



An elasto-plastic-damage model for initiation and propagation of spalling in rolling bearings

Fei Shen^a, Kun Zhou^{a,b,*}

^a School of Mechanical and Aerospace Engineering, Nanyang Technological University, 50 Nanyang Avenue, Singapore 639798, Singapore

^b Environment Process Modelling Centre, Nanyang Environment & Water Research Institute, Nanyang Technological University, 1 Cleantech Loop, Singapore 637141, Singapore

ARTICLE INFO

Keywords:

Spalling
Initiation and propagation
Rolling contact fatigue
Continuum damage mechanics

ABSTRACT

This study develops an elasto-plastic-damage model to investigate the spalling initiation and propagation behavior of bearings under rolling contact fatigue loading. The intrinsic coupling relationship between the fatigue damage and the elasto-plastic behavior of materials is considered by specifying the damage-related Helmholtz free energy and dissipative potential function in the thermodynamic framework. The damage evolution law is derived based on the plastic strain and the non-proportional cyclic stress. The developed model is implemented into finite element analysis for rolling contact of cylindrical bearings to evaluate the evolutions of the stress, strain and fatigue damage. The initiation and propagation processes of spalling in bearings are simulated. The predicted results of the spalling lifetime are compared with the experimental data from the literature. The damage-induced plastic deformation causes the shear stress-strain curve to possess hysteresis loops. It is observed that the fatigue damage competes with the plastic deformation to influence the shear stress that plays a major driving role in spalling. Moreover, the effects of the contact pressure and the coefficient of friction on the spalling process are studied.

1. Introduction

Spalling is a dominant failure mode of rolling contact fatigue (RCF) in bearings under severe loading [1–3]. It commonly starts with cracks initiated in the subsurface of inner and outer rings due to the formation and coalescence of micro-voids or micro-cracks. The cracks then propagate towards the contact surface, causing the removal of bulk materials. The spalling process is difficult to detect and measure through experimental methods. Therefore, it is of practical and academic significance to develop numerical approaches for integrity assessment of bearings under RCF loading.

RCF in bearings experiences several unique features due to the rolling between contacting components. The contact between rings and rollers causes severe stress concentration. High contact stress is localized in a small zone and plastic strain may occur in the subsurface of bearings. The stress at a material point in the ring varies during the rolling process. Furthermore, this cyclic stress exhibits non-proportional behavior, i.e., the variations of the stress components in one cycle are not consistent. Lubrication affects the RCF behavior of rolling bearings. Fine lubrication condition reduces the contact stress and the wear of rollers and rings, which postpones or even prevents the formation of spalling in bearings. The features mentioned above cause complicated

stress status in bearings so that the spalling in RCF is difficult to investigate.

The accurate prediction of the contact and subsurface stresses is the prerequisite for the further modeling of spalling. The contact problem of rollers and rings was commonly simplified as a two-dimensional contact problem [4–10]. The contact pressure was assumed to have a Hertzian distribution for calculation of the subsurface stresses in the rings. This simplification might be suitable for the cylinder-flat contact under an elastic deformation assumption. However, the contact pressure distribution in bearings is more complicated than the Hertzian distribution. More importantly, it changes gradually during the rolling process. Therefore, a more accurate numerical method is needed to evaluate the contact and subsurface stresses in bearings. Arakere et al. [11] used the finite element method to conduct stress modeling for ball bearings. Li et al. [12] used the same method to calculate the stress and strain distributions in cylindrical bearings.

Based on the obtained stress and strain results, the RCF lifetime of bearings can be predicted by probabilistic and deterministic models [10]. Lundberg and Palmgren [13] firstly proposed an empirical probabilistic model for RCF relating the bearing lifetime for 10% probability of failure to the maximum orthogonal shear stress. Extensive models based on this pioneering work have been developed to consider the

* Corresponding author.

E-mail address: kzhou@ntu.edu.sg (K. Zhou).

effects of plastic behavior, defect type and geometry, and coefficient of friction on the contact stress field and fatigue lifetime [10]. Based on numerous experimental results, the probabilistic models were proposed by directly correlating the survival probability of bearings with several controlling factors. However, the intrinsic mechanisms of spalling such as the initiation and propagation stages were not considered in the models. Several deterministic models have thus been proposed to predict the lifetime of spalling initiation and propagation [10]. Liu et al. [14] developed a critical plane method for RCF to predict the crack initiation lifetime of railroad wheels. Jiang and Sehitoglu [15] implemented a multiaxial fatigue damage model in an elasto-plastic finite element analysis of RCF to evaluate the crack initiation behavior. Zhou et al. [16,17] developed an analytical model to predict the total lifetime of bearings, in which the crack propagation lifetime was predicted using the Paris law. However, the deterministic models could not provide a detailed description of the fatigue damage process in spalling. The non-proportional behavior of cyclic stress was not considered in these models. Furthermore, the coupling relationship between the fatigue damage and stress-strain response was not included.

The continuum damage mechanics (CDM)-based approach is a powerful method to consider the fatigue damage process in fatigue problems. A damage variable (scalar or tensor) is introduced to represent the average material degradation at the macroscopic scale induced by microvoids and micro-cracks [18–24]. The coupling relationship between the fatigue damage and stress-strain response is considered in the thermodynamic framework. Fatigue damage weakens material properties such as the Young's modulus, yield limit and hardening modulus. Conversely, the evolution of fatigue damage is controlled by the history of stress and strain fields. The damage gradually increases with an increased number of cyclic loading. As the damage at a material point reaches a critical value, the material loses its loading capacity, indicating the initiation of a crack there. Afterwards, the crack propagates from the initiation point. Therefore, the CDM-based approach has been used to investigate the damage process in various fatigue problems [25–29]. Warhadpande et al. [9] evaluated the effect of plasticity on the RCF spalling through a Voronoi tessellation-based finite element model. The damage variable was included in a bilinear elasto-plastic constitutive law. The crack initiation and propagation were modeled by the separation of elements along their edges. Walvekar et al. [4] used a similar method to predict the RCF behavior of case-carburized steels in which the residual stress and inclusions induced by the carburizing process were considered. More recently, Li et al. [12] proposed an elastic damage model in conjunction with finite element modeling of cylindrical bearing rolling to evaluate the subsurface-initiated spalling.

It should be highlighted that the rolling process of bearings was not modeled in the CDM-based approaches [4,9]. The constant Hertzian contact pressure was used to calculate the subsurface stresses and therefore the variation of the contact pressure in the rolling process was not considered. Furthermore, these approaches might encounter a severe numerical convergence problem because the elements near the crack tips were not refined after the crack initiation caused by the separation of element edges. For the CDM-based approach in [12], elastic deformation was assumed and thus plastic deformation and relevant plastic damage could not be accounted for.

The present study proposes an elasto-plastic-damage model to investigate the initiation and propagation of spalling in bearings under RCF loading. A damage-coupled elasto-plastic constitutive model is used to calculate the stress and strain fields. The plastic strain and non-proportional cyclic stress are used to calculate the evolution of fatigue damage. The model considers the coupling relationship between the fatigue damage and the stress-strain response. The rolling process between the roller and ring is simulated in finite element analysis to obtain an accurate prediction of the varying surface contact pressure and subsurface stresses. The initiation and propagation processes of spalling are evaluated using the developed model and the predicted results are compared with experimental data.

2. Elasto-plastic-damage model

2.1. Damage-coupled elasto-plastic constitutive model

The thermodynamic framework is simply revisited to develop the constitutive model considering plastic deformation and damage. The Helmholtz free energy includes three contributions from elasticity, isotropic hardening and kinematic hardening behavior [19,20]:

$$\rho\psi = \frac{1}{2}\mathbf{C}(1-D)\epsilon_e : \epsilon_e + R_\infty \left(r + \frac{1}{b} \exp(-br) \right) + \frac{c}{3} \boldsymbol{\alpha} : \boldsymbol{\alpha}, \quad (1)$$

where the first term $\mathbf{C}(1-D)\epsilon_e$: $\epsilon_e/2$ denotes the damage-related elastic strain energy density in which \mathbf{C} is the fourth-order elastic tensor, ϵ_e is the elastic strain tensor, and D is a damage variable and a scalar because isotropic damage is used in the study; the second term $R_\infty(r + \exp(-br)/b)$ denotes the isotropic hardening-related energy in which the scalar r is an internal variable and its associated variable R represents the increment of the yield limit; the last term $c\boldsymbol{\alpha}$: $\boldsymbol{\alpha}/3$ denotes the kinematic hardening-related energy in which the tensor $\boldsymbol{\alpha}$ is also an internal variable and its associated tensor \mathbf{X} is the backstress describing the movement of the center of a yield surface. The variables R_∞ , b and c are material parameters. The status functions can be obtained according to the damage-related Helmholtz free energy [19,23,30]:

$$\boldsymbol{\sigma} = \rho \frac{\partial \psi}{\partial \boldsymbol{\epsilon}_e} = \mathbf{C}(1-D)\boldsymbol{\epsilon}_e, \quad (2)$$

$$R = \rho \frac{\partial \psi}{\partial r} = R_\infty(1 - \exp(-br)), \quad (3)$$

$$\mathbf{X} = \rho \frac{\partial \psi}{\partial \boldsymbol{\alpha}} = \frac{2}{3} c \boldsymbol{\alpha}, \quad (4)$$

$$Y = -\rho \frac{\partial \psi}{\partial D}, \quad (5)$$

where Y is the work-conjugate associated variable of the damage. Eqs. (2)–(4) will be used in the deduction of the constitutive model.

The evolution laws of these internal variables including plastic strain ϵ_p are determined by a dissipative potential function that fulfils the second thermodynamic principle. Similar to the Helmholtz free energy, the dissipative potential function used in this study includes three components [20]:

$$\varphi = \left[\left(\frac{\boldsymbol{\sigma}}{1-D} - \mathbf{X} \right)_{\text{eq}} - \sigma_Y - R \right] + \frac{3\gamma}{4c} \mathbf{X} : \mathbf{X} + F_D, \quad (6)$$

where the first term $(\boldsymbol{\sigma}/(1-D) - \mathbf{X})_{\text{eq}} - \sigma_Y - R$ is the von Mises yield criterion, with the subscript “eq” indicating the calculation of von Mises stress and σ_Y being the initial yield limit; and the second term $3\gamma\mathbf{X} : \mathbf{X}/4c$ and the third term F_D are the plastic and damage dissipative potentials, respectively. The evolution laws are written as [19,23,30]

$$\dot{\epsilon}_p = \dot{\lambda} \frac{\partial \varphi}{\partial \boldsymbol{\sigma}} = \frac{3}{2} \frac{\dot{\lambda}}{1-D} \frac{\mathbf{s}/(1-D) - \mathbf{X}}{(\mathbf{s}/(1-D) - \mathbf{X})_{\text{eq}}} = \frac{\dot{\lambda}}{1-D} \mathbf{r}, \quad (7)$$

$$\dot{r} = -\dot{\lambda} \frac{\partial \varphi}{\partial R} = \dot{\lambda}, \quad (8)$$

$$\dot{\boldsymbol{\alpha}} = -\dot{\lambda} \frac{\partial \varphi}{\partial \mathbf{X}} = \dot{\lambda} \left(\frac{3}{2} \frac{\mathbf{s}/(1-D) - \mathbf{X}}{(\mathbf{s}/(1-D) - \mathbf{X})_{\text{eq}}} - \frac{3\gamma}{2c} \mathbf{X} \right), \quad (9)$$

$$\dot{D} = \dot{\lambda} \frac{\partial \varphi}{\partial Y}, \quad (10)$$

where $\dot{\lambda}$ is the plastic multiplier; \mathbf{r} represents the direction of the plastic strain; \mathbf{s} is the deviator stress tensor, i.e., $\mathbf{s} = \boldsymbol{\sigma} - \text{trace}(\boldsymbol{\sigma})\mathbf{I}/3$; γ is a material parameter. The equivalent plastic strain is defined as $\dot{p} = \sqrt{2\dot{\epsilon}_p} : \dot{\epsilon}_p/3 = \dot{\lambda}/(1-D)$. Especially, several components of the backstress rather than one component are usually used to provide a more

accurate prediction of the mechanical behavior [31]. Therefore, the rate of the backstress is obtained according to Eqs. (4), (7) and (9) [31]:

$$\mathbf{X} = \sum_{k=1}^M \mathbf{X}^{(k)}, \quad (11)$$

$$\dot{\mathbf{X}}^{(k)} = (1 - D) \left(\frac{2}{3} c_k \dot{\epsilon}_p - \gamma_k \dot{p} \mathbf{X}^{(k)} \right), \quad (12)$$

where M is the number of backstress components. The evolution law of the isotropic hardening is written as [31]

$$\dot{R} = b(1 - D)(R_\infty - R)\dot{p}. \quad (13)$$

The isotropic hardening and the kinematic hardening are used together to model the cyclic hardening behavior of materials under RCF loading [32–34]. A similar model without damage has the capacity to describe the ratcheting behavior of materials under conventional cyclic loading [31,35–37], fretting loading [38] and RCF loading [39,40]. With the presence of fatigue damage, the interaction between the damage and ratcheting can be considered in the damage-coupled elasto-plastic constitutive model.

2.2. Damage evolution law

The history of stress and strain in bearings exhibits non-proportional behavior. Therefore, conventional multiaxial damage evolution models using the maximum equivalent stress or strain within one cycle cannot directly be used for RCF analysis. The whole history of stress and strain should be considered. Furthermore, plastic deformation in bearings also causes the damage increment, which is different from the case of pure elastic deformation. The stress-related variables, such as von Mises stress, damage equivalent stress and shear stress, are commonly used to govern the damage evolution under the elastic deformation condition. For the cases where plastic deformation is significant (low-cycle fatigue), the equivalent plastic strain is used to determine the fatigue damage evolution. In the present study, these two types of damage evolution laws are used together for RCF analysis.

A stress-based evolution law is used to determine the elastic damage evolution, in which the non-proportional behavior of cyclic stress is considered [12]:

$$\frac{dD_e}{dN} = \left(\frac{\Delta\tau_8}{\tau_r(1-D)} \right)^q, \quad (14)$$

$$\Delta\tau_8 = \frac{1}{2} \max_{t_k} \max_{t_l} \left(\frac{3}{2} (s_{ij,t_k} - s_{ij,t_l}) (s_{ij,t_k} - s_{ij,t_l}) \right), \quad (15)$$

where $\Delta\tau_8$ is the amplitude of the maximum octahedral shear stress calculated using stress components at all time steps within one cycle; s_{ij,t_k} and s_{ij,t_l} are the deviatoric stress components at times t_k and t_l within one cycle, respectively; τ_r and q are material parameters. A plastic strain-based evolution law is used to calculate the plastic damage evolution [9,23]:

$$\frac{dD_p}{dN} = \left(\frac{(\sigma_{\max}^*)^2}{2ES(1-D)^2} \right)^m \Delta p, \quad (16)$$

$$\sigma^* = \sigma_{\text{eq}} R_v^{1/2}, \quad (17)$$

where σ_{\max}^* is the maximum damage equivalent stress within one cycle; Δp is the equivalent plastic strain increment; R_v is the stress triaxiality, i.e., $R_v = 2(1 + \nu)/3 + 3(1 - 2\nu)(\sigma_H/\sigma_{\text{eq}})^2$ where σ_H is the hydrostatic stress; S and m are material parameters.

The total damage increment is calculated as the summation of the elastic and plastic damage increments [23]:

$$\frac{dD}{dN} = \frac{dD_e}{dN} + \frac{dD_p}{dN}. \quad (18)$$

The total damage is reduced to the elastic damage in the case of elastic deformation. The plastic damage becomes dominant compared to the elastic damage when the plastic strain is large. Therefore, the damage evolution law can be used to investigate the damage behavior in both high-cycle and low-cycle fatigue problems.

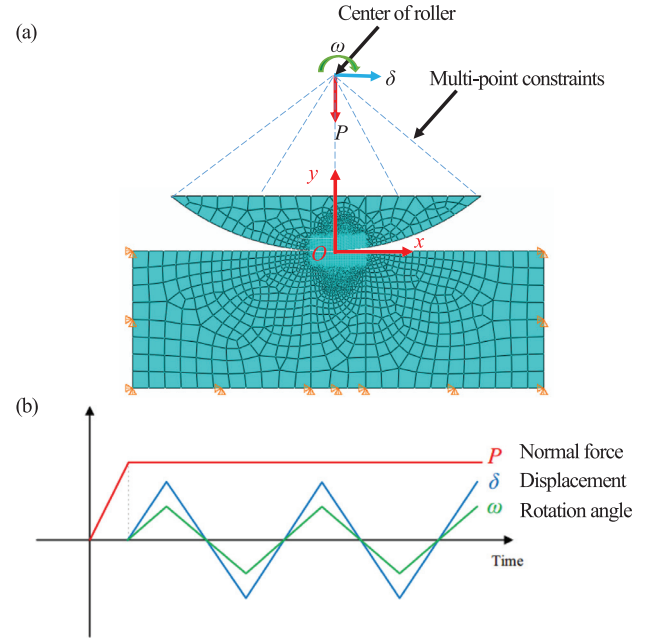


Fig. 1. (a) Finite element model of cylindrical bearings, and (b) loading history of the normal force, displacement along the x -direction and rotation angle applied on the roller center.

3. Numerical implementation

3.1. Finite element modeling of rolling contact in bearings

The finite element model of a cylindrical bearing is established in the commercial software package ABAQUS. The diameter of the roller is set as 9 mm. A portion of one roller and a section of the out ring are extracted from the entire bearing to model its rolling contact, as shown in Fig. 1(a). The curvature of the out-ring section is not considered. The plane-strain elements are used in the modeling. The contact zones in the roller and out ring are fine meshed by elements with the minimum dimension of around 0.015 mm, while the other zones are coarsely meshed. As the roller material Si_3N_4 is harder than the ring material AISI 52100, the circular surface of the roller is the main contact surface and the raceway surface of the out ring is the slave contact surface. The frictional behavior is defined by the Coulomb friction law. As it is difficult to determine the coefficient of friction in the bearing, several coefficients ranging from 0.02 to 0.1 are used to study their effect on the initiation and propagation of spalling. The origin point of a coordinate system is set at the center of the contact surface on the ring.

Fig. 1(b) shows the loading history applied on the roller. A small normal force is firstly applied on the roller center through multi-point constraints. This force then increases to cause the peak contact pressure to reach a given maximum value, namely the maximum peak contact pressure. The influence of the maximum peak contact pressure on the spalling is evaluated using the proposed model. Three maximum peak values ranging from 2000 to 2500 MPa are used in the modeling. The normal force remains unchanged after the maximum peak contact pressure is reached. The rolling process of the bearing is simplified as the repeated rolling of one roller on the out ring [12]. Therefore, the cyclic rotation and translational displacement are then applied on the roller center. The rotational angle and the displacement are 0.05 radians and 0.45 mm, respectively. The bottom surface and two side surfaces of the out ring are fixed in the rolling process.

The roller material Si_3N_4 is treated as an elastic material. The Young's modulus and Poisson's ratio are 300 GPa and 0.3, respectively. Since spalling mainly occurs in the ring, the fatigue damage of the

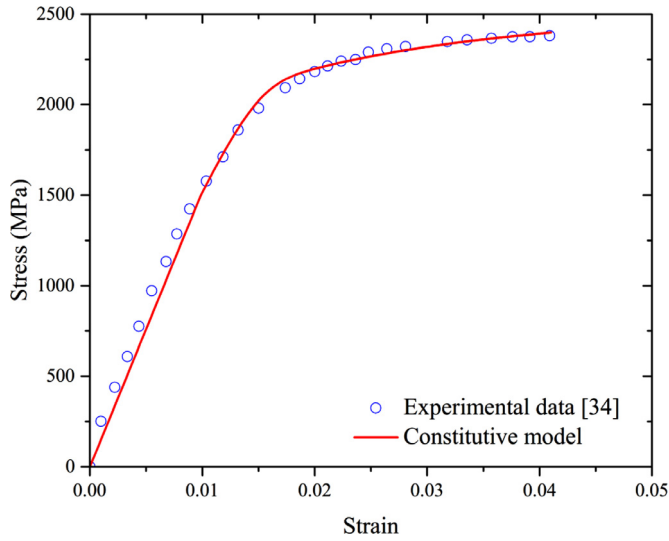


Fig. 2. Experimental and modeling data of the compressive stress-strain response of the bearing material AISI 52100.

roller is not considered. The ring material AISI 52100 exhibits damage-coupled elasto-plastic deformation behavior. The material density is 7800 kg/m^3 . The distributions of stress, strain and damage in the out ring are calculated by implementing the developed model into ABAQUS through the explicit and implicit subroutines. However, the results obtained by the explicit implementation are sensitive to the simulation parameters including the time step length. The results from the explicit implementation should be validated in comparison with those from the implicit implementation. It is noted that the implicit solver suffers a severe convergence problem in modeling the crack propagation even though it can be used to calculate the crack initiation [23,25]. Therefore, the spalling propagation is simulated through the explicit solver.

3.2. Determination of the material parameters

The material parameters are determined from compression and fatigue testing results. The rate equations in the constitutive model can be integrated into a closed form under the uniaxial condition. Damage is not considered in the determination of the material parameters because the virgin material is used for compression testing. The stress-strain relationship is written as

$$\sigma = \sigma_y + R_\infty (1 - \exp(-b\varepsilon_p)) + \sum_{k=1}^M \frac{c_k}{\gamma_k} (1 - \exp(-\gamma_k \varepsilon_p)). \quad (19)$$

Fig. 2 shows the model fitting of the experimental stress-strain response for the bearing material AISI 52100 from the literature [34]. In fact, the stress-strain curves were obtained under three different temperatures (room temperature, 200°C and 400°C). Since bearings are running with a high rotation speed and high-pressure oil lubrication, the temperature on the ring raceway can reach up to 150°C or even higher. Therefore, the stress-strain response under 200°C is used to determine the material parameters in the damage-coupled elasto-plastic model. Two backstress components ($M=2$) are used in the fitting. The parameter determination methods for the damage evolution law refer to Refs [9,12]. All the material parameters in the developed model are listed in Table 1.

Several linear kinematic hardening models were used to describe the stress-strain behavior of the material AISI 52100 [9,41,42]. The isotropic hardening law was not considered in these models. However, the kinematic hardening law alone could not provide an accurate prediction of the stress-strain response including the ratcheting behavior [31]. Therefore, a combined hardening model including both kinematic

Table 1

Material parameters of the elasto-plastic-damage model for AISI 52100.

E (MPa)	ν	σ_y (MPa)	R_∞ (MPa)	b	c_1 (MPa)	c_2 (MPa)
152,000	0.3	1500	315.87	56.09	593,303	4788.5
γ_1	γ_2	τ_r (MPa)	q	S (MPa)	m	
1000	55.9	4991	10.1	100	4.12	

and isotropic hardening behaviors has been used in some other studies [43–45]. The present study also uses the combined hardening model.

3.3. Damage accumulation and stress update algorithm

For the explicit implementation, the whole algorithm includes two parts: the calculation of stress and strain and the calculation of fatigue damage accumulation. The rate equations in the model are solved by an incremental method. The variables σ_n , ε_n , ε_n^p , $\mathbf{X}_n^{(k)}$, R_n , p_n and D_n at the time step n are updated at the next time step. The stress and strain are obtained at all the time steps within each loading cycle. However, the fatigue damage increases over cycles, which is updated at the end of each loading cycle. Within each cycle, the damage is treated as a constant. As the lifetime of bearings is on the order of millions of cycles, it is impossible to calculate the damage accumulation after each loading cycle due to tremendous computational cost. Therefore, the results of the stress, strain and damage are assumed to remain unchanged within ΔN cycles. The damage increment after ΔN cycles is calculated by simply multiplying the damage increment in one cycle by ΔN [9,20,30,46,47]:

$$D_{n+1} = D_n + \frac{dD}{dN} \Delta N, \quad (20)$$

where dD/dN is determined by the damage evolution law according to the results of stress and strain at the time step n . If the time step $n+1$ is not at the end of a loading cycle, the damage is not updated, i.e., $D_{n+1} = D_n$.

After the damage accumulation, the stress and strain can be obtained by using the damage-coupled constitutive model. The increment of the total strain $\Delta\varepsilon$ is known at the beginning of the time step $n+1$. The incremental equations of the strain, equivalent plastic strain and stress are written as [48]

$$\varepsilon_{n+1} = \varepsilon_n + \Delta\varepsilon, \quad (21)$$

$$\varepsilon_{n+1}^p = \varepsilon_n^p + \Delta\varepsilon^p = \varepsilon_n^p + \mathbf{r}_{n+1} \Delta p, \quad (22)$$

$$p_{n+1} = p_n + \Delta p, \quad (23)$$

$$\sigma_{n+1} = \sigma_n + C(1 - D_{n+1})(\Delta\varepsilon - \Delta\varepsilon^p), \quad (24)$$

where $\Delta\varepsilon^p$ is the increment of plastic strain. The increment of elastic strain is calculated as $\Delta\varepsilon^e = \Delta\varepsilon - \Delta\varepsilon^p$. The backstress components can be written as

$$\mathbf{X}_{n+1}^{(k)} = \mathbf{X}_n^{(k)} + (1 - D_{n+1}) \left(\frac{2}{3} c_k \Delta\varepsilon^p - \gamma_k \mathbf{X}_{n+1}^{(k)} \Delta p \right). \quad (25)$$

Thus, the total backstress is given as

$$\mathbf{X}_{n+1} = \sum_{k=1}^M \frac{1}{1 + (1 - D_{n+1}) \gamma_k \Delta p} \left(\mathbf{X}_n^{(k)} + \frac{2}{3} (1 - D_{n+1}) c_k \Delta\varepsilon^p \right). \quad (26)$$

The increment of the yield limit is written as

$$R_{n+1} = R_\infty (1 - \exp(-(1 - D_{n+1}) b p_{n+1})). \quad (27)$$

These variables at the time step $n+1$ should fulfill the yield function. Substituting Eqs. (21)–(27) into the von Mises yield function

leads to

$$\begin{aligned} \left(\frac{\mathbf{s}_{n+1}}{1-D_{n+1}} - \mathbf{X}_{n+1} \right)_{\text{eq}} &= \left(\left(\frac{\sigma_n}{1-D_{n+1}} + C\Delta\varepsilon \right)_{\text{dev}} - \sum_{k=1}^M \frac{1}{1+(1-D_{n+1})\gamma_k \Delta p} \mathbf{X}_n^{(k)} \right)_{\text{eq}} \\ &\quad - \left(3G + \sum_{k=1}^M \frac{1}{1+(1-D_{n+1})\gamma_k \Delta p} (1-D_{n+1})c_k \right) \Delta p \\ &= \sigma_Y + R_\infty (1 - \exp(-(1-D_{n+1})bp_{n+1})). \end{aligned} \quad (28)$$

The above equation is a nonlinear function of the variable Δp , which is calculated using the Newton–Raphson method. Afterwards, the variables σ_{n+1} , ε_{n+1} , ε_{n+1}^p , $\mathbf{X}_{n+1}^{(k)}$, R_{n+1} and p_{n+1} at the time step $n+1$ can be determined. The backward Euler method is applied in the stress update algorithm to ensure that the yield function be satisfied at all time steps.

The same algorithm for the damage evolution and stress update is used in the implicit implementation. Furthermore, the consistent tangent modulus for the stress update algorithm should be provided to calculate the strain increment at the next time step:

$$C_{\text{alg}} = \frac{d\sigma_{n+1}}{d\varepsilon_{n+1}}. \quad (29)$$

The results of the stress and fatigue damage obtained through the explicit and implicit implementations are compared. Moreover, the number of ΔN affects the prediction accuracy of the fatigue lifetime. Although ΔN should be small enough to obtain convergence results for the damage evolution and fatigue lifetime, a too small value of ΔN will result in a tremendous computational cost. Thus, a trade-off is required between the simulation convergence and the computational efficiency. A suitable value of ΔN is chosen through multiple attempts for the convergence prediction of spalling lifetime [25].

3.4. Modeling of spalling initiation and propagation

As the roller rolls cyclically on the out ring, the stress and strain distributions in two contacting components are calculated at each time step. The fatigue damage of the out ring is also obtained every ΔN cycles. The damage limit is set as 0.95 rather than 1 to avoid numerical singularity because the term $(1-D)$ acts as a denominator in the model. When the damage of an element reaches this value, spalling initiates and the lifetime of spalling initiation is recorded. The damage of this element remains unchanged after the initiation. The damage of other elements gradually reaches 0.95 during subsequent loading cycles. A piece of material is chipped out as a completely damaged element reaches the contact surface of the out ring.

4. Results and discussion

4.1. Comparison of the explicit and implicit implementation results

The loading condition with the maximum peak contact pressure of $p_{\text{max}} = 2500$ MPa and the coefficient of friction of $\mu = 0.05$ is simulated using the explicit and implicit implementations. The number ΔN is set to be 1×10^4 cycles for both numerical methods. Fig. 3(a) shows the results of the contact pressure distribution in the first cycle obtained by both methods. Two curves including the contact width are nearly identical, indicating the consistent of the stress calculation through two implementations. Fig. 3(b) illustrates the comparison between the damage evolution results at the spalling initiation position predicted by the explicit and implicit implementations. Slight discrepancy is observed for two damage evolution curves before 1.1×10^6 cycles, which is mainly attributed to the minor difference between the subsurface stress distributions obtained by two implementations. The maximum fatigue damage in this stage is around 0.3. However, as the fatigue damage accumulates over cycles, the discrepancy of two damage evolution results become large after 1.1×10^6 cycles. Afterwards, the fatigue damage increases rapidly since the plastic strain and the related plastic damage becomes

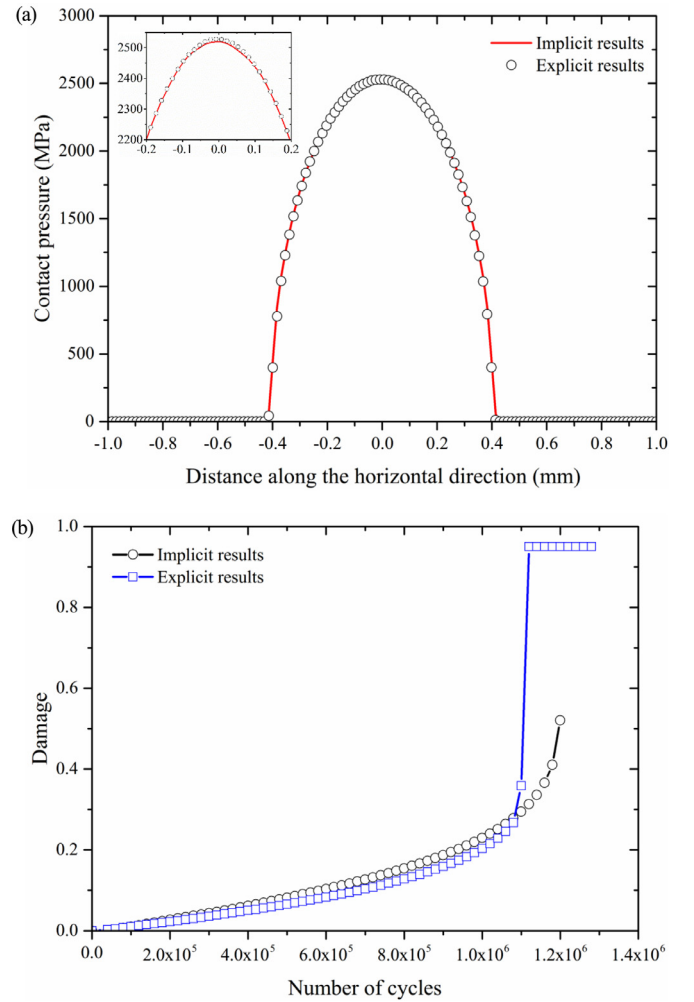


Fig. 3. Comparison between the results obtained by the implicit and explicit methods: (a) contact pressure distribution in the first cycle and (b) damage evolution at the spalling initiation position for the simulation case ($p_{\text{max}} = 2500$ MPa, $\mu = 0.05$).

dominant at a high damage level. The discrepancy between the spalling initiation lifetimes obtained by the explicit and implicit implementations is around 9%.

For the same simulation condition of $p_{\text{max}} = 2500$ MPa and $\mu = 0.05$, the smaller number of $\Delta N = 5 \times 10^3$ cycles is used for the explicit implementation to validate the convergence of the predicted fatigue lifetime. Fig. 4 depicts the comparison between the step-wise damage evolution results obtained by the explicit method under two different numbers of ΔN , and the result difference is negligible, indicating the result convergence. Therefore, the number ΔN is set to be 1×10^4 cycles for this simulation condition to save computational cost. Note that the suitable number of ΔN varies for different simulation conditions.

4.2. Initiation and propagation of spalling

Fig. 5 shows the damage distributions in the ring after different numbers of loading cycles under the loading condition of $p_{\text{max}} = 2500$ MPa and $\mu = 0.05$. The fatigue damage zone firstly occurs beneath the contact surface of the ring. The first element that reaches the damage value of 0.95 is located at the depth of around 0.2 mm, indicating the initiation of spalling at 1.1×10^6 cycles. The material near this element also experiences high damage, which cover a horizontal length of around 0.25 mm smaller than the displacement applied on the roller. If the applied displacement increases, the length covered by the damaged elements

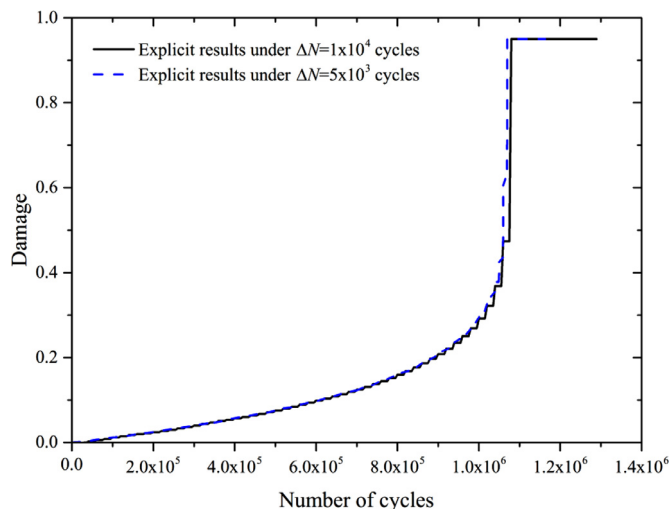


Fig. 4. Comparison between the step-wise damage evolution results obtained by the explicit method under two different numbers of ΔN (1×10^4 and 5×10^3 cycles) for the simulation case ($p_{max} = 2500$ MPa and $\mu = 0.05$).

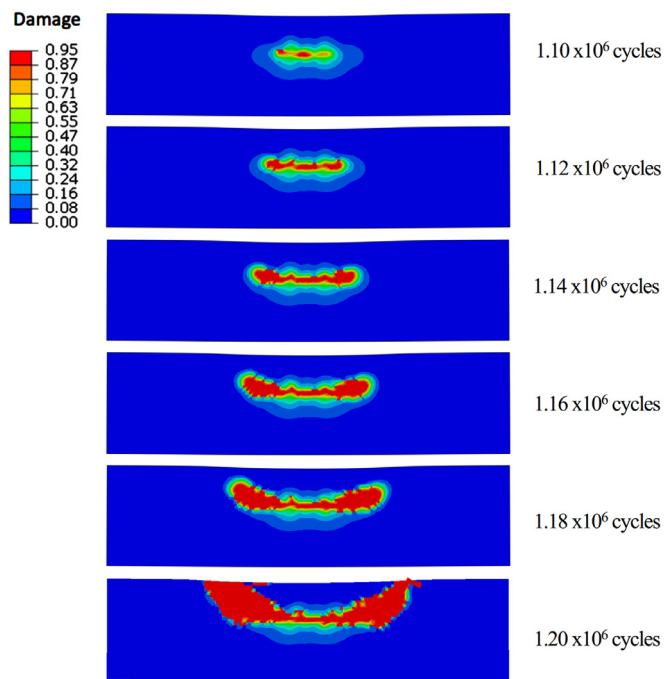


Fig. 5. Evolution of the damage distribution in the ring under the simulation condition of $p_{max} = 2500$ MPa and $\mu = 0.05$.

will increase. After the initiation of spalling, the damage of the elements near the initiation point also reaches 0.95 quickly. The propagation of spalling after 1.12×10^6 cycles is nearly along the horizontal direction. Then the damaged elements gradually move to the contact surface. A piece of material is finally chipped out after 1.2×10^6 cycles. The direction of the spalling propagation is around 21.8° with respect to the horizontal direction. The length of the spalling is about 0.87 mm. The fatigue lifetime of the spalling propagation is about 1×10^5 cycles, which is one order of magnitude less than the spalling initiation lifetime. The initiation of spalling consumes nearly 90% of the total fatigue lifetime of bearings. Therefore, it is critical to detect the initiation of spalling in bearings through non-destructive testing.

In the present study, the material of the cylindrical bearings is assumed to be perfect without the presence of inclusions. If inclusions are present in a material, they will cause stress concentration to provide

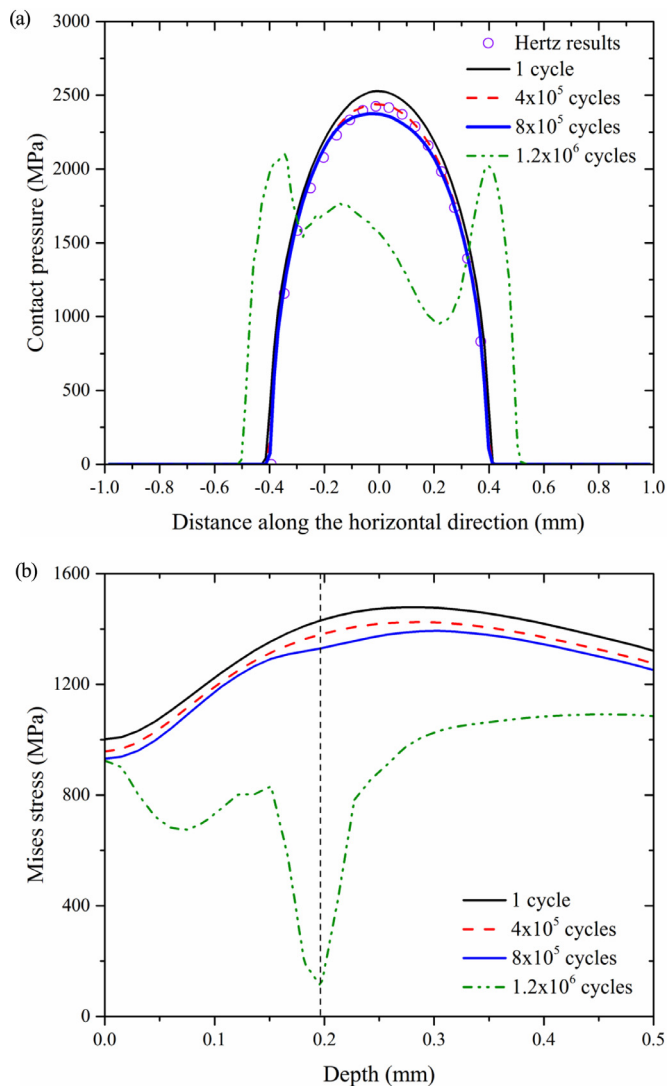


Fig. 6. (a) Evolution of the surface contact pressure distribution of the ring, and (b) evolution of the von Mises stresses along the depth direction of the ring ($p_{max} = 2500$ MPa and $\mu = 0.05$).

sites for crack initiation and greatly facilitate the initiation process. As a result, the initiation stage of spalling will account for only a small fraction of the total fatigue lifetime. An experimental study on the RCF behavior of GCr15 bearings revealed that subsurface micro-cracks initiated at the tips of inclusions made of ferric oxide and the crack initiation duration accounted for about 10% of the total fatigue lifetime [49].

The predicted results of the spalling lifetime are compared with the experimental results in [50]. In the experiment, the roller diameter of the cylindrical bearings made of AISI 52100 steel in the experiment was 9 mm, a value taken in the present simulation; the peak contact pressure was 2410 MPa and the total spalling lifetime was 1.53×10^6 cycles. In the simulation, the spalling lifetimes obtained under two peak contact pressures, i.e., 2250 and 2500 MPa, are 2.9×10^6 and 1.2×10^6 cycles, respectively. The experimental data of the peak contact pressure and the spalling lifetime fall in the range of the simulation data, indicating that the developed elasto-plastic-damage model can provide reasonable prediction of bearing spalling under RCF loading.

4.3. Contact pressure, subsurface stress and fatigue damage evolution

Fig. 6(a) plots the evolution of the contact pressure distribution on the ring. Since neither fatigue damage nor plastic deformation is present

within the first cycle, the small discrepancy between the contact pressure distribution and the Hertz theoretical solution during that cycle may be due to the finite dimensions of the ring in the modeling. As the number of loading cycles increases to 8×10^5 cycles, the peak contact pressure decreases from the initial value of 2500 to 2370 MPa. The maximum damage of 0.166 is located at a depth of about 0.2 mm beneath the surface during this cycle. Thus, the fatigue damage distribution has a negligible effect on the surface contact pressure. The normal force is carried by a layer of undamaged material at the contact surface. However, after 1.2×10^6 cycles, the contact pressure no longer exhibits a Hertzian distribution. The contact pressure in the contact zone of $-0.4 < x < 0.4$ mm decreases significantly. This is because the completely damaged elements reach the contact surface during the process of crack propagation. The elements in this contact zone have degraded material properties due to high fatigue damage. Furthermore, the contact width increases from around 0.8 to 1 mm. The pressure peaks occur at two contact edges.

Fig. 6(b) illustrates the evolution of the von Mises stresses along the depth direction of the ring. The start position is set as the origin point of the coordinate system as shown in Fig. 1. The von Mises stress decreases with an increasing fatigue damage. The maximum decrease in the von Mises stress happens at the point of spalling initiation, i.e., at a depth of 0.2 mm. The fatigue damage at this position after 8×10^5 cycles is about 0.13, causing the von Mises stress to decrease from the initial value of 1431 to 1330 MPa. As the number of loading cycles increases to 1.2×10^6 cycles, the fatigue damage reaches 0.95 and the material at this position completely loses the loading-carrying capacity, resulting in a dramatic decrease in the von Mises stress.

Fig. 7(a) illustrates the four stress components at the point of spalling initiation during the first cycle. The stress components demonstrate non-proportional variations over loading cycles. The compressive stress in the y-direction is much larger than the other three components but is not the driving force to cause crack initiation. The shear stress σ_{xy} with a relatively high value is critical for crack initiation under RCF loading.

Fig. 7(b) depicts the history of the damage and the equivalent plastic strain at the position of spalling initiation. The fatigue damage increases slowly from 0 to around 0.2 during the first 1×10^6 cycles, and then dramatically increases to 0.95 in the subsequent 1×10^5 cycles. Although the stress level is reduced due to the fatigue damage, the term $(1-D)$ in the denominator of the damage evolution law still causes the damage to increase rapidly under a high damage level. A similar evolution of the fatigue damage was also obtained by Li et al. [12]. The plastic deformation occurs as the number of loading cycles reaches 6.2×10^5 cycles, during which the fatigue damage is around 0.087. The fatigue damage causes the stress at the position of spalling initiation to reach the yield surface, causing damage-induced plastic deformation. The equivalent plastic strain demonstrates a slow increase before the damage reaches 0.95.

4.4. Competition between fatigue damage and plastic deformation

The influences of the fatigue damage and the plastic deformation on the shear stress σ_{xy} are evaluated. Fig. 8 shows the evolution of the shear stress-strain curve after different loading cycles. During the first cycle, there is no fatigue damage and plastic strain, and the shear stress-strain curve is linear. After 5×10^5 cycles, the fatigue damage increases to 0.066 without the occurrence of plastic deformation, and the curve remains linear. The maximum shear stress keeps unchanged and the minimum value decreases by about 30 MPa. However, the inclination angle of the curve decreases as compared with that for the first cycle because the fatigue damage weakens the material properties including the shear modulus. After 1×10^6 cycles, the fatigue damage increases to around 0.2 and the average equivalent plastic strain over this cycle is about 0.13. The damage-induced plastic deformation causes the curve to possess an obvious hysteresis loop. The hardening effect of the plastic deformation on the shear stress is dominant as compared with the weak-

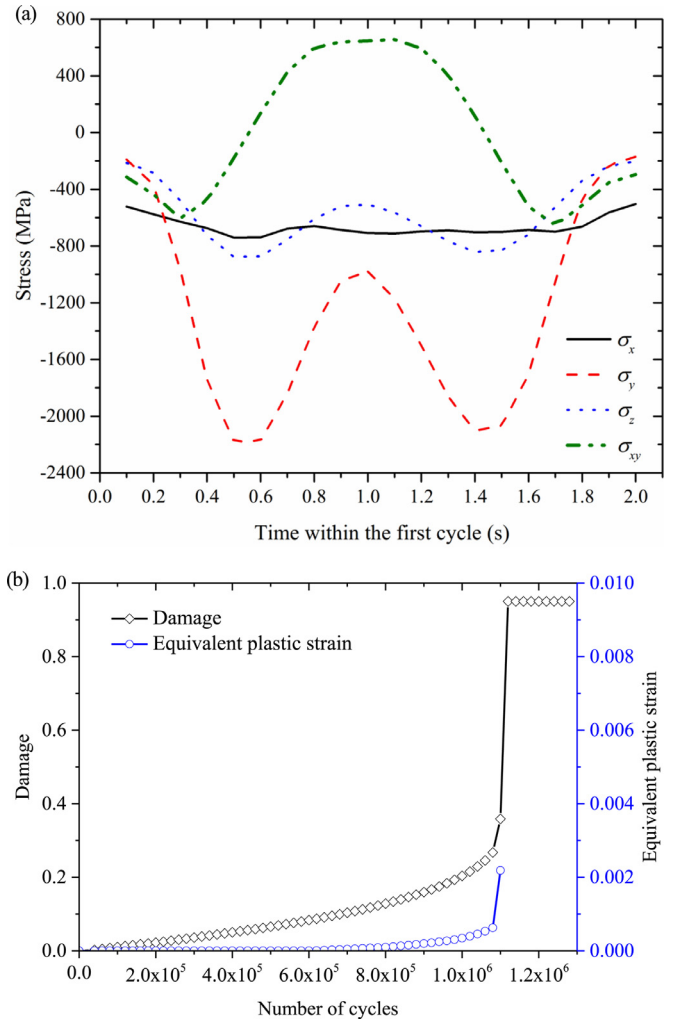


Fig. 7. (a) History of the stress components at the spalling initiation point within the first cycle, and (b) history of the fatigue damage and the equivalent plastic strain at the same point over loading cycles ($p_{max} = 2500$ MPa and $\mu = 0.05$).

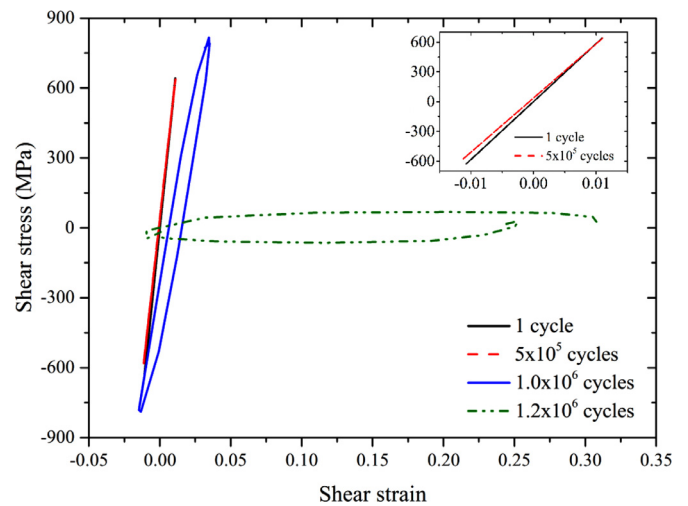


Fig. 8. Evolution of the shear stress-strain curves after different loading cycles ($p_{max} = 2500$ MPa and $\mu = 0.05$).

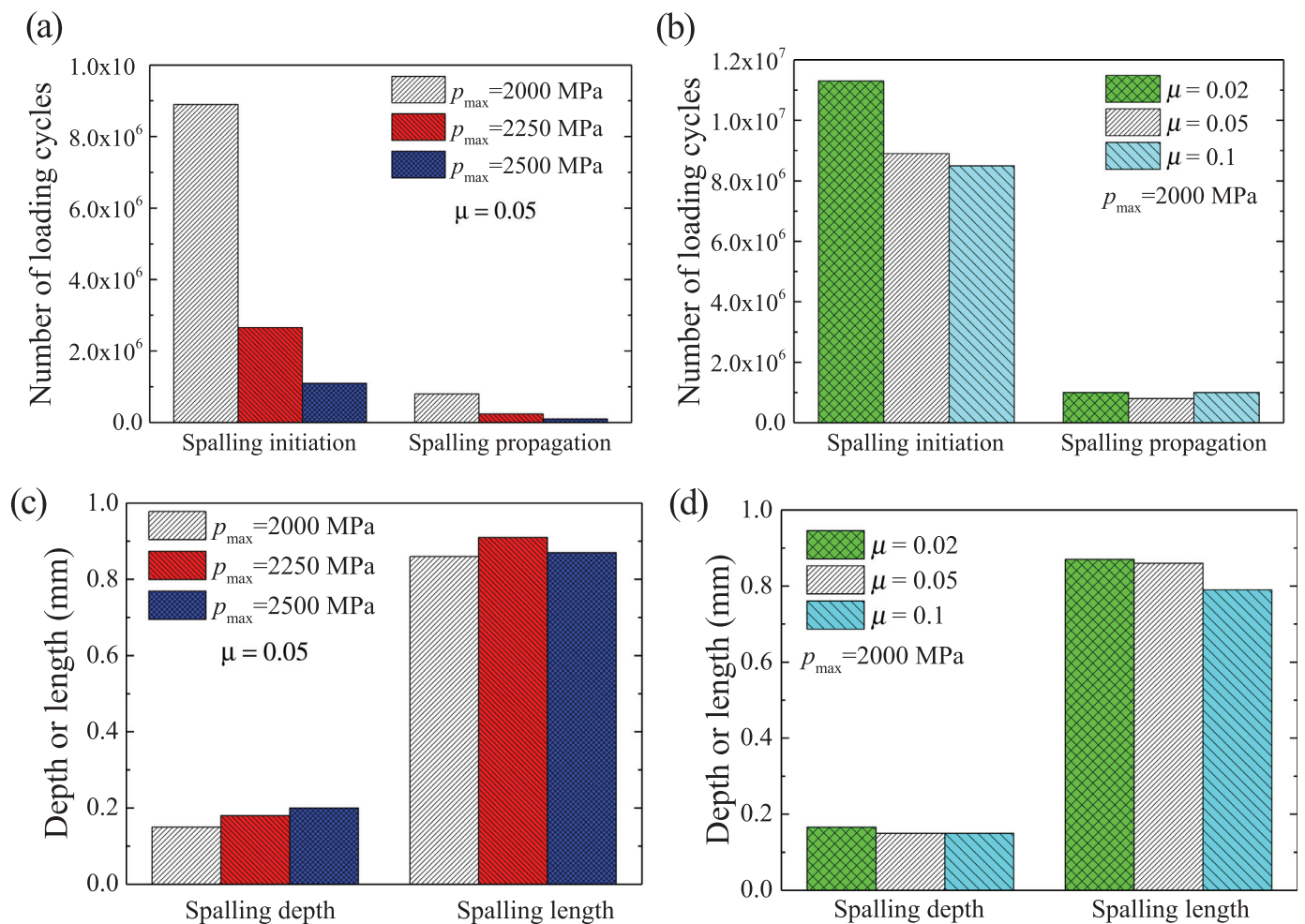


Fig. 9. Effects of (a) the peak contact pressure and (b) the coefficient of friction on the spalling initiation and propagation lifetimes; effects of (c) the peak contact pressure and (d) the coefficient of friction on the spalling depth and length.

ening effect of the fatigue damage. Therefore, the maximum shear stress increases to around 790 MPa. As the fatigue damage reaches 0.95 after 1.2×10^6 cycles, the weakening effect of the fatigue damage on the shear stress becomes dominant. The material completely loses its capacity of loading-carrying. The maximum shear stress reduces to around 60 MPa and the shear strain range during the cycle increases dramatically. The competition between the influences of the fatigue damage and the plastic deformation governs the evolution of the shear stress-strain response.

4.5. Effects of the contact pressure and coefficient of friction

The effects of the maximum peak contact pressure and the coefficient of friction on the initiation and propagation processes of spalling are studied. Fig. 9(a) and (b) illustrate the spalling initiation and propagation lifetimes obtained under different peak contact pressures and coefficients of friction. The spalling propagation lifetime is much less than the initiation lifetime under these five loading conditions. For the initiation lifetime, the peak contact pressure has a more significant influence than the coefficient of friction. The initiation lifetime under $p_{max} = 2500$ MPa is only one ninth of that under $p_{max} = 2000$ MPa. The coefficient of friction affects the tangential traction distribution on the surface contact zone. The maximum surface tangential tractions under $\mu = 0.02$ and 0.1 are around 36.8 and 188.5 MPa, respectively, both of which are close to but less than the values obtained by directly using the Coulomb friction law. The increase in the coefficient of friction causes the octahe-

dral shear stress range mentioned in Eq. (15) to increase from 880 to 929 MPa, leading to the small variation in the spalling lifetime.

Fig. 9(c) and (d) illustrate the spalling depths and lengths obtained under different peak contact pressures and coefficients of friction. Under higher contact pressure, the location of the maximum octahedral shear stress becomes deeper in the ring. As a result, the spalling depth increases from around 0.15 to 0.2 mm as the contact pressure increases from 2000 to 2500 MPa. The effect of the contact pressure on the spalling length is not obvious. As the coefficient of friction increases from 0.02 to 0.1, the tangential traction on the contact surface also increases, causing the maximum octahedral shear stress to occur close to the contact surface. Accordingly, the spalling depth slightly decreases.

5. Conclusions

The initiation and propagation of spalling in bearings under RCF loading are investigated through a numerical approach. An elasto-plastic-damage model is proposed to evaluate the stress-strain response of damaged materials and their damage evolution during the rolling process. Specifically, the damage-coupled elasto-plastic model is developed in the thermodynamic framework by specifying the damage-related Helmholtz free energy and dissipative potential function. The non-proportional behavior of the cyclic stress and the plastic strain are used to calculate the fatigue damage evolution.

The initiation and propagation processes of spalling are simulated through the numerical implementation of the developed model. The

surface contact pressure and fatigue damage evolution obtained by the explicit implementation are compared with those by the implicit implementation. The shear stress is the driving force for the spalling initiation. The spalling initiates beneath the contact surface of the ring. A piece of material is chipped out as the damage reaches the contact surface. The predicted spalling lifetime agrees well with the experimental data from the literature. The evolutions of contact pressure, subsurface stress, equivalent plastic strain and fatigue damage during the rolling process are obtained and their relationships are evaluated. It is found that the fatigue damage competes with the plastic deformation to affect the shear stress-strain response during the rolling process. Furthermore, the effects of the maximum peak contact pressure and the coefficient of friction on the initiation and propagation processes of spalling are studied. The peak contact pressure has more significant influence than the coefficient of friction on the spalling lifetime.

Acknowledgment

The authors acknowledge the financial support from the SMRT-NTU Smart Urban Rail Corporate Lab, Singapore. K. Z. acknowledges the financial support from the Nanyang Environment and Water Research Institute (Core Fund), Nanyang Technological University, Singapore.

References

- Bhadeshia H. Steels for bearings. *Prog Mater Sci* 2012;57(2):268–435.
- Rosado L, Forster NH, Thompson KL, Cooke JW. Rolling contact fatigue life and spall propagation of AISI M50, M50NiL, and AISI 52100, part I: experimental results. *Tribol Trans* 2009;53(1):29–41.
- Forster NH, Rosado L, Ogden WP, Trivedi HK. Rolling contact fatigue life and spall propagation characteristics of AISI M50, M50 NiL, and AISI 52100, part III: metal-lurgical examination. *Tribol Trans* 2009;53(1):52–9.
- Walvekar AA, Sadeghi F. Rolling contact fatigue of case carburized steels. *Int J Fatigue* 2017;95:264–81.
- Paulson NR, Sadeghi F, Habchi W. A coupled finite element EHL and continuum damage mechanics model for rolling contact fatigue. *Tribol Int* 2017;107:173–83.
- Guan J, Wang L, Mao Y, Shi X, Ma X, Hu B. A continuum damage mechanics based approach to damage evolution of M50 bearing steel considering residual stress induced by shot peening. *Tribol Int* 2018;126:218–28.
- He H, Liu H, Zhu C, Wei P, Sun Z. Study of rolling contact fatigue behavior of a wind turbine gear based on damage-coupled elastic-plastic model. *Int J Mech Sci* 2018;141:512–19.
- Wang W, Liu H, Zhu C, Wei P, Tang J. Effects of microstructure on rolling contact fatigue of a wind turbine gear based on crystal plasticity modeling. *Int J Fatigue* 2019;120:73–86.
- Warhadpande A, Sadeghi F, Kotzalas MN, Doll G. Effects of plasticity on subsurface initiated spalling in rolling contact fatigue. *Int J Fatigue* 2012;36(1):80–95.
- Sadeghi F, Jalalahmadi B, Slack TS, Raje N, Arakere NK. A review of rolling contact fatigue. *J Tribol* 2009;131(4):041403.
- Arakere NK, Branch N, Levesque G, Svendsen V, Forster NH. Rolling contact fatigue life and spall propagation of AISI M50, M50NiL, and AISI 52100, part II: stress modeling. *Tribol Trans* 2009;53(1):42–51.
- Li F, Hu W, Meng Q, Zhan Z, Shen F. A new damage-mechanics-based model for rolling contact fatigue analysis of cylindrical roller bearing. *Tribol Int* 2017;120:105–14.
- Lundberg G, Palmgren A. Dynamic capacity of roller bearing. *Acta Polytech. Mechanical Engineering Series*; 1952.
- Liu Y, Stratman B, Mahadevan S. Fatigue crack initiation life prediction of railroad wheels. *Int J Fatigue* 2006;28(7):747–56.
- Jiang Y, Sehitoglu H. A model for rolling contact failure. *Wear* 1999;224(1):38–49.
- Zhou RS, Cheng HS, Mura T. Micropitting in rolling and sliding contact under mixed lubrication. *J Tribol* 1989;111(4):605–13.
- Zhou RS. Surface topography and fatigue life of rolling contact bearings. *Tribol Trans* 1993;36(3):329–40.
- Lemaitre J. A continuous damage mechanics model for ductile fracture. *J Eng Mater Technol* 1985;107(1):83–9.
- Lemaitre J, Chaboche JL. *Mechanics of solid materials*. Cambridge: Cambridge university press; 1994.
- Lemaitre J, Desmorat R. *Engineering damage mechanics: ductile, creep, fatigue and brittle failures*. New York: Springer Science & Business Media; 2005.
- Voyiadjis GZ, Kattan PI. A comparative study of damage variables in continuum damage mechanics. *Int J Damage Mech* 2008;18(4):315–40.
- Ghosh A, Leonard B, Sadeghi F. A stress based damage mechanics model to simulate fretting wear of Hertzian line contact in partial slip. *Wear* 2013;307(1–2):87–99.
- Shen F, Hu W, Meng Q. A damage mechanics approach to fretting fatigue life prediction with consideration of elastic–plastic damage model and wear. *Tribol Int* 2015;82:176–90.
- Hojjati-Talemi R, Wahab MA. Fretting fatigue crack initiation lifetime predictor tool: using damage mechanics approach. *Tribol Int* 2013;60:176–86.
- Shen F, Hu W, Meng Q. A non-local approach based on the hypothesis of damage dissipation potential equivalence to the effect of stress gradient in fretting fatigue. *Int J Fatigue* 2016;90:125–38.
- Yao Y, He X, Keer LM, Fine ME. A continuum damage mechanics-based unified creep and plasticity model for solder materials. *Acta Mater* 2015;83:160–8.
- Do VNV, Lee C-H, Chang K-H. High cycle fatigue analysis in presence of residual stresses by using a continuum damage mechanics model. *Int J Fatigue* 2015;70:51–62.
- Kang G, Liu Y, Ding J, Gao Q. Uniaxial ratcheting and fatigue failure of tempered 42CrMo steel: damage evolution and damage-coupled visco-plastic constitutive model. *Int J Plast* 2009;25(5):838–60.
- Sun Y, Hu W, Shen F, Meng Q, Xu Y. Numerical simulations of the fatigue damage evolution at a fastener hole treated by cold expansion or with interference fit pin. *Int J Mech Sci* 2016;107:188–200.
- Shen F, Voyiadjis GZ, Hu W, Meng Q. Analysis on the fatigue damage evolution of notched specimens with consideration of cyclic plasticity. *Fatigue Fract Eng Mater Struct* 2015;38(10):1194–208.
- Chaboche JL. A review of some plasticity and viscoplasticity constitutive theories. *Int J Plast* 2008;24(10):1642–93.
- Bhattacharyya A, Pandkar A, Subhash G, Arakere N. Cyclic constitutive response and effective S–N diagram of M50 NiL case-hardened bearing steel subjected to rolling contact fatigue. *J Tribol* 2015;137(4):041102.
- Bhattacharyya A, Subhash G, Arakere N, Allison BD, McCoy B. Influence of residual stress and temperature on the cyclic hardening response of M50 high-strength bearing steel subjected to rolling contact fatigue. *J Eng Mater Technol* 2016;138(2):021003.
- Guo Y, Liu C. Mechanical properties of hardened AISI 52100 steel in hard machining processes. *J Manuf Sci Eng* 2002;124(1):1–9.
- Kang G. Ratcheting: recent progresses in phenomenon observation, constitutive modeling and application. *Int J Fatigue* 2008;30(8):1448–72.
- Ding J, Kang G, Zhu Y, Zhu M. Finite element analysis on bending fretting fatigue of 316L stainless steel considering ratcheting and cyclic hardening. *Int J Mech Sci* 2014;86:26–33.
- Karvan P, Varvani-Farahani A. Ratcheting assessment of 304 steel samples by means of two kinematic hardening rules coupled with isotropic hardening descriptions. *Int J Mech Sci* 2018;149:190–200.
- Shen F, Zhou K. Investigation on thermal response in fretting sliding with the consideration of plastic dissipation, surface roughness and wear. *Int J Mech Sci* 2018;148:94–102.
- Pandkar AS, Arakere N, Subhash G. Microstructure-sensitive accumulation of plastic strain due to ratcheting in bearing steels subject to rolling contact fatigue. *Int J Fatigue* 2014;63:191–202.
- Pandkar AS, Arakere N, Subhash G. Ratcheting-based microstructure-sensitive modeling of the cyclic hardening response of case-hardened bearing steels subject to rolling contact fatigue. *Int J Fatigue* 2015;73:119–31.
- Hahn GT, Bhargava V, Rubin CA, Chen Q, Kim K. Analysis of the rolling contact residual stresses and cyclic plastic deformation of SAE 52100 steel ball bearings. *J Tribol* 1987;109(4):577–95.
- Hahn GT, Bhargava V, Chen Q. The cyclic stress-strain properties, hysteresis loop shape, and kinematic hardening of two high-strength bearing steels. *Metall Trans A* 1990;21(2):653–65.
- Guo YB, Wen Q, Horstemeyer MF. An internal state variable plasticity-based approach to determine dynamic loading history effects on material property in manufacturing processes. *Int J Mech Sci* 2005;47(9):1423–41.
- Şimşir C, Dalgıç M, Lübben T, Irretier A, Wolff M, Zoch HW. The Bauschinger effect in the supercooled austenite of SAE 52100 steel. *Acta Mater* 2010;58(13):4478–91.
- Wolff M, Suhr B, Şimşir C. Parameter identification for an Armstrong–Frederick hardening law for supercooled austenite of SAE 52100 steel. *Comput Mater Sci* 2010;50(2):487–95.
- Zhang T, McHugh P, Leen S. Finite element implementation of multiaxial continuum damage mechanics for plain and fretting fatigue. *Int J Fatigue* 2012;44:260–72.
- Lemaitre J. *A course on damage mechanics*. 2nd ed. Springer Science & Business Media; 2012.
- Belytschko T, Liu WK, Moran B, Elkhodary K. *Nonlinear finite elements for continua and structures*. John Wiley & sons; 2013.
- Chen L, Chen Q, Shao E. Study on initiation and propagation angles of subsurface cracks in GCr15 bearing steel under rolling contact. *Wear* 1989;133(2):205–18.
- Harris TA, Barnsby RM. Life ratings for ball and roller bearings. *Proc Inst Mech Eng Part J* 2005;215(6):577–95.