Computational Analysis of Surface and Subsurface Initiated Fatigue Crack Growth due to Contact Loading

S. Glodež [1], B. Aberšek [1], G. Fajdiga [2] and J. Flašker [2]

Abstract: A computational model for simulation of surface and subsurface initiated fatigue crack growth due to contact loading is presented. The model is based on fracture mechanics theory where the required materials properties are obtained from common fatigue tests. For computational simulations an equivalent model of two contacting cylinders is used instead of simulating the actual contact of mechanical elements. The discretised model with the initial crack on or under the surface is then subjected to normal contact pressure, which takes into account the elasto-hydro-dynamic (EHD) lubrication conditions, and tangential loading due to friction between contacting surfaces. The model considers also the moving contact of mechanical elements and for the surface initiated crack also the fluid trapped in the crack. The virtual crack extension method, implemented in the finite element method is then used for simulating the fatigue crack growth from the initial crack up to the formation of the surface pit. The numerical results correspond well with available experimental data. The described model can be used for simulation of pitting phenomenon of contacting mechanical elements like gears, bearings, wheels, etc.

keyword: Contact fatigue, Fracture mechanics, Pitting, Computational analysis

1 Introduction

Mechanical elements subjected to rolling and sliding contact conditions like gears, bearings, wheels, etc fail by several mechanisms and the most prominent among these is surface pitting [Lester (1986)]. Although surface pitting is a well-known problem in engineering and many hypotheses and models have been proposed to-date, the general theory to realistically describe the complicated mechanism completely has yet to be established. Pitting strongly depends on surface finish, material microstructure and operating conditions, such as type of the contact, loading conditions, temperature, *etc.*, and it is obviously very difficult to take into account all the influential parameters.

Some early attempts to apply the fracture mechanics to the study of pit formation mechanism [Keer and Bryant (1983), Miller, Keer and Cheng (1985), Murakami and Kaneta (1987), Bower (1988)] led to some very comprehensive computational models to be proposed recently [Hanson and Keer (1992), Murakami, Sakae and Ichimaru (1994), Glodež, Ren and Fajdiga (2001), Flašker, Fajdiga, Glodež and Hellen (2001)]. These models from the onset assume that the crack is initiated on the surface due to significant friction forces. Under cyclic contact loading the crack then propagates at a shallow angle to the core by the combined action of shear forces and trapped fluid until the surface layer breaks away at a certain point.

In the case of high precision mechanical components with fine surface finish and good lubrication (small coefficient of friction) the fatigue crack is usually initiated under the contacting surface in the area of the largest contact stresses [Leng, Chen and Shao (1988), Cheng, Mura and Keer (1994), Glodež, Flašker and Ren (1997), Glodež, Ren and Flašker (1988)]. The mechanism of crack initiation and growth is here attributed to large shearing stresses at the onset of cracking and mixed mode fracture in the later stages of crack propagation when the crack approaches the surface.

The proposed model comprises both, surface and subsurface initiated fatigue crack. Its growth is then simulated using a two-dimensional computational model based on fracture mechanics where the required materials properties are obtained from common fatigue tests. It is recognised that surface pitting is a consequence of the dynamic character of contact loading and that it should be treated as such. However, here it is assumed that the fatigue cracks will always initiate in the region of the maximum contact stresses on or under the surface, dynamic

[1] University of Maribor, Faculty of Education
[2] University of Maribor, Faculty of Mechanical Engineering

or quasi-static. This enables quasi-static simulations of fatigue crack propagation for the worst loading case of cyclic load on top of the crack as it is described here.

For computational simulations of surface pitting it is advantageous to use an equivalent model of two contacting cylinders instead of simulating the actual contact of mechanical elements. Using the equivalent contact model of two cylinders the contact stresses can be determined by utilising the Hertzian contact theory [Johnson (1985)]. The equivalent cylinders have the same radii as are the curvature radii of the treated mechanical elements at the point of the actual contact.

2   Fatigue crack initiation under contact loading

The initiation of fatigue cracks represents one of the most important stages in the pitting process. Position and mode of fatigue crack initiation depends on the microstructure of a material, the type of the applied stress and micro- and macro-geometry of the specimen [Zhou, Cheng and Mura (1989)]. Cracks are usually initiated in the area of the largest stress. Depending on different combinations of rolling and sliding contact conditions the crack initiation periods can be very different and crack can be initiated either on or under the surface.

The majority of structural materials are polycrystalline, i.e. they consist of aggregates of grains, each of which has a particular crystal orientation. There are many possible stress concentrations due to grain boundaries, crystal voids, triple points, machining marks, inclusions and large notches, which separately, or in a variety of combinations, can upon loading immediately initiate a crack either on or under the surface [Miller (1993)]. The surface cracks can also be initiated as a consequence of a mechanical and thermal treatment of the material.

However, in high precision mechanical components with smooth surfaces and good lubrication, the largest contact stresses always appear at a certain depth under the contacting surfaces. Therefore, the fatigue cracks leading to pitting are here usually initiated under the surface along persistent slip band inside a crystal grain. The average length of the initial crack is in this case equal to the grain size D, *i.e.* $a_0 = D$. If shearing, principal and von Mises equivalent stresses are computed, the positions of their maxima under the contacting surfaces almost coincide. This leads to the conclusion that any of these can be used to determine the position of the crack, from which the
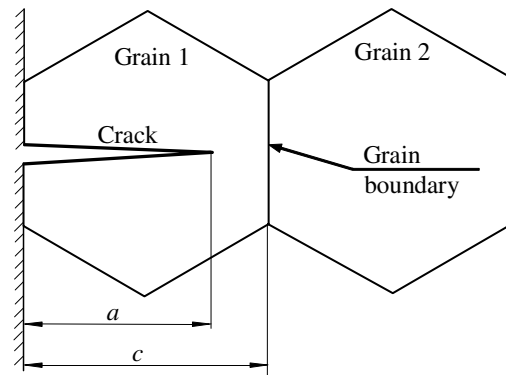

Figure 1 : Formulation of the short crack growth

likelihood of the fatigue crack initiation is the highest.

3   Fatigue crack propagation under contact loading

The presented model is based on the theory of the short fatigue crack growth, where the crack growth rate da/dN is proportional to the crack tip plastic displacement δ_{pl} [Navarro and Rios (1988), Sun, Rios and Miller(1991)]

$$\frac{da}{dN} = C_o \cdot (\delta_{pl})^{m_o} , \tag{1}$$

where C_o and m_o are material constants, that can be determined experimentally. It is known that short cracks do not behave in accordance with LEFM. However, in view of the numerical simulation, it is beneficial to express the plastic displacement δ_{pl} in terms of the stress intensity factor K. This relationship has been provided in the following form [Navarro and Rios (1988)]:

$$\delta_{pl} = \frac{2 \cdot (1-\nu)}{G\sqrt{\pi}} \cdot \frac{\sqrt{1-n^2}}{n} \cdot K(a) \cdot \sqrt{a} , \tag{2}$$

where G is the shear modulus and ν is the Poisson's ratio. Parameter $n = a/c$ describes relative position of crack tip to the grain boundary, where a is the crack length and c is the crack length together with the plastic zone ahead of the crack, which extends to the grain boundary, see Fig. 1. During short crack growth the material microstructure has a significant influence on crack propagation, which is characterised by successive blocking of the persistent slip bands by grain boundaries and subsequent initiation of the slip band in the following grain. This implies the discontinuous character of the crack growth process, since in each grain the crack growth rate decreases as the crack approaches the grain boundary. However, at the moment

when the stress concentration ahead of the crack is able to initiate the plastic slip in the next grain, the crack growth rate increases and the process is repeated. The complete procedure is fully described in [Glodež, Ren and Flašker (1998)].

The purpose of this study was only the determination of the functional relationship between the stress intensity factor K and crack length a, which is necessary to solve Eq. (2) and consequently Eq. (1). For the purpose of the fatigue crack growth simulation the virtual crack extension (VCE) method in the framework of the finite element method (FEM) has been applied. The VCE-method, as proposed by [Hellen (1975)], is based on the criterion of released strain energy dV per crack extension da ($G = dV/da$), which serves as a basis for determining the combined stress intensity factor K around the crack tip for the plane strain conditions. The complete procedure is described in [Glodež, Flašker and Ren (1997)].

Assuming the validity of the maximum energy release criterion, the crack will propagate in the direction corresponding to the maximum value of G, *i.e.* in the direction of the maximum stress intensity factor K. The computational procedure is based on incremental crack extensions, where the size of the crack increment is prescribed in advance. For each crack extension increment, the stress intensity factor is determined in several different possible crack propagation directions (Fig. 2) and the crack is actually extended in the direction of the maximum stress intensity factor, which requires local remeshing around the new crack tip. When the fatigue crack reaches the critical crack length a_c, *i.e.* when K exceeds the critical stress intensity factor K_c, the crack growth becomes uncontrolled. K_c has to be determined experimentally for a given material and boundary conditions, and is generally not available. However, K_c can be estimated from the critical stress intensity factor for Mode I fracture K_{Ic}, which is available for different materials from the literature [Glodež (1996), Ewalds and Wanhill (1991)]. For high quality steels for mechanical elements like gears and bearings K_{Ic} is in the range $K_{Ic} = 2000$ to 3000 Nmm$^{-3/2}$. In fatigue crack propagation, cracks usually grow such that the Mode I behaviour dominates, and at critical values the use of K_{Ic} is valid.

4 Parameters influencing the fatigue crack growth

The used computational model attempts to account for different parameters influencing the pitting pro-
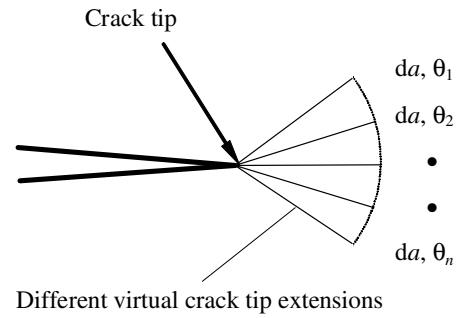

Figure 2 : Virtual crack extensions of the crack tip

cess: Hertzian contact pressure, friction between contacting surfaces, EHD-lubrication, moving contact of gear flanks, fluid trapped in the crack (only for surface initiated cracks).

4.1 Normal and tangential contact loading

The normal contact loading $p(x)$ and the maximum contact pressure p_0 in the middle of the contact zone can be determined using the Hertzian contact theory as described in [Johnson (1985)]. The distribution of tangential contact loading $q(x)$ due to the relative sliding of the contacting surfaces is here determined using the Coulomb friction law

$$q(x) = \mu \cdot p(x), \tag{3}$$

where μ is the coefficient of friction between the contacting surfaces.

4.2 Influence of the EHD-lubrication

In the proposed computational model, the normal contact loading distribution $p(x)$ considers also the influence of the elasto-hydro-dynamic (EHD) lubrication conditions, which appear in the contact of mechanical elements like gears and bearings. Under such conditions, the contacting surfaces are separated with a thin lubricant film. However, the presence of the viscous lubricant in the contact area of the sliding surfaces affects the contact pressure distribution in a way that is illustrated in Fig. 3. This pressure distribution can be determined experimentally or with the use of appropriate computational models [Williams (1994)]. A significant pressure spike develops in the outlet contact region and strongly depends on the lubricant's pressure-viscosity characteristic.
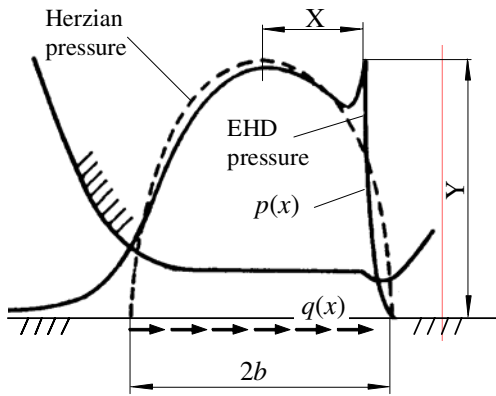
Figure 3 : Contact loading distribution

In all computations the dimensionless pressure spike amplitude Y and dimensionless pressure spike location X (see Fig. 3) have been determined using the following empirical equations [Hamrock, Lee and Houpert (1987)]:

$$Y = 0.267\, W^{-0.375}\, U^{0.174}\, G^{0.219} \tag{4}$$

$$X = 1 - 2.469\, W^{-0.941}\, U^{0.206}\, G^{-0.848} \tag{5}$$

Here, W, U and G are dimensionless load, speed and material parameters that are determined as follows:

$$W = \frac{F_N}{E^* R^*} \tag{6}$$

$$U = \frac{\eta_o u}{E^* R^*} \quad ; \quad \eta_o = \rho \cdot \nu \tag{7}$$

$$G = \alpha E^* \tag{8}$$

where F_N is the normal force per unit contact width, R^* and E^* are the equivalent radius and the equivalent Young's modulus of the equivalent contact model [Johnson (1985)], η_o is the dynamic viscosity at the atmospheric pressure, $u = (u_1 + u_2)/2$ is the mean surface velocity and u_1 and u_2 are respective surface velocities, ρ is the lubricant density, ν is the kinematic viscosity and α is the pressure-viscosity coefficient.

4.3 Influence of moving contact and lubricant trapped in the crack

For a more realistic simulation of the fatigue crack growth, it is necessary to consider the moving contact of the gear flanks. The moving contact can be simulated with different loading configurations, as shown in Fig. 4. In all configurations, the normal $p(x)$ and tangential

$q(x)$ contact loading distributions are of the same magnitude, however they are acting at different positions with respect to the crack.

The simulation of the surface initiated fatigue crack propagation also has to consider the influence of the lubricant pressure acting on the crack faces. The lubricant pressure is not constant but depends on the contact loading position, *i.e.* the contact pressure distribution position with respect to the crack, see Fig. 4.

5   Practical applications

The presented model has been used for simulation of pitting phenomenon on the equivalent model of two cylinders made from through-hardened steel 42CrMo4 (according to DIN standard), which is often used for high loaded mechanical elements. The average grain size of this material is $D \approx 0.05$ mm. The equivalent cylinders with radii $R_1 = R_2 = 10$ mm, Young's modulus $E_1 = E_2 = 2.06 \cdot 10^5$ N/mm^2 and Poisson ratio $\nu_x = \nu_y = 0.3$ was loaded with the maximum contact pressure $p_0 = 1550$ N/mm^2. In that case the half-length of the contact area is equal to $b = 0.274$ mm.

The normal contact loading distribution $p(x)$ has been determined using the Hertzian contact theory as described in section 4.1. The tangential loading $q(x)$ has been determined using Eq. (3), where the coefficient of friction $\mu = 0.04$ has been used. It is the average value for well-lubricated contact of mechanical elements like gears and bearings [Winter and Knauer (1988)].

The influence of EHD-lubrication on the normal loading distribution $p(x)$ has been estimated using eqs. (4) to (8) for the lubricant oil ISO-VG-220, with the kinematic viscosity $\nu_{40} = 220$ mm^2/s, density $\rho_{15} = 0.9$ kg/dm^3 and pressure-viscosity coefficient $\alpha = 0.18 \cdot 10^{-7}$ m^2/N. The mean surface velocity of the contacting surfaces has been taken as a constant value $u = 5$ m/s, which is common value for gears [Glodež (1996)]. Using these parameters, the dimensionless pressure spike amplitude Y and the dimensionless pressure spike location X (see Fig. 3) are: $X = 0.9462$ and $Y = 0.8146$.

The finite element mesh showed in Figure 5, and the boundary conditions as described above, have then been used in the subsequent numerical analysis. The simulation of pitting phenomenon has been studied for surface and subsurface initiated fatigue crack growth.
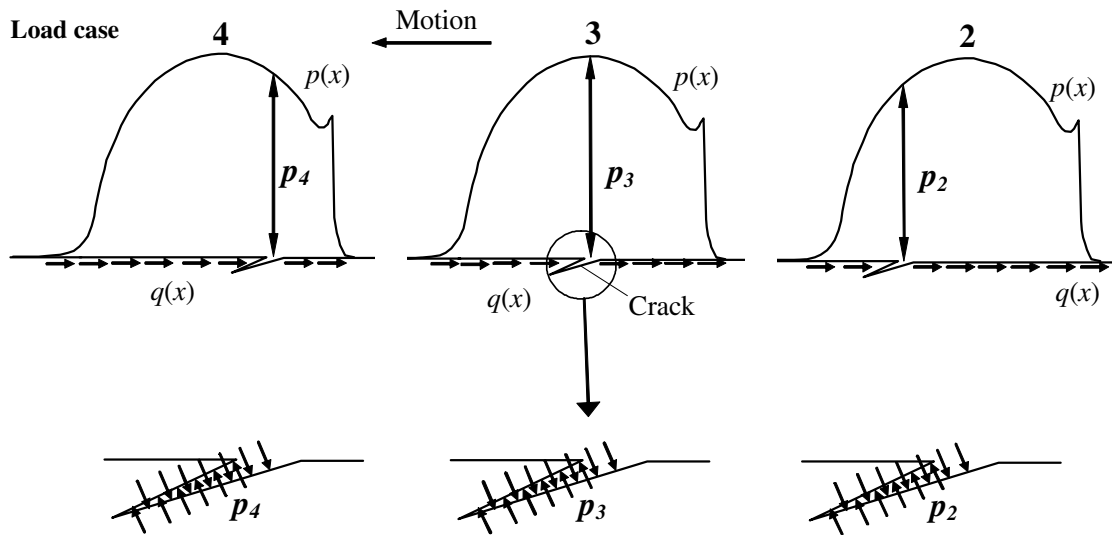
Figure 4 : Different load cases and lubricant pressure acting on the crack faces
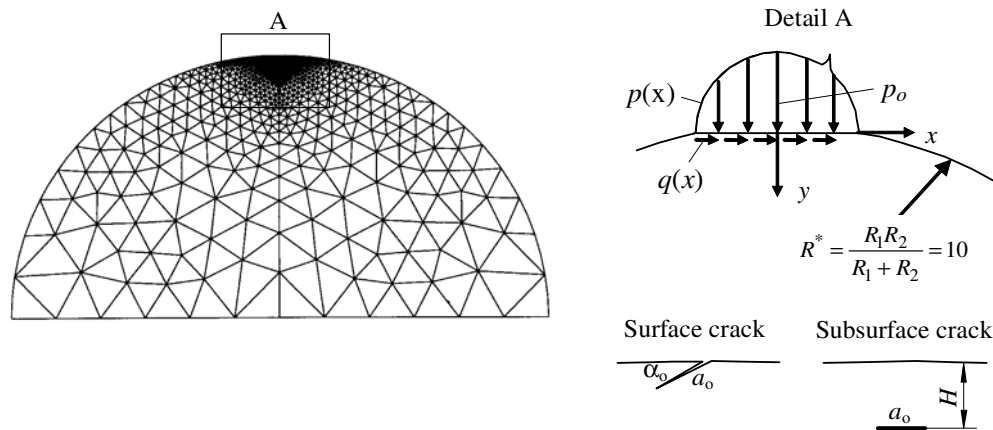

Figure 5 : FE model and configuration of the initial crack

5.1 Surface initiated fatigue crack growth

For the surface initiated crack it was assumed that the initial length of the crack is equal to a_o=15 μm, with the initial inclination angle towards the contact surface equal to α_o=22°. This configuration follows the metallographic investigations of initial cracks appearing on gears, which are a consequence of mechanical or heat treatment of the material as well as the running in process [Raymond (1992)]. In numerical computations, the crack increment was of size $\Delta a = 1.5$ μm. The stress intensity factor K was estimated in each crack increment for 30 different virtual crack tip extensions (see Fig. 2). Five different load cases have been considered in each computation for the purpose of simulating the effect of the moving contact (only three are presented on Fig. 4). For each crack in-

crement, the crack was actually extended in the direction of the recorded K_{max} from all calculated load cases. Fig. 6 shows the relationship between stress intensity factor K and the crack length a, and also the shape and magnitude of the surface pit. It can be seen that the computed stress intensity factor K is quite small at the beginning, but later increases as the crack propagates towards the contact surface. Numerical simulations have shown that at the moment when the crack reaches the vicinity of the contact surface, the stress intensity factor is extremely high. At that moment it can be expected that the material surface layer breaks away and the pit occurs on the surface. Because of the very small dimensions of surface pits, they can be termed as micro pitting.

Micro pitting as shown in Fig. 6 is not the final and most critical surface failure. Further operation of the mechani-
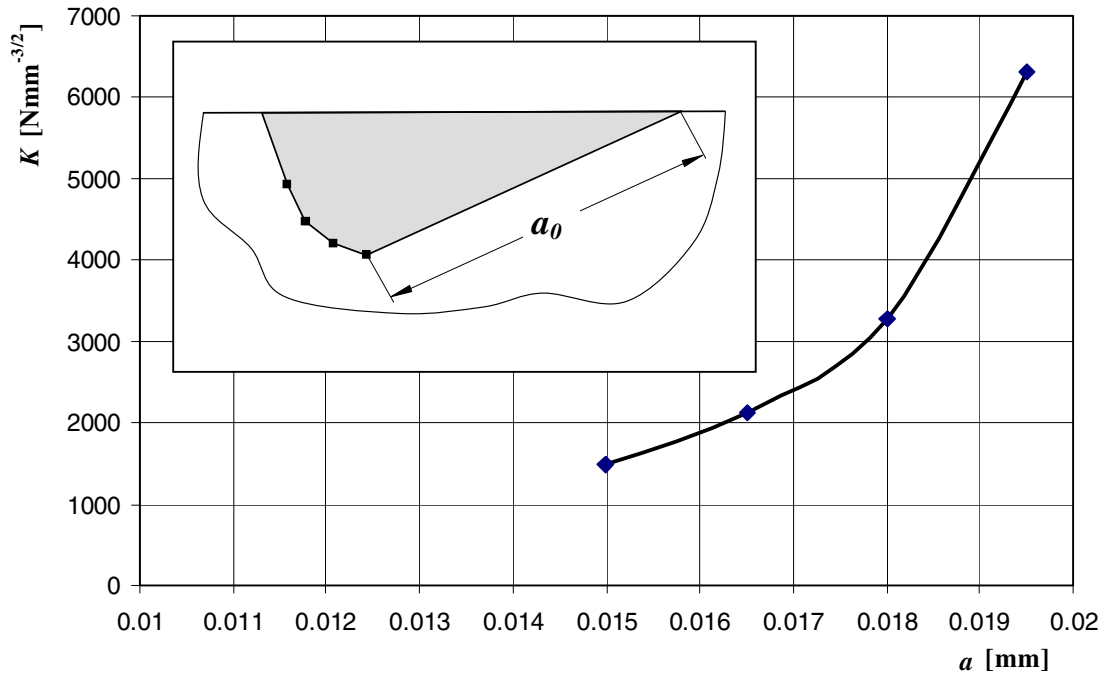
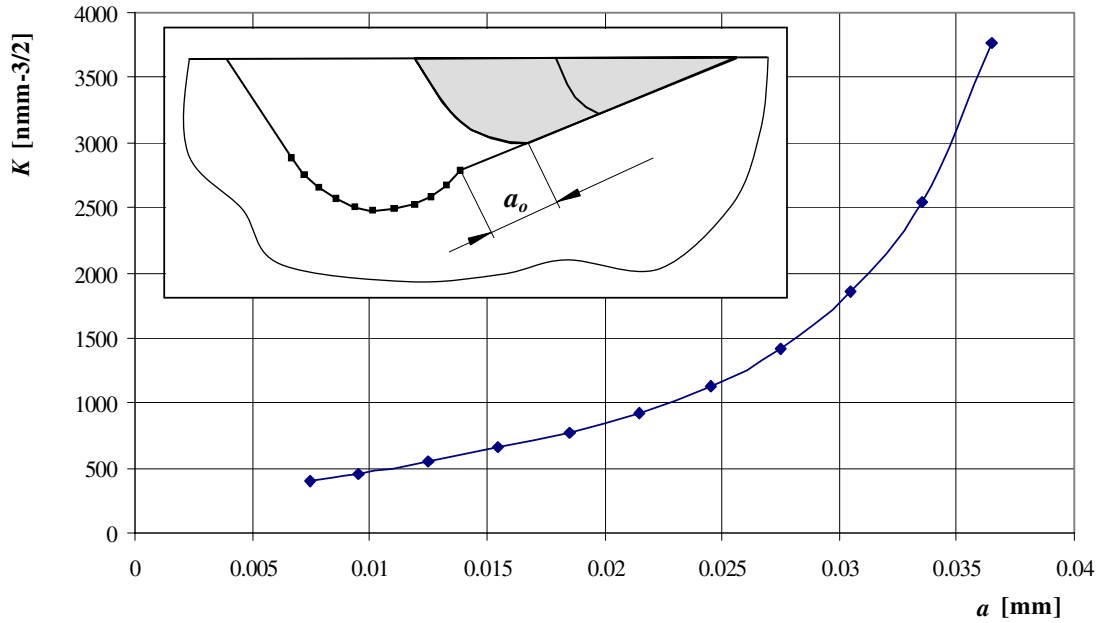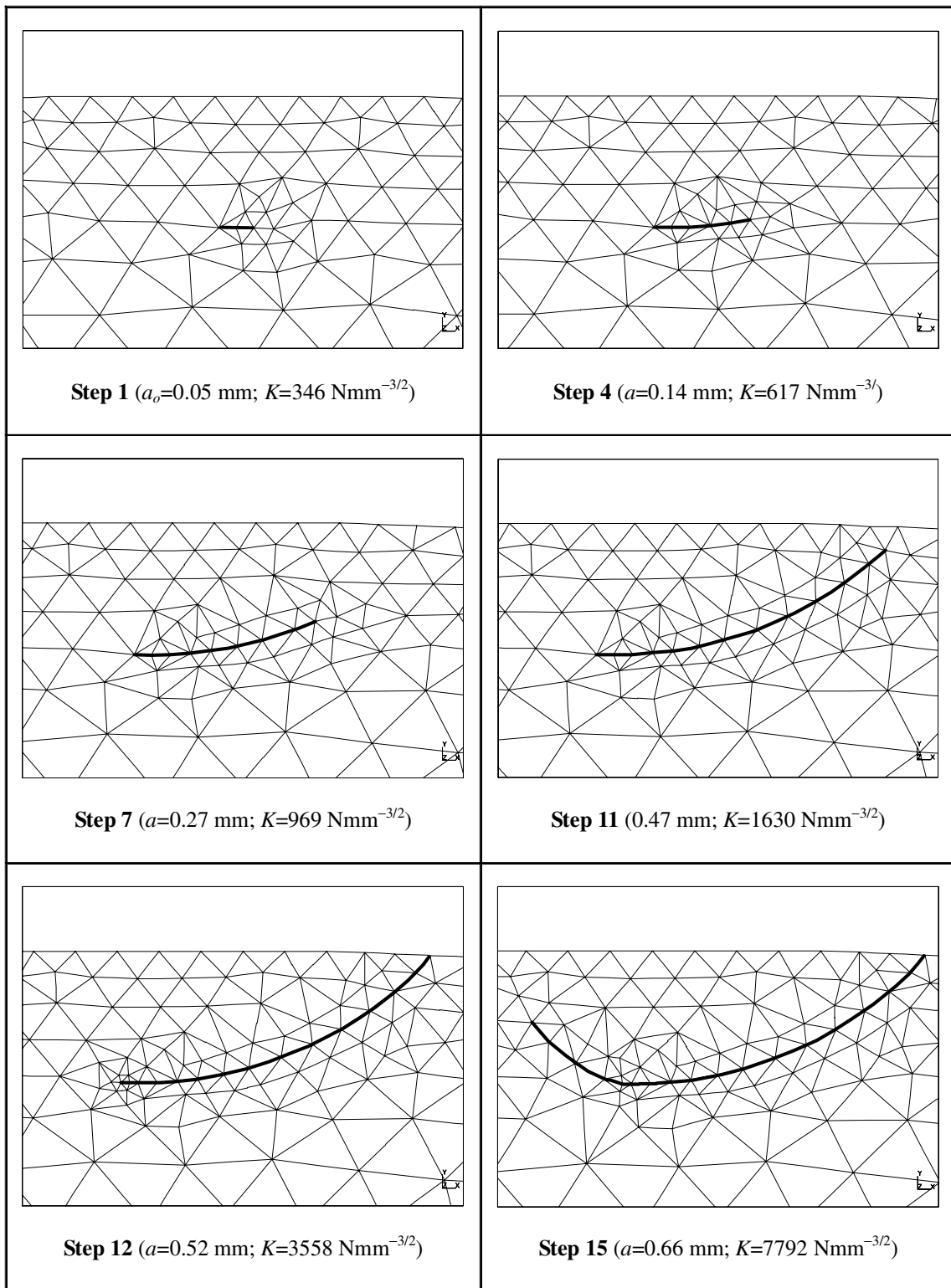**Figure 6** : Stress intensity factor for surface initiated crack growth

Figure 7 : Stress intensity factor for crack propagation from the existing micro pit

cal components results in the formation of larger pits, and consequently progressive pitting. In this respect, similar numerical simulations as described above have been continued for two more steps, as shown in Fig. 7. In these calculations, it was assumed that the initial crack of length a_o=7.5 μm started from the bottom of the existing surface pit. For all numerical computations, the crack increment was of size $\Delta a = 2$ μm.

Step 1 (a_o=0.05 mm; K=346 Nmm$^{-3/2}$)

Step 4 (a=0.14 mm; K=617 Nmm$^{-3/}$)

Step 7 (a=0.27 mm; K=969 Nmm$^{-3/2}$)

Step 11 (0.47 mm; K=1630 Nmm$^{-3/2}$)

Step 12 (a=0.52 mm; K=3558 Nmm$^{-3/2}$)

Step 15 (a=0.66 mm; K=7792 Nmm$^{-3/2}$)

Figure 8 : Numerical simulation of the subsurface initiated fatigue crack growth

Pit shape for surface initiated crack

Pit shape for subsurface initiated crack
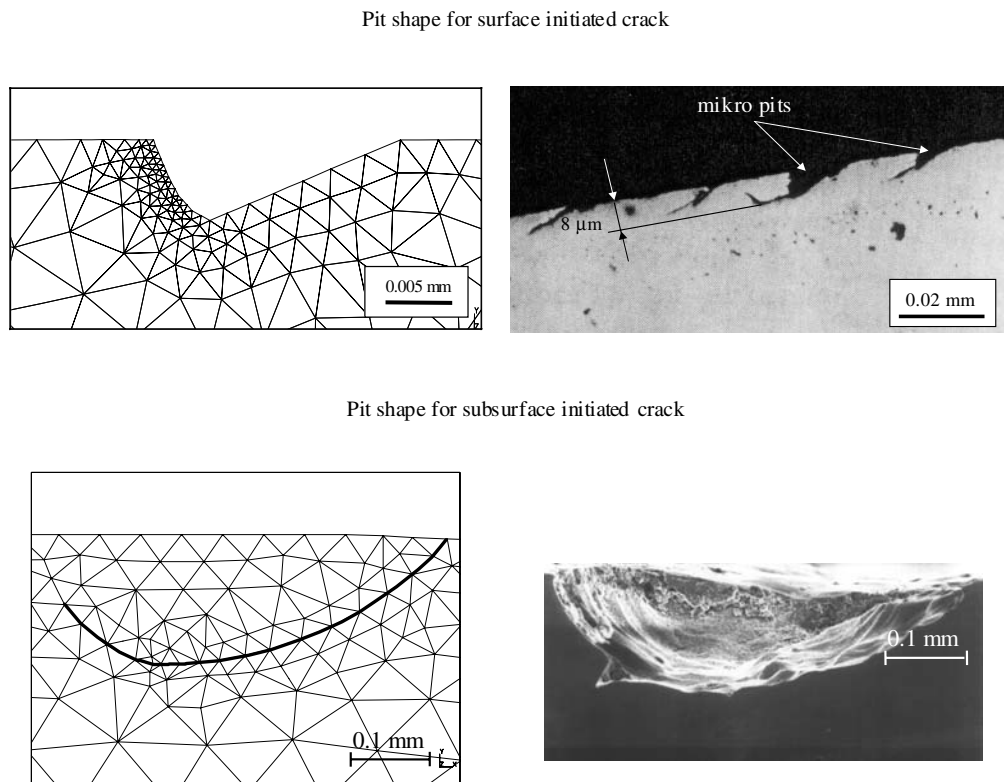
Figure 9 : Comparison of numerically and experimentally obtained pit shapes

5.2 *Subsurface initiated fatigue crack growth*

Following the assumptions made in Section 2, the initial crack of length $a_o = D$=0.05 mm has been positioned at the point of the maximum equivalent stress, *i.e.* at depth H=0.192 mm under the contact surface, see Fig. 5. The crack increment has been set to da=0.03 mm. Numerical analyses have shown, that upon initial crack appearance, the stress intensity factor is much higher in the crack tip, which coincides with the direction of applied friction force, which is the right crack tip for the treated case, see Fig. 8. Therefore it is safe to assume that the crack will start to propagate in this direction. At the present the computational model can follow only the growth of a single crack tip and it was necessary to assume that the crack will grow only in one direction until it reaches the surface. Only then will the crack start to propagate from the other initial crack tip.

Fig. 8 illustrates the simulation of the fatigue crack growth for data set p_o=1550 MPa and R^*=10 mm. The local finite element discretisation around the crack is shown for some typical crack extensions. Step 1 shows the initial crack, while steps 4-11 illustrate the right crack tip propagation towards the surface. Step 15 depicts the

crack propagating from the left crack tip after the initial crack has broken through to the surface. During the finite element analysis the computed stress intensity factor is very small at the beginning but later increases as the crack propagates towards the contact surface. Numerical simulations have shown that at the moment when the crack breaks through to the contact surface, the stress intensity factor in the other initial crack tip exceeds the critical stress intensity factor K_{Ic} for high quality steels, see chapter 3. This implies that when the subsurface crack first reaches the contact surface the corresponding crack length can be taken as the critical crack length a_c.

6   Concluding remark

Rolling contact fatigue cracks are one of the most common forms of failure in gears, railway track and bearings. There are two different types of contact fatigue crack: the crack may either be initiated at the contacting surfaces, and thereafter propagate at a shallow angle to the surface; alternatively, the cracks may be initiated at large non-metallic inclusions below the surface, in the region of the maximum cyclic shear stress. In both cases, the cracks continue to propagate in this region before either

detaching lumps of surface material, or changing direction to cause catastrophic failure.

A computational model based on fracture mechanics has been developed to demonstrate the concept of predictive damage modelling for contacting mechanical components like gears, bearings, wheels, etc. A simple equivalent model with applied Hertzian boundary conditions has been used for simulation of surface and subsurface initiated fatigue crack growth under conditions of rolling and sliding contact. The model also considers the EHD-lubrication conditions, the moving contact of mechanical elements and for the surface initiated crack the fluid trapped in the crack.

Simulation of the surface and subsurface fatigue crack propagation from the initial crack up to the formation of the surface pit and dependence of the stress intensity factor on the crack length are determined by the finite element method, where the required functional relationship between the stress intensity factor at the crack tip and the crack length is determined using the virtual crack extension method. On the basis of computed relationships between the stress intensity factor and crack length and with consideration of some particular material parameters, the service life of mechanical component can then be determined as described in [Glodež, Flašker and Ren (1997)].

Comparison of numerically predicted and experimentally recorded [Elstorpf (1993)] pit shapes show that they are in a very good agreement, see Fig. 9. However, the model should be improved with additional theoretical, numerical and foremost experimental research, since it relies mostly on experimentally determined material parameters.

References

Lester, E. A. (1986): *Systematic analysis of gear failures*, American Society for Metals.

Keer, L. M.; and Bryant, M. D. (1983): A pitting model for rolling contact fatigue. *ASME Jour. Lubr. Tech.*, vol. 105, pp. 198-205.

Miller, G. R.; Keer, L. M.;Cheng, H. S. (1985): On the mechanics of fatigue crack growth due to contact loading. *Proc. Roy. Soc. London*, vol. A397, pp. 197-209.

Murakami, Y.; Kaneta, M. (1987): Fracture mechanics approach to tribology problems. *The twentieth symposium of fracture mechanics, Lehigh University-Bethlehem*, pp. 668-687.

Bower, A. F. (1988): The influence of crack face friction and trapped fluid on surface initiated rolling contact fatigue cracks. *ASME Jour. Trib.*, vol. 110, pp. 704-711.

Hanson, M. T.; Keer, L. M. (1992): An analytical life prediction model for the crack propagation occurring in contact fatigue failure. *STLE Trib. Trans.*, vol. 35, pp. 451-461.

Murakami, Y.; Sakae, C.; Ichimaru, K. (1994): Three-dimensional fracture mechanics analysis of pit formation mechanism under lubricated rolling-sliding contact loading. *STLE Trib. Trans.*, vol. 37, pp. 445-454.

Glodež, S.; Ren, Z.; Fajdiga, G. (2001): Computational modelling of the surface fatigue crack growth on gear teeth flanks. *Commun. numer. methods eng.*, vol. 17, pp. 529-541.

Flašker, J.; Fajdiga, G.; Glode,ž S.; Hellen T.K. (2001): Numerical simulation of surface pitting due to contact loading. *Int. j. fatigue*, vol. 23, pp. 599-605.

Leng, X.; Chen, Q; Shao, E. (1998): Initiation and propagation of case crushing cracks in rolling contact fatigue. *Wear*, vol. 122, pp. 33-43.

Cheng, W.; Cheng, H. S.; Mura, T.; Keer, L. M. (1994): Micromechanics modelling of crack initiation under contact fatigue. *ASME J. Tribology*, vol. 116, pp. 2-8.

Glodež, S.; Flašker, J.; Ren, Z. (1997): A new model for the numerical determination of pitting resistance of gear teeth flanks. *Fatigue Fract Eng Mat&Struct*, vol. 20, pp. 71-82.

Glodež, S.; Ren, Z. and Flašker, J. (1998): Simulation of surface pitting due to contact loading. *Int. j. numer. methods eng.*, vol. 43, pp. 33-50.

Johnson, K. L. (1985): *Contact mechanics*. Cambridge University press.

Zhou, R. S.; Cheng, H. S.; Mura, T. (1989): Micropitting in Rolling and sliding contact under mixed lubrication. *ASME J. Tribology*, vol. 111, pp. 605-613.

Miller, K.J. (1993): Materials Science Perspective of Metal Fatigue Resistance. *Materials Science and Technology*, vol. 9, pp. 453-462.

Navarro, A.; Rios, E. R. (1988): Short and long fatigue crack growth-a unified model. *Philosophical Magazine A*, vol. 57, pp. 15-36.

Sun, Z.; Rios, E.R; Miller, K.J. (1991): Modelling small fatigue cracks interacting with grain boundaries. *Fatigue Fract Eng Mat&Struct,* vol. 14, pp. 277-291.

Glodež, S. (1996): *The fracture mechanics model of gear flanks fatigue*, PhD thesis, Faculty of Mechanical Engineering, University of Maribor (in Slovenian).

Ewalds, H.L.; Wanhill, R.J.H. (1991): *Fracture Mechanics*, Delftse U.M., Co-publication of Edward Arnold, Delft.

Hellen, T. K. (1975): On the method of virtual crack extensions. *International Journal for Numerical Methods in Engineering*, vol. 9, pp. 187-207.

Williams, J. A. (1994): *Engineering tribology*, Oxford University press.

Hamrock, B.J.; Lee, R.T; Houpert, L.G. (1987): Parametric study of performance in elastohydrodynamic lubricated line contact. *Fluid film lubrication–Osborne Reynolds centenary*, pp. 199-206.

Winter, H.; Knauer, G. (1990): Einfluss von Schmierstoff und Betriebstemperatur auf die Grübchentragfähigkeit einsatzgehärteter Zahnräder. *Antriebstechnik*, vol. 29, pp. 65-84 (in German).

Knauer, G. (1988): *Zur Grübchentragfähigkeit einsatzgehärteter Zahnräder*, Ph.D. thesis, TU Munich (in German).

Raymond, D. (1992): *Fundamentals of Gear Design*, Butterworths.

Elstorpff M.-G. (1993): Ph.D. thesis, TU Munich (in German).