Title
NEAR-THRESHOLD FATIGUE CRACK PROPAGATION: A PERSPECTIVE ON THE ROLE OF CRACK CLOSURE

Permalink
https://escholarship.org/uc/item/8mr83283

Authors
Suresh, S.
Ritchie, R.O.

Publication Date
1983-11-01

NEAR-THRESHOLD FATIGUE CRACK PROPAGATION: A PERSPECTIVE ON THE ROLE OF CRACK CLOSURE
S. Suresh and R.O. Ritchie
November 1983

TWO-WEEK LOAN COPY
This is a Library Circulating Copy which may be borrowed for two weeks. For a personal retention copy, call Tech. Info. Division, Ext. 6782.
DISCLAIMER

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor the Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or the Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or the Regents of the University of California.
NEAR-THRESHOLD FATIGUE CRACK PROPAGATION: A PERSPECTIVE ON THE ROLE OF CRACK CLOSURE

S. Suresh
Division of Engineering, Brown University
Providence, Rhode Island 02912

and

R. O. Ritchie
Materials and Molecular Research Division, Lawrence Berkeley Laboratory,
and Department of Materials Science and Mineral Engineering
University of California, Berkeley, CA 94720

November 1983

To be published in "Fatigue Crack Growth Threshold: Concepts," eds.
D. L. Davidson and S. Suresh, TMS-AIME, Warrendale, PA, 1983

This work was supported by the Director, Office of Energy Research,
Office of Basic Energy Sciences, Materials Science Division of the U.S.
Department of Energy under Contract No. DE-AC03-76SF00098.
NEAR-THRESHOLD FATIGUE CRACK PROPAGATION: A PERSPECTIVE ON THE ROLE OF CRACK CLOSURE

S. Suresh\(^1\) and R. O. Ritchie\(^2\)

\(^1\)Division of Engineering, Brown University, Providence, RI 02912, USA

\(^2\)Materials and Molecular Research Division, Lawrence Berkeley Laboratory, and Department of Materials Science and Mineral Engineering, University of California, Berkeley, CA 94720, USA

Abstract

In recent years, mechanistic and continuum studies on fatigue crack propagation, particularly at near-threshold levels, have highlighted a dominant role of crack closure in influencing growth rate behavior. In this paper we review and model the various sources of closure induced by cyclic plasticity, corrosion deposits, irregular fracture morphologies, viscous fluids and metallurgical phase transformations. It is shown that many of the commonly observed effects of mechanical factors, such as load ratio, microstructural factors, such as strength and grain size, and certain environmental conditions can be traced to the extrinsic influence of closure in modifying the effective driving force for crack extension.
The implications of such closure mechanisms are discussed in the light of constant and variable amplitude fatigue behavior, the existence of a threshold stress intensity for no fatigue crack growth and the validity of such threshold concepts for the case of short fatigue cracks.

Introduction

The elastic-plastic models of McClintock (1), presented twenty years ago, postulated the existence of a propagation threshold for long fatigue cracks. The condition that fatigue cracks cease to propagate when the extent of cyclic plastic zone becomes comparable to some characteristic microstructural size-scale (p), was derived from failure criteria based on the attainment of a critical local strain or cumulative damage over the characteristic dimension p ahead of the crack tip (1). In 1966, Frost and co-workers (2) reported experimental evidence supporting the existence of a fatigue threshold when the value of the parameter (σₐ₃a), where σₐ is the alternating stress and a the crack length, reached a critical value. With the advent of linear elastic fracture mechanics and its adoption to characterize the propagation of fatigue cracks, it soon became apparent, particularly from the work of Paris and co-workers (3), that the threshold for non-propagation of fatigue cracks can be denoted by a critical value of the stress intensity range ∆K, below which crack growth is not experimentally detectable.
The application of such fracture mechanics concepts to fatigue crack propagation, as originally postulated by Paris and co-workers (4), is based on the implicit assumption that the nominal value of crack driving force, taken for the case of linear elastic behavior as the range of stress intensity factor $\Delta K$, uniquely and autonomously characterizes the crack tip stress and strain fields. Although this is clearly a simplification, since the analysis is based on monotonic loading of a stationary crack, using this linear elastic fracture mechanics (LEFM) approach the value of the alternating stress intensity $\Delta K$ can then be taken as a correlator for fatigue crack extension such that:

$$\frac{da}{dN} = C(\Delta K)^m,$$  \hspace{1cm} (1)

where $\frac{da}{dN}$ is the crack propagation rate per cycle, and $C$ and $m$ are scaling constants (4). Experimental studies of fatigue crack propagation in engineering materials, however, have begun to reveal that the premise on which Eq. (1) is based may lead to overestimates of crack advance because of a phenomenon which has become known as crack closure.

Early observations of crack closure can be traced back to the work of Endo et al. (5) who suggested in 1969 that the slower crack growth rates of a low strength steel in water compared to air could be attributed to crack surface corrosion deposits in the former medium which prematurely closed the crack. Endo et al.'s results (5)
were presented in terms of conventional stress-life approach and LEFM terminology was not used. The first mechanistic justification for crack closure in terms of fracture mechanics concepts was presented by Elber (6) in 1971. Based on compliance measurements of fatigue cracks at high $\Delta K$ levels, Elber (6) proposed that premature contact between the crack faces can occur even during the tensile portion of the fatigue cycle because of the permanent residual displacements, arising from prior plastic zones, left in the wake of a growing fatigue crack. Since the crack cannot propagate whilst it remains closed, the consequence of such plasticity-induced crack closure (6) is to modify the actual stress intensity factor range experienced by the crack tip from a nominal value of:

$$\Delta K = K_{\text{max}} - K_{\text{min}} \quad (2)$$

based on global measurements of geometry, applied loads and crack size to some near tip effective value of:

$$\Delta K_{\text{eff}} = K_{\text{max}} - K_{\text{cl}} \quad (3)$$

Here, $K_{\text{max}}$ and $K_{\text{min}}$ are the maximum and minimum stress intensity factors, respectively, during the fatigue cycle and $K_{\text{cl}}$ is the stress intensity factor at which the two fracture surfaces first come into contact during the unloading portion of the fatigue cycle. It was postulated by Elber (6) that fatigue crack propagation rate is then
governed by the effective value of the stress intensity factor range $\Delta K_{\text{eff}}$, such that:

$$\frac{da}{dN} = C(\Delta K_{\text{eff}})^m = C(U\Delta K)^m,$$

where $U$ is an empirically-measured value defined as $\Delta K_{\text{eff}}/\Delta K$.

Since Elber's pioneering work, the role of crack closure in influencing fatigue crack growth has remained a topic of considerable research interest and practical significance, although it has not been until fairly recently that microscopic descriptions of closure have been developed. Despite the application of crack closure arguments to rationalize a number of fatigue characteristics in engineering materials, several independent experimental studies have revealed convincing evidence that the plasticity-induced crack closure phenomenon plays a role principally under plane stress loading conditions (e.g., refs. 7 and 8). Yet, there has long been considerable information in the literature indicating that marked crack closure can occur even in the near-threshold regime where predominantly plane strain conditions exist (e.g., ref. 9). Such apparently anomalous and conflicting experimental data led to the conclusion by Ritchie (10) that plasticity-induced crack closure must be questioned as an important mechanism for near-threshold fatigue crack growth. More recently, finite element studies by Newman (11) showed that plasticity-induced closure, modelled for plane strain conditions, was inadequate to explain the marked role of load ratio
for fatigue crack growth at near-threshold levels. The reasons for this apparent paradox have only recently become evident from the observations that Elber's mechanism involving cyclic plasticity is not the sole source of closure. Numerous independent experimental studies have clearly identified several other sources of crack closure, which can become of major importance particularly at near-threshold growth rates. These include, closure arising from i) corrosion deposits formed within the crack (oxide-induced crack closure), ii) irregular crack surface morphologies (roughness-induced crack closure), iii) viscous media penetrated inside the crack (viscous fluid-induced crack closure) and iv) phase transformations due to crack-tip plasticity (transformation-induced crack closure), as illustrated schematically in Figure 1.

In this paper, we summarize, in as quantitative terms as possible, the various micro-mechanisms underlying different crack closure processes and assess the significance of such mechanisms to the propagation of near-threshold fatigue cracks. The implications of crack closure phenomena are also discussed in the light of various aspects of fatigue behavior involving short cracks, variable amplitude loading, as well as the question of the measurement and very existence of a fatigue threshold.
Description of Closure Mechanisms

Plasticity-Induced Crack Closure

Plasticity-induced closure (Fig. 1a) is considered to arise from the elastic constraint in the wake of the crack tip of material elements permanently stretched within prior plastic zones, leading to an interference between mating crack surfaces (6). Based on experimental compliance measurements from fatigue tests conducted at growth rates above $10^{-6}$ mm/cycle in 2024-T3 aluminum alloy, Elber (6) proposed the following empirical relationship between the ratio of the closure stress intensity to maximum stress intensity ($K_{cl}/K_{\text{max}}$) and load ratio, $R$:

$$\frac{K_{cl}}{K_{\text{max}}} = 0.5 + 0.1R + 0.4R^2$$

(5)

It should be noted, however, that Eq. (5) pertains to crack growth at stress intensities much greater than $\Delta K_o$, where deformation conditions approach that of plane stress, especially at the higher load ratio values. However, since the mechanism results from the constraint of surrounding elastic material on the plastic zone enclave surrounding the crack, such closure will become insignificant at very high stress intensity levels once the material becomes fully plastic. Based on extensive crack closure measurements in aluminum alloys, Schijve (12) modified Elber's empirical model to account for
load ratio effects at the higher $\Delta K$ values by incorporating higher order terms in $R$ such that:

$$\frac{K_{cl}}{K_{max}} = 0.45 + 0.2R + 0.25R^2 + 0.1R^3 \quad (6)$$

For such plane stress crack growth, physically-based analytical models have been developed by Budiansky and Hutchinson (13) and by Kanninen and Atkinson (14), whereas Newman (15,16) has developed numerical models involving finite-element analyses of crack tip deformation to represent plasticity-induced closure under both plane stress and plane strain conditions.

Despite the lack of convincing analytical formulations for plane strain closure, Schmidt and Paris (9) used crack closure concepts* to account for the dependence of $\Delta K_0$ on $R$ (Fig. 2). The variation of threshold $\Delta K_0$ with $R$ was found to show two distinct regimes when the following assumptions were considered: a) the closure stress intensity $K_{cl}$ at the threshold is independent of $R$ and b) a constant effective threshold stress intensity range $\Delta K_{th}$ is necessary to produce fatigue crack growth. As illustrated schematically in Figure 2, and by experimental results on 2 1/4Cr-1Mo steels in Figure 3, at low load ratios where $K_{min} \leq K_{cl}$, the maximum stress intensity at the threshold ($K_{o,max}$) appears constant, since:

$$K_{o,max} = K_{cl} + \Delta K_{th}, \quad K_{min} \leq K_{cl} \quad (7)$$

---

*The exact nature of the crack closure process was not defined (9), although it was presumably related to plasticity-induced closure.
whereas at high R values where $K_{\text{min}} \geq K_c$, the threshold stress intensity range ($\Delta K_o$) appears constant, since:

$$\Delta K_o = \Delta K_{\text{th}}, \quad K_{\text{min}} \geq K_c.$$  \hspace{1cm} (8)

Such results imply that there exists a critical value of the load ratio, $R_{\text{cr}}$, above which crack closure effects become insignificant. In Figure 2,

$$R = R_{\text{cr}} \text{ when } K_{\text{min}} = K_c, \text{ such that } \Delta K_{\text{th}} = \Delta K_o.$$  \hspace{1cm} (9)

The nature of load ratio dependence of $\Delta K_o$, as predicted by the Schmidt and Paris model (9), has been found consistent with the experimental results obtained for aluminum alloys (9) and steels (17,18), as shown in Figure 3. The model, however, does not account for the significant effects of environmental and microstructural factors on near-threshold crack propagation rates and on the value of $\Delta K_o$, as indicated in Figure 3 by the difference in behavior between moist air and dry hydrogen. Moreover, there are existing experimental data, notably those by Lindley and Richards (7), which indicate that plasticity-induced closure plays a far less significant role in influencing fatigue crack propagation rates under plane strain conditions, such as those at near-threshold stress intensities. Thus, it is clear that the effects of various mechanical, environmental and microstructural variables on near-
threshold crack growth rates cannot be explained solely on the basis of plasticity-induced crack closure and that other sources of crack closure need to be taken into consideration.

Oxide-Induced Crack Closure

It has long been realized that corrosion products formed within growing cracks (Fig. 4) can significantly influence their subsequent crack extension rates. Endo et al. (5) published early observations of slower crack growth in aqueous media than in gaseous environments and attributed the effect to the presence of oxide layers between the fracture surfaces, although their results were presented only in terms of crack length vs. number of cycles data. A number of independent studies, notably those by Paris et al. (19), Lindley (20), Tu and Seth (21), Skelton and Haigh (22), Kitagawa and co-workers (23), Stewart (24) and Suresh and Ritchie (17,25,26), have, since then, suggested the possibility of crack closure due to corrosion debris influencing near-threshold growth rates and threshold $\Delta K_0$ values. Suresh, Ritchie and co-workers (25) reported the first experimental attempts to quantify the effects of corrosion deposits on fatigue crack propagation rates, especially in the near-threshold regime where the extent of oxide formation within the crack (estimated with the aid of scanning Auger spectroscopy) was found to be comparable to the scale of crack tip opening displacements.

The concept of oxide-induced crack closure (Fig. 1b) (17,24,25) is based on the phenomenon that at low load ratios, near-threshold
growth rates are significantly reduced in moist environments (such as air or water), compared to dry environments (such as hydrogen or helium gas), due to the presence of corrosion deposits on crack faces. Moist atmospheres lead to the formation of oxide layers within the crack, which are thickened at low load ratios by "fretting oxidation," i.e., a continual breaking and reforming of the oxide scale behind the crack tip due to the repeated contact between the fracture surfaces arising from Mode II displacements and plasticity-induced closure (17,25,26). Such oxide debris, whose formation is limited in dry, moisture-free environments or at high load ratios (where there is no plasticity-induced closure), provides a mechanism for enhanced closure by resulting in an earlier contact between the fracture surfaces through an increased closure stress intensity $K_{c1}$. The formation of corrosion deposits and the process of oxide-induced crack closure are promoted by a) small crack tip opening displacements (CTOD), such as in the near-threshold regime, which are comparable to the thickness of the excess debris within the crack (25), b) highly oxidizing media (such as water), where substantial thermal oxidation is possible even in the absence of fretting (18), c) low load ratios, which facilitate repeated contact (and hence fretting) between the fracture surfaces through small CTOD values (17,24-26), d) rough fracture surfaces, which at low $\Delta K$ values, promote relative sliding and rubbing between mating crack faces (25) and e) lower strength materials, where the plasticity-induced closure
mechanism (and hence fretting) is dominant and the base metal is soft enough to undergo considerable fretting damage (17,24-26).

The concept of oxide-induced crack closure has been successfully employed by a number of investigators to rationalize, either wholly or partially, the near-threshold corrosion fatigue characteristics in a wide range of engineering materials comprising ferritic-pearlitic (26), ferritic-bainitic (17), fully bainitic (25) and fully martensitic (26,27) steels, as well as nickel-base superalloys (28,29), copper (30) and aluminum alloys (31). In steels, the influence of oxide-induced closure is found to be inversely related to strength level. In lower strength steels (yield strength ≤ 1000 MPa), closure mechanisms appear to dominate over conventional corrosion fatigue processes, such as hydrogen embrittlement or anodic dissolution, in influencing near-threshold environmental behavior (25,26), whereas in ultrahigh strength steels the reverse is found to be true (32). In contrast to such behavior for steels, the work of Vasudevan and Suresh (31) has shown that the crack tip oxidation behavior of high strength aluminum alloys (e.g., 7075) at low ΔK levels is very sensitive to aging treatment, processing methods, and composition even for alloys which fall in a narrow range of yield strength values. This is presumably due to the dependence of oxidation on the extent of copper in solution as well as on the differences in the tenacity of the oxide layers produced in different aging conditions (31).
Order of magnitude estimates of the extent of oxide-induced crack closure have been reported by Suresh and Ritchie (26,33) by considering stress intensity solutions for a rigid wedge inside a linear elastic crack (Fig. 5). Assuming only a mechanical closure phenomenon arising from the presence of the excess oxide debris on the crack faces and ignoring plasticity and hysteresis effects, the closure stress intensity value \( K_{cl} \) at the crack tip can be obtained for plane strain conditions as (26,33):

\[
K_{cl} \bigg|_{x = 0} = \frac{d_0 \varepsilon}{4\sqrt{\pi x}(1 - v^2)},
\]

where \( d_0 \) is the maximum excess oxide thickness, \( 2\ell \) the location behind the tip corresponding to the thickest oxide formation and \( E/(1 - v^2) \) the effective Young's modulus in plane strain. However, due to the assumptions pertaining to the rigidity of the oxide wedge, the non-uniform thickness of the oxide deposits along the crack length and along the crack front and the uncertainties in the estimates of \( d_0, \ell \) and \( \delta_{max} \), Eq. (9) is realistically capable of providing only a crude description of the extent of oxide-induced closure (33).

**Roughness-Induced Crack Closure**

A third source crack closure can arise at near-threshold stress intensities due to the rough nature of the fracture surface (Fig. 1c) when crack tip opening displacements become comparable to the size-scale of the fracture surface asperities (Fig. 6). While it is
widely acknowledged that crack propagation in the Paris regime (i.e., typically $10^{-6} \leq \frac{da}{dN} \leq 10^{-3}$ mm/cycle) occurs by a striation mechanism induced by concurrent or alternating crack tip shear (Stage II in Forsyth's terminology (34)). There is growing evidence that near-threshold crack advance in certain microstructures takes place primarily along a single active slip system (Stage I in Forsyth's terminology (34)) (35-37). Such single-shear Stage I growth, which occurs primarily where maximum plastic zone sizes at near-threshold growth rates are typically smaller than the significant microstructural unit, e.g., the grain size, results in crystallographic or generally faceted fracture features, an irregular surface morphology and locally mixed-mode crack growth (Fig. 7). While an ideally elastic crack unloads reversibly and does not undergo any premature closure, the irreversible nature of inelastic crack tip deformation and the possibility of slip-step oxidation in moist environments, in reality, lead to mismatch between the fracture surface asperities during the unloading portion of the fatigue cycle (38). The occurrence of such non-uniform crack tip opening and closure and the resulting Mode II crack opening displacements have been experimentally verified by Davidson (39) using in situ observations of cyclic deformation in the scanning electron microscope. It was first suggested by Purushothaman and Tien (40) and by Walker and Beevers (41) that the interference between the mating fracture surface asperities behind the crack tip can lead to apparently higher (roughness-induced) crack closure loads and
consequently, lower crack propagation rates. Minakawa and McEvily (35) have, subsequently, presented a qualitative model schematically describing the Mode II and Mode I displacements accompanying near-threshold crack advance (Fig. 8).

Roughness-induced crack closure is promoted by a) small plastic zone sizes at the crack tip (e.g., typically less than a grain diameter), which induce single shear mechanisms, b) small crack tip opening displacements, which are of a size scale comparable to the asperity height (such as in the near-threshold regime) (25), c) coarse-grained materials and microstructures with coherent and shearable precipitates capable of inducing coarse planar slip (43-45), d) crack deflection mechanisms (38), induced by stress-state, environment, microstructure or load excursions and the resulting non-linear crack profiles (46), e) inelastic crack tip deformation and oxidation of slip steps in moist environments, both of which can lead to mismatch between the fracture surfaces during the decreasing portion of the fatigue loading cycle (38) and f) low load ratios where the minimum crack tip opening displacements may be substantially smaller than the size of the surface asperities (18,25).

Estimates of (roughness-induced) closure stress intensities were first reported by Purushothaman and Tien (40) who denoted the fracture surface undulation by:

$$h = h_0 \exp(\varepsilon_f)$$  \hspace{1cm} (11)
where \( h_0 \) and \( h \) are length dimensions describing the initial and final fracture surface roughness, and \( \varepsilon_f \) the true fracture ductility representative of the stress state ahead of the crack tip. By equating the change in asperity height \( \alpha(h - h_0) \), where \( \alpha \) is an empirical parameter of value less than unity, to the crack tip opening displacement, estimates of \( K_{c1} \) were derived (40). This model, however, does not incorporate the role of Mode II displacements, which are typical of near-threshold crack advance, in influencing roughness-induced closure.

In order to estimate the influence of Mode II displacements and fracture surface roughness on closure stress intensity levels, Suresh and Ritchie (47) developed a simple, two-dimensional geometric model. By assuming equal-sized asperities, the non-dimensional closure stress intensity at the point of first asperity contact was derived to be (47):

\[
\left( \frac{K_{c1}}{K_{\text{max}}} \right) = \sqrt{\frac{2\gamma x}{1 + 2\gamma x}},
\]

(12)

where \( \gamma \) is the non-dimensional surface roughness parameter taken as the ratio of the height, \( h \), to the width, \( w \), of the asperities and \( x \) the ratio of the Mode II to Mode I crack tip displacements \( (u_{II}/u_{I}) \). Although Eq. (11) must only be considered as a first order model for roughness-induced crack closure, it does provide quantitative estimates for the extent of such closure in terms of the degree of fracture surface mismatch and the relative magnitude of the crack tip
shear displacements. Such estimates for $K_{c1}/K_{max}$ as a function of $\gamma$ are shown in Fig. 9 (47), and indicate reasonably good agreement with experimental results derived from data (35,43) on a 1018 mild steel and a fully pearlitic rail steel.

**Viscous Fluid-Induced Crack Closure**

An additional source of closure can arise from the presence of a viscous fluid penetrated within a growing crack (Fig. 1d). Due to the approaching velocity of the crack walls, the presence of the fluid can give rise to a hydrodynamic wedging action counteracting the closing of the crack (Fig. 10). As experimentally demonstrated by Endo et al. (48,49) and more recently by Tzou et al. (50), this mechanism of viscous-fluid-induced crack closure becomes relevant for fatigue crack propagation in oil environments, where the rate of crack extension has been found to be sensitive to the kinematic viscosity ($\eta$) of oil. As shown in Figure 11, fatigue crack growth rates in chemically-inert oil environments at low load ratios tend to be in excess of those in moist air at near-threshold levels, due to a suppression of oxide-induced closure. However, such growth rates are slower than in moist air above $\sim 10^{-6}$ mm/cycle, due to a minimal effect of corrosion fatigue mechanisms (active path corrosion/hydrogen embrittlement), and are found to be faster in oils of higher viscosity (50) in the present case of a bainitic steel tested at 50 Hz in silicone and paraffin oils. At high load ratios,
however, the influence of viscosity on growth rate behavior largely disappears (50).

The hydrodynamic wedging effect of a viscous fluid was first quantitatively analyzed by Endo et al. for fluid fully penetrated inside the crack (48, 49). Although not presented in fracture mechanics terms of resultant stress intensities, these authors found that the pressure distribution due to the oil as a function of distance, $x$, along a crack of length, $a$, could be presented in terms of the mouth opening displacement, $h$, the density of the fluid, $\rho$, and the angular closing velocity of the crack walls $\left(\frac{d\Theta}{dt}\right) = \left(\frac{1}{a}\right)\left(\frac{dh}{dt}\right)$, viz (49):

$$p(x) = \frac{6n\rho a^2}{h^3} \frac{dh}{dt} \log\left(1 - \frac{x}{a}\right). \quad (13)$$

Converting this pressure into an effective stress intensity $K_p$ due to the presence of the viscous fluid using standard linear elastic $K_I$ solutions indicates that the magnitude of the fluid-induced closure should increase with increasing viscosity, yet the experimental data in Figure 11 show an opposite trend in that the highest growth rates are observed with the highest viscosity oils. This discrepancy can be explained by recognizing that the fluid pressure induced by the higher viscosity oils is reduced by the kinetic limitations of such oils in penetrating the crack. Accordingly, Tzou et al. (51) proposed an alternative quantitative description for viscous-fluid-induced crack closure based on the partial penetration of the oil.
inside the crack. Here the penetration distance $d$ is computed as a function of time $t$ from capillary flow analysis as:

$$d = \left[ \frac{\gamma_L \cos \beta}{3np} \int_0^t \bar{h} \, dt \right]^{\frac{1}{3}} ,$$  

(14)

where $\gamma_L$ is the surface tension, $\beta$ the wetting angle and $\bar{h}$ the average crack opening, such that the subsequent pressure distribution within the crack becomes (51):

$$p(x) = \frac{6np}{h^3} \frac{d\bar{h}}{dt} (d - x) ,$$  

(15)

where $v = \frac{d\bar{h}}{dt}$ is the approaching velocity of the crack walls (Fig. 10). The extent of viscous-fluid-induced closure is then determined numerically by superposition of the resultant $K_p$ values to the applied stress intensities from simultaneous solution of Eqs. (14) and (15). Based on the experimental data shown in Figure 11, this analysis indicated that in the lower viscosity oils (e.g., 5 to 75 cS), where the oil fully penetrated the crack, the extent of closure at $R = 0.05$ was characterized by $K_p/K_{\text{max}}$ values of between 0.25 and 0.50 at $\Delta K$ levels between 10 and 20 MPa$\sqrt{m}$. In the higher viscosity oils (e.g., $\geq 12500$ cS), however, where only partial penetration occurred the extent of closure was somewhat less in that $K_p/K_{\text{max}}$ values were estimated to lie between 0.22 and 0.43. Although the penetration rates of the fluids are much more rapid at high load.
ratios, no closure was predicted for $R = 0.75$ because of the much larger crack openings there.

Such closure is clearly promoted by low load ratios and is affected by such factors as oil viscosity and chemistry, test frequency and strength level/elastic modulus (through their influence on the extent of crack opening). However, it is difficult to simply indicate a trend for the latter variables since although the maximum extent of closure, i.e., $K_p$ values for full oil penetration, is enhanced by increasing viscosity, increasing frequency and decreasing crack opening, such factors concurrently act to minimize the penetration of the oil. In view of the numerical estimates of steady state values of $K_p$ under the equilibrium conditions of infinite time and full penetration (51), it would appear that closure induced by viscous fluids is a less potent mechanism than that generated by oxide debris or irregular crack paths.

**Phase Transformation-Induced Crack Closure**

In materials which undergo stress- or strain-induced phase transformations under cyclic loading, a further mechanism of fatigue crack closure can result from the so-called TRIP effect, i.e., arising from Transformation Induced Plasticity (Fig. 1e). In analogous fashion to the increase in fracture toughness observed under monotonic loading in both metals (52), e.g., metastable austenitic stainless and high strength TRIP steels, and ceramics (53), e.g., tetragonal zirconia, due to deformation-induced
martensitic transformations, the constraint of surrounding elastic material on regions which have transformed and thus undergone a positive volume change will place such regions under compression. Once the crack penetrates these transformed regions, generally corresponding to some process zone at the crack tip, they will act to close the crack and hence reduce the nominal stress intensity to some lower effective value at the tip.

Such transformation-induced closure will be promoted by metallurgical phase changes which show a large increase in volume (i.e., the martensitic transformation in steels is typically 4%), and by conditions which enhance the transformation, i.e., lower temperatures, higher strain rates, etc. However, analogous to the plasticity-induced mechanism, the closure effect arises from the constraint of surrounding elastic material on the transformed regions such that if the material is completely transformed prior to cracking, no closure effect will be felt. Furthermore, the transformation of regions remote from the tip will similarly not cause closure; the effect will only be experienced as the crack penetrates and is enveloped by the transformed region.

Rigorous analysis of the TRIP effect on crack closure has yet to be analyzed for deformation conditions involving deviatoric strains or cyclic inelasticity (54). However, McMeeking and Evans (55) and Budiansky et al. (56) have analyzed the effect under monotonic loading for purely stress-induced dilatant transformations in
ceramics and shown the reduction in stress intensity due to closure to be given by:

\[ K_I - K_{eff} = \frac{0.22}{(1 - \nu)} V_f \epsilon_T E \sqrt{r_t} \]  

(16)

where \( K_I \) and \( K_{eff} \) are the nominal (applied) and effective (near tip) stress intensities, \( E \) the elastic modulus, \( \epsilon_T \) the transformation strain, \( V_f \) the volume change and \( r_t \) is the width of the transformed zone. It is worth noting that the above analyses indicate a resistance curve effect in that the full influence of transformation-induced closure is only felt when the crack has penetrated the transformation region to a distance comparable with approximately 5 times the width of the zone \( r_t \) (55). This highlights a common characteristic of all closure mechanisms in that closure acts in the wake of the crack tip. Thus, for small cracks, which by definition have a limited wake (i.e., in the present case, of a length less than 5\( r_t \)) the extent of closure will be lessened and hence the effective driving force at the crack tip will be correspondingly larger.

**Implications of Crack Closure**

The implications of such fatigue crack closure mechanisms are now considered for the complex and often seemingly inconsistent nature of near-threshold fatigue crack growth. The understanding of the many sources of closure outlined above has enabled a much more
consistent explanation to be obtained for the effects of many mechanical, microstructural and environmental factors which have been documented over the years to influence near-threshold growth rates and the value of the fatigue threshold. Although in most instances, direct experimental closure measurements have confirmed a primary role of crack closure, a common trend can be seen in that where closure is dominant, the effect of these factors, whether they involve strength, grain size or moist environments, etc., will be most pronounced at low load ratios with long cracks. In fact, when assessed at high load ratios or with small cracks, the effect may well be severely diminished, non-existent or even different since under these conditions closure mechanisms cannot be so effective. We now examine several of the major variables which affect thresholds and discuss the respective implications of closure.

**Mechanical Variables**

**Load Ratio.** Of the mechanical factors influencing fatigue crack growth, clearly one of the most important is that of mean stress, characterized in terms of the load or stress ratio as \( R = \frac{K_{\text{min}}}{K_{\text{max}}} \). Numerous experimental investigations, reviewed in refs. 10 and 18, have shown that with increase in \( R \) crack growth rates for tension/tension loading are enhanced, although the effect generally predominates only at very high growth rates close to final instability (i.e., as \( K_{\text{max}} \) approaches \( K_{\text{IC}} \)) and at near-threshold levels (i.e., as \( \Delta K \) approaches \( \Delta K_{\text{o}} \)) (10). The former effect of load
ratio in accelerating growth rates close to instability is usually associated with the occurrence of "static" fracture modes, i.e., cleavage, intergranular and fibrous fracture, in addition to ductile striation growth (57). However, the near-threshold load ratio effect, shown in Figure 12 for SA542-3 steel tested in laboratory environment (18), occurs without dramatic changes in fractography and is attributed primarily to mechanisms of crack closure.

The analysis of Schmidt and Paris (9) reviewed in the section on plasticity-induced crack closure clearly shows the trend of closure effects on the threshold (Fig. 2) in that as the mean level is raised, the role of $K_c$ in modifying the effective $\Delta K$ is lessened. Studies on the role of load ratio in influencing $\Delta K_0$ values in dehumidified gaseous hydrogen, as compared to moist air atmospheres, show behavior similar to that predicted by Schmidt and Paris (9). The critical load ratio $R_{cr}$ above which $\Delta K_0$ values are insensitive to $R$, however, is lower in the hydrogen environments (Fig. 3). As $R_{cr}$ represents the load ratio above which closure effects are insignificant, such behavior can be attributed in this case to a much smaller effect of oxide-induced crack closure, consistent with the reduction in crack surface corrosion debris with increasing $R$. In hydrogen, which merely plays the role of a dry environment in low strength steels, oxide-induced closure is less significant (18). Such an interpretation of a lessening effect of closure with increasing $R$ and in dry environments is consistent with numerous experimental closure measurements of $K_c$. An example of such
results, relevant to the threshold data in Figure 3, is presented in Figure 13, taken from an assessment of $K_{c1}$ values at $\Delta K_0$ in a bainitic steel using an ultrasonic technique (26). It is apparent that the opening of the crack is most delayed at $R = 0.05$ in moist air where both plasticity-induced and oxide-induced closure can occur.

The load ratio dependence of fatigue thresholds is often quite different to that depicted in Figure 2 in completely inert or chemically very aggressive environments. In either case the reduction in $\Delta K_0$ with increase in $R$ is far less marked compared to moist air (58,18); observations which are again consistent with oxide-induced closure concepts. In inert environments, such as high vacuum, the shallow dependence of $\Delta K_0$ on $R$ has been attributed to the fact that little or no corrosion deposits are formed at any load ratio such that, unlike moist air, the magnitude of oxide-induced closure does not vary with $R$ (18). Similarly in aggressive environments, such as water (Fig. 14), fracture surface oxide deposits can form and thicken at all load ratios by natural thermally-activated mechanisms so that the extent of oxide-induced closure does not decay with increasing $R$ (18).

There are several other instances where the load ratio dependence of fatigue thresholds is decreased due to a reduced role of closure. In addition to behavior in inert or very aggressive environments (18) discussed above, the variation in $\Delta K_0$ with $R$ is significantly less marked in ultrahigh strength steels (59), where
oxide-induced closure effects are minimized (32), and in finer grained materials (60), where roughness-induced closure effects are minimized (37). Such behavior has been described in detail in ref. 18.

Behavior at negative load ratios is far less understood. However, the indications are that in certain materials near-threshold growth rates at $R = -1$ may marginally exceed corresponding growth rates at $R = 0$ (e.g., refs. 61, 62). Based on limited experiments closure measurements (62), it would appear that such behavior is associated with a reduction in crack closure due to "smoothening" of asperity contacts behind the crack tip from compressive contact between crack surfaces.

Cyclic Frequency. Studies by Schmidt and Paris (9) in 2024-T3 aluminum alloy first revealed that the threshold stress intensity value is significantly influenced by the cyclic frequency, although no definitive trend could be observed for the variation of $\Delta K_0$ with frequency. More recently, Bignonnet et al. (63) found lower $\Delta K_0$ values and correspondingly less oxide debris during low load ratio tests in a quenched and tempered low alloy steel tested at a frequency of 7 Hz as compared to 65 Hz. The mechanism by which increasing frequency enhances the formation of crack surface corrosion deposits (and hence promoted oxide-induced crack closure) was not specified (63). The observations of Bignonnet et al. (63) are consistent with the results for 7475-T7 aluminum alloy (64) where the $\Delta K_0$ values and oxide thickness at $\Delta K_0$ are found to be higher at
200 Hz than at 25 Hz for tests conducted in 95% humid air. Studies over a wider range of frequencies (5 to 500 Hz) in SA542-3 steel, however, have failed to substantiate this effect above 50 Hz, although oxide thicknesses and $\Delta K_0$ values were somewhat smaller at 5 Hz (65).

Microstructural Variables

It has long been known that microstructural variables, such as grain size, precipitate distribution and morphology, slip characteristics and duplex structures, can have a more significant influence on fatigue crack propagation behavior in the near-threshold regime than at higher growth rates (e.g., refs. 8-10,37,41,43,45-47,64). Furthermore, in most instances it has been shown that such microstructural effects on near-threshold crack growth can be significantly lessened for tests at high, as opposed to low, load ratios (10,37,43,45,47). The reasons for this prominence of microstructural-sensitive behavior for low load ratio, near-threshold conditions is again largely a consequence of closure and in some instances related mechanisms of crack deflection (38). We now examine the role of each of these metallurgical factors in turn:

Strength Level. Early studies of near-threshold fatigue behavior in ferrous alloys (59,67,68) indicated a large influence of strength in increasing ultra-low growth rates and decreasing $\Delta K_0$ (Fig. 15). This effect, however, is less prominent at high load ratios (59), for small cracks (69) and for non-ferrous alloys (70).
The major role of yield strength in influencing near-threshold fatigue crack growth in ferrous alloys is consistent with a primary role of closure. Recent studies (32) in steels over a very wide range of strength levels (i.e., 290 to 1740 MPa) have confirmed that the extent of crack surface corrosion debris and hence oxide-induced closure is sharply decreased with increasing strength at $R = 0.05$ (Fig. 16). This was attributed to a reduced role of plasticity-induced closure (which promotes crack surface contact to enhance fretting oxidation) in higher strength steels and by the fact that the oxide products formed will be less potent in causing more fretting debris on the harder substrate (24,25,27,32). Furthermore, it has also been pointed out that higher strength steels are more prone to environmental effects (10) and that the microstructures typically utilized for highest strength levels, i.e., tempered martensites, are generally of considerably finer scale than for lower strength structures, such that with the resulting smoother fracture surfaces, roughness-induced closure may also be lessened (66).

The effect of strength level on near-threshold crack growth in non-ferrous alloys is somewhat more complex in that, although behavior is similar to steels in copper alloys (30), there is no marked trend of increasing or decreasing $K_0$ values with change in strength level in precipitation hardened systems (70). In aluminum alloys, this has recently been interpreted in terms of aging treatment, processing methods, composition and slip mode, involving a mutual competition between resistance to corrosion fatigue mechanisms
and oxide-induced closure (which are promoted in overaged structures) and roughness-induced closure (which are promoted in underaged structures) (31, 44, 45, 76). Thus in view of the different hardening mechanisms responsible for a given yield strength in such alloys (e.g., shearing versus by-passing of precipitates) and the effect such slip characteristics can have on resultant closure mechanisms (e.g., 31, 45), it is inappropriate to identify a direct correlation of yield strength and near-threshold behavior in non-ferrous systems.

**Grain Size.** Numerous authors have shown that near-threshold crack growth resistance can be markedly enhanced through coarsening grain size in a wide range of materials including steels (e.g., 10, 43, 68), titanium alloys (e.g., 71-74) and aluminum alloys (e.g., 75, 76). Again closure mechanisms appear to play a dominant role since the beneficial effect of coarse grain sizes is generally much reduced for tests at high load ratios (e.g., 37, 43).

As discussed in the section on roughness-induced crack closure, the reduction in near-threshold growth rates in coarse grain materials has been attributed to enhanced crack closure (Fig. 17), specifically roughness-induced (25, 37, 43, 47, 72). This is due to the fact that when maximum plastic zone sizes are typically less than a grain diameter, the preferred single shear mode of crack extension promotes a more crystallographic or faceted crack path leading to a wedging of the crack from fracture surface asperity contact (35, 41, 47).
This effect is enhanced by mechanisms which promote greater crack deflection at each grain boundary, since they will not only induce a rougher fracture surface but also reduce local crack tip stress intensities from deflection mechanisms (38). As discussed below, planar slip materials (44,45,74,77) and duplex microstructures (46), consisting of both hard and soft phases, offer the best potential in this regard. However, the effect is likely to be much reduced in inert environment where the lack of oxidation of exposed fracture surface at the crack tip will aid reversibility of slip there, thereby reducing the extent of closure (76).

**Duplex Microstructures.** Duplex microstructures seem to provide the most effective means to promote roughness-induced crack closure. Here dispersions of hard and soft phases can lead to significant crack deflection resulting in the formation of tortuous crack paths. Notable examples of this effect have been shown in certain titanium alloys (e.g., 78), and particularly in dual-phase steels (46,79-81). For the latter class of materials, duplex ferritic-martensitic microstructures were first shown by Suzuki and McEvily (83) to provide a metallurgical means of increasing fatigue threshold values in mild steel while simultaneously increasing strength. More recent work (46) in Fe/2Si/0.1C steel employing coarse or fine fibrous martensite dispersions in a coarse ferrite matrix has shown that the threshold \( \Delta K_0 \) values can be raised to above 20 MPa\( \sqrt{\text{m}} \) (at \( R = 0.05 \)) in steels with yield strength in excess of 600 MPa. Based on quantitative metallography (46,83,84) and experimental closure measurements (46),
such unusually good fatigue properties could be traced to significant crack deflection which induces a meandering crack path morphology which in turn leads to a dominant role of roughness-induced closure at low load ratios (Fig. 18). Since closure is a primary factor in causing the superior fatigue crack growth resistance, it is totally consistent that the superior properties attained with the best duplex microstructures are not nearly as appreciable at high load ratios (46) or in short crack/fatigue limit tests (82). In fact, the fatigue threshold properties for the duplex structures in the latter two cases was found to be no better than that measured in conventionally heat-treated, normalized structures.

**Slip Characteristics.** In addition to the microstructural factors listed above, it has been shown, principally in titanium (71-74) and aluminum (44,45,76,77) alloys that microstructures which promote coarse planar slip, as opposed to wavy slip, generally give rise to superior near-threshold fatigue crack growth resistance. Such microstructures are found in materials of low stacking fault energy or in dispersion hardened systems employing small coherent precipitates, and under near-threshold fatigue conditions give rise to coarse faceted fracture surfaces which enhance roughness-induced crack closure (Fig. 17). A good example of this effect can be seen by comparing underaged with overaged microstructures (at the same strength level) in Al-Zn-Mg-Cu alloys (45). Here the underaged structures, which are hardened by small shearable coherent precipitates, show extensive planar slip, coarse faceted fracture
surfaces and hence greater roughness-induced closure compared to overaged structures, where the wavy slip characteristics, i.e., hardening through Orowan bypassing of large coherent precipitates, lead to a more planar fracture morphology. Since such roughness-induced closure mechanisms are active where the fracture surface roughness is comparable to the crack tip displacements, it is interesting to note that the beneficial near-threshold fatigue crack growth resistance of underaged compared to overaged structures is not retained at high propagation rates above $10^{-6}$ mm/cycle where the role of closure is severely diminished (45).

Environmental Variables*

Gaseous Environments. In general, environmentally-influenced crack extension at near-threshold levels can be considered in terms of a mutual competition of two basic mechanisms, namely corrosion fatigue processes, such as hydrogen embrittlement or active path corrosion which accelerate growth rates and the resultant closure mechanisms which decelerate growth rates. Where closure mechanisms dominate, such as at ultralow growth rates in lower strength, coarse-grained materials at low load ratios, environmentally-influenced fatigue behavior may be different than that conventionally reported for higher growth rates above typically $10^{-6}$ mm/cycle. For example, for lower strength steels tested at low load ratios at the

*In general, the precise influence of environment on crack growth is very specific to the particular material/environment system. Below we merely discuss the role of simple environments, e.g., air, water, hydrogen, etc., in affecting the behavior of engineering alloys.
frequencies typically utilized for low growth rate testing (i.e., above ~20 Hz), near-threshold growth rates have been observed to be faster in dry inert environments, such as helium gas, compared to seemingly more aggressive environments such as moist air (Fig. 19) (25,26). This follows because at such high frequencies in lower strength steels, the susceptibility to hydrogen embrittlement from external environments is kinetically limited by the entry of hydrogen, and correspondingly moist environments cause a deceleration in near-threshold growth rates due to enhanced oxide-induced closure from fretting corrosion deposits (Fig. 14) (17-27). As noted in the section on oxide-induced closure, this dominant role of closure has been verified by quantitative assessment of crack surface oxide film profiles (25,26,30-33) and by experimental closure measurements (26). In such lower strength ferrous systems, near-threshold crack growth has been consistently shown to be slower in moist gaseous atmospheres, such as moist air and wet gaseous hydrogen, compared to dry atmospheres, such as gaseous hydrogen, helium, argon, dry oxygen, etc., all due to an enhanced oxide formation (Figs. 14 and 19) (17,24-27). However, in such gaseous environments, crack surface corrosion deposits are formed predominately by fretting oxidation mechanisms, which are mainly operative at lower load ratios where crack tip opening displacements are small. Thus, at high load ratios, there is little enhancement of the oxide debris on fatigue fracture surfaces and correspondingly behavior becomes similar in both wet and dry environments (17,25).
This simple rationalization for near-threshold fatigue behavior in gaseous atmospheres in terms of oxide-induced closure becomes more complex in alloy/environment systems where conventional corrosion fatigue mechanisms, such as hydrogen embrittlement, assume greater importance. For example, the extrinsic role of external gaseous hydrogen in simply providing a dry environment which minimizes the oxide-induced closure effect in low strength steels at near-threshold levels (25) must be contrasted with the intrinsic role of hydrogen charging where accelerated near-threshold growth rates due to actual hydrogen embrittlement have been observed (83). However, the latter effect can be further complicated by the additional roughness-induced closure arising from the resulting intergranular fracture (83,84). Moreover, in ultrahigh strength steels (UTS ≥ 1000 MPa) tested in moist air and dry hydrogen environments, the extent of corrosion debris and hence oxide-induced closure is extremely limited at all load ratios and accordingly behavior is dominated by hydrogen embrittlement processes even for external environments (27,32). Similarly for 7075 aluminum alloys tested in moist air environments (31,45,85), the enhanced oxide-induced crack closure, observed in overaged (T7) microstructures, does not necessarily lead to slower growth rates in this structure compared to under- or peak-aged conditions. This is because in the latter structures, resistance to hydrogen embrittlement/active path corrosion processes tends to be greater (31) and furthermore the planar slip characteristics of
underaged structures can enhance additional roughness-induced closure (45).

**Aqueous Environments.** Behavior in aqueous environments, such as distilled water or sodium chloride solutions, is largely similar to that described for moist gaseous environments above with the exception that such aqueous solutions often tend to form more copious corrosion deposits thereby increasing the effect of oxide-induced crack closure. For example, the crack surface formation of voluminous calcareous corrosion debris is well known for corrosion fatigue in sea water and other aqueous salts solutions, and recent work has shown that the resulting closure effects can indeed retard crack advance (e.g., 86). Studies in distilled water (18) revealed an interesting point in that, unlike moist gaseous atmospheres where enhanced oxide debris only form at low load ratios due to fretting mechanisms, in the more oxidizing aqueous environment, thick oxide deposits could form at all load ratios due to natural thermal processes. Correspondingly, the influence of load ratio on near-threshold behavior was far less pronounced in this environment since a prominent retarding effect of oxide-induced closure was present at high as well as low load ratios (Fig. 14). On the other hand, certain aqueous environments can "accelerate" crack growth by removing crack surface oxidation deposits, as has been shown by studies of near-threshold growth in HCl solutions (23).

**Viscous Environments.** The role of viscous environments in influencing fatigue crack propagation rates can also be attributed,
at least in part, to closure mechanisms. As discussed above in the section on viscous fluid-induced crack closure, in addition to the suppression of oxide-induced closure in viscous environments, such as dehumidified oils, the presence of a viscous fluid within the crack walls can induce crack closure effects from the hydrodynamic wedging effect of the oil (48-51). This results in an effect of viscosity on growth rates (Fig. 11), although as discussed above the precise trend in behavior with increasing viscosity depends upon a balance between the internal pressure generated by the oil wedge (which is enhanced with increasing n) and the kinetic ability of the oil to flow into the crack (which is decreased with increasing n). In general though behavior in inert viscous environments will resemble that in other inert atmospheres, such as dry gaseous helium, such that the suppression of oxide-induced closure at near-threshold levels will lead to faster growth rates compared to moist air, whereas the opposite effect will occur at higher growth rates due to the suppression of conventional corrosion fatigue mechanisms in the inert fluid (Fig. 11).

Temperature. The influence of temperature on specific mechanisms of closure has not been studied extensively to date, presumably due to the difficulty of isolating the individual effect on closure in experiments at different temperatures. Yuen and co-workers (29), however, have attributed the enhanced oxide formation, and hence oxide-induced closure, to the occurrence of premature thresholds during elevated temperature fatigue of nickel-base
superalloys. Corresponding studies on the role of closure at low temperatures are complicated by the intrinsic effects of low temperature deformation characteristics on crack extension mechanisms (83). However, it is clear that oxide-induced closure is likely to be inhibited at low temperatures due to lower oxidation rates whereas phase transformation-induced closure may well be enhanced due to the greater driving force for the deformation-induced reaction.

Applications

The various crack closure phenomena, discussed thus far, not only play an important role in influencing the constant amplitude fatigue crack growth rates, but also have a marked effect on several applications such as the short crack problem, variable amplitude crack advance comprising overloads and underloads, the existence of a fatigue threshold and its uniqueness for a given material and testing procedure. We now examine each in turn.

Short Crack Behavior

Cracks are defined as being short i) when their length is small compared to relevant microstructural dimensions (a continuum mechanics limitation), ii) when their length is small compared to the scale of local plasticity (a linear elastic fracture mechanics limitation), and iii) when they are merely physically small (i.e., crack length \(\leq 0.5-1.0\) mm) (87). Recent studies have shown that all
three types of short cracks grow substantially faster than (or at least as fast as) the equivalent long cracks subjected to the same nominal stress intensity factor range (e.g., refs. 88-94). In particular, the first two types of short flaws are known to propagate at high velocities even below the threshold for long cracks, $\Delta K_0$ (88-90). The initially high growth rates of such short cracks are known to decelerate progressively (or even arrest completely, in some cases, to form so-called non-propagating cracks) before merging with the long crack data (88-90,92). Recent studies have suggested that part of the reason for the apparently faster growth of short cracks, even below the long crack threshold, stems from their limited wake and the consequent absence of appreciable crack closure (11,87,89,93). There is growing experimental evidence indicating that the progressive deceleration and/or arrest of cracks, which are of a length comparable to the scale of local plasticity or microstructure, and the apparently faster growth of physically-short flaws can be accounted for, at least partially, through arguments based on the development of crack closure with increasing crack length (11,38,87,89,93). Morris et al. (89) first showed that for short flaws (a ~ 50 to 500 \( \mu \)m) in titanium alloys, crack closure (presumably that induced by roughness) decreased with decreasing crack length. A further influence of roughness-induced closure was evident in the work of McCarver and Ritchie (93) who found that the threshold $\Delta K_0$ values for short cracks (a ~ 20 to 200 \( \mu \)m) in René 95 nickel-base superalloys were about 60% smaller than for long cracks.
(a~25 mm) at low load ratios, even though no differences were apparent at high load ratios where closure effects were minimal. Tanaka and Nakai (94) measured crack closure loads for small cracks emanating from notches in a low carbon steel and showed that the anomalies between short and long cracks can be eliminated when results are plotted in terms of $\Delta K_{\text{eff}}$ by factoring out crack closure (Fig. 20). Thus, it may be argued that one of the reasons that short flaws show apparently larger growth rates than long cracks, subjected to the same nominal $\Delta K$ level, is because of their limited wake and less closure.

**Variable-Amplitude Loading**

The various crack closure phenomena described in this paper for constant-amplitude fatigue crack growth also have a strong influence on the transient fatigue crack advance following variable amplitude loading in the form of overloads or underloads. It was first suggested by Elber (6) that plasticity-induced closure is the dominant mechanism controlling crack growth retardation due to single or block overloads. Several subsequent studies have attempted to correlate the retarded delay distance with the size of the enlarged plastic zone generated by overloads (e.g., 95-97). Recent work by Suresh (98-100) has shown that in addition to plasticity-induced closure, other forms of closure such as those due to fracture surface oxides or micro-roughness also play an important role in controlling retardation because post-overload crack propagation is effectively
governed by the micromechanisms of near-threshold fatigue. Figure 21, for example, shows a highly serrated crystallographic crack profile following an 80% overload in 7075 aluminum alloy at a baseline $\Delta K = 7.7$ MPa\(\sqrt{m}\), where a planar fracture surface will be observed in the absence of an overload (100).

Oxide-induced crack closure has also been suggested as a possible explanation for the transient growth characteristics arising from fatigue underloads (101). Suresh and Ritchie (101) demonstrated that even though cyclic underloads with $\Delta K$ values smaller than the threshold $\Delta K_0$ do not contribute to crack advance or transient growth due to changes in crack tip plasticity, they can result in enhanced oxidation on fracture surfaces. Results in a 2 1/4Cr-1Mo steel indicated (101) that whenever the cyclic CTOD corresponding to the fatigue underload was greater than the maximum thickness of the excess oxide formed prior to the application of the underload, a further enhancement in oxide thickness occurs during underload cycling (such as the case when a fatigue test is "parked" overnight at $\Delta K$ values below the threshold $\Delta K_0$).

Crack Growth and Threshold Measurement

The dominant role of crack closure in influencing propagation rates also implies that the conditions in the wake of the crack and prior loading history can have a bearing on the current crack propagation rates. This poses a serious question on the very uniqueness of the crack growth rate at low $\Delta K$ levels (where closure effects are
important) and on the value of $\Delta K_0$ for a given material, environment and testing conditions. Clearly too rapid a rate of load shedding will result in premature arrest and an apparently higher threshold $\Delta K_0$ value (e.g., 102), presumably from enhanced plasticity-induced closure (akin to retardation following a high-low block loading sequence). Similarly if the rate of load shedding is too slow, it is conceivable that enlarged oxide deposits can form, again leading to an apparently higher $\Delta K_0$ from enhanced oxide-induced closure. Evidence for such effects have been reported both for aluminum alloys (102) and steels (103). However, the degree of non-uniqueness for most normal load shedding rates (i.e., < 10% reduction in $\Delta K$ for each load step over increments of 1 mm of crack advance) is not as large as might be expected. This follows because the closure mechanisms which dominate near-threshold crack growth behavior are developed in the very near tip region, a short distance behind the crack tip, and thus will only affect subsequent crack extension over a similar distance ahead of the tip. Estimates based on oxide debris measurements suggest that the crack surface contact due to oxide-induced closure to be within roughly 5 $\mu$m from the crack tip (18,26,33). More macroscopic experiments where material was removed behind the crack tip similarly indicated that closure within 1 mm of the tip is the most significant in influencing near-threshold behavior (104). However, it is clear that, despite the uncertainty in the location of crack surface contact, the development of a threshold for no fatigue crack growth is intimately linked to the extent of crack closure (25-
27), and since the magnitude of such closure will be dependent upon variables such as crack size, load history and so forth, some degree of variability in threshold measurements must be expected.

Concluding Remarks

In this review we have critically examined the many sources of fatigue crack closure and their influence on crack growth behavior, especially in the near-threshold region, under various mechanical, microstructural and environmental conditions. An understanding of the role of such closure processes is also found essential to such phenomena as the behavior of short cracks and the transient crack growth behavior during variable amplitude loading, both for single overloads and block loading, above and below the threshold. In all these instances, closure provides a mechanism whereby the local driving force for crack advance, i.e., the near tip $\Delta K_{\text{eff}}$ value, differs from the nominally applied values, i.e., $\Delta K$, computed globally in terms of geometry and applied loads. This latter concept clearly is of great importance to the fracture mechanics interpretation of fatigue data since it implies that one can no longer assume a unique dependence of the growth rate, for a given material/environment system, solely in terms of particular values of $\Delta K$ and $K_{\text{max}}$. Furthermore, since the closure mechanisms which primarily contribute to this discrepancy between local (near-tip) and global (far field) stress intensities all act in the wake of the crack tip, their effect
is critically sensitive to crack length, implying a crack length
dependence on the local driving force even when nominal ΔK levels are
identical. This constitutes a breakdown in the fracture mechanics
similitude concept (106). As outlined in the previous section, it is
felt that this latter fact is one of the major causes of the short
crack problem where short flaws are observed to propagate below ΔK₀
at rates often far in excess of long cracks at the same nominal ΔK
levels. However, such "anomalous" behavior of short flaws is only
one example, albeit a major one, of problems of uniqueness arising
from the current practice of characterizing crack extension in terms
of nominal driving forces, e.g., ΔK, ΔJ, etc., when closure
mechanisms are clearly present. For example, the varying dependence
of ΔK₀ values with microstructure and environment at high and low
load ratios, the lack of uniqueness in ΔK₀ measurements when rates of
load shedding are radically changed, the varying response of long
crack da/dN tests with classical stress-strain/life behavior, and so
forth all follow from the presence of crack closure mechanisms.

In view of this situation, one should conclude that the
characterization of fatigue crack propagation data in terms of a
nominal characterizing parameter, such as ΔK, should be discouraged
since in the presence of closure it does not represent the actual
crack driving force. The problem with this approach, however, is
what to substitute in its place, since no fundamental continuum
mechanics characterizations exist to date for fatigue cracks, which
include a consideration of cyclic plasticity, an analysis of the non-
stationary crack tip fields, and a modification of these fields due
to closure behavior in the wake. Until such analyses are available,
the use of $\Delta K_{\text{eff}}$, representing closure-adjusted $\Delta K$ values, provides
probably the most fundamental approach, at least for academic
assessment of fatigue behavior. However, in view of the many
difficulties in obtaining reproducible $K_{\text{c1}}$ data, the use of $\Delta K_{\text{eff}}$
will clearly present numerous practical difficulties for the
standardization of experimentally-determined fatigue crack
propagation data for engineering purposes.

\textbf{Acknowledgements}

This work was supported partly by the Director, Office of Energy
Research, Office of Basic Energy Sciences, Materials Science Division
of the U.S. Department of Energy, under Contract No. DE-AC03-
76SF00098 and partly by the Materials Research Laboratory under NSF
Grant No. DMR-8216726. Thanks are due to Dr. S. M. Wulf for his
continued support, both at the University of California in Berkeley
and formerly at M.I.T., to J. L. Tzou and V. B. Dutta for providing
us access to their as yet unpublished work, and to Madeleine M.
Penton for her help in preparing this manuscript.
References


85. J. Petit, this volume.


100. S. Suresh and A. K. Vasudevan, this volume.


Figure 1 - Schematic illustration of the various mechanisms of fatigue crack closure.
Figure 2 - Schematic representation of the model of Schmidt and Paris (9) for the dependence of fatigue thresholds on load ratio R, showing a) definition of the nominal and effective stress intensity ranges, $\Delta K_0$ and $\Delta K_{th}$, respectively, b) the predicted variation in $\Delta K_0$ with R and c) the predicted variation of the (nominal) maximum stress intensity at threshold $K_{0,\text{max}}$ with R. $K_{cl}$ and $K_c$ are the stress intensities at closure and at final failure, respectively, $R_{cr}$ is the critical load ratio above which closure effects are minimal.
Figure 3 - Experimentally-measured variation of alternating and maximum threshold stress intensities ($\Delta K_0$ and $K_{0,max}$, respectively) with load ratio ($R$) in two 2\%Cr-1Mo steels, SA387 and SA542-3 ($\sigma_y = 290$ and 500 MPa, respectively). Tests in room temperature environments of moist air and dehumidified hydrogen gas. Sample points shown for dry inert gas atmospheres of argon and helium. Data from refs. 17 and 25.
Figure 4 - Macrograph of a fatigue fracture surface in X-60 HSLA steel ($\sigma_y = 414$ MPa) showing accumulation of oxide debris near $\Delta K_0$ at the center of the test piece. Note how the crack, which is fully arrested at the center, is still trying to propagate at the surfaces where the accumulation of oxide is less. The crack is propagating from bottom to top. After ref. 18.
Figure 5 - Idealization of the role of crack face oxide debris in influencing fatigue crack advance by oxide-induced crack closure, in terms of a rigid wedge within a linear elastic crack. After ref. 33.
Figure 6 - An example of roughness-induced crack closure from the variable amplitude fatigue studies of Schulte et al. (105) on underaged X-7075 aluminum alloy tested in vacuum. Here the characteristic crystallographic mode of crack advance gives rise to markedly serrated crack paths, thereby promoting closure through fracture surface asperity contact.
Near-Threshold: \( r_y < d_g \)  
(Stage I, Modes II+I)

Higher Growth Rates: \( r_y > d_g \)  
(Stage II, Mode I)

Figure 7 - Crack opening profiles and resulting crack path morphologies corresponding to a), b), c), near-threshold (Stage I) and d), e), f), higher growth rate (Stage II) fatigue crack propagation. b) and e) show sections through fatigue cracks in 1018 steel (after Minakawa and McEvily (35)) whereas c) and f) show sections in 7075-T6 aluminum alloy (after Louwaard (36)). \( r_y \) is the maximum plastic zone size, compared with the mean grain diameter \( d_y \). After ref. 47.
Figure 8 - Schematic illustration of the development of roughness-induced crack closure during "Stage I type" crack growth due to fracture surface asperity contact from Mode II and Mode I crack tip displacements, compared to the absence of such closure at higher strength intensities during "Stage II type" crack growth involving predominately Mode I displacements. After ref. 35.
Figure 9 - Predicted variation from Eq. (12) of the extent of roughness-induced closure, in terms of $K_{cl}/K_{max}$, with fracture surface roughness factor $\gamma$ as a function of $x$, the ratio of unloading Mode II and Mode I displacements ($u_{II}$ and $u_I$, respectively). Experimental data points derived from results on 1018 steel after Minakawa and McEvily (35) and on fully pearlitic rail steel after Gray et al. (43). After ref. 47.
Figure 10 - Idealization of the distribution of pressure $p(x)$ and resultant force ($F_p$) due to the presence of a viscous fluid inside a fatigue crack, of length ($a$), where the fluid has only partially penetrated the crack along a distance $d$. The resultant stress intensity arising from such internal oil pressure is a function of $a$, $d$, the average crack opening ($h$), the closing velocity of the crack walls ($v$) and the absolute viscosity ($\eta_p$). After ref. 51.
Figure 11 - Comparison of fatigue crack propagation behavior at $R = 0.05$ in a 2 1/4Cr-1Mo steel ($\sigma_y = 500$ MPa) tested at 50 Hz in high and low viscosity silicone and paraffin oils with behavior in gaseous environments of moist air, dry hydrogen and dry helium gases. After ref. 50.
Figure 12 - Influence of load ratio R on near-threshold and higher growth rates in 2 1/4Cr-1Mo steel tested at 50 Hz in moist air. Whereas above $10^{-5}$ mm/cycle, behavior is essentially insensitive to R, below $10^{-6}$ mm/cycle, growth rates are increased and $\Delta K_0$ values decreased with increasing positive load ratio from R = 0.05 to 0.75. After ref. 18.
Figure 13 - Experimental measurements of crack closure using an ultrasonics technique indicating the variation in ultrasonic signal transmitted across a near-threshold crack during a fatigue cycle at \( \Delta K_0 \). Data, after ref. 26, for a 2\%Cr-1Mo steel tested in moist air and dry hydrogen gas at \( R = 0.05 \) and 0.50.
Figure 14 - Variation of fatigue thresholds $\Delta K_0$ and $K_{0,\text{max}}$ with the maximum excess oxide thickness $d_0$, as a function of load ratio $R$, for a bainitic $2\frac{1}{2}$Cr-1Mo steel ($\sigma_Y = 500$ MPa) tested at 50 Hz in room temperature environments of moist air, dry gaseous hydrogen and distilled water. After ref. 18.
Figure 15 - Compilation from ref. 10 of fatigue threshold $\Delta K_0$ data as a function of yield strength $\sigma_y$ for steels, tested at low load ratios in room air.
Figure 16 - Variation of maximum excess oxide thickness, measured on the fracture surface at the threshold $\Delta K_0$ at $R = 0.05$, with yield strength $\sigma_y$ for a range of steels tested at 50 Hz in moist air. After ref. 32.
Figure 17 - Variation in experimentally measured crack closure, with $\Delta K$ for a 7475 aluminum alloy ($\sigma_y = 445-505$ MPa) tested in a) room air and b) vacuum. Closure data are compared between fine and coarser grained microstructures in the underaged and overaged conditions. After ref. 76.
Figure 18 - Experimental closure data as a function of $\Delta K$ for dual phase Fe/2Si/0.1C steel ($\sigma_Y \approx 600$ MPa) tested at $R = 0.05$ and 0.75. Certain morphologies of the duplex ferritic/martensitic microstructures promote crack deflection causing significant (roughness-induced) closure, e.g., for the IA and SQ conditions, compared to more linear crack paths in the IQ condition. After ref. 46.
Figure 19 - Fatigue crack propagation behavior in a martensitic 2⅓Cr-1Mo steel (σ_y = 769 MPa) tested at R = 0.05 and 0.75 in moist air and dehumidified gaseous helium, showing faster near-threshold crack growth rates in the inert gas at low load ratios only. After ref. 25.
Figure 20 - Variation in fatigue crack propagation rate with a) nominal and b) effective stress intensity range values, \( \Delta K \) and \( \Delta K_{\text{eff}} \), respectively, in a 0.17% C steel (\( \sigma_y = 194 \) MPa). Data points represent behavior of short cracks emanating from notches compared with solid line representing conventional long crack \( da/dN \) data. After ref. 94.
Figure 21 - Crystallographic crack advance following an 80% overload at a baseline $\Delta K = 5$ MPa$\sqrt{m}$ in an underaged 7075 aluminum alloy leading to a highly serrated crack profile and roughness-induced closure; this implies that near-threshold mechanisms can play a significant role in influencing post-overload growth. After ref. 100.
This report was done with support from the Department of Energy. Any conclusions or opinions expressed in this report represent solely those of the author(s) and not necessarily those of The Regents of the University of California, the Lawrence Berkeley Laboratory or the Department of Energy.

Reference to a company or product name does not imply approval or recommendation of the product by the University of California or the U.S. Department of Energy to the exclusion of others that may be suitable.