

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

186,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Fracture Mechanics Based Models of Structural and Contact Fatigue

Ilya I. Kudish

Additional information is available at the end of the chapter

<http://dx.doi.org/10.5772/48511>

1. Introduction

The subsurface initiated contact fatigue failure is one of the dominating mechanisms of failure of moving machine parts involved in cyclic motion. Structural fatigue failure may be of surface or subsurface origin. The analysis of a significant amount of accumulated experimental data obtained from field exploitation and laboratory testing provides undisputable evidence of the most important factors affecting contact fatigue [1]. It is clear that the factors affecting contact fatigue the most are as follows (a) acting cyclic normal stress and frictional stress (detailed lubrication conditions, surface roughness, etc.) which in part are determined by the part geometry, (b) distribution of residual stress versus depth, (c) initial statistical defect/crack distribution versus defect size, and location, (d) material elastic and fatigue parameters as functions of materials hardness, etc., (e) material fracture toughness, (f) material hardness versus depth, (g) machining and finishing operations, (h) abrasive contamination of lubricant and residual surface contamination, (i) non-steady cyclic loading regimes, etc. In case of structural fatigue the list of the most important parameters affecting fatigue performance is similar. None of the existing contact or structural fatigue models developed for prediction of contact fatigue life of bearings and gears as well as other structures takes into account all of the above operational and material conditions. Moreover, at best, most of the existing contact fatigue models are only partially based on the fundamental physical and mechanical mechanisms governing the fatigue phenomenon. Most of these models are of empirical nature and are based on assumptions some of which are not supported by experimental data or are controversial as it is in the case of fatigue models for bearings and gears [1]. Some models involve a number of approximations that usually do not reflect the actual processes occurring in material.

Therefore, a comprehensive mathematical models of contact and structural fatigue failure should be based on clearly stated mechanical principles following from the theory of elasticity, lubrication theory of elastohydrodynamic contact interactions, and fracture mechanics. Such

models should take into consideration all the parameters described in items (a)-(e) and beyond. The advantage of such comprehensive models would be that the effect of variables such as steel cleanliness, externally applied stresses, residual stresses, etc. on contact and structural fatigue life could be examined as single or composite entities.

The goal of this chapter is to provide fracture mechanics based models of contact and structural fatigue. Historically, one of the most significant problems in realization of such an approach is the availability of simple but sufficiently precise solutions for the crack stress intensity factors. To overcome this difficulty some problems of fracture mechanics will be analyzed and their solutions will be represented in an analytical form acceptable for the further usage in modeling of contact and structural fatigue. In particular, a problem for an elastic half-plane weakened with a number of subsurface cracks and loaded with contact normal and frictional stresses as well with residual stress will be formulated and its asymptotic analytical solution will be presented. The latter solutions for the crack stress intensity factors are expressed in terms of certain integrals of known functions. These solutions for the stress intensity factors will be used in formulation of a two-dimensional contact fatigue model. In addition to that, a three-dimensional model applicable to both structural and contact fatigue will be formulated and some examples will be given. The above models take into account the parameters indicated in items (a)-(e) and are open for inclusion of the other parameters significant for fatigue.

In particular, these fatigue models take into account the statistical distribution of inclusions/cracks over the volume of the material versus their size and the resultant stress acting at the location of every inclusion/crack. Some of the main assumptions of the models are that (1) fatigue process in any machine part or structural unit runs in a similar manner and it is a direct reflection of the acting cyclic stresses and material properties and (2) the main part of fatigue life corresponds to the fatigue crack growth period, i.e. fatigue crack initiation period can be neglected. Furthermore, the variation in the distribution of cracks over time due to their fatigue growth is accounted for which is absent in all other existing fatigue models. The result of the above fatigue modeling is a simple relationship between fatigue life and cyclic loading, material mechanical parameters and its cleanliness as well as part geometry.

2. Three-dimensional model of contact and structural fatigue

The approach presented in this section provides a unified model of contact and structural fatigue of materials [1, 2]. The model development is based on a block approach, i.e. each block of the model describes a certain process related to fatigue and can be easily replaced by another block describing the same process differently. For example, accumulation of new more advanced knowledge of the process of fatigue crack growth may provide an opportunity to replace the proposed here block dealing with fatigue crack growth with an improved one. Two examples of the application of this model to structural fatigue are provided.

2.1. Initial statistical defect distribution

It is assumed that material defects are far from each other and practically do not interact. However, in some cases clusters of nonmetallic inclusions located very close to each other

are observed. In such cases these defect clusters can be represented by single defects of approximately the same size. Suppose there is a characteristic size L_σ in material that is determined by the typical variations of the material stresses, grain and surface geometry. It is also assumed that there is a size L_f in material such that $L_d \ll L_f \ll L_\sigma$, where L_d is the typical distance between the material defects. In other words, it is assumed that the defect population in any such volume L_f^3 is large enough to ensure an adequate statistical representation. It means that any parameter variations on the scale of L_f are indistinguishable for the fatigue analysis purposes and that in the further analysis any volume L_f^3 can be represented by its center point (x, y, z) .

Therefore, there is an initial statistical defect distribution in the material such that each defect can be replaced by a subsurface penny-shaped or a surface semi-circular crack with a radius approximately equal to the half of the defect diameter. The usage of penny-shaped subsurface and semi-circular surface cracks is advantageous to the analysis because such fatigue cracks maintain their shape and their size is characterized by just one parameter. The orientation of these crack propagation will be considered later. The initial statistical distribution is described by a probabilistic density function $f(0, x, y, z, l_0)$, such that $f(0, x, y, z, l_0)dl_0dxdydz$ is the number of defects with the radii between l_0 and $l_0 + dl_0$ in the material volume $dxdydz$ centered about point (x, y, z) . The material defect distribution is a local characteristic of material defectiveness. The model can be developed for any specific initial distribution $f(0, x, y, z, l_0)$. Some experimental data [3] suggest a log-normal initial defect distribution $f(0, x, y, z, l_0)$ versus the defect initial radius l_0

$$f(0, x, y, z, l_0) = 0 \text{ if } l_0 \leq 0, \quad (1)$$

$$f(0, x, y, z, l_0) = \frac{\rho(0, x, y, z)}{\sqrt{2\pi}\sigma_{ln}l_0} \exp\left[-\frac{1}{2}\left(\frac{\ln(l_0) - \mu_{ln}}{\sigma_{ln}}\right)^2\right] \text{ if } l_0 > 0,$$

where μ_{ln} and σ_{ln} are the mean value and standard deviation of the crack radii, respectively.

2.2. Direction of fatigue crack propagation

It is assumed that the duration of the crack initiation period is negligibly small in comparison with the duration of the crack propagation period. It is also assumed that linear elastic fracture mechanics is applicable to small fatigue cracks. The details of the substantiation of these assumptions can be found in [1]. Based on these assumptions in the vicinity of a crack the stress intensity factors completely characterize the material stress state. The normal k_1 and shear k_2 and k_3 stress intensity factors at the edge of a single crack of radius l can be represented in the form [4]

$$k_1 = F_1(x, y, z, \alpha, \beta)\sigma_1\sqrt{\pi l}, \quad k_2 = F_2(x, y, z, \alpha, \beta)\sigma_1\sqrt{\pi l}, \quad (2)$$

$$k_3 = F_3(x, y, z, \alpha, \beta)\sigma_1\sqrt{\pi l},$$

where σ_1 is the maximum of the local tensile principal stress, F_1 , F_2 , and F_3 are certain functions of the point coordinates (x, y, z) and the crack orientation angles α and β with respect to the coordinate planes. The coordinate system is introduced in such a way that the x - and

y -axes are directed along the material surface while the z -axis is directed perpendicular to the material surface.

The resultant stress field in an elastic material is formed by stresses $\sigma_x(x, y, z)$, $\sigma_y(x, y, z)$, $\sigma_z(x, y, z)$, $\tau_{xz}(x, y, z)$, $\tau_{xy}(x, y, z)$, and $\tau_{zy}(x, y, z)$. Some regions of material are subjected to tensile stress while other regions are subjected to compressive stress. Conceptually, there is no difference between the phenomena of structural and contact fatigue as the local material response to the same stress in both cases is the same and these cases differ in their stress fields only. As long as the stress levels do not exceed the limits of applicability of the quasi-brittle linear fracture mechanics when plastic zones at crack edges are small the rest of the material behaves like an elastic solid. The actual stress distributions in cases of structural and contact fatigue are taken in the proper account. In contact interactions where compressive stress is usually dominant there are still zones in material subjected to tensile stress caused by contact frictional and/or tensile residual stress [1, 5].

Experimental and theoretical studies show [1] that after initiation fatigue cracks propagate in the direction determined by the local stress field, namely, perpendicular to the local maximum tensile principle stress. Therefore, it is assumed that fatigue is caused by propagation of penny-shaped subsurface or semi-circular surface cracks under the action of principal maximum tensile stresses. Only high cycle fatigue phenomenon is considered here. On a plane perpendicular to a principal stress the shear stresses are equal to zero, i.e. the shear stress intensity factors $k_2 = k_3 = 0$. To find the plane of fatigue crack propagation (i.e. the orientation angles α and β), which is perpendicular to the maximum principal tensile stress, it is necessary to find the directions of these principal stresses. The latter is equivalent to solving the equations

$$k_2(N, \alpha, \beta, l, x, y, z) = 0, \quad k_3(N, \alpha, \beta, l, x, y, z) = 0. \quad (3)$$

Usually, there are more than one solution sets to these equations at any point (x, y, z) . To get the right angles α and β one has to choose the solution set that corresponds to the maximum tensile principal stress, i.e. maximum of the normal stress intensity factor $k_1(N, l, x, y, z)$. That guarantees that fatigue cracks propagate in the direction perpendicular to the maximum tensile principal stress if α and β are chosen that way.

For steady cyclic loading for small cracks $k_{20} = k_2/\sqrt{l}$ and $k_{30} = k_3/\sqrt{l}$ are independent from the number of cycles N and crack radius l together with equation (3) lead to the conclusion that for cyclic loading with constant amplitude the angles α and β characterizing the plane of fatigue crack growth are independent from N and l . Thus, angles α and β are functions of only crack location, i.e. $\alpha = \alpha(x, y, z)$ and $\beta = \beta(x, y, z)$. For the most part of their lives fatigue cracks created and/or existed near material defects remain small. Therefore, penny-shaped subsurface cracks conserve their shape but increase in size.

Even for the case of an elastic half-space it is a very difficult task to come up with sufficiently precise analytical solutions for the stress intensity factors at the edges of penny-shaped subsurface or semi-circular surface crack of arbitrary orientation at an arbitrary location (x, y, z) . However, due to the fact that practically all the time fatigue cracks remain small and exert little influence on the material general stress state the angles α and β can be determined in the process of calculation of the maximum tensile principle stress σ . The latter is equivalent to solution of equations (3). As soon as σ is determined for a subsurface crack its normal stress

intensity factor k_1 can be approximated by the normal stress intensity factor for the case of a single crack of radius l in an infinite space subjected to the uniform tensile stress σ , i.e. by $k_1 = 2\sigma\sqrt{l/\pi}$.

2.3. Fatigue crack propagation

Fatigue cracks propagate at every point of the material stressed volume V at which $\max_T(k_1) > k_{th}$, where the maximum is taken over the duration of the loading cycle T and k_{th} is the material stress intensity threshold. There are three distinct stages of crack development: (a) growth of small cracks, (b) propagation of well-developed cracks, and (c) explosive and, usually, unstable growth of large cracks. The stage of small crack growth is the slowest one and it represents the main part of the entire crack propagation period. This situation usually causes confusion about the duration of the stages of crack initiation and propagation of small cracks. The next stage, propagation of well-developed cracks, usually takes significantly less time than the stage of small crack growth. And, finally, the explosive crack growth takes almost no time.

A relatively large number of fatigue crack propagation equations are collected and analyzed in [6]. Any one of these equations can be used in the model to describe propagation of fatigue cracks. However, the simplest of them which allows to take into account the residual stress and, at the same time, to avoid the usage of such an unstable characteristic as the stress intensity threshold k_{th} is Paris's equation

$$\frac{dl}{dN} = g_0 \left(\max_{-\infty < x < \infty} \Delta k_1 \right)^n, \quad l|_{N=0} = l_0, \quad (4)$$

where g_0 and n are the parameters of material fatigue resistance and l_0 is the crack initial radius. Notice, that in cases of loading and relaxation such as in contact fatigue $\Delta k_1 = k_1$.

Fatigue cracks propagate until they reach their critical size with radius l_c for which $k_1 = K_f$ (K_f is the material fracture toughness, i.e. to the radius of $l_c = (K_f/k_{10})^2$). After that their growth becomes unstable and very fast. Usually, the stage of explosive crack growth takes just few loading cycles.

It can be shown that the number of loading cycles needed for a crack to reach its critical radius is almost independent from the material fracture toughness K_f . This conclusion is supported by direct numerical simulations. For the further analysis, it is necessary to determine for a crack its initial radius l_{0c} , which after N loading cycles reaches the critical size of l_c . Solving the initial-value problem (4) one obtains the formula

$$l_{0c} = \left\{ l_c^{\frac{2-n}{2}} + N \left(\frac{n}{2} - 1 \right) g_0 \left[\max_{-\infty < x < \infty} \Delta k_{10} \right]^n \right\}^{\frac{2}{2-n}}, \quad (5)$$

where l_{0c} depends on N , x , y , and z . Obviously, for $n > 2$ and fixed x , y , and z the value of l_{0c} is a decreasing function of N . It is important to keep in mind that $l_{0c}(N, x, y, z)$ is minimal where $k_{10}(x, y, z)$ is maximal, which, in turn, happens where the material tensile stress reaches its maximum.

2.4. Crack propagation statistics

To describe crack statistics after the crack initiation stage is over it is necessary to make certain assumptions. The simplest assumptions of this kind are: the existing cracks do not heal and new cracks are not created. In other words, the number of cracks in any material volume remains constant in time. Based on a practically correct assumption that the defect distribution is initially scarce, the coalescence of cracks and changes in the general stress field are possible only when cracks have already reached relatively large sizes. However, this may happen only during the last stage of crack growth the duration of which is insignificant for calculation of fatigue life. Therefore, it can be assumed that over almost all life span of fatigue cracks their orientations do not change. This leads to the equation for the density of crack distribution $f(N, x, y, z, l)$ as a function of crack radius l after N loading cycles in a small parallelepiped $dx dy dz$ with the center at the point with coordinates (x, y, z)

$$f(N, x, y, z, l) dl = f(0, x, y, z, l_0) dl_0, \quad (6)$$

which being solved for $f(N, x, y, z, l)$ gives

$$f(N, x, y, z, l) dl = f(0, x, y, z, l_0) \frac{dl_0}{dl}, \quad (7)$$

where l_0 and dl_0/dl as functions of N and l can be obtain from the solution of (4) in the form

$$l_0 = \{l^{\frac{2-n}{2}} + N(\frac{n}{2} - 1)g_0[\max_{-\infty < x < \infty} \Delta k_{10}]^n\}^{\frac{2}{2-n}}, \quad (8)$$

$$\frac{dl_0}{dl} = \{1 + N(\frac{n}{2} - 1)g_0[\max_{-\infty < x < \infty} \Delta k_{10}]^n l^{\frac{n-2}{2}}\}^{\frac{n}{n-2}}.$$

Equations (7) and (8) lead to the expression for the crack distribution function f after N loading cycles

$$f(N, x, y, z, l) = f(0, x, y, z, l_0(N, l, y, z)) \{1 + N(\frac{n}{2} - 1)g_0[\max_{-\infty < x < \infty} \Delta k_{10}]^n l^{\frac{n-2}{2}}\}^{\frac{n}{n-2}}, \quad (9)$$

where $l_0(N, l, y, z)$ is determined by the first of the equations in (8).

Formula (9) leads to a number of important conclusions. The fatigue crack distribution function $f(N, x, y, z, l)$ depends on the initial crack distribution $f(0, x, y, z, l_0)$ and it changes with the number of applied loading cycles N in such a way that the crack volume density $\rho(N, x, y, z)$ remains constant. Because of crack growth the crack distribution $f(N, x, y, z, l)$ widens with respect to l with number of loading cycles N .

2.5. Local fatigue damage accumulation

Let us design the measure of material fatigue damage. If at a certain point (x, y, z) after N loading cycles radii of all cracks $l < l_c$ then there is no damage at this point and the material local survival probability $p(N, x, y, z) = 1$. On the other hand, if at this point after N loading cycles radii of all cracks $l \geq l_c$, then all cracks reached the critical size and the material at this point is completely damaged and the local survival probability $p(N, x, y, z) = 0$. Obviously,

the more fatigue cracks with larger radii l exist at the point the lower is the local survival probability $p(N, x, y, z)$. It is reasonable to assume that the material local survival probability $p(N, x, y, z)$ is a certain monotonic measure of the portion of cracks with radius l below the critical radius l_c . Therefore, $p(N, x, y, z)$ can be represented by the expressions

$$\begin{aligned}
 p(N, x, y, z) &= \frac{1}{\rho} \int_0^{l_c} f(N, x, y, z, l) dl \text{ if } f(0, x, y, z, l_0) \neq 0, \\
 p(N, x, y, z) &= 1 \text{ otherwise,} \\
 \rho &= \rho(N, x, y, z) = \int_0^{\infty} f(N, x, y, z, l) dl = \rho(0, x, y, z).
 \end{aligned} \tag{10}$$

Obviously, the local survival probability $p(N, x, y, z)$ is a monotonically decreasing function of the number of loading cycles N because fatigue crack radii l tend to grow with the number of loading cycles N .

To calculate $p(N, x, y, z)$ from (10) one can use the specific expression for f determined by (9). However, it is more convenient to modify it as follows

$$\begin{aligned}
 p(N, x, y, z) &= \frac{1}{\rho} \int_0^{l_{0c}} f(0, x, y, z, l_0) dl_0 \text{ if } f(0, x, y, z, l_0) \neq 0, \\
 p(N, x, y, z) &= 1 \text{ otherwise,}
 \end{aligned} \tag{11}$$

where l_{0c} is determined by (5) and ρ is the initial volume density of cracks. Thus, to every material point (x, y, z) is assigned a certain local survival probability $p(N, x, y, z)$, $0 \leq p(N, x, y, z) \leq 1$.

Equations (11) demonstrate that the material local survival probability $p(N, x, y, z)$ is mainly controlled by the initial crack distribution $f(0, x, y, z, l_0)$, material fatigue resistance parameters g_0 and n , and external contact and residual stresses. Moreover, the material local survival probability $p(N, x, y, z)$ is a decreasing function of N because l_{0c} from (5) is a decreasing function of N for $n > 2$.

2.6. Global fatigue damage accumulation

The survival probability $P(N)$ of the material as a whole is determined by the local probabilities of all points of the material at which fatigue cracks are present. It is assumed that the material fails as soon as it fails at just one point. It is assumed that the initial crack distribution in the material is discrete. Let $p_i(N) = p(N, x_i, y_i, z_i)$, $i = 1, \dots, N_c$, where N_c is the total number of points in the material stressed volume V at which fatigue cracks are present. Then based on the above assumption the material survival probability $P(N)$ is equal to

$$P(N) = \prod_{i=1}^{N_c} p_i(N). \tag{12}$$

Obviously, probability $P(N)$ from (12) satisfies the inequalities

$$[p_m(N)]^{N_c} \leq P(N) \leq p_m(N), \quad p_m(N) = \min_V p(N, x, y, z). \quad (13)$$

In (13) the right inequality shows that the survival probability $P(N)$ is never greater than the minimum value $p_m(N)$ of the local survival probability $p(N, x, y, z)$ over the material stressed volume V .

An analytical substantiation for the assumption that the first pit is created by the cracks from a small material volume with the smallest survival probability $p_m(N)$ is provided in [1]. Moreover, the indicated analysis also validates one of the main assumptions of the model that new cracks are not being created. Namely, if new cracks do get created in the process of loading, they are very small and have no chance to catch up with already existing and propagating larger cracks. The graphical representation of this fact is given in Fig. 1 [1]. In this figure fatigue cracks are initially randomly distributed over the material volume with respect to their normal stress intensity factor k_1 and are allowed to grow according to Paris' law (see (4)) with sufficiently high value of $n = 6.67 - 9$. In Fig. 1 the values of the normal stress intensity factor k_1 are shown at different time moments (k_0 and L_0 are the characteristic normal stress intensity factor and geometric size of the solid). These graphs clearly show that a crack with the initially larger value of the normal stress intensity factor k_1 propagates much faster than all other cracks, i.e. the value of its k_1 increases much faster than the values of k_1 for all other cracks, which are almost dormant. As a result of that, the crack with the initially larger value of k_1 reaches its critical size way ahead of other cracks. This event determines the time and the place where fatigue failure occurs initially. Therefore, in spite of formula (12) which indicates that all fatigue cracks have influence on the survival probability $P(N)$, for high values of n the material survival probability $P(N)$ is a local fatigue characteristic, and it is determined by the material defect with the initially highest value of the stress intensity factor k_1 . The higher the power n is the more accurate this approximation is.

Therefore, assuming that it is a very rare occurrence when more than one fatigue initiation/spall happen simultaneously, it can be shown that at the early stages of the fatigue process the material global survival probability $P(N)$ is determined by the minimum of the local survival probability $p_m(N)$

$$P(N) = p_m(N), \quad p_m(N) = \min_V p(N, x, y, z), \quad (14)$$

where the maximum is taken over the (stressed) volume V of the solid.

If the initial crack distribution is taken in the log-normal form (1) then

$$P(N) = p_m(N) = \frac{1}{2} \left\{ 1 + \operatorname{erf} \left[\min_V \frac{\ln l_{0c}(N, y, z) - \mu_{ln}}{\sqrt{2}\sigma_{ln}} \right] \right\}, \quad (15)$$

where $\operatorname{erf}(x)$ is the error integral [8]. Obviously, the local survival probability $p_m(N)$ is a complex combined measure of applied stresses, initial crack distribution, material fatigue parameters, and the number of loading cycles.

In cases when the mean μ_{ln} and the standard deviation σ_{ln} are constants throughout the material formula (15) can be significantly simplified

$$P(N) = p_m(N) = \frac{1}{2} \left\{ 1 + \operatorname{erf} \left[\frac{\ln \min_V l_{0c}(N, y, z) - \mu_{ln}}{\sqrt{2}\sigma_{ln}} \right] \right\}. \quad (16)$$

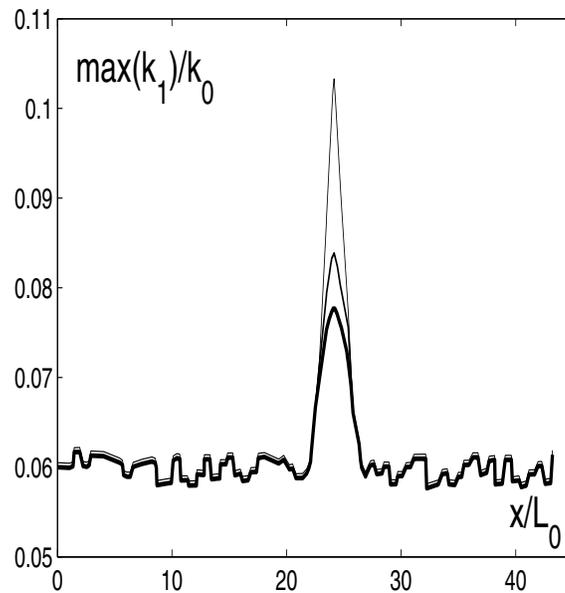


Figure 1. Illustration of the growth of the initially randomly distributed normal stress intensity factor k_1 with time N as initially unit length fatigue cracks grow (after Tallian, Hoeprich, and Kudish [7]). Reprinted with permission from the STLE.

To determine fatigue life N of a contact for the given survival probability $P(N) = P_*$, it is necessary to solve the equation

$$p_m(N) = P_* \quad (17)$$

2.7. Fatigue life calculation

Suppose the material failure occurs at point (x, y, z) with the probability $1 - P(N)$. That actually determines the point where in (16) the minimum over the material volume V is reached. Therefore, at this point in (16), the operation of minimum over the material volume V can be dropped. By solving (16) and (17), one gets

$$N = \left\{ \left(\frac{n}{2} - 1 \right) g_0 \left[\max_{-\infty < x < \infty} \Delta k_{10} \right]^n \right\}^{-1} \left\{ \exp \left[\left(1 - \frac{n}{2} \right) (\mu_{ln} + \sqrt{2} \sigma_{ln} \operatorname{erf}^{-1}(2P_* - 1)) \right] - l_c^{\frac{2-n}{2}} \right\}, \quad (18)$$

where $\operatorname{erf}^{-1}(x)$ is the inverse function to the error integral $\operatorname{erf}(x)$. Assuming that the material initially is free of damage, i.e., when $P(0) = 1$, one can simplify the latter equation. Discounting the very tail of the initial crack distribution, one gets $\max_V l_0 \leq l_c$. Thus, for well-developed cracks and, in many cases, even for small cracks, the second term in (5) for l_{0c} dominates the first one. It means that the dependence of l_{0c} on l_c and, therefore, on the material fracture toughness K_f can be neglected. Then equation (18) can be approximated by

$$N = \left\{ \left(\frac{n}{2} - 1 \right) g_0 \left[\max_{-\infty < x < \infty} \Delta k_{10} \right]^n \right\}^{-1} \left\{ \exp \left[\left(1 - \frac{n}{2} \right) (\mu_{ln} + \sqrt{2} \sigma_{ln} \operatorname{erf}^{-1}(2P_* - 1)) \right] \right\}. \quad (19)$$

Taking into account that in the case of contact fatigue k_{10} is proportional to the maximum contact pressure q_{max} and that it also depends on the friction coefficient λ and the ratio of residual stress q^0 and $q_{max} = p_H$ as well as taking into account the relationships $\mu_{ln} = \ln \frac{\mu^2}{\sqrt{\mu^2 + \sigma^2}}$, $\sigma_{ln} = \sqrt{\ln [1 + (\frac{\sigma}{\mu})^2]}$ [1] (where μ and σ are the regular initial mean and standard deviation) one arrives at a simple analytical formula

$$N = \frac{C_0}{(n-2)g_0p_H^n} \left(\frac{\sqrt{\mu^2 + \sigma^2}}{\mu} \right)^{\frac{n}{2}-1} \times \exp \left[\left(1 - \frac{n}{2} \right) \sqrt{2 \ln \left[1 + \left(\frac{\sigma}{\mu} \right)^2 \right]} \operatorname{erf}^{-1} (2P_* - 1) \right], \quad (20)$$

where C_0 depends only on the friction coefficient λ and the ratio of the residual stress q^0 and the maximum Hertzian pressure p_H . Finally, assuming that $\sigma \ll \mu$ from (20) one can obtain the formula

$$N = \frac{C_0}{(n-2)g_0p_H^n \mu^{\frac{n}{2}-1}} \exp \left[\left(1 - \frac{n}{2} \right) \frac{\sqrt{2}\sigma}{\mu} \operatorname{erf}^{-1} (2P_* - 1) \right]. \quad (21)$$

Also, formulas (20) and (21) can be represented in the form of the Lundberg-Palmgren formula (see [1] and the discussion there).

Formula (21) demonstrates the intuitively obvious fact that the fatigue life N is inversely proportional to the value of the parameter g_0 that characterizes the material crack propagation resistance. Equation (21) exhibits a usual for roller and ball bearings as well as for gears dependence of the fatigue life N on the maximum Hertzian pressure p_H . Thus, from the well-known experimental data for bearings the range of n values is $20/3 \leq n \leq 9$. Keeping in mind that usually $\sigma \ll \mu$, for these values of n contact fatigue life N is practically inverse proportional to a positive power of the mean crack size, i.e. to $\mu^{\frac{n}{2}-1}$. Therefore, fatigue life N is a decreasing function of the initial mean crack (inclusion) size μ . This conclusion is valid for any material survival probability P_* and is supported by the experimental data discussed in [1]. In particular, $\ln N$ is practically a linear function of $\ln \mu$ with a negative slope $1 - \frac{n}{2}$ which is in excellent agreement with the Timken Company test data [9]. Keeping in mind that $n > 2$, at early stages of fatigue failure, i.e. when $\operatorname{erf}^{-1} (2P_* - 1) > 0$ for $P_* > 0.5$, one easily determines that fatigue life N is a decreasing function of the initial standard deviation of crack sizes σ . Similarly, at late stages of fatigue failure, i.e. when $P_* < 0.5$, the fatigue life N is an increasing function of the initial standard deviation of crack sizes σ . According to (21), for $P_* = 0.5$ fatigue life N is independent from σ , however, according to (20), for $P_* = 0.5$ fatigue life N is a slowly increasing function of σ . By differentiating $p_m(N)$ obtained from (16) with respect to σ , one can conclude that the dispersion of $P(N)$ increases with σ .

The stress intensity factor k_1 decreases as the magnitude of the compressive residual stress q^0 increases and/or the magnitude of the friction coefficient λ decreases. Therefore, in (20) and (21) the value of C_0 is a monotonically decreasing function of residual stress q^0 and friction coefficient λ .

Being applied to bearings and/or gears the described statistical contact fatigue model can be used as a research and/or engineering tool in pitting modeling. In the latter case, some of the model parameters may be assigned certain fixed values based on the scrupulous analysis of steel quality and quality and stability of gear and bearing manufacturing processes.

In case of structural fatigue the Hertzian stress in formulas (20) and (21) should be replaced the dominant stress acting on the part while constant C_0 would be dependent on the ratios of other external stresses acting on the part at hand to the dominant stress in a certain way (see examples of torsional and bending fatigue below).

2.8. Examples of torsional and bending fatigue

Suppose that in a beam material the defect distribution is space-wise uniform and follows equation (1). Also, let us assume that the residual stress is zero.

First, let us consider torsional fatigue. Suppose a beam is made of an elastic material with elliptical cross section (a and b are the ellipse semi-axes, $b < a$) and directed along the y -axis. The beam is under action of torque M_y about the y -axis applied to its ends. The side surfaces of the beam are free of stresses. Then it can be shown (see Lurye [10], p. 398) that

$$\tau_{xy} = -\frac{2G\gamma a^2}{a^2+b^2}z, \quad \tau_{zy} = \frac{2G\gamma b^2}{a^2+b^2}x, \quad \sigma_x = \sigma_y = \sigma_z = \tau_{xz} = 0, \quad (22)$$

where G is the material shear elastic modulus, $G = E/[2(1+\nu)]$ (E and ν are Young's modulus and Poisson's ratio of the beam material), and γ is a dimensionless constant. By introducing the principal stresses σ_1 , σ_2 , and σ_3 that satisfy the equation $\sigma^3 - (\tau_{xy}^2 + \tau_{zy}^2)\sigma = 0$, one obtains that

$$\sigma_1 = -\sqrt{\tau_{xy}^2 + \tau_{zy}^2}, \quad \sigma_2 = 0, \quad \sigma_3 = \sqrt{\tau_{xy}^2 + \tau_{zy}^2}. \quad (23)$$

For the case of $a > b$ the maximum principal tensile stress $\sigma_1 = -\frac{2G\gamma a^2 b}{a^2+b^2}$ is reached at the surface of the beam at points $(0, y, \pm b)$ and depending on the sign of M_y it acts in one of the directions described by the directional cosines

$$\cos(\alpha, x) = \mp \frac{\sqrt{2}}{2}, \quad \cos(\alpha, y) = \pm \frac{\sqrt{2}}{2}, \quad \cos(\alpha, z) = 0, \quad (24)$$

where α is the direction along one of the principal stress axes. For the considered case of elliptic beam, the moments of inertia of the beam elliptic cross section about the x - and y -axes, I_x and I_z as well as the moment of torsion M_y applied to the beam are as follows (see Lurye [10], pp. 395, 399) $I_x = \pi ab^3/4$, $I_z = \pi a^3 b/4$, $M_y = G\gamma C$, $C = 4I_x I_z / (I_x + I_z)$. Keeping in mind that according to Hasebe and Inohara [11] and Isida [12], the stress intensity factor k_1 for an edge crack of radius l and inclined to the surface of a half-plane at the angle of $\pi/4$ (see (24)) is $k_1 = 0.705 |\sigma_1| \sqrt{\pi l}$, one obtains $k_{10} = \frac{1.41}{\sqrt{\pi}} \frac{|M_y|}{ab^2}$. Then, fatigue life of a beam under torsion follows from substituting the expression for k_{10} into equation (19)

$$N = \frac{2}{(n-2)g_0} \left\{ \frac{1.257ab^2}{|M_y|} \right\}^n g(\mu, \sigma), \quad (25)$$

$$g(\mu, \sigma) = \left(\frac{\sqrt{\mu^2 + \sigma^2}}{\mu^2} \right)^{\frac{n-2}{2}} \exp \left[\left(1 - \frac{n}{2} \right) \sqrt{2 \ln \left[1 + \left(\frac{\sigma}{\mu} \right)^2 \right]} \operatorname{erf}^{-1} (2P_* - 1) \right]. \quad (26)$$

Now, let us consider bending fatigue of a beam/console made of an elastic material with elliptical cross section (a and b are the ellipse semi-axes) and length L . The beam is directed along the y -axis and it is under the action of a bending force P_x directed along the x -axis which is applied to its free end. The side surfaces of the beam are free of stresses. The other end $y = 0$

of the beam is fixed. Then it can be shown (see Lurye [10]) that

$$\begin{aligned}\sigma_x = \sigma_z = 0, \quad \sigma_y &= -\frac{P_x}{I_z}x(L-y), \\ \tau_{xz} = 0, \quad \tau_{xy} &= \frac{P_x}{2(1+\nu)I_z} \frac{2(1+\nu)a^2+b^2}{3a^2+b^2} \left\{ a^2 - x^2 - \frac{(1-2\nu)a^2z^2}{2(1+\nu)a^2+b^2} \right\}, \\ \tau_{zy} &= -\frac{P_x}{(1+\nu)I_z} \frac{(1+\nu)a^2+\nu b^2}{3a^2+b^2} xz,\end{aligned}\quad (27)$$

where I_z is the moment of inertia of the beam cross section about the z -axis. By introducing the principal stresses that satisfy the equation $\sigma^3 - \sigma_y\sigma^2 - (\tau_{xy}^2 + \tau_{zy}^2)\sigma = 0$, one can find that

$$\begin{aligned}\sigma_1 &= \frac{1}{2}[\sigma_y - \sqrt{\sigma_y^2 + 4(\tau_{xy}^2 + \tau_{zy}^2)}], \quad \sigma_2 = 0, \\ \sigma_3 &= \frac{1}{2}[\sigma_y + \sqrt{\sigma_y^2 + 4(\tau_{xy}^2 + \tau_{zy}^2)}].\end{aligned}\quad (28)$$

The tensile principal stress σ_1 reaches its maximum $\frac{4|P_x|L}{\pi a^2 b}$ at the surface of the beam at one of the points $(\pm a, 0, 0)$ (depending on the sign of load P_x) and is acting along the y -axis - the axis of the beam. Based on equations (28) and the solution for the surface crack inclined to the surface of the half-space at angle of $\pi/2$ (see Hasebe and Inohara [11] and Isida [12]), one obtains $k_{10} = \frac{4.484}{\sqrt{\pi}} \frac{|P_x|L}{a^2 b}$. Therefore, bending fatigue life of a beam follows from substituting the expression for k_{10} into equation (19)

$$N = \frac{2}{(n-2)g_0} \left\{ \frac{0.395a^2b}{|P_x|L} \right\}^n g(\mu, \sigma), \quad (29)$$

where function $g(\mu, \sigma)$ is determined by equation (26).

In both cases of torsion and bending, fatigue life is independent of the elastic characteristic of the beam material (see formulas (25), (29), and (26)), and it is dependent on fatigue parameters of the beam material (n and g_0), the initial defect distribution (i.e. on μ and σ), the geometry of the beam cross section (a and b), and its length L and the applied loading (P_x or M_y).

In a similar fashion the model can be applied to contact fatigue if the stress field is known. A more detailed analysis of contact fatigue is presented below for a two-dimensional case.

3. Contact problem for an elastic half-plane weakened by straight cracks

A general theory of a stress state in an elastic plane with multiple cracks was proposed in [13]. In this section this theory is extended to the case of an elastic half-plane loaded by contact and residual stresses [1]. A study of lubricant-surface crack interaction, a discussion of the difference between contact fatigue lives of drivers and followers, the surface and subsurface initiated fatigue as well as fatigue of rough surfaces can be found in [1].

The main purpose of the section is to present formulations for the contact and fracture mechanics problems for an elastic half-plane weakened by subsurface cracks. The problems for surface cracks in an elastic lubricated half-plane are formulated and analyzed in [1]. The

problems are reduced to systems of integro-differential equations with nonlinear boundary conditions in the form of alternating equations and inequalities. An asymptotic (perturbation) method for the case of small cracks is applied to solution of the problem and some numerical examples for small cracks are presented.

Let us introduce a global coordinate system with the x^0 -axis directed along the half-plane boundary and the y^0 -axis perpendicular to the half-plane boundary and pointed in the direction outside the material. The half-plane occupies the area of $y^0 \leq 0$. Let us consider a contact problem for a rigid indenter with the bottom of shape $y^0 = f(x^0)$ pressed into the elastic half-plane (see Fig. 2). The elastic half-plane with effective elastic modulus E' ($E' = E/(1 - \nu^2)$, E and ν are the half-plane Young's modulus and Poisson's ratio) is weakened by N straight cracks. The crack faces are frictionless. Besides the global coordinate system we will introduce local orthogonal coordinate systems for each straight crack of half-length l_k in such a way that their origins are located at the crack centers with complex coordinates $z_k^0 = x_k^0 + iy_k^0$, $k = 1, \dots, N$, the x_k -axes are directed along the crack faces and the y_k -axes are directed perpendicular to them. The cracks are inclined to the positive direction of the x^0 -axis at the angles α_k , $k = 1, \dots, N$. All cracks are considered to be subsurface. The faces of every crack may be in partial or full contact with each other. The indenter is loaded by a normal force P and may be in direct contact with the half-plane or separated from it by a layer of lubricant. The indenter creates a pressure $p(x^0)$ and frictional stress $\tau(x^0)$ distributions. The frictional stress $\tau(x^0)$ between the indenter and the boundary of the half-plane is determined by the contact pressure $p(x^0)$ through a certain relationship. The cases of dry and fluid frictional stress $\tau(x^0)$ are considered in [1]. At infinity the half-plane is loaded by a tensile or compressive (residual) stress $\sigma_{x^0}^\infty = q^0$ which is directed along the x^0 -axis. In this formulation the problem is considered in [1].

Then the problem is reduced to determining of the cracks behavior. Therefore, in dimensionless variables

$$\begin{aligned} (x_n^{0'}, y_n^{0'}) &= (x_n^0, y_n^0)/\tilde{b}, \quad (p_n^{0'}, \tau_n^{0'}, p_n') = (p_n^0, \tau_n^0, p_n)/\tilde{q}, \\ (x_n', t') &= (x_n, t)/l_n, \quad (v_n', u_n') = (v_n, u_n)/\tilde{v}_n, \quad \tilde{v}_n = \frac{4\tilde{q}l_n}{E'}, \\ (k_{1n}^{\pm'}, k_{2n}^{\pm'}) &= (k_{1n}^\pm, k_{2n}^\pm)/(\tilde{q}\sqrt{l_n}) \end{aligned} \quad (30)$$

the equations of the latter problem for an elastic half-plane weakened by cracks and loaded by contact and residual stresses have the following form [1]

$$\begin{aligned} \int_{-1}^1 \frac{v_k'(t)dt}{t-x_k} + \sum_{m=1}^N \delta_m \int_{-1}^1 [v_m'(t)A_{km}^r(t, x_k) - u_m'(t)B_{km}^r(t, x_k)]dt \\ = \pi p_{nk}(x_k) + \pi p_k^0(x_k), \quad v_k(\pm 1) = 0, \\ \int_{-1}^1 \frac{u_k'(t)dt}{t-x_k} + \sum_{m=1}^N \delta_m \int_{-1}^1 [v_m'(t)A_{km}^i(t, x_k) - u_m'(t)B_{km}^i(t, x_k)]dt \\ = \pi \tau_k^0(x_k), \quad u_k(\pm 1) = 0, \end{aligned} \quad (31)$$

$$p_k^0 - i\tau_k^0 = -\frac{1}{\pi} \int_a^b [p(t)\overline{D}_k(t, x_k) + \tau(t)\overline{G}_k(t, x_k)]dt - \frac{1}{2}q^0(1 - e^{-2i\alpha_k}), \quad (32)$$

$$p_{nk}(x_k) = 0, \quad v_k(x_k) > 0; \quad p_{nk}(x_k) \leq 0, \quad v_k(x_k) = 0, \quad k = 1, \dots, N,$$

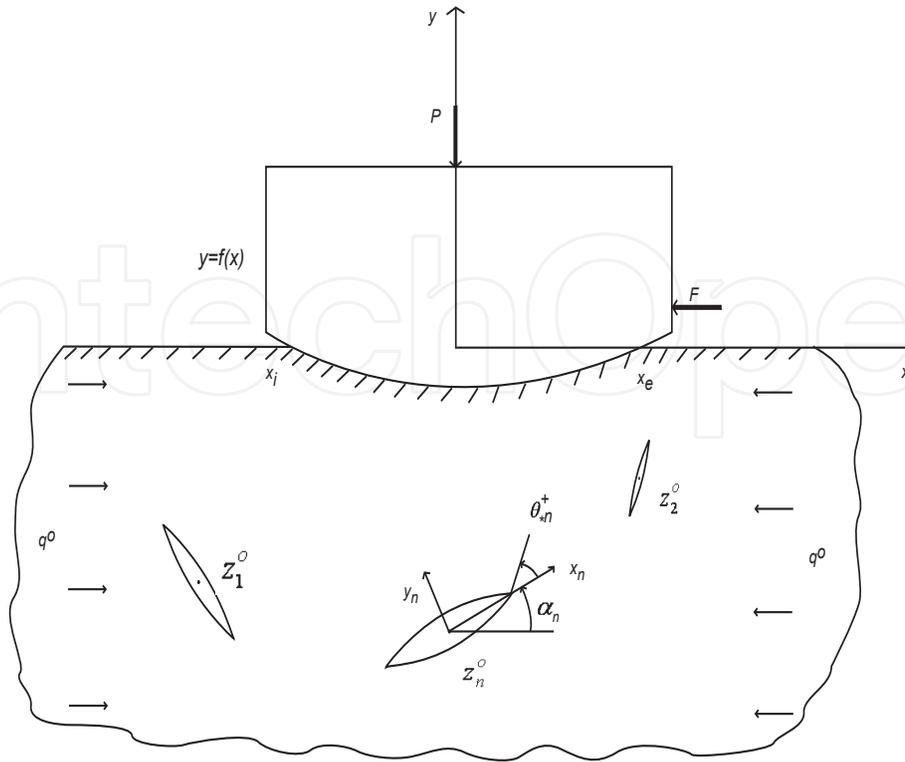


Figure 2. The general view of a rigid indenter in contact with a cracked elastic half-plane.

where the kernels in these equations are described by formulas

$$\begin{aligned}
 A_{km} &= \overline{R_{km} + S_{km}}, \quad B_{km} = -i(\overline{R_{km} - S_{km}}), \\
 (A_{km}^r, B_{km}^r, D_k^r, G_k^r) &= \text{Re}(A_{km}, B_{km}, \overline{D_k}, \overline{G_k}), \\
 (A_{km}^i, B_{km}^i, D_k^i, G_k^i) &= \text{Im}(A_{km}, B_{km}, \overline{D_k}, \overline{G_k}), \\
 D_k(t, x_k) &= \frac{i}{2} \left[-\frac{1}{t-X_k} + \frac{1}{t-\overline{X_k}} - \frac{e^{-2i\alpha_k}(\overline{X_k}-X_k)}{(t-\overline{X_k})^2} \right], \\
 G_k(t, x_k) &= \frac{1}{2} \left[\frac{1}{t-X_k} + \frac{1-e^{-2i\alpha_k}}{t-\overline{X_k}} - \frac{e^{-2i\alpha_k}(t-X_k)}{(t-\overline{X_k})^2} \right], \\
 R_{nk}(t, x_n) &= (1 - \delta_{nk})K_{nk}(t, x_n) + \frac{e^{i\alpha_k}}{2} \left\{ \frac{1}{X_n-\overline{T_k}} + \frac{e^{-2i\alpha_n}}{\overline{X_n}-T_k} \right. \\
 &\quad \left. + (\overline{T_k} - T_k) \left[\frac{1+e^{-2i\alpha_n}}{(\overline{X_n}-T_k)^2} + \frac{2e^{-2i\alpha_n}(T_k-X_n)}{(\overline{X_n}-T_k)^3} \right] \right\}, \\
 S_{nk}(t, x_n) &= (1 - \delta_{nk})L_{nk}(t, x_n) + \frac{e^{-i\alpha_k}}{2} \left[\frac{T_k-\overline{T_k}}{(\overline{X_n}-\overline{T_k})^2} + \frac{1}{\overline{X_n}-T_k} \right. \\
 &\quad \left. + \frac{e^{-2i\alpha_n}(T_k-X_n)}{(\overline{X_n}-T_k)^2} \right], \quad K_{nk}(t, x_n) = \frac{e^{i\alpha_k}}{2} \left[\frac{1}{T_k-X_n} + \frac{e^{-2i\alpha_n}}{\overline{T_k}-\overline{X_n}} \right],
 \end{aligned} \tag{33}$$

$$L_{nk}(t_k, x_n) = \frac{e^{-i\alpha_k}}{2} \left[\frac{1}{\overline{T_k - \overline{X_n}}} - \frac{T_k - X_n}{(\overline{T_k - \overline{X_n}})^2} e^{-2i\alpha_n} \right],$$

$$T_k = te^{i\alpha_k} + z_k^0, \quad X_n = x_n e^{i\alpha_n} + z_n^0, \quad k, n = 1, \dots, N,$$

where $v_k(x_k)$, $u_k(x_k)$, and $p_{nk}(x_k)$, $k = 1, \dots, N$, are the jumps of the normal and tangential crack face displacements and the normal stress applied to crack faces, respectively, a and b are the dimensionless contact boundaries, δ_k is the dimensionless crack half-length, $\delta_k = l_k/\tilde{b}$, δ_{nk} is the Kronecker tensor ($\delta_{nk} = 0$ for $n \neq k$, $\delta_{nk} = 1$ for $n = k$), i is the imaginary unit, $i = \sqrt{-1}$.

For simplicity primes at the dimensionless variables are omitted. The characteristic values \tilde{q} and \tilde{b} that are used for scaling are the maximum Hertzian pressure p_H and the Hertzian contact half-width a_H

$$p_H = \sqrt{\frac{E'P}{\pi R}}, \quad a_H = 2\sqrt{\frac{RP}{\pi E'}}, \quad (34)$$

where R can be taken as the indenter curvature radius at the center of its bottom.

To simplify the problem formulation it is assumed that for small subsurface cracks (i.e. for $\delta_0 \ll 1$, $\delta_0 = \max_{1 \leq k \leq N} \delta_k$) the pressure $p(x^0)$ and frictional stress $\tau(x^0)$ are known and are close to the ones in a contact of this indenter with an elastic half-plane without cracks. It is worth mentioning that cracks affect the contact boundaries a and b and the pressure distribution $p(x^0)$ as well as each other starting with the terms of the order of $\delta_0 \ll 1$.

Therefore, for the given shape of the indenter $f(x^0)$, pressure $p(x^0)$, frictional stress functions $\tau(x^0)$, residual stress q^0 , crack orientation angles α_k and sizes δ_k , and the crack positions z_k^0 , $k = 1, \dots, N$, the solution of the problem is represented by crack faces displacement jumps $u_k(x_k)$, $v_k(x_k)$, and the normal contact stress $p_{nk}(x_k)$ applied to the crack faces ($k = 1, \dots, N$). After the solution of the problem has been obtained, the dimensionless stress intensity factors k_{1k}^\pm and k_{2k}^\pm are determined according to formulas

$$k_{1n}^\pm + ik_{2n}^\pm = \mp \lim_{x_n \rightarrow \pm 1} \sqrt{1 - x_n^2} [v_n'(x_n) + iu_n'(x_n)], \quad 0 \leq n \leq N. \quad (35)$$

3.1. Problem solution

Solution of this problem is associated with formidable difficulties represented by the nonlinearities caused by the presence of the free boundaries of the crack contact intervals and the interaction between different cracks. Under the general conditions solution of this problem can be done only numerically. However, the problem can be effectively solved with the use of just analytical methods in the case when all cracks are small in comparison with the characteristic size \tilde{b} of the contact region, i.e., when $\delta_0 = \max_{1 \leq k \leq N} \delta_k \ll 1$. In this case, it can

be shown that the influence of the presence of cracks on the contact pressure is of the order of $O(\delta_0)$ and with the precision of $O(\delta_0)$ the crack system in the half-plane is subjected to the action of the contact pressure $p_0(x^0)$ and frictional stress $\tau_0(x^0)$ that are obtained in the absence of cracks. The further simplification of the problem is achieved under the assumption that cracks are small in comparison to the distances between them, i.e.

$$z_n^0 - z_k^0 \gg \delta_0, \quad n \neq k, \quad n, k = 1, \dots, N. \quad (36)$$

The latter assumption with the precision of $O(\delta_0^2)$, $\delta_0 \ll 1$, provides the conditions for considering each crack as a single crack in an elastic half-plane while the crack faces are loaded by certain stresses related to the contact pressure $p_0(x^0)$, contact frictional stress $\tau_0(x^0)$, and the residual stress q^0 . The crucial assumption for simple and effective analytical solution of the considered problem is the assumption that all cracks are subsurface and much smaller in size than their distances to the half-plane surface

$$z_k^0 - \bar{z}_k^0 \gg \delta_0, k = 1, \dots, N. \quad (37)$$

Essentially, that assumption permits to consider each crack as a single crack in a plane (not a half-plane) with faces loaded by certain stresses related to $p_0(x^0)$, $\tau_0(x^0)$, and q^0 .

Let us assume that the frictional stress $\tau_0(x^0)$ is determined by the Coulomb law of dry friction which in dimensionless variables can be represented by

$$\tau_0(x^0) = -\lambda p_0(x^0), \quad (38)$$

where λ is the coefficient of friction. In (38) we assume that $\lambda \geq 0$, and, therefore, the frictional stress is directed to the left. It is well known that for small friction coefficients λ the distribution of pressure is very close to the one in a Hertzian frictionless contact. Therefore, the expression for the pressure $p_0(x^0)$ in the absence of cracks with high accuracy can be taken in the form

$$p_0(x^0) = \sqrt{1 - (x^0)^2}. \quad (39)$$

Let us consider the process of solution of the pure fracture mechanics problem described by equations (31)-(33), (35), (38), (39). For small cracks, i.e. for $\delta_0 \ll 1$, the kernels from (32) and (33) are regular functions of t , x_n , and x_k and they can be represented by power series in $\delta_k \ll 1$ and $\delta_n \ll 1$ as follows

$$\begin{aligned} & \{A_{km}(t, x_k), B_{km}(t, x_k)\} \\ &= \sum_{j+n=0; j, n \geq 0}^{\infty} (\delta_k x_k)^j (\delta_m t)^n \{A_{kmjn}, B_{kmjn}\}, \end{aligned} \quad (40)$$

$$\{D_k(t, x_k), G_k(t, x_k)\} = \sum_{j=0}^{\infty} (\delta_k x_k)^j \{D_{kj}, G_{kj}\}. \quad (41)$$

In (40) and (41) the values of A_{kmjn} and B_{kmjn} are independent of δ_k , δ_m , x_k , and t while the values of $D_{kj}(t)$ and $G_{kj}(t)$ are independent of δ_k and x_k . The values of A_{kmjn} and B_{kmjn} are certain functions of constants α_k , α_m , x_k^0 , y_k^0 , x_m^0 , and y_m^0 while the values of $D_{kj}(t)$ and $G_{kj}(t)$ are certain functions of α_k , x_k^0 , and y_k^0 . Therefore, for $\delta_0 \ll 1$ the problem solution can be sought in the form

$$\{v_k, u_k\} = \sum_{j=0}^{\infty} \delta_k^j \{v_{kj}(x_k), u_{kj}(x_k)\}, p_{nk} = \sum_{j=0}^{\infty} \delta_0^j p_{nkj}(x_k) \quad (42)$$

where functions v_{kj} , u_{kj} , p_{nkj} have to be determined in the process of solution. Expanding the terms of the equations (31)-(33), and (35) and equating the terms with the same powers of δ_0 we get a system of boundary-value problems for integro-differential equations of the first kind

which can be easily solved by classical methods [1, 14]. We will limit ourselves to determining only the first two terms of the expansions in (42) in the case of Coulomb's friction law given by (38) and (39). Without getting into the details of the solution process (which can be found in [1]) for the stress intensity factors k_{1n}^{\pm} and k_{2n}^{\pm} we obtain the following analytical formulas [1]

$$\begin{aligned} k_1^{\pm} &= c_0^r \pm \frac{1}{2}\delta_0 c_1^r + \dots \text{ if } c_0^r > 0, \quad k_1^{\pm} = 0 \text{ if } c_0^r < 0, \\ k_1^{\pm} &= \frac{\sqrt{3}\delta_0}{9} c_1^r [\pm 7 - 3\theta(c_1^r)] \sqrt{\frac{1 \pm \theta(c_1^r)}{1 \pm 3\theta(c_1^r)}} + \dots \text{ if } c_0^r = 0 \text{ and } c_1^r \neq 0, \\ k_2^{\pm} &= c_0^i \pm \frac{1}{2}\delta_0 c_1^i + \dots, \end{aligned} \quad (43)$$

$$c_j = \frac{1}{\pi} \int_{-1}^1 [p(x)\overline{D}_j(x) + \tau(x)\overline{G}_j(x)] dx + \frac{\delta_{j0}}{2} q^0 (1 - e^{-2i\alpha}), \quad j = 1, 2,$$

$$c_j^r = \operatorname{Re}(c_j), \quad c_j^i = \operatorname{Im}(c_j),$$

where the kernels are determined according to the formulas

$$\begin{aligned} D_0(x) &= \frac{i}{2} \left[-\frac{1}{x-z^0} + \frac{1}{x-\overline{z}^0} - \frac{e^{-2i\alpha}(\overline{z}^0 - z^0)}{(x-\overline{z}^0)^2} \right], \quad G_0(x) = \frac{1}{2} \left[\frac{1}{x-z^0} \right. \\ &\quad \left. + \frac{1-e^{-2i\alpha}}{x-\overline{z}^0} - \frac{e^{-2i\alpha}(x-z^0)}{(x-\overline{z}^0)^2} \right], \quad D_1(x) = \frac{ie^{-i\alpha}}{2(x-\overline{z}^0)^2} \left[1 - e^{-2i\alpha} \right. \\ &\quad \left. - \frac{2e^{-2i\alpha}(\overline{z}^0 - z^0)}{x-\overline{z}^0} \right] + \frac{ie^{i\alpha}}{2} \left[-\frac{1}{(x-z^0)^2} + \frac{e^{-2i\alpha}}{(x-\overline{z}^0)^2} \right], \quad G_1(x) = \frac{e^{-i\alpha}}{2(x-\overline{z}^0)^2} \left[1 \right. \\ &\quad \left. - e^{-2i\alpha} - \frac{2e^{-2i\alpha}(x-z^0)}{x-\overline{z}^0} \right] + \frac{e^{i\alpha}}{2} \left[\frac{1}{(x-z^0)^2} + \frac{e^{-2i\alpha}}{(x-\overline{z}^0)^2} \right], \end{aligned} \quad (44)$$

and $\theta(x)$ is the step function ($\theta(x) = -1$ for $x < 0$ and $\theta(x) = 1$ for $x \geq 0$).

3.2. Comparison of analytical asymptotic and numerical solutions for small subsurface cracks

Let us compare the asymptotically (k_{1a}^{\pm} and k_{2a}^{\pm}) and numerically (k_{1n}^{\pm} and k_{2n}^{\pm}) obtained solutions of the problem for the case when $y^0 = -0.4$, $\delta_0 = 0.1$, $\alpha = \pi/2$, $\lambda = 0.1$, and $q^0 = -0.005$. The numerical method used for calculating k_{1n}^{\pm} and k_{2n}^{\pm} is described in detail in [1]. Both the numerical and asymptotic solutions are represented in Fig. 3. It follows from Fig. 3 that the asymptotic and numerical solutions are almost identical except for the region where the numerically obtained $k_1^+(x^0)$ is close to zero. The difference is mostly caused by the fact that the used asymptotic solution involve only two terms, i.e., the accuracy of these asymptotic solutions is $O(\delta_0^2)$ for small δ_0 . However, according to the two-term asymptotic solutions the maximum values of k_1^{\pm} differ from the numerical ones by no more than 1.4%. One can expect to get much higher precision if $\delta_0 < 0.1$ and $|y^0| \gg \delta_0$.

Therefore, formulas (43) and (44) provide sufficient precision for most possible applications and can be used to substitute for numerically obtained values of k_1^{\pm} and k_2^{\pm} .

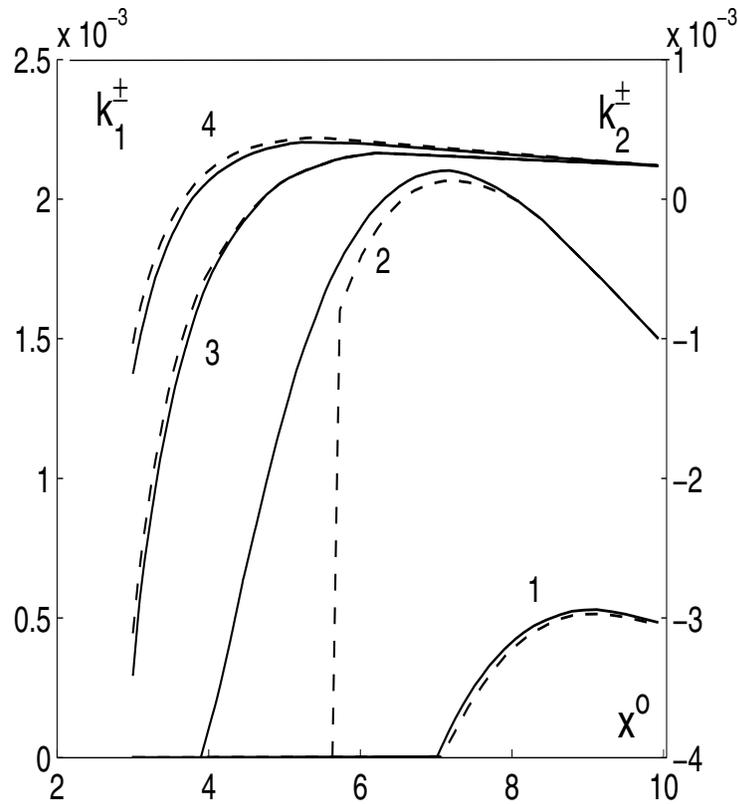


Figure 3. Comparison of the two-term asymptotic expansions k_{1a}^\pm and k_{2a}^\pm with the numerically calculated stress intensity factors k_{1n}^\pm and k_{2n}^\pm obtained for $y^0 = -0.4$, $\delta_0 = 0.1$, $\alpha = \pi/2$, $\lambda = 0.1$, and $q^0 = -0.005$. Solid curves are numerical results while dashed curves are asymptotical results. (k_{1n}^- group 1, k_{1n}^+ group 2, k_{2n}^- group 3, k_{2n}^+ group 4) (after Kudish [15]). Reprinted with permission of the STLE.

3.3. Stress intensity factors k_{1n}^\pm and k_{2n}^\pm behavior for subsurface cracks

Some examples of the behavior of the stress intensity factors k_{1n}^\pm and k_{2n}^\pm for subsurface cracks are presented below.

It is important to keep in mind that for the cases of no friction ($\lambda = 0$) and compressive or zero residual stress ($q^0 \leq 0$) all subsurface cracks are closed and, therefore, at their tips $k_{1n}^\pm = 0$.

Let us consider the case when the residual stress q^0 is different from zero. The residual stress influence on k_{1n}^+ results in increase of k_{1n}^+ for a tensile residual stress $q^0 > 0$ or its decrease for a compressive residual stress $q^0 < 0$ of the material region with tensile stresses. From formulas (43), (44), and Fig. 4 (obtained for $y_n^0 = -0.2$, $\alpha_n = \pi/2$, and $\delta_n = 0.1$) follows that for all x_n^0 and for increasing residual stress q^0 (see the curves marked with 3 and 5 that correspond to $\lambda = 0.1$, $q^0 = 0.04$, and $\lambda = 0.2$, $q^0 = 0.02$, respectively) the stress intensity factor k_{1n}^+ is a non-decreasing function of q^0 . Moreover, if at some material point $k_{1n}^+(q_1^0) > 0$ for some residual stress q_1^0 , then $k_{1n}^+(q_2^0) > k_{1n}^+(q_1^0)$ for $q_2^0 > q_1^0$ (compare curves marked with 1 and 2 with curves marked with 3 and 4 as well as with curves marked with 5 and 6, respectively). Similarly, for all x_n^0 when the magnitude of the compressive residual stress ($q^0 < 0$) increases (see curves marked with 4 and 6 that correspond to $\lambda = 0.1$, $q^0 = -0.01$ and

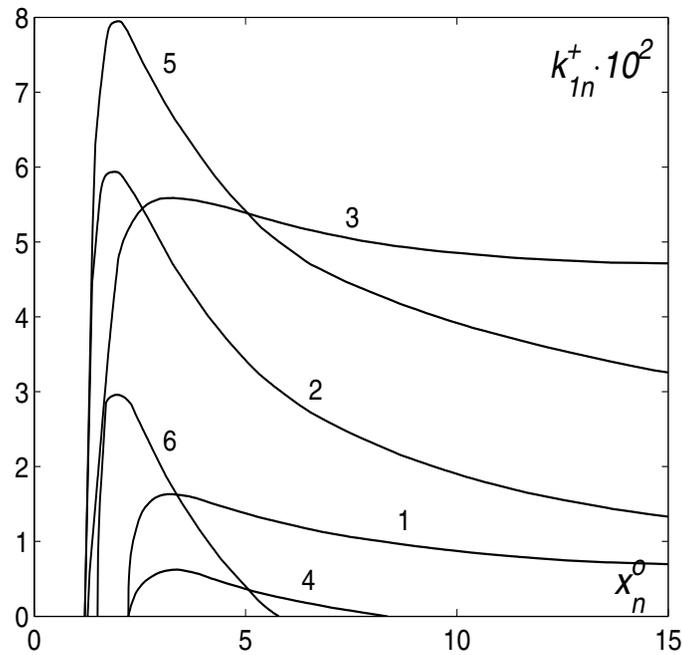


Figure 4. The dependence of the normal stress intensity factor k_{1n}^+ on x_n^0 for $\delta_n = 0.1$, $\alpha = \pi/2$, $y_n^0 = -0.2$ and different levels of the residual stress q^0 . Curves 1 and 2 are obtained for $q^0 = 0$ and $\lambda = 0.1$ and $\lambda = 0.2$, respectively. Curves 3 and 4 are obtained for $\lambda = 0.1$, $q^0 = 0.04$ and $q^0 = -0.01$, respectively, while curves 5 and 6 are obtained for $\lambda = 0.2$, $q^0 = 0.01$ and $q^0 = -0.03$, respectively (after Kudish [5]). Reprinted with permission of Springer.

$\lambda = 0.2$, $q^0 = -0.03$, respectively) and at some material point $k_{1n}^+(q_1^0) > 0$ for some residual stress q_1^0 then $k_{1n}^+(q_2^0) < k_{1n}^+(q_1^0)$ for $q_2^0 < q_1^0$. Based on the fact that the normal stress intensity factor k_{1n}^+ is positive only in the near surface material layer, we can make a conclusion that this layer increases in size and, starting with a certain value of tensile residual stress $q^0 > 0$, crack propagation becomes possible at any depth beneath the half-plane surface. For increasing compressive residual stresses, the thickness of the material layer where $k_{1n}^+ > 0$ decreases.

The analysis of the results for subsurface cracks following from formulas (43) and (44) shows [1] that the values of the stress intensity factors k_{1n}^\pm and k_{2n}^\pm are insensitive to even relatively large variations in the behavior of the distributions of the pressure $p(x^0)$ and frictional stress $\tau(x^0)$. In particular, the stress intensity factors k_{1n}^\pm and k_{2n}^\pm for the cases of dry and fluid (lubricant) friction as well as for the cases of constant pressure and frictional stress are very close to each other as long as the normal force (integral of $p(x^0)$ over the contact region) and the friction force (integral of τ over the contact region) applied to the surface of the half-plane are the same [1].

Qualitatively, the behavior of the normal stress intensity factors k_{1n}^\pm for different angles of orientation α_n is very similar while quantitatively it is very different. An example of that is presented for a horizontal ($\alpha_n = 0$) subsurface crack in Fig. 5 and for a subsurface crack perpendicular to the half-plane boundary ($\alpha_n = \pi/2$) in Fig 6 for $p(x^0) = \pi/4$, $\tau(x^0) = -\lambda p(x^0)$, $y_n^0 = -0.2$, $q^0 = 0$ for $\lambda = 0.1$ and $\lambda = 0.2$.

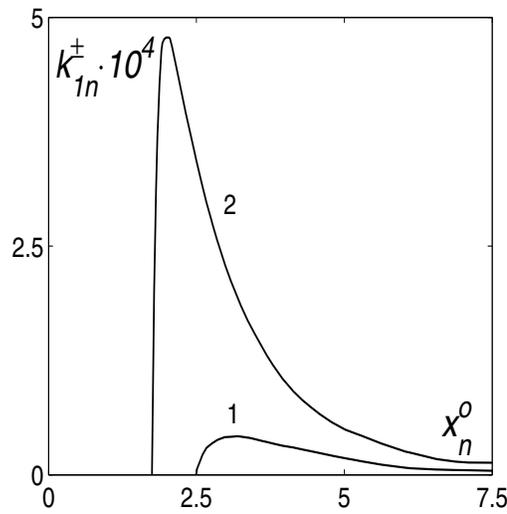


Figure 5. The dependence of the normal stress intensity factor k_{1n}^{\pm} on the coordinate x_n^0 for the case of the boundary of half-plane loaded with normal $p(x^0) = \pi/4$ and frictional $\tau(x^0) = -\lambda p(x^0)$ stresses, $y_n^0 = -0.2, \alpha_n = 0, q^0 = 0$: $\lambda = 0.1$ - curve marked with 1, $\lambda = 0.2$ - curve marked with 2 (after Kudish and Covitch [1]). Reprinted with permission from CRC Press.

For the same loading and crack parameters the behavior of the shear stress intensity factor k_{2n}^{\pm} is represented in Fig. 7 and 8. It is important to observe that the shear stress intensity factors k_{2n}^{\pm} are insensitive to changes of the friction coefficient λ .

Fig. 5 and 6 clearly show that for subsurface cracks with angle $\alpha_n = \pi/2$ for zero or tensile residual stress q^0 the normal stress intensity factors k_{1n}^{\pm} are significantly higher (by two orders of magnitude) than the ones for $\alpha_n = 0$. For $\alpha_n = 0$ and $\alpha_n = \pi/2$ the orders of magnitude of the shear stress intensity factors k_{2n}^{\pm} are the same (see Fig. 7 and 8). Moreover, from these graphs it is clear that the normal stress intensity factors k_{1n}^{\pm} are significantly influenced by the friction coefficient λ while the shear stress intensity factors k_{2n}^{\pm} are insensitive to the value of the friction coefficient λ .

Obviously, in the single-term approximation the behavior of k_{1n}^{0-} and k_{2n}^{0-} is identical to the one of k_{1n}^{0+} and k_{2n}^{0+} , respectively. Generally, the difference between k_{1n}^{0-} and k_{1n}^{0+} as well as between k_{2n}^{0-} and k_{2n}^{0+} is of the order of magnitude of $\delta_0 \ll 1$.

3.4. Lubricant-surface crack interaction. Stress intensity factors k_{1n}^{\pm} and k_{2n}^{\pm} Behavior

The process of lubricant-surface crack interaction is very complex and the details of the problem formulation, the numerical solution approach, and a comprehensive analysis of the results can be found in [1]. Therefore, here we will discuss only the most important features of this phenomenon. It is well known that the presence of lubricant between surfaces in contact

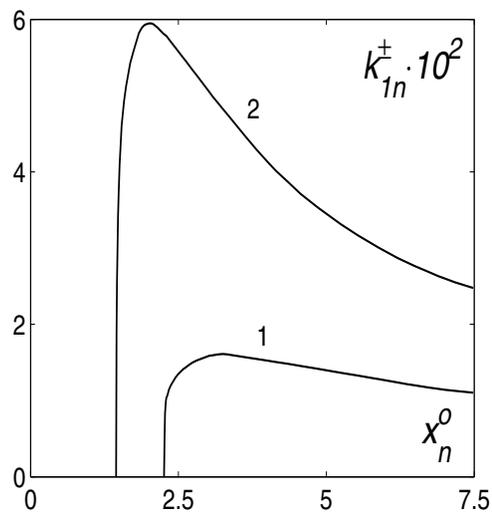


Figure 6. The dependence of the normal stress intensity factor k_{1n}^{\pm} on the coordinate x_n^0 for the case of the boundary of half-plane loaded with normal $p(x^0) = \pi/4$ and frictional $\tau(x^0) = -\lambda p(x^0)$ stresses, $y_n^0 = -0.2$, $\alpha_n = \pi/2$, $q^0 = 0$: $\lambda = 0.1$ - curve marked with 1, $\lambda = 0.2$ - curve marked with 2 (after Kudish and Covitch [1]). Reprinted with permission from CRC Press.

is very beneficial as it reduces the contact friction and wear and facilitates better heat transfer from the contact. However, in some cases the lubricant presence may play a detrimental role. In particular, in cases when the elastic solid (half-plane) has a surface crack inclined toward the incoming high contact pressure transmitted through the lubricant. Such a crack may open up and experience high lubricant pressure applied to its faces. This pressure creates the normal stress intensity factor k_{1n}^- far exceeding the value of the stress intensity factors k_{1n}^{\pm} for comparable subsurface cracks while the shear stress intensity factors for surface k_{2n}^- and and subsurface k_{2n}^{\pm} cracks remain comparable in value. That becomes obvious from the comparison of the graphs from Fig. 4-8 with the graphs from Fig. 10 and 9.

In cases when a surface crack is inclined away from the incoming high lubricant pressure the crack does not open up toward the incoming lubricant with high pressure and it behaves similar to a corresponding subsurface crack, i.e. its normal stress intensity factor k_{1n}^- in its value is similar to the one for a corresponding subsurface crack. It is customary to see the normal stress intensity factor for surface cracks which open up toward the incoming high lubricant pressure to exceed the one for comparable subsurface cracks by two orders of magnitude. In such cases the compressive residual stress has very little influence on the crack behavior due to domination of the lubricant pressure.

The possibility of high normal stress intensity factors for surface cracks leads to serious consequences. In particular, it explains why fatigue life of drivers is usually significantly higher than the one for followers [1]. Also, it explains why under normal circumstances fatigue failure is of subsurface origin [1].

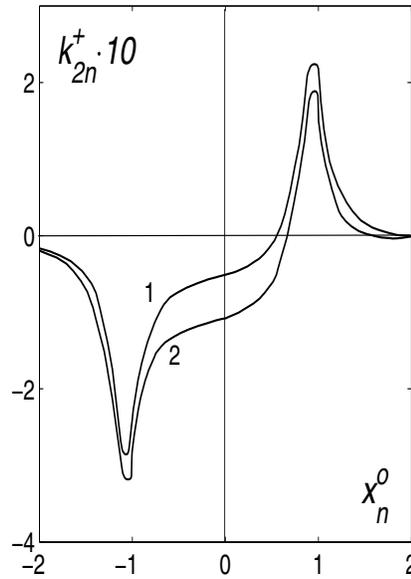


Figure 7. The dependence of the shear stress intensity factor k_{2n}^+ on the coordinate x_n^0 for the case of the boundary of half-plane loaded with normal $p(x^0) = \pi/4$ and frictional $\tau(x^0) = -\lambda p(x^0)$ stresses, $y_n^0 = -0.2, \alpha_n = 0, q^0 = 0$: $\lambda = 0.1$ - curve marked with 1, $\lambda = 0.2$ - curve marked with 2 (after Kudish and Covitch [1]). Reprinted with permission from CRC Press.

4. Stress intensity factors and directions of fatigue crack propagation

The process of fatigue failure is usually subdivided into three major stages: the nucleation period, the period of slow pre-critical fatigue crack growth, and the short period of fast unstable crack growth ending in material losing its integrity. The durations of the first two stages of fatigue failure depend on a number of parameters such as material properties, specific environment, stress state, temperature, etc. Usually, the nucleation period is short [1]. We are interested in the main part of the process of fatigue failure which is due to slow pre-critical crack growth. In these cases $\max(k_1^+, k_1^-) < K_f$, where K_f is the material fracture toughness.

During the pre-critical fatigue crack growth cracks remain small. Therefore, they can be modeled by small straight cuts in the material. For such small subsurface cracks it is sufficient to use the one-term asymptotic approximations $k_1^\pm = k_1$ and $k_2^\pm = k_2$ for the stress intensity factors from (43) and (44) which in dimensional variables take the form

$$\begin{aligned}
 k_1 &= \sqrt{l}[Y^r + q^0 \sin^2 \alpha]\theta[Y^r + q^0 \sin^2 \alpha], \quad k_2 = \sqrt{l}[Y^i - \frac{q^0}{2} \sin 2\alpha], \\
 Y &= \frac{1}{\pi} \int_{-a_H}^{a_H} [p(t)\overline{D}_0(t) + \tau(t)\overline{G}_0(t)]dt, \quad \tau = -\lambda p, \\
 \{Y^r, Y^i\} &= \{Re(Y), Im(Y)\}, \\
 D_0(t) &= \frac{i}{2} \left[-\frac{1}{t-X} + \frac{1}{t-\overline{X}} - \frac{e^{-2i\alpha}(\overline{X}-X)}{(t-\overline{X})^2} \right], \\
 G_0(t) &= \frac{1}{2} \left[\frac{1}{t-\overline{X}} + \frac{1-e^{-2i\alpha}}{t-X} - \frac{e^{-2i\alpha}(t-X)}{(t-\overline{X})^2} \right], \quad X = x + iy,
 \end{aligned}
 \tag{45}$$

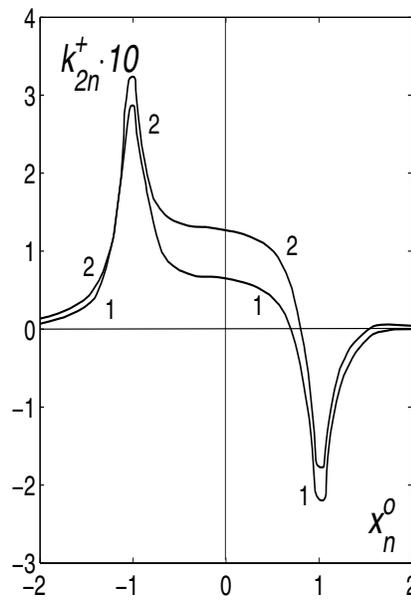


Figure 8. The dependence of the shear stress intensity factor k_{2n}^+ on the coordinate x_n^0 for the case of the boundary of half-plane loaded with normal $p(x^0) = \pi/4$ and frictional $\tau(x^0) = -\lambda p(x^0)$ stresses, $y_n^0 = -0.2$, $\alpha_n = \pi/2$, $q^0 = 0$: $\lambda = 0.1$ - curve marked with 1, $\lambda = 0.2$ - curve marked with 2 (after Kudish and Covitch [1]). Reprinted with permission from CRC Press.

where i is the imaginary unit ($i^2 = -1$), $\theta(x)$ is a step function: $\theta(x) = 0$, $x \leq 0$ and $\theta(x) = 1$, $x > 0$. It is important to mention that according to (45) for subsurface cracks the quantities of $k_{10} = k_1 l^{-1/2}$ and $k_{20} = k_2 l^{-1/2}$ are functions of x and y and are independent from l .

Numerous experimental studies have established the fact that at relatively low cyclic loads materials undergo the process of pre-critical failure while the rate of crack growth dl/dN (N is the number of loading cycles) in the predetermined direction is dependent on k_1^\pm and K_f . A number of such equations of pre-critical crack growth and their analysis are presented in [6]. However, what remains to be determined is the direction of fatigue crack growth.

Assuming that fatigue cracks growth is driven by the maximum principal tensile stress (see the section on Three-Dimensional Model of Contact and Structural Fatigue) we immediately obtain the equation

$$k_2^\pm(N, x, y, l, \alpha^\pm) = 0, \quad (46)$$

which determines the orientation angles α^\pm of a fatigue crack growth at the crack tips. Due to the fact that a fatigue crack remains small during its pre-critical growth (i.e. practically during its entire life span) and being originally modeled by a straight cut with half-length l at the point with coordinates (x, y) the crack remains straight, i.e. the crack direction is characterized by one angle $\alpha = \alpha^+ = \alpha^-$. This angle is practically independent from crack half-length l because $k_2^\pm = k_{20}^\pm \sqrt{l}$, where k_{20}^\pm is almost independent from l for small l (see (45)). The dependence of k_2^\pm on the number of loading cycles N comes only through the dependence of the crack half-length l on N . Therefore, the crack angle α is just a function of the crack location (x, y) .

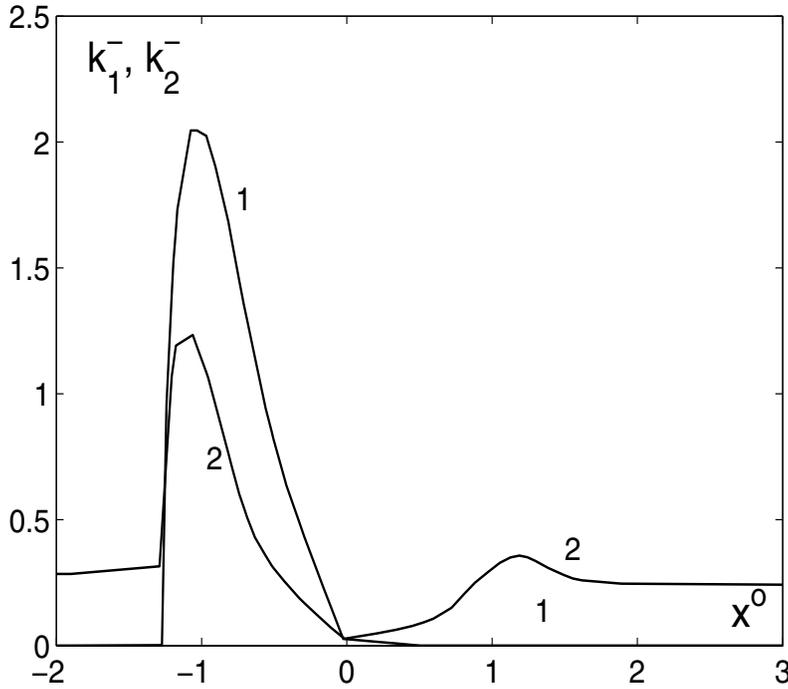


Figure 9. Distributions of the stress intensity factors k_1^- (curve 1) and k_2^- (curve 2) in case of a surface crack: $\alpha = 0.339837$, $\delta_0 = 0.3$, and other parameters as in the previous case of a surface crack (after Kudish [15]). Reprinted with permission of the STLE.

In particular, according to (45) and (46) at any point (x, y) there are two angles α_1 and α_2 along which a crack may propagate which are determined by the equation in dimensional variables

$$\tan 2\alpha = - \frac{2y \int_{-a_H}^{a_H} (t-x)T(t,x,y)dt}{\frac{\pi}{2}q^0 + \int_{-a_H}^{a_H} [(t-x)^2 - y^2]T(t,x,y)dt}, \quad T(t, x, y) = \frac{yp(t) + (t-x)\tau(t)}{[(t-x)^2 + y^2]^2}. \quad (47)$$

Along these directions k_1 reaches its extremum values. The actual direction of crack propagation α is determined by one of these two angles α_1 and α_2 for which the value of the normal stress intensity factor $k_1(N, x, y, l, \alpha)$ is greater.

A more detailed analysis of the directions of fatigue crack propagation can be found in [1].

5. Two-dimensional contact fatigue model

In a two-dimensional case compared to a three-dimensional case a more accurate description of the contact fatigue process can be obtained due to the fact that in two dimensions it is relatively easy to get very accurate formulas for the stress intensity factors at crack tips [1, 5]. The rest of the fatigue modeling can be done the same way as in the three-dimensional case with few simple changes. In particular, in a two-dimensional case of contact fatigue only subsurface originated fatigue is considered and cracks are modeled by straight cuts with half-length l . That gives the opportunity to use equations (45) and (47) for stress intensity

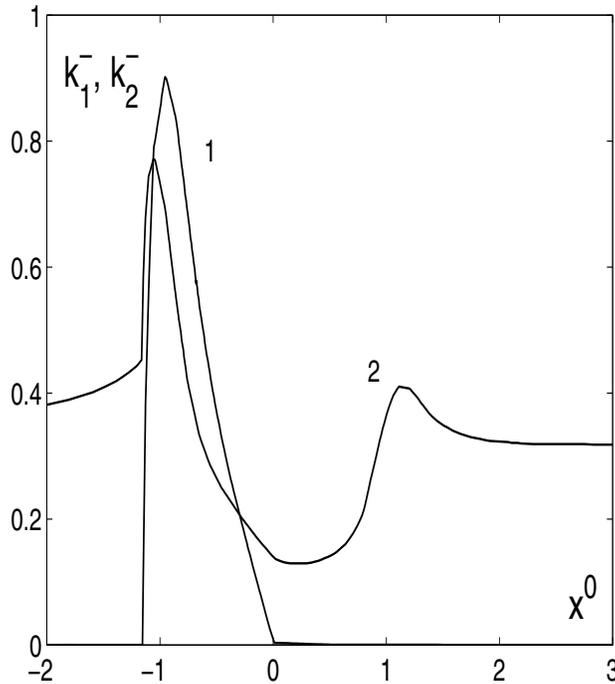


Figure 10. Distributions of the stress intensity factors k_1^- (curve 1) and k_2^- (curve 2) in case of a surface crack: $\alpha = \pi/6$, $\lambda = 0.1$, and $\eta^0 = -0.5$ (after Kudish and Burris [16]). Reprinted with permission from Kluwer Academic Publishers.

factors k_1 , k_2 , and crack angle orientations α . The rest of contact fatigue modeling follows exactly the derivation presented in the section on Three-Dimensional Model of Contact and Structural Fatigue. Therefore, the fatigue life N with survival probability P_* of a contact subjected to cyclic loading can be expressed in the form [1, 17]

$$N = \left\{ \left(\frac{n}{2} - 1 \right) g_0 \left[\max_{-\infty < x < \infty} k_{10} \right]^n \right\}^{-1} \left\{ \exp \left[\left(1 - \frac{n}{2} \right) (\mu_{ln} + \sqrt{2} \sigma_{ln} \operatorname{erf}^{-1}(2P_* - 1)) \right] - l_c^{\frac{2-n}{2}} \right\}, \quad (48)$$

where $\operatorname{erf}^{-1}(x)$ is the inverse function to the error integral $\operatorname{erf}(x)$ [8].

First, let us consider the model behavior in some simple cases. If $f(0, x, y, z, l_0)$ is a uniform crack distribution over the material volume V (except for a thin surface layer where $f = 0$). Then based on (11) it can be shown that $p(N, x, y, z)$ reaches its minimum at the points where k_{10} and the principal tensile stress reach their maximum values. This leads to the conclusion that the material local failure probability $(1 - p)$ reaches its maximum at the points with maximal tensile stress. Therefore, for a uniform initial crack distribution $f(0, x, y, z, l_0)$ the survival probability $P(N)$ from (16) is determined by the material local survival probability at the points at which the maximal tensile stress is attained.

However, the latter conclusion is not necessarily correct if the initial crack distribution $f(0, x, y, z, l_0)$ is not uniform over the material volume. Suppose, $k_{10}(x, y, z)$ is maximal at

the point (x_m, y_m, z_m) and at the initial time moment $N = 0$ at some point (x_*, y_*, z_*) there exist cracks larger than the ones at the point (x_m, y_m, z_m) , namely,

$$\int_0^{l_c} f dl_0 |_{(x_*, y_*, z_*)} < \int_0^{l_c} f dl_0 |_{(x_m, y_m, z_m)} .$$

Then after a certain number of loading cycles $N > 0$ the material damage at point (x_*, y_*, z_*) may be greater than at point (x_m, y_m, z_m) , where l_{0c} reaches its maximum value. Therefore, fatigue failure may occur at the point (x_*, y_*, z_*) instead of the point (x_m, y_m, z_m) , and the material weakest point is not necessarily is the material most stressed point.

If μ_{ln} and σ_{ln} depend on the coordinates of the material point (x, y, z) , then there may be a series of points where in formula (16) for the given number of loading cycles N the minimum over the material volume V is reached. The coordinates of such points may change with N . This situation represents different potentially competing fatigue mechanisms such as pitting, flaking, etc. The occurrence of fatigue damage at different points in the material depends on the initial defect distribution, applied stresses, residual stress, etc.

In the above model of contact fatigue the stressed volume V plays no explicit role. However, implicitly it does. In fact, the initial crack distribution $f(0, x, y, z, l_0)$ depends on the material volume. In general, in a larger volume of material, there is a greater chance to find inclusions/cracks of greater size than in a smaller one. These larger inclusions represent a potential source of pitting and may cause a decrease in the material fatigue life of a larger material volume.

Assuming that μ_{ln} and σ_{ln} are constants, and assuming that the material failure occurs at the point (x, y, z) with the failure probability $1 - P(N)$ following the considerations of the section on Three-Dimensional Model of Contact and Structural Fatigue from (48) we obtain formulas (19)-(21). Actually, fatigue life formulas (20) and (21) can be represented in the form of the Lundberg-Palmgren formula, i.e.

$$N = \frac{C_*}{p_H^n}, \quad (49)$$

where parameter n can be compared with constant c/e in the Lundberg-Palmgren formula [1]. The major difference between the Lundberg-Palmgren formula and formula (49) derived from this model of contact fatigue is the fact that in (49) constant C_* depends on material defect parameters μ, σ , coefficient of friction λ , residual stresses q^0 , and probability of survival P_* in a certain way while in the Lundberg-Palmgren formula the constant C_* depends only on the depth z_0 of the maximum orthogonal stress, stressed volume V , and probability of survival P_* .

Let us analyze formula (21). It demonstrates the intuitively obvious fact that the fatigue life N is inverse proportional to the value of the parameter g_0 that characterizes the material crack propagation resistance. So, for materials with lower crack propagation rate, the fatigue life is higher and vice versa. Equation (21) exhibits a usual for gears and roller and ball bearings dependence of fatigue life N on the maximum Hertzian pressure p_H . Thus, from the well-known experimental data for bearings, the range of n values is $20/3 \leq n \leq 9$. Keeping in mind that usually $\sigma \ll \mu$, for these values of n contact fatigue life N is practically inverse proportional to a positive power of the crack initial mean size, i.e., to $\mu^{n/2-1}$. Therefore, fatigue life N is a decreasing function of the initial mean crack (inclusion) size μ and $\ln N = -(n/2 - 1) \ln \mu + \text{constant}$. This conclusion is valid for any value of the material survival

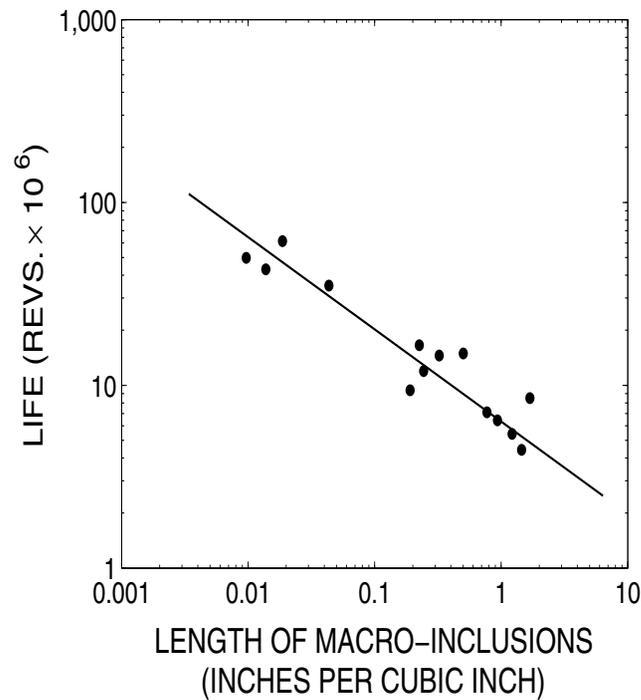


Figure 11. Bearing life-inclusion length correlation (after Stover and Kolarik II[18], COPYRIGHT The Timken Company 2012).

probability P_* and is supported by the experimental data obtained at the Timken Company by Stover and Kolarik II [18] and represented in Fig. 11 in a log-log scale. If $P_* > 0.5$ then $\text{erf}^{-1}(2P_* - 1) > 0$ and (keeping in mind that $n > 2$) fatigue life N is a decreasing function of the initial standard deviation of crack sizes σ . Similarly, if $P_* < 0.5$, then fatigue life N is an increasing function of the initial standard deviation of crack sizes σ . According to (20), for $P_* = 0.5$ fatigue life N is independent from σ , and, according to (21), for $P_* = 0.5$ fatigue life N is a slowly increasing function of σ . By differentiating $p_m(N)$ obtained from (16) with respect to σ , we can conclude that the dispersion of $P(N)$ increases with σ .

From (45) (also see Kudish [1]) follows that the stress intensity factor k_1 decreases as the magnitude of the compressive residual stress q^0 increases and/or the magnitude of the friction coefficient λ decreases. Therefore, it follows from formulas (45) that C_0 is a monotonically decreasing function of the residual stress q^0 and friction coefficient λ . Numerical simulations of fatigue life show that the value of C_0 is very sensitive to the details of the residual stress distribution q^0 versus depth.

Let us choose a basic set of model parameters typical for bearing testing: maximum Hertzian pressure $p_H = 2 \text{ GPa}$, contact region half-width in the direction of motion $a_H = 0.249 \text{ mm}$, friction coefficient $\lambda = 0.002$, residual stress varying from $q^0 = -237.9 \text{ MPa}$ on the surface to $q^0 = 0.035 \text{ MPa}$ at the depth of $400 \mu\text{m}$ below it, fracture toughness K_f varying between 15 and $95 \text{ MPa} \cdot \text{m}^{1/2}$, $g_0 = 8.863 \text{ MPa}^{-n} \cdot \text{m}^{1-n/2} \cdot \text{cycle}^{-1}$, $n = 6.67$, mean of crack initial half-lengths $\mu = 49.41 \mu\text{m}$ ($\mu_{ln} = 3.888 + \ln(\mu\text{m})$), crack initial standard deviation $\sigma = 7.61 \mu\text{m}$ ($\sigma_{ln} = 0.1531$). Numerical results show that the fatigue life is practically independent from the material fracture toughness K_f , which supports the assumption used for the derivation of formulas (19)-(21). To illustrate the dependence of contact fatigue life on some of the model

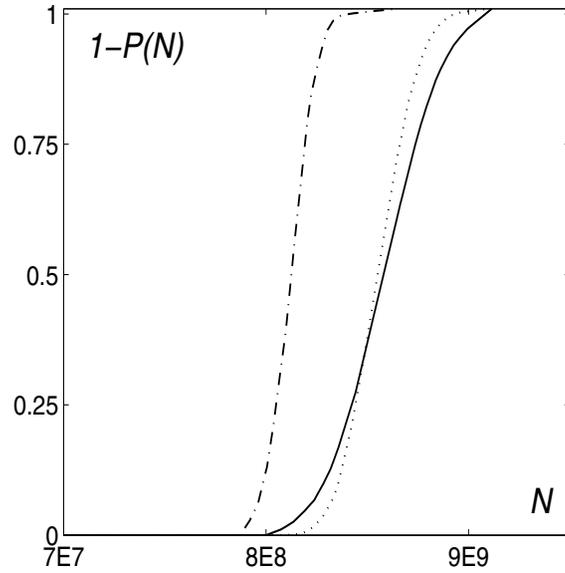


Figure 12. Pitting probability $1 - P(N)$ calculated for the basic set of parameters (solid curve) with $\mu = 49.41 \mu m$, $\sigma = 7.61 \mu m$ ($\mu_{ln} = 3.888 + \ln(\mu m)$, $\sigma_{ln} = 0.1531$), for the same set of parameters and the increased initial value of crack mean half-lengths (dash-dotted curve) $\mu = 74.12 \mu m$ ($\mu_{ln} = 4.300 + \ln(\mu m)$, $\sigma_{ln} = 0.1024$), and for the same set of parameters and the increased initial value of crack standard deviation (dotted curve) $\sigma = 11.423 \mu m$ ($\mu_{ln} = 3.874 + \ln(\mu m)$, $\sigma_{ln} = 0.2282$) (after Kudish [17]). Reprinted with permission from the STLE.

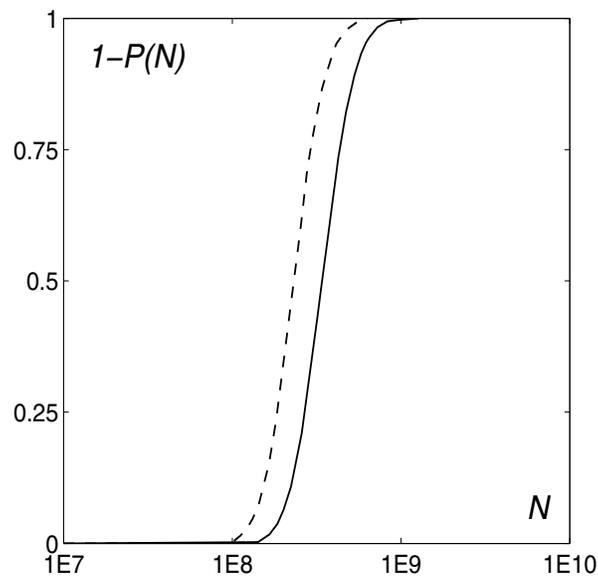


Figure 13. Pitting probability $1 - P(N)$ calculated for the basic set of parameters including $\lambda = 0.002$ (solid curve) and for the same set of parameters and the increased friction coefficient (dashed curve) $\lambda = 0.004$ (after Kudish [17]). Reprinted with permission from the STLE.

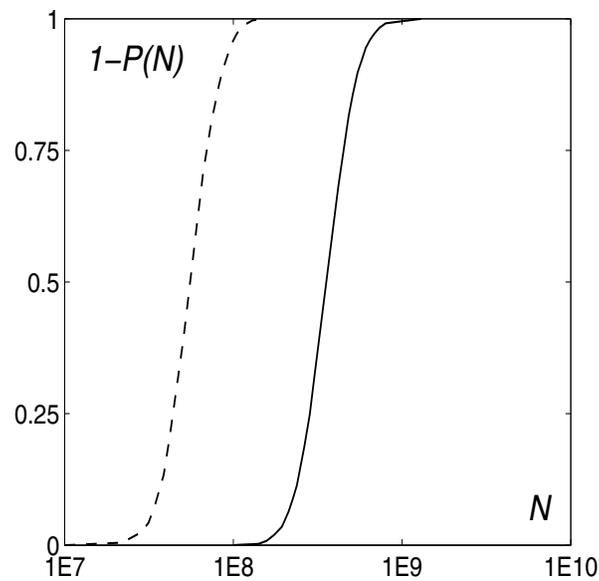


Figure 14. Pitting probability $1 - P(N)$ calculated for the basic set of parameters (solid curve) and for the same set of parameters and changed profile of residual stress q^0 (dashed curve) in such a way that at points where q^0 is compressive its magnitude is unchanged and at points where q^0 is tensile its magnitude is doubled (after Kudish [17]). Reprinted with permission from the STLE.

parameters, just one parameter from the basic set of parameters will be varied at a time and graphs of the pitting probability $1 - P(N)$ for these sets of parameters (basic and modified) will be compared. Figure 12 shows that as the initial values of the mean μ of crack half-lengths and crack standard deviation σ increase contact fatigue life N decreases. Similarly, contact fatigue life decreases as the magnitude of the tensile residual stress and/or friction coefficient increase (see Fig. 4 and 13). The results show that the fatigue life does not change when the magnitude of the compressive residual stress is increased/decreased by 20% of its base value while the tensile portion of the residual stress distribution remains the same. Obviously, that is in agreement with the fact that tensile stresses control fatigue. Moreover, the fatigue damage occurs in the region with the resultant tensile stresses close to the boundary between tensile and compressive residual stresses. However, when the compressive residual stress becomes small enough the acting frictional stress may supersede it and create new regions with tensile stresses that potentially may cause acceleration of fatigue failure.

μ [μm]	σ [μm]	$N_{15.9}$ [cycles]
49.41	7.61	$2.5 \cdot 10^8$
73.13	11.26	$1.0 \cdot 10^8$
98.42	15.16	$5.0 \cdot 10^7$
147.11	22.66	$2.0 \cdot 10^7$
244.25	37.62	$6.0 \cdot 10^6$

Table 1. Relationship between the tapered bearing fatigue life $N_{15.9}$ and the initial inclusion size mean and standard deviation (after Kudish [17]). Reprinted with permission from the STLE.

Let us consider an example of the further validation of the new contact fatigue model for tapered roller bearings based on a series of approximate calculations of fatigue life. The

main simplifying assumption made is that bearing fatigue life can be closely approximated by taking into account only the most loaded contact. The following parameters have been used for calculations: $p_H = 2.12 \text{ GPa}$, $a_H = 0.265 \text{ mm}$, $\lambda = 0.002$, $g_0 = 6.009 \text{ MPa}^{-n} \cdot \text{m}^{1-n/2} \cdot \text{cycle}^{-1}$, $n = 6.67$, the residual stress varied from $q^0 = -237.9 \text{ MPa}$ on the surface to $q^0 = 0.035 \text{ MPa}$ at the depth of $400 \mu\text{m}$ below the surface, fracture toughness K_f varied between 15 and $95 \text{ MPa} \cdot \text{m}^{1/2}$. The crack/inclusion initial mean half-length μ varied between 49.41 and $244.25 \mu\text{m}$ ($\mu_{ln} = 3.888 - 5.498 + \ln(\mu\text{m})$), the crack initial standard deviation varied between $\sigma = 7.61$ and $37.61 \mu\text{m}$ ($\sigma_{ln} = 0.1531$). The results for fatigue life $N_{15,9}$ (for $P(N_{15,9}) = P_* = 0.159$) calculations are given in the Table 1 and practically coincide with the experimental data obtained by The Timken Company and presented in Fig. 19 by Stover, Kolarik II, and Keener [?] (in the present text this graph is given as Fig. 11). One must keep in mind that there are certain differences in the numerically obtained data and the data presented in the above mentioned Fig. 11 due to the fact that in Fig. 11 fatigue life is given as a function of the cumulative inclusion length (sum of all inclusion lengths over a cubic inch of steel) while in the model fatigue life is calculated as a function of the mean inclusion length.

It is also interesting to point out that based on the results following from the new model, bearing fatigue life can be significantly improved for steels with the same cumulative inclusion length but smaller mean half-length μ (see Fig. 12). In other words, fatigue life of a bearing made from steel with large number of small inclusions is higher than of the one made of steel with small number of larger inclusions given that the cumulative inclusion length is the same in both cases. Moreover, bearing and gear fatigue lives with small percentage of failures (survival probability $P > 0.5$) for steels with the same cumulative inclusion length can also be improved several times if the width of the initial inclusion distribution is reduced, i.e., when the standard deviation σ of the initial inclusion distribution is made smaller (see Figure 12). Figures 13 and 14 show that the elevated values of the tensile residual stress are much more detrimental to fatigue life than greater values of the friction coefficient.

Finally, the described model is flexible enough to allow for replacement of the density of the initial crack distribution (see (1)) by a different function and of Paris's equation for fatigue crack propagation (see (4)) by another equation. Such modifications would lead to results on fatigue life varying from the presented above. However, the methodology, i.e., the way the formulas for fatigue life are obtain and the most important conclusions will remain the same.

This methodology has been extended on the cases of non-steady cyclic loading as well as on the case of contact fatigue of rough surfaces [1]. Also, this kind of modeling approach has been applied to the analysis of wear and contact fatigue in cases of lubricant contaminated by rigid abrasive particles and contact surfaces charged with abrasive particles [19] as well as to calculation of bearing wear and contact fatigue life [20].

In conclusion we can state that the presented statistical contact and structural fatigue models take into account the most important parameters of the contact fatigue phenomenon (such as normal and frictional contact and residual stresses, initial statistical defect distribution, orientation of fatigue crack propagation, material fatigue resistance, etc.). The models allows for examination of the effect of variables such as steel cleanliness, applied stresses, residual stress, etc. on contact fatigue life as single or composite entities. Some analytical results illustrating these models and their validation by the experimentally obtained fatigue life data for tapered bearings are presented.

6. Closure

The chapter presents a detailed analysis of a number of plane crack mechanics problems for loaded elastic half-plane weakened by a system of cracks. Surface and subsurface cracks are considered. All cracks are considered to be straight cuts. The problems are analyzed by the regular asymptotic method and numerical methods. Solutions of the problems include the stress intensity factors. The regular asymptotic method is applied under the assumption that cracks are far from each other and from the boundary of the elastic solid. It is shown that the results obtained for subsurface cracks based on asymptotic expansions and numerical solutions are in very good agreement. The influence of the normal and tangential contact stresses applied to the boundary of a half-plane as well as the residual stress on the stress intensity factors for subsurface cracks is analyzed. It is determined that the frictional and residual stresses provide a significant if not the predominant contribution to the problem solution. Based on the numerical solution of the problem for surface cracks in the presence of lubricant the physical nature of the "wedge effect" (when lubricant under a sufficiently large pressure penetrates a surface crack and ruptures it) is considered. Solution of this problem also provides the basis for the understanding of fatigue crack origination site (surface versus subsurface) and the difference of fatigue lives of drivers and followers. New two- and three-dimensional models of contact and structural fatigue are developed. These models take into account the initial crack distribution, fatigue properties of the solids, and growth of fatigue cracks under the properly determined combination of normal and tangential contact and residual stresses. The formula for fatigue life based on these models can be reduced to a simple formula which takes into account most of the significant parameters affecting contact and structural fatigue. The properties of these contact fatigue models are analyzed and the results based on them are compared to the experimentally obtained results on contact fatigue for tapered bearings.

7. References

- [1] Kudish, I.I. and Covitch, M.J., 2010. *Modeling and Analytical Methods in Tribology*. Boca Raton, London, New York: CRC Press, Taylor & Francis Group.
- [2] Kudish, I.I. 2007. Fatigue Modeling for Elastic Materials with Statistically Distributed Defects. *ASME J. Appl. Mech.* 74:1125-1133.
- [3] Bokman, M.A., Pshenichnov, Yu.P., and Pershtein, E.M. 1984. The Microcrack and Non-metallic Inclusion Distribution in Alloy D16 after a Plastic Strain. *Plant Laboratory, Moscow*, 11:71-74.
- [4] Cherepanov, G.P. 1979. *Mechanics of Brittle Fracture*. New York: McGraw-Hill.
- [5] Kudish, I.I. 1987. Contact Problem of the Theory of Elasticity for Pre-stressed Bodies with Cracks. *J. Appl. Mech. and Techn. Phys.* 28, 2:295-303.
- [6] Yarema, S.Ya. 1981. Methodology of Determining the Characteristics of the Resistance to Crack Development (Crack Resistance) of Materials in Cyclic Loading. *J. Soviet Material Sci.* 17, No. 4:371-380.
- [7] Tallian, T., Hoeprich, M., and Kudish, I. 2001. Author's Closure. *STLE Tribology Trans.* 44, No. 2:153-155.
- [8] *Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables*, Eds. M. Abramowitz and I.A. Stegun, National Bureau of Standards, 55, 1964.
- [9] Stover, J.D., Kolarik II, R.V., and Keener, D.M. 1989. The Detection of Aluminum Oxide Stringers in Steel Using an Ultrasonic Measuring Method. *Mechanical Working and Steel*

- Processing XXVII: Proc. 31st Mechanical Working and Steel Processing Conf.*, Chicago, Illinois, October 22-25, 1989, Iron and Steel Soc., Inc., 431-440.
- [10] Lurye, A.I. 1970. *Theory of Elasticity*. Moscow: Nauka.
- [11] Hasebe, N. and Inohara, S. 1980. Stress Analysis of a Semi-infinite Plate with an Oblique Edge Crack. *Ing. Arch.* 49:51-62.
- [12] Isida, M. 1979. Tension of a Half Plane Containing Array Cracks, Branched Cracks and Cracks Emanating from Sharp Notches. *Trans. Jpn. Soc. Mech. Engrs.* 45, No. 392:306-317.
- [13] Savruk M.P. 1981. *Two – – Dimensional Problems of Elasticity for Bodies with Cracks*. Kiev: Naukova Dumka.
- [14] Vorovich, I.I., Alexandrov, V.M., and Babesko, V.A. 1974. *Non-Classic Mixed Problems in the Theory of Elasticity*. Moscow: Nauka.
- [15] Kudish I.I. 2002. Lubricant-Crack Interaction, Origin of Pitting, and Fatigue of Drivers and Followers. *STLE Tribology Trans.* 45:583-594.
- [16] Kudish I.I. and Burris, K.W. 2004. Modeling of Surface and Subsurface Crack Behavior under Contact Load in the Presence of Lubricant. *Intern. J. Fract.* 125:125-147.
- [17] Kudish, I.I. 2000. A New Statistical Model of Contact Fatigue. *STLE Tribology Trans.* 43:711-721.
- [18] Stover, J.D. and Kolarik II, R.V. 1987. The Evaluation of Improvements in Bearing Steel Quality Using an Ultrasonic Macro-Inclusion Detection Method. *The Timken Company Technical Note*, (January 1987), 1-12.
- [19] Kudish, I.I. 1991. Wear and Fatigue Pitting Taking into Account Contaminated Lubricants and Abrasive Particles Indented into Working Surfaces, *Journal of Friction and Wear*, Vol. 12, No. 3:713-725.
- [20] Kudish, I.I. 1990. Statistical Calculation of Rolling Bearing Wear and Pitting. *Soviet J. Frict. and Wear* 11:71-86.