ABSTRACT

The present article reviews plasticity, roughness, and oxide-induced crack closure and its direct effect on fatigue crack propagation (FCP) problems. In such cases, plasticity-influenced crack closure significantly influences number of cycles to failure. Thus effective value of crack driving force for fatigue crack propagation problems is dependent upon assessment of plasticity effect both in front of crack tip and in the wake of the advancing crack in the form of appendages of plasticity affected material. Crack closure (through plasticity) is affected by plane stress and plane strain transition, and effect of both types of conditions can be implemented to predict FCP, by taking into consideration appropriate value of plasticity constraint factor. Crack growth under fatigue loading conditions subjected to constant amplitude loading, single overloads, block overloads, variable amplitude loading and random loading is influenced by crack closure effect strongly during transition from plane stress (PSS) to plane strain (PSN) conditions.

Cite this article as: Kumar A, Harmain GA. Revisiting the crack closure and its effect on fatigue life of components. Sigma J Eng Nat Sci 2024;42(2):590–599.
enhance deformation. Extrinsic effects were characterized by [1] as geometric effects, and contact shielding process.

The crack shape also induces stress and strain fields with local variation during propagation of the crack because of combined effect of several modes of loading. Local variation in stress and strain fields near crack tip may result in variation in crack driving force and in case of crack deflection or branching, can result in combined effects of modes of loading [5–9]. Crack driving force is also affected by oxide-influenced crack closure, roughness influenced crack closure, crack splitting by ligaments and plastic transformation influenced crack closure. Plasticity-induced crack closure, may causes transfer of load during certain part of cycles, and consistently effects the deformation field at crack-tip. However measurability and predictability of crack closure still presents a challenge due to host of variables including loading history, microstructure, crack length etc. the modes of loading i.e. Mode I, Modes II and III strongly influence the effective value of crack driving force.

Crack closure has been popular as the primary mechanism of near-threshold fatigue, as such, an attempt has been made in this study to review the genesis and prognose plasticity, roughness, and oxide-influenced crack closure and its influence on fatigue life of components subjected to a particular loading condition. A brief background of this research area followed by discussion on the measurability and predictability of crack closure has also been undertaken.

Background

Ewing and Humphrey provided the first findings on early fracture initiation and subsequent FCP during cyclic loading [10]. The evolution of fatigue cracks for a long period was obtained using optical microscope. The first use of fracture mechanics was introduced by Paris in the early 1960s, the SIF range to characterise fatigue crack propagation [11,12] proposed a simple transformation scheme, 

\[ \Delta K_{eff} = K_{max} - K_{cl} \]  

where \( \Delta K_{eff} \) is the SIF range, \( K_{max} \) is the maximum value of SIF and \( K_{cl} \) is the SIF at which the first fracture surface contact occurs. Equation 1 would not be able to compute \( \Delta K_{eff} \) entirely in terms of \( \Delta K_{eff} \) or another effective driving force and

\[ \Delta K_{eff} = \Delta K_{max} - K_{cl} \]  

In 1970s, as such short crack effect, which states that short cracks grow quicker than long cracks with same driving force, \( \Delta K \), and chemically short cracks may advance well below the threshold value \( K_{th} \) of a long crack [29,30]. The short crack phenomenon has been categorized into several types which includes (i) micro-structurally short crack (on the basis of configuration of flaw size with respect to grain size), (ii) mechanically short crack, (iii) chemically short crack, and (iv) physical short crack [31]. In several papers, reported in last four decades on the phenomenon of CC, the effect on FCP with respect to effective crack driving force (\( \Delta K_{eff} \)) has been extended to physically short crack problems. Despite the large number of research that indicates its relevance to CC, research reports have raised questions about the notion [32, 33–35]. Numerous studies have been conducted to model variety of crack problems involving crack propagation [36–38], contact stresses and fatigue cracks using various enriched techniques [39–42].

Plasticity-influenced CC under plane strain loading circumstances, as well as roughness-influenced and debris-influenced CC, have all been questioned. The role of flaws has been studied with respect to CC work of Louat [33]. They suggest that plasticity effect of crack tip does not produce CC unless fracture is entirely packed with asperities. CC caused by oxides, corrosion, and roughness were found to be significant in causing effective FCP. Although extensive studies have been conducted for clarification debates and attempts at clarification, several disagreements about the necessity of CC appear to remain unresolved. The following are the basic investigations:

- Is it possible to express \( da/dN \) entirely in terms of \( \Delta K_{eff} \) (or another effective driving force) and
- Is \( \Delta K_{eff} \) effectively calculated by

\[ \Delta K_{eff} = \frac{\Delta K_{max} - K_{cl}}{K_{cl} - K_{th}} \]

where \( \Delta K_{eff} \) is the SIF range, \( K_{max} \) is the maximum value of SIF and \( K_{cl} \) is the SIF at which the first fracture surface contact occurs. Equation 1 would not be able to compute \( \Delta K_{eff} \) in the latter circumstance because there is no known \( K_{cl} \). To determine effective \( \Delta K \) it is imperative to obtain the value of opening SIF (\( K_{op} \)) for a loading block with \( K_{max} \) and \( K_{cl} \).

For evaluation of CTOD at crack tip, PSS/PSN effect also has to be considered for estimation of effective SIF range in a 3D geometry [43]. The complexity of the problem will increase manifold if plasticity effects are considered owing to 3D geometry of fracture. Further limited investigation on plasticity-induced CC has been reported on high entropy alloys, nickel-based superalloys and some specials alloys of steel including 316LN steel, 304LN steels etc. [44,45].

Conventional and Fracture Mechanics Based Fatigue Life Assessment

The advent of new manufacturing processes and introduction of novel materials has resulted in need for determination of their properties such as fracture toughness,
endurance limit and crack growth rates. The structural integrity requirements as regards fatigue, creep and creep-fatigue interaction for new materials e.g. aluminium syntactic foams, composite metal foam, metal matrix syntactic foams and cellular materials needs to be investigated [46–51].

Conventionally fatigue life has been obtained using S-N (stress –Life) diagrams. In such investigations, a considerable portion of life of various engineering components has been reported in the initiation period. Subsequently crack length ‘a’ versus life ‘N’ has been used for obtaining life of various specimens [52] presented assessment of fatigue life w.r.t. various loading conditions. The fatigue life is also influenced by mean stress which may be obtained from Haigh diagram. For several ductile materials other than steel it is essential that a new set of threshold conditions may be defined on the basis of number of cycles to failure ($10^6$ cycles). Figure 1 is often referred to as a fatigue design diagram, and zero mean stress is used as a design value for fatigue limit.

Goodman’s design line and Gerber’s parabola also provide designer first-hand information on assessment of life. These are based on the idea that the constraints for alternating stress and static load should be linked by a smooth curve. The start of fatigue cracks is not fully known initially. Initiation of fatigue crack is obtained from post-mortem studies in several ways e.g. using persistent slip bands [53–56]. Design value of fatigue strength in related to material’s ability to withstand local stress state and, eventually, its plasticity resistance.

Crack Closure Measurements

Elber [57] used a crack opening extensometer to make the first estimate of closure by employing a near-tip strain gauge. However, it is challenging to measure closure values precisely. Figure 2 presents results based on data from investigation of Schmidt and Paris [58] as an example. This plot succinctly gives sharp change in compliance values: the high value corresponds to an open crack while the low value is corresponding to a closed crack. It had been left unlocked. Because of the hazy slope transition, until zero load or less, the crack does not entirely shut.

Opening load corresponds to the load acting on the crack at which crack is closed, while closure load is defined as the load corresponding to which, change in compliance occurs as a result of contact. An unzipping tendency attributed to plastic closure was observed by Schmidt and Paris [58] by analyzing load displacement curves recorded at various points along the fracture.

Using stereo-imaging in the SEM, Lankford and Davidson [59] mapped fracture opening displacements. Such measures can indicate the presence of closure, but touch cannot be immediately detected. However, using fractography closure loads can be estimated indirectly [60–62].

Figure 1. Schematic representation of Haigh diagram.

Figure 2. Estimating the closure load using Schmidt and Paris data.
though the external section of crack tip may close but the inside section of the crack tip usually remains open. When striations are prominent on fracture surface, Sunder and Dash [63] proposed experimental technique for measurement of opening and closure loads. In this technique $K_{	ext{min}}$ is decreased while $K_{\text{max}}$ is held constant, till striations become prominent on fracture surface, followed by gradual reduction of $K_{\text{max}}$ with increase in value of $K_{\text{min}}$ till striations completely disappear. By measuring the interspacing between the striations, opening and closing stress intensities can be measured under both load increasing as well as load reducing conditions. The intensity of closure stress does not have to be consistent. The majority of closure observations have been reported invoking the differences in compliance values [64]. The existence of closure can be estimated using compliance measurements over entire specimen, but usually the variation in compliance is negligible. As such, these measurements must be mapped very close to crack to record significant difference. Though the use of crack mouth opening gauges to estimate closing and opening stress intensities has been criticised owing to appreciable difference in the said values from those seen at the crack tip, the crack mouth opening gauge is the basic tool and is still in use [65]. Unzipping of the crack does not always occur in the same direction. As an example, in the case of 2024-T3 aluminium, by employing large number of strain gauges, it has been reported that the crack appeared to open initially at the tip and then gradually toward the mouth [58], while unzipping in reverse direction has been reported by Davidson [66] in 7091 aluminium alloy. However, usage of one strain gauge, instead of several gauges, particularly at the crack-tip region has been preferred more [57-58].

Slope discontinuities associated with load-displacement curves obtained by positioning single gauge in path of the crack, is commonly employed to estimate closure. A step forward from gauges is to obtain crack position with respect to the surfaces. Laser interferometry is another viable option for estimation of closure. This technique uses a scattering location; on either end of crack, for laser light, which produces two-point diffraction patterns. This technique is highly sensitive provided that scattering sites are not highly distorted by plasticity [62]. Irrespective of the approach that is used, the position of the discontinuity in slope, is to be determined. Alternatively an offset method as proposed by [53] can be utilized to determine the level of closure stress. The plot of load versus displacement, shown on the left of Figure 3, is a hysteresis loop. The plot on the right hand side is obtained by removing the mean value pertaining to slope of the linear part of the curve immediately over the closure values. The value of stress intensity factor at crack-tip is found at the point where non-linear zone meets the linear region as we ascend.

An alternate method was proposed by Donald [67] for determination of $K_{\text{max}}$, $K_{op}$. In absence of closure, the compliance ratio approach modifies the stress intensity values by dividing the measured average compliance by the compliance that would have existed in the absence of closure. Measurements of displacements at the crack are used to obtain these compliances. When fatigue damage occurs well below the closing force, there does not appear to be a methodical analysis of how a partial closure might affect displacement measurements in crack tip area, particularly in the elastic–plastic situation. Such a technique will necessitate solution for mode I elastic–plastic crack displacement field [68].

Originally, the fracture closure measures were meant to obtain the effective stress intensity range in a simple form such as:

$$\Delta K_{\text{eff}} = \Delta K_{\text{applied}} - (K_{op} - K_{min}) = K_{\text{max}} - K_{op}$$

Provided $K_{min} < K_{op}$, where $K_{op}$ represents initial stress intensity while $K_{min}$ represents the applied minimum stress intensity. This method has been reported to successfully shrink data obtained at several $R$-values into a single line in crack growth rate plot [69]. However, the effectiveness of this method can be to a great extent attributed to the expertise with which $K_{op}$ is obtained. Artificially induced closure experiments provided some answers to questions about this subjectivity.

For the cases where in closure was induced by insertion of shims into crack mouth, compliance change has been reported, even though the effect on threshold and crack propagation has been reported to be smaller than what would be predicted by employing stress intensity range as given by Equation 2. Numerous studies have been reported which have proposed modifications of Equation 2...
to provide more close and realistic approximation of stress intensity in light of crack closure [43].

**Plasticity Induced Crack Closure**

Vasudevan *et al.* [34], proposed the CC effect, of which the closure displacement of the crack tip plasticity is the most important factor. Depending on continuum plasticity, a theoretical model [70] was proposed to demonstrate that fracture closure can be important in PSS conditions. Since 1975, the relative importance of plasticity-induced CC has been greatly reduced [34], as contributions from oxidation closure, surface roughness and corrosion products have become prevalent. Study of crack propagation in various types of short cracks and fatigue retardation phenomenon associated with single and multiple overloads has revealed that contribution of plasticity-influenced CC is negligible, depending on both theoretical and experimental evidence [71]. As can be seen in Figure 4, there are two types of displacement in the plasticity affected area in the vicinity of crack tip: one ahead of advancing crack tip and other from the plasticity affected appendages in the wake of advancing crack. When displacement occurs at the crack-tip, the 'negative' half of the displacement results in the bulge of crack-tip, allowing the positive part to penetrate material.

It is commonly considered that the plasticity contribution to CC is mostly due to the cyclic plasticity affected volume of an advancing crack. The displacement on the wake should only include one half of the loop, while other half remains close to surface of the crack and forms an edge (Figure 4 (d)). Therefore, the $y$-displacement caused by crack closure is insignificant compared to $y$-displacement when crack remains open. It can be concluded that the propagation of the crack may be retarded through flow of material (out of plane), which can cause reduction in plasticity affected zone. When critical value of SIF is reached, crack propagation occurs with high intensity and crack closure does not affect the overall process of crack growth. The plasticity-influenced CC under planar stress is influenced by the wake of advancing crack. This wedge results in reduction of deformation (cyclic plastic) at the crack tip resulting in reduction of rate of fatigue crack growth (FCG) [70, 72].

In case of components subjected to PSN, there is limited plastic wedge behind the crack-tip. The plasticity effected material in the wake shears volume elements to rotate, resulting in the formation of a wedge around the crack tip [73,74]. There are several methods to represent the transfer of material to the crack-tip due to combined effect of plastic deformation. The simplest explanation is one based on existence of dislocations (Figure 4 (a)). The rotation of volume elements in this case is caused by necessary geometrical arrangement of the displacements on plasticity effected appendages in the wake (Figure 4 (a) and (b)). From a continuum mechanics point of view (Figure 4 (c) - (d)), the transfer of material to the crack-tip is explained by strain-core deformation [75].

Overloading results in extra plastic flow in immediate vicinity of crack-tip, resulting in an equivalent crack opening and internal stresses that are protected from the stress field. Crack growth is slowed or stopped as a result of overload. Overload causes cracking, although this is not the same as CC. The displacement caused by the displacement of the plastic zone must not exceed twice the displacement field that caused the crack. This crack tip shielding procedure should not be confused with CC. However, plastic flow through displacement protection processes (similar to HRR fields) helps to modify the stress field in the immediate vicinity of crack-tip and alters the driving force at the crack-tip. As a result, a clear difference should be made between CC in the wake of advancing crack and crack-tip stress adjustment front of a crack-tip. Similarly, various other residual stresses, such as cold deformation, shot peening, and machining-induced stresses, can also alter the stress field. In such internal stress zones, a migrating

![Figure 4](image-url)

*Figure 4.* Schematic representation of states of dislocation loop formation (a) elastic crack, (b) - (c) a head of crack tip (d) in the wake of crack in comparison to elastic crack.
crack may result in fatigue crack retardation. The advancing crack is influenced by both plasticity ahead of crack tip and the wake.

**Roughness- Influenced Crack Closure**

Asymmetry in crack-tip deformation generates non-conformity roughness of resulting fracture surfaces which may be far away from actual tip of a crack [76]. Argument that was previously utilized to explain the plasticity- influenced crack closure with PSN conditions has been extended for explanation of roughness-induced CC under same conditions. The explanation for plasticity-influenced CC may be given on the basis of slip process, which result in moment of material dislocations during crack propagation; this may be attributed to moment of dislocations during crack propagation of the crack flanks in the direction of crack propagation $\Delta u_1$ (Figure 5). Because of the considerations stated previously, the crystal orientation in the grains in the immediate vicinity of crack tip may vary on both sides of the flank. The degree of orientation is typically different on both sides and varies from grain to grain. As a result, there is still a mismatch between the two crack flanks. The precise mathematical explanation of retard displacement fields gives credibility to concepts provided for plasticity- influenced CC under PSN conditions [73]. The mismatch of fracture surfaces attributed to asymmetric dislocation arrangement (geometrically necessary dislocations) in the crack tip wake explains the occurrence of the roughness influenced CC far behind the fatigue crack tip [76].

In light of these findings, the asymmetry of the plasticity affects appendage in the advancing crack is a critical factor in controlling roughness-influenced CC. Hence roughness-influenced closure of crack is apparent in coarse-grained alloys approaching the threshold, under the condition that the monotonic plasticity size is similar or less than that of grains. By having a bigger monotonic plastic zone, which includes more grains than a smaller one, asymmetric deformation due to crystal anisotropy is minimized.

Fracture surfaces resulting from fatigue loading show variable roughness, the largest variation can be of the scale of grain size and smallest variation can be of the magnitude of lattice spacing which can be attributed to ledges associated with dislocations generated on fracture surface or at crack tip. However much smaller variations in near threshold region can be attributed to microstructural features of scale less than grain size and dislocation-dislocation interaction [18,77]. The magnitude of roughness in crack front direction and direction of propagation can be different owing to the fact that displacement of the crack surfaces varies locally. However, roughness based closure does not depend only on displacement of geometrically similar fracture surfaces [78–80]. Non identical plastically affected appendages of fractured surfaces leads to local contact at peaks of asperities. Schematic representation of several kinds of roughness based closure are depicted in Figure 6. Crack flank is bent locally in the vicinity of the kink by dislocations, forming a hump and the crack flanks come in contact at the top of the irregularity.

![Figure 5. Schematic illustration of roughness- influenced crack closure, ($\Delta u_1$ = roughness parameter between crack-tip and contact).](image)

![Figure 6. Schematic of various types of roughness based closure (a) Plasticity stretch effect dominated closure (b) Lateral displacement of crack surfaces dominated roughness closure.](image)
The mechanism of crack closure associated with roughness is a strong function of asymmetry associated with plastic deformation and size of plastic zone; in addition to ratio of deflection length to plastic zone size. Plasticity stretch effect (hump) effects the closure load in situations involving small plastic zone size with respect to deflection length. However, in cases where in plastic zone size is large with respect to deflection length with asymmetry in plastic deformation, a shift of crack flank dominated by roughness induced crack closure will occur [2].

**Oxide-Influenced Crack Closure**

The deposition of oxide layer on the fractured surface characterized as a wedge infiltrating crack causes the range of effective stress intensity to reduce. Crack closure which is a result of oxide-influenced closure occurring due corrosion of crack surface, is promoted under low load ratios in oxidizing environments. A lot of emphasis has been laid by many researchers [21,34] on the significance of crack closure as result of oxidation of fracture surface at room temperature and at higher temperatures. It is common for newly formed fracture surfaces to oxidize in the presence of an atmosphere and this effect is frequently more noticeable when the environment is moist. Prominent examples are the calcareous deposits and crack surface oxides formed as a result of corrosion under fatigue loading conditions in structured steels, tested undersea and distilled water conditions. Simple numerical modeling, based on the idea of rigid wedge in a linear crack, suggests that oxide-induced crack closure is a very strong function of oxide film thickness ‘d’ and the position of its maximum thickness from the crack tip ‘2z’ [21].

\[
K_{\text{ct}} \approx \frac{E' d}{4\sqrt{\pi z}}
\]  

(3)

where, \(E'\) = Young’s modulus of material.

Thin oxide layer formation on fresh surface of metals and alloys is of a few nanometers in thickness at ambient room temperature. Interaction between the resulting fracture surface and the roughness-influenced crack closure is effected by formation of oxide layer. Formation, tearing and reformation of oxide layer occur continuously in fretting process, resulting in thick oxide layer formation. Accumulation of oxide layer on virgin fracture surfaces is particularly notable at low \(R\)-ratios i.e., at or near the fracture surface threshold. As a result of their dark color, these thick oxide coatings produced on the fracture surfaces of steels may readily be visible to human eye. It is important to note that the oxide-influenced crack closure helps to explain a variety of phenomenon occurring near the threshold:

- The environmental effect; as in moist air, large \(\Delta K_{\text{th}}\) has been reported at low \(R\) ratios.
- Impact of loading procedure followed for pre-cracking, on the measured value of \(\Delta K_{\text{th}}\) [81].
- Unanticipated loading history effects, as an example, amplitudes below the threshold, may retard crack growth rate.

**CONCLUSION**

The influence of \(R\) ratio, and load history on FCP crack closure is required for ductile materials such as nickel based super alloys and 316LN steels etc. The genesis and prediction of plasticity-, roughness, and oxide-influenced CC have been given in detail. Since around 1980, CC has been popular as the primary mechanism of near-threshold fatigue. Study of several recently published papers pertaining to influence of CC on propagating crack reveals that Plasticity influenced closure is seen to be less important, in near-threshold fatigue. Plasticity based closure under PSS and PSN conditions directly influences fatigue life. Closure effects based on plane stress conditions can be viewed as a rigid insert between the crack surfaces. Under PSN conditions, the extent of plastic deformation is small in comparison with PSS resulting in reduction of life under PSN.

Roughness-based closure is influenced by direct asperity contact in the wake of an advancing crack. Due to plastic deformation CC is, in theory, foreseeable; nonetheless, there are numerous unanswered concerns for its exact determination, particularly the 3D effects. Further, the transition from the PSN to the PSS case, and the influence of cyclic and monotonic softening or hardening also requires further detailed study. A different approach for evaluating FCG data is needed to explain the \(\Delta K_{\text{th}}\) variation with respect to \(R\) and other phenomena related to the closure.

**AUTHORSHIP CONTRIBUTIONS**

Authors equally contributed to this work.

**DATA AVAILABILITY STATEMENT**

The authors confirm that the data that supports the findings of this study are available within the article. Raw data that support the finding of this study are available from the corresponding author, upon reasonable request.

**CONFLICT OF INTEREST**

The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

**ETHICS**

There are no ethical issues with the publication of this manuscript.

**REFERENCES**


[68] Weertman J. Why a complete solution has not been found of a mode i crack in an elastic plastic solid. Phys Status Solidi 1992;172:27–40. [CrossRef]


[73] Reinhard P, Riemelmoser FO. Visualization of the plasticity-induced crack closure under plane strain conditions. Eng Fract Mech 1998;60:315–322. [CrossRef]


