 S & T MEMO 1/66

THE INFLUENCE OF SHOT-PEENING AND SIMILAR SURFACE TREATMENTS ON THE FATIGUE PROPERTIES OF METALS

PART I*

BY

F. Sherratt

REPORT OF WORK UNDER CONSULTANCY AGREEMENT REF. Est.2/3B, HC/S/1300 FOR MINISTRY OF AVIATION

*AN APPENDIX TO PART 1, SECTION B, ON CASE-HISTORIES IS AVAILABLE AS S & T MEMO 2/66

ISSUED
FEBRUARY, 1966
BY
TECHNICAL INFORMATION AND LIBRARY SERVICES

MINISTRY OF AVIATION

U.D.C. Nos.621.787.6: 539.43

 S & T MEMO 1/66

THE INFLUENCE OF SHOT-PEENING AND SIMILAR SURFACE TREATMENTS ON THE FATIGUE PROPERTIES OF METALS

PART 1

BY

F. SHERRATT

Report of work under Consultancy Agreement Ref. Est.2/3 B, HC/S/1300 for Ministry of Aviation

Summary

Section A reviews existing knowledge about mechanical surface treatments and fatigue. Data show that processes such as shot peening induce residual compression, work harden the surface layer and alter the surface geometry.

Fatigue effects arise mainly from the residual compressive stress which improves fatigue strength in many cases by inhibiting the tensile propagation of fatigue cracks; it has little effect on crack initiation or propagation in a shearing mode. Guiding rules are given.

In Section B case histories giving experimental evidence for a variety of circumstances are reviewed.

It is concluded that in testing, specimen or component shape is all important.

(An appendix to Part 1, Section B, on Case-Histories is available as S & T Memo 2/66).

Section B: Direct Experimental Evidence

"The influence of Shot-peening and similar surface treatments on the fatigue properties of metals".

General Introduction

The fatigue strength of metal components can often be increased by the proper use of shot peening or a similar mechanical surface treatment, and these processes are regularly used in some sections of the mechanical engineering and aircraft industries. Their widespread adoption owes a great deal to the work of Almen and his associates, who used them on a production scale in America from the early 1940's, but the experimental basis goes back at least to the early 1930's, when Thum and Foppl showed quite clearly that controlled cold working could improve fatigue performance. In spite of this, a modern engineer deciding whether or not to specify shot-peening for a particular component is often not able to say with confidence that the treatment will improve matters, unless expensive tests have already been carried out on prototypes. Without considerable direct experience in the use of the particular process, he will have difficulty in drafting a specification likely to approach the optimum, and even an experienced user finds it difficult to guess the magnitude of any improvement. Finally, if the process is precisely specified, control at production level may be difficult, and reasonably rigorous inspection methods are not generally established. All these factors detract from the usefulness of the mechanical surface treatments, and hinder their general application to components which might well be substantially improved.

The difficulties fall broadly into two groups, and with this in mind the review has been divided into two main parts. In the first place, the whole question of why shot-peening improves fatigue strength at all has only been explained in a very rudimentary way. It is usually considered enough to point out that it induces residual compression in the surface, and that compressive mean stresses raise the allowable range of cyclic stress. This only pushes the problem one stage further away, and does very little to rationalise some of the known facts about peening, including the very important one that peening does not always improve fatigue strength. There is need, then, for a "general theory of shot-peening", or, being more realistic, a working hypothesis based as much as possible on modern fundamental fatigue knowledge, but relying on experimental, empirical rules to fill the gaps. This will be attempted in Part I, below, Section A being a general exposition, and Section B an examination of some test results with the working hypothesis in mind.

The second type of difficulty could be called purely practical, and is typically encountered when a surface treatment reaches the production line. Questions such as "What exposure limits are allowable ?", "How can we check exposure ?", "What happens at the edge of a peened region ?", and "Does a scratch through the peened surface ruin its effect ?" are all asked regularly, and often not answered satisfactorily. The answers to many of these problems have not yet been established, but the aim of Part II will be to report existing experience, and comment, again with the working hypothesis in mind.

General characteristics of shot-peening, etc.

The term "mechanical surface treatments" will be used below to mean those treatments which deliberately cold work the surface of a component, with the aim of improving its fatigue strength. This includes shot-peening, vapour-blasting, vacu-blasting, grit-blasting etc., and much of the discussion will apply equally well to cold-rolling and barrel-tumbling, although no attempt has been made to cover these comprehensively. The dominant "blasting" process, from a fatigue viewpoint, is shot-peening, where the particles are comparatively large, and the affected surface layer comparatively deep : "small-particle" processes, such as vapour-blasting, are important in volume of production, but do not rely so heavily on inducing a significant depth of cold working, and do not generally have such a pronounced fatigue effect as shot-peening.

In the shot peening operation, the treated surface is impacted by a stream of shot, which causes plastic deformation in a shallow skin. The immediate effects are that:-

A. A state of residual stress is set up, the thin surface skin having fairly high compressive stresses, balanced by smaller tensile stresses in the core.

B. Plastic flow in the surface material causes the usual changes of mechanical properties, such as work-hardening.

C. Changes occur in surface geometry, often a roughening of the immediate surface, sometimes accompanied by the formation of fine, shallow cracks.

Any attempt to explain the mechanism by which shot-peening improves fatigue strength must take account of all these factors, and of their interactions on each other. Some of the earliest experimental work in this field was aimed at separating the effects of residual compression from those of work hardening, so that the approach is by no means new, but the same division has been used in the present account, since it still seems the most promising line of attack.

Part I. Basic Principles of Mechanical Surface Treatment

Introduction

To an engineering designer, the most elementary problem in this sphere is to decide whether or not a certain component is even likely to be improved by mechanical surface treatment. Experience has led to a number of empirical rules which at least help to spot the "unfavourable" situation. Fuchs (1963) has summarized the ones for shot peening very well:-

"(a) Parts with section changes or other stress raisers can benefit more by shot peening than parts that are smooth.

(b) Parts with skins that are prone to fatigue damage (decarburization, corrosion attack, chromium plate) can gain far more than parts without such skins.

(c) Potential gains in fatigue life increase with the hardness of the steel, regardless of alloy content. At equal hardness levels, various steel alloys gain about equally.

(d) If failures occur at less than 1000 cycles, shot peening is not likely to help. If they occur at more than 100,000 cycles, shot peening is likely to increase the life by a large factor.

(e) Large and small parts can gain equally by peening.

(f) Fatigue failures almost always start at the surface."

Rules of this type must be the outcome of the rational, fundamental mechanism by which shot peening improves fatigue life: the implication is that a full understanding of this mechanism would simplify the rules, and widen their scope. Any complete explanation must be based on a complete understanding of the fatigue mechanism, and this does not yet exist, although major advances have been made in the last fifteen years or so. In Section A, below, an attempt is made to examine the implications of those advances, and develop a simplified mechanistic approach to the surface treatments, and in Section B the approach is applied to some of the published work in the field.

It is convenient to break down the immediate effects of mechanical surface treatments in the same way as was suggested in the general introduction, and Section A will consider in turn the effects on fatigue of:-

(i) Residual stress

(ii) Work hardening

(iii) Surface profile changes.

Section A: The Mechanism of Fatigue Strength Improvement

1. The basic fatigue process

1.1. Crack initiation and crack propagation

The mechanism which starts a crack is not necessarily the same as that which makes it grow. Failure to recognise this led many early fatigue theories into difficulties, and in the same way some of the suggestions made about surface treatments are of doubtful use. It is now apparent that many specimens and components contain cracks for a very high proportion of their lives, and all suggestions must take account of this.

1.2. Mechanism of crack initiation

The search for the basic mechanism which starts a fatigue crack is being continued mainly by electron microscopy, in association with the mathematical theory of crystal defects. References are numerous, but Thompson and Wadsworth (1958) summarized the situation up to a few years ago, and both Kennedy (1962) and Maclean (1962) use the results extensively in general textbook form. There is no doubt that a "crack" (or some form of sub-microscopic "hole") is formed by the movement within a crystal of dislocations (i.e. a line of mis-fitting atoms) and point defects (i.e. a vacancy in the crystal lattice, or an extra atom). A precise description of the mechanism would seem to be more interesting to a physicist than to a design engineer, but one of the basic differences between some of the competing theories has a very practical engineering significance. Put very simply, the mechanisms are of two forms:-

(a) Processes relying entirely on dislocation movement, for example Wood (1956).

(b) Processes in which the formation and diffusion of vacancies plays an important part, for example Mott (1955).

The engineering significance is that if process (a) is the dominant one, only the <u>range</u> of shear stress will affect the <u>initiation</u> of a crack, the mean or peak values being unimportant: on the other hand, the diffusion processes in (b) will be affected by mean stress. Therefore a simple answer to the question "Does diffusion of vacancies play a part in initiating a crack" is also an answer to the question "Does mean stress affect the initiation of a crack".

The real situation, of course, cannot be resolved in such simple terms, and there is little doubt that diffusion of vacancies is important in some circumstances, (for instance creep-fatigue) but not in others (e.g. low-stress fatigue at very low temperatures). Even the simple picture of dislocations gliding on fixed crystal planes under fluctuating shear stress is not complete, since the "climb" of dislocations from one plane to another is a process affected by time, temperature and sustained stress. Without getting involved in this complex subject we can, for our present purpose, note the following trends:-

(i) A fairly general acceptance of dislocation glide as the <u>primary</u> crack initiator, with slip on more than one plane as an almost essential condition, e.g. Mott (1958).

(ii) Recognition of diffusion as an important <u>secondary</u> process, negligible at low temperatures, low stresses and high cycling frequencies, but more prominent at high temperatures, high stresses and low frequencies.

(iii) A swing away from phenomenological explanations based on progressive work-hardening, such as Orowan's (1939).

1.3. Crack propagation

Experimental work on crack propagation has not been as intensive as the work on crack initiation, and has been carried out in the main by engineers, rather than by physicists. One result is that the central theories on the topic are phenomenological, like Head's (1953), and explanations at the atomic level are not nearly as well developed. Head's approach, based on Orowan's concept of plastic elements progressively work-hardening, links successfully with observed rates of crack growth, but in the present context, we are looking for the mechanism by which shot peening improves fatigue life, and it is reasonable to expect most help from speculations about the mechanics of crack growth, rather than from models displaying the same mathematical characteristics.

A basic difficulty is to decide when a crack stops 'initiating' and starts 'propagating', and it must be admitted that a crack has not yet been defined in a way which allows this distinction to be made with precision. Many writers find that the fatigue process has to be divided into more than two phases: for instance Lipsett et. al. (1959) observed five stages in aluminium sheet, although their loading conditions and material anisotropy seem to account for some of the changes. One of the most helpful observations is that of Stubbington and Forsyth (1961), writing particularly of corrosion fatigue in Al-Zn-Mg alloys: they distinguished three phases, i.e.

(a) Crack initiation.

(b) Stage I propagation (in the shear mode, i.e. at some appreciable angle to the maximum tensile stress.)

(c) Stage II propagation (in the tensile mode, i.e. in a plane normal to the maximum tensile stress.)

They found varying proportions of shear-mode and tensile-mode cracks according to stress level (rotating cantilever) or other variables, including cases when shear-mode cracks were "not apparent". The general idea, though, of a transition phase of shearing-type propagation, sandwiched between a shear-dominated initiation phase and a tension-dominated main propagation phase, seems logical. In another paper, Forsyth (1961) extends the idea, suggesting that the change to Stage II growth occurs when a crack meets a slip obstacle such as a grain boundary, that the criterion for growth in this stage is the maximum principal tensile stress, and that striations or beach markings on the fatigued surface may occur during this phase.

Forsyth and Ryder (1961) suggest that these striations are formed by internal voids opening ahead of the crack tip, and Holden (1961) reports X-ray diffraction studies which support this, but the exact mechanism is still doubtful.

Even without this knowledge the ideas are still useful, and we can add two further comments:-

(i) Measurements by Rolfe and Munse (1963) show that the range of maximum principal strain near the tip of a crack builds up to a steady value as the crack starts to propagate: this possibly provides the change in circumstances needed for the switch from shear-mode to tensile-mode propagation.

(ii) Observations suggest that extensive regions of shear-mode fracture can occur when either a mean compressive stress, or marked material anistropy is present.

1.4 Non-propagating cracks.

The work of Frost (1956-1963) and his colleagues on non-propagating cracks is of direct importance to our present theme. Extended experimental work, reported in many papers, has established quite firmly that when steep stress gradients (sharp notches) are present, cracks can be started at the most highly-stressed point and then fail to propagate as they reach regions of lower overall stress. The critical alternating stress required to propagate a crack is given by $\sigma^3 \ell = C$, where σ = alternating stress, ℓ = crack length or depth, C = a material constant. Extracts from the latest experimental results are given in Table 1.

Table 1. (after Frost, 1963).

Material	Tensile strength tons/sq.in.	Plain fatigue limit: tons/sq.in.	C
Mild steel	28	±13	5.5
Alloy steel	60	±32	5.5
Copper	14.5	± 4	0.6
4½% Cu-Al alloy	29	± 9	0.2

In his latest paper Frost refers to the ideas of Forsyth and Holden, accepting that the $\sigma^3 \ell$ relationship is probably restricted to tensile-mode propagation, and an important section of Holden's paper examines the significance of non-propagating cracks. Holden's finding that a minimum finite volume must suffer deformation for the crack to propagate, gives a new physical meaning to a very long-standing idea, i.e. Neuber's mathematical concept of a minimum element over which the stress must exceed some critical value.

This work has already shed much light on the notch and size effects, and will obviously be developed much further, but we have enough information for our present purpose, and can summarize the findings as follows.

1.5. The mechanism of fatigue

Frost (1963) describes the simplest possible fatigue situation in these terms:-

> "When a plain specimen is subjected to stresses greater than the plain fatigue limit, cyclic shear stresses cause cracks to develop on slip planes. Such slip plane cracks penetrate to a certain depth and then change their direction of growth from that of the crystallographic slip plane to being normal to the applied loading. They now propagate under the action of normal cyclic stresses by a mechanism distinct from that responsible for slip plane cracking."

We may extend this simple picture by adding:-

(i) When the crack initiates at a notch, the range of alternating shear stress will diminish as the crack starts to propagate away from the notch tip, slowing the rate of shear-mode progress, and possibly halting the crack before satisfying the conditions for continuing in the tensile mode.

(ii) Departure from a simple tension-compression cycle will modify the balance between range of alternating shearing stress and maximum principal tensile stress, probably modifying the balance between shear-mode and tensile-mode progress. In the extreme case of a completely compressive cycle the whole process of cracking may rely on shearing-mode.

(iii) Local variations in material properties, and especially anisotropy, may confuse the picture. For instance, propagation will presumably continue under the fastest available mechanism and directional properties could cause shearing on a particular plane to offer faster progress than tensile-mode on some other, stronger plane.

(iv) A macroscopic crack normal to the applied load is not certain evidence that tensile-mode propagation was responsible. When a high stress field is narrow in extent, a shearing crack could zig-zag on alternate 90° planes, giving a generally normal crack. McClintock (1952) and Peterson (1959) have both suggested this as a basic propagation mode, and while it must be regarded as superseded by Forsyth's suggestion, it could still happen in special circumstances.

One tacit assumption, of course, is that tensile-mode progress is inherently faster than shear-mode, unless special conditions intervene.

2. Mean Stresses and residual stresses

2.1. The influence of mean stress

To a single grain of material, a macroscopic residual stress must seem the same as an applied mean stress, and have much the same effect. The gross influence of mean stress on total life to fracture is well documented, e.g. Smith (1942) or Sines (1955). A tensile mean stress decreases the range of alternating stress for a given life, a compressive mean stress increases it, and a mean shear stress has little effect. These are generalizations, and exceptions have been reported: the only restriction we need to note at present, though, is that yield on a macroscopic scale at any part of the stress cycle will introduce special conditions.

The most direct simplifying assumption we can make about mean stress is that it has no effect on the initiation of a fatigue crack, and only changes the total life of a specimen by altering the propagation process. Some authorities do, in fact, accept this; Phillips and Frost (1963) state quite boldly that "..... only the alternating component of a load leads to the initiation of a surface fatigue crack", and strong evidence in support of this goes back at least to Gough's single crystal work of the 1930's. The real problem is no longer whether or not the statement is broadly true, but why it is not more widely accepted as true, and why it has made so little impact on fatigue design rules. Many hypotheses in fatigue, such as Gerber's and Goodman's rules for the effect of mean stress, Miner's "Law" for cumulative damage, and various suggestions for dealing with combined stresses, are obvious topics for an approach on these lines, but even in the few cases where such an approach has been tried, it has not been spectacularly successful.

One of the obstacles has certainly been that there was little reliable information about the extent of the 'propagation' phase, and although the occurrence of non-propagating cracks had been noted, detailed evidence was lacking. Both these gaps in knowledge are now being steadily filled. The existence of two possible modes of propagation, shear and tensile, has sometimes been acknowledged, but the possibility of a change in conditions affecting one mode more than the other does not seem to have been thoroughly examined. Again, the comments on diffusion of vacancies in para.1.2 above, suggest that the basic premise may itself break down when high stresses, high temperatures, and slow load cycling are present, i.e. in the circumstances where we would normally expect creep to play a part. Finally, marked anisotropy of the fatigued material could so change the direction of propagation of a crack, that simple correlations broke down.

All these factors confuse the issue, and no doubt other modifying influences operate in some special cases. If the original principle is to be hedged about with restrictions and alterations, though, it will soon loose its simplifying influence, and a deliberate compromise must be made between universal and precise applicability on the one hand, and ease of use on the other.

2.2. A guiding principle for the effect of mean stress

With this in mind, then, we may state:-

Rule 1. A fatigue crack forms in three stages, (a) Initiation, (b) Shear propagation, (c) Tensile propagation. In normal circumstances, a compressive mean stress will have no effect on crack initiation, and only a small effect on shear propagation, but will significantly increase the range of applied stress needed to sustain tensile propagation.

The points we need to remember when applying this are:-

(i) "Normal circumstances" are ones in which no macroscopic yield occurs during the loading cycle, and no significant amount of creep is expected.

(ii) Total life may be made up of varying proportions of stages (a), (b) and (c) in different circumstances, including the complete absence of any particular stage (e.g. when a treatment induces surface cracking from the outset).

(iii) Mean shear stresses of any sort will not affect any stage.

(iv) The important stresses are those experienced by the crack, and are referred to the direction in which the crack is propagating; thus a tensile mean stress pulls the crack faces apart, a compressive stress presses them together, and shearing stress tries to slide them over each other. These obvious statements become essential guides when a crack is changing direction in a complex stress field.

One part of this rule has been stated with no prior examination, i.e. that mean stress has a "small, secondary effect" on shear propagation. This has been included with the work of Findley et.al. (1956) in mind: they suggested a combined stress failure criterion based on the "range of cyclical shear stress", with a modifying term to allow for direct stress normal to the path of the crack. Without adopting the whole of this approach, the simple concept of compression forcing together the faces of a crack, so that friction between them carries a portion of the cyclical shear, seems a reasonable one. Any more rigorous application would need an experimental value for the "coefficient of normal stress effect", and if this proved to be zero the effect would simply drop out.

2.3. The special circumstances of the fatigue "limit"

A particularly careful approach must be made when considering the materials with a well-defined plain fatigue limit (i.e. for polished specimens), such as mild steel. Within the framework we have established, the simplest idea of a plain fatigue limit supposes that it is the minimum range of cyclical stress needed to initiate a crack, i.e. complete Stage I. If this were so, mean stresses would have no effect on the fatigue limit of polished specimens. Experimentally, this is not true, and it becomes necessary to treat the fatigue limit as the minimum stress needed to complete stages (a) and (b), to fit the observation that mean stress has a small effect on plain fatigue limits and a larger one on notched fatigue limits. The sharp distinction between stages (a) and (b) then tends to merge into a gentle transition, from a slip-induced surface roughening at the start, to the definite propagation of a shear crack as Stage (b) gets under way. A perfectly polished specimen, fatigued just under the limit, could thus run indefinitely for one of two reasons: either because no slip bands formed at all, or because they formed, but failed to achieve the minimum length of crack required to start tensile propagation. The second condition implies that the specimen will contain non-propagating cracks which have come to a halt in a reasonably uniform stress field, which is much more difficult to accept than the idea of cracks halting after moving away from a notch. Several ways round this difficulty can be suggested, including the complete rejection of a "perfectly polished" specimen as an experimental possibility, but much more needs to be known about the conditions controlling shearing propagation before a deeper analysis is possible. Since crack propagation is only an incidental concern of the present survey, we can note the existence of the difficulty, conclude that it does not conflict with the experimental observation that residual stresses have only a small effect on plain fatigue limits, and pass on.

2.4 Equating residual stresses and mean stresses

It is usual, and valid, to regard a residual stress as equivalent to a steady mean stress. Three factors must always be considered in using this approach, though: they are:-

(i) A residual stress system is rarely a uniform, uni-axial one, and care must be taken that all likely stresses have been considered.

(ii) The nature and magnitude of the residual stresses induced by a particular process are rarely known with any precision.

(iii) Application of cyclic loading to a residually stressed body is likely to cause progressive changes ("fading") in the residual stresses.

The second and third factors, in particular, are a serious obstacle to any quantitative predictions in this field. The labour and expense of the technique for measuring residual stress, most of them destructive, lie behind this lack of knowledge, and at the same time make it impossible to measure such stresses on production components. There is no doubt that a reliable, rapid, non-destructive "residual-stress-meter" would revolutionise this whole field, and a major effort to produce one (perhaps using X-ray diffraction, see Bolstad et.al. 1963) would certainly be worth while.

Published work gives some help in estimating the likely magnitude and extent of the residual stresses near the surface of simple components, made from common metals, and subjected to commonly-used surface treatments, and reference is made to some of this work in Section B. Shot peening, for instance, is generally agreed to induce biaxial compression of magnitude half the yield stress or more, in a layer of the order .005" to .020" (say one quarter the shot diameter). Rough guesses of this sort are often good enough for qualitative purposes, and so far the mechanism of the peening improvement is so little understood that few quantitative estimates are attempted. One of the main gaps in this sort of knowledge concerns the stresses induced in the vicinity of a notch, e.g. when a V-groove is peened with shot too large to impact the root of the notch directly: again, the component of residual stress in a direction perpendicular to the surface is not well understood, and this could be pertinent when the fatigue mechanism near a notch is being considered.

It is much more difficult to find experimental backing for predictions about the changes in residual stresses under cyclical loading, except for a general confirmation that "fading" can occur. The simplest theoretical approach with any hope of quantitative results is to apply a yielding criterion to the total residual-stress-plus-load-stress system (taking triaxial stresses into account), and calculate the maximum residual stress values which can exist without causing yield at some part of the load cycle: Rosenthal (1959) uses a maximum-shear-stress criterion in this way with some success. It is generally accepted, though, that alternating stresses are themselves a 'relaxing' influence on materials containing large numbers of dislocations (e.g. work-hardened metals), and that the residual stresses will fade to values even lower than those predicted by the simple application of static yield criteria. A comprehensive and reliable method for estimating, even approximately, the amount of fading that does occur ought certainly to be one of the major aims of future research in this field. To collect the fragments of evidence in this section is unlikely to help matters, and more definite reference will be made to them in Section B, when each can be related to a specific fatigue circumstance. For the present, an attempt must be made to formulate a guiding rule which is neither so vague as to be useless, nor so ambitious that it is not supported by experiment. Two helpful points on which there is "general agreement" are the increased likelihood of fading in soft materials as opposed to hard ones, and the much reduced fading of stress near a notch, where triaxiality of stress helps. Pronounced fading also seems more common when the initial residual stresses are high (say more than 75% of the yield stress), and the fatigue limit (or 10^8 strength) is high relative to the yield strength. One serious point of disagreement concerns whether the fading occurs rapidly in the first few loading cycles, or gradually over an extended fatigue life. A majority of reports (e.g. Rosenthal & Sines, 1951 or Elsesser 1957) show practical completion of the fading early in the life, but Pattinson and Dugdale (1962) surprisingly found that with a 4½% Cu-Al alloy little fading occurred before 10^6 cycles, followed by quite pronounced progressive fading as cycling was continued to 10^8, even for applied stresses less than half the fatigue limit. This means, unfortunately, that when such late fading is possible, short-life testing at high applied stresses may overestimate, rather than underestimate, the benefits of a particular treatment. Apart from this anomaly, which deserves more experimental investigation, the general picture is that high applied stresses cause more rapid fading, so that short lives benefit <u>less</u> from surface treatment than long ones.

2.5. A guiding principle for the effect of residual stress

<u>Rule 2</u>. Residual stresses affect fatigue life in the same way, and to the same extent, as the equivalent mean stresses. "Equivalence" must take account of stresses in more than a single direction, and allow for the potential fading of residual stress.

More specific points in applying this are:-

(i) Biaxial compression of magnitude at least 30% to 50% of the yield stress may be induced and retained in medium-hard to hard materials (say over 300 V.P.N.).

(ii) Similar treatments will induce residual compression approaching 100% of the yield stress in softer materials (say about 200 V.P.N.), but unless favourable notch conditions exist, early fatigue cycles will remove most of the residual stress.

(iii) High fatigue stresses, leading to failure in 1000 cycles or so, will cause severe fading in almost any material, under any conditions: so will prolonged exposure to temperatures appreciably higher than ambient.

(iv) The penetration (i.e. depth of residually stressed layer) of a treatment will be one of the significant variables, since crack propagation is involved. This will not simply be a question of deeper penetration giving more improvement, though, because of the connection with non-propagating cracks etc. In given circumstances there will normally be a minimum depth for any improvement at all to take place, and a maximum beyond which no extra improvement occurs, but these will vary widely according to material and shape of the component, etc.

3. The effect of work-hardening

Modern theories of work-hardening, based on the interaction stresses of massed dislocations, have tended to blur a little the sharp engineering distinction between "pure" work hardening on the one hand, and macroscopic residual stress on the other, since both arise from local imbalance between the atomic forces. Division into 'micro-' and 'macro-' residual stress is profitable in some circumstances, but an adequate idea of 'pure' work hardening in the present context comes from imagining a cylinder of material machined from the core of a bar which has been uniformly compressed beyond its yield point. In spite of differences in crystal orientation, etc., which cause local variation in yielding, with resulting local stresses, there will be no <u>systematic</u> residual stresses in such a cylinder, in contrast to the situation if, say, its surface alone had been rolled beyond yield. For engineering purposes, then, simple types of uniform plastic distortion cause work-hardening without residual stress, and it is the effect of this sort of work-hardening (artificial though the distinction is) that we need to consider.

3.1. Work hardening vs. residual stress

A simple approach to the problem is to assume that the effect on fatigue of a certain amount of work-hardening plus a given residual stress, is the same as the sum of the effects of each applied separately. Experiments to verify this are easy to devise, and work-hardened specimens without macroscopic residual stress can easily be obtained by suitable cutting, although there is some argument about whether residual stress is easily produced without work-hardening (or some other physical change). In spite of this, a fierce controversy is said to have been carried on in Germany during the 1930's about the relative contributions of these immediate effects, and although there is now little arguement about the subject, conclusive evidence one way or the other is difficult to find.

There is no doubt that work-hardening alone can have a significant effect on fatigue strength, although the magnitude and sign vary between materials. Some examples reported are a substantial improvement for mild steel and titanium, little, if any change for copper and a Ni-Cr alloy steel, and a possible decrease for a high strength aluminium alloy (Kaufman and D'Appolonia, 1955; Frost, 1958 and 1960; Teed 1952). Even if these effects could be simply added to that of residual stress, then, the situation would vary from one material to another, and there is evidence that the actual combined effect is not simply additive, but depends on the type of applied stress system, among other things. Thum is sometimes quoted as concluding that for bending fatigue 80% of any improvement comes from residual stress, and 20% from work-hardening, but that in torsion the proportions are reversed.

One possibility is that the confusion arises from neglect of directional effects, and the bending/torsion discrepancy supports this. Reverting to the fatigue crack mechanism, any preferred orientation of the crystal planes, or change in the effective grain size, such as cold-working produces, could have a strong effect on the initiation of a crack, since crystal slip is the dominant factor. Stage (b), with its emphasis on shearing planes, could also be strongly affected. If, in addition, the potential length of the shearing crack were a deciding factor in fatigue strength, work-hardening could make a decisive contribution. There is little solid evidence for this sort of arguement, but any future attempts to separate work-hardening and residual stress effects should be carefully designed to monitor variation of strength in different directions, with a particular watch on anisotropy of the original, untreated material (if wrought aluminium alloys are used, for instance).

At present the only circumstances where the work-hardening contribution of a comparatively shallow surface treatment is likely to be important is when plain specimens of a soft metal (say, mild steel) are being treated. Because of fading, residual stresses are unlikely to help much, and a major part of the small (perhaps 5-10%) improvement may be due to work-hardening, but these are in any case circumstances unlikely to interest a designer. The emphasis here must be on the "comparatively shallow" aspect of the surface treatments like shot peening. When gross deformation of the whole cross-section occurs, as in a cold-drawing operation, work-hardening must be more important than residual stress, and Harris (1961) quotes a case where the fatigue limit of an 18/8 Ni-Cr steel was more than doubled by cold work.

3.2. A guiding principle for the effect of work-hardening

<u>Rule 3</u>. In most practical applications of surface treatment, work-hardening is likely to be less important than residual stress. Its effect will be beneficial on

soft materials with high hardening coefficients, but may be slightly detrimental on less ductile metals. Anisotropy due to cold work should be considered a potential contributor to anomalous test results.

Until the effects of residual stress are more exactly predictable, efforts to isolate the work-hardening contribution are not likely to directly affect the practical uses of shot-peening, although they may help with fundamental fatigue knowledge.

4. The effect of surface profile

The detrimental effect on fatigue strength of roughening a surface is well known, and qualitative statements in general terms are easily made, the shortest maxim being "the smoother the better". The extensive literature on the subject does not include a general quantitative rule, although particular circumstances have sometimes been covered by formulae, usually relating fatigue limit to surface roughness (C.L.A. in micro-inches, say) and hardness, for a particular metal or group. In practice roughness rarely varies alone, and different grades of turned finish, for instance, differ in the amounts of cold-work and residual stress they introduce in thin layers near the surface, as well as producing different profiles. This problem of separating the effects, linked with the problem of whether it is even desirable to separate the effects, is more acute when shot-peening and such processes are being considered, since strong residual compression may overcome the effects of surface roughness.

4.1. Surface roughness produced by various treatments

This is one region where a clear qualitative distinction can be made between two of the main treatments, represented by shot-peening and vapour blasting. In general, a shot peened surface is much rougher than the original, untreated surface, so that as-peened components invariably rely on residual stress and work-hardening for any improvement. If environment and applied loading remove these by relaxation, there is a very real danger that the peening roughness will cause a marked reduction in fatigue strength. On the other hand, if the surface before treatment has sharp scratches and gouges, these may be locally rounded by vapour blasting, giving an improvement which does not depend on maintaining residual stress. This smoothing effect is relatively straightforward, but the roughening from shot peening needs a little more discussion.

4.2. The nature of the "as-peened" surface

Impacting a surface with shot of an appreciable size (e.g. .020" - .030") produces at best a smoothly dimpled surface, but usually the profile contains at least a proportion of sharper pits. There is evidence that in many cases shallow cracks are formed, perhaps penetrating .003" or so. When residual compression is maintained in the surface layer these do not have a disastrous effect on fatigue strength, although Coombs et. al. (1956) found that polishing off the surface layer after peening gave a further improvement on plain specimens of a spring steel. Hard evidence on the subject is not plentiful, but it is quite established that where the rough, as-peened surface cannot be accepted, a careful finish-machining operation after peening does not reduce the final strength, and may well improve it. The depth removed is likely to be critical, since less than .003" may not clean the surface, and more than .005" may remove a substantial part of the stressed layer. Mainly for this reason, development testing would certainly be needed before adopting peening-plus-machining in a given circumstance. For general purposes, it is as well to regard an as-peened surface as one already containing minute cracks, which do not propagate because of the residual compression.

4.3. A guiding principle for the effect of surface roughness

Rule 4. Changes in surface profile will depend mainly on size and sharpness of the impacting particles. Fine-particle processes such as vapour blasting will in general give a slight improvement on this count, but the surface left by shot-peening should be regarded as potentially cracked. Smoothing after peening is allowable, and careful machining in these circumstances can add to fatigue strength.

The effect of peening before applying an anti-corrosion coating, such as electrodeposited metal, is a particularly interesting case (see Section B). Peening changes the stresses at the junction between parent metal and coating, from tension to compression. Cracks from the coating are thus prevented from propagating into the core, and fatigue strength is substantially improved.

5. Summary of basic principles

The general conclusion, then, is that although other factors are present, the most important single effect of a mechanical surface treatment is to induce residual compression in a shallow layer. This improves fatigue strength by slowing down, or completely halting, the propagation of fatigue cracks, rather than by affecting the initiation of the cracks. When treatments produce no improvement it may be because the fatigue strength of the untreated component is limited by the condition for crack initiation (perfectly smooth specimens, for example) or because the compressive stresses have faded under the action of cyclic applied stresses.

One consequence of the basic premise is that the depth of the affected layer is a significant factor, and since this depends largely on the size of the particles used to impact the surface, processes may be broadly classified as either:-

(a) "Large-particle" processes, such as shot-peening, with shot sizes of .020" upwards, giving an affected zone about .010" to .050" deep.

or (b) "Small-particle" processes, such as vapour blasting, where typical penetrations are of the order .001" or .002".

The contribution made by changes in surface geometry depends on whether a large-particle, or a small-particle process is involved. The detrimental effect of the rough surface left by the former is generally nullified by the fairly deep layer of residual compression, and surface geometry is not a dominant factor. On the other hand, small-particle processes generally round off any sharp gouges or deep scratches in the surface, and the benefit this gives probably accounts for a significant part of the fatigue improvement.

One quick way of eliminating some situations where surface treatment is unlikely to help is based on the "fatigue-ratio" (fatigue limit divided by U.T.S.). Harris (1961) has suggested that 0.45 to 0.5 is an "ideal" value of this ratio for all steels, usually achieved only by avoiding inclusions, decarburisation, surface roughness etc., and that where values of 0.45 or so are already being achieved, further surface treatment is unlikely to give any improvement. This does not conflict with any of the ideas above, since the fatigue limit in these circumstances is decided by whether or not a crack can be initiated, a process which is not affected by residual compression. The best way of using this suggestion is probably to regard a fatigue ratio of 0.45 or so as the maximum attainable with any steel, either by the use of an optimum surface treatment, or by eliminating all stress raisers, and adopt whichever of these two processes is most expedient. A process like shot peening will generally be cheaper than one aiming at a highly polished surface, although the rough peened surface may be unacceptable for other reasons.

The empirical rules for shot-peening quoted in the introduction (Fuchs, 1963) do not conflict with the approach now being suggested. Taking some of the specific points, for instance, the emphasis on "parts with section changes or other stress raisers" has dual origins: firstly, the propagation phase is more important when stress raisers are present, and secondly the notch configuration will help to prevent fading of the residual stresses. "Skins prone to fatigue damage" are, in effect, skins in which cracks form easily under cyclic stress, and large improvements naturally come from a treatment which prevents these from propagating. Hard steels have a low fatigue ratio and high notch sensitivity because of the ease with which cracks propagate away from minute stress raisers in such steels, and the increased effectiveness of peening as hardness increases is to be expected. Finally, the complete lack of improvement at very short lives is linked with the rapid fading of residual stresses which occurs when high level cyclic stresses are imposed.

5.1. Re-statement of guiding rules

Collecting together the four rules given previously, we have:-

Rule 1. A fatigue crack forms in three stages, (a) Initiation, (b) Shear propagation, (c) Tensile propagation. In normal circumstances, a compressive mean stress will have no effect on crack initiation, and only a small effect on shear propagation, but will significantly increase the range of applied stress needed to sustain tensile propagation.

Rule 2. Residual stresses affect fatigue life in the same way, and to the same extent, as the equivalent mean stresses. "Equivalence" must take account of stresses in more than a single direction, and allow for the potential fading of residual stress.

Rule 3. In most applications of surface treatment, work-hardening is likely to be less important than residual stress. Its effect will be beneficial on soft materials with high hardening coefficients, but may be slightly detrimental on less ductile metals. Anisotropy due to cold work should be considered a potential contributor to anomalous test results.

Rule 4. Changes in surface profile will depend mainly on size and sharpness of the impacting particles. Fine-particle processes such as vapour-blasting will in general give a fatigue strength improvement on this count, but the surface left by shot-peening should be regarded as potentially cracked. Smoothing after peening is allowable, and careful machining in these circumstances can add to fatigue strength.

Section B. Direct Experimental Evidence

Introduction

A great deal of useful information about surface treatments has been obtained in the most direct manner, by fatigue testing specimens or components previously subjected to the treatment, and comparing the results with those from similar tests on untreated components. The overall circumstances of any particular test programme are made up from a very large number of separate factors, and inevitably some of these factors have been more thoroughly investigated than others. For instance the influence of shot-peening on small rotating-bending specimens of medium-hard steels has been investigated many times, and an exhaustive search through old test records would probably yield a mass of information on this topic alone. It is unlikely that the exercise would be rewarding, though, since the general pattern of behaviour in these circumstances is by now pretty generally agreed. On the other hand, the effect of surface treatment on components fatigue-loaded in direct tension, with various mean loads, has not been so thoroughly examined, in spite of its obvious importance in both the practical and fundamental fields.

The rules put forward in Section A have been derived so far by blending fundamental fatigue knowledge with general surface treatment "know-how", and must obviously be tested against existing direct evidence. Because the rotating-bending case has been so widely studied, it will have a prominent place in any such comparison, but to avoid giving an unbalanced picture the information has to be carefully selected, rather than comprehensive. In making a selection, some pattern has to be adopted, and the aim below has been to set out a series of case histories, chosen to test the validity of the rules in as wide a variety of circumstances as possible.

Because residual stresses are so important, ideal test results include residual stress measurements as well as fatigue tests. A lot of good evidence would have to be rejected, though, if this were an essential condition, and an attempt must be made to estimate the likely residual stresses for some of the reported conditions where no actual measurements were made. In any case, a reliable method for estimating these stresses, even approximately, would be of direct help in the application of surface treatments, and it will be useful to examine the information on this topic before starting to look at fatigue test results.

2. Residual stresses induced by shot-peening

2.1. The general stress pattern

Both shot peening and vapour blasting induce residual compression in a surface layer, but the depth of this layer is significantly greater for shot peening, and in fact one of the points to be made later is that the depth of penetration is largely controlled by the particle size. The fine-particle processes, then, can be regarded in the present context as special cases of shot-peening, and since most of the available measurements are on shot-peened surfaces, it is expedient to refer throughout to the effects of "peening", or /shot-peening".

Methods of measuring residual stresses may be classified as either dissection-type or X-ray type. In the former, the specimen has portions removed in a controlled sequence, and the resulting dimension changes are measured: X-ray methods rely on the distorsion of diffraction patterns, and generally give only the stresses at the immediate surface. Both methods have been used in peening investigations, but dissection techniques have supplied the bulk of the information, and it will be assumed in the discussion below that this was the method used unless X-ray diffraction is specifically mentioned.

A casual examination of some of the published residual stress measurements brings out one immediate difficulty. A residual stress system in a solid body is obviously triaxial, and even if the discussion is limited to the immediate surface region, two principal stresses are involved. In spite of this, many workers quote only one measured stress, usually along a principal geometrical axis in the surface. This may be enough when the results are to be applied to fatigue specimens with simple shapes, loaded in a simple manner, but the limitations of the information must be kept in mind when any more complex situation is being examined.

One of the simplest cases is that of a flat plate, peened on one side. This is a convenient specimen shape for dissection techniques, and the results can be applied fairly directly to the shot-peening of laminated automotive springs, a regular commercial application. Enough information has been published on this type of specimen for some broad conclusions to be drawn about the effect of different peening conditions. Most specimens have been rectangular, both length and breadth being much greater than the thickness, and the discussion will generally be concentrated on the longitudinal principal stress in a specimen clamped during the peening. When such a specimen is released after peening, it springs into a curve, convex towards the peened side, and the stress distribution obviously changes. The important distribution, though, is that in the clamped straight specimen, immediately after peening but before release, and this is the distribution considered below. Very massive components, or symmetrical components peened on both sides, do not need this distinction.

The form of the longitudinal residual stress pattern is shown in Fig.1, σ_S, σ_{Max} and σ_T being the surface, maximum and tensile stresses respectively. The penetration, Δ, is arbitrarily taken as the depth to the layer of zero residual stress, and a dimension Δ_{Max} is introduced, being the depth to the layer in which maximum residual compression occurs. This does not imply that the peak compression is never at the immediate surface, but does allow information about its position to be treated as one of the variables.

The general shape of this stress distribution is supported by reports such as Mattson (1956) and Lessells and Broderick (1956), both based on flat steel specimens about ¼" thick. Richards (1945) found the same general form in flat light alloy strips, and Föppl (1951) differed only in details when he examined shot-peened steel cylinders. The pattern, then, is simple and similar over a range of conditions, and variation occurs only in the magnitudes of σ_s, Δ etc.

The factors affecting σ_S, Δ etc. fall into two groups: firstly there are the properties of the peened material, including the form of the peened object: secondly the conditions of the applied process must be considered, covering such items as size of shot, blasting pressure or speed, and intensity of exposure. The upper limits of σ_S σ_{Max}, and to a certain extent Δ, will be decided by the properties of the peened material. These are sometimes referred to as "saturation" values, implying correctly that once they are achieved, continuing the peening brings no further benefit. It is not necessarily true, though, that continuing peening has no effect at all, since actual abrasion and damage to the surface can continue after the saturation level has been reached. One of the reasons for studying the effects of the second group of factors, i.e. the "process" variables, is to enable material saturation to be reached as easily and economically as possible, with a minimum of damage to the actual surface, but it would also be useful to have some idea of the stresses induced by less severe peening processes, which do not reach the saturation level. Any real situation will be concerned with the effects of both material and process variables, but it is convenient to deal with them in separate sections, bearing in mind that the former allow estimates to be made of the stress system which can reasonably be achieved, and the latter help to choose circumstances which achieve it.

2.2. The effect of the properties of the peened materials on the induced stress system.

Since plastic deformation is the basis of peening, yield strength and work-hardening capability are of obvious importance, and indentation hardness might be expected to correlate with response to peening. Experiments support these expectations fairly well, although correlation with yield stress is not precise: yield tests at high rates of strain might throw some light on this, but only static yield is usually quoted.

Taking the main characteristics of the stress distribution in turn, then, we have:-

(i) Magnitude of maximum compression, σ_{Max}.

Investigations like those of Mattson (1956) and Lessells and Brodrick (1956) have generally aimed at establishing a ratio of σ_{Max} to σ_Y (yield stress): both used flat plates of medium to high strength steels, and found ratios in the 0.5 to 0.6 region. Almen (1950) gives general support to this figure, and the reports of Foppl (1951) and Hempel (1959) show that cylindrical specimens give similar results, but there is a tendency in much of the literature to imply that lower strength steels may well have higher values of this ratio. Without quoting test results, Fuchs (1963) says that for steels with ultimate tensile strengths between 1000,000 and 3000,000 lb./sq.in., the residual compression is given by

$$\sigma_{Max} = 70{,}000 + 0.2 \text{ U.T.S.}$$

The reason for the correlation with U.T.S., rather than yield stress, is not obvious, but the rule is certainly successful when checked against the reports quoted above. Softer materials such as aluminium alloys, seem capable of developing residual compression at least up to the static yield strength. Richards (1945) peened flat strips of X76S-T, $\frac{1}{4}$" thick, and found compression up to 66,000 lb./sq.in. (c.f. 0.2% proof of 60,000 lb./sq.in.) and Waisman and Phillips (1954) report between 50,000 and 80,000 lb./sq.in. compression in 0.109" thick strips of 75S-T6.