Chapter 3
Fatigue Fracture

From the practical point of view, fatigue fracture is the most important damage process. Available statistics show that, including corrosive assistance, fatigue is a leading cause of material failures registered during a long-term performance of engineering components and structures [246]. From the micromechanical point of view, the fatigue process can be understood as a sequence of the following stages: nucleation of cracks, stable propagation of short (small) cracks, stable propagation of long cracks and unstable fracture [247].

This chapter is divided into four sections. In Section 3.1, morphological patterns reflecting all crack growth stages on the fracture surface are briefly described. Moreover, the topological methods widely utilized in fatigue research and the quantitative fractography are outlined. These topics are important for all subsequent sections of the chapter.

The second section is devoted to the propagation of fatigue cracks under the remote opening mode (mode I). Mechanisms of nucleation and growth of short cracks are briefly reported, although these initial stages of the fatigue process were not a special subject of our research. Inclusion of these topics was, however, inevitable in order to provide a self-contained description of fatigue micromechanisms. The growth of short cracks is governed by shear stress components in favourably oriented crystallographic slip systems that are inclined at about 45° with respect to the maximal principal stress. This means that the short cracks grow in a local mixed-mode I+II and the picture of dislocation emission from the crack tip as well as the related growth micromechanisms are completely different from those related to long cracks.

After a certain incipient period that incorporates crystallographic and transient growths of short cracks, the fatigue cracks incline towards a direction perpendicular to the maximal principal stress, i.e., nearly towards mode I loading of the crack tip. This means that the long cracks keep propagating so that the crack tip plasticity is produced in the opening loading mode. Thus, the second section provides the reader with a micromechanical interpretation of all important phenomena accompanying this, most frequent, type of fatigue crack growth. Knowledge of these micromechanisms is essential not only for materials scientists and mechanical engineers, but also for technologists attempting to design structural materials exhibiting better resistance to fatigue crack propagation.

When the shear components of the applied stress are dominant, both short and long cracks can grow macroscopically under shear loading modes II or III. As usual, however, pure shear-mode crack propagation persists only for a limited number of loading cycles and the cracks incline or branch to get loaded in mode I. This leads to a local mixed-mode I+II, I+III or I+II+III crack propagation. The factory roof formation under torsion loading is an instructive example of such behaviour. Therefore, Section 3.3 refers to both shear-mode and mixed-mode crack growth. The first subsections introduce theoretical models and experimental results concerning crack growth under pure-shear and torsional loading. Since a combined cyclic bending-torsion is applied to many structural components, a rather extended part of the third section reports on the results of fracture tests performed under this kind of loading.

The final Section, 3.4, is devoted to the application of quantitative fractography to failure analysis. The fracture morphology can purvey a direct link between damage micromechanisms and both initiation and propagation of cracks. This section directly outlines how the knowledge of fracture micromechanisms can help to identify the reasons for failures of structural components in service. Therefore, its content can be useful for machine designers, especially for those working in the transport industry.

In general, Chapter 3 attempts to convince the reader of how useful the unified nano- micro- meso- macroscopic approach can be when trying to describe and interpret the behaviour of fatigue cracks.

3.1 Quantitative Fractography

Micromechanisms of fatigue crack propagation can be advantageously studied by means of fractographic tools. Indeed, the fracture surface morphology reflects many important stages of both stable and unstable fatigue crack propagation. Moreover, quantitative fractography is a powerful tool for failure analysis. Consequently, all sections of this fatigue chapter more or less refer to morphological patterns and fractographical results. Therefore, the quantitative fractography section was placed at the very beginning of this chapter.

The term "fractography" was first used by Carl A. Zapffe for the procedure of descriptive analysis of fracture surfaces in 1945 [248]. The output of this analysis is a set of numerical characteristics (number, shape, size, orientation, distribution) related to morphological patterns or parameters (roughness, fractality, texture) of the global surface topography. In both cases, the accuracy of these data is determined by knowledge of space coordinates of points on the fracture surface investigated. A very intensive development of quantitative fractography is obviously directly associated with increasing accuracy of measuring methods as well as with a rapidly growing capacity and computing rate of computers. During the second half of the last century, two-dimensional fractography in the scanning electron microscope was widely developed. In the last 20 years, however, an extended utilization of computer-aided topography techniques has enabled enormous progress in three-dimensional methods. Similarly, a number of descriptive concepts, distinguished by both the extent and the quality of utilized parameters, have been developed in the area of quantitative fractography.

Nevertheless, two various approaches can be distinguished here. The first utilizes various parameters of roughness, fractality or texture of fracture surfaces in order to find relationships between the fracture topology on one side, and the loading mode or the crack growth rate on the other side. In the first subsection, therefore, definitions of basic topological parameters are outlined. A rather different problem, usually demanding extensive research experience, constitutes the correct identification of morphological patterns as fracture facets, ridges, beach marks, tire tracks or fatigue striations. Thus, the second approach deals with the quantification of morphological patterns and their relationships to the loading parameters.

3.1.1 Topological Analysis

A general description of fracture surfaces by means of topographical parameters should be able to provide topological characteristics as well as morphological patterns as special cases. This might, in principle, be achieved by an analysis of a global set and relevant subsets of topological data. However, the basic problem here is represented by a high variability of fracture surface topology measured at different resolutions as well as an extreme complexity of the related microreliefs. In spite of a rather long history, no general definition of the surface topology and, consequently, no universal methodology of its quantification has been commonly accepted up to now [249–254].

In order to acquire a sufficiently wide and relevant set of topological parameters, advanced three-dimensional topological methods are to be employed. A great majority of results presented in this book was obtained by application of two methods that are based on different physical principles. Stereophotogrammetry is a method that makes use of the stereoscopic principles in order to obtain topological data of the fracture surface under investigation. Inputs to the method are two images of the analyzed region taken from different angles of view (so-called stereoimages or the stereopair) and some additional parameters that characterize a projection used during their acquisition. Usually, a scanning electron microscope (SEM) equipped with a eucentric holder is employed and the stereopair is obtained by tilting the specimen in the

SEM chamber by an angle that depends on the local roughness of the surface. The stereopair is processed via a matching algorithm in order to find corresponding points on both images (homologous points) and the relative z-coordinates of these points are calculated. The 3D model of the depicted surface area usually consists of ten to twenty thousand non-equidistant points and so the Delaunay triangulation must be performed [255].

Optical chromatography represents another method useful for a 3D reconstitution of the fracture surface micromorphology. The profilometer MicroProf FRT, Fries Research & Technology GmbH, makes use of the chromatic aberration of the optical lens. Different light monochromatic components are focused at different heights from a reference plane at the output of the optical fibre. The light intensity exhibits a maximum at the wavelength exactly focused on the surface and the height of the surface irregularities is deduced by using a calibration table. This optical method was usually employed only for verification of selected results obtained by stereophotogrammetry.

According to their mathematical basis, recently used topological parameters can be divided into five main categories. This classification is based on published works [250–253] and, in particular, on the work of Petropoulos *et al.* [256]. First two categories represent vertical (altitudinal) and length roughness parameters which characterize vertical and horizontal distributions of surface points, respectively. The third group involves hybrid parameters simultaneously describing more than one of the above-mentioned aspects. The fourth and fifth groups respectively consist of spectral and fractal characteristics of the fracture surface.

3.1.1.1 Roughness Parameters

In this brief overview only vertical, length, hybrid, spectral and fractal parameters are mentioned in more detail. The description of other parameters can be found elsewhere [214, 249–251]. For the sake of simplicity, the assumption of non-overlapping surface elements is accepted hereafter. This means that each pair of coordinates $(x_i,\, y_i)$ that determine the location of a point on the reference plane perpendicular to the macroscopic fracture surface is uniquely related to one altitudinal coordinate z_i. Because of a discrete data set of points utilized here to quantify surface topography, only a discrete form of definition of individual parameters is presented. Integral definition of many topological characteristics can be found, e.g., in [251, 252].

Vertical Parameters

Most important altitudinal parameters are characteristics associated with the probability of realization of the value z_i in terms of central moments, defined as

$$\mu_k = \frac{1}{n} \sum_{i=0}^{n-1} (z_i - \langle z \rangle)^k,$$

where k is the order of the central moment, n is the range of the analyzed set of data and $\langle z \rangle$ is the arithmetic average height:

$$\langle z \rangle = \frac{1}{n} \sum_{i=0}^{n-1} z_i.$$

With respect to properties of $\langle z \rangle$, the first central moment is zero. The second moment (variance) yields information on the width of the distribution. This parameter is connected with the standard deviation R_q by the relation

$$R_q = \sqrt{\mu_2}.$$

The characteristic of the third central moment, the skewness, describes a symmetry of the function $p(z)$. It is defined as

$$\gamma_1 = \frac{\mu_3}{\mu_2^{3/2}} = \frac{\mu_3}{R_q^3}.$$

A normalized form of the fourth central moment is the kurtosis:

$$\gamma_2 = \frac{\mu_4}{\mu_2^2} - 3 = \frac{\mu_4}{R_q^4} - 3.$$

This parameter becomes zero for the Gaussian distribution and a negative value indicates a more flat distribution. Another parameter often used in the literature is the arithmetic roughness (the centre line average):

$$R_a = \frac{1}{n} \sum_{i=0}^{n-1} |z_i - \langle z \rangle|.$$

This parameter, however, is usually nearly proportional to the standard deviation R_q [250].

The second group of vertical parameters involves characteristics based on extreme values of the set Z [257]. The most widely used parameters are the highest height R_p ($R_p = z_{\max} - \langle z \rangle$) and the maximum depth R_v ($R_v = \langle z \rangle - z_{\min}$) which are the maximal and minimal values of the relative coordinates associated with a selected altitudinal level $\langle z \rangle$, respectively. Another parameter is the vertical range, R_z, representing their sum $R_z = z_{\max} - z_{\min}$ (Figure 3.1).

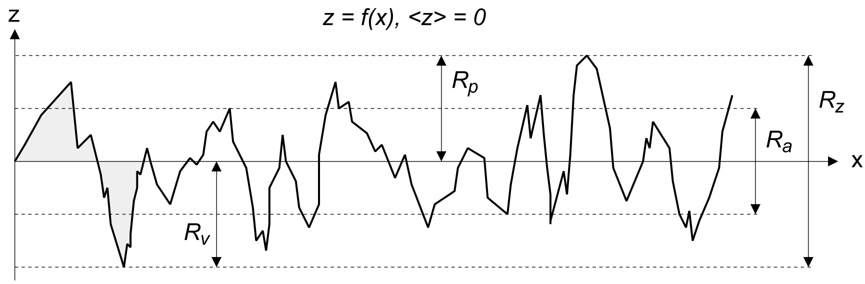

Figure 3.1 Graphical definition of basic extreme vertical parameters: the maximum hight R_p, the maximum depth R_v and the vertical range R_z. For comparison, the arithmetic roughness R_a is also depicted. The *grey area* marks the first element of the profile

Length Parameters

Length parameters describe the distribution of specific altitudinal levels of the surface in the horizontal plane x–y. These parameters are particularly useful for quality control in manufacturing processes. Their application in the fractography is rather rare. The most commonly used length parameters are, for example, the average spacing of profile elements on the mean line, S_m, the number of profile intersections with the mean line n_0 or the number of peaks per the length unit m_0.

By evaluation of the parameter m_0, for example, the peak is counted only when its horizontal distance from the previously counted peak is higher than $\frac{1}{10}$ of the vertical range R_z [251]. Consequently, the relation $x_i^p - x_{i-1}^p > \frac{1}{10} R_z$ must be fulfilled, where x_i^p is the x-coordinate of the i-th peak and x_{i-1}^p is that of the previously counted peak.

Hybrid Parameters

Hybrid parameters can be understood as a combination of altitudinal and length characteristics [250, 252]. In quantitative fractography, the linear roughness R_L and the area roughness R_A have been used for some time. These dimensionless characteristics, sometimes respectively called the relative profile length and the relative surface area, are defined as

$$R_L = \frac{L}{L'},$$

$$R_A = \frac{S}{S'},$$

where L is the fracture profile length, S is the area of the fracture surface, L' is the profile projection length and S' is the surface area projection into the macroscopic fracture plane. In the case of the profile composed of z randomly oriented linear segments $R_L = \frac{\pi}{2}$, whereas for the fracture surface composed of randomly oriented (nonoverlapping) facets $R_A = 2$ [214]. The area roughness R_A can be roughly assessed by means of the linear roughness R_L as

$$R_A = \left(\frac{4}{\pi}\right)(\langle R_L \rangle - 1) + 1.$$

The average slope Δ_a and its standard deviation Δ_q are defined in the following manner:

$$\Delta_a = \frac{1}{n-1} \sum_{i=0}^{n-2} \frac{|z_{i+1} - z_i|}{(x_{i+1} - x_i)},$$

$$\Delta_q = \left[\frac{1}{n-1} \sum_{i=0}^{n-2} \left(\frac{|z_{i+1} - z_i|}{(x_{i+1} - x_i)} - \Delta_a\right)\right]^{\frac{1}{2}}.$$

It should be noted that the parameters R_L and R_A also provide information about the angular distribution of surface elements [258].

Spectral Parameters

Spectral character of the profile can be described by means of the autocorrelation function which is a quantitative measure of similarity of the "original" surface to its laterally shifted version [251, 253]. Thus, the autocorrelation function expresses the level of interrelations of surface points to neighbouring ones. In the case of the fracture profile described by a set of n equidistant points (or $m \times n$ points obtained by sampling using constant steps in the directions of coordinate axes x, y), the autocorrelation function is defined by the following relations:

$$R(p) = \frac{1}{(n-p)} \sum_{i=0}^{n-p-1} (z_i - \langle z \rangle)(z_{i+p} - \langle z \rangle),$$

$$R(p,q) = \frac{1}{(m-p)(n-q)} \sum_{i=0}^{m-p-1} \sum_{j=0}^{n-q-1} (z_{i,j} - \langle z \rangle)(z_{i+p,j+q} - \langle z \rangle),$$

where p and q are shifts in the directions of x and y. The autocorrelation function has the following properties:

1. $R(0) = \mu_2$ or $R(0,0) = \mu_2$;
2. $R(p) = R(-p)$ or $R(p,q) = R(-p,-q)$;
3. $R(0) \geq |R(p)|$ or $R(0,0) \geq |R(p,q)|$ which means that the autocorrelation function attains a maximum for zero shifts.

With respect to the first attribute, the autocorrelation function is often normalized so that $R(0) = 1$ or $R(0,0) = 1$. The normalized autocorrelation function is usually denoted as $r(p)$ or $r(p,q)$:

$$-1 \leq r(p) = \frac{R(p)}{\mu_2} \leq 1, \qquad -1 \leq r(p,q) = \frac{R(p,q)}{\mu_2} \leq 1.$$

Autocorrelation lengths β_p and β_q are defined as shifts p, q corresponding to a drop of the autocorrelation function to a given fraction of its initial value. The fractions $\frac{1}{10}$ and $\frac{1}{e}$ are most frequently utilized [251,253]. Consequently, the surface points more distant than β_p, β_q can be assumed to be uncorrelated. This means that the related part of the fracture surface was created by another, rather independent, process of surface generation.

The character of the spectral surface can also be described in the Fourier space. The most important characteristic is the power spectral density

$$G(\omega_p) = |F(\omega_p)|^2,$$

$$G(\omega_p, \omega_q) = |F(\omega_p, \omega_q)|^2,$$

where ω_p and ω_q are space frequencies in the directions of coordinate axes x and y [251,254]. Functions $F(\omega_p)$ and $F(\omega_p, \omega_q)$ represent relevant Fourier transforms of the fracture surface:

$$F(\omega_p) = \frac{1}{n} \sum_{k=0}^{n-1} z_k \exp\left\{-\mathrm{i}2\pi\left(\frac{k\omega_p}{n}\right)\right\}, \tag{3.1}$$

$$F(\omega_p, \omega_q) = \frac{1}{mn} \sum_{k=0}^{m-1} \sum_{l=0}^{n-1} z_{k,l} \exp\left\{-\mathrm{i}2\pi\left(\frac{k\omega_p}{m} + \frac{l\omega_q}{n}\right)\right\}, \tag{3.2}$$

where i is the imaginary unit [254]. Equations 3.1 and 3.2 define the so-called discrete Fourier transform (DFT). The conventional factors $\frac{1}{n}$ and $\frac{1}{mn}$ might differ for various applications. Instead of the highly computationally demanding DFT the fast Fourier transform is often utilized.

Fractal Parameters

Fractal geometry is a mathematical discipline introduced by Mandelbrot [259] in the early 1980s. It is widely utilized as a suitable tool for the description of jagged natural objects of complicated geometrical structure. Fundamental properties of fractal objects are so-called self-similarity or self-affinity which mean an invariance with respect to scale changes. As a measure of the fractality the Hausdorff (fractal) dimension D_H is often used. The metrics of D_H can be determined by means of the Hausdorff measure

$$\Gamma_H^d = \lim_{\varepsilon \to 0} \inf_{U_i} \sum_i (\text{diam } U_i)^d \,. \tag{3.3}$$

When calculating the Hausdorff measure, the object is covered by cells U_i. The diameter of each cell meets the following condition: diam $U_i = \sup \{|x - y| : x, y \in U_i\} \leq \varepsilon$. Consequently, one searches the cell network minimizing the sum in Equation 3.3 for an infinitely small diameter of covering cells ($\varepsilon \to 0$). There is only a single value of D_H fulfilling the conditions $\Gamma_H^d = 0$ for each $d > D_H$ and $\Gamma_H^d = \infty$ for each $d < D_H$. This value is called the Hausdorff (fractal) dimension of the object. In the case of a smooth (Euclidean) object $D_H = d_T$, where d_T is the topological dimension, whereas $d_T < D_H \leq (d_T + 1)$ holds for the fractal object. In general, D_H is a rational number exceeding the topological dimension. A higher D_H-value means a higher segmentation of the object. As an example of the fractal object, Von Koch's curve is depicted in Figure 3.2 along with several first steps of its construction.

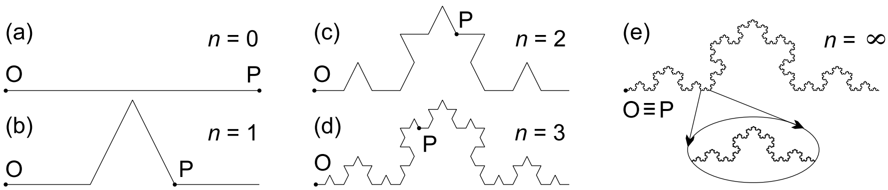

Figure 3.2 Von Koch's curve: (a) fractal initiator, (b) first iteration, (c) second iteration, (d) third iteration, and (e) final fractal ($D_H \approx 1.262$)

As can be seen from Figure 3.2, the length of the curve increases with increasing number of iterations and, for the final fractal, it becomes infinite. On the other hand, the area under the curve remains finite and practically unchanged. The infinite length of the fractal curve means that the marked points O and P retain the same distance during all iterations. Paradoxically, however, they coincide in the case of the final fractal ($\lim_{n\to\infty} |OP| = 0$). Real natural objects exhibit a statistical self-similarity rather than a perfect deterministic one. This means that the self-similarity does not hold for the object itself but only for its statistical parameters (average, variation, etc.) [260].

With respect to difficulties connected with the calculation of the Hausdorff dimension Γ_H^d directly according to Equation 3.3, various other simpler solutions were derived. The most widely used methods are depicted in Figure 3.3. By application of those methods to real objects, however, deviations from the theoretical fractal dependencies are usually observed [249, 261, 262]. A sigmoidal trend, obtained when calculating the parameter D_H, can serve as a typical example.

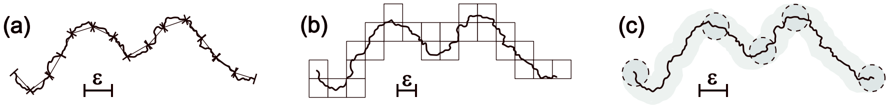

Figure 3.3 Some computation methods of the of fractal dimension: (a) perimeter method, (b) computation of squares, and (c) Minkowski method

Calculation of the of areas fractal dimension ($d_T = 2$) is more complicated and, as usual, it is performed either by means of space versions of curve methods [263, 264] or using the area-perimeter method. The latter method analyzes the fractal dimension of boundary curves of "islands" created by intersections of the horizontal plane with the object surface [264, 265].

Self-affinity is a more general form of self-similarity. A regular object exhibiting self-affinity is invariant to the transformation

$$x' \to \lambda_x x, \qquad y' \to \lambda_y y, \qquad z' \to \lambda_z z,$$

where $\lambda_y \propto \lambda_x^{\nu_y}$ and $\lambda_z \propto \lambda_x^{\nu_z}$ so that $\lambda_z \propto \lambda_y^{\nu_z/\nu_y}$. The ratio $H = \nu_z/\nu_y$ is called the Hurst exponent (the exponent of self-affinity), $H \in \langle 0; 1 \rangle$. When all contraction coefficients are equal ($\lambda_x = \lambda_y = \lambda_z$), the same transformation describes the self-similarity (see Figure 3.2(e)). In the case of isotropic surfaces, $\lambda_x = \lambda_y$ and $H = \nu_z$. This relation also refers to the arbitrary self-affine plane curve [260]. Again, the natural objects exhibit a statistical self-affinity rather than a deterministic one. Many experiments reveal that the fracture surfaces of most materials exhibit such a property [260, 266, 267]. Indeed, the self-similarity is usually preserved in the horizontal plane x–y (the area-perimeter method is based on that assumption), whereas the self-affinity is associated with the z-coordinate.

The Hurst exponent also yields information on a degree of internal randomness. When the object can be described by the Hurst exponent $H > H^*$, where $H^* = \frac{d_T}{d_T+1}$, the trend in the local site x (e.g., low or high z-values) is most probably followed by a similar trend in every other site $x + \Delta x$ (the persistence or the long-term memory). On the other hand, $H < H^*$ means an opposite tendency (antipersistence or short-term memory). The former type is typical for brittle fractures whereas the latter is typical for ductile ones [260].

As a rule, the Hurst exponent is calculated by means of the so-called variable bandwidth method [267, 268]. First, the profile is divided into k

movable windows (size ε) and, for each window, the local value of R_q is determined. Then the dependence of the average (global) characteristic

$$W = \frac{1}{k}\sum_{i=0}^{k-1} R_{qi}$$

on ε is established. Finally, the Hurst exponent is calculated according to the relation $W(\varepsilon) \propto \varepsilon^H$ as a slope of the exponential function plotted in bi-logarithmic coordinates (similarly to the fractal dimension).

3.1.2 Morphological Patterns in Fatigue

The morphological features in fatigue are different when looking at the fracture surface under low and high magnifications. Therefore, they can simply be separated into macroscopic and microscopic patterns [149, 187, 269, 270].

Macroscopic Patterns

The fractographic investigation of fatigue fracture surfaces usually starts visually or by low-magnification optical microscopy. Because of a change from a stage I slip plane that is inclined to the fracture plane macroscopically perpendicular to the principal loading axis, one can often distinguish a boundary between stage I and stage II crack propagations. It should be noted that the stage I fracture never extends beyond about three grains around the crack initiation site at the surface. The transition from stage I to stage II is usually accompanied by ridges parallel to the crack propagation direction. These ridges are formed by local shears that merge many plateaus corresponding to different fracture plains in individual grains.

One of the most important features usually found on fatigue fracture surfaces is beach marks (arrest marks), which are centred around the point of fatigue crack origin. These patterns correspond to stage II of crack propagation, and occur as a result of changes in loading or frequency or by oxidation of fracture surfaces during periods of crack arrest from intermittent service of the component.

The final-fracture zone can usually be identified by a fibrous morphology, which is different to that of the stage II crack growth. The size of this zone depends on the magnitude of loading, and its shape depends on the mode of external loading. Therefore, the transition boundary (line) between the stage II and final fracture zones can be related to the value of critical stress intensity factor K_c, associated with the fast (unstable) fracture. In ductile materials, shear lips (plane-stress) at approximately 45° to the fracture surface appear at the end of the final-fracture zone related to the free surface of the component. These lips enable us to differentiate clearly the final fracture zone from that of the initial crack growth stage.

Microscopic Patterns

During stage II of fatigue crack propagation, striations are formed in the way shown in Section 3.2.3. According to the model relevant for the Paris–Erdogan region, the striation spacing should be equal to the local crack growth rate, which is particularly true for ductile and homogeneous materials. Fatigue striations often bow out in the direction of crack propagation and generally tend to align perpendicular to the principal (macroscopic) crack growth direction. However, the crack locally propagates along multiple plateaus (facets) so that the local crack propagation directions in individual facets are usually different. The SEM picture of a striation field which documents this phenomenon is presented in Figure 3.4. Moreover, the plateaus can lie at different elevations with respect to each other and join by tear ridges or walls that also contain striations.

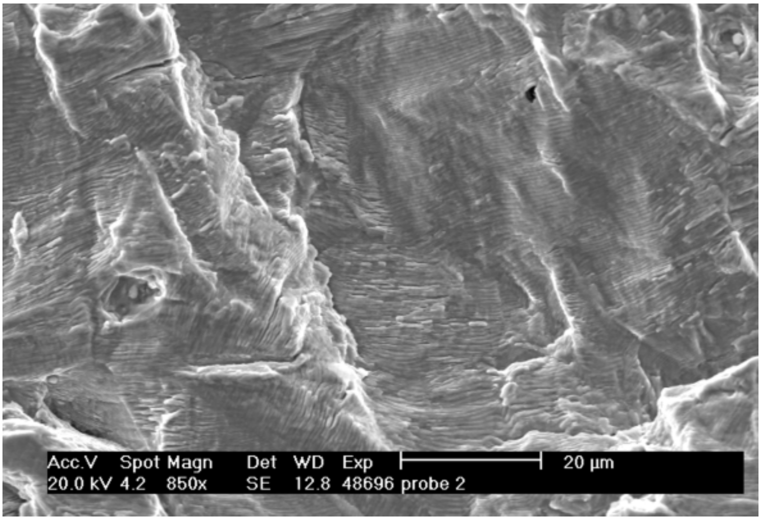

Figure 3.4 The striation field in the austenitic steel

Therefore, when comparing the macroscopic crack growth rate with the striation spacing for a given crack length, one has to consider an average of local crack growth rates indicated within the related striation field. This can be done by assuming the average of projections of local crack growth rates, perpendicular to the striations on individual facets, to the macroscopic propagation direction by using the formula

$$\bar{s} = \frac{1}{n} \sum_{i=1}^{n} s_i \cos \alpha_i, \tag{3.4}$$

where n is the number of facets, s_i is the striation spacing on the i-th facet and α_i are the related projection angles (see Figure 3.5)

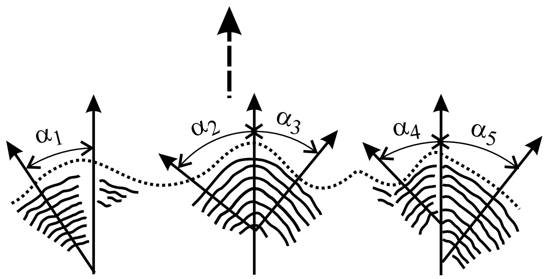

Figure 3.5 The scheme of local inclination angles of striation fields. The macroscopic crack growth direction is marked by the *arrow*

On striation facets which are at an angle to the average plane of crack growth, so-called fissures can be formed [271, 272]. They appear as short (secondary) cracks related only to one surface of the local crack, are regularly spaced, and penetrate a distance below the fracture surface which is much larger than the variations in fracture surface topography. The fissures form at striations most likely due to their stress concentration effect. Because the formation of fissures causes local crack branching, the further striations created on the inclined facet start to be immediately shielded especially from the local tensile stress parallel to the fracture surface. As the crack front moves away from the fissure, the local tensile stress builds up to the point where a new fissure is formed, and the process is repeated. Therefore, the fissure spacing does not directly correspond to the local crack growth rate, similarly to the spacing of false striations created in the near-threshold region (Section 3.2).

There are also other periodic patterns, sometimes observed on fatigue fracture surfaces generated by a shear loading, known as fibrous patterns and tire tracks. Both these markings are produced by a contact wear of crack-flank asperities. The fibrous patterns are, unlike striations, parallel to the crack growth direction and can be clearly seen in Figure 3.6. The tire tracks, resembling the tracks left by the tread pattern of a tire, are the result of particles or protrusions on the fracture surface being successively impressed into the mating surface part during the closing portion of the loading cycle. They are nicely depicted especially in Figure 3.52 (Section 3.3). An appearance of both fibrous patterns and tire tracks always indicates a presence of either remote or local shear mode II at the crack front. The local shear mode can also be induced in the case of a pure remote mode I by inclinations of crack front elements from the main crack plane perpendicular to the remote loading direction (see also Section 3.3). The direction of the tire tracks and the change in the spacing of indentations can also indicate both the magnitude and the type of displacements that occurred during the fracture process, such as lateral movement from shear or torsional loading. There is, again, no simple correspondence between the spacing of tire tracks (or fibrous patterns) and the crack growth rate.

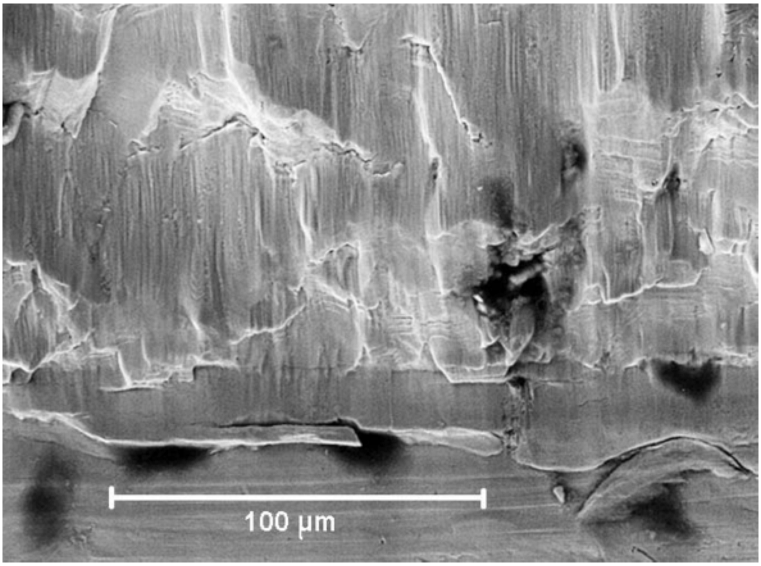

Figure 3.6 Fibrous patterns and tire tracks formed under the shear (mode II) crack propagation in austenitic steel. The crack growth direction is from the bottom to the top

Sometimes, periodic microstructural phases (pearlite, martensite laths, etc.) or slip traces can also be observed on fracture surfaces. Such patterns, obviously, have nothing to do with the crack growth rate. Thus, the fields of fatigue striations are the only relevant patterns that can be directly correlated with the rate of crack front propagation. A clear distinction between the striations and other periodical features is not trivial. Therefore, when performing a reconstitution of the fatigue process based on the fracture surface micromorphology, one should possess a sufficient level of fractographic experience.

3.2 Opening Loading Mode

There is common agreement within the international scientific community concerning a principal physical difference between the driving force of the fatigue (stable) crack growth and that of the brittle (unstable) fracture. While the latter is directly associated with a drop in the elastic energy (elastic strain) and the critical value of the stress intensity factor K_c, the driving force in fatigue is directly related to the range of the cyclic plastic strain at the crack tip. Because the maximum K-value during the stable crack growth lies below K_c, the crack growth can proceed only when supported by the work of external cyclic forces. Similarly, the stable growth in the case of stress corrosion cracking presumes the assistance of the chemical driving force. In 1963, Paris and Erdogan [273] proved that the diagram $\mathrm{d}a/\mathrm{d}N$ *vs* ΔK for so-called long cracks in the small-scale yielding range (the high-cycle fatigue) retains the advantage of LEFM, namely a satisfactory invariance in the shape and size of cracked solids. It might seem to be surprising that the linear elastic parameter also allows us to describe successfully the rate of plastic processes at the crack tip. Several years later, however, Rice [274] brought to light a theoretical reason justifying the present opinion: the small-scale cyclic plasticity (the cyclic plastic zone) at the crack tip is, indeed, controlled by the value of ΔK.

The local value of ΔK at the crack tip is determined by both external and internal stresses resulting from external forces and local plastic deformations (or generally from microstructural defects), respectively. The tensile and compressive elastic energies associated with internal stresses are mutually compensated all over the bulk and, therefore, they do not contribute to a global tensile elastic energy of stressed solids. Consequently, they cannot be released to support the unstable fracture process. At the same time, limited amounts of elastic energy that can be released by relaxation of local tensile internal stresses in small volumes adjacent to the crack tip can significantly influence neither the onset of brittle fracture nor its crack growth rate. On the other hand, the level of internal stresses can substantially influence the stable fatigue growth rate in each step of the crack advance, since the emission of dislocations from the crack tip occurs at very low stress intensity factors. This means that even very small changes of local K-values at the crack tip can considerably modify the stable crack growth rate. Thus, unlike in the case of unstable fracture, the internal stresses created by dislocation configurations and secondary phases are to be considered as an important additional factor affecting the fatigue crack propagation rate. It is particularly this difference that elucidates a much higher complexity of shielding (or anti-shielding) effects accompanying the fatigue crack growth when compared to brittle fracture. Unlike in brittle fracture, for example, the contact shielding in fatigue also occurs under the opening loading mode, which causes so-called crack closure phenomena. Moreover, many important phenomena associated especially with the small scale yielding, e.g., the existence of fatigue thresholds or a delay in ratcheting, can be explained only when the behaviour of discrete dislocations is taken into account. This underlines the necessity to utilize the multiscale approaches to the fatigue crack growth phenomenon which is documented in Chapter 3. Unfortunately, those fundamental diversities of stable and unstable crack growths still do not seem to be widely understood, particularly among “classical” mechanical engineers dealing solely with continuum mechanics.

There is, however, another important factor strongly influencing the stable crack growth rate in contrast to the unstable one. This is the effect of environment, even that of the air. Simply, the stronger the chemical interaction between the environment and the material, the higher the fatigue crack growth rate. For a majority of metallic materials, the crack growth rates in air are two orders in magnitude higher than those in vacuum [275]. Although the detailed mechanisms of chemical processes are not examined here, the micromechanical reasons for their strong influence are clearly outlined hereafter. On the other hand, the growth rate of unstable (brittle) cracks remains unaffected by the environment because there is not enough time available to produce a chemical damage during the fast fracture.

Let us emphasize that the mechanism of cyclic plasticity immediately elucidates why the fatigue cracks avoid propagating in hard (or brittle) microstructural phases. In such phases, the movement of dislocations is strongly limited and, therefore, the fatigue crack is always repelled to a softer material [276, 277]. The same qualitative explanation holds for an increase in the fatigue limit with an increasing hardness (or ultimate strength). Indeed, the fatigue limit is related to a critical stress under which the microstructurally short cracks still remain arrested [149, 247]. These cracks are smaller in high-strength materials (finer microstructure) so that the stress necessary for their further growth is higher. The fundamental difference in micromechanisms of brittle and fatigue crack propagation also has a consequence in the following well–known phenomenon: whilst a continuous increase in the materials ductility along the path of propagating brittle cracks (induced, e.g., by a temperature gradient) causes their arrest, this is generally not the case of fatigue cracks.

Large deformations inside the plastic zone in both the Paris–Erdogan and the near fracture regions of long–crack propagation can be described by classical continuum theories (Figure 3.7). However, these regions are of less engineering importance than those of the short crack growth and the long–crack threshold. Here, on the other hand, the plastic zone is relatively small and the dislocation activity is confined to one or two favourable slip systems. Under such conditions, the experimentally observed phenomena can be sufficiently elucidated only when taking the discrete nature of plasticity into account. Therefore, Sections 3.2.1, 3.2.2 and 3.2.3 are devoted to discrete dislocation models of cyclic plasticity and fatigue crack growth.

The results presented here are essential for a sufficient grasp of both the crack closure effects and the unified model of the crack-tip shielding, as de-

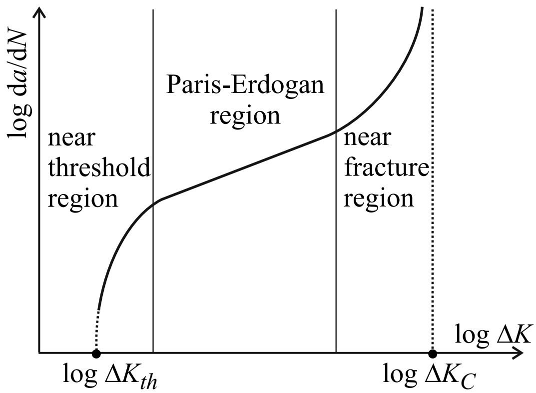

Figure 3.7 Scheme of the crack-growth rate *vs* ΔK dependence for long fatigue cracks

scribed in Sections 3.2.4 and 3.2.5. The unified model was developed as a multiscale concept involving three basic levels; micro – crystal defects, meso – grain (phase) microstructure and macro – continuum. The discrete dislocation theory and the size ratio effect constitute links between these three levels. It is important to emphasize, however, that a full comprehension of the crack growth threshold phenomenon requires an insight into the atomistics of dislocation emissions from the crack tip. Section 3.2.6 presents results of an extended application of the unified model to identify the shielding components as well as the intrinsic resistance to the near-threshold fatigue crack growth in various metallic materials. Finally, Section 3.2.7 is devoted to the influence of shielding effects on the crack growth rate in the Paris–Erdogan region.

3.2.1 Discrete Dislocation Models of Mechanical Hysteresis

3.2.1.1 Hysteresis Loop

The process of cyclic plastic deformation controls both the crack initiation and the rate of fatigue crack propagation. The Nabarro–Cottrell analysis [156] represents a rather simplified model of mechanical hysteresis but it is very useful for understanding the physical background of that process. This analysis was applied to provide an insight into the early stage of cyclic softening as well as to the micromechanism of ratcheting (the cyclic creep) in polycrystalline materials at room temperature [149, 278, 279].

A polycrystalline material at the onset of cyclic softening can be considered to be a nearly elastic aggregate containing a small number of perfectly plastic grains. The macroscopic stress-strain response of such a system of n grains corresponds to a composite of n_{el} elastic grains with Young's modulus E (a major phase) and n_{pl} plastic grains with Young's modulus $E_2 \ll E$ (a minor phase). Denoting the relative number of plastic grains as $\alpha = n_{pl}/n$ ($n \approx n_{el}$), the composite modulus E_c can then be expressed as

$$E_c = E\left(1 - 1.9\frac{\alpha}{1+\alpha} + 0.9\frac{\alpha^2}{(1+\alpha)^2}\right) \approx E(1-2\alpha),$$

in analogy to a porosity influence in ceramic materials [149, 280].
The aggregate strain can be expressed as

$$\begin{array}{ll} \varepsilon = \sigma/E & \text{for } \sigma < \sigma_0, \\ \varepsilon = \varepsilon_e + \varepsilon_p = \sigma/E + 2\alpha(\sigma - \sigma_0)/E & \text{for } \sigma > \sigma_0, \end{array} \tag{3.5}$$

where σ_0 is the yield stress of the composite. The stress-strain response of such aggregate is depicted in Figure 3.8 in $\tau - \gamma$ shear coordinates ($E \to G, \sigma \to \tau, \varepsilon \to \gamma$).

By raising the applied tensile stress τ from zero, the strain remains elastic up to the point A in Figure 3.8(a), where $\tau = \tau_0$. At this moment the first dislocations are emitted from Frank–Read (F-R) sources and create pile-ups at grain boundaries. These pile-ups hinder further dislocation emissions by producing an increasing back stress to F-R sources. Note that, in engineering applications, this stress is usually declared as a residual stress that remains in the material after unloading. Thus, the dislocations are emitted from the F-R source and produce plastic strain as long as the increasing effective shear stress $\tau - \tau_0$ compensates the back stress. At the point B the applied stress stops increasing ($\tau = \tau_{\max}$), and the following equilibrium equation holds:

$$\tau_{\max} - \tau_0 - \tau_B = 0, \tag{3.6}$$

where τ_B is the back stress.

Now we start to reduce the tensile stress from $\tau_{\max} \to \tau_{\max} - \Delta\tau$. Before a sufficiently high value of $\Delta\tau$ is reached, the dislocations cannot move back and the reverse deformation proceeds in an elastic manner (the segment BC in Figure 3.8(a)). The backward motion of dislocations begins only after a stress reversal at the point C, where the sum of the applied (reversed) stress and the back stress becomes equal to the compressive yield stress:

$$\tau_{\max} - \Delta\tau - \tau_B + \tau_0 = 0. \tag{3.7}$$

Combining Equations 3.6 and 3.7 one obtains

$$\Delta\tau = 2\tau_0.$$

This simple relation reveals that, for the unloading strain path, a doubled yield stress is to be taken into account. The value of the compressive yield stress $\tau_{\max} - 2\tau_0$ corresponding to the point C is lower than that of the tensile

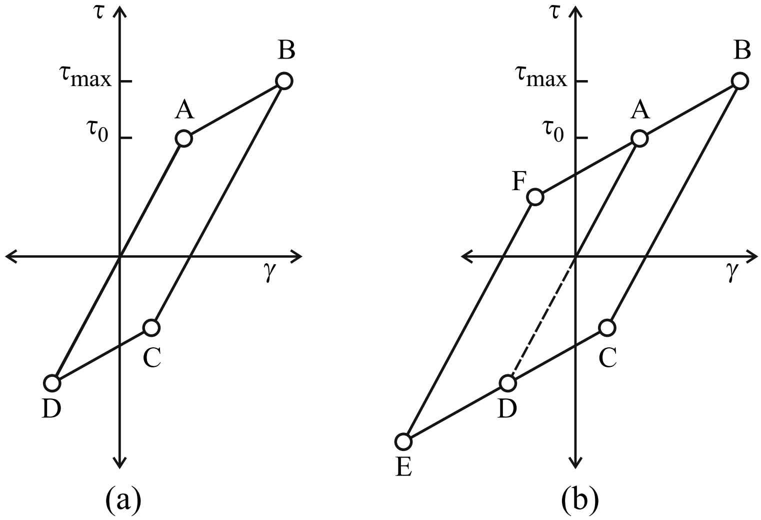

Figure 3.8 Scheme of the mechanical hysteresis behaviour in the frame of the Nabarro–Cottrell analysis: (a) the asymmetric loading cycle, and (b) the symmetric loading cycle

yield stress. This is the well-known Bauschinger effect. Beyond the point C, the dislocations move backwards while gradually reducing the back stress up to the point D, where $\tau_B = 0$. If the applied compressive stress starts to be reduced now, the whole process is repeated in the frame of a closed hysteresis loop. Such a loop is also created when the reverse deformation starts before reaching the point B or proceeds beyond the point D. An example of the latter case for a symmetrical loading is shown in Figure 3.8(b). Indeed, an inverse back stress is created within the segment DE, which causes an appropriate reduction of the tensile yield stress at the point F, similarly to that of the compressive yield stress at the point C. Let us finally emphasize that such a behaviour presumes a totally reversible dislocation slip.

As a rule, the elastic parts of real hysteresis loops in engineering materials are somewhat shorter, while the plastic ones are longer and curved. A main reason for that fact constitutes a statistical distribution of values of the yield stress in various grains [281]. It should also be noted that all the discrete dislocation models presented in this book utilize only dislocations inevitable for simulation of cyclic plasticity phenomena, the so-called geometrically necessary dislocations. There are also other dislocations forming the dislocation structure in real polycrystals, the so-called statistically stored dislocations. For more details about the geometrically necessary dislocations, see [8].

3.2.1.2 Ratcheting

The ratcheting (cyclic creep) is a process of gradual plastic elongation or contraction of a sample during cycling loading of a constant nominal stress amplitude. The first experimental observation of the ratcheting at room temperature was reported by Kennedy [282]. This process can lead to a substantial reduction of fatigue life in comparison with the strain-controlled loading of the same stress range [283, 284]. For example, failures of helicopter airscrews can be induced by ratcheting [285]. When loadings of a high asymmetry near the fatigue limit are applied, the cyclic microcreep can cause undesirable shape changes of precisely manufactured components as spiral springs or turbine blades.

In most practical cases, the prescribed load (work) level is achieved only after a certain period of time (the so-called ramp-loading). During that starting period, the plastic strain is increasing and the ratcheting process can only appear when the plastic strain range exceeds a certain critical value $\Delta\varepsilon_{pc}$ which is characteristic for a particular material. This means that, during the loading, no ratcheting occurs when either the highest achieved value of the plastic strain range $\Delta\varepsilon_p$ is less than $\Delta\varepsilon_{pc}$ or a cyclic hardening starts before reaching $\Delta\varepsilon_{pc}$. In the case of cyclically softening materials such behaviour is observed even when the work load is reached just in the first loading cycle. The scheme of a typical hysteresis behaviour during the asymmetric loading of cyclically softening materials before and after the onset of ratcheting is shown in Figure 3.9. However, remarkable elongations may occur even when a symmetrical loading cycle is applied (the cyclic ratio $R = \sigma_{\min}/\sigma_{\max} = -1$) [286, 287].

The value of $\Delta\varepsilon_{pc}$ is a material characteristic and depends on the cyclic ratio (decreases with increasing R) – see e.g., [278, 287]. The elongation or contraction per one loading cycle (the ratcheting rate) increases with increasing both the cyclic softening rate and the cycle asymmetry [149, 288].

As first described in [279], the analysis based on the discrete dislocation model of the hysteresis loop can be utilized to elucidate all the experimentally observed phenomena. Obviously, a closed hysteresis loop means that no ratcheting can take place. According to the above-mentioned Nabarro–Cottrell analysis, the incomplete closure of the hysteresis loop must be caused by micromechanisms producing an irreversibility of the cyclic plastic strain. During initial stages of a global cyclic softening, the plastic deformation is gradually transferred from pile-ups in most favourably oriented grains into the adjacent grains [289]. Simultaneously, a cyclic hardening starts to take place in some of the already plasticized grains. At the end of the global cyclic softening stage, all grains become plastic and the global response changes to the cyclic hardening. During the global cyclic softening stage the density of dislocations increases by a cooperative operation of Frank–Read sources and, in this way, primary dislocation networks and loop patches are successively created. In their vicinity, a gradual increase of internal stresses results in an activation of secondary slip systems. Schemes of both the hysteresis loop and the primary slip system in Figure 3.10(a) correspond to a moment just before the activation of the secondary slip. The interaction of primary and secondary dislocations leads to a formation of sessile dislocations, i.e., either to the Lommer–Cottrell barriers (in fcc metals) or to the [001] dislocations (in bcc metals). This means a start of the slip irreversibility, since the primary dislocations become locked in between the secondary (newly created) barrier and the primary (microstructural) barrier. Consequently, they cannot return to the Frank–Read source during the reverse half-cycle. Moreover, the secondary slip always reduces the back stress of the pile-up (regardless to a creation of the secondary barrier). This corresponds to an increase in the compressive yield stress. These micromechanisms are schematically depicted in Figure 3.10(b) ($\Delta\tau_B$ is the increment of the compressive yield stress, GB is the grain boundary, SD is the sessile dislocation and SS is the secondary slip system). Both processes substantially reduce the reverse plasticity, and leave a residual tensile plastic strain γ_{rb}, i.e., initiate the ratcheting [278, 290]. The first experimental observation of a direct connection between the secondary slip activity and the ratcheting process was reported by Lorenzo and Laird [291]. Due to superimposed internal and external stresses, the creation of barriers is related to a peak of the applied true stress. If the peak is tensile (compressive), the ratcheting causes the elongation (contraction) of the specimen.

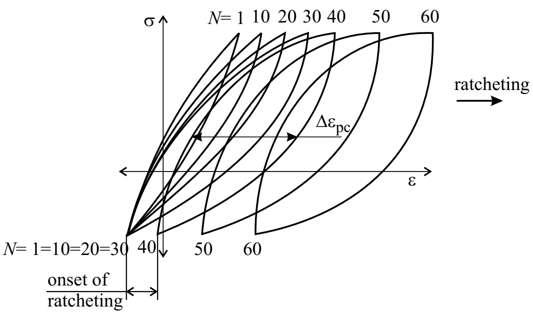

Figure 3.9 The scheme of the hysteresis behaviour of cyclic softening materials before and after the onset of ratcheting. The ratcheting starts after reaching the critical range of plastic strain $\Delta\varepsilon_{pc}$

In order to quantify the residual plastic deformation associated with the loop disclosure, let us consider the reverse slip of the pile-up in more detail. The Burgers vector density $B(x)$ within the pile-up can be expressed as

$$B(x) = \frac{\tau}{\pi c_0} \frac{x}{\sqrt{a^2 - x^2}},$$

where $c_0 = \mu/[2\pi(1-\nu)]$, ν is the Poisson's ratio, x is the coordinate along the pile-up and a is the length of the pile-up. In the first half-cycle the slip

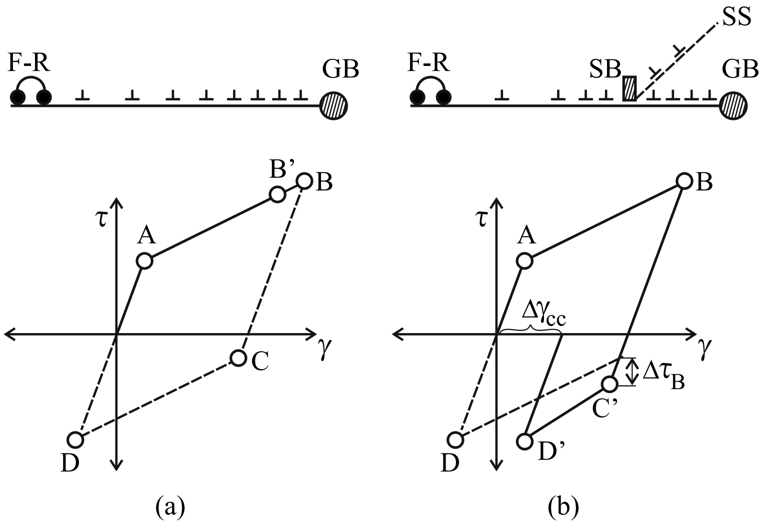

Figure 3.10 Incomplete closure of hysteresis loops produced by the secondary slip mechanism: (a) the moment just before a secondary slip activation (point B'), and (b) disclosure of the loop caused by the secondary barrier

of pile-up dislocations emitted from the F-R source is restricted only by the microstructural barrier. Consequently, the related shear strain

$$\gamma_r \sim \int_0^a B(x)x\mathrm{d}x = \frac{\tau}{\pi c_0}\int_0^a \frac{x^2}{\sqrt{a^2-x^2}}\mathrm{d}x. \tag{3.8}$$

When the maximum applied stress $\tau_{\max}$ is reached, the secondary slip creates the sessile dislocation. One can assume that the probability $P(x)$ of its location at the point x inside the pile-up is proportional to the density $B(x)$, i.e., $P(x) = DB(x)$. The probability of finding the dislocation barrier anywhere in between the F-R source and the primary obstacle (grain boundary) must be equal to 1, and, therefore, the following relation holds:

$$1 = \int_0^a P(x)\mathrm{d}x \Rightarrow 1 = \frac{\tau D}{\pi c_0}\int_0^a \frac{x}{\sqrt{a^2-x^2}}\mathrm{d}x \Rightarrow P(x) = \frac{1}{a}B(x). \tag{3.9}$$

After the creation of the secondary barrier, only dislocations located between the F-R source and that barrier can move back to the F-R source. These dislocations are positioned in the range $(0, y)$, where y is the coordinate of the secondary obstacle occurring with the probability $P(y)$. This means that, according to Equations 3.9 and 3.8, the total reverse shear displacement can be expressed as

$$\gamma_{rb} \sim \int_0^a \int_0^y P(y)B(x)x\mathrm{d}x\mathrm{d}y = \frac{\tau}{\pi c_0 a} \int_0^a \frac{y}{\sqrt{a^2-y^2}} \int_0^y \frac{x^2}{\sqrt{a^2-x^2}} \mathrm{d}x\mathrm{d}y. \quad (3.10)$$

After integrating Equations 3.8 and 3.10 one obtains

$$\gamma_{rb}/\gamma_r = 0.41.$$

This result shows that the secondary barrier prohibits at least a half of dislocations from returning back to the F-R source in the unloading half-cycle, which leads to a disclosure of the hysteresis loop. Obviously, this demands a sufficient density of primary dislocations to be generated during the cyclic softening in order to allow the creation of secondary barriers in the preferentially oriented large grains. From a macroscopic point of view it corresponds to a critical amount of a global cyclic plastic strain, i.e., the critical plastic strain range $\Delta\varepsilon_{pc}$ associated with the onset of ratcheting during the ramp-loading (generally during the global cyclic softening stage). This is a micromechanical interpretation of initial delays in the cyclic creep process observed for many metallic materials and various cyclic ratios [286,288]. The existence of ratcheting in case of symmetrical loading can be elucidated in a similar way. Due to a transverse contraction of the specimen (component), the peaks of the true stress come up in the tensile half-cycles. Because the creation of the first barriers is related to these peaks, the cyclic creep follows the tensile direction and elongates the specimen.

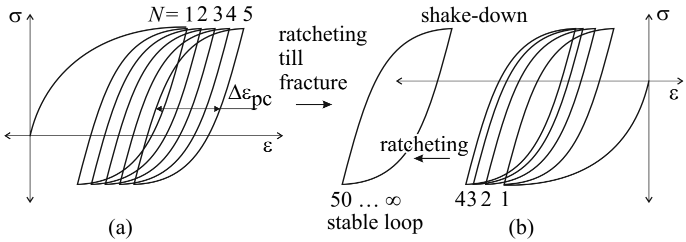

Figure 3.11 The scheme of the hysteresis behaviour when the critical plastic strain range is exceeded in all loading cycles: (a) the loading starts in the tensile direction, and (b) the loading starts in the compressive direction (the shake-down behaviour)

When the work load is reached just in the first loading cycle and, simultaneously, the critical value $\Delta\varepsilon_{pc}$ is exceeded, the kinetics of ratcheting strongly depends on the starting loading direction and the plastic strain range in the following cycles [292]. Such multislip plastic behaviour can already be described by continuum plasticity theories involving kinematic hardening [293,294]. The starting tensile (compressive) half-cycle produces the elongation (contraction) after a completion of the first cycle – see Figure 3.11. The ratcheting process proceeds only when the plastic strain range also remains sufficiently high in the next cycles (Figure 3.11(a)). This case corresponds to exceeding the so-called shakedown limit, usually expressed in terms of a critical applied stress value. When the ratcheting is oriented into the compressive direction, however, it eventually stops after a certain number of cycles due to a decreasing true stress range (an increasing diameter of the specimen) – see Figure 3.11(b). Such behaviour is known as a plastic shakedown. When the material hardens and the plastic strain range is relatively small, a rather slow ratcheting, reversed to that of the first half-cycle, is usually observed (Figure 3.12). As a rule, however, the loop stabilizes and the ratcheting shakes down after a certain number of cycles. This backward ratcheting is caused by high residual stresses which are created in the first half-cycle. During a certain period of a cyclic plastic deformation, these stresses gradually become relaxed by multislip.

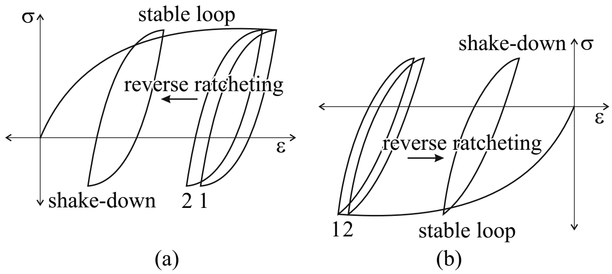

Figure 3.12 The scheme of the shake-down behaviour when the critical plastic strain range is exceeded only in the first loading cycle: (a) starting in the tensile direction, and (b) starting in the compressive direction

In order to assess the ratcheting rate in the initial softening stage, the difference between the plastic strains produced in the loading and unloading parts of the cycle is to be determined. Let us consider that during each loading cycle the reverse slip is restricted to one half due to the secondary slip in $\Delta\alpha$ grains. This means that the reversed plasticity will be reduced proportionally to $\Delta\alpha/2$. By using Equation 3.5, and respecting the reduction of the reversed plasticity, it leads to the following result:

$$\begin{aligned}\Delta\gamma_{cc} = \gamma_p - \gamma_{pR} &\approx \pm\left(\frac{2\alpha}{\mu}(\tau_{\max}-\tau_0) - \frac{2\alpha-\Delta\alpha}{\mu}(\tau_{\max}-\tau_0)\right) = \\ &= \pm\frac{\Delta\alpha}{\mu}(\tau_{\max}-\tau_0) = \pm\frac{\Delta\alpha}{\mu}\left(\frac{2}{1-R}\tau_a - \tau_0\right).\end{aligned} \tag{3.11}$$

The positive sign holds for $R > -1$ (elongation) and the negative sign for $R < -1$ (contraction). According to Equation 3.11 the ratcheting rate increases with both the increasing cycle asymmetry and the plastic strain range. Indeed, one can assume $\Delta\alpha \propto \Delta\varepsilon_p$, where $\Delta\varepsilon_p$ is the push-pull plastic strain range. This is in general agreement with experimental observations as well as with continuum plasticity models [295]. It should be emphasized, however, that the continuum mechanics is unable to interpret the two, already discussed, micromechanically induced phenomena: the critical plastic strain range related to the onset of ratcheting at the end of a delay and the tensile cyclic creep in the case of symmetrical loading.

Equation 3.11 was quantitatively verified by a simulation of initial ratcheting stages in ultra-high-strength steels [296]. Because $\Delta\alpha \approx k\Delta\varepsilon_p$ and the push-pull stresses σ and strains ε are proportional to the shear quantities τ and γ, Equation 3.11 can be rewritten in terms of σ and ε as

$$\frac{\Delta\varepsilon_{cc}}{\Delta\varepsilon_p} \approx \frac{k}{E}(\sigma_{\max} - \sigma_0). \tag{3.12}$$

According to Equation 3.12, the dependence $\Delta\varepsilon_{cc}/\Delta\varepsilon_p$ *vs* $\sigma_{\max}$ should be linear. Thus, the linear interpolation of experimental points is plotted in Figure 3.13 along with associated scatter ranges.

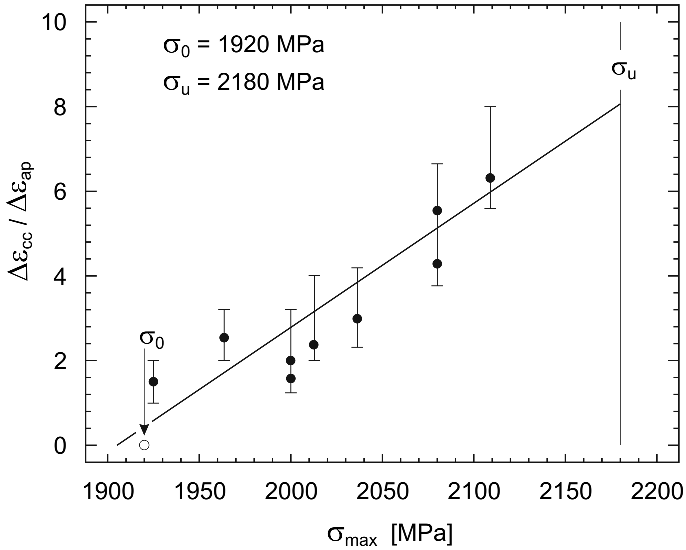

Figure 3.13 The normalized ratcheting rate as a function of the maximum loading stress. The *straight line* corresponds to the theoretical model

The plastic strain range $\Delta\varepsilon_p$ in the experiment was of the order of 10^{-5}. The extrapolated value $\sigma_0 = 1905\,\mathrm{MPa\,m}^{1/2}$ for the zero ratcheting rate agrees well with the experimental one ($\sigma_0 = 1920\,\mathrm{MPa\,m}^{1/2}$). The constant k is of the order of 10^4 and $\Delta\alpha \approx 10^{-1}$. This means that, in every loading cycle, the reversed dislocation slip was suppressed approximately in one of each ten grains. This seems to be a plausible result as well.

Let us finally note that similar irreversible micromechanisms also operate in the cyclic plastic zone ahead of an advancing fatigue crack front (see the next subsection), and contribute to the roughness-induced shielding (see Section 3.2.4).

3.2.2 Nucleation and Growth of Short Cracks

3.2.2.1 Mechanisms of Crack Nucleation

Fatigue cracks in metallic materials are nucleated by local interactions of dislocation slip bands or pile-ups with microstructural heterogeneities and defects as grain boundaries, phase boundaries, large secondary phase particles or free surfaces [149, 192, 297]. These interactions are induced by elastic mismatch strains evolving at boundaries of grains and microstructural phases with different stress-strain characteristics during the external loading. Although the related high elastic incompatibility stresses can be relaxed by plastic deformation (dislocation movements) in a nearly entire volume of a deformed solid, relatively high long-range stresses of dislocation arrangements in slip bands and pile-ups remain locally at microstructural boundaries. Since the cohesive strength (fracture energy) of incoherent boundaries can be very low, these stresses often lead to microcrack initiation at the weakest sites. Discrete dislocation models [298–300] revealed that the peak stresses could be higher than the cohesion stress, which depends on the surface and grain boundary energies. A high hydrostatic component of the long-range stresses intensifies a diffusion of interstitials into the stressed regions, thereby reducing the cohesive strength of the boundary. Following the molecular dynamics computations of Van der Ven and Ceder [301], a fraction of 40% of oxygen (hydrogen) atoms in a $\{111\}$ aluminium plane induces a relative reduction in the cohesive strength by a factor of two (three). In order to propagate further the nucleated microcrack, however, a sufficient amount of irreversible plasticity must be produced at its front. This is conditioned by a permanent flux of oxygen to avoid a repeating recovering of newly created fracture surfaces (see hereafter). Therefore, an intergranular propagation from the bulk interior along the grain or phase boundaries by repeating decohesion mechanism is usually observed only in strong corrosive environments. On the other hand, a transgranular crack propagation from the internal nucleus at an inclusion was often observed in specimens with a reinforced surface layers or in an ultra-high cycle regime (see Section 3.3.4).

It should be emphasized, however, that the fatigue cracks preferentially initiate at the surface [149, 192]. Indeed, the free surface is a boundary that separates media with an extremely high difference in properties (lack of interatomic bonds in air). In this particular case, the movement of dislocations towards the free surface is caused by attractive mirror forces (see also Section 2.3). During the first loading cycles progressive changes in the dislocation structure start to proceed. The damage processes begins at the sites of cyclic strain localization, usually called persistent slip bands (PSBs) and results in the formation of sharp surface slip patterns called persistent slip markings (PSMs). PSMs consist of extrusions and intrusions which develop on the initially flat surface at emerging PSBs (see Figure 3.14). While the amplitude of reversed plastic slip in the interior of PSBs or pile-ups is highly constrained by surrounding elastic matrix, this is not the case for the parts near the PSMs on the free surface where high local plastic deformations appear. Indeed, atomic force microscopy revealed that these strains can be two orders of magnitude higher than the applied macroscopic ones (e.g., [302]).

The slip localization in surface thin bands can be explained by several mechanisms. For example, cycling of precipitation hardened alloys induces formation of thin slip bands in zones with easily shareable (coherent) precipitates [303]. In ductile fcc and bcc polycrystals subjected to cyclic loading the PSBs are often formed within the surface grains. The PSBs consist of hard walls (high density of edge dislocation dipoles) and soft channels (low density of screw dislocations which glide and cross-slip) creating the well-known ladder-like structure (e.g., [304, 305]). This rather regular dislocation arrangement lowers the internal energy and enables high local shear deformations [306, 307]. Although the ladder-like structure is usually not observed in metals and alloys possessing a low stacking-fault energy, the PSMs occur in these materials [302].

The amount of plastic slip inside the slip bands and PSBs can be modelled by considering an elongated bulk inclusion embedded in the matrix which mimics the whole polycrystal [308]. By considering analytical solutions derived by Eshelby for bulk inclusions [309] one can show that the primary slip in the vicinity of the free surface is higher than that in the bulk, where the deformation is constrained by the grain boundary surrounded by a hard elastic material [310]. The slip within the band can be defined as the ratio between the displacement along the most active Burgers vector and the slip band thickness, h. Then the shear deformation within the surface slip bands can be expressed as

$$\gamma_p = r(1-\nu)\frac{L}{h}\left(1+\frac{h}{L}\right)^2 \frac{0.5\Sigma - \tau_{crs}}{G}, \tag{3.13}$$

where L is the length of the slip band (usually the grain size), $r = 1.9$, ν is the Poisson's ratio, G is the shear modulus, Σ is the applied stress and τ_{crs} is the critical resolved shear stress in the channel [310]. The factor 0.5 corresponds to a highest possible Schmid factor of the slip band inclined 45° from the loading axis. Equation 3.13 was recently verified by finite element calculations [311]. Since in large grains the values of L are tens of microns and the values of h are units or tens of nanometers, the ratio $h/L \approx 0.005 \ll 1$. With regard to Equation 3.13 this means that the surface plastic deformation in the slip band is proportional to L/h, i.e., it increases with increasing aspect ratio of the band. Clearly, the Eshelby inclusion elongated in the direction of the applied stress experiences lower back stress. Thus, the ratio of the strain ε_L localized in slip bands at the surface and the applied (nearly elastic) strain ε can reach values as high as $\varepsilon_L/\varepsilon \approx 200$. Note that the localization ratio related to the inner end of the slip band (impinged by a grain boundary) is about ten times lower [311].

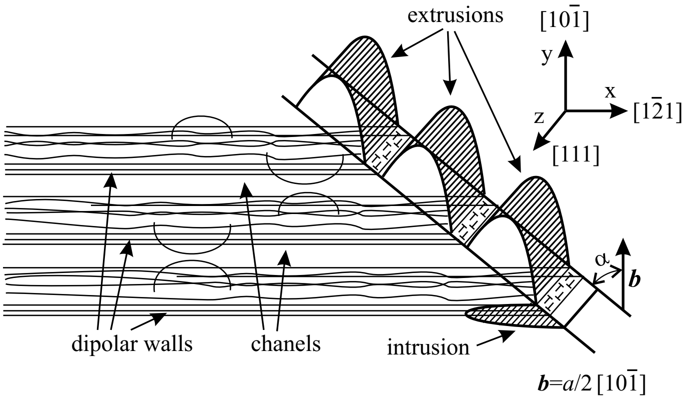

Figure 3.14 A scheme of extrusion and intrusion patterns at the intersection of the persistent slip band with a free surface

The high local reversed plasticity in slip bands and channels of PSBs produces a surface microroughness in the form of extrusions and intrusions (e.g., [312]). The movement of screw dislocations with jogs generates a surplus of vacancies in the channels. The counterbalancing flux of atoms inside the channels causes extrusions at the free surface, the volume of which is much higher than that of intrusions, as shown in Figure 3.14. On the other hand, volumes near the surface and close to the outer boundary of the channels become depleted by atoms (or enriched by vacancies). This usually leads to formation of thin intrusions next to the extrusions [313]. Such a model of combined movements of dislocations and point defects well reflects geometrical proportions of extrusions and intrusions [314]. However, there are also many other older models of the surface relief evolution (e.g., [315, 316]).

High stress concentrations around extrusions and, particularly, at the tip of intrusions leads to a formation of short surface cracks which start to propagate along the slip bands or PSBs inside the bulk. This growth is controlled by irreversible emission and absorption of dislocations at the fronts of these cracks. Owing to the localization ratio and Equation 3.13, such a damage process is most probable in the largest surface grains along slip planes with the highest Schmid factors. One should also note that the average level of cyclic plasticity is raising when transferring from a high-cycle regime to a low-cycle one. Consequently, the concentration of nucleated surface cracks follows the same trend.

3.2.2.2 Propagation of Short Cracks

There is no clear notion about the moment of transition from the vacancy-assisted intrusion growth into the dislocation based propagation of the related short crack. Nevertheless, there are some plausible models explaining possible mechanisms of initial growth stages (e.g., [149, 312, 317]). Since the crack nuclei are strictly aligned with the slip planes of PSBs, their further advance proceeds along these planes. This means that the cracks propagate under the shear stress coupled with the tensile normal stress (the mixed-mode I+II) that can somewhat facilitate the emission of dislocations from the crack tip by reducing the ideal shear strength (see Section 1.1). During cyclic loading in air or other corrosive environments, a passive oxide layer always forms on metallic surfaces. If the oxide layer is fractured as a consequence of deformation in the underlying polycrystal, the passivation is lost and a re-oxidation occurs. The process of repeated fracture and re-oxidation is a central principle of slip-oxidation models that are widely used to elucidate the mechanism of propagation of both short and long cracks [149, 317].

The simplest model assuming the propagation along a single slip plane is depicted in Figure 3.15. In the tensile half-cycle, the edge dislocations are emitted from (or absorbed at) the crack tip which generates a new fracture surface on one of the crack flanks. The size of the new surface is equal to the number of dislocations times the Burgers vector and the length of the crack front. When assuming an immediate surface oxidization, the dislocations returning during the unloading cycle cannot remove the new surface. Instead of that, they will form a new fracture surface on the other crack flank, the size of which is, again, equal to the number of returning dislocations times the Burgers vector and the crack front length. Under a constant loading amplitude the number of returning dislocations to the crack tip is nearly equal to the number of dislocations generated during loading. Hence, the crack extension per cycle is equal to the number of dislocations generated at the crack tip times the Burgers vector. In other words it is equal to the cyclic crack tip opening displacement. In this way, the crack advances during each loading cycle.

In general, the crack growth rate $\mathrm{d}a/\mathrm{d}N$ can be assumed to be nearly proportional to the frequency of the oxide-layer fractures which is, again, proportional to the plastic strain γ_p in the slip band. Consequently, the crack growth rate can be assumed to be proportional to $(L/h)^m$, where $m \in (0.3, 0.8)$ is commonly accepted [311, 317]. This also means that the growth rate of short cracks is higher in the long PSBs embedded in large surface grains.

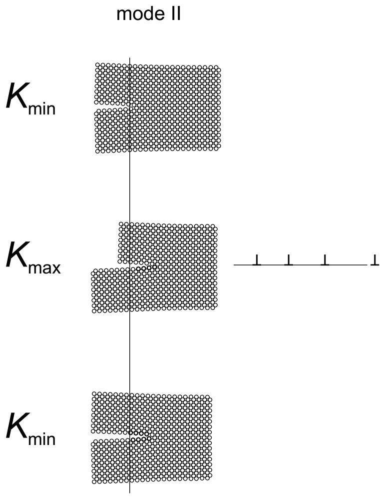

Figure 3.15 The single-slip model of short crack propagation

When the growing short crack approaches a grain boundary, the emission of crack tip dislocations starts to be restricted by the back stress from a creating pile-up. As a result, the growth rate rapidly decreases and, eventually, the crack could be arrested at the grain boundary. In general, the crack growth of these so-called microstructurally short cracks (MSCs) becomes retarded by various microstructural barriers of different strength. The maximal length of MSCs is determined by the distance b_s of the strongest barriers such as grain or phase boundaries. As was sufficiently verified particularly by the group of K. J. Miller in Sheffield [318], the fatigue limit corresponds to a maximal stress still not high enough to propagate the longest short cracks, i.e., to transfer it through the boundary of the surface grain to adjacent bulk grains. The distance b_s also corresponds well to the maximal size of non-damaging cracks in the well known Kitagawa–Takahashi diagram [319]. In specimens fractured close to the fatigue limit, therefore, many small surface cracks arrested at grain or phase boundaries can be found.

Unlike in the case of long cracks, a description of MSC propagation in terms of ΔK does not make too much sense. Indeed, the relative size of the plastic zone r_p/a determined by the zone of emitted dislocations at the tip is not small enough to fulfil the conditions of small-scale yielding and plane strain. Moreover, there is a rapid change in the T-stress during the crack propagation from the surface towards the grain boundary. Therefore, the growth rate of MSCs is often described in the form

$$\mathrm{d}a/\mathrm{d}N = A\Delta\varepsilon_p^l(b_s - a)^k,$$

where $\Delta\varepsilon_p$ is the applied plastic strain range, and A, l and k are material parameters ($k \leq 1$) [247]. This relationship implies a proportionality between the applied and the local plastic strains and reflects well the retardation and the subsequent arrest of MSCs at the barriers ($\mathrm{d}a/\mathrm{d}N = 0$ for $a = b_s$). It should be emphasized that the MSCs can grow at substantially lower applied stresses than the long cracks, i.e., well below the fatigue limit. This is mainly caused by an absence of crack closure effects (see Section 3.2.5 for more details). Indeed, the crack flanks of MSCs are ideally flat so that there is no roughness-induced closure. A highly elongated shear-mode plastic zone lies within the confines of one or two slip planes and no plastic blunting of the crack tip occurs. Therefore, there is also a lack of plasticity-induced crack closure. In the very first growth stages, the closure level might even be negative due to the fact that the Peierls–Nabarro stress prevents individual dislocations from the backward motion within the plastic zone.

For a further advancement of MSCs, the applied stress must be raised to reinitiate the crack in the adjacent grains. Indeed, the superposition of the remote and local stresses can also create sufficiently high local plastic strains in the slip planes with lower Schmid factors. When the crack overcomes the grain boundaries, it starts to be spatially tortuous. This means that the crack flanks become rougher and the growing friction forces decelerate the crack growth under the shear mode II. To avoid friction, the crack attempts to incline to the plane perpendicular to the direction of the applied stress in order to get a more open crack tip by a higher mode I loading component. The cracks occurring in such a transient stage are called physically short cracks (PSCs). As a rule, MSCs start to be PSCs after subsequent passing through one or two grain boundaries. The plastic zone size of a growing PSC steadily increase to embrace more than a single grain. While the crack closure effect increases, the influence of microstructure on the crack growth rate decreases. The PSC completely converts into a long crack (LC) after passing more than ten grain boundaries. At this growth stage, the crack tip plastic zone already embraces several grains and the microstructural influence becomes negligible.

The dependence of the crack growth rate on the range of the stress intensity factor ΔK during all the stages of MSC, PSC and LC is schematically depicted in Figure 3.16. An enormous variation of crack growth rate due to interactions with individual grain boundaries is typical for the MSC stage. The threshold range ΔK_{th} indicates the limit below which the long cracks do not propagate.

3.2.3 Discrete Dislocation Models of Mode I Growth of Long Cracks

3.2.3.1 Near-threshold Crack Tip Plasticity

In the near threshold region of long fatigue cracks, the range of the stress intensity factor (SIF) is very small (units of $\mathrm{MPa\,m^{1/2}}$), and the plastic zone is

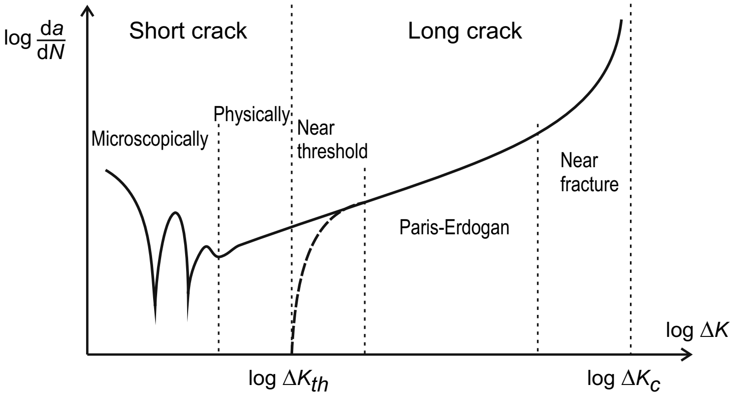

Figure 3.16 The scheme of the crack growth rate of short and long cracks showing the characteristic ranges of stress intensity factors. The near-threshold behaviour of long cracks is indicated by the *bold dashed line*

bounded to a close vicinity of the crack tip. Therefore, the internal dislocation sources are likely not present, and the dislocations can be assumed to be generated at the crack tip. Atomistic models developed by Rice and Thomson [134] and Rice [139] for atomically sharp cracks show that the spontaneous emission of dislocations occurs when the SIF reaches a critical value k_e. This value can be estimated as

$$k_e^2 = \frac{(1+(1-\nu)\tan\Phi)}{(1+\cos\theta)\sin^2\theta}\frac{16G\gamma_{us}}{(1-\nu)}.$$

Here γ_{us} is the unstable stacking fault energy and θ, Φ are angles characterizing the activated slip system. According to [139] and [320], the predicted values of k_e are $1.3\,\mathrm{MPa\,m^{1/2}}$ for Fe, $0.3\,\mathrm{MPa\,m^{1/2}}$ for Al and $0.5\,\mathrm{MPa\,m^{1/2}}$ for Cu. However, the stress fields produced by emitted dislocations shield the crack tip, so that the value of the local SIF k_{loc} becomes significantly lower than that of K_{rem} transmitted from the remote loading. One can write

$$k_{\mathrm{loc}} = K_{\mathrm{rem}} + \sum_i k_{di}, \tag{3.14}$$

where $k_{di} < 0$ is the SIF originated from i-th single dislocation [320]. In the crack tip plasticity models, the movement of an already emitted, individual, dislocation is determined by the Peach–Koehler force

$$\mathrm{d}\mathbf{f} = (\sigma \cdot \mathbf{b}) \times \mathrm{d}\mathbf{l},$$

where $\mathbf{b}$ is the Burgers vector, σ is the stress tensor and $\mathrm{d}\mathbf{f}$ is the force on the line segment $\mathrm{d}\mathbf{l}$. The stress tensor is defined by interaction with other dislocations, free surfaces (e.g., a crack) and remote stresses. The dislocation remains at rest whenever the slip component of the Peach–Koehler force is less

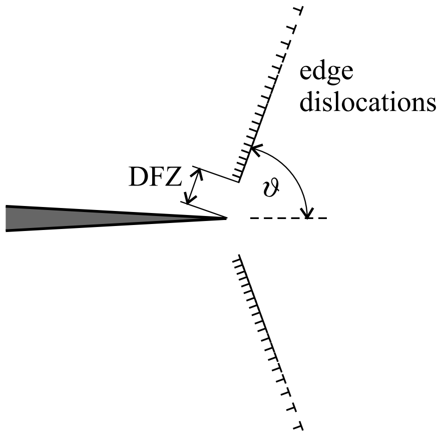

Figure 3.17 Configuration of geometrically necessary dislocations emitted from the crack tip during a loading sequence (DFZ – the dislocation free zone)

than the lattice resistance τ_0 ($\tau_0 \approx G/2000$ for bcc metals). The dislocations may return to the crack, where they disappear by decreasing the crack tip opening displacement (CTOD).

The mode I crack tip plasticity can be numerically simulated as motion of edge dislocations, which are parallel to the crack front (plane strain) [9]. Dislocations are symmetrically emitted in pairs and the angle between the slip planes and the crack propagation plane $\vartheta = 70°$ – see Figure 3.17. This value meets the results of both theoretical and experimental investigations [22]. When K_{rem} is increased from its zero value to $K_{\max}$ and then unloaded back to zero, the model shows the following main features of dislocation behaviour:

1. $K_{\max} < k_e$: no dislocations are generated and the material behaves elastically.
2. $k_e \le K_{\max} < 1.3k_e$: several pairs of positive (shielding) dislocations are generated and pushed into the bulk so that the k_{loc} becomes lower than K_{rem}. Upon unloading the back stress never reaches the negative friction stress $-\tau_0$. Consequently, the material behaves elastically in all further cycles.
3. $1.3k_e \le K_{\max} < 3.5k_e$: many pairs of positive dislocations are generated and, during the unloading sequence, some of them return to the crack again. This means that the crack-tip cyclic plasticity appears first within this range of $K_{\max}$ values. No negative dislocations (with opposite Burgers vectors) are emitted during unloading.
4. $3.5k_e \le K_{\max} < 4k_e$: during the unloading some pairs of negative (anti-shielding) dislocations are emitted but they immediately annihilate with their positive counterparts (their mutual distance is less than $10\mathbf{b}$).

5. $4k_e \leq K_{\max}$: some pairs of negative dislocations, generated during the unloading half-cycle, remain stable.

The fully developed dislocation configurations look like inverse pile-ups (Figure 3.17). A dislocation free zone (DFZ) is found in the immediate vicinity of the crack tip. Such zones were often observed in thin foils (e.g., [321]) and their existence is an exclusive consequence of the discrete nature of plasticity. Similarly, the cyclic plasticity starts to operate only when $K_{\max} \approx 2k_e$ is reached (the point 3). This condition can be interpreted as an intrinsic threshold $K_{in,th}$ of the stationary mode I fatigue crack. The coefficient 2 was chosen here instead of 1.3 (the lower limit), because the emitted dislocations blunt the crack tip. As a consequence, the real critical values for spontaneous dislocation emission ($K_{in,th} \approx 2.6\,\mathrm{MPa\,m}^{1/2}$ for Fe, $0.6\,\mathrm{MPa\,m}^{1/2}$ for Al and $1\,\mathrm{MPa\,m}^{1/2}$ for Cu) are expected to be somewhat higher than those calculated for the atomically sharp cracks. Thus, the experimentally verified existence of the intrinsic threshold can be nicely interpreted by the discrete dislocation model.

3.2.3.2 Crack Growth in Near-threshold Region

In order to study the development of dislocation configurations and related details of the crack front geometry during the loading history, the discrete dislocation model of growing fatigue crack was developed by Riemelmoser *et al.* [320, 322]. It is a blunting model very similar to the original concept of Pelloux [323], but the plastic deformation is modelled as motion of edge dislocations.

Blunting and growing of the modelled crack is schematically shown in Figure 3.18. The crack growth increment per emission of one dislocation pair is $\Delta a = b\cos\theta$, where b is the magnitude of the Burgers vector.

The crack resharpens again when the emitted dislocations return to the crack tip (or crack flanks). As usual, an assumption of full irreversibility of the fracture process is accepted when modelling the fatigue crack growth in the air. It means that, due to the environmental assistance (in particular that of the oxygen), once created free surfaces do not re-weld. Consequently, after the unloading the length of the crack remains the same as it was at the moment of the preceding maximum load. Thus, in principle, only one dislocation that was emitted, and returned back to the crack tip causes an elementary fatigue crack advance. This is the microstructural interpretation of the previously mentioned fact that the range of local cyclic plasticity can be considered to be the fatigue crack driving force.

A result of the numerical analysis of a growing crack is demonstrated in Figure 3.19. In the first loading half-cycle many dislocations are emitted as a slip band. During unloading m of them return to the crack, where they annihilate on the free surface (Figure 3.19(b)). In the second cycle the same m dislocations are generated and annihilated again, while the crack propa-

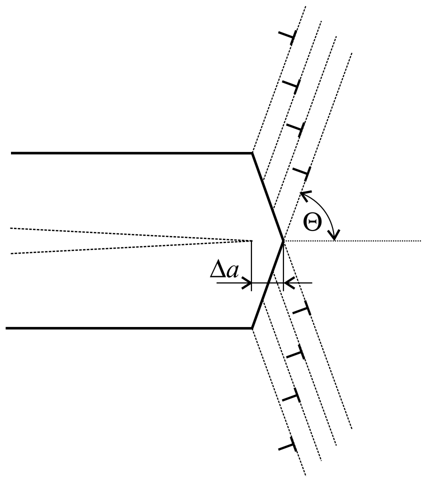

Figure 3.18 Scheme of the crack tip advancement produced by its blunting due to dislocation emission. Δa is the crack advancement caused by emission of four dislocation pairs

gates. The same occurs in many further cycles, until the crack advancement becomes about 200 Burgers vectors. Then one further pair of dislocations remains in the bulk, since the repulsive stress from the slip band, acting at the advanced crack tip, has become somewhat lower (Figure 3.19(c)). This happens every 200 Burgers vector distances after the crack advancement is so that a dislocation wall parallel with the crack is formed, see Figure 3.19(d). When the influence of the first slip band on the moving-away crack tip becomes small enough, the newly generated dislocations can pass the wall to form a second slip band, as also shown in Figure 3.19(d). In the following unloading half-cycle the dislocations in the wall experience the back stress from the dislocations in the secondary band. Many of them return to the crack tip and disappear. This situation is depicted in Figure 3.19(e). From that moment on, the whole process starts to repeat periodically which leads to a dislocation arrangement as shown in Figure 3.19(f). The width of each periodicg999 segment is about 2500 Burgers vectors (0.6 mm). The segments are accompanied by surface steps (adjacent to the slip bands) that can be visible on fracture surfaces as so-called subcritical (near-threshold) striations. Their spacing of about 0.6 mm is, unlike that of supercritical striations (see hereafter), fully independent of the crack growth rate.

Thus, the growing near-threshold fatigue crack leaves a plastically deformed band, the static plastic zone, along its flanks behind the crack tip. In real crystals, this zone consists of both the geometrically necessary dislocations and the statistically stored ones. Although only geometrically necessary dislocations are considered here, the dislocation arrangements observed in the electron microscopy look very similar, e.g., [324]. Such a periodic configura-

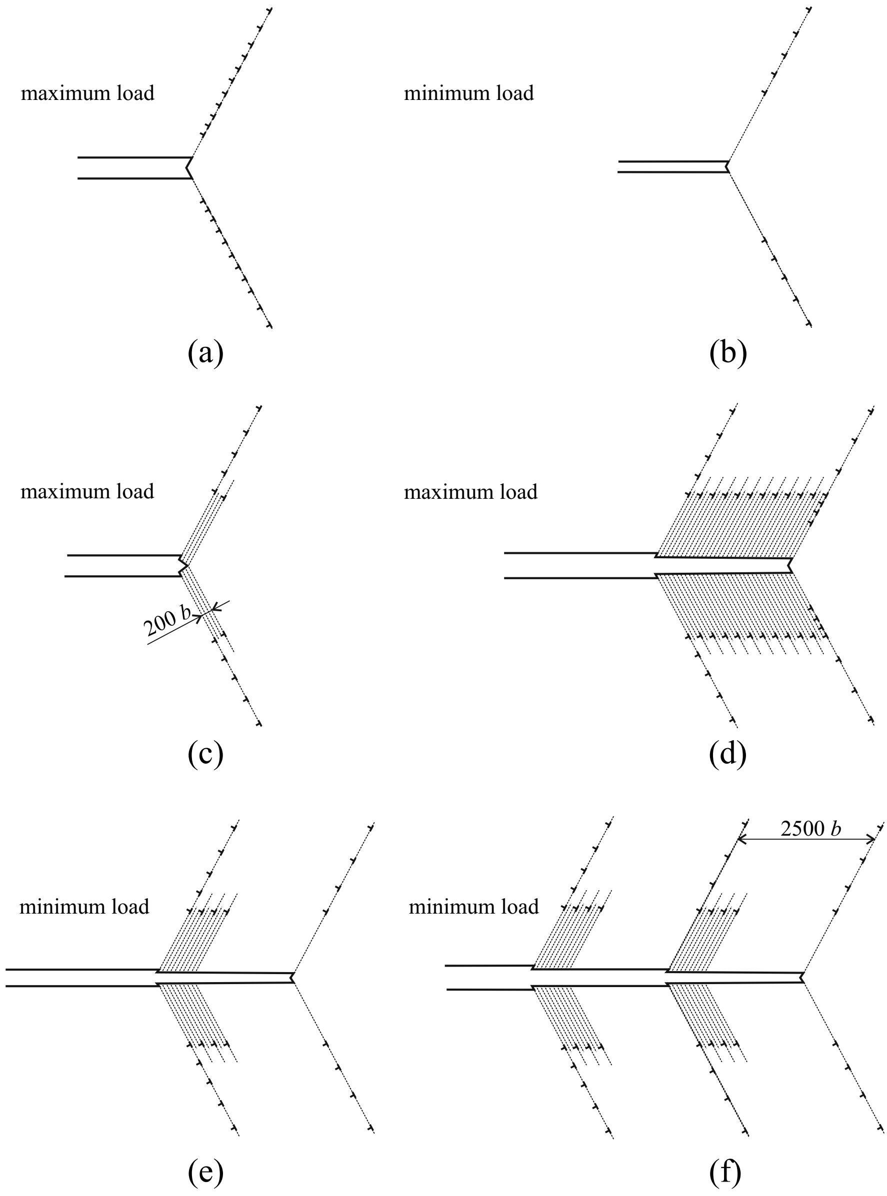

Figure 3.19 Prediction of geometrically necessary dislocation arrangements in the wake of a growing near-threshold crack according to the discrete dislocation model: (a) the first growing stage, (b) after releasing the load some dislocations disappear to free surfaces, (c) generation of the first wall dislocation, (d) generation of the second slip band, (e) vanishing of a part of wall dislocations in the unloading phase, and (f) creation of a periodic dislocation structure in the crack wake

tion of slip bands, remaining in the wake of a propagating crack, causes a tilting of the crystal lattice near crack flanks. This can produce a significant contact shielding of the crack tip (see Section 3.2.4).

The model can also well reproduce the intrinsic crack growth curve [9] near the threshold. On the other hand, the continuum plasticity models (e.g., [274]) are not able to reproduce correctly any of above-mentioned phenomena. The difference in prediction of the near-threshold crack growth behaviour by the continuum approach and the discrete dislocation model is schematically depicted in Figure 3.20.

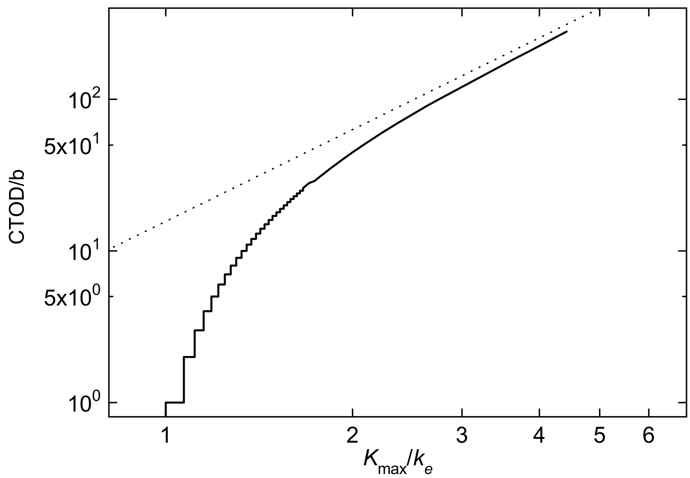

Figure 3.20 The scheme of the near-threshold crack growth rate according to the continuum approach (*dotted line*) and the discrete dislocation model (*solid line*)

3.2.3.3 Crack Growth in Paris–Erdogan Region

When the crack proceeds to reach the Paris–Erdogan region, the SIF at its tip keeps raising steadily. The ability of dislocation emission increases and, consequently, the width of periodic segments decreases. Due to a much higher number of emitted dislocations, simultaneously, the crack advance per cycle becomes much larger. After reaching the Paris–Erdogan region, both these phenomena lead to a vanishing of the periodic crack-wake dislocation structure shown in Figure 3.19(f). Many dislocation slip bands are emitted during each loading cycle, which results in a significant crack blunting. The related qualitative change in the micromechanism of crack propagation is schematically depicted in Figure 3.21. In order to make this demonstration clear, only four slip bands are assumed to be emitted per each cycle.

During the loading phase, large slip bands are alternatively emitted on both sides of the crack tip. This is shown in Figure 3.21(a,b). Owing to their repulsive stress, the subsequently emitted bands appear shorter than the previous ones (Figure 3.21(c,d)). During unloading m dislocations of the second band return to the crack, where they annihilate on the free surface. This leads to resharpening of the near-tip part of crack flanks (Figure 3.21(e)). The same sequence is carried on in further loading cycles while leaving periodic patterns on crack flanks which are known as (supercritical) striations. Their morphology on both fracture surfaces is repeating in each cycle, which is clearly seen in Figure 3.21(f), where the crack flanks are pressed together during a compressive part of the loading cycle. Here the microscopic tortuosity of the crack path is suppressed in order to highlight important periodic patterns. Indeed, such created striations successfully mark locations of the crack front and their spacing corresponds to a local rate of the crack growth (microscopic growth rate). Therefore, they play a significant role in the failure analysis (see Section 3.4). Naturally, the schematically described "Paris" crack also leaves the geometrically necessary dislocations in individual bands on both sides of the crack flanks.

Another simple model for crack advance was proposed by Neumann [325]. It assumes an alternating shear only on two slip planes, primary and secondary, which intersect the crack front as shown in Figure 3.22. The work hardening on primary slip planes leads to alternating shear on secondary planes. In the reverse half-cycles a majority of dislocations returns back to the crack tip and disappears. Some of dislocations remain in the crack wake to create the residual plastic zone and saw-tooth supercritical striations. This model is also applicable to ductile metal single crystals as well as to the near threshold growth in polycrystals. In the latter case, however, the repulsive forces of remaining shear-band dislocations do not allow emission of the shear bands again in many subsequent loading cycles. During these cycles the crack proceeds along a finer zig-zag path (without leaving observable saw-tooth striations) in a way similar to that described in connection with Figures 3.19(c,d). It should be noted that the original Neumann's model did not assume any dislocations left behind the advancing crack tip.

In engineering materials, the crack path exhibits significant tortuosity on many scales owing to different microstructural barriers to dislocation motion. The largest fluctuations are of the order of the grain size. Especially in the near threshold region much smaller deflections can occur which can be induced by dislocation-dislocation interaction or by microstructural features much smaller than the grain size [10]. The smallest deviations are of the order of lattice spacing. They are caused by remaining ledges formed by dislocations generated at the tip or entered on the fracture surface. Therefore, the crack path geometry is usually simply drawn in a zig-zag manner as in Figure 3.22(c). Let us emphasize that the extent of created dislocation bands (the plastic zone size) in relation to the length of largest deflected crack segments constitutes a very important parameter. It determines a symmetry

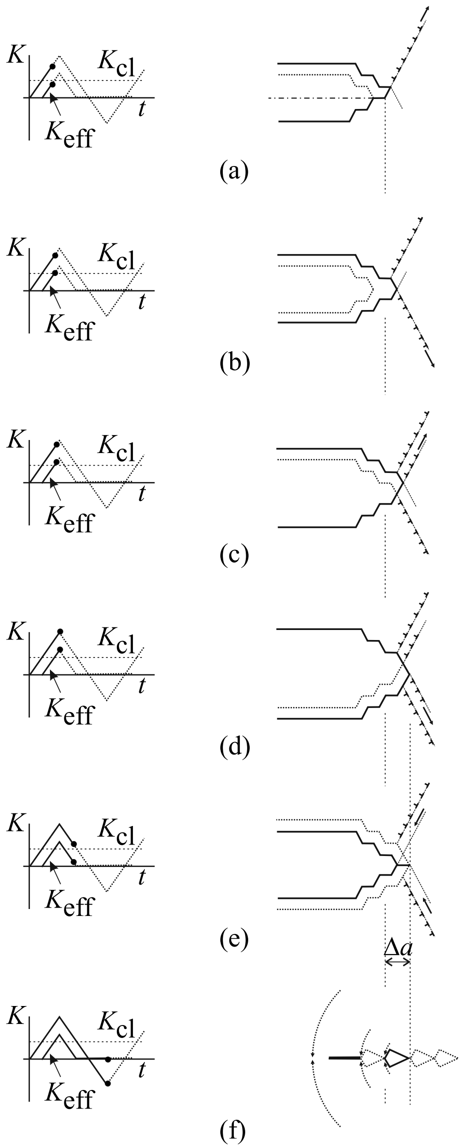

Figure 3.21 A scheme of the crack tip advance by dislocation emission during one loading cycle in the Paris–Erdogan region. The points on the K *vs* t diagram show subsequent loading stages. In the last compressive stage the crack flanks are pressed together which highlights ductile striations

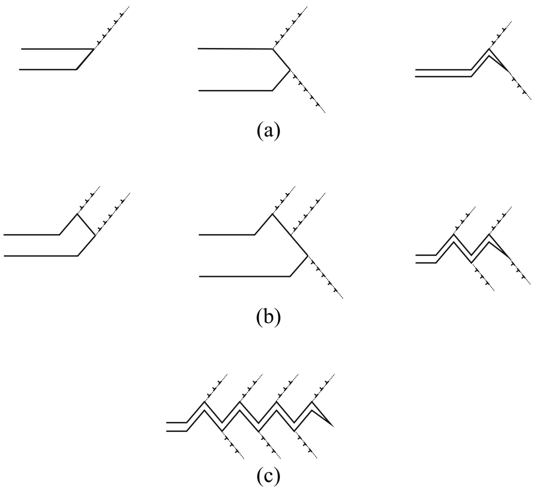

Figure 3.22 Alternating slip model after Neumann: (a) the first loading cycle, (b) the second loading cycle, and (c) the resulting crack path leaving saw-tooth striations on fracture surfaces

of the crack-wake dislocation arrangements and decides about the level of roughness-induced crack closure (see the next subsection).

3.2.4 Crack Closure Mechanisms

The fatigue crack growth behaviour, particularly near the threshold of the applied SIF, is affected by both the chemical composition and the microstructure. Various shielding mechanisms that are used to explain the diversity in the threshold values can be, according to Ritchie [165], categorized as intrinsic and extrinsic (see also Chapter 2). A general scheme of extrinsic mechanisms is depicted in Figure 3.23. In this figure the applied (external) ΔK consists of two basic components: the closure range ΔK_{cl} and the effective range ΔK_{eff}. The range ΔK_{cl} is divided into three parts (from the left to the right): the oxide-induced closure, the roughness-induced closure and the plasticity-induced closure. The range ΔK_{eff} consists of two parts: the geometrically-induced shielding ΔK_{br} (crack kinking and branching) and the intrinsic range ΔK_{in}. While all the closure mechanisms operate in the crack wake, the crack kinking and branching produce shielding in front of the crack tip.

Thus, besides the geometrical shielding that has already been described in the brittle-fracture section, several closure (or contact shielding) mechanisms become important in fatigue.

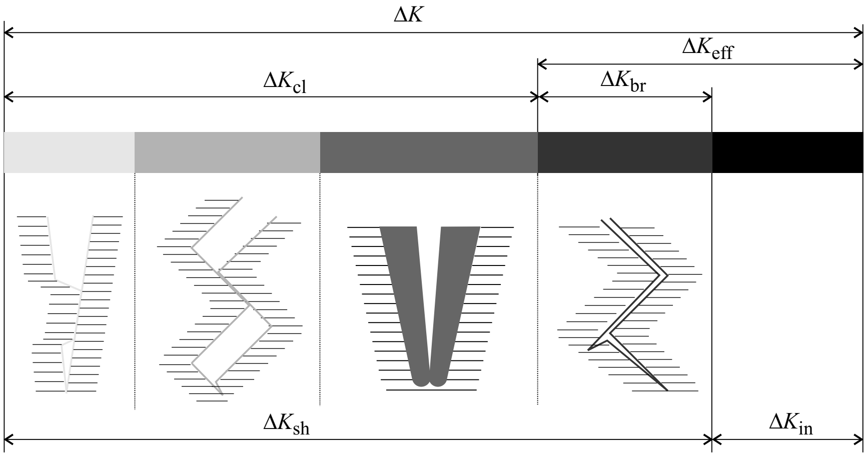

Figure 3.23 A general scheme of extrinsic mechanisms occurring during the fatigue crack propagation; ΔK – the applied SIF range, ΔK_{cl} – the closure SIF range, ΔK_{eff} – the effective SIF range, ΔK_{in} – the intrinsic SIF range

As recognized by Elber [326], the crack remains to be closed not only in the compressive part of the loading cycle but, to some extent, also during its tensile part. This means that only the $\Delta K_{eff} = \Delta K - K_{cl}$ part of the applied ΔK is effective as a crack driving force of the fatigue crack growth – see Figure 3.24. When the minimum applied SIF is higher than K_{cl}, the crack remains steadily open (Figure 3.25). However, this happens only for very high cyclic ratios ($R > 0.6$).

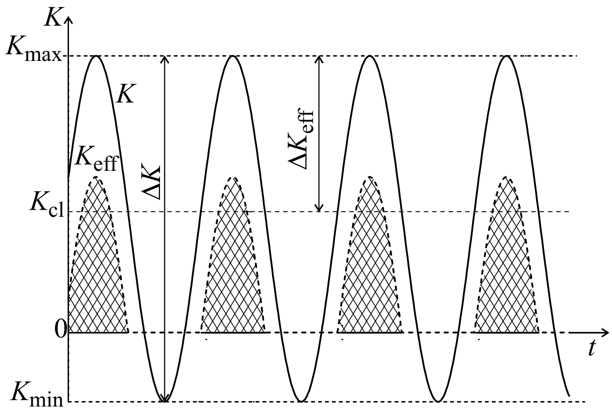

Figure 3.24 Applied and effective stress intensity factors when $K_{\min} < K_{cl}$. K is the applied SIF (*solid line*), K_{cl} is the closure SIF level and K_{eff} is the effective SIF (*dotted line*). The *dashed areas* show ranges, where the crack tip is open

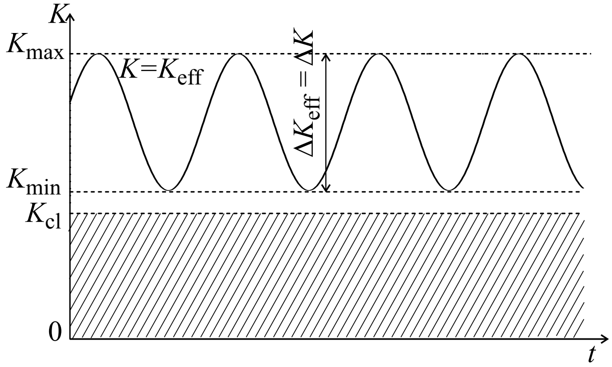

Figure 3.25 Applied and effective stress intensity factors when $K_{\min} > K_{cl}$ ($\Delta K_{eff} = \Delta K$), and the crack remains steadily open during the loading. The *dashed area* shows the K-range of a potential crack closure

In metallic materials fatigued in air or vacuum, both the roughness-induced crack closure (RICC) and the plasticity-induced crack closure (PICC) are the most important closure mechanisms. There can also be other shielding contributions [165] such as, for example, the oxide-induced closure (OICC) and the crack-wake bridging that occur in corrosive environments and composite materials, respectively. It should be emphasized that there is no experimental tool for measuring individual closure components; only the total crack closure level can be measured using the methods based on extension or electrical resistance (e.g., [327]).

When the crack front branches (splits) during the crack propagation, the effective driving force becomes further reduced to ΔK_{in} as marked in Figure 3.23. An additional reduction can be caused by the crack deflection and meandering. However, this fully holds only for brittle materials as high-strength alloys or ceramics (see Chapter 2 for details). Another shielding (or anti-shielding) component can be associated with large particles or microvoids embedded in the matrix. These effects are important only in nodular ductile irons (see Section 3.2.7) or in particle reinforced composites (see Chapter 2). The value of $\Delta K_{in,th}$ represents the intrinsic threshold or the intrinsic resistance to the fatigue crack growth. Because a partial recovery of newly created surfaces usually takes place in engineering materials during the unloading cycle, the real values of $\Delta K_{in,th}$ are somewhat higher than those of 0.5–2.6 $\mathrm{MPa\,m^{1/2}}$ resulting from the discrete dislocation model. However, they can be almost an order of magnitude lower than the measured fatigue thresholds ΔK_{th}, as shown in Section 3.2.6.

Note that there is no experimental method available for measuring the geometrical shielding component and, consequently, the values of the intrinsic threshold $\Delta K_{in,th}$. Nevertheless, they can be assessed theoretically by employing rather standard materials data in the frame of the unified analytical model of extrinsic shielding, as presented hereafter.

3.2.4.1 Plasticity-induced Crack Closure

The well known reason for the PICC in thin solids (the plane stress condition) is the plastic wedge in the wake of the crack front; see Figure 3.26. This wedge is created by tensile necking within the crack-tip plastic zone. Experimental data lead mostly to the conclusion that the ratio of the closure stress level to the maximum stress (the so-called closure ratio) $\sigma_{cl}/\sigma_{\max} \approx 0.6$ is rather independent of the material. There is, however, a moderate dependence on the cyclic ratio R. Theoretical models published in the last 30 years confirm these results [192]. In the case of thin specimens, however, the LEFM concept loses its validity, which constitutes a considerable disadvantage.

In a rather more important case of plane strain (thick solids), the PICC can also be detected, albeit to a much lower extent. Indeed, the closure ratio $K_{cl}/K_{\max} \approx 0.25$. However, the mechanism of that phenomenon remained unclear for rather a long time. Under the plane strain, indeed, the necking associated with a transverse contraction is not allowed by definition. The problem was solved by Pippan *et al.* in 2004 [10]. A static plastic zone of a constant height is always formed by both the geometrically necessary and the statistically stored crack-wake dislocations; see Figure 3.27 (compare also Section 3.2.3). Unlike those statistically stored, the geometrically necessary dislocations have a long-range tilting effect as shown in Figure 3.28. Only two narrow bands of dislocations, forming low-angle tilt boundaries, are considered here since the effect of the real distribution can simply be obtained by integrating over many such dislocation bands. The tilting of the volume elements in the plastic wake is associated with shear in the direction of the crack propagation. This shear causes a transfer of the material to the crack tip – a reason for the PICC under plane strain conditions.

The dislocation bands shield the crack tip from the remote SIF range ΔK, as generally expressed by Equation 3.14. In the case of an unloaded crack, one dislocation band produces the local shielding stress intensity factors k_1^s and k_2^s at the crack tip as

$$\begin{pmatrix} k_1^s \\ k_2^s \end{pmatrix} = Gb\sqrt{r}\frac{1}{D}\begin{pmatrix} c_1 \\ c_2 \end{pmatrix}, \tag{3.15}$$

where r is the distance of a single band of dislocations from the crack flanks, D is the spacing between individual dislocations in the array and c_I and c_{II} are constants of about -1 [10]. In the near tip regime, the displacements at the crack flanks are given by the well known LEFM equations

$$\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \frac{2(1-\nu)}{G}\sqrt{\frac{|x_c|}{2\pi}}\begin{pmatrix} k_2^s \\ k_1^s \end{pmatrix}, \tag{3.16}$$

where x_c is the distance of the crack-wake contact of crack flanks from the crack tip. Indices 1 and 2 correspond to the shear mode II and the opening mode I, respectively. Note that an overlap of the crack flanks (PICC) takes

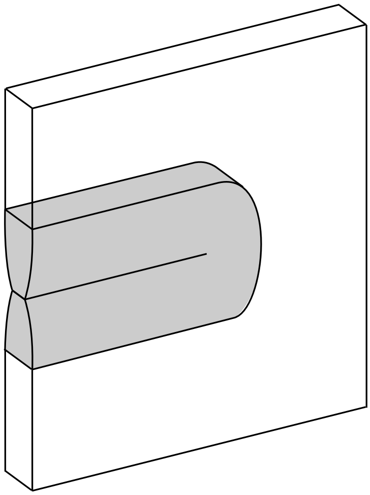

Figure 3.26 The scheme of the plastic wedge in the crack-wake causing the plasticity-induced crack closure in thin specimens (the plane stress condition)

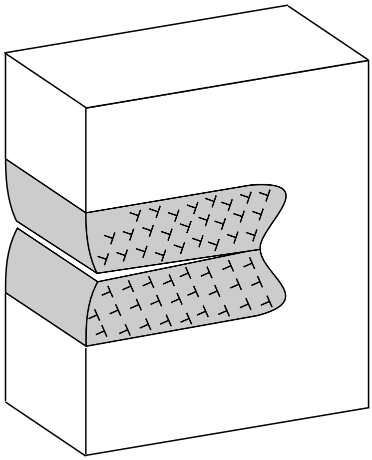

Figure 3.27 The scheme of geometrically necessary crack-wake dislocations causing the plasticity-induced crack closure in thick solids (the plane strain condition)

place since k_1^s and u_2 are negative. Dislocation arrays on both sides of the crack flanks contribute to PICC. Consequently, the closure distance for PICC can be assessed by combining Equations 3.15 and 3.16 to get

$$\delta_{cl,p} \approx \frac{4(1-\nu)b}{D}\sqrt{\frac{|x_c|}{2\pi}}. \tag{3.17}$$

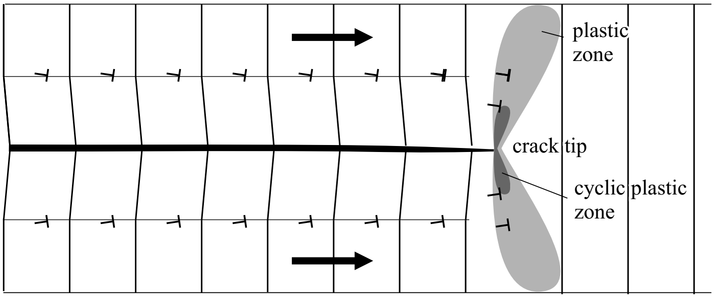

Figure 3.28 Two narrow dislocation bands approximating the real arrangement of geometrically necessary dislocations

The spacing D is inversely proportional to the number n of geometrically necessary dislocations per unit extension of the crack. That number must be proportional to the mean strain at the distance r of the dislocation arrays from the crack flanks. The mean strain is proportional to the yield stress σ_y divided by the Young's modulus $E = 2G(1+\nu)$. That reasoning leads to

$$D \approx 2BG\,(1+\nu)/\sigma_y, \tag{3.18}$$

where $B \approx 10^{-10}\,$m can be assumed to be rather material independent [11].

The maximal applied SIF can be expressed as

$$K_{I\,\max} \approx \sigma_y(2\pi r_p)^{1/2} \tag{3.19}$$

and the maximal opening displacement as

$$\delta_{\max} = 2u_2(|x_c|, K_{I,\max}). \tag{3.20}$$

A half-size of the static plastic zone r_p constitutes a plausible estimate of the distance of dislocation strips from the crack flanks, i.e., $r \approx r_p/2$. By combining Equations 3.17, 3.18, 3.20 and 3.19, the closure ratio reads

$$\frac{\delta_{cl,p}}{\delta_{\max}} = 2C, \tag{3.21}$$

where $C = b/(10\sqrt{\pi}B) \approx 0.1$ is a dimensionless constant [204]. Equation 3.21 is in agreement with the experimentally determined value of the closure ratio of about 0.25, nearly independent of the material.

3.2.4.2 Roughness-induced Crack Closure

The existence of RICC is connected with asperities on fracture surfaces in the wake of the tortuous crack front. Suresh and Ritchie [328] proposed a simple two-dimensional LEFM model describing clearly the essential mechanism of that phenomenon. Irreversible operation of the local mode II at the front of the tortuous crack causes a mutual horizontal shifting of fracture surfaces and their premature contact. Although this geometrical scheme is principally correct, the model does not yield an explicit relation to the microstructure, since the averaged crack path inclination angle used in that analysis does not change with the microstructure coarseness. Thus, in fact, this model is not very useful for a quantitative assessment of the role of the microstructure in the RICC. In order to assess the dependence of the mode II displacement on the mean grain size, Wang and Mueller [258,329] proposed a more recent concept that deals with an irreversible dislocation pile-up adjacent to the crack tip and restricted by a grain boundary. This one-parameter approach requires a complicated analysis of the fracture surface roughness and, moreover, the assumption of a planar slip mode is necessary for its validity.

It should be emphasized that all the previous analytical models could not take the so-called long-range component of the RICC into account because, as mentioned above, that important component was described only recently [10].

Maximal Long-range Component

When the arrangement of crack-wake dislocations becomes asymmetric as in the extreme case in Figure 3.29, the shear displacements of both crack flanks are different. This shear misfit causes the long-range RICC on rough fracture surfaces. Indeed, the asymmetric arrangements of crack wake dislocations can produce the RICC far behind the crack tip, in contrast to shear displacements induced by the irreversible slip (see the next paragraph). Thus, both the asymmetry of crack-wake dislocations and the roughness of fracture surfaces can be assumed to be necessary conditions for the appearance of long-range RICC.

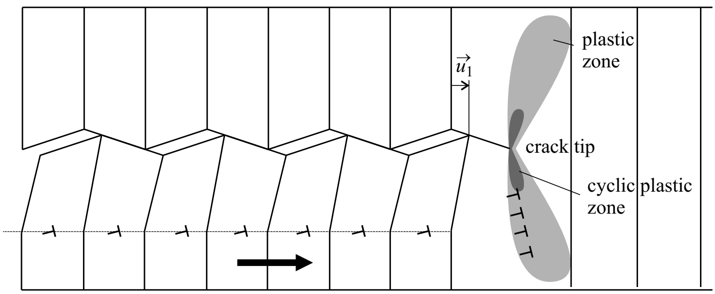

Figure 3.29 The scheme of a shear misfit and the roughness-induced crack closure produced by an asymmetric arrangement of crack-wake dislocations

As follows from theoretical and experimental results [11,168,204], the necessary conditions for RICC are well fulfilled when $S_R > 1$, where $S_R = d/r_p$ is the ratio of the characteristic microstructural distance d (the mean grain size or the interparticle spacing) and the plastic zone size r_p (see also Chapter 2). On the other hand, when the plastic zone embraces several grains or particles, i.e., $S_R \ll 1$, the phenomenon of RICC can be neglected. The reason for these opposite effects is schematically elucidated in Figures 3.30 and 3.31.

In the case of $S_R > 1$, an alternating (zig-zag) single-slip accompanied by dislocation pile-ups represents a strongly prevalent crack propagation mechanism in all metallic materials [192,194]. The static plastic zone is constrained within individual grains containing the crack front. The grain boundaries or secondary phase particles constitute obstacles for crack growth since the atomistic structure of these defects is quite different from that of the matrix, where the dislocations possess a minimum energy and can slide in an optimal manner. Additionally, there is no possibility for their further gliding on the same slip planes in the adjacent grain or inside the secondary particle. Therefore, the crack often grows along grain boundaries and secondary phases or, eventually, even intergranular or interface propagation appears. This obviously leads to a microscopically tortuous crack path associated with a high surface roughness (Figure 3.30). The deviations from the straight growth direction are comparable to the grain size, particle size or their spacing. Such kind of crack advance must also be associated with highly asymmetric crack-wake plasticity varying from grain to grain. The distance of the characteristic dislocation strip from the crack flanks is lower than the grain size. It means that, even when the strips regularly alternate their positions towards the crack flanks, the interaction of dislocation configurations in adjacent grains (reducing the shear asymmetry) can be nearly neglected along the whole length of the crack flanks. This is schematically depicted in Figure 3.30), where only the short segments of the cracks flanks (marked bold) experience the same shear. All the above-mentioned effects result in the maximal level of the RICC.

In the case of $S_R \ll 1$, on the other hand, grain boundaries and secondary phases do not constitute obstacles for the crack growth owing to a large-scale crack-tip plastic deformation embracing many grains or particles (Figure 3.31). The plastic zone is not restricted (constrained) by barriers on either side of the crack flanks and, obviously, there is no reason for the asymmetry of the crack-wake plasticity. The symmetry of the large-scale multislip deformation results in rather straight crack propagation, where crack advance during one cycle (or the CTOD) is higher than the microtortuosity of the crack path (the width and the height of asperities on the fracture surface). The distance of the characteristic dislocation strips from the crack flanks is higher than the average grain size. All those effects result in a negligible level of the RICC.

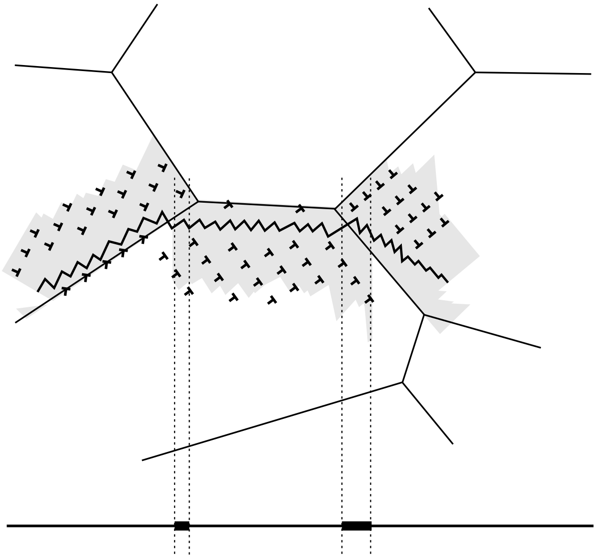

Figure 3.30 The scheme of asymmetric dislocation configurations produced in the case of high size ratios when the plastic zone size (*the grey region*) is smaller than the mean grain size. The highlighted short segments on the projected crack path correspond to nearly symmetric dislocation arrangements

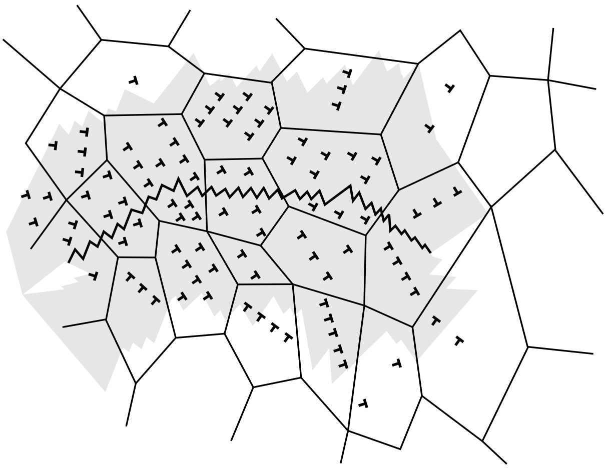

Figure 3.31 The scheme of symmetric dislocation configurations produced in the case of low size ratios (the plastic zone is larger than the mean grain size). The dislocation arrangement is nearly symmetric along the whole crack path

Thus, the statistical concept of the size ratio effect, as already described in Chapter 2, also holds valid for the RICC phenomenon. On the other hand, the size ratio effect is not so significant for the PICC since the dislocation bands on both sides of the crack flanks contribute to that effect. Nevertheless, a small influence might still be expected when $S_R > 1$ since, due to the asymmetry of dislocation strips, the total closure effect might be somewhat less than that described by Equation 3.21. In our further considerations, however, this small inaccuracy will be neglected.

The maximum possible level of RICC can be determined by assuming only one characteristic dislocation band that produces the local mode II SIF k_2^s and the crack-flank shear displacement u_1 according to Equations 3.15 and 3.16, respectively. The simple scheme of RICC in Figure 3.32 relates the closure distance $\delta_{cl,rl}$ to both the shear displacement u_1 and the surface roughness $R_A = 1/\cos\theta$ (in this 2D model $R_L = R_A$) as follows:

$$\delta_{cl,rl} = |u_1| \sqrt{R_A^2 - 1}. \tag{3.22}$$

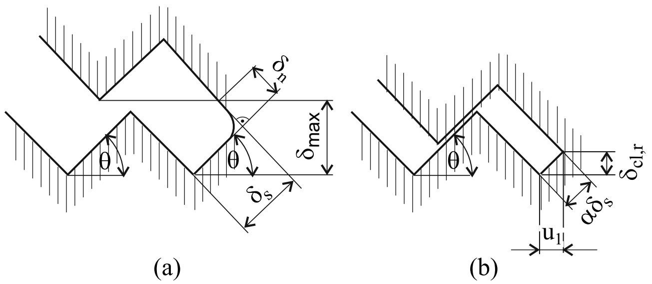

Figure 3.32 The scheme of the roughness-induced crack closure mechanism: (a) the crack tip opening at a maximal tensile load, and (b) the shear shift after unloading

Then a combination of Equations 3.16, 3.18, 3.19, 3.20 and 3.22 yields

$$\frac{\delta_{cl,rl}}{\delta_{\max}} = C\sqrt{(R_A^2 - 1)}. \tag{3.23}$$

This is the final expression for a maximum possible long-range closure ratio in metallic materials.

Maximal Short-range Component

Besides the long-range component of RICC, a short-range component has to be considered as well. As already mentioned, this component is associated with the local mode II induced by the zig-zag crack path, and it is created by an irreversible slip at the crack tip – see Figure 3.32. It was the mechanism that was solely considered in the previous RICC models [168,328,329]. Unlike the long-range component, the overlap of crack flanks caused by the short-range mechanism quickly decays at larger distances from the crack tip. It completely disappears at distances comparable to the mean grain size.

During the loading part of the cycle, the crack tip is opened in a local mixed-mode reaching the maximum CTOD denoted as $\delta_{\max}$ at the moment of the peak stress (Figure 3.32(a)). This displacement is considered to be composed of both the reversible normal component δ_n and the irreversible shear component δ_s defining the local mixed-mode 1+2. Since the shear displacement created in the loading phase is not totally compensated during the unloading phase, the crack flanks remain shifted (Figure 3.32(b)). The main reason for this irreversibility, namely the interaction of dislocations of primary and secondary slip systems, has already been described in Section 3.2.1. Thus, the maximum value of the ratio of backward/forward shear displacements (the irreversibility level) is about 0.5. Obviously, this value is associated with sufficiently long pile-ups (a high probability of activating the secondary slip), i.e., with the size ratio $S_R \geq 1$. On the other hand, $S_R \ll 1$ means no residual shift of crack flanks, i.e., zero irreversibility level. Indeed, the related multislip mechanism cannot produce the residual shear shift, which is apparent from Figure 3.32. Moreover, the crack advance per cycle is comparable to (or higher than) the width of the asperities. Thus, by following the scheme in Figure 3.32(b), and assuming the irreversibility parameter $\alpha \approx 0.5\eta$, one can easily derive the expression at Equation 3.24 for the maximal short-range closure ratio:

$$\frac{\delta_{cl,rs}}{\delta_{\max}} = \frac{\eta \delta_s \sin\vartheta}{2(\delta_n \cos\vartheta + \delta_s \sin\vartheta)}. \tag{3.24}$$

In the framework of the saw-tooth (zig-zag) approximation of the crack flanks [330] it holds that

$$\frac{\delta_n}{\delta_s} = \frac{k_1}{k_2} = \sqrt{\frac{2}{3}} \cot\frac{\vartheta}{2}, \tag{3.25}$$

where k_1 and k_2 are the local SIFs induced by the remote K_I at the deflected crack tip. Note that the linear relationship $\delta \propto k$ is used here (as in Equation 3.16), instead of a quadratic one in the previous models. This is because the real contact point, determined by a sum of the closure effects, lies in the crack-wake. Because $R_A = 1/\cos\theta$, the final relation for the maximal short-range RICC ratio can be obtained by combining Equations 3.24 and 3.25 as follows:

$$\frac{\delta_{cl,rs}}{\delta_{\max}} = \frac{3\eta\,(R_A - 1)}{2\sqrt{6} + 6\,(R_A - 1)}. \tag{3.26}$$

The real short-range closure effect can, obviously, be composed of both the mode II and the mode III contributions (twisted crack segments). The corresponding shear displacements seem to influence the crack closure more or less independently. Hence, the RICC is determined predominantly by those sides of the fracture surface that exhibit the highest local values of either mode II or mode III. These values are associated with the highest and the steepest asperities and, therefore, the standard deviation D_R of the surface roughness seems to be also an important roughness parameter [329]. Although the quantities D_R and R_A are usually well correlated, the level of RICC according to Equation 3.26 can be underestimated in the case of very irregular fracture profiles of extremely high D_R (highly anisotropic and heterogeneous materials).

It should be noted that the roughness-induced crack closure is caused not only by the lateral displacement of two geometrically identical fracture surfaces. Due to local plasticity the shapes of the two rough surfaces are not identical, which leads to a premature contact at the top of the asperities. Such a mechanism, however, can dominate the RICC only when the plastic zone size is very small in comparison with the deflection length (see [10] for more details).

Let us finally mention that stress levels, corresponding to the moments of crack closure and crack opening, are practically equal. Most probably, however, there is a subtle difference related to the fact that the crack propagation is slightly postponed beyond the opening stress level (see [327] for details). However, this inconsistency practically does not affect the effective range ΔK_{eff}.

3.2.5 Unified Model of Crack-tip Shielding

The total crack closure effect represents a sum of the PICC and the RICC components ($\delta_{cl} = \delta_{cl,p} + \delta_{cl,rl} + \delta_{cl,rs}$). Because of the statistics of the size ratio in engineering materials, the parameter η determines the weight of the RICC components similarly to that of the geometrical shielding term in the brittle fracture case (see Chapter 2). Regarding that fact and Equations 3.21, 3.23 and 3.26, the total contact shielding ratio reads

$$\frac{\delta_{cl}}{\delta_{\max}} = \frac{K_{cl}}{K_{\max}} = C\eta\sqrt{R_S^2 - 1} + \frac{3\eta\,(R_S - 1)}{2\sqrt{6} + 6\,(R_S - 1)} + 2C. \tag{3.27}$$

Equation 3.27 reflects many interesting phenomena that were observed in connection with the fatigue crack behaviour:

1. Both terms of the RICC must approach zero for microscopically straight (planar) cracks, i.e., the microstructurally short fatigue cracks in polycrystals or shear cracks in single crystals. This obviously corresponds to $R_S = 0$ in Equation 3.27.

2. The total RICC level must also approach zero in the near fracture region of fatigue crack growth. In that region, clearly, the mean size ratio $S_{Rm} \to 0$ and $\eta \to 0$. Thus, the crack closure ratio is independent of the material and fully determined by the PICC.
3. The maximal level of RICC is detected in the near-threshold region, where, obviously, $S_{Rm} \to 1$ and $\eta \to 1$ is to be expected.

Equation 3.27 also reflects all the transient crack growth phenomena observed after overloads or during the random loading. For example, a sudden increase in $\delta_{\max}$ following an overload cycle increases the total level of the closure displacement δ_{cl}, and reduces the effective crack driving force ΔK_{eff} in the next cycles. The crack growth rate decelerates and remains lower as long as the crack front propagates through the plastic zone created by the overload cycle. After passing that zone, the crack growth rate approaches the previous "steady-state" value. Note that this transient effect is mainly due to the sudden change in the PICC because the overload causes a simultaneous increase in the size ratio (the parameter η). This keeps the level of the RICC close to that before the overload.

The effective crack driving force ΔK_{eff} can be expressed by means of Equation 3.27 and relations $K_{\mathit{eff}} = K_{\max} - K_{cl}$, $K_{\max} = \Delta K/(1-R)$. After some rearrangements one obtains

$$\begin{aligned}\Delta K_{\mathit{eff}} &= \left(1 - C\eta\sqrt{R_S^2 - 1} - \frac{3\eta\,(R_S - 1)}{2\sqrt{6} + 6\,(R_S - 1)} - 2C\right)\frac{\Delta K}{1-R} = \\ &= \Omega_1 \frac{\Delta K}{1-R}.\end{aligned} \tag{3.28}$$

The main advantage of the relation at Equation 3.28 is a possibility of its direct comparison with experimental ΔK_{eff} data, as shown in Section 3.2.6. Although a very moderate decrease in C with decreasing R is to be expected due to a slight drop in the crack-wake dislocation density, the constant C must lie within the range of 0.2 – 0.3. Therefore, in fact, the parameter $S_{Rc} \in (0.1, 1)$ is the only one real fitting parameter in this analysis. Anyway, both dimensionless parameters C and S_{Rc} possess a clear physical meaning and their values must lie within narrow ranges for all metallic materials. The relation at Equation 3.28 predicts that, for a particular material and a fixed applied ΔK, the effective driving force decreases with:

1. decreasing cyclic ratio;
2. increasing surface roughness;
3. decreasing plastic zone size (decreasing η).

This is in agreement with experimental observations [192]. Equation 3.28 also predicts that, in nanomaterials, the RICC level must be negligible in the whole fatigue crack growth region (not only near fracture). Indeed, the value of the intrinsic (effective) fatigue threshold range is lower than $3\,\mathrm{MPa\,m}^{1/2}$

(see Section 3.2.3). Therefore, assuming $K_{\max,th} \approx 5\,\mathrm{MPa\,m^{1/2}}$ as a minimum value for $R < 0.5$, and considering $\sigma_y \approx 1000\,\mathrm{MPa}$, one obtains $r_p \approx 2500\,\mathrm{nm}$ in the near-threshold region. For nanomaterials with $d_m < 100\,\mathrm{nm}$ it means that $S_{Rm} < 0.04 \ll S_{Rc} \approx 0.5 \rightarrow \eta \rightarrow 0$ and $(\delta_{cl}/\delta_{\max})_{RICC} \rightarrow 0$. In nanomaterials, consequently, a negligible level of the RICC ratio applies even to the near-threshold region. Recent experimental results [331] confirm this conclusion.

The values of ΔK_{eff} (or the effective resistance to the crack growth) can be measured by indicating the contact of crack flanks using extensometrical or resistometrical methods. Some authors have used a much simpler procedure based on the fact that, for cycle asymmetries $R > 0.7$, the crack remains steadily open. However, this method was recently challenged by Kondo [332], who indicated a decrease in ΔK_{eff} with increasing $K_{\max}$ for $R > 0.7$ in high-strength steels. This is, most probably, caused by an increasing environmental effect. Related extreme crack tip opening displacements (and associated lattice dilatations) enable a massive penetration of hydrogen into the plastic zone, which results in a reduction of the intrinsic matrix resistance.

When the crack tip opening displacement (CTOD) is much less than the deflections (kinks) or branches (splits) of the crack front, the geometrical shielding can take place [168,192,328]. These geometrical irregularities induce a reduction of the crack driving force (both the k_I and the k_{eff}) as has already been discussed in Chapter 2 in detail. The contact- and geometrical-shielding components can be unified on the basis of the size ratio statistics. With regard to Equations 2.23 and 3.28, the full effect of the geometrical shielding can be included in the model as follows:

$$\begin{aligned}\Delta K_{in} &= \left(1 - C\eta\sqrt{R_S^2 - 1} - \frac{3\eta\,(R_S - 1)}{2\sqrt{6} + 6\,(R_S - 1)} - 2C\right) \times \\ &\quad \times \left[\left(1 - \eta + \eta\sqrt{\frac{\bar{g}_{eff,r}}{R_S}}\right)(1 - A_B) + 0.5A_B\right]\frac{\Delta K}{1 - R} = \\ &= \Omega_1\Omega_2\frac{\Delta K}{1 - R}.\end{aligned} \tag{3.29}$$

Equation 3.29 for calculation of the intrinsic crack driving force ΔK_{in} represents the most general form of the unified model. This relation holds in both the near-threshold and the Paris–Erdogan regions, but its application is limited when special shielding and anti-shielding effects are induced by large secondary phase particles (see Section 3.2.7). In cases of very high cycle asymmetry ($R > 0.6$), the crack closure level lies below the minimum SIF value ($K_{cl} < K_{\min}$) – see Figure 3.25. Under such circumstances no closure occurs and, obviously, Equation 3.29 remains no longer valid.

A typical diagram of relative contributions of individual mechanisms to crack tip shielding ($\Delta\bar{K}_{br} = \Delta K_{br}/\Delta K, \Delta\bar{K}_{in} = \Delta K_{in}/\Delta K, \Delta\bar{K}_{cl} = \Delta K_{cl}/\Delta K$) in dependence on both the applied ΔK (constant d_m) and the mean grain size d_m (constant ΔK) according to Equation 3.29 is schematically depicted in Figures 3.33 and 3.34 for a material of particular chemical composition. Note that the decreasing (increasing) influence of the RICC mechanism with increasing ΔK (d_m) almost completely determines the related dependencies of global shielding or intrinsic driving force.

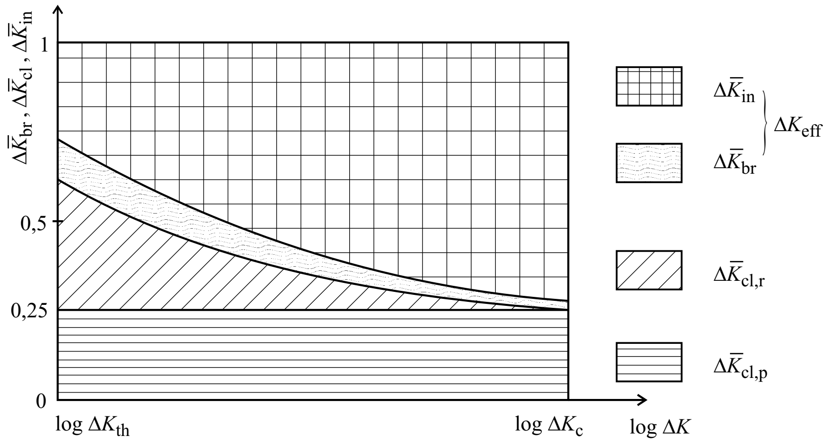

Figure 3.33 Diagrams of relative participation of individual mechanisms in the crack tip shielding as functions of the applied ΔK (constant d_m)

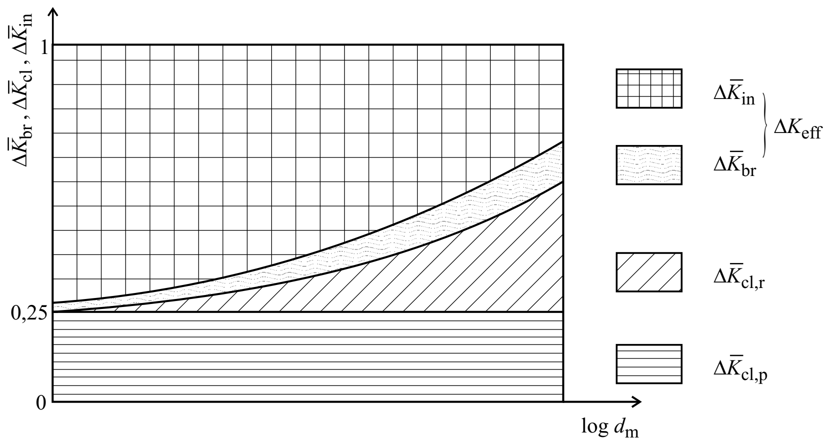

Figure 3.34 Diagrams of relative participation of individual mechanisms in the crack tip shielding as functions of the mean grain size d_m (constant ΔK)

By applications of Equation 3.29 to metallic materials, however, one has to pay attention to the following problem. Unlike in brittle fracture, the fatigue crack growth mechanism in metallic materials is closely related to a cyclic movement of dislocations that is controlled by the shear stress. As shown by Pippan [333], the deflections like kinks or double-kinks do not cause any reduction in the maximal shear stress at the crack tip. Thus, that kind of shielding can be neglected in a majority of metallic materials, perhaps except for those exhibiting a quasi-brittle fracture mechanism. On the other hand, the crack branching causes both the reduction of the local shear stress and the increase in the resistance to the crack growth (the doubled fracture energy) [168,169]. Although the branching process is not typical of the fatigue cracks in metallic materials, its occurrence must always be explicitly reflected when applying the unified model to a determination of both the shielding components and the intrinsic thresholds in metallic materials.

For a material of a particular chemical and phase composition, the values of the intrinsic thresholds should be practically independent of either the microstructure coarseness or the cyclic ratio. A slight dependence on the former might result from a difference in the resistance of grain or phase boundaries to fatigue crack penetration. As already mentioned, a small change in the crack-wake dislocation density with variation in R is also expected. These effects are, most probably, the main reasons why there is some scatter in the calculated values of $\Delta K_{\mathit{eff},th}$ or $\Delta K_{in,th}$ (see the following subsection).

3.2.6 Applications of the Unified Model

Equations 3.28 and 3.29 of the unified model can be simply applied to the assessment of the intrinsic thresholds and the shielding components in engineering metallic materials. Indeed, only a few rather standard material characteristics and loading parameters are necessary: σ_y, d_m, R_S, ΔK and R. A question important for practical measurements of the surface roughness R_S arises in connection with a semi-fractal character of the fracture surface morphology (e.g., [334]). As already mentioned in Section 3.2.3, almost all cracks in metallic materials can be assumed to be tortuous within a considerably wide range of scaling. Consequently, one should be able to decide about the fracture surface roughness R_S of a particular material under specific loading conditions. Fortunately, the size ratio effect gives a rather clear answer to that question: the asperities smaller than about one half of the static plastic zone size do not contribute to the RICC. This statement is justified by the fact that the size of asperities is closely related to the characteristic microstructural parameter. Thus, for example, $K_{\max,th} \approx 5\,\mathrm{MPa\,m}^{1/2}$ is generally appropriate for the near-threshold region. When considering the yield stress $\sigma_y \approx 500\,\mathrm{MPa}$, the facet sizes of the order of units of μm might still be significant. In the Paris–Erdogan region ($K_{\max} \approx 20 - 30\,\mathrm{MPa\,m}^{1/2}$), however, the finest plausible scale constitutes only tens of microns.

The unified concept of crack tip shielding was applied in several studies of the intrinsic thresholds in aluminium alloys, α-titanium, titanium alloys and steels [168, 173, 335]. The geometrical shielding effects were neglected for almost all investigated metals ($\Delta K_{in,th} \approx \Delta K_{\mathit{eff},th}$), which means that, in most cases, Equation 3.28 was used to determine the intrinsic thresholds. An exception to the rule was the duplex steel, where a pronounced interphase cracking and crack branching took place. The analyses are briefly reported hereafter. Although some results were slightly corrected according to the current state of the theoretical concept, the changes do not have any impact on conclusions stated in the works published earlier.

3.2.6.1 Aluminium Alloys

The crack closure contribution in the near-threshold region has been experimentally studied using underaged and overaged compact tension samples of a 7475 aluminium alloy in air and vacuum [336]. Different thermomechanical treatments have produced microstructures with grain sizes of 18 µm and 80 µm. The experimental and theoretical data are displayed in Table 3.1.

Table 3.1 Experimental and calculated fatigue threshold data for 7475 aluminium alloy in air and vacuum (Λ is the percentage of extrinsic shielding)

Envir.	ΔK_{th}^{exp} [$\mathrm{MPa\,m}^{1/2}$]	d_m [µm]	R_S	σ_y [MPa]	η	$\Delta K_{\mathit{eff},th}$ [$\mathrm{MPa\,m}^{1/2}$]	Λ [%]
Air	2.60	18	1.30	505	1.00	1.2	54
	1.70	18	1.21	455	1.00	0.89	48
	2.70	80	1.90	451	1.00	1.14	58
	2.20	80	1.25	445	1.00	1.08	51
Vacuum	4.00	18	1.30	505	0.97	1.87	53
	2.90	18	1.21	455	0.99	1.52	48
	8.80	80	1.90	451	0.94	1.85	79
	4.10	80	1.25	445	1.00	2.02	51

The values of $\Delta K_{\mathit{eff},th}$ calculated according to Equation 3.28 for the tests in air are practically identical and independent of the grain size. They lie only slightly below the measured data of about $1.4\,\mathrm{MPa\,m}^{1/2}$ [336] and are in a good agreement with the range of $0.75 - 1.0\,\mathrm{MPa\,m}^{1/2}$ reported by Pippan [337] for aluminium alloys. The calculated values $\Delta K_{\mathit{eff},th}$ for the vacuum tests are also in good agreement with the experiment. They are distinctly higher than those in air, which can be reasonably elucidated by more extended re-welding of newly created surfaces in the vacuum.

3.2.6.2 Titanium Alloys

Experimental near-threshold crack growth data for various α-titanium grades [338] are collected in Table 3.2. The large scatter of measured values of ΔK_{th} is mainly a consequence of a wide distribution of grain sizes. The average measured value $\Delta K_{eff,th} = 2.1\,\mathrm{MPa\,m}^{1/2}$ [339] lies within the computed range of intrinsic thresholds $\Delta K_{eff,th} \in (2.0, 3.8)\,\mathrm{MPa\,m}^{1/2}$.

Table 3.2 Experimental and calculated fatigue threshold data for α titanium grades

ΔK_{th}^{exp} [$\mathrm{MPa\,m}^{1/2}$]	R	d_m [µm]	σ_y [MPa]	η	$\Delta K_{eff,th}$ [$\mathrm{MPa\,m}^{1/2}$]	Λ [%]
6.00	0.07	40	430	0.95	2.03	66
5.00	0.35	40	430	0.89	2.56	49
5.30	0.07	35	260	0.68	2.25	58
4.30	0.35	35	260	0.48	3.02	30
7.00	0.07	230	220	0.96	2.35	66
5.80	0.35	230	220	0.91	2.91	50
6.00	0.07	20	630	0.96	2.02	66
4.30	0.35	20	630	0.95	2.08	52
10.00	0.07	210	580	1	3.23	68
8.00	0.35	210	580	0.99	3.71	54

Experimental and theoretical threshold data for two grades of the Ti-2.5%Cu alloy are displayed in Table 3.3. The microstructures consisted of coarse lamellar colonies ($d_m = 580\,\mu\mathrm{m}$) and a fine Widmanstätten microstructure ($d_m = 10\mu\mathrm{m}$), respectively [329]. The calculated intrinsic threshold values of $3.3\,\mathrm{MPa\,m}^{1/2}$ are very close to the average measured value $\Delta K_{eff,th} = 3.2\,\mathrm{MPa\,m}^{1/2}$ for Ti-6Al-4V alloys [340].

Table 3.3 Experimental and calculated fatigue threshold data for Ti-2.5%Cu alloy

ΔK_{th}^{exp} [$\mathrm{MPa\,m}^{1/2}$]	R	d_m [µm]	σ_y [MPa]	η	$\Delta K_{eff,th}$ [$\mathrm{MPa\,m}^{1/2}$]	Λ [%]
9.00	0.1	580	420	1.00	3.32	63
7.00	0.1	10	499	0.65	3.15	55

3.2.6.3 ARMCO Iron

Experimental and calculated data for different grades of the ARMCO iron [341, 342] are summarized in Table 3.4. Besides an enormous variance in the grain size there is also a wide range of R ratios. The measured effective thresholds of 2.75 MPa m$^{1/2}$ do not depend on the mean grain size. The theoretical data are presented only for loading cases, where a non-zero closure was predicted. The calculated $\Delta K_{eff,th}$ values in Figure 3.35 lie close to the experimental dotted and dashed line.

The dominant contribution of RICC explains well the strong dependence $\Delta K_{th}(d_m, R)$ observed in the experiment. Note that $\Delta K_{th} \approx \Delta K_{in,th} = \Delta K_{eff,th}$ for $R > 0.55$, which means no closure during cycling.

Table 3.4 Experimental and calculated fatigue threshold data for ARMCO iron

ΔK_{th}^{exp} [MPa m$^{1/2}$]	R	d_m [µm]	σ_y [MPa]	η	$\Delta K_{eff,th}$ [MPa m$^{1/2}$]	Λ [%]
4.5	0.1	2	530	0.19	2.78	38
2.9	0.7	2	530			
5.3	0.1	20	240	0.5	2.49	53
2.9	0.7	20	240			
6.8	0.1	90	150	0.55	3.03	55
3.65	0.55	90	150	0.44	3.64	0
2.8	0.7	90	150			
2.9	0.8	90	150			
8.7	0.1	410	108	0.76	2.96	66
3.2	0.7	410	108			
10.3	0.1	3550	96	0.99	2.37	77
3.6	0.7	3550	96	0.99	2.49	31

3.2.6.4 Duplex Steel

The fatigue threshold values of duplex ferritic-martensitic steels can reach 20 MPa m$^{1/2}$ [343], which is much higher than the threshold range of single phase alloys [344]. The high thresholds of multi phase alloys are typically attributed to extreme crack deflection and branching of the crack front [343, 345].

Hot rolled bars made of the austenitic-ferritic duplex stainless steel SAF 2507 (equivalent to UNS S32750) with a diameter of 80 mm (AD) were also used in our experiments [171]. The tests were performed at three different temperatures 150°C, −50°C and at room temperature ($R = 0.1$ and the frequency $f = 10$ Hz). The microstructure was evaluated using the optical microscope Olympus PMG3 after etching and the values of R_S were measured by the profilometer MicroProf FRT. The path of propagating cracks was found to be very tortuous and branched. The propagation directions change in the austenite or ferrite phase and, in particular, at the austenite/ferrite interface. Crack branching could be identified by secondary cracks

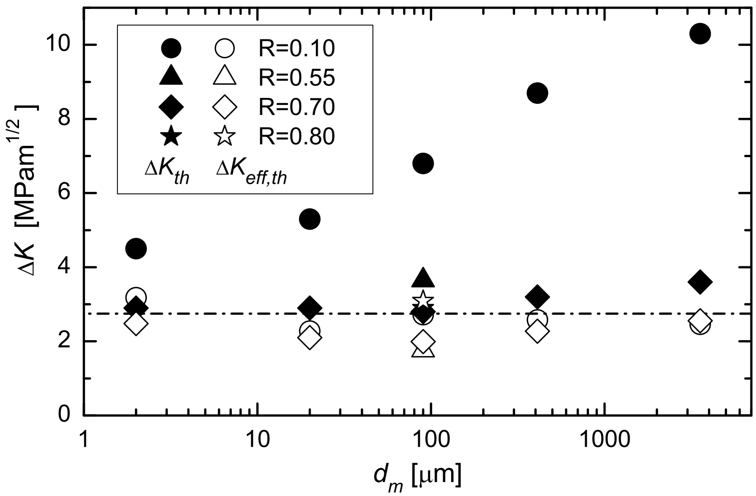

Figure 3.35 The dependence of both measured thresholds ΔK_{th} and calculated effective thresholds $\Delta K_{eff,th}$ on the mean grain size of the ARMCO iron for various cyclic ratios. Reprinted with permission from ASTM International. (see page 265)

in all specimens (see Figure 3.36). The values of the area fraction A_B were assessed by a careful counting of both the number and the size of secondary cracks. Because of a quasi-brittle character of the crack path, the shielding effect induced by crack deflections was taken into account as well.

Measured values of R_S and A_B for all investigated temperatures are listed in Table 3.5.

Table 3.5 Experimental fatigue threshold data for the duplex steel

t [°C]	ΔK_{th}^{exp} [MPa m$^{1/2}$]	A_B	R_S	σ_y [MPa]	$\Delta K_{eff,th}^{exp}$ [MPa m$^{1/2}$]
150	10.83	10	1.55	484	4.92
20	8.64	20	1.43	625	3.90
−50	14.85	30	1.56	738	5.84

Calculated values of both the statistical factor η and $\Delta K_{eff,th}$ are displayed in Table 3.6. There is good agreement between the calculated and the measured effective threshold values (compare with Table 3.5).

The effective threshold values for −50°C and 150°C are higher than the effective value of 2.75 MPa m$^{1/2}$ measured and calculated for the ferritic single-phase steel. This disagreement can be almost eliminated by calculating $\Delta K_{in,th}$ using Equation 3.29 (see Table 3.6). The remaining slight deviations can be attributed to: (1) a different intrinsic resistance of single-phase and

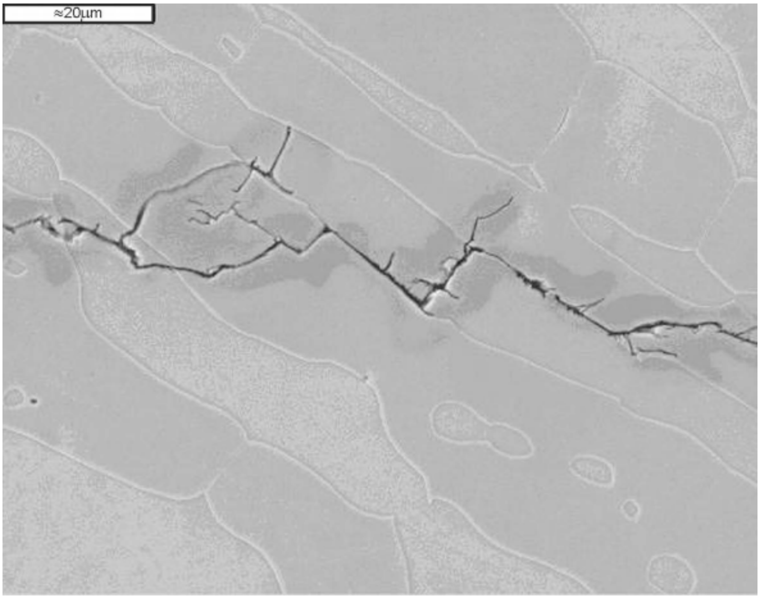

Figure 3.36 Tortuous and branched geometry of fatigue cracks in the duplex steel

Table 3.6 Calculated fatigue threshold data for the duplex steel

t [°C]	η	$\Delta K_{eff,th}$ [MPa m$^{1/2}$]	$\Delta K_{in,th}$ [MPa m$^{1/2}$]	Λ [%]
150	0.58	4.6	3.6	67
20	0.92	2.9	1.9	78
−50	0.72	5.3	3.6	76

duplex matrices; (2) imperfections in experimental methods; (3) an approximate character of the theoretical concept and (4) the temperature influence. The last column in Table 3.6 shows a contribution of the total extrinsic component (closure + shielding) of almost 80%, which is much higher than that in all other investigated materials, including austenite steel [346]. This confirms the decisive role of the extrinsic toughening induced by the duplex microstructure.

The difference in intrinsic thresholds of materials with various chemical composition is, in principle, determined by differences in their values of the critical SIF for dislocation emission (see Section 3.2.3). However, the environmental effects can also play an important role because, during experimental procedures, elementary crack advances cannot be avoided.

3.2.7 Influence of Shielding on Crack Growth Rate

When a long crack propagates under the constant applied stress range, both the plastic zone size and the crack opening displacement linearly increase with an increasing crack length. Consequently, the size ratio S_R decreases and the relative influence of surface asperities becomes smaller. Particularly in the near-fracture region the crack advancement per cycle usually becomes higher than the size of crack-wake asperities. As a consequence, the supercritical fatigue striations, created cycle-by-cycle by partially reversible large-scale blunting beyond the near-threshold region, usually remain as the most striking morphological patterns on fracture surfaces (see Section 3.4 for more details). It obviously means that the role of the roughness-induced shielding diminishes. Nevertheless, the length fraction of deviated segments along the crack path can be reasonably estimated by the mean value $\bar{\eta}$ during the crack growth.

When assuming an ideally straight crack propagating in a quasi-brittle matrix ($R = 0$), the simplest "intrinsic" form of the well known Paris–Erdogan relation

$$\left(\frac{\mathrm{d}a}{\mathrm{d}N}\right)_s = A(\Delta K_{in})^n \tag{3.30}$$

applies, where A and n are materials constants. Obviously, no shielding appears because $\eta \to 0$, $g \to 1$, $C \to 0$ $A_B \to 0$ and $R_S = 1$ in Equation 3.29, i.e., $\Delta K = \Delta K_{in}$. This intrinsic behaviour is assumed to be reproduced for all further crack types which means, at least, the same chemical composition and crystallography of the matrix. On the other hand, the presence of shielding reduces the crack driving force and, consequently, the substitution $\Delta K \to \Omega_1 \Omega_2 \Delta K$ is to be made in Equation 3.30. Moreover, an extended crack path due to the crack tortuosity is to be considered. The ratio of the mean crack growth rates along the straight and the tortuous paths must be inverse to that of their lengths. The length l_t of the tortuous crack path is determined by the mean value of its linear (profile) roughness $\bar{R}_L$. In a general case, the length l_m of a partially tortuous crack path can be determined by the mean values of $\bar{R}_L$ and $\bar{\eta}$ as follows:

$$\left(\frac{\mathrm{d}a}{\mathrm{d}N}\right)_s \bigg/ \left(\frac{\mathrm{d}a}{\mathrm{d}N}\right)_m = l_m/l_s \approx 1 - \bar{\eta} + \bar{\eta}\bar{R}_L.$$

This leads to

$$\left(\frac{\mathrm{d}a}{\mathrm{d}N}\right)_m \approx \frac{A}{1 - \bar{\eta} + \bar{\eta}\bar{R}_L} \left(\Omega_1 \Omega_2\right)^n \Delta K^n = \frac{\Omega_1^n \Omega_2^n}{1 - \bar{\eta} + \bar{\eta}\bar{R}_L} \left(\frac{\mathrm{d}a}{\mathrm{d}N}\right)_s. \tag{3.31}$$

When introducing typical values $\bar{\eta} = 0.8$, $g = 0.9$, $\bar{R}_L = 1.2$, $\bar{R}_S = 1.3$, $A_B = 0.1$, $C = 0.1$ and $n = 3$ into Equation 3.31, the growth rate for a partially tortuous crack is an order of magnitude lower than that of the ideally straight one, which nearly covers the scatter observed in experiments [192].

Additional types of shielding and anti-shielding effects can appear in some special materials [177]. When analyzing the fatigue crack growth rate in ductile irons, for example, such effects are induced by graphite noduli that may occupy 25% of the fracture surface, as documented in Figure 3.37. Because this number is about twice as high as that of the volume fraction (13%), the crack path is strongly affected by the presence of nodules. The difference is clearly visible in Figure 3.38, where the left part of the picture corresponds to fatigue fracture whereas the right part corresponds to the ductile final fracture (closer to the volume fraction of noduli). This means that the interparticle distance of graphite noduli, and not the grain or the phase size, is the characteristic microstructural parameter in ductile irons. Both the stress concentration around the noduli and the crack initiation at nodule/matrix interface are to be particularly considered, as expressed in the following summary of shielding and anti-shielding interactions.

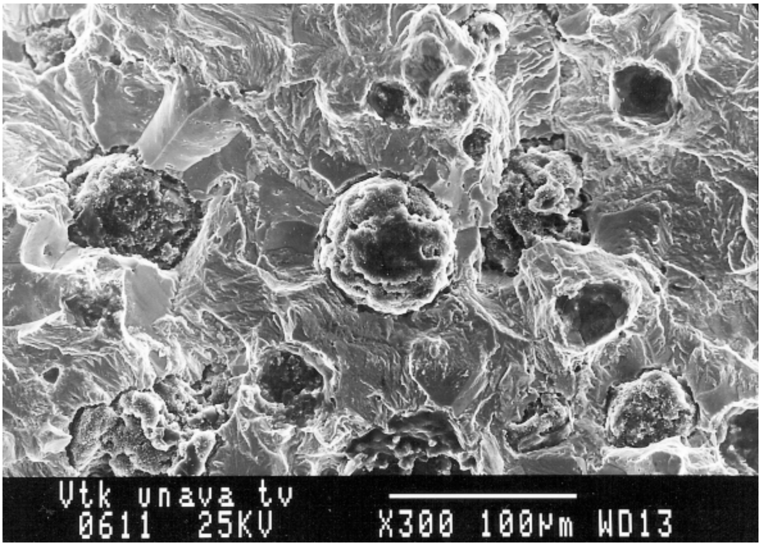

Figure 3.37 The fatigue fracture morphology of the ferritic ductile iron showing a high participation of graphite noduli in the fracture process

Shielding effects:

1. A decrease in ΔK due to a higher level of the roughness-induced crack closure.
2. An extension of the main crack path due to the crack front deflections.
3. A reduction in ΔK due to the crack front deflection and branching. Because of a quasi-brittle fracture mechanism controlled by noduli, the shielding effect caused by crack deflections is to be also taken into account.

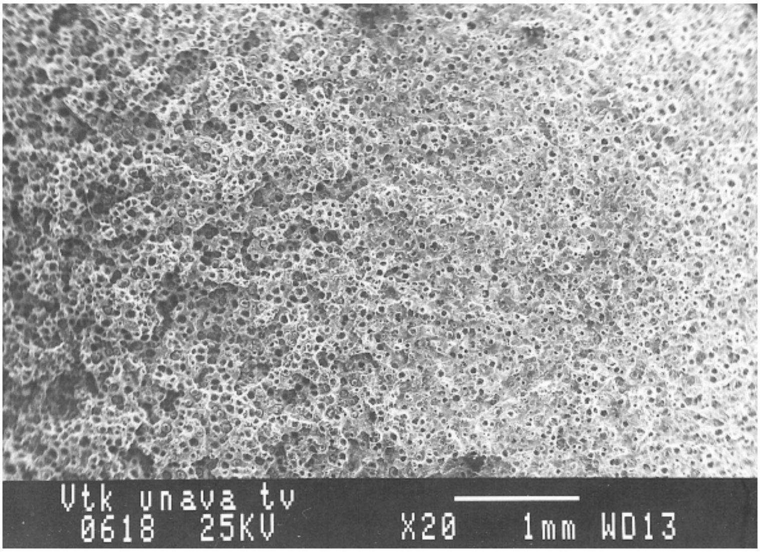

Figure 3.38 The overview fractograph documenting a more frequent presence of graphite noduli on fatigue fracture surfaces (*the left half*) than it would correspond to their bulk concentration (*the right half* – the ductile final fracture)

4. A reduction in ΔK owing to a network of microcracks connected with noduli that are located within the plastic (process) zone at the crack tip.
5. An increase in the material resistance to the crack growth caused by an additional work needed for a creation of the microcrack network;

   Anti-shielding effects:
6. An increase in ΔK due to the difference between elastic moduli E_G (graphite) and E_F (ferrite): $E_G \ll E_F \rightarrow$ the stress concentration around the graphite nodule (the local effect).
7. An increase in ΔK caused by the interaction between the main crack and the microcracks at noduli in front of the main crack.
8. An increase in the crack driving force ΔG caused by the difference between the elastic moduli E_{FDI} (ferritic ductile iron) and E_F: $E_{FDI} < E_F$ (the bulk effect).
9. A reduction of the crack path due to a coalescence of the main crack front with the microcracks.

Note that the effects described by points 4–9 are not considered in Equation 3.31. Therefore, Equation 3.31 can be simply generalized as

$$\left(\frac{\mathrm{d}a}{\mathrm{d}N}\right)_t = \frac{\prod_{i=1}^{n} \Omega_i}{1 - \bar{\eta} + \bar{\eta} R_L} \left(\frac{\mathrm{d}a}{\mathrm{d}N}\right)_s, \tag{3.32}$$

where n is the number of shielding effects involved ($n = 8$ for FDI, point No. 4 irrelevant). Obviously, the shielding effects correspond to $\Omega_i < 1$, whereas $\Omega_i > 1$ applies to the anti-shielding ones.

Equation 3.32 shows a general structure of the Paris–Erdogan law in terms of shielding effects. By using that equation, in particular, a variation of fatigue crack growth rates in engineering materials of the same matrix but with different microstructures can be understood in a quantitative manner.

The difference in intrinsic crack growth rates (given by difference in parameters A and n) is determined mainly by a diverse environmental influence of air on rewelding of newly created fracture surfaces in various materials. In the ultra-high vacuum, the intrinsic crack growth rates of various steels, aluminium and copper alloys are practically identical when normalized to the Young's modulus [275].

3.3 Shear and Mixed-mode Loading

When either a torsion or a pure shear loading is applied to a cracked solid, the crack front can propagate under local shear modes II and III. Therefore, the first subsection deals with basic theoretical concepts of mode II and mode III propagations of long fatigue cracks. A special emphasis is devoted to the micromechanism of mode III crack growth which seem to be much more complicated than that of the mode II. Indeed, the theoretical analysis and recent experiments show that the macroscopic mode III crack propagation can be produced by local mode II or mixed-mode II+III micromechanisms. As a consequence, mode III crack growth rate is often found to be lower than that of the mode II which is documented in the Section 3.3.3.

Pure shear-mode crack propagation usually persists only for a limited number of loading cycles and the cracks incline (or branch) to get a support of the mode I loading. Such a behaviour is predominantly caused by attempting to avoid friction stresses at shear crack flanks that reduce the driving force. This leads to a local mixed-mode I+II, I+III or I+II+III crack propagation as is documented in Sections 3.3.2 and 3.3.3 dedicated to crack propagation under cyclic torsion and pure shear. Typical products of this complicated crack path are so-called factory roofs that represent one of the most remarkable morphological patterns in fatigue. The related model, based on LEFM mixed-mode analysis and 3D fractography, reveals that the opening loading mode dominates the process of factory roof formation. However, the most important result of this analysis seems to be a definition of generalized conditions of mode I branching (kinking) from the shear-mode crack propagation.

The last two parts are devoted to fatigue life of steel specimens subjected to a combined bending-torsion loading. Many engineering components such as shafts, piston rods or gear wheels operate under this type of loading. An application of stress-based multiaxial criteria to predictions of fatigue life and comparison with experimental data obtained on virgin and nitrided specimens are presented. In addition, an extended fractographical analysis of nitrided specimens reveals changes in the fracture surface roughness and the geometry of so-called fish-eye cracks with varying proportion of bending and torsion loading components.

3.3.1 Models of Shear-mode Crack Growth

While the principal micromechanisms of fatigue crack growth in modes I and II can be clearly demonstrated by simple dislocation models, this is not the case in a pure mode III crack propagation. Nevertheless, besides a local twisting of the mode III crack front to get a mode I support [347], a micromechanical interpretation of the crack growth under the out-of-plane (mode III) cyclic shear is possible by means of alternating in-plane (mode II) shears acting either at a microscopically tortuous crack front [348] or between particles cracked near the crack front [348, 349]. Models based on such concepts are described hereafter.

3.3.1.1 Basic Assumptions

Plastic deformation in metals, alloys and intermetallics is usually caused by generation and movement of dislocations. Only in special cases do twinning, grain boundary sliding and diffusion of vacancies give an additional contribution to plastic deformation. In fact, a large amount of dislocations, including those statistically stored, on different slip systems is active during plastic deformation. However, only geometrically necessary ones are sufficient to demonstrate the local fracture processes. Deformations at the crack tip under modes I, II and III produced by geometrically necessary dislocations are schematically depicted in Figures 3.39 and 3.15, where the formation of the new fracture surface is illustrated on the atomistic scale.

As has already been described in Section 3.2, mode I deformation at the crack tip is caused by a symmetric generation of edge dislocations at the crack tip. A new fracture surface is created that is proportional to the Burgers vector times the number of generated dislocations and the crack front length. If one assumes that the new fracture surface will be immediately oxidized so that the dislocations cannot annihilate the newly generated fracture surface during unloading, the crack will propagate with a rate proportional to the cyclic crack tip opening displacement.

Mode II deformation is caused by edge dislocations in the plane of the crack. This process has already been described in the previous section (see Figure 3.15). The crack extension per cycle is, again, proportional to the

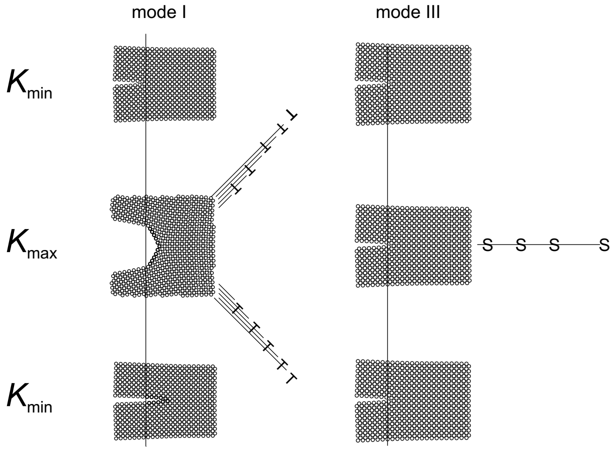

Figure 3.39 Atomistic scheme of fatigue crack advances in the mode I and the mode III during a single loading cycle

number of generated dislocations at the crack tip times the Burgers vector, i.e., to the cyclic crack tip "opening" displacement.

In the case of an idealized mode III loading screw dislocations are generated at (or near) the crack tip as shown in Figure 3.39. Although these dislocations form a plastic zone in a similar way to edge dislocations in the mode II case, they do not create new fracture surfaces along the crack front since their Burgers-vector is parallel to the crack front. Thus, in a planar sample, the movement of the screw dislocations generates ledges on the surface within the plastic zone. Each dislocation generates a surface ledge that is equal to the Burgers vector times the distance of its movement. If one assumes that the surface ledge cannot annihilate during the reverse motion of dislocations during unloading, a new "fracture surface" is generated only within the plastic zone at the surface of the sample. However, along the crack front no crack extension should take place. During further loading the surface crack (the ledge in the plastic zone) will grow in a mode II along the mode III crack front.

In summary, the cyclic movement of dislocations under idealized modes I and II generates a new fracture surface along the whole crack front in each cycle. On the other hand, the screw dislocations in a pure mode III generate only ledges on the surface of the sample which may propagate as "local" mode II cracks. The situation along the crack front is similar to the movement of dislocations near a free surface: the movement of the dislocation will not generate a ledge (a surface step) if its Burgers vector is parallel to the surface. One can imagine that, in reality, mutual interactions of dislocations (including those statistically stored) might generate some microcracks ahead of the crack front in the case of all modes I, II and III. This might lead to local advances of the main crack front by a coalescence of microcracks with

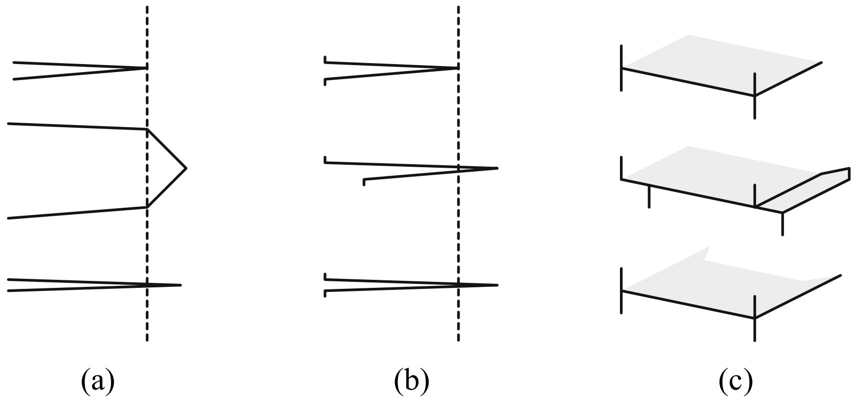

Figure 3.40 The geometric scheme of fatigue crack advance under: (a) mode I, (b) mode II, and (c) mode III during a single loading cycle

the main crack front. However, the crack front advance per cycle produced in such a way must be negligible when compared with the straightforward crack growth caused by the irreversible movement of geometrically necessary dislocations in modes I and II.

The basic difficulty with a pure mode III mechanism in homogeneous materials can also be simply understood when following the macroscopic crack growth schemes drawn in Figure 3.40. During one loading cycle, new surfaces are created ahead of both mode I and mode II fatigue crack fronts by non-zero components of shear displacements parallel to the crack growth direction. Environmental degradation of a newly created surface and irreversibility of dislocation movement are two commonly accepted reasons for an incomplete recovery of atomic bonds at the crack tip during the reversed loading. On the other hand, no shear displacements creating such new surfaces are produced by pure mode III loading.

3.3.1.2 In-plane Shear Models of Out-of-plane Shear Crack Growth

It should be emphasized that a macroscopic mode III crack front propagation does not necessarily need to be produced by pure mode III displacements. Indeed, pure mode II cracking micromechanisms can be exclusively responsible for such crack front advance. In homogeneous materials this demands only a natural assumption of a microscopically tortuous crack front. The crack fronts in engineering materials are never absolutely straight except for single crystals provided that the crack plane is identical with the shear plane and, simultaneously, the direction of the "straight" crack front is exactly parallel to the Burgers vector. In all other cases the crack fronts have a serrated shape at least at the atomic level. Therefore, a global out-of-plane shear deformation will cause fatigue crack propagation of the serrated flanks as depicted in Figure 3.41. In case (a) the small displacement ranges Δu_{III} mean that the sum of edge components of dislocations emitted from the more horizontal parts of the crack front does not exceed a few total Burgers vectors. This corresponds to a loading near the local mode II+III crack growth threshold. In case (b) higher displacement ranges Δu_{III} cause crack propagation along the whole crack front.

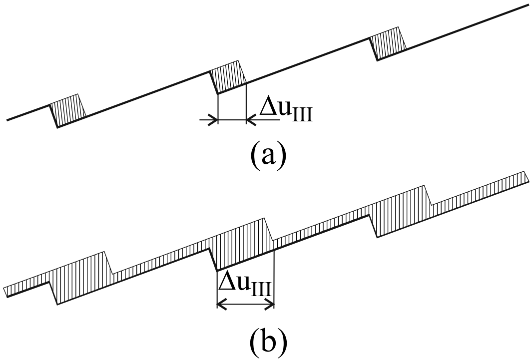

Figure 3.41 Scheme of deformation-controlled mode III growth of a serrated crack front: (a) small out-of-plane strain, and (b) large out-of-plane strain. Reprinted with permission from John Wiley & Sons, Inc. (see page 265)

Let us now consider a simple model of a macroscopically straight but microscopically tortuous crack front under a pure macroscopic mode III loading as shown in Figure 3.42. The triangular microledges are loaded in a mixed-mode II+III, but the out-of-plane shear stress vector can be resolved into two pure mode II (in-plane) components, perpendicular to the legs of the triangle. This enables an alternating step by step growth of the crack front segments under a pure mode II mechanism and leads to a gradual smoothing of the front. This effect may decelerate the macroscopic mode III crack growth or even cause its arrest. Such a behaviour, already reported by Tschegg [350–352], was originally attributed only to the friction and clinching of spatially tortuous crack flanks. The front propagates in the x direction and, indeed, it remains parallel to the direction of the macroscopic "mode III" crack front. This can elucidate a mode III-like fatigue crack growth from a circumferential mode I pre-crack under torsion [353, 354]. When applying the alternating mode II advance of the crack front segments to the circumferential pre-crack described above [348], the resulting shape of the growing crack always becomes qualitatively very similar to that experimentally observed by Murakami [355]. After a certain number of cycles, the shape of the crack front gives an impression of a mode III-like crack propagation (see [348] for more details).

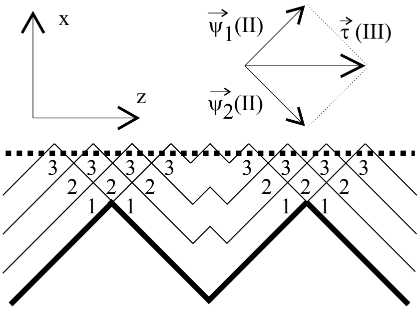

Figure 3.42 Scheme of the alternating pure mode II mechanism of a gradual advance of a microscopically tortuous (macroscopically straight) crack front in a macroscopic mode III. The *thin lines* indicate positions of the crack front after a specific number of fatigue cycles. Reprinted with permission from John Wiley & Sons, Inc. (see page 265)

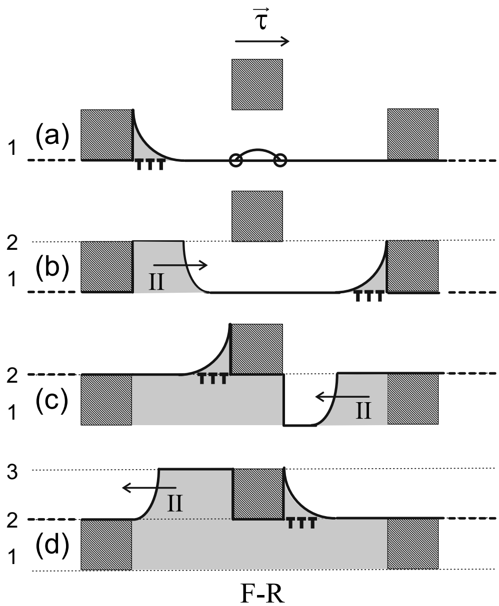

Figure 3.43 Scheme of a mode II mechanism ensuring a macroscopic mode III crack advance (*grey areas*) between secondary phase particles (*hatched squares*). Fracture of the particle-matrix interface is assumed when the crack front approaches the particles

Even a microscopically straight crack front can propagate in a macroscopically pure mode III when considering the assistance of cracks related to secondary phase particles [348, 349]. Indeed, the cracks formed by fracture of particles or particle-matrix interfaces can spread in a pure mode II along the "mode III" crack front as depicted in Figure 3.43. The particles in the vicinity of the crack front are depicted by hatched squares. Frank–Read (F-R) sources adjacent to the main crack front generate dislocation pile-ups at the particle-matrix interfaces and produce coplanar interface cracks with crack fronts nearly perpendicular to the main crack front. Consequently, these cracks can extend under mode II loading along the crack front. The mean crack propagation rate in such a model depends on the total length of the microscopic mode II crack front, given by both the size and the line concentration of particles, associated with a unit length of the "mode III" crack front.

Let us note, however, that another dominant damage mechanism might be responsible for fatigue crack propagation in some metallic materials. One can imagine that, inside the whole cyclic plastic zone ahead of the crack front, many secondary particles can be separated from the matrix during one loading cycle. Such microcracks can then be mutually jointed by shearing under the local mode II to coalesce with the main crack front by breaking remaining ligaments (in a local mixed-mode I+II+III). In this way, rather large crack advances in a nearly pure mode III through the whole process zone can be accomplished during each of several loading cycles and, consequently, the crack growth rate under mode III must not necessarily be lower than that under mode II. Although such a mechanism might be primarily expected to operate in a low cycle fatigue, one cannot exclude that it could also be efficient in a high-cycle fatigue.

3.3.2 Propagation of Cracks under Cyclic Torsion

The great majority of shear loading experiments was performed by cyclic torsion of cylindrical specimens. Many of these studies revealed that there is a big difference in the crack propagation mode in smooth and notched specimens (e.g., [192,350,356,357]). In smooth specimens, the maximal shear and normal stresses are equal in magnitude. Therefore, there is competition between the planes of the maximal shear (parallel and perpendicular to the specimen axis) and those of the maximal tensile stress (main planes, inclined at 45° to the specimens axis) with respect to the fracture process. During the first few cycles, this competition is, as a rule, decided in a favour of the latter planes. Thus, the cracks start to propagate under an opening mode I.

In the circumferentially notched specimens, however, the concentration of shear stresses in the plane perpendicular to the specimen axis stabilizes the crack growth in this plane. Therefore, the cracks propagate under shear modes II+III at least for a certain portion of the fatigue life to an extent which increases with increasing level of torsion loading. In quasistatic and very low-cycle fatigue regions, the cracks can propagate under a dominant shear mode until the final fracture. In these cases, an intensive cyclic deformation within a large plastic zone (embracing many grains and microstructural elements) quickly produces a network of microcracks at grain boundaries or secondary-phase particles. This damage zone is extended ahead of the crack front rather coplanar with the notch plane (e.g., [358]). A coalescence of these microcracks and their shear interconnections with the main crack front keep the crack path in the plane perpendicular to the specimen axis. It is to be expected that the coalescence is controlled predominantly by local mode II shearing processes already described in the frame of the particle-assisted model of mode III crack growth (Figure 3.43). In the high-cycle region, on the other hand, the plastic zone size is comparable to the characteristic microstructural distance. Therefore, the microstructurally-induced tortuosity of the crack front soon causes the initial shear-mode propagation to transfer into the zig-zag mode I growth by formation of factory roofs (e.g., [359, 360]).

In the next subsections, the results of fractographical studies of both smooth and notched specimens made of low-alloyed high-strength steel are presented [348]. Moreover, a theoretical model of factory roof formation is briefly described according to a more extended analysis reported in [361].

3.3.2.1 Smooth Specimens

Initiation and propagation of fatigue surface cracks in smooth cylindrical specimens made of a low-alloy, Cr-Al-Mo steel BS 970/1-83 ($\sigma_y = 850\,\mathrm{MPa}$), were investigated by means of optical and scanning electron microscopes. Pure torsion fatigue tests (number of cycles to fracture $N_f \in (10^4, 10^6)$ cycles) were interrupted after defined numbers of cycles, the specimens were statically fractured in liquid nitrogen and their fracture surfaces were analyzed in SEM.

A formation of a network of microcracks macroscopically perpendicular and parallel to the specimen axis was detected during the first loading stage. This initiation stage was followed by growth of the perpendicular microcracks preferentially in the mode II along the surface and their coalescence.

A typical shape of microcracks that macroscopically lie in the plane perpendicular to the specimen axis is shown in Figure 3.44. However, the real crack plane related to stage I is inclined to the macroscopic plane at an angle of 45°. The crack front at the end of the stage I exhibits, similarly to the scheme in Figure 3.42, a microscopically rough zig-zag geometry. The depth of stage I cracks was in the range of 10–30 μm. The plane of the stage II crack is inclined at 50° to the opposite direction and twisted at 20° so that it gets an additional mode I support. As a rule, the crack fronts at the end of stage II were smoother than those of stage I. All such cracks were rather

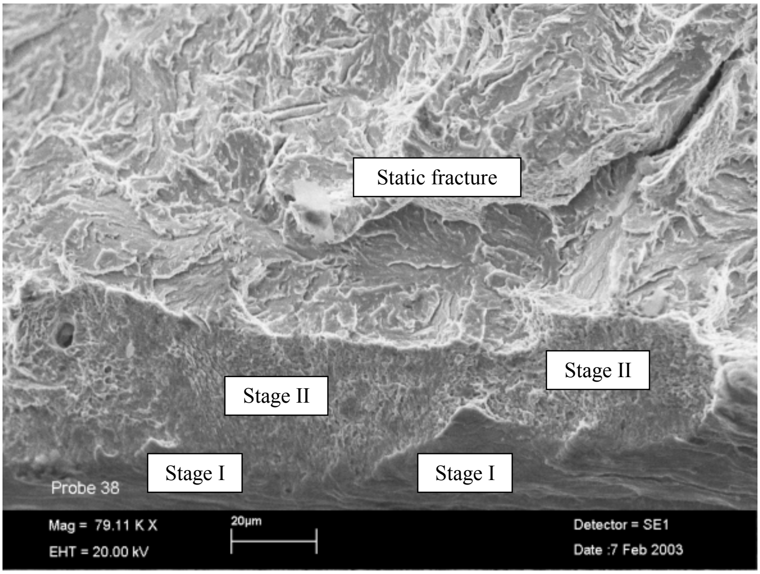

Figure 3.44 A typical shape of a surface microcrack on the fracture surface macroscopically perpendicular to the specimen axis. Both the stage I and stage II planes are inclined at about 45° to the macroscopical fracture plane. Reprinted with permission from John Wiley & Sons, Inc. (see page 265)

slowly propagating in a mode I+III to a depth of nearly 100 µm while simultaneously growing and coalescing along the specimen surface under modes II or I+II. As a result, a shallow circumferential macrocrack has developed round the whole specimen as depicted in Figure 3.45. A much higher growth rate of a mode II or I+II crack front segments in comparison with those of the mode I+III can, however, be partially attributed to the interaction and coalescence of surface microcracks. In some cases, however, long mode I branches were also observed leading to a deep and extremely tortuous surface macrocrack propagating into the interior of the specimen along 45° planes of the maximum tensile stress.

In general, the fractographical analysis revealed that:

1. mode III crack growth was always supported by a mode I component due to the propagation in planes of maximum tensile stress (short stage I cracks) or even with additional twisting of the crack plane (stage II cracks);
2. the mode II (or I+II) crack growth rate was much higher than that of the mode I+III which often resulted in a continuous circumferential crack.

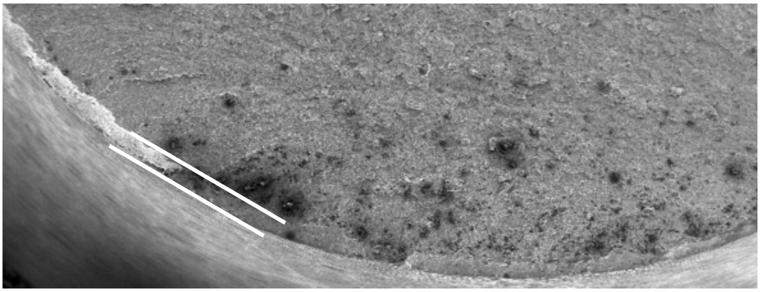

Figure 3.45 SEM picture of a circumferential crack developed after coalescence of surface microcracks in modes II or I +II (depth about 100 μm). White lines mark a segment of the continuous fatigue crack. Reprinted with permission from John Wiley & Sons, Inc. (see page 265)

3.3.2.2 Notched Specimens

A typical way in which the cracks emanating from circumferential notches of cylindrical specimens further propagate into the specimen interior under torsional loading is by formation of factory roofs. The factory roof (F-R) is one of the most extraordinary fractographical patterns ever observed in fatigue and fracture of metallic materials. The roughness (or visibility) of F-R particularly depends on the applied cyclic shear stress amplitude though a significant influence of both the material microstructure and the material yield strength was also observed (e.g., [349, 353, 356, 357, 362]). In spite of the fact that first reports on the F-R appeared in the early 1950s, their formation mechanism was only just beginning to be understood in 2006 when the experimental work of Matake *et al.* [363] appeared. This study revealed that there are three stages of F-R formation: (1) initiation and growth of surface semi-elliptical microcracks under shear loading modes II+III, (2) their interaction, coalescence and growth in the local mixed-mode I+II+III by forming mode I branches (tilted and twisted segments) and (3) growth of the periodic main crack under the prevalent mode I loading. The theoretical work [364] was focused on the problem of friction and shielding phenomena associated with a simple saw-tooth model of F-R patterns. However, neither a detailed geometry of F-R nor any quantitative rules of their formation were reported. Therefore, many principal questions concerning the phenomenon of F-R still remained unsolved:

1. What is the characteristic 3D picture of the F-R?
2. Which physically based relationships control the initiation and growth of the F-R?
3. What is the kinetics of the F-R formation?
4. Why the F-Rs are not observed in the region of a very low cycle fatigue?
5. Are the F-Rs observed only in the case of torsional loading?

Answers to these questions are the main subject of the paper [361]. Some of the most important results are mentioned hereafter.

3D Topography of Factory Roof

The three-dimensional micromorphology of F-R was investigated by stereophotogrammetry and optical chromatography. The fracture surfaces of V-notched cylindrical specimens made of a high-strength low-alloy steel were generated by a reversed torsion loading. Three applied values of the torsion moment $M_{t1} = 13\,\mathrm{Nm}$, $M_{t2} = 17.9\,\mathrm{Nm}$ and $M_{t3} = 22.7\,\mathrm{Nm}$ led to fatigue lives of $N_{f1} = 8.31 \times 10^5, N_{f2} = 2.44 \times 10^5$ and $N_{f3} = 1.58 \times 10^4$ cycles, respectively. Unlike in the latter case (low-cycle fatigue), distinct F-R patterns were observed in both high-cycle regimes. The macrophotograph of the fracture surface with highlighted region of investigated F-R patterns is shown in Figure 3.46.

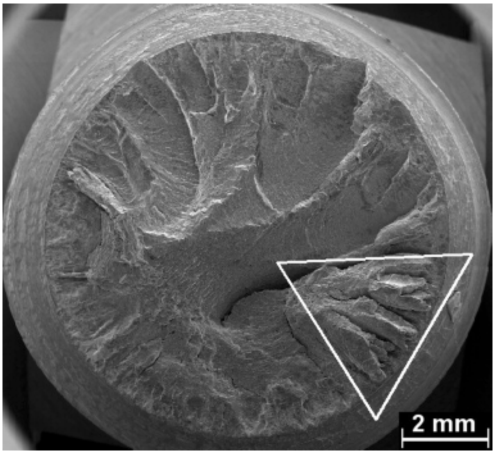

Figure 3.46 Fracture surface with highlighted region of investigated factory-roof patterns ($M_t = 17.9\,\mathrm{Nm},\ N_f = 2.44 \times 10^5$ cycles)

Figure 3.46 documents not only a complexity but also a certain regularity of F-R patterns. Indeed, the lamellar-like F-R structure consists of nearly parallel elongated "mountain-like" massifs (segments) joined by rather narrow valleys with secondary cracks. Near the surface, the massif is usually split into two smaller segments which constitute the initial stage of the F-R formation. Note that there are several F-R patterns which significantly differ in their size: The smaller the F-R patterns the finer is their lamellar structure.

The profiles of F-R topology in nearly tangential directions are plotted in Figure 3.47 as obtained by stereophotogrammetry. The profiles well document the general geometrical features: the slopes (hillsides) of the embryonic

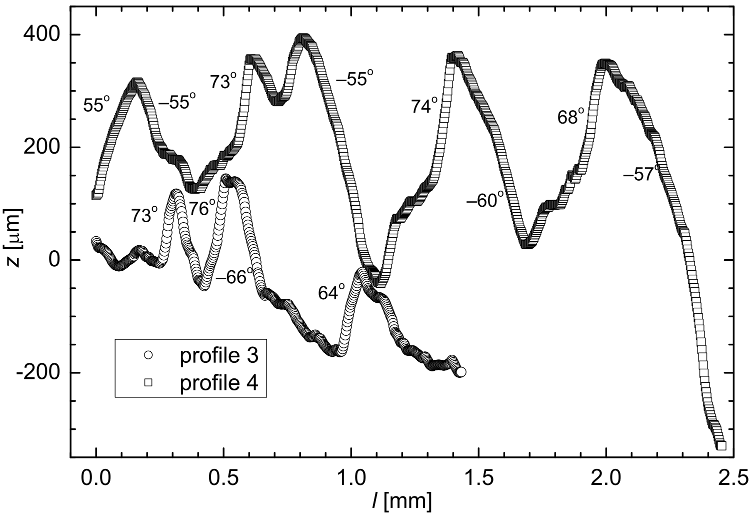

Figure 3.47 Topological profiles of F-R patterns in tangential direction

triangle-like segments gradually decrease starting from an initial angle α of 65–77° near the surface (profile 3), continuing to somewhat lower angles of 60–70° near their conjunction to one main segment (profile 4) to finally reach 40–60° at the top of the main ridge and near the centre of the fracture surface.

A repeated contact of fracture surfaces and the related bending loading initiates secondary cracks in the valleys and also contributes to a further mode I propagation of the main F-R crack inside the specimen. Therefore, numerous wear traces (fibrous patterns or tire tracks) could be found on the SEM pictures of the fracture surface.

Initiation of Factory-roof Patterns

Mixed-mode II+III exists at all points of the semi-elliptical crack front except for two points on the surface (pure mode II) and one point at its centre (pure mode III). However, nearly straight crack fronts of such semi-elliptical cracks change to a highly tortuous profile by inclinations towards a mode I loading rather soon. This change is accomplished by tilting (branching) and twisting of the crack front segments. There are several possible reasons for such behaviour that are particularly related to the crack advance in the radial direction. Indeed, the notch stress concentration decreases and, moreover, there is also a very limited ability of a pure mode III segment to propagate in that direction [348]. Interactions of the crack front with microstructural barriers are accompanied by an increase in the roughness of crack flanks. This leads to an increase in the friction stress (shear closure) and to the reduction of the mode II+III crack driving force.

In order to get a maximum support of the opening mode I to avoid the above-mentioned problems, mode II segments rotate around the axis parallel to the crack front. Such a rotation is relatively easy even for large tilted segments since their planes intersect the main crack plane along the line (curve). On the other hand, twisting around the axis perpendicular to the crack front provides mode III crack segments with mode I support. Since the planes of the twisted elements and the main crack intersect only at a single point, the size of the twisted elements is very limited and the twisting can occur only on microscopic ledges at the main crack front. This means that the formation of mode I branches at the mode II crack front is easier than that at the mode III crack front.

Thus, the F-R formation starts by the creation of mode I branches at particular sites along the elliptical mixed-mode II+III front. A detailed mathematical analysis of mode I branching based on LEFM is given in [361]. The main aim of this theoretical analysis was to predict the most probable sites of mode I branching at the semi-elliptical crack front. These sites are considered to be associated with branches of a pure local mode I stress intensity factor K_I at the semi-elliptical crack front. The angles X and Θ of such branches respectively define the twist and the tilt of the branch with respect to the semi-elliptical crack front as shown in Figure 3.48. Simultaneously, these branches were found to be loaded by a maximal effective stress intensity factor $K_{eff} = (K_I^2 + K_{II}^2 + (1-\nu)^{-1}K_{III}^2)^{1/2}$. The polar angle φ defines the position of the branch at the semi-elliptical crack front as also depicted in Figure 3.48.

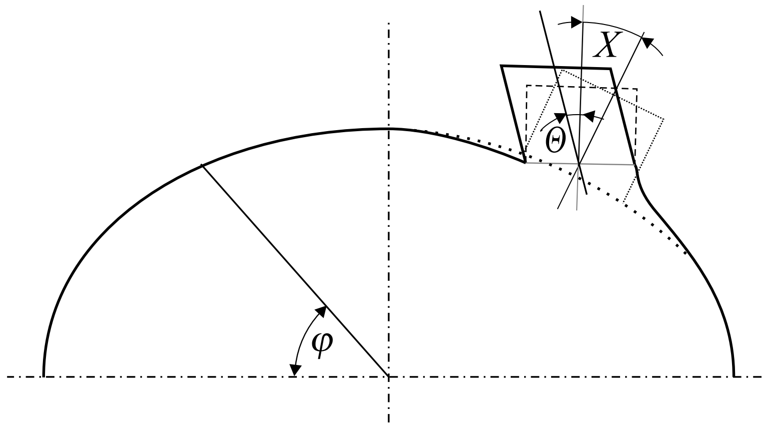

Figure 3.48 Scheme of a branched element at the semi-elliptical crack front

The analysis revealed that the tilt angle $\Theta = 71.6°$ keeps the same value within the whole range of $\varphi \in \langle 0°, 180° \rangle$ and the twist angle $X = 19.5°$ remains constant within the range $\varphi \in \langle 0°, 54° \rangle$ and $\varphi \in \langle 126°, 180° \rangle$. Starting from the critical polar angle $\varphi_c = 54°$, the twist angle X decreases to zero which corresponds to $\varphi = 90°$. This result holds well for all semi-elliptical cracks exhibiting aspect ratios in the range $b/a \in (0.6, 0.8)$ which contains a great majority of the experimentally observed semi-elliptical cracks. Only the critical polar angle changes in the range $\varphi_c \in \langle 45°, 60° \rangle$.

It should be emphasized that the superposition of tilting and twisting raises the values of $K_{I,\max}$ (or $K_{\mathit{eff},\max}$) for the mode I branch significantly above the original K_{eff} values for the semi-elliptical crack front just before the initiation of the mode I branch. This very important result can be understood in terms of a synergy effect of both the mode II and the mode III loading on the creation of the mode I branch. For all the semi-elliptical cracks, the branches of maximal values of $K_{I,\max}$ (or $K_{\mathit{eff},\max}$) lie in the ranges $\varphi_{\max} \in (20°,\ 30°)$ and $(150°,\ 160°)$. This ranges determines two segments on the semi-elliptical crack front that correspond to the maximal probability of the creation of the F-R nuclei (mode I branches).

The initiation and further propagation of F-R nuclei is conditioned by exceeding the mode I threshold for the HSLA steel on the mode I branches. For cracks with $b/a \in (0.6, 0.8)$, the computed maximal values of K_I are in the range $K_{\max} \in (6.4, 7.1)\,\mathrm{MPa\,m^{1/2}}$. These values are higher than the threshold amplitude $\Delta K_{th}/2 = 4.6\,\mathrm{MPa\,m^{1/2}}$ but still sufficiently close. This result, along with the good prediction of initiation sites, confirms the plausibility of the theoretical approach. Moreover, it can be used to define generalized conditions of mode I branching (kinking) from the shear-mode propagation:

1. the first branch forms at that site of the crack front where the value of ΔK_I on its facet would be maximal;
2. the branching appears at the moment when this maximal ΔK_I-value exceeds that of the threshold ΔK_{Ith} related to the applied cyclic ratio.

The geometrical and microstructural parameters related to both the crack front and the crack wake (the level of the roughness-induced closure) of the mode I branch might be somewhat different from those of a crack in the standard specimen used for the determination of ΔK_{Ith}. This difference could be a reason for some deviations from the rule (2). Nevertheless, the above-mentioned conditions give a general frame to a quantitative understanding of the transition from the shear-mode propagation to that under the opening mode.

Model of Factory-roof Formation

On the basis of the above-mentioned results, the kinetics of F-R formation can be qualitatively assessed. Individual stages of F-R formation in terms of gradual positions of F-R crack front are shown in Figure 3.49, where a 3D image of the investigated F-R is depicted according to the stereophotogrammetric reconstruction. The process starts with the creation of mode I branches at well defined sites of semi-circular cracks. The positions of mode I branches corresponding to the ranges $\varphi_{\max} \in (20^\circ,\ 30^\circ)$ and $(150^\circ,\ 160^\circ)$ are marked on the crack fronts of all deduced semi-elliptical cracks. One can see that there is a good correspondence with practically all the real initiation sites of F-R nuclei. In the cases of considerable size difference of neighbouring semi-elliptical cracks, the mode I branch appears first on the larger semi-ellipse. When a large semi-ellipse is adjacent to a small one, a short stage of a backward growth towards both the critical site at the smaller semi-ellipse and the remaining semi-elliptical crack front may appear due to their mutual interaction. This accelerates both the initiation of the second mode I branch and the coalescence process.

After the coalescence, the initial F-R crack front consists of the local spatial ledges and branches that form the embryonic massifs connected by the remaining fronts of semi-elliptical cracks. The contact bending loading on these massifs starts to produce the secondary cracks adjacent to the nuclei and, later on, to the main valleys (or hilltops on the mating fracture surface). These mode I cracks spread in planes inclined at 45° to the macroscopic plane of the maximum shear stress, and eventually approach the advancing F-R crack front.

In order to reduce the line tension of such a tortuous front, the embryonic massifs expand in both the radial direction (inside the specimen bulk) and the tangential direction (along the semi-elliptical fronts). In this way, both the width and the height of the nuclei increase while forming the local U-shaped valleys of decreasing width. Finally, the embryonic segments link-up to form the main massifs and to terminate the local valleys. From that moment on, the main front of a nearly saw-tooth profile propagates further in the direction of a maximum increase of the crack driving force, i.e., more or less in the radial direction. For simple geometrical reasons, the F-R crack front has to contract during propagation towards the specimen centre. Consequently, the main massifs are brought mutually closer and their heights and widths decrease. The extinction of F-R patterns near the specimen centre precedes the final fracture of the specimen. A geometrical model of F-R patterns [361] also revealed that both the height h and the width w of the F-R patterns decrease with decreasing distance (or increasing density) of individual segments.

When considering all the presented theoretical and experimental results, the first three basic questions seem to be answered in a satisfactory manner. The last two questions can also be answered in a rather simple way. The density of initiated semi-elliptical cracks increases with decreasing number of cycles to failure (increasing applied stress range). For the above-mentioned geometrical reasons, the higher the density, the lower the size of the F-R patterns. In the very low-cycle and quasistatic regions, moreover, the microcrack coalescence in the damage zone start to dominate the fracture process. Both these facts mean that the F-R practically vanishes when approaching

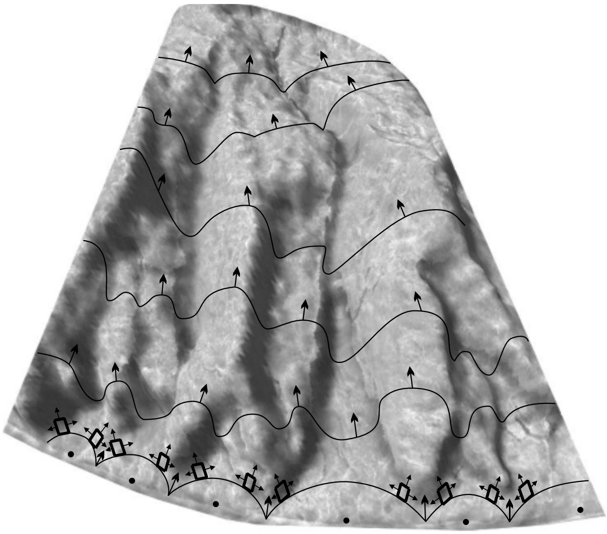

Figure 3.49 The 3D model of the F-R formation

these regions. It should be noted that some authors emphasize another reason for the vanishing of F-R: the shear displacements in the low-cycle regime become very large and the related strong wear can destroy the F-R morphology. However, the relevancy of this reason is substantially weakened by the theoretical analysis performed by Vaziri and Nayeb-Hashemi [359] that predicts sliding of F-R walls rather than their abrasive wear in the low-cycle region (see Section 3.3.3 for more details).

Let us finally remark that the biaxial stress state induced by torsion is not the only kind of loading that produces the F-R patterns. Indeed, these patterns are also developing under pure shear loading, i.e., under a uniaxial stress state [365] (see also the Section 3.3.3).

3.3.3 Propagation of Cracks under Cyclic Shear

An investigation of shear-mode cracks in the case of torsional loading is very difficult and ineffective especially in the high-cycle fatigue region. Indeed, there is only a very short initial period of shear mode crack growth during which almost the whole crack front grows either in the mixed-mode I+II+III (smooth specimens) or II+III (notched specimens). Therefore, it is practically impossible to study the pure modes II and III separately in order to distinguish their growth mechanisms. Under this kind of loading, moreover, an appropriate identification of threshold values ΔK_{IIth} and ΔK_{IIIth} is also impracticable. However, good experiments enabling the pure mode II or mode

III crack tip loading are very difficult to arrange. One of the main problems is to avoid a parasitic mode I vibrational loading. To our knowledge, the first experimental device of this kind was based on four-point bending of angular beams [366, 367]. However, pure mode II or mode III loadings could be realized only in the very centre of the beams so that the extent of the related crack growth was very limited. Diverse experimental arrangements were employed in [365, 368]. The basic idea was to realize simultaneous pure mode II and mode III crack propagations in one specimen. Two different devices were utilized for low-cycle and high-cycle fatigue experiments. Hereafter, a brief description of these experiments and obtained results are presented.

3.3.3.1 Low-cycle Fatigue

Experimental Arrangement

A loading scheme of a special cylindrical specimen utilized in the experiments is introduced in Figure 3.50. Two very sharp circumferential notches were machined by a lathe tool and additionally subjected to a compressive loading in order to produce a crack-like notch. Then, the specimen was annealed in order to remove residual stresses at the crack tip induced by compressive loading.

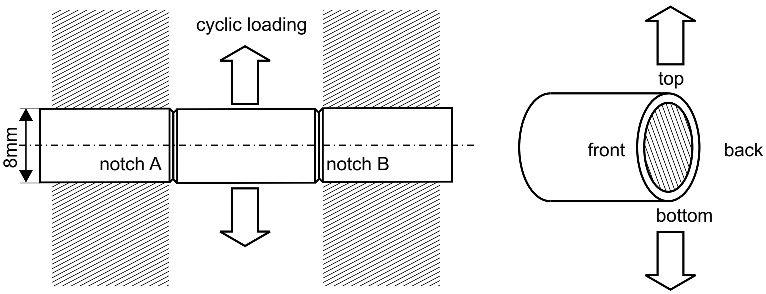

Figure 3.50 The loading scheme of the two-notch specimen. Reprinted with permission from Elsevier B.V. (see page 265)

Specimens were then placed into a fixed rigid holder in the machine, where both side parts of the specimen could be gripped tightly. The whole middle part of the specimen (in between both circumferential notches) was gripped by a moving part of the machine and loaded strictly uniaxial, in tension and compression. The clamping device was stabilized in order to prevent any vibrations and movements out of the tensile axis. A more detailed description of the testing device can be found in [368]. In this way, pure shear loading was transferred to the crack-like notches: At both the top and the bottom sites of two round notches a pure mode II loading was applied, whereas the cracks starting from both middle-site segments were subjected to a pure mode III loading. Thus, four sites of both mode II and mode III growing cracks could be analyzed in a single specimen. Other sites along the notch root experienced a mixed-mode II+III loading.

The constant applied cyclic displacement produced a large-scale cyclic yielding in the ligament that corresponded to a displacement of the crack flaws of the order of a few microns. Such a loading produced crack advance of several hundreds microns during about a hundred of cycles applied in the experiments.

An austenitic steel X5NiCrTi26-15, used e.g., in the aviation industry or as turbine blade material (yield strength of 600 MPa), was selected as an experimental material. This material was nearly free of inclusions or precipitates so that the particle-induced mechanism of mode III propagation (Figure 3.51) was not expected to be dominant.

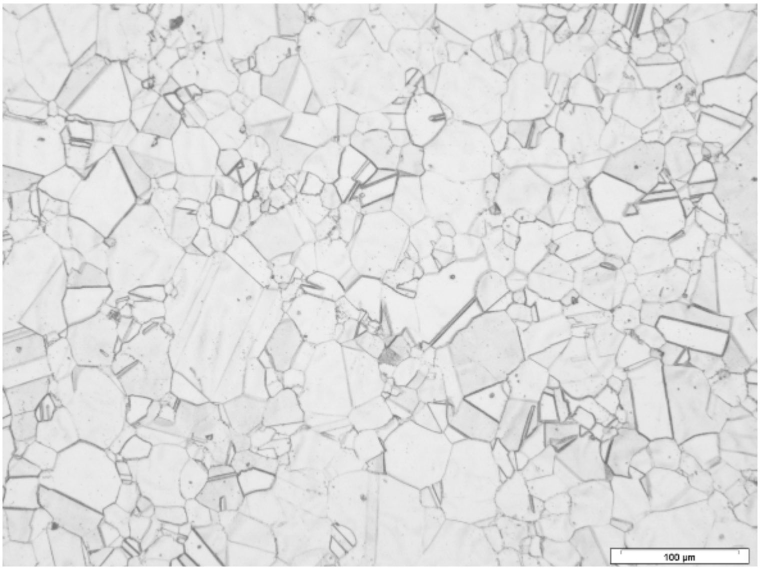

Figure 3.51 The microstructure of the austenitic steel along the crack growth plain. Reprinted with permission from Elsevier B.V. (see page 265)

Crack Growth Data

Two specimens were fatigued by 200 cycles and one specimen by 100 cycles. After the fatigue tests the specimens were broken by cyclic tensile loading. Fractographic observation permitted the determination of the crack extension caused by mode II and mode III cyclic loading, which was clearly distinguishable from the crack propagation caused by the subsequent cyclic tensile loading. The fatigue crack advance under macroscopically pure modes II and III was measured at appropriate sites of fracture surfaces along short crack-front segments of 100 μm. In spite of a relatively high scatter of individual measured fatigue crack lengths, a clear difference in averaged remote mode II and mode III final crack lengths was identified. In the case of 200 applied cycles, the averaged length of 420 μm corresponded to mode II but only 200 μm was determined for mode III. The respective lengths of 210 μm (mode II) and 120 μm (mode III) were measured in the case of 100 applied cycles. Since the fatigue crack path was only a tenth of the total crack length, a nearly constant crack growth rate could be assumed in individual tests. Thus, the remote mode II crack growth rate was approximately twice higher than that of the mode III. It should be emphasized that the numerical analysis for the circumferential crack subjected to pure remote shear revealed that the value of CTOD_{III} was more than 1.4 times higher than that of CTOD_{II} in the small-scale yielding case for the same applied strain range [369]. By assuming the dependence $\frac{\mathrm{d}a}{\mathrm{d}N} \propto \mathrm{CTOD}^2$ this means that, in fact, the crack growth rate under mode III must have been about five times lower than that under mode II.

Fractographical Observations

A detailed 2D fractographical investigation by means of SEM was performed in order to identify real local fracture modes. Examples of a typical fracture surface produced during the remote mode II loading is shown in Figure 3.52. Here the directions of both the applied shear stress and the macroscopic crack growth are vertical (from the bottom to the top). Many facets covered by striations, mostly nearly perpendicular to the growth direction, could be observed all over the fatigue fracture surface. The occurrence of striations was related to the presence of a small opening mode I at the crack tip that was induced here by local inclinations of the crack front from the plane of maximum shear strain. Sometimes the crack front advanced in pure mode II as can be easily deduced from tire tracks typical for shear mode presence that could be identified in many parts of the fracture surface (see Figure 3.52). The direction of these periodic patterns indicates the expected propagation of the shear crack front from bottom to top.

Many secondary cracks, mostly perpendicular to the growth direction, were also found all over the fracture surface [368]. These cracks are considered to be preferentially created at the corners of asperities left in the crack wake which, under the applied mode II shear, are loaded by cyclic bending.

All observed morphological features confirm that pure mode II and combined mode I+II are the dominating microscopical fracture mechanisms caus-

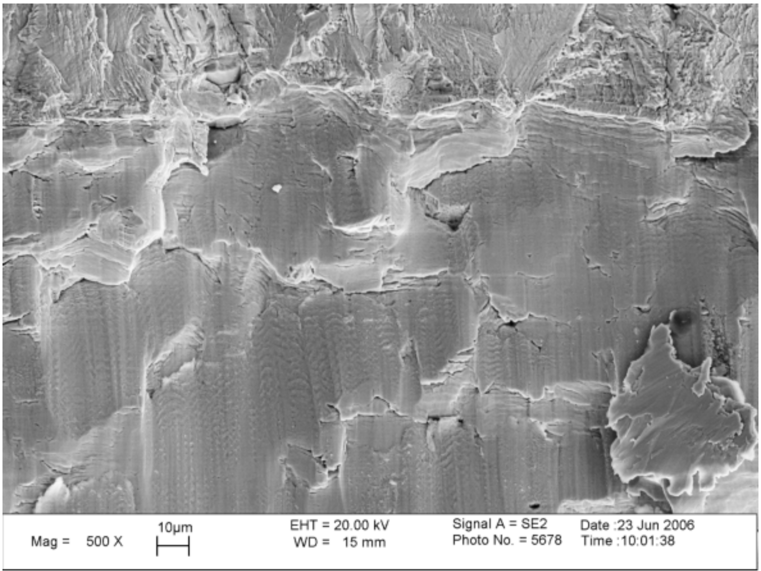

Figure 3.52 Tire tracks and striations (*at the top*) on the remote mode II fracture surface

ing rather straightforward advance of the crack front under remote mode II loading.

A typical example of fracture surfaces produced under remote mode III is displayed in Figure 3.53. Here the direction of the applied shear stress is horizontal. The fracture morphology was found to be completely different from that created by the remote mode II. Practically all striation fields confirmed the dominance of mode II or II+III propagation in accordance with the models in Figures 3.42 and 3.43, often supported by mode I due to kinking and twisting of the crack plane. Indeed, a prevalent direction of striations is nearly vertical which means that the crack front propagated horizontally, i.e., under local mode II. A careful examination of the mode III fractographs uncovered some regions looking like mode I+III crack propagation, where the crack formed striations parallel to the external applied mode III loading. Due to the crystallographic nature of plasticity, the macroscopic shear displacement parallel to the crack front is expected to be always locally associated with small mode I and mode II displacements in addition to the mode III ones. Indeed, there must practically always be a certain deviation of the Burgers vectors from the crack front direction.

The secondary cracks generated under remote mode III are more frequent and noticeable than those under remote mode II loading (see Figure 3.53). These cracks, mostly parallel to the growth direction, were created preferentially as corner cracks at the crack-wake asperities bent by external mode III shear. Coalescence of such cracks causes crumbling of asperities and falling

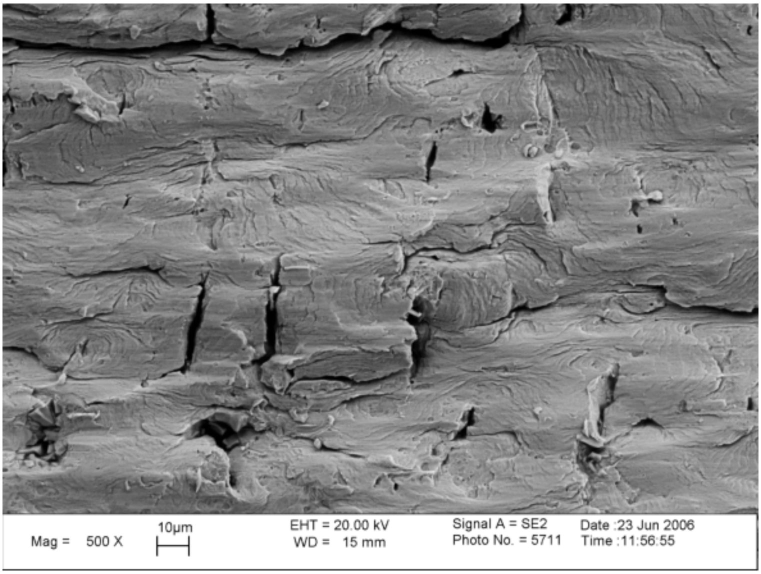

Figure 3.53 Striations on the remote mode III fracture surface revealing the dominance of local mode II or I+II crack propagation mechanisms. Reprinted with permission from Elsevier B.V. (see page 265)

of the fracture surface. As a result, large voids are left in between mating surfaces as was previously detected by Slámečka and Pokluda [370] by using stereophotogrammetrical methods after torsional fatigue fractures. Secondary cracks also sporadically initiated at particles of secondary phases near the main crack front. These cracks then propagated as mode II or I+II cracks in between the particles along the main crack front – see Figure 3.53 (left). Indeed, the striations emanating from the cracked particles indicate a horizontal direction of the crack front propagation in accordance with the model in Figure 3.43.

In summary, the fracture morphologies generated by remote mode II and mode III loadings are significantly different. However, the observed morphological features revealed that the dominating growth mechanisms in both cases are similar: pure mode II and the combined mode I+II. This is in agreement with the models in Figures 3.42 and 3.43. Some indication of the local mode I+III crack propagation under the remote mode III loading was also found. Thus, the local fracture modes caused a complicated zig-zag advance of the crack front unlike the rather straightforward growth under the remote mode II. As a consequence, the crack front propagation rate under remote mode III was significantly lower than that in the remote mode II case.

3.3.3.2 High-cycle Fatigue

Experimental Arrangements

In the high-cycle regime, the shear displacements are much lower than those in low-cycle fatigue and, consequently, small parasitic mode I vibrations and bendings could play a much more important role. Therefore, two other original testing setups (cells) with a higher stiffness were designed and utilized to assure a pure remote shear loading. The loading scheme of the first cell is depicted in Figure 3.54. The construction of the specimen holder and its orientation with respect to the loading axis provided pure mode II loading at the "top" and "bottom" sites of the specimen and pure mode III loading at the "front" and "back" sites. At all other points along the crack front mixed-mode II+III was applied. A circumferential V-notch was machined by a lathe tool at the specimen mid-length and a pre-crack was introduced by a blade mechanism. Finally the specimens were compressed by 20 kN to sharpen the pre-crack and subsequently annealed. The resulting pre-crack is depicted in Figure 3.55. Specimens made of the ferritic steel (0.01 %C) with outer diameter of 8 mm and inner diameter of 4 mm were loaded by different ranges of the nominal ligament shear stress τ_n (the cyclic ratio R = 0.1) After the shear mode tests, the specimens were rapidly fractured in liquid nitrogen. Specimens made of austenitic steel X5NiCrTi26-15 were also tested by means of this experimental set-up. These specimens were fractured by cyclic tensile loading.

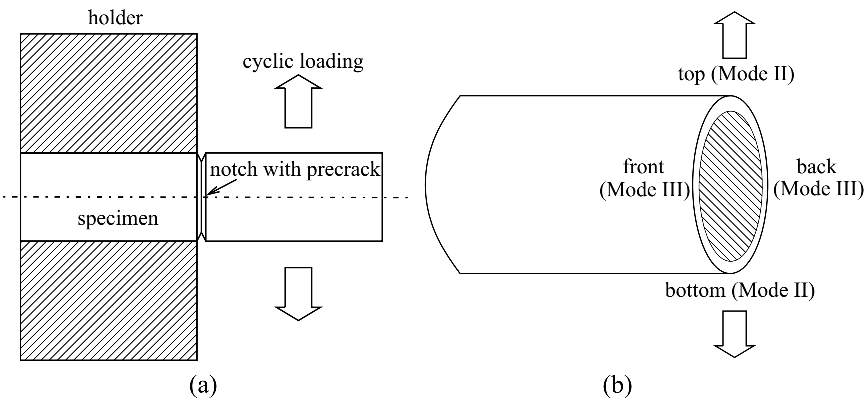

Figure 3.54 The experimental setup: (a) the loading scheme, and (b) the loading modes operating at different specimen sites

The second special cell for loading specimens made of austenitic steel was manufactured to enable higher loadings (see Figure 3.56). The pre-crack was introduced by means of a compressive load of 200 kN. After the shear mode tests, the specimens were also fractured by cyclic tensile loading.

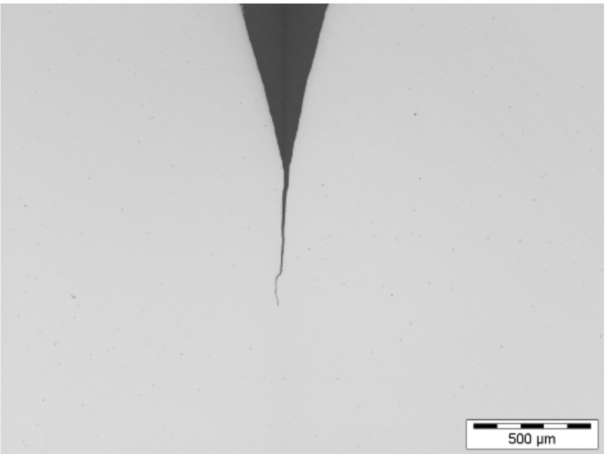

Figure 3.55 The shape of the pre-crack after compression

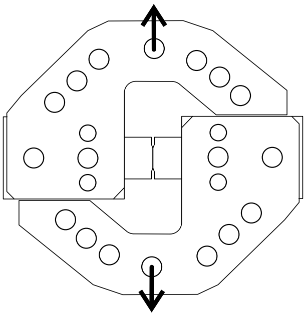

Figure 3.56 The scheme of the loading cell for higher loading forces

Numerical K-calibration of Specimens

In order to determine the mode II and mode III stress intensity factors at the crack tip, a numerical analysis was performed by means of the ANSYS code. Although the loaded specimen was modelled as rotationally symmetric, a full linear–elastic 3D solution had to be used owing to a different symmetry of the loading. In the first step, the stress-strain field along the crack front loaded by the remote shear stress of 180 MPa was determined by utilizing a rough finite-element network. This field was used to create a ring-like submodel with a very fine finite-element network that embraced only the pre-crack region. The submodel surface was loaded by the rimstrains computed for the same surface on the rough model.

In this way, very precise values of the stress intensity factors K_I, K_{II} and K_{III} could be determined for many points along the circular crack front. Mutual shear displacements of crack flanks were calculated at four points

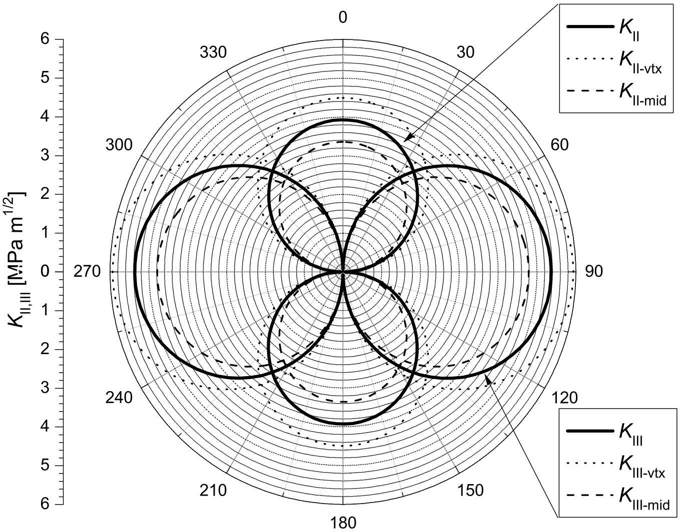

Figure 3.57 Values of K_{II} and K_{III} along the crack front in polar coordinates. The *dashed lines* correspond to the middle points of the finite elements, the *dotted lines* are related to their vertexes and the *full lines* represent the average values

near the crack front. The values of K_I, K_{II} and K_{III} were determined by an extrapolation to the crack front for all applied nominal shear stress ranges. The ratio of maximal values in pure shear modes II and III was found to be $K_{III\,\max}/K_{II\,\max} = 1.37$ while the values of $K_{I\,\max}$ were negligible (for more results see [365]). This can be seen from the polar diagram of maximal values of $K_{III\,\max}$ and $K_{II\,\max}$ in the loading cycle (for $\tau_n = 180\,\mathrm{MPa}$) that is plotted in Figure 3.57. It should be emphasized that such determined numerical values of $K_{III\,\max}$ were found to be in excellent agreement with calculations performed using the asymptotic method [371]. Let us mention that recent results of elasto-plastic analysis have shown that the ratio $\mathrm{CTOD}_{III\,\max}/\mathrm{CTOD}_{II\,\max}$ in the large-scale yielding case can be even higher than 1.4 [369].

Crack Growth Data

The spatial shear crack path was determined by an SEM identification of the fracture surface morphology in selected rectangular regions corresponding to pure mode II and mode III loading of both austenitic and ferritic steels. The fracture morphology of pure mode II and III shear cracks is shown in Figure 3.58 for the austenitic steel. The areas corresponding to the pre-crack, the shear crack propagation and the final tensile fracture are marked as well. Practically all the mode II shear cracks were globally inclined from the shear plane. Averaged deflection angles were found to be $60\pm16°$ (austenite) and $53\pm15°$ (ferrite). This means that just after the start of shear mode cycling, the mode II cracks branched to the opening mode I. In terms of the previously defined branching condition, the value of $\Delta K_{Ith} \approx 5.8$ for the austenitic steel at $R = 0.1$ must have already been exceeded at the mode I branches. Indeed, the crack flanks near the pre-crack tip are microscopically tortuous, as can be seen in Figure 3.55. Consequently, a bending moment at interlocked asperities has, most probably, produced an additional local mode I loading at the pre-crack tip. The fracture morphology of mode III cracks typically consisted of factory-roof patterns (Figure 3.59).

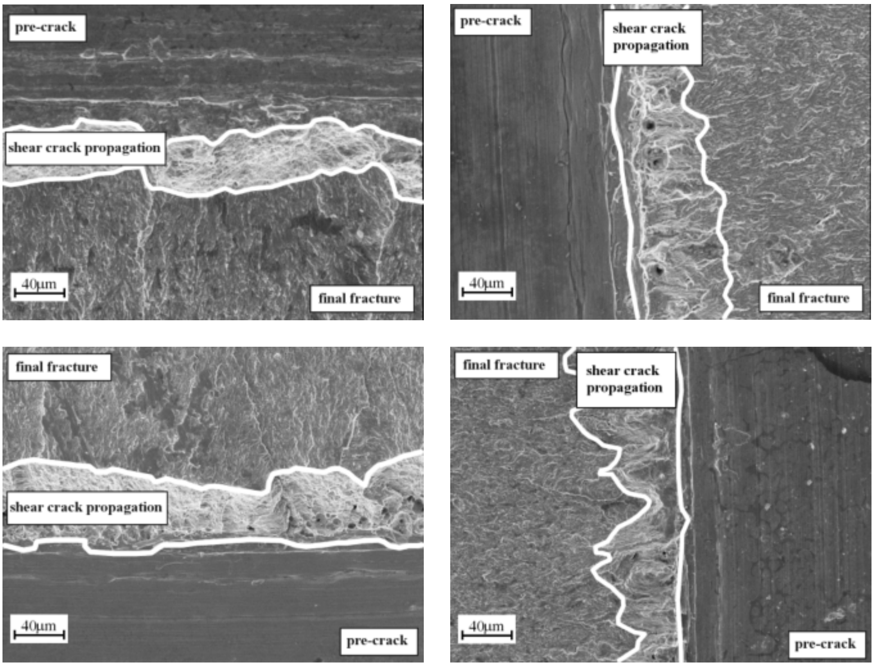

Figure 3.58 SEM pictures of fracture surfaces showing propagation of shear cracks. The *left* (*right*) pictures correspond to remote mode II (mode III) cracks

Profiles of fracture micromorphology in the direction of applied stress for both mode II and mode III crack propagation stages are depicted in Figure 3.60. As expected, the crack path in the remote mode III case clearly reveals the formation of factory roofs. The related interlocking of crack flanks caused by the F-R asperities must lead to a dramatic reduction of the crack tip driving force especially in the near-threshold crack growth region [359]. This is schematically shown in Figure 3.61 where the interlocked F-R asperities are associated with a low effective crack driving force $\Delta K_{eff,III}$ for low applied values of τ. When the applied shear stress or ΔK_{III} becomes sufficiently

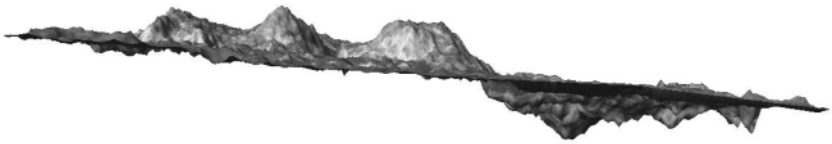

Figure 3.59 3D picture of the factory-roof morphology created by a pure mode III shear (viewed parallel to the fracture surface)

high (or the FR patterns sufficiently small) the sliding of fracture surfaces in the crack wake rapidly becomes dominant. This is accompanied by a sudden increase in the effective driving force in the near-fracture region of the crack growth rate diagram (or in the low cycle region). On the other hand, the microrelief of mode II crack growth is very smooth so that the effective driving force is high. Consequently, the crack growth rate under mode II is expected to be significantly higher than that under mode III.

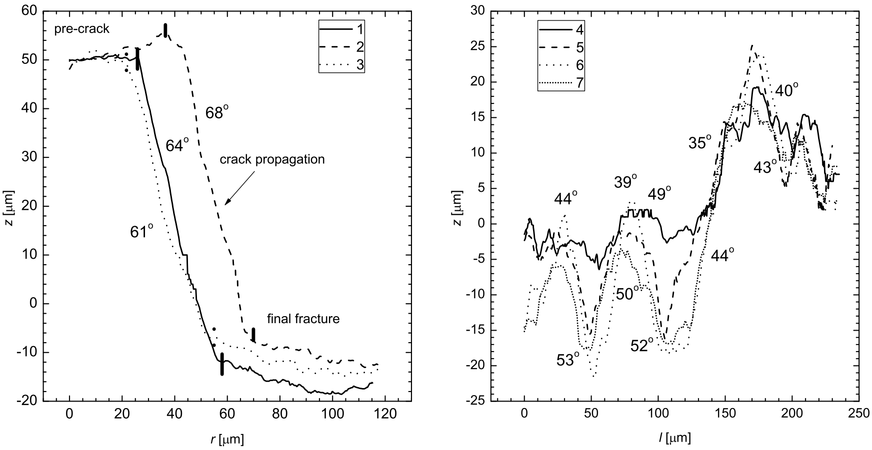

Figure 3.60 Profiles of fracture morphology in the direction of applied stress corresponding to mode II (*left*) and mode III (*right*) crack propagation. The smooth profiles 1, 2 and 3 show propagation paths of three elements of the mode II crack-front (extended in the tangential direction) along the radial direction r towards the specimen centre. The profiles 4, 5, 6 and 7 correspond to consecutive propagation stages of the mode III crack front extended in the tangential direction l

Since the length of the shear mode cracks was an order lower than that of the pre-crack, a nearly constant crack growth rate during the shear propagation could be assumed. Therefore, the crack growth rate was calculated simply by dividing the total length of shear cracks by the corresponding num-

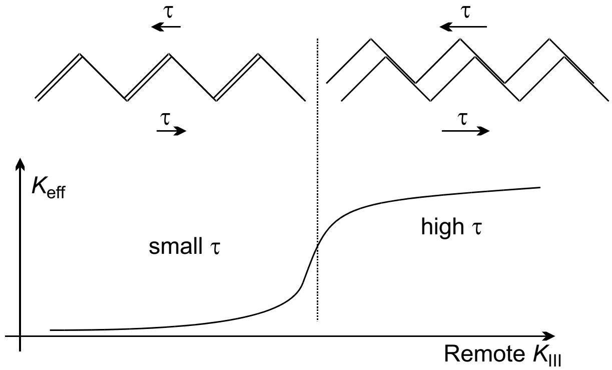

Figure 3.61 The scheme of the interlocked factory-roof asperities in the case of a small applied shear stress (*the left part*, wedging) and a high applied shear stress (*the right part*, sliding)

bers of cycles. The near-threshold crack growth curves for both the mode II and the mode III propagation in the austenitic steel are plotted in Figure 3.62. The related regression curves follow the Klesnil–Lukas relationship

$$\frac{\mathrm{d}a}{\mathrm{d}N} = A\left(\Delta K^n - \Delta K_{th}^n\right),$$

where A and n are material constants [289]. While the exponents $n_{II} = 5.6$ and $n_{III} = 5.1$ are nearly identical, the constants $A_{II} = 1.131$ and $A_{III} = 2.721$ [MPa m$^{1/2}$, m] are significantly different as well as the calculated thresholds $\Delta K_{IIth} = 3.5\,\mathrm{MPa\,m}^{1/2}$ and $\Delta K_{IIIth} = 4.7\,\mathrm{MPa\,m}^{1/2}$.

According to a general condition of mode I branching defined in Section 3.3.2, this effect can appear only when the mode I threshold value is exceeded at the mode I branch. With the help of the local bending moment induced by the crack-wake asperities near the crack tip, the value $\Delta K_{Ith} \approx 6\,\mathrm{MPa\,m}^{1/2}$ for the austenitic steel ($R = 0$) [372] could already be exceeded by applying ΔK_{II} values very close to the remote mode II threshold $\Delta K_{IIth} \approx 3.7\,\mathrm{MPa\,m}^{1/2}$. One can also clearly see that, for the same value of the applied stress intensity range, the crack growth rates for the mode II loading are about six times higher than those for the mode III loading. This is in agreement with results achieved in the low-cycle fatigue region and confirms the diversity of the mechanisms of mode II and mode III crack propagation in metals within the whole range of the fatigue life.

Also in the case of the ferritic steel, the threshold values $\Delta K_{IIth} = 1.2\,\mathrm{MPa\,m}^{1/2}$ and $\Delta K_{IIIth} \approx 2.0\,\mathrm{MPa\,m}^{1/2}$ were found to be different. This is documented in Figure 3.63, although further experiments must be performed in order to obtain more precise results.

To prevent mode I branching in order to observe a sufficiently long mode II and III crack growth near the threshold, the pre-crack flanks should be very smooth, particularly in the close vicinity of the pre-crack tip. This might be

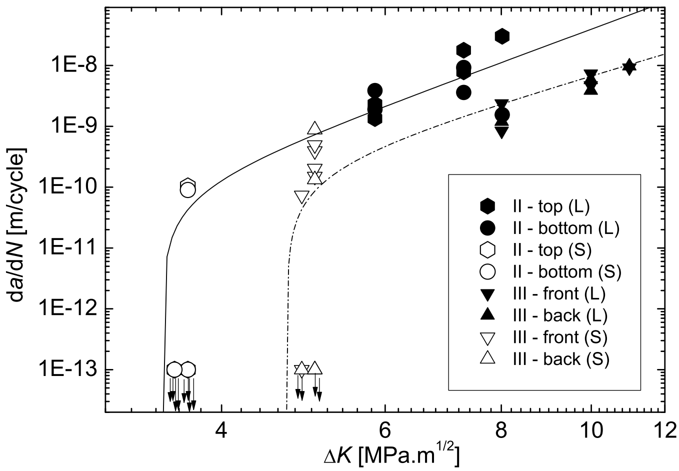

Figure 3.62 The crack growth curves for the austenitic steel in the near-threshold region

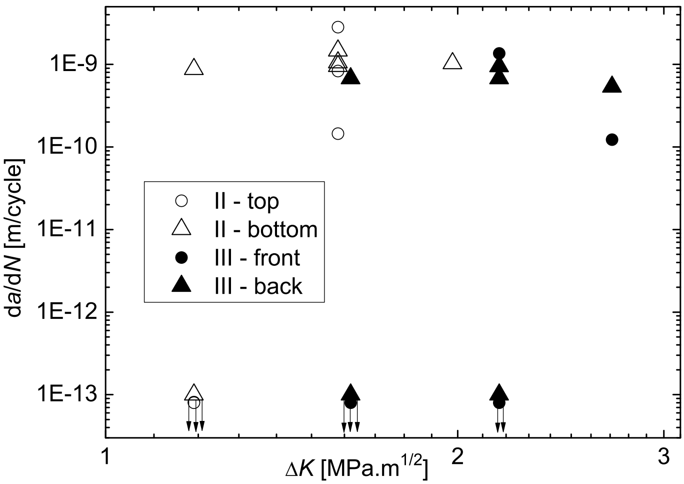

Figure 3.63 The crack growth data for the ferritic steel close to the fatigue threshold

achieved by generation of an additional short pre-crack by cyclic compressive loading. In this case, the intrinsic values $\Delta K_{IIth,in}$ and $\Delta K_{IIIth,in}$ could be determined by a sequential cyclic shear mode loading. Results of first experiments of this kind are plotted in Figure 3.64. For mode II loading, open stars and squares are associated with planar shear crack propagation with the deflection angle less than 15°, while solid stars and squares correspond to branching angles higher than 40°. The latter symbols can be found only for $\Delta K_{II} = 7.3\,\mathrm{MPa\,m^{1/2}}$ which is a higher value than that of ΔK_{Ith}. Due to the smoothing and sharpening of the crack tip by cyclic compression, there is also a significant shift of both mode II and mode III thresholds towards lower values ($\Delta K_{IIth,in} \approx 2.9\,\mathrm{MPa\,m}^{1/2}$ and $\Delta K_{IIIth,in} \approx 4\,\mathrm{MPa\,m}^{1/2}$).

These results do not confirm the identity $\Delta K_{IIth,in} = \Delta K_{IIIth,in} \approx 9.6\,\mathrm{MPa\,m}^{1/2}$ as obtained by Murakami *et al.* [354] for carbon steel under cyclic torsion. It should be noted, however, that the methodology used in the latter case was rather complicated and, in our opinion, not correct. Indeed, the extremely high threshold values were the result of a wrong presumption that the branching of the mode II+III cracks occurred at the first joint sites of the factory-roof massifs. In fact, the branching starts much sooner: at sites of embryonic semi-ellipses that are well defined by the theoretical analysis in Section 3.3.2. Consequently, the real threshold values should be substantially lower.

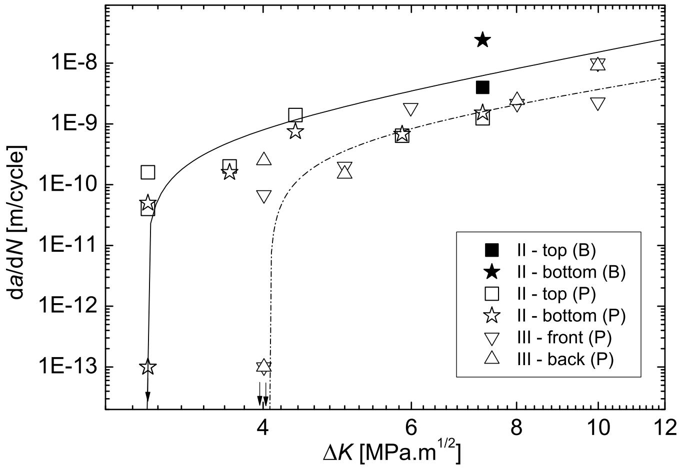

Figure 3.64 The crack growth curves for the specimens with an additional pre-crack made by cyclic compression in the austenitic steel

The results achieved in both the low and high-cycle region show that the crack growth rate under mode III is about five times lower than that under the mode II for the same applied strain or stress ranges. This fact was already respected in empirical codes for mixed-mode crack growth in the small scale yielding case [373, 374] by utilizing the effective stress intensity factor approximately as

$$K_{\mathit{eff}}^2 = K_I^2 + \xi K_{II}^2 + \eta K_{III}^2, \qquad (\eta < \xi < 1).$$

Thus, it seems that a similar assumption can be also extended to the large scale yielding regime by using the effective crack opening displacement

$$\mathrm{COD}_{\mathit{eff}} = \mathrm{COD}_I + \xi^* \mathrm{COD}_{II} + \eta^* \mathrm{COD}_{III}, \qquad (\eta^* < \xi^* < 1).$$

Let us finally emphasize that, up to now, the experiments on mode II and mode III crack propagation were accomplished only for a relatively small number of engineering materials. In some metallic materials, the dominant fatigue crack growth mechanism must not necessarily be based on the environmentally prevented recovery of the newly created surfaces ahead of the crack front. Indeed, the accumulation of damage due to the creation of microcracks inside the cyclic plastic zone could also control the crack propagation rate in some metallic materials. This could lead to nearly identical crack growth rates under both mode II and III loading cases. Therefore, the problem of shear-mode fatigue crack propagation still remains a challenge for further research.

3.3.4 Crack Growth and Fatigue Life under Combined Bending-torsion Loading

Combined cyclic bending-torsion (CCBT) is a kind of multiaxial fatigue loading that is probably mostly employed in scientific experiments. At the same time, many engineering components such as, for example, shafts, piston rods or gear wheels also operate under CCBT. Basically, there are two types of CCBT observed in practice and utilized in experiments: in-phase and out-of-phase loadings. In the next subsections, some results achieved by us and our co-workers using the in-phase variant of CCBT are presented.

3.3.4.1 Fatigue Life

Stress State under In-phase Bending-torsion Loading

In experimental investigations sinusoidal loading is usually applied. In this case, the time-dependence of non-zero components of the stress tensor can generally be expressed as

$$\sigma_{ij}(t) = \sigma_{ijm} + \sigma_{ija}\sin(\omega_{ij}t - \varphi_{ij}), \qquad i,j = x,y,z, \tag{3.33}$$

where σ_{ijm} is the mean stress, σ_{ija} is the stress amplitude, ω_{ij} is the angular velocity and φ_{ij} is the initiation phase.

Because of the biaxial character of bending-torsion, the loading trajectories can be plotted in 2D diagrams [312]. In the cases of pure bending, pure torsion and, generally, in-phase loading, the directions of principal stresses remain time-independent. This means that the ratio of torsion/bending stresses is constant along the whole loading trajectory ($\tau/\sigma = \text{const}$). In the case of out-of-phase loading, on the other hand, only the amplitude ratio remains constant ($\tau_a/\sigma_a = \text{const}$) and the principal stress directions are time dependent.

The stress tensor at an arbitrary point inside the cylindrical specimen under in-phase bending-torsion (Figure 3.65) can be described as

$$\hat{\sigma}(t) = \begin{bmatrix} 0 & 0 & \tau_{xz}(t) \\ 0 & 0 & \tau_{yz}(t) \\ \tau_{zx}(t) & \tau_{zy}(t) & \sigma_{zz}(t) \end{bmatrix} = \begin{bmatrix} 0 & 0 & -\frac{M_t(t)}{I_p}y \\ 0 & 0 & \frac{M_t(t)}{I_p}x \\ -\frac{M_t(t)}{I_p}y & \frac{M_t(t)}{I_p}x & \frac{M_b(t)}{I_y}x \end{bmatrix},$$

$$I_p = \frac{\pi d^4}{32}, \qquad I_y = \frac{\pi d^4}{64},$$

where the components denoted τ (σ) represent shear (normal) stresses, M_t (M_b) is the torsion (bending) moment, d is the specimen diameter, I_p is the polar inertial moment and I_y is the inertial moment round the y-axis ([279]).

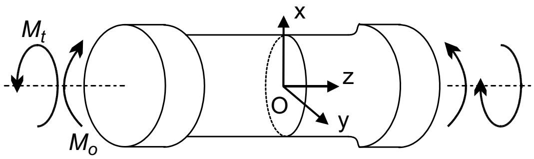

Figure 3.65 A scheme of the cylindrical specimen loaded by bending-torsion moments

In the case of in-phase symmetric loading one can consider $\omega_\sigma = \omega_\tau = \omega$, $\sigma_m = \tau_m = 0$ and $\varphi_\sigma = \varphi_\tau = 0$ so that Equation 3.33 becomes very simple:

$$\sigma_{ij}(t) = \sigma_{ija}\sin(\omega t), \qquad i, j = x, y, z.$$

This means that only stress amplitudes are sufficient for further reasoning and the time dependence can be neglected.

Since the short crack growth near the surface of rather smooth specimens occupies a significant part of their fatigue life, the stress components at the surface ($x = \pm d_2$, $y = 0$) are the most important parameters, especially in the case of high-cycle fatigue. At these points the principal stresses can be determined by solving the equation

$$\begin{vmatrix} -\sigma_{ia} & 0 & 0 \\ 0 & -\sigma_{ia} & \tau_a \\ 0 & \tau_a & \sigma_a - \sigma_{ia} \end{vmatrix} = 0,$$

where $\sigma_{zza} = \sigma_a$ and $\tau_{yza} = \tau_{zya} = \tau_a$. Hence, in accordance with the convention $\sigma_1 > \sigma_2 > \sigma_3$, one obtains

$$\sigma_{1a} = \frac{1}{2}\left(\sigma_a + \sqrt{\sigma_a^2 + 4\tau_a^2}\right), \quad \sigma_{2a} = 0, \quad \sigma_{3a} = \frac{1}{2}\left(\sigma_a - \sqrt{\sigma_a^2 + 4\tau_a^2}\right),$$

which means that the combined bending-torsion loading induces the biaxial state of stress. One can easily show that the directional cosines c_{ij} related to i and j axes of the principal coordinate system can be expressed as

$$c_{ji}^2 = \frac{(\sigma_{kk} - \sigma_j)(\sigma_{mm} - \sigma_j) - \sigma_{km}^2}{3\sigma_j^2 - 2\sigma_j I_1 + I_2}, \, i \neq k \neq m,$$

where J_1 and J_2 are the first and second invariant of the stress tensor, respectively:

$$J_1 = \sigma_{xx} + \sigma_{yy} + \sigma_{zz}, \qquad J_2 = \sigma_{xx}\sigma_{yy} + \sigma_{xx}\sigma_{zz} + \sigma_{yy}\sigma_{zz} - (\tau_{xy}^2 + \tau_{xz}^2 + \tau_{yz}^2).$$

The maximal shear stresses, crucial for the short crack initiation, act in planes inclined at $\pm 45°$ towards the principal axis, and their magnitude is

$$\tau_{12} = \frac{1}{2}\left|\sigma_1 - \sigma_2\right|, \quad \tau_{13} = \frac{1}{2}\left|\sigma_1 - \sigma_3\right|, \quad \tau_{23} = \frac{1}{2}\left|\sigma_2 - \sigma_3\right|.$$

For selected bending-torsion loading cases, the orientation of principal stresses and maximal shear stresses at surface points is depicted in Figure 3.66.

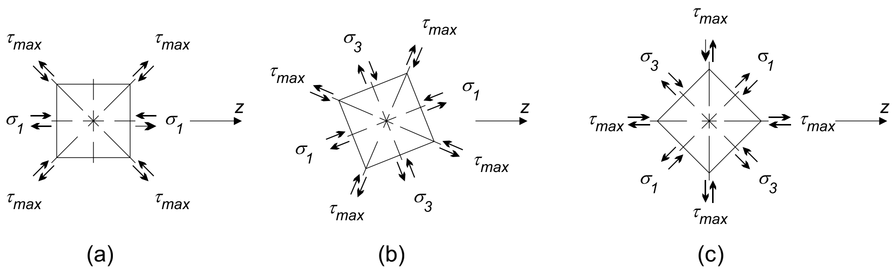

Figure 3.66 Principal stresses and maximal shear stresses at surface points of the cylindrical specimen loaded under: (a) pure bending, (b) in-phase combined bending-torsion ($\tau_a = \sigma_a$), and (c) pure torsion. The z-axis is parallel to the specimen axis

Let us finally emphasize that, in general, the stresses in planes not too much inclined at those in Figure 3.66 are very similar in magnitude. In the case of uniaxial tension, for example, the stresses in planes inclined at 10° are reduced by only 3% [375]. This is particularly important with respect to the crystallographic nature of short crack propagation in fatigue.

Stress-based Criteria of Multiaxial Fatigue Life

The experimental determination of fatigue life curves for complicated multiaxial loading cases represents, as a rule, a very difficult and expensive task. Therefore, many criteria of multiaxial fatigue life based on data available from uniaxial push-pull, bending or torsion tests were developed during the second half of the last century. In general, these criteria can be categorized as based on stress, strain or energy [375]. The stress-based criteria are appropriate for the large class of materials and components that operate near or below the fatigue limit (high-cycle fatigue).

The stress-based criteria can be written in the general form

$$af(\tau_a) + bg(\sigma_a, \sigma_n, \sigma_m) = 1, \tag{3.34}$$

where a and b are parameters that can be determined from two independent simple fatigue tests and $f(\tau_a)$ and $g(\sigma_a, \sigma_n, \sigma_m)$ are functions of applied shear and normal stresses, respectively. In the case of combined bending-torsion loading, for example, the fatigue limits in fully reversed pure torsion t_{-1} and pure bending f_{-1}, corresponding to a prescribed number of cycles to failure, are preferably used to obtain parameters a and b. In order to assess the efficiency of a particular criterion, the error index I is widely utilized as

$$I = 100(\mathrm{LHS} - 1) \quad [\%],$$

where LHS means the left-hand side of Equation 3.34. The ideal prediction leads to $I = 0$. If $I > 0$, the criterion gives conservative (safe) results since it predicts the specimen (or component) failure under lower loads than observed in the experiment. In other words, the applied stress amplitudes τ_a and σ_a lead to higher than prescribed number of cycles to failure.

The oldest criterion was suggested by Gough and Pollard [376, 377]. This criterion is based on the Huber–Hencky–Mises equivalent stress and its simplest form can be written as

$$a\tau_a^2 + b\sigma_a^2 = 1,$$

where $a = 1/t_{-1}^2$ and $b = 1/f_{-1}^2$.

Another classical empirical criterion, as proposed by Sines [378], uses the amplitude of the second stress invariant $J_{2,a}$ and the mean hydrostatic stress σ_m:

$$a\sqrt{J_{2,a}} + b\sigma_m = 1.$$

In the case of bending-torsion fatigue, $a = 1/t_{-1}$ and $b = \sqrt{3}f_{-1}/(\sigma_u t_{-1})$, where σ_u is the ultimate tensile strength. The criterion reported by Crossland [379] is of the same form while $b = 3/f_{-1} - \sqrt{3}/t_{-1}$ and $\sigma_m \to \sigma_{ma}$, where σ_{ma} is the amplitude of the mean hydrostatic stress.

More sophisticated criteria suggested by Findley, McDiarmid, Matake, Dang Van and Spagnoli are based on identification of the stress acting on specific planes within the component bulk. These planes are termed critical planes and are defined as one or more planes subjected to maximum damage. This concept respects the fact that the fatigue crack initiates in slip systems of maximal shear stresses and, in the short crack stage, it propagates along these planes with an enhancing support of normal stresses that open the crack tip. Particularly in smooth specimens or components, these stages occupy a majority of the fatigue life [247]. Thus, the fatigue life is controlled by the combination of shear and normal stresses acting on a critical plane. Findley [380] suggested that the maximal normal stress σ_n (the sum of the mean stress σ_m and the stress amplitude σ_a) on the critical plane might have a linear influence on the allowable alternating shear stress τ_a:

$$(a\tau_a + b\sigma_n)_{\max} = 1, \tag{3.35}$$

where $a = 1/\sigma_a$ and $b = k/\sigma_a$. The damage leading to failure is expected to be related to the critical plane that is associated with the largest term in the left-hand side of Equation 3.35. The constant k must be determined experimentally by performing tests under pure bending and pure torsion. In the case of in-phase proportional bending-torsion loading, a simple relation $\tan 2/\Theta = k^{-1}$ incorporates the angle Θ between the direction of the first principal stress σ_1 and the critical plane [375]. However, the Findley criterion incorrectly predicts a dependence of the torsion fatigue limit when the mean torsion stress is superimposed [381].

In the McDiarmid criterion [382], the critical plane is defined as the plane on which the amplitude of the shear stress reaches its maximum (not the plane on which the damage quantity is maximized). The torsion fatigue strength $t_{A,B}$ is introduced either for case A (cracking parallel to the surface) or case B (cracking inwards from the surface). The bending-torsion loading refers to case A and the McDiarmid criterion becomes

$$\frac{\tau_{a,\max}}{t_{-1}} + \frac{\sigma_n}{2\sigma_u} = 1.$$

The Matake criterion [363] can be formally expressed in the form of the Findley relation at Equation 3.35. However, the critical plane is assumed to be the same as that in the McDiarmid criterion (the maximum shear stress amplitude) which leads to $a = 1/t_{-1}$ and $b = 2/f_{-1} - 1/t_{-1}$ (see also Equation 3.36).

The criterion of Carpinteri and Spagnoli [383] is an example of a quadratic form

$$a\tau_{a,\max}^2 + b\sigma_n^2 = 1,$$

where $a = 1/t_{-1}^2$ and $b = 1/f_{-1}^2$.

Criteria of Dang Van [384], Papadopoulos (integral approach) [385] and Gonçalves *et al.* [386] are examples of advanced approaches. These criteria also take, besides the critical plane, the distribution of local microstructural stresses and possible slip systems into account. Dang Van has proposed an endurance limit criterion based on the concept of microstresses within a critical volume of the material. This criterion can be written as a time-dependent cumulative combination of the local shear stress $\tau_L(t)$ and the local hydrostatic stress $\sigma_{h,L}(t)$:

$$d = \max\left[\frac{\tau_L(t)}{b - a\sigma_{h,L}(t)}\right].$$

The critical damage is reached when the condition $d \geq 1$ becomes fulfilled.

In the Papadopoulos criterion both the shear stress and the normal stress are integrated over all slip planes:

$$\sqrt{\frac{5\kappa^2}{8\pi^2}\int\limits_{\varphi=0}^{2\pi}\int\limits_{\psi=0}^{\pi}\int\limits_{\chi=0}^{2\pi}\left(T_a\left(\varphi,\psi,\chi\right)\right)^2 \mathrm{d}\chi \sin\psi \mathrm{d}\psi \mathrm{d}\varphi + \left(3-\sqrt{3}\kappa\right)\cdot\sigma_{h,\max} \leq f_{-1}},$$

where T_a is the amplitude of resolved stress, φ, ψ and χ are the Euler angles and $\sigma_{h,\max}$ is the maximum value of the mean hydrostatic stress.

The Gonçalves criterion is expressed as

$$\frac{\kappa-1}{\sqrt{2}\left(1-\frac{1}{\sqrt{3}}\right)}\sqrt{\sum_{i=1}^{5} d_i} + \frac{\sqrt{3}-\kappa}{\sqrt{3}-1}\sigma_{1,\max} \leq f_{-1},$$

where the parameters d_i can be determined from minimum and maximum values of the deviatoric stress tensor as

$$d_i = \frac{1}{2}\left(\max\, s_i(t) - \min\, s_i(t)\right).$$

Fatigue Life of Steel Specimens under Bending-torsion

Comparison of prediction efficiency of individual multiaxial criteria in the case of combined bending-torsion fatigue of smooth specimens was reported in the work of Major *et al.* [387]. One set of specimens was subjected to a nitriding procedure, a technological treatment that results in a harder surface layer containing compressive residual stresses. Hereafter, these specimens will be called "nitrided" ones unlike the "virgin" ones without the surface layer.

The cylindrical specimens were made of high-strength low-alloy Cr-Al-Mo steel (equivalent to EN 37CrAlMo6) of the chemical composition given in Table 3.7. When compared with commonly used nitrided steels, the investigated material contains a slightly higher percentage of carbon, which offers higher strength but, at the same time, it introduces larger differences in microstructure caused by local variations in the nitriding process. After the heat treatment consisting of annealing (920°C, 25 min, air), quenching (930°C, 25 min, oil), and tempering (650°C, 25 min, air), the microstructure consisted of the sorbitic phase. The heat treatment resulted in yield strength $\sigma_y = 840\,\mathrm{MPa}$ and ultimate tensile strength $\sigma_u = 950\,\mathrm{MPa}$.

Table 3.7 Chemical composition of EN 37CrAlMo6 steel (wt%)

C	Mn	Cr	Mo	V	Cu	Al	W	Si	P	S
0.357	0.468	1.49	0.194	0.01	0.072	1.4	0.032	0.292	0.006	0.006

Nitriding is a prominent industrial technology that is used to enhance key properties of engineering components, especially surface hardness, fatigue strength and wear and corrosion resistance [388–390]. The plasma nitriding process is characterized by adsorption of nitrogen in the form of N^+, NH^+, NH^{2+} and NH^{3+} ions. The most important characteristics determining the depth of the nitrided layer are the nitriding time and the nitriding temperature. The nitrided layer consists of (1) the thicker subsurface diffusion layer and (2) the thin brittle surface compound layer composed of different iron nitride phases (often called the white layer). The higher strength of the nitrided layer together with associated compressive residual stresses causes subsurface crack nucleation and, therefore, improvement of fatigue strength. The micropulse plasma nitriding procedure was applied in two steps – cleaning (30 min) and nitriding (8 h) resulting in the depth of the diffusion layer $h_{dl} = 200\,\mu\mathrm{m}$ and the thickness of the white layer $h_{wl} = 3\,\mu\mathrm{m}$; see Table 3.8 for the details of this procedure. The basic tensile mechanical properties of the plasma nitrided specimens were measured as follows: yield strength $\sigma_y = 870\,\mathrm{MPa}$ and ultimate strength $\sigma_u = 1020\,\mathrm{MPa}$. This means that a slight improvement of both characteristics was achieved.

Table 3.8 Parameters of the nitriding process

Step	Temperature	Time	Atmosphere			Pressure	U	Pulse
	[°C]	[h]	N_2	H_2	CH_4	[mbar]	[V]	[µs]
Cleaning	510	0:30	20	2	-	0.7	800	120
Nitriding	515	32:00	21	7	0.4	2.6	530	150

The fatigue life of nitrided specimens was found to be significantly longer than that of virgin specimens, as can be clearly seen in Figure 3.67. The curves are plotted by using the McDiarmid criterion for the number of cycles

Table 3.9 Average error indices I_{avr} and their absolute values $I_{ABS,avr}$ of selected multiaxial criteria for virgin and nitrided specimens

Criterion	Specimens Indices [%]	Virgin $I_{ABS,avr}$	Virgin I_{avr}	Nitrided $I_{ABS,avr}$	Nitrided I_{avr}
Dang Van		7.35	−3.03	8.45	−0.35
Crossland		7.28	−5.42	8.65	−2.47
Sines		11.85	−11.23	12.32	12.82
McDiarmid		7.08	−3.02	8.65	1.72
Findley		7.08	−3.38	8.62	−0.87
Matake		7.12	−3.04	8.32	−0.23
Spagnoli		10.5	4.33	8.56	3.35
Papadopoulos (integral approach)		7.36	−3.04	8.46	−0.35
Papadopoulos (critical plane)		11.08	−6.30	16.56	−6.70
Gonçalves *et al.*		8.62	4.14	13.92	5.20

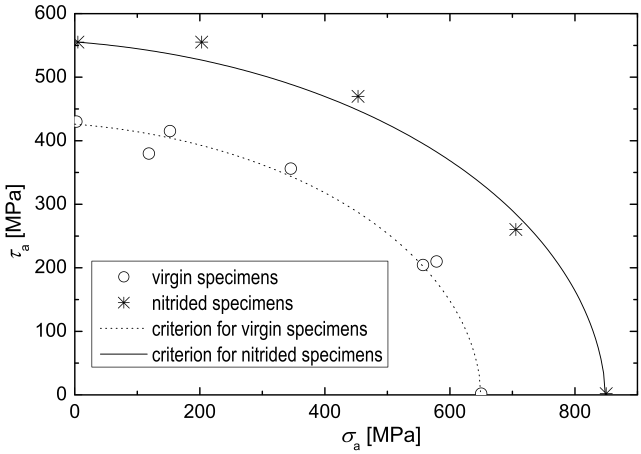

Figure 3.67 The constant fatigue life diagram ($N_f = 5 \times 10^5$ cycles) for bending-torsion loading of virgin and nitrided specimens according to the McDiarmid criterion

$N_f = 5 \times 10^5$ and the experimental data correspond to a fatigue life of $N_f = 5 \times 10^5 \pm 2 \times 10^5$. This conclusion also holds for push-pull fatigue tests reported elsewhere [390].

The error indices, calculated for selected multiaxial criteria by averaging the experimental data, are shown in Table 3.9. They reveal that the McDiarmid criterion was the most successful in fatigue life prediction for virgin specimens whereas the Matake criterion was the best for nitrided ones, although both criteria provide slightly non-conservative results. Nevertheless, the related errors are less than 10% which is safely below the safety factors commonly utilized in fatigue strength analyses applied in engineering practice.

3.3.4.2 Topology of Bending-torsion Fracture Surfaces

Cyclic torsion loading leads to a very complicated crack path exhibiting local crack arrests, a branch/twist crack morphology or the factory roof [347,361,375,391,392]. An interaction between both mating fracture surfaces (some combination of sliding, climbing, sticking, slipping and deforming) often makes even a qualitative understanding of fatigue crack propagation difficult [347,375]. On the other hand, a high amount of opening loading mode I or sometimes high shear amplitudes generate a fracture surface that appears macroscopically flat. Despite all mentioned studies, the crucial problem in biaxial fatigue topography is still a significant lack of experimental data. This shortcoming could be partially reduced by a study devoted to the fracture surface topology formed by a combined bending-torsion loading in both the low-cycle fatigue (LCF) and the high-cycle fatigue (HCF) regimes [393].

Experimental Procedure and Results

Experimental settings were based on Matake's critical plane criterion [381, 383]. According to fatigue life N_f, the virgin cylindrical specimens made of the EN 37CrAlMo6 steel were divided into LCF and HCF, respectively. The differences in N_f within both LCF ($N_f \approx 10^4$) and HCF ($N_f \approx 10^6$) sets were well within an order of magnitude, which is a scatter typical for fatigue life data.

Topological 3D data of 0.5×0.5 mm regions, selected on the fracture surfaces at a distance of 0.8 mm from the specimen surface (the crack initiation site), were obtained by means of stereophotogrammetry. Using the Delaunay triangulation, two sets of profiles were traced for all analyzed regions: the first in the crack propagation direction (referred to in the following as the y-direction) and the second in the perpendicular direction (referred to as the x-direction) thus representing different positions of the progressing fatigue crack front.

The following parameters were found to be sufficiently sensitive to different topological aspects of the extracted fracture surface profiles, corresponding to a different loading mixture: the standard deviation R_q of vertical z-coordinates, the number of peaks of the profile per unit length m_0 and the Hurst exponent H. It was expected that these parameters might be used for a wide range of fracture surfaces of engineering components made of materials with different microstructures. It should be emphasized that the arithmetic roughness R_a exhibited similar trends as R_q, and the fractal dimension D_D did not reveal any significant susceptibility to differentiate the loading cases.

Average values of analyzed parameters, together with their standard deviations (error bars), are plotted in Figure 3.68 for profiles oriented in the

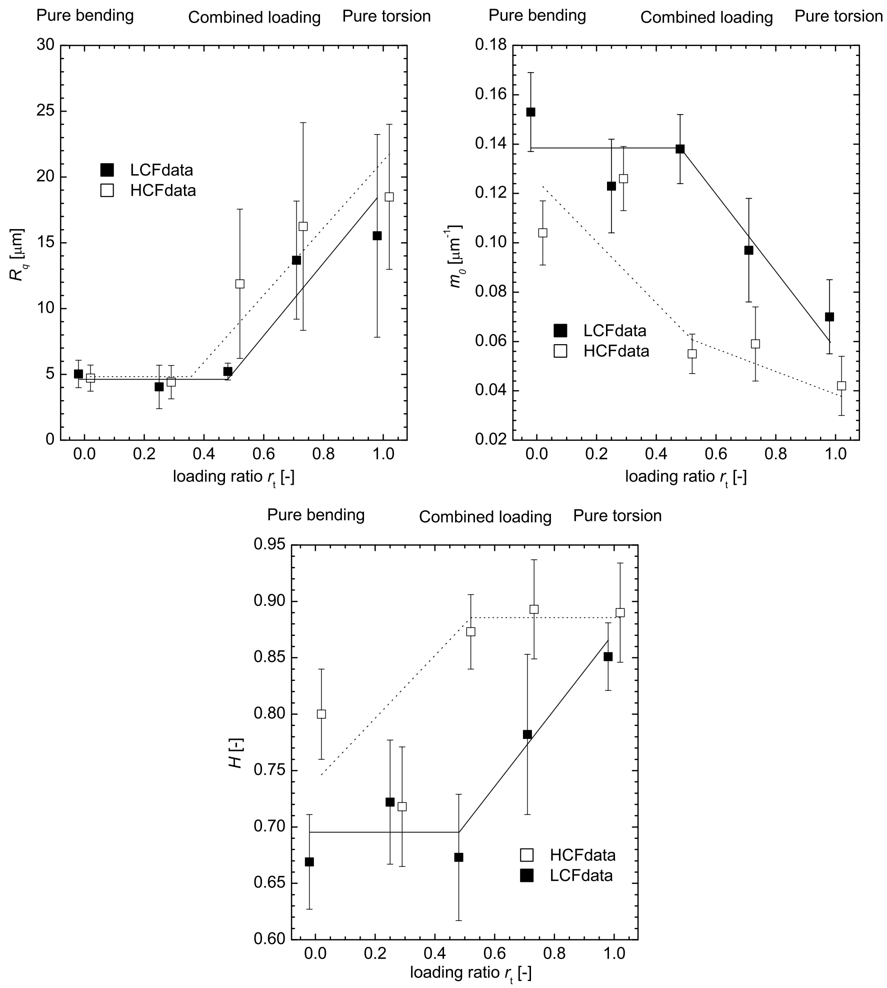

Figure 3.68 Averaged values of topological parameters as functions of the loading ratio r_t. Reprinted with permission from Elsevier B.V. (see page 265)

y-direction as functions of the loading ratio r_t. This ratio is defined as

$$r_t = \tau_a/(\tau_a + \sigma_a).$$

Discussion of Results

Since the parameters R_q and R_a describe both macroscopic and microscopic levels of the surface roughness, it is not surprising that lower values $R_q < 10\mu\text{m}$ correspond to low torsion components of the loading ($r_t < 0.5$) in both LCF and HCF. Fracture surfaces corresponding to a higher torsion component ($r_t > 0.5$) exhibited substantially higher R_q values.

Calculated values of the parameter m_0 decrease with an increase in the torsion component which is likely associated with frictional contact of both fracture surfaces during the fatigue process. Indeed, the contact friction must be more intensive for loading regimes of high r_t. Generally higher values of m_0 for LCF can be understood in terms of a higher roughness level. On the other hand, the values of H increase with increasing loading ratio r_t. Moreover, they are generally higher in the case of HCF. This behaviour, the opposite of that of m_0, can also be attributed to the frictional contact of fracture surfaces.

The most important result lies in the existence of a critical value of the loading ratio $r_t \approx 0.5$ that corresponds to a significant change of all examined topological characteristics. This means that the topography of analyzed fracture surfaces substantially changes when the torsion component becomes higher than the bending one. This result might be very important for the assessment of the loading type in failure analysis, provided that it remains true for other metallic materials.

Let us finally note that the experimental results obtained for the x-direction were similar to those for the y-direction. The only additional trend, observed only in the case of LCF, was a slight decrease (increase) in m_0 (H) with measured distance. In terms of the Hurst exponent, for example, this means that the fractality of the crack front decreases during its propagation. Such behaviour might be elucidated by a gradual merging of many initial surface microcracks that nucleate in the case of LCF.

3.3.5 Formation of Fish-eye Cracks under Combined Bending-torsion Loading

The high strength of the nitrided layer and the compressive residual stresses introduced within the diffusion zone are the main causes of an improvement of fatigue strength. Since such a layer hinders dislocation motion, the predominant failure mechanism in the high-cycle fatigue (HCF) region is subsurface fatigue crack growth (e.g., [394,395]). As a rule, the cracks initiate at internal inclusions within the core region and propagate in a near vacuum by forming so-called fish-eye cracks. The interior of these cracks looks bright to the naked eye or in the optical microscope, whereas the outside region seems to be grey. The difference in colour is, most probably, caused by different fracture micromechanisms. The bright morphology of the fish-eye crack is produced by a cyclic contact of the mating fracture surfaces under conditions of suppressed atmospheric effects. When the crack front approaches the low-toughness nitrided layer, local through-the-layer brittle cracking creates a connection to the surface and, subsequently, a penetration of the atmosphere to the inside of the fish-eye [394]. This changes the growth mechanism to an environmentally assisted one which leads to a different surface roughness and to the optical contrast.

Fish-eye cracks are also a typical feature of fractures in the ultra-high-cycle fatigue (UHCF) region [396–398]. The transition of the crack initiation site from the surface into the interior causes the stepped, double or multi-stage fatigue curve appearance. The initiation sites in the UHCF region are often non-metallic inclusions, although the crack nucleation near finely dispersed segregations or phase interfaces was sometimes observed. In the former case, the dark-looking fracture morphology in the vicinity of the inclusion, termed as the "optically dark area" (ODA), is usually present. The formation of the ODA is, most probably, attributed to slow intermittent fatigue crack growth assisted by the internal hydrogen trapped by the inclusion [399].

When the maximum tensile loading exceeds the yield strength of the material, the low-toughness nitrided layer breaks during the first few loading cycles due to a high strain mismatch at the layer/matrix interface [390]. Such initiated surface cracks can easily penetrate to the specimen bulk owing to the related hard/soft transition direction [198]. Thus, this damage mechanism is typical for the quasi-static region.

The geometric characteristics of the fish-eye cracks were, to some extent, studied by several researchers (e.g., [388, 389, 396, 400]). In general, the fish-eye cracks were found to be approximately of either a circular or an elliptical shape depending on the type of loading (push-pull, rotating bending, plane bending) and the presence of high compressive residual stresses introduced by some of the surface hardening procedures. A restricted propagation of the crack towards the nitrided layer is considered to be the reason why there is an increase of the fish-eye crack radial size with increasing distance of the inclusion from the specimen surface, which in turn is reported to be a function of the level of applied stress inside the nitrided layer and the compressive residual stress profile. In the UHCF region, the ODA is of particular interest due to the still insufficiently understood interplay of cyclic strain and internal hydrogen. The size of the ODA was found to increase with decreasing stress level or increasing fatigue life.

The residual stress effect on the crack initiation site under high-cycle push-pull loading was studied in [390]. The size of the inclusion required for fatigue crack initiation was computed as a function of inclusion depth. The local shift of the loading asymmetry caused by the presence of a compressive residual stress field was accounted for by using an equation describing the dependence of the fatigue threshold ΔK_{th} on the cyclic ratio R [401]. For the investigated nitrided steel, this dependence was determined by using precracked flat specimens with nitrided layers of the thickness of approximately a quarter of the specimen width and, therefore, the measured fatigue thresholds represented averaged layer/core values. When identifying the inclusion with a critical crack at the fatigue limit, the critical inclusion size d_c could be estimated as $d_{c1} = 3.79\,\mathrm{mm}$ near the specimen surface, $d_{c2} = 0.87\,\mathrm{mm}$

at depth $h = 0.3\,\mathrm{mm}$ inside the diffusion zone and $d_{c3} = 0.068\,\mathrm{mm}$ at depth $h = 0.4\,\mathrm{mm}$ [390] at the outer layer/matrix boundary. In spite of the fact that the averaged values of ΔK_{th} might have been inaccurate, the calculated critical size of the inclusion near the surface exceeded by two orders of magnitude the inclusion size of tens of microns found in the investigated steel as well as in most commercial steels. This is a clear reason why there was no crack initiation observed at (or near) the surface.

All the above-mentioned studies were, however, conducted on fish-eye cracks generated by the uniaxial loading regime or torsion and no information is currently available for the case of multiaxial loading. Therefore, the motivation for further research was to study the fish-eye cracks developing under symmetrical bending, symmetrical torsion and their synchronous in-phase combinations. The results of such an investigation [402] are presented in the following subsection.

3.3.5.1 Experimental

The specimens were made of high-strength low-alloy Cr-Al-Mo steel specified in the previous subsection. In the HCF regime, both the compressive residual stresses and the hardness of the nitrided layer induce the internal fatigue fish-eye crack initiation at the inclusion-matrix interface and influence a subsequent crack propagation. The dependence of the residual stress level inside the layer on the distance from the free surface is depicted in Figure 3.69. This dependence was obtained by means of precise X-ray measurement performed on samples made of an equivalent high-strength steel that was nitrided using the technology described above [403]. One can see that the compressive residual stresses reach a maximum of 750 MPa close to the surface and, up to about 0.2 mm depth, they start to decay steeply. Beyond the depth of 0.2 mm corresponding to the microstructurally distinguishable diffusion layer, a rather slow decrease of residual stresses continues to reach a zero value at 0.7 mm. The maximum value of the residual compressive stress of about 800 MPa was also confirmed by less precise measurement on the investigated nitrided steel [404].

Plasma nitrided specimens were tested at room temperature using a Polish resonance testing machine MZGS-200. The sinusoidal symmetrical bending and torsion as well as the synchronous in-phase bending-torsion were applied at a frequency $f = 29\,\mathrm{Hz}$. Experimental settings were based on the Matake critical plane criterion

$$\tau_a + \left(\frac{2\tau_c}{f_{-1}} - 1\right)\sigma_n = t_{-1}. \tag{3.36}$$

This criterion was used to estimate both bending and torsion loading components in order to constrain the fatigue life into the HCF region. The details of experimental settings and fatigue life data are published elsewhere [393].

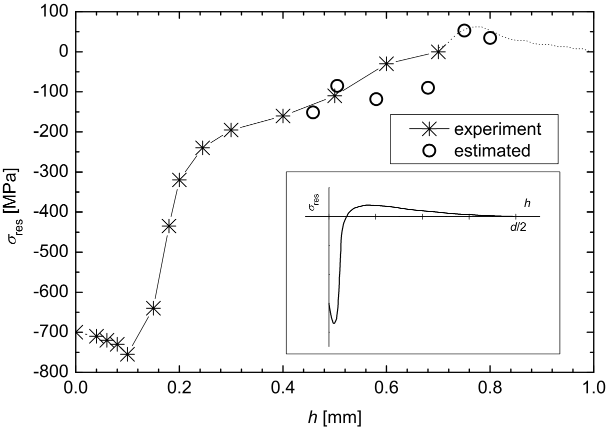

Figure 3.69 The experimental profile and the theoretical estimation of residual stresses in the nitrided layer. The scheme in the *inset* shows a general shape of the residual stress curve along the specimen radius. Reprinted with permission from Elsevier B.V. (see page 265)

3.3.5.2 Fish-eye Crack Geometry

Fish-eye cracks were studied using a chromatic optical profilometer FRT MicroProf 100, Fries Research & Technology GmbH. In the first step, the distance h of the centre of the inclusion from the free specimen surface in the radial direction was measured. After that, the maximal projection plane was found by tilting the specimen mounted on the moving x-y table and fish-eye crack sizes in radial (R_{tl}, R_{tr}) and tangential (R_{ru}, R_{rd}) directions were recorded; see Figure 3.70 for the nomenclature. While the inclination of the crack plane was almost negligible in the radial direction, the inclination angles in the tangential direction varied in the range of $(0°, 45°)$, where $0°$ nearly corresponds to pure bending ($r_t = 0$) and $45°$ approximately stands for pure torsion ($r_t = 1$). This actually means that in all cases of non-zero torsion component the crack front propagates under the local mode I loading. In several cases, the diameter of the inclusion was also estimated in the scanning electron microscope (SEM) by analyzing either the inclusion itself or its imprint found on the fracture surface. The average diameter was found to be $37 \pm 14\,\mu$m.

The observed fish-eye cracks were, in general, of an elliptical shape; see Figure 3.71. In the case of fish-eyes initiated close to the nitrided layer, the upper radial dimension R_{ru} (crack propagation towards the free specimen surface) was found to be smaller than the lower radial dimension R_{rd} (crack growth into the specimen core). This behaviour, which is understandable in terms of the retardation effect of the surface nitrided layer, is consistent with observations reported in the work of DelaCruz *et al.* [388] where, however, only fish-eye cracks generated under reversed bending were studied.

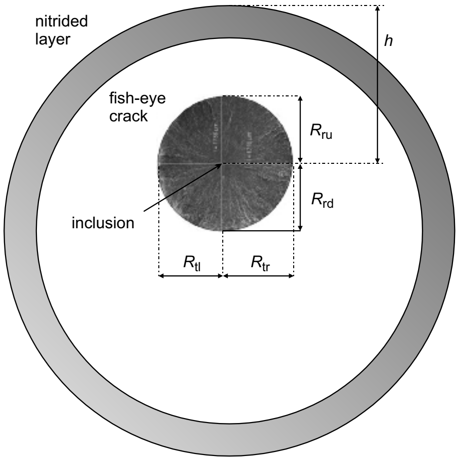

Figure 3.70 A scheme showing the nomenclature for the fish-eye crack characterization

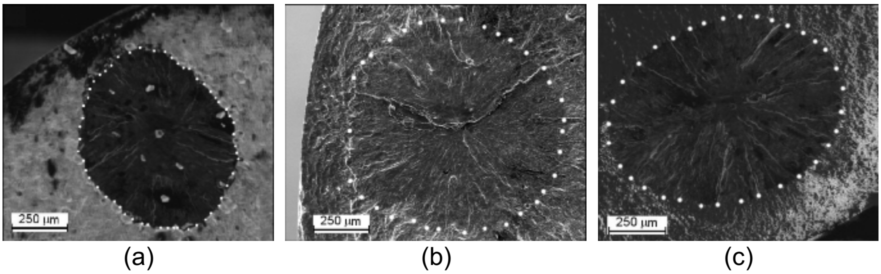

Figure 3.71 An example of the SEM microphotographs of fish-eye cracks for different loading regimes: (a) $r_t = 0$, (b) $r_t = 0.5$, and (c) $r_t = 1$. Reprinted with permission from Elsevier B.V. (see page 265)

The fish-eye crack shape was characterized by the radial asymmetry, S_r, the average size of the fish-eye crack (free of the influence of the nitrided layer), R_{avr}, and the elliptical coefficient Q. These quantities were defined as

$$S_{\mathrm{r}} = \frac{R_{\mathrm{ru}}}{R_{\mathrm{rd}}},$$

$$R_{\mathrm{avr}} = \frac{1}{4}\left(2R_{\mathrm{rd}} + R_{\mathrm{tl}} + R_{\mathrm{tr}}\right),$$

$$Q = \frac{R_{\mathrm{tl}} + R_{\mathrm{tr}}}{2R_{\mathrm{rd}}}.$$

All of the fish-eye cracks studied were nucleated at internal non-metallic inclusions. The distance h of the inclusion centre from the specimen free surface was in the range of $420 - 1040\,\mu\mathrm{m}$. The crack growth towards the free surface was always, at least partially, influenced by the compressive residual stress field ($h - R_{ru} < 0.6$ mm). Moreover, some fish-eye cracks (5 out of 26) grew into the compound layer ($h - R_{ru} < 0.2$ mm).

The effect of the free surface and the bending stress gradient on the fish-eye crack growth in the radial direction without the influence of the nitrided layer was analyzed using the analytical expressions for K_{IA} and K_{IB} at two opposite radial crack fronts A and B of the elliptical crack in a shaft under bending. The results revealed that, almost independently of the depth h in the range of (0.6, 1.2) mm, the radial asymmetry, S_r, was found to be in the range of (1.20, 1.22) [402]. Note that the influence of the nitrided zone on fish-eye cracks initiating outside the range of compressive stresses ($h > 0.7$ mm) must be nearly negligible. Because the stress gradient is also similar in the case of torsion loading, most of the measured values of S_r for the fish-eye cracks initiated at the depth $h > 0.7$ mm should lie close to $S_r = 1.2$ for all loading regimes.

3.3.5.3 Results

The dependence of the radial asymmetry, S_r, on the distance of the inclusion from the specimen surface, h, is shown in Figure 3.72. As mentioned above, the fatigue crack growth in the radial direction is influenced by bending-torsion loading component gradients, the free surface and the residual compressive stresses, which are prominent especially for the fish-eye cracks initiated near the nitrided layer. Obviously, when the inclusion depth is not very large ($h < 0.7$ mm), the residual compressive stresses decelerate the crack front growing towards the free surface as evidenced by the radial asymmetry values $S_r < 1$. On the other hand, as was predicted by the growth simulation, the radial asymmetry reaches $S_r = 1.2$ in the range $h \in (0.7, 0.8)$ mm. With an increasing inclusion depth $h > 0.8$ mm, however, the general tendency seems to be the growth in a rather symmetrical fashion $S_r \approx 1$. This might be a consequence of small tensile residual stresses that must balance the compressive ones beyond the depth $h = 0.7$ mm (see also hereafter).

An increase in the average fish-eye size, R_{avr}, with increasing inclusion depth and fatigue life, N_f, is depicted in Figure 3.73. The first tendency is in agreement with the result reported in [388] and follows from a simple geometrical consideration. Indeed, the fish-eye dimension R_{ru} is, more or less, determined by the distance from the initiation site to the nitrided layer. The second dependence can be understood in terms of the increasing equivalent stress amplitude with the decreasing number of cycles to failure N_f.

The elliptical coefficient, Q, exhibits no correlation with the number of cycles to failure (Figure 3.74(a)) in accordance with the result reported in [400]. On the other hand, the loading regime has a certain effect, as is demonstrated in Figure 3.74(b). Despite a large scatter in the data, the elliptical coefficient increases with increasing loading ratio (the regression line). In Figure 3.75, a

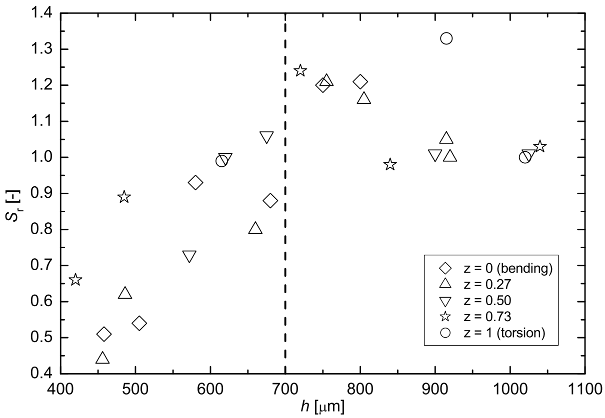

Figure 3.72 The dependence of the radial asymmetry on the inclusion depth. Reprinted with permission from Elsevier B.V. (see page 265)

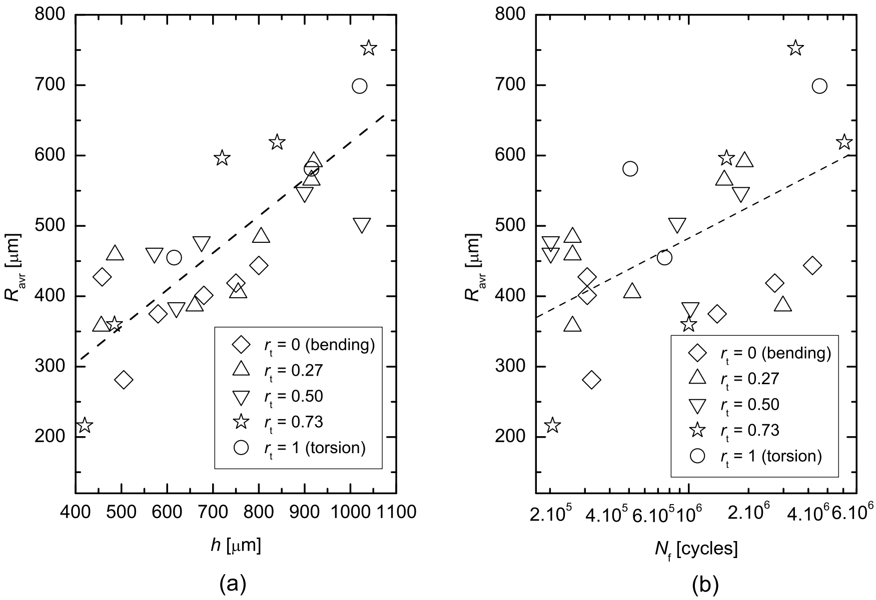

Figure 3.73 The dependence of average fish-eye crack size on: (a) the inclusion depth, and (b) the fatigue life

corrected coefficient $Q' = (R_{tl} + R_{tr})/(2R_{rd})$, where the lower radial dimension R_{rd} (free of the influence of the nitrided layer) replaced the upper radial dimension R_{ru}. This correction eliminates the effect of the nitrided layer on the fish-eye crack growth, diminishes the data scatter and accentuates the trend observed in Figure 3.74b. The dependence Q *vs* r_t means that with a higher portion of the torsion loading component the crack growth rate in the tangential direction becomes more rapid than that in the radial direction. This effect is, most probably, caused by a different influence of the stress gradients on the crack growth in tangential and radial directions.

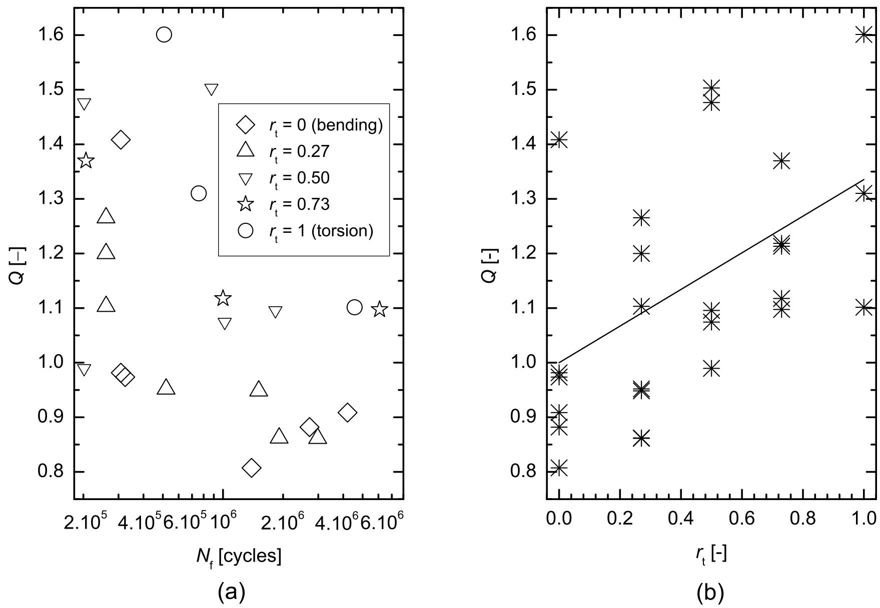

Figure 3.74 The dependence of the elliptical coefficient on: (a) the fatigue life, and (b) the loading ratio

3.3.5.4 Estimation of Residual Stresses

The experimental data obtained under pure bending could be used for an approximate assessment of the residual stresses in the innermost part of the nitrided layer and the adjacent core region. In the pure bending case, the direction of internal stresses is equal to that of the principal stress. The orientation of the residual stress vector is opposite to that of the main stress during the tensile loading half-cycle and identical to it during the compressive half-cycle. This means that the residual stress σ_{res} shifts the S-N (Wöhler) curve for a symmetric loading ($R = -1$) of the virgin material (without

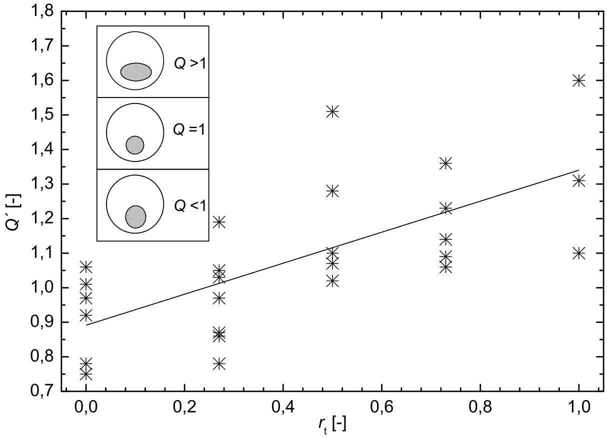

Figure 3.75 The dependence of the corrected elliptical coefficient Q' on the loading ratio

the nitrided layer) to a curve for a lower cyclic ratio $R < -1$. This fact is employed in the proposed method described as follows:

1. measurement of the initiation depth h_k of k-th fish-eye;
2. calculation of the amplitude of the bending stress σ_{ak} that corresponds to the depth h_k;
3. identification of the number of cycles to failure N_{fk} corresponding to the k-th bending experiment;
4. determination of the bending amplitude σ_{avk} (the surface stress amplitude) corresponding to N_{fk} on the Wöhler curve for virgin specimens;
5. utilization of both the ratio σ_{ak}/σ_{avk} and the relationship generally describing the shift of S-N curves to a determination of the mean stress $\sigma_m = \sigma_{res}$ corresponding to the depth h_k;
6. application of this procedure to fish-eyes of various depths h in order to obtain the dependence $\sigma_{res}(h)$.

This procedure was applied to obtain an approximate dependence $\sigma_{res}(h)$ for comparison with the measured values in Figure 3.69. In Figure 3.76, the S-N curve for virgin specimens [390] is plotted by a solid line along with that for nitrided specimens (the dashed line). The solid stars refer to the calculated stresses σ_{ak} at the depths h_k. For depths $h_k < 0.7$ mm, these points lie in between the S-N curves for virgin and nitrided specimens while, for $h_k > 0.7$ mm, they lie slightly below the virgin S-N curve. The well known Soderberg ($\sigma_a/\sigma_{av} = (1 - \sigma_m/\sigma_y)$) and Goodman ($\sigma_a/\sigma_{av} = (1 - \sigma_m/\sigma_u)$) relations for the S-N curve shift were adopted for calculation of residual stresses. All the experimental and computed data are displayed in Table 3.10. Since the Soderberg relation generally gives conservative results unlike the Goodman approximation that provides nonconservative estimates [192], the mean values of both approaches are plotted in Figure 3.69 (open symbols) as the theoretical assessment of residual stresses. One can see a rather good agreement between the calculated and measured data within the innermost part of the nitrided layer. In the adjacent core region, the computed data predicted small tensile residual stresses. This is to be expected owing to a necessary balance of residual stresses through the specimen radius $d/2$. Indeed, the integral of the residual stresses along the radius must be zero, as is schematically shown in the inset of Figure 3.69.

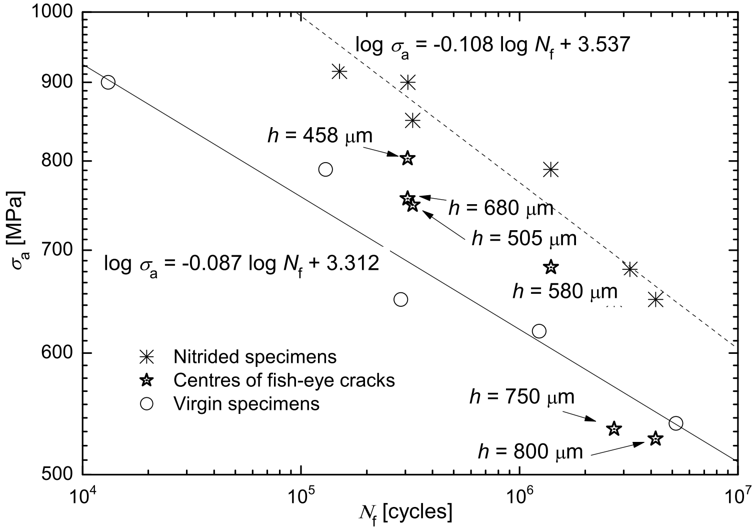

Figure 3.76 Wöhler curves for virgin and nitrided specimens and the data points corresponding to the analyzed fish-eye cracks. Reprinted with permission from Elsevier B.V. (see page 265)

Table 3.10 Results of the residual stress analysis

N_f [cycles]	h_k [µm]	σ_{ak} [MPa]	σ_{avk} [MPa]	σ_m [MPa] Soderberg	σ_m [MPa] Goodman
4.20×10^6	800	527.6	548.0	31.2	35.3
2.71×10^6	750	535.3	569.2	50.0	56.5
1.39×10^6	580	682.2	602.9	−110.5	−125.0
3.24×10^5	505	749.0	684.0	−79.9	−90.3
3.08×10^5	680	756.0	687.0	−84.3	−95.4
3.08×10^5	458	803.0	687.0	−141.8	−160.4

Although the estimation of the residual stresses seems to be plausible, one should take some limitations of the proposed method into account. First, the fatigue resistance of the layer increases with decreasing depth h due to the hardening effect of the nitriding process. This means that the virgin S-N curve certainly does not apply for $h_k < 0.2\,\mathrm{mm}$, i.e., for the depth region corresponding to the hard diffusion zone. For the values $h_k > 0.3\,\mathrm{mm}$, however, the microstructurally induced hardening effect becomes rather negligible and the hardening component induced by residual stresses starts to dominate. Consequently, the applicability of the method is restricted to fish-eyes initiating within that region. Since the initiation depth decreases with increasing σ_a, the depth condition $h_k > 0.3\,\mathrm{mm}$ also implies a limitation of the applied loading. Another load limit is given by the yield stress σ_y due to the related breaking of the nitrided zone. On the other hand, a different geometry of the crack growth stage in the virgin and nitrided specimens does not represent any serious problem. Indeed, the number of cycles to failure in both types of specimens are determined by the crack initiation stages where the inclusion-assisted nucleation mechanisms are identical. However, this is not necessarily true for low-strength materials, where the crack initiation mechanism in virgin specimens could be different (persistent slip bands, grain boundaries, etc.). This is a further limitation of the proposed method. Let us finally mention that the data scatter of S-N curves for both the virgin and nitrided specimens does not cause a significant complication, as one can observe in Figure 3.76.

The main results of the fish-eye analysis can be summarized in the following points:

1. The nitrided layer reduces the crack growth rate. The retardation effect of the residual compressive stresses is prominent especially for the fish-eye cracks initiated on inclusions at depths $h < 0.7\,\mathrm{mm}$, as evidenced by the radial asymmetry $S_r < 1$.
2. The average size of the fish-eye crack increases with increasing distance from the free surface and the number of cycles to failure.
3. The elliptical coefficient Q increases with an increasing portion of the torsion loading component. The transient value of the loading ratio can be estimated to be $r_t \approx 0.4$.
4. An assessment of residual stresses can be made by a combination of fractographical and strength analyses related to the fish-eye centre. The applicability of this theoretical method is restricted to the innermost part of the nitrided layer and to the adjacent core region.

3.4 Failure Analysis

Linear–elastic fracture mechanics can be successfully applied in numerical procedures predicting the residual fatigue life of structural components containing cracks longer than about 0.5 mm. According to defect-tolerant design approaches to fatigue, the useful fatigue life is the number of cycles to propagate the largest undetected crack to an unstable fracture. The discovery of fatigue striations in the early 1960s has promoted a development of quantitative retrogressive methods that can be very useful for the reconstruction of conditions under which the failure process occurred.

In failure analysis, the re-estimation of the stress amplitude, the cyclic ratio and the number of cycles to failure is of main practical interest [246]. First attempts to assess those parameters from the fracture surface were published in the early 1960s [405, 406]. These simple approaches allowed the assessment of applied loading amplitude (assumed to be a constant) but the value of cyclic ratio had to be anticipated. Similar concepts were described in the early 1970s [407, 408]. Several years later two different methods enabling an assessment of both $\Delta\sigma$ and R were reported by Uchimoto *et al.* [409] and Pokluda and Staněk [410, 411]. The method proposed by the latter authors is applicable even to cases of stationary random loading. Nowadays, failure analysis can be supported by laboratory devices that are able to simulate the fatigue process on real components under realistic loading spectra. Methods based on overload markings or image analysis can better identify the function $s(\mathrm{d}a/\mathrm{d}N)$ that defines the difference between striation spacings and real crack growth rate during the whole period of stable crack propagation [412, 413]. However, these time consuming and expensive methods do not necessarily provide much more accurate results. The reason lies in a lack of knowledge about a precise fracture mechanical description of crack growth rate with respect to the complexity of both applied loading spectra and fatigue crack paths in structural components [192, 347].

As a useful example, the method [411] which was developed by one of the authors during his employment in the military research institute is presented here. This method was widely used in case studies in order to assess fatigue loading parameters as well as the number of cycles spent by a long fatigue crack during its propagation from a small initial size to the final fracture (e.g., [414, 415]. The original method was recently improved by utilizing a more sophisticated relation describing the whole range of the long crack growth rate [416]. The applicability of the method is demonstrated in a case study concerning the fatigue failure of a compressor blade.

3.4.1 Theoretical Background

As was shown in Section 3.1, both the subcritical (stage II) crack propagation and the unstable fracture are usually reflected on the fracture surface in a characteristic way. In the reconstruction procedure, the following assumptions are utilized: (1) an approximate 1:1 relation between the mean projected striation spacing $\bar{s}$ and the crack growth rate in the Paris–Erdogan region of stage II crack growth; (2) the morphological boundary between stable/unstable crack growth, associated with the the crack length a_c, can be simply related to the cyclic fracture toughness K_c. Moreover, the basic crack-growth rate equation proposed by Forman *et al.* [417] is used in a modified form allowing a better accommodation to the near-threshold region:

$$\frac{\mathrm{d}a}{\mathrm{d}N} = A\frac{\Delta K^n - \Delta K_{th}^n (1-R)^{nm}}{K_c (1-R) - \Delta K}, \tag{3.37}$$

where ΔK is the range of applied stress intensity factor, ΔK_{th} is the threshold factor for $R = -1$, R is the cyclic ratio and A, m, n are experimental constants. Equation 3.37 reduces to the Klesnil–Lukas relation [418] for $\Delta K_{th} \approx \Delta K \ll K_c$. As a rule, $n \in (2, 5)$ and $m \in (0.3, 0.5)$ for metallic materials. The left-hand side of Equation 3.37 may be replaced by the mean striation spacing $\bar{s}$ ($\mathrm{d}a/\mathrm{d}N \approx \bar{s}$) that should be measured closely in the range of $a/a_c \in (0.3, 0.6)$ on the fracture surface (usually corresponding to the Paris–Erdogan region). The value $\bar{s}$ should be determined according to the relation at Equation 3.4 at k points corresponding to the same crack length a on the fracture surface. It should be emphasized, however, that the limit $k = 1$ is also permissible which substantially extends the applicability of the method. For many reasons, indeed, one can find only a single facet covered by striations. On the other hand, the accuracy of such an assessment might be lower.

The moment of unstable fracture corresponds to the cyclic fracture toughness K_c and to the critical crack length $a = a_c$. Therefore, one can write

$$K_c = \frac{\Delta K(a_c)}{1-R}. \tag{3.38}$$

Since $\Delta K = \Delta\sigma\sqrt{\pi a} f(a, a/W)$ (W is the width of the component and $f(a, a/W)$ is the shape function), a numerical solution of Equations 3.37 and 3.38 gives the estimates of loading parameters $\Delta\sigma$ and R that caused failure of the structural component. It should be noted that comprehensive handbooks of stress intensity factors are available today (e.g., [419]).

Substituting the obtained values $\Delta\sigma$ and R into the relation

$$N_c = \int_0^{N_c} \mathrm{d}N = \frac{1}{A}\int_{a_0}^{a_c} \frac{K_c (1-R) - \Delta K}{\Delta K^n - \Delta K_{th}^n (1-R)^m}\, \mathrm{d}a, \tag{3.39}$$

one can calculate the number of cycles N_c, associated with the fatigue crack propagation in the range of $\langle a_0, a_c\rangle$, where a_0 is the initial crack length (usually a minimum detectable length of inspection methods).

For the method described above, nearly constant values of $\Delta\sigma$ and R during the fatigue process are assumed. In engineering practice, however, this kind of loading occurs rather rarely (only about 10%). Roughly 40% of applied loading cases are stationary or quasi-stationary spectra (highly variable ΔK and nearly constant R). In these cases, the striation spacings $\bar{s}_1$ and $\bar{s}_2$ for at least two different crack lengths a_1 and a_2 are to be measured on the fracture surface within the Paris–Erdogan region [411]. The reason lies in an uncertain stress range at the moment of the final fracture corresponding to the crack length a_c. The values $\Delta\sigma_{rms}$ and R can then be obtained by a simultaneous numerical solution of two Equations 3.37 for a_1 ($\bar{s}_1$) and a_2 ($\bar{s}_2$), where $\Delta\sigma_{rms}$ means the root mean square value of the applied spectrum loading. About 50% of all exploitation loading spectra exhibit a highly variable R, i.e., a non-stationary behaviour. In these cases the method presented becomes useless, similar to most other approaches.

Thus, the following steps should be performed to accomplish the failure analysis:

1. experimental determination of Equation 3.37 in the laboratory by using samples made of the material of the fractured component;
2. selection of appropriate shape function(s) $f(a, a/W)$ for the fractured component and crack lengths a and a_c;
3. measurement of $\bar{s}$ for the selected length(s) a on the fracture surface;
4. measurement of the critical length a_c on the fracture surface;
5. numerical solution of Equations 3.37 and 3.38 to estimate $\Delta\sigma$ and R;
6. anticipation of the shape function $f(a, a/W)$ in the whole range of (a_0, a_c);
7. numerical solution of Equation 3.39 to assess N_c.

3.4.2 Case Study

Apart from many other applications, this method was used to reconstitute conditions of a surprisingly quick fatigue failure of a compressor blade in an aircraft engine after a general repair. The engine was tested in a stand placed on the ground inside a special semi-natural cave. Blocks of a constant loading amplitude were applied by changing engine frequencies to simulate the service regime. The blade was made of the high-strength alloy Ti-4Al-3Mo-1.5Zr ($\sigma_y = 1060\,\mathrm{MPa}$). A careful investigation in the scanning electron microscope revealed that the fatigue crack was initiated at the leading edge of the blade. Some silicon micro-particles were found to be stuck in the blade surface along with many micro-craters near the crack initiation site. Thus, the crack initiation was a result of impacting silicon particles coming, most probably, from a dust whirled up from the ground.

The shape function for the cracked compressor blade can be written as

$$\begin{aligned} f(a) = {} & 0.56419(1 - 1.1 \times 10^{-4}a^3) \times \\ & \times (1.67687 - 0.43573a + 0.0819338a^2 - 0.0065158a^3 + 0.00018858a^4). \end{aligned}$$

This form was obtained by compliance measurements performed on an identically designed blade containing crack growing from the leading edge [420]. The influence of a variable blade width W is already respected by coefficients in the shape function (a in [mm]).

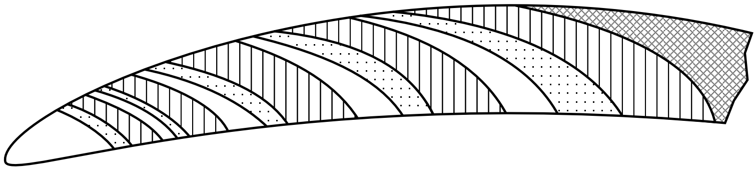

Figure 3.77 The scheme of growth curves marking different loading regimes on the fracture surface of the compressor blade

Table 3.11 Measured and calculated values

a [mm]	$\bar{s}$ [nm]	R	$\Delta\sigma$ [MPa]	N_c
8	45	0.345	518	2.01×10^6
10	120	0.265	582	1.28×10^6

The macromorphology of the fracture surface exhibited a sequence of characteristic regions and growth curves corresponding to different loading blocks; see Figure 3.77. Bioden-carbon replicas of fracture surfaces had to be prepared for observations using the transmission electron microscope in order to find some facets covered by fatigue striations. The critical crack length $a_c = 17\,\mathrm{mm}$ corresponded to the final fracture. Three-point bending specimens made of the blade material were used for measuring parameters in Equation 3.37: $\Delta K_{th} = 5\,\mathrm{MPa\,m}^{1/2}$, $K_c = 80\,\mathrm{MPa\,m}^{1/2}$, $A = 4.582 \times 10^{-36}$ [Pa,m], $n = 4.9$ and $m = 0.3$. Measured mean striation spacing $\bar{s}$ related to two different crack lengths is shown in Table 3.11 along with calculated values of R, $\Delta\sigma$ and N_c.

Computed values of $\Delta\sigma$ lie below the fatigue limit $\Delta\sigma_c \approx 600\,\mathrm{MPa}$ (for $R \approx 0.3$) of the blade material. The initial crack length $a_0 = 0.5\,\mathrm{mm}$ was chosen to be nearly equal to the size of silicon particles. The computed residual fatigue life $N_c \approx 10^6$ cycles also shows that the stable crack propagation ran at a stress below the fatigue limit of the material. The crack initiation stage was very low due to the impact of silicon particles. When neglecting the initiation stage and assuming the blade vibration frequency of several hundreds of Hz, the duration of the whole fracture process was no longer than several days.

It was realized later on that the test was performed during a week when, due to a service cut-off, the ground floor in the cave remained uncleaned.