mech. Trans ASME* (1951), **73**, 183.
18. Mindlin, R. D. and Deresiewicz, H. "Elastic spheres in contact under varying oblique forces", *J. Appl. Mech. Trans ASME* (1953), **75**, 327.
19. Tomlinson, G. A. and Gough, H. J. "An investigation of the fretting corrosion of closely fitting surfaces". *Proc. I Mech. E.* (1939), **141**, 223.
20. Smith, J. O. and Liu, C. K. "Stresses due to tangential and normal loads on an elastic solid with application to some contact stress problems", *J. App. Mech.*, June 1953.
21. Lipson, C. and Juvinal, R. C. *Handbook of Stress and Strength.* Macmillan, New York, 1963.
22. Meldahl, A. "Contribution to the Theory of Lubrication of Gears and of the Stressing of the Lubricated Flanks of Gear Teeth" Brown Boveri Review. Vol. 28, No. 11, Nov. 1941.

23. Dowson, D., Higginson, G. R. and Whitaker, A. V. "Stress distribution in lubricated rolling contacts", *I. Mech. E. Fatigue in Rolling Contact* (1964).
24. Crook, A. W. "Lubrication of Rollers". Parts I–IV. *Phil. Trans. Roy. Soc.*-Ser. A250 (1958); *Ibid.*, 254 (1962) and 255 (1963).
25. Dawson, P. H. "Rolling contact fatigue crack initiation in a 0.3% carbon steel", *Proc. I. Mech. E.* (1968/69), **183**, Part 1.
26. Scott, D. "The effect of materials properties, lubricant and environment in rolling contact fatigue", *I. Mech. E. Fatigue in Rolling Contact* (1964).
27. Sherratt, F. "The influence of shot-peening and similar surface treatments on the fatigue properties of metals", B.S.S.M. Sub-Surface Stresses, 1970 (unpublished).
28. Muro, H. "Changes of residual stress due to rolling contacts", *J. Soc. Mat. Sci. Japan* (July 1969), **18**, 615–619.
29. Timoshenko, S. P. and Goodier, J. N. *Theory of Elasticity*, 2nd edn. McGraw-Hill, 1951.
30. Almen, J. O. "Surface deterioration of gear teeth", *Trans. A.S.M.* (Jan. 1950).
31. Akaoka, J. "Some considerations relating to plastic deformation under rolling contact, *Rolling Contact Phenomena*. Elsevier.
32. Hearn E. J. "A three-dimensional photoelastic analysis of the stress distribution in double helical epicyclic gears", *Strain* (July 1979).
33. Roark, R. J. and Young, W. C. *"Formulas for Stress and Strain"*, 5th edn., McGraw-Hill, 19
34. Heindlhofer, K. *Evaluation of Residual Stresses*. McGraw-Hill, 1948.
35. Rosenthal, and Norton, "A method of measuring tri-axial residual stresses in plates", *Journal Welding Society*, (Welding supplement, 1945), **24**, 295–307.
36. Waisman, and Phillips, "Simplified measurement of residual stresses", *S.E.S.A.*, **II** (2), 29–44.
37. Mathar, J. "Determination of initial stresses by measuring the deformations around drilled holes", *Trans A.S.M.E.*, **56** (4), (1934), 249–254.
38. Bathgate, R. G. "Measurement of non-uniform bi-axial residual stresses by the hole drilling method", *Strain*. **4** (2) (1968), 20–27.
39. Beaney, E. M. "The Air Abrasive Centre-Hole Technique for the Measurement of Residual Stresses". B.S.S.M. Int'l Conference: "Product Liability and Reliability", September 1980 (Birmingham).
40. Procter, E. "An Introduction to Residual Stresses, their Measurement and Reliability Aspects". B.S.S.M. Int'l Conference: "Product Liability and Reliability", September 1980 (Birmingham).
41. Beaney, E. M. and Procter, E. "A critical evaluation of the centre-hole technique for the measurement of residual stresses". C.E.G.B. report (UK), RD/B/N2492, Nov. 1972.
42. Milbradt, K. P. "Ring method determination of residual stresses" *S.E.S.A.*, **9** (1), 63–74.
43. Durelli, A. J. and Tsao, C. H. "Quantitative evaluation of residual stresses by the stresscoat drilling technique" *S.E.S.A.*, **9** (1), 195–207.
44. French, D. N. and Macdonald, B. A. "Experimental methods of X-ray analysis", *S.E.S.A.*, **26** (2), 456–462.
45. Kirk, D. "Theoretical aspects of residual stress measurement by X-ray diffractometry", *Strain*, **6** (2) (1970).
46. Kirk, D. "Experimental features of residual stress measurement by X-ray diffractometry", *Strain*, **7** (1) (1971).
47. Andrews, K. W. *et al.* "Stress measurement by X-ray diffraction using film techniques", *Strain*, **10** (3) (July 1974), 111–116.
48. Abuki, S. and Cullity, B. D. "A magnetic method for the determination of residual stresses", *S.E.S.A.*, **28** (1), 217–223.
49. Sines, G. and Carlson, R. "Hardness measurements for the determination of residual stresses", *A.S.T.M. Bull.*, no. 180 (February 1952), pp. 35–37.
50. Noranha, P. J. and Wert, J. J. "An ultrasonic technique for the measurement of residual stress", *Journal Testing and Evaluation*, **3** (2) (1975), 147–172.
51. Kino, G. S. *et al.* "Acoustic techniques for measuring stress regions in materials", Electric Power Research Institute NP-1043 Project 609-1. Interim Report Apr. 1979.
52. Hearn, E. J. and Golsby, R. J. "Residual stress investigation of a noryl telecommunication co-ordinate selector switch", *S.E.S.A.*, 3rd Int. Congress Experimental Mechanics, Los Angeles, 1973.
53. Denton, A. A. "The use of spark machining in the determination of local residual stress" *Strain*, **3** (3) (July 1967), 19–24.
54. Moore, M. G. and Evans, W. D. "Mathematical correction for stress in removed layers in X-ray diffraction residual stress analysis", *S.A.E. Trans.*, **66** (1958), 340–345.
55. Scaramangas, A. A. *et al.* "On the correction of residual stress measurements obtained using the centre-hole method", *Strain*, **18** (3) (August 1982), 88–96.
56. Huang, T. C. *Bibliography on Residual Stresses*. Soc. Auto. Eng., New York, 1954.
57. Peterson, R. E. *Stress Concentration Design Factors*, John Wiley & Sons, New York, 1953.
58. Perry, C. C. "Stress and strain concentration", Vishay Lecture and Series, Vishay Research and Education, Michigan, U.S.A.
59. Brown, A. F. C. and Hickson, V. M. "A photoelastic study of stresses in screw threads", *Proc. I. Mech. E.*, **1B**, 1952–3.
60. Lipson, C. and Juvinal, R. C. *Handbook of Stress and Strength*. Macmillan, New York, 1963.

61. Heywood, R. B. *Photoelasticity for Designers*. Pergamon, Oxford, 1969.
62. Hetenyi, M. "A photoelastic study of bolt and nut fastenings", *Trans. A.S.M.E. J. Appl. Mechs*, **65** (1943).
63. Leven, M. M. "Stresses in keyways by photoelastic methods and comparison with numerical solution", *Proc. S.E.S.A.*, **7** (2) (1949).
64. Roark, R. J. *Formulas for Stress and Strain*, 5th edition. McGraw-Hill, New York, 1975.
65. Neuber, N. *Theory of Notch Stresses*. Edwards, Michigan, 1946.
66. Hearn, E. J. "A new look at the bending strength of gear teeth", *Experimental Mechanics S.E.S.A.*, October 1980.
67. Hearn, E. J. "A new design procedure for helical gears", *Engineering*, October 1978.
68. Dolan and Broghamer
69. Jacobson, M. A. "Bending stresses in spur gear teeth: proposed new design factors. ...", *Proc. I. Mech. E.*, 1955.

Examples

Example 10.1

(a) Two parallel steel cylinders of radii 50mm and 100mm are brought into contact under a load of 2 kN. If the cylinders have a common length of 150 mm and elastic constants of $E = 208$ GN/m^3 and $\nu = 0.3$ determine the value of the maximum contact pressure. What will then be the magnitude and position of the maximum shear stress?

(b) How would the values change if the larger cylinder were replaced by a flat surface?

Solution (a)

For contacting parallel cylinders eqn. (10.9) gives the value of the maximum contact pressure (or compressive stress) as

$$\sigma_c = -p_0 = -0.591\sqrt{\frac{P}{L\Delta}\left(\frac{1}{R_1} + \frac{1}{R_2}\right)}$$

where

$$\begin{aligned}\Delta &= \frac{1}{E_1}[1 - \nu_1^2] + \frac{1}{E_2}[1 - \nu_2^2] \\ &= \frac{2}{E}[1 - \nu^2] \text{ for similar materials} \\ &= \frac{2 \times 0.91}{208 \times 10^9}\end{aligned}$$

$\therefore$ Max. contact pressure

$$\begin{aligned}p_0 &= 0.591\sqrt{\frac{2 \times 10^3 \times 208 \times 10^9}{150 \times 10^{-3} \times 2 \times 0.91}\left(\frac{1}{50} + \frac{1}{100}\right)10^3} \\ &= 0.591 \times 21.38 \times 10^7 \\ &= \mathbf{126.4\ MN/m^2}\end{aligned}$$

Maximum shear stress $= 0.295 p_0 = \mathbf{37.3\ MN/m^2}$

occurring at a depth $\mathbf{d = 0.786b}$

with (from eqn. (10.8)) $$b = 1.076\sqrt{\frac{P\Delta}{L\left(\frac{1}{R_1}+\frac{1}{R_2}\right)}}$$

$$= 1.076\sqrt{\frac{2\times 10^3\times 2\times 0.91}{150\times 10^{-3}\times 208\times 10^9\times 30}}$$

$$= 1.076\times 0.624\times 10^{-4}$$

$$= 0.067 \text{ mm}$$

$\therefore$ Depth of max shear stress $= 0.786\times 0.067 = \mathbf{0.053\ mm}$

(*b*) Replacing the 100 mm cylinder by a flat surface makes $\dfrac{1}{R_2} = 0$ and

$$\text{contact pressure } p_0 = 0.591\sqrt{\frac{2\times 10^3\times 208\times 10^9}{150\times 10^{-3}\times 2\times 0.91}\left(\frac{1}{50}\right)10^3}$$

$$= 0.591\times 17.48\times 10^7$$

$$= \mathbf{103.2\ MN/m^2}$$

with max shear stress $= 0.295\times 103.2 = \mathbf{30.4\ MN/n^2}$

and $b = 0.082$ mm

$\therefore$ Depth of max shear stress $= 0.786\times 0.082 = \mathbf{0.064\ mm.}$

Example 10.2

(a) What will be the maximum compressive stress set up when two spur gears transmit a torque of 250 N m? One gear has 150 teeth on a pitch circle diameter of 200 mm whilst the second gear has 200 teeth. Both gears have a common face-width of 200 mm. Assume $E = 208$ GN/m^2 and $\nu = 0.3$ for both gears.

(b) How will this value change if the spur gears are replaced by helical gears of $17\frac{1}{2}^\circ$ pressure angle and 30° helix?

Solution (a)

(*a*) From eqn. (10.21) the maximum compressive stress at contact is

$$\sigma_c = -p_0 = -0.475\sqrt{K}$$

with $$K = \frac{W}{F_w d}\left[\frac{m+1}{m}\right]$$

$$= \frac{250}{100\times 10^{-3}}\times\frac{1}{200\times 10^{-3}\times 200\times 10^{-3}}\left[\frac{\frac{200}{150}+1}{\frac{200}{150}}\right]$$

$$= 109375$$

$$\therefore \qquad \sigma_c = -0.475\sqrt{109375} = \mathbf{-157.1\ MN/m^2}$$

the negative sign indicating a compressive stress value.

Solution (b)

For the helical gears, eqn. (10.22) gives

$$\sigma_c = -p_0 = C\sqrt{\frac{K}{m_p}}$$

and for the given pressure angle and helix values Table 10.2 gives values of

$$C = 0.435 \text{ and } m_p = 1.53$$

$$\therefore \sigma_c = -0.435\sqrt{\frac{109375}{1.52}} = \mathbf{-116.3\ MN/m^2}$$

Example 10.3

A rectangular bar with shoulder fillet is subjected to a uniform bending moment of 100 Nm. Its dimensions are as follows (see Fig. 10.22) $D = 50$ mm; $d = 25$ mm; $r = 2.5$ mm; $h = 10$ mm.

Determine the maximum stress present in the bar for static load conditions. How would the value change if (a) the moment were replaced by a tensile load of 20 kN, (b) the moment and the tensile load are applied together.

Solution

For applied moment

From simple bending theory, nominal stress (related to smaller part of the bar) is:

$$\sigma_{\text{nom}} = \frac{My}{I} = M \times \frac{d}{2} \times \frac{12}{hd^3} = \frac{6M}{hd^2}$$

$$= \frac{6 \times 100}{10 \times 10^{-3} \times (25 \times 10^{-3})^2} = 96\ \text{MN/m}^2$$

Now from Fig. 10.22 the elastic stress concentration factor for $D/d = 2$ and $r/d = 0.1$ is:

$$K_t = 1.85$$

$$\therefore \qquad \text{Maximum stress} = 1.85 \times 96 = \mathbf{177.6\ MN/m^2}.$$

(a) For tensile load

Again for smallest part of the bar

$$\sigma_{\text{nom}} = \frac{P}{hd} = \frac{20 \times 10^3}{10 \times 10^{-3} \times 25 \times 10^{-3}} = 80\ \text{MN/m}^2$$

and from Fig. 10.2, $\qquad K_t = 2.44$

$$\therefore \qquad \text{Maximum stress} = 2.44 \times 80 = \mathbf{195.2\ MN/m^2}.$$

(b) For combined bending and tensile load

Since the maximum stresses arising from both the above conditions will be direct stresses in the fillet radius then the effects may be added directly, i.e. the most adverse stress condition will arise in the bending tensile fillet when the maximum stress due to combined tension and bending will be:

$$\sigma_{\max} = K_t \sigma_{b_{\text{nom}}} + K_t' \sigma_{d_{\text{nom}}}$$

$$= 177.6 + 195.2 = \mathbf{372.8\ MN/m^2}$$

Example 10.4

A semi-circular groove of radius 3 mm is machined in a 50 mm diameter shaft which is then subjected to the following combined loading system:

(a) a direct tensile load of 50 kN,
(b) a bending moment of 150 Nm,
(c) a torque of 320 Nm.

Determine the maximum value of the stress produced by each loading separately and hence estimate the likely maximum stress value under the combined loading.

Solution

For the shaft dimensions given, $D/d = 50/(50-6) = 1.14$ and $r/d = 3/44 = 0.068$.

(a) For tensile load

$$\text{Nominal stress } \sigma_{\text{nom}} = \frac{P}{A} = \frac{50 \times 10^3}{\pi \times (22 \times 10^{-3})^2} = 32.9\ \text{MN/m}^2.$$

From Fig. 10.23 $\quad K_t = 2.51$

$$\therefore \quad \text{Maximum stress} = 2.51 \times 32.9 = \mathbf{82.6\ MN/m^2}.$$

(b) For bending

$$\text{Nominal stress } \sigma_{\text{nom}} = \frac{32M}{\pi d^3} = \frac{32 \times 150}{\pi \times (44 \times 10^{-3})^3} = \mathbf{18\ MN/m^2}$$

and from Fig. 10.24, $\quad K_t = 2.24$

$$\therefore \quad \text{Maximum stress} = 2.24 \times 18 = \mathbf{40.3\ MN/m^2}.$$

(c) For torsion

$$\text{Nominal stress } \tau_{\text{nom}} = \frac{16T}{\pi d^3} = \frac{16 \times 320}{\pi \times (44 \times 10^{-3})^3} = 19.1\ \text{MN/m}^2$$

and from Fig. 10.25, $\quad Kt_s = 1.65$

$$\therefore \quad \text{Maximum stress} = 1.65 \times 19.1 = \mathbf{31.5\ MN/m^2}$$

(d) For the combined loading the direct stresses due to bending and tension add to give a total maximum direct stress of $82.6 + 40.3 = 122.9$ MN/m^2 which will then act in conjunction with the shear stress of 31.5 MN/m^2 as shown on the element of Fig. 10.49.

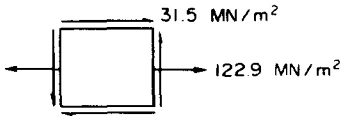


Fig. 10.49.

Then either by Mohrs circle or the use of eqn. (13.11)† the maximum principal stress will be

$$\sigma_1 = \mathbf{130.5\ MN/m^2}.$$

With a maximum shear stress of $\tau_{max} = \mathbf{69\ MN/m^2}$.

Example 10.5

Estimate the bending strength of the shaft shown in Fig. 10.50 for two materials

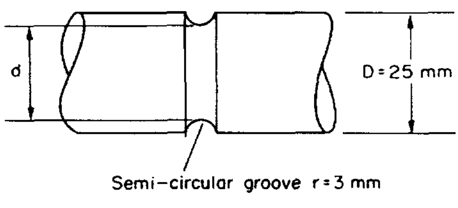


Fig. 10.50.

(a) Normalised 0.4% C steel with an unnotched endurance limit of 206 MN/m^2
(b) Heat-treated $3\frac{1}{2}$% Nickel steel with an unnotched endurance limit of 480 MN/m^2.

Solution

From the dimension of the figure

$$\frac{D}{d} = \frac{25}{19} = 1.316 \quad \text{and} \quad \frac{r}{d} = \frac{3}{19} = 0.158$$

$$\therefore \text{ From Fig. 10.24} \qquad K_t = 1.75$$

From Fig. 10.37 for notch radius of 3 mm

$$q = 0.93 \text{ for normalised steel}$$

$$q = 0.97 \text{ for nickel steel (heat-treated)}$$

$\therefore$ From eqn. (10.32) for the normalised steel

$$K_f = 1 + q(K_t - 1)$$
$$= 1 + 0.93(1.75 - 1) = 1.698$$

† E.J. Hearn, *Mechanics of Materials 1*, Butterworth-Heinemann, 1997.

and the fatigue strength

$$\sigma_f = \frac{206}{1.698} = \mathbf{12.3\ MN/m^2}$$

and for the nickel steel

$$K_f = 1 + 0.97(1.75 - 1) = 1.728$$

and the fatigue strength

$$\sigma_f = \frac{480}{1.728} = \mathbf{277.8\ MN/m^2}$$

N.B. Safety factors should then be applied to these figures to allow for service loading conditions, etc.

Problems

10.1 (B). Two parallel steel cylinders of radii 100 mm and 150 mm are required to operate under service conditions which produce a maximum load capacity of 3000 N. If the cylinders have a common length of 200 mm and, for steel, $E = 208$ GN/m^2 and $\nu = 0.3$ determine:

(a) the maximum contact stress under peak load;
(b) the maximum shear stress and its location also under peak load.

[99.9 MN/m^2; 29.5 MN/m^2; 0.075 mm]

10.2 (B). How would the answers for problem 10.1 change if the 150 mm radius cylinder were replaced by a flat steel surface? [77.4 MN/m^2; 22.8 MN/m^2; 0.097 mm]

10.3 (B). The 150 mm cylinder of problem 10.1 is now replaced by an aluminium cylinder of the same size. What percentage change of results is obtained?

For aluminium $E = 70$ GN/m^2 and $\nu = 0.27$. [−29.5%; −29.5%; +41.9%]

10.4 (B). A railway wheel of 400 mm radius exerts a force of 4500 N on a horizontal rail with a head radius of 300 mm. If $E = 208$ GN/m^2 and $\nu = 0.3$ for both the wheel and rail determine the maximum contact pressure and the area of contact.

[456 MN/m^2; 14.8 mm^2]

10.5 (B). What will be the contact area and maximum compressive stress when two steel spheres of radius 200 mm and 150 mm are brought into contact under a force of 1 kN? Take $E = 208$ GN/m^2 and $\nu = 0.3$.

[751 MN/m^2; 2.01 mm^2]

10.6 (B). Determine the maximum compressive stress set up in two spur gears transmitting a pinion torque of 160 Nm. The pinion has 100 teeth on a pitch circle diameter of 130 mm; the gear has 200 teeth and there is a common face-width of 130 mm. Take $E = 208$ GN/m^2 and $\nu = 0.3$. [222 MN/m^2]

10.7 (B). Assuming the data of problem 10.6 now relate to a pair of helical gears of 30 helix and 20° pressure angle what will now be the maximum compressive stress?

[161.4 MN/m^2]